



Bioestatística quantitativa aplicada

Edison Capp
Otto Henrique Nienov
Organizadores

Caroline Darski
Charles Francisco Ferreira
Cristiana Palma Kuhl
Fernanda Dapper Machado
Fernanda Vargas Ferreira
Hellen Meiry Grosskopf Werka
Johanna Ovalle Diaz
Marina Petter Rodrigues
Michele Strelow Moreira
Nadine de Souza Ziegler
Paula Barros Terraciano
Pedro Henrique Comerlato
Sinara Santos

Universidade Federal do Rio Grande do Sul
Faculdade de Medicina
Programa de Pós-Graduação em Ciências da Saúde:
Ginecologia e Obstetrícia

Bioestatística Quantitativa Aplicada

Porto Alegre 2020
UFRGS

U58b Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação em Ciências da Saúde: Ginecologia e Obstetrícia Bioestatística quantitativa aplicada/ Universidade Federal do Rio Grande do Sul; organizadores: Edison Capp e Otto Henrique Nienov – Porto Alegre: UFRGS, 2020.

260p.

ISBN: 978-65-86232-43-1

E-Book: 978-65-86232-44-8

1. Epidemiologia e Bioestatística 2. Estatística 3. SPSS I. Capp, Edison, org. II. Nienov, Otto Henrique, org. III Título.

NLM: WA950

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
(Bibliotecária Shirlei Galarça Salort – CRB10/1929)

Endereço:

PPG em Ciências da Saúde: Ginecologia e Obstetrícia

FAMED – UFRGS

Rua Ramiro Barcellos, 2400/2º andar

CEP 900035-003 – Porto Alegre – RS

Telefone: +55 51 3308 5607

E-mail: ppggo@ufrgs.br

Editoração e diagramação: Edison Capp

Capa: Edison Capp, imagens: www.freepik.com/starline

Edison Capp
Otto Henrique Nienov
Organizadores

Caroline Darski
Charles Francisco Ferreira
Cristiana Palma Kuhl
Fernanda Dapper Machado
Fernanda Vargas Ferreira
Hellen Meiry Grosskopf Werka
Johanna Ovalle Diaz
Marina Petter Rodrigues
Michele Strelow Moreira
Nadine de Souza Ziegler
Paula Barros Terraciano
Pedro Henrique Comerlato
Sinara Santos

Universidade Federal do Rio Grande do Sul

Reitor: Prof. Rui Vicente Oppermann

Vice-Reitora: Profa. Jane Fraga Tutikian

Pró-Reitor de Pós-Graduação: Prof. Celso Giannetti Loureiro Chaves

Faculdade de Medicina

Diretora: Profa. Lúcia Maria Kliemann

Vice-Diretor: Prof. Luciano Zubaran Goldani

Departamento de Ginecologia e Obstetrícia

Chefe: Profa. Solange Garcia Accetta

Chefe Substituta: Profa. Suzana Arenhart Pessini

PPG em Ciências da Saúde: Ginecologia e Obstetrícia

Coordenador: Prof. Edison Capp

Coordenador substituto: Prof. Eduardo Pandolfi Passos

Hospital de Clínicas de Porto Alegre

Diretora Presidente: Profa. Nadine Clausell

Diretor Médico: Prof. Milton Berger

Coordenador GENS: Prof. José Geraldo Lopes Ramos

Coordenadora GPPG: Profa. Patrícia Ashton Prolla

Serviço de Ginecologia e Obstetrícia

Chefe: Profa. Maria Celeste Osório Wender

Serviço de Mastologia

Chefe: Prof. Jorge Villanova Biazús

Centro de Estudos do Serviço de Ginecologia e Obstetrícia

Presidente: Prof. Jorge Villanova Biazús

Fundação Médica do Rio Grande do Sul

Presidente: Prof. Fernando Gomes

“Daqui a um tempo não muito remoto, se entenderá que, para uma iniciação completa como um cidadão eficiente será necessário poder calcular, pensar em médias e máximos e mínimos, como agora se é capaz de ler e escrever”.

H.G. Wells
Mankind in the Making

Autores

Caroline Darski

Graduação em Fisioterapia, Universidade Federal do Rio Grande do Sul (UFRGS). Mestrado no Programa de Pós-Graduação em Ciências da Saúde Ginecologia e Obstetrícia (PPGGO), UFRGS. Doutoranda PPGGO, UFRGS.

Charles Francisco Ferreira

Graduando em Toxicologia Analítica, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA). Graduação em Ciências Biológicas: Bacharelado, Universidade Estadual do Oeste do Paraná (UNIOESTE). Mestre e Doutor em Ciências Biológicas: Neurociências, UFRGS. Pós-Doutorado no PPGGO, UFRGS. Pesquisador do Escritório de Projetos, Hospital Moinhos de Vento (HMV).

Cristiana Palma Kuhl

Graduação em Fisioterapia, Centro Universitário Metodista IPA. Mestrado no Programa de Pós-Graduação em Medicina: Ciências Cirúrgicas, UFRGS. Doutoranda PPGGO, UFRGS

Edison Capp

Graduação em Medicina, UFRGS. Residência Médica em Ginecologia e Obstetrícia, Hospital Materno-Infantil Presidente Vargas, Porto Alegre. Doutor em Medicina pela Universidade Ludwig-Maximilian, Munique, Alemanha. Professor do Departamento de Ginecologia e Obstetrícia e do PPGGO, FAMED, UFRGS.

Fernanda Dapper Machado

Graduação em Farmácia, Universidade Feevale. Mestrado PPGGO. Doutoranda PPGGO, UFRGS.

Fernanda Vargas Ferreira

Graduação em Fisioterapia, Universidade Federal de Santa Maria, UFSM. Doutorado e Pós-Doutorado PPGGO, UFRGS. Professora do Curso de Fisioterapia, Universidade Federal do Pampa, UNIPAMPA.

Hellen Meiry Grosskopf Werka

Graduação em Medicina, Universidade Federal de Pelotas. Residente de Patologia Cirúrgica, Hospital de Clínicas de Porto Alegre, HCPA. Mestranda, PPGGO, UFRGS.

Johanna Ovalle Diaz

Graduação em Medicina, Universidade Militar Nueva Granada, Bogotá, Colombia, UMNG. Graduação em Urologia, Fundação Universitaria Sanitas, Bogotá, Colombia. Pós-Graduação em Urologia Infantil e Reconstrutiva, HCPA. Mestranda PPPGO, UFRGS.

Marina Petter Rodrigues

Graduação em Fisioterapia, UFRGS. Especialização em Fisioterapia Pélvica, Faculdade Inspirar, Porto Alegre. Mestrado PPGGO. Doutoranda PPGGO, UFRGS.

Michele Strelow Moreira

Graduação em Licenciatura Plena em Matemática, Universidade Católica de Pelotas. Graduação em Farmácia, Universidade Federal de Pelotas. Pós-Graduanda em Farmácia Hospitalar e Clínica, Faculdade Ciências da Saúde Moinhos de Vento. Mestranda PPGGO, UFRGS.

Nadine de Souza Ziegler

Graduação em Medicina pela Universidade Federal de Santa Maria. Especialista em Ginecologia e Obstetrícia e Medicina Fetal pela Febrasgo. Especialista em Ultrassom em Ginecologia e Obstetrícia pelo Conselho Brasileiro de Radiologia. Mestranda PPGGO, UFRGS.

Otto Henrique Nienov

Graduação em Toxicologia Analítica, UFCSPA. Graduação em Biomedicina, Universidade FEEVALE. Mestrado, Doutorado e Pós-Doutorado PPGGO, UFRGS.

Paula Barros Terraciano

Graduação em Ciências Biológicas pela PUCRS. Mestrado e Doutorado em Ciências Veterinárias- Biotécnicas da Reprodução Animal, UFRGS. Mestrado em Ciências da Saúde: Ginecologia e Obstetrícia-UFRGS. Professora no PPGGO, UFRGS. Embriologista no Serviço de Ginecologia e Obstetrícia, HCPA.

Pedro Henrique Comerlato

Graduação em Medicina,UFCSPA. Residência Médica em Medicina Interna e em Medicina Intensiva, HCPA/UFRGS. Doutorando PPG Endocrinologia, UFRGS.

Sinara Santos

Graduação em Psicologia, Universidade Luterana do Brasil (ULBRA). Especialista em Psicologia Hospitalar, Hospital de Clínicas de Porto Alegre. Mestrado PPGGO. Doutoranda PPGGO, UFRGS.

Sumário

Apresentação	13
1 Conceitos básicos em estatística e epidemiologia.....	17
Projeto de pesquisa.....	22
Delineamento de pesquisa.....	23
Estudos observacionais	24
Estudo transversal	25
Estudo de coorte	25
Estudo de caso-controle	26
Estudo ecológico	27
Estudos experimentais	27
Ensaio clínico randomizado	27
Ensaio de campo.....	28
Ensaio comunitário.....	28
Estudos pré-clínicos.....	29
Estudo <i>in vivo</i>	29
Estudo <i>in vitro</i>	29
Estudo <i>in silico</i>	29
Conceitos básicos em estatística.....	29
População	30
Amostra.....	30
Indivíduos.....	30
Observação	31
Variável.....	31
Medidas de tendência central	33
Medidas de distribuição.....	34
2 Guia Básico do Microsoft Office Excel	41
Como montar um banco de dados	42
Conhecendo um banco de dados.....	42
Acessando as funções estatísticas.....	43
Distribuição de frequência	43
Gráficos	47
Viés em Pesquisa.....	48
Tipos de erro.....	49
Viés de seleção	50
Viés de informação (ou classificação).....	51
Viés de confirmação.....	54
Fator de confundimento ou de confusão.....	54

3 Primeiros passos no SPSS	57
Conhecendo o SPSS	57
Construindo um banco de dados no SPSS.....	57
Acessando um banco de dados existente.....	62
O que é um arquivo de sintaxe?	63
Os resultados no visualizador de saídas.....	64
Executando uma análise.....	64
Estatísticas descritivas no SPSS	65
Representação gráfica das análises descritivas	66
Outliers: o que fazer?	68
Distribuição de frequências no SPSS.....	68
Representação gráfica das frequências	69
Exportação do arquivo de saída para Word.....	70
Exportando um banco de dados do SPSS para o Excel	70
Importando um banco de dados do Excel para o SPSS	70
4 Transformando e manipulando dados no SPSS	73
Por que transformar e manipular dados?	73
Transformar variável quantitativa em categórica	74
Calcular a diferença entre duas datas	77
Combinando variáveis categóricas.....	78
Calculando variável através de outras existentes.....	81
Como criar uma variável a partir da contagem de outras	82
Classificação de casos a partir de uma ou mais variáveis.....	83
Restringindo a análise a um subconjunto de casos.....	84
Análises simultâneas em diferentes subconjuntos	87
Inclusão de novos casos ou novas variáveis	89
5 Teste de Hipóteses	93
Criando as hipóteses	93
As hipóteses estatísticas.....	94
Testando as hipóteses	94
Formas do teste de hipóteses.....	95
Erro do teste de hipóteses	98
Nível de significância.....	99
Valor de p ou P-valor.....	99
Etapas do teste de hipótese	100
6 Métodos estatísticos para desfechos qualitativos	101
Análise de dados categóricos para uma variável	101
Análise de dados categóricos para duas variáveis	103
Teste de Qui-quadrado de associação	105

Resíduo padronizado ajustado	106
Teste de Qui-quadrado de homogeneidade	108
Teste de Qui-quadrado de ajustamento	108
Complementos para variáveis nominais.....	110
Fi e V de Cramer	110
Coeficiente de contingência	111
Lambda	111
Coeficiente de Incerteza	111
Complementos para variáveis ordinais.....	111
Gama.....	112
d de Somers.....	112
Tau-b de Kendall	112
Tau-c de Kendall.....	112
Complementos para variáveis nominais por intervalo	113
Eta	113
Estatísticas de Cochran e Mantel-Haenszel	113
Variáveis ordinais com mais de duas categorias	113
Teste U de Mann-Whitney.....	113
Variáveis com duas categorias medidas em dois momentos	114
Testes de McNemar e de Wilcoxon	114
Variáveis em escala de medida ordinal	117
Teste de Kruskal-Wallis.....	118
Teste de Friedman	119
Variáveis em escala de medida nominal ou ordinal dicotomizadas	120
Teste Q de Cochran	120
Análise de concordância	121
Teste W de Kendall	121
Índice Kappa de Cohen	121
7 Desfechos quantitativos: amostras independentes	125
Comparação de amostras	125
Testes paramétricos e não paramétricos.....	126
Definindo a normalidade.....	128
Teste de normalidade para uma amostra.....	129
Teste de normalidade para duas ou mais amostras	131
Testes para amostras independentes.....	132
Teste t de Student.....	132
Teste U de Mann-Whitney.....	135
Análise de variância com um fator – ANOVA (<i>one-way</i>).....	137
Análise de variância com dois fatores – ANOVA <i>two-way</i>	143
Análise de covariância – ANCOVA.....	150
Teste de Kruskal-Wallis.....	153

8 Desfechos quantitativos: amostras pareadas.....	157
Teste t pareado.....	157
Teste de Wilcoxon.....	159
Análise de variância para medidas repetidas (ANOVA-MR).....	161
Equações de estimativas generalizadas (GEE).....	166
Teste de Friedman.....	170
Tamanho de efeito e sua implicação no cálculo amostral.....	172
9 Correlações, risco, razão de chances e avaliação de testes diagnósticos...177	
Correlações bivariadas.....	178
Coeficiente de correlação de Pearson.....	180
Coeficiente de correlação de Spearman.....	180
Coeficiente de correlação de Kendall.....	180
Risco relativo e razão de chance.....	183
Avaliação de testes diagnósticos.....	188
Curva ROC.....	190
10 Regressão linear simples e múltipla.....197	
Análise de regressão.....	197
Construindo modelos de regressão.....	198
Regressão linear.....	202
Regressão linear simples.....	208
Regressão linear múltipla.....	212
11 Modelos lineares generalizados.....217	
Regressão logística.....	217
Regressão logística binária.....	219
Regressão de Poisson.....	226
Análise de sobrevivência.....	228
Regressão de Cox.....	228
Resolução dos Exercícios.....233	
Capítulo 1.....	233
Capítulo 2.....	234
Capítulo 3.....	237
Capítulo 4.....	242
Capítulo 5.....	247
Capítulo 6.....	247
Capítulo 7.....	251
Capítulo 8.....	253
Capítulo 9.....	254
Capítulo 10.....	257
Capítulo 11.....	260

Apresentação

A realização de um curso de pós-graduação *stricto sensu* é entusiasmante e desafiadora. Assim, nada melhor que as dicas e sugestões de pares, de colegas que já estão nesta empreitada há algum tempo e até mesmo daqueles que já a concluíram, para facilitar a tarefa.

Este livro é despretensioso e originou-se a partir da disciplina “Bioestatística Quantitativa Aplicada” com a participação dos demais autores atuando como monitores. Foi assistida por mais de 100 alunos de pelo menos 12 diferentes programas de pós-graduação da UFRGS e de outras universidades. As aulas atravessaram os limites estaduais, tendo participantes do Sul ao Norte do Brasil, e nacionais, com alunos assistindo às aulas remotamente da Argentina e do Canadá. Foi toda ministrada de maneira remota e, apesar de tantos participantes, funcionou surpreendentemente bem. Houve oportunidade de esclarecimento de dúvidas e, mesmo sendo a primeira experiência de alguns com atividades remotas, as participações foram feitas com etiqueta que deve servir de exemplo para outras atividades neste formato.

O objetivo é ensinar conceitos estatísticos utilizando o SPSS como ferramenta. Bioestatística Quantitativa Básica foi escrito de modo que pudesse ser usado por vários níveis de leitores, desde os iniciantes até aqueles com conhecimento mais avançados em SPSS. Desta forma, para proporcionar um melhor entendimento, tanto da teoria quanto da prática, o livro conta com bancos de dados disponibilizados pelo link <https://bit.ly/bancosdedados>, que foram utilizados nos exemplos e na aplicação de exercícios práticos em cada capítulo.

Agradecimentos a todos que participaram da realização desta atividade.

Edison Capp
Otto Henrique Nienov
Organizadores

Apresentação FAMED

É com muita satisfação que apresentamos esse novo livro de estatística, produzido a partir do nosso Programa de Pós-Graduação em Ciências da Saúde: Ginecologia e Obstetrícia. A produção técnica, independente da área, não pode prescindir de um desenvolvimento baseado na qualidade e no estrito rigor ao método científico. Uma informação sem qualidade pode ser muito mais danosa do que a falta de informação. E é dentro desse espírito de formação de bons pesquisadores que essa publicação foi produzida, almejando levar ao maior número possível de alunos de graduação e pós-graduação e demais membros das equipes de pesquisa, através de uma linguagem acessível, algumas das ferramentas necessárias para a produção de evidências solidamente embasadas. Aos autores, o agradecimento da Faculdade de Medicina da Universidade Federal do Rio Grande do Sul!

Lúcia Maria Kliemann
Diretora FAMED/UFRGS

1 Conceitos básicos em estatística e epidemiologia

*Caroline Darski
Cristiana Kuhl
Edison Capp
Otto Henrique Nienov*

A pesquisa científica é a aplicação prática de um conjunto de procedimentos objetivos, utilizados por um pesquisador (investigador), para o desenvolvimento de um experimento ou estudo, com a finalidade de produzir um novo conhecimento que pode ser integrado àqueles preexistentes. Constitui-se, portanto, de etapas ordenadamente dispostas, de maneira lógica e racional, as quais o pesquisador deverá conhecer para aplicá-las convenientemente. Estas etapas, de maneira sucinta, incluem desde a escolha do tema a ser pesquisado, o planejamento da investigação, o desenvolvimento do método (que deve ser adequado para responder o problema da pesquisa), a coleta e a tabulação dos dados, a análise dos resultados, a elaboração das conclusões, até a divulgação/publicação do trabalho final.

Portanto, para a realização de uma pesquisa com o rigor científico e metodológico apropriado, pressupõe-se que o pesquisador escolha um tema de sua preferência e seu domínio, defina o problema a ser investigado, elabore um plano de trabalho consistente e, após a execução operacional desse plano, compile e analise os resultados obtidos, e escreva um relatório final, o qual deve ser redigido de forma clara, concisa e responder o objetivo do estudo. Todos estes procedimentos deverão ser executados em etapas.

A definição do tema é o primeiro passo para a elaboração do protocolo de pesquisa. O pesquisador deverá se perguntar: "O que, de fato, quero estudar?". Dentro do tema proposto, o investigador deverá selecionar a questão da pesquisa, a qual corresponde a uma parte delimitada do assunto escolhido. No estabelecimento da questão da pesquisa, o pesquisador deverá partir do assunto geral, o qual deverá ser desmembrado em tópicos específicos, em partes e, então, escolherá uma ou duas dessas partes para elaborar o protocolo de pesquisa.

Uma vez selecionado o tema, a definição do problema é o passo seguinte e de sua correta formulação, dependerá o sucesso da pesquisa. Toda pesquisa nasce de algum problema, elaborado na forma de uma pergunta de pesquisa, que norteará o projeto. É preciso desenvolver uma ideia clara do problema a ser resolvido. Tendo em mente a ideia, você começa a formular hipóteses que explicam a origem desse problema. Portanto, a pergunta de pesquisa estabelecida deve ser estruturada, clara e factível de ser respondida. Lembre-se sempre: todos os procedimentos propostos para a realização da pesquisa deverão ser planejados no sentido de solucionar ou esclarecer o problema proposto. A ordem correta de raciocínio é: “Qual é a questão que necessita de investigação e/ou solução?” “O que ela causa?” “No que a minha pesquisa irá contribuir para solucioná-la”? As características de uma boa questão de pesquisa estão resumidas no quadro 1, no qual estas características básicas são representadas pelo acrônimo FINER: factível, interessante, nova (inovadora), ética e relevante.

Quadro 1 - Características de uma boa questão de pesquisa (adaptado de Hulley e cols., 2008).

Característica	Requisitos
Factível	<ul style="list-style-type: none"> - Número adequado de sujeitos - Domínio técnico - Viável em tempo e custos - Escopo manejável
Interessante	<ul style="list-style-type: none"> - Para o investigador
Nova (Inovadora)	<ul style="list-style-type: none"> - Confirma ou refuta achados anteriores - Expande os achados anteriores
Ética	<ul style="list-style-type: none"> - Cumpre as normas da Resolução 466/2012 e as demais resoluções do Conselho Nacional de Saúde (CNS).
Relevante	<ul style="list-style-type: none"> - Para conhecimento científico - Para as diretrizes clínicas de saúde - Para direcionamentos futuros de pesquisa

Portanto, antes do desenvolvimento do projeto de pesquisa, questione se a maneira como a pesquisa está sendo planejada é passível de ser respondida, se você possui domínio suficiente do assunto e experiência para realizá-la e se o tempo e os recursos (sejam financeiros, materiais e/ou humanos) disponíveis serão suficientes. Abaixo, estão descritos detalhadamente os aspectos do acrônimo FINER, que podem ajudar a responder as questões citadas:

Factível: é necessário conhecer não somente seus limites, mas também os recursos disponíveis, antes de se iniciar o estudo. Isso evita o gasto de tempo e de recursos materiais e financeiros. Exemplos de perguntas que devem ser feitas antes de se iniciar um projeto de pesquisa: É possível ser executado? O número de sujeitos é adequado? Domínio técnico e recursos humanos estão disponíveis? Há viabilidade de tempo, custos, espaço físico e recursos materiais?

Interessante: é importante que o problema da pesquisa seja de interesse do pesquisador. O crescimento profissional é um motivo importante e deve ser construído ao longo da sua carreira. Cada pesquisa é um avanço nesta trajetória. No entanto, o conhecimento deve ser a principal motivação do pesquisador.

Inovadora: toda boa pesquisa deve produzir novos conhecimentos, ou, pelo menos, questionar, confirmar ou refutar algum achado anterior, ou, ainda, se os resultados obtidos podem ser replicados ou generalizados em outras populações e cenários. Além disso, uma pesquisa científica não precisa ser necessariamente inédita ou inovadora, porém repetir estudos cujos resultados já estão bem estabelecidos pela comunidade científica, não é recomendado, uma vez que desperdiçam trabalho, tempo e recursos.

Ética: neste aspecto, precisamos considerar se a pesquisa está delineada de modo que não comprometa a integridade, a autonomia e o bem-estar do participante da pesquisa. A Resolução do Conselho Nacional da Saúde N° 466, de 12 de dezembro de 2012, considera que toda pesquisa envolvendo seres humanos, direta ou indiretamente, envolve risco, podendo o dano eventual ser imediato ou tardio, no indivíduo ou à coletividade. Desta forma, toda e qualquer pesquisa envolvendo seres humanos deve ser submetida

à apreciação de um Comitê de Ética em Pesquisa (CEP). As pesquisas envolvendo seres humanos devem atender às exigências éticas e científicas fundamentais, que implicam consentimento livre e esclarecido dos indivíduos pesquisados, proteção de grupos vulneráveis (como gestantes e comunidades indígenas) e dos legalmente incapazes (como crianças e pessoas com deficiência intelectual), respeitando o princípio da autonomia. Potenciais riscos e benefícios deverão ser ponderados, individuais ou coletivos, comprometendo-se, a pesquisa, com o máximo de benefícios e a redução de danos, seguindo o princípio da beneficência. Por fim, estima-se a relevância social, com vantagens significativas para os sujeitos e minimização do ônus para os vulneráveis, o que garante a igual consideração dos interesses envolvidos sem perder o sentido de sua destinação sócio-humanitária, conforme os princípios da justiça e de equidade. A prática que envolve o uso de animais é regulada pelo Conselho Nacional de Controle de Experimentação Animal (CONCEA), que orienta as Comissões de Ética no Uso de Animais (CEUA), pesquisadores e docentes na utilização de animais em aulas e pesquisas.

Relevante: sem dúvida, este é o requisito mais importante quando se elabora uma questão de pesquisa. Se a pergunta de pesquisa não for relevante, ela pode tornar-se desinteressante para o pesquisador e sua equipe. Sua relevância está no fato de mostrar e justificar como o estudo pretendido pode ser inserido em um contexto mais amplo. Deve-se questionar por que essa questão é importante e como suas respostas poderão contribuir para o conhecimento científico. A questão de pesquisa deve ser cuidadosamente revisada e estruturada e, caso haja necessidade, o projeto deve ser adequado para que possa ser respondida.

Considerando que os critérios do método FINER descrevem os importantes aspectos da questão em geral, um formato útil para usar no desenvolvimento de uma questão de pesquisa específica é o acrônimo PICOT: considera a população (P) de interesse, a intervenção (I) em estudo, a comparação (C) (grupo ou com o que a intervenção está sendo comparada) e o resultado do interesse (O). Também podemos incluir no acrônimo o delineamento do estudo (T), ou seja, o tipo de estudo que melhor responderá a questão de pesquisa. A abordagem PICOT ajuda a gerar uma questão que

auxilia na construção da estrutura do estudo e, posteriormente, no desenvolvimento do protocolo, citando os critérios de inclusão e exclusão e identificando os grupos de pacientes a serem incluídos. Conhecer a população de interesse, a intervenção e o resultado esperado também pode ajudar o pesquisador a identificar uma ferramenta de medição de resultados apropriada.

A questão de pesquisa é o ponto de partida de qualquer projeto de pesquisa, ou seja, é quando o pesquisador define um problema e como procurará abordá-lo. A questão de pesquisa, como mencionado, pode ser construída através do acrônimo PICOT (exemplos no Quadro 2), no qual:

P = população, paciente ou problema/com quem a pergunta se relaciona;

I = intervenção/exposição/pode ser uma terapia, teste diagnóstico, fator prognóstico ou interesse;

C = comparador/não exposição/pode ser outra intervenção, teste diagnóstico, placebo ou cuidado usual padrão;

O = desfecho/especificar claramente no que você está interessado;

T = delineamento/decidir o melhor desenho do estudo que possa responder a sua questão de pesquisa”.

Quadro 2 - Exemplos de PICOT.

		Qual o efeito de antitussígenos na melhora em sinais e sintomas respiratórios e na qualidade de vida de pacientes com tosse?	Qual o desempenho de aprendizagem em crianças com deficiência visual comparado a crianças sem deficiência visual?	Qual o impacto do tabagismo no desenvolvimento do câncer de pulmão em jovens adultos?
Item	Descrição	Exemplo 1	Exemplo 2	Exemplo 3
P	População	Pacientes com tosse	Crianças	Jovens adultos
I	Intervenção/ Exposição	Antitussígenos	Deficiência visual	Hábito de fumar
C	Comparador	Não fazer nada/ placebo	Sem deficiência visual	Sem hábito de fumar
O	Outcome/ Desfecho	Supressão da tosse, qualidade de vida	Aprendizagem	Câncer de pulmão
T	Delineamento	Ensaio clínico randomizado	Estudo de casos e controles	Estudo de coorte

Quanto mais definida a população de interesse e, portanto, quanto melhor definidos os critérios de inclusão e exclusão, maior o efeito na interpretação e subsequente generalização dos resultados da pesquisa. Restringir a população do estudo pode limitar o viés e aumentar a validade interna do estudo. No entanto, essa abordagem limitará a validade externa do estudo e, portanto a generalização dos achados para outros cenários ou populações. Por outro lado, uma população de estudo menos definida pode melhorar a representatividade da amostra, mas aumentar o viés e reduzir a validade interna do estudo. Uma pergunta de pesquisa mal elaborada implicará na escolha equivocada do desenho do estudo e, potencialmente, dificultará a chance de se obter um achado com relevância clínica (positiva ou negativa), o que poderá prejudicar a publicação. Sem dedicar tempo e recursos suficientes para o desenvolvimento da questão de pesquisa, a qualidade do estudo e os resultados subsequentes podem ser comprometidos. Durante o estágios iniciais de qualquer pesquisa, é fundamental formular uma pergunta de pesquisa que seja clinicamente relevante e ética.

Projeto de pesquisa

A elaboração do projeto corresponde à etapa mais importante e de maior complexidade da pesquisa, pois, do correto delineamento (desenho) depende o sucesso na obtenção das respostas esperadas pela questão da pesquisa. Delinear uma pesquisa é, em última análise, planejar a realização de sua parte científica operacional, tanto experimental como observacional; ou seja, é escrever corretamente um projeto onde estarão previstas todas as etapas de sua realização.

Com uma boa questão de pesquisa, se parte para o planejamento e elaboração do projeto de pesquisa para que esse seja submetido nas plataformas necessárias. A parte estatística se encontra inserida dentro da seção Método (também chamada de Material e Métodos ou Metodologia) do seu projeto de pesquisa, no qual devem constar seu cálculo de amostra e quais análises estatísticas são esperadas para responder a questão de pesquisa, conforme observado na figura 1.



Figura 1 - Partes de um projeto de pesquisa.

A compreensão do método estatístico, suas características e, principalmente, suas limitações, é fundamental para uma análise crítica responsável. A análise estatística adequada evita erros nos resultados e, conseqüentemente, na interpretação dos dados.

Delineamento de pesquisa

Os estudos podem ser classificados quanto ao período de tempo durante o qual os dados foram registrados em relação ao tempo no qual o estudo começou. São considerados retrospectivos quando se parte do efeito para a causa, ou seja, os dados sobre os pacientes já existem (ex. em banco de dados ou prontuários eletrônicos). São considerados prospectivos quando se parte da causa e se procura o efeito, com acompanhamento de pacientes ao longo do tempo (*follow up*, estudo com seguimento, sequencial). Estudos prospectivos geralmente são mais caros e sujeitos a perdas de participantes.

O delineamento do projeto precisa ser definido e deve ser apresentado explicitamente como item. Os estudos podem ser divididos em observacionais e experimentais, conforme o esquema da figura 2.

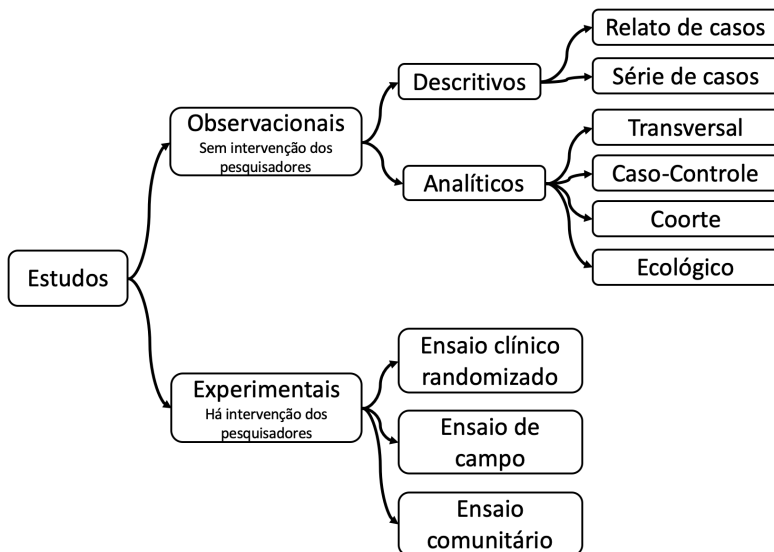


Figura 2 - Delineamentos de pesquisa.

Estudos observacionais

São aqueles em que não há intervenção do pesquisador. O pesquisador simplesmente observa o paciente, as características da doença ou transtorno, e sua evolução, sem interferir ou modificar qualquer aspecto que esteja sob observação. Incluem estudos de coorte, estudos de caso-controle e estudos transversais. Podem ser classificados em descritivos ou analíticos. Os estudos descritivos não testam hipóteses (as hipóteses surgem durante o projeto), portanto não fazem associações. Eles descrevem uma condição de indivíduo ou pequeno grupo de indivíduos. São geralmente usados para assuntos ainda não muito conhecidos como doenças novas, manifestações raras ou associações de doenças, caracterizando seus aspectos semiológicos, etiológicos, fisiopatológicos e epidemiológicos. São considerados estudos descritivos estudos de casos e séries de casos onde a diferença entre eles é o número de indivíduos. O primeiro pode variar entre 1 a 9 indivíduos e o segundo tem um mínimo de 10 participantes.

Os estudos analíticos, por sua vez, testam hipóteses (definidas *a priori*). Estão relacionados a distribuição de eventos na população, estimando incidência e prevalência. O investigador introduz um fator de exposição ou de causalidade, e avalia-o utilizando ferramentas estatísticas. A precisão dos dados está diretamente associada com a representatividade dos dados na população (daí a importância do cálculo amostral). São considerados estudos analíticos os estudos transversal, de coorte, caso-controle e ecológico.

Estudo transversal

É o “retrato” da situação no momento em que o pesquisador observa. A população ou amostra é estudada em um único momento. É utilizado para avaliar se existe relação entre as variáveis da pesquisa e é importante para determinação de prevalência de doenças e acurácia de testes diagnósticos. O estudo transversal é um estudo barato e fácil de ser realizado e gera hipóteses de associação ao risco, porém apresentam desvantagens como a impossibilidade de estabelecer relações causais por não provarem a existência de uma sequência temporal entre a exposição ao fator e o subsequente desenvolvimento da doença (não há possibilidade de determinação do que é exposição e do que é efeito). São exemplos de estudos transversais: estudo ecológico (as medidas são observadas em grandes populações), estudo de prevalência (para determinar a prevalência de uma característica em uma população de interesse é medida a proporção entre o número de pessoas acometidas pela doença e as que permanecem saudáveis), estudo de sensibilidade e especificidade (para avaliar um novo teste diagnóstico é utilizado um padrão de referência, representado por um teste ideal conhecido como “padrão-ouro”).

Estudo de coorte

É um estudo longitudinal, ou seja, envolve acompanhamento do cenário ao longo do tempo (como se fosse um filme). Um grupo de sujeitos com uma característica comum é acompanhado para verificar a ocorrência ou não de determinado desfecho. Os indivíduos são distribuídos em expostos e não expostos a um fator em estudo e são acompanhados durante um determinado período de tempo para verificar a incidência de uma

doença e fatores de risco. Compara-se a proporção dos que ficaram doentes dentre os expostos e a proporção dos que ficaram doentes entre os não expostos. Esse modelo de estudo é do tipo analítico e longitudinal, e pode ser prospectivo ou retrospectivo (coorte histórica). As vantagens desse estudo são: sequência temporal dos acontecimentos, a doença é claramente estabelecida, é ideal para ver incidência, tem alta precisão e pode medir diversos resultados. Como desvantagens, o estudo de coorte é um estudo caro e demorado, não é útil para doenças raras e há perdas durante o seguimento (se espera 20% de perda dos pacientes).

Estudo de caso-controle

Também chamado de estudo caso-referência, é um estudo onde se parte de indivíduos com doença (casos) e sem doença (controle) para investigar a associação entre um evento de interesse e alguns preditores, ou seja, o parâmetro a ser estudado é a exposição ou não a um fator (Figura 3). É um estudo importante para analisar doenças raras e situações de surtos ou agravos desconhecidos. É um estudo mais rápido que o Coorte, analisa vários preditores e é um estudo inicial para novas hipóteses, mas suas informações geralmente são incompletas, possui um importante viés de memória e fatores de confusão e não avalia frequência dos eventos. São modelos de estudo analíticos, longitudinais e, geralmente, retrospectivos (casos já tiveram o diagnóstico antes do início do estudo), porém podem ser prospectivos em algumas situações (quando os casos são diagnosticados após o início do estudo).

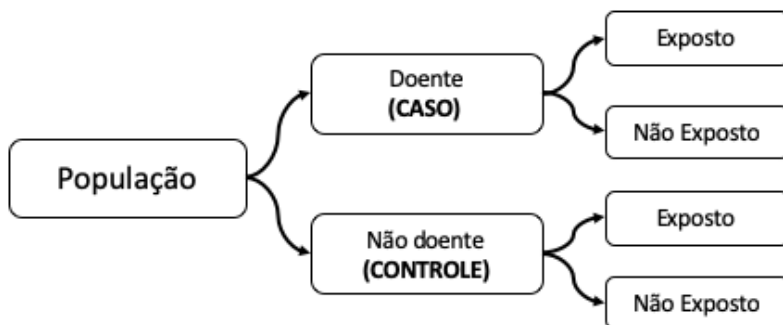


Figura 3 - Estudo do tipo caso-controle.

Estudo ecológico

É um estudo onde a unidade de análise é a população ou um grupo de pessoas pertencentes a uma área geográfica definida. São avaliados como os contextos social e ambiental podem afetar a saúde da população de interesse. Esse estudo compara variáveis globais quase sempre por meio de correlação entre indicadores de condições de vida e indicadores de situação de saúde. É um estudo que avalia efeitos contextuais e gera e testa novas hipóteses, porém há dificuldade de controlar fatores de confusão e não há possibilidade de associação entre fator e doença no indivíduo.

Estudos experimentais

Há intervenção do pesquisador. Podem ser classificados como ensaio clínico randomizado, ensaio de campo e ensaio comunitário.

Ensaio clínico randomizado

É um estudo prospectivo usado para comparar determinada intervenção com outra (Figura 5). É considerado o padrão-ouro para teste de eficácia de uma intervenção. As fases de seleção e inclusão dos pacientes são importantes para a alocação aleatória dos grupos. O mascaramento do ensaio clínico randomizado pode ser aberto (não ocorre), cego (apenas a equipe de investigação sabe qual foi o tratamento instituído e os pacientes desconhecem), duplo-cego (a equipe, os investigadores e tampouco os pacientes sabem a que grupos os mesmos pertencem), triplo-cego (equipe de investigação, pacientes e estatístico desconhecem qual é o grupo controle e o grupo experimental) ou quadruplo-cego (além da equipe de investigação, dos pacientes e do estatístico, o investigador que está escrevendo a discussão sobre os resultados, inicialmente de forma genérica quanto à identidade dos grupos, também não sabe qual é o grupo controle e o grupo experimental) (Figura 6). Como vantagens, esse estudo apresenta balanço dos fatores prognósticos entre os grupos, permite coleta de informações detalhadas e o cegamento dos participantes. Como desvantagens, tem-se a generalização externa reduzida por critérios de exclusão, é um tipo de estudo demorado com custo elevado, geralmente necessita de amostras grandes e os indivíduos podem não aderir às intervenções alocadas.

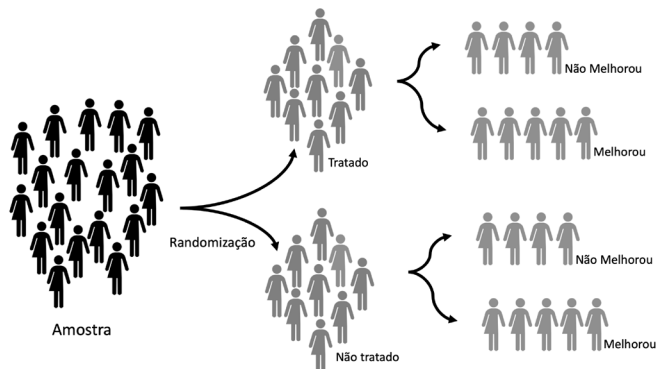


Figura 5 - Ensaio clínico randomizado.

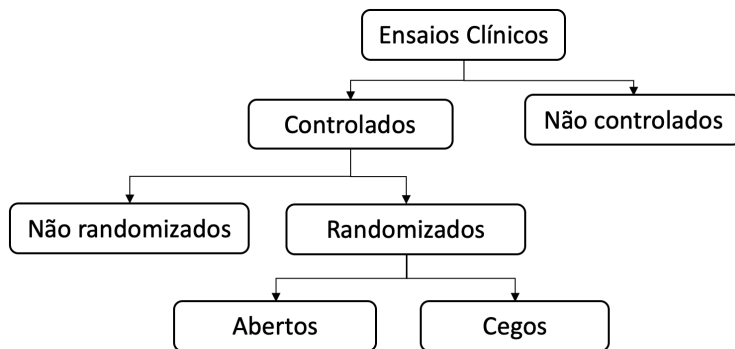


Figura 6 - Tipos de ensaios clínicos randomizados.

Ensaio de campo

É semelhante ao ensaio clínico, mas a população não é de pacientes e sim de indivíduos com risco de doenças. Os dados são coletados na população em geral.

Ensaio comunitário

Envolve a intervenção em nível de comunidades. Avalia a eficácia e efetividade de intervenções que busquem prevenção primária através de modificação de fatores de risco na população. É um estudo conduzido dentro de um contexto socioeconômico de uma população naturalmente formada. Uma importante limitação deste estudo é a dificuldade de isolar uma comunidade (difícil controle de fatores externos).

Estudos pré-clínicos

Estudo *in vivo*

Refere-se à experimentação realizada em animais vivos. São importantes para definição do perfil farmacológico e toxicológico de substâncias, testes de segurança, processos farmacocinéticos e estudos de pré-formulação e formulação.

Estudo *in vitro*

Refere-se a executar um determinado procedimento em um ambiente controlado fora de um organismo vivo. Frequentemente são utilizadas culturas de células. Uma das desvantagens é que dificilmente se consegue replicar as condições celulares de um organismo, levando a resultados que podem não corresponder às condições *in vivo*.

Estudo *in silico*

É feito através de técnicas computacionais (programas/softwares) que simulam eventuais efeitos e/ou interações.

Existe uma hierarquia na força da evidência conforme o delineamento dos estudos, sendo que quanto maior o nível de evidência, mais importante é a informação que o estudo apresenta para a tomada de decisão (Figura 7).

A definição do delineamento de pesquisa a ser utilizado é importante para saber quais testes estatísticos serão usados. Técnicas estatísticas insuficientes ou inadequadas podem ser uma grande fraqueza de um projeto de pesquisa. A qualidade e a adequação do método estatístico utilizado pode implicar no sucesso ou no fracasso de um artigo.

Conceitos básicos em estatística

A estatística é a ciência que se preocupa com a organização e descrição dos dados (estatística descritiva) e com análises e interpretações destes (estatística indutiva ou inferente). Tem como objetivo tirar conclusões sobre populações com base nos resultados observados em amostras extraídas dessas populações.



Figura 7 - Níveis de evidência.

População

É o conjunto de elementos (pessoas, animais,...) com pelo menos uma característica em comum. Esta característica comum deve delimitar claramente quais os elementos que pertencem à população e quais os elementos que não pertencem (através de critérios de inclusão e exclusão). Uma população pode ser finita e pequena, sendo fácil de conhecer todos os seus elementos. Porém, na maioria das vezes, é finita, mas incontável ou mesmo infinita. Nestes dois últimos casos, para conhecer uma população, a estatística lança mão de um recurso que é coletar uma amostra desta população e caracterizar alguma variável da população, ou seja, tirar conclusões sobre a população a partir de resultados obtidos em amostras (inferência estatística).

Amostra

É um subconjunto de uma população, definido segundo algum critério estabelecido pelo observador, sobre o qual se conduz algum estudo com objetivo de se fazer inferências estatísticas sobre toda a população. Esta deve ser representativa de toda a população de interesse.

Indivíduos

São os objetos descritos por um conjunto de dados. Podem ser pessoas, animais ou coisas.

Observação

É o valor de uma variável de um elemento particular de uma amostra ou população.

Variável

É a característica de interesse de elementos individuais de uma população ou amostra que nos interessa averiguar estatisticamente. Geralmente, correspondem às colunas no banco de dados. Podem ser qualitativas ou quantitativas, conforme figura 8.

As **variáveis qualitativas** são medidas originalmente em categorias. A descrição da característica de interesse resulta em valor não numérico. Exemplo: sexo (masculino = 0, feminino = 1), profissão, religião, município, região. São comumente resumidas pelo número de indivíduos em cada grupo (frequência absoluta, "n") e porcentagem de indivíduos em cada grupo (frequência relativa, "n%"). São classificadas em:

- Nominais: não existe ordenação nas possíveis respostas. Ex: sexo, estado civil;
- Ordinais: existe uma certa ordem nas possíveis respostas, ou seja, podem ser arranjados em alguma ordem. Ex: grau de escolaridade, hierarquias, classe de renda familiar.

As variáveis qualitativas são avaliadas por distribuição das frequências (descrição numérica), isto é, conhecer qual o comportamento dessa variável, analisando a ocorrência de suas possíveis realizações. Como mencionado anteriormente, são expressas em frequência absoluta e frequência relativa (tabela 1). Não se pode calcular média, variância ou desvio-padrão de variáveis qualitativas ou categóricas. A representação gráfica pode ser feita em gráficos em barras e de composição de setores (retângulos ou pizza).

Tabela 1 - Exemplo de variáveis qualitativas.

Grau de instrução	Frequência	Proporção	Porcentagem
Fundamental	12	0,333	33,33%
Médio	18	0,500	50,00%
Superior	6	0,167	16,67%
Total	36	1,000	100,00%

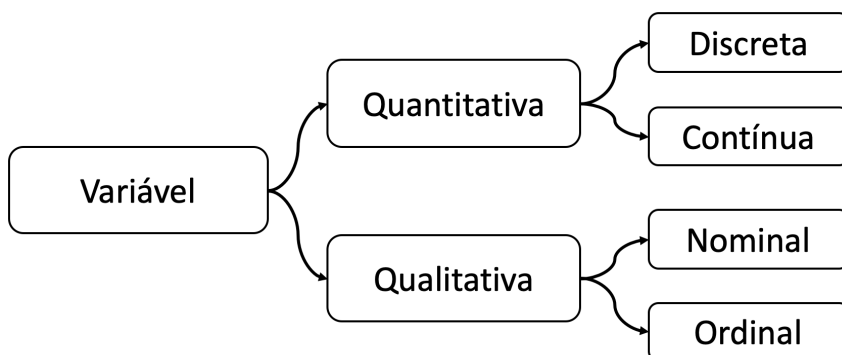


Figura 8 - Variáveis quantitativas e qualitativas.

As **variáveis quantitativas** podem ser mensuradas através de escalas, isto é, através de unidades de medida. Exemplos: idade (medida em anos, ou meses), massa (em kg), estatura (em cm). São expressas por medidas de tendência central (média, mediana e moda), de dispersão (variância, desvio padrão, amplitude), e em certos casos, a relação entre duas variáveis (massa e estatura). São classificadas em:

- Discretas: quando assumem apenas valores pertencentes a um conjunto finito ou enumerável de números inteiros. São obtidas mediante alguma forma de contagem. Ex: número de filhos;
- Contínuas: quando os possíveis valores estão dentro de um intervalo, aberto ou fechado, dos números reais. Ex: massa corporal, estatura, triglicerídeo sérico.

As variáveis quantitativas são expressas usando medidas descritivas (ex: média, desvio padrão, valor mínimo, valor máximo) e demonstradas em gráficos (ex: histograma, gráficos de dispersão e *box plot*).

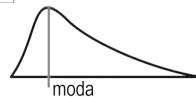
Medidas de tendência central

Servem para verificar a distribuição dos dados brutos (ou frequências) sobre o eixo de variação da variável em questão.

- Média: é o resultado do somatório de todos os valores da amostra dividido pelo número de valores (n);
- Moda (ou modas): é definida como o valor (ou valores) de máxima frequência. Uma amostra pode ser amodal (não ter moda), unimodal (apenas um valor com repetição expressiva), bimodal (dois valores tem repetições expressivas);
- Mediana: é o valor que está em posição intermediária na distribuição de valores, isto é, tem tantos elementos acima quanto abaixo de sua posição. Para determinar a mediana, primeiramente deve-se ordenar os "n" valores em ordem crescente. Quando "n" é ímpar, a mediana é um valor real da amostra. Quando "n" é par, a mediana é um valor fictício entre dois valores reais da amostra.

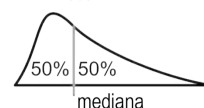
ID	P1	P2	P3	P4	P5	P6
Idade (em anos)	25	22	35	28	35	33

MODA: 35 anos de idade



MEDIANA:

22 25 28 33 35 35
 $28 + 33 = 61$
 $61/2 = 30,5$ anos de idade



MÉDIA:

$25 + 22 + 35 + 28 + 35 + 33 = 178$
 $178/6 = 29,67$ anos de idade

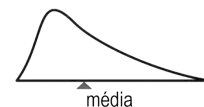


Figura 9 - Exemplos de mediana, moda e média.

Em distribuições normais (chamadas de simétricas ou paramétricas) a média e a mediana são próximas e os dados são expressos em média e desvio padrão. Em distribuições não normais (chamadas assimétricas ou não paramétricas) a média e a mediana são distantes e os dados são expressos em mediana e quartis ou valores mínimo e máximo.

Medidas de distribuição

Servem para indicar como os dados se distribuem em torno do valor central (média ou mediana). Portanto, caracterizam o grau de variação existente em um conjunto de valores.

- Amplitude: é definida como a diferença entre o maior e o menor valores do conjunto de dados;
- Variância: é definida como a medida do espalhamento dos valores em torno da média;
- Desvio padrão: é a raiz quadrada positiva da variância. É expressa na mesma unidade da variável, por isso é de maior interesse que a variância nas aplicações práticas, sendo mais realístico para efeito de comparação de dispersões. O desvio padrão é uma medida que indica a dispersão dos dados dentro de uma amostra com relação à média. Assim, quando se calcula o desvio padrão juntamente com a média de diferentes grupos, obtêm-se mais informações para avaliar e diferenciar seus comportamentos;
- Erro padrão: é uma medida de variação de uma média amostral em relação à média da população. Sendo assim, é uma medida que ajuda a verificar a confiabilidade da média amostral calculada. Através do erro padrão, pode-se estimar um intervalo de confiança (IC95%) para a média populacional a partir da média amostral calculada. A média das amostras \bar{x} encontra-se no centro desse intervalo, e o intervalo é $\bar{x} \pm$ intervalo de confiança.

O desvio padrão trata de um índice de dispersão da amostra em relação à média, enquanto o erro padrão é uma medida que ajuda a avaliar a confiabilidade da média calculada.

- Coeficiente de variação: é definido como o quociente entre o desvio padrão e a média, sendo frequentemente expresso em porcentagem;
- Percentil: são os valores de uma variável quantitativa que dividem os dados ordenados em grupos de modo que uma determinada porcentagem esteja acima e outra porcentagem esteja abaixo;
- Quartis: são valores da variável que dividem a distribuição em quatro partes iguais, cada uma delas com 25% dos dados totais ordenados. Mediana divide as observações em duas partes iguais. São utilizados em amostras com distribuições não normais (Figura 10).

Medidas de assimetria

São importantes para descrição visual quando a distribuição dos dados é normal ou não normal. Em uma distribuição estatística, a assimetria (*skewness*) é o quanto sua curva de frequência se desvia ou se afasta da posição simétrica.

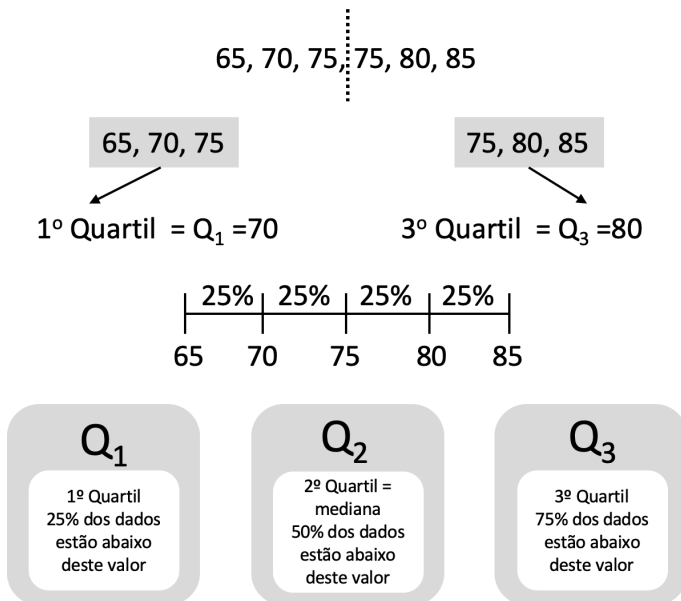


Figura 10 - Intervalos interquartis.

Curtose

É o menor ou maior grau de "achatamento" da distribuição ou curva de frequência considerada em relação a uma curva normal representativa da distribuição.

Referências

Bonita R, Beaglehole R, Kjellström T. Epidemiologia básica. 2. ed. São Paulo: Santos, 2010. 213 p.

Callegari-Jacques SM. Bioestatística: princípios e aplicações. Porto Alegre: ArtMed, 2011. 255p.

Farrugia P, Petrisor BA, Farrokhyar F, Bhandari M. Practical tips for surgical research: Research questions, hypotheses and objectives. Canadian Journal of surgery. Journal Canadien de Chirurgie. 2010;53(4):278-281.

Guimarães LSP, Hirakata VN, Camey AS, Nunes LN, Mancuso ACB. Os principais delineamentos na epidemiologia: ensaios clínicos (Parte II). Revista HCPA. 2013;33(3/4):295-302

Hulley SB, Cumming SR, Browner WS, Grady DG, Hearst NB, Newman TB. Delineando a pesquisa clínica: uma abordagem epidemiológica. Artmed, Porto Alegre; 2008. 384 p.

Mancuso ACB, Camey AS, Nunes LN, Hirakata VN, Guimarães LSP. Os principais delineamentos na epidemiologia: ensaios clínicos (Parte I). Revista HCPA. 2013;33(3/4):286-294.

Nunes LN, Camey AS, Guimarães LSP, Mancuso ACB, Hirakata VN. Os principais delineamentos na epidemiologia. Revista HCPA. 2013;33(2):178-183.

Regatieri CV, Alves AV, Rocha E. Medical acumen and scientific approach meet the FINER method. Arq. Bras. Oftalmol. 2019;82(4): 1-5.

Riva JJ, Malik KM, Burnie SJ, Endicott AR, Busse JW. What is your research question? An introduction to the PICOT format for clinicians. J Can Chiropr Assoc. 2012 Sep;56(3):167-71.

Exercícios sugeridos

Escolha a alternativa correta:

1. Para realizar um estudo epidemiológico, um investigador selecionou 33 mulheres com câncer de endométrio e 33 mulheres sem a doença. As mulheres foram pareadas por idade e paridade. Todas as mulheres foram questionadas sobre o uso de estrogênio e progesterona em algum momento de sua vida. O desenho de estudo utilizado foi:

- a) Estudo de coorte.
- b) Estudo caso-controle.
- c) Estudo ecológico.
- d) Estudo transversal.
- e) Ensaio clínico randomizado.

2. Um pesquisador está tentando determinar os fatores de risco para o câncer de mama. A amostra selecionada foi de 48.000 mulheres, acompanhados ao longo de 5 anos, para ver quais desenvolveriam a doença. Qual tipo de estudo é esse?

- a) Estudo de coorte.
- b) Estudo de caso-controle.
- c) Ensaio clínico randomizado.
- d) Estudo de caso.
- e) Estudo ecológico.

3. Um grupo de 500 mulheres é observado em um dado momento e mede-se a prevalência de Síndrome do Ovário Policístico (SOP). Que tipo de estudo é esse?

- a) Estudo caso-controle.
- b) Revisão sistemática e meta-análise.
- c) Estudo de coorte.
- d) Estudo ecológico.
- e) Estudo transversal.

4. Um pesquisador atende uma paciente com Síndrome de Rokitsky, uma doença muito rara que provoca alterações no útero e na vagina, e escreve um artigo sobre esta paciente, especialmente sobre a sua história, aspectos físicos e curso clínico. Que tipo de estudo é esse?

- (a) Estudo caso-controle.
- (b) Série de casos.
- (c) Estudo de coorte.
- (d) Estudo de caso.
- (e) Estudo transversal.

5. Para avaliar a relação entre alterações no percentil do Índice de Massa Corporal (IMC), através da curva de Atalah, e resultados perinatais, foi realizado um estudo com 1.475 mulheres. Os dados referentes ao peso gestacional, características sociodemográficas e resultados perinatais foram coletados de forma retrospectiva, por meio de prontuários, cartão pré-natal e entrevistas realizadas, em um momento do período pós-parto. As mulheres foram classificadas, de acordo com a curva de Atalah, em: baixo peso, peso adequado, sobrepeso e obesidade. Isto é um:

- a) Estudo transversal.
- b) Estudo de coorte.
- c) Ensaio clínico randomizado.
- d) Estudo caso-controle.
- e) Série de casos.

6. Um estudo envolvendo 32 parturientes, realizado em um hospital universitário, buscou determinar o desfecho materno e neonatal de acordo com a aplicação ou não da estimulação elétrica nervosa transcutânea (TENS) para alívio da dor do parto antes da analgesia combinada raqui-peridural (ACRP). Essas pacientes foram aleatoriamente designadas para receber ou não a TENS antes da ACRP. A análise dos resultados incluiu a intensidade da dor medida pela escala visual analógica (EVA), o intervalo entre a avaliação inicial e a instalação da ACRP, a duração do trabalho de parto, as taxas de parto por cesárea e por parto vaginal instrumental (fórceps) e, frequência de hipóxia neonatal. Isto representa um:

- a) Estudo transversal.
- b) Estudo de coorte.
- c) Ensaio clínico randomizado.
- d) Estudo caso-controle.
- e) Série de casos.

7. Um estudo teve como objetivo avaliar a frequência de deleção homozigótica dos genes GSTM1 e GSTT1 e suas combinações entre pacientes com câncer de mama e indivíduos saudáveis, associando-os à suscetibilidade da doença. Assim, 39 mulheres diagnosticadas com câncer de mama confirmadas por exame patológico e 39 mulheres saudáveis sem evidência de câncer e sem histórico familiar prévio de câncer de mama foram convidadas a participar. Todas responderam a um questionário e foram submetidos à coleta de amostras de sangue. O DNA foi extraído do sangue e a genotipagem foi realizada por reação em cadeia da polimerase (PCR). Isto caracteriza um:

- a) Estudo transversal.
- b) Estudo de coorte.
- c) Ensaio clínico randomizado.
- d) Série de casos.
- e) Estudo de caso-controle.

8. Um estudo avaliou as alterações de HDL-Colesterol induzidas pela cirurgia bariátrica em obesos não diabéticos com níveis baixos ou normais de HDL-Colesterol antes da cirurgia. Neste estudo de coorte prospectivo, 90 indivíduos foram avaliados quanto ao sexo (masculino ou feminino), idade (em anos), técnica cirúrgica (bypass gástrico ou gastrectomia em manga), tempo de cirurgia (em meses), grau de obesidade (sobrepeso, obesidade graus I, II e III), perda de peso (em kg), circunferência da cintura (em cm), níveis séricos de glicose (em mg/dL), triglicerídeos (em mg/dL), HDL-Colesterol (em mg/dL), não HDL-Colesterol (em mg/dL), LDL-Colesterol (em mg/dL), além do gasto energético estimado (em METs). Responda às seguintes questões:

a) Estruture a questão de pesquisa no acrônimo PICOS;

b) Das variáveis avaliadas, quais são consideradas quantitativas e quais são consideradas qualitativas? Classifique-as em nominal ou ordinal se qualitativas e em contínua ou discreta se quantitativas.

9. Neste exercício, estruture a questão de pesquisa do seu projeto de pesquisa no formato PICOT (*patients, interventions, comparisons, outcomes, study design*). Avalie se sua questão está bem definida, se é factível, interessante, inovadora, ética e relevante. Também descreva as possíveis variáveis do estudo e classifique-as quanto à sua natureza, se qualitativa (nominal ou ordinal) ou quantitativa (contínua ou discreta). Este é um bom exercício prático para aplicar os conceitos vistos neste capítulo.

2 Guia Básico do Microsoft Office Excel

Fernanda Dapper Machado

Fernanda Vargas Ferreira

Edison Capp

Otto Henrique Nienov

A primeira versão do Excel foi lançada em 1985 para o Mac e, em 1987, para o Windows. Posteriormente, com sua incorporação no Microsoft Office, suas versões gradativamente tornaram esta empresa a líder no desenvolvimento desse tipo de *software*. Atualmente, este programa de edição de planilhas eletrônicas tem sido empregado, por exemplo, para organização geral de planilhas para uso doméstico, acadêmico ou empresarial, realização de análises estatísticas ou financeiras e construção e manipulação de tabelas e gráficos.

Uma das primeiras dúvidas acerca da aplicabilidade desse programa se encontra em como “iniciar um banco de dados”. Entretanto, é fundamental identificar previamente como são as características da população ou amostra que são passíveis de representação numérica, obtidas como resultado de medições e contagens. Tais características são denominadas de variáveis e podem ser classificadas em qualitativas e quantitativas.

Variáveis qualitativas são resultado de uma classificação por atributos ou tipos, por exemplo, cor dos olhos. Adicionalmente, tais variáveis se subdividem em a) nominal, que ocorre por meio de “categorias” como sexo (masculino ou feminino) ou estado civil e b) ordinal, quando há uma ordenação, por exemplo, grau de satisfação e escolaridade.

Similarmente, variáveis quantitativas refletem números de uma contagem ou mensuração, como número de filhos e idade. Essas variáveis podem ser classificadas em a) contínuas, se os valores podem assumir qualquer valor em um determinado intervalo (ex.: massa, estatura), e b) discreta, cujos valores podem ser enumerados, por exemplo, número de ligações em um celular e número de discentes de uma disciplina ou em uma escola (página 30, Figura 8).

Como montar um banco de dados

No banco de dados, atribua um título para cada variável, o qual ficará disposto na primeira linha das colunas. Os dados referentes a cada caso, paciente ou objeto de pesquisa ficarão dispostos nas linhas, na respectiva célula ou casela. Lembre-se, sempre, de destinar uma variável para a identificação do caso, a qual pode ser feita por uma sequência numérica, facilitando a localização de informações no caso de serem identificados equivocados de digitação e também a organização do banco de dados.

É importante montar um “dicionário” do banco de dados para facilitar a identificação das variáveis do estudo em caso de compartilhamento ou quando você for revisitar o banco de dados futuramente. O dicionário deve identificar as variáveis, seu significado, a unidade de medida, a definição das siglas utilizadas e os rótulos dos valores das variáveis categóricas.

Conhecendo um banco de dados

A aplicação prática da estatística está relacionada com o problema exposto pelo estudo, ou seja, pela pergunta da pesquisa. Por isso, é fundamental o planejamento das análises estatísticas que serão realizadas para responder a pergunta antes de iniciarmos as coletas dos dados. A estatística faz parte do planejamento e envolve questões fundamentais, como o cálculo do tamanho da amostra necessário para responder a pergunta de pesquisa. Neste capítulo, vamos aprender a utilizar as funções básicas do Excel para caracterizar uma amostra e/ou descrever os seus resultados. Para isso, utilizaremos um banco de dados existente, o “Banco de dados 1.xls” (disponível em <https://bit.ly/bancosdedados>), que nos fornece dados de alunos de uma disciplina.

O primeiro passo ao se trabalhar com um banco de dados é conhecê-lo. Sempre é necessário ter uma identificação das variáveis dispostas nas colunas. Cada linha corresponde a um caso, usualmente a primeira coluna é reservada para a variável de identificação, onde cada caso receberá um código próprio. O restante das colunas corresponderá às variáveis do estudo (respostas de um questionário ou medidas laboratoriais, por

exemplo). Verifique a natureza das variáveis: se qualitativa ou quantitativa. Para isso, consulte o dicionário, que está disponível em caso de dúvidas na identificação das variáveis.

Acessando as funções estatísticas

As funções estatísticas do Excel podem ser acessadas no menu "Fórmulas", na aba "Inserir função", onde é possível selecionar a categoria "Estatística", que apresentará todas as funções estatísticas disponíveis no programa. Também é possível acessar através da ferramenta de acesso rápido "Auto soma", onde estão disponíveis ferramentas básicas de soma, média, valor mínimo e máximo. Na aba "Biblioteca de funções", em "Mais funções", pode-se selecionar "Estatística", que dará acesso às funções estatísticas do programa.

Distribuição de frequência

A frequência corresponde simplesmente à contagem do número de ocorrências, podendo ser expressa pela frequência absoluta (n) e frequência relativa ($n\%$). Para exemplificarmos, no "Banco de dados 1", vamos verificar a frequência absoluta e relativa da variável "sexo". A variável "sexo" é uma variável qualitativa dicotômica, ou seja, apresenta duas categorias: sexo masculino e sexo feminino. Será necessário montar uma tabela que indique nas linhas as categorias da variável "sexo" e nas colunas as frequências absoluta e relativa. A tabela resultante será semelhante ao Quadro 1. Na célula ou casela reservada para a frequência absoluta do sexo masculino, utilizaremos a função CONT.SE. Para isso, clique em "Fórmulas", "Estatística" e selecione "CONT.SE". A janela resultante trará uma breve instrução de como utilizar a função.

Quadro 1 - Exemplo para criação de tabela de frequência.

Sexo	Frequência Absoluta (n)	Frequência relativa ($n\%$)
Masculino		
Feminino		
Total		

A função CONT.SE faz a contagem de elementos de acordo com um critério pré-estabelecido. Essa função permite a opção de contar as células vazias e apresenta outras opções, como contar número, valores e vazios. Neste caso, em "Intervalo", é possível selecionar o intervalo de células e contar aquelas que não estão em branco. Selecione todas as células correspondentes à coluna "sexo". Cuide para não selecionar o nome da variável na contagem. Em "Critérios", como queremos selecionar apenas os indivíduos do sexo masculino, inserimos "Masculino". Na própria janela, o programa indica a quantidade de indivíduos do sexo masculino que serão incluídos no intervalo selecionado. Ao clicar em "Ok", a célula selecionada será preenchida com a quantidade (valor absoluto) de indivíduos do sexo masculino no intervalo selecionado. É possível ainda adicionar mais critérios à função, como por exemplo, contar quantos indivíduos do sexo masculino estão no doutorado. Agora, basta repetir o processo para preencher os indivíduos do sexo feminino. Para calcular o total, utilize a função CONTAR.VALORES. Esta função conta o número de células que não estão vazias em um intervalo. Para isso, basta informar o intervalo desejado e o programa retorna com a quantidade de sujeitos daquele intervalo.

Para calcular o número relativo, divida o número absoluto de cada categoria pelo número total da categoria. Essa proporção multiplicada por 100 (%) corresponderá à frequência de cada categoria (sexo masculino e sexo feminino). Para montar a tabela do exercício, é necessário fazer essa etapa para os sexos masculino e feminino. Para obtermos o número relativo total, basta utilizar a função SOMA. Na "Página Inicial" é possível configurar o número de casas após a vírgula, clicando em " \rightarrow .000" ou " \leftarrow .000", conforme necessário. Na interpretação dos resultados, observamos que, dos 115 alunos da disciplina, 82,6% ($n = 95$) são do sexo feminino.

Outra maneira de se obter esses resultados é utilizando a função FREQUÊNCIA, que calcula a frequência absoluta em que os valores ocorrem dentro de um intervalo de valores e, em seguida, retorna uma matriz vertical de números. Ou seja, exerce o mesmo que vimos na função CONT.SE. No entanto, pelo fato de FREQUÊNCIA retornar uma matriz, deve ser inserida como uma fórmula matricial.

Para utilizar essa função, precisamos atribuir códigos para as categorias da variável. Neste exemplo, vamos utilizar os códigos "1" para o sexo masculino e "2" para o sexo feminino. Primeiro, selecione a coluna correspondente à variável "Sexo", esse passo é importante para evitar que a próxima etapa modifique outras variáveis do banco de dados. Vamos utilizar a ferramenta "Localizar e Selecionar", disponível na página inicial. Selecionamos "Substituir" e, no campo "Localizar" escrevemos "Masculino" e, no campo "Substituir por", o código "1" correspondente. Em seguida, clique em "Substituir tudo". O programa sinaliza a quantidade de substituições realizadas. Em seguida, realize a mesma etapa substituindo a categoria "Feminino". Agora, na coluna da variável "Sexo", você verá apenas os códigos 1 e 2 atribuídos aos sexos masculino e feminino, respectivamente.

Ao realizar a substituição, na tabela em que utilizamos a função CONT.SE, os valores calculados serão alterados. Isso ocorre porque o critério apresentado na função já não existe mais. Podemos inserir na função CONT.SE os novos critérios numéricos.

Para utilizar a função FREQUÊNCIA, devemos criar uma tabela similar à anterior mas com os códigos, 1 e 2. Na célula ao lado do código 1 insira a função FREQUÊNCIA. Na janela de argumentos da função, o primeiro argumento que a função FREQUÊNCIA nos pede é a "Matriz_dados". Neste caso, selecione o intervalo de dados (toda a coluna da variável "sexo", exceto a primeira célula que contém o nome da variável). Na "Matriz_bin" será necessário informar o intervalo referente a coluna com os códigos digitados, portanto selecione as células em que foram inseridos os códigos 1 e 2. A resposta somente retornará o equivalente ao código 1. Para atribuir o valor do código 2, seleciona-se o intervalo de células onde queremos exibir os resultados e, na barra de fórmulas, pressione as teclas de atalho "Ctrl + Shift + Enter". Isso irá retornar o valor do código 2, portanto, teremos os valores das duas categorias.

A análise descritiva permite ao pesquisador uma melhor compreensão dos dados por meio de medidas de resumo, identificando tendências, variabilidade e valores extremos. É importante para conhecermos as principais medidas estatísticas.

Para exemplificar, vamos realizar a análise descritiva de medidas de tendência central e de dispersão da variável quantitativa "Idade". Para isso, construa uma tabela para a variável com os dados de média (função MEDIA), mediana (função MED), moda (função MODO.UNICO), valor mínimo (função MÍNIMO), valor máximo (função MÁXIMO), desvio padrão (função DESVPAD.A), variância (função VAR.A), quartis (função QUARTIL.INC) e percentis (função PERCENTIL.INC). Para preencher esta tabela, utilize as respectivas funções. Para as funções QUARTIL.INC e PERCENTIL.INC também é necessário informar os quartis ou percentis, incluindo os valores extremos (mínimo e máximo). No caso dos quartis, como abordado no capítulo anterior, dividimos os dados ordenados em quatro partes iguais; assim, temos: valor mínimo = 0, primeiro quartil (25%) = 1, valor mediano (50%) = 2, terceiro quartil (75%) = 3 e, valor máximo = 4. Para os percentis, dividimos os dados ordenados em k percentil de valores em um intervalo, em que k está no intervalo de 0 a 1. Ou seja, se determinarmos k como sendo 0,33, teremos um percentil de 33%. É muito importante observar os valores mínimo e máximo, comparar média e mediana, para verificar se não há inconsistências nos dados e, assim, identificar erros de digitação no banco de dados. Na interpretação dos dados, para variáveis com distribuição normal, apresenta-se os resultados de média e desvio padrão e, para as variáveis com distribuição não-normal, apresenta-se os dados de mediana e valores mínimo e máximo ou quartis.

Existe uma forma subjetiva para avaliar a distribuição dos dados comparando-se os valores de média e mediana. Valores muito próximos ou iguais de média e mediana indicam uma tendência central, ou seja, distribuição normal. No caso do exercício proposto, os valores de média e mediana são próximos, porém não pode-se afirmar a normalidade sem um teste estatístico adequado, como o de Shapiro-Wilk.

No banco de dados do exercício, podemos observar que há uma célula sem valor para a variável "Idade", ou seja, com valor omisso. Esse valor também pode ser chamado de "*missing value*" ou perda de dado. Quando há valores faltantes é importante mencionar o número de dados analisados na interpretação dos dados nas tabelas e/ou nos gráficos. Isso pode ser sinalizado ao lado do nome da variável na tabela com um símbolo (*) ou

indicando o "n", ou no rótulo da tabela. O Excel não sinaliza valores faltantes utilizando as funções descritas.

Há ainda outra opção para a análise de dados disponível no menu "Dados", na "Análise de dados". Esta ferramenta nem sempre está disponível no Excel. Para disponibilizar a ferramenta, basta clicar no menu "Arquivo", em "Opções", selecione "Suplementos", na opção "Gerenciar" marque a opção "Suplementos do Excel" e clique em "Ir". Na janela "Suplementos", selecione "Ferramentas de Análise" e clique "Ok".

A opção "Análise de dados" resulta em uma janela onde seleciona-se a opção "Estatística Descritiva" e, em seguida, clique em "Ok". Na janela "Estatística descritiva", no campo "Intervalo de dados", é necessário selecionar os dados referentes à variável de interesse, no caso, "Idade". Atenção! Ao utilizar essa função, devemos selecionar inclusive a célula com o nome da variável e marcar a opção "Rótulo na primeira linha". Pode-se optar por exportar os dados para uma nova planilha ou para uma célula selecionada ("Intervalo de saída"). Selecione as opções de "Resumo Estatístico" e "Nível de confiabilidade para média". O Excel assume a distribuição normal dos dados e só calcula nível de confiabilidade para média, sendo essa uma das limitações do uso do programa para execução de análises estatísticas. Ao clicar em "Ok", teremos uma tabela com a análise descritiva da variável "Idade". No item "Contagem", o Excel vai retornar apenas o número de células com valores computados na análise, facilitando a identificação de dados faltantes.

Utilizando-se essa opção, o Excel apresenta também dados de curtose e assimetria, que referem-se à distribuição da variável. Se os valores de assimetria forem próximos a zero, a distribuição é aproximadamente simétrica ou normal. A curtose refere-se à nitidez do pico da curva central.

Gráficos

O Excel oferece suporte para vários tipos de gráficos com a finalidade de ajudar a exibir os resultados. A representação gráfica deve obedecer requisitos fundamentais para ser realmente útil:

- Simplicidade: o gráfico deve ser destituído de detalhes, bem como de traços desnecessários;
- Clareza: deve possibilitar uma correta interpretação dos valores representativos da amostra em estudo;
- Veracidade: deve expressar a verdade sobre a amostra em estudo.

No menu "Inserir", na opção "Gráficos", é possível criar diferentes modelos de gráfico. Como exercício, vamos criar um gráfico com os resultados de frequência da variável "Sexo", que analisamos anteriormente. Para isso, selecione as células com as categorias "Masculino" e "Feminino", bem como as células com os valores de frequência absoluta obtidos na tabela 1. Clique na opção de "Gráfico de Pizza", e selecione a primeira opção, gráfico 2D. Esse processo já resultará no gráfico, onde é possível editar detalhes como título do gráfico, alterar as cores, inserir o rótulo dos dados (ao clicar com o botão direito sobre a área do gráfico, clique em "Adicionar Rótulo de Dados". Clicando com o botão direito sobre o valor absoluto do rótulo inserido, nas "Opções de rótulo", é possível inserir o valor relativo). O programa também sugere modelos prontos de gráficos. Naturalmente, esses dados podem ser apresentados através de outros tipos de gráficos, como em barras, colunas, como visto no capítulo anterior.

O Excel ainda tem disponível uma infinidade de outras opções, como correlação, teste t, ANOVA, regressão. Lembrando que o Excel não é um software de estatística e sim de planilha, portanto pode haver discrepância ou variação nos resultados, diminuindo a confiabilidade dos dados obtidos. Comparativamente, o Excel fornece uma redução na redundância de dados, enquanto o SPSS é mais rápido e oferece um desempenho confiável. O Excel não fornece a construção de função, ou seja, precisa de muito trabalho para executar uma tarefa específica. O SPSS possui ferramentas de manipulação de dados integradas para gravação e transformação de variáveis.

Viés em Pesquisa

Para falar de viés, precisamos falar da validade de um estudo, que pode ser interna ou externa. A validade interna reflete as conclusões de uma investigação em relação à amostra estudada. Já a validade externa é relativa à extrapolação dessas conclusões para a população de onde a amostra foi retirada ou, até mesmo, para outras populações.

A validade externa gira em torno da questão da capacidade de generalização dos resultados da pesquisa se manterem verdadeiros em outras situações ou para outras populações. Ou seja, as conclusões do estudo podem ser generalizadas para a população de interesse? A amostra representa essa população?

A validade interna reflete a qualidade do delineamento, da condução e da análise da pesquisa. Pode ser definida como a ausência de erro sistemático nos achados do estudo. Devemos nos perguntar até que ponto os resultados do estudo estão distorcidos em decorrência de erros metodológicos na concepção do estudo?

Tipos de erro

Erro Aleatório: trata-se de uma divergência entre o valor da amostra, que é medida pelo estudo, e o valor verdadeiro, encontrado na população. Este tipo de erro ocorre exclusivamente em decorrência do acaso, diminuindo a precisão (confiabilidade) do estudo e tendendo a diminuir a medida de associação. O erro aleatório pode ser consequência da variação biológica natural, de erro de amostragem e/ou erros de medida. O erro de amostragem decorre geralmente da falta de representatividade da amostra, que não contempla toda a variabilidade da população. A melhor forma de reduzir o erro de amostragem é aumentar o tamanho amostral. De fato, não há como garantir a eliminação do erro aleatório, pois os estudos geralmente são conduzidos com uma pequena amostra da população. Na estatística, uma forma de se evitar esse erro é computar a significância estatística e os intervalos de confiança.

Erro Sistemático: trata-se da diferença sistemática entre a medida real de efeito e a medida obtida no estudo. Decorre de problemas no delineamento, planejamento, condução ou análise

do estudo, não podendo ser corrigido com aumento do tamanho de amostra, nem durante a análise dos dados. Pode ser evitado na elaboração do estudo, corrigindo o processo de seleção da amostra e coleta das informações. Do ponto de vista estatístico, viés é quando o valor médio da medida de associação obtido a partir de um número infinito de estudos não corresponde ao valor verdadeiro de associação. Em outras palavras, viés refere-se à diferença entre a média encontrada de uma estimativa computada após múltiplas análises randômicas e o valor verdadeiro do parâmetro a ser estimado. Do ponto de vista epidemiológico, é o desvio dos resultados ou inferências em relação à verdade, ou o processo que leva a esse desvio. Se refere a qualquer tendência na coleta, análise, interpretação, publicação ou revisão de dados que possa conduzir a conclusões sistematicamente diferentes da verdade.

Viés de seleção

Ocorre quando há uma diferença sistemática no momento da seleção ou alocação dos participantes da pesquisa, resultando em diferenças entre os indivíduos alocados e não alocados para o estudo. Por exemplo, quando os grupos de comparação não provêm da mesma população-base do estudo, impactando na medida de associação entre exposição e desfecho, ou seja, são grupos de sujeitos não comparáveis.

O viés de seleção pode ocorrer em todos os tipos de delineamentos epidemiológicos, mas é uma grande preocupação nos estudos do tipo caso-controle, quando a seleção de casos e de controles é baseada em critérios diferentes e estes estão relacionados à exposição. De uma forma geral, este viés é minimizado em estudos prospectivos, pois a seleção dos indivíduos expostos e não-expostos é feita antes do registro da resposta.

A seletividade da amostra pode render resultados distorcidos porque a natureza das relações entre variáveis nos casos incluídos na amostra não é necessariamente igual às relações que observaríamos em uma amostra aleatória.

Pode ser causada pela forma de randomização (em ensaios clínicos randomizados), de seleção ou admissão, pela

perda de seguimento ou não-resposta dos participantes. Nestes dois últimos casos, a preocupação é a probabilidade de que as perdas possam ter ocorrido de forma seletiva e, portanto, seja relacionada à exposição, à resposta ou ambos. Pode-se controlar esse viés através de planejamento e amostragem aleatória (cada indivíduo tem probabilidade conhecida e diferente de zero de fazer parte da amostra). Pode-se ter diferentes tipos de viés de seleção, que estão resumidos no Quadro 2.

Nos estudos de coorte dificilmente ocorre viés de seleção, pois os grupos são determinados de acordo com o fator de exposição e não há desfecho algum no momento da seleção. No entanto, é comum ocorrerem vieses do tipo de perda de seguimento. Alguns indivíduos expostos a determinado fator de risco em um estudo no qual estes seriam acompanhados após a alta hospitalar podem evoluir a óbito, não comparecendo ao seguimento. Essa perda está relacionada tanto com exposição quanto com desfecho.

Viés de informação (ou classificação)

O viés de informação em estudos de coorte é mais frequente no momento de aferir o desfecho. Se os investigadores conhecem a exposição, podem ficar sugestionados ao aferir o desfecho. Erros na mensuração da exposição ou do desfecho que ocorrem no momento da coleta de informações podem distorcer a classificação tanto da exposição, quanto do desfecho nos sujeitos da pesquisa, em outras palavras, os métodos utilizados não medem corretamente o desfecho ou exposição. Pode ocorrer através de uso de meios inadequados para obter informações.

Deriva de variação individual, variação entre observadores, instrumentos de medição ou erro de aferição ou calibração. Pode ser controlado através da padronização do método, treinamento dos avaliadores e utilização de instrumento de pesquisa adequado. Se sua ocorrência for igual entre os grupos é chamado de viés não-diferencial e pode ser responsável pela discrepância entre resultados de diferentes estudos epidemiológicos. Existem vários tipos de viés da informação, resumidos no Quadro 3.

Quadro 2 - Tipos de viés de seleção (adaptado de Bonita e cols., 2010).

Viés de Seleção	Definição
Alocação	O investigador escolhe um método não-randomizado para distribuir sujeitos nos grupos de estudo.
Detecção	Resultado da falha em detectar um caso da doença, um possível fator causal ou um efeito de interesse.
Participação	Quando fatores afetam a participação dos sujeitos da pesquisa, como exclusões inadvertidas ou recusa em participar.
Amostragem	A amostra não representa o espectro de características da população alvo. Alguns indivíduos têm maior chance de serem selecionados. Por exemplo, um estudo sobre desnutrição em áreas rurais, onde os investigadores evitam as moradias isoladas e examinam apenas os moradores de regiões próximas à estrada.
Trabalhador Saudável/ Auto-seleção/ Afiliação	Indivíduos com determinadas características (ligada à doença ou exposição) têm maior probabilidade de entrar no estudo. Pessoas empregadas são relativamente mais saudáveis que as desempregadas, aposentadas ou incapacitadas. Provavelmente menos pessoas empregadas se candidatam a participar de um estudo.
Não-respondentes	Pode ser encontrado em estudos transversais, onde os indivíduos não localizados apresentam tipicamente maiores prevalências da maioria das doenças do que os que são entrevistados. Em estudos de caso-controle, por exemplo, pode ocorrer se a taxa de não-resposta for diferente entre casos e controles e se a não resposta estiver associada com a exposição.
Perdas de acompanhamento/ seguimento	Pode ocorrer em estudos de coorte ou de intervenção. Por exemplo, os indivíduos que deixaram de ser acompanhados podem ser diferentes daqueles efetivamente estudados.
Hospitalização/ Admissão/ Diagnóstico (viés de Berkson)	Pacientes com determinadas características podem ser mais frequentemente internados que outros sem tal característica com a mesma doença.
Prevalência/ Incidência	São excluídos os sujeitos curados e falecidos (prevalência) ou os sujeitos com coorte de início (incidência). Também chamado de viés de sobrevivência seletiva ou de Neymann.
Sobrevida Seletiva	Após o diagnóstico, o sujeito muda os hábitos associados à doença. Qualquer característica associada com a duração da doença afetará os resultados em um estudo de casos prevalentes.
Viés de sobrevivência	Refere-se às situações nas quais você procura pelos erros analisando os acertos. O termo surgiu na Segunda Guerra Mundial, quando o matemático Wald resolveu um problema de blindagem na fuselagem dos aviões. Inicialmente, os locais de blindagem eram decididos pela frequência dos locais de marcas de tiros dos aviões que retornavam à base. Wald entendeu que, se os aviões com marcas de tiros naqueles locais retornavam, era porque os outros locais não atingidos que precisariam do reforço.

Quadro 3 - Tipos de viés de Informação (adaptado de Bonita e cols., 2010).

Viés de Informação	Definição
Recordação/Memória	Quando o estudo é conduzido de forma retrospectiva, os casos podem lembrar-se melhor de sua história de exposição do que os indivíduos sem a carga da doença. Por exemplo, pacientes com refluxo gástrico podem se lembrar melhor de certas características da dieta no passado do que indivíduos saudáveis.
Registro	Indivíduos com doenças mais graves tendem a ter registros mais completos sobre exposições, o que pode resultar em medidas errôneas de associação.
Aferição/ Mensuração/ Instrumento	Ocorre quando os métodos de aferição (instrumentos) empregados diferem entre os grupos, favorecendo resultados inadequados a um subgrupo de pacientes. Pode ocorrer por diferenças no método, no preparo dos observadores, medidas basais ou dados de seguimento, falta de validade dos instrumentos de coleta de dados, subjetividade de escalas de medidas ou uso de técnicas inadequadas.
Publicação	Tendência de que resultados publicados sejam sistematicamente diferentes da realidade. A não publicação dos resultados pode ser por decisão do autor, do financiador do estudo ou de editores por não submeterem ou aceitarem trabalhos com resultados negativos ou desfavoráveis.
Deteção/ Diagnóstico	Características ligadas à exposição podem aumentar a chance de que uma doença seja detectada. Por exemplo, mulheres em uso de anticoncepcionais podem fazer exame citopatológico com mais frequência, portanto têm maior risco de deteção de alguma alteração citopatológica.
Entrevistador/ Observador	Quando o entrevistador coleta de modo diferente as informações de doentes e saudáveis ou expostos e não-expostos. Ou ainda, quando há direcionamento da resposta do participante. Pode ser contornado com o cegamento dos investigadores.
Duração	Quando os casos mais moderados e mais indolentes de uma doença são detectados desproporcionalmente em programas populacionais de rastreamento.
Exame tardio	Ocorre quando casos moderados de uma doença que progride lentamente são preferencialmente detectados em um inquérito.
Espectro de doença	Quando somente um espectro limitado da doença é incluído.
Tempo Ganho	Ocorre quando o rastreamento detecta uma doença mais precocemente na sua história natural do que teria acontecido de outra maneira.
Extensão	Ocorre quando o espectro completo de um tumor em particular é composto de cânceres que variam de agressivos a indolentes.

Viés de confirmação

É a tendência de que, uma vez adotada uma hipótese pelo pesquisador, este busque exemplos que a confirme. É a preferência que temos por confirmar nossas hipóteses a ter de refutá-las. Pode levar à pesquisa tendenciosa, polarização de opiniões e interpretação tendenciosa.

Fator de confundimento ou de confusão

Acontece quando a associação observada ocorre devido a um terceiro fator (variável de confusão, podendo ser um fator de risco ou de proteção) que se relaciona com a variável de exposição e com o desfecho, mas não faz parte da cadeia causal da doença (Figura 1).

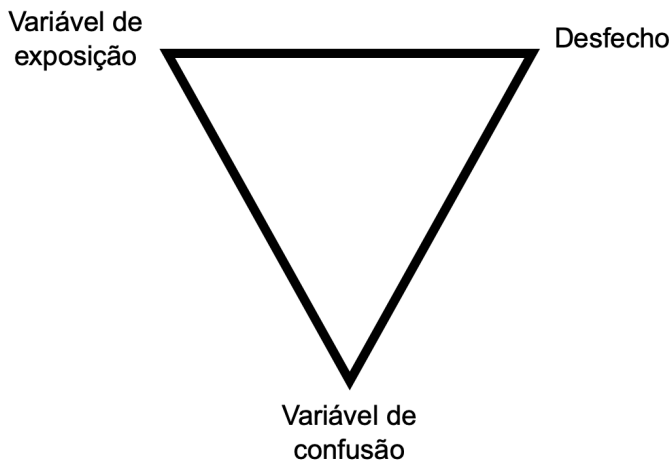


Figura 1. Fator de confusão se relaciona com a variável de exposição e com o desfecho, mas não faz parte da cadeia causal da doença.

Pode também ser definido como a existência de outra exposição que esteja associada com o desfecho. Ocorre devido a uma diferença não aleatória na distribuição dos fatores de risco entre

os grupos, ou seja, a variável de confundimento está distribuída desigualmente entre os grupos comparados. Isso pode alterar a estimativa de risco (subestimar ou superestimar), gerar falso-negativo e/ou falso-positivo e, inclusive, inverter a direção de associação.

O fator de confusão não é considerado um viés, pois não resulta de erro sistemático relacionado ao delineamento de pesquisa. Pode ser controlado tanto no delineamento quanto na análise estatística, através de identificação prévia, presença de grupo controle, randomização, pareamento ou emparelhamento, restrição de categorias, critérios de inclusão e exclusão, estratificação e análise multivariável.

Referências

Bonita R, Beaglehole R, Kjellström T. Epidemiologia básica. 2. ed. São Paulo: Santos, 2010. 213p.

Guerrero, H. Excel Data Analysis: Modeling and Simulation. 2. ed. Springer Nature Switzerland AG. 358 p.

Guia básico/intermediário. PET Civil UFRGS, 2018, Disponível em: <https://www.ufrgs.br/petcivil/wp-content/uploads/2018/06/Apostila-Excel-2018-1-site.pdf>.

Medronho RA, Bloch KV, Luiz RR, Werneck GL. Epidemiologia. 2. ed. São Paulo: Atheneu, 2008. 676p.

Schoenbach VJ, Rosamond WD. Understanding the Fundamentals of Epidemiology an evolving text. Chapel Hill, North Carolina. 2000. 584 p.

Exercícios sugeridos

1. No "Banco de dados 1.xls", qual a frequência do nível de pós-graduação entre os alunos da disciplina? Interprete os resultados e construa o gráfico (pizza, barras ou retângulo).
2. No "Banco de dados 2.xls", que traz dados antropométricos e de pressão arterial de uma amostra de sujeitos com obesidade, realize a análise descritiva, construa a tabela e interprete os resultados de idade, Índice de Massa Corporal (IMC), circunferência da cintura e pressões arteriais sistólica e diastólica. Suponha distribuição não normal dos dados para interpretação.
3. No "Banco de dados 2.xls", qual a frequência dos graus de obesidade? E de hipertensos e não hipertensos? Interprete os resultados e construa os gráficos (pizza, barras ou retângulos). Suponha distribuição não normal dos dados para interpretação.

3 Primeiros passos no SPSS

Hellen Meiry Grosskopf Werka
Johanna Ovalle Diaz
Edison Capp
Otto Henrique Nienov

Conhecendo o SPSS

O *software* IBM® SPSS (*Statistical Package for Social Sciences*) é um pacote estatístico amplamente utilizado no ambiente acadêmico. É uma ferramenta para análise de dados, que utiliza técnicas estatísticas básicas e avançadas. Além disso, permite executar estatísticas descritivas, correlação e regressão, categorizar variáveis, manipular dados e criar gráficos e tabelas.

É uma ferramenta paga, adquirida no site da IBM (<https://www.ibm.com>), mas também está disponível para a comunidade acadêmica da UFRGS (<https://www.ufrgs.br/documentacaoti/spss-tutorial-de-instalacao/>), sendo necessária a instalação e ativação da VPN (<https://www.ufrgs.br/documentacaoti/vpn/>). Através da VPN (*Virtual Private Network*) é criado um canal que permite o tráfego de dados de forma segura e o acesso a rede interna da UFRGS, possibilitando o acesso em casa, por exemplo.

Apesar de ser um *software* estatístico de fácil manuseio, usar o SPSS sem conhecimento algum de estatística pode ser arriscado. O SPSS apresenta um complexo conjunto de opções para cada teste, muitos dos quais necessitam de atenção na hora de serem utilizados e interpretados. Pensando nisso, este capítulo busca descrever aspectos básicos para aqueles que não têm experiência com o SPSS.

Construindo um banco de dados no SPSS

O primeiro passo é conhecer o SPSS. Para isso, vamos construir um banco de dados. Abra o SPSS, clicando no atalho da versão instalada na área de trabalho do seu computador. Na caixa

de diálogo é possível abrir arquivos existentes que, por padrão possuem a extensão “.sav”, ou podemos inserir novos dados criando um banco de dados.

Selecione a opção “Inserir Dados” e, em seguida, clique em “Ok”. Quando você executa o SPSS desta forma, ele irá apresentar uma planilha em branco. Nessa janela de edição de dados, onde inserimos os dados e executamos as funções estatísticas, há duas planilhas: uma chamada “Visualização da variável” e outra “Visualização dos dados”. A primeira corresponde a um dicionário das variáveis, onde iremos criar as variáveis do banco de dados e inserir as informações que a definem, como sua descrição, identificação dos rótulos e sua natureza (quantitativa ou qualitativa); enquanto a segunda corresponde ao banco de dados. Lembre-se que o banco de dados é uma estrutura organizada que permite a extração de informações. As linhas correspondem aos dados de uma unidade experimental (que pode ser uma pessoa, animal ou objeto) e as colunas correspondem as variáveis criadas na “Visualização da variável”. Caso você esteja utilizando o SPSS, é possível abrir um novo banco de dados clicando em “Arquivo”, “Novo” e, em seguida, em “Dados”.

Desta forma, o primeiro passo para inserir os dados no SPSS é criar a variável utilizando a planilha “Visualização da variável” do editor de dados e então entrar com os dados na planilha “Visualização dos dados”. Na planilha de “Visualização dos dados”, clicando duas vezes sobre a coluna “var”, você será direcionado para a primeira linha da planilha “Visualização da variável”. Isso significa que, a primeira coluna da “Visualização dos dados” corresponde à primeira linha na “Visualização da variável”. Portanto, está será nossa primeira variável. É recomendado que a primeira coluna da matriz corresponda a uma variável de identificação. Pode ser o número do questionário, número do caso, ou ainda código do registro. Isso facilita a localização de informações no caso de serem identificados equívocos de digitação.

Nesta planilha de “Visualização da variável”, vamos detalhar como a variável será criada. Na primeira coluna (“Nome”), digite-se o nome da variável desejada (por exemplo, “ID”). Esse nome será apresentado no início da coluna na planilha “Visualização dos

dados” quando você for inserir os dados e serve para identificar as variáveis. Existem algumas regras gerais para o nome das variáveis, como utilizar oito dígitos no máximo e não utilizar espaço em branco, nem os símbolos “-”, “. ” e “/”. Também, o SPSS não permite que o nome de uma variável comece com um número. Por isso, caso você tenha variáveis repetidas, insira o número ao final do nome da variável, por exemplo. Se for violada alguma dessas regras, o SPSS informará que o nome da variável é inválido.

Seguindo na descrição da variável, na coluna “Tipo”, é definido o tipo de dado. Você pode ter diferentes tipos de dados. Clique na célula “Tipo” e, na janela “Tipo de variável”, você pode selecionar o tipo de dados que constará naquela variável. O tipo mais comum (padrão) é o “Numérico”. Caso a variável seja numérica, como por exemplo o número de prontuário, selecione a opção “Numérico”. Importante! As variáveis qualitativas, em que as categorias foram descritas em códigos, também são do tipo “Numérico”. Caso a variável seja alfa-numérica (texto) selecione a opção “Sequência”. Uma variável de texto é simplesmente uma linha de caracteres que representa uma informação que não se deseja analisar, como o nome do participante. Há outras opções específicas, como “Data”. Esta deve ser classificada, na coluna “Medir”, como quantitativa (“Escala”) quando se deseja calcular diferença entre datas para obter-se outras variáveis, como idade, tempo de acompanhamento, etc. Preferencialmente use sempre a modalidade “Numérico” para variáveis categóricas (qualitativas), como sexo, estado civil, etc., atribuindo um código para as categorias. O SPSS não reconhece caracteres (palavras), sendo necessário codificar as categorias da variável, que serão identificadas na coluna “Valores”. Você deve atribuir a cada categoria da variável um número e, então, informar ao SPSS que número representa cada categoria. Selecione o tipo de variável desejado e clique em “Ok”.

Na coluna “Largura”, é possível definir o tamanho do valor da variável. Por exemplo, se o número do caso tiver quatro dígitos, definimos como a largura sendo igual a quatro. Na coluna “Decimais”, é possível definir o número de casas decimais após a vírgula que iremos atribuir à variável. A largura não pode ser menor do que o decimal. Na coluna “Colunas”, definimos o tamanho do nome da variável.

Como o nome de uma variável é restrito a oito caracteres, o SPSS fornece uma maneira para que possamos identificar as variáveis de uma forma mais detalhada. Na coluna "Rótulo", é possível digitar um nome mais completo (por extenso) para a variável, identificando a variável e especificando a unidade de medida. Por exemplo, para a variável "ID" é possível digitar a expressão "Identificação do caso". Em "Valores", atribuímos os valores da variável de interesse, ou seja, o valor das categorias (códigos). Clique na célula "Valor" e, em seguida, na janela "Rótulos de valor", insira o valor da categoria ("Valor") e o nome da categoria ("Rótulo"). Clique em "Adicionar" para inserir o valor do rótulo e, em seguida, em "Ok". É conveniente registrar no banco de dados os nomes das categorias de variáveis categóricas. Por exemplo, para a variável "SEXO", poderíamos atribuir o valor "1" para o rótulo "masculino" e "2" para "feminino". Esses códigos informam ao SPSS que todos os casos da variável atribuídos ao código "1" pertencem ao sexo masculino, da mesma forma que os casos atribuídos ao código "2" pertencem ao sexo feminino. Importante! A utilização de códigos numéricos para as categorias de uma variável qualitativa não a tornam quantitativa. Também é possível alterar as informações de valor e rótulo em "Alterar" ou removê-los em "Remover".

Existem situações em que não temos todos os dados coletados (*missing value*) para os casos ou estes se apresentam como não-resposta ou, ainda, respostas que não desejamos considerar para o tratamento estatístico, como por exemplo, respostas não corretas, que não se aplicam, etc. A omissão de dados pode ocorrer por vários motivos: em questionários longos onde os participantes podem acidentalmente esquecer de responder alguma questão; em fichas de coleta onde o investigador pode esquecer de registrar algum dado; em estudos que utilizam registros de prontuários onde os dados podem não estar apresentados ou; em um experimento que falha para determinado caso e não se obtém o valor do dado. Para isso, utilizamos a coluna "Ausente". Na janela "Valores ausentes", pode-se registrar o código de não-resposta na opção "Valores ausentes discretos", utilizando preferencialmente 9, 99, 999, etc. Clique em "Ok" para adicionar o valor. Entretanto, a melhor opção para não-resposta é deixar o espaço em branco no banco de dados,

selecionando a opção “Sem valores ausentes”, pois variáveis podem assumir o código de não-resposta como um valor válido.

Em “Alinhar”, pode-se selecionar o alinhamento dos dados na planilha à direita, à esquerda ou no centro. Em “Medir”, deve-se selecionar como serão analisadas estatisticamente as variáveis: a) “Escala” refere-se a variáveis quantitativas, sejam discretas ou contínuas; b) “Ordinal” refere-se a variáveis qualitativas ordinais e; c) “Nominal” a variáveis qualitativas nominais. Por fim, na coluna “Função”, é definida a funcionalidade da variável. Como padrão, as variáveis são classificadas como “Entrada”.

Além da janela de edição de dados, o SPSS fornece uma janela de visualização de resultados, onde irão ser apresentados os resultados das análises, e a janela de edição de sintaxe, que trará os comandos das funções estatísticas.

Quando você for construir um novo conjunto de dados, você deve entrar com os dados de uma forma lógica. Para exemplificar, vamos construir um banco de dados fictício, com características demográficas de uma amostra aleatória. Crie as variáveis na planilha de “Visualização da variável” e, uma vez criada a variável, insira os dados na planilha de “Visualização dos dados” (Quadro 1). Neste banco de dados, a variável “ID” corresponde à identificação dos casos, “NOME” ao nome dos sujeitos, “SEXO” ao sexo (onde, 1 = sexo masculino e 2 = sexo feminino), “DN” à data de nascimento, “IDADE” à idade (em anos) e “ESTATURA” à estatura (em metros). Essas informações devem ser inseridas na planilha de “Visualização da variável”. Lembre-se que cada linha na planilha de “Visualização da variável” representa uma variável e, para particularizar uma variável, basta configurar as informações nas colunas.

À medida que as informações são registradas no banco de dados, é importante salvá-las. Para isso, clique em “Arquivo” e, em seguida, em “Salvar como...”. Na janela “Salvar dados como”, se escolhe o caminho desejado para salvar o arquivo e cria-se um nome para o banco de dados, que terá automaticamente a terminação “.sav”. Entretanto, você poderá salvar dados em outros formatos, como o do Microsoft Office Excel. Para isso, clique em “Salvar como tipo” e localize o formato “.xls”.

Quadro 1. Exemplo de banco de dados.

ID	NOME	SEXO*	DN	IDADE	ESTATURA
1	A	1	12/11/1980	39	1,65
2	B	2	06/01/1975	45	1,92
3	C	1	13/07/1988	31	1,54
4	D	2	01/04/2001	19	1,74
5	E	2	13/10/1976	43	1,68
6	F	1	12/09/1978	41	1,72

*1 = sexo masculino; 2 = sexo feminino.

Acessando um banco de dados existente

Se você irá trabalhar com um banco de dados existente, é importante que você saiba como abrir esse arquivo de dados no SPSS. O procedimento é bastante simples. Para isso, faça o download do “Banco de dados 1.sav” em seu computador (disponível em <https://bit.ly/bancosdedados>). Este banco de dados nos fornece dados de alunos de uma disciplina. Ao abrir o SPSS, na caixa de diálogo, é possível abrir arquivos existentes que, por padrão, possuem a extensão “.sav”, na opção “Abrir uma origem de dados existente”. Localize o arquivo “Banco de dados 1.sav” e clique em “Abrir” para abrir o banco de dados. Também pode-se “Colar” o banco de dados, gerando um arquivo de sintaxe, que corresponde ao comando da ação. Para executar o comando, selecione-o no arquivo de sintaxe e clique em “Executar seleção”. Caso o SPSS esteja aberto, clique em “Arquivo”, “Abrir” e, em seguida, clique em “Dados” para abrir o arquivo “Banco de dados 1.sav”. No mesmo caminho é possível abrir arquivos de sintaxe e de saída.

Importante! Ao acessar um banco de dados existente, verifique se as variáveis estão corretamente classificadas antes de iniciar qualquer análise estatística. Observe se as variáveis estão classificadas corretamente e se as variáveis qualitativas possuem os valores dos rótulos atribuídos. Variáveis classificadas incorretamente ou sem seus rótulos identificados, podem impossibilitar as análises e/ou dificultar a sua interpretação.

O que é um arquivo de sintaxe?

Um arquivo de sintaxe é simplesmente um arquivo de texto que contém uma linguagem de comandos para executar análises estatísticas e manipulação de dados no SPSS, ou seja, é uma criptografia do comando. Você pode abrir uma janela de sintaxe e digitar comandos diretamente, mas é mais fácil deixar que as caixas de diálogo façam uma parte ou todo o trabalho para você. A sintaxe é muito útil quando você ainda está coletando o seu banco de dados ou realizando análises semelhantes muitas vezes. Nessas situações, ao criar e salvar o arquivo de sintaxe, você poderá executá-la futuramente com o banco de dados atualizado, não necessitando executar comando por comando. É interessante, por exemplo, informar a função de cada comando, caso você venha executar diferentes análises. Desta forma, você também pode editar o arquivo de sintaxe para descrever os comandos. Basta clicar em uma das linhas da sintaxe, inserir um asterisco (*), escrever a sentença e inserir um asterisco novamente para fechar a sentença. Uma vez que os comandos tenham sido gerados ou digitados, é necessário executá-los. Para isso, selecione o comando e clique em "Executar seleção".

Assim como o banco de dados, é possível abrir e salvar um arquivo de sintaxe. Para abrir um arquivo de sintaxe salvo, clique em "Arquivo", "Abrir" e, em seguida, "Sintaxe...". Uma caixa de diálogo padrão para abrir arquivos é exibida. Selecione um arquivo de sintaxe. Se nenhum arquivo de sintaxe for exibido, certifique-se de que a opção "Sintaxe (*.sps)" esteja selecionada como o "Tipo de arquivo" que você deseja visualizar. Selecione o arquivo e clique em "Abrir". Para salvar o arquivo de sintaxe, clique em "Arquivo", "Salvar como" e escolha o destino para salvá-lo, dando origem a um arquivo do tipo ".sps".

Para exemplificar, no "Banco de dados 1.sav", vamos analisar a frequência da variável "SEXO". No menu "Analisar", em "Estatísticas descritivas", clique em "Frequências...". Selecione a variável "SEXO" e mova-a para a lista "Variável(is)". Clique em "Gráficos". Na caixa de diálogo "Gráficos", selecione "Gráficos de pizza". No grupo "Valores do Gráfico", selecione "Frequências". Clique em "Continuar" e, em seguida, em "Colar" para copiar a

sintaxe criada como resultado das seleções da caixa de diálogo para o arquivo de Sintaxe. Na janela de edição de sintaxe estará o comando para executar a análise estatística solicitada. Para executar a sintaxe, selecione o comando e clique em "Executar seleção". Abrirá uma nova janela contendo o arquivo de saída, ou seja, o resultado da distribuição da frequência de sexo entre os alunos da disciplina.

Os resultados no visualizador de saídas

O objetivo ao executar o arquivo de sintaxe é gerar um arquivo de saída, que trará os gráficos, tabelas e resultados da análise estatística. Portanto, os resultados da análise estatística aparecem em uma nova janela de resultados, que poderá ser salva, dando origem a um arquivo do tipo ".spv". Também podemos editar o arquivo de saída. Para isso, clique em "Inserir" e, em seguida, em "Novo texto".

É possível também editar os gráficos no arquivo de saída. Para isso, basta dar um duplo clique naquele que se deseja editar. No lado esquerdo da janela de saída existe um diagrama de árvore ilustrando a estrutura da saída. Esse diagrama é útil quando são realizadas inúmeras análises, pois nos fornece uma forma fácil e rápida de acessar partes específicas da saída. A estrutura de árvore é praticamente autoexplicativa, listando cada procedimento executado como um cabeçalho principal.

Executando uma análise

Da mesma forma que precisamos conhecer o banco de dados no Microsoft Office Excel, precisamos também conhecer no SPSS. Tome algum tempo para examinar o banco de dados e explorar os dados. Uma boa forma de checar o banco de dados é realizar algumas estatísticas descritivas e construir alguns gráficos.

Estatísticas descritivas no SPSS

Uma boa forma de explorar os dados é realizar uma análise descritiva. Existem duas formas de realizá-la: 1) utilizando o "Banco de dados 1.sav", clique em "Analisar", "Estatísticas descritivas" e, em seguida, em "Frequências...". Localize no painel da esquerda a variável de interesse (por exemplo, "IDADE") e clique na seta para mover para a lista de "Variável(is)". Qualquer análise escolhida será realizada em todas as variáveis listadas na lista de "Variável(is)". Clique em "Estatísticas...". Como a idade é uma variável quantitativa, iremos assinalar as medidas de tendência central ("Média", "Mediana" e "Moda"), de dispersão ("Desvio padrão", "Variância", valores "Mínimo" e "Máximo") de variabilidade ("Quartis") e de forma ("Assimetria" e "Curtose"). Clique em "Continuar" e, após, em "OK" para gerar o arquivo de saída, ou em "Colar" para gerar o arquivo de sintaxe. Os resultados da análise estatística serão apresentados na janela de resultados (arquivo de saída); 2) clique em "Analisar", "Estatísticas descritivas" e, em seguida, em "Descritivos...". Localize no painel da esquerda a variável de interesse (por exemplo, "IDADE") e clique na seta para mover para a lista de "Variável(is)". Clique em "Opções...". Assinale as mesmas medidas de análise ("Média", "Desvio padrão", "Variância", "Mínimo", "Máximo", "Curtose" e "Assimetria"). Clique em "Continuar" e, após, em "OK" para gerar o arquivo de saída, ou em "Colar" para gerar o arquivo de sintaxe. Os resultados da análise estatística serão apresentados na janela de resultados (arquivo de saída).

Agora, com os resultados obtidos, é preciso observá-los e interpretá-los. Primeiro, observe o número de casos válidos e omissos (*missing value*). Em seguida, observe os valores máximo e mínimo. Verifique se há alguma inconsistência nos valores encontrados. Essa inspeção visual é importante para localizarmos possíveis erros de digitação no banco de dados. Compare também os valores de média, mediana e moda obtidos. Note que o SPSS nos indica quando há mais de uma moda na amostra, enquanto que o Microsoft Office Excel nos fornece apenas um valor de moda. Para verificar se uma distribuição é normal, podemos olhar para os valores da assimetria e da curtose (os valores da assimetria e da curtose deverão ser zero em uma distribuição normal), comparar os

valores de média e mediana ou visualizar graficamente a distribuição dos dados através de um histograma. Mas, como mencionado no Capítulo 1 e, abordado no Capítulo 7, existem testes que avaliam de forma mais confiável a normalidade de distribuição dos dados.

Abaixo da tabela de análise estatística, também é apresentada uma tabela com as frequências das idades. Para isso, na opção “Analisar”, “Estatísticas descritivas”, “Frequências...”, deve-se selecionar “Exibir tabelas de frequências”. Por padrão, o SPSS produz uma distribuição de frequências de todos os escores no formato de tabela.

Após observar minuciosamente os dados, é hora de interpretá-los. Se considerarmos uma distribuição normal dos dados, utilizaremos os valores de média e desvio padrão; assim, a interpretação dos dados seria que os alunos da disciplina apresentam idade média de 30 anos, ou então, 30 ± 6 anos de idade. Se assumirmos uma distribuição não normal dos dados, utilizaremos os valores de mediana e quartis ou de mediana e valores mínimo e máximo; assim, a interpretação dos dados seria que os alunos da disciplina apresentam mediana de idade de 29 (25-33) anos ou de 29 (21-55) anos.

Representação gráfica das análises descritivas

Podemos também apresentar os resultados das análises descritivas através de gráficos. O SPSS tem recursos para construir diferentes tipos de gráficos. Um destes gráficos, muito utilizado é o histograma. Um histograma apresenta barras ao longo de uma escala de mesmo intervalo e a altura de cada barra é a contagem de valores de uma variável quantitativa que corresponde ao intervalo.

Existem duas formas de obtermos o histograma: 1) em “Analisar”, “Estatísticas descritivas”, “Frequências”, ao solicitarmos a estatística descritiva (por exemplo, para a variável “IDADE”), clique na opção “Gráficos...”. Neste caso, como temos uma variável quantitativa, podemos construir um “Histograma”. Lembre-se que o histograma é um gráfico de frequência que tem como objetivo ilustrar como uma determinada amostra ou população de dados está distribuída. Na mesma janela, selecione

a opção "Mostrar curva de normalidade no histograma". Uma curva normal sobreposta em um histograma ajuda a avaliar se os dados são normalmente distribuídos. Clique em "Continuar" e, em seguida, em "OK" para gerar o arquivo de saída, ou em "Colar" para gerar o arquivo de sintaxe; 2) na janela de edição de dados, clique em "Gráficos" e, em seguida, em "Criador de gráfico...". Clique em "Ok". Na janela "Criador de gráfico", na "Galeria", existem diferentes tipos de gráficos. Selecione a opção do "Histograma" e, escolha o primeiro modelo gráfico, arrastando-o para o local indicado (área grande acima da "Galeria"). Use o modelo como seu ponto inicial. A visualização prévia do gráfico usa dados de exemplo. Para definir as variáveis dos eixos "X" e "Y", selecione a variável (por exemplo, "IDADE") em "Variáveis" e arraste-a para o eixo "X". Adicionalmente, você pode adicionar um título ao gráfico, clicando em "Títulos/Notas de rodapé". Basta inseri-lo em "Conteúdo" e clicar em "Aplicar". Por fim, clique em "Ok" para gerar o arquivo de saída ou em "Colar" para gerar o arquivo de sintaxe.

Em seguida, verifique no arquivo de saída o gráfico gerado. É possível editar o gráfico, clicando duas vezes sobre o gráfico no arquivo de saída. Na janela "Editor de gráfico", você poderá fazer edições no gráfico. Clicando com o botão direito do mouse, você pode acrescentar a curva de distribuição dos dados, clicando em "Mostrar curva de distribuição". Além disso, é possível alterar o nome dos rótulos ("X" e "Y") e formatar a área do gráfico, inclusive as cores e o plano de fundo.

Outra forma de representar os resultados é através de um gráfico do tipo *box plot*. Através deste recurso gráfico é possível observar como as variáveis estão distribuídas em relação à homogeneidade dos dados, valores de tendência central (Q1: primeiro quartil, Q2: mediana e Q3: terceiro quartil), valores máximo e mínimo e valores atípicos (*outlier*), se existirem. Lembre-se que os quartis dividem as observações em quatro grupos de tamanhos iguais. Quando a caixinha (*box*) é muito pequena, significa que os dados estão muito concentrados em torno da mediana, e se a caixinha for grande, significa que os dados são mais heterogêneos.

Para isso, clique em “Gráficos” e, em seguida, em “Criador de gráfico...”. É possível redefinir o tipo de gráfico clicando em “Redefinir”. Na galeria, selecione a opção “Diagrama em caixa”. Em seguida, selecione o primeiro modelo de gráfico e arraste para o local indicado. Selecione a variável (por exemplo, “IDADE”) e arraste para o eixo “Y”.

Clique em “Ok” para gerar o arquivo de saída ou em “Colar” para gerar o arquivo de sintaxe. Verifique o arquivo de saída. Há um valor atípico representado pelo número “100”. Isso indica que há um caso, que corresponde ao caso 100, que se diferencia de todos os outros em relação à variável “IDADE”.

Outliers: o que fazer?

Valores atípicos distorcem a média e inflacionam o desvio padrão, sendo necessária uma análise criteriosa dos dados para detectá-los. Entender os *outliers* é fundamental em uma análise de dados por pelo menos dois aspectos: 1) os *outliers* podem viesar negativamente todo o resultado de uma análise e 2) o comportamento dos *outliers* pode representar justamente o que está sendo procurado. Mas, o que fazer com os *outliers*? Pode-se excluir as observações discrepantes da amostra de dados, quando o dado discrepante é fruto de um erro de imputação de dados (por isso é importante verificar os valores mínimo e máximo), então ele precisa ser removido ou corrigido da amostra; recomenda-se realizar a análise com e sem o *outlier* para ver seu impacto sobre os resultados e; realize, se possível, uma análise separada apenas com os *outliers*: esta abordagem é útil quando se quer investigar casos extremos.

Distribuição de frequências no SPSS

Como já mencionado, por padrão, o SPSS produz uma distribuição de frequências de todos os escores no formato de tabela. Essa distribuição de frequências fornece exibições estatísticas e gráficas que são úteis para descrever as variáveis, incluindo variáveis quantitativas (escala) e qualitativas (nominal ou ordinal). As tabulações e porcentagens fornecem uma descrição útil dos dados, especialmente para as variáveis com categorias. Lembre-se que é necessário utilizar códigos numéricos para codificar variáveis categóricas.

Para isso, na janela de edição de dados do "Banco de dados 1.sav", clique em "Analisar", "Estatísticas Descritivas" e, em seguida, em "Frequências...". É possível redefinir as opções de teste clicando em "Redefinir". Localize no painel da esquerda a variável de interesse (por exemplo, "SEXO") e clique na seta para mover para a lista "Variável(is)". Mantenha somente a opção "Exibir tabelas de frequência" selecionada. Clique em "Ok" para gerar o arquivo de saída ou em "Colar" para gerar o arquivo de sintaxe. Verifique no arquivo de saída a análise de frequência gerada.

Na primeira tabela, temos o número de casos válidos e omissos (*missing value*). Na segunda tabela, constam as frequências absoluta e relativa (percentual), além das frequências percentuais válida e cumulativa para as categorias da variável "SEXO" e o total da variável. Observados os resultados, podemos interpretá-los: dos 115 alunos da disciplina, 82,6% ($n = 95$) são do sexo feminino

Representação gráfica das frequências

Gráficos são uma boa forma de apresentar, ilustrar e resumir resultados. No entanto, lembre-se que a forma de apresentar as variáveis qualitativas é através de gráficos em barra ou de pizza. Um gráfico de pizza exhibe a contribuição de partes para um todo e cada fatia corresponde a um grupo (categoria) que é definido por uma variável de agrupamento. Um gráfico de barras exhibe a contagem para cada valor ou categoria distinta como uma barra separada, permitindo comparar categorias visualmente.

Da mesma forma que obtivemos os gráficos para a variável quantitativa ("IDADE"), existem duas formas de obtermos os gráficos para as variáveis qualitativas, correspondendo aos mesmos caminhos. Na primeira forma de obtermos o gráfico, seguindo com a variável "SEXO", é através da opção "Analisar", "Estatísticas Descritivas", "Frequências...". Na opção "Gráficos", é possível selecionar "Gráfico de barras" ou "Gráfico de pizza", expressando os valores em "Frequências" ou "Porcentagens". Clique em "Continuar", e em seguida em "Ok" para gerar o arquivo de saída ou em "Colar" para gerar o arquivo de sintaxe.

Na segunda opção para obtermos o gráfico, na janela de edição de dados, clique em "Gráficos" e, em seguida, em

“Criador de gráfico...”. Clique em “Ok”. Clique em “Redefinir”, se necessário. Na galeria, selecione a opção “Barra” e selecione o primeiro modelo de gráfico, arrastando-o para o local indicado. Selecione a variável (por exemplo, “SEXO”) e arraste-a para o eixo “X”. Na opção “Propriedades do elemento...”, em “Estatístico”, selecione a opção “Contagem” para expressar o valor absoluto ou “Porcentagem” para expressar o valor percentual. Clique em “Aplicar”. Ao final, clique em “Ok” para gerar o arquivo de saída ou em “Colar” para gerar o arquivo de sintaxe. Para alterar o tipo de gráfico, como por exemplo, um gráfico de “Pizza/Polar”, basta “Redefinir” as opções no “Criador de gráfico” e selecionar a nova opção gráfica. Da mesma forma, no arquivo de saída, é possível editar o gráfico clicando duas vezes sobre o gráfico.

Exportação do arquivo de saída para Word

É possível exportar as tabelas e gráficos gerados no arquivo de saída para o Word, além da opção de copiar e colar. Para isso, no arquivo de saída, clique em “Arquivo” e, em seguida, em “Exportar”. Na janela “Exportar saída”, em “Objetos a exportar”, é possível escolher exportar tudo, apenas uma seleção ou tudo que está visível. Em “Documento”, “Tipo”, escolha o tipo de arquivo para o qual se deseja exportar os resultados (neste caso, no formato Word). Em “Nome do arquivo”, digite o local e o nome do arquivo que será salvo. Clique em “Ok” para gerar o arquivo de saída ou em “Colar” para gerar o arquivo de sintaxe. Pronto! O arquivo com as análises estatísticas será gerado no Word.

Exportando um banco de dados do SPSS para o Excel

Para exportar um banco de dados do SPSS para o Excel, clique em “Arquivo” e, em seguida, em “Salvar como...”. Na janela “Salvar dados como”, crie um nome para o banco de dados, em “Nome do arquivo” e, em “Salvar como tipo”, escolha “Excel (.xls)”. Clique em “Salvar”. Um novo arquivo com a terminação “.xls” será salvo.

Importando um banco de dados do Excel para o SPSS

Antes de importar um banco de dados do Microsoft Office Excel para o SPSS, é importante revisarmos alguns aspectos:

lembre-se que cada linha corresponde a um caso e que cada coluna corresponde a uma variável; o nome da variável pode ter acento, mas não pode conter espaços (pode-se substituir por “_”) e deve conter 8 dígitos no máximo; o SPSS não entende as variáveis enquanto nomes, devendo ser substituídas por códigos numéricos. Como visto no Capítulo 2, é possível substituir as categorias da variável por códigos utilizando a ferramenta “Localizar e selecionar” do Microsoft Office Excel.

Para exemplificar a importação do banco de dados, vamos utilizar o “Banco de dados 1.xls” (disponível em <https://bit.ly/bancosdedados>). Ao abrir o SPSS, na caixa de diálogo, é possível abrir arquivos existentes na opção “Abrir uma origem de dados existente”. Na janela seguinte (“Abrir dados”), é possível selecionar em “Arquivos do tipo” a terminação “Excel (*.xls, *.xlsx, *.xlsm)” e importar bancos de dados nestes formatos. Selecione o arquivo “Banco de dados 1.xls” e clique em “Abrir”. Na janela “Abrindo origem de dados do Excel”, caso a primeira linha da planilha do Excel corresponda aos nomes das variáveis (como é o caso), é importante você indicar a opção “Ler nomes de variável a partir da primeira linha de dados”. Após, clique em “Ok”. Pronto! O banco de dados será importado do Microsoft Office Excel para o SPSS. Na mesma janela, é possível selecionar ainda a aba da planilha do Excel que se deseja importar (caso haja mais de uma) e qual a faixa de valores dessa planilha (intervalo de casos), indicando em “Planilha” e “Faixa”, respectivamente. Caso o SPSS esteja aberto, esse mesmo processo de importação pode ser feito clicando em “Arquivo”, “Abrir” e, em seguida, em “Dados”, seguindo os mesmos passos descritos.

Após a importação dos dados, verifique a organização e a forma de apresentação dos dados. Note que na planilha “Visualização da variável” é necessário preencher as informações das colunas referentes às variáveis, como tipo de variável, número de casas decimais após a vírgula, descrição do rótulo e valores dos rótulos. Lembre-se também de verificar se a variável está classificada corretamente (ordinal, nominal ou escala).

Deixamos aqui uma sugestão. Como mencionado no Capítulo 2, o Excel é um software de planilha, enquanto o SPSS

é um software de análise estatística. Desta forma, é indicado que você construa o banco de dados diretamente no SPSS, pois 1) evita erros de importação, 2) considere seu tempo precioso (com a importação será necessário revisar e especificar todas as informações das variáveis) e, além disso, 3) você irá utilizar o SPSS de qualquer forma para análise dos dados.

Referências

IBM®. IBM SPSS software. Disponível em <https://www.ibm.com/analytics/spss-statistics-software>.

Field A. Descobrimos a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Exercícios sugeridos

1. Este primeiro exercício propõe, com base nas variáveis do seu projeto de pesquisa descritas no exercício 9 do Capítulo 1, a elaboração do banco de dados no SPSS. Este é um bom exercício prático para você se familiarizar e conhecer o software estatístico.

2. Com relação ao “Banco de dados 1.sav”, que fornece informações de alunos de uma disciplina:

- a) Qual a frequência do nível mestrado entre os alunos da disciplina?
- b) Dentre os programas de pós-graduação, qual(is) o(s) mais frequente(s) entre os alunos da disciplina?
- c) Como é a distribuição do conhecimento dos alunos em estatística? E em epidemiologia?
- d) A maioria dos alunos tem experiência no uso do SPSS?

3. No “Banco de dados 2.sav”, que traz dados antropométricos e de pressão arterial de sujeitos com obesidade, realize a análise descritiva para as variáveis idade, Índice de Massa Corporal (IMC), circunferência da cintura e medidas de pressão arterial. Construa os histogramas e interprete os resultados obtidos supondo distribuição não normal dos dados.

4. No “Banco de dados 2.sav”, realize a análise de frequências para as variáveis grau de obesidade e hipertensão. Construa os gráficos e interprete os resultados.

4 Transformando e manipulando dados no SPSS

*Marina Petter Rodrigues
Michele Strelow Moreira
Edison Capp
Otto Henrique Nienov*

Por que transformar e manipular dados?

O banco de dados é uma estrutura que organiza e armazena informações relacionadas a unidade experimental (pessoa, animal ou objeto), obtidas através de um instrumento de coleta (como fichas de coleta, planilhas, prontuários, etc). Um banco de dados precisa ser planejado. Cada linha corresponde a uma unidade experimental (ou caso) e contém os dados sobre aquela unidade. Cada coluna corresponde a uma variável e contém um único tipo de informação que aparece para cada caso do estudo. Uma boa forma de elaborar o banco de dados é avaliar se as variáveis contidas nele são suficientes para responder os objetivos propostos pelo projeto de pesquisa.

Precisamos considerar todas as formas de avaliar os dados que serão coletados. No entanto, algumas vezes, os dados com os quais começamos nem sempre são organizados da melhor forma para sua análise. Por exemplo, você pode querer criar uma variável categórica a partir de uma variável quantitativa, ou criar uma variável de tempo a partir da diferença entre duas datas, ou combinar variáveis categóricas em uma única variável, ou criar uma variável através de um cálculo com outras variáveis existentes, ou, ainda, criar uma variável a partir da contagem de outras.

Além disso, há casos em que os arquivos de dados não estão organizados da forma ideal para sua análise. Por exemplo, você pode querer classificar os casos com base no valor de uma ou mais variáveis, ou restringir sua análise a um subconjunto de casos, ou executar análises simultâneas em diferentes subconjuntos, ou, ainda, incluir novos casos ou novas variáveis no banco de dados.

Nesse sentido, o SPSS fornece ferramentas de manipulação de dados integradas para gravação e transformação de variáveis.

Para demonstrar essas ferramentas, neste capítulo será utilizado o “Banco de dados 3.sav” (disponível em <https://bit.ly/bancosdedados>), que contém dados demográficos, antropométricos, de pressão arterial e perfil lipídico de uma amostra de sujeitos obesos. Além disso, os sujeitos foram classificados de acordo com a frequência e duração das atividades cotidianas (caminhadas, atividades físicas moderadas e vigorosas) realizadas nos últimos 7 dias em cinco categorias: sedentários, insuficientemente ativos B e A, ativos ou muito ativos, de acordo com o Questionário Internacional de Atividade Física (IPAQ) versão curta. Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Transformar variável quantitativa em categórica

Uma abordagem muito utilizada é a de categorizar as variáveis quantitativas, ou seja, transformar uma variável quantitativa em uma variável qualitativa ordinal. A categorização significa escolher dois ou mais valores contínuos e agrupá-los na mesma categoria. A transformação de variáveis quantitativas em categóricas pode ser utilizada para facilitar a caracterização da amostra.

Para exemplificar, iremos categorizar a variável “circunferência da cintura” (“CC”) utilizando os quartis. Lembre-se que os quartis são pontos de corte na escala da variável de tal forma que, cada grupo formado a partir destes pontos de corte terá 1/4 dos casos, ou seja, 25% do tamanho total da amostra. Podemos categorizar variáveis quantitativas de diferentes formas.

Uma das formas de categorização é utilizando a opção “Recodificar em variáveis diferentes”, do menu “Transformar”. Para isso, primeiro precisamos calcular os quartis para definirmos os pontos de corte. Como visto no capítulo 3, obtemos os quartis no menu “Analisar”, clicando em “Estatísticas descritivas” e, em

seguida, em "Frequências". Na janela "Frequências", localize na tela da esquerda a variável de interesse (por exemplo, "circunferência da cintura") e clique na seta para mover para a lista de "Variável(is)". Clique em "Estatísticas" e assinale a opção "Quartis". Clique em "Continuar" e, após, em "OK" ou em "Colar" para gerar o arquivo de sintaxe. Lembre-se que, se você colar a sintaxe, é preciso selecioná-la para rodar o comando clicando em "Executar seleção". Os resultados da análise aparecerão no arquivo de saída e, a partir dos valores dos quartis obtidos, iremos categorizar a variável "circunferência da cintura" conforme mostrado no quadro 1. O número de categorias de agrupamento é um número maior do que o número de pontos de corte. Portanto, neste exemplo, a nova variável categorizada terá quatro categorias, sendo que cada uma irá conter 25% do tamanho total da amostra.

Quadro 1. Pontos de corte para categorização da variável "circunferência da cintura".

CATEGORIA	INTERVALO DE VALORES
1	Até 104,00 cm
2	104,01 cm a 111,00 cm
3	111,01 cm a 120,00 cm
4	Mais de 120,01 cm

Definidos os pontos de corte, vamos categorizar a variável. Clique em "Transformar" e, em seguida, em "Recodificar em variáveis diferentes". Na janela "Recodificar em variáveis diferentes", localize na tela da esquerda a variável "Circunferência da cintura" e clique na seta para mover para a tela "Variável numérica -> Variável de saída". Atribua um novo nome para a variável de saída, como por exemplo "CCREC", em "Nome". Informe também uma descrição para a variável que será criada, como por exemplo "Circunferência da cintura recodificada", em "Rótulo". Clique em "Alterar". Criada a nova variável, precisamos informar as quatro categorias definidas pelos valores dos quartis. Para isso, clique em "Valores antigo e novo". Clique em "Faixa, MAIS BAIXO até o valor" e digite o valor obtido para o primeiro

quartil, no caso "104,00", que corresponderá ao menor valor de corte. Em "Novo valor", digite "1", pois, conforme apresentado no quadro 1, é a primeira categoria. Clique em "Adicionar". Na tela "Antigo -> Novo" estará indicado o valor correspondente a categoria 1. Em seguida, vamos informar nossa segunda categoria. Como corresponde a uma faixa de valores, clique em "Faixa", e indique o intervalo do valor, neste caso, "104,01" até "111,00". Em "Novo valor", digite "2", pois corresponde a nossa segunda categoria. Clique em Adicionar. Na tela "Antigo -> Novo" estará indicado o valor correspondente a categoria 2. Repita a ação, indicando na "Faixa" o intervalo da terceira categoria, neste caso, "111,01" até "120,00". E, em "Novo valor", digite "3". Clique em "Adicionar". Por fim, precisamos indicar o último ponto de corte. Selecione a opção "Faixa, valor até o MAIS ALTO" e digite o valor "120,01". Em "Novo valor", digite "4". Clique em "Adicionar". Pronto! As quatro categorias com os respectivos pontos de corte foram informadas. Clique em "Continuar" e, em seguida, em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. Lembre-se que, se você colar a sintaxe, é preciso selecioná-la para rodar o comando clicando em "Executar seleção". A nova variável "CCREC", correspondente à variável "CC" categorizada, automaticamente será incluída no final do banco de dados. É importante informar os valores dos rótulos da variável criada em "Valores" na planilha "Visualização da variável".

Outra forma de categorizar uma variável quantitativa é através da opção "Armazenamento visual", no menu "Transformar". Na janela de "Armazenamento visual", localize na tela da esquerda a variável de interesse (por exemplo, "circunferência da cintura") e mova para a lista "Variáveis para armazenar". Clique em "Continuar". Na janela seguinte, em "Variável armazenada", defina um novo nome para a variável, como por exemplo "CCQUARTIS", pois também vamos categorizá-la baseado nos valores dos quartis. Perceba que há um histograma que exhibe a distribuição da variável selecionada. Clique em "Fazer pontos de corte". Escolha a opção "Percentis iguais baseados em casos verificados". Preencha uma das seguintes opções: ou "Número de pontos de corte:" n° de grupos (categorias) - 1, no caso, $4 - 1 = "3"$; ou "Largura (%):" % de casos em cada grupo, no caso como

são quatro grupos de mesmo tamanho, $100/4 = "25"$. Clique em "Aplicar". Os valores exibidos na grade representam os pontos de corte definidos, que são os valores terminais superiores de cada categoria. As linhas verticais no histograma também indicam os locais dos pontos de corte. Por padrão, esses valores de ponto de corte são incluídos nas categorias correspondentes. Por exemplo, a primeira categoria irá incluir todos os valores menores ou iguais a 104,00 cm. É possível inserir os valores dos rótulos automaticamente clicando em "Fazer rótulos". Isso gera automaticamente rótulos de valores descritivos para cada categoria. Você também pode inserir manualmente ou alterar pontos de corte e rótulos diretamente na grade, alterar locais de pontos de corte arrastando e soltando as linhas dos pontos de corte no histograma e excluir pontos de corte arrastando as linhas de pontos de corte para fora do histograma. Clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. Uma mensagem de alerta indicará que "As especificações de armazenamento criarão 1 variável". Clique em "Ok". A nova variável "CCQUARTIS" é exibida na janela de edição de dados. Como a variável é incluída no final do arquivo, ela será exibida na coluna à extrema direita na planilha "Visualização de Dados" e na última linha na planilha "Visualização de Variável".

Calcular a diferença entre duas datas

Uma das tarefas mais comuns envolvendo datas é calcular o período de tempo entre duas datas. Para exemplificar, iremos criar uma variável "Idade" a partir das variáveis de "Data de nascimento" ("DN") e de "Data da coleta de dados" ("DATACOL"), utilizando a função "Datediff (datetime2,datetime1,"unit")". Para isso, no menu "Transformar", clique em "Calcular variável...". Na janela "Calcular variável", em "Variável de destino", digite o nome da nova variável, como por exemplo "IDADE". Clique em "Tipo e rótulo" e defina uma expressão para a variável, como "Idade (em anos)", ou utilize a própria expressão como rótulo, clicando em "Usar expressão como rótulo". Na tela "Grupo de função", selecione "Aritmética de data" e, na tela abaixo ("Funções e variáveis especiais"), selecione a opção "Datediff". Ao selecionar a opção, clique na seta para inseri-la na tela de

"Expressão numérica". A função será inserida na expressão. A função não é concluída até você inserir os argumentos, representados por pontos de interrogação na função colada. O número de pontos de interrogação indica o número mínimo de argumentos requeridos para concluir a função. Leia atentamente a descrição da função "*Datediff (datetime2,datetime1,\"unit\")*": será calculada a diferença entre a data da coleta ("*datetime2*") e a data de nascimento ("*datetime1*") para obtermos a idade em anos ("*years*"). Desta forma, destaque o primeiro ponto de interrogação na função colada, localize na tela da esquerda a variável "Data da coleta de dados" e clique na seta para inseri-la na "Expressão numérica" (a variável selecionada deve ficar no lugar de "*datetime2*"). Em seguida, destaque o segundo ponto de interrogação na função colada, localize na mesma tela a variável "Data de nascimento" e clique na seta para inseri-la (a variável selecionada deve ficar no lugar de "*datetime1*"). Em "*units*" (que corresponde ao terceiro ponto de interrogação), digite, entre aspas, a unidade de tempo em que a variável será expressa, no caso, "*years*" (utilize "*months*" para meses, "*days*" para dias, etc.). Clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. A variável será incluída no final do banco de dados.

Combinando variáveis categóricas

Em algumas situações pode ser interessante combinar duas variáveis categóricas para obtermos uma variável categórica combinada. Para exemplificar, iremos combinar as variáveis "Sexo" ("*SEXO*") e "Hipertensão" ("*HAS*") para criarmos a variável "*SEXOHAS*". Sabendo que a variável "Sexo" é categorizada da seguinte forma: 1 = Sexo masculino e 2 = Sexo feminino e, a variável "Hipertensão" é categorizada em: 0 = Não hipertenso e 1 = Hipertenso, a variável "*SEXOHAS*" poderá apresentar as categorias apresentadas no Quadro 2.

Quadro 2. Combinação das categorias das variáveis “SEXO” e “HAS” para criação das categorias da variável “SEXOHAS”.

CATEGORIA	VALOR DOS RÓTULOS
1	Sexo masculino hipertenso
2	Sexo masculino não hipertenso
3	Sexo feminino hipertenso
4	Sexo feminino não hipertenso

No menu “Transformar”, clique em “Calcular variável...”. Lembre-se que é possível redefinir os comandos de teste na opção “Redefinir”. Em seguida, na janela “Calcular variável”, informe o nome da nova variável, como por exemplo “SEXOHAS”, em “Variável de destino”. Clique em “Tipo e rótulo” e defina uma expressão para a variável, como “Combinação de sexo e hipertensão”, ou utilize a própria expressão como rótulo, clicando em “Usar expressão como rótulo”. Na tela “Expressão numérica”, digite “1”. Em seguida, vamos informar a combinação de valores de rótulos que define a primeira categoria da variável, conforme apresentado no quadro 2. Clique em “Se...(condição de seleção de caso opcional)”. Selecione a opção “Incluir se o caso atender à condição”. Em seguida, localize na tela da esquerda a variável “SEXO”, selecione e insira a variável dentro da tela “Incluir se o caso atender à condição”. Digite ao lado da expressão SEXO utilizando o teclado virtual localizado no centro da janela, “= 1” (correspondendo ao “Sexo masculino”). Agora, precisamos inserir o restante da combinação que define a primeira categoria. Insira “&” e localize na tela da esquerda a variável “HAS”. Selecione e insira a variável dentro da tela “Incluir se o caso atender à condição”. Na mesma tela, ao lado da expressão HAS, digite utilizando o teclado virtual localizado no centro da janela, “= 1” (correspondendo ao “Hipertenso”). Após esse procedimento a expressão na janela deve ser a seguinte: “SEXO = 1 & HAS = 1”, ou seja, a categoria 1 = Sexo masculino hipertenso. Clique em “Continuar”. Em seguida, clique em “Ok” ou em “Colar” para gerar o arquivo de sintaxe. A variável será incluída no final do banco de dados.

No entanto, temos somente a primeira categoria da nova variável informada. Ainda precisamos incluir as outras três categorias da variável "SEXOHAS". Para isso, procedese de maneira análoga, alterando o código na tela "Expressão numérica" para "2", "3" e "4" e a expressão da tela "Incluir se o caso atender à condição" para "SEXO = 1 e HAS = 0", "SEXO = 2 e HAS = 1" e "SEXO = 2 e HAS = 0", respectivamente.

Na janela de edição de dados existe um atalho onde é possível acessar de forma rápida os comandos acessados recentemente, em "Chamar novamente as caixas de diálogo usadas recentemente". Caso contrário, clique novamente em "Calcular variável...", no menu "Transformar". Na tela "Expressão numérica", digite "2". Clique em "Se..." para alterar a expressão. Na tela "Incluir se o caso atender à condição", informe utilizando o teclado virtual localizado no centro da janela, a expressão "SEXO = 1 & HAS = 0", ou seja, a categoria 2 = Sexo masculino não hipertenso. Clique em "Continuar". Em seguida, clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. Uma mensagem perguntará se você deseja "Alterar variável existente?". Clique em "OK". Você verá que a variável "SEXOHAS" apresenta a primeira e a segunda categorias.

Em seguida, na tela da "Expressão numérica", informe o dígito "3". Clique em "Se..." para alterar a expressão. Na tela "Incluir se o caso atender à condição", informe utilizando o teclado virtual localizado no centro da janela, a expressão "SEXO = 2 & HAS = 1", ou seja, a categoria 3 = Sexo feminino hipertenso. Clique em "Continuar". Em seguida, clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. A variável "SEXOHAS" apresenta três categorias. Por fim, para obtermos todas as categorias da nova variável "SEXOHAS", na tela "Expressão numérica", digite "4". Clique em "Se..." e, na tela "Incluir se o caso atender à condição", informe a expressão "SEXO = 2 e HAS = 0", ou seja, a categoria 4 = Sexo feminino não hipertenso. Clique em "Continuar". Em seguida, clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. Pronto! As quatro categorias da variável "SEXOHAS" foram criadas.

Calculando variável através de outras existentes

Na opção "Calcular variável...", além de uma ampla variedade de funções matemáticas e equações complexas (como logaritmo decimal "Log10" e exponencial "Exp", localizadas na função "Aritmética") podemos calcular uma nova variável utilizando variáveis existentes no banco de dados. Como vimos até agora, o resultado de qualquer uma das funções de cálculo irá produzir uma nova coluna na janela de edição de dados. Para exemplificar, iremos criar uma variável "Índice de Massa Corporal", a partir das variáveis "Massa corporal" e "Estatura", através da fórmula: $IMC = \text{Massa corporal (em kg)} / \text{Estatura}^2 \text{ (em metros)}$.

No menu "Transformar", clique em "Calcular variável". Na caixa de diálogo "Calcular variável", clique em "Redefinir" e, em seguida, em "Variável de destino", digite o nome da nova variável, por exemplo, "IMC". Se você digitar um nome que já existe, o SPSS irá informar e perguntar "Alterar variável existente?". Se você clicar em "Ok", o SPSS substituirá os dados existentes na coluna com os resultados da função utilizada. Se você clicar em "Cancelar", você precisará renomear a nova variável. Clique em "Tipo e rótulo" e defina uma expressão para a variável, como "Índice de Massa Corporal (em kg/m²)", ou utilize a própria expressão como rótulo, clicando em "Usar expressão como rótulo". Localize na tela da esquerda a variável "Massa corporal", selecione e clique no botão de seta para copiá-la para a caixa de texto "Expressão numérica". No teclado virtual localizado no centro da janela, clique em "/" (função de divisão). Em seguida, localize na tela da esquerda a variável "Estatura", selecione e utilize o botão de seta para inseri-la na caixa de texto "Expressão numérica". No teclado virtual localizado no centro da janela, clique em "*" (função de exponenciação) e em "2". A função final será "MASSA/ESTATURA ** 2". No mesmo teclado, temos funções de adição ("+"), subtração ("-"), multiplicação ("*") e símbolos de menor que ("<"), menor que ou igual a ("<="), maior que (">"), maior que ou igual a (">="), igual a ("=") e não igual a ("!="). Clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. Pronto! A variável "IMC" foi criada e automaticamente incluída no final do banco de dados.

Como criar uma variável a partir da contagem de outras

Muitas vezes nos deparamos com situações onde é necessário contar em quantas variáveis apareceu algum código. Por exemplo, quantas medicações o indivíduo toma, ou quantas comorbidades o paciente possui, ou ainda, quantas variáveis possuem dados omissos (*missing values*). Para isso, no comando "Valores de contagem dentro de casos", no menu "Transformar", é possível criarmos uma variável cuja informação seja uma contagem a partir de outras.

Para exemplificar, iremos contar quantos sujeitos praticam atividade física através da variável "IPAQ". Originalmente, a variável possui cinco categorias válidas: 0 = Sedentário, 1 = Insuficientemente ativo B, 2 = Insuficientemente ativo A, 3 = Ativo e, 4 = Muito ativo. No entanto, digamos que o investigador considere apenas as categorias "3" e "4" como aquelas que representam quem realmente pratica atividade física. No menu "Transformar", clique em "Valores de contagem dentro de casos...". Na janela "Ocorrência de contagem de valores dentro de casos", em "Variável de destino", digite o nome da variável que irá armazenar a contagem, como por exemplo, "PRATAF". Em "Rótulo de destino", digite uma descrição para a variável, como por exemplo "Prática de atividade física". Na tela da esquerda, selecione as variáveis (neste caso, "IPAQ") e utilize o botão da seta para inclui-las na tela "Variáveis numéricas". É possível incluir mais de uma variável na tela "Variáveis numéricas" e definir para cada uma o valor de contagem. Clique em "Definir valores...". Como definimos as categorias "3" e "4" como sendo as que definem a prática de atividade física, selecione a opção "Faixa" e informe os valores "3" até "4". Clique em "Adicionar". A faixa de valores informada será incluída na tela "Valores a contar". Clique em "Continuar". Em seguida, clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. A nova variável "PRATAF" é exibida na janela de edição de dados. Como a variável é incluída no final do arquivo, ela será exibida na coluna à extrema direita na planilha "Visualização de Dados" e na última linha na planilha "Visualização de Variável".

Classificação de casos a partir de uma ou mais variáveis

A classificação de casos (ordenação de linhas do arquivo de dados) é frequentemente útil e, às vezes, necessária para determinados tipos de análise. Para reordenar a sequência de casos no arquivo de dados com base no valor de uma ou mais variáveis utilizamos a opção "Classificar casos" do menu "Dados". Na tela "Classificar casos", selecione a variável de interesse segundo a qual o arquivo deve ser ordenado e mova-a, através do botão da seta, para o quadro "Classificar por". Se você selecionar várias variáveis de classificação, a ordem em que elas aparecem no quadro "Classificar por" determina a ordem na qual os casos são classificados. Desta forma, o arquivo será primeiramente ordenado pelos valores da primeira variável escolhida e depois pela segunda variável. A segunda ordenação é feita para os valores comuns da primeira variável. Logo abaixo do quadro "Classificar por", existe um quadro chamado "Ordem de classificação", que permite escolher se a ordenação será "Crescente" ou "Decrescente".

Para exemplificar, faça uma ordenação segundo a variável "Sexo", selecionando uma ordem crescente, e segundo a variável "Idade", selecionando uma ordem decrescente. Mova a variável "Sexo" para o quadro "Classificar por" e escolha a opção "Crescente" no quadro "Ordem de classificação". Em seguida, mova a variável "Idade" para o quadro "Classificar por" e escolha a opção "Decrescente" no quadro "Ordem de classificação". Clique em "Ok", para executar a ordenação, ou "Colar", para gerar o arquivo de sintaxe. Atenção! Após a execução deste comando a posição dos indivíduos nas linhas fica completamente alterada, pois o indivíduo na primeira linha do banco de dados após a ordenação pode não ser o primeiro caso digitado. Para que esta informação não se perca é essencial que exista uma variável com a identificação dos casos ("ID").

Uma outra alternativa, mais prática para fazer esta ordenação, é clicar com o botão direito do mouse sobre o nome da variável na "Planilha de dados" e selecionar a ordenação ("Classificar em ordem crescente" ou "Classificar em ordem decrescente") na caixa de comando. Perceba que nesta caixa também é possível "Cortar", "Copiar", "Colar", "Eliminar" e "Inserir" uma variável.

Restringindo a análise a um subconjunto de casos

É possível restringir sua análise a um subgrupo específico com base nos critérios que incluem variáveis e expressões complexas ou também selecionar uma amostra aleatória de casos. Os critérios usados para definir um subgrupo podem incluir valores de variáveis e intervalos, intervalos de data e hora, números de casos (linhas), expressões aritméticas e funções. Essa seleção de parte do arquivo de dados pode ser uma necessidade temporária, como quando queremos verificar alguma condição em determinado grupo, ou permanente, como quando estamos interessados em estudar apenas um segmento específico da amostra. O SPSS possui várias formas de seleção de dados. Para fazer qualquer tipo de seleção, utilizamos a opção "Selecionar casos..." no menu "Dados".

Na janela "Selecionar casos", estão presentes cinco opções diferentes para seleção, descritas no quadro "Selecionar":

1. "Todos os casos": opção utilizada por padrão e inclui todas as observações do banco de dados, ou seja, nenhuma seleção é executada. Clique em "Ok" para selecionar todos os caso, ou em "Colar" para gerar o arquivo de sintaxe.

2. "Se a condição for cumprida": através dessa opção, podemos definir expressões condicionais para seleção de casos, inclusive expressões aritméticas e funções. Para isso, indicamos a condição clicando em "Se...". Na caixa de diálogo "Selecionar casos: Se", a expressão condicional pode usar nomes de variáveis existentes, constantes, operadores aritméticos, operadores lógicos, operadores relacionais e funções. É possível digitar e editar a expressão na caixa de texto apenas como texto em uma janela de saída. É possível também usar o teclado da calculadora, a lista de variáveis e a lista de funções para colar elementos na expressão. Desta forma, você pode mover para o quadro superior as variáveis que estão na tela da esquerda, bem como selecionar as funções da tela "Grupo de função" e inserir uma função específica da tela "Funções e variáveis

especiais” no mesmo quadro. Por exemplo, para avaliar apenas os sujeitos ativos (“3”) e muito ativos (“4”) da variável “Classificação IPAQ”, selecione a variável “IPAQ” na tela da esquerda, insira-a no quadro superior com o botão da seta e informe, utilizando o teclado virtual localizado no centro da janela, a primeira parte da expressão: “IPAQ = 3”. Para acrescentar a outra categoria da mesma variável “IPAQ”, insira o símbolo “|” e adicione o restante da expressão: “IPAQ = 4”. Após esse procedimento, o quadro superior deve conter a seguinte expressão: “IPAQ = 3 | IPAQ = 4”, ou seja, sujeitos ativos (“3”) e muito ativos (“4”) da variável “IPAQ” serão selecionados. Nota: utilize o símbolo “&” quando forem selecionadas categorias de diferentes variáveis. Clique em “Continuar”. Em seguida, clique em “Ok” ou em “Colar” para gerar o arquivo de sintaxe. Na janela de edição de dados, na planilha “Visualização de dados”, você perceberá algumas alterações após a seleção: os casos (linhas) que não foram selecionadas apresentam-se tachadas com uma linha diagonal no meio do número da linha; a barra localizada na parte inferior da janela apresenta a mensagem “Filtro Ativado”; além disso, uma coluna de nome “filter_\$” é adicionada. Essa nova coluna apresenta valor “0” para as linhas que não foram selecionadas e valor “1” para as linhas que foram selecionadas.

3. “Amostra aleatória de casos”: podemos selecionar uma porcentagem ou número exato de casos. A seleção é feita aleatoriamente. Para isso, clicamos em “Amostra”. Na caixa de diálogo “Selecionar casos: Amostra aleatória”, é possível selecionar uma das seguintes alternativas para o tamanho da amostra: a) “Aproximadamente”: em que será selecionada uma porcentagem especificada pelo investigador, gerando uma amostra aleatória aproximadamente da porcentagem especificada de casos; b) “Exatamente”: neste caso, um número de casos especificado pelo usuário será selecionado. Você deve também especificar o número de casos a partir dos quais deve ser gerada a amostra. Este segundo número deve ser menor ou igual ao número total de casos no arquivo de

dados. Se o número exceder o número total de casos no arquivo de dados, a amostra conterá proporcionalmente menos casos do que o número solicitado.

4. "Com base na hora ou faixa de casos": usamos essa opção quando estamos interessados em selecionar uma faixa específica de valores, por exemplo, os casos do número 50 ao 100, ou também podemos fazer seleções baseadas em datas. Para isso, informe o intervalo em "Faixa...". Na caixa de diálogo "Selecionar casos: Faixa", é possível selecionar um intervalo de números de observações (linhas). Na "Observação", no "Primeiro caso", insira o número de observação inicial (número da linha na janela de edição de dados). Se você não especificar um valor de "Último caso", todos os casos a partir do número de observação inicial serão selecionados. No "Último caso", insira o número de observação final (número da linha na janela de edição de dados). Se você não especificar um valor de "Primeiro Caso", todos os casos a partir do número de observação final serão selecionados.

5. "Usar variável de filtro": nesta opção, uma variável é escolhida no banco de dados e usada como filtro. Todos os casos para os quais a variável filtro assume o valor "0" não serão selecionados.

Na mesma janela, no quadro "Saída", você pode:

1. "Não incluir na filtragem os casos não selecionados": os casos (linhas) que não são selecionados não são incluídos nas análises posteriores, porém, permanecem na janela de dados. Caso você queira usar os casos não selecionados, na mesma janela, no quadro "Selecionar", opte por incluir "Todos os casos" na seleção. Se você selecionar uma amostra aleatória ou se selecionar casos com base em uma expressão condicional, isso gerará uma variável denominada "filter_\$" com um valor de "1" para casos selecionados e um valor de "0" para casos não selecionados;

2. "Copiar casos selecionados para um novo conjunto de dados": nesta opção, o SPSS cria uma nova janela de dados

com os casos selecionados, ou seja, os casos selecionados são copiados para um novo conjunto de dados, deixando o conjunto de dados original não afetado. Os casos não selecionados não são incluídos no novo conjunto de dados e são deixados em seu estado original no conjunto de dados original. O nome da nova janela de dados deve ser fornecido pelo usuário na lacuna “Nome do conjunto de dados”.

3. “Excluir os casos não selecionados”: os casos (linhas) não selecionados são apagados (excluídos) da janela de dados. Os casos excluídos podem ser recuperados apenas saindo do arquivo sem salvar nenhuma mudança e, em seguida, reabrindo o arquivo. A exclusão de casos será permanente se você salvar as mudanças no arquivo de dados. Neste caso deve-se tomar o cuidado de salvar o banco de dados com outro nome, em “Analisar”, “Salvar como...”.

Lembre-se: para retornar todo o conjunto de dados, basta selecionar a opção “Todos os casos”, no quadro “Selecionar”.

Análises simultâneas em diferentes subconjuntos

Em algumas situações, o pesquisador pode querer analisar diferentes grupos de uma amostra de forma idêntica, mas em particular para cada grupo. Por exemplo, vamos supor que o pesquisador queira analisar separadamente as medidas pressóricas para homens e mulheres, pois verificou que essas variáveis se comportam de forma completamente diferente entre os sexos. Não faz sentido, portanto, apresentar a análise destes dados considerando ambos os sexos. Ou seja, o que pretendemos fazer, daqui para frente, são duas análises idênticas, uma para cada sexo.

A opção “Dividir arquivo...”, no menu “Dados”, permite repetir qualquer análise sobre vários grupos de casos. Essa função permite especificar uma variável de grupo (lembre-se que essas variáveis são utilizadas para especificar categorias da unidade experimental). Qualquer procedimento posterior no SPSS será executado em cada uma das categorias especificadas pela variável de grupo.

Por padrão, o SPSS considera todos os casos juntos na análise, sem separação por grupos. Por esse motivo, a opção selecionada na janela "Dividir arquivo" é "Analisar todos os casos, não criar grupos". Para executar a análise para as categorias de uma determinada variável, podemos optar por "Comparar grupos" ou "Organizar saída por grupos" e, então, o quadro "Grupos baseados em" fica disponível. Movemos para o quadro "Grupos baseados em" a(s) variável(is) que definirá(ão) os grupos para os quais a análise deve ser repetida. Se mais do que uma variável for selecionada, os grupos serão definidos pela combinação das categorias de todas as variáveis. Logo abaixo, podemos ainda escolher se o banco de dados deve ser ordenado pela variável que definirá os grupos ("Classificar o arquivo agrupando variáveis") ou se o banco de dados já está ordenado pela variável que definirá os grupos ("O arquivo já foi classificado"). Se o banco não estiver ordenado corretamente, o comando não será executado de forma adequada.

Se você selecionar "Comparar grupos", os resultados de todos os grupos de divisão de arquivo serão incluídos na mesma tabela. Se você selecionar "Organizar saída por grupos", duas tabelas dinâmicas serão criadas, uma para cada grupo. Se o processamento de divisão de arquivo estiver ativo, a mensagem "Dividir por" aparecerá na barra de status na parte inferior da janela de edição dos dados.

Para exemplificar, clique em "Dividir arquivo...", no menu "Dados", marque a opção "Comparar grupos", selecione na tela da esquerda a variável "Sexo" e mova-a com o botão da seta para o quadro "Grupos baseados em". Em seguida, clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe. Lembre-se: após a execução desse comando, todas as análises feitas gerarão dois resultados, um para o sexo masculino e outro para o sexo feminino.

Note que a única mudança observada na janela de edição de dados é a mensagem "Dividir por SEXO" na barra inferior direita, ou a ordenação dos casos pela variável que definiu os grupos, caso o banco de dados ainda não estivesse ordenado. Faça uma análise descritiva para a variável "Colesterol total", por exemplo, para verificar como os resultados serão apresentados no arquivo de saída. Para isso, clique em "Analisar", "Estatísticas descritivas..."

e “Frequências...”. Na janela “Frequências”, selecione na tela da esquerda a variável “Colesterol total” e mova-a para a lista de “Variável(is)”. Clique em “Estatísticas” e solicite medidas de tendência central e de dispersão. Em seguida, clique em “Ok” ou em “Colar” para gerar o arquivo de sintaxe. Veja que os resultados são apresentados em um único bloco, mas separando em primeiro o sexo masculino e em segundo o sexo feminino.

Atenção! Quando você tiver terminado com o comando de divisão de arquivo, lembre-se de desligá-lo. Retorne à caixa de diálogo “Dividir arquivo” e selecione a opção “Analise todos os casos, não crie grupos”.

Inclusão de novos casos ou novas variáveis

Muitas vezes os dados se encontram em diferentes bancos de dados e em algum momento precisamos uni-los para poder cruzar alguma variável que está num banco com outra de outro banco ou ainda unificar em um mesmo banco, casos de diferentes bancos de dados com as mesmas variáveis. O SPSS fornece uma opção para incluir um banco de dados dentro de outro sem que seja necessário digitar todas as variáveis novamente. Nesse caso, os dois bancos precisam estar no formato SPSS, ou seja, no formato “.sav”. Esse procedimento é feito pela opção “Mesclar arquivos” no menu “Dados”, que permite a união de bancos de dados que tenham 1) as mesmas variáveis com diferentes sujeitos (“Adicionar casos”) ou 2) variáveis diferentes dos mesmos sujeitos (“Adicionar variáveis”).

Para a primeira opção, em “Dados”, selecione “Adicionar casos” em “Mesclar arquivos”. Neste caso, é imprescindível que os dois bancos de dados tenham as mesmas variáveis, com o mesmo nome e o mesmo formato. Para exemplificar, vamos juntar o “Banco de dados 3.sav” e o “Banco de dados 3 casos novos.sav” (disponível em <https://bit.ly/bancosdedados>), que possuem diferentes indivíduos cada um, mas o mesmo conjunto de variáveis. Na janela “Adicionar casos”, é possível escolher uma das seguintes opções: 1) “Um conjunto de dados aberto”: caso o outro banco de dados (“Banco de dados 3 casos novos.sav”) já

esteja aberto no SPSS, ou 2) "Um arquivo de dados externo do SPSS Statistics": se o banco de dados não estiver aberto. Clique em "Navegar" para localizar o arquivo, selecione o banco de dados e clique em "Continuar". Após selecionado o banco de dados, na caixa de diálogo seguinte, serão listadas as "Variáveis sem par". Isso significa que existem variáveis no primeiro banco de dados que não constam no segundo banco de dados. As variáveis seguidas de "*" estarão no primeiro banco de dados e aquelas seguidas de "+" estarão no segundo banco de dados e ambas não serão incluídas no novo banco de dados. Se você deseja utilizar as variáveis sem par em análises futuras, é recomendado que você salve o banco de dados original antes de adicionar os novos casos. As variáveis comuns serão automaticamente incluídas na tela "Variáveis no novo conjunto de dados ativo". Se você quiser identificar qual o banco de origem de cada caso no novo banco, clique em "Indicar origem de caso como variável". Ao clicar em "Ok" ou em "Colar", o comando será executado e os novos casos serão adicionados ao final do banco de dados. Se você solicitou "Indicar origem de caso como variável", haverá uma coluna indicando casos antigos ("0") e novos ("1").

Na segunda opção, em "Dados", selecione "Adicionar variáveis" em "Mesclar arquivos". Neste caso, é imprescindível que os dois bancos de dados não possuam as mesmas variáveis, mas possuam os mesmos sujeitos e que os casos estejam ordenados. Para exemplificar, vamos juntar as variáveis do "Banco de dados 3.sav" e do "Banco de dados 3 questionário.sav" (disponível em <https://bit.ly/bancosdedados>). Neste último, temos variáveis de um questionário de sintomas neuropáticos. Primeiro, precisamos nos certificar que os casos estejam ordenados em ambos os bancos de dados. Para isso, como vimos neste capítulo é possível ordenar os casos clicando com o botão direito do mouse sob a coluna de identificação dos casos (variável "ID") e selecionar "Classificar em ordem crescente". Da mesma forma vista anteriormente, é possível escolher, na janela "Adicionar variáveis": 1) "Um conjunto de dados aberto": caso o outro banco de dados ("Banco de dados 3 questionário.sav") já esteja aberto no SPSS, ou 2) "Um arquivo de dados externo do SPSS Statistics": se o banco de dados não estiver aberto. Clique em "Navegar" para localizar o

arquivo, selecione o banco de dados e clique em "Continuar". Na caixa de diálogo seguinte, selecione a opção "Combinar casos em variáveis chave em arquivos classificados", e opte por uma das seguintes opções: 1) "Ambos os arquivos fornecem casos": se cada banco possui uma linha para cada indivíduo da amostra, 2) "O conjunto de dados não ativo é uma tabela vinculada" ou 3) "O conjunto de dados ativo é uma tabela vinculada": se um dos bancos possuir mais de um registro (linha) por caso, estes devem estar identificados pela mesma variável chave. Opte pela primeira opção e selecione a variável "ID" que identifica o número dos casos (deve estar na tela "Variáveis excluídas"). Clique no botão da seta para inseri-la na caixa "Variáveis-chave". Na janela "Novo conjunto de dados ativo", as variáveis seguidas de "*" estão no primeiro banco de dados e aquelas seguidas de "+" estão no segundo banco de dados e correspondem as variáveis que serão incluídas. Em seguida, clique em "Ok" ou em "Colar". Uma mensagem de advertência aparecerá: "A correspondência vinculada falhará se os dados não forem classificados na ordem crescente de variáveis-chave". Clique em "Ok". Note que as variáveis do questionário serão automaticamente incluídas no final do banco de dados.

Referências

Field A. Descobrimo a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Exercícios sugeridos

1. Utilizando o "Banco de dados 3.sav":

a) Categorize a variável "IMC", criada neste capítulo, em graus de obesidade, onde: Obesidade grau I = 30,00 a 34,99 kg/m²; Obesidade grau II = 35,00 a 39,99 kg/m² e; Obesidade grau III = >40,0 kg/m².

b) Crie uma variável de Pressão Arterial Média (PAM), através das medidas pressóricas, utilizando a expressão: $PAM = ((2 * PAD) + PAS) / 3$. Onde: PAS = Pressão Arterial Sistólica e PAD = Pressão Arterial Diastólica.

c) Crie uma variável de LDL-Colesterol (LDL), utilizando a Fórmula de Friedewald: $LDL = CT - HDL - (TG / 5)$. Onde: CT = Colesterol total, HDL = HDL-Colesterol e TG = Triglicerídeos.

d) Após categorizar e calcular as variáveis, faça uma análise descritiva das variáveis pressóricas (PAS, PAD e PAM) e lipídicas (Colesterol total, Triglicerídeos, HDL-Colesterol e LDL-Colesterol), selecionando somente os indivíduos com obesidade grau III. Construa uma tabela apresentando os resultados e interprete-os supondo uma distribuição normal dos dados.

e) Faça a mesma descritiva separando por sexo, sem discriminar o grau de obesidade (incluindo todos os casos). Construa uma tabela apresentando os resultados e interprete-os supondo uma distribuição normal dos dados.

2. Utilizando o "Banco de dados 1.sav", inclua os novos casos do "Banco de dados 1 casos novos.sav".

3. Em seguida, utilizando o "Banco de dados 1.sav" com os novos casos incluídos, divida o banco de dados pelo nível de pós-graduação e responda:

a) Qual a média de idade do nível mestrado e do nível doutorado?

b) Qual a frequência de sexo feminino no nível mestrado e no nível doutorado?

c) Em que período se encontram a maioria dos estudantes do nível mestrado e do nível doutorado?

5 Teste de Hipóteses

*Nadine Ziegler
Pedro Henrique Comerlato
Sinara Santos
Edison Capp
Otto Henrique Nienov*

Criando as hipóteses

Toda pesquisa se origina de um problema, estruturado na forma de uma pergunta. Para responder essa pergunta, o pesquisador estabelece objetivos e formula hipóteses, suposições, que se presume serem verdadeiras. Ao ser formulada essa afirmação, é necessário saber se os dados, provenientes de uma amostra aleatória extraída da população de interesse, contrariam ou não a hipótese inicialmente assumida. Para tal, utilizamos um teste de hipóteses para verificar a veracidade da hipótese.

Cada hipótese apresenta uma probabilidade de erro, pois ela não representa a totalidade dos casos existentes de uma determinada condição. Lembre-se que a pesquisa é realizada em uma pequena amostra, portanto está sujeita a um erro. O erro não pode ser evitado, mas sua probabilidade pode ser controlada ou mensurada, obtendo-se uma medida de validade das conclusões do estudo.

Para testar a hipótese, existem métodos que irão verificar se ela é verdadeira ou não e, assim, o pesquisador pode tomar sua decisão a respeito das hipóteses previamente formuladas. As variáveis terão uma unidade métrica para verificar se o teste foi executado adequadamente e, tão logo, se a hipótese pode ser confirmada. Desta forma, testar a hipótese permite verificar se os dados amostrais trazem evidências que apoiem ou não a hipótese formulada e, para isso, é necessário transformar a hipótese de estudo em uma hipótese estatística.

As hipóteses estatísticas

As hipóteses estatísticas são suposições ou afirmações feitas sobre um parâmetro, ou parâmetros, na população. Em qualquer teste de hipóteses existem duas hipóteses: a hipótese nula ou de nulidade (H_0) e a hipótese alternativa (H_A ou H_1).

A hipótese nula é a primeira hipótese a ser apresentada em um projeto de pesquisa e estabelece uma ausência de diferença entre os parâmetros avaliados. Ou seja, a média da população amostrada (μ_A) é igual a média da população referência (μ_0). A hipótese alternativa corresponde a hipótese que o pesquisador quer confirmar e é contrária à hipótese nula. Ou seja, é diferente.

A hipótese testada sempre será H_0 . Portanto, é importante destacar que entender qual é a nossa H_0 e qual é a nossa H_A é fundamental, pois irá refletir na interpretação das análises realizadas pelo SPSS, que rejeita ou não a hipótese de nulidade. Se aceitarmos H_0 , rejeitamos H_A . Se rejeitarmos H_0 , aceitamos H_A .

Testando as hipóteses

Em muitas situações, o interesse do pesquisador ao tomar sua decisão, envolve a veracidade ou não de uma hipótese estatística, como por exemplo, se dois grupos têm a mesma média ou não; ou se um parâmetro populacional tem um valor em particular ou não; ou se uma intervenção é melhor do que a outra ou não; ou, simplesmente, se quer determinar o comportamento da distribuição de uma variável na população. Desta forma, o teste de hipóteses é um procedimento estatístico que permite tomar decisão (aceitar ou rejeitar a H_0) entre duas ou mais hipóteses (H_0 ou H_A), associado à um risco máximo de erro, utilizando os dados observados de um determinado estudo.

Os testes de hipótese nos fornecem subsídios para realizar esse tipo de estudo e, a escolha do teste adequado, é feita em função do estudo que se deseja realizar. O objetivo de um teste de hipóteses é verificar se são verdadeiras as afirmações sobre os parâmetros de uma ou mais populações, nos permitindo verificar

se os dados amostrais trazem evidência que apoiem ou não uma hipótese estatística formulada.

Como os dados obtidos não são de toda a população, apenas temos uma amostra representativa, precisamos estabelecer um risco máximo admitido para o erro de afirmar que existe uma diferença, quando na verdade ela não existe. Ou seja, rejeitar ou não H_0 considerando um nível de significância (α). O nível de significância corresponde ao risco máximo admitido, geralmente de 0,050 (5%) ou 0,010 (1%). Para um nível de significância de 5% temos um valor crítico de 1,96. Isto é, um desvio de até 1,96 erros padrão é admitido como não significativo. A região crítica ou região de rejeição é o conjunto de valores assumidos pela variável aleatória ou estatística de teste para os quais a H_0 é rejeitada. Seu complementar é a região de aceitação.

Além disso, os testes de hipóteses podem ser classificados em: paramétricos (usados em dados que apresentam distribuição normal) e não paramétricos (usados em dados com distribuição não conhecida ou que apresentam distribuição não normal). Os testes de normalidade serão abordados no capítulo 7.

Formas do teste de hipóteses

O pesquisador deve decidir o tipo de teste a ser realizado antes de coletar os dados da pesquisa. Preferencialmente, quando estiver elaborando o projeto de pesquisa e formulando suas hipóteses.

A maioria dos testes de hipóteses envolvendo médias é do tipo bilateral, ou seja, testa a hipótese de nulidade de ausência de diferença contra a hipótese alternativa de que existe uma diferença entre as médias (Figura 1). O método bilateral é utilizado para verificarmos se a média da amostra não é igual a média da hipótese. Assim, se supormos que a média da hipótese é descrita como sendo 100, a H_A irá alegar que a média da amostra não é 100. O método bilateral apresenta dois pontos de corte, que são determinados pelo nível de significância escolhido. Assim, se optarmos por um α de 1%, os dois extremos da curva normal representarão cada um 0,5%; se optarmos por um α de 5%, as duas extremidades da curva normal compreenderão cada uma 2,5%.

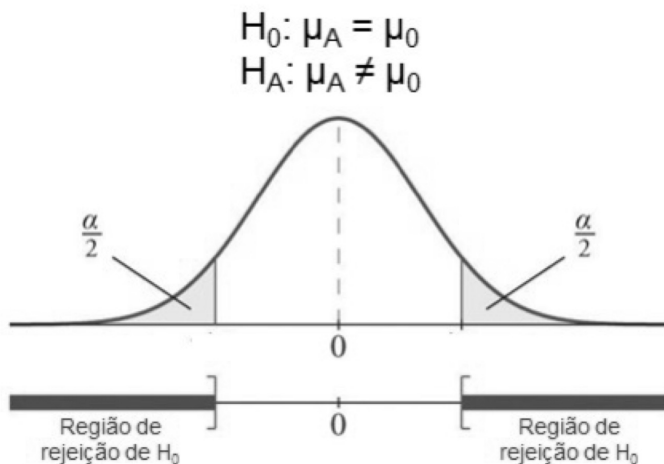


Figura 1. Teste de hipótese bilateral. O μ_0 representa o valor numérico específico considerado nas H_0 e H_A .

Se a pontuação calculada estiver entre os dois fins, aceitamos a H_0 e rejeitamos a H_A . Isso ocorre porque a pontuação será na região de não rejeição de H_0 . Se a pontuação estiver fora deste intervalo, então rejeitamos a H_0 e aceitamos a H_A porque está fora do intervalo. Portanto, a média da amostra é na verdade diferente da média da H_0 , que é a média reivindicada.

Para exemplificar, vamos supor que um laboratório de análises clínicas compre de um determinado fabricante, um kit para dosagem sérica de ácido fólico cujo limite de detecção especificado é de 24 ng/mL. Os valores de normalidade da concentração de ácido fólico sérico variam de 3 a 17 ng/mL de acordo com testes não paramétricos, conforme consta nas instruções do fabricante. Em um determinado dia, o laboratório recebeu um grande lote do kit e a equipe do controle de qualidade do laboratório resolveu verificar se o lote atende as especificações. Logo, a H_0 assume que o lote atende as especificações ($\mu = 24$) e a H_A assume que o lote não atende as especificações do fabricante ($\mu \neq 24$).

Mas, há situações em que para o pesquisador somente haverá interesse se μ_A for menor ou maior do que μ_0 . Nestes casos,

teremos um teste de hipótese unilateral à esquerda ou à direita, respectivamente. No método unilateral à esquerda, assim como à direita, o valor crítico do teste é determinado pelo ponto de corte do α escolhido (0,050, por exemplo).

Quando se está interessado apenas na diferença negativa entre as médias, temos o teste unilateral à esquerda, ou seja, a região de significância encontra-se na cauda esquerda da curva (Figura 2a). O método unilateral à esquerda é usado se quisermos determinar se a média da amostra é menor que a média da hipótese. Para um α de 5%, o valor crítico passa a ser -1,64. Qualquer valor abaixo do ponto de corte no teste unilateral à esquerda representa a região de rejeição de H_0 , ou seja, H_0 é falsa. Se a pontuação estiver acima do ponto de corte significa que este se encontra na região de não rejeição de H_0 e, logo, H_0 é verdadeira.

Por exemplo, suponhamos que a distribuição de probabilidade da massa corporal dos homens entre 20 e 30 anos em uma população obesa seja normal com média de 110 kg e desvio padrão de 15 kg. Um pesquisador propõe um tratamento que consiste na associação de um programa de atividade física, dieta "low-carb" e ingestão de um determinado medicamento. Sua hipótese é de que com o tratamento a massa corporal média da população da faixa etária em estudo reduzirá em um período de três meses, ou seja, a H_0 assume que a massa corporal se manterá após os três meses de tratamento ($\mu = 110$) e a H_A assume que a massa irá diminuir após os três meses de tratamento ($\mu < 110$).

Por outro lado, quando se está interessado apenas na diferença positiva entre as médias, temos o teste unilateral à direita, ou seja, a região de significância encontra-se na cauda direita da curva (figura 2b). O método unilateral à direita é usado se quisermos determinar se uma média da amostra é maior que a média da hipótese. Para um α de 5%, o valor crítico passa a ser +1,64. Qualquer valor acima do ponto de corte no teste unilateral à direita representa a região de rejeição de H_0 , ou seja, H_0 é falsa. Se a pontuação estiver abaixo do ponto de corte significa que este se encontra na região de não rejeição de H_0 , e aceitamos a hipótese de nulidade como sendo verdadeira.

Vamos supor que um fabricante de lâmpadas para microscópico afirme que a durabilidade da lâmpada seja de 38.500 horas. A equipe técnica de produção decide verificar se a incorporação de um novo componente no processo de fabricação aumenta a durabilidade da lâmpada. Sendo assim, a H_0 assume que a durabilidade da lâmpada é a mesma com a incorporação do componente novo ($\mu = 38.500$) e a H_A assume que a durabilidade da lâmpada é maior com a incorporação do componente novo ($\mu > 38.500$).

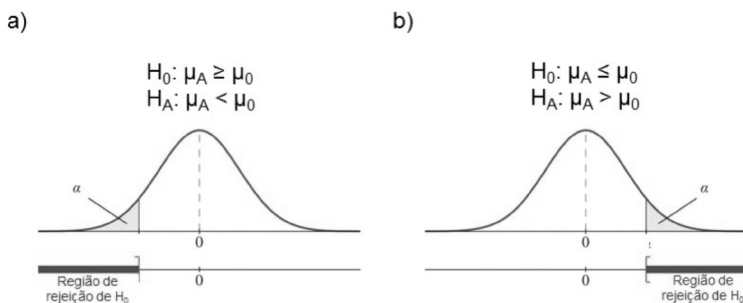


Figura 2. Teste de hipótese unilateral à esquerda (a) e à direita (b). O μ_0 representa o valor numérico específico considerado nas H_0 e H_A .

Erro do teste de hipóteses

Todo teste de hipótese tem sua conclusão sujeita à erro, pois é baseado em estatísticas de amostras extraídas de uma população. Existem dois tipos de erros:

- Erro tipo I: rejeita-se H_0 quando está é verdadeira, ou seja, é o erro de afirmar que existe uma diferença quando na verdade não existe;
- Erro tipo II: não se rejeita H_0 quando está é falsa, ou seja, é o erro de aceitar H_0 quando não se deveria. Neste caso, afirma-se uma igualdade quando o correto seria afirmar uma diferença.

A probabilidade de se cometer o erro tipo I é denominada de α , enquanto a probabilidade de se cometer um erro do tipo II é denominada de β . É muito difícil calcular sua probabilidade, por

isso o erro tipo II é utilizado quando se deseja calcular o tamanho amostral necessário para se responder o objetivo da pesquisa. Calculando o tamanho da amostra reduzimos a probabilidade deste erro acontecer. Podemos também determinar o poder da nossa amostra após a realização da pesquisa. A capacidade de um teste identificar diferenças que realmente existem, ou seja, de rejeitar H_0 quando é realmente falsa, é denominada de poder do teste $(1-\beta)$. O quadro 1 resume os tipos de erros e suas respectivas probabilidades.

Quadro 1. Tipos de erros e respectivas probabilidades.

	H_0 é verdadeira	H_0 é falsa
Rejeita-se H_0	Decisão errada: erro tipo I Probabilidade: α	Decisão correta Probabilidade: $1-\beta$ (Poder do teste)
Não se rejeita H_0	Decisão correta Probabilidade: $1-\alpha$	Decisão errada: erro tipo II Probabilidade: β

Ao diminuir a probabilidade de ocorrer um erro do tipo I, ou seja, diminuindo o valor de α , aumenta-se a probabilidade de ocorrer um erro do tipo II. Para um tamanho fixo da amostra, não há como controlar simultaneamente ambos os erros. Convencionou-se que o erro mais sério seria o erro do tipo I.

Nível de significância

O nível de significância, denominado pelo α , corresponde a probabilidade máxima permissível de se cometer o erro tipo I. Se o valor de p é "igual" ou "menor que" α , dizemos que os dados são estatisticamente significativos no nível α . Comumente utilizamos valores de α de 0,050 (5%). Desta forma, o pesquisador assume que o risco na decisão de rejeitar a H_0 pode estar incorreto em 5% das vezes.

Valor de p ou p-valor

A partir dos dados, pode-se calcular o valor da estatística do teste. A probabilidade relacionada a este valor de estatística do teste chama-se valor de p ou p-valor. Assim, o valor de p

descreve a probabilidade que mede a força da evidência contra a H_0 . A interpretação do valor de p vai determinar o significado estatístico do estudo em questão. Por exemplo, se temos um valor de $p \leq 0,050$ (estatisticamente significativo ao nível de 5%), isso nos fornece evidência contra H_0 , logo, H_A é verdadeira (existe diferença); se temos um $p > 0,050$ (sem significado estatístico), isso não nos fornece evidência contra H_0 , logo, H_A não é verdadeiro (não existe diferença). Atenção! Não significa que não haja diferença. É uma evidência daquele estudo e, talvez, a amostra seja pequena.

Etapas do teste de hipótese

- 1) Primeiro, identifique o(s) parâmetro(s) de interesse;
- 2) Em seguida, estabeleça as hipóteses H_0 ($H_0: \mu_A = \mu_0$) e H_A ($H_A: \mu_A \neq \mu_0$ ou $\mu_A < \mu_0$ ou $\mu_A > \mu_0$);
- 3) Escolha o α que determinará a região de rejeição e determine o valor crítico do teste;
- 4) Estabeleça uma estatística apropriada de teste e calcule o valor da estatística;
- 5) Decida se a H_0 deve ou não ser rejeitada;
- 6) Conclua.

Referências

Callegari-Jacques SM. Bioestatística: princípios e aplicações. Porto Alegre: ArtMed, 2011. 255p.

Exercícios sugeridos

1. Formule as hipóteses (H_0 e H_A) do seu projeto de pesquisa e descreva o tipo de teste de hipótese que será realizado. Este é um bom exercício prático para aplicar os conceitos vistos neste capítulo.

6 Métodos estatísticos para desfechos qualitativos

*Pedro Henrique Comerlato
Sinara Santos
Nadine Ziegler
Edison Capp
Otto Henrique Nienov*

As variáveis qualitativas ou categóricas não possuem valor numérico, mas são definidas através de categorias, classificando indivíduos, objetos, animais. Como visto no capítulo 1, as variáveis qualitativas podem ser divididas em nominais (aquelas em que não existe ordenação entre as categorias, como cor dos olhos, cor da pele, etc) ou ordinais (aquelas em que existe uma ordenação entre as categorias, como classe social, escolaridade, etc).

Análise de dados categóricos para uma variável

Existem situações em que o objetivo do pesquisador é verificar se existem diferenças significativas na distribuição de frequência entre as categorias de uma mesma variável. Neste caso, precisamos considerar dois aspectos: 1) o número de categorias de resposta da variável e 2) o tamanho da amostra. Lembre-se que, se a variável apresenta somente duas categorias de resposta, independente se ordinal ou nominal, esta é dicotômica. Desta forma, quando a variável apresenta duas categorias e se deseja analisar a distribuição de frequência entre as categorias de resposta em uma amostra pequena (< 25 casos), recomenda-se o uso do teste Binomial. Por outro lado, quando se tem uma amostra maior (≥ 25 casos), independente do número de categorias de resposta da variável, recomenda-se o uso do teste de Qui-quadrado para proporção (ou de aderência).

Para exemplificar o teste Binomial, no “Banco de dados 3 Resolvido.sav” (disponível em <https://bit.ly/bancosdedados>) vamos verificar se existe diferença entre a distribuição das categorias da variável “Sexo”. Para isso, vamos supor que o banco de dados

apresente apenas 20 casos. Clique em "Analisar", "Testes não paramétricos", "Caixas de diálogo legadas" e, em seguida, selecione "Binômio...". Na janela "Teste binomial", selecione a variável "Sexo" na tela da esquerda e a insira, com o auxílio do botão da seta, para a "Lista de variável de teste". Na opção "Proporção de teste", o pesquisador pode editar o valor de proporção que deseja confrontar. Caso o pesquisador não tenha informações suficientes para alterar a proporção testada, recomenda-se deixar o valor de 50% (0,5), que admite como H_0 a igualdade de distribuição entre as categorias da variável. Clique em "Exato..." e, na janela "Testes exatos", selecione a opção "Monte Carlo". Esta opção realiza os testes exatos ao invés dos assintóticos, que são uma aproximação. Por fim, clique em "Continuar" e, em "Ok" ou "Colar".

No arquivo de saída, teremos apenas um quadro (*Binomial Test*) que apresenta informações sobre as categorias da variável, a frequência absoluta (N), a proporção observada, a proporção testada (0,5) e o p-valor associado ao teste bicaudal (*Exact Sig 2-tailed*). Conforme o resultado do teste Binomial, existe diferença significativa na distribuição de proporção entre as categorias de resposta da variável "Sexo" ($p < 0,001$), devendo o pesquisador rejeitar a H_0 de que a distribuição de proporção é igual entre as categorias. Ou seja, a proporção de casos do sexo masculino ($n = 95, 21,0\%$) e do sexo feminino ($n = 367, 79,0\%$) na amostra é significativamente diferente ($p < 0,001$).

Para o teste do Qui-quadrado de proporção, utilizando o mesmo banco de dados, vamos verificar se existe diferença na distribuição de proporção entre as categorias da variável "Grau de Obesidade". No menu "Analisar", "Testes não paramétricos", "Caixas de diálogo legadas", selecione a opção "Qui-quadrado...". Na janela "Teste do Qui-quadrado", selecione na tela da esquerda a variável "Grau de obesidade" e a insira, com o auxílio do botão da seta, na "Lista de variável de teste". Clique em "Exato..." e selecione a opção "Monte Carlo". Em seguida, clique em "Continuar" e, por fim, em "Ok" ou "Colar".

Na primeira tabela informada no arquivo de saída, temos a frequência absoluta observada (N) para as categorias da variável "Grau de obesidade", a frequência absoluta esperada (assumindo

proporções iguais entre as categorias) e o resíduo (valor residual). O resíduo representa a diferença entre os valores observados e esperados. No segundo quadro (*Test Statistics*), é apresentado o resultado do teste do Qui-quadrado, o grau de liberdade (df) e o p-valor do teste (*Asymp. Sig.*). O grau de liberdade representa o número de informações independentes sobre uma estatística e é obtido por: (número de linhas da tabela de contingência – 1) x (número de colunas da tabela de contingência – 1). O resultado do teste indica que existe diferença significativa na distribuição de proporção entre as categorias da variável ($p < 0,001$). Além disso, no mesmo quadro, no item a., o SPSS indica que zero (0) células têm valor esperado menor que 5 sujeitos, ou seja, o pressuposto para aplicação do teste do Qui-quadrado foi atendido.

Mesmo para amostras maiores que 25 casos em variáveis dicotômicas, em que uma das categorias apresenta frequência relativa observada menor que 20% do valor esperado ou frequência absoluta menor que 5, é recomendado o teste Binomial. Quando a variável testada apresenta mais de duas categorias de resposta, independente do tamanho da amostra, a opção de análise é o teste do Qui-quadrado para proporção. No entanto, se alguma das categorias de resposta apresentar frequência absoluta menor que 5, devido amostragem pequena, recomenda-se reduzir o número de categorias de resposta ou dicotomizar a variável.

Análise de dados categóricos para duas variáveis

Quando o interesse do pesquisador é avaliar a associação entre duas variáveis categóricas, ou realizar uma comparação de uma variável categórica entre dois (ou mais) grupos, ou ainda, verificar se a amostra apresenta as mesmas proporções de outro estudo, utilizamos os testes de Qui-quadrado (X^2) de associação, heterogeneidade e de ajustamento, respectivamente. O teste de Qui-quadrado é um teste não paramétrico, ou seja, não depende dos parâmetros populacionais, como média e variância. O teste baseia-se na comparação das proporções, isto é, as possíveis divergências entre as frequências observadas e esperadas para um certo evento. No SPSS, os comandos para a realização dos testes de associação e heterogeneidade são os mesmos. O teste de ajustamento é encontrado nos testes não paramétricos. O

quadro 1 apresenta as hipóteses dos testes de Qui-quadrado de associação, heterogeneidade e de ajustamento.

Quadro 1. Hipóteses dos testes de Qui-quadrado de associação, heterogeneidade e de ajustamento.

Teste	Hipóteses
Teste de Qui-quadrado de independência ou associação (estudos transversal e de coorte)	H_0 : As variáveis são independentes/ não estão associadas.
	H_A : As variáveis não são independentes/estão associadas.
Teste de Qui-quadrado de comparação de proporção ou de heterogeneidade (estudo caso-controle)	H_0 : A proporção das respostas é a mesma entre os grupos.
	H_A : A proporção das respostas é diferente entre os grupos.
Teste de Qui-quadrado de ajustamento ou aderência	H_0 : As proporções observadas se ajustam às proporções esperadas.
	H_A : As proporções observadas não se ajustam às proporções esperadas.

O teste é utilizado para verificar se a frequência com que um determinado evento é observado se desvia significativamente ou não da frequência com que é esperado. Ou, ainda, para comparar a distribuição de diversos eventos em diferentes amostras, a fim de avaliar se as proporções observadas destes eventos mostram ou não diferenças significativas ou se as amostras diferem significativamente quanto às proporções desses eventos.

No teste de Qui-quadrado, aspectos como o tamanho da amostra e o número de categorias de resposta da variável, influenciam a escolha do teste estatístico (Quadro 2). Uma opção para analisar a associação entre duas variáveis qualitativas em escala de medida nominal de grupos independentes é o teste do Qui-quadrado de Pearson, que exige variáveis dicotômicas, resultando em uma tabela de contingência 2x2. Quando temos uma tabela de contingência 2x2, o teste do Qui-quadrado de Pearson tende a produzir valores significativos que são muito pequenos e podem levar ao erro do tipo I. Portanto, existe uma correção de teste, a correção de continuidade de Yates. Outra opção para o teste do Qui-

quadrado de Pearson é a estatística de Razão de Verossimilhança (*Likelihood Ratio*), que é baseada na teoria de Verossimilhança, ou seja, na razão das frequências observadas para as esperadas e é indicado quando a amostra é pequena. Caso a variável qualitativa esteja em escala de medida ordinal e apresente mais de duas categorias de resposta, a opção de teste é a de associação de tendência linear. Porém, quando a amostra é pequena e se observa percentual maior que 25% de células com tamanho menor que 5 sujeitos esperados, recomenda-se o teste Exato de Fisher.

Quadro 2. Regra para encontrar o p -valor da tabela do teste de Qui-quadrado.

Primeiro, temos que olhar o percentual de células com tamanho menor que 5 sujeitos esperados.	
Se o percentual for maior que ou igual a 25%, devemos olhar para o teste Exato de Fisher (<i>Exact Sig 2-sided</i>).	
Se o percentual for menor que 25%	E tivermos uma tabela de contingência 2x2, devemos olhar para o teste de correção de continuidade de Yates (<i>Asymp. Sig. 2-sided</i>).
	E tivermos uma tabela de contingência diferente de 2x2, devemos olhar para o teste de Qui-quadrado de Pearson (<i>Asymp. Sig. 2-sided</i>).

Teste de Qui-quadrado de associação

O teste de Qui-quadrado de associação (ou independência) é realizado em estudos com delineamento transversal ou de coorte. Neste caso, conhecemos de antemão o tamanho amostral final, mas não a quantidade de sujeitos em cada categoria do nosso desfecho. Semelhante ao teste de proporção, o teste de associação é baseado na comparação entre os valores observados e os valores esperados.

Para exemplificar, vamos verificar se a variável "Sexo" está associada com a variável "Hipertensão", no "Banco de dados 3 Resolvido.sav". No menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Insira a variável desfecho em "Coluna(s)", no caso "Hipertensão", e a variável independente em "Linha(s)", no caso "Sexo". Utilize o botão de seta para inseri-las. Clique em "Exato..." e selecione "Monte Carlo". Em seguida, clique

em "Continuar". Clique em "Estatísticas..." e em "Qui-quadrado". Esta é a opção para pedir que ele realize o teste de X^2 . Em seguida, clique em "Continuar". Clique em "Células...", marque "Observado" em "Contagens", selecione "Linha", "Coluna" e "Total" em "Porcentagens", e "Padronizado ajustado" em "Residuais". Em seguida, clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, na primeira tabela (*Case Processing Summary*), temos as frequências absoluta (N) e relativa (percentual) dos casos válidos, omissos e total. Na segunda tabela, é apresentada a tabela de contingência, neste caso, no formato 2x2. Podemos observar que 78,6% ($n = 363$) dos obesos da amostra são hipertensos, independente do sexo, e a maioria da amostra é composta pelo sexo feminino ($n = 367$, 79,4%). Dentre os hipertensos, 77,4% ($n = 281$) são do sexo feminino. Se formos observar os sexos separadamente, 86,3% ($n = 82$) dos homens são hipertensos, enquanto que 76,6% das mulheres são hipertensas.

Com base nestes dados, será que a hipertensão está associada ao sexo? Para responder esta pergunta, precisamos olhar para o p -valor (Quadro 2) que consta na terceira tabela (*Chi-Square Tests*) do arquivo de saída.

Neste exemplo, temos menos de 25% das células com tamanho menor que 5 sujeitos esperados e uma tabela de contingência 2x2. Portanto, devemos considerar o valor de p expresso pelo teste de correção de continuidade de Yates, que foi igual a 0,054. Essas proporções, em nível de associação global, não são significativamente diferentes conforme o teste de correção de continuidade de Yates. Consideramos este valor de p como sendo limítrofe para associação.

Resíduo padronizado ajustado

Além de considerarmos o p -valor do teste, precisamos avaliar os resíduos ajustados na associação das variáveis. Desta forma, verificada a associação global entre as variáveis pelo teste do Qui-quadrado, pode-se verificar se há associação local entre as categorias, olhando-se os resíduos ajustados. O resíduo de uma célula é o valor observado menos o valor esperado, dividido

por uma estimativa do seu erro padrão. O resíduo padronizado resultante é expresso em unidades de desvio padrão acima ou abaixo da média. Quanto maior for o resíduo ajustado, maior a associação entre as categorias.

O resíduo ajustado tem distribuição normal com média zero e desvio padrão igual a 1. Desta forma, caso o resíduo ajustado seja maior que 1,96, em valor absoluto, pode-se dizer que há evidências de associação significativa entre as duas categorias naquela casela (tabela de contingência). Olhando para os resíduos ajustados do exemplo, observamos que os maiores valores (positivos) indicam associação entre sexo masculino e ser hipertenso (2,1), bem como há associação entre sexo feminino e não ser hipertenso (2,1). Porém, precisamos lembrar que pelo teste de correção de continuidade de Yates, não há associação entre sexo e hipertensão ($p = 0,054$).

Os resultados das análises podem ser apresentados no formato de tabela (Tabela 1). Esta tabela deve conter todas as informações necessárias para sua compreensão. O título deve descrever o conteúdo da tabela, bem como a forma de apresentação dos dados. O corpo da tabela é composto por linhas e colunas. Nas linhas encontram-se as variáveis de estudo (variáveis dependentes) e nas colunas os grupos (variáveis independentes) e o p-valor. No rodapé, coloca-se as notas da tabela, como abreviações, nível de significância e teste estatístico utilizados.

Tabela 1. Comparação da frequência de hipertensão entre os sexos. Os números são apresentados como números absolutos e porcentagens para variáveis qualitativas. As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Sexo masculino (n = 95)	Sexo feminino (n = 367)	p-valor
HAS (n, %)	82 (86,3)	281 (76,6)	0,054 (a)

HAS, hipertensão arterial; n = número absoluto.

*Significância estatística considerada de $p \leq 0,050$; (a) Teste de correção de continuidade de Yates.

O teste de Qui-quadrado apresenta algumas limitações, conforme comentado anteriormente. Principalmente em amostras pequenas, onde o erro do valor de Qui-quadrado é alto, este deverá ser substituído pelo teste exato de Fisher quando o percentual de células com tamanho menor que 5 sujeitos esperados for igual a ou maior que 25%. Por exemplo, na tabela de teste de Qui-quadrado, no arquivo de saída, constará uma mensagem no rodapé que nos informará da existência de uma (ou mais) célula(s) com valor esperado menor que 5. Neste caso, é preferível usar o teste exato de Fisher. Outra solução, seria realizar a análise para todas as categorias exceto a categoria com valor esperado menor que 5, usando a função "Selecionar casos" em "Dados". Ou então, caso haja grande número de categorias na variável, poderia-se recodificar a variável em menor número de categorias, utilizando a função "Transformar".

Teste de Qui-quadrado de homogeneidade

O teste de Qui-quadrado de homogeneidade é utilizado nos estudos do tipo caso-controle ou em estudos em que o tamanho amostral de cada grupo foi pré-estabelecido antes do momento da coleta. Lembre-se que os comandos para a realização desse teste são os mesmos encontrados no teste de Qui-quadrado de associação. A diferença é na interpretação dos resultados.

Teste de Qui-quadrado de ajustamento

O teste de Qui-quadrado de ajustamento (ou aderência) é usado quando queremos verificar se a proporção encontrada em nossa amostra é igual a uma proporção referencial (baseado em algum livro, artigo, teoria). Para exemplificar, vamos utilizar dados fictícios de um experimento utilizando como referência a experiência de hereditariedade com ervilhas de cheiro (*Pisum sativa*) de Gregor Johan Mendel.

Em seus experimentos, Mendel utilizou apenas plantas de linhagens puras e realizou uma série de cruzamentos. Como resultado desses cruzamentos, Mendel estabeleceu a teoria genética simples que sugere que o resultado de um determinado cruzamento com ervilha seria constituído por descendentes de três

tipos, denotados por "AA", "aa" e "Aa" com probabilidades de $\frac{1}{4}$, $\frac{1}{4}$ e $\frac{1}{2}$, respectivamente. Um pesquisador realizou um experimento similar e encontrou as frequências absolutas de descendentes em 1000 ervilhas descritas no quadro 3. Mas, será que os resultados observados no experimento se ajustam aos achados encontrados pela teoria de Mendel?

Quadro 3. Contagens de descendentes relatadas em experimento com 1000 ervilhas.

Categoria	Gene	Frequência absoluta (n)
1	"AA"	194
2	"aa"	279
3	"Aa"	527
Total		1000

O primeiro passo é contruir o banco de dados no SPSS a partir dos dados fornecidos na tabela. Na planilha "Visualização da variável", crie uma variável que identifique os grupos, por exemplo, variável "Gene", e informe as categorias descritas no quadro: 1 = "AA", 2 = "aa" e 3 = "Aa", sendo esta do tipo nominal. Em seguida, crie uma variável "Peso", do tipo escala, onde iremos inserir os valores absolutos das categorias. Para a construção deste banco, vamos utilizar a função "Ponderar casos...", no menu "Dados". A função permite criarmos uma replicação simulada e, assim, não precisamos informar as 1000 linhas no banco de dados. Indicadas as variáveis do banco de dados, informe as categorias e os pesos de cada categoria na planilha de "Visualização de dados".

Com os dados inseridos, vamos ponderar os casos. Clique em "Dados", "Ponderar casos...". Na janela "Ponderar casos", selecione a opção "Ponderar casos por" e selecione, com o auxílio do botão da seta, a variável "Peso". Clique em "Ok" ou "Colar". Após, realizamos o teste de Qui-quadrado de ajustamento na seção não paramétrica. No menu "Analisar", "Testes não paramétricos", clique em "Uma amostra...". Na janela "Testes não paramétricos de uma amostra", na aba "Objetivo", selecione "Customizar análise". Na aba "Campos", insira a variável "Gene" dentro da tela "Testar campos". Na aba "Configurações", clique em "Customizar testes" e selecione a

opção “Comparar probabilidades observadas às hipotetizadas (teste Chi-Quadrado)”. Clique em “Opções...”. Selecione “Customizar probabilidade esperada” e na tabela “Probabilidades esperadas”, digite a categoria e a probabilidade esperada: categoria 1 (AA): $\frac{1}{4}$ (0,25); categoria 2 (aa): $\frac{1}{4}$ (0,25) e; categoria 3 (Aa): $\frac{1}{2}$ (0,50). Clique em “Ok” e, em seguida, em “Executar” ou “Colar”.

No resultado do teste de Qui-quadrado de ajustamento (*Hypothesis Test Summary*), a comparação com a proporção esperada, ou seja, de 25%, 25% e 50%, foi significativa ($p < 0,001$). Concluímos assim, que os valores observados diferem estatisticamente dos valores esperados. Como o próprio teste indica, rejeita-se a H_0 de que as proporções observadas se ajustam às proporções esperadas. Ao realizarmos uma análise de frequência, observa-se que as três categorias no estudo possuem os seguintes percentuais: “AA” = 19,4%, “aa” = 27,9% e “Aa” = 52,7%, ou seja, são diferentes da proporção esperada de 25%, 25% e 50%.

O SPSS ainda possibilita produzir outras análises estatísticas complementares a análise de teste do Qui-quadrado, que nos indica somente a existência ou não de associação. Estes complementos nos indicam a força e a direção desta associação.

Complementos para variáveis nominais

Para variáveis nominais é possível selecionar Coeficiente de Contingência, F_i (coeficiente) e V de Cramer, Λ (lambdas simétricos e assimétricos e tau de Goodman e Kruskal), e Coeficiente de Incerteza.

F_i e V de Cramer

O F_i e o V de Cramer são medidas do grau de associação entre duas variáveis categóricas. O F_i é uma medida de associação baseada em Qui-quadrado que envolve dividir a estatística Qui-quadrado pelo tamanho da amostra e obter a raiz quadrada do resultado, e é utilizado em tabelas de contingência 2x2. O V de Cramer é uma medida de associação baseada em Qui-quadrado e é utilizado quando uma das variáveis apresenta mais de duas categorias de resposta. Se uma das duas variáveis categóricas apresenta mais do que duas categorias, o V de Cramer é mais adequado do que o

Fi. Para ambos os testes, os valores são analisados em um intervalo entre 0 e 1, onde o valor 1 indica a máxima relação entre as variáveis e 0 a ausência de relação. Há indicações para considerar como significativos valores de $p \leq 0,010$.

Coefficiente de contingência

O Coeficiente de Contingência é uma medida de associação com base em Qui-quadrado e também é utilizada para medir a força da relação entre as variáveis, porém é mais conservador, o que leva a produzir valores inferiores ao Fi e ao V de Cramer. O valor obtido também varia entre 0 e 1, onde 0 indica nenhuma associação entre as variáveis e valores próximos a 1 indicam alto grau de associação entre as variáveis de linha e de coluna. O valor máximo possível depende do número de linhas e de colunas em uma tabela. Há indicações para considerar como significativos valores de $p \leq 0,010$.

Lambda

O Lambda é a medida do teste de Goodman e Kruskal e indica a redução proporcional no erro obtido quando valores da variável independente são usados para prever valores da variável dependente, ou seja, quando uma variável é utilizada para prever a outra. Um valor de 1 significa que uma variável prevê perfeitamente a outra, e um valor de 0 indica a ausência da capacidade de previsão.

Coefficiente de Incerteza

O Coeficiente de Incerteza também é uma medida de associação que indica a redução proporcional de erro quando os valores de uma variável são utilizados para prever valores da outra variável. Por exemplo, um valor de 0,78 indica que o conhecimento de uma variável reduz em 78% o erro ao prever valores da outra variável. O SPSS fornece ambas as versões simétrica e assimétrica do coeficiente de incerteza.

Complementos para variáveis ordinais

Para tabelas nas quais as linhas e colunas apresentam valores ordenados, podemos utilizar Gama (ordem zero para tabelas de dois fatores e condicional para tabelas de 3 a 10 fatores), Tau-b de Kendall e Tau-c de Kendall. Para categorias de coluna de predição a partir de categorias de linha, utilize d de Somers.

Gama

O Gama é uma medida de associação simétrica entre duas variáveis ordinais que varia entre -1 e 1. Valores próximos de 1 indicam um relacionamento forte entre duas variáveis; valores próximos de -1 indicam um relacionamento forte inverso entre duas variáveis; valores próximos a 0 indicam pouco ou nenhum relacionamento. Para tabelas de duas vias, gamas de ordem zero são exibidas. Para tabelas de 3 a "n" vias, gamas condicionais são exibidas.

d de Somers

O d de Somers é utilizado como medida de associação entre duas variáveis ordinais que varia de -1 a 1. Os valores próximos de 1 indicam um forte relacionamento entre as duas variáveis; valores próximos de -1 indicam um relacionamento forte inverso entre duas variáveis; valores próximos a 0 indicam pouco ou nenhum relacionamento. O d de Somers é uma extensão assimétrica de Gamma que difere apenas na inclusão do número de pares não relacionados à variável independente. Uma versão simétrica dessa estatística também é calculada.

Tau-b de Kendall

O Tau-b de Kendall é uma medida de correlação não paramétrica para duas variáveis ordinais. O sinal do coeficiente indica a direção do relacionamento e seu valor absoluto indica a intensidade, com os valores absolutos maiores indicando os relacionamentos mais fortes. Os valores possíveis variam de -1 a 1, mas um valor de -1 ou 1 pode ser obtido apenas a partir de tabelas de contingência 2x2.

Tau-c de Kendall

O Tau-c de Kendall é uma medida de associação não paramétrica entre uma variável ordinal e outra nominal. O sinal do coeficiente indica a direção do relacionamento e seu valor absoluto indica a intensidade, com os valores absolutos maiores indicando os relacionamentos mais fortes. Os valores possíveis variam de -1 a 1, mas um valor de -1 ou 1 pode ser obtido apenas a partir de tabelas de contingência 2x2.

Complementos para variáveis nominais por intervalo

Quando uma variável é categórica e a outra é quantitativa, selecione a opção Eta. A variável categórica deve ser codificada de forma numérica.

Eta

A opção Eta é recomendada para analisar a relação entre uma variável nominal e uma intervalar (quantitativa). Trata-se de uma medida de associação que varia de 0 a 1, com 0 indicando nenhuma associação entre as variáveis de linha e de coluna e valores próximos de 1 indicando um alto grau de associação. O Eta é apropriado para uma variável dependente medida em uma escala de intervalo e uma variável independente com um número limitado de categorias. Dois valores de Eta são obtidos: um aborda a variável de linha como a variável de intervalo, e o outro aborda a variável de coluna como a variável de intervalo.

Estatísticas de Cochran e Mantel-Haenszel

As estatísticas de Cochran e Mantel-Haenszel podem ser utilizadas para testar a independência entre uma variável de fator dicotômico e uma variável de resposta dicotômica, condicional aos padrões de covariáveis definidos por uma ou mais variáveis de camada (controle). Enquanto outras estatísticas são calculadas camada por camada, as estatísticas de Cochran e Mantel-Haenszel são calculadas uma vez para todas as camadas.

Variáveis ordinais com mais de duas categorias

Teste U de Mann-Whitney

O teste U de Mann-Whitney é uma opção interessante quando o pesquisador pretende comparar o comportamento de uma variável qualitativa ordinal com mais de duas categorias de resposta entre dois grupos independentes. Para exemplificar, no " Banco de dados 3 Resolvido.sav", vamos verificar se existe diferença entre os sexos (variável "Sexo") quanto ao grau de obesidade (variável "Grau de obesidade").

No menu "Analisar", "Testes não paramétricos", "Caixas de diálogo legadas", clique em "2 amostras independentes...". Na caixa de diálogo "Testes de duas amostras independentes", insira na "Lista de variável teste" a variável qualitativa em escala ordinal "Grau de obesidade" e, em "Variável de agrupamento", insira a variável qualitativa dicotômica "Sexo" que divide os grupos de comparação. Clique na opção "Definir grupos...". Na janela "Duas amostras independentes", insira os códigos numéricos que definem os grupos: 1 = sexo masculino e 2 = sexo feminino. Clique em "Continuar". Retornando a caixa de diálogo principal, certifique-se de que em "Tipo de teste", o teste "U de Mann-Whitney" esteja selecionado. Clique em "Exato...". Selecione "Monte Carlo". Em seguida, clique em "Continuar" e, por fim, em "Ok" ou "Colar".

No arquivo de saída, na primeira tabela (*Ranks*), é apresentada a frequência absoluta (N) em cada grupo, a média dos postos (*Mean Rank*) e a soma dos postos (*Sum of Ranks*). Analisando a média dos postos é possível observar que o sexo masculino apresenta valor superior ao sexo feminino, o que indica que o sexo masculino contém as posições mais elevadas. Os resultados produzidos pelo teste U de Mann-Whitney (segunda tabela) indicam que existe uma diferença significativa entre os sexos quanto à distribuição das categorias da variável "Grau de obesidade" ($p < 0,001$). Realizando uma análise de frequências, observa-se que os sujeitos do sexo masculino apresentam maior frequência de obesidade grau III (85,3%) do que o sexo feminino (58,6%, $p < 0,001$).

Variáveis com duas categorias medidas em dois momentos

Testes de McNemar e de Wilcoxon

Para comparar a distribuição de frequência de variáveis qualitativas medidas em dois momentos, ou seja, grupos dependentes, utiliza-se o teste de McNemar ou o teste de Wilcoxon. Porém, quando se tem amostras pequenas (< 25 casos) é recomendado utilizar o teste Binomial. Em ambos, é exigido que a variável seja dicotômica e apresente apenas duas categorias de resposta.

O teste de McNemar é um teste não paramétrico utilizado para analisar proporções entre duas variáveis dicotômicas relacionadas (pareadas). É particularmente aplicado aos estudos “antes e depois”, onde o próprio sujeito é utilizado como seu próprio controle. Para o teste é exigido que a variável seja dicotômica (nominal). O teste assume como H_0 que não existe diferença antes e após o tratamento e, como H_A , que existe diferença antes e após o tratamento. Para tabelas quadradas maiores, o teste de simetria de McNemar-Bowker é relatado. No teste de Wilcoxon, é possível dizer qual dos pares é “maior”, isto é, determinar o sentido da diferença dentro do par, e ordenar essas diferenças no sentido de seu valor absoluto. Para o teste é exigido que a variável apresente apenas duas categorias de resposta.

Para exemplificar, vamos utilizar o “Banco de dados 4.sav” (disponível em <https://bit.ly/bancosdedados>). Este banco de dados contém a avaliação de sujeitos com obesidade no momento pré-operatório (1) e em dois momentos após a cirurgia bariátrica (2 e 3). Portanto, temos três medidas distintas, uma para cada ponto no tempo (pareadas), para as variáveis massa corporal, circunferência da cintura, medidas pressóricas, hipertensão e atividade física avaliada pelo IPAQ versão curta. Além disso, temos descritos sexo, técnica cirúrgica e estatura dos participantes da pesquisa. Os mesmos sujeitos também foram avaliados quanto a alterações de sensibilidade nos membros inferiores (neuropatia periférica) através do *Michigan Neuropathy Screening Instrument* (MNSI), por dois examinadores. Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Para exemplificar, vamos comparar a hipertensão avaliada nos tempos 1 (pré-operatório) e 2 (pós-operatório). Existem duas formas de análise. Na primeira, é possível gerar os testes de McNemar e de Wilcoxon no mesmo local no SPSS. No menu “Analisar”, “Testes não paramétricos”, “Caixa de diálogo legadas”, clique em “2 amostras relacionadas...”. Na caixa de diálogo “Testes de duas amostras relacionadas”, selecione e insira as duas variáveis que iremos comparar (“HAS 1” e “HAS 2”) em “Pares de teste”. Em “Tipo de teste”, selecione o teste desejado (“McNemar” e “Wilcoxon”). Neste campo, também é possível realizar o teste Binomial (< 25 casos), selecionando “Sinal”. Clique

em "Exato..." e selecione "Monte Carlo". Em seguida, clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No teste de Wilcoxon, o primeiro quadro (*Ranks*) é apresentado o número de elementos compostos positivos, negativos e iguais, ou seja, aqueles que se tornaram hipertensos, não hipertensos e os que se mostraram inalterados. Também temos a média das posições (*Mean Rank*) e a soma das posições (*Sum of Ranks*) para as posições positivas e negativas. No segundo quadro (*Test Statistics*) é apresentado o escore z (Z) e o p-valor (*Asymp. Sig 2-tailed*) associado ao teste de Wilcoxon.

No mesmo arquivo de saída, logo abaixo, temos os resultados do teste de McNemar. No primeiro quadro, são apresentadas a distribuição de frequência absoluta do cruzamento das variáveis nos momentos pré- (1) e pós-cirúrgicos (2) semelhante a tabela de contingência 2x2, porém sem os valores totais. No segundo quadro (*Test Statistics*), é apresentado o número de elementos (N), a estatística associada ao teste do Qui-quadrado (*Chi-Square*) com correção de continuidade e o p-valor referente ao teste de McNemar (*Asymp. Sig.*). Em ambos os testes, foi significativa a redução no número de hipertensos ($p < 0,001$).

Na segunda opção, gera-se somente o resultado do teste de McNemar. No entanto, obtém-se uma tabela de contingência mais completa. No menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Na janela "Tabulações cruzadas", insira a variável "HAS 1" em "Linhas" e a variável "HAS 2" em "Coluna(s)". Utilize o botão das setas para inseri-las. Clique em "Exato..." e selecione "Monte Carlo". Em seguida, clique em "Continuar". Clique em "Estatísticas..." e selecione "McNemar". Em seguida, clique em "Continuar". Clique em "Células...", marque "Observado" em "Contagens", selecione "Linha", "Coluna" e "Total" em "Porcentagens" e, "Padronizado ajustado" em "Residuais". Em seguida, clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, é apresentada a tabela de contingência e o quadro contendo os resultados da estatística do teste. Neste caminho não é produzida a correção de continuidade e isto pode produzir p-valores um pouco diferentes. Podemos observar que, dos

147 sujeitos (75,4%) com hipertensão no pré-operatório (tempo 1), 60 (40,8%) continuaram tendo hipertensão no pós-operatório (tempo 2) e, 87 (59,2%) tiveram a resolução da sua hipertensão. Essa redução no número de hipertensos foi significativa ($p < 0,001$). Note que o p-valor encontrado foi “,000”, ou seja, $p < 0,001$. Além disso, podemos apresentar os resultados da análise em uma tabela (Tabela 2).

Tabela 2. Avaliação da hipertensão pré e pós-operatória em sujeitos obesos submetidos à cirurgia bariátrica. Os números são apresentados como números absolutos e porcentagens para variáveis qualitativas. As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Pré-operatório (n = 195)	Pós-operatório (n = 195)	p-valor
HAS (n, %)	147 (75,4)	70 (35,9)	<0,001* (a)

HAS, hipertensão arterial; n = número absoluto.

*Significância estatística considerada de $p \leq 0,050$; (a) Teste de McNemar.

Variáveis em escala de medida ordinal

Existem situações em que as análises incluem duas variáveis qualitativas, sendo uma representada pela variável dependente em escala de medida ordinal e com mais de três categorias de resposta, e a outra variável independente em escala de medida ordinal ou nominal, porém com três ou mais grupos de comparação. Nestes casos, o teste de Qui-quadrado de associação não é o mais recomendado, sendo indicado os testes de Kruskal-Wallis ou de Friedman.

A escolha do teste depende da forma como os grupos de comparação estão organizados: em situações em que os grupos são independentes, deve-se utilizar o teste de Kruskal-Wallis; já onde um único grupo tem três ou mais medidas repetidas, deve-se utilizar o teste de Friedman. Como veremos nos capítulos 7 e 8, o teste de Kruskal-Wallis é indicado como a opção não paramétrica para o teste ANOVA, e o teste de Friedman é indicado como a opção não paramétrica para ANOVA de medidas repetidas (ANOVA-MR) ou para Equações de Estimativas Generalizadas

(GEE). Mas, é importante lembrar que neste capítulo estamos trabalhando com variáveis categóricas.

Teste de Kruskal-Wallis

Para exemplificar o uso do teste de Kruskal-Wallis, no “Banco de dados 3 Resolvido.sav”, vamos utilizar as variáveis “Grau de obesidade” (variável independente que apresenta três categorias ordinais) e a variável “Classificação IPAQ” (variável dependente que apresenta cinco categorias ordinais). Vamos supor que os três grupos de grau de obesidade (graus I, II e III) tiveram seus valores de atividade física categorizados em cinco níveis do IPAQ (sedentário, insuficientemente ativo B e A, ativo e muito ativo). O objetivo será verificar se a distribuição das categorias de atividade física avaliadas pelo IPAQ difere entre os graus de obesidade.

No menu “Analisar”, “Testes não paramétricos”, “Caixas de diálogo legadas”, clique em “K amostras independentes...”. Na caixa de diálogo “Testes para diversas amostras independentes”, insira a variável dependente “Classificação IPAQ” em “Lista de variável de teste”, e a variável independente “Grau de obesidade” em “Variável de agrupamento”, com o auxílio do botão da seta. Clique em “Definir faixa”, para atribuir os valores. Insira em “Mínimo” o código que representa a menor categoria (no caso, 1 = Grau I) e, em “Máximo”, o código que representa a maior categoria da variável independente (no caso, 3 = Grau III). Clique em “Continuar”. Em “Exato...” selecione “Monte Carlo”. Em seguida, clique em “Continuar”. Retornando a caixa de diálogo principal, certifique-se de que em “Tipo de teste” a opção “H de Kruskal-Wallis” esteja selecionada. Por fim, clique em “Ok” ou “Colar”.

No arquivo de saída, no primeiro quadro (*Ranks*), é apresentada a média das posições da variável “Classificação IPAQ” para cada grau de obesidade. Note que as médias são muito similares entre os três grupos. No segundo quadro (*Test Statistics*), é apresentado o valor de $p = 0,850$ que indica que não existe diferença significativa entre os graus de obesidade no nível de atividade física avaliada pelo IPAQ.

Supondo que houvesse uma diferença estatisticamente significativa entre os graus de obesidade no nível de atividade

física, apesar de se saber que existe essa diferença, o teste de Kruskal-Wallis não indica entre quais grupos a diferença é significativa. Neste caso, é necessária uma análise *post hoc*. Uma opção é realizar comparações “par a par” pelo teste U de Mann-Whitney. Porém, as múltiplas comparações deverão aumentar as chances de erro do tipo I, sendo necessário dividir o valor crítico de “p” de 0,050 pelo número de comparações. Assim, dividindo o valor de 0,05 por 3, neste exemplo, temos agora um valor crítico de “p” para rejeitar a H_0 de 0,016, ou seja, para considerar uma diferença significativa entre os grupos será necessário um p-valor maior ou igual a 0,016. Deve-se observar que testes *post hoc* são necessários para identificar quais dos pares do grupo diferem.

Teste de Friedman

Quando os dados de k amostras correspondentes (pareadas) se apresentam em escala ordinal (ranqueamento), isto é, o número de casos é o mesmo para cada uma delas, pode-se utilizar o teste de Friedman para comprovar que as k amostras foram extraídas da mesma população. O teste de Friedman é uma alternativa não paramétrica para testar diferenças entre duas ou mais amostras dependentes. É considerado uma extensão do teste Binomial, para duas amostras dependentes, quando existem mais de duas amostras dependentes. Se $k = 2$, o teste de Friedman fornece um resultado equivalente ao teste Binomial.

Para exemplificar, vamos considerar que um único grupo de sujeitos obesos teve a atividade física avaliada em três momentos distintos, ou seja, queremos verificar se as distribuições das categorias da classificação do IPAQ se modifica ao longo do tempo em toda a amostra. Neste caso, é recomendado o teste de Friedman. Para isso, vamos utilizar o “Banco de dados 4.sav”, que apresenta medidas repetidas para a variável “Classificação IPAQ”.

No menu “Analisar”, “Testes não paramétricos”, “Caixas de diálogo legadas”, clique em “K amostras relacionadas...”. Na janela “Testes para várias amostras relacionadas”, insira em “Variáveis de teste”, com o auxílio do botão da seta, as três variáveis que representam os momentos de comparação (“IPAQ 1”, “IPAQ 2” e “IPAQ 3”). Certifique-se de que a opção “Friedman” esteja

selecionada em "Tipo de teste". Clique na opção "Exato..." e selecione "Monte Carlo". Em seguida, clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, no primeiro quadro (*Ranks*), é apresentada a média das posições para cada momento avaliado. Note que as duas últimas possuem as maiores médias das posições, que são similares entre si. No quadro seguinte (*Test Statistics*), é apresentado o *p*-valor associado ao teste de Friedman. Com $p < 0,001$, pode-se afirmar que existe diferença significativa na distribuição de frequência das categorias de atividade física avaliada pelo IPAQ entre os momentos pré- (1) e pós-operatórios (2 e 3). Porém, semelhante ao teste de Kruskal-Wallis, o teste de Friedman não indica entre quais momentos a diferença é significativa, sendo necessário a aplicação de outro teste (*post hoc*) para localizar as diferenças. Para o teste de Friedman, as comparações "par a par" devem ser feitas pelo teste de Wilcoxon.

Variáveis em escala de medida nominal ou ordinal dicotomizadas

Teste Q de Cochran

O teste de McNemar para duas amostras pode ser estendido para situações que envolvam mais de duas amostras pareadas. Esta extensão é denominada de teste Q de Cochran. É idêntico ao teste de Friedman, mas é aplicável quando todas as respostas são dicotomizadas, em nível nominal ou ordinal. O teste permite comparar três ou mais conjuntos correspondentes de frequências ou proporções para verificar se estas se diferem significativamente entre si. O teste Q de Cochran testa a H_0 de que diversas variáveis dicotômicas relacionadas têm a mesma média. As variáveis são medidas no mesmo indivíduo ou em indivíduos correspondentes.

O teste Q de Cochran é realizado no mesmo local dos testes de Friedman e W de Kendall. Para exemplificar, no "Banco de dados 4.sav", vamos comparar as três medidas de hipertensão ("HAS 1", "HAS 2" e "HAS 3"). No menu "Analisar", "Teste não paramétricos", "Caixas de diálogo legadas", clique em "K amostras relacionadas...". Na janela "Testes para várias amostras relacionadas", insira as variáveis em "Variáveis de teste". Utilize o

botão da seta para inseri-las. Marque a opção “Q de Cochran”. Clique em “Exato...” e selecione “Monte Carlo”. Em seguida, clique em “Continuar”. Por fim, clique em “Ok” ou “Colar”.

Na primeira tabela (*Frequencies*), na primeira coluna, são apresentados, em valor absoluto, os sujeitos sem hipertensão (código 0) e, na segunda coluna, os sujeitos com hipertensão (código 1), avaliados nos três períodos (“HAS 1”, “HAS 2” e “HAS 3”). Na segunda tabela (*Test Statistics*), temos o p-valor associado ao teste Q de Cochran. Considerando que o teste mostrou um p-valor inferior a 0,001, podemos dizer que há diferença estatisticamente significativa na proporção de sujeitos hipertensos entre os três períodos (pré- e pós-operatórios). Como nos outros testes, o teste Q de Cochran nos force o p-valor associado à análise global. A partir desta análise, teremos que fazer um teste comparando cada par de períodos (HAS 1 x HAS 2, HAS 1 x HAS 3 e HAS 2 x HAS 3), ou seja, uma análise *post hoc*. No caso do teste Q de Cochran, é recomendado análise complementar pelo teste de McNemar.

Análise de concordância

Teste W de Kendall

Na mesma janela “Testes para várias amostras relacionadas”, existe a opção de teste W de Kendall. Trata-se de uma normalização do teste de Friedman. Sua interpretação é feita como a do coeficiente de concordância (Índice Kappa), que é uma medida de acordo entre os avaliadores quando as classificações forem ordinais e apresenta três ou mais níveis de classificação. Cada caso é um avaliador, e cada variável é um item ou caso que está sendo avaliado. Para cada variável, a soma dos ranqueamentos é calculada. O teste W de Kendall varia entre 0 (nenhuma concordância) e 1 (concordância total). Quanto maior o valor de Kendall, mais forte é a concordância.

Índice Kappa de Cohen

Existem situações em que há a necessidade de avaliarmos a concordância entre duas variáveis, tendo elas sido obtidas por diferentes avaliadores ou avaliadas por meio de dois métodos diferentes. Por exemplo, uma imagem de radiografia do tórax pode ser avaliada por dois radiologistas em relação à presença

(ou não) de doença pulmonar obstrutiva crônica (DPOC) e o pesquisador precisa saber qual o grau de concordância entre os avaliadores, pela subjetividade da interpretação de cada avaliador. Ou ainda, a medida de força muscular na DPOC pode ser avaliada por meio de dinamômetros de cabo e dinamômetros eletrônicos, e o pesquisador deseja saber o grau de concordância entre os métodos.

Quando a variável a ser analisada apresenta-se como categórica, podendo ter duas ou mais categorias de resposta, o grau de concordância pode ser obtido pelo coeficiente de concordância Kappa. O coeficiente Kappa é calculado baseado na comparação entre a proporção de concordância observada e a esperada ao acaso. O kappa é baseado em uma tabela quadrada na qual os valores da linha e da coluna representam a mesma escala. No SPSS, é necessário que as duas variáveis estejam lado a lado, com um indivíduo em cada linha. O Kappa não será calculado se o tipo de armazenamento de dados (sequência de caracteres ou numérico) não for o mesmo para as duas variáveis. Para a variável de sequência de caracteres, ambas as variáveis devem ter o mesmo comprimento definido. Qualquer célula que possuir valores observados para uma variável, mas não para outra, é designada a uma contagem de 0. Lembre-se que o Kappa informa somente o quanto os métodos concordam, mas não informa o quanto é eficiente para detectar os doentes ou não doentes. De acordo com Landis & Koch, os valores de Kappa podem ser classificados conforme o quadro 4.

Quadro 4. Parâmetros sugeridos por Landis & Koch para interpretar o índice de concordância de Kappa.

Valor de Kappa	Interpretação da concordância
0,00	Sem concordância
0,00-0,19	Discreta
0,20-0,39	Regular
0,40-0,59	Moderada
0,60-0,79	Substancial
0,80-1,00	Quase perfeita

Para realizar a análise no SPSS, vamos avaliar o grau de concordância dos dois avaliadores para o instrumento MNSI, do "Banco de dados 4.sav". No menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Na janela "Tabulações cruzadas", selecione as variáveis categóricas em que será avaliada a concordância, no caso, a avaliação do instrumento MNSI pelos dois avaliadores ("MNSI 1" e "MNSI 2"). Insira variável "MNSI 2" em "Coluna(s)" e a outra variável "MNSI 1" em "Linha(s)". Clique na opção "Exato..." e selecione "Monte Carlo". Em seguida, clique em "Continuar". Clique em "Estatísticas..." e selecione "Capa". Em seguida, clique em "Continuar". Clique em "Células...", marque "Observado" nas contagens, "Total" nas porcentagens e marque "Padronizado ajustado" nos resíduos. Clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, na primeira tabela (*Case Processing Summary*), estão apresentados os números absoluto (N) e relativo (percentual) de casos válidos, omissos e total. Na segunda tabela encontramos a tabela cruzada das respostas atribuídas por cada examinador e, na terceira tabela (*Symmetric Measures*), o resultado do índice Kappa. O índice de Kappa, neste exemplo, pode ser interpretado como uma concordância substancial entre os avaliadores (0,740), sendo significativa esta concordância ($p < 0,001$).

Referências

Field A. Descobrimo a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Chagas, E.F.B. Módulo 3: Estatística Analítica II - Análise de dados categóricos e teste diagnóstico no SPSS. 2016.

Exercícios sugeridos

1. Com base no "Banco de dados 1 Resolvido.sav", responda:

a) Existe diferença na proporção de sexo entre os alunos?

b) Existe diferença na proporção do nível de pós-graduação entre os alunos?

c) Existe diferença no nível de conhecimento em estatística entre os alunos de mestrado e doutorado?

d) Existe diferença no nível de conhecimento em epidemiologia entre os alunos de mestrado e doutorado?

e) Existe diferença na experiência do uso do SPSS entre os alunos de mestrado e doutorado?

2. Com base no "Banco de dados 4.sav", responda:

a) Comparando-se as técnicas cirúrgicas (bypass gástrico, RYGB, e gastrectomia vertical, SG), há alguma diferença na hipertensão avaliada no tempo 1 (pré-operatório)? E, na avaliada no tempo 2 (pós-operatório)? E, no tempo 3 (pós-operatório)? Construa a tabela e interprete os resultados das frequências encontradas. Nota: temos três variáveis de desfecho: hipertensão pré (HAS 1) e pós-operatórias (HAS 2 e HAS 3).

b) E, separando-se as técnicas cirúrgicas, a frequência de hipertensos diminui ao longo do tempo, ou seja, pré- (1) e pós-operatórios (2 e 3) no by-pass gástrico (RYGB)? E na gastrectomia vertical (SG)? Construa as tabelas e interprete os resultados das frequências encontradas. Nota: temos uma única variável de desfecho (hipertensão) avaliada ao longo do tempo (pré- e pós-operatórios).

7 Desfechos quantitativos: amostras independentes

Caroline Darski

Cristiana Kuhl

Paula Barros Terraciano

Otto Henrique Nienov

Comparação de amostras

Pensemos que um fabricante de lâmpadas para microscópio óptico queira trocar um dos componentes na expectativa de obter maior durabilidade e potência. Antes de iniciar a produção em grandes quantidades, é preciso que o fabricante tenha informações suficientes que demonstrem que o novo componente realmente fará com que a lâmpada apresente maior durabilidade e potência. Pensando nisso, o fabricante reuniu, aleatoriamente, 15 unidades da lâmpada produzidas com o componente antigo (lâmpada A) e 15 unidades da lâmpada produzidas com o novo componente (lâmpada B). Em seguida, registrou a durabilidade e a potência de cada uma delas em uma planilha e calculou as respectivas médias. Comparando os dois modelos de lâmpadas, o fabricante identificou que a lâmpada B apresentou maiores durabilidade (38.500 horas *versus* 38.000 horas, respectivamente) e potência (30W *versus* 25W, respectivamente) médias do que as registradas para a lâmpada A. Essa informação é suficiente para concluirmos que a nova lâmpada tem melhor desempenho?

A resposta é não! Lembre-se que o experimento demonstrou que, somente naquela amostra, a lâmpada B apresentou um melhor desempenho do que a lâmpada A. E, se fossemos repetir o experimento em uma nova amostra aleatória, com outras 15 unidades de cada lâmpada, será que obteríamos o mesmo resultado? Como poderíamos verificar se realmente existe diferença significativa entre a durabilidade e a potência dos dois modelos de lâmpadas?

Através dos testes de hipóteses, que nos fornecem subsídios para verificar se uma hipótese é verdadeira ou não. No exemplo do fabricante, seria necessário utilizar um teste de hipóteses para testar se existem diferenças significativas na durabilidade e na potência entre os dois modelos de lâmpadas. Se o fabricante assumir que a duração e a potência médias das lâmpadas apresentam uma distribuição normal, utilizaríamos testes paramétricos; se o fabricante verificar que a duração e a potência médias das lâmpadas apresentam uma distribuição não normal, empregariamos testes não paramétricos e expressariamos os valores de durabilidade e potência das lâmpadas em mediana. Note que no exemplo, estamos nos referindo às variáveis quantitativas. Por isso, antes de falarmos sobre teste de normalidade e testes paramétricos e não paramétricos, vamos lembrar o que são variáveis quantitativas. As variáveis quantitativas são aquelas que podem ser mensuradas através de escalas quantitativas, isto é, escalas que têm unidades de medida. Podem ser classificadas em discretas (quando assumem apenas valores pertencentes a um conjunto finito ou enumerável de números) ou contínuas (quando os possíveis valores estão dentro de um intervalo, aberto ou fechado, dos números reais).

Testes paramétricos e não paramétricos

A principal diferença entre os testes paramétricos e não paramétricos está na distribuição da variável de estudo. Quando é possível assumir a normalidade de distribuição da variável, podem ser utilizados testes paramétricos. Caso contrário, pode-se transformar a variável de modo a obter a normalidade (logaritmo, raiz quadrada, exponencial, etc.). No SPSS, na opção "Calcular variável..." do menu "Transformar", é possível obter o logaritmo decimal ("Log10") e exponencial ("Exp") através da função "Aritmética". Se, ainda assim, a distribuição obtida não for normal, utilizam-se testes não paramétricos.

A escolha adequada do teste (paramétrico ou não paramétrico) a ser utilizado é de fundamental importância, pois conclusões erradas podem ser obtidas se optarmos pelo teste inadequado. Além disso, a apresentação dos dados deve estar de

acordo com a suposição de distribuição da variável: se distribuição normal, utilizamos média e desvio padrão; se distribuição não normal, expressamos os dados em mediana e quartis ou valores mínimo e máximo. Na comparação de grupos, além da distribuição da variável estudada (em cada um dos grupos), a escolha do teste estatístico também dependerá do número de grupos a serem comparados e do tipo de amostras (independentes ou pareadas). Amostras (ou grupos) são consideradas independentes quando não há relação entre os sujeitos, ou seja, quando se caracterizam por indivíduos diferentes e não pareados. Já amostras (ou grupos) pareados correspondem àqueles que estão relacionados, ou seja, correspondem aos mesmos indivíduos ou a um controle emparelhado (Figura 1).

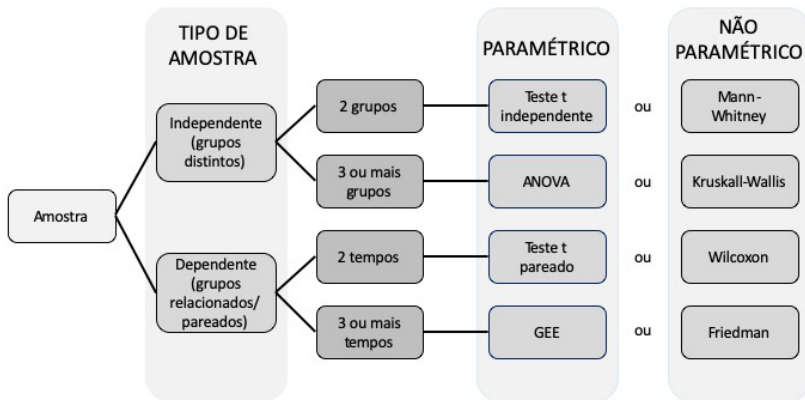


Figura 1. Testes paramétricos e não paramétricos para amostras independentes e pareadas (Mancuso e cols., 2017).

Os testes paramétricos são utilizados para comparar médias, supondo-se a normalidade da variável estudada dentro de cada grupo. Assim, devemos primeiramente verificar a hipótese de normalidade da distribuição, para então decidir qual análise estatística utilizar na comparação dos grupos. Os testes não paramétricos não fazem suposições sobre os parâmetros da distribuição, ou seja, não assumem uma distribuição normal da variável de estudo. São os mais apropriados quando não se conhece a distribuição dos dados na população.

Desta forma, os testes não paramétricos são menos exigentes, pois dispensam a suposição de normalidade e, logo, são mais eficientes do que os testes paramétricos quando os dados não possuem distribuição normal. E, quando normal, sua eficiência é levemente inferior. Também são úteis quando, na comparação, a distribuição é normal em alguns grupos e assimétrica em outros. Por isso, são testes de aplicação mais ampla do que os testes paramétricos. Geralmente, as probabilidades resultantes são exatas, com exceção de casos onde as amostras são grandes. Além disso, são mais fáceis de aplicar e exigem menor número de comandos.

Uma das grandes vantagens dos testes paramétricos é a possibilidade de testar interações, com fatores fixos, aleatórios e covariáveis. Em contrapartida, a obtenção, utilização e interpretação das tabelas do arquivo de saída são, em geral, mais complexas. Os testes não paramétricos não permitem testar interações, o que restringe seu uso em modelos mais complexos e, também, não levam em consideração a magnitude dos dados, pois, na maior parte das vezes, os dados são transformados em ordens ou sinais, ocasionando muitas vezes um “desperdício” de informações. Quando todas as exigências do modelo estatístico são satisfeitas, o teste paramétrico é mais poderoso. Portanto, para se obter a mesma eficiência no teste não paramétrico é necessário um tamanho amostral maior do que seria necessário com um teste paramétrico.

Definindo a normalidade

A distribuição de probabilidades de uma variável é uma função que descreve a chance com que uma variável pode assumir certos valores ao longo de um subconjunto de possíveis valores. A normalidade de distribuição da variável pode ser verificada através da: 1) inspeção visual utilizando um histograma ou um gráfico quantil-quantil (Q-Q Plot). O Q-Q Plot é um gráfico de dispersão dos quantis (pontos estabelecidos em intervalos regulares que dividem os dados ordenados em “q” subconjuntos de dados de dimensão essencialmente igual) calculados na amostra (quantis amostrais) contra os quantis da distribuição normal (quantis teóricos), sendo útil para verificar a adequação de distribuição de frequência dos dados à

uma distribuição de probabilidades. Quanto mais próximos os pontos estão de uma linha reta, mais próximos estão da distribuição normal; 2) comparação entre os valores de média e mediana e; 3) observação dos valores de assimetria e curtose. Lembre-se que a assimetria é uma medida da distorção da distribuição de probabilidade de uma variável aleatória sobre sua média e, a curtose, informa a altura da distribuição dos dados em relação a uma curva de Gauss. Se ambas forem zero (0), os dados são perfeitamente simétricos. No entanto, essas formas de avaliar a distribuição são consideradas subjetivas.

Para testar a distribuição normal dos dados e evitar a subjetividade, diversos testes estatísticos podem ser utilizados, entre eles o teste de Kolmogorov-Smirnov e o teste de Shapiro-Wilk. Assumimos como H_0 que a distribuição da variável não é assimétrica, enquanto que a H_A assume que a distribuição da variável é assimétrica. O teste de Shapiro-Wilk é considerado como de melhor potência, mas o SPSS nos fornece ambos os testes. Se a significância do teste for superior a 0,050 ($p > 0,050$), as variáveis seguem uma distribuição normal e, logo, aplicam-se testes paramétricos e utiliza-se média e desvio padrão. Se a significância do teste for inferior ou igual a 0,050 ($p \leq 0,050$), as variáveis seguem uma distribuição não normal e, por conseguinte, expressa-se em mediana e percentis ou valores mínimo e máximo e utiliza-se testes não paramétricos.

Atenção! É recomendada a realização do teste de normalidade somente quando, na variável estudada, o tamanho amostral de cada grupo for maior que 15 sujeitos. Caso contrário, deve-se proceder com análises não-paramétricas.

Teste de normalidade para uma amostra

Mesmo quando queremos caracterizar uma amostra, sem considerar grupos, é necessário verificar a normalidade de distribuição das variáveis para aquela amostra. Para exemplificar, vamos verificar a normalidade das variáveis "Idade" e "Massa corporal" do "Banco de dados 2.sav" (disponível em <https://bit.ly/bancosdedados>). O banco de dados apresenta dados de sujeitos com obesidade. Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

No menu "Analisar", "Estatísticas descritivas, clique em "Explorar". Na janela "Explorar", selecione, na tela da esquerda, as variáveis "Idade" e "Massa corporal" e as insira no campo "Lista dependente" com auxílio do botão da seta. Clique em "Estatísticas..." e marque as opções "Descritivos" e "Percentis". Clique em "Continuar". Clique em "Diagramas..." e selecione as opções "Histograma" e "Diagramas de normalidade com testes". Clique em "Continuar". Importante! Caso haja valores omissos, clique em "Opções" e selecione a opção "Excluir casos por par". Quando incluímos, na "Lista dependente", mais de uma variável e uma destas apresenta *missing value*, a normalidade será testada sem incluir os casos omissos em ambas as variáveis, mesmo nas que não apresentam dado faltante. Ou seja, o SPSS considera o valor omissos em todas as variáveis de estudo. Clique em "Continuar". Em seguida, clique em "Ok" ou "Colar".

Antes de verificarmos os testes de normalidade, no quadro das análises descritivas (*Descriptives*), compare os valores de média e mediana e observe os valores de assimetria e curtose. Também verifique os histogramas e a curva de normalidade. É possível afirmar que essas variáveis não possuem uma distribuição normal? Talvez sim, mas lembre-se que essas formas de verificar a normalidade são subjetivas. Para evitar essa subjetividade, vamos checar os valores dos testes de normalidade.

Conforme comentado anteriormente, o SPSS realiza dois testes de normalidade: o de Kolmogorov-Smirnov e o de Shapiro-Wilk, sendo este último considerado mais poderoso. Portanto, no quadro dos testes (*Tests of Normality*), vamos utilizar os valores do teste de Shapiro-Wilk. Lembre-se que, se a significância do teste for superior a 0,050, as variáveis de estudo apresentam uma distribuição normal e, se a significância do teste for inferior ou igual a 0,050, as variáveis de estudo seguem uma distribuição não normal. Neste caso, ambas as variáveis ("Idade" e "Massa corporal") apresentam distribuição não normal ($p < 0,001$). Logo, podemos afirmar que, a mediana de idade foi de 37 (31-45) anos e a mediana da massa corporal foi de 111,1 (100,5-126,4) kg, onde temos "mediana (percentis 25-75)".

Teste de normalidade para duas ou mais amostras

Quando comparamos grupos, como no exemplo do fabricante de lâmpadas para microscópio óptico, deve-se testar a normalidade em ambos os grupos. Para exemplificar, vamos verificar o comportamento da variável "Idade" entre as categorias da variável "Sexo" no "Banco de dados 2.sav". No menu "Analisar", "Estatísticas descritivas", clique em "Explorar...". Lembre-se que é possível redefinir as opções de teste clicando em "Redefinir". Em seguida, na janela "Explorar", selecione na tela da esquerda e insira a variável dependente ("Idade") na "Lista dependente", e a variável independente ("Sexo", que corresponde aos grupos) na "Lista de fator" com auxílio do botão da seta. Clique em "Estatísticas..." e marque as opções "Descritivos" e "Percentis". Clique em "Continuar". Clique em "Diagramas..." e selecione as opções "Histograma" e "Diagramas de normalidade com testes". Clique em "Continuar". Em "Opções", selecione "Excluir casos por par" e clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, temos as tabelas com as análises descritivas (*Descriptives*), os percentis (*Percentiles*), o teste de normalidade (*Tests of Normality*) e os gráficos (*Histogram*) para cada grupo, ou seja, para o sexo masculino e para o sexo feminino. Observando os valores de teste de Shapiro-Wilk, conclui-se que em ambos os grupos a variável apresenta distribuição não normal, com medianas de idade no sexo masculino de 37,5 (31,0-44,0) anos e no sexo feminino de 37,0 (31,0-45,0) anos.

Se fôssemos aplicar um teste estatístico para avaliar se as medianas de idade diferem entre os sexos, o teste de escolha seria não paramétrico neste exemplo. O teste paramétrico seria aplicado se em ambos os grupos a variável apresentasse distribuição normal, ou seja, se a significância do teste for superior a 0,050 e, logo, expressaríamos a idade em média e desvio padrão. Caso a variável apresentasse distribuição normal em um grupo e não normal noutro, consideraríamos teste não paramétrico, expressando, em ambos os grupos, as medidas em mediana e quartis ou valores mínimo e máximo.

Testes para amostras independentes

Neste item, são abordados os testes paramétricos e não paramétricos para amostras independentes. Os testes indicados para amostras relacionadas serão abordados no capítulo 8.

As amostras (ou grupos) são consideradas independentes quando não há relação entre os sujeitos, ou seja, quando se caracterizam por indivíduos diferentes (grupos distintos) e não pareados.

Atenção! Antes de decidir qual teste utilizar (paramétrico ou não paramétrico) é necessário verificar a normalidade de distribuição da variável pelo teste de Shapiro-Wilk.

Teste t de Student

O teste t de Student para amostras independentes compara as médias de uma variável quantitativa paramétrica (variável dependente) em dois grupos independentes (variável categórica independente, também chamada de fator, que apresenta duas categorias). Ou seja, quando os valores observados em um grupo não influenciam e nem se relacionam com os valores observados noutro grupo e apresentam distribuição normal. A H_0 do teste t de Student para amostras independentes assume que as médias dos dois grupos são iguais, enquanto que a H_A assume que as médias dos dois grupos são diferentes.

Uma das suposições do teste t de Student para amostras independentes é de que as variâncias dos dois grupos sejam iguais. Quando comparamos médias, primeiramente devemos nos preocupar com a variabilidade encontrada em cada grupo. É preciso levar em conta se os grupos possuem ou não variâncias iguais, ou seja, se os grupos são homogêneos quanto à variabilidade. Para isso, o SPSS realiza o teste de Levene que compara as variâncias e, com base no resultado deste teste, escolhe-se o resultado do teste t supondo variâncias iguais ou variâncias diferentes. Em outras palavras, a H_0 assume que as variâncias são iguais (homocedasticidade) e, a H_A , assume que pelo menos uma das variâncias se difere (heteroscedasticidade).

Na interpretação, se o teste de Levene for significativo, ou seja, se a significância do teste for inferior a 0,050 ($p < 0,050$), decidimos que os grupos possuem variâncias diferentes. Assim, devemos considerar o segundo teste t (*Equal variances not assumed*) como válido (localizado na segunda linha da coluna *Sig.(2-tailed)*). Se o teste de Levene não for significativo, ou seja, se a significância do teste for superior ou igual a 0,050 ($p \geq 0,050$), assumimos que não há diferença significativa entre a homogeneidade dos dois grupos. Assim consideramos o primeiro teste t (*Equal variances assumed*) como válido (localizado na primeira linha da coluna *Sig.(2-tailed)*).

Para exemplificar, no “Banco de dados 3 casos novos.sav” (disponível em <https://bit.ly/bancosdedados>), comparamos as médias de “Massa corporal” entre as categorias da variável “Sexo”. Lembre-se: antes de optarmos pelo teste, paramétrico ou não paramétrico, é preciso testar a normalidade de distribuição da variável entre os dois grupos pelo teste de Shapiro-Wilk. Neste caso, ambos os grupos apresentaram significância do teste superior a 0,050. Portanto, a variável de estudo (“Massa”) segue uma distribuição normal e podemos utilizar o teste t de Student para amostras independentes, visto que há apenas dois grupos a serem comparados.

Para executar o teste, no menu “Analisar”, “Comparar médias”, clique em “Teste T de amostras independentes...”. Na caixa de diálogo “Teste T de amostras independentes”, localize e selecione a variável “Massa corporal” na tela da esquerda, e a insira em “Variáveis de teste”. Em seguida, localize e selecione a variável “Sexo” na mesma tela, e a insira em “Variável de agrupamento”. Perceba que é necessário informar os códigos que definem os grupos (“? ?”). Para isso, clique em “Definir grupos...”. Na janela “Definir grupos”, defina as categorias correspondentes. Selecione os grupos que queremos comparar clicando em “Usar valores especificados” e indique os códigos numéricos “1” para “Sexo masculino” e “2” para “Sexo feminino”. Clique em “Continuar”. Em seguida, clique em “Ok” ou “Colar”.

No arquivo de saída, no primeiro quadro (*Group Statistics*), encontram-se as análises descritivas da variável “Massa corporal” para os dois grupos, com a frequência absoluta (N), a média

(Mean), o desvio padrão (Std. Deviation) e o erro padrão da média (Std. Error Mean). Observe que a média do sexo masculino (134,43) é maior que a média do sexo feminino (107,76). Será que a média do sexo masculino difere significativamente da média do sexo feminino?

Para responder essa pergunta, precisamos verificar o p-valor do teste t de Student para amostras independentes. Lembre-se que, primeiro, precisamos checar o teste de Levene para homogeneidade, verificando se a variabilidade entre cada grupo é semelhante ou não. Se o teste de Levene for significativo ($p < 0,050$) decidimos que os grupos possuem variâncias diferentes e optamos pelo valor do teste t não assumindo igualdade de variâncias (segunda linha) como válido. Se o teste de Levene não for significativo ($p \geq 0,050$) assumimos que não há diferença significativa entre a homogeneidade dos dois grupos e optamos pelo valor do teste t assumindo igualdade de variâncias (primeira linha) como válido.

No exemplo, o teste de Levene apresentou um p-valor de 0,353, ou seja, assumimos que não há diferença significativa entre a homogeneidade dos dois grupos e optamos pelo p-valor do teste t da primeira linha, que corresponde a $p < 0,001$. Assim, a média de massa corporal do sexo masculino ($134,43 \pm 17,56$ kg) foi estatisticamente diferente da média de massa corporal do sexo feminino ($107,76 \pm 11,60$ kg; $p < 0,001$). Também podemos apresentar os resultados em uma tabela (tabela 1).

Tabela 1. Comparação das medidas de massa corporal entre sujeitos obesos do sexo masculino e feminino. Os dados são representados por média \pm desvio padrão. As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Sexo masculino (n = 9)	Sexo feminino (n = 21)	P-valor
Massa corporal (em kg)	134,43 \pm 17,56	107,76 \pm 11,60	<0,001*

n = número absoluto.

*Teste t de Student independente. Significância estatística considerada de $p \leq 0,050$.

Atenção! Apesar da distribuição normal pelo teste de Shapiro-Wilk, temos nove sujeitos no sexo masculino. Neste caso, recomenda-se proceder com análise não paramétrica, pois o tamanho amostral de um dos grupos é menor que 15 sujeitos.

Teste U de Mann-Whitney

O teste U de Mann-Whitney é um teste não paramétrico para a comparação de dois grupos independentes, substituindo o teste t de Student independente. A constatação de diferença pelo teste indica que as duas amostras diferem na variável de estudo. Para exemplificar, no "Banco de dados 3 casos novos.sav" (disponível em <https://bit.ly/bancosdedados>), vamos comparar as medianas de "HDL-Colesterol" entre as duas categorias da variável "Sexo". Lembre-se: antes de optarmos pelo teste, paramétrico ou não paramétrico, é preciso testar a normalidade de distribuição da variável entre os dois grupos pelo teste de Shapiro-Wilk. Neste caso, a variável "HDL-Colesterol" tem distribuição não normal no sexo feminino (significância do teste inferior ou igual a 0,050) e normal no sexo masculino (significância do teste superior a 0,050). Portanto, será utilizado o teste U de Mann-Whitney, visto que há dois grupos independentes e em um deles a variável possui distribuição não normal.

No menu "Analisar", "Testes não paramétricos", clique em "Amostras independentes...". Na janela "Testes Não Paramétricos: Duas ou Mais Amostras Independentes", na aba "Objetivo", a opção padrão do SPSS é "Compara automaticamente distribuições entre grupos". Clique na aba "Campos" e, selecione e insira a(s) variável(is) a ser(em) testada(s) (que devem estar declaradas como escalar) no campo "Testar campos" com o auxílio do botão da seta. Em seguida, selecione e insira, com o auxílio do botão da seta, a variável categórica (que deve ser declarada como numérica) em "Grupos".

Como mencionado, o teste não paramétrico é muito simples. Você define se as amostras são independentes ou pareadas e o restante (a escolha de qual teste utilizar) é automaticamente feito pelo SPSS. Para isso, na aba "Configurações", você deve marcar a opção "Escolhe automaticamente os testes com base nos dados". Desta forma, quando se deseja comparar grupos independentes, o

SPSS irá escolher automaticamente se fará o teste para dois grupos (teste U de Mann-Whitney) ou para três ou mais grupos (teste de Kruskal-Wallis), baseado no número de categorias da variável informada no campo "Grupos" da aba "Campos".

Atenção! É necessário classificar corretamente as variáveis quanto ao seu tipo (escalar, ordinal ou nominal) para o teste ser aplicado adequadamente. Reveja os conceitos no capítulo 1. Você também pode definir o teste, selecionando a opção "Mann-Whitney U (2 amostras)" em "Customizar testes". Por este caminho, é possível alterar o nível de significância (α) em "Opções de teste". Por fim, clique em "Executar" ou "Colar" para obter o resultado do teste estatístico.

O teste utilizado é informado no arquivo de saída, junto com a hipótese de nulidade, o p-valor e a decisão de teste (padrão baseado em uma significância de 5%). Perceba que não obtemos a análise descritiva, somente o resultado do teste de hipótese (*Hypothesis Test Summary*). A mediana e percentis ou valores mínimo e máximo serão extraídos das tabelas com as análises descritivas (*Descriptives*) e percentis (*Percentiles*) solicitados no teste de normalidade. Conforme é indicado pelo teste de hipótese, rejeita-se H_0 de que a distribuição dos níveis séricos de HDL-Colesterol seja igual entre as categorias da variável "Sexo". Ou seja, com p-valor igual a 0,009, há diferença significativa nos níveis séricos de HDL-Colesterol entre os sexos masculino e feminino. A mediana para o sexo masculino foi de 35,0 (34,0-42) mg/dL e para o sexo feminino foi de 47,0 (44,0-52,0) mg/dL. Podemos apresentar os dados em uma tabela (Tabela 2).

Tabela 2. Comparação das medidas de HDL-Colesterol entre sujeitos obesos do sexo masculino e feminino. Os dados são representados por mediana e percentis (25-75). As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Sexo masculino (n = 9)	Sexo feminino (n = 21)	P-valor
HDL-Colesterol (mg/dL)	35,0 (34,0-42)	47,0 (44,0-52,0)	0,009*

HDL, lipoproteína de alta densidade; n = número absoluto.

*Teste U de Mann-Whitney. Significância estatística considerada de $p \leq 0,050$.

Ao darmos um duplo clique no teste de hipóteses no arquivo de saída, uma nova janela de “Visualizador de modelo” mostrará um gráfico contendo a frequência da variável de cada grupo.

Análise de variância com um fator – ANOVA (one-way)

Quando comparamos grupos, além de considerar a distribuição e a relação entre os grupos, também devemos levar em conta o número de grupos ou fatores a serem comparados. Quando a análise envolve comparar três ou mais grupos, supondo a normalidade dos dados pelo teste de Shapiro-Wilk, podemos realizar uma análise de variância (ANOVA). A ANOVA é uma generalização do teste t de Student, onde as variáveis categóricas devem ser selecionadas como fatores do modelo. Cada nível do fator pode ter um efeito diferente sobre a variável dependente. Além de considerarmos a normalidade e o número de grupos a serem comparados (três ou mais), a ANOVA possui outros dois pré-requisitos: a homogeneidade (ou homocedasticidade) das variâncias (pelo teste de Levene) e a independência das observações (definida pelo delineamento do estudo).

O objetivo da ANOVA é verificar se há diferença significativa entre pelo menos duas médias dos grupos. Para isso, a ANOVA considera como H_0 que não há diferença entre as médias dos grupos ($\mu_A = \mu_B = \dots = \mu?$) e, como H_A , que há diferença entre pelo menos duas das médias dos grupos ($\mu_A \neq \mu?$). Desta forma, se não rejeitamos a H_0 ($p > 0,050$), concluímos a análise de que não há evidências significativas de diferença entre as médias dos grupos. Por outro lado, se rejeitamos a H_0 ($p \leq 0,050$), devemos prosseguir com a análise, comparando os grupos dois a dois (testes *post hoc*), para identificar quais médias são significativamente diferentes, já que pelo menos duas das médias dos grupos diferem entre si.

A ANOVA também produz um valor chamado F (*F-statistics* ou *F-ratio*). Esse valor de F compara a quantidade de variância sistemática dos dados com a quantidade de variância não sistemática. Em outras palavras, o valor de F indica a razão entre o modelo e seu erro. O valor de F vai nos dizer o quanto as médias desses grupos não são iguais. Para a comparação de médias entre os grupos, um valor de F elevado significa que há alguma diferença

entre esses grupos capaz de ser expressa adequadamente por meio de um modelo de regressão. No entanto, o valor de F só nos indica se houve diferença global entre as médias, mas não nos informa quais pares de grupos possuem médias significativamente diferentes. Para isso, precisamos dos testes *post hoc*.

Como exemplo, vamos comparar as médias da variável "Estatura" entre os níveis da variável "Circunferência da cintura recodificada", do "Banco de dados 3 Resolvido.sav" (disponível em <https://bit.ly/bancosdedados>). A variável "Circunferência da cintura recodificada" apresenta quatro categorias: "1" = até 104,00 cm; "2" = 104,01-111,00 cm; "3" = 111,01-120,00 cm; e "4" = mais de 120,01 cm e, a variável "Estatura" apresenta distribuição normal ($p > 0,050$) em todas as quatro categorias. Lembre-se: antes de optarmos pelo teste, paramétrico ou não paramétrico, é preciso testar a normalidade de distribuição da variável entre os grupos pelo teste de Shapiro-Wilk.

Existem dois caminhos para a ANOVA one-way no SPSS. No primeiro, no menu "Analisar", "Comparar médias", clique em "ANOVA a um fator...". Na janela "ANOVA a um fator", selecione, na tela da esquerda, a variável "Estatura" como "Variável dependente" e a variável "Circunferência da cintura recodificada" em "Fator", com o auxílio do botão da seta. Clique em "Post Hoc" (comparação par a par). Na opção "Variâncias iguais presumidas", selecione "Tukey", "Bonferroni", "S-N-K" e "Duncan". Todos esses testes *post hoc* supõem variâncias iguais, mas sugere-se o teste de Tukey, por ser mais rigoroso, ou de Bonferroni, por ser mais conservador. Clique em "Continuar". Clique em "Opções" e marque, em "Estatísticas", as opções "Descritivo" e "Homogeneidade do teste de variância", e selecione a opção de "Diagrama de médias". Clique em "Continuar" e, em seguida, em "Ok" ou "Colar".

No arquivo de saída, na primeira tabela (*Descriptives*), temos a análise descritiva da variável dependente ("Estatura") pelos níveis do fator ("Circunferência da cintura recodificada"). São apresentados: frequência absoluta (N), média (*Mean*), desvio padrão (*Std. Deviation*), erro padrão (*Std. Error*), intervalo de confiança de 95% para a média (*95% Confidence Interval for Mean*) e os valores mínimo (*Minimum*) e máximo (*Maximum*) da

variável dependente para cada nível do fator. Em seguida, temos o quadro do teste de Levene para homogeneidade de variância (*Test of Homogeneity of Variances*). Podemos verificar que, no exemplo, o teste de Levene foi significativo ($p = 0,001$), ou seja, o pré-requisito de homogeneidade não foi satisfeito (os grupos possuem variâncias diferentes) e a análise não pode prosseguir.

Apesar do teste de Shapiro-wilk ter resultado em um p-valor não significativo (o que indica a normalidade de distribuição dos dados), o teste de Levene foi significativo, ou seja, menor que 5%. Neste caso, teríamos que utilizar outro método estatístico para analisar nossos dados. Uma alternativa é utilizar testes não-paramétricos, como o teste de Kruskal-Wallis, que ainda será abordado neste capítulo.

Mas, suponhamos que o teste de Levene tenha apresentado p-valor não significativo ($p \geq 0,050$), ou seja, os dados têm homogeneidade de variância (pressuposto da homogeneidade). Uma vez atendido o pressuposto, podemos seguir com a análise da ANOVA *one-way*. Ao observar a tabela da ANOVA, verificamos que o p-valor do teste foi significativo, ou seja, existe diferença de estatura entre os níveis da circunferência da cintura ($p < 0,001$) e, o valor de F elevado (53,493), indica que há alguma diferença entre esses grupos. No entanto, precisamos verificar quais pares de grupos possuem médias significativamente diferentes. Para isso, devemos analisar os testes *post hoc*.

O quadro seguinte (*Multiple Comparisons*) apresenta os testes de Tukey HSD e Bonferroni com a comparação par a par das categorias (níveis) do fator para a variável dependente. Como mencionado, dentre os testes *post hoc*, sugere-se o teste de Tukey por ser o mais rigoroso. Por isso, iremos interpretar os dados para o teste *post hoc* de Tukey. A interpretação para os outros testes é similar. Além disso, no quadro, temos a comparação par a par para cada nível fixo do fator ((I) Circunferência da cintura recodificada) com os outros níveis ((J) Circunferência da cintura recodificada), a diferença entre as médias da variável dependente ("Estatura") para cada par de nível comparado (*Mean Difference (I-J)*), o erro padrão (*Std. Error*), o nível de significância (*Sig.*) e o intervalo de confiança (*95% Confidence Interval*).

Analisando estes resultados, podemos concluir, com 95% de confiança, que as categorias 2 e 3 não diferem em média na estatura ($p = 0,718$). A categoria 1 difere significativamente da categoria 2 com $p = 0,005$ e, das categorias 3 e 4, com $p < 0,001$. As categorias 2 e 3 se diferem significativamente da categoria 4 ($p < 0,001$) e, esta última, se diferencia de todas ($p < 0,001$). Uma forma de simplificar a representação destas informações é fazer uma tabela utilizando um sistema de letras, onde letras distintas representam médias significativamente diferentes (Tabela 3). Ou seja, primeiro colocamos os níveis (1, 2, 3 e 4) alinhados por ordem crescente em relação às médias. Em seguida, traduzimos os resultados da tabela de comparação par a par. Para o exemplo, a categoria 1, por ser a de menor média, recebe a letra "a" e as categorias que também não se diferenciam significativamente dela recebem a letra "a" (a categoria 1 se difere de todas); as categorias 2 e 3 são as únicas que não se diferem entre si, por isso recebem a mesma letra "b"; a categoria 4 se difere de todas as categorias, por isso recebe a letra "c".

Tabela 3. Comparação das medidas de estatura entre as categorias de circunferência da cintura de sujeitos obesos. Os dados são representados por média e desvio padrão. As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

Níveis de circunferência da cintura	Frequência absoluta (n)	Média (em metros)	Desvio padrão
1 (até 104,00 cm)	126	1,60 ^a	0,06
2 (104,01 a 111,00 cm)	110	1,63 ^b	0,06
3 (111,01 a 120,00 cm)	124	1,64 ^b	0,08
4 (mais de 120,01 cm)	102	1,72 ^c	0,09
Total	462	1,64	0,08

Letras distintas representam média de estatura entre os níveis de circunferência da cintura significativamente diferentes.

n = número absoluto.

*Teste ANOVA one-way. Significância estatística considerada de $p \leq 0,050$;

Os resultados também podem ser apresentados por gráficos. A figura 2 apresenta a comparação das médias da estatura entre os quatro níveis de circunferência da cintura, possibilitando a análise visual dos resultados descritos na tabela. Note que é possível perceber que a categoria 1 se difere de todas; as categorias 2 e 3 são as únicas que não diferem entre si; e a categoria 4 se difere de todas as categorias.

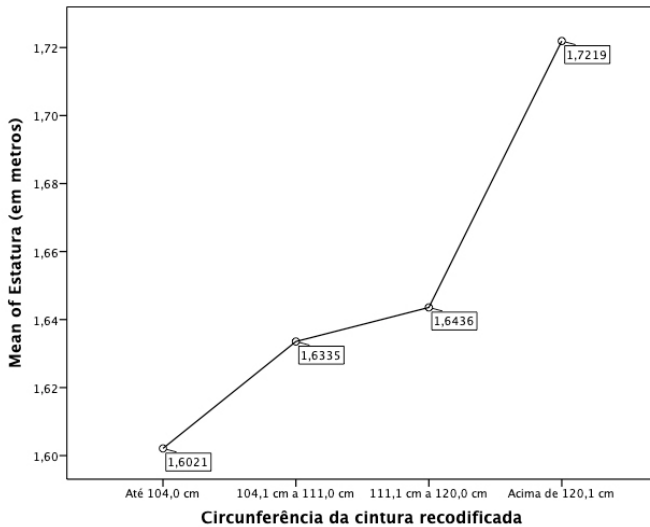


Figura 2. Comparação das médias da estatura entre os quatro níveis de circunferência da cintura.

Na segunda opção, além da ANOVA *one-way*, podemos realizar a ANOVA *two-way* (acrescentando dois fatores) e a ANCOVA (adicionando uma covariável). Estes testes serão abordados a seguir neste capítulo.

No menu "Analisar", "Modelo linear geral", clique em "Com uma única variável...". Na caixa de diálogo "Univariável", selecione a variável "Estatura" como "Variável dependente" e a variável "Circunferência da cintura recodificada" em "Fator(es) fixo(s)". Clique em "Post Hoc" e insira a variável "Fator(es)" em "Testes *post hoc* para". Na opção "Variâncias iguais presumidas" selecione "Tukey", "Bonferroni", "S-N-K" e "Duncan". Clique

em "Continuar". Em "Opções", marque as opções "Estatísticas descritivas", "Estimativas de tamanho de efeito", "Potência observada" e "Testes de homogeneidade". Clique em "Continuar" e, em seguida, em "Ok" ou "Colar".

Como visto da primeira maneira, nas primeiras tabelas temos a análise descritiva da variável dependente ("Estatura") pelos níveis do fator ("Circunferência da cintura recodificada") e, em seguida, temos o teste de Levene para homogeneidade de variância. Como podemos verificar, o pré-requisito de homogeneidade não foi satisfeito (os grupos possuem variâncias diferentes) e a análise não pode prosseguir. Se o requisito fosse cumprido, seguiríamos os mesmos passos de interpretação dos dados vistos anteriormente.

Desta forma, vamos apresentar as diferenças entre os dois caminhos de teste de ANOVA. A primeira diferença é que obtemos um coeficiente de correlação (*Partial Eta Square*). O coeficiente de correlação é uma medida de tamanho de efeito, que pode ser interpretado como o percentual da variância total da variável dependente explicado pela variação entre as categorias formadas pelas variáveis independentes. Neste exemplo, 25,9% da variância total da "Estatura" pode ser explicada pelos níveis do fator "Circunferência da cintura recodificada". A segunda diferença corresponde ao poder do teste (*Observed Power*). Relacionado ao nível de significância, o poder do teste indica a probabilidade de rejeitar corretamente a hipótese de nulidade. Assim, quanto maior a potência do teste, maior o grau de confiança nos resultados. Neste exemplo, essa estatística assume o valor máximo (igual a 1).

Outra diferença, é que não são construídos os gráficos. É necessário criá-los no menu "Gráficos", "Caixas de diálogo legendas", "Linha...". Na janela "Gráficos de linha", selecione as opções "Simples" e "Resumos para grupos de casos". Clique em "Definir". Na janela seguinte, clique em "Outra estatística (por exemplo, média)" e selecione, na tela da esquerda, a variável "Estatura" e a insira em "Variável". Clique em "Alterar estatística..." e indique "Média dos valores". Clique em "Continuar". Insira a variável "Circunferência da cintura recodificada" em "Eixo da categoria". Em seguida, clique em "Ok" ou "Colar". Assim, no arquivo de saída, teremos o mesmo gráfico apresentado na figura 2.

Análise de variância com dois fatores – ANOVA two-way

A ANOVA two-way trabalha com dois fatores, ou seja, duas variáveis explicativas categóricas. O objetivo desta análise é verificar se o comportamento dos fatores em conjunto (interação) pode interferir na variável de estudo que, assim como na ANOVA one-way, deve ser quantitativa paramétrica. Desta forma, podemos analisar os efeitos principais dos dois fatores de forma independente (isoladamente) e o efeito da interação dos mesmos (em conjunto). O efeito principal corresponde ao efeito direto de uma variável independente sobre a variável dependente e o efeito de interação reflete o efeito conjunto de duas ou mais variáveis independentes sobre a variável dependente. Os fatores podem ser de efeitos fixos (são variáveis onde todas as categorias estão representadas) ou de efeitos aleatórios (são as variáveis em que apenas algumas das categorias, em relação ao todo, são selecionadas aleatoriamente).

Para entender as interações, vamos observar o exemplo ilustrado por Guimarães e Hirakata (2012), utilizando os fatores (variáveis categóricas), grupo e tempo, para explicar o comportamento da variável dependente ansiedade, apresentada por um escore (Figura 3). Na figura A1, não há efeito dos fatores, grupo e tempo, sobre a variável ansiedade, ou seja, não há relação entre os fatores e a variável dependente. Na figura A2, não há interação, mas há um efeito do grupo, mostrando que o Tempo 2 possui um escore maior que os demais tempos e separadamente certificamos que os grupos possuem médias diferentes, sendo que o Grupo A é sempre maior que o Grupo B. Na figura A3, temos o efeito (diferença) da variável grupo, mas não há efeito no tempo, diferente da figura A4 em que há efeito do tempo, mas não do grupo. Nas duas últimas figuras, observamos a dependência de uma variável em relação à outra. Na figura A5, há divergência dos grupos em relação ao terceiro tempo (interação) e, na figura A6, há um cruzamento entre os grupos, observando que no Tempo 1 o grupo B possui maiores médias de escores de ansiedade, se igualando no tempo 2 e reduzindo os valores médios no Tempo 3 (interação).

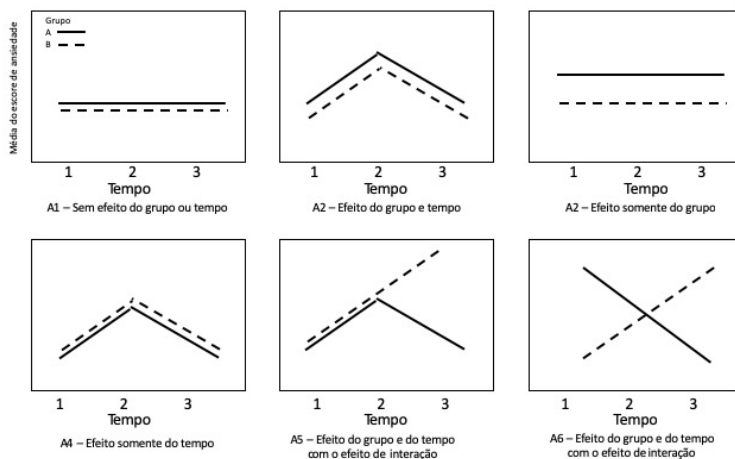


Figura 3. Interação entre os fatores, grupo e tempo, para explicar o comportamento da variável dependente “escore de ansiedade”. (Guimarães e Hirakata, 2012).

Assim como a ANOVA *one-way*, a ANOVA *two-way* também apresenta pré-requisitos para o teste. São estes: 1) o número de grupos a serem comparados (três ou mais); 2) a normalidade de distribuição dos dados (pelo teste de Shapiro-wilk); 3) a homogeneidade (ou homocedasticidade) das variâncias (pelo teste de Levene) e; 4) a independência das observações (definida pelo delineamento do estudo). A ANOVA de três ou mais vias segue os mesmos pré-requisitos, comandos e interpretações da ANOVA *two-way*.

Para exemplificar, supondo normalidade dos dados, vamos verificar a relação entre as médias da variável “Estatura” (variável dependente) entre os níveis dos fatores “Circunferência da cintura recodificada” e “Classificação IPAQ” (ambos fatores de efeito fixo), no “Banco de dados 3 Resolvido.sav” (disponível em <https://bit.ly/bancosdedados>). Vamos seguir os mesmos passos vistos na segunda forma para realizar a ANOVA *one-way*, mas incluindo dois fatores. Lembre-se: antes de optarmos pelo teste, paramétrico ou não paramétrico, é preciso testar a normalidade de distribuição da variável entre os grupos pelo teste de Shapiro-Wilk. Neste exemplo, apesar da normalidade da variável “Estatura” entre

os grupos da variável "Circunferência da cintura recodificada", temos uma distribuição não normal da variável entre os grupos da "Classificação IPAQ". Portanto, não poderíamos realizar a ANOVA two-way. Por isso, para o exemplo, iremos supor que o pressuposto de normalidade tenha sido cumprido.

No menu "Analisar", "Modelo linear geral", clique em "Com uma única variável...". Na janela "Univariável", selecione, na tela da esquerda, a variável "Estatura" como "Variável dependente" e, agora, vamos adicionar dois fatores ao invés de um, no caso, as variáveis "Circunferência da cintura recodificada" e "Classificação IPAQ" em "Fator(es) fixo(s)". Clique em "Diagramas" e insira no "Eixo horizontal" a variável "CCREC" e em "Linhas separadas" a variável "IPAQ". Clique em "Adicionar". Perceba que está sinalizado no quadro "Diagrama" a interação entre os dois fatores ("CCREC*IPAQ"). Em seguida, clique em "Continuar". Em "Post Hoc", selecione os dois "Fator(es)" e insira-os em "Testes post hoc para". Na opção "Variâncias iguais presumidas" selecione "Tukey", "Bonferroni", "S-N-K" e "Duncan". Clique em "Continuar". Clique em "Opções", localize e selecione a variável da interação ("CCREC*IPAQ") e a insira dentro do quadro "Exibir médias para". Marque as opções "Estatísticas descritivas", "Estimativas de tamanho de efeito", "Potência observada" e "Testes de homogeneidade" em "Exibir". Clique em "Continuar" e, em seguida, em "Colar", para gerar o arquivo de sintaxe.

Atenção! Será necessário adicionar alguns comandos (grifados no Quadro 1) ao algoritmo manualmente no arquivo de sintaxe para obtermos a análise de *post hoc*. Neste comando, os níveis do primeiro fator serão fixos e, o segundo fator com um nível fixo é comparado par a par com os outros níveis, e vice-versa, semelhante ao descrito no quadro da ANOVA *one-way* (*Multiple Comparisons*), para verificar a diferença das médias de estatura. Após adicionar o comando, selecione todo o algoritmo e execute-o.

Quadro 1. Comando para a análise *post hoc*, fixando os níveis de um dos fatores e comparando par a par os níveis de outro fator.

```
UNIANOVA ESTATURA BY CCREC IPAQ
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/POSTHOC=CCREC IPAQ(SNK TUKEY DUNCAN BONFERRONI)
/PLOT=PROFILE(CCREC*IPAQ)
/EMMEANS=TABLES(CCREC*IPAQ) COMPARE (CCREC) ADJ(BONFERRONI)
/EMMEANS=TABLES(CCREC*IPAQ) COMPARE (IPAQ) ADJ(BONFERRONI)
/PRINT=OPOWER ETASQ HOMOGENEITY DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESGIN=CCREC IPAQ CCREC*IPAQ.
```

No arquivo de saída, no primeiro quadro (*Between-Subjects Factors*), são apresentados os fatores, os níveis (categorias) de cada fator e a frequência absoluta (N) de cada nível. No quadro seguinte (*Descriptive Statistics*), temos a análise descritiva (média e desvio padrão) da variável dependente ("Estatura") para cada nível dos fatores ("Circunferência da cintura recodificada" e "Classificação IPAQ").

Em seguida, temos o teste de Levene para homogeneidade de variância (*Levene's Test of Equality of Error Variances*). Podemos verificar que, no exemplo, o teste de Levene foi significativo ($p = 0,010$), ou seja, o pré-requisito de homogeneidade não foi satisfeito (os grupos possuem variâncias diferentes) e a análise não pode prosseguir.

Mas, imaginemos que o pré-requisito de homogeneidade tenha sido satisfeito (os grupos não possuem variâncias diferentes) e a análise pode prosseguir. No quadro seguinte (*Tests of Between-Subjects Effects*), são testados o efeito da interação (em conjunto) e os efeitos principais (separadamente). A leitura deve ser realizada de baixo para cima, ou seja, devemos primeiramente avaliar o efeito da interação e posteriormente, se não houver significância estatística na interação, os efeitos principais de cada fator sobre a

variável dependente. Com p-valor igual a 0,345, a interação não foi significativa, afirmado pelo valor de $F(1,116)$. Neste caso, trabalhamos os efeitos principais, como na ANOVA *one-way*, olhando para os testes *post hoc* (que se encontram ao final da análise).

Também podemos observar o coeficiente de correlação (*Partial Eta Square*) e o poder do teste (*Observed Power*). Neste exemplo, 2,9% da variância total na estatura pode ser explicado pela interação dos níveis (categorias) da variável circunferência da cintura e IPAQ e essa estatística assume o valor de 0,649 (grau de confiança nos resultados). Se o p-valor mostrasse uma interação significativa, isso significaria que há um efeito conjunto dos fatores ("Circunferência da cintura" e "IPAQ") com a variável dependente ("Estatura"). Sendo assim, não é correto analisar separadamente os efeitos principais dos fatores, somente a interação.

Também é apresentado um quadro (*Estimates*) com os níveis dos fatores e as médias (*Mean*), erro padrão (*Std. Error*) e intervalo de confiança de 95% (*95% Confidence Interval*) da variável dependente para cada nível.

Vamos supor que a interação entre os fatores tenha sido significativa e, assim, para completar a análise, precisamos analisar o teste de *post hoc* de Bonferroni (*Pairwise Comparisons*). O raciocínio é o mesmo visto na interpretação do teste de *post hoc* de Tukey visto na ANOVA *one-way*. O objetivo é verificar quais categorias (níveis) diferem significativamente entre si na comparação par a par. Para essa análise ocorrer, foi necessário inserir os comandos no algoritmo do arquivo de sintaxe.

Em um primeiro quadro, há a comparação média das estaturas entre os níveis (categorias) da circunferência da cintura para cada nível do IPAQ (variável fixa). Para interpretar os resultados, vamos somente olhar para os indivíduos classificados como sedentários no IPAQ. Conforme a comparação par a par do teste de *post hoc* de Bonferroni, nos sujeitos sedentários, a categoria 4 (acima de 120,01 cm) da circunferência da cintura se diferenciou significativamente, na estatura média, em comparação a todas as outras categorias ($p < 0,001$). Essa interpretação deve ser feita para cada conjunto de comparação.

No segundo quadro, temos a comparação contrária, ou seja, a comparação média das estaturas entre as categorias (níveis) do IPAQ para cada categoria (nível) da circunferência da cintura (variável fixa). Para interpretar os resultados, vamos observar a última categoria da circunferência da cintura. Conforme a comparação par a par do teste de *post hoc* de Bonferroni, nos sujeitos com mais de 120,01 cm de circunferência da cintura (categoria 4), os sujeitos insuficientemente ativos A se diferenciam significativamente, na estatura média, em comparação aos sujeitos ativos ($p = 0,031$). Essa mesma interpretação deve ser feita para cada conjunto de comparação. Da mesma forma vista na ANOVA *one-way*, podemos sumarizar os resultados em uma tabela (Tabela 4) ou gráfico (Figura 4).

Tabela 4. Interação entre fatores, circunferência da cintura e IPAQ, para explicar o comportamento da variável dependente estatura. Os dados são representados por média e desvio padrão. As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

Categories CC	1 (até 104,0 cm)	2 (104,01-111,0 cm)	3 (111,01-120,0 cm)	4 (mais de 120,01 cm)
Classificação IPAQ	Média ± DP	Média ± DP	Média ± DP	Média ± DP
Sedentário (n = 60)	1,57 ± 0,05 ^{aA}	1,60 ± 0,05 ^{aA}	1,64 ± 0,10 ^{bA}	1,73 ± 0,10 ^{bAB}
Insuf. B (n = 140)	1,62 ± 0,07 ^{aA}	1,64 ± 0,06 ^{aA}	1,65 ± 0,07 ^{aA}	1,71 ± 0,09 ^{bAB}
Insuf. A (n = 89)	1,60 ± 0,04 ^{aA}	1,64 ± 0,07 ^{abA}	1,66 ± 0,07 ^{bA}	1,77 ± 0,06 ^{cA}
Ativo (n = 157)	1,59 ± 0,06 ^{aA}	1,63 ± 0,07 ^{aA}	1,63 ± 0,07 ^{aA}	1,70 ± 0,09 ^{bB}
Muito ativo (n = 16)	1,57 ± 0,07 ^{aA}	1,66 ± 0,08 ^{abA}	1,64 ± 0,03 ^{abA}	1,73 ± 0,05 ^{bAB}

Letras minúsculas distintas representam médias significativamente diferentes entre categorias de circunferência da cintura, fixando IPAQ (→). Letras maiúsculas distintas representam médias significativamente diferentes entre IPAQ, fixando categorias de circunferência da cintura (↓). CC, circunferência da cintura; IPAQ, questionário internacional de atividade física; DP, desvio padrão, n, número absoluto

*Teste ANOVA two-way. Significância estatística considerada de $p \leq 0,050$.

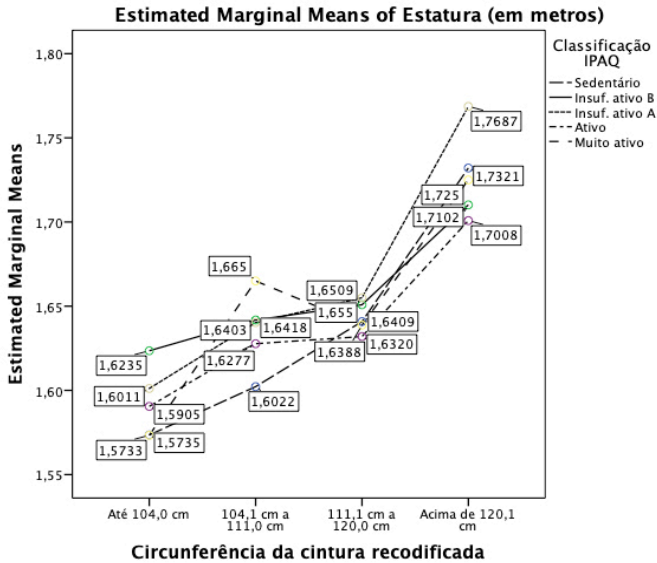


Figura 4. Interação entre fatores, circunferência da cintura e IPAQ, para explicar o comportamento da variável dependente estatura.

Atenção! Lembre-se que, como o p-valor foi igual a 0,345, a interação não foi significativa. Desta forma, devemos ver os efeitos principais (isoladamente), como na ANOVA one-way, para cada fator, olhando para os testes *post hoc* de Tukey (que se encontram ao final da análise).

Os quadros apresentados (*Multiple comparisons*) são idênticos a ANOVA one-way: temos a comparação par a par para cada nível fixo do fator ((I) Circunferência da cintura recodificada ou (I) Classificação IPAQ) com os outros níveis ((J) Circunferência da cintura recodificada ou (J) Classificação IPAQ), a diferença entre as médias da variável dependente ("Estatura") para cada par de nível comparado (*Mean Difference (I-J)*), o erro padrão (*Std. Error*), o nível de significância (*Sig.*) e o intervalo de confiança (95% *Confidence Interval*).

Sendo assim, na interpretação referente aos níveis de circunferência da cintura, podemos dizer que, com 95% de confiança, as categorias 2 e 3 são as únicas que não diferem em média para a estatura ($p = 0,714$). Com relação as categorias do

IPAQ, podemos concluir, com 95% de confiança, que nenhuma categoria do IPAQ difere em média para a estatura ($p > 0,050$).

Análise de covariância – ANCOVA

A análise de covariância (ANCOVA) tem o mesmo objetivo da ANOVA, que é testar os efeitos principais e a interação de variáveis independentes (categóricas) sobre a variável dependente (contínua). No entanto, na ANCOVA, podemos controlar estes efeitos para outras variáveis independentes contínuas, as quais chamamos de covariáveis, ou ainda, variável auxiliar ou concomitante. A covariável complementa o controle local e, na grande maioria das situações, simplesmente o substitui. Desta forma, a ANCOVA permite que se faça um “ajuste” do efeito de uma variável resposta que sofreu influência de uma variável ou uma causa de variação não controlada. Obrigatoriamente, a covariável necessita estar correlacionada com a variável resposta para que se possa fazer uso de tal análise. Quando a ANCOVA é realizada com mais de uma covariável, chama-se de análise de ANCOVA².

Para exemplificar, no “Banco de dados 4.sav” (disponível em <https://bit.ly/bancosdedados>), vamos verificar se há diferença na massa corporal após a cirurgia bariátrica (“Massa corporal 3”) entre os sexos (“Sexo”), supondo a normalidade de distribuição dos dados, e utilizando a massa corporal basal (“Massa corporal 1”) como controle (covariável). Ou seja, vamos corrigir essa comparação pela massa corporal basal dos indivíduos, para saber se este não é um fator confundidor. Iremos assim, verificar se as variáveis sexo (“Sexo”) e massa corporal basal (“Massa corporal 1”) possuem efeito na variável dependente massa corporal final (“Massa corporal 3”). Seguiremos o mesmo caminho de análise visto na ANOVA *one-way*, mas acrescentando uma covariável. Lembre-se: antes de optar pelo teste, paramétrico ou não paramétrico, é preciso testar a normalidade de distribuição da variável entre os grupos pelo teste de Shapiro-Wilk.

Este banco de dados contém a avaliação de sujeitos com obesidade no momento pré-operatório (1) e em dois momentos após a cirurgia bariátrica (2 e 3). Portanto, temos três medidas distintas, uma para cada ponto no tempo (pareadas), para as

variáveis massa corporal, circunferência da cintura, medidas pressóricas, hipertensão e atividade física avaliada pelo IPAQ versão curta. Além disso, temos descritos sexo, técnica cirúrgica e estatura dos sujeitos. Estes sujeitos também foram avaliados quanto a alterações de sensibilidade nos membros inferiores (neuropatia periférica) através do MNSI, por dois examinadores. Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Primeiro, temos que realizar uma ANOVA *one-way* para nos certificarmos que a variável independente ("Sexo") não tenha efeito sobre a covariável ("Massa corporal 1"). Para isso, o p-valor do teste deve ser não significativo, ou seja, maior que 0,050. Porém, no exemplo, há efeito do sexo sobre a massa corporal basal ($p < 0,001$), evidenciado pelo valor de F elevado (59,458). Vamos supor que não haja efeito, para podermos prosseguir com a ANCOVA.

No menu "Analisar", "Modelo linear geral", clique em "Com uma única variável...". Na janela "Univariável", selecione, na tela da esquerda, a variável "Massa corporal 3" como "Variável dependente", a variável "Sexo" como "Fator(es) fixo(s)" e a variável "Massa corporal 1" como "Covariável". Em "Modelo", é necessário que a opção "Fatorial completo" esteja selecionada (padrão). Clique em "Opções" e, selecione e insira a variável "Sexo" para "Exibir médias para". Clique em "Comparar efeitos principais" e selecione na aba "Ajuste do intervalo de confiança" o teste de "Bonferroni". Marque as opções: "Estatísticas descritivas", "Estimativas de tamanho de efeito", "Potência observada" e "Testes de homogeneidade" em "Exibir". Clique em "Continuar", e em seguida, clique em "Ok" ou "Colar".

No arquivo de saída, no primeiro quadro (*Descriptive Statistics*) observamos a análise descritiva da variável dependente ("Massa corporal 3") discriminada pelas categorias da variável "Sexo", com as médias (*Mean*), desvio padrão (*Std. Deviation*) e frequência absoluta (N). Assim como na ANOVA, temos o teste de Levene para homogeneidade (*Levene's Test of Equality of Error Variances*). Com $p = 0,066$, aceitamos a hipótese de nulidade de homogeneidade das variâncias entre as categorias de sexo.

Nos próximos quadros, temos a análise de covariância (*Tests of Between-Subjects Effects*) e o teste de *post hoc* de Bonferroni (*Pairwise Comparisons*). Com $p = 0,560$, não há diferença significativa entre os sexos em relação à massa corporal final, ajustando a análise pela massa corporal inicial ($p < 0,001$). Além disso, no quadro "*Estimates*", são apresentadas novas médias controladas, ou seja, as médias estimadas para os sexos e corrigidas pela massa corporal inicial (fixada em 112,99). Perceba que as novas médias de massa final (74,44 kg para o sexo masculino e 73,19 kg para o sexo feminino) diferem para as médias "reais" sem o controle da massa corporal inicial (85,26 kg para o sexo masculino e 71,15 para o sexo feminino).

Para verificar a interação dos fatores, precisamos desfazer as ações realizadas em "Opções" e, em "Modelo", marcar a opção "Personalizado", onde deve-se solicitar o efeito principal do fator ("Sexo") e da covariável ("Massa corporal 1"), e a interação ("Massa corporal 1*Sexo"). Para isso, basta selecionar ambas as variáveis em "Fatores e covariáveis", marcar o "Tipo: interação" e, com auxílio do botão da seta, inserir no "Modelo". Clique em "Continuar", e em seguida, clique em "Ok" ou "Colar".

No arquivo de saída, é possível observar que não há interação entre sexo e a massa corporal basal ($p = 0,054$). Ou seja, as categorias apresentam o mesmo comportamento. Se houvesse interação ($p \leq 0,050$), as categorias (sexos masculino e feminino) apresentariam comportamentos de massa corporal basal diferentes, por exemplo, um seria o inverso do outro.

Podemos demonstrar os resultados da ANCOVA através de um gráfico (Figura 5). Para construção do gráfico, clique em "Gráficos", "Caixas de diálogo legadas" e selecione a opção "Dispersão/ponto...". Clique em "Dispersão simples" e em "Definir". Insira a variável "Massa corporal 3" no "Eixo Y" e "Massa corporal 1" no "Eixo X". Selecione a variável "Sexo" e insira em "Colunas". Clique em "Ok" ou "Colar".

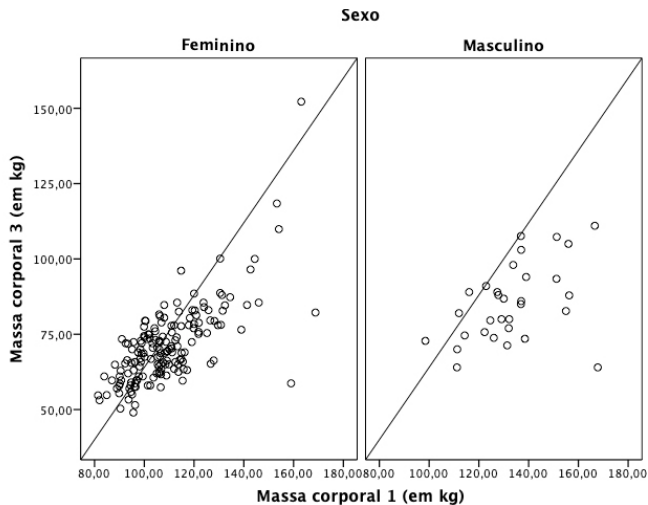


Figura 5. Gráfico de dispersão para a massa corporal basal versus a massa corporal final, discriminado pelo sexo.

O gráfico obtido é do tipo “scatter plot” (gráfico de dispersão) para a massa corporal basal versus a massa corporal final, discriminado pelo sexo. A linha de tendência foi desenhada com um duplo clique na imagem do gráfico no arquivo de saída. Com o botão direito, clique sobre o gráfico na janela “Editor de gráfico” e selecione “Adicionar linha de referência da equação”. Analisando, então, a associação entre a massa corporal final e basal discriminado por sexo, observamos que em ambos os sexos o comportamento é o mesmo. Assim, para a ANCOVA, não há interação dos fatores.

Teste de Kruskal-Wallis

Quando os pré-requisitos necessários para o teste paramétrico não podem ser cumpridos para a variável de estudo, não se pode confiar no resultado de uma análise de variância, pois a probabilidade de se cometer um erro do tipo I afasta-se perceptivelmente de α . O teste de Kruskal-Wallis substitui a ANOVA one-way quando isto ocorre. Juntamente com o teste U de Mann-Whitney, o teste de Kruskal-Wallis é um dos testes não paramétricos mais utilizados. É utilizado para comparação de variáveis ordinais ou contínuas (intervalares) entre três ou mais grupos independentes. Quando constatada uma diferença significativa, temos indicação

de que pelo menos duas amostras diferem na variável de interesse. Para verificar qual dos pares se diferenciam, o teste de Kruskal-Wallis apresenta uma análise de *post hoc* de Dunn.

Para exemplificar, no "Banco de dados 3 casos novos.sav" (disponível em <https://bit.ly/bancosdedados>), vamos comparar as medianas da variável "Massa corporal" nos quatro grupos da variável "Classificação IPAQ". Para tanto, será utilizado o teste de Kruskal-Wallis para amostras independentes, visto que há quatro grupos a serem comparados e a variável possui distribuição não normal em um dos grupos. Lembre-se: antes de optarmos pelo teste, paramétrico ou não paramétrico, é preciso testar a normalidade de distribuição da variável entre os grupos pelo teste de Shapiro-Wilk.

O teste é executado da mesma forma que se realiza o teste U de Mann-Whitney. No menu "Analisar", "Testes não paramétricos", clique em "Amostras independentes...". Na aba "Objetivo", consta como padrão "Compara automaticamente distribuições entre grupos". Na aba seguinte ("Campos"), localize e insira, na tela da esquerda, a(s) variável(is) a ser(em) testada(s) (que deve estar declarada como escalar) no campo "Testar campos" e a variável categórica (que deve ser declarada como numérica) em "Grupos", com o auxílio do botão da seta.

Como mencionado no teste U de Mann-Whitney, a comparação não paramétrica é muito simples. Você define se as amostras são independentes ou pareadas e o restante (a escolha de qual teste utilizar) é automaticamente feito pelo SPSS. Para isso, na aba "Configurações", você deve marcar a opção "Escolhe automaticamente os testes com base nos dados". Desta forma, quando se deseja comparar grupos independentes, o SPSS irá escolher automaticamente se fará o teste para dois grupos (teste U de Mann-Whitney) ou para três ou mais grupos (teste de Kruskal-Wallis), baseado no número de categorias da variável informada no campo "Grupos" da aba "Campos".

Atenção! É necessário classificar corretamente as variáveis quanto ao seu tipo (escalar, ordinal ou nominal) para o teste ser aplicado adequadamente. Reveja os conceitos no capítulo 1. Você também pode definir o teste manualmente, selecionando a opção

“Kruskal-Wallis 1-way ANOVA (amostras k)” em “Customizar testes”. Por este caminho, é possível alterar o nível de significância (α) em “Opções de teste”. Por fim, clique em “Executar” ou “Colar” para obter o resultado do teste estatístico.

O teste utilizado é informado no arquivo de saída, junto com a hipótese de nulidade, o p-valor e a decisão do teste (baseada em uma significância de 5%). Conforme informado pelo teste de hipótese (*Hypothesis Test Summary*), com $p = 0,346$, deve-se reter H_0 de que a distribuição de massa corporal entre as categorias do IPAQ é a mesma. Ou seja, podemos afirmar que não há diferença significativa na massa corporal entre as categorias de atividade física avaliadas pelo IPAQ ($p = 0,346$). Podemos, ainda, descrever os dados em uma tabela (Tabela 5).

Tabela 5. Comparação da massa corporal entre as categorias do IPAQ. Os dados são representados por mediana e quartis (25-75). As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

Categorias do IPAQ	Sedentário (n = 4)	Insuf. Ativo B (n = 8)	Insuf. Ativo A (n = 8)	Ativo (n = 10)	P-valor
Massa corporal (em kg)	126,0 (120,5-131,5)	1120,0 (100,8-131,0)	108,9 (99,5-118,0)	110,5 (101,0-119,0)	0,346*

IPAQ, questionário internacional de atividade física; n, número absoluto.

*Teste de Kruskal-Wallis. Significância estatística considerada de $p \leq 0,050$.

Supondo que o teste de Kruskal-Wallis fosse significativo, isso indicaria que pelo menos duas categorias possuem distribuições diferentes. Mas, quais são as categorias? Para obter esse resultado, deve-se clicar duas vezes sobre o resultado da hipótese (*Hypothesis Test Summary*), no arquivo de saída. Na janela “Visualizador do modelo”, no campo inferior direito, em “Visualização”, selecione a opção “Comparações por pares”. O novo resultado apresentado, informa uma tabela com as comparações par a par e um gráfico que indicará as diferenças entre os pares. A comparação par a par corresponde ao teste de *post hoc* de Dunn (ou Dunn-Bonferroni). Na tabela, na primeira coluna, estará a comparação par a par e, nas

duas últimas colunas, o p-valor e p-valor ajustado (este segundo deve ser levado em consideração). Em algumas situações, é possível que o teste global indique que há diferença entre os grupos, porém, na comparação par a par, nenhuma seja significativa, já que esses são testes independentes.

Referências

Callegari-Jacques SM. Bioestatística: princípios e aplicações. Porto Alegre: ArtMed, 2011. 255p.

Field A. Descobrimo a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Guimarães L.S.P, Hirakata V.N. Uso do Modelo de Equações de Estimções Generalizadas na análise de dados longitudinais. Revista HCPA. 2012;32(4):503-511.

Mancuso ACB, Nune LN, Guimarães LSP, Camey AS, Hirakata VN. Métodos Estatísticos para Desfechos Quantitativos utilizando o SPSS v.18, 2017. Disponível em <https://sites.google.com/a/hcpa.edu.br/bioestatcursos/home/ii>. Acessado em 13/07/2020.

Exercícios sugeridos

1. No “Banco de dados 1 Resolvido.sav”, a idade é diferente entre os alunos do nível mestrado e doutorado?
2. No “Banco de dados 3 Resolvido.sav”, há diferenças nas medidas pressóricas (PAS, PAD e PAM) e lipídicas (colesterol total, triglicerídeos, HDL-Colesterol e LDL-Colesterol) entre os sexos? Construa a tabela apresentando os dados da análise descritiva, o teste estatístico utilizado e o p-valor encontrado e, interprete os resultados para um nível de significância de 5%.
3. No “Banco de dados 3 casos novos.sav”, há diferenças nas medidas pressóricas (PAS e PAD) e lipídicas (colesterol total, triglicerídeos e HDL-Colesterol) entre os sexos? Construa a tabela apresentando os dados da análise descritiva, o teste estatístico utilizado e o p-valor encontrado e, interprete os resultados para um nível de significância de 5%.

8 Desfechos quantitativos: amostras pareadas

*Fernanda Vargas Ferreira
Fernanda Dapper Machado
Edison Capp
Otto Henrique Nienov*

Neste capítulo, iremos abordar os testes paramétricos e não paramétricos indicados para amostras pareadas. Atenção! Antes de decidir qual teste utilizar (paramétrico ou não paramétrico) é necessário verificar a normalidade de distribuição da variável pelo teste de Shapiro-Wilk.

Os testes para amostras pareadas (relacionadas ou emparelhadas), que são consideradas independentes quanto à relação entre os indivíduos, são utilizados quando se objetiva comparar variáveis que tenham alguma dependência, seja por serem do mesmo sujeito de pesquisa ou de um controle emparelhado (por exemplo, por sexo, idade, etc.).

Teste t pareado

É um teste análogo do teste t de Student, para amostras paramétricas relacionadas, ou seja, é utilizado quando a suposição de normalidade é aceita para ambos os grupos pareados. Alguns autores defendem que a normalidade das diferenças entre os pares também seja um pressuposto do teste. Por exemplo, quando o interesse é comparar duas medições (no tempo) em um único grupo, onde é realizado entre elas uma intervenção, a análise recomendada é o teste t pareado.

O teste t pareado avalia se as médias de duas medidas relacionadas são estatisticamente diferentes uma da outra. Neste caso, a hipótese de nulidade de um teste t para amostras pareadas é de que a média das diferenças entre as medidas relacionadas é igual a zero, ou seja, não há diferença entre as medidas.

Além da comparação de tempo, métodos também podem ser comparados pelo teste t pareado, quando estes estiverem na mesma unidade e atenderem aos pressupostos do teste. Por exemplo, o cálculo da medida de consumo energético utilizando o Questionário de Frequência Alimentar comparado à média do Recordatório de 24 horas.

Para exemplificar, no "Banco de dados 4.sav" (disponível em <https://bit.ly/bancosdedados>), vamos comparar as medidas de massa corporal avaliadas em dois momentos ("Massa corporal 1" e "Massa corporal 2") no sexo masculino. Para isso, é necessário dividir o arquivo, no menu "Dados", "Dividir arquivo", "Organizar saída por grupos" baseados pela variável "Sexo". Em seguida, precisamos verificar a distribuição de normalidade das variáveis pelo teste de Shapiro-Wilk, como visto no capítulo 7, no menu "Analisar", "Estatísticas descritivas", "Explorar...".

Com o arquivo dividido, observa-se uma distribuição normal (significância do teste de Shapiro-Wilk superior a 0,05) para o sexo masculino, sendo apropriado o uso do teste t pareado para comparar as médias de massa corporal e, para o sexo feminino, observa-se uma distribuição não-normal (significância do teste de Shapiro-Wilk inferior ou igual a 0,05). Logo, para comparar as medianas de massa corporal no sexo feminino, utilizaremos o teste de Wilcoxon, que será abordado a seguir.

Para realizar o teste t pareado, no menu "Analisar", "Comparar médias", clique em "Teste T de amostras em Pares". Na janela "Teste T de amostras em pares", insira em "variáveis em pares" as variáveis "Massa corporal 1" ("Variável1") e "Massa corporal 2" ("Variável2"). Em seguida, clique em "Ok" ou "Colar".

No arquivo de saída, encontramos os resultados para ambos os sexos. No entanto, considerando a distribuição não normal para o sexo feminino, iremos apenas interpretar o valor do teste para o sexo masculino. Na primeira tabela (*Paired Samples Statistics*), temos a análise descritiva, com média (*Mean*), frequência absoluta (*N*), desvio padrão (*Std. Deviation*) e erro padrão da média (*Std Error Mean*). Em seguida (*Paired Samples Correlations*), é apresentado o coeficiente de correlação de

Pearson ($r = 0,885$) e o p-valor ($p < 0,001$). Os testes de correlação serão abordados no capítulo 9.

E, finalmente, na terceira tabela (*Paired Samples Test*), encontra-se o resultado do teste t pareado. São apresentados média, desvio padrão, erro padrão da média e intervalos de confiança calculados a partir da diferença entre as medidas de massa corporal. A coluna "t" traz o cálculo da diferença média dividida pelo seu desvio-padrão e, em "df", temos o grau de liberdade do teste. O valor de p do teste t de amostras pareadas se encontra na última coluna da tabela (*Sig. (two-tailed)*).

Na interpretação, com p-valor inferior a 0,05, há diferença significativa entre as médias de massa corporal no sexo masculino, ou seja, a massa corporal após a cirurgia bariátrica ($96,6 \pm 14,3$ kg) diminui significativamente em comparação à massa corporal basal ($133,3 \pm 16,7$ kg, $p < 0,001$). Podemos apresentar o resultado em uma tabela (Tabela 1).

Tabela 1. Comparação da massa corporal antes e após a cirurgia bariátrica no sexo masculino. Os dados são representados por média \pm desvio padrão. As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Tempo 1 (n = 31)	Tempo 2 (n = 31)	P-valor*
Massa corporal (em kg)	133,3 \pm 16,7	96,6 \pm 14,3	<0,001

n, número absoluto.

*Teste t pareado. Significância estatística considerada de $p \leq 0,050$.

Teste de Wilcoxon

O teste de Wilcoxon é uma opção de teste não paramétrico para variáveis contínuas, quando se deseja comparar duas amostras emparelhadas, sendo substituto do teste t pareado. A hipótese de nulidade deste teste é que a mediana das diferenças entre os pares é igual a zero. Este teste avalia a diferença entre a quantidade de pares com diferenças positivas e a quantidade de pares com diferenças negativas. O teste de Wilcoxon não é recomendado quando há muitos pares com diferença igual a

zero, pois estes acabam por reduzir o tamanho da amostra, sendo excluídos da análise.

Como as variáveis “Massa corporal 1” e “Massa corporal 2” apresentaram distribuição não normal no sexo feminino pelo teste de Shapiro-Wilk, utilizaremos o teste de Wilcoxon para avaliar a diferença entre as medianas de massa corporal avaliadas antes e após a cirurgia bariátrica. Para isso, no “Banco de dados 4.sav”, no menu “Analisar”, “Testes não paramétricos”, clique em “Amostras relacionadas”. Na janela “Testes Não Paramétricos: Duas ou Mais Amostras Relacionadas”, na aba “Objetivo”, a opção padrão do SPSS é “Compara automaticamente dados observados a hipotetizados”. Clique na aba “Campos”, selecione e insira as variáveis a serem testadas (“Massa corporal 1” e “Massa corporal 2”) em “Testar Campos”, com o auxílio do botão da seta. É importante destacar que somente variáveis classificadas como escalares poderão ser adicionadas ao teste.

Assim como mencionado nos testes para amostras independentes do capítulo 7, o teste não paramétrico é muito simples. Na aba “Configurações”, deve-se marcar a opção “Escolhe automaticamente os testes com base nos dados” e o SPSS irá escolher automaticamente se realizará o teste para dois grupos (teste de Wilcoxon) ou para três ou mais grupos (teste de Friedman) baseado no número de variáveis no campo “Testar campos”. Você também pode definir o teste, selecionando a opção “Wilcoxon (2 amostras)” em “Customizar testes”. Por este caminho, é possível alterar o nível de significância (α) em “Opções de teste”. Note que também temos as opções de testes para dados binários (dicotômicos): testes de McNemar e Q de Cochran, vistos no capítulo 6. Por fim, clique em “Executar” ou “Colar” para obter o resultado do teste estatístico.

O teste utilizado é informado na saída, bem como a hipótese de nulidade, o p-valor e a decisão baseada em uma significância de 5%. Novamente, teremos os resultados de teste para ambos os sexos, mas iremos considerar apenas o resultado para o sexo feminino. Perceba que não obtemos a análise descritiva, somente o resultado do teste de hipótese (*Hypothesis Test Summary*). A mediana e percentis ou valores mínimo e máximo serão extraídos

das tabelas com as análises descritivas (*Descriptives*) e percentis (*Percentiles*) solicitados no teste de normalidade. Conforme é indicado pelo teste de hipótese, rejeita-se H_0 de que a diferença entre as medianas de massa corporal seja igual a zero (0). Ou seja, com p-valor inferior a 0,001, há diferença significativa nas medianas de massa corporal entre os tempos pré- e pós-operatório no sexo feminino. A mediana de massa corporal basal foi de 106,5 (98,0-115,7) kg e a do pós-operatório foi de 78,4 (71,2-87,2) kg. Podemos apresentar os dados em uma tabela (Tabela 2).

Tabela 2. Comparação da massa corporal antes e após a cirurgia bariátrica no sexo feminino. Os dados são representados por mediana e percentis (25-75). As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Tempo 1 (n = 164)	Tempo 2 (n = 164)	P-valor
Massa corporal (em kg)	106,5 (98,0-115,7)	78,4 (71,2-87,2)	<0,001*

n, número absoluto.

*Teste de Wilcoxon. Significância estatística considerada de $p \leq 0,050$.

Ao darmos um duplo clique no teste de hipóteses no arquivo de saída, uma nova janela de “Visualizador de modelo” mostrará um histograma contendo os deltas (ou diferenças) da massa corporal, o percentual do delta, o número de diferenças negativas, positivas e iguais a zero e, apesar de não indicarem o tamanho de efeito, essas informações podem indicar a direção do efeito.

Análise de variância para medidas repetidas (ANOVA-MR)

A análise de variância para medidas repetidas (ANOVA-MR) é uma extensão da ANOVA abordada no capítulo anterior, que pode ser utilizada para análises longitudinais. Medidas repetidas é o termo utilizado quando o sujeito analisado participa de todas as condições do experimento. Na ANOVA-MR, a variável resposta é acompanhada ao longo do tempo, criando-se um modelo longitudinal com medidas repetidas.

Para este modelo, é necessário respeitar os seguintes pressupostos: a) independência, no sentido de que cada participante é independente, não interferindo no resultado do outro, assim como as variáveis não interferem umas nas outras; b) normalidade, avaliada em cada medida no tempo da variável pelo teste de Shapiro-Wilk; c) homocedasticidade, que sugere a igualdade de variância, indicada pelo teste de Levene; e d) esfericidade, medida pela esfericidade de Mauchly e correções. Este último pré-requisito assume que a relação entre os pares das condições experimentais nas quais as medidas foram realizadas é similar. A esfericidade requer que a variância das diferenças de todos os pares seja constante. Se $p \geq 0,05$, não se rejeita a H_0 de que as variâncias sejam iguais e, se $p < 0,05$, a esfericidade não pode ser assumida e, assim, existem diferenças significativas entre as variâncias.

Alguns estatísticos também defendem a conferência da normalidade dos resíduos em análises de ANOVA-MR. Os resíduos são automaticamente salvos no banco de dados como variáveis e, podemos fazer um teste de normalidade conforme vimos no capítulo 7.

Para exemplificar, no "Banco de dados 4.sav", vamos verificar se há diferença entre as medidas de massa corporal pré- e pós-operatórias ("Massa corporal 1", "Massa corporal 2" e "Massa corporal 3") em relação à atividade física basal ("Classificação do IPAQ 1"), supondo a normalidade dos dados. Para isso, não iremos considerar o banco de dados dividido pelo sexo. No menu "Dados", clique em "Dividir arquivo" e selecione "Analisar todos os casos, não criar grupos".

Para a análise de ANOVA-MR, procede-se da seguinte forma: No menu "Analisar", "Modelo Linear Geral", clique em "Medidas Repetidas". Na janela dos "Fator(es) de definição de medida...", em "Número de Níveis", informe o fator (por exemplo, "Massa") e o número de vezes que o fator se repete (neste caso, "3"). Clique em "Adicionar" e, em seguida, em "Definir". Na próxima janela ("Medidas repetidas"), aparecerão três pontos de interrogação ("?_1)", "?_2)", ("?_3)"), na tela "Variáveis dentro de assuntos (fator1)". Esses pontos devem ser substituídos pelas variáveis "Massa 1", "Massa

2" e "Massa 3", respectivamente. Em "Fator(es) entre assuntos", inserimos o fator "Classificação IPAQ 1". Em "Diagramas", no "Eixo Horizontal" vamos inserir "Massa" e, em "Linhas Separadas", o fator "IPAQ1". Clique em "Adicionar" e, então, em "Diagramas", teremos a interação "Massa*IPAQ1". Clique em "Continuar". Em "Post Hoc", selecione em "Fator(es)" o "IPAQ1" e insira em "Testes post hoc para". Na opção "Variâncias iguais presumidas", selecione "Tukey", "Bonferroni", "S-N-K" e "Duncan". Estes testes correspondem ao *post hoc*. Clique em "Continuar". Em seguida, clique em "Opções" e envie para "Exibir médias para" as variáveis ("IPAQ1" e "Massa") e a interação ("IPAQ1*Massa"). Marque a opção "Compartilhar efeitos principais" e selecione na aba "Ajuste do intervalo de confiança" o teste de "Bonferroni". Marque, em "Exibir", as opções "Estatísticas descritivas", "Estimativas de tamanho de efeito", "Potência observada" e "Testes de homogeneidade". Clique em "Continuar" e, por fim, em "Ok" ou "Colar".

No arquivo de saída, na primeira tabela (*Descriptive Statistics*), temos a análise descritiva das medidas repetidas de massa corporal para cada categoria do fator, com valores de média (*Mean*), desvio padrão (*Std. Deviation*) e absoluto (N). Em seguida, é apresentado o teste de Levene (*Levene's Test of Equality of Error Variances*). Se o teste de Levene for significativo ($p \leq 0,05$), os grupos possuem variâncias diferentes. Se o teste de Levene não for significativo ($p > 0,05$), assumimos igualdade de variâncias. Ou seja, não há diferença significativa entre a homogeneidade dos dois grupos. No exemplo, assumimos a igualdade.

Em seguida, temos o teste de esfericidade de Mauchly (*Mauchly's Test of Sphericity*), que nos fornece a estimativa de ϵ , que corresponde à medida de esfericidade. Quando há esfericidade, o valor de ϵ é igual a 1. Valores menores indicam desvio de esfericidade, como neste caso, em que temos um valor de 0,721. Este teste é altamente influenciado pelo tamanho da amostra (falha em demonstrar esfericidade em amostras pequenas e superestima em amostras grandes). Podemos verificar que nesse exemplo o teste de Mauchly foi significativo ($p < 0,001$), ou seja, existe diferença entre as variâncias. Quando isso ocorre, o próprio teste oferece um ajuste, que são as correções de Greenhouse-Geisser e de Huynh-Feldt. A correção de Greenhouse-Geisser

é mais rígida que a de Huynh-Feldt, sendo a preferencialmente utilizada. De um modo geral, recomenda-se o uso de Huynh-Feldt quando $\epsilon > 0,75$ e de Greenhouse-Geisser quando o $\epsilon < 0,75$.

Na ANOVA-MR são incluídos os resultados das análises uni e multivariadas que, em alguns casos, podem apresentar-se discordantes. Quando há esfericidade, os resultados da análise univariada têm mais poder e, do contrário, pode-se usar os resultados multivariados quando não há esfericidade. No quadro seguinte (*Tests of Between-Subjects Effects*), temos a análise de variância da ANOVA-MR, com os valores da esfericidade assumida (*Sphericity Assumed*) e corrigida (Greenhouse-Geisser e Huynh-Feldt) para a interação e o fator. Para ambos, são expressos o grau de liberdade (df), a média quadrada (*Mean Square*), o valor de F (F) e de significância (Sig.). Como visto no capítulo 7, também obtemos o coeficiente de correlação (*Partial Eta Square*), que é uma medida de tamanho de efeito, e o poder do teste (*Observed Power*), que indica a probabilidade de rejeitar corretamente a hipótese de nulidade. No exemplo, rejeita-se a H_0 do fator porque há pelo menos uma diferença entre as medidas repetidas ($p < 0,001$), mas para a interação aceitamos a H_0 de que não existe diferença ($p = 0,518$). Ou seja, na interação das massas corporais com o IPAQ, não houve diferença. Quando a interação é estatisticamente significativa, ela indica que o padrão de mudança entre os tempos é diferente entre os grupos.

Nos quadros seguintes (*Pairwise Comparisons*), são identificadas quais as medidas que diferem entre si, na avaliação *post-hoc*. As análises *post-hoc* só devem ser verificadas quando a ANOVA-MR tiver um valor de $p \leq 0,05$. Lembre-se: os testes *post-hoc* consistem na análise por pares dos grupos. A escolha do teste *post-hoc* é definida levando-se em consideração diversos fatores. Considerando os testes utilizados na análise, os testes de Bonferroni e Tukey são considerados bastante conservadores, porém, o teste de Tukey requer que os grupos tenham tamanhos iguais e permite análise por grupos. Os dois fazem a comparação por pares com intervalo de confiança 95%. O teste de Tukey é indicado quando há esfericidade. Quando tal pressuposto não é assumido, utiliza-se Bonferroni. Já os testes de Duncan e S-N-K são considerados liberais, porém também exigem grupos de

tamanhos iguais e não fornecem a comparação por pares com intervalo de confiança de 95%.

Sendo assim, como mencionado, os valores de p não foram significativos para a interação. No gráfico (Figura 1), é possível visualizar que as linhas da massa corporal seguem de forma semelhante nos cinco grupos da classificação do IPAQ. Em outras palavras, a massa corporal (efeito principal) diminui significativamente conforme o tempo ($p < 0,001$), porém, comparando os grupos (categorias do IPAQ), a massa corporal diminui da mesma forma (sem diferença estatisticamente significativa).

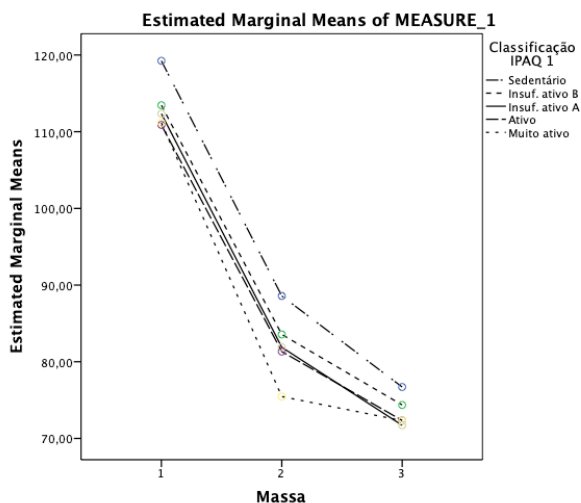


Figura 1. Comparação das medidas de massa corporal pré- e pós-operatórias em relação às categorias de atividade física basal avaliadas pelo IPAQ.

Uma amostra com muitos valores omissos limita o uso da ANOVA-MR, pois sujeitos com uma medida faltante são excluídos da análise. Ao se reportar os resultados de uma análise de ANOVA-MR, é importante incluir as médias, desvio padrão, valores de p para efeitos principais e interações.

Equações de estimativas generalizadas (GEE)

A análise de Equações de Estimativas Generalizadas (GEE) é utilizada para dados pareados e longitudinais. Esta análise tem sido considerada melhor que a ANOVA-MR, que possui pré-requisitos difíceis de serem cumpridos, além de outras limitações, tais como: o tempo, que é considerado uma variável categórica; os dados devem estar balanceados, ou seja, não deve existir nenhuma falha na coleta de dados de nenhum indivíduo e; as covariáveis ou os fatores envolvidos devem ser fixos, não podendo variar com o tempo.

A análise de GEE se baseia na metodologia dos modelos lineares generalizados, permitindo que sejam trabalhados desfechos categóricos ou quantitativos. As covariáveis ou fatores estudados podem ser avaliados novamente, mudando seus valores conforme o passar do tempo, acrescentando-se no modelo essa variação. Além disso, a variável tempo pode ser usada como uma covariável, ou seja, pode ser analisado o acréscimo de uma unidade de tempo no efeito do desfecho. Dessa forma, o método de análise de dados longitudinais pelo GEE possibilita a análise de desfechos contínuos, mesmo quando a variável não apresenta distribuição normal ou esfericidade.

A aplicação da GEE no SPSS exige uma reestruturação do banco de dados para um formato chamado de dados empilhados. No exemplo, iremos transpor as medidas de massa corporal dos casos, que se encontram nas colunas, para as linhas. Na execução da GEE, recomenda-se reformular o banco de dados separadamente para cada variável a ser analisada. Porém, é possível reformular e incluir várias variáveis para serem analisadas de uma só vez. Na reformulação, o SPSS multiplica cada caso de acordo com o número de tempos a serem avaliados, incluindo todas as avaliações em uma única variável.

Para exemplificar, no "Banco de dados 4.sav", será utilizada GEE para comparar as três medidas de massa corporal ("Massa corporal 1", "Massa corporal 2" e "Massa corporal 3") no sexo masculino, pois as variáveis apresentaram distribuição normal pelo teste de Shapiro-Wilk. Mas primeiro, é preciso reestruturar o banco de dados. Para reestruturar o banco de dados, no menu

“Dados”, clique em “Reestruturado”. Na janela “Assistente de reestruturação de dados”, haverá três opções de reestruturação: a primeira transforma um banco horizontal em vertical (opção que será utilizada), a segunda opção modifica um banco vertical em horizontal e a terceira troca de lugares colunas e linhas, transpondo o banco. Selecione a opção “Reestruturar variáveis selecionadas em casos” e clique em “Próximo”.

Na etapa 2 de 7, especifica-se o número de grupos de variáveis que serão reestruturados. Se somente um, selecione a primeira opção (neste caso) ou, se mais de um, a segunda opção e indique o número de grupos de variáveis a serem reestruturados. Clique em “Próximo”. Como visto no capítulo 3, no banco de dados, deve-se ter uma variável que identifique os casos da amostra, de preferência numérica, para organizar o banco de dados e facilitar a localização dos casos. Assim, em “Identificação de grupo de caso”, selecione “Usar variável selecionada” e indique a variável referente à identificação dos casos (“Identificação do paciente”). Em “Variável de destino”, selecione “trans 1” e insira as três medidas de massa corporal (“Massa 1”, “Massa 2” e “Massa 3”), que corresponderão ao grupo de variáveis. Caso seja reestruturado mais de um grupo de variáveis, alternamos para “trans 2” e selecionamos as medidas da variável a serem reestruturadas. Em seguida, repetimos o processo em “trans 3” e, assim, sucessivamente. No último campo da etapa 3 de 7, devem ser inseridas as variáveis fixas, que permanecem constantes no tempo. Neste caso, a variável “Sexo” será selecionada. Em seguida, clique em “Próximo”. Na etapa 4 de 7, indica-se o número de variáveis auxiliares que serão utilizadas. Neste caso, somente uma. Clique em “Próximo”.

Como teremos uma variável auxiliar, precisamos informar o tipo de valor de índice que será utilizado e, pode-se editar o nome e o rótulo dessa variável. Há duas escolhas para os valores da variável índice: 1) colocar uma sequência numérica (valores de índice 1, 2, 3) e; 2) utilizar os nomes das variáveis que estamos reestruturando. A segunda opção é indicada quando trabalhamos com a reestruturação de mais de uma variável. O nome da variável auxiliar pode ser editado no campo “Editar o nome e o rótulo da variável de índice”. Indique o nome da variável índice como “Tempo”, por exemplo. Clique em “Próximo”.

Na etapa 6 de 7, o padrão do SPSS é reestruturar o banco de dados em um novo arquivo. Em relação a valores omissos, a opção padrão é considerar todos os sujeitos, mesmo que o desfecho tenha a casela vazia. Uma opção é descartar os valores omissos do banco de dados. Clique em "Próximo". Por fim, na última etapa, é possível salvar a sintaxe desse comando ("Colar a sintaxe gerada pelo assistente em uma janela de sintaxe"). Clique em "Concluir". Selecione o comando no arquivo de sintaxe e execute-o.

No banco de dados, o resultado é reestruturado em duas variáveis, uma para o desfecho ("trans1") e outra indicando o tempo (1, 2, 3), além das variáveis fixadas ("Identificação do paciente" e "Sexo"). Agora, cada sujeito apresenta três linhas no banco de dados. É importante salvar esse banco de dados reestruturado com outra denominação, para não se perder os dados originais do banco de dados.

Com o banco de dados reestruturado, podemos realizar a análise de GEE. Para isso, vamos novamente dividir o arquivo por sexo, no menu "Dados", "Dividir arquivo" e "Organizar a saída por grupos" baseados pela variável "Sexo". Em seguida, no menu "Analisar", "Modelos Lineares Generalizados", selecione "Equações de estimativas generalizadas". Na janela "Equações de estimativa generalizadas", na aba "Repetido", defina a variável que identifica o sujeito ("Identificação do paciente") em "Variáveis de assunto", e insira a variável "Tempo" em "Variáveis dentro do assunto".

Na aba "Tipo de Modelo", defina a distribuição da variável dependente como "Linear" em "Resposta de escala", pois estamos trabalhando com variável de distribuição normal (teste de Shapiro-Wilk). Na aba "Resposta", selecione como "Variável dependente" a "Massa 1 [trans1]". Na aba "Preditores", selecione a variável categórica ("Tempo") em "Fatores" e, se necessário, pode-se inserir as variáveis contínuas em "Covariáveis". Na aba "Modelo", selecione a variável "Tempo" em "Fatores e covariáveis", selecione o "Tipo: fatorial", e a insira em "Modelo". Caso haja outras variáveis no modelo, pode-se criar uma variável de interação entre as variáveis ("Por*") na opção "Criar termo aninhado" e incluí-la no modelo. Na aba "Médias EM", selecione os "Fatores e interações" e insira em "Exibir médias para",

selecionando a opção de comparação “Por par” em “Contraste”. Em “Ajuste para comparações múltiplas”, selecione a opção “Bonferroni”. Clique em “Ok” ou “Colar”.

No arquivo de saída, encontramos os resultados para ambos os sexos. No entanto, considerando a distribuição não normal para o sexo feminino, iremos apenas interpretar o valor do teste para o sexo masculino. No primeiro quadro (*Case Processing Summary*), são indicados os valores absoluto (N) e relativo (Percent) de casos incluídos, excluídos e total. No segundo quadro (*Categorical Variable Information*), temos os valores absoluto (N) e relativo (Percent) para cada medida no tempo da massa corporal. Também são apresentados os valores mínimo e máximo, a média e o desvio padrão para a variável dependente (*Continuous Variable Information*).

No quadro seguinte (*Tests of Model Effects*), o teste de Qui-quadrado de Wald fornece o valor de p para efeitos principais e interação. Neste exemplo, somente o efeito principal foi avaliado, pois não houve interação com outra variável. Esse efeito principal foi significativo ($p < 0,001$), ou seja, ao longo do tempo houve uma mudança significativa da massa corporal no sexo masculino.

A partir desse resultado, precisamos verificar a análise de *post hoc* para definir onde ocorre essa mudança e como ocorre. No quadro “*Estimates*”, são apresentados a média (*Mean*), erro padrão (*Std. Error*) e intervalo de confiança de 95% para cada medida de massa corporal no tempo (1, 2 e 3). Em seguida (*Pairwise Comparisons*), temos os resultados do teste de Bonferroni. Fixando-se os tempos, há diferença significativa entre as três medidas de massa corporal. Com p-valor menor que 0,001, há diferença significativa entre cada medida de massa corporal nos três tempos para o sexo masculino. Ou seja, cada medida de massa corporal no tempo é diferente uma da outra. Lembre-se: quando a interação for significativa, os efeitos principais não devem ser analisados. Podemos apresentar os resultados da análise em uma tabela (Tabela 3), utilizando o sistema de letras sobrescritas.

Tabela 3. Comparação da massa corporal antes e após a cirurgia bariátrica no sexo masculino. Os dados são representados por média \pm desvio padrão. As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Tempo 1 (n = 31)	Tempo 2 (n = 31)	Tempo 3 (n = 31)	P-valor
Massa corporal (em kg)	133,3 \pm 16,7 ^a	96,62 \pm 14,3 ^b	85,3 \pm 12,7 ^c	<0,001*

Letras sobrescritas distintas representam média de massa corporal significativamente diferentes.

n, número absoluto.

* GEE. Significância estatística considerada de $p \leq 0,050$.

Teste de Friedman

O teste de Friedman substitui a análise GEE quando a variável estudada não apresentar os pré-requisitos para tal. É utilizado para a comparação de variáveis contínuas entre três ou mais tempos (amostras pareadas). Assim como no teste para amostras independentes, uma diferença significativa é a indicação de que pelo menos dois tempos diferem na variável de interesse. Como nos outros testes não paramétricos, o teste de Friedman também é baseado no ranqueamento dos dados.

Para exemplificar essa análise, no “banco de dados 4.sav”, vamos comparar as medianas de massa corporal (“Massa corporal 1”, “Massa corporal 2” e “Massa corporal 3”) para o sexo feminino, uma vez que o teste de Shapiro-Wilk indicou uma distribuição não normal. Para isso, vamos novamente dividir o arquivo por sexo, no menu “Dados”, “Dividir arquivo” e “Organizar a saída por grupos” baseados pela variável “Sexo”. Em seguida, no menu “Analisar”, “Testes não-paramétricos”, clique em “Amostras relacionadas”.

Na janela “Testes Não Paramétricos: Duas ou Mais Amostras Relacionadas”, na aba “Objetivo”, a opção padrão do SPSS é “Compara automaticamente dados observados e hipotetizados”. Clique na aba “Campos”, selecione e insira as variáveis a serem testadas (“Massa corporal 1”, “Massa corporal 2” e “Massa corporal 3”) em “Testar Campos”, com o auxílio do botão da seta. É importante destacar que somente variáveis classificadas como escalares poderão ser adicionadas ao teste.

Na aba “Configurações”, deve-se marcar a opção “Escolhe automaticamente os testes com base nos dados” e o SPSS irá escolher automaticamente se realizará o teste para dois grupos (teste de Wilcoxon) ou para três ou mais grupos (teste de Friedman) baseado no número de variáveis no campo “Testar campos”. Você também pode definir o teste, selecionando a opção “Friedman (amostras k)” em “Customizar testes”. Por este caminho, é possível alterar o nível de significância (α) em “Opções de teste”. Por fim, clique em “Executar” ou “Colar” para obter o resultado do teste estatístico.

O teste utilizado é informado na saída, bem como a hipótese de nulidade, o p-valor e a decisão baseada em uma significância de 5%. Novamente, teremos os resultados de teste para ambos os sexos, mas iremos considerar apenas o resultado para o sexo feminino. Perceba que não obtemos a análise descritiva, somente o resultado do teste de hipótese (*Hypothesis Test Summary*). A mediana e percentis ou valores mínimo e máximo serão extraídos das tabelas com as análises descritivas (*Descriptives*) e percentis (*Percentiles*) solicitados no teste de normalidade. Conforme é indicado pelo teste de hipótese, rejeita-se H_0 de que a distribuição das medianas de massa corporal seja a mesma. Ou seja, com p-valor inferior a 0,001, há diferença significativa nas medianas de massa corporal entre os tempos pré- e pós-operatórios no sexo feminino. A mediana de massa corporal basal foi de 106,5 (98,0-115,7) kg e, nos períodos pós-operatórios, as medianas foram de 78,4 (71,2-87,2) kg e de 69,3 (63,0-77,8) kg.

Como o resultado do teste de Friedman foi menor que 0,001, ou seja, pelo menos duas medidas diferem entre si, precisamos verificar na análise de *post-hoc* quais medianas de massa corporal são diferentes. Para isso, clique duas vezes sobre o teste de hipótese no arquivo de saída. Na janela de “Visualizador de modelo”, no campo inferior direito, em “Visualização”, deve-se selecionar a opção “Comparações por pares”. Neste caso, a massa corporal foi diferente em todas as comparações por pares. Podemos apresentar os dados em uma tabela (Tabela 4), utilizando o sistema de letras sobrescritas.

Tabela 4. Comparação da massa corporal antes e após a cirurgia bariátrica no sexo feminino. Os dados são representados por mediana e percentis (25-75). As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Tempo 1 (n = 164)	Tempo 2 (n = 164)	Tempo 3 (n = 164)	P-valor
Massa corporal (em kg)	106,5 (98,0-115,7) ^a	78,4 (71,2-77,8) ^b	69,3 (63,0-77,8) ^c	<0,001*

Letras sobrescritas distintas representam medianas de massa corporal significativamente diferentes.

n, número absoluto.

*Teste de Friedman. Significância estatística considerada de $p \leq 0,050$.

Vamos supor que, considerando a perda de massa corporal significativa em ambos os sexos, o pesquisador resolva comparar as medidas entre os sexos masculino e feminino. Podemos comparar os dados da forma que estão apresentados: média (distribuição normal) para o sexo masculino e mediana (distribuição não-normal) para o sexo feminino? A resposta é não, pois a distribuição dos dados é diferente, logo, o teste de hipótese e as medidas são diferentes. Neste caso, deve-se optar pela análise não paramétrica e expressar os dados em ambos os sexos pela mediana e percentis.

Tamanho de efeito e sua implicação no cálculo amostral

Determinar o tamanho da amostra, que é um subconjunto da população, é crucial para o desenvolvimento de um estudo por aspectos de viabilidade, custos, estatística, além de possibilitar maior precisão nos resultados. Para o cálculo do tamanho amostral, considere: qual a pergunta da pesquisa? Qual a variável primária? Como será a análise estatística? Qual é o resultado de estimativa no grupo controle? Qual a diferença a ser detectada e o grau de confiança?

Adicionalmente, deve-se entender que o cálculo propiciará a quantidade de sujeitos necessários para responder à pergunta de pesquisa, sendo essencial ainda identificar quatro fatores: 1) tamanho de efeito que pode ser obtido utilizando a média e a variância; 2) distribuição da amostra; 3) poder do teste e; 4) nível de significância.

Para se calcular o tamanho da amostra, pode-se: a) realizar um estudo-piloto de tamanho pequeno definido arbitrariamente e viável de ser investigado em curto espaço de tempo; b) comparar duas médias quando o desfecho é medido através de escores de escalas; c) comparar duas proporções (por exemplo, estudos transversais, coorte); d) risco relativo (RR), que estima a magnitude da associação entre a exposição ao fator de risco e o desfecho, indicando quantas vezes a ocorrência do desfecho nos expostos é maior do que aquela entre os não-expostos; e) razão de chances (*Odds Ratio*), empregado em estudos de caso-controle, é uma estimativa do risco relativo entre o percentual de exposição nos controles (pode-se usar dados da população geral sadia) e o número de controles por cada caso.

Além disso, outros três pontos devem ser destacados no cálculo do tamanho da amostra: a) estudos longitudinais (coortes prospectivas e ensaios clínicos), devido às características de seguimento dos voluntários por determinado tempo, podem apresentar perdas (por desistência, morte, exclusão), recomendando por consequência, a correção desse cálculo da amostra inicial com aumento de até 30%; b) mediante mais de um desfecho de interesse, o tamanho da amostra deve ser calculado para cada um deles, adotando-se o maior valor e; c) se houver modificação no delineamento de pesquisa (por exemplo, de transversal para estudo de coorte) será necessário realizar um novo cálculo amostral.

Similarmente, caso a coleta de dados já tenha ocorrido, é possível calcular o tamanho do efeito, a fim de se verificar se a amostra obtida foi válida. Apenas o valor de p ser significativo não garante que o efeito é significativo ou importante. É essencial sinalizar que o tamanho do efeito deve ser entendido como uma estatística descritiva que serve como complemento ao teste de significância estatística da hipótese de nulidade. Ele propicia realizar a medição de uma potencial significância real de um efeito em uma intervenção por meio da descrição do tamanho dos efeitos observados, independentemente de um possível efeito enganoso em função do tamanho da amostra. O tamanho do efeito também é útil para se comparar resultados de estudos com diferentes escalas de medidas. O tamanho do efeito é medido em

unidades de desvio padrão, descrevendo a distância entre dois valores de média.

O cálculo do tamanho de efeito baseado em diferenças padronizadas entre as médias engloba as opções d de Cohen, g de Hedges e Δ de Glass. Por sua vez, se as variáveis independentes possuem mais de duas categorias ou são contínuas, pode-se empregar Epsilon quadrado (ϵ^2), Eta quadrado (η^2), Eta quadrado parcial ($\eta^2 p$), Eta quadrado generalizado ($\eta^2 G$) e f de Cohen. Já quando as variáveis independente e dependente são categóricas, sugere-se ϕ (f), V de Cramer (ou f_c), λ de Goodman-Kruskal e w de Cohen.

Dessa forma, com os dados já coletados, poder-se-ia aplicar a avaliação do tamanho de efeito d de Cohen, obtido pela divisão da diferença das médias pelo desvio-padrão, cuja interpretação seria: tamanho de efeito grande (valores superiores ou iguais a 0,8), tamanho de efeito médio (entre 0,8 a 0,2) e tamanho de efeito pequeno (inferiores a 0,2). Quando a variância dos grupos é similar, pode-se utilizar o desvio padrão de um dos grupos. Caso as variâncias sejam diferentes, utiliza-se o desvio padrão agrupado, calculado como a raiz-quadrada da média dos desvios padrão dos dois grupos. O teste d de Cohen é apropriado para dados de distribuição normal e não indicado para amostras pequenas por superestimar os resultados.

O g de Hedge é utilizado nas amostras pequenas em substituição ao d de Cohen. Este teste também utiliza médias e desvios padrão, porém faz um ajuste para tamanho de amostra. Para estudos experimentais, o tamanho de efeito é calculado pelo Δ de Glass, que é obtido da mesma maneira do teste d de Cohen, porém utilizando o desvio padrão do grupo controle.

O Eta quadrado é um índice de força de associação entre um fator e a variável dependente, sendo a proporção da variação total atribuída ao fator, calculado como a razão da variância do fator sobre a variância total. Os valores obtidos variam de 0 a 1. No SPSS, pode-se obter no menu "Analisar", "Comparar médias", e em "Médias, nas "Opções", selecione "Tabelas ANOVA" e "Eta". A partir do Eta quadrado pode se obter o f de Cohen, dividindo a

raiz quadrada do quadrado de Eta sobre um menos Eta quadrado. Essa conversão é apropriada quando os tamanhos dos grupos são semelhantes. Para o efeito f de Cohen, um valor de 0,1 é considerado um tamanho de efeito pequeno, 0,25 é considerado médio e 0,4 é considerado um efeito grande.

Referências

Barton, B, Peat, J. Medical Statistics: A guide to SPSS. Data analysis and Critical Appraisal. Second edition. John Wiley & Sons, Ltda. 2014. Disponível em: www.wiley.com/go/barton/medicalstatistics2.e

Coutinho ESF, Cunha GM. Conceitos básicos de epidemiologia e estatística para a leitura de ensaios clínicos controlados. Revista Brasileira de Psiquiatria, v.27, n.2, p.146-51, 2005.

Espírito-Santo H, Daniel F. Calcular e apresentar tamanhos do efeito em trabalhos científicos (1): As limitações do $p < 0,05$ na análise de diferenças de médias de dois grupos. Revista Portuguesa de Investigação Comportamental e Social, v.1, n. 1, p. 3-16, 2015.

Field, A. Discovering statistics using SPSS. Third Edition. SAGE, 2009.

Guimarães LSP, Hirakata VN. Uso do Modelo de Equações de Estimções Generalizadas na análise de dados longitudinais. Revista HCPA, v. 32, n.4, p. 503-511, 2012.

Lindenau J Dal-Ri, Guimarães LSP. Calculando o tamanho de efeito no SPSS. Revista HCPA, v. 32, n.3, p. 363-381, 2012.

Miot HA. Tamanho da amostra em estudos clínicos e experimentais. Jornal Vascular Brasileiro, v.10, n.4, p 275-278, 2011.

Pallant, J. SPSS Survival Manual. Mc Graw Hill Open University Peress. 2007.

Pereira DA de A et al. Aspectos metodológicos na construção de projetos de pesquisa em Nutrição Clínica. Revista de Nutrição, v. 27, n. 5, p. 597-604, 2014.

Exercícios propostos

1. No "banco de dados 4.sav", há diferença nas medidas de circunferência da cintura e pressórica (PAS e PAD) entre os três tempos avaliados? Construa a tabela apresentando os dados da análise descritiva, o teste estatístico utilizado e o p-valor encontrado e, interprete os resultados considerando um nível de significância de 5%.

9 Correlações, risco, razão de chances e avaliação de testes diagnósticos

*Johanna Ovalle Diaz
Hellen Meiry Grosskopf Werka
Edison Capp
Otto Henrique Nienov*

Pensemos na seguinte situação hipotética: um professor de uma disciplina de bioestatística percebe que alguns dos seus alunos não estão apresentando um bom desempenho na realização das tarefas e, notando que estes não estavam dedicando tempo suficiente para o estudo do conteúdo, decide fazer um experimento com a turma. O professor solicita aos alunos que relatem o tempo que cada um dedicou em casa ao estudo do conteúdo e, em seguida, monta uma tabela com os dados de tempo de estudo (expresso em horas) e o desempenho obtido nas tarefas. A seguir, explica aos alunos que, se existe um relacionamento entre o tempo dedicado aos estudos e o desempenho nas tarefas, isso poderia facilmente ser observado em um gráfico (Figura 1).

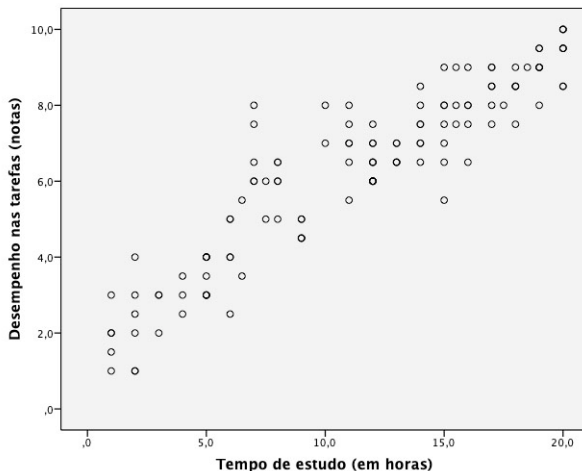


Figura 1. Relação entre o desempenho nas tarefas e o tempo de estudo dos alunos.

Utilizando o eixo X para o tempo de estudo (em horas) e o eixo Y para o desempenho nas tarefas (notas), o professor marca, no gráfico, a nota e o tempo de estudo de cada aluno. Visualmente parece haver alguma relação do tempo de estudo e o desempenho nas tarefas, pois quanto maior o tempo de estudo, melhor tende a ser o desempenho do aluno. Mas, como confirmar e quantificar essa relação?

Correlações bivariadas

Muitas vezes, o objetivo do estudo não é apenas observar se dois grupos possuem a mesma média, mas também se variam conjuntamente, se há alguma relação. Nestes casos, a estatística de interesse é o coeficiente de correlação. A correlação mede como as ordens de variáveis ou de ranqueamento estão relacionadas, ou seja, representa um relacionamento entre duas variáveis. Pode ser analisada por gráficos de dispersão e por coeficientes de correlação. No entanto, antes de calcularmos o coeficiente de correlação, é necessário verificar se os dados possuem valores discrepantes, que podem resultar em conclusões equivocadas.

Os gráficos de dispersão são outro método para avaliar a correlação entre duas variáveis de forma visual. Organiza-se um gráfico cartesiano de pontos e cada ponto corresponde a um sujeito do estudo, cujo posicionamento no gráfico representa os valores para as variáveis localizadas no eixo X e Y, respectivamente. O comportamento dos pontos como conjunto permite interpretar se existe ou não correlação entre as variáveis.

A principal vantagem da correlação é que conseguimos demonstrar o comportamento do relacionamento: se positivo ou negativo (inverso). Além disso, podemos avaliar qualitativamente quanto a sua intensidade, conforme apresentado no quadro 1. O valor do intervalo de coeficientes de correlação varia de -1 (relacionamento negativo ou inverso perfeito) e +1 (relacionamento positivo ou direto perfeito). Um valor de zero ("0") indica nenhum relacionamento. No gráfico, quando existe uma correlação perfeita (-1 ou +1), os pontos formam uma linha reta inclinada. Quando positiva (ou direta), X e Y variam no mesmo sentido, ou seja, os

dois aumentam. Quando negativa (ou inversa), X e Y variam em sentidos opostos, ou seja, quando X aumenta, Y diminui.

Quadro 1. Avaliação qualitativa do grau de correlação entre duas variáveis.

Coeficiente de correlação	Intensidade da correlação
0,000	Nula ou inexistente
0,001 a 0,299	Fraca
0,300 a 0,599	Regular ou moderada
0,600 a 0,899	Forte
0,900 a 0,999	Muito forte
1,000	Plena ou perfeita

O que a correlação procura entender é como uma variável se comporta em um cenário onde outra está variando, visando identificar se existe alguma relação entre a variabilidade de ambas. Portanto, devemos ser cuidadosos ao interpretar coeficientes de correlação porque eles não dão indicação da direção da causalidade. As opções de testes estatísticos variam de acordo com a natureza das variáveis de estudo (Quadro 2).

Quadro 2. Tipos de correlação baseado na forma e comportamento das variáveis.

Correlação	Coeficiente	Tipo de variável
Pearson	r	Quantitativas normais
Spearman	ρ	Quantitativas não normais e qualitativas ordinais
Kendall	τ	Quantitativas não normais e qualitativas ordinais

Coeficiente de correlação de Pearson

O coeficiente de correlação de Pearson, também chamado de correlação linear ou r de Pearson, denominado pela letra r , descrito em 1896 por Karl Pearson, é uma medida de associação linear (simétrica), ou seja, permite avaliar a força de associação entre duas variáveis lineares. Isso quer dizer que as variáveis quantitativas devem estar normalmente distribuídas.

Duas variáveis podem estar perfeitamente relacionadas, mas se o relacionamento não for linear, o coeficiente de Pearson não é o mais apropriado para medir a associação entre as variáveis de estudo.

Coeficiente de correlação de Spearman

O coeficiente de correlação de Spearman, denominado pela letra grega ρ (ρ), descrito em 1904 por Charles Spearman, é uma estatística não paramétrica, que permite avaliar variáveis quantitativas se não estiverem normalmente distribuídas ou se possuírem categorias ordenadas, isto é, variáveis qualitativas ordinais. Uma vantagem do ρ de Spearman é que este não é afetado pelos valores atípicos ou extremos.

Coeficiente de correlação de Kendall

O coeficiente de correlação tau-b de Kendall, denominado pela letra grega tau (τ), descrito em 1938 por Maurice Kendall, é outra correlação não paramétrica, que permite associação para variáveis ordinais. Uma vantagem sobre o coeficiente de Spearman é a possibilidade de ser generalizado para um coeficiente de correlação parcial. Deve ser usada ao invés do coeficiente de Spearman quando temos um conjunto pequeno de dados com um grande número de postos empatados. Isso significa que, se você ordenar todos os escores e muitos deles apresentarem o mesmo posto, o tau de Kendall deve ser usado.

Quando uma das duas variáveis é do tipo nominal, não se pode fazer uma estatística de correlação. O que se costuma fazer nestas situações é aplicar um teste de Qui-quadrado de associação. No entanto, é possível calcular o coeficiente de correlação entre

duas variáveis, uma numérica e outra nominal, quando esta for dicotômica. O primeiro passo é codificar a variável nominal. Então, para se calcular a correlação, procede-se com o cálculo do coeficiente de correlação de Pearson da forma usual, que é chamado neste caso de coeficiente de correlação ponto-biserial.

Para exemplificar, vamos utilizar a situação hipotética mencionada no início deste capítulo, na qual um professor deseja verificar se existe uma relação entre o tempo de estudo (em horas) e desempenho nas tarefas (notas) da turma. Para isso, vamos utilizar o “Banco de dados 5.sav” (disponível em [https:// bit.ly/ bancosdedados](https://bit.ly/bancosdedados)). Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Primeiro, é necessário verificar a normalidade das variáveis “Tempo de estudo” e “Desempenho nas tarefas” através do teste de Shapiro-Wilk, visto no Capítulo 7. Observa-se que as variáveis apresentam distribuição não normal, logo, realizaremos uma correlação de Spearman. O procedimento para calcular os coeficientes de correlação de Pearson, r de Spearman e tau-b de Kendall são os mesmos, assim como a interpretação é feita da mesma forma.

No menu “Analisar”, “Correlacionar”, clique em “Bivariável...”. Na janela “Correlações de bivariável”, localize na tela da esquerda as variáveis “Tempo de estudo” e “Desempenho nas tarefas”, selecione-as e, com o auxílio do botão da seta, movas para a lista de “Variáveis”. O padrão do SPSS é a correlação de Pearson. Em “Coeficientes de correlação”, marque a opção “Spearman”. Em “Teste de Significância”, é possível selecionar se o teste é ou não de uma ou duas caudas (visto no Capítulo 5). Se a direção da associação for conhecida, ou seja, quando você tiver uma hipótese direcional, selecione “Unilateral”. Caso contrário, se você não souber a natureza do relacionamento, selecione “Bilateral”. Por fim, clique em “OK” ou “Colar”.

No arquivo de saída, é apresentado uma matriz dos coeficientes de correlação (*Correlations*) para as duas variáveis, contendo o valor do coeficiente de correlação r de Spearman, o p-valor e o tamanho da amostra (N). Para cada par de variáveis é apresentado um coeficiente de correlação. Cada variável é perfeitamente correlacionada com ela mesma, tendo um coeficiente

igual a 1,000. O coeficiente da correlação bivariada é encontrado na intersecção das duas variáveis de estudo. Os coeficientes de correlação significativos no nível de significância de 5% são identificados com um único asterisco e, os significativos no nível de significância de 1%, são identificados com dois asteriscos.

Desta forma, temos uma correlação positiva muito forte entre o tempo de estudo e o desempenho nas tarefas, com um coeficiente de $\rho = 0,915$ e $p\text{-valor} < 0,001$. O $p\text{-valor}$ significativo da correlação de Spearman indica a rejeição da hipótese de nulidade de que o coeficiente de correlação seja igual a zero, evidenciando uma relação significativa entre as variáveis testadas. Logo, o professor conclui que, quanto maior o tempo de estudo do aluno, melhor tende a ser o seu desempenho nas tarefas da disciplina.

Podemos apresentar o resultado da correlação através de um gráfico de dispersão. No menu "Gráficos", "Caixas de diálogo legadas", clique em "Dispersão/Ponto". Na janela "Dispersão/Ponto", marque a opção "Dispersão simples" e clique em "Definir". Na janela seguinte ("Diagrama de dispersão simples"), localize na tela da esquerda a variável "Desempenho nas tarefas" e a mova para o "Eixo Y" e, insira no "Eixo X", a variável "Tempo de estudo". Por fim, clique em "OK" ou em "Colar".

Novamente, no arquivo de saída, temos o gráfico de dispersão (Figura 1), onde percebemos que há uma relação significativa entre as variáveis testadas. Os gráficos de dispersão são úteis para visualizarmos a força de um relacionamento entre duas variáveis. O eixo X representa a variável independente e o eixo Y representa a variável dependente. Além disso, o gráfico nos permite verificar como as variáveis estão distribuídas e onde estão os valores atípicos.

Um gráfico de dispersão pode utilizar análise de regressão para estimar a força e direção do relacionamento entre as variáveis dependente e independente. Ao clicarmos duas vezes sobre o gráfico, abre-se uma janela de "Editor de gráfico", onde é possível editá-lo. Nesta janela, é possível "Adicionar uma linha de ajuste no total", clicando com o botão direito sobre o gráfico, que permite analisar o comportamento da dispersão dos dados. Também se obtém o coeficiente de determinação (R^2), que corresponde

ao quadrado do coeficiente de correlação e informa quanto da variação da variável dependente é explicada pela variável independente. No exemplo, a variabilidade do desempenho nas tarefas é explicada em 84,4% pelo tempo de estudo dos alunos. Além de tabelas e gráficos, podemos apresentar os resultados da correlação entre duas variáveis através de setas de oposto, equivalente e contrapeso (Figura 2).

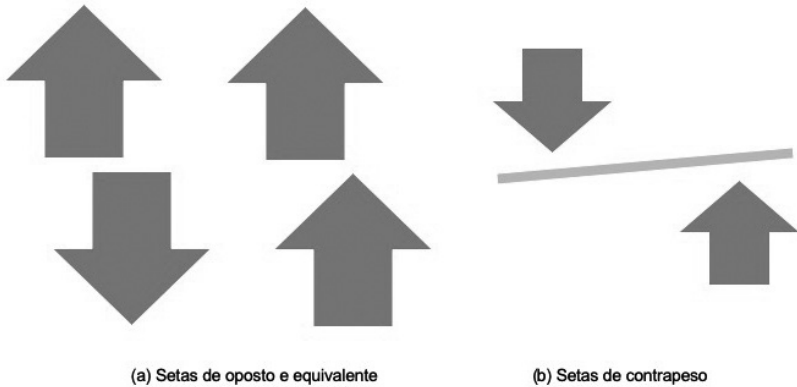


Figura 2. Apresentação de correlação entre duas variáveis.

Risco relativo e razão de chance

Os testes do Qui-quadrado indicam se duas variáveis dicotômicas estão significativamente relacionadas entre si. No entanto, não fornecem uma estatística para descrever a força do relacionamento. Para tabelas 2x2, também é possível calcular uma medida da intensidade da associação entre a presença de um fator e a ocorrência de um evento, ou seja, quanto da probabilidade de ocorrência da variável dependente (doença ou sem doença) deve-se à sua relação com a variável independente (exposição ou não exposição).

O risco relativo (RR) e a razão de chance (*Odds Ratio*, OR) representam as medidas desta associação. Ambas as estatísticas são estimativas de risco e, como tal, descrevem a probabilidade de que os indivíduos expostos a um determinado fator tenham uma doença em comparação com os indivíduos que não estão

expostos ao mesmo fator. A escolha do uso de RR ou de OR depende do desenho do estudo.

Quando o numerador corresponde a eventos que ocorrem de maneira prospectiva, como em estudos experimentais e de coorte prospectiva, a medida de associação denomina-se RR. O RR corresponde a medida da força da associação entre a exposição e o evento (desfecho) e mede a incidência ou risco da doença (variável dependente) acontecer entre os expostos e os não expostos (à variável independente) num período de tempo.

Outras análises podem ser realizadas em estudos que avaliam o efeito de um tratamento. A Redução do Risco Absoluto (RRA) representa a diferença do risco entre os expostos e não expostos e equivale a quantos eventos a menos se produzem ao expor ou não os sujeitos. A Redução do Risco Relativo (RRR) é a diferença do risco entre os expostos e não expostos, mas com denominador constituído apenas pelos não expostos. É possível calcular também o Número Necessário a Tratar (NNT), que equivale ao número de indivíduos que se deve tratar para evitar o evento indesejado.

Quando o numerador corresponde a eventos que ocorreram, como em estudos de caso-controle e de coorte retrospectiva, a medida de associação denomina-se OR e é definido como a probabilidade de que um evento ocorra dividido pela probabilidade de que ele não ocorra. O OR pode ser utilizado como uma estimativa.

Como mencionado, para a análise dos RR e OR é necessário realizar a análise da distribuição de frequência em tabelas de contingência 2x2. Portanto, é necessário que as variáveis sejam qualitativas dicotômicas. A tabela 1 indica a estrutura de organização da tabela de contingência 2x2 para as análises de RR e OR. Como no SPSS a tabela de contingência é organizada de acordo com a ordem numérica dos códigos atribuídos para as categorias de resposta, é necessário verificar a apresentação da tabela antes de realizar a análise. Caso necessário, recodifica-se as variáveis, como visto no Capítulo 4, para obtermos a tabela de contingência corretamente.

Tabela 1. Tabela de contingência 2x2 para as análises de RR e OR.

	Doentes	Não-doentes	Total
Expostos	A	B	A+B
Não-expostos	C	D	C+D
Total	A+C	B+D	A+B+C+D

Onde:

$$RR = A*(A+B)/C*(C+D)$$

$$RRA = [A/(A+B)] - [C/(C+D)]$$

$$RRR = (1 - RR)*100$$

$$NNT = 1/RRA$$

$$OR = (A/B)/(C/D) = (A*D)/(B*C)$$

Como os resultados são decorrentes de uma razão, o valor de nulidade para esses é de 1, ou seja, se o intervalo de confiança incluir um valor inteiro igual a 1, não será possível supor que o fator está associado ao evento. Logo, significa ausência de associação e o risco de acontecer a doença é idêntico, independe da presença ou não da exposição. Se o resultado é > 1 , há uma associação possivelmente causal, ou seja, o risco aumenta com a presença da exposição. Se < 1 , há uma associação possivelmente de proteção, ou seja, o risco diminui na presença da exposição. A interpretação é a mesma, no entanto, no RR avaliamos o risco, enquanto que no OR avaliamos a probabilidade ou chance de o evento ocorrer.

Para exemplificar, vamos utilizar o “Banco de dados 6.sav” (disponível em [https:// bit.ly/bancosdedados](https://bit.ly/bancosdedados)), que contém variáveis qualitativas dicotômicas como tabagismo, câncer de pulmão, obesidade, hipertensão arterial e diabetes mellitus; variáveis quantitativas como massa corporal, IMC, glicemia de jejum, HDL-Colesterol; além de dois escores de neuropatia periférica, calculados através dos instrumentos MNSI e NDS (*Neuropathy Disability Score*). Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Vamos supor que um estudo de coorte acompanhou 100 indivíduos com hábito de fumar e 100 indivíduos sem o hábito

de fumar para observar o desenvolvimento de câncer no pulmão. Pensando na estrutura de organização da tabela de contingência 2x2, a doença é o câncer no pulmão, a exposição é o hábito de fumar, o total de indivíduos de cada grupo é 100 e a amostra total do estudo é 200. Ao final do estudo, observou-se que 33 tabagistas e 7 não-tabagistas desenvolveram câncer no pulmão (Tabela 2). Portanto, iremos verificar o risco de ocorrência de câncer no pulmão entre tabagistas e não tabagistas.

Tabela 2. Tabela de contingência 2x2 para avaliação do risco de ocorrência de câncer no pulmão em indivíduos com hábito de fumar.

	Com câncer de pulmão	Sem câncer de pulmão	Total
Tabagista	33	67	100
Não-tabagista	7	93	100
Total	40	160	200

Primeiro, devemos lembrar que no SPSS a tabela de contingência é organizada de acordo com a ordem numérica dos códigos atribuídos pelo pesquisador. Por isso, é necessário verificar se a combinação de códigos obedece a estrutura de organização da tabela de contingência 2x2. Para isso, no menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Na janela de "Tabulações cruzadas", localize na tela da esquerda a variável "Tabagismo" (exposição) e a insira na "Linha(s)". Em seguida, na mesma tela, selecione a variável "Câncer de pulmão" (desfecho) e a mova para "Coluna(s)". Clique em "Ok" ou "Colar".

No arquivo de saída, é possível perceber que a tabela de contingência 2x2 obtida não corresponde ao formato correto. Portanto, para analisar o RR e o OR precisamos recodificar as variáveis tabagismo ("0" = "Tabagista" e "1" = "Não tabagista") e câncer no pulmão ("0" = "Com câncer" e "1" = "Sem câncer"), para obtermos a estrutura de organização correta da tabela de contingência 2x2.

Como visto no Capítulo 4, no menu "Transformar", clique em "Recodificar em variáveis diferentes...". Na caixa de diálogo "Recodificar em variáveis diferentes", localize na tela da esquerda a

variável "Tabagismo" e mova, com o auxílio do botão da seta, para a lista de "Variável numérica -> Variável de saída". Em seguida, informe um novo nome e rótulo para a variável em "Nome" e "Rótulo", respectivamente. Clique em "Alterar". Clique em "Valores antigo e novo". Na janela seguinte, em "Valor", informe o valor antigo "0" e insira o novo valor "1". Clique em "Adicionar". Em seguida, informe o valor antigo "1" e insira o novo valor "0". Clique em "Adicionar". Por fim, clique em "Continuar" e, em "OK" ou "Colar". A nova variável aparecerá automaticamente no final da planilha "Visualização de dados". Informe o valor dos novos rótulos na planilha "Visualização da variável". Agora, repita o processo para recodificar a variável "Câncer de pulmão".

Com as variáveis recodificadas, vamos realizar a análise de RR e OR. No menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Na janela de "Tabulações cruzadas", localize na tela da esquerda a nova variável "Tabagismo" (exposição) e a insira na "Linha(s)". Em seguida, na mesma tela, selecione a nova variável "Câncer de pulmão" (desfecho) e a mova para "Coluna(s)". Em "Exato", marque a opção de "Monte Carlo". Clique "Continuar". Para realizar o cálculo do RR e do OR no SPSS, clique em "Estatísticas" e marque as opções "Qui Quadrado" e "Risco". Clique "Continuar". Em "Células", marque as opções "Observado" em "Contagens", "Linha" e "Total" em "Porcentagens" e, para "Residuais", marque a opção "Padronizado ajustado". Clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, temos a tabela de contingência 2x2 (*Crosstabulation*), que fornece os dados necessários para o cálculo do RR e OR. O teste de correção de continuidade de Yates indica associação significativa entre tabagismo e câncer de pulmão ($p < 0,001$), evidenciada pelos resíduos. Independentemente do tipo de estudo, o SPSS produz a análise de ambos RR e OR, e compete ao pesquisador selecionar o tipo de resultado que está adequado ao seu estudo.

O SPSS, além do cálculo do RR e OR, também produz seus respectivos intervalos de confiança de 95% (IC95%). Para um p -valor $\leq 0,05$, o IC95% representa o intervalo onde deve estar o parâmetro, ou seja, o RR e OR verdadeiros. Se o valor 1, que se refere à nulidade

da associação, não estiver contido no IC95%, temos uma confiança que na população de onde nossa amostra foi extraída, o achado da amostra é significativo. Desta forma, ser tabagista aumenta o risco de câncer no pulmão em 4,71 vezes (IC95%: 2,19-10,15; $p < 0,001$) e aumenta a chance de câncer no pulmão em 6,54 vezes (IC95%: 2,73-15,68; $p < 0,001$) em comparação a não ser tabagista.

Avaliação de testes diagnósticos

O teste diagnóstico tem como objetivo determinar se um indivíduo tem ou não a doença. Para avaliar a acurácia deste teste, o pesquisador pode fazer uso dos parâmetros de sensibilidade e especificidade. Uma forma eficiente de demonstrar a relação antagônica entre a sensibilidade e a especificidade são as Curvas de Características de Operação do Receptor (Curvas ROC). A Curva ROC é uma ferramenta poderosa para medir e especificar problemas no desempenho dos testes diagnósticos, pois permite avaliar a variação da sensibilidade e especificidade para diferentes valores de corte.

Para isso, é necessário um teste de referência, também chamado de padrão ouro, que é capaz de diferenciar um indivíduo doente de um não doente, pois acredita-se que este tenha a medida "verdadeira". É utilizado como comparação na avaliação da qualidade (performance) de testes índices, que são medidos numa escala numérica (variável quantitativa) e, assim, a curva ROC permite estabelecer um ponto de corte que define os resultados deste teste como sendo positivo e negativo (variável dicotômica).

A reprodutibilidade, repetibilidade ou precisão de um teste diagnóstico consiste na consistência dos resultados quando o exame se repete pelo mesmo examinador ou por outros examinadores. A comparação dos resultados pode ser apresentada através do índice de concordância Kappa, visto no Capítulo 6.

A acurácia ou validade do teste refere-se à probabilidade de um teste determinar o verdadeiro valor do que está sendo medido (ser positivo nos doentes e negativo nos não doentes). Como mencionado, para avaliar a acurácia de um teste diagnóstico o pesquisador pode fazer uso dos índices de sensibilidade e especificidade.

A sensibilidade é a probabilidade de um teste detectar os sujeitos verdadeiramente positivos (doentes). Testes sensíveis são utilizados para reduzir possibilidades diagnósticas, como por exemplo, em triagens sorológicas em bancos de sangue para prevenção de transmissão de infecções nas quais a não detecção de casos acarretará risco para a população. Nestes casos, o ponto de corte deverá ser estabelecido tendo como objetivo alcançar maior sensibilidade do teste para que não ocorram resultados falsos negativos.

A especificidade é a probabilidade de um teste detectar os sujeitos verdadeiramente negativos (não doentes). Testes específicos são utilizados para confirmar a presença de doença, como por exemplo, em indicações cirúrgicas, onde deve-se evitar resultados falsos positivos. Nestes casos, o ponto de corte deve ser definido de tal forma que aumente a especificidade do teste.

Para se realizar as análises de sensibilidade e especificidade é necessário que utilizemos dois métodos ou instrumentos diferentes, mas que tenham o mesmo objetivo, sendo um o padrão ouro. O teste ideal, com 100% de sensibilidade e especificidade, raramente existe na prática. Isso porque, na maioria das vezes, a tentativa de melhorar a sensibilidade diminui a especificidade e vice-versa.

O valor preditivo positivo (VPP) mede a proporção de doentes entre os positivos pelo teste, ou seja, a probabilidade de doença quando o teste é positivo. Quanto mais específico, melhor o VPP (redução de falsos positivos). O valor preditivo negativo (VPN) mede a proporção de saudáveis entre os negativos pelo teste, ou seja, a probabilidade de ausência de doença quando o teste é negativo. Quanto mais sensível, melhor o VPN (redução de falsos negativos).

A razão de verossimilhança positiva (RVP) é a probabilidade de um resultado positivo em alguém com a doença dividida pela probabilidade do mesmo resultado em alguém sem a doença, ou seja, sensibilidade/1-especificidade. A razão de verossimilhança negativa (RVN) é a probabilidade de um resultado negativo em alguém com a doença dividida pela probabilidade do mesmo resultado em alguém sem a doença, ou seja, 1-sensibilidade/especificidade. Desta forma, os resultados dos testes padrão e índice podem ser organizados numa tabela de contingência 2x2 (Tabela 3).

Tabela 3. Tabela de contingência 2x2 para avaliar a relação de teste-doença.

		Padrão-ouro		Total
		Doente	Não doente	
Teste índice	Positivo	A (Verdadeiro positivo)	B (Falso positivo)	A+B
	Negativo	C (Falso negativo)	D (Verdadeiro negativo)	C+D
Total		A+C	B+D	A+B+C+D

Onde:

$$\text{ACURÁCIA} = A+D/A+B+C+D$$

$$\text{SENSIBILIDADE} = A/A+C$$

$$\text{ESPECIFICIDADE} = D/B+D$$

$$\text{VPP} = A/A+B$$

$$\text{VPN} = D/C+D$$

$$\text{RVP} = [A/(A+C)]/[B/(B+D)]$$

$$\text{RVN} = [C/(A+C)]/[D/(B+D)]$$

Curva ROC

Quando o resultado do teste é expresso como uma variável quantitativa, deve-se procurar um ponto de corte ("*cutoff point*"), cujo valor determinará a interpretação do teste como sendo positivo ou negativo. Para isso, é utilizada a curva ROC, que avalia o desempenho dos testes e a variação da sensibilidade e especificidade para diferentes valores de corte num gráfico que apresenta a proporção dos verdadeiros positivos (sensibilidade), no eixo Y, versus a proporção dos falsos positivos (1-especificidade), no eixo X. Desta forma, a Curva ROC é um gráfico de sensibilidade.

Neste tipo de representação gráfica, a área abaixo da curva (AUC) define a probabilidade do teste em detectar os verdadeiros positivos e os verdadeiros negativos, ou seja, expressa a acurácia do teste. A curva ROC é uma forma de representar a relação antagônica entre a sensibilidade e a especificidade. Para construir uma curva ROC traça-se um diagrama que represente a sensibilidade em

função da proporção de falsos positivos (1-especificidade) para um conjunto de valores de corte e , a linha diagonal, corresponde a um teste que é positivo ou negativo, aleatoriamente (Figura 3).

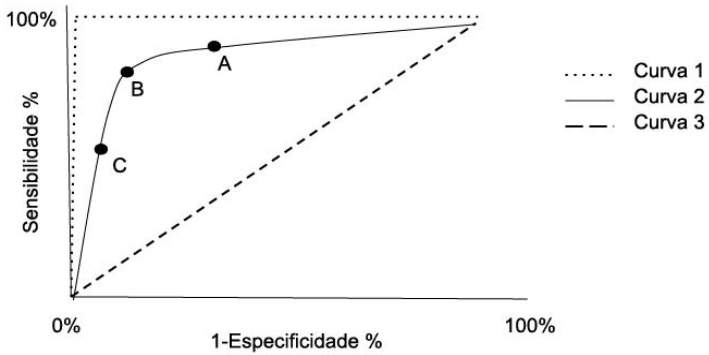


Figura 3. Curva ROC. 1, AUC 100%; 2, AUC 85%; 3, AUC 50%. Pontos de corte A, B e C com sensibilidade e especificidade respectivas.

Na curva 1, é representado o ponto de maior acurácia do teste (100%). No entanto, como mencionado, as características de sensibilidade e especificidade geralmente apresentam um comportamento antagônico. A curva 2 representa uma acurácia de 85%, que significa que em 85% das vezes em que for utilizado, o teste discriminará os verdadeiros positivos e negativos, mas também significa que em 15% das vezes fornecerá um falso resultado. A curva 3, expressa a acurácia de 50% e significa que o teste é tão bom quanto qualquer escolha ao acaso. Em geral, uma acurácia maior ou igual a 70% é considerada satisfatória.

O valor escolhido como ponto de corte influenciará as características do teste. Na curva 2, quanto menor o ponto de corte (ponto A), maior é a sensibilidade, mas menor é a especificidade. Neste caso, se pretendemos um teste muito sensível e menos específico, escolhe-se este ponto para obter uma menor proporção de falsos negativos e uma maior proporção de falsos positivos. O ponto de corte localizado no ponto B permite pôr em evidência os valores para os quais existe maior otimização da sensibilidade em função da especificidade. Já quanto maior o ponto de corte (ponto C), maior especificidade, porém, menor sensibilidade.

Neste caso, se pretendemos um teste muito específico e menos sensível, escolhe-se o ponto C, obtendo-se uma menor proporção de falsos positivos e uma maior proporção de falsos negativos.

O valor do ponto de corte é definido como um valor que pode ser selecionado, arbitrariamente, pelo pesquisador entre os valores possíveis para a variável de decisão. Acima do ponto de corte, o paciente é classificado como positivo (teste positivo, presença de doença) e, abaixo, é classificado como negativo (teste negativo, ausência de doença). Para cada ponto de corte são calculados valores de sensibilidade e especificidade, dispostos no gráfico.

Para exemplificar, no "Banco de dados 6.sav" (disponível em [https:// bit.ly/bancosdedados](https://bit.ly/bancosdedados)), vamos construir a curva ROC do NDS, utilizando como padrão ouro o MNSI, ambos instrumentos de rastreamento para neuropatia periférica. O SPSS não realiza os cálculos de sensibilidade, especificidade, VPP e VPN, porém fornece as informações na tabela de contingência 2x2 que permitem realizar os cálculos.

Para calcular esses valores, a priori, foi definido um ponto de corte de 3,0 para classificar os sujeitos como sendo neuropatas ("0") e não neuropatas ("1") pelo teste índice (NDS) e, para o padrão ouro (MNSI), um ponto de corte de 2,5 é utilizado para classificar os sujeitos da mesma forma. Os valores foram atribuídos desta forma para obtermos a correta estrutura de organização da tabela de contingência 2x2.

No menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Insira na "Linha(s)" a variável que classifica os sujeitos pelo NDS e, em "Coluna(s)", a variável que classifica os sujeitos pelo MNSI. Em "Exato...", selecione "Monte Carlo". Clique em "Continuar". Em "Estatísticas...", selecione as opções "Qui-quadrado" e "Capa". Clique em "Continuar". Em "Células...", marque "Observado" em "Contagens", "Total" em "Porcentagens" e, em "Residuais", selecione "Padronizado ajustado". Clique em "Continuar".

No arquivo de saída, os resultados do teste Exato de Fisher indicam haver associação significativa entre os dois testes ($p < 0,001$). Porém, o teste de Kappa indica uma concordância regular entre os métodos (0,330), apesar de significativa ($p < 0,001$). Estes resultados nos indicam uma baixa sensibilidade do método índice (NDS) para rastreamento de neuropatia periférica com o ponto de corte 3,0.

Na tabela de contingência estão disponíveis as informações necessárias para os cálculos da sensibilidade, especificidade, VPP e VPN:

$$\text{ACURÁCIA} = 15+143/200 = 0,79 = 79,0\%$$

$$\text{SENSIBILIDADE} = 15/(15+40) = 0,272 = 27,2\%$$

$$\text{ESPECIFICIDADE} = 143/(143+2) = 0,986 = 98,6\%$$

$$\text{VPP} = 15/(15+2) = 0,882 = 88,2\%$$

$$\text{VPN} = 143/(143+40) = 0,781 = 78,1\%$$

Como esperado, o NDS apresenta uma baixa sensibilidade. Como se trata de um teste de triagem, seria interessante utilizarmos um ponto de corte que apresente uma maior sensibilidade. Lembre-se que, no caso do rastreamento da neuropatia periférica, os valores do NDS produzem o *screening* positivo a partir de pontos de corte. Desta forma, as modificações dos pontos de corte podem produzir desempenhos diferentes quanto aos valores de sensibilidade, especificidade, VPP e VPN.

Uma forma de se analisar o desempenho dos pontos de corte é através da Curva ROC. Para que seja possível realizar está análise no SPSS, é necessário que uma variável associada ao método de referência (padrão ouro) para o diagnóstico seja qualitativa em escala dicotômica (positivo ou negativo) e, a variável que representa o teste índice que terá os pontos de corte analisados, seja do tipo quantitativa.

No menu "Analisar", clique na opção "Curva ROC". Na janela "Curva ROC", localize na tela da esquerda as variáveis e insira, com o auxílio do botão da seta, a variável quantitativa ("NDS") em "Variável de teste", e mova a variável qualitativa relacionada ao critério de referência ("NEUROMNSI") em "Variável de estado". Na

opção "Valor da variável de estado", insira o valor "0" que se refere ao valor positivo para o teste de rastreamento ("0" = "Neuropata"). Em "Exibir", selecione as opções "Curva ROC" para gerar o gráfico; a opção "Com linha de referência diagonal" para gerar a linha que representa a hipótese de nulidade; "Erro padrão e intervalo de confiança" para gerar o intervalo de confiança associado ao erro padrão e; "Coordenar pontos da curva ROC" para gerar a tabela com valores de sensibilidade e especificidade nos diferentes pontos de corte. Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, no primeiro quadro (*Case Processing Summary*), é indicado o número de casos positivos e negativos conforme a variável de referência. No quadro seguinte (*Area Under the Curve*), temos o valor de AUC (Area), o erro padrão, o p-valor e o IC95%. Lembre-se que a AUC é uma medida de desempenho que serve para determinar a acurácia do teste. Um teste totalmente incapaz de discriminar indivíduos doentes e não doentes, teria uma área sob a curva de 0,500 (50%), correspondendo a H_0 . Acima de 0,700 (70%) é considerado desempenho satisfatório. No exemplo, o valor de AUC de 0,831 indica que os valores do NDS têm uma probabilidade de 83% de detectar os resultados positivos (IC95%: 0,769-0,893; $p < 0,001$).

No terceiro quadro (*Coordinates of the Curve*), é possível verificar a relação dos diferentes pontos de corte de sensibilidade e 1-especificidade. A sensibilidade é a proporção de neuropatas com os resultados do teste maior que o ponto de corte e 1-especificidade é a proporção de não neuropatas com os resultados do teste maior que o de ponto de corte. O ponto de corte de equilíbrio será escolhido de acordo com o objetivo do teste. O gráfico gerado permite complementar a análise para a tomada de decisão (Figura 4).

O menor valor de corte é o valor mínimo de teste observado menos 1 e, o maior valor de corte, é o valor máximo de teste observado mais 1. Todos os outros valores de corte são as médias de dois valores de teste observados ordenados consecutivos. No ponto de corte, "0" é equivalente a assumir que todos são neuropatas. No ponto de corte "6" é equivalente a assumir que todos são não neuropatas. Ambos os extremos são insatisfatórios. É preciso escolher um ponto de corte que devidamente equilibre as necessidades de sensibilidade e especificidade.

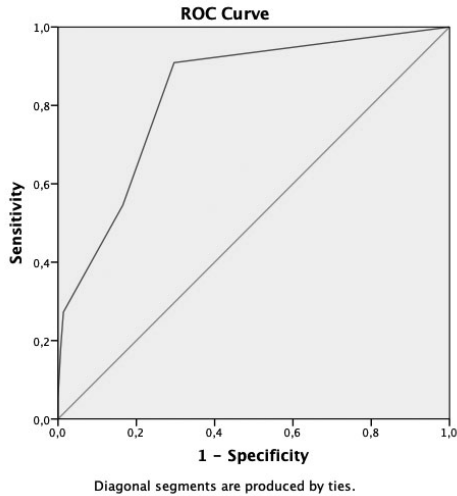


Figura 4. Curva ROC para avaliação da sensibilidade e 1-especificidade do teste índice (NDS) em comparação ao padrão ouro (MNSI).

Se considerarmos o ponto de corte de 2,5, temos o valor de sensibilidade de 0,273 (27,3%) e 1-especificidade de 0,014 (1,4%). Assim, aproximadamente 27,3% dos sujeitos com neuropatia periférica seriam corretamente identificadas e 1,4% dos sujeitos sem neuropatia periférica poderiam ser incorretamente identificadas como neuropatas (98,6% de especificidade). Por outro lado, se considerarmos o ponto de corte de 1,5, teríamos valores de sensibilidade de 0,545 e 1-especificidade de 0,166, ou seja, 54,5% dos neuropatas estariam classificados corretamente e 16,6% dos não neuropatas estariam classificados incorretamente como neuropatas (83,4% de especificidade). Note que os valores deste quadro representam as melhores orientações para os quais devemos considerar os pontos de cortes, mas não inclui as estimativas de erro. Portanto, não há garantia da exatidão da sensibilidade ou especificidade para um dado ponto de corte no quadro.

A escolha do ponto de corte será decidida pela necessidade de aumentar a sensibilidade ou a especificidade: uma maior sensibilidade para detectar os doentes aumentará as chances de falsos positivos, enquanto que uma maior especificidade para

determinar quem não esteja doente, que pode aumentar as chances de falsos negativos. Logo, dependerá do pesquisador a seleção do ponto de corte de equilíbrio entre sensibilidade e especificidade.

Referências

Barton, B, Peat, J. Medical Statistics: A guide to SPSS. Data analysis and Critical Appraisal. 2. ed. John Wiley & Sons, Ltda. 2014. 411 p.

Callegari-Jacques SM. Bioestatística: princípios e aplicações. Porto Alegre: ArtMed, 2011. 255p.

Campbell MJ, Machin D & Walters SJ. Medical Statistics, a Textbook for the Health Sciences. 4. ed. 2007. 344 p.

Chagas, E.F.B. Módulo 3: Estatística Analítica II - Análise de dados categóricos e teste diagnóstico no SPSS. 2016. 44 p.

Field A. Descobrendo a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Schoenbach VJ & Rosamond W. Understanding the Fundamentals of Epidemiology: An Evolving Text. Chapel Hill, North Carolina. 2000. 584 p.

Exercícios sugeridos

1. Utilizando o "Banco de dados 6.sav":

- a) Há correlação entre massa corporal e glicemia de jejum? Realize o teste estatístico, interprete os resultados e construa o gráfico.
- b) Há correlação entre massa corporal e HDL-Colesterol? Realize o teste estatístico, interprete os resultados e construa o gráfico.

2. Utilizando o "Banco de dados 6.sav":

- a) Há correlação entre IMC e glicemia de jejum? Realize o teste estatístico, interprete os resultados e construa o gráfico.
- b) Há correlação entre IMC e HDL-Colesterol? Realize o teste estatístico, interprete os resultados e construa o gráfico.

3. Utilizando o "Banco de dados 6.sav":

- a) Qual a chance de ser obeso e ter hipertensão? Realize o teste estatístico e interprete os resultados.
- b) Qual a chance de ser obeso e ter diabetes? Realize o teste estatístico e interprete os resultados.

10 Regressão linear simples e múltipla

*Michele Strelow Moreira
Marina Petter Rodrigues
Charles Francisco Ferreira
Otto Henrique Nienov*

Como normalmente existem muitas possíveis variáveis explicativas (ou independentes) em um estudo, fica difícil analisá-las de uma só vez. Por esta razão, é comum, primeiro, buscar identificar fatores associados a um determinado desfecho (resposta) realizando análises univariadas, seguidas por uma análise multivariada. Mas, nem sempre é possível ou necessário realizá-la.

A análise univariada avalia isoladamente a relação entre cada possível variável independente e a variável desfecho, sem levar em conta as demais. A análise bivariada inclui métodos de análise de duas variáveis, podendo ser ou não estabelecida uma relação de causa e efeito entre elas. São exemplos de métodos de análise bivariada, o teste de Qui-quadrado e os coeficientes de correlação (e.g. Correlação de Pearson, Correlação de Spearman).

A análise multivariada refere-se a modelos de regressão múltipla que buscam explicar uma variável desfecho com base em um conjunto de variáveis independentes. Em modelos estatísticos, denomina-se a variável dependente (desfecho) como aquela em que se tem interesse em analisar.

Análise de regressão

Em um estudo, quais os fatores que mais importam? O que podemos ignorar? Como esses fatores interagem uns com os outros? E, o mais importante, quão certos estamos sobre esses fatores? A análise de regressão ajuda a responder essas questões, visto que o coeficiente de cada variável preditora descreve a contribuição relativa de cada variável ao desfecho.

A análise de regressão é uma forma de prever algum resultado, relacionada à variável de resposta ou dependente (principal fator que você está tentando entender ou prever), a partir de uma ou mais variáveis preditoras, explicatórias ou independentes (fatores que você suspeita terem algum impacto no desfecho). Desta forma, conseguimos avaliar o impacto de cada variável independente sobre o desfecho.

Você pode estar se perguntado se, por exemplo, uma correlação não poderia fazer isso. A correlação pode ser uma ferramenta bastante útil, mas ela não nos informa sobre o poder preditivo das variáveis. Na análise de regressão, ajustamos um modelo preditivo aos nossos dados e, então, usamos esse modelo para prever valores da variável dependente a partir de uma ou mais variáveis independentes.

Quando queremos avaliar a influência de um conjunto de fatores sobre doenças ou outras características de interesse, a análise de regressão é o método estatístico utilizado, estabelecendo uma equação que simula os relacionamentos entre a variável dependente e os fatores que se deseja investigar (variáveis independentes).

Ainda podemos classificar a regressão conforme a quantidade de fatores preditores que se deseja investigar (variáveis independentes). Quando existe apenas uma variável independente, ou seja, uma única variável preditora, a definimos como sendo simples. Por outro lado, em regressões múltiplas, mais de uma variável independente é utilizada para prever o desfecho.

Construindo modelos de regressão

O modelo de regressão começa quando um pesquisador deseja descrever matematicamente a relação entre algumas variáveis preditoras (independentes) e a variável desfecho (dependente). Em um projeto de pesquisa, normalmente muitas variáveis são julgadas importantes para responder uma pergunta. Contudo, no modelo de regressão, incluímos apenas algumas delas. O pesquisador deve tentar eliminar as variáveis que não estão relacionadas e incluir apenas aquelas que ele acredite terem algum relacionamento verdadeiro, sendo possível considerar muitos modelos diferentes ao longo do percurso.

Mas, como determinar as variáveis que farão parte do modelo de regressão? Geralmente se busca o modelo mais parcimonioso que explique os dados. No entanto, para a obtenção deste “modelo mais parcimonioso” final, algumas etapas devem ser percorridas. Deixamos aqui alguns tópicos que podem auxiliar você na elaboração do modelo final de regressão:

1) Tamanho de amostra para regressão: existe uma divergência sobre o tamanho da amostra ideal para o modelo de regressão, a depender do número de variáveis preditoras. Alguns autores indicam o tamanho amostral de 10 indivíduos para cada preditor (variável independente) no modelo, outros apontam 15 indivíduos por variável independente. Assim, se você tivesse cinco variáveis independentes no modelo, seria necessário 50 ou 75 casos, respectivamente. O tamanho da amostra necessário irá depender do tamanho de efeito que estamos tentando detectar.

2) Número de variáveis incorporadas no modelo: um modelo pouco especificado tende a produzir estimativas não reais, assim como um modelo muito especificado pode produzir estimativas menos precisas. Se um excesso de variáveis independentes forem incluídas, maiores se tornam os erros e mais dependente será o modelo dos dados observados. Portanto, um modelo com as variáveis independentes mais pertinentes terá menor risco de viés e apresentará as estimativas mais precisas. Minimizando o número de variáveis independentes, o modelo matemático final terá maior probabilidade de ser generalizável.

3) Avaliação da multicolinearidade: variáveis independentes altamente correlacionadas entre si são denominadas multicolineares. Muitos tipos de modelos de regressão pressupõem que a multicolinearidade não deve estar presente no conjunto de dados, por causar problemas na classificação de variáveis com base em sua importância ou dificultar o trabalho na seleção da variável independente mais importante. A multicolinearidade indica que duas variáveis independentes (fatores preditores) analisados possuem explicação para o mesmo segmento do desfecho analisado.

4) Análise univariada: na prática, existe um grande número de variáveis explicativas (independentes). Logo, a análise univariada pode servir como critério de seleção das variáveis que entrarão em um modelo de regressão. Como se trata de uma etapa inicial e não-definitiva da análise de dados, podemos ser menos rigorosos e adotar níveis de significância maiores que o usual (por exemplo, $p < 0,150$, $p < 0,200$ ou $p < 0,250$) para não correr o risco de desprezar variáveis importantes. Denominamos estas de variáveis tendenciosas, ou seja, que apresentam um p-valor limítrofe para associação. Considere, também, aspectos como a análise dos resíduos do teste de Qui-quadrado, para verificar a indicação dessa associação; a direção (positiva ou inversa) e a intensidade (fraca ou forte) dos coeficientes de correlação; o valor do índice F da Análise de Variância (ANOVA), onde um valor elevado significa que há alguma diferença entre os grupos capaz de ser expressa adequadamente por meio de um modelo de regressão.

5) Aspectos encontrados na literatura: pesquise o que outros autores fizeram e incorpore essas descobertas na construção do seu modelo. Antes de iniciar a análise de regressão, desenvolva uma ideia de quais são as variáveis importantes, juntamente com suas relações, sinais de coeficiente e magnitudes de efeito. Com base nos resultados de outras pessoas, fica mais fácil coletar os dados corretos e especificar o melhor modelo de regressão sem a necessidade de usar mineração de dados. Em seguida, verifique se o modelo obtido se alinha com a teoria e faça ajustes. Por exemplo, com base na teoria, você pode incluir uma variável preditora no modelo mesmo que seu valor-p não seja significativo.

6) Variáveis de controle: insira no modelo variáveis que você julga importantes, inclusive para controlar variáveis do próprio modelo, mesmo estas não sendo significativas.

7) Teste modelos diferentes: após selecionadas as variáveis candidatas para o modelo, a última etapa é modelar juntas todas essas variáveis e testar modelos diferentes. Dessa forma, as variáveis mais associadas ao desfecho de interesse

são selecionadas para o modelo final. Esta etapa consiste em avaliar simultaneamente, um modelo de cada vez, o efeito das variáveis selecionadas sobre a resposta. Modelos mais simples geralmente produzem previsões mais precisas. Existem diferentes formas de abordagem, conforme apresentado no quadro 1.

Quadro 1. Formas de abordagem do modelo de regressão.

Abordagem	Descrição
Método <i>ENTER</i> (forma manual)	É o método mais indicado para se testar hipóteses, uma vez que é o pesquisador que determina a relevância das variáveis no modelo. Pode-se iniciar com a variável mais relevante e avalia-se a inclusão, uma a uma, das seguintes, ou, se a amostra é suficientemente grande (pelo menos 10 casos com o desfecho menos frequente por variável), pode-se começar a modelagem com todas as variáveis selecionadas, retirando-se, uma a uma, a menos relevante.
Método <i>STEPWISE</i> (forma automática)	É indicado quando não existem pesquisas de base para auxiliar na escolha das variáveis do modelo. No entanto, pode ser influenciado por variações aleatórias dos dados. As variáveis são selecionadas automaticamente, sem o controle do pesquisador, tanto para inclusão quanto para exclusão, de uma maneira sequencial, baseada apenas em critérios estatísticos. Há duas modelagens:
	<i>FORWARD</i> : as variáveis são adicionadas uma a uma ao modelo, sendo a primeira variável candidata aquela mais significativa e, assim sucessivamente. <i>BACKWARD</i> : todas as variáveis selecionadas pelo pesquisador são incluídas, sendo retirada, uma a uma, a variável menos significativa. Esta abordagem permite a comparação de modelos que talvez não fossem gerados pelo método manual.

É comum que variáveis estatisticamente significativas na análise univariada percam a importância na multivariada. Isso porque, quando analisamos o fator isoladamente, não estamos levando em conta outras características que podem estar relacionadas a este fator (viés). Embora seja menos comum, é possível também que uma variável não-significativa na univariada passe a ser significativa na multivariada, principalmente quando houver interação entre os fatores. A existência de interação significa que o efeito de um fator sobre a resposta depende de outro fator.

Regressão linear

A regressão linear estima os coeficientes da equação linear, com o objetivo prever o comportamento de uma variável dependente (quantitativa) em função de uma (simples) ou mais (múltipla) variáveis independentes (quantitativas ou qualitativas binárias). Assumimos a relação entre as variáveis por meio de uma reta (Figura 1) e, assim, utilizamos o resultado da função dessa reta para estimar valores, quando conhecemos as variáveis que afetam o seu comportamento.

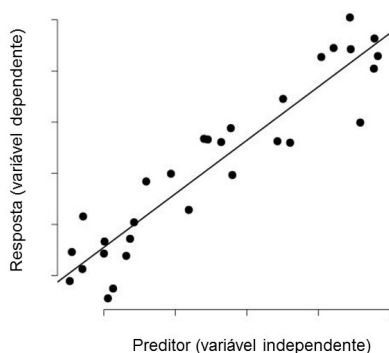


Figura 1. Reta da regressão linear.

No eixo Y da representação gráfica, temos a variável dependente (desfecho) que estamos tentando descobrir e, no eixo X da representação gráfica, as variáveis independentes (preditoras) que exercem influência sobre a variável dependente. A variável Y deve ter distribuição normal ou aproximadamente normal. Os parâmetros de distribuição β_0 e β_1 são denominados de coeficientes da regressão: β_0 é o intercepto em Y da equação de regressão (é o valor de Y quando X é igual a 0) e β_1 refere-se à inclinação da reta de regressão (indica a mudança na média de Y quando X é acrescido de uma unidade). O erro (ϵ) representa as influências não controladas, ou seja, influências que a variável dependente possui além da exercida pelas variáveis no modelo. Conforme mencionado, o modelo para essa relação entre as variáveis é linear, sendo representado pela equação:

$$Y = \beta_0 + \beta_1 * X_1 + \dots + \epsilon$$

Onde:

Y = variável dependente da equação de regressão linear

β_0 = intercepto em Y da equação de regressão linear

β_1 = inclinação da reta de regressão linear

ε = erro da equação de regressão linear.

Tanto na regressão linear simples quanto na múltipla, as suposições do modelo ajustado precisam ser validadas para que os resultados sejam confiáveis. Essa validação é feita através da análise dos resíduos, um conjunto de técnicas utilizadas para investigar a adequabilidade de um modelo de regressão com base nos resíduos. Esses resíduos representam o erro que está presente no modelo, correspondendo a diferença entre o valor previsto e o observado, ou seja, a distância do ponto até a reta que foi prevista no modelo. Se o modelo se ajusta bem aos dados da amostra, todos os resíduos devem ser pequenos (se o modelo aderir perfeitamente aos dados, todos os pontos estarão sobre a linha de regressão e todos os resíduos serão iguais a zero). Se o modelo não tiver uma boa aderência aos dados da amostra, os resíduos serão grandes. Além disso, se qualquer caso destacar-se por ter um grande resíduo, ele poderá ser um valor atípico.

Os pré-requisitos para a realização dessa análise de resíduos são:

1) Normalidade dos resíduos: o erro (diferença entre a variável dependente e a estimação feita pelo modelo) deve ter distribuição normal. A distribuição normal dos resíduos é essencial para que os resultados do ajuste do modelo de regressão sejam confiáveis. Como visto no capítulo 7, podemos verificar essa suposição por meio de uma inspeção visual (histograma de resíduos padronizados e gráficos de probabilidade normal comparando a distribuição de resíduos padronizados com uma distribuição normal, também chamado de P-P plot) e de testes estatísticos (Shapiro-Wilk e Kolmogorov-Smirnov) para os resíduos.

2) Homocedasticidade: a variância do erro experimental (ε) para observações distintas deve ser constante, ou seja, deve ser homocedástico. Isso pode ser visto em um gráfico de resíduos do tipo *scatterplot* (Figura 2). Neste gráfico, é apresentada a relação dos valores preditos no eixo X e

dos valores residuais no eixo Y, onde a distribuição destes resíduos não deve exibir nenhum padrão óbvio. Desta forma, se os pontos estão aleatoriamente distribuídos, sem nenhum comportamento ou tendência, temos indícios de que a variância dos resíduos é homoscedástica. Já a presença de um “funil”, é um indicativo da presença de heterocedasticidade. Além disso, os testes de Breusch-Pagan e de Goldfeld-Quandt também podem ser utilizados para testar se os resíduos são homoscedásticos. Baseado no teste multiplicador de Lagrange, o teste de Breusch-Pagan é bastante utilizado para testar a H_0 de que as variâncias dos erros são iguais versus a H_A de que as variâncias dos erros são uma função multiplicativa de uma ou mais variáveis, sendo que esta(s) variável(eis) pode(m) pertencer ou não ao modelo em questão. É indicado para grandes amostras e quando a suposição de normalidade nos erros é assumida. O teste de Goldfeld-Quandt também é utilizado para testar a homoscedasticidade dos resíduos, mas limita-se por exigir que a amostra seja relativamente grande. Se não houver homoscedasticidade, será necessário transformar os dados, como por exemplo, logaritmo.

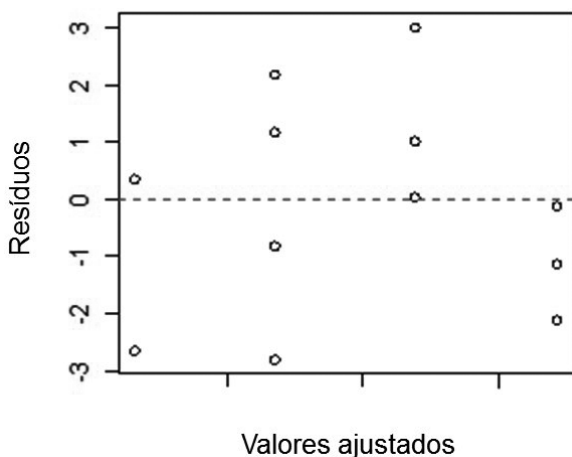


Figura 2. Gráfico de *scatterplot* dos resíduos versus valores ajustados.

3) Multicolinearidade: o valor do erro para uma observação deve ser independente dos valores das variáveis do modelo e do erro das outras observações, ou seja, não devemos ter uma relação entre as variáveis independentes. Esse diagnóstico é importante quando utilizarmos duas ou mais variáveis independentes no modelo de regressão múltipla. Desta forma, se não houver nenhum relacionamento entre elas, dizemos que são ortogonais. Na prática, é muito difícil que as variáveis de entrada sejam ortogonais e, felizmente, a falta de ortogonalidade não é grave. Mas, se as variáveis forem muito correlacionadas, as inferências baseadas no modelo de regressão podem ser errôneas ou pouco confiáveis. Um dos diagnósticos para multicolinearidade pode ser baseado na estatística K , uma análise de estrutura em função dos autovalores, onde um $K < 10$ indica que não há problemas de multicolinearidade; enquanto que $10 < K < 100$ indica uma multicolinearidade moderada e; para $K > 100$, temos uma multicolinearidade severa. Outro diagnóstico pode ser realizado pelo Fator de Inflação de Variância (VIF), que deve ser menor que 10. O VIF quantifica a extensão da correlação entre um preditor e os outros preditores em um modelo. Se maiores que 10, há indicação de multicolinearidade, ou seja, é difícil avaliar com precisão a contribuição dos preditores para um modelo. Ou conforme o índice de Tolerância ($1/VIF$), que deve ser maior que 0,2.

4) Autocorrelação: os resíduos devem ser independentes, isto é, não correlacionados. Para verificar esta suposição, utilizamos o teste de Durbin-Watson, que testa a suposição de independência dos erros. No entanto, diferentemente dos testes de hipóteses vistos anteriormente, devemos analisar os valores da estatística D , apresentados na tabela 1. Seguindo a linha do respectivo tamanho amostral (N) e a coluna do número de variáveis independentes consideradas (X), temos, para cada nível de significância (α): se D é maior que $D-U$, aceita-se H_0 (resíduos não correlacionados); se D menor que $D-L$, rejeita-se H_0 (resíduos correlacionados) ou; se D estiver entre $D-L$ e $D-U$, o teste é inconclusivo.

Tabela 1. Valores da estatística D do testes de Durbin-Watson.

Observações		Número de variáveis independentes (X)							
		1		2		3		4	
N	α	D-L	D-U	D-L	D-U	D-L	D-U	D-L	D-U
15	0,050	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97
	0,010	0,81	1,07	0,70	1,25	0,59	1,46	0,49	1,70
20	0,050	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83
	0,010	0,95	1,15	0,86	1,27	0,77	1,41	0,68	1,57
25	0,050	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77
	0,010	1,05	1,21	0,98	1,30	0,90	1,41	0,83	1,52
30	0,050	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74
	0,010	1,13	1,26	1,07	1,34	1,01	1,42	0,94	1,51
40	0,050	1,44	1,54	1,39	1,60	1,34	1,66	1,39	1,72
	0,010	1,25	1,34	1,20	1,40	1,15	1,46	1,10	1,52
50	0,050	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72
	0,010	1,32	1,40	1,28	1,45	1,24	1,49	1,20	1,54
60	0,050	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73
	0,010	1,38	1,45	1,35	1,48	1,32	1,52	1,28	1,56
80	0,050	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74
	0,010	1,47	1,52	1,44	1,54	1,42	1,57	1,39	1,60
100	0,050	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76
	0,010	1,52	1,56	1,50	1,58	1,48	1,60	1,46	1,63

5) Pontos influentes e valores atípicos (também denominados *outliers*): o valor atípico é um caso que difere substancialmente da maioria dos dados, que pode ser vista em gráficos de dispersão. Além de procurar valores atípicos olhando para os erros do modelo, também é possível buscar certos casos que influenciam os parâmetros do modelo. Um ponto influente é uma observação que pode influenciar em qualquer parte da análise de regressão, como a previsão de uma resposta ou o resultado de testes de hipóteses. Os valores extremos

têm potencial para serem influentes, podendo introduzir tendenciosidade no modelo, pois eles irão afetar os valores dos coeficientes de regressão estimados. Mas, temos que investigar para avaliar o quanto eles são influentes. Se um valor extremo for influente, ele interfere sobre a função de regressão ajustada (a inclusão ou não do ponto modifica substancialmente os valores ajustados). Por outro lado, uma dada observação pode ser considerada um valor atípico e não ser um ponto influente. Da mesma forma, podemos ter pontos que influenciam na análise de regressão, mas não são valores atípicos. Esse tipo de análise pode ajudar a determinar se o modelo de regressão é estável por toda a amostra ou se ele pode estar sendo influenciado somente por poucos casos.

Verificados os pré-requisitos da análise de resíduos, prosseguimos com a análise da regressão linear. Para verificar se a regressão é significativa, devemos inicialmente observar o resultado da ANOVA. Se o p -valor $> 0,050$, não temos evidência para dizer que o modelo de regressão linear é importante para explicar a variável desfecho, ou seja, as variáveis independentes não exercem influência na variável dependente. Por outro lado, se o p -valor $\leq 0,050$, podemos dizer que ao menos uma das variáveis do modelo é importante para explicar a variável desfecho, ou seja, pelo menos uma variável independente exerce influência na variável dependente.

Para cada parâmetro estimado é realizado um teste de significância (teste t de Student), onde a H_0 é de que o respectivo coeficiente é igual a zero e, a H_A é de que o coeficiente é diferente de zero, sendo esta última significativa quando p -valor $\leq 0,050$.

Constatada a significância do modelo, também verificamos o coeficiente de determinação (R^2). O R^2 é um indicador para a análise do ajuste do modelo adotado, indicando a proporção da variabilidade de Y que pode ser explicada pela variabilidade das variáveis X . O R^2 pode variar de 0 a 1: quanto mais próximo de zero, indica que o modelo não explica a variabilidade dos dados de desfecho ao redor de sua média; quanto mais próximo de um, indica que o modelo explica toda a variabilidade dos dados de desfecho ao redor de sua média.

Regressão linear simples

Para melhor compreender a análise de regressão, começaremos construindo uma regressão linear simples, onde há apenas uma variável independente. Logo, a equação resume-se a:

$$Y = \beta_0 + \beta_1 * X_1 + \dots + \varepsilon$$

No “Banco de dados 2.sav” (disponível em [https:// bit.ly/bancosdedados](https://bit.ly/bancosdedados)), vamos verificar se a estatura depende da idade, supondo que a estatura tem distribuição normal. Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis (Mínimo? Máximo? Média? Mediana? Intervalo interquartilico? Teste de normalidade?).

No menu “Analisar”, “Regressão”, clique em “Linear...”. Na janela “Regressão linear”, selecione a variável “Estatura” e a insira em “Dependente” e, a variável “Idade”, em “Independente(s)”. Em “Estatísticas”, selecione as opções “Estimativas” e “Intervalo de confiança” em “Coeficiente de regressão”; “Ajuste do modelo”, “Descritivos” e “Diagnóstico de colinearidade” (para avaliar a multicolinearidade) e; em “Residuais”, “Durbin-Watson” (para verificar a autocorrelação) e “Diagnóstico por caso” (para analisar os valores atípicos além de três desvios padrão). Clique em “Continuar”. Em “Diagramas”, defina os valores preditos padronizados (ZPRED) no “X” e os resíduos padronizados (ZRESID) no “Y”. ZPRED são formas padronizadas dos valores previstos pelo modelo e ZRESID são as diferenças padronizadas entre os dados observados e os valores que o modelo prevê. Marque as opções de “Histograma” e “Diagrama de probabilidade normal” em “Diagramas residuais padronizados”. Clique em “Continuar”.

É necessário salvar os resíduos e os valores previstos da análise de regressão. Para isso, clique em “Salvar” e selecione “Não padronizados” em “Valores previstos”, “Não padronizado” em “Residuais” e, no “Intervalos de previsão”, selecione “Média”. Os valores preditos não padronizados correspondem ao valor que o modelo prediz para a variável dependente. Os residuais não padronizados indicam a diferença entre um valor observado e o valor predito pelo modelo. Os intervalos de predição para média indicam limites inferior e superior (duas variáveis) para o intervalo de predição da resposta média predita. Clique em “Continuar”. Por fim, clique em “Ok” ou “Colar”.

A análise de regressão apresenta uma série de quadros no arquivo de saída. No primeiro (*Descriptive Statistics*), temos os valores de média (*Mean*) e desvio padrão (*Std. Deviation*) e tamanho da amostra (*N*). Em seguida (*Correlations*), o coeficiente de correlação de Pearson. É importante verificarmos se a relação entre a variável dependente e as variáveis independentes é de fato linear. Conforme o gráfico (Figura 3), é visível a relação linear, sendo negativa fraca ($r = -0,133$, $p = 0,002$).

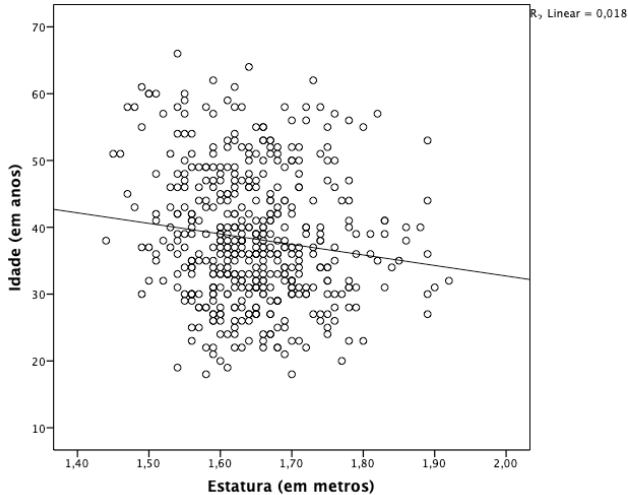


Figura 3. Correlação entre idade (em anos) e estatura (em metros).

Antes de validar os pré-requisitos da análise dos resíduos, vamos verificar a análise da regressão. O primeiro quadro que devemos analisar é o da ANOVA, que mostra o índice F. Esta análise é bastante familiar, pois foi vista no capítulo 7. Neste exemplo, temos como a H_0 de que a idade não explica a variação da estatura ($R^2 = 0$), enquanto que na H_A a idade explica a variação da estatura. O valor do índice F (8,459) é significativo, portanto, podemos assumir que o modelo explica uma quantidade significativa da variação da estatura em função da variabilidade da idade. Com p -valor igual a 0,004, o teste global mostra que há regressão entre as variáveis.

Cada variável é multiplicada por uma constante (β), que corresponde ao valor B para a constante (intercepto). O teste t verifica se essa constante é (H_0) ou não (H_A) igual a zero. No segundo

quadro a ser analisado (*Coefficients*), a variável "Idade" possui $\beta = -0,001$, sendo este significativamente diferente de zero pelo valor do índice t ($p = 0,004$). Assim, para cada unidade a mais na idade, a estatura média é menor em 0,001 metros.

Por fim, verificamos a qualidade do modelo (*Model Summary*), através do coeficiente de correlação (R), que é idêntico ao coeficiente de correlação de Pearson; do coeficiente de determinação (R^2) e; do coeficiente de determinação ajustado (R^2_{adj}), que penaliza por cada variável colocada no modelo, baixando o coeficiente de determinação quando o modelo é construído com mais de uma variável. Estes servem como um parâmetro para comparar modelos de regressão.

Neste exemplo, temos que a correlação entre as variáveis é de 0,133 e a proporção da variabilidade de estatura que pode ser explicada pela variabilidade da idade é de 1,8% (R^2). A regressão linear simples mostrou que a idade prevê a estatura, [$F(1,467) = 8,459$, $p = 0,004$, $R^2 = 0,018$]. A estatura, em metros, corresponde a $1,690 - 0,001 \times (\text{idade})$, sendo a idade medida em anos.

Para validar a regressão, devemos realizar a análise de resíduos. Lembre-se que a variável dependente é a estatura, pois é ela que depende da idade, e não o contrário. Os gráficos podem ajudar na validação das suposições de normalidade, linearidade e igualdade das variâncias. Além disso, são úteis para detectar valores discrepantes, observações atípicas e casos influentes.

O primeiro passo é a verificação do gráfico de *scatterplot* dos resíduos versus valores preditos (Figura 4). Conforme o gráfico, é visível que os pontos estão aleatoriamente distribuídos, sem nenhum comportamento ou tendência (sem comportamento cônico) e, assim, temos indícios de que a variância dos resíduos é homoscedástica.

A colinearidade (ou multicolinearidade) é a situação indesejável em que uma variável independente é uma função linear de outras variáveis independentes. Para avaliar a multicolineariedade, o quadro "*Coefficients*" apresenta o índice de Tolerância e o VIF, ambos quantificados em 1, ou seja, não há problemas de multicolinearidade. Para verificar a autocorrelação,

o quadro “*Model Summary*” traz a estatística D de Durbin-Watson, que foi de 2,038. Assim, para $N > 100$ e X igual a 1 (modelo com uma única variável independente), temos D maior que D-U, ou seja, podemos afirmar que não há autocorrelação.

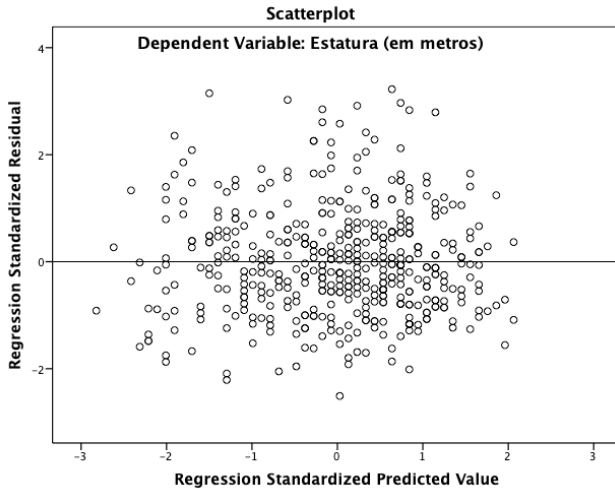


Figura 4. Gráfico de *scatterplot* dos resíduos versus valores preditos.

No quadro “*Casewise Diagnostics*”, temos indicado três casos com valores atípicos (acima de três desvios-padrões). Neste caso, devemos realizar a análise de regressão com e sem os valores atípicos, para verificar a influência destes no modelo. Por fim, verificamos a normalidade dos resíduos. Pode-se utilizar um histograma ou um gráfico P-P *plot* normal (Figura 5). No gráfico P-P *plot*, é possível observar que os pontos geralmente seguem a reta normal (diagonal) sem desvios fortes. Isso indica que os resíduos aparentemente estão normalmente distribuídos, mas o ideal é confirmar com o teste de normalidade de Shapiro-Wilk.

Para solicitar o teste de normalidade dos resíduos, no menu “Analisar”, “Estatística descritivas”, clique em “Explorar...” e utilize a variável de resíduos criada na análise de regressão. O quadro “*Test of Normality*” apresenta o teste de normalidade com p-valor inferior a 0,001. Desta forma, rejeitamos a hipótese de normalidade dos resíduos, ou seja, o teste de regressão não é válido pois não cumpriu o pré-requisito de normalidade dos resíduos.

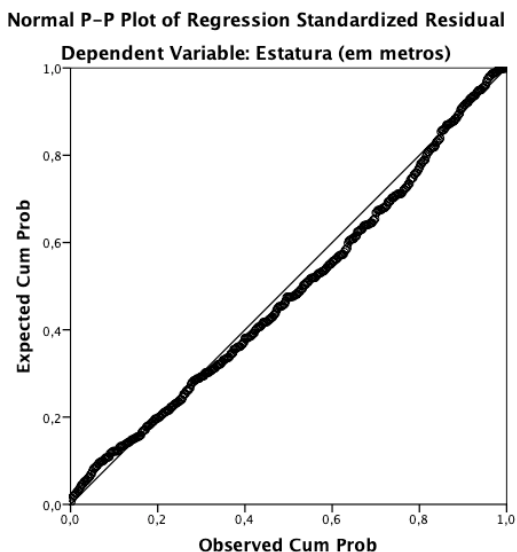


Figura 5. Gráfico P-P *plot* para verificar a normalidade dos resíduos.

Regressão linear múltipla

Existem muitos testes estatísticos que avaliam o grau de relação entre duas variáveis, como os coeficientes de correlação, por exemplo, mas estes não são capazes de explorar relações multivariadas, como mencionado. Comumente temos um desfecho que não depende apenas de uma variável, mas de várias. Nesses casos podemos utilizar a regressão linear múltipla, ou seja, com mais de uma variável independente, adotando a seguinte equação: $Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_k * X_k + \varepsilon$.

Assim, Y é a variável de resultado (desfecho), β_1 é o coeficiente do primeiro predictor (variável independente X_1), β_2 é o coeficiente do segundo predictor (variável independente X_2), β_k é o coeficiente do k -ésimo predictor (variável independente X_k) e ε é a diferença entre o valor previsto e o observado de Y . Como na regressão linear simples, na múltipla também devemos validar as suposições do modelo ajustado para que os resultados sejam confiáveis, através da análise de resíduos.

Para exemplificar, no “Banco de dados 2.sav”, vamos verificar se a estatura depende da idade e do sexo, supondo que a estatura tem distribuição normal. No menu “Analisar”, “Regressão”, clique

em "Linear...". Na janela "Regressão linear", selecione a variável "Estatura" e a insira em "Dependente" e, as variáveis "Idade" e "Sexo", como "Independente(s)". Em "Estatísticas", selecione as opções "Estimativas" e "Intervalo de confiança" em "Coeficiente de regressão"; "Ajuste do modelo", "Descritivos" e "Diagnóstico de colinearidade" (para avaliar a multicolinearidade) e; em "Residuais", "Durbin-Watson" (para verificar a autocorrelação) e "Diagnóstico por caso" (para analisar os valores atípicos além de três desvios padrão). Clique em "Continuar".

Em "Diagramas", defina o valor previsto (ZPRED) no "X" e o resíduo ajustado (ZRESID) no "Y". Marque as opções de "Histograma" e "Diagrama de probabilidade normal" em "Diagramas residuais padronizados". Clique em "Continuar". É necessário salvar os resíduos e os valores previstos da análise de regressão. Para isso, clique em "Salvar" e selecione "Não padronizados" em "Valores previstos", "Não padronizado" em "Residuais" e, no "Intervalos de previsão", selecione "Média". Clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No exemplo, note que a variável "Sexo" não é quantitativa, mas sim, qualitativa. Nestes casos, é comum a codificação indicadora, na qual as categorias são representadas por "1" ou "2", conforme rótulos calibrados para esta variável, por exemplo. Para a análise de regressão, deve-se tratá-la como as demais variáveis independentes.

Da mesma forma, no arquivo de saída, a análise de regressão apresenta vários quadros. No primeiro (*Descriptive Statistics*), temos os valores de média (*Mean*) e desvio padrão (*Std. Deviation*) e tamanho da amostra (*N*). Em seguida (*Correlations*), o coeficiente de correlação de Pearson. É importante verificarmos se a relação entre a variável dependente e as variáveis independentes é de fato linear. Conforme o gráfico (Figura 6), é visível a relação linear negativa fraca com a idade ($r = -0,133$, $p = 0,002$) e negativa forte com o sexo ($r = -0,628$, $p < 0,001$).

Seguindo a interpretação vista no exemplo anterior, primeiramente devemos analisar o teste global (ANOVA) e o teste t para os coeficientes. Com p-valor menor que 0,001, para ambos, o teste global mostra que há regressão entre as variáveis e, pelo teste t, os coeficientes são significativamente diferentes de zero.

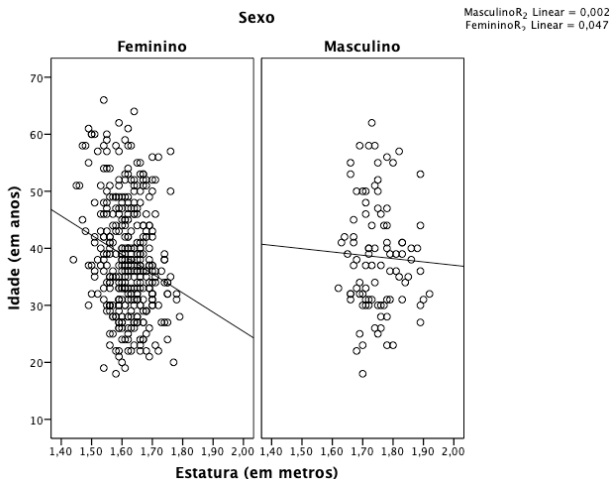


Figura 6. Correlação entre sexo, idade (em anos) e estatura (em metros).

Conforme a equação, 1,927 é a estatura constante, independentemente da idade e do sexo. O coeficiente angular da idade é negativo, indicando dependência inversa, sendo de $-0,001$ a variação na estatura em relação à unidade de variação na idade. Já o coeficiente para a variável sexo representa diferenças entre médias para cada grupo, isto é, $-0,131$ representa o efeito linear do grupo não omitido, que recebeu o código um (sexo masculino).

Em seguida, analisamos o coeficiente de determinação ajustado. Temos que a correlação entre as variáveis é de $0,643$ e a proporção da variabilidade de estatura que pode ser explicada pela idade e pelo sexo é de $41,1\%$ (R^2_{adj}). Com isso, concluímos que foi utilizada a regressão linear múltipla para verificar se idade e sexo são capazes de prever a estatura. A análise resultou em um modelo estatisticamente significativo [$F(2,466) = 164,519$; $p < 0,001$; $R^2 = 0,411$]. A idade ($\beta = -0,140$; $t = -3,947$; $p < 0,001$) e o sexo ($\beta = -0,629$; $t = -17,745$; $p < 0,001$) são preditores da estatura.

No entanto, lembre-se que é necessário validarmos os resultados pela análise dos resíduos. De acordo com o VIF e o índice de Tolerância apresentados no quadro "Coefficients", não há presença de multicolinearidade. A estatística D de Durbin-Watson está apresentada em "Model Summary" e foi igual a $2,051$. Isto é,

para $N > 100$ e X igual a 2, temos D maior que $D-U$, logo, podemos afirmar que não há autocorrelação. Analisando o *scatterplot* (Figura 7), sugere-se a homogeneidade de variâncias, pois não há um parâmetro de distribuição das variáveis. Verificada a normalidade dos resíduos pelo teste de Shapiro-Wilk ($p = 0,197$), podemos afirmar que o modelo de regressão é válido para esta análise.

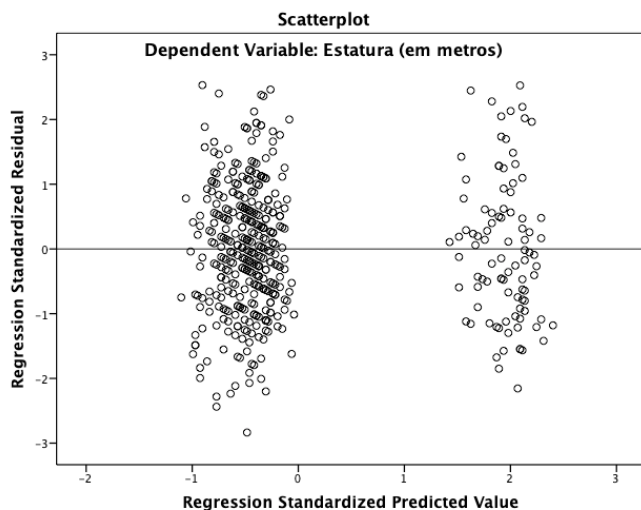


Figura 7. Gráfico de *scatterplot* dos resíduos versus valores preditos.

Você pode agrupar variáveis independentes em blocos e especificar métodos de entrada diferentes para diferentes subconjuntos de variáveis. Quando pretendemos testar diferentes modelos, de forma manual, seleciona-se, na janela "Regressão linear", o "Método inserir" (*Enter*) e inclui-se uma variável por vez em cada "Bloco", utilizando as opções "Anterior" e "Próximo" para definir a ordem de entrada das variáveis no modelo (conforme Quadro 1, página 201). No arquivo de saída, obtém-se os resultados para cada modelo enumerado.

Também pode-se construir os modelos de forma automática (*stepwise*), utilizando os modelos "Avançar" (*Forward*) (as variáveis são inseridas sequencialmente no modelo, sendo a primeira variável aquela com a maior correlação positiva ou negativa com a variável dependente e, assim, sucessivamente),

“Retroceder” (*Backward*)” (todas as variáveis são inseridas na equação e, em seguida, todas as variáveis selecionadas pelo pesquisador são incluídas, sendo retirada, uma a uma, a variável menos significativa), “Remover” (todas as variáveis em um bloco são removidas em um único passo) e “Por etapa” (em cada passo, a variável independente fora da equação que tiver o menor índice F será inserida, se esse índice for suficientemente pequeno e, as variáveis que já estiverem na equação da regressão serão removidas se o índice F for suficientemente grande) no “Método”.

Referências

Callegari-Jacques SM. Bioestatística: princípios e aplicações. Porto Alegre: ArtMed, 2011. 255p.

Field A. Descobrimo a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Harrell , FE. General Aspects of Fitting Regression Models. In: Regression Modeling Strategies. Springer Series in Statistics. Cham: Springer International Publishing; 2015:13-44.

Heumann C, Schomaker M, Shalabh. Linear Regression. In: Introduction to Statistics and Data Analysis. Cham: Springer International Publishing; 2016:249-295.

Montgomery DC. Fitting Regression Models. In: Design and Analysis of Experiments Eighth Edition. Vol 2. ; 2012:449-475.

Exercícios sugeridos

1. No “Banco de dados 2.sav”, verifique se a pressão arterial sistólica depende da massa corporal, supondo distribuição normal da variável dependente. Realize a análise estatística, interprete os resultados e verifique os pré-requisitos da análise de resíduos.

2. No “Banco de dados 2.sav”, verifique se a pressão arterial sistólica depende da circunferência da cintura e do sexo, supondo distribuição normal da variável dependente. Realize a análise estatística, interprete os resultados e verifique os pré-requisitos da análise de resíduos.

11 Modelos lineares generalizados

*Sinara Santos
Nadine Ziegler
Pedro Henrique Comerlato
Edison Capp
Otto Henrique Nienov*

Os modelos lineares generalizados correspondem a um conjunto de análises de regressão que são uma generalização dos modelos lineares gerais. Enquanto que os modelos lineares gerais são usados para modelos de regressão e variância com uma distribuição normal e uma função de ligação identidade, os modelos lineares generalizados permitem a utilização de variáveis com distribuição assimétrica.

Para definir qual modelo linear generalizado tem de ser aplicado, deve-se considerar o tipo de variável dependente (Quadro 1) e o tipo de delineamento do estudo (Quadro 2).

Quadro 1. Análise de regressão de acordo com o tipo de variável dependente.

Tipo de variável	Tipo de regressão
Qualitativa nominal	Regressão logística binária
	Regressão de Poisson
	Regressão logística politômica ou multinominal
Qualitativa ordinal	Regressão logística ordinal

Regressão logística

A regressão logística é muito similar à regressão linear, pois ambas avaliam a relação entre uma variável dependente (variável desfecho ou de resposta) e uma ou mais variáveis preditoras (variável

independente ou de exposição). O quadro 3 elenca as principais diferenças e semelhanças entre as regressões linear e logística.

Quadro 2. Análise de regressão de acordo com o delineamento do estudo.

Tipo de estudo	Tipo de regressão
Caso-controle	Regressão logística
ECR	Regressão de Poisson
Transversal	
Coorte	Regressão de Cox

Em relação as semelhanças entre a regressão linear e a logística, existe uma boa razão para que não possamos aplicar a regressão linear diretamente a uma situação onde a variável de saída é dicotômica: uma das hipóteses da regressão linear é que o relacionamento entre as variáveis seja linear. Portanto, para o modelo de regressão linear ser válido, os dados observados devem ter um relacionamento linear. Quando a variável de saída é dicotômica, essa hipótese é normalmente violada.

Quadro 3. Comparação entre regressão linear e regressão logística.

Características	Regressão linear	Regressão logística
Variável dependente	Analisa respostas quantitativas	Analisa desfechos categóricos
Variáveis independentes	Podem ser categóricas (também nomeadas como fatores) e/ou contínuas (nomeadas como covariáveis)	
Interpretação dos resultados	Obtém-se uma associação em relação a uma função linear	Associação é quantificada por meio de uma razão da probabilidade de ocorrência
Reta da regressão	Reta	Curva

Regressão logística binária

A regressão logística binária determina a probabilidade de ocorrência das categorias de uma variável dicotômica, ou seja, prevê qual de duas categorias é mais provável, dado um conjunto de informações. Descomplicando, isso quer dizer que é possível prever a qual de duas categorias uma pessoa é provável de pertencer dado certas informações. A figura 1 descreve algumas considerações pertinentes para a análise de regressão logística binária, relacionadas a variável dependente e aos preditores do modelo.

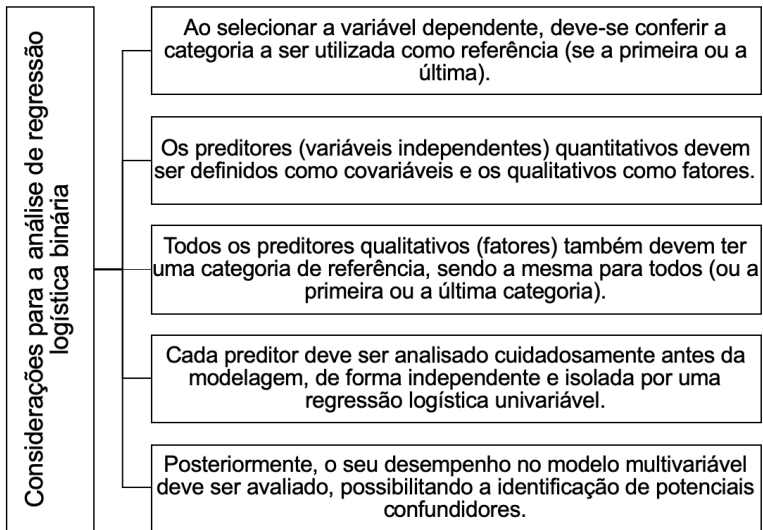


Figura 1. Considerações referentes a variável dependente e aos preditores (variáveis independentes) para a análise de regressão logística binária.

Na sua forma simples, quando existe um único previsor X , a equação da regressão logística, a partir da qual a probabilidade da variável Y é prevista, define-se em:

$$P(Y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \varepsilon)}}$$

Onde $P(Y)$ é a probabilidade de Y ocorrer, e é a base dos logaritmos naturais e os demais coeficientes da equação formam uma combinação linear muito semelhante à regressão linear simples. A parte da equação entre parênteses é idêntica à regressão linear, pois existe uma constante (β_0), uma variável previsor (X) e um coeficiente (ou peso) agregado ao previsor (β_1).

Da mesma forma que para a regressão linear, é possível estender essa equação para incluir diversas variáveis predictoras. Quando isso ocorre, a equação toma a seguinte forma:

$$P(Y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon)}}$$

A regressão logística, ao contrário da regressão linear, não pressupõe normalidade. Mas, da mesma forma, deve-se investigar a multicolinearidade entre as variáveis independentes e a presença de valores atípicos. Também se recomenda um tamanho amostral mínimo para a análise de regressão, dependendo do número de variáveis predictoras. Alguns autores indicam 10 casos para cada previsor no modelo, outros 15 casos por previsor.

Para exemplificar, faremos uma regressão logística binária univariável utilizando o "Banco de dados 7.sav" (disponível em <https://bit.ly/bancosdedados>), supondo um delineamento de estudo do tipo caso-controle, em que o critério de seleção foi a ocorrência ou não de polineuropatia periférica (PNP) (onde, "0" = "Não neuropata" e "1" = "Neuropata") após a cirurgia bariátrica. Pretende-se determinar se a idade (variável quantitativa) e o sexo (variável qualitativa) (onde, "1" = "Masculino" e "2" = "Feminino") possuem associação com o desfecho PNP. Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Atenção! Deve-se ter cuidado com as tabelas de contingência contendo células vazias. Incluir estas variáveis no modelo causará resultados numéricos indesejados e a melhor estratégia para resolver este problema é o agrupamento desta variável em um conjunto menor de categorias.

Como visto no capítulo 10, para definirmos um modelo de regressão, podemos utilizar como critério de seleção as variáveis que se mostraram estatisticamente significativas ou com p-valor limítrofe para associação na análise univariada. No exemplo, houve uma diferença significativa entre as medianas de idade dos sujeitos com e sem PNP (50 anos versus 34 anos, respectivamente; $p < 0,001$) pelo teste U de Mann-Whitney e, uma indicação de associação de PNP com o sexo masculino (61,8%; $p < 0,001$) pelo teste de correção de continuidade de Yates. Essas associações nos motivam a avaliar o impacto da idade e do sexo na PNP.

Primeiro, precisamos verificar a colinearidade e a presença de valores atípicos. Essa verificação é feita da mesma forma que na regressão linear. No menu "Analisar", "Regressão", clique em "Linear...". Na janela "Regressão linear", selecione a variável "Polineuropatia periférica" na tela da esquerda e a insira em "Dependente" e, as variáveis "Idade" e "Sexo" como "Independente(s)". Clique em "Estatísticas" e selecione as opções "Diagnóstico de colinearidade" e "Diagnóstico por caso". Clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, verifique os parâmetros de avaliação de colinearidade, índice de Tolerância e VIF. Lembre-se que o VIF deve ser menor que 10 e o índice de Tolerância maior que 0,2. Neste caso, temos um VIF de 1,177 e um índice de Tolerância de 0,850, ou seja, não há colinearidade entre as variáveis. Na análise, também, foi identificado um *outlier* (caso 47). Nestes casos, o recomendado é realizar a análise de regressão em um modelo sem e o outro com o caso, para verificar a influência deste sobre o resultado.

Verificados a colinearidade e valores atípicos, podemos prosseguir para a regressão logística binária univariável, avaliando primeiro a idade (variável quantitativa). No menu "Analisar", "Modelos lineares generalizados", clique em "Modelos lineares generalizados...". Na janela "Modelos lineares generalizados", precisaremos definir alguns parâmetros. Na aba "Tipo de Modelo", selecione a opção "Logística binária". Na aba "Resposta", selecione e insira em "Variável dependente" a variável "Polineuropatia periférica". Na mesma aba, em "Tipo de variável dependente (apenas distribuição binomial)", marque

"Binário". Clique em "Categoria de referência..." e selecione a opção "Primeiro (menor valor)", onde "0" = "Não neuropata". Clique em "Continuar". Na aba "Preditores", selecione na tela da esquerda e insira em "Covariáveis" a variável "Idade" (variável quantitativa). Na aba "Modelo", selecione em "Fatores e covariáveis" a variável "Idade" e mova-a para "Modelo" marcando como "Tipo: Efeito Principal". Na aba "Estatísticas", marque as opções "Razão de probabilidade" em "Estatísticas de quiquadrado", "Probabilidade do perfil" em "Tipo de intervalo de confiança" e "Incluir estimativas de parâmetro exponencial" em "Imprimir". Por fim, clique em "Ok" ou "Colar".

Nos resultados da análise, no arquivo de saída, os primeiros quadros mostram as informações do modelo, as estatísticas descritivas da variável dependente e da variável preditora. O primeiro quadro (*Model Information*) mostra as informações gerais do modelo. O segundo quadro (*Case Processing Summary*) apresenta o número de casos incluídos, excluídos e total. Em seguida (*Categorical Variable Information*), temos a proporção absoluta e relativa para as categorias da variável dependente (PNP), onde 27,5% ($n = 55$) da amostra é neuropata. No quarto quadro (*Continuous Variable Information*), temos o valor absoluto (N), média (*Mean*), desvio padrão (*Std. Deviation*) e os valores mínimo (*Minimum*) e máximo (*Maximum*) para a covariável (Idade).

Após, encontramos no quadro "*Goodness of Fit*" diferentes medidas para avaliar a qualidade de ajuste do modelo, úteis para comparar modelos concorrentes. Em todas, grandes valores indicam falta de ajuste. O resultado do "*Omnibus Test*" nos indica se o modelo de regressão foi significativo. Neste caso, com $p < 0,001$, temos evidência que o modelo da regressão é estatisticamente significativo.

No quadro "*Tests of Model Effects*", temos o teste de Quiquadrado de *Likelihood Ratio* para a significância dos preditores. Esse verifica a existência de diferenças entre o modelo com o preditor e o modelo nulo (intercepto). Se significativo (neste caso tem um $p < 0,001$), o teste informa que o coeficiente do preditor difere de zero. Se isto ocorrer, pode-se dizer que o preditor está contribuindo de modo significativo para o modelo.

Constatada a significância do preditor, o quadro "*Parameter Estimates*" apresenta os parâmetros estimados. Encontramos o coeficiente B (que corresponde a constante β), o erro padrão (*Std. Error*), o limite inferior (95% *Profile Likelihood Confidence Interval: Lower*) e superior (95% *Profile Likelihood Confidence Interval: Upper*) do intervalo de confiança para B. Também é apresentado o teste Qui-quadrado de Wald, que testa as mesmas hipóteses do teste Qui-quadrado de *Likelihood Ratio*. No entanto, utilizamos os resultados do teste Qui-quadrado de *Likelihood Ratio* pois é mais poderoso que o teste de Wald. Além disso, a estatística de Wald deve ser vista com cautela porque quando o coeficiente de regressão (β) é grande, o erro padrão tende a ficar inflacionado, resultando em uma estatística de Wald subestimada. A inflação do erro padrão aumenta a probabilidade de que o previsor seja rejeitado quando na realidade ele contribui para o modelo (erro do Tipo II).

Nas últimas colunas, encontra-se o OR, expresso pelo valor de $\text{Exp}(B)$ e seu intervalo de confiança. O mais importante para a interpretação da regressão logística é o valor da $\text{Exp}(B)$, que é um indicador da mudança nas probabilidades resultantes da mudança de uma unidade no previsor. O OR ($\text{Exp}(B)$) indica a presença de um fator de risco ou proteção (Quadro 4) e é interpretado da mesma forma vista no capítulo 9.

Quadro 4. Interpretação dos valores de OR.

Valor de OR	Interpretação
$\text{Exp}(B) > 1$	Indica que à medida que o previsor aumenta, aumentam as chances de ocorrência do desfecho.
$\text{Exp}(B) = 1$	O efeito é nulo.
$\text{Exp}(B) < 1$	Indica que à medida que o previsor aumenta, as chances de o desfecho ocorrer diminuem.

Conclui-se que a idade é um preditor de PNP. Usando a primeira categoria da variável PNP como referência, PNP igual à zero (não neuropata), podemos afirmar que a chance de ter PNP aumenta em 41,5% (IC95%: 28,6%-60,3%) para cada ano de incremento na idade ($p < 0,001$).

Em seguida, vamos avaliar o sexo (variável qualitativa dicotômica). No menu "Analisar", "Modelos lineares generalizados", clique em "Modelos lineares generalizados...". Na aba "Tipo de modelo", selecione a opção "Logística binária". Na aba "Resposta", selecione na tela da esquerda e insira em "Variável dependente" a variável "Polineuropatia periférica". Na mesma aba, em "Tipo de variável dependente (apenas distribuição binominal)", marque "Binário". Clique em "Categoria de referência..." e selecione a opção "Primeiro (menor valor)", onde "0" = "Não neuropata". Clique em "Continuar". Na aba "Preditores", selecione na tela da esquerda e insira em "Fatores" a variável "Sexo" (variável dicotômica). Como se trata de uma variável categórica, também precisamos informar a categoria de referência. Clique em "Opções", localizado abaixo da tela de "Fatores" e selecione como "Crescente" (assume que a categoria de referência é a última em todos os fatores). A opção "Decrescente" assume que a categoria de referência é a primeira em todos os fatores e, "Usar ordem de dados" assume que a categoria de referência é a que tem maior probabilidade de ocorrência do desfecho. Clique em "Continuar". Na aba "Modelo", selecione em "Fatores e covariáveis" a variável "Sexo" e mova-a para "Modelo" marcando como "Tipo: Efeito Principal". Na aba "Estatísticas", marque as opções "Razão de probabilidade" em "Estatísticas de qui-quadrado", "Probabilidade do perfil" em "Tipo de intervalo de confiança" e "Incluir estimativas de parâmetro exponencial" em "Imprimir". Por fim, clique em "Ok" ou "Colar".

Os resultados dos primeiros quadros são idênticos aos obtidos para a variável "Idade", porém com ajustes para um preditor categórico. O quadro "*Omnibus Test*" nos indica que o modelo de regressão foi significativo ($p < 0,001$). A principal diferença está nos parâmetros estimados da tabela "*Parameter Estimates*", onde teremos um parâmetro para cada categoria da variável. Os resultados estão em comparação ao sexo feminino, cuja categoria é "2". Por isso, em "[SEXO=2]" o coeficiente está zerado, assim como o OR ($\text{Exp}(B)$) é um ("1"), pois a categoria está sendo comparada com ela mesma.

Assim, conclui-se que o sexo é um preditor significativo para a ocorrência de PNP ($p < 0,001$). A probabilidade de ter PNP é 11,4 (IC95%: 5,6-24,4) vezes maior para o sexo masculino em comparação ao sexo feminino.

Considerando o mesmo exemplo, será que um modelo multivariado (com as variáveis "Sexo", "Idade" e "Massa corporal") melhor prediz a ocorrência de PNP? Como já mencionado, podemos utilizar como critério de seleção as variáveis que se mostraram estatisticamente significativas ou com p-valor limítrofe para associação na análise univariada. Para a massa corporal, obteve-se um p-valor limítrofe para associação ($p = 0,062$) pelo teste U de Mann-Whitney.

Modelos com múltiplas variáveis analisam a relação entre o desfecho e as variáveis preditoras controlando a influência de outras variáveis. No menu "Analisar", "Modelos lineares generalizados", clique em "Modelos lineares generalizados...". Na aba "Tipo de modelo", selecione "Logística binária". Na aba "Resposta", selecione na tela da esquerda e insira em "Variável dependente" a variável "Polineuropatia periférica". Na mesma aba, em "Tipo de variável dependente (apenas distribuição binominal)", marque "Binário". Clique em "Categoria de referência..." e selecione a opção "Primeiro (menor valor)", onde "0" = "Não neuropata". Clique em "Continuar". Na aba "Preditores", selecione em "Fatores" a variável "Sexo" e, em "Covariáveis", as variáveis "Idade" e "Massa corporal". Como a variável "Sexo" é uma variável categórica, precisamos informar a categoria de referência. Clique em "Opções", localizado abaixo da tela de "Fatores" e selecione como "Crescente" (assume que a categoria de referência é a última em todos os fatores). Na aba "Modelo", selecione em "Fatores e covariáveis" as variáveis "Sexo", "Idade" e "Massa corporal" e as insira em "Modelo", marcando "Tipo: Efeitos principais". Na aba "Estatísticas", marque as opções "Razão de probabilidade" em "Estatísticas de quiquadrado", "Probabilidade do perfil" em "Tipo de intervalo de confiança" e "Incluir estimativas de parâmetro exponencial" em "Imprimir". Por fim, clique em "Ok" ou "Colar".

Novamente, no arquivo de saída, temos os mesmos quadros vistos na análise univariável, com a descrição do modelo, das variáveis e do ajuste do modelo. Assim como na análise univariável, o quadro "Omnibus Test" nos indica que o modelo de regressão foi significativo ($p < 0,001$). Entretanto, na análise multivariável, a tabela "Tests of Model Effects" ganha maior importância, pois no modelo multivariável ela não se equivale à tabela "Omnibus Test".

A significância de um modelo multivariado na tabela “*Omnibus Test*” denota a significância de pelo menos um dos preditores (análise global do modelo), para tanto, deve-se checar se todas as variáveis independentes devem permanecer no modelo. Para o exemplo em questão, os preditores “Sexo” e “Idade” possuem coeficiente significativamente diferente de zero, mas o “Massa corporal” não. Neste momento, deve-se ponderar se este será o modelo final ou se algum preditor deve sair ou entrar antes da análise dos coeficientes, pois os mesmos são influenciados pelo conjunto de variáveis independentes.

Constatada a significância dos preditores, a tabela “*Parameter Estimates*” apresenta os parâmetros estimados. Concluindo, pode-se afirmar que, controlando por sexo e massa corporal, as chances de ter PNP aumentam em 38,0% (IC95%: 25,6%-56,1%, $p < 0,001$) para cada aumento de um ano de idade. Controlando por sexo e idade, a massa corporal não prediz a ocorrência de PNP ($p = 0,727$). E, controlando-se o efeito da idade e da massa corporal, as chances de PNP em homens aumentam em 7,8 (IC95%: 2,5-26,2, $p < 0,001$) vezes em relação as mulheres.

Regressão de Poisson

O modelo de regressão de Poisson é adequado quando os dados das variáveis dependentes são contáveis, ou seja, quando se descreve o número de vezes que um evento ocorre em um espaço de observação finito. Tem por característica a análise de dados contados na forma de proporções ou razões de contagem, ou seja, leva em consideração o total de pessoas com uma determinada doença.

Esta análise de regressão é utilizada em estudos com delineamento transversal (razão de prevalências) ou longitudinal (razão de incidências), inclusive em ensaios clínicos randomizados, cujos desfechos (variáveis dependentes) são variáveis binárias ou dicotômicas do tipo óbito/não óbito ou doença/não doença, por exemplo. As variáveis independentes (fatores em estudo) podem ser categóricas ou contínuas. É imprescindível que, ao se utilizá-la para avaliar desfechos dicotômicos, seja aplicado um ajuste robusto nas variâncias.

Para exemplificar, no “Banco de dados 7.sav” (disponível em <https://bit.ly/bancosdedados>), vamos avaliar a magnitude da associação (através da razão de prevalências) entre o sexo e a presença de PNP, supondo um estudo de delineamento transversal. No menu “Analisar”, “Modelos lineares generalizados”, clique em “Modelos lineares generalizados...”. Como o teste está localizado no mesmo local onde foi realizada a regressão logística, clique em “Redefinir”. Na aba “Tipo de modelo”, defina a distribuição da variável dependente como “Linear de log de Poisson”. Na aba “Resposta”, selecione a variável dependente “Polineuropatia periférica” na tela da esquerda e a insira em “Variável dependente”. Na aba “Preditores”, selecione a(s) variável(is) independente(s), da mesma forma vista na regressão logística: se variável categórica, insira em “Fatores”; se contínua, mova para “Covariáveis”. Como a variável “Sexo” é uma variável dicotômica, insira em “Fatores” e defina, em “Opções”, a categoria de referência como “Crescente”. Na aba “Modelo”, selecione a variável “Sexo” em “Fatores e Covariáveis” e a insira em “Modelo” marcando a opção “Tipo: Efeitos principais”. Na aba “Estimativa”, marque “Estimador robusto” em “Matriz de covariância” e, na aba “Estatísticas”, selecione a opção “Incluir estimativas de parâmetro exponencial”. Clique em “Ok” ou “Colar”.

Nos resultados da análise, os primeiros quadros mostram as informações utilizadas na análise: a variável dependente, a distribuição e a função de ligação, a matriz de correlação, o número total de sujeitos e suas discriminações pelas variáveis independentes (fatores). O primeiro quadro que temos que verificar é o “*Omnibus Test*”, que nos indica se o modelo de regressão foi significativo. Neste caso, com $p < 0,001$, temos evidência de que o modelo é estatisticamente significativo. Na tabela “*Test of Model Effects*”, temos os resultados dos efeitos principais e da interação entre as variáveis, caso solicitada no modelo. Pelo teste de Qui-quadrado de Wald, neste exemplo, o efeito do sexo foi significativo ($p < 0,001$), ou seja, há uma diferença nas prevalências de PNP entre homens e mulheres.

Em seguida, avalia-se o valor da razão de prevalências no quadro “*Parameter Estimates*”. Neste quadro, tem-se os resultados dos coeficientes calculados pelo modelo para cada uma das variáveis,

juntamente com erros-padrão, a exponencial do coeficiente (razão de prevalências) e intervalos de confiança para cada variável. Como foi indicado a categoria de referência como sendo "Crescente", a última categoria ("2" = "Feminino") foi determinada como a categoria de referência. Assim, temos que, com uma razão de prevalências de 4,608, nos homens a chance de apresentar PNP é 4,6 (IC95%: 3,0-7,2; $p < 0,001$) vezes maior em relação às mulheres.

Análise de sobrevida

Em estudos que se caracterizam por tempo de seguimento (longitudinais), além da variável de desfecho, como ocorrência de um evento, também precisamos levar em consideração o tempo médio para o evento ocorrer. Neste sentido, a análise de sobrevida possui como variável dependente o tempo de ocorrência de determinado evento. Por exemplo, a taxa de sobrevida entre pacientes submetidos a certo medicamento ou procedimento. O estado do paciente (ter ou não o evento) é uma variável dicotômica (binária) que define a situação individual de sua participação no seguimento. Este recebe um código zero ("0") se o evento não ocorrer e um ("1") se o evento ocorrer.

Uma vantagem dessa análise é permitir a utilização da informação de todos os participantes até o momento em que desenvolvem o evento ou são censurados. Ou seja, se o paciente sem o evento desistir ou abandonar o estudo no meio do seguimento, registra-se o tempo de permanência do sujeito no estudo e coloca-se o código zero ("0") (referente a não ocorrência do evento). Esses pacientes com código zero, pois não desenvolveram o evento, são chamados de censurados. As informações a partir de sujeitos censurados contribuem positivamente para a estimação do modelo.

Regressão de Cox

A regressão de Cox permite modelar, de forma uni ou multivariada, as variáveis que influenciam no tempo de ocorrência de um determinado evento. O Modelo de Cox é dito como "semi-paramétrico", pois assume que as taxas de falhas são proporcionais, ou seja, o risco de falha das variáveis é constante ao longo do tempo.

A regressão é muito utilizada para verificar inicialmente se a variável independente possui uma relação com a variável desfecho e avaliar o efeito das variáveis explicativas. Para realizar a regressão de Cox univariável são necessárias três variáveis: tempo, estado do paciente e variável independente. Já para a multivariável, utilizam-se duas ou mais variáveis independentes ou explicativas (categóricas e contínuas).

A interpretação dos coeficientes obtidos dá-se pela razão de taxas de falha ou RR. A interpretação do RR é similar à do OR da regressão logística. Quando o evento medido é óbito, a função de risco hazard ratio (HR), também chamada taxa de risco (de falha), representa a taxa instantânea de morte de um indivíduo no intervalo de tempo t a $t+1$, sabendo que sobreviveu até o momento t . Valores altos do HR indicam grande potencial de que o evento ocorra.

Para exemplificar, no “Banco de dados 7.sav” (disponível em <https://bit.ly/bancosdedados>), iremos medir o tempo na ocorrência da PNP (o evento é ter ou não PNP). Inicialmente, realizaremos uma análise de regressão de Cox univariável com a variável independente “Hipertensão” (variável categórica).

No menu “Analisar”, “Sobrevivência”, clique em “Regressão de Cox...”. Na janela “Regressão de Cox”, selecione na tela da esquerda a variável “Tempo pós-bariátrica” e a insira em “Tempo”, com o auxílio do botão da seta. Em seguida, na mesma tela, selecione a variável “Polineuropatia periférica” e a mova para “Status”. Clique em “Definir evento” e indique o código da categoria referente a ocorrência do evento em “Valor único” (“1” = “Neuropata”). Clique em “Continuar”. Ainda na mesma janela, insira o fator ou a covariável de interesse em “Covariáveis” (neste caso, “Hipertensão”). Nos modelos multivariados coloca-se mais de uma variável nesse quadro. Como temos uma variável independente categórica, clique em “Categórico...”. Selecione para a direita somente a variável “Hipertensão” e indique se a menor (ou a maior) categoria é a de referência. Neste caso, nossa referência será a primeira (“0” = “Não hipertenso”). Não é possível classificar uma categoria intermediária como referência. Não esqueça de clicar em “Alterar” se mudar de categoria. Clique em “Continuar”.

Em seguida, clique em "Opções" e selecione "CI para exp(B)". Clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

Quando trabalhamos com variáveis categóricas, indicamos uma das suas categorias para ser a referência. A categoria "Não hipertenso" da variável "Hipertensão" é a categoria de referência, neste exemplo. O SPSS mostra em um quadro (*Categorical Variable Codings*) no arquivo de saída, as variáveis com suas categorias de referência. O valor zero ("0") na coluna "1" mostra qual categoria é a de referência. O quadro "*Variables in the Equation*" possui as informações necessárias para podermos interpretar a relação entre as variáveis estudadas. A primeira informação encontrada é do parâmetro β (B) e seu erro-padrão (SE). A exponencial do parâmetro B (Exp(B)) e seu intervalo de confiança (IC95%) são encontrados nas últimas colunas. A Exp(B) foi de 2,780, ou seja, indivíduos hipertensos apresentam taxa de risco para PNP 2,78 (IC95%: 1,63-4,73, $p < 0,001$) vezes maior do que indivíduos não hipertensos.

Em seguida, vamos realizar uma análise de regressão de Cox univariável com a variável independente "Massa corporal" (variável quantitativa). No menu "Analisar", "Sobrevivência", clique em "Regressão de Cox...". Primeiro, na janela "Regressão de Cox", vamos "Redefinir" as opções de teste. Após, selecione na tela da esquerda a variável "Tempo pós-bariátrica" e a insira em "Tempo", com o auxílio do botão da seta. Em seguida, na mesma tela, selecione a variável "Polineuropatia periférica" e a mova para "Status". Clique em "Definir evento" e indique o código da categoria referente a ocorrência do evento em "Valor único" ("1" = "Neuropata"). Clique em "Continuar". Ainda na mesma janela, insira o fator ou a covariável de interesse em "Covariáveis" (neste caso, "Massa corporal"). Em seguida, clique em "Opções" e selecione "CI para exp(B)". Clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

Quando temos uma variável contínua como variável independente, a sua interpretação é relacionada ao acréscimo de uma unidade. Neste caso, a massa corporal obteve uma Exp(B) igual a 1,078. Assim, interpretamos que para cada aumento de 1 kg na massa corporal, o risco de PNP aumenta em 7,8% (IC95%: 5,9-9,7, $p < 0,001$). As interpretações para modelo multivariável seguem da mesma forma.

Referências

Barton, B, Peat, J. Medical Statistics: A guide to SPSS. Data analysis and Critical Appraisal. Second edition. John Wiley & Sons, Ltda. 2014.

Field A. Descobrimdo a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Exercícios sugeridos

1. No "Banco de dados 7.sav", supondo um estudo caso-controle, em que o critério de seleção foi a ocorrência ou não de polineuropatia periférica ("0" = "Não neuropata" e "1" = "Neuropata") após a cirurgia bariátrica, verifique se o tipo de cirurgia realizada ("1" = Bypass gástrico, RYGB e "2" = gastrectomia vertical, SG) possui associação com o desfecho. Realize a análise estatística e interprete os resultados.

2. No "Banco de dados 7.sav", supondo um estudo transversal, avalie a magnitude da associação (através da razão de prevalências) entre o sexo e a presença de hipertensão. Realize a análise estatística e interprete os resultados.

Resolução dos Exercícios

Capítulo 1

Respostas corretas:

1. b) Trata-se de um estudo caso-controle, onde temos um grupo de indivíduos com doença (câncer de endométrio) e outro sem doença (controle) pareados por idade e paridade.
2. a) Trata-se de um estudo de coorte, onde temos um grupo de mulheres que é acompanhado para verificar a ocorrência e os fatores de risco para o câncer de mama.
3. e) Trata-se de um estudo transversal, em que a amostra é estudada em um único momento para avaliar a prevalência de Síndrome do Ovário Policístico.
4. d) Trata-se de um estudo de caso, onde temos o relato de uma doença rara (Síndrome de Rokitansky) em uma paciente.
5. a) Trata-se de um estudo transversal, em que os dados são coletados retrospectivamente em um dado momento.
6. c) Trata-se de um ensaio clínico randomizado, onde temos um grupo com intervenção (estimulação elétrica nervosa transcutânea associada a analgesia combinada raqui-peridural) e outro sem intervenção (analgesia combinada raqui-peridural).
7. e) Trata-se de um estudo caso-controle, onde temos um grupo de indivíduos com doença (câncer de mama) e outro sem doença (controle).
8. a)
 - P = obesos não diabéticos submetidos à cirurgia bariátrica;
 - I = obesos com nível sérico de HDL-Colesterol baixo;
 - C = obesos com nível sérico de HDL-Colesterol normal;
 - O = alterações no nível sérico de HDL-Colesterol após a cirurgia bariátrica;
 - S = estudo de coorte prospectivo.

Lembre-se que um estudo de coorte é um estudo observacional, longitudinal, analítico, prospectivo ou retrospectivo. Neste caso, os grupos devem ser comparáveis e, preferencialmente, representativos da mesma população (obesos não diabéticos submetidos à cirurgia bariátrica). Os sujeitos devem ser similares nas outras características que são consideradas de risco para a doença, à exceção do fator de exposição em estudo (HDL-Colesterol baixo antes da cirurgia bariátrica). A comparação ocorre entre um grupo exposto (HDL-Colesterol baixo) e outro não exposto (HDL-Colesterol normal) ao longo do tempo, indo de encontro às consequências desta exposição (nível do HDL-Colesterol após a cirurgia bariátrica). Neste exemplo, a cirurgia bariátrica não corresponde à intervenção. Teríamos esta situação em um ensaio clínico randomizado, onde os sujeitos obesos não diabéticos seriam randomizados em dois grupos, um com intervenção (tratamento cirúrgico) e outro sem intervenção (tratamento clínico) e, ao final, avaliaríamos o nível sérico de HDL-Colesterol nestes grupos.

b) As variáveis quantitativas correspondem a idade, tempo de cirurgia, perda de massa corporal, circunferência da cintura, níveis séricos de glicose, triglicérides, HDL-Colesterol, não HDL-Colesterol, LDL-Colesterol, além do gasto energético estimado, sendo todas contínuas. As variáveis qualitativas correspondem ao sexo (nominal), técnica cirúrgica (nominal) e grau de obesidade (ordinal).

9. Resolução não aplicável.

Capítulo 2

1. Tabela 1

Nível	Número Absoluto (n)	Número relativo (n%)
Mestrado	60	52,2
Doutorado	55	47,8
Total	115	100,0

Distribuição do nível de Pós-Graduação entre alunos da GIN026

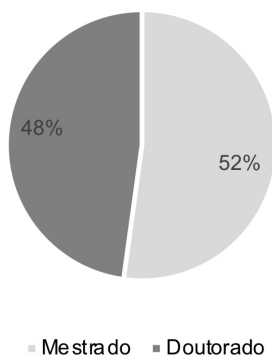


Figura 1

Interpretação: dos 115 alunos da disciplina, 52,2% (n=60) são do nível mestrado.

2.

Tabela 1. Características antropométricas e de pressão arterial de sujeitos com obesidade. Os dados quantitativos estão apresentados em mediana e valores mínimo e máximo.

Tabela 2

	Obesos (n=469)
Idade (anos)	37 (18 - 66)
IMC (kg/m ²)	41,34 (34,04 - 77,18)
Circunferência da cintura (cm)	111 (74 - 166)
PAS (mmHg)*	129 (96 - 180)
PAD (mmHg)*	86 (53 - 126)

*Para as variáveis pressóricas, 468 indivíduos foram avaliados. IMC: Índice de Massa Corporal; PAS: Pressão Arterial Sistólica; PAD: Pressão Arterial Diastólica; n: número absoluto.

Interpretação: Nos sujeitos com obesidade foram encontradas:
 - Mediana de idade de 37 anos, com valor mínimo de 18 anos e máximo de 66 anos.

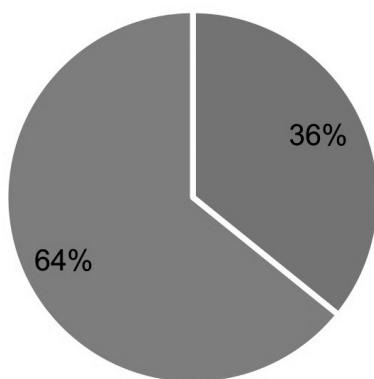
- Mediana de IMC de 41,34 kg/m², com valor mínimo de 34,04 kg/m² e máximo de 77,18 kg/m².
- Mediana de circunferência da cintura de 111 cm, com valor mínimo de 74 cm e máximo de 166 cm.
- Mediana de pressão arterial sistólica de 129 mmHg, com valor mínimo de 96 mmHg e máximo de 180 mmHg.
- Mediana de pressão arterial diastólica foi de 86 mmHg, com valor mínimo de 53 mmHg e máximo de 126 mmHg.

3. Tabela 3

Grau de Obesidade	Número absoluto (n)	Número relativo (n%)
Grau II	168	35,8
Grau III	301	64,2
Total	469	100,0

Interpretação: dos 469 sujeitos com obesidade, 64,2% (n=301) apresentam obesidade grau III.

Frequência dos graus de obesidade



■ Grau II ■ Grau III

Figura 2

Tabela 4

Hipertensão	Número absoluto (n)	Número relativo (n%)
Hipertenso	392	83,8
Não hipertenso	76	16,2
Total	468	100,0

Frequência de hipertensão em sujeitos obesos

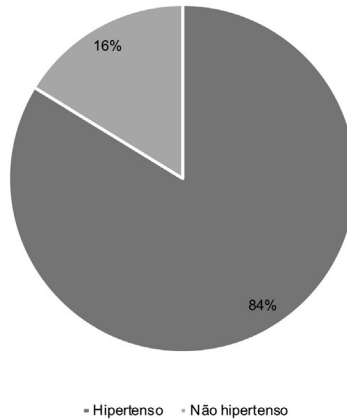


Figura 3

Interpretação: dos 468 sujeitos com obesidade, 83,8% ($n = 392$) são hipertensos. A hipertensão foi avaliada em 468 sujeitos da amostra.

Capítulo 3

1. Resolução não aplicável.

2.a.

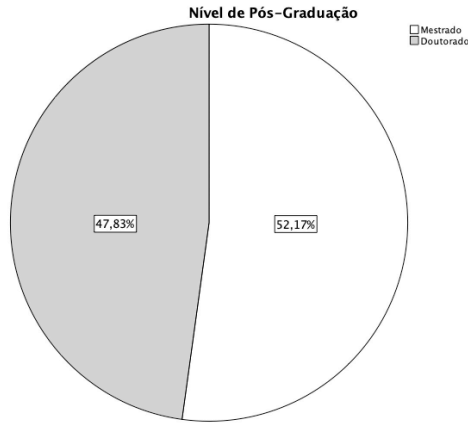


Figura 1

Interpretação: dos 115 alunos da disciplina, 52,2% (n = 60) estão no nível do mestrado.

b.

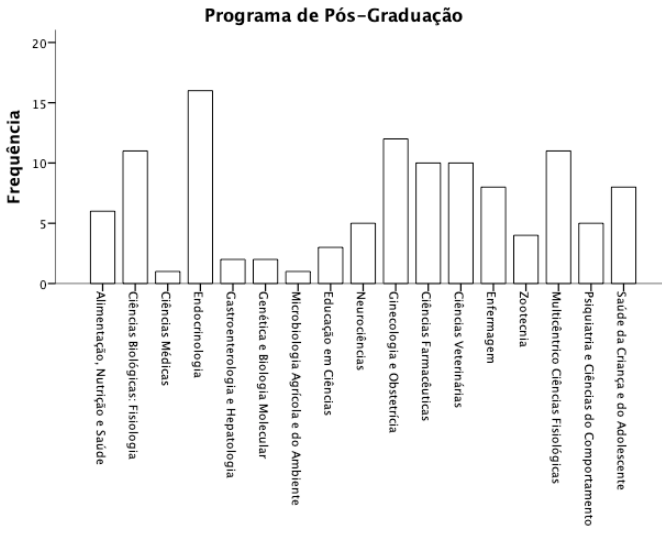


Figura 2

Interpretação: o programa de pós-graduação de maior frequência é o da Endocrinologia (n = 16, 13,9%).

C.

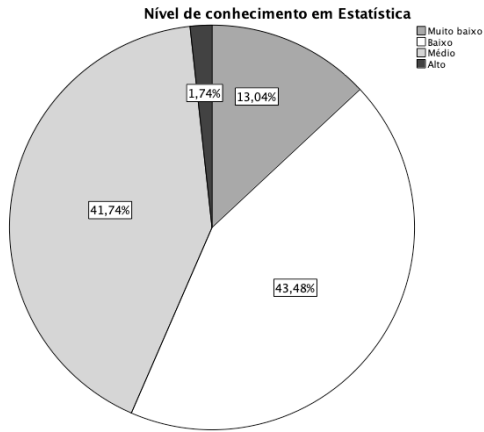


Figura 3

Interpretação: dos alunos da disciplina, 43,5% (n = 50) consideram seu nível de conhecimento em estatística baixo e 41,7% (n = 48) o consideram médio.

d.

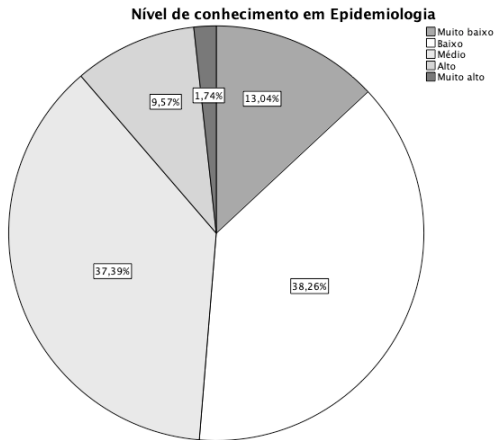


Figura 4

Interpretação: dos alunos da disciplina, 38,3% (n = 44) consideram seu nível de conhecimento em epidemiologia baixo e 37,4% (n = 43) o consideram médio.

e.



Figura 5

Interpretação: dos 115 alunos da disciplina, 62,6% (n = 72) não possuem experiência no SPSS.

3.

Nos sujeitos com obesidade (n = 469) foram encontradas:

- Mediana de idade de 37 (18-66) anos ou 37 (31-45) anos.
- Mediana de IMC de 41,34 (35,04-77,18) kg/m² ou 41,34 (38,69-44,79) kg/m².
- Mediana de circunferência da cintura de 111 (74-166) cm ou 111 (104-120) cm.
- Mediana de pressão arterial sistólica de 129 (96-180) mmHg ou 129 (119-138) mmHg.*
- Mediana de pressão arterial diastólica foi de 86 (53-126) mmHg ou 86 (79-93) mmHg.*

*Para as variáveis pressóricas, 468 indivíduos foram avaliados.

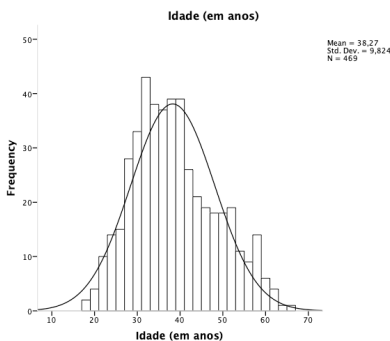


Figura 6

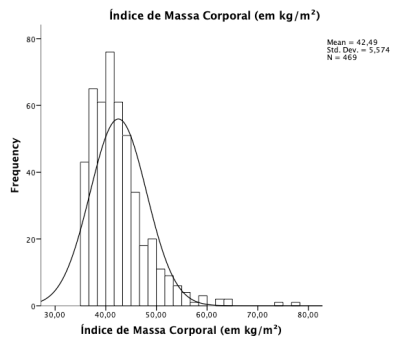


Figura 7

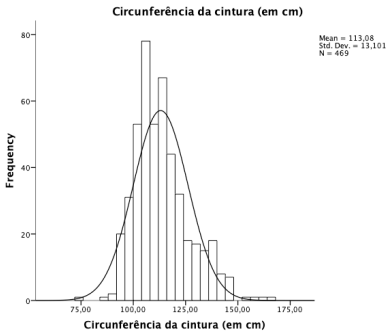


Figura 8

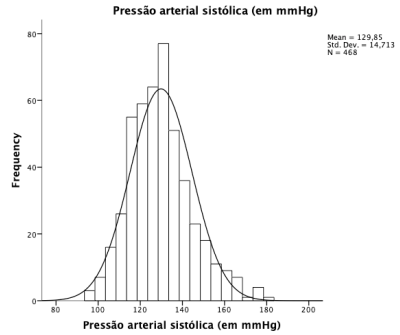


Figura 9

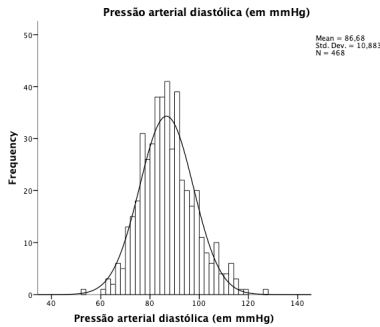


Figura 10

4. Dos 469 sujeitos com obesidade, 64,2% ($n = 301$) apresentavam obesidade grau III e, dos 468 sujeitos com obesidade avaliados para hipertensão, 83,8% ($n = 392$) eram hipertensos.

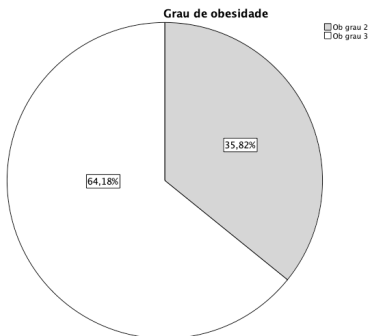


Figura 11

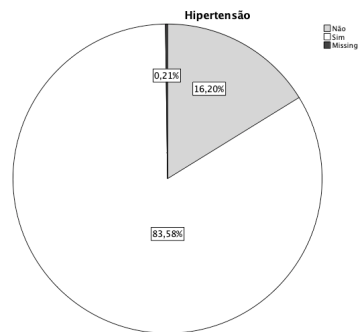


Figura 12

Capítulo 4

1.

a) Utilize a opção "Recodificar em variáveis diferentes", no menu "Transformar". Na janela ""Recodificar em variáveis diferentes", localize na tela da esquerda a variável "IMC" e clique na seta para mover para a tela "Variável numérica -> Variável de saída". Atribua um nome e informe uma descrição para a variável de saída. Clique em "Alterar". Em seguida, informe os valores e as categorias descritas no exercício em "Valores antigo e novo". Adicione todas as categorias, uma a uma, na tela "Antigo -> Novo". Clique em "Continuar" e, em seguida, em "Ok" ou em "Colar" para gerar o arquivo de sintaxe.

b) Utilize a opção "Calcular variável", no menu "Transformar". Na caixa de diálogo "Calcular variável", defina a "Variável de destino" (PAM) e informe uma descrição para a variável em "Tipo e rótulo". Localize na tela da esquerda as variáveis que correspondem a expressão informada no exercício. Utilize o teclado virtual, localizado no centro da janela, para construir a expressão numérica: " $((2 * PAD) + PAS) / 3$ ". Clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe.

c) Utilize a opção "Calcular variável", no menu "Transformar". Na caixa de diálogo "Calcular variável", defina a "Variável de destino" (LDL) e informe uma descrição para a variável em "Tipo e rótulo". Localize na tela da esquerda as variáveis que correspondem a Fórmula de Friedewald informada no exercício. Utilize o teclado virtual, localizado no centro da janela, para construir a expressão numérica: " $CT - HDL - (TG / 5)$ ". Clique em "Ok" ou em "Colar" para gerar o arquivo de sintaxe.

d) Utilize a opção "Selecionar casos..." no menu "Dados" para informar a condição correspondente à categoria de obesidade grau III.

Interpretação: Nos sujeitos com obesidade grau III, as médias pressóricas (PAS*, PAD* e PAM*) e de perfil lipídico (Colesterol total, Triglicerídeos, HDL-Colesterol e LDL-Colesterol) foram 131,0 ± 14,9 mmHg, 87,4 ± 10,9 mmHg, 101,9 ± 11,5 mmHg, 192,7 ± 34,8 mg/dL, 150,7 ± 86,7 mg/dL, 49,1 ± 13,3 mg/dL e 113,5 ±

30,7 mg/dL, respectivamente. *Avaliadas em 295 dos 296 sujeitos com obesidade grau III, ou seja, há um caso omissivo.

Tabela 1. Avaliação de dados pressóricos e de perfil lipídico de sujeitos com obesidade grau III. Os dados quantitativos estão apresentados em média \pm DP.

Variável	Sujeitos com obesidade grau III (n=296)
PAS (mmHg)*	131,0 \pm 14,9
PAD (mmHg)*	87,4 \pm 10,9
PAM (mmHg)*	101,9 \pm 11,5
Colesterol total (mg/dL)	192,7 \pm 34,8
Triglicerídeos (mg/dL)	150,7 \pm 86,7
HDL-Colesterol (mg/dL)	49,1 \pm 13,3
LDL-Colesterol (mg/dL)	113,5 \pm 30,7

* Para as variáveis pressóricas, 295 indivíduos foram avaliados. PAS, pressão arterial sistólica; PAD, pressão arterial diastólica; PAM, pressão arterial média; n, número absoluto.

e) Neste exercício, primeiro, opte por selecionar "Todos os casos" em "Selecionar casos..." no menu "Dados". Em seguida, no menu "Dados", em "Dividir arquivo...", selecione a opção "Comparar grupos" e insira a variável "Sexo" na tela "Grupos baseados em".

Interpretação: Nos sujeitos do sexo masculino com obesidade, as médias pressóricas (PAS*, PAD*, PAM*) e de perfil lipídico (Colesterol total, Triglicerídeos, HDL-Colesterol e LDL-Colesterol) foram 137,5 \pm 15,2 mmHg, 89,8 \pm 11,4 mmHg, 105,7 \pm 12,0 mmHg, 195,2 \pm 34,7 mg/dL, 181,4 \pm 107,0 mg/dL, 40,9 \pm 8,7 mg/dL e 118,0 \pm 31,5 mg/dL, respectivamente. *Avaliadas em 94 dos 95 sujeitos do sexo masculino, ou seja, há um caso omissivo. Nos sujeitos do sexo feminino com obesidade, as médias pressóricas (PAS*, PAD*, PAM*) e de perfil lipídico (Colesterol total, Triglicerídeos, HDL-Colesterol e LDL-Colesterol) foram 127,8 \pm 14,1 mmHg, 85,9 \pm 10,5 mmHg, 99,9 \pm 11,0 mmHg, 191,7 \pm 35,0 mg/dL, 136,2 \pm 69,0 mg/dL, 51,5 \pm 12,7 mg/dL e 113,0 \pm 30,4 mg/dL, respectivamente.

Tabela 2. Avaliação de dados pressóricos e de perfil lipídico em homens e mulheres com obesidade. Os dados quantitativos estão apresentados em média \pm DP.

Variável	Sexo masculino (n=95)	Sexo feminino (n=367)
PAS (mmHg)*	137,5 \pm 15,2	127,8 \pm 14,1
PAD (mmHg)*	89,8 \pm 11,4	85,9 \pm 10,5
PAM (mmHg)*	105,7 \pm 12,0	99,9 \pm 11,0
Colesterol total (mg/dL)	195,2 \pm 34,7	191,7 \pm 35,0
Triglicérides (mg/dL)	181,4 \pm 107,0	136,2 \pm 69,0
HDL-Colesterol (mg/dL)	40,9 \pm 8,7	51,5 \pm 12,7
LDL-Colesterol (mg/dL)	118,0 \pm 31,5	113,0 \pm 30,4

* Para as variáveis pressóricas, 94 indivíduos foram avaliados no sexo masculino.

PAS, pressão arterial sistólica; PAD, pressão arterial diastólica; PAM, pressão arterial média; n, número absoluto.

2.

Clique em "Mesclar arquivos", no menu "Dados", e selecione "Adicionar casos". Selecione o "Banco de dados 1 casos novos. sav" e adicione os novos casos ao banco de dados original.

3.

No menu "Dados", em "Dividir arquivo...", selecione a opção "Comparar grupos" e insira a variável "Nível de pós-graduação" na tela "Grupos baseados em".

a) A idade média dos alunos no nível mestrado foi de 28,9 anos e do nível doutorado foi de 32,4 anos*. *Avaliada em 56 dos 57 alunos, ou seja, há um caso omissos.

b) Dos alunos do nível mestrado, 88,1% (n = 59) são do sexo feminino, enquanto que do nível doutorado, 73,7% (n=42) são do sexo feminino.

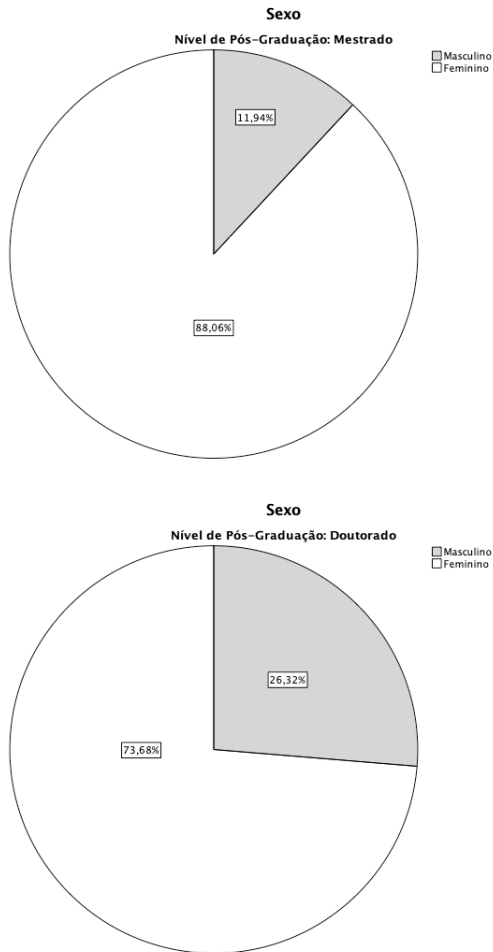


Figura 1

c) Em ambos os níveis de pós-graduação, a maior parte dos alunos se encontram no primeiro ano de estudos. Correspondem a 68,7% (n = 46) dos alunos do mestrado e 35,1% (n = 20) dos alunos do doutorado.

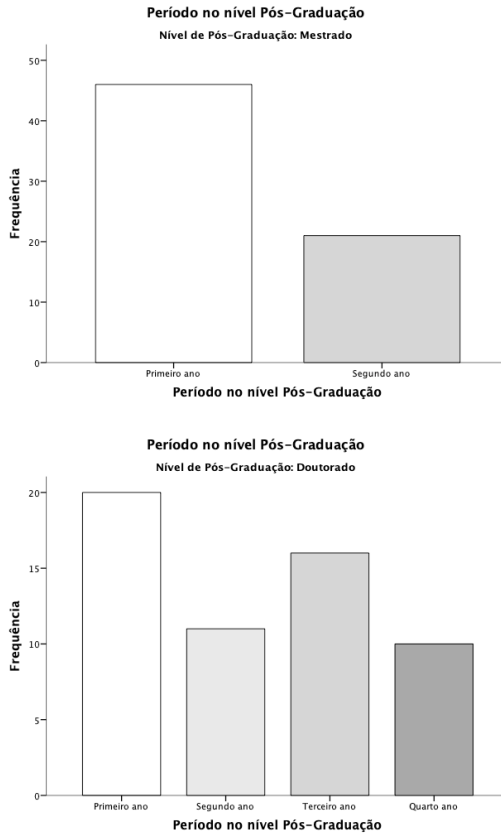


Figura 2

Capítulo 5

1. Resolução não aplicável.

Capítulo 6

1.

a) Análise de dados categóricos para uma variável qualitativa dicotômica com mais de 25 casos. Utiliza-se teste de Qui-quadrado de proporção.

Interpretação: Os resultados indicam que há diferença significativa na proporção das categorias da variável sexo ($p < 0,001$) entre os alunos. Dos 124 alunos, houve uma maior frequência do sexo feminino (81,5%; $p < 0,001$).

b) Análise de dados categóricos para uma variável qualitativa dicotômica com mais de 25 casos. Utiliza-se teste de Qui-quadrado de proporção.

Interpretação: Os resultados indicam que não há diferença entre os níveis de pós-graduação ($p = 0,418$) entre os alunos. Dos 124 alunos, não houve diferença na proporção de alunos no nível mestrado e doutorado (54,0% e 46,0%, respectivamente; $p = 0,418$).

c) Análise de dados categóricos para duas variáveis. Utiliza-se teste de Qui-quadrado de associação (ou independência). Para a escolha do p-valor do teste do Qui-quadrado, considere que temos duas células com contagem menor que 5 elementos (25,0%), por isso utilizamos o teste Exato de Fisher.

Interpretação: Não houve diferença no nível de conhecimento em estatística entre os alunos do nível mestrado e do nível doutorado ($p = 0,487$). Dos 67 (54,0%) alunos do nível mestrado, 47,8% relataram nível de conhecimento baixo e 38,8% relataram nível de conhecimento médio em estatística. No nível doutorado ($n = 57$, 46,0%), 40,4% relataram nível de conhecimento baixo e 43,9% relataram nível de conhecimento médio em estatística.

d) Análise de dados categóricos para duas variáveis. Utiliza-se teste de Qui-quadrado de associação (ou independência). Para a escolha do p-valor do teste do Qui-quadrado, considere que

temos duas células com contagem menor que 5 elementos (20%) e uma tabela de contingência diferente de 2x2, por isso utilizamos o teste de Qui-quadrado de Pearson.

Interpretação: Não houve diferença no nível de conhecimento em epidemiologia entre os alunos do nível mestrado e do nível doutorado ($p = 0,199$). Dos 67 (54,0%) alunos do nível mestrado, 35,8% relataram nível de conhecimento baixo e 46,3% relataram nível de conhecimento médio em epidemiologia. No nível doutorado ($n = 57$, 46,0%), 42,1% relataram nível de conhecimento baixo e 29,8% relataram nível de conhecimento médio em estatística.

e) Análise de dados categóricos para duas variáveis. Utiliza-se teste de Qui-quadrado de associação (ou independência). Para a escolha do p-valor do teste do Qui-quadrado, considere que temos nenhuma célula com menos de 5 elementos e uma tabela de contingência 2x2, por isso utilizamos o teste de correção de continuidade de Yates.

Interpretação: Houve diferença significativa na experiência com o SPSS entre os alunos do nível mestrado e do nível doutorado ($p = 0,018$). Dos 67 (54,0%) alunos do nível mestrado, 74,6% relataram não ter experiência com o SPSS, enquanto que no nível doutorado 52,6% relataram não ter experiência. Olhando os resíduos ajustados, observamos que há indicação de associação entre nível mestrado e não ter experiência no SPSS, bem como há associação entre nível de doutorado e ter experiência no SPSS.

2.

a) Análise de dados categóricos para duas variáveis. Utiliza-se o teste de Qui-quadrado de associação (ou independência). Todas as tabelas de contingência são no formato 2x2 e apresentam nenhuma célula com menos de 5 elementos, por isso utilizamos o p-valor do teste de correção de continuidade de Yates.

Interpretação: Não houve diferença entre os tipos de cirurgia na hipertensão avaliada no período pré-operatório (tempo 1) ($p = 0,908$). No tempo 1, haviam 80 (76,2%) hipertensos no grupo bypass gástrico (RYGB), enquanto que no grupo gastrectomia vertical (SG)

havia 67 (74,4%) hipertensos. Não houve diferença entre os tipos de cirurgia na hipertensão avaliada no período pós-operatório (tempo 2) ($p = 0,512$). No tempo 2, havia 35 (33,3%) hipertensos no grupo RYGB, enquanto que no grupo SG havia 35 (38,9%) hipertensos. Houve diferença entre as técnicas cirúrgicas na hipertensão avaliada no período pós-operatório (tempo 3) ($p = 0,033$). No tempo 3, havia 25,7% ($n = 27$) de hipertensos no grupo RYGB, enquanto que no grupo SG havia 41,1% ($n = 37$) de hipertensos. O tipo RYGB parece contribuir mais para a melhora da hipertensão a longo prazo. Olhando os resíduos ajustados, observamos que há indicação de associação entre RYGB e não ter HAS no tempo 3, bem como há associação entre SG e ter HAS no tempo 3.

Tabela 1. Comparação da hipertensão avaliada nos períodos pré (1) e pós-operatórios (2 e 3) em sujeitos obesos submetidos à cirurgia bariátrica dos tipos bypass gástrico e gastrectomia vertical. Os números são apresentados como números absolutos e porcentagens para variáveis qualitativas. As unidades de medida e número de pacientes são apresentados com as variáveis correspondentes.

	RYGB (n = 105)	SG (n = 90)	P-valor (a)
HAS 1 (n, %)	80 (76,2)	67 (74,4)	0,908
HAS 2 (n, %)	35 (33,3)	35 (38,9)	0,512
HAS 3 (n, %)	27 (25,7)	37 (41,1)	0,033*

RYGB, bypass gástrico; SG, gastrectomia vertical; HAS, hipertensão arterial; n = número absoluto.

*Significância estatística considerada de $p \leq 0.050$; (a) Teste de correção de continuidade de Yates.

b) Neste exercício, considere que temos uma única variável de desfecho (hipertensão) avaliada ao longo do tempo (períodos pré- e pós-operatórios). O primeiro passo é dividir o banco de dados pelas técnicas cirúrgicas. Note que podemos realizar uma análise global ("HAS 1" x "HAS 2" x "HAS 3") e uma análise *post hoc*, de comparação par a par ("HAS 1" x "HAS 2", "HAS 1" x "HAS 3" e "HAS 2" x "HAS 3"). Uma das formas de fazer essa comparação par a par é através do teste de McNemar. O teste de McNemar para duas amostras pode ser estendido para situações que envolvam mais de duas amostras. Esta extensão é

denominada de teste Q de Cochran para k amostras relacionadas. É idêntico ao teste de Friedman, mas é aplicável quando todas as respostas são dicotômicas, como neste caso.

Interpretação: No bypass gástrico (RYGB), dos 76,2% (n = 80) sujeitos que apresentavam hipertensão no período pré-operatório (tempo 1), 63,7% (n = 51) resolveram no pós-operatório (tempo 2). Na gastrectomia vertical (SG), dos 74,4% (n = 67) que apresentavam hipertensão no tempo 1, 53,7% (n = 36) resolveram no tempo 2. A hipertensão diminuiu significativamente em ambas as técnicas cirúrgicas ($p < 0,001$). No RYGB, dos 76,2% (n = 80) que apresentavam hipertensão no tempo 1, 72,5% (n = 58) resolveram no pós-operatório (tempo 3). No SG, dos 74,4% (n = 67) que apresentavam hipertensão no tempo 1, 49,3% (n = 33) resolveram no tempo 3. A hipertensão diminuiu significativamente em ambas as técnicas cirúrgicas ($p < 0,001$). No RYGB, dos 33,3% (n=35) que apresentavam hipertensão no tempo 2, 40,0% (n = 14) resolveram no tempo 3. No SG, dos 38,9% (n = 35) que apresentavam hipertensão no tempo 2, 25,7% (n = 9) resolveram no tempo 3. Comparando os tempos 2 e 3, a hipertensão não mostrou-se significativamente diferente no RYGB ($p = 0,115$) e no SG ($p = 0,824$).

Tabela 2. Evolução da hipertensão arterial em sujeitos obesos submetidos à cirurgia bariátrica do tipo bypass gástrico. Os números são apresentados como números absolutos e porcentagens para variáveis qualitativas. As unidades de medida e número de pacientes são apresentados com as variáveis correspondentes.

	Grupo RYGB (n = 105)		
	Tempo 1	Tempo 2	Tempo 3
HAS (n, %) [‡]	80 (76,2) ^a	35 (33,3) ^b	27 (25,7) ^b

Nas linhas, os valores com diferentes letras sobrescritas (^{a,b}) indicam diferenças significativas ($p \leq 0,050$) entre os tempos avaliados com base no teste de McNemar ([‡]).

RYGB, bypass gástrico; HAS, hipertensão arterial; n = número absoluto.

Tabela 3. Evolução da hipertensão arterial em sujeitos obesos submetidos à cirurgia bariátrica do tipo gastrectomia vertical. Os números são apresentados como números absolutos e porcentagens para variáveis qualitativas. As unidades de medida e número de pacientes são apresentados com as variáveis correspondentes.

	Grupo SG (n = 90)		
	Tempo 1	Tempo 2	Tempo 3
HAS (n, %) [‡]	67 (74,4) ^a	35 (38,9) ^b	37 (41,1) ^b

Nas linhas, os valores com diferentes letras sobrescritas (^{a,b}) indicam diferenças significativas ($p \leq 0,050$) entre os tempos avaliados com base no teste de McNemar ([‡]).

SG, gastrectomia vertical; HAS, hipertensão arterial; n = número absoluto.

Capítulo 7

1.

Para a resolução do exercício, primeiro realize o teste de normalidade de Shapiro-Wilk. Em seguida, baseado no teste, opte pelo teste paramétrico ou não paramétrico. Como a significância do teste foi inferior ou igual a 0,050 ($p \leq 0,050$), a variável segue uma distribuição não normal. Como trata-se de dois grupos independentes, o teste não paramétrico indicado é o teste U de Mann-Whitney. O teste de hipótese indica que a hipótese de nulidade deve ser rejeitada, ou seja, a distribuição da idade entre as categorias do nível de pós-graduação não é igual. Os resultados devem ser expressos em mediana e percentis ou valores mínimo e máximo.

Interpretação: A mediana de idade dos alunos de mestrado (28 anos) é significativamente diferente da mediana de idade dos alunos de doutorado (31 anos, $p = 0,001$).

2. Para a resolução do exercício, primeiro realize o teste de normalidade de Shapiro-Wilk. Em seguida, baseado no teste, opte pelo teste paramétrico ou não paramétrico. Como a significância do teste foi inferior ou igual a 0,050 ($p \leq 0,050$), as variáveis seguem uma distribuição não normal. Como trata-se de

dois grupos independentes, o teste não paramétrico indicado é o teste U de Mann-Whitney. Os resultados devem ser expressos em mediana e percentis ou valores mínimo e máximo.

Interpretação: Comparando os sexos, houve diferença significativa entre as medidas de pressão arterial sistólica ($p < 0,001$), diastólica ($p = 0,003$) e média ($p < 0,001$), e nos níveis séricos de triglicerídeos ($p < 0,001$) e HDL-Colesterol ($p < 0,001$). Para as variáveis pressóricas, 94 sujeitos do sexo masculino foram avaliados.

Tabela 1. Comparação das medidas pressóricas e lipídicas entre os sexos. Os dados são representados por mediana e percentis (25 - 75). As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Sexo masculino (n = 95)	Sexo feminino (n = 367)	P-valor ^a
PAS (mmHg)**	134,5 (126,0-148,0)	127,0 (118,0-136,0)	< 0,001*
PAD (mmHg)**	89,0 (82,0-96,0)	85,0 (79,0-92,0)	0,003*
PAM (mmHg)**	104,0 (97,7-112,3)	99,3 (92,7-105,5)	< 0,001*
Colesterol total (mg/dL)	196,0 (174,0-213,0)	188,0 (170,0-211,0)	0,220
Triglicerídeos (mg/dL)	147,0 (115,0-219,5)	125,0 (89,0-165,5)	<0,001*
HDL-Colesterol (mg/dL)	41,0 (36,0-45,0)	49,0 (43,0-58,0)	<0,001*
LDL-Colesterol (mg/dL)	117,6 (96,2-139,2)	110,0 (91,9-130,4)	0,059

PAS, pressão arterial sistólica; PAD, pressão arterial diastólica; PAM, pressão arterial média; n = número absoluto.

^aTeste U de Mann-Whitney.

*Significância estatística considerada de $p \leq 0,050$.

** Foram avaliados 94 sujeitos do sexo masculino para as variáveis pressóricas.

3. Para a resolução do exercício, primeiro realize o teste de normalidade de Shapiro-Wilk. Em seguida, baseado no teste, opte pelo teste paramétrico ou não paramétrico. Como a significância do teste foi inferior ou igual a 0,050 ($p \leq 0,050$), as variáveis de pressão arterial sistólica e os níveis séricos de triglicerídeos e

HDL-Colesterol seguem uma distribuição não normal, enquanto as variáveis de pressão arterial diastólica e de colesterol total seguem uma distribuição normal ($p > 0,050$). No entanto, temos menos de 15 sujeitos no grupo do sexo masculino ($n = 9$). Neste caso, deve-se proceder com análise não paramétrica. Como trata-se de dois grupos independentes, o teste não paramétrico indicado é o teste U de Mann-Whitney. Os resultados devem ser expressos em mediana e percentis ou valores mínimo e máximo.

Interpretação: Comparando os sexos, houve diferença significativa apenas entre os níveis séricos de HDL-Colesterol ($p = 0,009$).

Tabela 2. Comparação das medidas pressóricas e lipídicas entre os sexos. Os dados são representados por mediana e percentis (25 - 75). As unidades de medida e número de pacientes são apresentadas com as variáveis correspondentes.

	Sexo masculino (n=9)	Sexo feminino (n=21)	P-valor ^a
PAS (mmHg)	134,0 (125,0-142,0)	130,0 (121,0-141,0)	0,372
PAD (mmHg)	81,0 (80,0-92,0)	88,0 (83,0-93,0)	0,304
Colesterol total (mg/dL)	198,0 (149,0-284,0)	202,0 (169,0-210,0)	0,965
Triglicerídeos (mg/dL)	116,0 (105,0-289,0)	129,0 (78,0-150,0)	0,449
HDL-Colesterol (mg/dL)	35,0 (34,0-42,0)	47,0 (44,0-52,0)	0,009*

PAS, pressão arterial sistólica; PAD, pressão arterial diastólica; n = número absoluto.

^aTeste U de Mann-Whitney.

*Significância estatística considerada de $p \leq 0,050$.

Capítulo 8

1.

Interpretação: Comparando as medidas de circunferência da cintura, houve diferença significativa entre todos os tempos avaliados ($p < 0,001$). Nas medidas pressóricas sistólica e diastólica, não houve diferença significativa entre os tempos 2 e 3 ($p = 0,426$ e $p = 100$, respectivamente).

Tabela 1: Comparação das medidas de circunferência da cintura e pressóricas em sujeitos com obesidade avaliados antes e após a cirurgia bariátrica. Os dados são representados por mediana e percentis (25-75). As unidades de medida e número de sujeitos são apresentadas com as variáveis correspondentes.

	Tempo 1 (n = 195)	Tempo 2 (n = 195)	Tempo 3 (n = 195)	P-valor
Circunferência da cintura (cm)	110,0 (104,0-118,0) ^a	89,0 (83,0-97,0) ^b	84,0 (77,0-90,0) ^c	<0,001*
Pressão arterial sistólica (mmHg)	128 (119,0-137,0) ^a	118,0 (110,0-127,0) ^b	118,0 (110,0-129,0) ^c	<0,001*
Pressão arterial diastólica (mmHg)	86,0 (79,0-93,0) ^a	75,0 (69,0-82,0) ^b	73,0 (67,0-81,0) ^c	<0,001*

Nas linhas, os valores com diferentes letras sobrescritas (^{a,b,c}) indicam diferenças significativas ($p \leq 0,050$) entre os tempos avaliados com base no teste de post hoc de Dunn-Bonferroni. n, número absoluto.

*Teste de Friedman. Significância estatística considerada de $p \leq 0,05$.

Capítulo 9

1.

Primeiro, é necessário realizar o teste de normalidade de Shapiro-Wilk, onde verifica-se distribuição não normal para as variáveis. Portanto, será aplicado uma correlação de Spearman.

a) Interpretação: Existe uma correlação positiva moderada entre massa corporal e glicemia de jejum, com valor de $\rho = 0,507$ e p-valor < 0,001.

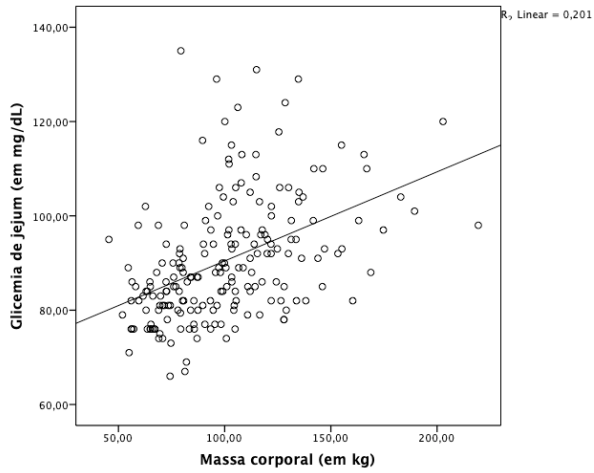


Figura 1. Correlação positiva (direta) moderada entre massa corporal e glicemia de jejum.

b) Interpretação: Existe uma correlação negativa (inversa) fraca entre massa corporal e HDL-Colesterol, com valor de $\rho = -0,266$ e p-valor $< 0,001$.

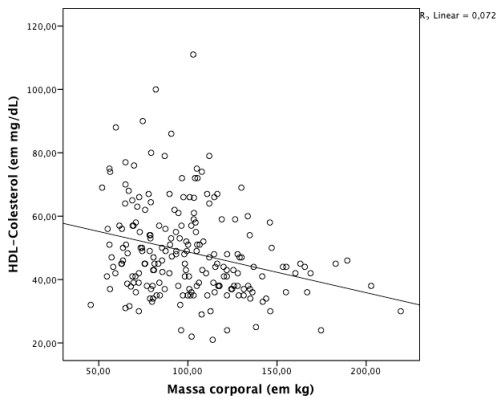


Figura 2. Correlação negativa (inversa) fraca entre massa corporal e HDL-Colesterol.

2.

Primeiro, é necessário realizar o teste de normalidade de Shapiro-Wilk, onde verifica-se distribuição não normal para as variáveis. Portanto, será aplicado uma correlação de Spearman.

a) Interpretação: Existe uma correlação positiva moderada entre IMC e glicemia de jejum, com valor de $\rho = 0,549$ e p-valor $< 0,001$.

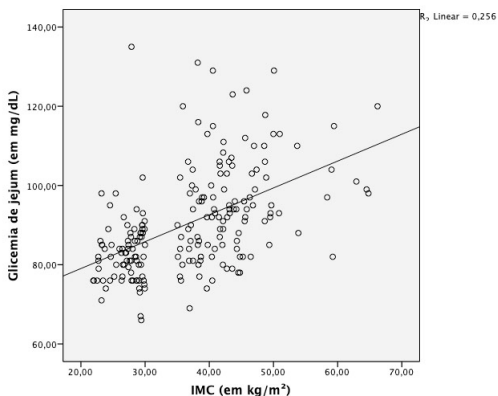


Figura 3. Correlação positiva (direta) moderada entre IMC e glicemia de jejum.

b) Interpretação: Existe uma correlação negativa (inversa) fraca entre IMC e HDL-Colesterol, com valor de $\rho = -0,214$ e p-valor $< 0,001$.

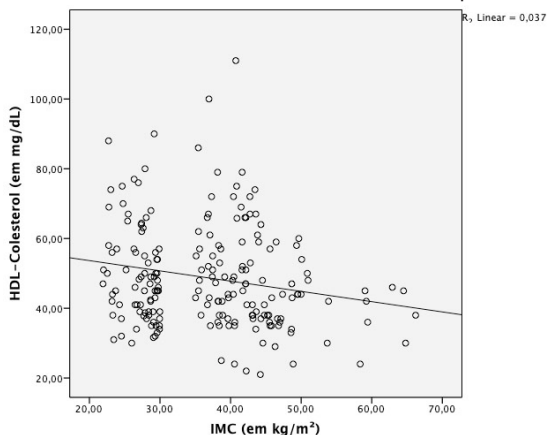


Figura 4. Correlação negativa (inversa) fraca entre IMC e HDL-Colesterol.

3.

Primeiro, é necessário recodificar as variáveis “Obesidade”, “Hipertensão arterial” e “Diabetes Mellitus” para obter-se a correta estrutura de organização da tabela de contingência 2x2.

a) Interpretação: A análise do teste de correção de continuidade de Yates indica associação significativa entre obesidade e hipertensão ($p < 0,001$). Os resíduos evidenciam essa indicação de associação. Ser obeso aumenta a chance de ter hipertensão em 3,82 vezes (IC95%: 2,10-6,93) em comparação a não ser obeso.

b) Interpretação: A análise do teste de correção de continuidade de Yates indica associação significativa entre obesidade e diabetes ($p = 0,001$). Os resíduos evidenciam essa indicação de associação. Ser obeso aumenta a chance de ter diabetes em 3,87 vezes (IC95%: 1,81-8,29) em comparação a não ser obeso.

Capítulo 10

1.

Na análise, foram incluídos 468 casos, havendo um valor omissão. Há uma correlação positiva fraca entre a pressão arterial sistólica e a massa corporal ($r = 0,244$, $p < 0,001$).

Interpretação: O teste F da ANOVA é significativo, portanto, podemos assumir que o modelo explica uma quantidade significativa da variação da pressão arterial sistólica em função da variabilidade da massa corporal. Com p-valor menor que 0,001, o teste global mostra que há regressão entre as variáveis. A variável massa corporal possui um $\beta = 0,172$, sendo este significativo ($p < 0,001$). Assim, para cada unidade a mais na massa corporal, a pressão arterial sistólica aumenta em 0,172 mmHg. Com uma correlação entre as variáveis de 0,244, a proporção da variabilidade de pressão arterial sistólica que pode ser explicada pela variabilidade da massa corporal é de 6,0%.

A regressão linear simples mostrou que a massa corporal prevê a pressão arterial sistólica [$F(1,466) = 29,581$, $p < 0,001$, $R^2 = 0,060$]. A pressão arterial sistólica, em mmHg, corresponde a $110,013 + 0,172 \times$ (massa corporal), sendo a idade medida em anos.

Para finalizar a regressão, vamos realizar a análise de resíduos para verificar se o modelo é válido. O primeiro diagnóstico é a verificação do gráfico de resíduos *versus* valores preditos. Conforme o gráfico (Figura 1), é visível que os pontos apresentam um comportamento cônico e, assim, temos indícios de que a variância dos resíduos é heteroscedástica. Para a multicolineariedade, o índice de Tolerância e o VIF estão ambos quantificados em 1, ou seja, não há problemas de multicolineariedade. Para autocorrelação, a estatística D de Durbin-Watson foi de 1,983. Assim, para $n > 100$ e X igual a 1, temos D maior que D-U, ou seja, podemos afirmar que não há autocorrelação. Temos três valores atípicos (casos 22, 366 e 386). O indicado é repetir o modelo de regressão sem estes casos e verificar os resultados da análise de regressão. O teste de Shapiro-Wilk indica que os resíduos são não normais. Assim, considerando que os pressupostos de normalidade dos resíduos e de homoscedasticidade não foram cumpridas, o teste de regressão não é válido.

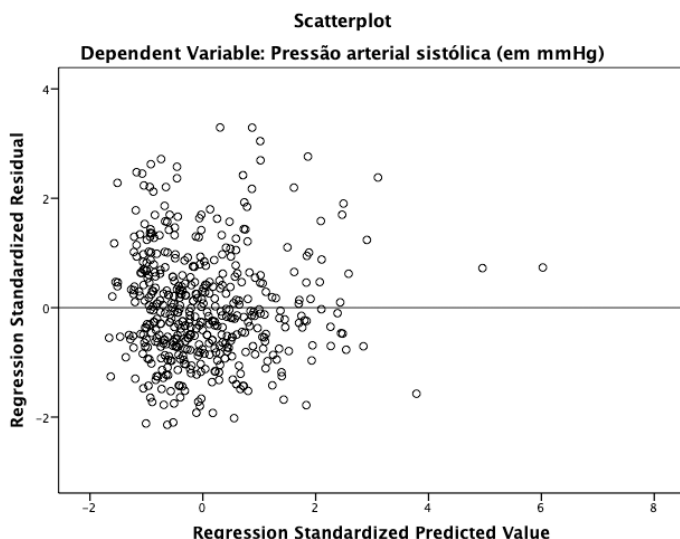


Figura 1. Gráfico de *scatterplot* dos resíduos *versus* valores preditos.

2.

Na análise, foram incluídos 468 casos, havendo um valor omissão. Há uma correlação positiva fraca entre a pressão arterial sistólica e a massa corporal ($r = 0,252$, $p < 0,001$) e negativa fraca entre a pressão arterial sistólica e o sexo ($r = -0,264$, $p < 0,001$).

Interpretação: O teste F da ANOVA é significativo, portanto, podemos assumir que o modelo explica uma quantidade significativa da variação da pressão arterial sistólica em função da variabilidade da massa corporal e do sexo. Com p-valor menor que 0,001, o teste global mostra que há regressão entre as variáveis. Conforme a equação, 121,525 é a pressão arterial sistólica constante, independentemente da circunferência da cintura e do sexo. A variável circunferência da cintura possui um $\beta = 0,182$, sendo este significativo ($p = 0,002$) e o sexo possui um $\beta = -6,776$, sendo este também significativo ($p < 0,001$). Assim, para cada unidade a mais na circunferência da cintura, a pressão arterial sistólica aumenta em 0,182 mmHg e a diferença média entre os sexos é de -6,776, tendo como referência o sexo masculino. Com uma correlação entre as variáveis de 0,299, a proporção da variabilidade de pressão arterial sistólica que pode ser explicada pela circunferência da cintura e pelo sexo é de 8,5%.

Foi utilizada a regressão linear múltipla para verificar se circunferência da cintura e sexo são capazes de prever a pressão arterial sistólica. A análise resultou em um modelo estatisticamente significativo [$F(2,465) = 22,757$; $p < 0,001$; $R^2_{\text{adj}} = 0,085$]. A circunferência da cintura ($\beta = 0,182$; $t = 3,166$; $p = 0,002$) e o sexo ($\beta = -6,776$; $t = -3,613$; $p < 0,001$) são preditores da pressão arterial sistólica.

Para finalizar a regressão, vamos realizar a análise de resíduos para verificar se o modelo é válido. O primeiro diagnóstico é a verificação do gráfico de resíduos *versus* valores preditos. Conforme o gráfico (Figura 2), é visível que os pontos não apresentam um comportamento cônico e, assim, temos indícios de que a variância dos resíduos é homoscedástica. Para a multicolineariedade, o índice de Tolerância apresenta valor de 0,756 e o VIF de 1,324, ou seja, não há problemas de multicolinearidade. Para autocorrelação, a estatística D de Durbin-Watson foi de 1,974. Assim, para $n > 100$ e X igual a 1, temos D maior que D-U, ou seja, podemos afirmar que não há autocorrelação. Temos um valor atípico (caso 386). O indicado é repetir o modelo de regressão sem este caso e verificar os resultados da análise de regressão. O teste de Shapiro-Wilk indica que os resíduos são não normais. Assim, considerando que o pressuposto de normalidade dos resíduos não foi cumprido, o teste de regressão não é válido.

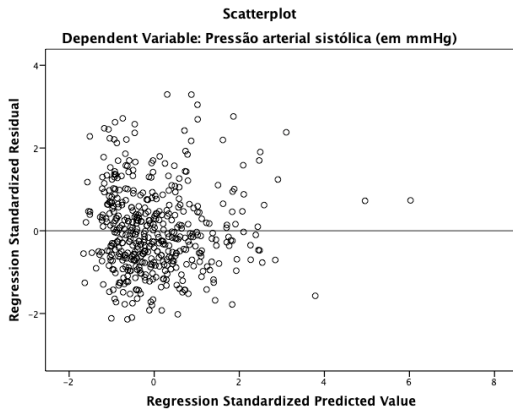


Figura 2. Gráfico de scatterplot dos resíduos versus valores preditos.

Capítulo 11

1.

Primeiro, em uma análise univariada, o teste de correção de continuidade de Yates mostrou uma indicação de associação de polineuropatia periférica com a técnica do tipo RYGB (70,9%, $p = 0,002$). Em seguida, realiza-se a regressão logística binária univariável.

Interpretação: Conclui-se que o tipo de cirurgia é um preditor significativo para a ocorrência de PNP ($p = 0,001$). A probabilidade de ter PNP é 3,0 (IC95%: 1,7-6,0) vezes maior nos sujeitos submetidos ao RYGB em comparação ao SG.

2.

Primeiro, em uma análise univariada, o teste de correção de continuidade de Yates mostrou uma indicação de associação de hipertensão com o sexo masculino (38,5%, $p = 0,002$). Em seguida, realiza-se a regressão de Poisson.

Interpretação: Conclui-se que, com uma razão de prevalências de 2,372, a chance de ter hipertensão nos homens é 2,4 (IC95%: 1,4-3,9; $p = 0,006$) vezes maior em relação às mulheres.

ISBN: 978-65-86232-44-8

TST



9 786586 232448



Faculdade de
Medicina
Universidade Federal
do Rio Grande do Sul

120
anos

PPG
GG

PPG em Ciências da Saúde:
Ginecologia e Obstetria

HOSPITAL DE
CLÍNICAS
PORTO ALEGRE RS



fundmed
Pesquisa, Ensino e Inovação