Resumo dos resultados de Estatística II

Versão 1.3

Insper

Quatro casos estudados para a variável de interesse X

Caso 1. $X \sim \mathrm{N}(\mu, \sigma^2)$ e σ^2 conhecida

Caso 2. $X \sim \mathrm{N}(\mu, \sigma^2)$

Caso 3. $X \sim \mathrm{Bernoulli}(p)$

Caso 4. X com "qualquer" distribuição e $\mathrm{Var}[X]$ conhecida

Caso 1. $X \sim \mathrm{N}(\mu, \sigma^2)$ e σ^2 conhecida

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \to \mu \qquad \bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \qquad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$\Pr\{|\bar{X} - \mu| < \epsilon\} = \gamma$$

$$\Pr\left\{ \underbrace{-\frac{\epsilon}{\sigma/\sqrt{n}}}_{-z} < \underbrace{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}_{=Z\sim\mathrm{N}(0,1)} < \underbrace{\frac{\epsilon}{\sigma/\sqrt{n}}}_{z} \right\} = \gamma \qquad \epsilon = z\frac{\sigma}{\sqrt{n}} \qquad n = \left(\frac{z}{\epsilon}\right)^2 \sigma^2$$

$$\Pr\left\{ \bar{X} - z\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z\frac{\sigma}{\sqrt{n}} \right\} = \gamma$$

$$\mathrm{IC}[\mu; \gamma] = \left(\bar{x} - z\frac{\sigma}{\sqrt{n}} \; ; \; \bar{x} + z\frac{\sigma}{\sqrt{n}} \right)$$

Caso **2**. $X \sim \mathrm{N}(\mu, \sigma^2)$: aprendendo μ

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mu \qquad\qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})$$

$$\Pr\{|\bar{X} - \mu| < \epsilon\} = \gamma$$

$$\Pr\left\{ \underbrace{-\frac{\epsilon}{S/\sqrt{n}}}_{-t} < \underbrace{\frac{\bar{X} - \mu}{S/\sqrt{n}}}_{=T\sim t_{n-1}} < \underbrace{\frac{\epsilon}{S/\sqrt{n}}}_{t} \right\} = \gamma \qquad\qquad \epsilon = t\frac{S}{\sqrt{n}}$$

$$\Pr\left\{ \bar{X} - t\frac{S}{\sqrt{n}} < \mu < \bar{X} + t\frac{S}{\sqrt{n}} \right\} = \gamma$$

$$\mathrm{IC}[\mu;\gamma] = \left(\bar{x} - t\frac{s}{\sqrt{n}} \; ; \; \bar{x} + t\frac{s}{\sqrt{n}} \right)$$

Caso **2**. $X \sim \mathrm{N}(\mu, \sigma^2)$: aprendendo σ^2

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}) \to \sigma^2 \qquad\qquad Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$\Pr\{q_1 < Q < q_2\} = \gamma$$

$$\Pr\left\{\frac{(n-1)S^2}{q_2} < \sigma^2 < \frac{(n-1)S^2}{q_1}\right\} = \gamma$$

$$\mathrm{IC}[\sigma^2;\gamma] = \left(\frac{(n-1)s^2}{q_2}\ ;\ \frac{(n-1)s^2}{q_1}\right)$$

$$\Pr\left\{S\sqrt{\frac{(n-1)}{q_2}} < \sigma < S\sqrt{\frac{(n-1)}{q_1}}\right\} = \gamma$$

$$\mathrm{IC}[\sigma;\gamma] = \left(s\sqrt{\frac{(n-1)}{q_2}}\ ;\ s\sqrt{\frac{(n-1)}{q_1}}\right)$$

Caso 3. $X \sim \mathrm{Bernoulli}(p)$

$$\hat{p} \to p \qquad \hat{p} \overset{\mathrm{TLC}}{\approx} \mathrm{N}\left(p, \frac{p(1-p)}{n}\right) \qquad \Pr\{|\hat{p} - p| < \epsilon\} = \gamma$$

$$\Pr\Bigg\{ \underbrace{-\frac{\epsilon}{\sqrt{p(1-p)/n}}}_{-z} < \underbrace{\frac{\hat{p} - p}{\sqrt{p(1-p)/n}}}_{=Z \overset{\mathrm{TLC}}{\approx} \mathrm{N}(0,1)} < \underbrace{\frac{\epsilon}{\sqrt{p(1-p)/n}}}_{z} \Bigg\} = \gamma$$

$$\epsilon = z\sqrt{\frac{p(1-p)}{n}} \qquad n = \left(\frac{z}{\epsilon}\right)^2 p(1-p)$$

Pior caso $(p = 0,5)$: $n = \frac{1}{4}\left(\frac{z}{\epsilon}\right)^2$ Piloto: $n = \left(\frac{z}{\epsilon}\right)^2 \hat{p}_{\text{piloto}}(1 - \hat{p}_{\text{piloto}})$

Caso **3**. $X \sim \text{Bernoulli}(p)$

$$\Pr\left\{\hat{p} - z\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + z\sqrt{\frac{p(1-p)}{n}}\right\} = \gamma$$

Conservador:

$$\text{IC}[p;\gamma] = \left(\hat{p} - z\sqrt{\frac{1}{4n}}\ ;\ \hat{p} + z\sqrt{\frac{1}{4n}}\right)$$

Otimista:

$$\text{IC}[p;\gamma] = \left(\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ ;\ \hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

Caso 4. X com "qualquer" distribuição e $\mathrm{Var}[X]$ conhecida

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mathrm{E}[X] = \mu \qquad \sigma^2 = \mathrm{Var}[X]$$

$$\bar{X} \overset{\mathrm{TLC}}{\approx} \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \qquad \Pr\{|\bar{X} - \mu| < \epsilon\} = \gamma$$

$$\Pr\left\{\underbrace{-\frac{\epsilon}{\sigma/\sqrt{n}}}_{-z} < \underbrace{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}_{= Z \overset{\mathrm{TLC}}{\approx} \mathrm{N}(0,1)} < \underbrace{\frac{\epsilon}{\sigma/\sqrt{n}}}_{z}\right\} = \gamma \qquad \epsilon = z\frac{\sigma}{\sqrt{n}} \qquad n = \left(\frac{z}{\epsilon}\right)^2 \sigma^2$$

$$\Pr\left\{\bar{X} - z\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z\frac{\sigma}{\sqrt{n}}\right\} = \gamma$$

$$\mathrm{IC}[\mu; \gamma] = \left(\bar{x} - z\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + z\frac{\sigma}{\sqrt{n}}\right)$$

Viés

Se $\hat{\theta}$ é um estimador para o parâmetro θ, definimos

$$\text{Viés}[\hat{\theta}] = \mathrm{E}[\hat{\theta}] - \theta.$$

Se $\text{Viés}[\hat{\theta}] = 0$, então dizemos que $\hat{\theta}$ é um estimador não enviesado (sinônimos: “não viciado” ou “não viesado” (sic)) para o parâmetro θ.

Se $\text{Viés}[\hat{\theta}] \neq 0$, então dizemos que $\hat{\theta}$ é um estimador enviesado (sinônimos: “viciado” ou “viesado” (sic)) para o parâmetro θ.

Intuição: um estimador enviesado erra em média o “alvo” (o valor do parâmetro θ) conforme replicamos o experimento um número muito grande de vezes.

Eficiência relativa

Intuição: a variância de um estimador é uma medida da variabilidade das estimativas em relação ao parâmetro estimado conforme replicamos o experimento um número muito grande de vezes.

Se temos dois estimadores **não enviesados** $\hat{\theta}_1$ e $\hat{\theta}_2$ para um único parâmetro θ, dizemos que $\hat{\theta}_1$ é relativamente mais eficiente do que $\hat{\theta}_2$ se

$$\mathrm{Var}[\hat{\theta}_1] < \mathrm{Var}[\hat{\theta}_2].$$

Erro Quadrático Médio (EQM)

Dado um estimador $\hat{\theta}$ para um parâmetro θ, definimos seu Erro Quadrático Médio por

$$\mathrm{EQM}[\hat{\theta}] = \mathrm{E}[(\hat{\theta} - \theta)^2].$$

É fácil provar que

$$\mathrm{EQM}[\hat{\theta}] = \mathrm{Var}[\hat{\theta}] + \mathrm{Viés}^2[\hat{\theta}].$$

O Erro Quadrático Médio combina os dois aspectos (Viés e Variância) do estimador.

Dado um conjunto de estimadores para um mesmo parâmetro θ, escolhemos (quando for possível fazer a comparação) aquele com o menor Erro Quadrático Médio.

Consistência

Dizemos que um estimador $\hat{\theta}$ para um parâmetro θ é consistente se a probabilidade de $\hat{\theta}$ errar o "alvo" (o valor do parâmetro θ) fica cada vez menor conforme o tamanho da amostra cresce.

Critério: se

$$\lim_{n \to \infty} \mathrm{EQM}[\hat{\theta}] = 0,$$

então $\hat{\theta}$ é um estimador consistente para θ.

Máxima Verossimilhança

Suponha que a variável de interesse X do nosso problema tem distribuição Bernoulli(θ), em que $0 < \theta < 1$. Isto significa que X é uma variável aleatória discreta, assumindo os valores 0 e 1 com probabilidades

$$\Pr_\theta\{X = 1\} = \theta \qquad \text{e} \qquad \Pr_\theta\{X = 0\} = 1 - \theta.$$

O índice θ nas probabilidades acima é um lembrete de que tais probabilidades dependem do valor do parâmetro θ.

A função (massa) de probabilidade de X é definida por

$$f(x \mid \theta) = \Pr_\theta\{X = x\} = \theta^x (1 - \theta)^{1-x}, \quad \text{para} \quad x = 0, 1.$$

Verifique que esta função de probabilidade "funciona" substituindo os dois valores possíveis de x.

Máxima Verossimilhança

Uma amostra aleatória $X_1, \ldots, X_n$ é constituída por variáveis aleatórias independentes tais que cada X_i tem a mesma distribuição da variável de interesse X. A probabilidade conjunta das X_i's é dada por

$$\Pr_\theta\{X_1 = x_1, \ldots, X_n = x_n\} = \prod_{i=1}^{n} \Pr_\theta\{X_i = x_i\} = \prod_{i=1}^{n} f(x_i \mid \theta).$$

Se em nosso experimento observarmos a amostra $\mathbf{x} = (x_1, \ldots, x_n)$, a probabilidade conjunta de $\mathbf{x}$, vista como uma função do parâmetro θ, é denominada função de verossimilhança:

$$L(\theta \mid \mathbf{x}) = \Pr_\theta\{X_1 = x_1, \ldots, X_n = x_n\} = \prod_{i=1}^{n} f(x_i \mid \theta).$$

Máxima Verossimilhança

O método de Estimação por Máxima Verossimilhança diz que devemos escolher como estimativa do parâmetro θ aquele valor que maximize a probabilidade da amostra observada.

Cada valor do parâmetro θ é um "mundo possível" e o mundo mais verossímil é aquele maximiza a probabilidade de observarmos a amostra que temos em mãos. Lembre-se do experimento com os dois envelopes feito em classe.

Verifique que a função de verossimilhança do modelo Bernoulli é dada por

$$L(\theta \mid \mathbf{x}) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}.$$

Máxima Verossimilhança

Para fins de maximização, em geral é muito mais fácil trabalharmos com a função de log-verossimilhança, definida por

$$\ell(\theta \mid \mathbf{x}) = \ln L(\theta \mid \mathbf{x}) = \ln \prod_{i=1}^{n} f(x_i \mid \theta) = \sum_{i=1}^{n} \ln f(x_i \mid \theta).$$

Supondo que a função de verossimilhança não é igual a zero para nenhum θ, encontrar um ponto de extremo da função de verossimilhança é equivalente a encontrar um ponto de extremo da função de log-verossimilhança, pois pela regra de cadeia

$$\frac{d\ell(\theta \mid \mathbf{x})}{d\theta} = \frac{1}{L(\theta \mid \mathbf{x})} \frac{dL(\theta \mid \mathbf{x})}{d\theta}.$$

Logo, $\frac{dL(\theta|\mathbf{x})}{d\theta} = 0$ se e somente se $\frac{d\ell(\theta|\mathbf{x})}{d\theta} = 0$.

Máxima Verossimilhança

Para o modelo Bernoulli, verifique que a função de log-verossimilhança é dada por

$$\ell(\theta \mid \mathbf{x}) = \left(\sum_{i=1}^{n} x_i\right) \ln\theta + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-\theta).$$

Zerando a derivada de $\ell(\theta \mid \mathbf{x})$ para encontrarmos o ponto de máximo, temos que

$$\frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{(n - \sum_{i=1}^{n} x_i)}{1-\theta} = 0,$$

o que nos leva ao estimador de máxima verossimilhança

$$\hat{\theta}_{\mathrm{MV}} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X},$$

que nada mais é do que a proporção amostral.

Máxima Verossimilhança

Para garantir que temos um ponto de máximo, verifique que a segunda derivada de $\ell(\theta \mid \mathbf{x})$ é negativa no ponto $\theta = \hat{\theta}_{\text{MV}}$.

Agora, suponha que nosso parâmetro de interesse não seja θ, mas sim a variância da variável de interesse

$$\sigma^2 = \text{Var}[X] = \theta(1-\theta) = g(\theta).$$

Uma das características mais interessantes e úteis dos estimadores de máxima verossimilhança é a propriedade de invariância: se $\hat{\theta}_{\text{MV}}$ é o estimador de máxima verossimilhança para θ, então $g(\hat{\theta}_{\text{MV}})$ é o estimador de máxima verossimilhança para $g(\theta)$.

Assim, pela propriedade de invariância, o estimador de máxima verossimilhança para a variância da variável de interesse é

$$\widehat{\sigma^2}_{\text{MV}} = g(\hat{\theta}_{\text{MV}}) = \hat{\theta}_{\text{MV}}(1-\hat{\theta}_{\text{MV}}) = \bar{X}(1-\bar{X}).$$

Máxima Verossimilhança

Quando a variável de interesse X é contínua, com densidade $f(x \mid \theta)$ tal que

$$\Pr_\theta\{X \leq x\} = \int_{-\infty}^{x} f(x \mid \theta)\, d\theta,$$

a função de verossimilhança é definida de maneira análoga, em termos da densidade conjunta $f(x_1, \ldots, x_n \mid \theta)$, por

$$L(\theta \mid \mathbf{x}) = f(x_1, \ldots, x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta).$$

Neste caso, o método de estimação por máxima verossimilhança diz que devemos escolher como estimativa do parâmetro θ aquele valor que maximize a densidade da amostra observada.

Máxima Verossimilhança

Para desenvolver sua técnica, resolva os exemplos uniparamétricos de máxima verossimilhança dos slides das "Aulas 13, 14 e 15" e da Lista 3.

1. Geométrica: slide 1 (Lista 3, ex. 8).
2. Pareto (slide 17).
3. Exponencial: slide 31.
4. Beta$(\alpha + 1, 1)$: slide 34 (Lista 3, ex. 7)
5. Poisson: slide 35.
6. Bernoulli: slide 39.
7. Weibull: Lista 3, ex. 6.
8. Binomial$(2, \theta)$: Lista 3, ex. 9.

Roteiro geral para testar hipóteses

1. Interprete o enunciado e escreva as hipóteses nula H_0 e alternativa H_A. Interprete os erros do tipo I e II.
2. Escolha a estatística do teste de maneira que, supondo que H_0 é verdadeira, sua distribuição não dependa de parâmetros desconhecidos.
3. Supondo que H_0 é verdadeira, quais valores da estatística do teste desfavorecem H_0 e favorecem H_A? Tais valores definem a Região Crítica RC do teste. Fixando o nível de significância α, construa a RC.
4. Calcule o valor observado da estatística do teste.
5. Se o valor observado da estatística do teste pertencer à RC, rejeite H_0. Caso contrário, não rejeite H_0.

Teste Z

Suponha que a variável de interesse X tem distribuição $\mathrm{N}(\mu, \sigma^2)$, em que a variância populacional σ^2 é **conhecida**, e que temos uma amostra aleatória $X_1, \ldots, X_n$.

Uma hipótese é qualquer afirmação sobre os valores dos parâmetros do modelo.

Suponha que queremos testar a hipótese nula $H_0 : \mu = \mu_0$ contra a hipótese alternativa $H_A : \mu > \mu_0$. Neste caso, a hipótese alternativa H_A é denominada unilateral.

Supondo que H_0 é verdadeira, a estatística do teste não pode depender de parâmetros desconhecidos.

A estatística do Teste Z é

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \;\Big|\; H_0 \sim \mathrm{N}(0, 1).$$

Teste Z

Ao testar a hipótese nula H_0, podemos tomar duas decisões erradas.

O erro do tipo I é a decisão de rejeitar H_0 quando H_0 é verdadeira.

O erro do tipo II é a decisão de não rejeitar H_0 quando H_0 é falsa.

É importante não confundir as decisões erradas com as respectivas probabilidades:

$$\alpha = \Pr\{\text{Rejeitar } H_0 \mid H_0 \text{ é verdadeira}\};$$

$$\beta = \Pr\{\text{Não rejeitar } H_0 \mid H_0 \text{ é falsa}\}.$$

Teste Z

	H_0 é verdadeira	H_0 é falsa
Rejeitar H_0	Erro do tipo I	Decisão correta
Não rejeitar H_0	Decisão correta	Erro do tipo II

Note que a probabilidade do erro do tipo II é de fato uma função do parâmetro μ, pois em geral há infinitos valores de μ para os quais a hipótese nula é falsa, cada um dos quais pode fornecer um valor diferente para β.

A função $\pi(\mu) = \Pr\{\text{Rejeitar } H_0 \mid \mu\}$ é denominada função poder do teste.

Teste Z

Quais valores da estatística do Teste Z desfavorecem H_0 e favorecem H_A?

Em primeiro lugar, observe que $\mathrm{E}[Z \mid H_0] = 0$.

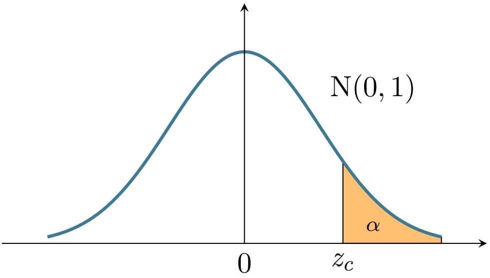
Por outro lado, para algum $\mu_1 > \mu_0$, temos que

$$\begin{aligned}
\mathrm{E}[Z \mid \mu = \mu_1] &= \mathrm{E}\left[\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \,\middle|\, \mu = \mu_1\right] \\
&= \underbrace{\mathrm{E}\left[\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \,\middle|\, \mu = \mu_1\right]}_{=0} + \underbrace{\frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}}_{>0} > 0.
\end{aligned}$$

Este raciocínio nos leva à seguinte regra prática: a região crítica está na cauda indicada pela direção da hipótese alternativa H_A.

Teste Z

Fixando o nível de significância α, encontramos a região crítica $\mathrm{RC} = [z_c; \infty)$.



$$\begin{aligned}\alpha &= \Pr\{\text{Rejeitar } H_0 \mid H_0\} = \Pr\{Z \in \mathrm{RC} \mid H_0\} \\ &= \Pr\{Z \geq z_c \mid H_0\}.\end{aligned}$$

Teste Z

O valor observado da estatística do teste é

$$z_{\text{obs}} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

e o teste fica definido pela seguinte regra de decisão.

Se $z_{\text{obs}} \in \text{RC}$, então rejeito H_0; caso contrário, não rejeito H_0.

Teste Z

O valor-p é o menor nível de significância para o qual rejeitaríamos H_0.

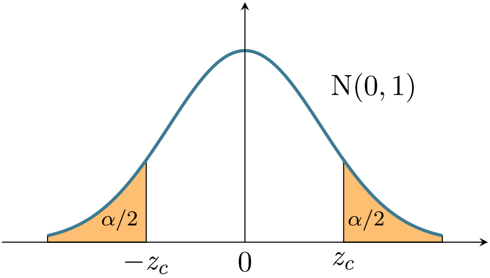
$$\text{valor-}p = \Pr\{Z \geq z_{\text{obs}} \mid H_0\}.$$

Em termos do valor-p, a regra de decisão do teste pode ser escrita da seguinte maneira.

Se o valor-p é menor do que o nível de significância α, então rejeito H_0; caso contrário, não rejeito H_0.

Teste Z

Quando a hipótese alternativa é bilateral da forma $H_A : \mu \neq \mu_0$, a região crítica $\text{RC} = (-\infty; -z_c] \cup [z_c; \infty)$.



Teste Z

No caso bilateral, se $z_{\text{obs}} > 0$, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{Z \geq z_{\text{obs}} \mid H_0\}.$$

Se $z_{\text{obs}} < 0$, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{Z \leq z_{\text{obs}} \mid H_0\}.$$

No caso bilateral, o teste é equivalente à seguinte regra de decisão baseada no intervalo de confiança para μ com coeficiente de confiança $\gamma = 1 - \alpha$.

Se $\mu_0 \notin \text{IC}[\mu; \gamma = 1 - \alpha]$, então rejeito H_0; caso contrário, não rejeito H_0.

$$\text{IC}[\mu; \gamma] = \left(\bar{x} - z\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + z\frac{\sigma}{\sqrt{n}}\right)$$

Teste t

Suponha que a variável de interesse X tem distribuição $\mathrm{N}(\mu, \sigma^2)$, em que a variância populacional σ^2 é **desconhecida**, e que temos uma amostra aleatória $X_1, \ldots, X_n$.

O Teste t é similar ao Teste Z, com a diferença de que estimamos a variância populacional utilizando a variância amostral S^2.

Suponha que queremos testar a hipótese nula $H_0 : \mu = \mu_0$ contra a hipótese alternativa unilateral $H_A : \mu < \mu_0$.

A estatística do Teste t é

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \,\Big|\, H_0 \sim t_{(n-1)}.$$

Teste t

Quais valores da estatística do Teste t desfavorecem H_0 e favorecem H_A? Suponha que $n > 2$.

Em primeiro lugar, observe que $\mathrm{E}[T \mid H_0] = 0$.

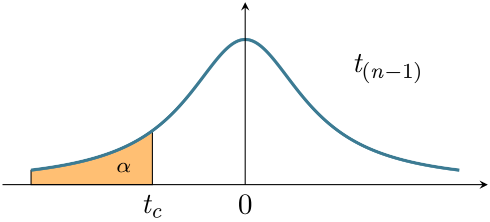
Por outro lado, para algum $\mu_1 < \mu_0$, temos que

$$\begin{aligned}\mathrm{E}[T \mid \mu = \mu_1] &= \mathrm{E}\left[\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \,\middle|\, \mu = \mu_1\right] \\ &= \underbrace{\mathrm{E}\left[\frac{\bar{X} - \mu_1}{S/\sqrt{n}} \,\middle|\, \mu = \mu_1\right]}_{=0} + \underbrace{(\mu_1 - \mu_0)}_{<0} \times \underbrace{\mathrm{E}\left[\frac{1}{S/\sqrt{n}} \,\middle|\, \mu = \mu_1\right]}_{>0} < 0.\end{aligned}$$

Regra prática: a região crítica está na cauda indicada pela direção da hipótese alternativa H_A.

Teste t

Fixando o nível de significância α, encontramos a região crítica $\mathrm{RC} = (-\infty; t_c]$.



$$\begin{aligned}\alpha &= \Pr\{\text{Rejeitar } H_0 \mid H_0\} = \Pr\{T \in \mathrm{RC} \mid H_0\} \\ &= \Pr\{T \leq t_c \mid H_0\}.\end{aligned}$$

Teste t

O valor observado da estatística do teste é

$$t_{\text{obs}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

e o teste fica definido pela seguinte regra de decisão.

Se $t_{\text{obs}} \in \text{RC}$, então rejeito H_0; caso contrário, não rejeito H_0.

Teste t

O valor-p é o menor nível de significância para o qual rejeitaríamos H_0.

Inspecionando a figura anterior, temos que

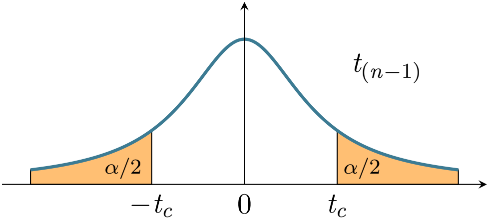
$$\text{valor-}p = \Pr\{\, T \leq t_{\text{obs}} \mid H_0 \}.$$

Em termos do valor-p, a regra de decisão do teste pode ser escrita da seguinte maneira.

Se o valor-p é menor do que o nível de significância α, então rejeito H_0; caso contrário, não rejeito H_0.

Teste t

Quando a hipótese alternativa é bilateral da forma $H_A : \mu \neq \mu_0$, a região crítica $\mathrm{RC} = (-\infty; -t_c] \cup [t_c; \infty)$.



Teste t

No caso bilateral, se $t_{\text{obs}} > 0$, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{\, T \geq t_{\text{obs}} \mid H_0 \}.$$

Se $t_{\text{obs}} < 0$, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{\, T \leq t_{\text{obs}} \mid H_0 \}.$$

No caso bilateral, o teste é equivalente à seguinte regra de decisão baseada no intervalo de confiança para μ com coeficiente de confiança $\gamma = 1 - \alpha$.

Se $\mu_0 \notin \text{IC}[\mu; \gamma = 1 - \alpha]$, então rejeito H_0; caso contrário, não rejeito H_0.

$$\text{IC}[\mu; \gamma] = \left(\bar{x} - t \frac{s}{\sqrt{n}} \; ; \; \bar{x} + t \frac{s}{\sqrt{n}} \right)$$

Teste para a variância

Suponha que a variável de interesse X tem distribuição $\mathrm{N}(\mu, \sigma^2)$ e que temos uma amostra aleatória $X_1, \ldots, X_n$.

Suponha que queremos testar a hipótese nula $H_0 : \sigma^2 = \sigma_0^2$ contra a hipótese alternativa unilateral $H_A : \sigma^2 > \sigma_0^2$.

A estatística do teste para a variância é

$$Q = \frac{(n-1)S^2}{\sigma_0^2} \, \Big| \, H_0 \sim \chi^2_{(n-1)}.$$

Teste para a variância

Quais valores da estatística do teste desfavorecem H_0 e favorecem H_A?

Em primeiro lugar, observe que $\mathrm{E}[Q \mid H_0] = n - 1$.

Por outro lado, para algum $\sigma_1^2 > \sigma_0^2$, temos que

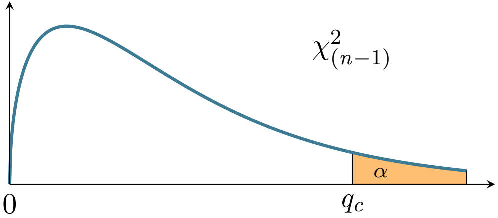
$$\mathrm{E}[Q \mid \sigma^2 = \sigma_1^2] = \mathrm{E}\left[\frac{(n-1)S^2}{\sigma_0^2} \,\middle|\, \sigma^2 = \sigma_1^2\right]$$

$$= \frac{\sigma_1^2}{\sigma_0^2} \times \mathrm{E}\left[\frac{(n-1)S^2}{\sigma_1^2} \,\middle|\, \sigma^2 = \sigma_1^2\right] = \underbrace{\frac{\sigma_1^2}{\sigma_0^2}}_{>1} \times (n-1) > (n-1).$$

Regra prática: a região crítica está na cauda indicada pela direção da hipótese alternativa H_A.

Teste para a variância

Fixando o nível de significância α, encontramos a região crítica $\mathrm{RC} = [q_c; \infty)$.



$$\alpha = \Pr\{\text{Rejeitar } H_0 \mid H_0\} = \Pr\{Q \in \mathrm{RC} \mid H_0\}$$
$$= \Pr\{Q \geq q_c \mid H_0\}.$$

Teste para a variância

O valor observado da estatística do teste é

$$q_{\text{obs}} = \frac{(n-1)s^2}{\sigma_0^2}$$

e o teste fica definido pela seguinte regra de decisão.

Se $q_{\text{obs}} \in \text{RC}$, então rejeito H_0; caso contrário, não rejeito H_0.

Teste para a variância

O valor-p é o menor nível de significância para o qual rejeitaríamos H_0.

Inspecionando a figura anterior, temos que

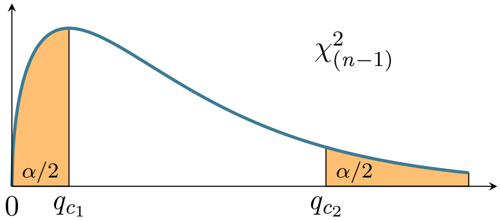
$$\text{valor-}p = \Pr\{Q \geq q_{\text{obs}} \mid H_0\}.$$

Em termos do valor-p, a regra de decisão do teste pode ser escrita da seguinte maneira.

Se o valor-p é menor do que o nível de significância α, então rejeito H_0; caso contrário, não rejeito H_0.

Teste para a variância

Quando a hipótese alternativa é bilateral da forma $H_A : \sigma^2 \neq \sigma_0^2$, a região crítica $\text{RC} = [0; q_{c_1}] \cup [q_{c_2}; \infty)$.



Teste para a variância

No caso bilateral, se q_{obs} cair na cauda da direita, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{Q \geq q_{\text{obs}} \mid H_0\}.$$

Se q_{obs} cair na cauda da esquerda, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{Q \leq q_{\text{obs}} \mid H_0\}.$$

No caso bilateral, o teste é equivalente à seguinte regra de decisão baseada no intervalo de confiança para σ^2 com coeficiente de confiança $\gamma = 1 - \alpha$.

Se $\sigma_0^2 \notin \text{IC}[\sigma^2; \gamma = 1 - \alpha]$, então rejeito H_0; caso contrário, não rejeito H_0.

$$\text{IC}[\sigma^2; \gamma] = \left(\frac{(n-1)s^2}{q_2} \; ; \; \frac{(n-1)s^2}{q_1} \right)$$

Teste t pareado

Suponha que temos uma variável de interesse X (defina $\mu_X = \mathrm{E}[X]$) e uma variável de interesse Y (defina $\mu_Y = \mathrm{E}[Y]$).

No Teste t pareado estamos interessados na variável $\Delta = Y - X$ e iremos supor que $\Delta \sim \mathrm{N}(\mu_\Delta, \sigma^2_\Delta)$.

Segue imediatamente que $\mu_\Delta = \mathrm{E}[\Delta] = \mu_Y - \mu_X$.

Para cada unidade amostral, iremos observar os valores de X e de Y, obtendo a amostra aleatória dos pares

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n).$$

Defina as diferenças dos pares $\Delta_i = Y_i - X_i$, para $i = 1, \ldots, n$.

Defina

$$\bar{\Delta} = \frac{1}{n} \sum_{i=1}^{n} \Delta_i \qquad \text{e} \qquad S^2_\Delta = \frac{1}{n-1} \sum_{i=1}^{n} (\Delta_i - \bar{\Delta})^2.$$

Teste t pareado

Suponha que queremos testar a hipótese nula $H_0 : \mu_\Delta = \delta_0$ contra a hipótese alternativa unilateral $H_A : \mu_\Delta < \delta_0$.

A estatística do Teste t pareado é

$$T = \frac{\bar{\Delta} - \delta_0}{S_\Delta/\sqrt{n}} \Big| H_0 \sim t_{(n-1)}.$$

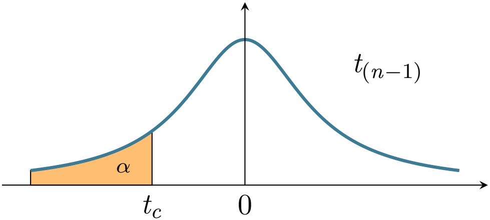
Portanto, o Teste t pareado nada mais é do que um Teste t em que as observações são as diferenças Δ_i.

Quais valores da estatística do Teste t pareado desfavorecem H_0 e favorecem H_A?

Repetindo o argumento usado no Teste t, obtemos a regra prática: a região crítica está na cauda indicada pela direção da hipótese alternativa H_A.

Teste t pareado

Fixando o nível de significância α, encontramos a região crítica $\mathrm{RC} = (-\infty; t_c]$.



$$\begin{aligned}\alpha &= \Pr\{\text{Rejeitar } H_0 \mid H_0\} = \Pr\{T \in \mathrm{RC} \mid H_0\} \\ &= \Pr\{T \le t_c \mid H_0\}.\end{aligned}$$

Teste t pareado

O valor observado da estatística do teste é

$$t_{\text{obs}} = \frac{\bar{\delta} - \delta_0}{s_\Delta / \sqrt{n}}$$

e o teste fica definido pela seguinte regra de decisão.

Se $t_{\text{obs}} \in \text{RC}$, então rejeito H_0; caso contrário, não rejeito H_0.

Teste t pareado

O valor-p é o menor nível de significância para o qual rejeitaríamos H_0.

Inspecionando a figura anterior, temos que

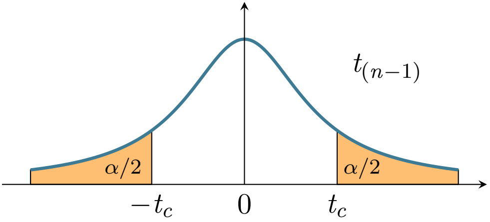
$$\text{valor-}p = \Pr\{\, T \leq t_{\text{obs}} \mid H_0 \}.$$

Em termos do valor-p, a regra de decisão do teste pode ser escrita da seguinte maneira.

Se o valor-p é menor do que o nível de significância α, então rejeito H_0; caso contrário, não rejeito H_0.

Teste t pareado

Quando a hipótese alternativa é bilateral da forma $H_A : \mu_\Delta \neq \delta_0$, a região crítica $\mathrm{RC} = (-\infty; -t_c] \cup [t_c; \infty)$.



Teste t pareado

No caso bilateral, se $t_{\text{obs}} > 0$, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{\, T \geq t_{\text{obs}} \mid H_0 \}.$$

Se $t_{\text{obs}} < 0$, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{\, T \leq t_{\text{obs}} \mid H_0 \}.$$

No caso bilateral, o teste é equivalente à seguinte regra de decisão baseada no intervalo de confiança para μ_Δ com coeficiente de confiança $\gamma = 1 - \alpha$.

Se $\delta_0 \notin \text{IC}[\mu_\Delta; \gamma = 1 - \alpha]$, então rejeito H_0; caso contrário, não rejeito H_0.

$$\text{IC}[\mu_\Delta; \gamma] = \left(\bar{\delta} - t \frac{s_\Delta}{\sqrt{n}} \; ; \; \bar{\delta} + t \frac{s_\Delta}{\sqrt{n}} \right)$$

Duas populações: Teste para as variâncias

Suponha que temos duas populações: na primeira delas iremos observar valores de uma variável de interesse X com distribuição $\mathrm{N}(\mu_1, \sigma_1^2)$, obtendo uma amostra aleatória $X_1, \ldots, X_{n_1}$, e na segunda iremos observar valores de uma variável de interesse Y com distribuição $\mathrm{N}(\mu_2, \sigma_2^2)$, obtendo uma amostra aleatória $Y_1, \ldots, Y_{n_2}$.

Iremos supor que as duas amostras são independentes e queremos testar a hipótese nula $H_0 : \sigma_1^2 = \sigma_2^2$ contra a hipótese alternativa unilateral $H_A : \sigma_1^2 > \sigma_2^2$.

A estatística do teste para as variâncias é

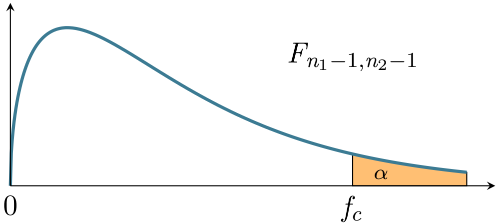
$$F = \frac{S_1^2}{S_2^2} \;\Big|\; H_0 \sim F_{n_1-1, n_2-1},$$

em que $n_1 - 1$ é o número de graus de liberdade do numerador e $n_2 - 1$ é o número de graus de liberdade do denominador.

Regra prática: a região crítica está na cauda indicada pela direção da hipótese alternativa H_A.

Duas populações: Teste para as variâncias

Fixando o nível de significância α, encontramos a região crítica $\mathrm{RC} = [f_c; \infty)$.



$$
\begin{aligned}
\alpha &= \Pr\{\text{Rejeitar } H_0 \mid H_0\} = \Pr\{F \in \mathrm{RC} \mid H_0\} \\
&= \Pr\{F \geq f_c \mid H_0\}.
\end{aligned}
$$

Duas populações: Teste para as variâncias

O valor observado da estatística do teste é

$$f_{\text{obs}} = \frac{s_1^2}{s_2^2}$$

e o teste fica definido pela seguinte regra de decisão.

Se $f_{\text{obs}} \in \text{RC}$, então rejeito H_0; caso contrário, não rejeito H_0.

Duas populações: Teste para as variâncias

O valor-p é o menor nível de significância para o qual rejeitaríamos H_0.

Inspecionando a figura anterior, temos que

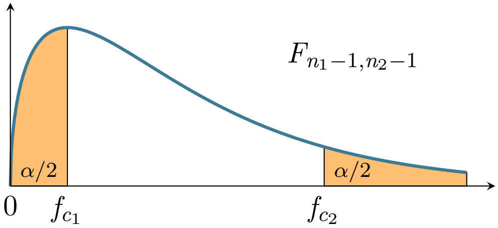
$$\text{valor-}p = \Pr\{F \geq f_{\text{obs}} \mid H_0\}.$$

Em termos do valor-p, a regra de decisão do teste pode ser escrita da seguinte maneira.

Se o valor-p é menor do que o nível de significância α, então rejeito H_0; caso contrário, não rejeito H_0.

Duas populações: Teste para as variâncias

Quando a hipótese alternativa é bilateral da forma $H_A : \sigma_1^2 \neq \sigma_2^2$, a região crítica $\mathrm{RC} = [0; f_{c_1}] \cup [f_{c_2}; \infty)$.



Duas populações: Teste para as variâncias

No caso bilateral, se f_{obs} cair na cauda da direita, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{F \geq f_{\text{obs}} \mid H_0\}.$$

Se f_{obs} cair na cauda da esquerda, o valor-p será dado por

$$\text{valor-}p = 2 \cdot \Pr\{F \leq f_{\text{obs}} \mid H_0\}.$$

Duas populações: Teste para as variâncias

Intervalo de confiança para a razão σ_1^2/σ_2^2.

$$\mathrm{IC}\left[\frac{\sigma_1^2}{\sigma_2^2};\gamma\right] = \left(\frac{1}{f_2}\frac{s_1^2}{s_2^2}\,;\;\frac{1}{f_1}\frac{s_1^2}{s_2^2}\right).$$

No caso bilateral, o teste para as variâncias é equivalente à seguinte regra de decisão baseada no intervalo de confiança para σ_1^2/σ_2^2 com coeficiente de confiança $\gamma = 1 - \alpha$.

Se $\sigma_1^2/\sigma_2^2 = 1 \notin \mathrm{IC}[\sigma_1^2/\sigma_2^2; \gamma = 1 - \alpha]$, então rejeito H_0; caso contrário, não rejeito H_0.

Duas populações: Teste t

Suponha que temos duas populações: na primeira delas iremos observar valores de uma variável de interesse X com distribuição $\mathrm{N}(\mu_1, \sigma_1^2)$, obtendo uma amostra aleatória $X_1, \ldots, X_{n_1}$, e na segunda iremos observar valores de uma variável de interesse Y com distribuição $\mathrm{N}(\mu_2, \sigma_2^2)$, obtendo uma amostra aleatória $Y_1, \ldots, Y_{n_2}$.

Iremos supor que as duas amostras são independentes.

Suponha que queremos testar a hipótese nula $H_0 : \mu_1 - \mu_2 = \delta_0$ contra a hipótese alternativa unilateral $H_A : \mu_1 - \mu_2 > \delta_0$.

Duas populações: Teste t

A estatística do teste dependerá do caso considerado.

Caso 1: variâncias σ_1^2 e σ_2^2 conhecidas.

$$Z = \frac{(\bar{X} - \bar{Y}) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \Bigg| H_0 \sim \mathrm{N}(0, 1).$$

Caso 2: variâncias σ_1^2 e σ_2^2 desconhecidas e iguais.

$$T = \frac{(\bar{X} - \bar{Y}) - \delta_0}{\sqrt{S^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \Bigg| H_0 \sim t_{(n_1+n_2-2)},$$

na qual

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

Duas populações: Teste t

Caso 3: variâncias σ_1^2 e σ_2^2 desconhecidas e diferentes.

$$T = \frac{(\bar{X} - \bar{Y}) - \delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \,\Bigg|\, H_0 \approx t_{(k)},$$

na qual

$$k = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}.$$

Note que em geral k não é um número inteiro, sendo necessário arredondar para o inteiro mais próximo para encontrar o número de graus de liberdade.

Abreviações para as suposições mais frequentes

AA = A população é suficientemente grande para que o modelo de observáveis independentes e identicamente distribuídos (amostra aleatória simples com reposição) seja adequado.

Note que não fazemos esta suposição quando utilizamos uma *correção de população finita*.

Quando o problema envolver duas populações, passe a afirmativa para o plural.

Abreviações para as suposições mais frequentes

$X \sim \mathrm{N}(\mu, \sigma^2) =$ A variável de interesse tem distribuição normal.

Verifique cuidadosamente se esta suposição foi realmente utilizada.

Se fizermos esta suposição, significa que não iremos utilizar o TLC.

Muito cuidado para não errar: apenas sob esta suposição podemos utilizar as distribuições t, χ^2 e F.

Em um problema com duas populações teremos uma segunda variável de interesse Y.

Abreviações para as suposições mais frequentes

TLC = O tamanho da amostra é suficientemente grande para que possamos utilizar o Teorema do Limite Central (TLC) para aproximar a distribuição da média amostral.

A regra prática é que a aproximação do TLC é adequada quando a amostra tem tamanho $n \geq 30$.

Quando o problema envolver duas populações, passe a afirmativa para o plural.

2AI = As duas amostras são independentes.

Utilizada em problemas com duas populações.

Cuidado para não confundir problemas de duas populações com problemas pareados, nos quais há apenas uma população, mas fazemos duas observações para cada unidade amostral.