



# 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](http://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

## 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 ( $\mu_A$ ) é igual a média da população referência ( $\mu_0$ ). 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 ( $\alpha$ ). 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  $\alpha$  de 1%, os dois extremos da curva normal representarão cada um 0,5%; se optarmos por um  $\alpha$  de 5%, as duas extremidades da curva normal compreenderão cada uma 2,5%.

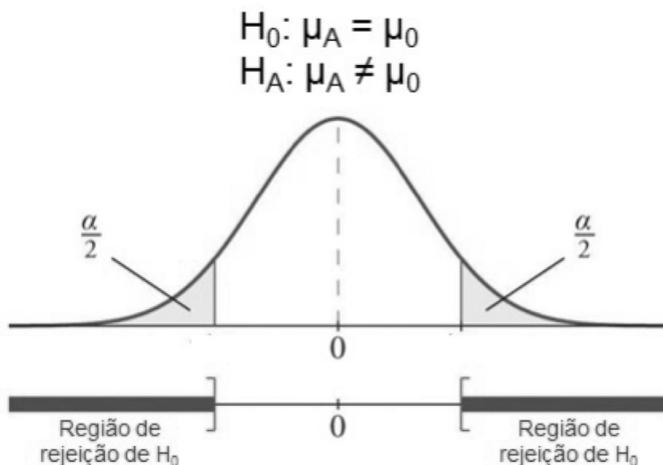


Figura 1. Teste de hipótese bilateral. O  $\mu_0$  representa o valor numérico específico considerado nas  $H_0$  e  $H_A$ .

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  $\mu_A$  for menor ou maior do que  $\mu_0$ . 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  $\alpha$  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  $\alpha$  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  $\alpha$  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 ( $\mu > 38.500$ ).

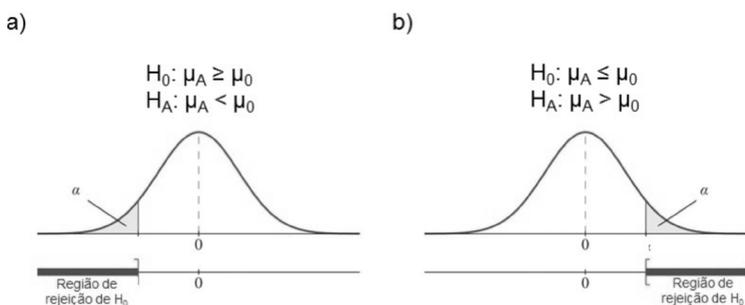


Figura 2. Teste de hipótese unilateral à esquerda (a) e à direita (b). O  $\mu_0$  representa o valor numérico específico considerado nas  $H_0$  e  $H_A$ .

## 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  $\alpha$ , enquanto a probabilidade de se cometer um erro do tipo II é denominada de  $\beta$ . 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: $\alpha$	Decisão correta Probabilidade: $1-\beta$ (Poder do teste)
Não se rejeita $H_0$	Decisão correta Probabilidade: $1-\alpha$	Decisão errada: erro tipo II Probabilidade: $\beta$

Ao diminuir a probabilidade de ocorrer um erro do tipo I, ou seja, diminuindo o valor de  $\alpha$ , 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  $\alpha$ , corresponde a probabilidade máxima permissível de se cometer o erro tipo I. Se o valor de  $p$  é "igual" ou "menor que"  $\alpha$ , dizemos que os dados são estatisticamente significativos no nível  $\alpha$ . Comumente utilizamos valores de  $\alpha$  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  $\alpha$  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.