Matemàtiques III del grau de Matemàtiques i de la doble titulació GMIT.
Taules de contrastos d'hipòtesi més usuals II: Dues mostres

En aquest document recollim els contrastos d'hipòtesis paramètrics més usuals per a dues mostres que es poden portar a terme "a mà." Per a cada contrast donam: les condicions, l'estadístic de contrast, la regió crítica, l'interval de confiança i el p-valor.

En la definició dels estadístics hem emprat la notacions següents:

- Z: Distribució normal estàndard $N(0, 1)$.
- t_n: Distribució t de Student amb n graus de llibertat.
- χ^2_n: Distribució khi-quadrat amb n graus de llibertat.
- F_{n_1,n_2}: Distribució F de Fisher amb n_1 i n_2 graus de llibertat.
- X_α: Indica l'α-quantil de la variable aleatòria X, és a dir (si X és contínua, que és sempre el cas en aquest document), el valor on la funció de distribució de X val α: $P(X \leq X_\alpha) = \alpha$.

  Recordau la traducció als quantils de les propietats de simetria de Z, t i F:

  - Simetria de la normal: $z_\alpha = -z_{1-\alpha}$.
  - Simetria de la t de Student: $t_{n,\alpha} = -t_{n,1-\alpha}$.
  - Permutació dels graus de llibertat de la F de Fisher: $F_{n_1,n_2,\alpha} = \frac{1}{F_{n_2,n_1,1-\alpha}}$.

Els contrastos paramètrics amb `R` els estudiarem més endavant a una lliçó del manual de `R`.

Tipus de contrast i condicions				
Hip. nul·la	Condicions	Mostra	Hip. alt.	Cas
$H_0 : \mu_1 = \mu_2$ Cas independent	σ_1 i σ_2 conegudes. Poblacions normals o n_1 i n_2 grans.	Dues m.a.s. independents de mides n_1 i n_2	$H_1 : \mu_1 \neq \mu_2$	**I**
			$H_1 : \mu_1 < \mu_2$	**II**
			$H_1 : \mu_1 > \mu_2$	**III**
	σ_1 i σ_2 desconegudes i $\sigma_1 = \sigma_2$. Poblacions normals o n_1 i n_2 grans.	Dues m.a.s. independents de mides n_1 i n_2	$H_1 : \mu_1 \neq \mu_2$	**IV**
			$H_1 : \mu_1 < \mu_2$	**V**
			$H_1 : \mu_1 > \mu_2$	**VI**
	σ_1 i σ_2 desconegudes i $\sigma_1 \neq \sigma_2$. Poblacions normals o n_1 i n_2 grans.	Dues m.a.s. independents de mides n_1 i n_2	$H_1 : \mu \neq \mu_2$	**VII**
			$H_1 : \mu_1 < \mu_2$	**VIII**
			$H_1 : \mu_1 > \mu_2$	**IX**
$H_0 : \mu_1 = \mu_2$ Cas dependent	Dues poblacions normals dependents o n gran. σ_d coneguda.(1)	Dues m.a.s. dependents de mida n	$H_1 : \mu_d \neq 0$	**X**
			$H_1 : \mu_d < 0$	**XI**
			$H_1 : \mu_d > 0$	**XII**
	Dues poblacions normals dependents. σ_d desconeguda.(1)	Dues m.a.s. dependents de mida n	$H_1 : \mu_d \neq 0$	**XIII**
			$H_1 : \mu_d < 0$	**XIV**
			$H_1 : \mu_d > 0$	**XV**
	Dues poblacions dependents, n gran. σ_d desconeguda.(1)	Dues m.a.s. dependents de mida n	$H_1 : \mu_d \neq 0$	**XVI**
			$H_1 : \mu_d < 0$	**XVII**
			$H_1 : \mu_d > 0$	**XVIII**
$H_0 : p_1 = p_2$ Cas independent	Poblacions Bernoulli, n_1 i n_2 grans, molts èxits i fracasos.	Dues m.a.s. independents de mides n_1 i n_2	$H_1 : p_1 \neq p_2$	**XIX**
			$H_1 : p_1 < p_2$	**XX**
			$H_1 : p_1 > p_2$	**XXI**
$H_0 : p_a = p_d$ Cas dependent	Poblacions Bernoulli, n_1 i n_2 grans, molts casos discordants.	Dues m.a.s. dependents de mida n	$H_1 : p_a \neq p_b$	**XXII**
			$H_1 : p_a < p_b$	**XXIII**
			$H_1 : p_a > p_b$	**XXIV**
$H_0 : \sigma_1^2 = \sigma_2^2$ Cas independent	Poblacions normals.	Dues m.a.s. independents de mides n_1 i n_2	$H_1 : \sigma_1^2 \neq \sigma_2^2$	**XXV**
			$H_1 : \sigma_1^2 < \sigma_2^2$	**XXVI**
			$H_1 : \sigma_1^2 > \sigma_2^2$	**XXVII**

(1) σ_d és la desviació típica de la variable $D = X_1 - X_2$.

Detalls del test				
Cas	Estadístic	Regió crítica	Interval confiança	p-valor
I	$Z=\frac{\overline{X}_1-\overline{X}_2}{\widetilde{S}}$ és $N(0,1)$ (vegeu (a))	$\{Z\leq -z_{1-\frac{\alpha}{2}}\}\cup\{Z\geq z_{1-\frac{\alpha}{2}}\}$	$\left]\overline{X}_1-\overline{X}_2-z_{1-\frac{\alpha}{2}}\widetilde{S},\overline{X}_1-\overline{X}_2+z_{1-\frac{\alpha}{2}}\widetilde{S}\right[$	$2P(Z\geq\|z\|)$
II		$\{Z\leq z_{\alpha}\}$	$\left]-\infty,\overline{X}_1-\overline{X}_2-z_{\alpha}\widetilde{S}\right[$	$P(Z\leq z)$
III		$\{Z\geq z_{1-\alpha}\}$	$\left]\overline{X}_1-\overline{X}_2-z_{1-\alpha}\widetilde{S},+\infty\right[$	$P(Z\geq z)$
IV	$T=\frac{\overline{X}_1-\overline{X}_2}{\widetilde{S}_{1,2}}$ és t_m (vegeu (b,c))	$\{T\leq -t_{m,1-\frac{\alpha}{2}}\}\cup\{T\geq t_{m,1-\frac{\alpha}{2}}\}$	$\left]\overline{X}_1-\overline{X}_2-t_{m,1-\frac{\alpha}{2}}\widetilde{S}_{1,2},\overline{X}_1-\overline{X}_2+t_{m,1-\frac{\alpha}{2}}\widetilde{S}_{1,2}\right[$	$2P(t_m>\|T\|)$
V		$\{T\leq t_{m,\alpha}\}$	$\left]-\infty,\overline{X}_1-\overline{X}_2-t_{m,\alpha}\widetilde{S}_{1,2}\right[$	$P(t_m\leq T)$
VI		$\{T\geq t_{m,1-\alpha}\}$	$\left]\overline{X}_1-\overline{X}_2-t_{m,1-\alpha}\widetilde{S}_{1,2},+\infty\right[$	$P(t_m\geq T)$
VII	$T=\frac{\overline{X}_1-\overline{X}_2}{\widetilde{S}_{1,2}}$ és t_f (vegeu (d,e))	$\{T\leq -t_{f,1-\frac{\alpha}{2}}\}\cup\{T\geq t_{f,1-\frac{\alpha}{2}}\}$	$\left]\overline{X}_1-\overline{X}_2-t_{f,1-\frac{\alpha}{2}}\widetilde{S}_{1,2},\overline{X}_1-\overline{X}_2+t_{f,1-\frac{\alpha}{2}}\widetilde{S}_{1,2}\right[$	$2P(t_f>\|T\|)$
VIII		$\{T\leq t_{f,\alpha}\}$	$\left]-\infty,\overline{X}_1-\overline{X}_2-t_{f,\alpha}\widetilde{S}_{1,2}\right[$	$P(t_f\leq T)$
IX		$\{T\geq t_{f,1-\alpha}\}$	$\left]\overline{X}_1-\overline{X}_2-t_{f,1-\alpha}\widetilde{S}_{1,2},+\infty\right[$	$P(t_f\geq T)$
X	$Z=\frac{\overline{D}}{\frac{\sigma_D}{\sqrt{n}}}$ és $N(0,1)$ (vegeu (f))	$\{Z\leq -z_{1-\frac{\alpha}{2}}\}\cup\{Z\geq z_{1-\frac{\alpha}{2}}\}$	$\left]\overline{D}-z_{1-\frac{\alpha}{2}}\frac{\sigma_D}{\sqrt{n}},\overline{D}+z_{1-\frac{\alpha}{2}}\frac{\sigma_D}{\sqrt{n}}\right[$	$2P(Z\geq\|z\|)$
XI		$\{Z\leq z_{\alpha}\}$	$\left]-\infty,\overline{D}-z_{\alpha}\frac{\sigma_D}{\sqrt{n}}\right[$	$P(Z\leq z)$
XII		$\{Z\geq z_{1-\alpha}\}$	$\left]\overline{D}-z_{1-\alpha}\frac{\sigma_D}{\sqrt{n}},+\infty\right[$	$P(Z\geq z)$
XIII	$T=\frac{\overline{D}}{\frac{\widetilde{S}_D}{\sqrt{n}}}$ és t_{n-1} (vegeu (f))	$\{T\leq -t_{n-1,1-\frac{\alpha}{2}}\}\cup\{T\geq t_{n-1,1-\frac{\alpha}{2}}\}$	$\left]\overline{D}-t_{n-1,1-\frac{\alpha}{2}}\frac{\widetilde{S}_D}{\sqrt{n}},\overline{D}+t_{n-1,1-\frac{\alpha}{2}}\frac{\widetilde{S}_D}{\sqrt{n}}\right[$	$2P(t_{n-1}>\|T\|)$
XIV		$\{T\leq t_{n-1,\alpha}\}$	$\left]-\infty,\overline{D}-t_{n-1,\alpha}\frac{\widetilde{S}_D}{\sqrt{n}}\right[$	$P(t_{n-1}\leq T)$
XV		$\{T\geq t_{n-1,1-\alpha}\}$	$\left]\overline{D}-t_{n-1,1-\alpha}\frac{\sigma_D}{\sqrt{n}},+\infty\right[$	$P(t_{n-1}\geq T)$
XVI	$Z=\frac{\overline{D}}{\frac{\widetilde{S}_D}{\sqrt{n}}}$ és $N(0,1)$ (vegeu (f))	$\{Z\leq -z_{1-\frac{\alpha}{2}}\}\cup\{Z\geq z_{1-\frac{\alpha}{2}}\}$	$\left]\overline{D}-z_{1-\frac{\alpha}{2}}\frac{\widetilde{S}_D}{\sqrt{n}},\overline{D}+z_{1-\frac{\alpha}{2}}\frac{\widetilde{S}_D}{\sqrt{n}}\right[$	$2P(Z\geq\|z\|)$
XVII		$\{Z\leq z_{\alpha}\}$	$\left]-\infty,\overline{D}-z_{\alpha}\frac{\widetilde{S}_D}{\sqrt{n}}\right[$	$p(Z\leq z)$
XVIII		$\{Z\geq z_{1-\alpha}\}$	$\left]\overline{D}-z_{1-\alpha}\frac{\sigma_D}{\sqrt{n}},+\infty\right[$	$P(Z\geq z)$
XIX	$Z=\frac{\widehat{p}_1-\widehat{p}_2}{\sqrt{\widehat{p}\widehat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ és $N(0,1)$ (vegeu (g,h))	$\{Z\leq -z_{1-\frac{\alpha}{2}}\}\cup\{Z\geq z_{1-\frac{\alpha}{2}}\}$	$\left]\widehat{p}_1-\widehat{p_2}-z_{1-\frac{\alpha}{2}}\sqrt{\widehat{p}\widehat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)},\right.$ $\left.\widehat{p}_1-\widehat{p_2}+z_{1-\frac{\alpha}{2}}\sqrt{\widehat{p}\widehat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right[$	$2P(Z\geq\|z\|)$
XX		$\{Z\leq z_{\alpha}\}$	$\left]-\infty,\widehat{p}_1-\widehat{p_2}-z_{\alpha}\sqrt{\widehat{p}\widehat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right[$	$P(Z\leq z)$
XXI		$\{Z\geq z_{1-\alpha}\}$	$\left]\widehat{p}_1-\widehat{p_2}-z_{1-\alpha}\sqrt{\widehat{p}\widehat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)},+\infty\right[$	$P(Z\geq z)$
XXII	$Z=\frac{\widehat{p}_{1\bullet}-\widehat{p}_{\bullet 1}}{\sqrt{\frac{b+d}{n^2}}}$ és $N(0,1)$ (vegeu (i))	$\{Z\leq -z_{1-\frac{\alpha}{2}}\}\cup\{Z\geq z_{1-\frac{\alpha}{2}}\}$	$\left]\widehat{p}_{1\bullet}-\widehat{p}_{\bullet 1}-z_{1-\frac{\alpha}{2}}\sqrt{\frac{b+d}{n^2}},\right.$ $\left.\widehat{p}_{1\bullet}-\widehat{p}_{\bullet 1}+z_{1-\frac{\alpha}{2}}\sqrt{\frac{b+d}{n^2}}\right[$	$2P(Z\geq\|z\|)$
XXIII		$\{Z\leq z_{\alpha}\}$	$\left]-\infty,\widehat{p}_{1\bullet}-\widehat{p}_{\bullet 1}-z_{\alpha}\sqrt{\frac{b+d}{n^2}}\right[$	$P(Z\leq z)$
XXIV		$\{Z\geq z_{1-\alpha}\}$	$\left]\widehat{p}_{1\bullet}-\widehat{p}_{\bullet 1}-z_{1-\alpha}\sqrt{\frac{b+d}{n^2}},+\infty\right[$	$P(Z\geq z)$
XXV	$F=\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2}$ és F_{n_1-1,n_2-1}	$\{F\leq F_{n_1-1,n_2-1,\frac{\alpha}{2}}\}\cup$ $\{F\geq F_{n_1-1,n_2-1,1-\frac{\alpha}{2}}\}$	$\left]\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2}F_{n_1-1,n_2-1,\frac{\alpha}{2}},\right.$ $\left.\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2}F_{n_1-1,n_2-1,1-\frac{\alpha}{2}}\right[$	$2\min\{P(F_{n_1-1,n_2-1}\leq F),$ $P(F_{n_1-1,n_2-1}\geq F)\}$
XXVI		$\{F\leq F_{n_1-1,n_2-1,\alpha}\}$	$\left]0,\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2}F_{n_1-1,n_2-1,1-\alpha}\right[$	$P(F_{n_1-1,n_2-1}\leq F)$
XXVII		$\{F\geq F_{n_1-1,n_2-1,1-\alpha}\}$	$\left]\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2}F_{n_1-1,n_2-1,\alpha},+\infty\right[$	$P(F_{n_1-1,n_2-1}\geq F)$

(a) $\widetilde{S} = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

(b) $\widetilde{S}_{1,2} = \sqrt{\dfrac{(n_1 - 1)\widetilde{S}_1^2 + (n_2 - 1)\widetilde{S}_2^2}{n_1 + n_2 - 2} \cdot \left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$

(c) $m = n_1 + n_2 - 2$

(d) $\widetilde{S}_{1,2} = \sqrt{\dfrac{\widetilde{S}_1^2}{n_1} + \dfrac{\widetilde{S}_2^2}{n_2}}$

(e) $f = \left\lfloor \dfrac{\left(\frac{\widetilde{S}_1^2}{n_1} + \frac{\widetilde{S}_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{\widetilde{S}_1^2}{n_1}\right)^2 + \frac{1}{n_2-1}\left(\frac{\widetilde{S}_2^2}{n_2}\right)^2} \right\rfloor - 2$

(f) $\overline{D}$ i $\widetilde{S}_D$ són la mitjana i la desvació típica mostrals de $D = X_1 - X_2$

(g) $\widehat{p} = \dfrac{n_1\widehat{p}_1 + n_2\widehat{p}_2}{n_1 + n_2}$

(h) $\widehat{q} = 1 - \widehat{p}$

(i) Per fer el contrast, hem de construir la taula següent:

		Mostra després			
		Èxit	Fracàs	Freqüència	Proporció
Mostra abans	Èxit	a	b	$a+b$	$\widehat{p}_{1\bullet} = \frac{a+b}{n}$
	Fracàs	d	c	$c+d$	$\widehat{p}_{2\bullet} = \frac{c+d}{n}$
	Freqüència	$a+d$	$b+c$	n	
	Proporció	$\widehat{p}_{\bullet 1} = \frac{a+d}{n}$	$\widehat{p}_{\bullet 2} = \frac{b+c}{n}$		1

Aleshores, l'estadístic de contrast també es pot escriure

$$Z = \frac{\frac{b}{n} - \frac{d}{n}}{\sqrt{\frac{b+d}{n^2}}}$$