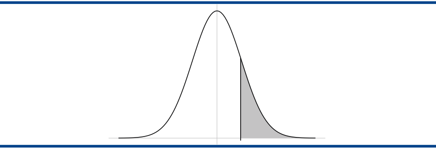
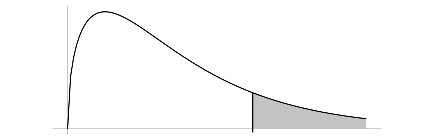
Esenciais USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

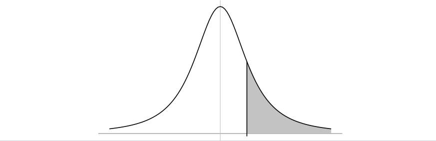
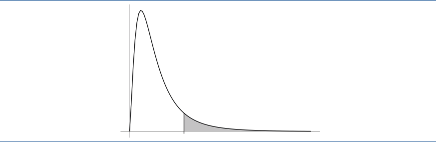
MATEMÁTICAS E ESTATÍSTICA II

conceptos e fórmulas da estatística no Grao de Farmacia

José Carlos Díaz Ramos, Antonio Gómez Tato, Enrique Macías Virgós (Departamento de Matemáticas)

Distribucións de probabilidade

Normal $z = N(0,1)$ (simétrica)	
χ^2 de Pearson (non simétrica)	

t-Student (simétrica)	
F de Fisher-Snedecor $F_{n,m,\,\alpha} = \frac{1}{F_{m,n,\,1-\alpha}}$	

Para unha variable aleatoria X con distribución D, D_α denota a abscisa para a que $P(X \geq D_\alpha) = \alpha$

Contrastes de independencia e homoxeneidade para datos categóricos

Estatístico $\sum_{i,j} \frac{(n_{ij}-\widehat{E_{ij}})^2}{\widehat{E_{ij}}} \sim \chi^2_{(f-1)(c-1)}$

Rexión crítica $(\chi^2_{(f-1)(c-1),\,\alpha},\, \infty)$

Contrastes de independencia $H_0\colon p_{ij} = p_{i\cdot}p_{\cdot j}$, $\forall i \in \{1,\dots,f\}, \forall j \in \{1,\dots,c\}$

Contrastes de homoxeneidade $H_0\colon p_{1j} = p_{2j} = \cdots = p_{fj}$, $\forall j \in \{1,\dots,c\}$

X / Y	B_1	$\dots$	B_j	$\dots$	B_c	Σ
A_1	n_{11} $\widehat{E_{11}} = \frac{n_{1\cdot}n_{\cdot 1}}{n}$ $\frac{(n_{11}-\widehat{E_{11}})^2}{\widehat{E_{11}}}$		$\vdots$		n_{1c} $\widehat{E_{1c}} = \frac{n_{1\cdot}n_{\cdot c}}{n}$ $\frac{(n_{1c}-\widehat{E_{1c}})^2}{\widehat{E_{1c}}}$	$n_{1\cdot}$
$\vdots$						
A_i	$\dots$		n_{ij} $\widehat{E_{ij}} = \frac{n_{i\cdot}n_{\cdot j}}{n}$ $\frac{(n_{ij}-\widehat{E_{ij}})^2}{\widehat{E_{ij}}}$		$\dots$	$n_{i\cdot}$
$\vdots$						
A_f	n_{f1} $\widehat{E_{f1}} = \frac{n_{f\cdot}n_{\cdot 1}}{n}$ $\frac{(n_{f1}-\widehat{E_{f1}})^2}{\widehat{E_{f1}}}$		$\vdots$		n_{fc} $\widehat{E_{fc}} = \frac{n_{f\cdot}n_{\cdot c}}{n}$ $\frac{(n_{fc}-\widehat{E_{fc}})^2}{\widehat{E_{fc}}}$	$n_{f\cdot}$
Σ	$n_{\cdot 1}$	$\dots$	$n_{\cdot j}$	$\dots$	$n_{\cdot c}$	n

Regresión e correlación

Modelo de recta de regresión $Y = \alpha + \beta X + \epsilon$

$\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$

$\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$

$s_X^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^n x_i^2 - \overline{x}^2$

$s_Y^2 = \frac{1}{n}\sum_{i=1}^n (y_i - \overline{y})^2 = \frac{1}{n}\sum_{i=1}^n y_i^2 - \overline{y}^2$

$s_{XY} = \frac{1}{n}\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{n}\sum_{i=1}^n x_i y_i - \overline{x}\,\overline{y}$

Coeficiente de correlación $r = \frac{s_{XY}}{s_X s_Y}$

Coeficiente de determinación $r^2 = \frac{s_{XY}^2}{s_X^2 s_Y^2} = \frac{s_R^2}{s_Y^2} = \frac{SS_R}{SS_Y}$

$a = \hat{\alpha} = \overline{y} - b\,\overline{x}$

$b = \hat{\beta} = \frac{s_{XY}}{s_X^2}$

Análise de varianza (ANOVA)

Contraste de hipóteses $H_0\colon \rho = 0 \quad H_1\colon \rho \neq 0$ (equivalentemente $H_0\colon \beta = 0 \quad H_1\colon \beta \neq 0$) Rexión crítica $(F_{1,\,n-2,\,\alpha},\, \infty)$

Variabilidade	g.l.	SS	MS	Cociente
Regresión	1	$SS_R = \sum_{i=1}^n (\hat{Y}_i - \overline{Y})^2 = nr^2 s_Y^2$	$MS_R = \frac{SS_R}{1}$	$F_{1,\,n-2} \sim \frac{MS_R}{MS_E}$
Erro	$n-2$	$SS_E = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = n(1-r^2)s_Y^2$	$MS_E = \frac{SS_E}{n-2}$	
Total	$n-1$	$SS_Y = \sum_{i=1}^n (Y_i - \overline{Y})^2 = ns_Y^2$		

Media dunha poboación

Estimador puntual $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Varianza poboacional coñecida

Estatístico $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim Z$

INTERVALOS DE CONFIANZA

Inecuación $\left|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right| \leq z_{\alpha/2}$ Fórmula $X \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Varianza poboacional descoñecida

Estatístico $\frac{\overline{X}-\mu}{s_{n-1}/\sqrt{n}} \sim t_{n-1}$

INTERVALOS DE CONFIANZA

Inecuación $\left|\frac{\overline{X}-\mu}{s_{n-1}/\sqrt{n}}\right| \leq t_{n-1,\,\alpha/2}$ Fórmula $X \pm t_{n-1,\,\alpha/2}\frac{s_{n-1}}{\sqrt{n}}$

CONTRASTES DE HIPÓTESES

$H_0\!: \mu = \mu_0$ $H_1\!: \mu \neq \mu_0$

Rexión crítica $\left(-\infty,\ -t_{n-1,\,\alpha/2}\right) \cup \left(t_{n-1,\,\alpha/2},\ +\infty\right)$

$H_0\!: \mu \leq \mu_0$ $H_1\!: \mu > \mu_0$

Rexión crítica $\left(t_{n-1,\,\alpha},\ +\infty\right)$

Varianza dunha poboación

Media poboacional coñecida

Estimador puntual $s_\mu^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$ Estatístico $\frac{ns_\mu^2}{\sigma^2} \sim \chi_n^2$

INTERVALOS DE CONFIANZA

Inecuación $\chi^2_{n,\,1-\alpha/2} \leq \frac{ns_\mu^2}{\sigma^2} \leq \chi^2_{n,\,\alpha/2}$ Fórmula $\left[\frac{ns_\mu^2}{\chi^2_{n,\,\alpha/2}},\ \frac{ns_\mu^2}{\chi^2_{n,\,1-\alpha/2}}\right]$

Media poboacional descoñecida

Estimador puntual $s_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ Estatístico $\frac{(n-1)s_{n-1}^2}{\sigma^2} \sim \chi^2_{n-1}$

INTERVALOS DE CONFIANZA

Inecuación $\chi^2_{n-1,\,1-\alpha/2} \leq \frac{(n-1)s_{n-1}^2}{\sigma^2} \leq \chi^2_{n-1,\,\alpha/2}$ Fórmula $\left[\frac{(n-1)s_{n-1}^2}{\chi^2_{n-1,\,\alpha/2}},\ \frac{(n-1)s_{n-1}^2}{\chi^2_{n-1,\,1-\alpha/2}}\right]$

CONTRASTES DE HIPÓTESES

$H_0\!: \sigma^2 = \sigma_0^2$ $H_1\!: \sigma^2 \neq \sigma_0^2$

Rexión crítica $\left(0,\ \chi^2_{n-1,\,1-\alpha/2}\right) \cup \left(\chi^2_{n-1,\,\alpha/2},\ +\infty\right)$

$H_0\!: \sigma^2 \leq \sigma_0^2$ $H_1\!: \sigma^2 > \sigma_0^2$

Rexión crítica $\left(\chi^2_{n-1,\,\alpha},\ +\infty\right)$

$H_0\!: \sigma^2 \geq \sigma_0^2$ $H_1\!: \sigma^2 < \sigma_0^2$

Rexión crítica $\left(0,\ \chi^2_{n-1,\,1-\alpha}\right)$

Proporcións nunha poboación

Estimador puntual $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$ Estatístico $\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \sim Z$

INTERVALOS DE CONFIANZA

Inecuación $\left|\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}}\right| \leq z_{\alpha/2}$ Fórmula $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

CONTRASTES DE HIPÓTESES

$H_0\!: p = p_0$ $H_1\!: p \neq p_0$

Rexión crítica $\left(-\infty,\ -z_{\alpha/2}\right) \cup \left(z_{\alpha/2},\ +\infty\right)$

$H_0\!: p \leq p_0$ $H_1\!: p > p_0$

Rexión crítica $\left(z_\alpha,\ +\infty\right)$

Comparación da varianza de dúas poboacións

Estatístico $\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{n_1-1,\,n_2-1}$

$H_0\!: \sigma_1^2 = \sigma_2^2$ $H_1\!: \sigma_1^2 \neq \sigma_2^2$

Rexión crítica $\left(0,\ \frac{1}{F_{n_2-1,\,n_1-1,\,\alpha/2}}\right) \cup \left(F_{n_1-1,\,n_2-1,\,\alpha/2},\ +\infty\right)$

Comparación da media de dúas poboacións

Varianzas poboacionais coñecidas

Estatístico $\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim Z$

INTERVALOS DE CONFIANZA

Inecuación $\left|\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\right| \leq z_{\alpha/2}$

Fórmula $(\overline{X}-\overline{Y}) \pm z_{\alpha/2}\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}$

CONTRASTES DE HIPÓTESES

$H_0\!: \mu_1-\mu_2 = (\mu_1-\mu_2)_0$ $H_1\!: \mu_1-\mu_2 \neq (\mu_1-\mu_2)_0$

Rexión crítica $\left(-\infty,\ -z_{\alpha/2}\right) \cup \left(z_{\alpha/2},\ +\infty\right)$

$H_0\!: \mu_1-\mu_2 \leq (\mu_1-\mu_2)_0$ $H_1\!: \mu_1-\mu_2 > (\mu_1-\mu_2)_0$

Rexión crítica $\left(z_\alpha,\ +\infty\right)$

Varianzas poboacionais descoñecidas pero iguais

Estatístico $\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$

Cuasivarianza conxunta $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

INTERVALOS DE CONFIANZA

Inecuación $\left|\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{s_p\sqrt{1/n_1+1/n_2}}\right| \leq t_{n_1+n_2-2,\,\alpha/2}$

Fórmula $(\overline{X}-\overline{Y}) \pm t_{n_1+n_2-2,\,\alpha/2}\, s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$

CONTRASTES DE HIPÓTESES

$H_0\!: \mu_1-\mu_2 = (\mu_1-\mu_2)_0$ $H_1\!: \mu_1-\mu_2 \neq (\mu_1-\mu_2)_0$

Rexión crítica $\left(-\infty,\ -t_{n_1+n_2-2,\,\alpha/2}\right) \cup \left(t_{n_1+n_2-2,\,\alpha/2},\ +\infty\right)$

$H_0\!: \mu_1-\mu_2 \leq (\mu_1-\mu_2)_0$ $H_1\!: \mu_1-\mu_2 > (\mu_1-\mu_2)_0$

Rexión crítica $\left(t_{n_1+n_2-2,\,\alpha},\ +\infty\right)$

Varianzas poboacionais descoñecidas

Estatístico $\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{s_1^2/n_1+s_2^2/n_2}} \sim t_\gamma$ Graos de liberdade $\gamma \leq \frac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}}$

INTERVALOS DE CONFIANZA

Inecuación $\left|\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{s_1^2/n_1+s_2^2/n_2}}\right| \leq t_{\gamma,\,\alpha/2}$

Fórmula $(\overline{X}-\overline{Y}) \pm t_{\gamma,\,\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$

CONTRASTES DE HIPÓTESES

$H_0\!: \mu_1-\mu_2 = (\mu_1-\mu_2)_0$ $H_1\!: \mu_1-\mu_2 \neq (\mu_1-\mu_2)_0$

Rexión crítica $\left(-\infty,\ -t_{\gamma,\,\alpha/2}\right) \cup \left(t_{\gamma,\,\alpha/2},\ +\infty\right)$

$H_0\!: \mu_1-\mu_2 \leq (\mu_1-\mu_2)_0$ $H_1\!: \mu_1-\mu_2 > (\mu_1-\mu_2)_0$

Rexión crítica $\left(t_{\gamma,\,\alpha},\ +\infty\right)$

Comparación de proporcións de dúas poboacións

Estatístico $\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \sim Z$

INTERVALOS DE CONFIANZA

Inecuación $\left|\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1+\hat{p}_2(1-\hat{p}_2)/n_2}}\right| \leq z_{\alpha/2}$

Fórmula $(\hat{p}_1-\hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

CONTRASTES DE HIPÓTESES

$H_0\!: p_1-p_2 = (p_1-p_2)_0 \neq 0$ $H_1\!: p_1-p_2 \neq (p_1-p_2)_0$

Rexión crítica $\left(-\infty,\ -z_{\alpha/2}\right) \cup \left(z_{\alpha/2},\ +\infty\right)$

$H_0\!: p_1-p_2 \leq (p_1-p_2)_0 \neq 0$ $H_1\!: p_1-p_2 > (p_1-p_2)_0$

Rexión crítica $\left(z_\alpha,\ +\infty\right)$

Valor nulo cero

Estatístico $\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim Z$ Proporción ponderada $\hat{p} = \frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}$

CONTRASTES DE HIPÓTESES

$H_0\!: p_1 = p_2$ $H_1\!: p_1 \neq p_2$

Rexión crítica $\left(-\infty,\ -z_{\alpha/2}\right) \cup \left(z_{\alpha/2},\ +\infty\right)$

$H_0\!: p_1 \leq p_2$ $H_1\!: p_1 > p_2$

Rexión crítica $\left(z_\alpha,\ +\infty\right)$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
XUNTA DE GALICIA
RAGC/ REAL ACADEMIA GALEGA DE CIENCIAS