

PAC3 (Pràctica)

Josep Andreu Miralles

31/10/2020

EXERCICI 1

Cal que importeu a R el fitxer `datos_clin.csv` tenint en compte que el separador decimal és la coma (",") i el de camps és ";". En aquest fitxer trobareu les següents variables referides a dades clíniques de pacients abans i després de fer un tractament:

Suposarem que la variable SBP1 segueix una distribució normal:

- Trobeu els intervals de confiança per a la mitjana de la SBP1 amb nivells de confiança del 99% i del 90%. Quin és més gran? Quin ens dona un interval de confiança que amb més seguretat contindrà la mitjana? Compareu els resultats.
- Trobeu un interval de confiança al 90% per a la proporció d'individus amb un nivell de pressió sistòlica abans de fer el tractament superior a 130. Aquest apartat l'heu de fer aplicant les fórmules de l'interval de confiança que es troben als apunts de teoria.

Solució

a) Trobeu els intervals de confiança per a la mitjana de la SBP1 amb nivells de confiança del 99% i del 90%. Quin és més gran? Quin ens dona un interval de confiança que amb més seguretat contindrà la mitjana? Compareu els resultats.

```
library(readr)
dades<-read_csv("datos_clin.csv",header=TRUE,sep=";",dec=",")
t.test(dades$SBP1,alternative="two.sided",mu=0.0,conf.level=0.99)
```

```
##
## One Sample t-test
##
## data:  dades$SBP1
## t = 36.246, df = 110, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 99 percent confidence interval:
##  98.2005 113.5112
## sample estimates:
## mean of x
## 105.8559
```

```
t.test(dades$SBP1,alternative="two.sided",mu=0.0,conf.level=0.90)
```

```
##  
## One Sample t-test  
##  
## data: dades$SBP1  
## t = 36.246, df = 110, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 0  
## 90 percent confidence interval:  
## 101.0113 110.7004  
## sample estimates:  
## mean of x  
## 105.8559
```

Interval de confiança al 99% i 90% per a la mitjana de SBP1 és (98.2005 , 113.5112) i (101.0113 , 110.7004) respectivament. L'interval més gran corresponent al del 99%, ja que en aquest cas tenim tan sols una probabilitat de l'1% de que la mitjana poblacional no pertanyi a l'interval calculat, per tant la probabilitat d'equivocar-nos es menor en l'interval del 99% que no en el del 90%.

b) Trobeu un interval de confiança al 90% per a la proporció d'individus amb un nivell de pressió sistòlica abans de fer el tractament superior a 130. Aquest apartat l'heu de fer aplicant les fórmules de l'interval de confiança que es troben als apunts de teoria.

```
library(readr)  
dades<-read_csv("datos_clin.csv",header=TRUE,sep=";",dec=",")  
dades_SBP1<-dades$SBP1  
dades_SBP1<-na.omit(dades_SBP1)  
dades_SBP1m130<-dades_SBP1[dades$SBP1 > 130 & !is.na(dades$SBP1)]  
n_SBP1<-length(dades_SBP1)  
n_SBP1m130<-length(dades_SBP1m130)  
n_SBP1
```

```
## [1] 111
```

```
n_SBP1m130
```

```
## [1] 28
```

```
prop_SBP1m130<-n_SBP1m130/n_SBP1  
prop_SBP1m130
```

```
## [1] 0.2522523
```

```
int_sup<-prop_SBP1m130+(1.645*sqrt((prop_SBP1m130*(1-prop_SBP1m130))/n_SBP1))  
int_inf<-prop_SBP1m130-(1.645*sqrt((prop_SBP1m130*(1-prop_SBP1m130))/n_SBP1))  
int<-c(int_inf,int_sup)  
int
```

```
## [1] 0.1844414 0.3200632
```

EXERCICI 2

Es vol fer un estudi sobre l'ús de la biblioteca que fan els estudiants de la UOC. Per això es pren una mostra de 100 estudiants i es comprova si s'han connectat a la biblioteca durant el darrer mes. En cas afirmatiu es mira quants cops ho han fet. Dels 100 estudiants escollits de forma aleatòria 35 s'havien connectat a la biblioteca i la resta no.

- a) Obteniu un interval de confiança del 95% per a la proporció d'individus que s'havien connectat a la biblioteca. Amb aquests resultats es pot admetre que la meitat dels estudiants s'han connectat a la biblioteca?

D'altra banda, obtenim ara les dades de quants cops s'han connectat els 35 estudiants que sí que ho havien fet:

2, 4, 1, 1, 8, 2, 3, 6, 7, 10, 1, 1, 3, 4, 11, 6, 20, 1, 8, 9, 22, 14, 16, 4, 5, 3, 3, 6, 9, 10, 6, 7, 8, 5, 6

- b) Trobeu l'interval de confiança del 90% per a la mitjana del nombre de connexions entre els que es van connectar.
c) Interpreteu l'interval de confiança.

Solució

- a) Obteniu un interval de confiança del 95% per a la proporció d'individus que s'havien connectat a la biblioteca. Amb aquests resultats es pot admetre que la meitat dels estudiants s'han connectat a la biblioteca?

```
prop.test(35,100,alternative="two.sided", p=0.5, conf.level = 0.95, correct = FALSE)
```

```
##
## 1-sample proportions test without continuity correction
##
## data: 35 out of 100, null probability 0.5
## X-squared = 9, df = 1, p-value = 0.0027
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
## 0.2636425 0.4474556
## sample estimates:
## p
## 0.35
```

L'interval de confiança és (0.2636425 , 0.4474556), com veiem 0.50 queda fora de l'interval, per la qual cosa no es pot admetre que la meitat dels estudiants s'hagin connectat a la biblioteca.

- b) Trobeu l'interval de confiança del 90% per a la mitjana del nombre de connexions entre els que es van connectar.

```
dades2<-c(2, 4, 1, 1, 8, 2, 3, 6, 7, 10, 1, 1, 3, 4, 11, 6, 20, 1, 8, 9, 22, 14, 16, 4, 5, 3, 3, 6, 9, 10, 6, 7, 8, 5, 6)
t.test(dades2,alternative="two.sided",mu=0.0,conf.level=0.90)
```

```
##
## One Sample t-test
##
## data:  dades2
## t = 7.6298, df = 34, p-value = 7.239e-09
## alternative hypothesis: true mean is not equal to 0
## 90 percent confidence interval:
##  5.159543 8.097600
## sample estimates:
## mean of x
##  6.628571
```

L'interval de confiança al 90% per a la mitjana és (5.159543 8.097600)

c) Interpreteu l'interval de confiança.

Indica que dins del rang (5.159543 8.097600) hi ha una 90% de probabilitats que s'hi trobi la mitjana poblacional.

EXERCICI 3

Disposem d'una mostra dels treballadors d'una empresa i, de cara a encarregar els uniformes, volem estimar quin és l'altura mitjana dels homes. Realitzem el següent test amb R:

```
hombres <- datos[datos$sexo=="Hombre", ] t.test(x=hombresaltura, conf.level = 0.90)
OneSamplet-testdata : hombresaltura t = 118.03, df = 17, p-value < 2.2e-16 alternative hypothesis: true mean is not equal to 0 90 percent confidence interval: 176.4384 181.7172 sample estimates: mean of x 179.0778
```

A partir d'aquests resultats respongueu les següents preguntes: a) Quina és la mitjana de l'altura dels homes en la mostra? b) Quin estadístic hem utilitzat per obtenir l'interval de confiança i quin valor de l'estadístic hem obtingut? c) Quin és l'interval de confiança per a la mitjana de l'altura dels homes d'aquesta empresa? Amb quin nivell de confiança? d) Quants homes hi havia en la mostra?

Solució

a) Quina és la mitjana de l'altura dels homes en la mostra?

La mitjana mostral és 179.0778.

b) Quin estadístic hem utilitzat per obtenir l'interval de confiança i quin valor de l'estadístic hem obtingut?

Hem utilitzat l'estadístic t-student. I el resultat ens indica que el resultat de t és igual a 118.03.

c) Quin és l'interval de confiança per a la mitjana de l'altura dels homes d'aquesta empresa? Amb quin nivell de confiança?

L'interval de confiança és (176.4384 181.7172), amb un nivell de de confiança del 90%.

d) Quants homes hi havia en la mostra?

En la mostra hi havia 18 homes.