

Estatística

Formulário

Versão 1.0

Teoria da probabilidade

Propriedades

$$P(A) + P(\bar{A}) = 1$$

acontecimentos complementares

$$P(\emptyset) = 0$$

conjunto vazio

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Reunião

Probabilidade condicional

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Acontecimentos independentes

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

$$P(A \cap B) = P(A|B) \cdot P(B) = P(A) \cdot P(B)$$

Teorema de Bayes

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B|A_1) \cdot P(A_1) + \dots + P(B|A_N) \cdot P(A_N)}$$

Variáveis aleatórias Distribuições de probabilidade

Distribuições discretas

$$0 \leq P(y) \leq 1$$

função de probabilidade

$$\sum_y P(y) = 1$$

$$F(a) = P(Y \leq a) = \sum_{y \leq a} P(y)$$

função distribuição

Distribuições contínuas

$$P(a \leq X \leq b) = \int_a^b f(x) \cdot dx$$

função densidade de probabilidade

$$\int_{-\infty}^{+\infty} f(x) \cdot dx = 1$$

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) \cdot du$$

função distribuição

$$\frac{dF(x)}{dx} = f(x)$$

Média populacional

$$\mu_Y = E(Y) = \sum_y y \cdot P(y)$$

variáveis discretas

$$\mu_X = E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx$$

variáveis contínuas

Mediana	$F^{-1}\left(\frac{1}{2}\right) = \eta$	
Variância populacional	$\sigma_Y^2 = Var(Y) = \sum_y (y - \mu_Y)^2 \cdot P(y)$	variáveis discretas
	$\sigma_X^2 = Var(X) = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \cdot f(x) \cdot dx$	variáveis contínuas
Outros parâmetros	$\mu'_i = \sum_y y^i \cdot P(Y)$	momentos na origem (variáveis discretas)
	$\mu'_i = \int_{-\infty}^{+\infty} x^i \cdot f(x) \cdot dx$	momentos na origem (variáveis contínuas)
	$\mu_i = \sum_y (y - \mu_Y)^i \cdot P(Y)$	momentos centrados (variáveis discretas)
	$\mu_i = \int_{-\infty}^{+\infty} (x - \mu_X)^i \cdot f(x) \cdot dx$	momentos centrados (variáveis contínuas)
	$\mu_2 = \mu'_2 - (\mu'_1)^2$	
	$\mu_3 = \mu'_3 - 3 \cdot \mu'_2 \cdot \mu'_1 + 2 \cdot (\mu'_1)^3$	momentos centrados a partir de momentos na origem
	$\mu_4 = \mu'_4 - 4 \cdot \mu'_3 \cdot \mu'_1 + 6 \cdot \mu'_2 \cdot (\mu'_1)^2 - 3 \cdot (\mu'_1)^4$	
Coeficiente de assimetria	$\tau = \frac{\mu_3}{\sigma^3}$	
Coeficiente de kurtose	$\kappa = \frac{\mu_4}{\sigma^4} - 3$	
Transformações lineares	$W = a + b \cdot Z$	Transformação
	$E(W) = E(a + b \cdot Z) = a + b \cdot E(Z)$	valor esperado
	$Var(W) = Var(a + b \cdot Z) = b^2 \cdot Var(Z)$	variância
Transformações não lineares	$W = \phi(Z)$	Transformação
	$E(W) = E[\phi(Z)] \approx \phi[E(Z)]$	valor esperado
	$Var(W) = Var[\phi(Z)] \approx \left[\frac{d\phi}{dZ} \right]_{Z=\mu_Z}^2 \cdot Var(Z)$	variância

Distribuições conjuntas de probabilidade

2 distribuições discretas	$P(y_1, y_2) : \sum_{y_1} \sum_{y_2} P(y_1, y_2) = 1$	função conjunta de probabilidade
	$P_{Y_1}(y_1) = \sum_{y_2} P(y_1, y_2)$	funções de probabilidade marginais
	$P_{Y_2}(y_2) = \sum_{y_1} P(y_1, y_2)$	
	$F(a) = P(Y \leq a) = \sum_{y \leq a} P(y)$	função distribuição
2 distribuições contínuas	$f(x_1, x_2) : \iint_{\mathbb{R}^2} f(x_1, x_2) \cdot dx_1 \cdot dx_2 = 1$	função conjunta de densidade de probabilidade
	$f_{X_1}(x_1) = \int_{\mathbb{R}} f(x_1, x_2) \cdot dx_2$	Funções densidade de probabilidade marginais
	$f_{X_2}(x_2) = \int_{\mathbb{R}} f(x_1, x_2) \cdot dx_1$	
	$\frac{dF(x)}{dx} = f(x)$	
1 distribuição discreta, 1 distribuição contínua	$f(x, y) : \sum_y \left[\int_{\mathbb{R}} f(x, y) \cdot dx \right] = 1$	função conjunta de probabilidade
	$f_X(x) = \sum_y f(x, y)$	função densidade de probabilidade marginal
	$P_Y(y) = \int_{\mathbb{R}} f(x, y) \cdot dx$	função de probabilidade marginal
	$\frac{dF(x)}{dx} = f(x)$	
Covariância	$\gamma_{Y_1, Y_2} = \sum_{y_1} \sum_{y_2} (y_1 - \mu_{Y_1}) \cdot (y_2 - \mu_{Y_2}) \cdot P(y_1, y_2)$	2 variáveis discretas
	$\gamma_{X_1, X_2} = \iint_{\mathbb{R}} (x_1 - \mu_{X_1}) \cdot (x_2 - \mu_{X_2}) \cdot f(x_1, x_2) \cdot dx_1 \cdot dx_2$	2 variáveis contínuas
	$\gamma_{X, Y} = \sum_y (y - \mu_Y) \cdot \int_{\mathbb{R}} (x - \mu_X) \cdot f(x, y) \cdot dx$	1 variável discreta, 1 variável contínua
Coeficiente de correlação	$\rho_{WZ} = \frac{\gamma_{WZ}}{\sigma_W \cdot \sigma_Z}$	para quaisquer variáveis
Transformações lineares	$V = a + b \cdot W + c \cdot Z$	Transformação
	$E(V) = E(a + b \cdot W + c \cdot Z) = a + b \cdot E(W) + c \cdot E(Z)$	valor esperado

$$\begin{aligned} Var(V) &= Var(a + b \cdot W + c \cdot Z) \\ &= b^2 \cdot Var(W) + c^2 \cdot Var(Z) + 2 \cdot b \cdot c \cdot Cov(Z, W) \end{aligned} \quad \text{variância}$$

Soma de N variáveis

$$S = X_1 + X_2 + \dots + X_N \quad \text{Transformação}$$

$$E(S) = E(X_1) + E(X_2) + \dots + E(X_N) \quad \text{valor esperado}$$

$$Var(S) = Var(X_1) + \dots + Var(X_N) + 2 \cdot \sum_{i=1}^{N-1} \sum_{j=i+1}^N Cov(X_i, X_j) \quad \text{variância}$$

Transformações não lineares

$$V = \phi(W, Z) \quad \text{Transformação}$$

$$E(V) = E[\phi(W, Z)] \approx \phi[E(W), E(Z)] \quad \text{valor esperado}$$

$$\begin{aligned} Var(V) \approx & \left[\frac{\partial \phi}{\partial W} \right]_{W=\mu_W, Z=\mu_Z}^2 \cdot \sigma_W^2 + \left[\frac{\partial \phi}{\partial Z} \right]_{W=\mu_W, Z=\mu_Z}^2 \cdot \sigma_Z^2 + \\ & + 2 \cdot \left[\frac{\partial \phi}{\partial W} \right]_{W=\mu_W, Z=\mu_Z}^2 \cdot \left[\frac{\partial \phi}{\partial Z} \right]_{W=\mu_W, Z=\mu_Z}^2 \cdot \gamma_{WZ} \end{aligned} \quad \text{variância}$$

Distribuições discretas

Distribuição Binomial B(n,p)

$$P(y) = C_y^n \cdot p^y \cdot q^{n-y} \quad \text{função de probabilidade}$$

$$\mu = n \cdot p \quad \text{média}$$

$$\sigma^2 = n \cdot p \cdot q \quad \text{variância}$$

Distribuição Binomial Negativa BN(r,p)

$$P(y) = C_y^{y+r-1} \cdot p^r \cdot q^y \quad \text{função de probabilidade}$$

$$\mu = \frac{r \cdot q}{p} \quad \text{média}$$

$$\sigma^2 = \frac{r \cdot q}{p^2} \quad \text{variância}$$

Distribuição Geométrica G(p)
(caso particular da distribuição Binomial Negativa em que r=1)

$$P(y) = p \cdot q^y \quad \text{função de probabilidade}$$

$$\mu = \frac{q}{p} \quad \text{média}$$

$$\sigma^2 = \frac{q}{p^2} \quad \text{variância}$$

Distribuição Hipergeométrica H(M,p,M,q,n)

$$P(y) = \frac{C_y^{M-p} \cdot C_{n-y}^{M-q}}{C_n^M} \quad \text{função de probabilidade}$$

$$\mu = n \cdot p \quad \text{média}$$

$$\sigma^2 = n \cdot p \cdot q \cdot \frac{M-n}{M-1}$$

variância

Distribuição de Poisson	$P(y) = \frac{e^{-\lambda} \cdot \lambda^y}{y!}$	função de probabilidade
Poisson(λ)	$\mu = \lambda$	média
	$\sigma^2 = \lambda$	variância

Distribuições contínuas

Distribuição Uniforme U(a,b)	$f(x) = \begin{cases} 1/(b-a) & a \leq x \leq b \\ 0 & \text{outros } x \end{cases}$	função densidade de probabilidade
	$F(x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \leq x \leq b \\ 1 & x > b \end{cases}$	função distribuição de probabilidade
	$\mu = (a+b)/2$	média
	$\sigma^2 = \frac{(b-a)^2}{12}$	variância

Distribuição Exponencial Negativa EN(β)	$f(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}} & \text{outros } x \end{cases}$	função densidade de probabilidade
	$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\frac{x}{\beta}} & \text{outros } x \end{cases}$	função distribuição de probabilidade
	$\mu = \beta$	média
	$\sigma^2 = \beta^2$	variância

Distribuição Normal N(μ, σ^2)	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$	função densidade de probabilidade
	$f(x)$ não é integrável analiticamente	função distribuição de probabilidade
	$\mu = \mu$	média
	$\sigma^2 = \sigma^2$	Variância

Distribuição Normal Padronizada N(0,1)	$Z = \frac{X - \mu}{\sigma}$ $f(z) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{z^2}{2}}$ $f(x)$ não é integrável analiticamente $\mu = 0$ $\sigma^2 = 1$	sendo X uma variável $N(\mu, \sigma^2)$ função densidade de probabilidade função distribuição de probabilidade média variância	definição
Distribuição Lognormal LN(μ_V, σ_V^2)	$V = \ln X$ $f(x) = \frac{1}{x \cdot \sqrt{2 \cdot \pi \cdot \sigma_V^2}} \cdot e^{-\frac{1}{2\sigma_V^2}(\ln x - \mu_V)^2}$ $f(x)$ não é integrável analiticamente $\mu = e^{\mu_V + \sigma_V^2/2}$ $\sigma^2 = e^{2(\mu_V + \sigma_V^2)} - e^{2\mu_V + \sigma_V^2}$ $\eta = e^{\mu_V}$ $\zeta = e^{(\mu_V - \sigma_V^2)}$	sendo V uma variável $N(\mu_V, \sigma_V^2)$ função densidade de probabilidade função distribuição de probabilidade média variância mediana moda	definição
Distribuição do Qui-Quadrado χ_{GL}^2	$X = \sum_{i=1}^{GL} Z_i^2$ $f(x) = \frac{1}{2^{GL/2} \cdot \Gamma(GL/2)} \cdot e^{-x/2} \cdot x^{(GL/2)-1}$ $f(x)$ não é integrável analiticamente $\mu = GL$ $\sigma^2 = 2 \cdot GL$	sendo Z_i variáveis $N(0,1)$ função densidade de probabilidade função distribuição de probabilidade média variância	definição
Distribuição t de Student t_{GL}	$X = \frac{Z}{\sqrt{V/GL}}$ $f(x) = \frac{\Gamma((GL+1)/2)}{\sqrt{GL \cdot \pi} \cdot \Gamma(GL/2)} \cdot \left(1 + \frac{x^2}{GL}\right)^{-(GL+1)/2}$ $f(x)$ não é integrável analiticamente $\mu = 0$	sendo Z e V variáveis $N(0,1)$ e χ_{GL}^2 função densidade de probabilidade função distribuição de probabilidade média	definição

$$\sigma^2 = GL/(GL - 2)$$

variância

Distribuição F

F_{GL_1,GL_2}

$$X = \frac{\chi_{GL_1}^2/GL_1}{\chi_{GL_2}^2/GL_2}$$

definição

$$f(x) = \frac{\Gamma((GL_1 + GL_2)/2)}{\Gamma(GL_1/2) \cdot \Gamma(GL_2/2)} \cdot \frac{(GL_1/GL_2)^{GL_1/2} \cdot x^{(GL_1/2)-1}}{(GL_2 + GL_1 \cdot x)^{(GL_1+GL_2)/2}}$$

função densidade de probabilidade

$f(x)$ não é integrável analiticamente

função distribuição de probabilidade

$$\mu = \frac{GL_2}{GL_2 - 1}$$

média

$$\sigma^2 = \frac{2 \cdot GL_2 \cdot (GL_1 + GL_2 - 2)}{GL_1 \cdot (GL_2 - 2)^2 \cdot (GL_2 - 4)}$$

variância

$$F_{GL_1,GL_2}(1 - \alpha) = \frac{1}{F_{GL_2,GL_1}(\alpha)}$$

propriedade

Estatística descritiva

Média

$$\bar{x} = \frac{1}{N} \cdot \sum_{i=1}^N x_i$$

dados não agrupados

$$\bar{x} = \sum_{k=1}^K f_k \cdot x_k$$

dados discretos agrupados

$$\bar{x} = \sum_{k=1}^K f_k \cdot M_k$$

dados contínuos agrupados

Mediana

$$Med = x_{(N+1)/2}$$

n.º ímpar de valores

$$Med = (x_{N/2} + x_{N/2+1})/2$$

n.º par de valores

$$Med = LI + \frac{0.5 - fa^-}{fa^+ - fa^-} \cdot \Delta = LI + \frac{0.5 - fa^-}{f} \cdot \Delta$$

dados agrupados

Moda

$$Mod = LI + \frac{d_1}{d_1 + d_2} \cdot \Delta$$

dados agrupados

Erro absoluto médio

$$EAM = \frac{1}{N} \cdot \sum_{i=1}^N |x_i - \bar{x}|$$

dados não agrupados

$$EAM = \sum_{k=1}^K f_k \cdot |x_k - \bar{x}|$$

dados discretos agrupados

$$EAM = \sum_{k=1}^K f_k \cdot |M_k - \bar{x}|$$

dados contínuos agrupados

Erro quadrático médio	$EQM = \frac{1}{N} \cdot \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{1}{N} \cdot \sum_{i=1}^N x_i^2 - \bar{x}^2$	dados não agrupados
	$EQM = \sum_{k=1}^K f_k \cdot (x_k - \bar{x})^2 = \sum_{k=1}^K f_k \cdot x_k^2 - \bar{x}^2$	dados discretos agrupados
	$EQM \approx \sum_{k=1}^K f_k \cdot (M_k - \bar{x})^2 = \sum_{k=1}^K f_k \cdot M_k^2 - \bar{x}^2$	dados contínuos agrupados
	$EQM \approx \sum_{k=1}^K f_k \cdot (M_k - \bar{x})^2 - \frac{\Delta^2}{12}$	correcção de Sheppard
Variância amostral	$s^2 = \frac{N}{N-1} \cdot EQM$	Estimador não enviesado de
Momentos na origem de ordem 1, 2, ...	$m'_i = \frac{1}{N} \sum_{n=1}^N x_n^i$	dados não agrupados
	$m'_i = \sum_{i=1}^K f_i x_k^i$	dados discretos agrupados
	$m'_i \approx \sum_{i=1}^K f_i M_k^i$	dados contínuos agrupados
	$m'_1 = \bar{x}$	notas
Momentos centrados de ordem 1, 2, ...	$m_i = \frac{1}{N} \sum_{n=1}^N (x_n - \bar{x})^i$	dados não agrupados
	$m_i = \sum_{k=1}^K f_k (x_k - \bar{x})^i$	dados discretos agrupados
	$m_i = \sum_{k=1}^K f_k (M_k - \bar{x})^i$	dados contínuos agrupados
	$m_1 = 0 \quad m_2 = EQM$	
	Para o cálculo de momentos centrados a partir de momentos na origem, podem utilizar-se as expressões apresentadas para os momentos populacionais	Notas
Estimadores não enviesados dos momentos centrados	$k_2 = s^2 = \frac{N}{N-1} \cdot EQM$	2.ª ordem
	$k_3 = \frac{N^2}{(N-1)(N-2)} m_3$	3.ª ordem
	$k_4 = \frac{N(N^2 - 2N + 3)}{(N-1)(N-2)(N-3)} m_4 - \frac{3N(2N-3)}{(N-1)(N-2)(N-3)} m_2^2$	4.ª ordem
Coeficiente de assimetria	$CA = \frac{m_3}{m_2^{3/2}}$	enviesado

	$CA' = \frac{k_3}{k_2^{3/2}}$	enviesado (menos que o anterior)
	$CA'' = \frac{\sqrt{N(N-1)}}{N-2} CA$	não enviesado para populações normais
Coeficiente de kurtose (achatamento)	$CK = \frac{m_4}{m_2^2} - 3$	Enviesado
	$CK' = \frac{k_4}{k_2^2} - 3$	enviesado (menos que o anterior)
	$CK'' = \frac{N-1}{(N-2)(N-3)} ((N+1)CK + 6)$	não enviesado para populações normais
Recta de regressão	$y = a' + b \cdot x$ $b = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}$ $a' = \bar{y} - b \cdot \bar{x}$	Declive ordenada na origem
Covariância amostral	$c_{XY} = \frac{1}{N-1} \cdot \sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})$	
Coeficiente de correlação amostral	$r_{XY} = \frac{c_{XY}}{s_X \cdot s_Y}$	

Distribuições amostrais

Média amostral	$\bar{X} = \frac{1}{N} \cdot \sum_{i=1}^N X_i$	
	$\mu_{\bar{X}} = \mu_X$	Média
	$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{N}$	variância (amostras aleatórias simples)
	$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{N} \cdot \frac{M-N}{M-1}$	variância (populações finitas, amostragem sem reposição)
Média amostral de uma distribuição Normal	$X \rightarrow N(\mu_X, \sigma_X^2)$	distribuição de X
	$\bar{X} \rightarrow N\left(\mu_X, \frac{\sigma_X^2}{N}\right)$	Distribuição de \bar{X}

Estimação pontual

Propriedades dos estimadores	$Enviesamento_{\hat{\Theta}} = E(\hat{\Theta}) - \theta = \mu_{\hat{\Theta}} - \theta$	não-enviesamento
	$Eficiência_{\hat{\Theta}} = E((\hat{\Theta} - \theta)^2) = \sigma_{\hat{\Theta}}^2 + (Enviesamento_{\hat{\Theta}})^2$	eficiência
	$Eficiência relativa_{\hat{\Theta}_1/\hat{\Theta}_2} = \frac{Eficiência_{\hat{\Theta}_1}}{Eficiência_{\hat{\Theta}_2}}$	eficiência relativa
	$\lim_{N \rightarrow \infty} (\hat{\Theta} - \theta) < \delta) = 1 \quad \forall \delta > 0$	consistência
Estimação pela máxima verosimilhança	$Max : L(\theta_1, \theta_2, \dots, \theta_R)$	maximizar L
	$L = \prod_{i=1}^N P(Y_i = y_i \theta_1, \theta_2, \dots, \theta_R)$	distribuições discretas
	$L = \prod_{i=1}^N f(x_i \theta_1, \theta_2, \dots, \theta_R)$	distribuições contínuas

Intervalos de confiança/Testes de hipóteses

Valor esperado	Populações Normais com variância conhecida ou amostras de grande dimensão	
μ	$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$	distribuição
	$\bar{x} \pm z(\alpha/2) \cdot \frac{\sigma}{\sqrt{n}}$	intervalo de confiança
	Se a amostra for grande, $\sigma \approx s$ (embora desconhecido)	notas
	Populações Normais com variância desconhecida	
	$\frac{\bar{x} - \mu}{s/\sqrt{n}} \rightarrow t_{n-1}$	Distribuição
	$\bar{x} \pm t_{n-1}(\alpha/2) \cdot \frac{s}{\sqrt{n}}$	intervalo de confiança
	Se a amostra for pequena, a aproximação $\sigma \approx s$ não é válida	Notas

Proporção binomial	Amostras de grande dimensão	
p	$\frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}} \rightarrow N(0,1)$	Distribuição
	$\hat{p} \pm z(\alpha/2) \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$	intervalo de confiança

Valido sob as condições da aproximação da distribuição binomial à distribuição normal: $n \geq 20$ e $n \cdot p \geq 7$. Fora destas condições deve recorrer-se ao cálculo exacto através da distribuição binomial.

Na amostragem sem reposição, se a dimensão da população, N, não for muito maior do que a dimensão da amostra, n, deve-se ter em consideração o factor $(N - n)/(N - 1)$ no cálculo do erro padrão

Notas

Variância

Populações Normais

σ^2

$$\frac{(n-1) \cdot s^2}{\sigma^2} \rightarrow \chi_{n-1}^2$$

Distribuição

$$\left[\frac{(n-1) \cdot s^2}{\chi_{n-1}^2(\alpha/2)}, \frac{(n-1) \cdot s^2}{\chi_{n-1}^2(1-\alpha/2)} \right]$$

intervalo de confiança

Diferença entre valores esperados

Amostras independentes de populações Normais com variâncias conhecidas e diferentes ou amostras independentes de grande dimensão com variâncias diferentes

$\mu_A - \mu_B$

$$\frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \rightarrow N(0,1)$$

distribuição

$$(\bar{x}_A - \bar{x}_B) \pm z(\alpha/2) \cdot \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}$$

intervalo de confiança

Se a amostra for grande, $\sigma \approx s$ (embora desconhecido)

notas

Amostras independentes de populações Normais com variâncias conhecidas e iguais ou amostras independentes de grande dimensão com variâncias iguais

$$\frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sigma \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \rightarrow N(0,1)$$

distribuição

$$(\bar{x}_A - \bar{x}_B) \pm z(\alpha/2) \cdot \sigma \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$$

intervalo de confiança

Se a amostra for grande

$$\sigma^2 \approx s^2 = \frac{(n_A - 1) \cdot s_A^2 + (n_B - 1) \cdot s_B^2}{n_A + n_B - 2}$$

notas

Amostras independentes de populações Normais com variâncias desconhecidas e diferentes

$$\frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} \rightarrow t_{GL}$$

distribuição

$$(\bar{x}_A - \bar{x}_B) \pm t_{GL}(\alpha/2) \cdot \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

intervalo de confiança

$$GL = \frac{\left(\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}\right)^2}{\frac{\left(\frac{s_A^2}{n_A}\right)^2}{n_A - 1} + \frac{\left(\frac{s_B^2}{n_B}\right)^2}{n_B - 1}}$$

notas

Amostras independentes de populações Normais com variâncias desconhecidas e iguais

$$\frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{s \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \rightarrow t_{GL}$$

distribuição

$$(\bar{x}_A - \bar{x}_B) \pm t_{GL}(\alpha/2) \cdot s \cdot \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$$

intervalo de confiança

$$s^2 = \frac{(n_A - 1) \cdot s_A^2 + (n_B - 1) \cdot s_B^2}{n_A + n_B - 2} \quad GL = n_A + n_B - 2$$

notas

Diferença de proporções binomiais

p_A-p_B

Amostras independentes de grande dimensão

$$\frac{(\hat{p}_A - \hat{p}_B) - (p_A - p_B)}{\sqrt{\frac{p_A \cdot q_A}{n_A} + \frac{p_B \cdot q_B}{n_B}}} \rightarrow N(0, 1)$$

Distribuição

$$(\hat{p}_A - \hat{p}_B) \pm z(\alpha/2) \cdot \sqrt{\frac{\hat{p}_A \cdot \hat{q}_A}{n_A} + \frac{\hat{p}_B \cdot \hat{q}_B}{n_B}}$$

intervalo de confiança

Ver notas para a proporção binomial

Notas

Razão de variâncias

σ^2_A/σ^2_B

Populações Normais

$$\frac{s_A^2/\sigma_A^2}{s_B^2/\sigma_B^2} \rightarrow F_{GL_1, GL_2}$$

Distribuição

$$\left[\frac{s_A^2/s_B^2}{F_{GL_1, GL_2}(\alpha/2)}, \frac{s_A^2/s_B^2}{F_{GL_1, GL_2}(1 - \alpha/2)} \right]$$

intervalo de confiança

$$GL_1 = n_A - 1$$

$$GL_2 = n_B - 1$$

Notas

Proporção binomial

p

A utilizar no cálculo de intervalos de confiança quando não for possível ou desejável a aproximação à distribuição Normal

$$p_i : P(Y \geq \hat{Y}) = \frac{\alpha}{2} \Leftrightarrow \sum_{y=\hat{Y}}^n C_y^n \cdot p_i^y \cdot (1 - p_i)^{n-y} = \frac{\alpha}{2}$$

limite inferior

$$p_s : P(Y \leq \hat{Y}) = \frac{\alpha}{2} \Leftrightarrow \sum_{y=0}^{\hat{Y}} C_y^n \cdot p_s^y \cdot (1 - p_s)^{n-y} = \frac{\alpha}{2}$$

limite superior

Outras expressões

Cálculo combinatório

$$P_n = n!$$

Permutações

$$A_k^n = \frac{n!}{(n-k)!}$$

Arranjos

$$C_k^n = \binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$$

Combinações

Média geométrica

$$MG = \sqrt[N]{x_1 \cdot x_2 \cdot \dots \cdot x_N}$$

Média harmónica

$$MH = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}}$$

Notas

- ◆ Durante o cálculo deve, sempre que possível, trabalhar as expressões analiticamente e efectuar os cálculos todos no final.
- ◆ Sempre que possível, deve trabalhar com 5 algarismos significativos e nunca com menos de 3 algarismos significativos, de modo a não introduzir desvios significativos nos resultados. O ideal é guardar resultados intermédios na memória da máquina de calcular e trabalhar com a precisão que aquela permitir.