

Conceitos e análises estatísticas com R e JASP

Luis Anunciação (PUC-Rio), PhD

Contents

1	Prefácio	5
1.1	Atenção	7
1.2	A proposta	7
1.3	Objetivo	8
1.4	Público-alvo	8
1.5	Formato do livro	9
1.6	Como usar este livro	9
1.7	Pesquisas e dados	10
1.8	O R e os pacotes	11
1.9	JASP	11
1.10	Outros recursos	12
1.11	Capa	12
1.12	Versão do livro	13
1.13	Autor	13
1.14	Revisão técnica	13
1.15	Agradecimentos	14

Chapter 1

Prefácio

CONCEITOS E ANÁLISES ESTATÍSTICAS COM R E JASP



1.1 Atenção

Atenção: Leia com cuidado. Este livro ainda está em sua fase de revisão.
Última modificação: 04 March, 2021 às 08:55

1.2 A proposta

Este livro nasceu como um dos principais e mais frutíferos resultados das aulas de graduação e pós-graduação ministradas por mim em alguns locais, mas com maior intensidade na PUC-Rio, UFRJ e IBNeuro. Por bastante tempo, nas aulas de estatística aplicada à Psicologia e Bioestatística, eu recorri a diferentes livros que, cada qual a sua maneira, apresentavam conceitos de pesquisa, técnicas estatísticas e análise de dados.

No entanto, acabei percebendo (ou tendo a impressão) de que eles apresentam a estatística por diferentes atalhos pedagógicos, (1) sugerindo que pesquisa e estatística eram áreas distantes, (2) que toda estatística podia ser resumida por testes de hipóteses independentes entre si e que (3) situações envolvendo dados reais não tinham tanto interesse. No geral, parece-me que para eles apresentarem a estatística na ciência, era necessário se distorcer pesadamente a ciência da estatística.

Em função disso, nos últimos anos, eu fui sentindo necessidade de apresentar os conceitos de pesquisa e técnicas estatísticas de forma integrada, contanto com dados reais e seguindo por uma metodologia de aula que pudesse ser pragmática, mas sem reforçar vícios inadequados sobre conceitos de estatística.

Conciliar essas condições em um único livro de maneira adequada é bastante improvável. Dessa forma, esse livro opta por uma abordagem majoritariamente pragmática, mas que evita se distanciar de conceitos teóricos. O pragmatismo é fundamental para que o estudante consiga, rapidamente, entender os procedimentos relacionados à análise de dados e implementar técnicas estatísticas para tomar decisões. Quão antes o estudante entender a utilidade da estatística para resolver problemas, maior é a probabilidade dele vir a gostar da área. Por sua vez, os aspectos teóricos são os alicerces para que o estudante possa perceber também que a utilidade que a estatística tem na ciência só é possível por ela ser uma disciplina sólida e robusta e que veio se aprimorando nas últimas décadas.

Isso posto, este livro é fruto de um grande esforço que tem a proposta de ser um manual técnico, em que são apresentados conceitos de pesquisa e análises estatísticas realizadas no R e no JASP e com especial aplicação em Psicologia e Bioestatística. Em cada capítulo, o estudante terá a oportunidade de acessar:

1. Uma pesquisa científica, explicitando o problema e as hipóteses que a guiaram

2. O artigo publicado com os resultados
3. A base de dados em formato R ou CSV para reprodução das análises
4. O conjunto de procedimentos estatísticos utilizados
5. Recursos extras para aprofundamento em tópicos específicos
6. Exercícios que auxiliem no entendimento do conteúdo, quase sempre retirados de provas externas

Com isso, o livro oferece ao estudante um ambiente em que ele possa resolver um problema real, utilizando as técnicas e métodos estatísticos como ferramentas para tomada de decisão. Apesar do foco ser mais no problema de pesquisa do que nas ferramentas analíticas, em todos os capítulos, a aplicação da estatística na ciência será reforçada pela apresentação de alguns conceitos da ciência da estatística.

Espero que este livro possa ser útil a estudantes de graduação e pós-graduação, agradável a leitores de estatística como de Psicologia e um recurso importante para outros docentes que, eventualmente, precisem de um material de apoio.

1.3 Objetivo

O livro tem como objetivos (1) apresentar, (2) discutir e (3) operacionalizar conceitos de pesquisa e análises estatística de dados a partir de pesquisas publicadas e dados reais. Espera-se que qualquer o estudante consiga realizar todas as ações descritas no decorrer dos capítulos de maneira guiada e intuitiva. As sintaxes utilizadas no ambiente R e as telas de execução do JASP são integralmente disponíveis.

1.4 Público-alvo

Este livro foi desenvolvido de maneira mais focada a estudantes de Psicologia e Bioestatística. As pesquisas e exemplos utilizados são mais aderentes a essas duas áreas. No entanto, como parte dos conceitos e análises implementadas no livro são interdisciplinares, espera-se que estudantes de áreas como educação, administração e economia também possam também ter proveito do livro.

1.5 Formato do livro

O livro foi pensado para ter uma estrutura (1) organizada, (2) linear e (3) formada por capítulos autossuficientes escritos para responder questões específicas. Acredito que, assim, ele possa atender tanto estudantes interessados em ler a obra inteira, como aqueles que buscam informações mais específicas sobre um tópico particular.

Esse formato adotado tende a gerar uma percepção diferente entre aqueles que consultarem apenas um capítulo ou outro e aqueles que lerem o conteúdo por completo. Há uma maior chance disso ocorrer em capítulos sobre testes estatísticos. Uma vez que diversos testes estatísticos são casos particulares de outros, alguns assuntos que parecem destoantes em uma leitura inicial, tornam-se articulados em outros capítulos.

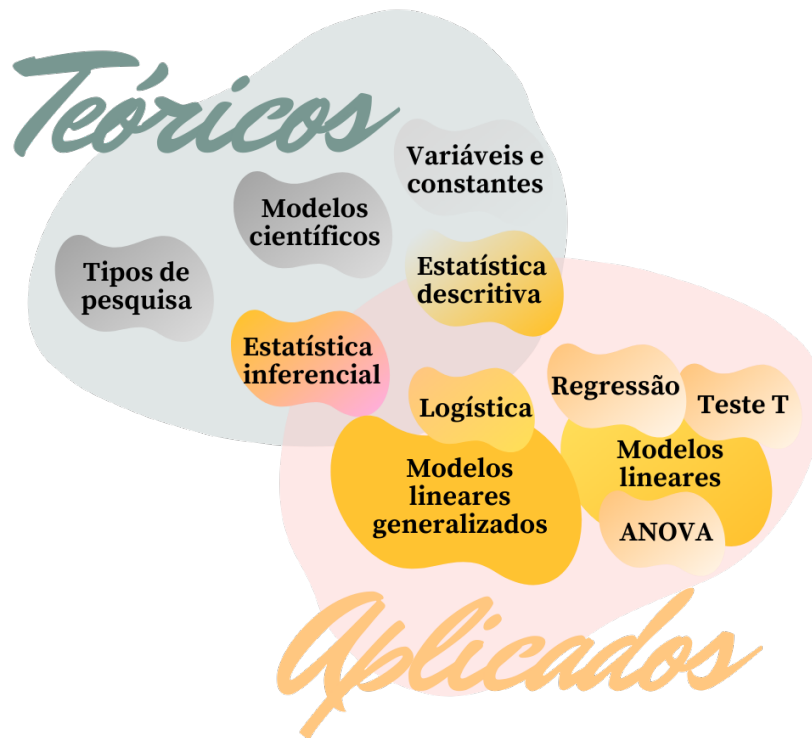
Muitos capítulos recebem o nome de testes de hipóteses (ex: Teste T ou Regressão). Isso foi intencional e visa auxiliar estudantes que precisem apenas de informações pontuais, bem como tende a enfraquecer a ideia de uma relação ponto a ponto tipicamente feita entre testes estatísticos e delineamentos específicos.

1.6 Como usar este livro

O livro é formado por dois componentes: capítulos teóricos e capítulos voltados à análise de dados. Os capítulos teóricos reúnem alguns conceitos fundamentais de pesquisa e estatística, tais como tipos de variáveis, delineamento de pesquisa e técnicas de amostragem. Estes capítulos foram escritos pensando em alunos de graduação do curso de Psicologia. Tenho a impressão que esses capítulos serão pouco acessados, apesar de importantes.

Os capítulos analíticos são focados em testes de hipóteses e contam com uma metodologia direta ao ponto, em que atividades similares às realizadas nos artigos são demonstradas. Estes capítulos foram desenvolvidos para estudantes de pós-graduação. Acredito que esses capítulos serão bastante acessados.

A figura abaixo diagrama os dois componentes de forma aproximada.



1.7 Pesquisas e dados

Neste livro, as seguintes pesquisas e seus materiais são utilizados:

- “Depression and Anxiety Symptoms in a Representative Sample of Undergraduate Students in Spain, Portugal, and Brazil”
- “Confirmatory analysis and normative tables for the Brazilian Ages and Stages Questionnaires: Social-Emotional”
- Psychometric properties of a short-term visual memory test (MEMORE)"
- “A relação entre o nível de Empreendedorismo (TEG) e os aspectos sociodemográficos dos Taxistas cooperados da cidade de Santo André/São Paulo, Brasil”

- “Avaliação psicométrica em português do indicador de dor crônica de Helsinki em cães com sinais crônicos de osteoartrite”
- “Aspects Related to Body Image and Eating Behaviors in Healthy Brazilian Undergraduate Students”
- “Parent-reported diagnosis of Attention Deficit Hyperactivity Disorder and psychostimulant use among children and adolescents: a population-based nationwide study”
- “Resilience and vulnerability in adolescents with primary headaches: a cross-sectional population-based study”

As bases são *Open Science*. Isso significa que elas são gratuitas e universalmente acessíveis para finalidades acadêmicas. Em cada capítulo, as bases irão aparecer na seção “Pesquisa”, da seguinte maneira:

A base desta pesquisa está disponível em formato **R (Rdata)** e em **CSV**, que é lido pelo JASP. Clique na opção desejada.

Base R: Base R

Base JASP: Base CSV

As bases em R tem formato .RData e as bases para o JASP tem formato .CSV.

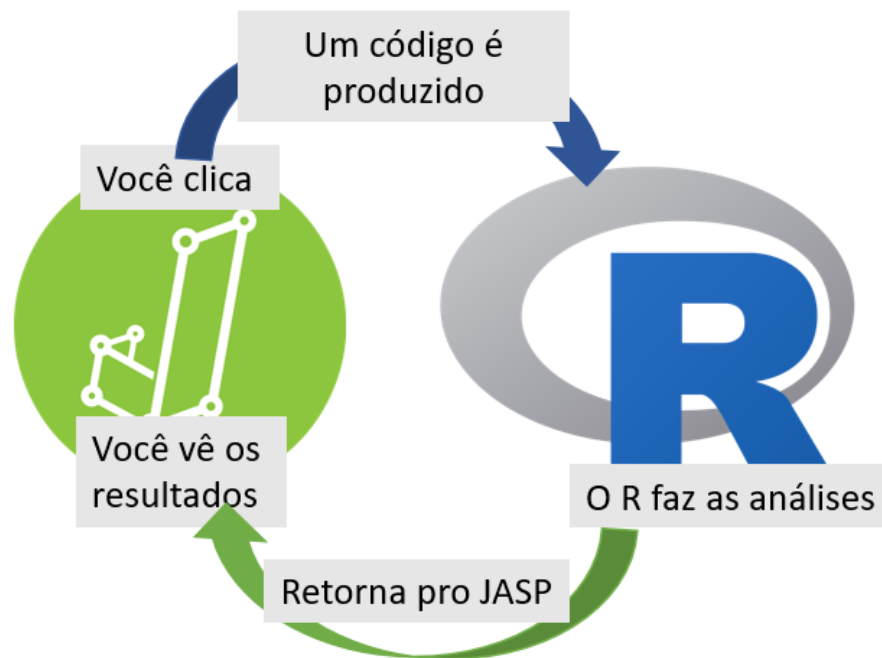
1.8 O R e os pacotes

O livro é integralmente desenvolvido pelo recurso de “programação letrada” no R Markdown, ou seja, ele entrelaça aspectos textuais e linhas de código. Em todos os capítulos, as funções nativas do R e do Tidyverse serão utilizadas. Caso alguém queira reproduzir as análises, será necessário apenas executar as linhas de código disponíveis no decorrer do livro.

O `tidyverse` costuma ter atualizações frequentes. Caso um alerta de `deprecated` seja apresentado, isso significa que a função utilizada foi parcialmente desativada, o que não costuma impactar nas análises.

1.9 JASP

O JASP é um programa gratuito que tem sido cada vez mais utilizado em Psicologia. Ele é feito integralmente por código aberto e sua interface é bastante amigável e intuitiva. Ao instalar o JASP, o R também será instalado em seu computador e ficará no pano de fundo. Dessa maneira, todas as ações feitas por *Point and Click* no JASP, serão convertidas em linhas de código no R e apresentadas de maneira dinâmica no JASP.



Em todos os capítulos, telas do JASP serão apresentadas para que seja possível a reprodução integral de algumas análises. Da mesma forma que qualquer pacote estatístico, o JASP é atualizado frequentemente. Esse livro contou com a versão 0.14.1 e espero que futuras atualizações não comprometam a proposta do livro.

1.10 Outros recursos

Em cada um dos capítulos, aplicações da estatística e referências bibliográficas serão apresentadas. Tenha em mente que há um debate intenso em diferentes conceitos de estatística, da mesma forma que muitas condições computacionais podem aparecer durante a execução das análises propostas. Eu recomendo fortemente a comunidade stackoverflow como um recurso pedagógico para auxiliar em ambos os cenários.

Questões relacionadas aos capítulos são listadas de forma a conectar o conteúdo do livro com exigências balizadas por critérios externos, tal como o ENADE e bancas de concurso.

1.11 Capa

Por tradição, livros de Ciência de Dados e Estatística utilizam a imagem de algum animal na capa. Há livros com cachorros, papagaios, peixes, carangueijos,

lagartos, etc. Esse livro não foge dessa regra e tem como capa a Jolie, a minha cachorrinha com a Anna. Ela foi indispensável para o atraso ao término deste livro.

1.12 Versão do livro

Como todos os livros, este também tem uma história de desenvolvimento. A tabela abaixo apresenta a versão, a data de lançamento e algumas características importantes.

Versão	Data	Características
Beta 1	Fevereiro, 2020	Primeira versão. Baixa revisão textual e dos conceitos estatísticos. Erros são esperados. A utilização deve ser feita apenas de maneira incipiente

1.13 Autor

Luis Anunciação é doutor em Psicometria pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), com intercâmbio na University of Oregon, mestre em saúde pública pela Universidade do Estado do Rio de Janeiro e especialista em Neuropsicologia (IBNeuro) e Bioestatística (Johns Hopkins University). Atualmente, é professor do Departamento de Psicologia da PUC-Rio, coordenador da ANOVA e psicometrista da Nila Press, uma editora especializada no desenvolvimento de instrumentos psicológicos.

1.14 Revisão técnica

Este livro contou com a revisão técnica do Dr. J. Landeira-Fernandez (Doutor em Neurociências pela University of California - UCLA, Pesquisador 1A do CNPq) e Ms. Regina Albanense (Graduada em Matemática e Estatística e Mestre em Ciências). Dr. Landeira foi fundamental em tópicos envolvendo aspectos de pesquisa e análises estatísticas e Ms. Regina foi fundamental para revisão da modelagem matemática e revisão conceitual.

1.15 Agradecimentos

Nenhum homem é uma ilha. Este livro só foi possível graças a um conjunto de pessoas que auxiliaram e fizeram uma profunda revisão do texto. Meus sinceros agradecimentos a (ao):

J. Landeira-Fernandez, PUC-Rio
Regina Albanense, CONRE
Cristiano Fernandes, PUC-Rio
Danilo Assis Pereira, IBNeuro
Anna Carolina de Almeida Portugal, UFRJ
Emanuel Cordeiro, UFPE
Alunos da PUC-Rio, UFRJ, IBNeuro e ANOVA