Formulário

Fórmula de Sturges	$K \approx 1 + 3.222\ \log_{10}(n);\quad K =$ número de classes
Quantil de ordem p	$q_p = \begin{cases} \frac{x_{(np)}+x_{(np+1)}}{2} & \Leftarrow np \text{ inteiro} \\ x_{(k+1)} & \Leftarrow np \text{ não inteiro } (k \text{ é a parte inteira de } np) \end{cases}$
Variância amostral	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\Big(\sum_{i=1}^{n}(X_i)^2 - \frac{1}{n}\Big(\sum_{i=1}^{n}X_i\Big)^2\Big)$
Variância amostral ponderada	$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$
Covariância amostral	$S_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})$ $= \frac{1}{n-1}\Big(\sum_{i=1}^{n} X_iY_i - \frac{1}{n}\Big(\sum_{i=1}^{n} X_i\Big)\Big(\sum_{i=1}^{n} Y_i\Big)\Big)$
Coeficiente correlação amostral (Pearson)	$R_P = \frac{S_{X,Y}}{S_X S_Y}$
Distibuição Binomial $B(n,p)$	$f(x) = \begin{cases} \binom{n}{x}p^x(1-p)^{n-x} & \text{para } x = 0,1,\ldots,n;\ 0 \le p \le 1 \\ 0 & \text{caso contrário,} \end{cases}$ $E(X) = np;\ V(X) = np(1-p)$
Distibuição de Poisson $P(\lambda)$	$f(x) = \begin{cases} e^{-\lambda}\frac{\lambda^x}{x!} & \text{para } x = 0,1,\ldots;\ \lambda > 0, \\ 0 & \text{caso contrário.} \end{cases}$ $E(X) = V(X) = \lambda$

Inferência Estatística

Notação	$P(Z > z_{\alpha/2}) = \alpha/2,\ Z \sim N(0,1);\quad P(T > t_{(\alpha/2,n)}) = \alpha/2,\ T \sim t_n$
Erro padrão de $\overline{X}$	$se = \dfrac{s}{\sqrt{n}}$
Erro padrão de $\overline{X}_1 - \overline{X}_2$	$se^* = \sqrt{se_1^2 + se_2^2} = \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$
Erro padrão de $\widetilde{P}$	$se_{\widetilde{P}} = \sqrt{\dfrac{\tilde{p}(1-\tilde{p})}{n+4}}$
Graus de liberdade	$gl \approx \dfrac{(se_1^2 + se_2^2)^2}{\frac{se_1^4}{n_1-1} + \frac{se_2^4}{n_2-1}};\quad gl^* = n_1 + n_2 - 2$

Populações normais	
IC para a média (variância conhecida)	$\left(\overline{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\ ,\ \overline{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right);\quad \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
IC para a média (variância desconhecida)	$\left(\overline{x} - t_{(\alpha/2,n-1)}se\ ,\ \overline{x} + t_{(\alpha/2,n-1)}se\right);\quad \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$
IC para a diferença de médias (variâncias conhecidas)	$\overline{x}_1 - \overline{x}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}};\quad \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$
IC para a diferença de médias (variância desconhecidas)	$\overline{x}_1 - \overline{x}_2 \pm t_{(\alpha/2,gl)}\ se^*;\quad \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \dot\sim t_{gl}$

Populações não normais (grau aproximado)	
IC para a média (variância desconhecida)	$\left(\overline{x} - z_{\alpha/2}\ se\ ,\ \overline{x} + z_{\alpha/2}\ se\right)$; $\quad \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \mathrel{\dot\sim} N(0,1)$
IC para a diferença de médias (variância desconhecidas)	$\overline{x}_1 - \overline{x}_2 \pm z_{\alpha/2} \sqrt{se_1^2 + se_2^2}$; $\quad \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \mathrel{\dot\sim} N(0,1)$
IC para p (Wald)	$\left(\hat{p} - z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}\right)$; $\quad \dfrac{\hat{P} - p}{\sqrt{p(1-p)/n}} \mathrel{\dot\sim} N(0,1)$
IC para p (Agresti-Coull grau de confiança aproximado: 0.95)	$\tilde{p} \pm 1.96\sqrt{\dfrac{\tilde{p}(1-\tilde{p})}{n+4}}$; $\quad \tilde{p} = \dfrac{y+2}{n+4}$
IC para $p_1 - p_2$ (Agresti-Coull)	$\tilde{p_1} - \tilde{p_2} \pm z_{\alpha/2}\sqrt{\dfrac{\tilde{p_1}(1-\tilde{p_1})}{n_1+2} + \dfrac{\tilde{p_2}(1-\tilde{p_2})}{n_2+2}}$; $\quad \tilde{p_1} = \dfrac{y_1+1}{n_1+2}$; $\quad \tilde{p_2} = \dfrac{y_2+1}{n_2+2}$

Estatísticas dos testes do qui-quadrado	
Notação	O_i : frequência observada da categoria i e_i : frequência esperada da categoria i (sob H_0)
Ajustamento	$X^2 = \sum_{i=1}^{k} \dfrac{(O_i - e_i)^2}{e_i} \mathrel{\dot\sim} \chi^2_{k-1}$
Homogeneidade Independência	$X^2 = \sum_{i=1}^{k} \dfrac{(O_i - e_i)^2}{e_i} \mathrel{\dot\sim} \chi^2_{(r-1)(c-1)}$ r : número de linhas; c : número de colunas; k : número de células