Aula 8 - Testes de hipóteses

PhD. Wagner Hugo Bonat

Laboratório de Estatística e Geoinformação-LEG
Universidade Federal do Paraná

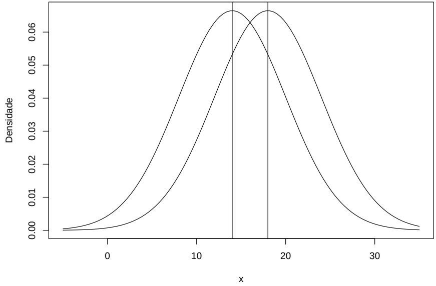
1/2017

Exemplo 8.1

- Suponha que, entre pessoas sadias, a concentração de certa substância no sangue se comporta segundo um modelo Normal com média 14 unidades/ml e desvio padrão 6 unidades/ml. Pessoas sofrendo de uma doença específica êm concentração méda da substância alterada para 18 unidades/ml. Admitimos que o modelo Normal com desvio padrão 6 unidades/ml, continua representado de forma adequada a concentração da substância em pessoas com a doença.

Exemplo 8.1

```
> dnorm1 <- function(x) {dnorm(x, mean = 14, sd = 6)}
> dnorm2 <- function(x) {dnorm(x, mean = 18, sd = 6)}
> curve(dnorm1, from = -5, to = 35, ylab = "Densidade")
> abline(v = 14)
> curve(dnorm2, from = -5, to = 35, add = TRUE)
> abline(v = 18)
```



Exemplo 8.2

- Deseja-se estudar a telerância de um equipamento eletrônico ao número de impactos termo-elétricos. Pelas características de fabricação do equipamento, é possível admitir que a probabilidade de falha seja constante, isto é, após cada impacto, existe uma probabilidade p que ele falhe. Representando por X a variável número de impactos anteriores à falaha, pretende-se verificar se o modelo Geométrico com $p = 0.4$ pe adequado para caracterizar essa variável.

Exemplo 8.1 continuação

- Interesse geral $\mu = 14$?
- Distribuição da média amostral $N(\mu, 36/30)$.
- Critério para decidir sobre o valor de μ.
- Valor crítico, digamos x_c tal que se a média for maior que x_c concluímos que a amostra pertence a população com média 18.
- Erros associados.

Hipóteses

- Hipótese simples:
  H_0 : O tratamento não é eficaz ($\mu = 18$);
  H_1 : O tratamento é eficaz ($\mu = 14$).
- Hipótese unilateral:
  H_0 : O tratamento não é eficaz ($\mu = 18$);
  H_1 : O tratamento é eficaz ($\mu < 18$).
- Hipótese bilateral:
  H_0 : O tratamento não é eficaz ($\mu = 18$);
  H_1 : O tratamento é eficaz ($\mu \neq 18$).

Erros ao realizar um teste de hipótese

- Erro Tipo I: rejeitar H_0, quando H_0 é verdadeira.
- Erro Tipo II: não rejeitar H_0 quando H_0 é falsa.

		Situação	
		H_0 Verdadeira	H_0 Falsa
Decisão	Rejeitar H_0	Erro Tipo I	Sem erro
	Não rejeitar H_0	Sem error	Erro Tipo II

- $\alpha = \text{P}(\text{erro tipo I}) = \text{P}(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira})$;
- $\beta = \text{P}(\text{error tipo II}) = \text{P}(\text{não rejeitar } H_0 \mid H_0 \text{ falsa})$.
- $\alpha = \text{P}$(concluir que o tratamento é eficaz quando na verdade não é);
- $\beta = \text{P}$(concluir que o tratamento não é eficaz quando na verdade ele é).

Laboratório de Estatística e Geoinformação

Valor crítico

- α é chamado nível de significância do teste.
- Supnha que $n = 30$ observação e a média amostral é $\bar{X}$.
- Supondo α conhecido podemos determinar o valor crítico x_c.

$$\alpha = P(\text{Erro tipo I}) = P(\text{rejeitar} \quad H_0|H_0 \quad \text{Verdadeira})$$

$$= P(\bar{X} < x_c|\mu = 18) = P(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{x_c - 18}{6/\sqrt{30}})$$

$$= P(Z < z_c)$$

com $Z \sim N(0, 1)$.

Laboratório de Estatística e Geoinformação

Obtendo o valor crítico

- Dado α encontramos z_c na Tabela normal padrão.
- Obtemos x_c

$$z_c = \frac{x_c - 18}{6/\sqrt{30}} \rightarrow x_c = 18 + z_c \frac{6}{\sqrt{30}}.$$

- Supondo $\alpha = 0.05$ temos

$$0.05 = P(Z < z_c) \rightarrow z_c = -1.64;$$

logo

$$x_c = 18 - 1.64 \frac{6}{\sqrt{30}} = 16.20.$$

Regiões de aceitação e rejeição

- Região crítica ou região de rejeição

$$RC = \{x \in \Re : x < 16.20\}.$$

- Região de aceitação (RA) é o complemento de RC.

Teste de hipótese bilateral

- Defina as hipóteses $H_0 : \mu = \mu_0$ e $H_1 : \mu \neq \mu_0$.
- Defina a região crítica

$$RC = \{x \in \Re | x < x_{c1} \quad \text{ou} \quad x > x_{c2}\}$$

tal que

$$P(\bar{X} < x_{c1}) = \alpha/2 \quad \text{e} \quad P(\bar{X} > x_{c2}) = \alpha/2.$$

Laboratório de Estatística e Geoinformação

Etapas de um teste de hipóteses

- Estabelecer as hipóteses nula e alternativa.
- Definir a forma da região crítica, com base na hipótese alternativa.
- Identificar a distribuição do estimador e obter sua estimativa.
- Fixar α e obter a região crítica.
- Concluir o teste com base na estimativa e na região crítica.

Exemplo 8.3

- Um pesquisador deseja estudar o efeito de certa substância no tempo de reação de seres vivos a um certo tipo de estímulo. Um experimento é desenvolvido com cobaias que são inoculadas com a substância e submetidas a um estímulo elétrico, com seus tempos de reação (em segundos) anotados. Os seguintes valores foram obtidos: 9.1; 9.3; 7.2; 7.5; 13.3; 10.9; 7.2; 9.9; 8.0; 8.6. Admite-se que o tempo de reação segue, em geral, o modelo Normal com média 8 e desvio padrão $\sigma = 2$ segundos. O pesquisador desconfia que o tempo médio sofre alteração por influência da substância. Efetue um teste de hipótese para verificar se o pesquisador tem razão. E calcule a probabilidade do erro tipo II supondo que $\mu = 9$.

Laboratório de Estatística e Geoinformação

Exemplo 8.4

- Um relatório de uma companhia afirma que 40% de toda a água obtida, através de poços artesianos no nordeste é salobra. Há muitas controvérsias sobre essa afirmação, alguns dizem que a proporção é maior, outros que é menor. Para dirimir as dúvidas, 400 poços foram sorteados e observou-se, em 120 deles água salobra. Qual seria a conclusão, ao nível de 3%?

Teste para a média com variância desconhecida

- Variância amostral $S^2 = (\sum_{i=1}^{n} X_i^2 - n\bar{X}^2)/(n-1)$.
- Defina a estatística $T = \frac{\bar{X}-\mu}{\sqrt{S^2/n}}$.
- T tem distribuição t-Student com $(n-1)$ graus de liberdade.

Laboratório de Estatística e Geoinformação

Exemplo 8.5

- Deseja-se investigar se uma certa moléstica que ataca o rim altera o consumo de oxigênio desse orgão. Para indivíduos sadios, admite-se que esse consumo tem distribuição Normal com média 12 cm^3/min. Os valores medidos em cinco pacientes com a moléstia foram: $14.4; 12.9; 15.1; 13.7; 13.5$. Qual seria a conclusão, ao nível de 1% de significância?

Laboratório de Estatística e Geoinformação

Intervalo de confiança para μ com variância desconhecida

- $IC(\mu, \gamma) = [\bar{X} - t_{\gamma/2}\frac{S}{\sqrt{n}}; \bar{X} + t_{\gamma/2}\frac{S}{\sqrt{n}}]$.
- Exemplo 8.6: Para os dados do exemplo anterior obtenha um intervalo com 90% de confiança para a média populacional.
- A TABELA T JÁ DA O VALOR BILATERAL!!!
- $[13.90 - 2.132\sqrt{0.67/5}; \; 13.90 + 2.132\sqrt{0.67/5}]$
- $[13.09; 14.71]$

Laboratório de Estatística e Geoinformação

Nível descritivo

- Em geral o α é pré-fixado.
- Podemos calcular supondo que a hipóese nula é verdadeira a probabilidade de se obter estimativas mais extremas do que a que está fornecida pela amostra.
- Exemplo 8.7 - Uma associação de defesa do consumidor desconfia que embalagens de 450 gramas de um certo tipo de biscoito estão abaixo do peso. Para verificar tal afirmação, foram coletados ao acaso 80 pacotes em vários supermercados, obtendo-se uma média de peso de 447 gramas. Admitindo-se que o peso dos pacotes segue o modelo Normal com desvio padrão 10 grams, que conclusão pode ser tirada através do nível descritivo?

Teste Qui-Quadrado

- Exemplo 8.9 - No exemplo 8.2, definimos X como sendo o número de impactos anteriores à falha em um equipamento eletrônico. Uma amostra de 80 ensaios foi obtida, cada ensaio representando os testes feitos até a interrupção por falha no equipamento, resultado 80 observações da variável de interesse. Pretende-se verificar se o modelo Geométrico com $p = 0.4$ é adequado.
- Exemplo 8.10 Tarefa de casa.

Teste Qui-Quadrado

- Exemplo 8.10 - A tabela abaixo contém os resultados obtidos por estudantes do ensino médio, em um exame com questões nas disciplinas de física e matemática. Deseja-se testar se existe dependência entre as notas dessas duas disciplinas que, para efeito de apresentação na tablea e analise de comportamento foram classificadas nas categorias alta, média e baixa.

Física/Matemática	Alta	Média	Baixa	Total
Alta	56	71	12	139
Média	47	163	38	248
Baixa	14	42	85	141
Total	117	276	135	528

- As notas de física e matemática são independentes?

Exercícios recomendados

- Seção 8.2 - 1 a 6.
- Seção 8.3 - 1 a 6.
- Seção 8.4 - 1 a 4.
- Seção 8.5 - 1 a 7.