Pràctica 8

Contrastos d'hipòtesis

8.1 Contrasts sobre la mitjana μ

Suposem en tots els casos que tenim una mostra normal.

8.1.1 Amb σ coneguda

Considerem les dades del fitxer `x1.txt`. Hi ha $n = 20$ nombres, que posarem a un vector `x`. Suposem que les dades provenen d'una distribució normal amb desviació estàndar $\sigma = 0.9$.

Contrast unilateral: Volem fer el contrast de les hipòtesis

$$\begin{cases} H_0 : \mu = 3.5 \\ H_1 : \mu > 3.5 \end{cases}$$

amb nivell de significació $\alpha = 0.05$.

Procediment: Si és certa H_0,

$$Z = \frac{\overline{X} - 3.5}{\sigma} \sqrt{20} \sim N(0, 1).$$

Per al vector `x` calculem la mitjana empírica `xm<-mean(x)` i la mitjana empírica estandaritzada:

```
z<-(xm-3.5)/0.9*sqrt(20)
```

Obtenim $z = 1.328225$, és un valor raonable? Per poder contestar aquesta pregunta calculem el p-valor Prob$[Z > \mathrm{z}]$.

```
p<-1-pnorm(z)
```

Aquest valor ($p = 0.09205182$) representa la probabilitat que en una mostra de mida $n = 20$ d'una $N(\mu = 3.5, \sigma = 0.9)$ obtinguem un valor de la mitjana mostral $\overline{x}$ igual o superior al valor obtingut a partir de la mostra `x`. Per tant, si aquesta probabilitat és gran suposem que H_0 és certa i si és petita considerem H_1 certa.

Com de petita ha de ser aquesta probabilitat per decidir-nos per H_1? La resposta ens la dóna el **nivell de significació** α:

si $p \leq \alpha$ decidim a favor de H_1, i en cas contari suposem H_0 certa.

En el nostre cas hem de comparar el nivell de significació $\alpha = 0.05$ amb el p-valor que hem obtingut $p = 0.09205182$, com que $p > \alpha$ considerem H_0 certa.

Observació: Si haguéssim considerat el nivell de significació $\alpha = 0.1$, la nostra decisió hauria sigut rebutjar H_0 a favor de H_1.

Contrast bilateral: Per fer el contrast

$$\begin{cases} H_0 : \mu = 3.5 \\ H_1 : \mu \neq 3.5 \end{cases}$$

amb nivell de significació $\alpha = 0.05$ seguim el mateix procediment que abans, però ara el p-valor serà Prob$[|Z| > \text{z}]$ que es calcula

```
p<-2*(1-pnorm(abs(z)))
```

i dóna 0.1841036; seguint els mateixos passos que abans comparem `p` amb α. I per $\alpha = 0.05$ o 0.1 suposem H_0 certa.

8.1.2 Amb σ desconeguda

Considerem les dades del fitxer `x2.txt`. Suposem que les dades provenen d'una distribució normal amb σ desconeguda. En aquest cas, tenim una funció en `R` que ens realitza el contrast.

Contrast unilateral: Per fer el contrast de les hipòtesis

$$\begin{cases} H_0 : \mu = 12 \\ H_1 : \mu < 12 \end{cases}$$

amb nivell de significació $\alpha = 0.05$, utilitzem

```
t.test(x2,mu=12, alternative="less", conf.level=1-0.05)
```

En canvi, per fer el contrast

$$\begin{cases} H_0 : \mu = 12 \\ H_1 : \mu > 12 \end{cases}$$

utilitzem

```
t.test(x2,mu=12, alternative="greater", conf.level=1-0.05)
```

Contrast bilateral: Igualment, amb nivell de significació $\alpha = 0.05$, per fer el contrast de

$$\begin{cases} H_0 : \mu = 12 \\ H_1 : \mu \neq 12 \end{cases}$$

fem servir

```
t.test(x2,mu=12, alternative="two.sided",conf.level=1-0.05)
```

8.2 Contrasts sobre una proporció

En una mostra de $n = 32$ observacions, s'ha observat que en $N = 14$ d'elles s'ha produit un cert esdeveniment A.

Contrast unilateral: Amb nivell de significació $\alpha = 0.1$, realitzem el següent contrast sobre la probabilitat $p = \text{Prob}(A)$:

$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p < 0.5 \end{cases}$$

Pel teorema central del límit, per n prou gran, i sota la hipòtesi nul·la ($p = 0.5$) $p_0 = 0.5$,

$$Z = \frac{N - n\,0.5}{\sqrt{n\,0.5\,(1 - 0.5)}} = \frac{\frac{N}{n} - 0.5}{\sqrt{0.5\,(1 - 0.5)}}\sqrt{n} \sim N(0, 1)$$

Per la mostra donada calculem,

$$z = \frac{14 - 32 * 0.5}{\sqrt{32 \ * 0.5 \ * (1 - 0.5)}} = -0.7071068$$

El p-valor, calculant amb la llei $N(0, 1)$, s'obté fent

```
pnorm(-0.7071068)
```

Com que el resultat ($p = 0.2397501$) és més gran que $\alpha = 0.1$, acceptem H_0.

Contrast bilateral: En una mostra de $n = 29$ observacions, s'ha observat que en $N = 17$ d'elles s'ha produit un cert esdeveniment A. Amb nivell de significació $\alpha = 0.05$, realitzeu el següent contrast sobre $p = \text{Prob}(A)$:

$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p \neq 0.5 \end{cases}$$

Com abans, calculem

$$z = \frac{17 - 29\,0.5}{\sqrt{29\,0.5\,(1 - 0.5)}} = 0.9284767$$

Tractant-se d'un contrast bilateral, el p-valor serà

$$\text{Prob}(|Z| > 0.9284767) = 2 * \text{Prob}(Z > 0.9284767) = 0.3531604$$

que hem de comparar amb $\alpha = 0.05$, com que és més gran, hem d'acceptar H_0.

Amb aquestes mateixes dades, si tinguéssim un test unilateral,

$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p > 0.5 \end{cases}$$

el p-valor seria:

$$\text{Prob}(Z > 0.9284767) = 0.1765802$$

Com que una altra vegada el p-valor és més gran que α acceptem H_0.

8.3 Contrasts sobre variància σ^2

Suposem que tenim una mostra normal i volem fer una comparació de la variància mostral amb un σ_0^2 donat.

Test bilateral:

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2, \\ H_1 : \sigma^2 \neq \sigma_0^2. \end{cases}$$

Suposant H_0 certa,

$$Q = \frac{n\,S^2}{\sigma_0^2} = \frac{(n-1)\,\widetilde{S}^2}{\sigma_0^2} \sim \chi^2(n-1)$$

Posant en el lloc de la variable $\widetilde{S}^2$ el valor $\widetilde{s}^2$ observat a la mostra, calculem

$$q = \frac{(n-1)\,\widetilde{s}^2}{\sigma_0^2},$$

Per al test bilateral calculem el p-valor com

```
p<- min(pchisq(q,n-1),1-pchisq(q,n-1))
```

i, donat el nivell de significació α, si $p < \alpha/2$, cal rebutjar H_0.

Test unilateral a l'esquerra:

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2, \\ H_1 : \sigma^2 < \sigma_0^2. \end{cases}$$

calculem el p-valor com

```
p<- pchisq(q,n-1)
```

i, donat el nivell de significació α, si $p < \alpha$, cal rebutjar H_0.

Test unilateral a la dreta:

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2, \\ H_1 : \sigma^2 > \sigma_0^2. \end{cases}$$

calculem el p-valor com

```
p<- 1-pchisq(q,n-1)
```

i, donat el nivell de significació α, si $p < \alpha$, cal rebutjar H_0.

8.4 Contrasts de dues mostres normals independents

8.4.1 Coincideixen les variàncies?

Test bilateral:

$$\begin{cases} H_0 : \sigma_x^2 = \sigma_y^2, \\ H_1 : \sigma_x^2 \neq \sigma_y^2. \end{cases}$$

```
var.test(x, y, alternative="two.sided")
```

Test unilateral a l'esquerra:

$$\begin{cases} H_0 : \sigma_x^2 = \sigma_y^2, \\ H_1 : \sigma_x^2 < \sigma_y^2. \end{cases}$$

```
var.test(x, y, alternative="less")
```

Test unilateral a la dreta:

$$\begin{cases} H_0 : \sigma_x^2 = \sigma_y^2, \\ H_1 : \sigma_x^2 > \sigma_y^2. \end{cases}$$

```
var.test(x, y, alternative="greater")
```

8.4.2 Coincideixen les mitjanes?

Aquests tests requereixen del supòsit que les dues variàncies siguin iguals

Test bilateral:

$$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x \neq \mu_y \end{cases}$$

```
t.test(x,y,alternative="two.sided", var.equal=TRUE)
```

Test unilateral a l'esquerra:

$$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x < \mu_y \end{cases}$$

```
t.test(x,y,alternative="less", var.equal=TRUE)
```

Test unilateral a la dreta:

$$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x > \mu_y \end{cases}$$

```
t.test(x,y,alternative="greater", var.equal=TRUE)
```

8.4.3 La diferència $\mu_x - \mu_y$ és un valor d prefixat?

Test bilateral:

$$\begin{cases} H_0 : \mu_x - \mu_y = d \\ H_1 : \mu_x - \mu_y \neq d \end{cases}$$

```
t.test(x,y,mu=d,alternative="two.sided", var.equal=TRUE)
```

Test unilateral a l'esquerra:

$$\begin{cases} H_0 : \mu_x - \mu_y = d \\ H_1 : \mu_x - \mu_y < d \end{cases}$$

```
t.test(x,y,mu=d,alternative="less", var.equal=TRUE)
```

Test unilateral a la dreta:

$$\begin{cases} H_0 : \mu_x - \mu_y = d \\ H_1 : \mu_x - \mu_y > d \end{cases}$$

```
t.test(x,y,mu=d,alternative="greater", var.equal=TRUE)
```

8.5 Contrast d'una mostra de dades aparellades

Test bilateral:

$$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x \neq \mu_y \end{cases}$$

```
t.test(x,y,alternative="two.sided",paired=TRUE)
```

Test unilateral a l'esquerra:

$$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x < \mu_y \end{cases}$$

```
t.test(x,y,alternative="less",paired=TRUE)
```

Test unilateral a la dreta:

$$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x > \mu_y \end{cases}$$

```
t.test(x,y,alternative="greater",paired=TRUE)
```

8.6 Problemes

1. Després d'un tractament contra l'obesitat, els pesos en Kg de vuit dones eren

$$58, 50, 60, 65, 64, 62, 56, 57.$$

   Suposeu normalitat i $\alpha = 0.1$.

   a) Podem afirmar que la mitjana teòrica μ és 61, sabent que $\sigma = 3$?
   b) Podem afirmar que la mitjana teòrica μ és 61, si σ és desconeguda?
   c) Podem afirmar que la variància teòrica σ^2 és superior a 10?

2. En unes eleccions on participen dos candidats, A i B, es va dur a terme un sondeig d'opinió amb 1000 votants seleccionats a l'atzar. D'ells, 615 prefereixen A.

   Podem afirmar que la proporció de votants que prefereixen A és superior al 60%?

3. A partir de les dades de l'arxiu `tterreny.txt`, contrasteu les hipòtesis següents amb nivell de significació $\alpha = 0.05$.:

   a) El consum mitjà a 120 Km/h és de 12 litres.
   b) La mitjana de la velocitat màxima és de 155 Km/h.
   c) La mitjana del consum urbà és inferior de 12.2 litres.
   d) A partir de les dades de les variables `Consum90` i `Consum120`, podem acceptar com a vàlida l'afirmació que $\mu_{120} = \mu_{90}$?
   e) A partir de les dades de les variables `Consum120` i `ConsumUrba`, podem acceptar com a vàlida l'afirmació que $\mu_{120} = \mu_U$?
   e) A partir de les dades de les variables `Consum120` i `ConsumUrba`, podem acceptar com a vàlida l'afirmació que $\mu_{120} = \mu_{90} + 2$?

4. Un inversionista ha de decidir entre dos tipus de valors. Té la informació següent sobre el rendiment dels valors, expressat en $\%$ del preu.

Valor A	7.8	10.3	7.9	8.7	9.2	8.9
Valor B	9.2	9.1	11.1	8.8	9.6	

Suggereixen aquestes dades que el valor B produeix diferent rendibilitat que el valor A?

5. S'ha sembrat un cereal en 10 parcel·les sense adob. Després d'un any, s'ha sembrat en les mateixes parcel·les el mateix cereal utilitzant un adob. Els resultats obtinguts han estats els següents:

Sense adob	5.4	5.8	5.5	5.7	6.0	5.3	5.9	5.8	6.8	5.8
Amb adob	6.5	5.9	6.6	6.1	5.8	6.1	5.7	6.2	6.2	6.4

Hi ha diferència significativa entre els dos cultius? És cert que la mitjana de la collita utilitzant adob és superior a la mitjana de la collita sense utilitzar adob?

6. Un fabricant desitja comparar el rodatge de dues classes de pneumàtics A i B. Per tant, es seleccionen aleatòriament cinc pneumàtics de cada classe i posen un de cada classe a les rodes posteriors de cinc automòbils, que recorren una distància preestablerta, registrant-se la quantitat desgastada de cada pneumàtic. Les dades obtingudes són les següents:

Automòbil	1	2	3	4	5
Pneumàtic A	10.6	9.8	12.3	9.7	9.1
Pneumàtic B	10.2	9.4	11.8	9.1	8.3

Hi ha diferència significativa entre les dues classes de pneumàtics? ($\alpha = 10\%$).

7. Les dades de la taula següent representen el sou actual i el sou inicial, en euros, de 10 empleats d'una empresa.

Homes		Dones	
Sou actual	Sou inicial	Sou actual	Sou inicial
57000	27000	21450	12000
40200	18750	21900	13200
45000	21000	21900	9750
32100	13500	27900	12750
36000	18750	24000	13500

a) Podem afirmar que la mitjana teòrica del sou actual dels homes és de 42000 euros? (Nivell de significció $\alpha = 0.05$).

b) Les dones, han tingut un augment de sou significatiu? (Nivell de significació $\alpha = 0.05$).

c) Podem afirmar que el sou mitjà inicial dels homes és més alt que el de les dones?

8. Es varen provar dos tractaments antireumàtics, administrant-los a 10 i 15 pacients, respectivament. Els resultats després del tractament (a més puntuació, més eficàcia), varen ser:

Tractament A: $15 \quad 71 \quad 21 \quad 17 \quad 40 \quad 42 \quad 10 \quad 23 \quad 35 \quad 28,$

Tractament B: $21 \quad 20 \quad 42 \quad 25 \quad 16 \quad 52 \quad 65 \quad 40 \quad 43 \quad 35 \quad 18 \quad 56 \quad 29 \quad 32 \quad 44.$

Amb nivell de significació $\alpha = 0.1$, hi ha diferències significatives entre els dos tractaments?

9. S'han mesurat les longituds (en mm) de les ales en 7 espècimens d'insectes nascuts en laboratori (L) i 7 de capturats (C). Els resultats són:

L:	6.7	1.9	6.4	4.8	2.6	4.9	6.7
C:	6.2	3.7	4.5	6.2	6.0	5.3	3.5

Podem afirmar que les mitjanes teòriques de les longituds són iguals en les dues poblacions? Nivell de significació $\alpha = 0.05$.

10. El test de personalitat de Eysenck té dues formes, A i B. Comproveu que totes dues mesuren l'extroversió de forma equivalent a partir dels següents resultats sobre 12 individus:

Individu	1	2	3	4	5	8	7	8	9	10	11	12
Forma A	12	18	21	10	15	27	31	6	15	13	8	10
Forma B	9	10	18	5	21	20	27	7	11	10	7	7

Considereu nivell de significació $\alpha = 0.05$.