# Lista A2 - Inferência Estatística

Professor: Philip Thompson
Monitores: Eduardo Adame & Ezequiel Braga

## Instruções

- É recomendável justificar bem os passos e argumentos usados. Parte da avaliação consiste em verificar a explanação matemática *além* da conclusão final.

## Questões

*Questão* 1. Seja $X$ variável aleatória com distribuição exponencial com parâmetro $\beta$. Suponha que queiramos testar as hipóteses $H_0 : \beta \geq 1$ versus $H_1 : \beta < 1$. Considere o teste $\delta$ em que

$\delta(X)$ rejeita $H_0$ sse $X \geq 1$.

a) Determine a função poder do teste $\delta$.

b) Compute o tamanho do teste $\delta$.

*Questão* 2. Suponha que $\boldsymbol{X} = (X_1, \ldots, X_n)$ seja uma amostra de uma distribuição uniforme no intervalo $[0, \theta]$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 &: \quad \theta \geq 2 \\ H_1 &: \quad \theta < 2. \end{aligned}$$

Seja $Y_n = \max\{X_1, \ldots, X_n\}$ e considere o teste $\delta$ em que

$\delta(\boldsymbol{X})$ rejeita $H_0$ sse $Y_n \leq 1.5$.

a) Determine a função poder do teste $\delta$.

b) Compute o tamanho do teste $\delta$.

*Questão* 3. Suponha que a distribuição amostrada seja uma Bernoulli com parâmetro $p \in (0, 1)$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 &: \quad p = 0.2 \\ H_1 &: \quad p \neq 0.2. \end{aligned}$$

Suponha que tenhamos uma amostra $\boldsymbol{X} = (X_1, \ldots, X_{20})$ e denote $Y_n = \sum_{i=1}^{20} X_i$. Considere o teste $\delta$ em que

$\delta(\boldsymbol{X})$ rejeita $H_0$ sse $Y_n \leq 1$ ou $Y_n \geq 7$.

a) Determine a função poder do teste $\delta$ nos pontos $p \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$. Plote-a nestes pontos.

b) Compute o tamanho do teste $\delta$.

*Questão* 4. Suponha que tenhamos apenas uma amostra $X$ da distribuição uniforme no intervalo $[\theta - 0.5, \theta + 0.5]$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 : &\quad \theta \leq 3 \\ H_1 : &\quad \theta \geq 4. \end{aligned}$$

Construa um teste $\delta$ para o qual a função poder seja $\pi_\delta(\theta) = 0$ para $\theta \leq 3$ e $\pi_\delta(\theta) = 1$ para $\theta \geq 4$.

*Questão* 5. Suponha que tenhamos uma amostra $\boldsymbol{X} = (X_1, \ldots, X_n)$ de uma distribuição uniforme no intervalo $[\theta - 0.5, \theta + 0.5]$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 : &\quad 3 \leq \theta \leq 4 \\ H_1 : &\quad \theta < 3 \text{ or } \theta > 4. \end{aligned}$$

Seja $Y_n = \max\{X_1, \ldots, X_n\}$. Considere os testes:

- $\delta(\boldsymbol{X})$ rejeita $H_0$ sse $Y_n \leq 2.9$ ou $Y_n \geq 4$.
- $\delta_*(\boldsymbol{X})$ rejeita $H_0$ sse $Y_n \leq 2.9$ ou $Y_n \geq 4.5$.

Mostre que

a) Prove que $\pi_{\delta_*}(\theta) = \pi_\delta(\theta)$ para $\theta \leq 4$.

b) Prove que $\pi_{\delta_*}(\theta) < \pi_\delta(\theta)$ para $\theta > 4$.

c) Qual dos dois testes é melhor?

*Questão* 6. Suponha que tenhamos uma amostra $\boldsymbol{X} = (X_1, \ldots, X_n)$ de uma distribuição normal $\mathcal{N}(\mu, 1)$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 : &\quad \mu \leq \mu_0 \\ H_1 : &\quad \mu > \mu_0. \end{aligned}$$

Seja $Z = \sqrt{n}(\bar{X} - \mu_0)$ e considere os testes

$\delta(\boldsymbol{X})$ rejeita $H_0$ sse $Z \geq c$.

Mostre que

a) Mostre que a função poder $\pi_\delta(\mu) = \mathbb{P}_\mu(Z \geq c)$ é crescente.

b) Ache o valor crítico $c$ tal que o teste tenha tamanho $\alpha_0$.

c) Justificando, ache uma fórmula para o valor-p do teste numa realização onde $Z = z$.

*Questão* 7. Suponha que tenhamos uma amostra única $X$ de uma distribuição Cauchy centrada em $\theta$, isto é, com pdf

$$f(x|\theta) = \frac{1}{\pi[1 + (x - \theta)^2]}.$$

Queremos testar as hipóteses:

$$\begin{aligned} H_0 &: \quad \theta \leq \theta_0 \\ H_1 &: \quad \theta > \theta_0. \end{aligned}$$

Considere o teste

$\delta_c(X)$ rejeita $H_0$ sse $X \geq c$.

Mostre que

a) Mostre que a função poder $\pi_{\delta_c}(\theta) = \mathbb{P}_\theta(X \geq c)$ é crescente.

b) Ache o valor crítico $c$ tal que o teste tenha tamanho $0.05$.

c) Justificando, ache uma fórmula para o valor-p do teste numa realização onde $X = x$.

*Questão* 8. Suponha que tenhamos uma amostra $\boldsymbol{X} = (X_1, \ldots, X_n)$ de uma distribuição normal $\mathcal{N}(\mu, 1)$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 &: \quad \mu = \mu_0 \\ H_1 &: \quad \mu \neq \mu_0. \end{aligned}$$

Considere o teste

$\delta(\boldsymbol{X})$ rejeita $H_0$ sse $\bar{X} \leq c_1$ ou $\bar{X} \geq c_2$.

Denote por $\pi_\delta(\mu)$ a função poder deste teste.

a) Determine os valores $c_1 < c_2$ tais que $\pi_\delta(\mu_0) = 0.1$ e a função $\pi_\delta(\mu)$ seja simétrica em relação ao ponto $\mu = \mu_0$.

b) Determine o menor valor de $n$ para o qual vale o item a) e $\pi_\delta(\mu_0 - 1) = \pi_\delta(\mu_0 + 1) \geq 0.95$.

*Questão* 9 (Intervalo de confiança para a variância). Suponha que tenhamos uma amostra $\boldsymbol{X} = (X_1, \ldots, X_n)$ de uma distribuição normal $\mathcal{N}(\mu, \sigma^2)$ com $(\mu, \sigma^2)$ desconhecidos. Construa um intervalo de confiança para $\sigma^2$ com nível de confiança $1 - \alpha$ com $\alpha \in (0, 1)$. DICA: você deve usar um pivot visto em aula.

*Questão* 10. Suponha que tenhamos uma amostra de tamanho 9 de uma distribuição normal $\mathcal{N}(\mu, \sigma^2)$ com $(\mu, \sigma^2)$ desconhecidos tais que $\bar{X}_9 = 22$ e $\sum_{i=1}^{9}(X_i - \bar{X}_9)^2 = 72$.

a) Conclua o resultado do teste das hipóteses à seguir com nível de significância $0.05$:

$$\begin{aligned} H_0 : &\quad \mu \leq 20 \\ H_1 : &\quad \mu > 20. \end{aligned}$$

b) Calcule o valor-p do teste do item a) para o correspondente valor da estatística obtida.

c) Construa o intervalo com confiança 0.95 correspondente para $\mu$.

*Questão* 11. Considere os dados do exercício anterior.

a) Conclua o resultado do teste das hipóteses à seguir com nível de significância $0.05$:

$$\begin{aligned} H_0 : &\quad \mu = 20 \\ H_1 : &\quad \mu \neq 20. \end{aligned}$$

b) Calcule o valor-p do teste do item a) para o correspondente valor da estatística obtida.

c) Construa o intervalo com confiança 0.95 correspondente para $\mu$.

*Questão* 12 (Teste de hipóteses para a variância). Suponha que tenhamos uma amostra $\boldsymbol{X} = (X_1, \ldots, X_n)$ de uma distribuição normal $\mathcal{N}(\mu, \sigma^2)$ com $(\mu, \sigma^2)$ desconhecido. Queremos testar as hipóteses:

$$\begin{aligned} H_0 : &\quad \sigma^2 \leq \sigma_0^2 \\ H_1 : &\quad \sigma^2 > \sigma_0^2. \end{aligned}$$

Seja $S_n^2 := \sum_{i=1}^n (X_i - \bar{X}_n)^2$ e considere o teste

$\delta(\boldsymbol{X})$ rejeita $H_0$ sse $V := S_n^2/\sigma_0^2 \geq c$.

Denote por $\pi_\delta(\mu, \sigma^2)$ a função poder deste teste. Para um nível de significância $\alpha_0 \in (0, 1)$, determine o valor de $c$ tal que:

- $\pi_\delta(\mu, \sigma^2) = \alpha_0$ se $\sigma^2 = \sigma_0^2$.
- $\pi_\delta(\mu, \sigma^2) < \alpha_0$ se $\sigma^2 < \sigma_0^2$.
- $\pi_\delta(\mu, \sigma^2) > \alpha_0$ se $\sigma^2 > \sigma_0^2$.

*Questão* 13 (Teste de hipóteses para a variância). Considere novamente o exercício anterior. Obtenha uma formula para o valor-p do teste $\delta$ num valor de estatística $V = v$. DICA: o quantil usado deverá ser da distribuição $\chi^2$.

*Questão* 14 (O teste-t como teste de razão de verossimilhança). Suponha que tenhamos uma amostra $\boldsymbol{X} = (X_1, \ldots, X_n)$ de uma distribuição normal $\mathcal{N}(\mu, \sigma^2)$ com $(\mu, \sigma^2)$ desconhecido. Queremos testar as hipóteses:

$$\begin{aligned} H_0 : &\quad \mu \leq \mu_0 \\ H_1 : &\quad \mu > \mu_0. \end{aligned}$$

Recorde que a verossimilhança é dada por

$$f_n(\boldsymbol{x}|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.\right]$$

A *razão de verossimilhança* é a estatística definida por:

$$\Lambda(\boldsymbol{x}) := \frac{\max_{(\mu,\sigma^2):\mu>\mu_0} f(\boldsymbol{x}|\mu, \sigma^2)}{\max_{(\mu,\sigma^2)} f(\boldsymbol{x}|\mu, \sigma^2)}.$$

Recorde que definimos o teste-t usando a estatística

$$U(\boldsymbol{X}) := \sqrt{n}\frac{\bar{X}_n - \mu_0}{\hat{\sigma}},$$

onde $\hat{\sigma}^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$.

Mostre que o teste $\delta(\boldsymbol{X})$ em que, para algum $k < 1$,

$\delta(X)$ rejeita $H_0$ sse $\Lambda(\boldsymbol{X}) \leq k$,

é exatamente o teste-t em que rejeita-se $H_0$ sse $U(\boldsymbol{X}) \geq c$, desde que

$$c = \sqrt{\left(\frac{1}{k^{1/2}} - 1\right)(n-1)}.$$

*Questão* 15. Suponha que tenhamos duas amostras independentes, $\boldsymbol{X} = (X_1, \ldots, X_m)$ de uma distribuição normal $\mathcal{N}(\mu_1, \sigma^2)$ e $\mathbf{Y} = (Y_1, \ldots, Y_n)$ de uma distribuição normal $\mathcal{N}(\mu_2, \sigma^2)$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 : &\quad \mu_1 = \mu_2 \\ H_1 : &\quad \mu_1 \neq \mu_2. \end{aligned}$$

Dado $\alpha_0 \in (0, 1)$, recorde o teste do livro texto tal que

$\delta(\boldsymbol{X})$ rejeita $H_0$ sse $U \geq c$,

onde $c = T^{-1}_{m+n-2}(1 - \alpha_0/2)$ e $U$ é dado por (9.6.3).

Mostre que o valor-p deste teste na realização $U = u$ observada é $2[1 - T_{m+n-2}(|u|)]$.

*Questão* 16. Suponha que tenhamos duas amostras independentes, $\boldsymbol{X} = (X_1, \ldots, X_6)$ de uma distribuição normal $\mathcal{N}(\mu_1, \sigma_1^2)$ e $\mathbf{Y} = (Y_1, \ldots, Y_{10})$ de uma distribuição normal $\mathcal{N}(\mu_2, \sigma_2^2)$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 : &\quad \sigma_1^2 \leq \sigma_2^2 \\ H_1 : &\quad \sigma_1^2 > \sigma_2^2. \end{aligned}$$

Dado $\alpha_0 \in (0, 1)$, recorde o teste do livro texto tal que

$\delta(\boldsymbol{X})$ rejeita $H_0$ sse $V \geq c$,

onde $c = G_{15,9}^{-1}(1-\alpha_0)$ e $V$ é dado por (9.7.4).

Suponha que observamos $\sum_{i=1}^{16} X_i = 84$, $\sum_{i=1}^{16} X_i^2 = 563$, $\sum_{i=1}^{10} Y_i = 18$, $\sum_{i=1}^{10} Y_i^2 = 72$. Tomando $\alpha_0 = 0.05$, escreva a conclusão do teste argumentando de duas maneiras diferentes: comparando $V = v$ com $c$ e comparando o valor-p em $V = v$ com $\alpha_0$.

*Questão* 17. Prove o Teorema 9.7.5 do livro texto.

*Questão* 18. Mostre $\sum_{i=1}^{n} (c_1 x_i + c_2)^2 = c_1^2 \sum_{i=1}^{n} (c_1 x_i + c_2)^2 + n(c_1 \bar{x} + c_2)^2$.

*Questão* 19. Considere os dados $(y_1, x_1), \ldots, (y_n, x_n)$ em $\mathbb{R} \times \mathbb{R}$ e o vetor $(\hat{\beta}_0, \hat{\beta}_1)$ a solução do método de mínimos quadrados

$$(\hat{\beta}_0, \hat{\beta}_1) \in \underset{(\beta_0, \beta_1)}{\operatorname{argmin}} Q(\beta_0, \beta_1),$$

onde

$$Q(\beta_0, \beta_1) := \sum_{i=1}^{n} \left(y_i - (\beta_0 + \beta_1 x_i)\right)^2.$$

Defina

$$s_x^2 := \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

Mostre que:

a) $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{s_x^2}$.

b) $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) y_i}{s_x}$.

c) $\hat{\beta}_0 = \hat{\beta}_1 \bar{x} - \hat{\beta}_0$.

*Questão* 20. No problema de regressao linear visto em aula, mostre que, condicionalmente à $(X_1, \ldots, X_n)$,

a) $\mathbb{E}[\hat{\beta}_1] = \beta_1$.

b) $\mathbb{E}[\hat{\beta}_0] = \beta_0$.

c) $\mathbb{V}(\hat{\beta}_1) = \frac{\sigma^2}{s_x^2}$.

d) $\mathbb{V}(\hat{\beta}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{s_x^2}\right)$.

e) $\mathsf{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\bar{X}\sigma^2}{s_x^2}$.