Aula de Exercícios - Testes de Hipóteses

Organização: Airton Kist *Digitação*: Guilherme Ludwig

Testes de Hipóteses

Exemplo

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 $\sigma = 10$.
Defina: *Erro do tipo I* – dizer que os habitantes da ilha são descendentes de **B** quando, na realidade, são de **A**.
Erro do tipo II – dizer que são de **A** quando são de **B**.

Testes de Hipóteses

Exemplo

(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 for $\mu_B = 178$? E $\mu_B = 180$? E $\mu_B = 181$? Coloque num gráfico os pares $(\mu_B, P(\text{Erro II}|\mu_b))$

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 329.

Testes de Hipóteses

(a) Note que H_0 : moradores são de A. Nossa região crítica é dada por $RC = \{\bar{X} > 176\}$, a região em que rejeitamos H_0. Note também que $\text{Var}(\bar{X})$ é σ^2/n. Então

$$P(\text{erro tipo I}) = P(\bar{X} > 176|\mu = 175) =$$

$$P\left(Z > \sqrt{100}\frac{176-175}{10}|\mu = 175\right) = 1 - \Phi(1) = 0.159$$

De modo análogo, a probabilidade do erro de tipo II é dada por

$$P(\text{erro tipo II}) = P(\bar{X} \leq 176|\mu = 175) =$$

$$P\left(Z \leq \sqrt{100}\frac{176-175}{10}|\mu = 175\right) = \Phi(-1) = 0.159$$

Testes de Hipóteses

(b) Queremos fixar c para que $P(\bar{X} > c|\mu{=}175) = 0.05$. Para isto, basta que

$$P(\bar{X} > c|\mu{=}175) = P\left(Z > \sqrt{100}\frac{c-175}{10}\right) = 1 - \Phi(c - 175) = 0.05$$

Note que $\Phi(z) = 0.95 \Leftrightarrow z = 1.64$, então temos que $(c - 175) = 1.64$ ou $c = 176.64$. Nossa regra de decisão agora é classificar o indivíduo como descendente de B se sua altura for superior a 176.64. Note que, agora, temos um erro do tipo II de

$$P(\bar{X} \leq 176.64|\mu = 177) = \Phi\left(\sqrt{100}\frac{176.64 - 177}{10}\right) = 0.359$$

Testes de Hipóteses

(c) Se $\sigma_A = 5$, o valor de c para que o erro do tipo I seja 0.05 é

$$P\left(Z > \sqrt{100}\frac{c - 175}{5}\right) = 1 - \Phi(2(c - 175)) = 0.05$$

Que é dado por $2(c - 175) = 1.64$ ou $c = 175.82$. O novo erro do tipo II agora é

$$P(\bar{X} \leq 175.82 | \mu = 177) = \Phi\left(\sqrt{100}\frac{175.82 - 177}{10}\right) = 0.119$$

Testes de Hipóteses

(d)

- Se $\mu_B = 178$, então

$$P(\bar{X} \leq 175.82|\mu = 178) = \Phi\left(\sqrt{100}\frac{175.82 - 178}{10}\right) = 0.014$$

- Se $\mu_B = 180$, então

$$P(\bar{X} \leq 175.82|\mu = 180) = \Phi\left(\sqrt{100}\frac{175.82 - 180}{10}\right) = 1.45 \times 10^{-5}$$

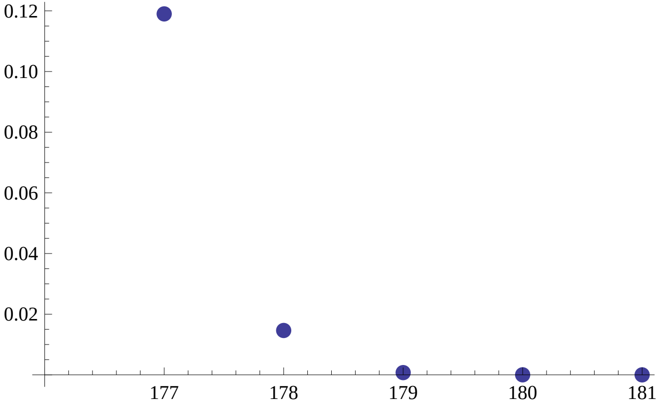
- Se $\mu_B = 181$, então

$$P(\bar{X} \leq 175.82|\mu = 181) = \Phi\left(\sqrt{100}\frac{175.82 - 181}{10}\right) = 1.11 \times 10^{-7}$$

Note que $\beta(\mu) = 1 - P(\text{erro tipo II}|\mu)$ é chamado de *função poder*.

Testes de Hipóteses

(d) O gráfico dos pares $(\mu_B, P(\text{Erro II}|\mu_b))$ é



Testes para média, σ conhecido

Exemplo

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 perdidos por acidentes, que foi de 50 horas. Você diria, no nível de 5%, que há evidência de melhoria?
Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 334.

Testes para média, σ conhecido

Queremos testar a hipótese que μ, o número médio de horas perdidas com acidentes de trabalho, tenha permanecido o mesmo. Ou seja, $H_0 : \mu = 60$ vs. $H_1 : \mu < 60$. Note que sob H_0, $\bar{X} \sim N(60, 400/9)$

Com um nível $\alpha = 0.05$, temos que a hipótese será rejeitada se $3(\bar{X} - 60)/20 < c$. Para a normal padrão, $P(Z < c) = 0.05 \Leftrightarrow c = -1.64$. Então a região crítica é $3(\bar{X} - 60)/20 < -1.64$ ou simplesmente $\bar{X} < 49.06$.

Como a média observada $\bar{x} = 50$ é superior a 49.06, não rejeitamos a hipótese nula a 5% de significância.

Testes para média, σ conhecido

Exemplo

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 seis 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? *Fonte: Morettin & Bussab*, Estatística Básica 5ª edição, pág 334.

Testes para média, σ conhecido

Queremos testar a hipótese que μ, o índice médio de nicotina dos cigarros, seja maior que 23. Ou seja, $H_0 : \mu \geq 23$ vs. $H_1 : \mu < 23$.

Com um nível $\alpha = 0.10$, temos que a hipótese será rejeitada se $\sqrt{6}(\bar{X} - 23)/\sqrt{4.86} < c$. Para a normal padrão, $P(Z < c) = 0.10 \Leftrightarrow c = -1.28$. Então a região crítica é $\sqrt{6}(\bar{X} - 23)/\sqrt{4.86} < -1.28$ ou simplesmente $\bar{X} < 21.85$.

Como a média observada é $\bar{x} = 24.16$ é superior a 21.85, não rejeitamos a hipótese nula a 10% de significância.

Testes para Proporções

Exemplo

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. Moste como o fabricante poderia refutar a acusação. Utilize um nível de significância de 10%.
Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 337.

Testes para Proporções

O fabricante não quer rejeitar a hipótese $p = 0.2$ em favor da hipótese $p > 0.2$. A região crítica é, portanto, da forma $\hat{p} > c$. Utilizando a aproximação normal, temos que c, sob H_0, é dado por

$$1-\Phi\left(\sqrt{50}\frac{c-0.2}{\sqrt{0.2\cdot 0.8}}\right) = 0.1 \Leftrightarrow c = 1.28\frac{\sqrt{0.2\cdot 0.8}}{\sqrt{50}}+0.2 = 0.2724$$

Como podemos ver, a proporção de itens defeituosos obtida pelo consumidor não é significativamente diferente da probabilidade de 20% anunciada pelo vendedor, a 10% de significância, pois não é superior a 0.2724 (contra 0.27 observado).

Testes para Proporções

Exemplo

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%.

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 337.

Testes para Proporções

A proporção observada de equipamentos sem defeito é de 0.875. A hipótese a ser testada é $H_0 : p = 0.9$ contra $H_1 : p < 0.9$. A região crítica é, portanto, da forma $\hat{p} < c$. A 5% de significância, temos que c é dado por

$$\Phi\left(\sqrt{200}\frac{c-0.9}{\sqrt{0.1\cdot 0.9}}\right) = 0.05 \Leftrightarrow c = -1.64\frac{\sqrt{0.1\cdot 0.9}}{\sqrt{200}} + 0.9 = 0.8652$$

Conclui-se que, a 5% de significância, a hipótese nula não é rejeitada.

Testes para Proporções

A 1% de significância, temos que c é dado por

$$\Phi\left(\sqrt{200}\frac{c-0.9}{\sqrt{0.1\cdot 0.9}}\right)=0.01 \Leftrightarrow c=-2.32\frac{\sqrt{0.1\cdot 0.9}}{\sqrt{200}}+0.9=0.8507$$

Conclui-se que, a 1% de significância, a hipótese nula não é rejeitada.

P-Valor

Exemplo

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 a probabilidade de significância $\hat{\alpha}$ do teste.

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 343.

P-Valor

A probabilidade de significância $\hat{\alpha}$ é mais comumente conhecida como *p-valor*. Observe que sob H_0, $\bar{X} \sim N(50, 25)$. Temos que $\bar{x} = 52$. A probabilidade de significância (ou p-valor) é obtida calculando-se a probabilidade do valor observado na estatística do teste, ou seja,

$$P\left(\bar{X} > 52\right) = \left(Z > \frac{52-50}{5}\right) = 1 - \Phi(2/5) \approx 0.34$$

Devemos interpetar o p-valor como "observados os dados, quão verossímil é a hipótese nula?" e, neste caso, ela é bastante verossímil (a probabilidade de observarmos $\bar{X} > 52$, dado H_0 verdadeira, é de 0.34) e portanto não rejeitamos H_0.

P-Valor

Exemplo

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, os quais apresentaram 20.5 horas como tempo médio de aprendizado. Usando o p-valor, você diria que a nova técnica é melhor que a anterior?

Fonte: Morettin & Bussab, Estatística Básica 5^a edição, pág 344.

P-Valor

Defina a hipótese nula H_0 : tempo continua igual ($\mu = 25$) vs. $H_1 : \mu < 25$. Observe que sob H_0, $\bar{X} \sim N(25, 6.25)$. Temos que $\bar{x} = 20.5$. A probabilidade de significância (ou p-valor) é obtida calculando-se a probabilidade do valor observado na estatística do teste, ou seja,

$$P\left(\bar{X} < 20.5\right) = \left(Z < \frac{20.5 - 25}{2.5}\right) = \Phi(-1.8) \approx 0.036$$

Neste caso, o p-valor é de apenas 3.6%, o que nos diz que para qualquer nível de significância maior que 3.6%, rejeitamos a hipótese nula $\Rightarrow$ a nova metodologia é melhor que a anterior.

Testes para duas populações

Exemplo

Duas máquinas, A e B, são usadas para empacotar pó de café. A experiência passada garante que o desvio padrão para ambas é de $10g$. Porém, suspeita-se que elas tem médias diferentes. Para verificar, sortearam-se duas amostras: uma com 25 pacotes da máquina A e outra com 16 pacotes da máquina B. As médias foram, respectivamente, $\bar{x}_A = 502.74g$ e $\bar{x}_B = 496.60g$. Com esses números, e com nível de 5%, qual seria a conclusão do teste $H_0 : \mu_A = \mu_B$?

Fonte: Morettin & Bussab, Estatística Básica 5ª edição, pág 352.

Testes para duas populações

Defina D como $\bar{X} - \bar{Y}$. Temos que

$$\mathbb{E}(D) = \mathbb{E}(\bar{X} - \bar{Y}) = \mathbb{E}(\bar{X}) - \mathbb{E}(\bar{Y}) = \mu_A - \mu_B$$

e

$$\text{Var}(D) = \text{Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) - 2\text{Cov}(\bar{X}, \bar{Y})$$

mas $\text{Cov}(\bar{X}, \bar{Y}) = 0$ pois as máquinas são independentes. Então

$$\text{Var}(D) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) = \frac{100}{25} + \frac{100}{16} = \frac{41}{4}$$

Testes para duas populações

Agora note que a hiṕotese $H_0 : \mu_A = \mu_B$ é equivalente a $H_0 : \mu_A - \mu_B = 0$. Podemos testar essa hipótese através de D. Para o teste bilateral (isto é, $H_1 : \mu_A \neq \mu_B$), temos que a 5% de significância, $P(\{Z < -c\} \cup \{Z > c\}) = 0.05 \Leftrightarrow c = 1.96$, e portanto a região crítica do teste é dada por

$$\left|\bar{X} - \bar{Y}\right| > 1.96 \cdot 3.2 = 6.27$$

e como $\bar{x} - \bar{y} = 6.14$, não rejeitamos a hipótese nula.