

ESTADÍSTICA AMB R

ANÀLISI D'UNA VARIABLE

Variable Contínua	<ul style="list-style-type: none">○ Resums numèrics (summary statistics) <i>summary() # Exemple summary(bmi)</i> <i>min() # mínim</i> <i>max() # màxim</i> <i>mean() # mitjana (mean, average)</i> <i>median() # mediana (median)</i> <i>sd() # desviació típica (standard deviation)</i> <i>IQR() # rang interquartíl·lic (interquartile rang)</i> <i>quantile(,) # percentil Ex. percentil 95%: quantile(x, 0.95)</i> ○ Gràfic de punts <i>plot() # Exemple plot(bmi)</i> ○ Histograma <i>hist() # Exemple hist(bmi)</i> ○ Diagrama de Caixa (box plot / box-and-whisker plot) <i>boxplot() # Exemple boxplot(bmi)</i>
Variable Categòrica	<ul style="list-style-type: none">○ Taula de freqüències (frequency table) <i>table() # Taula freq. absolutes</i> <i>prop.table(table()) # Taula freq. relatives</i> <i>100*prop.table(table()) # Taula percentatges</i> <i># Exemples:</i> <i>table(tractament)</i> <i>prop.table(table(tractament))</i> <i>100*prop.table(table(tractament))</i> ○ Diagrama de barres (bar plot) <i>barplot(table()) # Exemple barplot(table(tractament))</i> ○ Diagrama de sectors (pie chart) <i>pie(table()) # Exemple pie(table(tractament))</i>

RELACIÓ ENTRE DUES VARIABLES

	Relació entre dues variables
Contínua & contínua	<ul style="list-style-type: none">○ Coeficient de correlació <code>cor()</code> # Exemple: <code>x<-c(2, 4, 1, 3, 6, 5)</code> <code>y<-c(3, 5, 2, 2, 6, 3)</code> <code>cor(x,y)</code>○ Recta de regressió <code>lm(y~x)</code>○ Diagrama de dispersió i recta de regressió <code>plot(x,y)</code> <code>abline(lm(y~x))</code>
Contínua & categòrica	<ul style="list-style-type: none">○ Resums numèrics de la variable contínua per a cada categoria de la variable categòrica <code>tapply(<contínua>, <categòrica>, <funció>)</code> # Exemple: <code>tapply(bmi, tractament, summary)</code> <code>tapply(bmi, tractament, mean)</code>○ Diagrames de caixes múltiples <code>boxplot(<contínua>~<categòrica>)</code> # Exemple: <code>boxplot(bmi~tractament)</code>
Categòrica & categòrica	<ul style="list-style-type: none">○ Taula de contingència <code>table(,)</code> # taula freq. absolutes <code>prop.table()</code> # proporció total <code>prop.table(,1)</code> # proporció fila <code>prop.table(,2)</code> # proporció columna <code>100*prop.table(,1)</code> # percentatge fila # Exemples: <code>taula<-table(tractament, diagnostic)</code> <code>prop.table(taula, 1)</code>○ Diagrames de barres apilats <code>barplot(table())</code> # Exemple: <code>barplot(table(tractament, diagnostic))</code>

Resum proves d'hipòtesis amb R

<i>y</i> variable continua <i>x</i> variable categòrica	Test de normalitat: Shapiro-Wilk H0: les dades segueixen una distribució normal H1: les dades no segueixen una distribució normal <i>shapiro.test(y)</i>	
	Si p-valor de Shapiro >0.05 Les dades segueixen una distribució normal	Si p-valor de Shapiro <0.05 Les dades NO segueixen una distribució normal
Test per a una mitjana H0: mitjana=valor predeterminat H1: mitjana≠valor predeterminat	Test t per a una mostra <i>t.test(y, mu=valor)</i>	Test de Wilcoxon per a una mostra <i>wilcox.test(y, mu=valor)</i>
Test d'igualtat de dues mitjanes H0: mitjana1=mitjana2 H1: mitjana1≠ mitjana2	Test t per a mostres independents (cal fer prèviament el test d'igualtat de variàncies) <i>t.test(y~x, var.equal=T)</i> # si les variàncies són iguals <i>t.test(y~x, var.equal=F)</i> # si les variàncies són diferents	Test de Wilcoxon per a mostres independents <i>wilcox.test(y~x)</i>
Test d'igualtat de dues mitjanes amb dades aparellades H0: mitjana1=mitjana2 H1: mitjana1≠ mitjana2	Test t per a dades aparellades $d < -y_1 - y_2$ <i>t.test(d, mu=0)</i>	Test de Wilcoxon per a dades aparellades <i>wilcox.test(y1, y2, paired=TRUE)</i>
Test d'igualtat de més de dues mitjanes H0: mitjana1 = mitjana2 = mitjana3 = ... = mitjanak H1: alguna de les mitjanes és diferent	ANOVA d'un factor <i>aov(y~x)</i> Post-hoc analysis: <i>TukeyHSD(aov)</i>	Test de Kruskal-Wallis <i>kruskal.test(y~x)</i>
Test d'igualtat de dues variàncies H0: variància1= variància2 H1: variància1≠ variància2	Prova F d'igualtat de variàncies <i>var.test(y~x)</i>	

Test d'igualtat de proporcions H0: proporció1= proporció2 H1: proporció1≠ proporció2	Prova d'igualtat de dues proporcions <i>prop.test(table(x1,x2))</i> # x1 i x2 són factors amb 2 categories
Test d'independència de dos variables categòriques H0: X i Y són independents H1: X i Y estan relacionades	Prova xi-quadrat d'independència de dos factors <i>chisq.test(table(x1,x2))</i> # x1 i x2 són variables categòriques