Capítulo 12

Testes de Hipóteses

12.1 Introdução

Vimos no Capítulo 10 que um dos problemas a serem resolvidos pela Inferência Estatística é o de testar uma hipótese. Isto é, feita determinada afirmação sobre uma população, usualmente sobre um parâmetro dessa, desejamos saber se os resultados experimentais provenientes de uma amostra contrariam ou não tal afirmação. Muitas vezes, essa afirmação sobre a população é derivada de teorias desenvolvidas no campo substantivo do conhecimento. A adequação ou não dessa teoria ao universo real pode ser verificada ou refutada pela amostra. O objetivo do teste estatístico de hipóteses é, então, fornecer uma metodologia que nos permita verificar se os dados amostrais trazem evidências que apóiem ou não uma hipótese (estatística) formulada.

Neste capítulo iremos introduzir o procedimento básico de teste de hipótese sobre um parâmetro de uma população. A idéia central desse procedimento é a de supor verdadeira a hipótese em questão e verificar se a amostra observada é "verossímil" nessas condições. No capítulo seguinte daremos alguns testes para comparação de parâmetros de duas populações.

12.2 Um Exemplo

Vamos introduzir a idéia de teste de uma hipótese por meio de um exemplo hipotético que, partindo de uma situação simples, será gradualmente ampliado para atender à situação geral do teste de hipóteses.

Exemplo 12.1. Uma indústria usa, como um dos componentes das máquinas que produz, um parafuso importado, que deve satisfazer a algumas exigências. Uma dessas é a resistência à tração. Esses parafusos são fabricados por alguns países, e as especificações técnicas variam de país para país. Por exemplo, o catálogo do país A afirma que a resistência média à tração de seus parafusos é de 145 kg, com desvio padrão de 12 kg. Já para o país B, a média é de 155 kg e desvio padrão 20 kg.

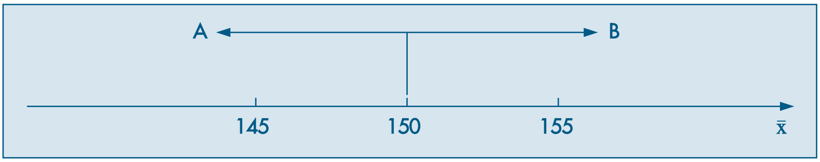
Um lote desses parafusos, de origem desconhecida, será leiloado a um preço muito convidativo. Para que a indústria saiba se faz ou não uma oferta, ela necessita saber qual país produziu tais parafusos. O edital do leiloeiro afirma que, pouco antes do leilão, será divulgada a resistência média $\bar{x}$ de uma amostra de 25 parafusos do lote. Qual regra de decisão deve ser usada pela indústria para dizer se os parafusos são do país A ou B?

Uma resposta que ocorre imediatamente é a que considera como país produtor aquele para o qual a média da amostra mais se aproximar da média da população. Assim, uma possível regra de decisão seria:

Se $\bar{x} \leqslant 150$ (o ponto médio entre 145 e 155), diremos que os parafusos são do país A; caso contrário, isto é, $\bar{x} > 150$, são do país B.

Na Figura 12.1 ilustramos essa regra de decisão.

Figura 12.1: Regra de decisão para o Exemplo 12.1.



Suponha que, no dia do leilão, fôssemos informados de que $\bar{x} = 148$; de acordo com nossa regra de decisão, diríamos que os parafusos são de origem A. Podemos estar enganados nessa conclusão? Ou, em outras palavras, é possível que uma amostra de 25 parafusos de origem B apresente média $\bar{x} = 148$? Sim, é possível. Então, para melhor entendermos a regra de decisão adotada, é interessante estudarmos os tipos de erros que podemos cometer e as respectivas probabilidades.

Podemos cometer dois tipos de erros, e vamos numerá-los para facilitar a linguagem:

Erro de tipo I: dizer que os parafusos são de A quando na realidade são de B. Isso ocorre quando uma amostra de 25 parafusos de B apresenta média $\bar{x}$ inferior ou igual a 150 kg.

Erro de tipo II: dizer que os parafusos são de B, quando na realidade eles são de A. Isso ocorre quando uma amostra de 25 parafusos de A apresenta média $\bar{x}$ superior a 150 kg.

Para facilitar ainda mais, vamos definir duas hipóteses também numeradas:

H_0: os parafusos são de origem B. Isso equivale a dizer que a resistência X de cada parafuso segue uma distribuição com média $\mu = 155$ e desvio padrão $\sigma = 20$.

H_1: os parafusos são de A, isto é, a média $\mu = 145$ e o desvio padrão $\sigma = 12$.

Finalmente, vamos indicar por RC a região correspondente aos valores menores que 150, ou seja,

$$RC = \{y \in \mathbb{R} | y \leqslant 150\}.$$

Com as notações indicadas acima, a probabilidade de se cometer cada um dos erros pode ser escrita:

$$P(\text{erro I}) = P(\bar{X} \in RC | H_0 \text{ é verdadeira}) = \alpha$$

e

$$P(\text{erro II}) = P(\overline{X} \notin RC \mid H_1 \text{ é verdadeira}) = \beta.$$

Quando H_0 for verdadeira, isto é, os parafusos forem de B, sabemos do TLC que $\overline{X}$ terá distribuição aproximadamente normal, com média 155 e desvio padrão igual a $20/\sqrt{25} = 4$, isto é,

$$\overline{X} \sim N(155{,}16).$$

Denotando por Z a v.a. com distribuição $N(0,1)$, temos

$$\begin{aligned} P(\text{erro I}) &= P(\overline{X} \in RC \mid H_0 \text{ é verdadeira}) \\ &= P(\overline{X} \leqslant 150 \mid \overline{X} \sim N(155{,}16)) \\ &= P\left(Z \leqslant \frac{150-155}{4}\right) \\ &= P(Z \leqslant -1{,}25) = 0{,}10565 = 10{,}56\% = \alpha. \end{aligned}$$

De modo análogo, quando H_1 for a alternativa verdadeira, teremos que a v.a. $\overline{X}$ é tal que, aproximadamente,

$$\overline{X} \sim N(145;\ 5{,}76).$$

Teremos, então,

$$\begin{aligned} P(\text{erro II}) &= P(\overline{X} \notin RC \mid H_1 \text{ é verdadeira}) \\ &= P(\overline{X} > 150 \mid \overline{X} \sim N(145;\ 5{,}76)) \end{aligned}$$

$$= P\left(Z > \frac{150-145}{2{,}4}\right) = P(Z > 2{,}08) = 0{,}01876 = 1{,}88\% = \beta.$$

Observando esses dois resultados, notamos que, com a regra de decisão adotada, estaremos cometendo o erro de tipo I com maior probabilidade do que o erro de tipo II. De certo modo, essa regra de decisão privilegia a afirmação de que os parafusos são de A. No Quadro 12.1 ilustramos as conseqüências que podem advir da regra de decisão adotada.

Quadro 12.1: Resumo do teste H_0: $\mu = 155$, H_1: $\mu = 145$, com RC = $]-\infty, 150]$.

<table>
<tr><td rowspan="3">Origem Real dos Parafusos</td><td colspan="2">Decisão</td></tr>
<tr><td colspan="2">RC 150 $\overline{X}$</td></tr>
<tr><td>A</td><td>B</td></tr>
<tr><td>A</td><td>Sem erro</td><td>Erro tipo II
$\beta = 1{,}88\%$</td></tr>
<tr><td>B</td><td>Erro tipo I
$\alpha = 10{,}56\%$</td><td>Sem erro</td></tr>
</table>

Desse quadro, podemos notar que, se os parafusos forem realmente de B (segunda linha) e a amostra tiver média superior a 150 (segunda coluna), diremos que são de B, e não cometeremos erro algum. Por outro lado, se a média $\bar{x}$ for inferior a 150 (primeira coluna), devemos dizer que são de A, e estaremos cometendo um erro cuja probabilidade nesse caso é de 10,56%. De modo análogo, teremos uma interpretação para o caso de os parafusos serem realmente de A (primeira linha).

Para cada regra de decisão adotada, isto é, se escolhermos um valor $\bar{x}_c$ em vez de 150 no Quadro 12.1, apenas as probabilidades α e β mudarão. Se $\bar{x}_c$ for escolhido menor que 150, notamos que α diminuirá e β aumentará. Logo, deve existir um ponto em que α seja igual a β, ou seja, uma regra de decisão em que a probabilidade de errar contra A seja a mesma que errar contra B. Mostre que esse ponto é $\bar{x}_c$ = 148,75, e nesse caso $\alpha = \beta$ = 5,94%.

Do exposto acima constatamos que, escolhido um valor de $\bar{x}_c$, podemos achar as probabilidades α e β de cometer cada tipo de erro. Mas também podemos proceder de modo inverso: fixar um dos erros, digamos α, e encontrar a regra de decisão que irá corresponder à probabilidade de erro de tipo I igual a α.

Por exemplo, fixemos α em 5%, e vejamos qual a regra de decisão correspondente. Temos

$$
\begin{aligned}
5\% = P(\text{erro I}) &= P(\bar{X} \leqslant \bar{x}_c | \bar{X} \sim N(155{,}16)) \\
&= P(Z \leqslant -1{,}645),
\end{aligned}
$$

mas da transformação para a normal padrão sabemos que

$$-1{,}645 = \frac{\bar{x}_c - 155}{4},$$

ou seja, $\bar{x}_c$ = 148,42. Então, a regra de decisão será:

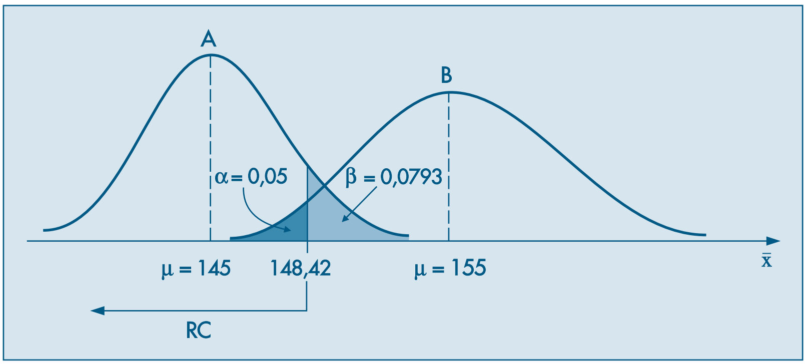
Se $\bar{x}$ for inferior a 148,42, dizemos que o lote é de A; caso contrário, dizemos que é de B.

Com essa regra, a probabilidade do erro de tipo II será

$$
\begin{aligned}
\beta = P(\text{erro II}) &= P(\bar{X} > 148{,}42 | \bar{X} \sim (145;\ 5{,}76)) \\
&= P(Z > 1{,}425) = 7{,}93\%.
\end{aligned}
$$

Veja a ilustração na Figura 12.2.

Figura 12.2: Ilustração dos erros de tipo I e II para o Exemplo 12.1.



Esse segundo tipo de procedimento é bastante utilizado, porque usualmente a decisão que devemos tomar não é apenas entre duas possíveis populações. Os parafusos poderiam ser produzidos por outros países além daqueles citados e, portanto, com outras características quanto à resistência média. Suponha, ainda, que interessa à indústria fazer uma proposta apenas no caso de o parafuso ser de origem B. Qual a regra de decisão que deve adotar?

A hipótese que nos interessa agora é:

H_0: os parafusos são de origem B (μ = 155 e σ = 20).

Caso essa não seja a hipótese verdadeira, a alternativa é muito mais ampla e pode ser expressa como:

H_1: os parafusos não são de origem B (μ e σ desconhecidos).

Aqui não podemos especificar os parâmetros sob a hipótese alternativa H_1, pois se não forem de origem B, os parafusos podem ser de vários outros países, cada um com suas próprias especificações. Alguns países podem ter técnicas mais sofisticadas de produção e, portanto, produzir com resistência média superior a 155. Outros, como no exemplo dado, com resistência menor. A especificação da hipótese alternativa depende muito do grau de informação que se tem do problema. Por exemplo, vamos admitir que a indústria do país B para esse caso seja a mais desenvolvida, e nenhum outro país possa produzir uma resistência média superior à dela. Então, nossa hipótese alternativa seria mais explícita:

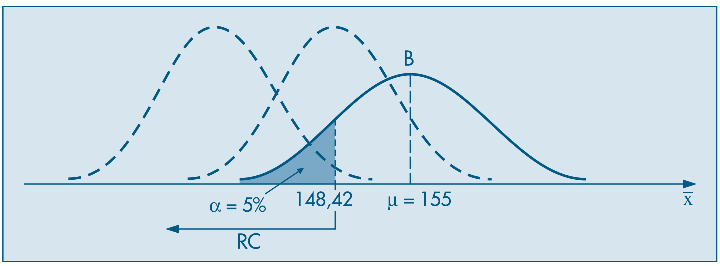
H_1: os parafusos não são de origem B ($\mu < 155$ e σ qualquer).

Isso significa que só iremos desconfiar de H_0 se $\bar{x}$ for muito menor do que 155. Ou seja, a nossa regra de decisão deverá ser semelhante à vista anteriormente. Como os parâmetros sob a hipótese alternativa são muitos, a melhor solução para construir a regra de decisão é fixar α, a probabilidade do erro de tipo I (rejeitar H_0 quando ela for verdadeira). Se fixarmos novamente α = 0,5, e nesse caso a regra de decisão depende apenas das informações de H_0, a regra de decisão será a mesma anterior:

Se $\bar{x}$ for superior a 148,42, diremos que o lote é de origem B; caso contrário, diremos que não é de origem B.

Com essa regra de decisão e com a hipótese alternativa mais ampla, não podemos encontrar β, pois não temos um único parâmetro μ como alternativa e nada sabemos sobre σ. Então, não podemos controlar o erro de tipo II. As implicações dessa regra de decisão estão resumidas na Figura 12.3 e no Quadro 12.2.

Figura 12.3: Teste $H_0 : \mu = 155$ vs $H_1 : \mu < 155$, com RC = $]-\infty; 148,42]$.



Quadro 12.2: Resumo do teste H_0: $\mu = 155$, H_1: $\mu < 155$, com RC = $]-\infty, 148{,}42]$.

Origem Real dos Parafusos	Decisão	
	RC 148,42 $\bar{x}$	
	não B	B
B	Erro tipo I, α = 5%	Sem erro
não B	Sem erro	Erro tipo II, β = ?

Podemos reescrever as hipóteses nessa situação da seguinte maneira:

$$H_0 : \mu = 155$$
$$H_1 : \mu < 155$$

O cálculo de β depende do valor de μ, que não é especificado. Mas podemos considerar a seguinte e importante função.

Definição. A função *característica de operação* (função CO) do teste acima é definida como

$$\beta(\mu) = P(\text{aceitar } H_0 \,|\, \mu) = P(\bar{X} > 148{,}42 \,|\, \mu).$$

Ou seja, $\beta(\mu)$ é a probabilidade de aceitar H_0, considerada como uma função de μ.

Usualmente, considera-se a função $\pi(\mu) = 1 - \beta(\mu)$, que é a probabilidade de se rejeitar H_0, como função de μ. Essa função é chamada *função poder do teste* e será estudada abaixo com certo detalhe. Nesses casos consideramos que σ é o mesmo para todos os valores de μ.

Admitamos, agora, que não exista razão alguma para acreditarmos que a resistência média dos parafusos de B seja maior ou menor do que a de outros países. Isso irá nos levar a duvidar que os parafusos não são de B, se a média observada for muito maior ou muito menor do que 155. Esta situação corresponde à seguinte hipótese alternativa:

H_1: os parafusos não são de origem B ($\mu \neq 155$).

Aqui, a regra de decisão deverá indicar dois pontos $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, tais que:

Se $\bar{x}$ estiver entre $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, diremos que os parafusos são de origem B; se $\bar{x}$ estiver fora do intervalo, diremos que não são de origem B.

Fixado α, a probabilidade do erro I, existirão muitos valores que satisfazem a essa condição. Daremos preferência àquelas soluções $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, simétricas em relação à média. Veja a Figura 12.4.

Voltando ao nosso problema, e fixado α em 5%, temos

$$0{,}05 = P(\text{erro I}) = P(\bar{X} < \bar{x}_{c_1} \text{ ou } \bar{X} > \bar{x}_{c_2} \,|\, \bar{X} \sim N(155{,}16))$$
$$= P(Z < -1{,}96 \text{ ou } Z > 1{,}96),$$

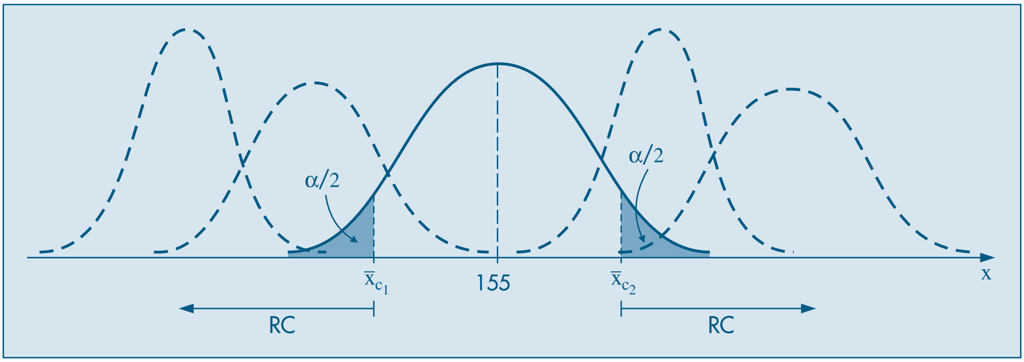
e daqui encontramos

$$-1{,}96 = (\bar{x}_{c_1} - 155)/4 \;\Rightarrow \bar{x}_{c_1} = 147{,}16$$

e

$$1{,}96 = (\bar{x}_{c_2} - 155)/4 \;\Rightarrow \bar{x}_{c_2} = 162{,}84.$$

Figura 12.4: Teste $H_0 : \mu = 155$ vs $H_1 : \mu \neq 155$.



Portanto, nesse caso, a região de rejeição da hipótese H_0 é (veja o Quadro 12.3)

$$RC = \{\bar{x} \in \mathbb{R} \,|\, \bar{x} < 147{,}16 \text{ ou } \bar{x} > 162{,}84\}.$$

Do apresentado nesta seção, vemos que, dependendo do grau de informação que se tem do problema, podemos ter regras de decisão unilaterais ou bilaterais. Na seção seguinte iremos dar os passos para a construção de um teste de hipótese.

Quadro 12.3: Resumo do teste H_0: $\mu = 155$, H_1: $\mu \neq 155$, com RC = $]-\infty, 147{,}16] \cup [162{,}84, +\infty[$.

Origem Real dos Parafusos	Decisão	
	RC: $\bar{x}$ < 147,16	RC: $\bar{x}$ > 162,84
	B	não B
B	Sem erro	Erro tipo II, β = ?
não B	Erro tipo I, α = 5%	Sem erro

Problemas

1. Para decidirmos se os habitantes de uma ilha são descendentes da civilização A ou B, iremos proceder do seguinte modo:
   (i) selecionamos uma amostra de 100 moradores adultos da ilha, e determinamos a altura média deles;
   (ii) se essa altura média for superior a 176, diremos que são descendentes de B; caso contrário, são descendentes de A.

   Os parâmetros das alturas das duas civilizações são:

   A : $\mu = 175$ e $\sigma = 10$;

   B : $\mu = 177$ e $\sigma = 10$.

   Definamos: Erro de tipo I — dizer que os habitantes da ilha são descendentes de B quando, na realidade, são de A.

   Erro de tipo II — dizer que são de A quando, na realidade, são de B.

   (a) Qual a probabilidade do erro de tipo I? E do erro de tipo II?

(b) Qual deve ser a regra de decisão se quisermos fixar a probabilidade do erro de tipo I em 5%? Qual a probabilidade do erro de tipo II, nesse caso?

(c) Se $\sigma_A = 5$, como ficariam as respostas de (b)?

(d) Quais as probabilidades do erro de tipo II, nas condições da questão (b), se a média $\mu_B = 178$? E $\mu_B = 180$? E $\mu_B = 181$? Coloque num gráfico os pares $(\mu_B, P(\text{erro II} \mid \mu_B))$.

2. Fazendo o teste

$$H_0 : \mu = 1.150\ (\sigma = 150) \text{ contra } H_1 : \mu = 1.200\ (\sigma = 200),$$

e $n = 100$, estabeleceu-se a seguinte região crítica:

$$RC = [1.170, +\infty[.$$

(a) Qual a probabilidade α de rejeitar H_0 quando verdadeira?

(b) Qual a probabilidade β de aceitar H_0 quando H_1 é verdadeira?

(c) Qual deve ser a região crítica para que $\alpha = \beta$?

3. Nas situações abaixo, escolha como hipótese nula, H_0, aquela que para você leva a um erro de tipo I mais importante. Descreva quais os dois erros em cada caso.

(a) O trabalho de um operador de radar é detectar aeronaves inimigas. Quando surge alguma coisa estranha na tela, ele deve decidir entre as hipóteses:

1. está começando um ataque;
2. tudo bem, apenas uma leve interferência.

(b) Num júri, um indivíduo está sendo julgado por um crime. As hipóteses sujeitas ao júri são:

1. o acusado é inocente;
2. o acusado é culpado.

(c) Um pesquisador acredita que descobriu uma vacina contra resfriado. Ele irá conduzir uma pesquisa de laboratório para verificar a veracidade da afirmação. De acordo com o resultado, ele lançará ou não a vacina no mercado. As hipóteses que pode testar são:

1. a vacina é eficaz;
2. a vacina não é eficaz.

4. Se, ao lançarmos três vezes uma moeda, aparecerem 3 coroas, decidimos rejeitar a hipótese de que a moeda é "honesta". Quais as probabilidades de erro de tipo I e erro de tipo II, se $p = 2/3$?

5. A variável X, custo de manutenção de um tear, pode ser considerada como tendo distribuição normal de média μ e desvio padrão 20 unidades. Os valores possíveis de μ podem ser 200 ou 210. Para verificar qual dos dois valores é o mais provável, usar-se-á uma amostra de 25 teares. Defina:

(a) Uma hipótese a ser testada.

(b) Uma regra de decisão e encontre as probabilidades dos erros de tipo I e II.

12.3 Procedimento Geral do Teste de Hipóteses

A construção de um teste de hipóteses, para um parâmetro populacional, pode ser colocada do seguinte modo. Existe uma variável X associada a dada população e tem-se uma hipótese sobre determinado parâmetro θ dessa população. Por exemplo, afirmamos que o verdadeiro valor de θ é θ_0. Colhe-se uma amostra aleatória de elementos dessa população, e com ela deseja-se comprovar ou não tal hipótese.

Como já vimos anteriormente, iniciamos nossa análise explicitando claramente qual a hipótese que estamos colocando à prova e a chamamos de *hipótese nula*, e escrevemos

$$H_0 : \theta = \theta_0.$$

Em seguida, convém explicitar também a hipótese que será considerada aceitável, caso H_0 seja rejeitada. A essa hipótese chamamos de *hipótese alternativa*, e a sua caracterização estatística irá depender do grau de conhecimento que se tem do problema estudado. A alternativa mais geral seria

$$H_1 : \theta \neq \theta_0.$$

Poderíamos, ainda, ter alternativas da forma

$$H_1 : \theta < \theta_0 \quad \text{ou} \quad H_1 : \theta > \theta_0,$$

dependendo das informações que o problema traz.

Qualquer que seja a decisão tomada, vimos que estamos sujeitos a cometer erros. Para facilitar a linguagem, introduzimos as definições:

Erro de tipo I: rejeitar a hipótese nula quando essa é verdadeira. Chamamos de α a probabilidade de cometer esse erro, isto é,

$$\alpha = P(\text{erro do tipo I}) = P(\text{rejeitar } H_0 | H_0 \text{ é verdadeira}).$$

Erro de tipo II: não rejeitar H_0 quando H_0 é falsa. A probabilidade de cometer esse erro é denotada por β, logo

$$\beta = P(\text{erro do tipo II}) = P(\text{não rejeitar } H_0 | H_0 \text{ é falsa}).$$

O objetivo do teste de hipóteses é dizer, usando uma estatística $\hat{\theta}$, se a hipótese H_0 é ou não aceitável. Operacionalmente, essa decisão é tomada através da consideração de uma região crítica RC. Caso o valor observado da estatística pertença a essa região, rejeitamos H_0; caso contrário, não rejeitamos H_0. Esta região é construída de modo que P($\hat{\theta} \in$ RC$|H_0$ é verdadeira) seja igual a α, fixado *a priori*. RC recebe o nome de *região crítica ou região de rejeição* do teste. Um fato importante a ressaltar é que a região crítica é sempre construída sob a hipótese de H_0 ser verdadeira. A determinação do valor de β já é mais difícil, pois usualmente não especificamos valores fixos para o parâmetro sob a hipótese alternativa. Mais adiante trataremos dessa situação, ao considerarmos o poder de um teste.

A probabilidade α de se cometer um erro de tipo I (ou de primeira espécie) é um valor arbitrário e recebe o nome de *nível de significância* do teste. O resultado da amostra é tanto mais significante para rejeitar H_0 quanto menor for esse nível α. Ou seja, quanto menor for α, menor é a probabilidade de se obter uma amostra com estatística pertencente à região crítica, sendo pouco verossímil a obtenção de uma amostra da população para a qual H_0 seja verdadeira. Usualmente, o valor de α é fixado em 5%, 1% ou 0,1%.

A fixação do valor de α envolve uma questionável arbitrariedade. Neste sentido há um modo alternativo de se proceder, que será considerado na seção 12.8.

12.4 Passos para a Construção de um Teste de Hipóteses

Vimos nas seções anteriores o procedimento que se deve usar para realizar um teste de hipóteses. Daremos abaixo uma seqüência que pode ser usada sistematicamente para qualquer teste de hipóteses.

Passo 1. Fixe qual a hipótese H_0 a ser testada e qual a hipótese alternativa H_1.

Passo 2. Use a teoria estatística e as informações disponíveis para decidir qual estatística (estimador) será usada para testar a hipótese H_0. Obter as propriedades dessa estatística (distribuição, média, desvio padrão).

Passo 3. Fixe a probabilidade α de cometer o erro de tipo I e use este valor para construir a região crítica (regra de decisão). Lembre que essa região é construída para a estatística definida no passo 2, usando os valores do parâmetro hipotetizados por H_0.

Passo 4. Use as observações da amostra para calcular o valor da estatística do teste.

Passo 5. Se o valor da estatística calculado com os dados da amostra não pertencer à região crítica, não rejeite H_0; caso contrário, rejeite H_0.

Procuraremos, sempre que fizermos teste de hipóteses, distinguir bem esses cinco passos. Finalmente um comentário sobre H_0 e o erro de tipo I. Devemos tomar como H_0 aquela hipótese, que, rejeitada, conduza a um erro de tipo I mais importante de evitar. Vejamos um exemplo devido a Neyman (1978). Suponha um experimento para se determinar se um produto A é ou não cancerígeno. Após realizado o teste, podemos concluir: (i) A é cancerígeno ou (ii) A não é cancerígeno. Cada uma dessas conclusões pode estar errada e temos os dois tipos de erro já mencionados, dependendo de qual hipótese seja H_0. Do ponto de vista do usuário do produto, a hipótese a ser testada deve ser

$$H_0 : \text{A é cancerígeno,}$$

pois a probabilidade de erro na rejeição dessa hipótese, se ela for verdadeira, deve ser um valor muito pequeno. Outros exemplos estão contidos no Problema 3.

12.5 Testes sobre a Média de uma População com Variância Conhecida

Vejamos, agora, uma aplicação dos cinco passos definidos na seção anterior, para testar a hipótese de que a média de uma população μ seja igual a um número fixado μ_0, supondo-se a variância σ^2 dessa população conhecida.

Exemplo 12.2. Uma máquina automática para encher pacotes de café enche-os segundo uma distribuição normal, com média μ e variância sempre igual a 400 g^2. A máquina foi regulada para μ = 500 g. Desejamos, periodicamente, colher uma amostra de 16 pacotes e verificar se a produção está sob controle, isto é, se $\mu = 500$ g ou não. Se uma dessas amostras apresentasse uma média $\bar{x} = 492$ g, você pararia ou não a produção para regular a máquina?

Vejamos como testar essa hipótese.

Passo 1. Indiquemos por X o peso de cada pacote; então, $X \sim N(\mu, 400)$. E as hipóteses que nos interessam são:

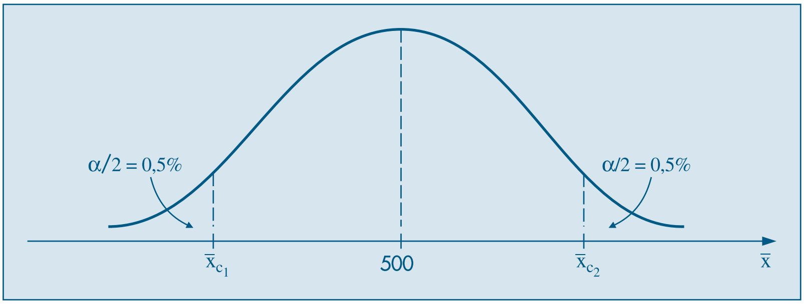
$$H_0 : \mu = 500 \text{ g},$$
$$H_1 : \mu \neq 500 \text{ g},$$

pois a máquina pode desregular para mais ou para menos.

Passo 2. Pela afirmação do problema, $\sigma^2 = 400$ será sempre a mesma; logo, para todo μ, a média $\bar{X}$ de 16 pacotes terá distribuição $N(\mu, 400/16)$, de modo que o desvio padrão (ou erro padrão) de $\bar{X}$ é $\sigma_{\bar{x}} = 5$. Em particular, se H_0 for verdadeira, $\bar{X} \sim N(500,25)$.

Passo 3. Vamos fixar $\alpha = 1\%$; pela hipótese alternativa, vemos que H_0 deve ser rejeitada quando $\bar{X}$ for muito pequena ou muito grande (dizemos que temos um teste bilateral). Portanto, nossa região crítica será como a da Figura 12.5.

Figura 12.5: Região crítica para o teste $H_0 : \mu = 500$ vs $H_1 : \mu \neq 500$ do Exemplo 12.2.



Da tabela da curva normal padronizada obtemos que

$$z_1 = -2{,}58 = (\bar{x}_{c_1} - 500)/5 \Rightarrow \bar{x}_{c_1} = 487{,}1,$$
$$z_2 = 2{,}58 = (\bar{x}_{c_2} - 500)/5 \Rightarrow \bar{x}_{c_2} = 512{,}9.$$

Segue-se que a região crítica é

$$RC = \{\bar{x} \in \mathbb{R} \mid \bar{x} \leqslant 487{,}1 \text{ ou } \bar{x} \geqslant 512{,}9\}.$$

Passo 4. A informação pertinente da amostra é sua média, que nesse caso particular é $\bar{x}_0 = 492$.

Passo 5. Como $\bar{x}_0$ não pertence à região crítica, nossa conclusão será não rejeitar H_0. Ou seja, o desvio da média da amostra para a média proposta por H_0 pode ser considerado como devido apenas ao sorteio aleatório dos pacotes.

A situação analisada não é muito realista: conhecer a variância da população. O caso mais geral, de média e variância desconhecidas, será tratado na seção 12.10.

Problemas

6. Sabe-se que o consumo mensal *per capita* de um determinado produto tem distribuição normal, com desvio padrão 2 kg. A diretoria de uma firma que fabrica esse produto resolveu que retiraria o produto da linha de produção se a média de consumo *per capita* fosse menor que 8 kg. Caso contrário, continuaria a fabricá-lo. Foi realizada uma pesquisa de mercado, tomando-se uma amostra de 25 indivíduos, e verificou-se que $\sum_{i=1}^{25} X_i = 180$ kg, onde X_i representa o consumo mensal do i-ésimo indivíduo da amostra.
   (a) Construa um teste de hipótese adequado, utilizando $\alpha = 0{,}05$, e com base na amostra colhida determine a decisão a ser tomada pela diretoria.
   (b) Qual a probabilidade β de se tomar uma decisão errada se, na realidade, a média populacional for $\mu = 7{,}8$ kg?
   (c) Se a diretoria tivesse fixado $\alpha = 0{,}01$, a decisão seria a mesma? (Justifique sua resposta.)
   (d) Se o desvio da população fosse 4 kg, qual seria a decisão, com $\alpha = 0{,}05$? (Justifique sua resposta.)
7. A associação dos proprietários de indústrias metalúrgicas está muito preocupada com o tempo perdido com acidentes de trabalho, cuja média, nos últimos tempos, tem sido da ordem de 60 horas/homem por ano e desvio padrão de 20 horas/homem. Tentou-se um programa de prevenção de acidentes, após o qual foi tomada uma amostra de nove indústrias e medido o número de horas/homens perdidas por acidente, que foi de 50 horas. Você diria, no nível de 5%, que há evidência de melhoria?
8. O salário médio dos empregados das indústrias siderúrgicas de um país é de 2,5 salários mínimos, com um desvio padrão de 0,5 salários mínimos. Uma indústria é escolhida ao acaso e desta é escolhida uma amostra de 49 empregados, resultando um salário médio de 2,3 salários mínimos. Podemos afirmar que esta indústria paga salários inferiores à média nacional, com o nível de 5%?
9. Uma companhia de cigarros anuncia que o índice médio de nicotina dos cigarros que fabrica apresenta-se abaixo de 23 mg por cigarro. Um laboratório realiza 6 análises desse índice, obtendo: 27, 24, 21, 25, 26, 22. Sabe-se que o índice de nicotina se distribui normalmente, com variância igual a 4,86 mg^2. Pode-se aceitar, no nível de 10%, a afirmação do fabricante?

12.6 Teste para Proporção

Vamos usar os passos descritos na seção 12.4 para mostrar a construção do teste para proporções.

Passo 1. Temos uma população e uma hipótese sobre a proporção p de indivíduos portadores de certa característica. Esta hipótese afirma que essa proporção é igual a certo valor p_0. Então,

$$H_0 : p = p_0.$$

O problema fornece informações sobre a alternativa, que pode ter uma das três formas abaixo:

(i) $H_1 : p \neq p_0$ (teste bilateral);
(ii) $H_1 : p > p_0$ (teste unilateral à direita); e
(iii) $H_1 : p < p_0$ (teste unilateral à esquerda).

Passo 2. Como vimos na seção 10.9, a estatística $\hat{p}$, a proporção amostral, tem uma distribuição aproximadamente normal, a saber,

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right).$$

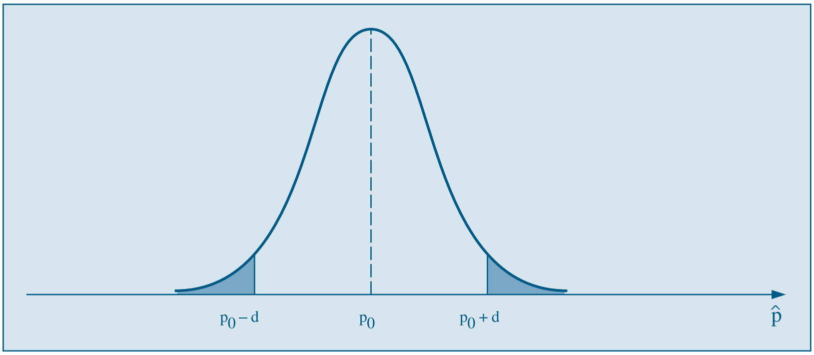
Passo 3. Fixado um valor de α, devemos construir a região crítica para p, sob a suposição de que o parâmetro definido por H_0 seja o verdadeiro. Ou seja, podemos escrever

$$\hat{p} \sim N\left(p_0, \frac{p_0(1-p_0)}{n}\right),$$

e, conseqüentemente, teremos a região crítica da Figura 12.6, supondo a alternativa (i) acima; sendo que $d = Z(1-\alpha/2)\sqrt{p_0(1-p_0)/n}$ e $Z(p)$ é o p-quantil da normal padrão.

O quarto e quinto passos irão depender da amostra, e o procedimento está descrito no exemplo seguinte.

Figura 12.6: Região crítica para o teste $H_0 : p = p_0$ vs $H_1 : p \neq p_0$.



Exemplo 12.3. Uma estação de televisão afirma que 60% dos televisores estavam ligados no seu programa especial da última segunda-feira. Uma rede competidora deseja contestar essa afirmação e decide usar uma amostra de 200 famílias para um teste. Qual deve ser o procedimento adotado para avaliar a veracidade da afirmação da estação? No passo 4 a seguir daremos o resultado da amostra, pois é importante ficar claro que esse resultado não deve influenciar a escolha da alternativa.

Passo 1. Vamos colocar à prova a afirmação da estação, isto é,

$$H_0 : p = 0{,}60.$$

Sabemos que, se essa hipótese não for verdadeira, espera-se uma proporção menor, nunca maior. A estação divulgaria o máximo possível. Isso nos leva à hipótese alternativa

$$H_1 : p < 0{,}60.$$

Passo 2. A estatística a ser usada é $\hat{p}$, a proporção de 200 famílias que assistiram ao programa na última segunda-feira, e da teoria sabemos que

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{200}\right).$$

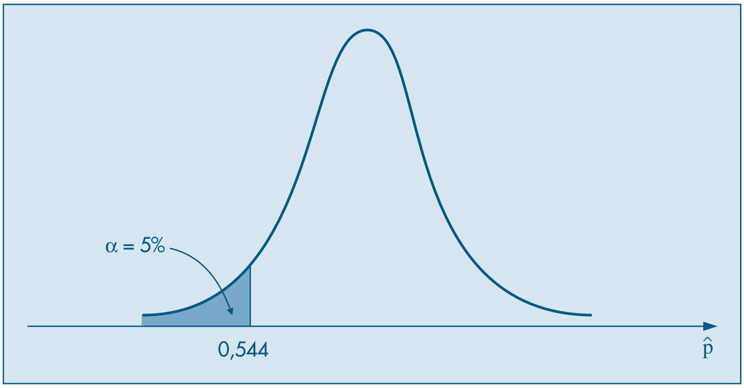
Passo 3. Fixaremos $\alpha = 0{,}05$ e sob a suposição que H_0 seja verdadeira,

$$\hat{p} \sim N(0{,}60,\ 0{,}24/200),$$

o que irá fornecer a região crítica (veja a Figura 12.7)

$$RC = \{\hat{p} \in \mathbb{R} \mid \hat{p} \leqslant 0{,}544\}.$$

Figura 12.7: Região crítica para o teste $H_0: p = 0{,}60$ vs $H_1: p < 0{,}60$ do Exemplo 12.3.



De fato, devemos achar o valor $\hat{p}_c$, tal que $P(\hat{p} \leqslant \hat{p}_c) = 0{,}05$, e usando a aproximação normal acima, teremos

$$P\left(Z \leqslant \frac{\hat{p}_c - 0{,}60}{\sqrt{0{,}24/200}}\right) = 0{,}05,$$

o que implica

$$\frac{\hat{p}_c - 0{,}60}{\sqrt{0{,}24/200}} = -1{,}645,$$

o valor –1,645 sendo obtido da normal padronizada. Segue-se que $\hat{p}_c = 0{,}544$, correspondendo à região crítica acima.

Passo 4. Admitamos que, da pesquisa feita com as 200 famílias, obtivemos 104 pessoas que estavam assistindo ao programa. A proporção da amostra será $\hat{p} = 104/200 = 0{,}52$.

Passo 5. Do resultado do passo anterior, vemos que $0{,}52 \in RC$; portanto, somos levados a rejeitar H_0. Isto é, há evidências que a audiência do programa de segunda-feira não foi de 60% e sim inferior a esse número.

Problemas

10. Uma pessoa gaba-se de adivinhar qual será o resultado do lance de uma moeda, mas é preciso que os presentes não o perturbem com pensamentos duvidosos. Para testar tal capacidade, lançou-se uma moeda perfeita 6 vezes, e o adivinhador acertou 5. Qual seria sua conclusão?

11. O consumidor de um certo produto acusou o fabricante, dizendo que mais de 20% das unidades fabricadas apresentam defeito. Para confirmar sua acusação, ele usou uma amostra de tamanho 50, onde 27% das peças eram defeituosas. Mostre como o fabricante poderia refutar a acusação. Utilize um nível de significância de 10%.

12. Um fabricante garante que 90% dos equipamentos que fornece a uma fábrica estão de acordo com as especificações exigidas. O exame de uma amostra de 200 peças desse equipamento revelou 25 defeituosas. Teste a afirmativa do fabricante, nos níveis de 5% e 1%.

13. Os produtores de um programa de televisão pretendem modificá-lo se for assistido regularmente por menos de um quarto dos possuidores de televisão. Uma pesquisa encomendada a uma empresa especializada mostrou que, de 400 famílias entrevistadas, 80 assistem ao programa regularmente. Com base nos dados, qual deve ser a decisão dos produtores?

12.7 Poder de um Teste

Vimos que, na construção de um teste de hipóteses, procuramos controlar o erro de tipo I, fixando sua probabilidade de ocorrência, α, e construindo a região crítica de modo que $P(\text{RC}|H_0 \text{ verdadeira}) = \alpha$. Ou seja, admitindo que H_0 seja verdadeira, estamos admitindo conhecido(s) o(s) parâmetro(s) que define(m) a distribuição da estatística usada no teste.

Por outro lado, a probabilidade do erro do tipo II, na maioria dos casos, não pode ser calculada, pois a hipótese alternativa usualmente especifica um conjunto de valores para o parâmetro. Voltemos ao exemplo da seção anterior.

Exemplo 12.2. (continuação) No exemplo da máquina de encher pacotes de café, a v.a. X, que descrevia o peso de cada pacote, tinha uma distribuição normal com média μ e variância 400, de modo que a média amostral $\bar{X} \sim N(500, 25)$, sob a hipótese H_0. Esse fato foi utilizado para determinar a região crítica $\text{RC} = \{\bar{x} \in \mathbb{R} \mid \bar{x} < 487{,}1 \text{ ou } \bar{x} > 512{,}9\}$ e nossa regra de decisão para verificar se a máquina estava ou não produzindo sob controle foi:

Se $\bar{x} \in RA$, a máquina está sob controle; se $\bar{x} \in RC$, não está,

onde RA é a região de aceitação do teste, isto é, o complementar de RC em relação a $\mathbb{R}$ e, portanto, dada no nosso caso por $\text{RA} = \{\bar{x} \in \mathbb{R} | 487{,}1 \leqslant \bar{x} \leqslant 512{,}9\}$.

A probabilidade β do erro de tipo II não pode ser calculada, a menos que se especifique um valor alternativo para μ. Segue-se que a função característica de operação do teste é dada por

$$\begin{aligned}\beta(\mu) &= P(\text{aceitar } H_0|\mu) = P(\bar{X} \in RA|\mu) \\ &= P(487{,}1 \leqslant \bar{X} \leqslant 512{,}9)|\mu).\end{aligned}$$

Por exemplo, se a máquina se desregular para μ = 505, teremos

$$\beta(505) = P(\overline{X} \in \text{RA}\,|\,\mu = 505) = P(-3{,}58 \leqslant Z \leqslant 1{,}58) = 94{,}28\%,$$

usando o fato que agora $\overline{X} \sim N(505, 25)$. Lembre-se de que supomos que $\sigma^2 = 400$, sempre!

Para qualquer outro valor do parâmetro μ podemos encontrar o respectivo valor de β, para a regra de decisão adotada. No Quadro 12.4 temos as decisões que podemos tomar e suas respectivas implicações.

Quadro 12.4: Decisões possíveis para o teste $H_0 : \mu = 500$ *versus* $H_1 : \mu \neq 500$

Decisão	Valor real do parâmetro	
	$H_0 : \mu = 500$	$H_1 : \mu \neq 500$
a máquina está sob controle: $\mu = 500$	$P(RA \mid H_0) = 0{,}99$	$P(RA \mid H_1) = \beta$ depende de valor alternativo de μ
a máquina não está sob controle: $\mu \neq 500$	$P(RC \mid H_0) = 0{,}01$	$P(RC \mid H_1) = 1 - \beta$ depende de valor alternativo de μ

Observe, por exemplo, que $1 - \beta(500) = P(\text{rejeitar } H_0|\mu = 500) = \alpha = 0{,}01$.

A quantidade $1 - \beta(\mu)$ é usualmente chamada de *poder* ou *potência do teste*, e é a probabilidade de rejeitar a hipótese H_0, dado um valor qualquer de μ, especificado ou não pela hipótese alternativa, e será denotado por $\pi(\mu)$. No nosso exemplo,

$$\pi(\mu) = P(\text{rejeitar } H_0|\mu) = P(\overline{X} < 487{,}1 \text{ ou } \overline{X} > 512{,}9\,|\,\mu).$$

Na Tabela 12.1 temos alguns valores de $\beta(\mu)$ e de $\pi(\mu)$, para diferentes valores de μ, e na Figura 12.8 a representação gráfica da determinação dessa probabilidade. Observe que quanto maior for a distância entre o valor fixado em $H_0(\mu = 500)$ e o valor atribuído para a hipótese alternativa, maior será a probabilidade de tomar a decisão correta. Na Figura 12.9 temos o gráfico de $\pi(\mu)$ para os valores de μ da Tabela 12.1.

Tabela 12.1: Valores de $\beta(\mu)$ e $\pi(\mu)$, usando a regra de decisão $RC = \{\bar{x} \in I\!R \,|\, \bar{x} \leqslant 487{,}1 \text{ ou } \bar{x} \geqslant 512{,}9\}$

Verdadeiro valor de μ		$\pi(\mu)$ (em %)	$\beta(\mu)$ (em %)
À esquerda de 500	À direita de 500		
500	500	1,0	99,0
498	502	1,7	98,3
495	505	5,7	94,3
492	508	16,4	83,6
490	510	28,1	71,9
487	513	49,0	51,0
485	515	66,3	34,7
480	520	92,1	7,9
475	525	99,2	0,8

Figura 12.8: Determinação do poder para o teste do Exemplo 12.2.

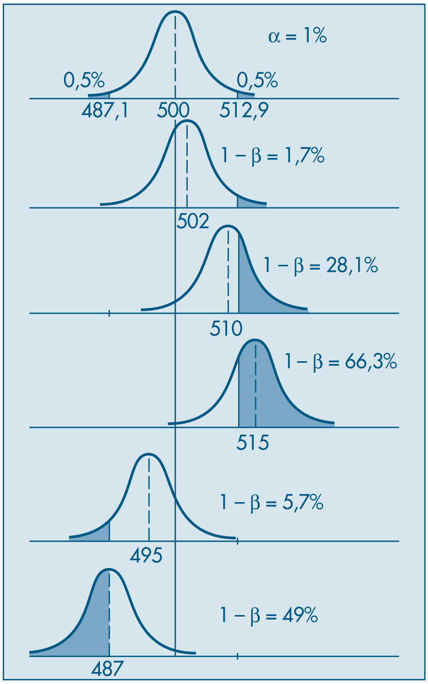
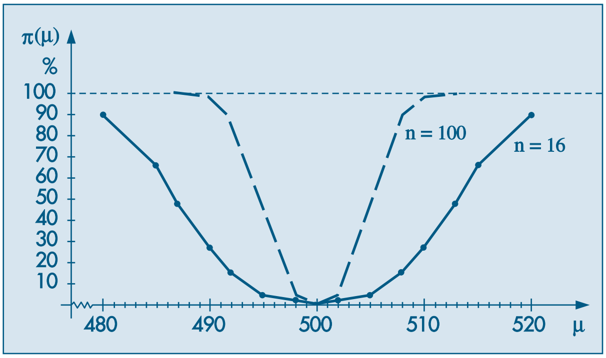


Figura 12.9: Curva de poder para o Exemplo 12.2.



As seguintes propriedades de $\pi(\mu)$ são facilmente verificadas:

(i) $\pi(-\infty) = \pi(+\infty) = 1$;

(ii) $\pi(500) = \alpha$;

(iii) π decresce para $\mu < 500$ (isto é, $d\pi/d\mu < 0$ para $\mu < 500$) e π cresce para $\mu > 500$ (isto é, $d\pi/d\mu > 0$, para $\mu > 500$).

Vemos que $\pi(\mu)$ indica a probabilidade de uma decisão correta, para as diversas alternativas do parâmetro e pode ser usada para decidir entre dois testes para uma mesma hipótese.

Exemplo 12.4. Se, no Exemplo 12.2, a amostra colhida fosse de 100 pacotes em vez de 16, e mantivéssemos o mesmo nível de significância $\alpha = 1\%$, a nova região crítica seria

$$RC = \{\bar{x} \in \mathbb{R} \mid \bar{x} \leqslant 494{,}8 \quad \text{ou} \quad \bar{x} \geqslant 505{,}2\}.$$

Construindo a função poder para esse teste, obtemos a curva tracejada na Figura 12.9. Verifique essas afirmações.

Observando as duas curvas na Figura 12.9, notamos que para todos os valores sob a hipótese alternativa, a probabilidade de uma decisão correta é maior para amostras de tamanho 100 do que de tamanho 16. Dizemos, nesse caso, que o teste baseado em amostras de tamanho 100 é *mais poderoso* do que o teste baseado em amostras de tamanho 16. Esse fato está de acordo com a intuição de que um teste com amostras maiores deve levar a melhores resultados.

De modo geral, se quisermos testar

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0,$$

e determinada a RC do teste, baseada na estatística $\hat{\theta}$, podemos dar a seguinte definição geral.

Definição. A função poder (ou potência) do teste de H_0 contra H_1 é definida por

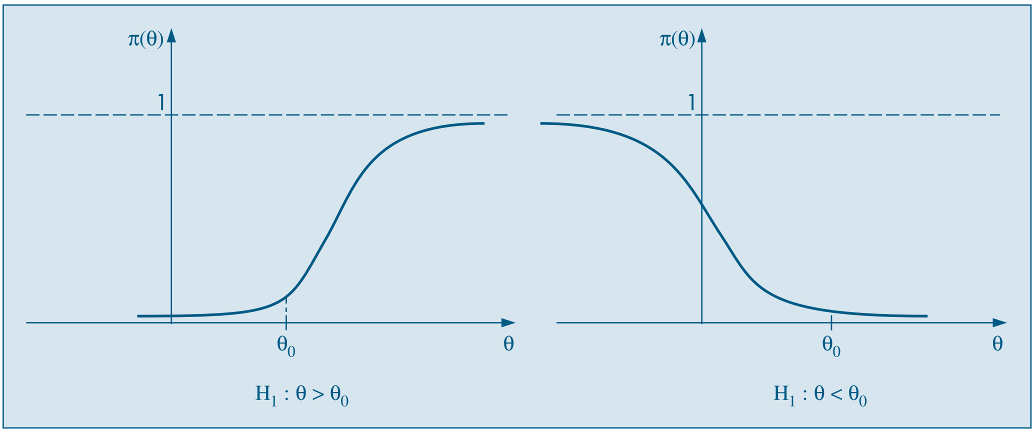
$$\pi(\theta) = P(\hat{\theta} \in RC|\theta),$$

ou seja, é a probabilidade de rejeitar a hipótese nula, como função de θ.

O gráfico dessa função é semelhante àqueles da Figura 12.9, e $\pi(\theta)$ tem as propriedades (i)–(iii) acima, substituindo 500 por θ_0.

Se tivermos hipóteses alternativas unilaterais, da forma $H_1 : \theta < \theta_0$ ou $H_1 : \theta > \theta_0$, obteremos os gráficos da Figura 12.10.

Figura 12.10: Curvas de poder para alternativas unilaterais.



Nos exemplos anteriores fixamos o tamanho da amostra, n, e o nível de significância, α. Suponha que queiramos determinar o tamanho da amostra e os limites da RC, para alcançarmos dado poder para determinado valor do parâmetro. No Exemplo 12.2 poderíamos, por exemplo, fixar $\pi(510) = 0,80$ e $\pi(500) = 0,05$ (o nível de significância). Dados esses valores, podemos determinar n e a RC. Veja o Problema 33.

Problemas

14. Suponha que estejamos testando $H_0 : p = 0,5$ contra $H_1 : p \neq 0,5$, e que, para uma amostra de tamanho $n = 10$, decidimos pela região crítica $RC = \{0, 1, 2, 8, 9, 10\}$.
    (a) Determine o nível de significância α.
    (b) Calcule o poder do teste para $p = 0,2, 0,4, 0,6, 0,8$. Faça um gráfico do poder como função de p.
    (c) Qual o poder do teste para $p = 0,5$?
15. Sendo X o custo de manutenção de um tear, sabe-se que $X \sim N(\mu, 400)$. Para testar a hipótese $H_0 : \mu = 200$, contra a alternativa $H_1 : \mu > 200$, será usada uma amostra de 25 teares.
    (a) Fixando-se $\alpha = 5\%$, encontre a correspondente RC.
    (b) Atribuindo-se valores arbitrários para μ, esboce a função poder do teste.
    (c) Para que valores de μ o poder será maior do que 50%?

12.8 Valor-*p*

O método de construção de um teste de hipóteses, descrito nas seções anteriores, parte da fixação do nível de significância α. Pode-se argumentar que esse procedimento pode levar à rejeição da hipótese nula para um valor de α e à não-rejeição para um valor menor. Outra maneira de proceder consiste em apresentar a *probabilidade de significância* ou *nível descritivo* ou ainda *valor-p* do teste. Os passos são muito parecidos aos já apresentados; a principal diferença está em não construir a região crítica. O que se faz é indicar a probabilidade de ocorrer valores da estatística mais extremos do que o observado, sob a hipótese de H_0 ser verdadeira.

Exemplo 12.5. Voltemos ao Exemplo 12.3, onde

$$H_0 : p = 0,60.$$

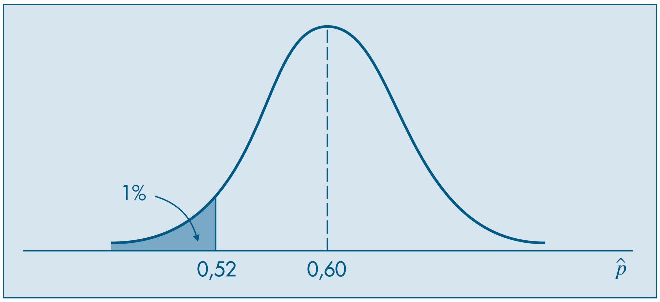
Como vimos, admitindo essa hipótese verdadeira, $\hat{p} \sim N(0,60; 0,24/200)$. Colhida a amostra obtivemos $\hat{p}_0 = 104/200 = 0,52$. Portanto, podemos calcular qual a probabilidade de ocorrerem valores de $\hat{p}$ mais desfavoráveis para H_0 do que esse. É evidente que quanto menor for $\hat{p}$, maior será a evidência contra $H_0 : p = 0,60$. Assim, calculemos

$$P(\hat{p} < 0,52 \mid p = 0,60) = P\left(Z < \frac{\sqrt{200}(0,52 - 0,60)}{\sqrt{0,24}}\right)$$
$$= P(Z < -2,30) = 0,01 = 1\%.$$

Esse resultado mostra que, se a audiência do programa fosse de 60% realmente, a probabilidade de encontrarmos uma amostra de 200 famílias com 52% ou menos de audiência é de 1%. Isso sugere que, ou estamos diante de uma amostra rara de ocorrer, 1 em 100, ou então a hipótese formulada não é aceitável. Nesse caso, somos levados a essa segunda opção, ou seja, os dados da amostra sugerem que a hipótese H_0 deve ser rejeitada.

O procedimento está ilustrado na Figura 12.11. O valor-p do teste será $\hat{\alpha} = 0{,}01$.

Figura 12.11: Determinação do valor-p para o Exemplo 12.5.



Exemplo 12.6. Um antibiótico A traz em sua bula a seguinte citação: "Nas broncopneumonias, a ação antiinflamatória de A é colocada em evidência pelo estudo dos parâmetros ventilatórios em duplo-cego contra placebo. Durante o tratamento com A pode-se observar uma melhora significativa em relação ao placebo, da capacidade vital ($p < 0{,}05$) e o VEMS($p < 0{,}001$) e do débito respiratório máximo ($p < 0{,}001$)".

Esse exemplo ilustra o uso cada vez mais difundido em muitas áreas aplicadas do conceito de valor-p. As afirmações do tipo "$p < 0{,}05$" acima referem-se a esse conceito. Vale a pena comentar um pouco sobre "estudos duplo-cego", mencionados acima. Nesse tipo de estudo, um número n de indivíduos é dividido em dois grupos de tamanhos aproximadamente iguais; a seleção dos indivíduos que vão pertencer a cada grupo é aleatória. Os indivíduos de um grupo recebem o tratamento (o antibiótico A, no caso), e os do outro grupo recebem placebo (uma substância inóqua). Os pesquisadores que acompanham o experimento não sabem quem recebeu tratamento e quem recebeu placebo, o mesmo acontecendo com os pacientes, daí o nome duplo-cego.

Podemos considerar probabilidades de significância bilaterais. Um procedimento é tomar o valor-p bilateral como sendo igual a duas vezes o valor-p unilateral. Esta prática é razoável quando a distribuição da estatística do teste, sob H_0, for simétrica.

Exemplo 12.7. Uma companhia de serviços de ônibus intermunicipais planejou uma nova rota para servir vários locais situados entre duas cidades importantes. Um estudo preliminar afirma que a duração das viagens pode ser considerada uma v.a. normal, com média igual a 300 minutos e desvio padrão 30 minutos. As dez primeiras viagens realizadas nessa nova rota apresentaram média igual a 314 minutos. Esse resultado comprova ou não o tempo médio determinado nos estudos preliminares?

Passo 1. Indicando por X a duração de cada viagem e por $\mu = E(X)$, queremos testar

$$H_0 : \mu = 300,$$
$$H_1 : \mu \neq 300.$$

Passo 2. Amostras de dez viagens terão média $\bar{X} \sim N(\mu, \sigma^2/10)$.

Passo 3. Sob a hipótese de que H_0 é verdadeira, e pelo fato de σ^2 ser conhecido (σ = 30), teremos

$$\bar{X} \sim N(300, 900/10).$$

Passo 4. Como o valor observado $\bar{x}_0 = 314$, podemos encontrar a probabilidade de ocorrerem amostras com valores de $\bar{X}$ mais extremos do que esse:

$$P(\bar{X} > 314) = P\left(Z > \frac{314 - 300}{9{,}49}\right) = P(Z > 1{,}48) = 0{,}07.$$

Como a distribuição de $\bar{X}$ é normal, portanto simétrica, tomamos $\hat{\alpha} = 0{,}14$. Nosso problema consiste em decidir se essa probabilidade corresponde ou não à chance de ocorrer um evento raro. Por ser uma probabilidade não muito pequena, podemos concluir que não existe muita evidência para rejeitar H_0. Assim, os estudos preliminares parecem estar corretos.

Um problema que pode ocorrer com o procedimento acima, de dobrar a probabilidade, é que o valor de $\hat{\alpha}$ pode ser maior do que um. Por isso, às vezes é preferível anunciar o valor do valor-*p* unilateral e a direção segundo a qual a observação afasta-se de H_0. No exemplo, o resultado indica que a chance de ocorrerem amostras com médias iguais ou superiores a 314 é 7%, que é um valor ainda não pequeno. Para outro método, ver o Problema 43.

Se indicarmos genericamente por $\hat{\alpha}$ o valor-*p*, rejeitaremos H_0 para aqueles níveis de significância α maiores do que $\hat{\alpha}$. No Exemplo 12.7, rejeitaremos H_0, por exemplo, se $\alpha = 0{,}10$, mas não a rejeitaremos se $\alpha = 0{,}05$ ou $\alpha = 0{,}01$. Ou seja, se o nível descritivo for muito pequeno, como o caso $\hat{\alpha} < 0{,}01$ do Exemplo 12.6, há evidências de que a hipótese não seja válida. Como vimos nesse exemplo, a probabilidade de significância é muitas vezes denotada por *p* na literatura (*p*-value).

Em nosso procedimento de testar uma hipótese estamos usando uma escala de evidências sugerida por Fisher (1954). Suponha que estejamos testando H_0 contra H_1 e, como vimos, rejeitamos H_0 se o valor-*p* $\hat{\alpha}$ for "bastante pequeno". A Tabela 12.2, extraída de Efron e Gous (1997), ilustra a escala de Fisher, contra H_0 (ou a favor de H_1).

Tabela 12.2: Escala de significância de Fisher.

valor-*p*	0,10	0,05	0,025	0,01	0,005	0,001
Natureza da evidência	marginal	moderada	substancial	forte	muito forte	fortíssima

Assim, um valor de $\hat{\alpha} = 0{,}01$ indica uma evidência forte contra a validade de H_0, $\hat{\alpha} = 0{,}05$ indica uma evidência moderada etc. É interessante notar que Fisher tomou como ponto de referência o valor 0,05: valores do valor-*p* menores do que 0,05 indicam que devemos rejeitar a hipótese nula. As considerações feitas por Fisher referiam-se a testes do qui-quadrado (veja o Capítulo 14).

Problemas

16. Suponha que queiramos testar $H_0 : \mu = 50$ contra $H_1 : \mu > 50$, onde μ é a média de uma normal $N(\mu, 900)$. Extraída uma amostra de $n = 36$ elementos da população, obtemos $\bar{x} = 52$. Calcule o valor-*p* $\hat{\alpha}$ do teste.

17. Os novos operários de uma empresa são treinados a operarem uma máquina, cujo tempo X (em horas) de aprendizado é anotado. Observou-se que X segue de perto a distribuição $N(25, 100)$. Uma nova técnica de ensino, que deve melhorar o tempo de aprendizado, foi testada em 16 novos empregados, o quais apresentaram 20,5 horas como tempo médio de aprendizado. Usando o valor-*p*, você diria que a nova técnica é melhor que a anterior?

12.9 Teste para a Variância de uma Normal

Um teste sobre a variância desconhecida de uma variável, com distribuição normal, irá usar a distribuição qui-quadrado, introduzida na seção 7.6.

Considere a média amostral $\bar{X}$ e a variância amostral S^2, ambas obtidas de uma amostra de tamanho n, $(X_1, ..., X_n)$ de $X \sim N(\mu, \sigma^2)$. A soma

$$\left(\frac{X_1 - \mu}{\sigma}\right)^2 + \dots \left(\frac{X_n - \mu}{\sigma}\right)^2$$

terá distribuição $\chi^2\,(n)$, pois cada $(X_i - \mu)/\sigma$ terá distribuição $N(0,1)$. Logo, se definirmos

$$\hat{\sigma}^2_* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2, \tag{12.1}$$

vemos que

$$Y = \frac{n\hat{\sigma}^2_*}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \tag{12.2}$$

tem distribuição $\chi^2(n)$. Observe que o estimador $\hat{\sigma}^2_*$ é muito parecido com o estimador $\hat{\sigma}^2$, definido em (11.6), com μ tomando o lugar de $\bar{X}$. É muito importante conhecer a distribuição de $\sum_{i=1}^{n}(X_i - \bar{X})^2$, para se ter a distribuição de S^2, que será usada no teste desta seção. Note inicialmente que

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}\{(X_i - \bar{X}) + (\bar{X} - \mu)\}^2$$

$$= \sum_{i=1}^{n}(X_i - \bar{X})^2 + 2(\bar{X} - \mu)\sum_{i=1}^{n}(X_i - \bar{X}) + n(\bar{X} - \mu)^2,$$

e de $\sum_i (X_i - \bar{X}) = 0$, vem que

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 + n(\bar{X} - \mu)^2. \tag{12.3}$$

Dividindo ambos os membros por σ^2, e reescrevendo (12.3) de forma conveniente, teremos

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2=\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2+\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\right)^2. \tag{12.4}$$

O primeiro membro da expressão (12.4) tem distribuição $\chi^2(n)$, como vimos acima. O último termo de (12.4) tem distribuição $\chi^2(1)$. Seria, então, razoável supor que o primeiro termo do segundo membro tenha distribuição $\chi^2(n-1)$. A comprovação desse fato exige recursos fora do alcance deste livro, mas podemos resumir o resultado da seguinte maneira.

Teorema 12.1. Seja $(Z_1, ..., Z_n)$ uma amostra aleatória simples retirada de uma população $N(0,1)$. Então:

(i) $\bar{Z}$ tem distribuição $N(0,1/n)$;

(ii) as variáveis $\bar{Z}$ e $\sum_{i=1}^{n}(Z_i-\bar{Z})^2$ são independentes; e

(iii) $\sum_{i=1}^{n}(Z_i-\bar{Z})^2$ tem distribuição $\chi^2(n-1)$.

Corolário 12.1. A variável aleatória $(n-1)S^2/\sigma^2$ tem distribuição $\chi^2(n-1)$.

Prova. De fato,

$$\frac{(n-1)S^2}{\sigma^2}=\frac{n-1}{\sigma^2}\,\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2=\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2=\sum_{i=1}^{n}(Z_i-\bar{Z})^2,$$

bastando escrever $(X_i-\bar{X})/\sigma=(X_i-\mu)/\sigma-(\bar{X}-\mu)/\sigma$.

A expressão (12.4) e a própria definição de χ^2 garantem uma propriedade muito útil: a soma de duas v.a. independentes, cada uma com distribuição χ^2, é uma v.a. também com distribuição χ^2:

$$\chi^2(p)+\chi^2(q)=\chi^2(p+q).$$

Voltemos ao nosso problema original. Queremos testar

$$H_0: \sigma^2=\sigma_0^2,$$
$$H_1: \sigma^2\neq\sigma_0^2.$$

Nossas suposições são que $X_i \sim N(\mu, \sigma^2)$, $i = 1, ..., n$ e os X_i são independentes. A estatística do teste será, sob H_0,

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1). \tag{12.5}$$

Como temos um teste bilateral, a região crítica será da forma $RC=(0, \chi_1^2] \cup [\chi_2^2, +\infty)$, tal que

$$P(\chi^2\in RC|H_0)=P(0<\chi^2<\chi_1^2 \text{ ou } \chi^2>\chi_2^2)=\alpha,$$

sendo α o nível de significância do teste, fixado *a priori*.

Observado o valor s_0^2 da estatística S^2, obteremos o valor $\chi_0^2 = \dfrac{(n-1)s_0^2}{\sigma_0^2}$. Se $\chi_0^2 \in$ RC, rejeitamos H_0; caso contrário, aceitamos H_0.

Exemplo 12.8. Uma das maneiras de manter sob controle a qualidade de um produto é controlar sua variabilidade. Uma máquina de encher pacotes de café está regulada para enchê-los com média de 500 g e desvio padrão de 10 g. O peso de cada pacote X segue uma distribuição $N(\mu, \sigma^2)$. Colheu-se uma amostra de 16 pacotes e observou-se uma variância de $S^2 = 169$ g². Com esse resultado, você diria que a máquina está desregulada com relação à variância?

Estamos interessados em testar, então,

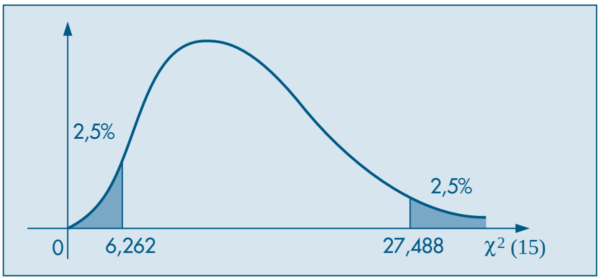
$$H_0 : \sigma^2 = 100,$$
$$H_1 : \sigma^2 \neq 100.$$

A estatística para realizar o teste é (12.5), com $n = 16$. Fixado o nível de significância α em 5%, teremos da Tabela IV que a região crítica é dada por RC $= \{\chi^2 : 0 \leqslant \chi^2 \leqslant 6{,}262$ ou $\chi^2 \geqslant 27{,}488\}$. Veja a Figura 12.12. O valor observado da estatística é

$$\chi_0^2 = \frac{(n-1)s_0^2}{\sigma_0^2} = \frac{(15)(169)}{100} = 25{,}35.$$

Como $\chi_0^2 \notin$ RC, somos levados a aceitar H_0, isto é, a máquina está sob controle quanto à variância.

Figura 12.12: Região crítica para o teste do Exemplo 12.8.



A construção do IC(σ^2; γ) é feita a partir da expressão

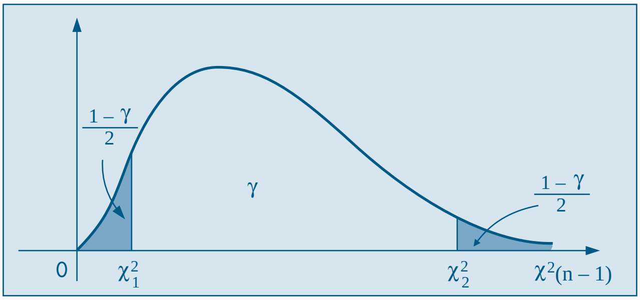
$$P\left(\chi_1^2 \leqslant \frac{(n-1)S^2}{\sigma^2} \leqslant \chi_2^2\right) = \gamma, \tag{12.6}$$

que permite obter a seguinte desigualdade:

$$\frac{(n-1)S^2}{\chi_2^2} \leqslant \sigma^2 \leqslant \frac{(n-1)S^2}{\chi_1^2}, \tag{12.7}$$

que será o IC procurado. Veja a Figura 12.13.

Figura 12.13: Valores críticos para a construção de um intervalo de confiança para a variância.



Exemplo 12.9. Os dados abaixo referem-se às vendas diárias, em reais, durante uma semana, de carros de uma revendedora. Construir um IC(σ^2; 90%).

$$\text{Vendas} : 253,\ 187,\ 96,\ 450,\ 320,\ 105.$$

Inicialmente, calculamos a variância amostral, que é $s_0^2 = 18.460$; em seguida, os valores χ_1^2 e χ_2^2 que satisfaçam (12.6):

$$P(1{,}145 \leqslant \chi^2(5) \leqslant 11{,}070) = 0{,}90.$$

Substituindo em (12.7) obtemos

$$\text{IC}(\sigma^2;\ 0{,}90) = [8.338;\ 80.611].$$

Problemas

18. De uma população $X \sim N(50, 100)$ retira-se uma amostra de dez elementos e calculam-se os valores de $\hat{\sigma}_*^2$ e S^2. Encontre os valores pedidos abaixo, com a maior precisão possível.
   (a) Se $P(\hat{\sigma}_*^2 > a) = 10\%$, encontre o valor de a.
   (b) Sabendo-se que $P(S^2 < a) = 5\%$ e $P(S^2 > b) = 5\%$, encontre a e b.
   (c) $P(S^2 < 163{,}16) = \alpha$, encontre α.
   (d) $P(S^2 > 100) = \alpha$, encontre α.
   (e) $P(S^2 < 18) = \alpha$, encontre α.
   (f) Se o valor observado de S^2 foi 180, qual a probabilidade de encontrar uma amostra que produza um S^2 maior do que o observado?

19. Observou-se a produção mensal de uma indústria durante vários anos, verificando-se que ela obedecia a uma distribuição normal, com variância 300. Foi adotada uma nova técnica de produção e, durante 24 meses, observou-se a produção mensal. Após esse período, constatou-se que $\bar{x} = 10.000$ e $s^2 = 400$. Há razões para se acreditar que a variância mudou, ao nível de 20%?

20. Numa linha de produção, é muito importante que o tempo gasto numa determinada operação não varie muito de empregado para empregado.
   (a) Que parâmetro estatístico poderia ser usado para avaliar esse fato? Por quê?

(b) Se 11 empregados apresentam os tempos abaixo para realizar essa operação, qual seria a estimativa para a parâmetro acima?

125	135	115	120	150	130
125	145	125	140	130	

12.10 Teste sobre a Média de uma Normal com Variância Desconhecida

Vimos, na seção 12.5, como testar a média de uma normal, supondo que a variância seja conhecida. Comentamos que essa não é uma suposição realista, logo iremos supor agora que temos uma v.a. X, com distribuição normal, com média μ e variância σ^2 desconhecidas.

No Capítulo 7 introduzimos a distribuição t de Student. Veremos, a seguir, como ela pode ser usada para testar hipóteses sobre μ nessa situação.

Consideremos a estatística

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}. \qquad (12.8)$$

Inicialmente, dividamos o numerador e denominador pelo desvio padrão σ da população, e teremos

$$\frac{((\sqrt{n}(\bar{X}-\mu)/\sigma)}{(S/\sigma)}$$

O numerador $Z = (\sqrt{n}\,(\bar{X}-\mu))/\sigma$ tem distribuição $N(0, 1)$, como já foi visto. O quadrado do denominador pode ser escrito como

$$\frac{(n-1)S^2}{\sigma^2}\Big/(n-1) = \frac{Y}{n-1},$$

onde $Y = (n-1)S^2/\sigma^2$. Mas, como foi visto na seção anterior, se os X_i forem normais, Y tem distribuição $\chi^2(n-1)$; logo, a estatística (12.8) é o quociente entre uma v.a $N(0, 1)$ e a raiz quadrada de uma v.a $\chi^2(n-1)$, dividida pelo número de graus de liberdade, e pelo Teorema 7.1 temos que

$$\frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t(n-1). \qquad (12.9)$$

Observe que Z e Y são independentes, pois $\bar{X}$ e S^2 são independentes, pelo Teorema 12.1 (ii).

Estamos, agora, em condições de testar as hipóteses

$$H_0 : \mu = \mu_0,$$
$$H_1 : \mu \neq \mu_0.$$

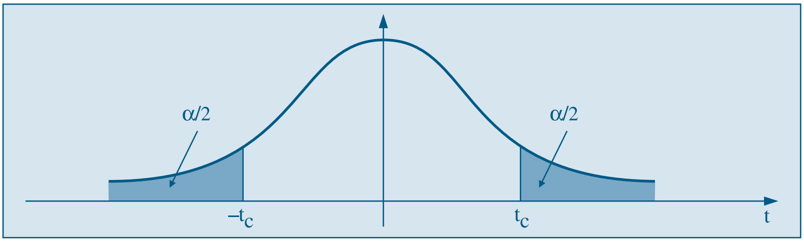
A hipótese alternativa poderia ser $\mu > \mu_0$ ou $\mu < \mu_0$, o que mudaria apenas a região de rejeição de bilateral para unilateral (à direita ou à esquerda, respectivamente).

A estatística a ser usada é

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{S}, \tag{12.10}$$

que sabemos agora ter uma distribuição t de Student com $(n - 1)$ graus de liberdade. Fixado o valor de α, podemos usar a Tabela V e encontrar o valor t_c, tal que $P(|T| < t_c) = 1 - \alpha$. Veja a Figura 12.14.

Figura 12.14: Valores críticos para o teste t.



Colhida a amostra de n indivíduos, calculamos os valores $\bar{x}_0$ e s_0^2 das estatísticas $\overline{X}$ e S^2, respectivamente, e depois o valor $t_0 = \sqrt{n}(\bar{x}_0 - \mu_0)/s_0$ de T. Se o valor dessa estatística for inferior a $-t_c$, ou superior a t_c, rejeita-se H_0. Caso contrário, aceita-se H_0.

Para a construção de intervalos de confiança, temos que

$$P\left(-t_\gamma < \frac{\sqrt{n}(\overline{X} - \mu)}{S} < t_\gamma\right) = \gamma,$$

da qual segue o intervalo de confiança

$$\text{IC}(\mu; \gamma) = \overline{X} \pm t_\gamma \frac{S}{\sqrt{n}}, \tag{12.11}$$

muito parecido com aquele da variância conhecida.

Exemplo 12.10. Um fabricante afirma que seus cigarros contêm não mais que 30 mg de nicotina. Uma amostra de 25 cigarros fornece média de 31,5 mg e desvio padrão de 3 mg. No nível de 5%, os dados refutam ou não a afirmação do fabricante?

Passo 1. As hipóteses aqui são:

$$H_0 : \mu = 30,$$
$$H_1 : \mu > 30.$$

Passo 2. Supondo que X, a quantidade de nicotina por cigarro, tenha distribuição $N(\mu, \sigma^2)$, a estatística

$$T = \frac{\sqrt{25}(\overline{X} - 30)}{S}$$

terá distribuição $t(24)$.

Passo 3. Por ser um teste unilateral, devemos procurar o valor t_c tal que

$$P(T > t_c) = 0{,}05.$$

Da Tabela V, obtemos t_c = 1,711, ou seja, a região crítica para a estatística T é RC=[1,711; +∞[.

Passo 4. O valor observado da estatística é

$$t_0 = \frac{5(31{,}5 - 30)}{3} = 2{,}5.$$

Passo 5. Como t_0 pertence à região crítica, rejeitamos H_0, ou seja, há evidências de que os cigarros contenham mais de 30 g de nicotina.

Outra maneira de proceder é calcular o valor-p, ou seja,

$$\hat{\alpha} = P(T > t_0|H_0) = P(T > 2{,}5|H_0) = 0{,}01.$$

Esse valor pequeno de $\hat{\alpha}$ leva à rejeição de H_0.

Para construir um IC(μ; 0,95), verificamos na Tabela V que o valor t_γ = 2,064 e, portanto,

$$\text{IC}(\mu;\ 0{,}95) = 31{,}5 \pm (2{,}064)\ 3/\sqrt{25},$$

ou seja,

$$\text{IC}(\mu;\ 0{,}95) =]30{,}26;\ 32{,}74[.$$

Antes de encerrar este capítulo cabe uma observação. Quando aceitamos uma hipótese, estamos concluindo que temos algum conhecimento sobre a distribuição da variável de interesse. Já quando rejeitamos a hipótese, a distribuição da variável não fica especificada. A construção de intervalos de confiança desempenha um papel importante nessa situação. Ressaltamos, também, que temos usado a expressão “aceitamos” a hipótese, quando o mais correto talvez fosse “não rejeitamos” a hipótese.

Problemas

21. Da população $X \sim N(50, 100)$ retirou-se uma amostra casual simples de tamanho $n = 10$, calculando-se o valor de $\overline{X}$, S e o respectivo valor de t.
    (a) Se $P(|\overline{X} - 50| < tS/\sqrt{10}) = 90\%$, encontre o valor de t.
    (b) Se $\overline{X} = 48$ e $S^2 = 120$, qual a probabilidade de encontrar um valor de t menor que o produzido por essa amostra?
    (c) Se $S^2 = 120$, calcule a $P(|\overline{X} - 50| < 2)$.

22. O tempo médio, por operário, para executar uma tarefa, tem sido 100 minutos, com um desvio padrão de 15 minutos. Introduziu-se uma modificação para diminuir esse tempo, e, após certo período, sorteou-se uma amostra de 16 operários, medindo-se o tempo de execução de cada um. O tempo médio da amostra foi 85 minutos, e o desvio padrão foi 12 minutos. Estes resultados trazem evidências estatísticas da melhora desejada? Em caso

afirmativo, estime o novo tempo médio de execução. (Apresente as suposições teóricas usadas para resolver o problema.)

23. Estamos desconfiados de que a média das receitas municipais *per capita* das cidades pequenas (0 – 20.000 habitantes) é maior do que a das receitas do estado, que é de 1.229 unidades. Para comprovar ou não essa hipótese, sorteamos dez cidades pequenas, e obtivemos os seguintes resultados: 1.230; 582; 576; 2.093; 2.621; 1.045; 1.439; 717; 1.838; 1.359.

    Obs.: Para facilitar os cálculos, informamos que a soma das observações é 13.500, e a soma dos quadrados das observações é 22.335.650 ($13.500^2 = 182.250.000$).

    (a) Mostre que o teste de hipótese usado, com $\alpha = 0{,}05$, levará à aceitação de que a média das cidades pequenas é igual à do estado.

    (b) Você não acha estranha essa conclusão quando observa que a média da amostra obtida é bem maior do que a média do estado? Como você explicaria isso?

24. Deseja-se estimar qual a porcentagem média da receita familiar gasta com alimentação pelos moradores de uma grande vila industrial. Para isso, selecionou-se uma amostra de 16 famílias, que apresentou os seguintes resultados:

41	44	35	42	34	22	42	42
38	62	29	63	38	45	48	40

    (a) Dê um IC de 95% para a porcentagem média de todas as famílias de moradares da vila.

    (b) Que suposição você fez para responder a pergunta anterior?

12.11 Problemas e Complementos

25. A precipitação pluviométrica anual numa certa região tem desvio padrão $\sigma = 3{,}1$ e média desconhecida. Para os últimos 9 anos, foram obtidos os seguintes resultados: 30,5; 34,1; 27,9; 35,0; 26,9; 30,2; 28,3; 31,7; 25,8.

    (a) Construa um teste de hipóteses para saber se a média da precipitação pluviométrica anual é maior que 30,0 unidades. Utilize um nível de significância de 5%.

    (b) Discuta o mesmo problema, considerando σ desconhecido.

    (c) Supondo que, na realidade, $\mu = 33{,}0$, qual a probabilidade de tirarmos uma conclusão errada?

26. Supõe-se que determinado tipo de indústria deva ter, em média, 30 empregados. Para testar tal hipótese, colhe-se uma amostra de 50 indústrias, cujo resultado está abaixo. Caso rejeite a hipótese, dê um intervalo de confiança para a verdadeira média (suponha que $s^2 = \sigma^2$).

Nº de empregados	Freqüência
25 ⊢ 35	8
35 ⊢ 45	10
45 ⊢ 55	13
55 ⊢ 65	10
65 ⊢ 75	9

27. Uma fábrica de automóveis anuncia que seus carros consomem, em média, 11 litros por 100 km, com desvio padrão de 0,8 litro. Uma revista resolve testar essa afirmação e

analisa 35 automóveis dessa marca, obtendo 11,3 litros por 100 km como consumo médio (considerar distribuição normal). O que a revista pode concluir sobre o anúncio da fábrica, no nível de 10%?

28. Um dos maiores problemas de uma grande rede de vendas a varejo é a adequação do estoque declarado com o real existente. Decidiu-se fazer a verificação através de procedimentos amostrais. Indicando por X o total em unidades monetárias de cada produto em estoque, verificou-se que $X \sim N(\mu, 400)$. Serão sorteados 4 produtos. O total X de cada um será verificado e calcular-se-á a média $\overline{X}$, que será a estatística de decisão. Numa determinada filial, o valor declarado de μ é 50. Havendo falta, esse parâmetro deve ser 45; no caso de excesso, 58.
   (a) Defina H_0 e H_1.
   (b) Descreva os erros do tipo I e II.
   (c) Fixando $\alpha = 10\%$, qual a regra de decisão para julgar se o estoque está correto ou não?
   (d) Calcule o erro β.
   (e) Qual o significado de α e β nesse problema?

29. Seja X uma v.a. com distribuição binomial, com $n = 15$. Considere $H_0 : p \geqslant 0{,}5$ contra $H_1 : p < 0{,}5$, com RC = $\{0, 1, 2\}$.
   (a) Calcule a probabilidade do erro de tipo I.
   (b) Calcule a probabilidade do erro de tipo II quando $p = 0{,}3$.
   (c) Esboce o gráfico do poder do teste.

30. O custo X de manutenção de teares segue uma distribuição normal, $X \sim N(\mu, 400)$. Durante muito tempo, o parâmetro μ tem sido adotado como igual a 200. Suspeita-se que esse parâmetro aumentou, e só nos interessa saber se o novo parâmetro superior a 210. Assim, queremos planejar um teste em que $\alpha = 5\%$ (quando $\mu = 200$) e $\beta = 10\%$ (quando $\mu = 210$).
   (a) Qual deve ser o tamanho da amostra?
   (b) Qual a RC nesse caso?

31. O número médio diário de clientes de um posto de gasolina tem sido 250, com um desvio padrão de 80 clientes. Durante uma campanha de 25 dias, em que os clientes recebiam um brinde, o número médio de clientes foi 280, com um desvio padrão de 50. Você diria que a campanha modificou a distribuição do número de clientes do posto? Descreva as suposições feitas para a resolução do problema.

32. A receita média, em porcentagem, dos quase 600 municípios de um estado tem sido 7%. O governo pretende melhorar esse índice e, para isso, está estudando alguns incentivos. Para verificar os efeitos desses incentivos, sorteou 10 cidades e estudou quais seriam as porcentagens investidas neles. Os resultados foram, em porcentagem, 8, 10, 9, 11, 8, 12, 16, 9, 12, 13. Admitindo-se que esses números realmente venham a ocorrer, os dados trazem evidência de melhoria? Caso altere a média do estado, dê um intervalo de confiança para a nova média.

33. Para o problema anterior, construa IC(σ^2; 90%) e descreva as suposições consideradas para obtenção da resposta.

34. A prefeitura de uma cidade quer estimar a proporção p dos moradores favoráveis à mudança do horário comercial, com o intuito de economizar combustível. Essa proporção deverá ser estimada com um erro máximo de 5%, a um nível de 90% de confiança.

(a) Que tamanho deverá ter a amostra se a proporção p esperada deve estar entre 20% e 50%? (Justifique a resposta.)

(b) Numa amostra de 400 moradores, 160 foram favoráveis à mudança; qual seria o intervalo de confiança para p, nesse caso, com $\gamma = 0,95$?

35. Numa pesquisa realizada com 2.000 proprietários de carros na cidade de São Paulo, 800 responderam que pretendem mudar de carro no decorrer do próximo ano. Dê um IC de 90% para a proporção de todos os proprietários de carros de São Paulo que pretendem mudar de carro no próximo ano.

36. Um fabricante de um certo tipo de aço especial afirma que seu produto tem um severo serviço de controle de qualidade, traduzido pelo desvio padrão da resistência à tensão, que não é maior do que 5 kg por cm^2. Um comprador, querendo verificar a veracidade da afirmação, tomou uma amostra de 11 cabos e submeteu-a a um teste de tensão. Os resultados foram os seguintes: $\bar{x} = 263$ e $s^2 = 48$. Estes resultados trazem alguma evidência contra a afirmação do fabricante? Use $\alpha = 0,05$.

37. Um escritório de investimento acredita que o rendimento das diversas ações movimentadas por ele foi de 24%. Mais ainda, a nova estratégia definida deve garantir uma maior uniformidade nos rendimentos das diversas ações. No passado, o desvio padrão do rendimento era da ordem de 5%. Para verificar as duas hipóteses, tomaram-se 8 empresas ao acaso, obtendo-se os seguintes rendimentos (dados em %): 23,6; 22,8; 25,7; 24,8; 26,4; 24,3; 23,9 e 25. Quais seriam as conclusões?

38. Sendo X o número de sucessos em $n = 10$ provas de Bernoulli, queremos testar $H_0 : p = 0,6$.

(a) Se o teste for unilateral e rejeitarmos H_0 para valores pequenos de X, determine $\hat{\alpha}$ se o valor observado de X for 3.

(b) Determine $\hat{\alpha}$ se o teste for bilateral, na situação de *(a)*, isto é, $X = 3$.

39. Considere a situação do problema anterior e suponha que o valor observado seja $X = 6$. O que acontece no caso *(b)* do problema anterior? O resultado $X = 6$ suporta ou não H_0?

40. Valor-*p* bilateral. Vimos no texto um procedimento para determinar $\hat{\alpha}$ no caso bilateral. Outra possibilidade é fazer as probabilidades nas duas caudas complementares em termos da distância à média (ou mediana) da distribuição sob H. Assim, se x for o valor observado de X e m for a média da distribuição, colocamos

$$\hat{\alpha} = P(X \geqslant x) + P(X \leqslant m - (x - m)),$$

se x estiver na cauda superior e

$$\hat{\alpha} = P(X \leqslant x) + P(X \geqslant m + (m - x)),$$

se x estiver na cauda inferior.

Calcule $\hat{\alpha}$ usando esse critério para os Problemas 41 e 42.