ESTATÍSTICA E DELINEAMENTO

FORMULÁRIO

Descrição	Fórmula
Covariância amostral	$cov_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})\,y_i$
Propriedades covariância de variáveis aleatórias	$Cov[a+bX, c+dY] = bd\,Cov[X,Y]$; $Cov[X,X]=V[X]$
	$Cov[X \pm Y\,,\,Z] = Cov[X\,,Z] \pm Cov[Y\,,\,Z]$
Regressão Linear Simples	
Estimador do declive da recta β_1	$\hat{\beta}_1 = \frac{Cov_{xY}}{s_x^2} = \sum_{i=1}^{n} c_i Y_i,$ com $c_i = \frac{x_i-\overline{x}}{(n-1)\,s_x^2}.$
Estimador da ordenada na origem β_0	$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{x} = \sum_{i=1}^{n} d_i Y_i,$ com $d_i = \frac{1}{n} - \frac{(x_i-\overline{x})\,\overline{x}}{(n-1)\,s_x^2}.$
Variância dos estimadores $\hat{\beta}_1$ e $\hat{\beta}_0$	$Var(\hat{\beta}_1) = \frac{\sigma^2}{(n-1)\,s_x^2}$, $Var(\hat{\beta}_0) = \sigma^2\left(\frac{1}{n} + \frac{\overline{x}^2}{(n-1)\,s_x^2}\right).$
Covariância dos estimadores $\hat{\beta}_1$ e $\hat{\beta}_0$	$Cov(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\overline{x}\,\sigma^2}{(n-1)\,s_x^2}.$
Intervalo de predição a $(1-\alpha)\times 100\%$ para	$\Big]\,(b_0+b_1x) - t_{\frac{\alpha}{2};n-2}\cdot\sqrt{QMRE\cdot\left[1+\frac{1}{n}+\frac{(x-\overline{x})^2}{(n-1)\,s_x^2}\right]}\;,$
observação individual de Y, dado X=x:	$(b_0+b_1x) + t_{\frac{\alpha}{2};n-2}\cdot\sqrt{QMRE\cdot\left[1+\frac{1}{n}+\frac{(x-\overline{x})^2}{(n-1)\,s_x^2}\right]}\,\Big[.$
Valor do efeito alavanca	$h_{ii} = \frac{1}{n} + \frac{(x_i-\bar{x})^2}{(n-1)\,s_x^2}$
Propriedades de Matrizes	$(\mathbf{A}\pm\mathbf{B})^t = \mathbf{A}^t \pm \mathbf{B}^t \quad (\mathbf{AB})^t = \mathbf{B}^t\mathbf{A}^t \quad (\mathbf{A}^t)^{-1} = (\mathbf{A}^{-1})^t$
Propriedades da Multinormal	Se $\vec{\mathbf{W}} \cap \mathcal{N}_n(\vec{\boldsymbol{\mu}}, \boldsymbol{\Sigma})$, $\vec{\mathbf{a}}$ vector $k\times 1$ (não aleatório)
	e $\mathbf{C}$ matriz $k\times n$ (não aleatória, de característica k)
	então $\mathbf{C}\vec{\mathbf{W}} + \vec{\mathbf{a}} \cap \mathcal{N}_k(\mathbf{C}\vec{\boldsymbol{\mu}} + \vec{\mathbf{a}}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^t)$
Regressão Linear Múltipla	
Equação do modelo	$\vec{\mathbf{Y}} = \mathbf{X}\vec{\boldsymbol{\beta}} + \vec{\boldsymbol{\epsilon}}$
Vector dos estimadores dos parâmetros	$\hat{\vec{\boldsymbol{\beta}}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\vec{\mathbf{Y}}.$
Matriz de projecção ortogonal	$\mathbf{H} = \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t.$
Vector dos valores estimados de Y	$\hat{\vec{\mathbf{Y}}} = \mathbf{X}\hat{\vec{\boldsymbol{\beta}}} = \mathbf{H}\vec{\mathbf{Y}}.$
Matriz de (co-)variâncias dos estimadores $\hat{\beta}_i$	$Var(\hat{\vec{\boldsymbol{\beta}}}) = \sigma^2.(\mathbf{X}^t\mathbf{X})^{-1}.$
Distribuição dos resíduos	$E_i \cap \mathcal{N}\left(0, \sigma^2\cdot(1-h_{ii})\right)$ com $h_{ii} = \mathbf{H}_{(i,i)}.$
AIC (Critério de Informação de Akaike)	$AIC = n\,\ln\left(\frac{SQRE_k}{n}\right) + 2(k+1).$
IC a $(1-\alpha)\times 100\%$ para combinações lineares	$\Big]\,\vec{\mathbf{a}}^t\vec{\mathbf{b}} - t_{\frac{\alpha}{2};n-(p+1)}\cdot\hat{\sigma}_{\vec{\mathbf{a}}^t\hat{\vec{\boldsymbol{\beta}}}}\;,\;\vec{\mathbf{a}}^t\vec{\mathbf{b}} + t_{\frac{\alpha}{2};n-(p+1)}\cdot\hat{\sigma}_{\vec{\mathbf{a}}^t\hat{\vec{\boldsymbol{\beta}}}}\,\Big[.$
dos parâmetros: $\vec{\mathbf{a}}^t\vec{\boldsymbol{\beta}} = \sum_{i=0}^{p} a_i\beta_i.$	com $\hat{\sigma}_{\vec{\mathbf{a}}^t\hat{\vec{\boldsymbol{\beta}}}} = \sqrt{QMRE\cdot\vec{\mathbf{a}}^t(\mathbf{X}^t\mathbf{X})^{-1}\vec{\mathbf{a}}}.$
Teste de Ajustamento Global	$F = \frac{n-(p+1)}{p}\cdot\frac{R^2}{1-R^2}.$
Teste aos Modelos Encaixados (teste F parcial)	$F = \frac{(SQRE_s - SQRE_c)/(p-k)}{(SQRE_c)/(n-(p+1))}.$
Resíduos (internamente) estandardizados	$R_i = \frac{E_i}{\sqrt{QMRE\cdot(1-h_{ii})}}$
Distância de Cook	$D_i = R_i^2\cdot\left(\frac{h_{ii}}{1-h_{ii}}\right)\cdot\frac{1}{p+1}$
R^2 modificado	$R^2_{mod} = 1 - \frac{QMRE}{QMT}$

Descrição	Fórmula
ANOVA	
Um factor	$\hat{\mu}_1 = \overline{Y}_{1\cdot}$; $\hat{\alpha}_i = \overline{Y}_{i\cdot} - \overline{Y}_{1\cdot}$; $E_{ij} = Y_{ij} - \overline{Y}_{i\cdot\cdot}$ $V[\hat{\mu}_1] = \frac{\sigma^2}{n_1}$; $V[\hat{\alpha}_i] = \sigma^2 \left[\frac{1}{n_1} + \frac{1}{n_i}\right]$ $SQF = \sum_{i=1}^{k} n_i \, (\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})^2$; $SQRE = \sum_{i=1}^{k} (n_i - 1)\, S_i^2$
Dois factores Factorial, com ou sem interacção	
(delineamento equilibrado, n_c obs. por célula)	$SQA = b\, n_c \sum_{i=1}^{a} (\overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot\cdot\cdot})^2$
(delineamento equilibrado, n_c obs. por célula)	$SQB = a\, n_c \sum_{j=1}^{b} (\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdot\cdot\cdot})^2$
Dois factores Factorial, sem interacção	$\hat{Y}_{ijk} = \overline{Y}_{i\cdot\cdot} + \overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdot\cdot\cdot}$
Dois factores Factorial, com interacção	$\hat{\mu}_{11} = \overline{Y}_{11\cdot}$; $\hat{\alpha}_i = \overline{Y}_{i1\cdot} - \overline{Y}_{11\cdot}$; $\hat{\beta}_j = \overline{Y}_{1j\cdot} - \overline{Y}_{11\cdot}$ $(\hat{\alpha\beta})_{ij} = (\overline{Y}_{ij\cdot} + \overline{Y}_{11\cdot}) - (\overline{Y}_{i1\cdot} + \overline{Y}_{1j\cdot})$ $\widehat{V[\hat{\alpha}_i]} = QMRE \left(\frac{1}{n_{i1}} + \frac{1}{n_{11}}\right)$ $\widehat{V[\hat{\beta}_j]} = QMRE \left(\frac{1}{n_{1j}} + \frac{1}{n_{11}}\right)$ $V[\widehat{(\hat{\alpha\beta})_{ij}}] = QMRE \left(\frac{1}{n_{ij}} + \frac{1}{n_{11}} + \frac{1}{n_{1j}} + \frac{1}{n_{i1}}\right)$ $E_{ijk} = Y_{ijk} - \overline{Y}_{ij\cdot}$ $SQRE = \sum_{i=1}^{a} \sum_{j=1}^{b} (n_{ij} - 1)\, S_{ij}^2$
Dois factores Hierarquizados	$g.l.(SQA) = a - 1$ $g.l.\,(SQB(A)) = \sum_{i=1}^{a} (b_i - 1)$ $g.l.(SQRE) = n - \sum_{i=1}^{a} b_i$ $SQRE = \sum_{i=1}^{a} \sum_{j=1}^{b_i} (n_{ij} - 1)\, S_{ij}^2$
(del. equilibrado, n_c obs. por célula/folha)	$SQA = n_c \sum_{i=1}^{a} b_i \, (\overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot\cdot\cdot})^2$
Testes de Tukey	
(n_c repetições em cada um de m níveis/células)	$q_{\alpha(m,\nu)} \sqrt{QMRE/n_c}$ (com $\nu = gl(SQRE)$).
Teste de Bartlett	
(m níveis/células)	$K = \frac{(n-m) \ln QMRE - \sum_i \left[\sum_j\right] (n_{i[j]}-1) \ln S^2_{i[j]}}{C} \sim \chi^2_{m-1}$
$[\cdot]$ indica que pode ou não ser preciso	onde $C = 1 + \frac{1}{3(m-1)} \left(\sum_i \left[\sum_j \right] \frac{1}{n_{i[j]}-1} - \frac{1}{n-m} \right)$