Notas de aulas pelos slides (para A2)

Distribuição de média e variância amostrais

- Se $X_1,\dots,X_n \sim Normal(\mu,\sigma^2)$, então

$$\bar{X}_n \sim Normal\left(\mu, \frac{\sigma^2}{n}\right)$$

$$S_n^2 \sim Gama\left(\frac{n-1}{2}, \frac{n}{2\sigma^2}\right)$$

- Seja $Y \sim$ Qui-quadrado(m) e $Z \sim Normal(0,1)$ e

$$X = \frac{Z}{\sqrt{\frac{Y}{m}}}$$

então X é uma **t de Student** com n graus de liberdade, e

$$f_X(x) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)}\left(1+\frac{x^2}{m}\right)^{-\frac{m+1}{2}}, \; x\in(-\infty,\infty)$$

- Se $\hat{\sigma}' = \sqrt{\frac{(\sum(x_i-\bar{X}_n)^2)^2}{n-1}}$ (estimador), então

$$\frac{\sqrt{n}\,(\bar{X}_n-\mu)}{\hat{\sigma}'} \sim Student(n-1)$$

Intervalos de confiança

- Seja $\mathbf{X}=\{X_1,\dots,X_n\}$, cada uma com p.d.f. $f(x|\theta)$, e considere $g(\theta)$. Sejam $A(\mathbf{X})$ e $B(\mathbf{X})$ estatísticas tais que

$$P\{A(\mathbf{X}) < g(\theta) < B(\mathbf{X})\} \geq \gamma$$

Então $I(\mathbf{X}) = (A(\mathbf{X}), B(\mathbf{X}))$ é um **intervalo de confiança** de $100.\gamma\,\%$

- Definimos intervalos de **confiança unilateral**, pondo $A(\mathbf{X}) = -\infty$ ou $B(\mathbf{X}) = \infty$ na definição acima.

- Seja $\mathbf{X}=\{X_1,\dots,X_n\}$, cada uma com p.d.f $f(x|\theta)$ e $V(\mathbf{X},\theta)$ uma v.a. com **mesma** distribuição para todo $\theta\in\Omega$.
Diz-se que $V(X,\theta)$ é uma **quantidade pivotal**.

- Seja $\mathbf{X}=\{X_1,\dots,X_n\}$, tome $0<\gamma<1$ e $\gamma_1<\gamma_2$ de modo que $\gamma_2-\gamma_1=\gamma$, e seguindo certas condições,

$$A(\mathbf{X}) = r(G^{-1}(\gamma_1), \mathbf{X})$$
$$B(\mathbf{X}) = r(G^{-1}(\gamma_2), \mathbf{X})$$

são os limites de um intervalo de confiança $100.\gamma\,\%$ para $g(\theta)$.

Testes de hipóteses

- Seja Ω um espaço de parâmetros e defina $\Omega_0, \Omega_1 \subset \Omega$ de modo que $\Omega_0 \cup \Omega_1 = \Omega$ e $\Omega_0\cap\Omega_1=\emptyset$

definimos

$$H_0 := \theta\in\Omega_0$$
$$H_1 := \theta\in\Omega_1$$

H_0 é a **hipótese nula** e H_1 é a **hipótese alternativa**. Se H_0, dizemos que rejeitamos a hipótese nula. Por outro lado, se $\theta\in\Omega_0$ dizemos que não rejeitamos ou falhamos em rejeitar H_0.

- Hipótese simples e hipótese compostas.

- Uma hipótese da forma $H_0: \theta\leq\theta_0$ (ou $\theta\geq\theta_0$) é dita **unilateral**, enquanto hipóteses da forma $H_0: \theta\neq\theta_0$ são ditas **bilaterais**.

- O conjunto

$$S_1 := \{x : |\bar{X}_n - \mu_0| \geq c\},$$

é chamado de **região crítica** do teste.

- Seja δ um procedimento de aceitação/rejeição. A **função poder** é definida

$$\pi(\theta|\delta) := P(\mathbf{X}\in S_1|\theta) = P(T\in R|\theta), \; \theta\in\Omega$$

- **Tipos de erros**

Nome	Erro cometido
Erro tipo I	Rejeitar H_0 quando ela é **verdadeira**.
Erro tipo II	Falhar em rejeitar H_0 quando ela é **falsa**.

- Dizemos que um teste, δ, tem **tamanho** ou **nível de significância** $\alpha(\delta)$, com

$$\alpha(\delta) := \sup_{\theta\in\Omega_0} \pi(\theta|\delta)$$

- Para cada t, seja δ_t o teste que rejeita H_0 se $T\geq t$. Então, quando $T=t$, o **p-valor** vale

$$p(t) := \sup_{\theta\in\Omega_0}\pi(\theta|\delta_t) = \sup_{\theta\in\Omega_0} P(T\geq t|\theta),$$

ou seja, o p-valor é o tamanho do teste δ_t.

Razões de verossimilhanças

- **Intervalos de confiança e testes são equivalentes**

- Se um conjunto aleatório $w(\mathbf{X})$ satisfaz

$$P(g(\theta_0)\in w(\mathbf{X})|\theta=\theta_0)\geq\gamma, \; \forall\theta_0\in\Omega$$

então chamamos $w(\mathbf{X})$ de um **conjunto de confiança** para $g(\theta)$.

- Se $w(\mathbf{X})$ é um conjunto de confiança para uma função de interesse $g(\theta)$. Então $\forall g_0$ assumido por $g(\theta)$, $\exists\ \delta_{g_0}$, de nível α_0 que rejeita

$$H_0: g(\theta) = g_0 \iff g(\theta_0) = g_0 \in w(\mathbf{X}).$$

- A estatística

$$\Lambda(x) = \frac{\sup_{\theta \in \Omega_0} f_n(x \mid \theta)}{\sup_{\theta \in \Omega} f_n(x \mid \theta)},$$

é chamada uma estatística de razão de verossimilhança. Um teste de razão de verossimilhança, δ_K é um teste que rejeita H_0 se $\Lambda(x) \leq K$ para K constante.

- Pelo Teorema de Wilks, a medida que a amostra aumenta $(n \to \infty)$, então

$$-2 \log \Lambda(x) \xrightarrow{d} \chi^2(K).$$

Teste t

- Dizemos que δ é não-viesado se e somente se para $\theta \in \Omega_0$ e $\theta' \in \Omega_1$, vale

$$\pi(\theta \mid \delta) \leq \pi(\theta' \mid \delta)$$

ou seja, se a função poder é pelo menos tão grande no espaço onde H_0 é falsa (Ω_1) quanto no espaço em que H_0 é verdadeira (Ω_0).

- Um teste δ_c que rejeita H_0 se $U \geq c$ $(U \leq c)$, com $c = T^{-1}(1-\alpha_0; n-1)$ é chamado teste t de tamanho α_0.

- Seja $U = u$ e $T(\cdot\,; n-1)$ a f.d.a. de distribuição t de Student com $n-1$ graus. Para a hipótese $H_0: \mu \geq \mu_0$, $H_1: \mu < \mu_0$, então o p-valor vale $T(u; n-1)$.

- Teste t pareado

- Considere $\mathbf{X} = \{X_1, \dots, X_m\}$ e $\mathbf{Y} = \{Y_1, \dots, Y_n\}$, para estudar a diferença das médias para duas amostras. Modelando em distribuição normal $X_i \sim \text{Normal}(\mu_1, \sigma_1^2)$ e $Y_i \sim \text{Normal}(\mu_2, \sigma_2^2)$, sob a premissa de homogeneidade $\sigma_1^2 = \sigma_2^2 = \sigma^2$, podemos testar a hipótese

$$H_0: \mu_1 \leq \mu_2$$
$$H_1: \mu_1 > \mu_2$$

computando a estatística

$$U = \frac{\sqrt{m+n-2}\ (\bar{X}_m - \bar{Y}_n)}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)(S_X^2 + S_Y^2)}}$$

- o teste t é não-viesado e pode ser computado como um teste de razão de verossimilhança.

Testes para igualdade de variâncias

- Seja $Y \sim$ Qui-quadrado(m) e $W \sim$ Qui-quadrado(n), então

$$X = \frac{Y/m}{W/n}$$

tem distribuição F com m e n graus de liberdade.

- $X \sim F(m,n) \Rightarrow \frac{1}{X} \sim F(m,n)$

- $Y \sim \text{Student}(n) \Rightarrow Y^2 \sim F(1,n)$

- Suponha $X_i \sim \text{Normal}(\mu_1, \sigma_1^2)$ e $Y \sim \text{Normal}(\mu_2, \sigma_2^2)$, queremos testar a diferença das variâncias.

$$H_0: \sigma_1^2 \leq \sigma_2^2,$$
$$H_1: \sigma_1^2 > \sigma_2^2$$

para isso basta computar a estatística.

- O teste F de homogeneidade (igualdade de variâncias) é o teste δ_c que rejeita H_0 se $V \geq c$, para uma constante positiva c.

- Seja $V = \dfrac{S_X^2/(m-1)}{S_Y^2/(n-1)}$, então:

$$\frac{\sigma_2^2}{\sigma_1^2} V \sim F(m-1, n-1)$$

Regressão Linear

- Sejam (x_i, y_i) n pontos e queremos a reta

$$y_i = \beta_0 + \beta_1 x_i$$

os valores que minimizam a soma de quadrados

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

- Podemos construir um modelo estatístico explícito para a relação entre as variáveis X e Y:

$$Y_i = \beta_0 + \sum \beta_j x_{ij} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

- Sobre as premissas: conhecimento do estimador, normalidade, linearidade na média, variância comum e independencia. Os estimadores de máxima verossimilhança para $\theta = (\beta_0, \beta_1, \sigma^2)$ são

$$\hat{\beta}_{0_{EMV}} = \bar{Y} - \hat{\beta}_{1_{EMV}} \bar{X}$$

$$\hat{\beta}_{1_{EMV}} = \frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}$$

$$\hat{\sigma}^2_{EMV} = \frac{1}{n} \sum \left(Y_i - \left(\hat{\beta}_{0_{EMV}} - \hat{\beta}_{1_{EMV}} \bar{x} \right) \right)^2$$

- Distribuições dos estimadores:

$$\hat{\beta}_{0_{EMV}} \sim Normal\left(\beta_0, \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{s_x^2}\right)\right)$$

$$\hat{\beta}_{1_{EMV}} \sim Normal\left(\beta_1, \frac{\sigma^2}{s_x^2}\right)$$

$$Cov\left(\hat{\beta}_{0_{EMV}}, \hat{\beta}_{1_{EMV}}\right) = \frac{-\bar{X}\sigma^2}{s_x^2}$$

onde $s_x = \sqrt{\sum (X_i - \bar{X})^2}$

- Intervalos de confiança para uma regressão linear:

$$\hat{\beta}_0 \pm \hat{\sigma}' c \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_x^2}} \quad e \quad \hat{\beta}_1 \pm c \frac{\hat{\sigma}'}{S_x}$$

$$\hat{\beta}_0 + \hat{\beta}_1 x_{pred} \pm c\hat{\sigma}' \sqrt{\frac{1}{n} + \frac{(x_{pred} - \bar{x})^2}{S_x^2}}$$

onde $c = T^{-1}\left(1 - \frac{\alpha_0}{2}; n-2\right)$ e

$$\hat{\sigma}' := \sqrt{\frac{\sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2}{n-2}}$$

- Para testar uma hipotese para o coeficiente angular, testamos

$$H_0 : \beta_1 = \beta^*$$
$$H_1 : \beta_1 \neq \beta^*$$

podemos utilizar a estatística

$$U_1 = s_x \frac{\hat{\beta}_1 - \beta^*}{\hat{\sigma}'}$$

- Se $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_{pred}$, então o EQM é

$$E\left[(\hat{Y} - Y)^2\right] = \sigma^2 \left(1 + \frac{1}{n} + \frac{(x_{pred} - \bar{x})^2}{S_x^2}\right)$$

- A probabilidade de $Y = \hat{\beta}_0 + \hat{\beta}_1 x_{pred}$ estar no intervalo

$$\hat{Y} \pm T^{-1}\left(1 - \frac{\alpha_0}{2}; n-2\right) \hat{\sigma}' \sqrt{1 + \frac{1}{n} + \frac{(x_{pred} - \bar{x})^2}{S_x^2}}$$

é $1 - \alpha_0$.