

UOC – Estudis d’Informàtica Estadística.

Primavera 2020 PAC4. Part de R.

Data de proposta	Data de lliurament
15/4/20	21/4/20

Heu de resoldre els problemes següents amb R i heu d’entregar la solució en format pdf **indicant les funcions de R i les sortides que R us dóna**:

Problema Cal que importeu a R el fitxer “PIS_MUN” que ja heu utilitzat a la PAC1. En aquest fitxer hi trobareu informació sobre el nombre de contractes de lloguer i el lloguer mitjà contractual (mitjana dels preus dels lloguers per contracte) als municipis de Catalunya i als districtes de la ciutat de Barcelona. Concretament trobareu les variables.

- COD = codi postal del municipi (o “08019” en cas dels districtes de Barcelona),
- MUNDIS = Nom del municipi o districte de la ciutat de Barcelona,
- MUN = Nom del municipi o “Barcelona” en el cas dels seus districtes,
- PROV = Província la que pertany el municipi/districte.
- NAny = Nombre de contractes signats el corresponent any al municipi/districte (on Any pot ser 2018, 2017, 2016 o 2015),
- MAny = Lloguer mitjà mensual dels contractes signats el corresponent any al municipi/districte.

En aquesta PAC suposarem que les variables corresponen a una mostra de ubicacions.

Heu de respondre les següents preguntes en un fitxer pdf que cal lliurar en aquesta activitat. Suposarem que les variables Nany i MAny segueixen una distribució normal.

- 1) Trobeu intervals de confiança per a la mitjana del lloguer mitjà mensual dels contractes signats els anys 2015, 2016, 2017 i 2018 amb nivell de confiança del 90%. Hi trobeu qualche diferència? Indicació: mirau-vos el mòdul 4. Distribucions de probabilitat i inferència estadística amb R-Commander, pàgines 28, 29, 30 i 31.
- 2) Trobeu un interval de confiança al 95% per a la mitjana del lloguer mitjà mensual dels contractes signats l’any 2016 als municipis/districtes de la província de Barcelona. Feu el mateix per les altres tres províncies. Hi ha

diferències entre els distints grups? Indicació: mirau-vos el mateix que en l'apartat anterior.

- 3) Trobeu un interval de confiança al 95% per la proporció de municipis/districtes on durant l'any 2017 tenen més de 100 contractes signats. Aquest apartat l'heu de fer aplicant les fórmules de l'interval de confiança que es troben a les notes d'estudi. Indicació: mirau-vos les pàgines 21 i 22 del mòdul 6 de les notes d'estudi.

Resposta

Primer carreguem les dades:

```
dades=read.csv("PIS_MUN.csv",header=TRUE,dec=",",sep=";")
```

- 1) Per calcular els intervals de confiança demanats, fem servir la funció `t.test`:

```
t.test(dades$M2015,conf.level = 0.9)
```

```
##
## One Sample t-test
##
## data: dades$M2015
## t = 30.683, df = 955, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 90 percent confidence interval:
## 236.8904 263.7552
## sample estimates:
## mean of x
## 250.3228
```

```
t.test(dades$M2016,conf.level = 0.9)
```

```
##
## One Sample t-test
##
## data: dades$M2016
## t = 31.407, df = 955, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 90 percent confidence interval:
## 258.6165 287.2319
## sample estimates:
```

```
## mean of x
## 272.9242
```

```
t.test(dades$M2017,conf.level = 0.9)
```

```
##
## One Sample t-test
##
## data: dades$M2017
## t = 32.364, df = 955, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 90 percent confidence interval:
## 275.8787 305.4532
## sample estimates:
## mean of x
## 290.666
```

```
t.test(dades$M2018,conf.level = 0.9)
```

```
##
## One Sample t-test
##
## data: dades$M2018
## t = 33.306, df = 955, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 90 percent confidence interval:
## 296.0063 326.7936
## sample estimates:
## mean of x
## 311.4
```

Observam que el lloguer mitjà ha anat pujant des de l'any 2015 fins a l'any 2018.

- 2) L'interval de confiança per a la mitjana del lloguer mitjà mensual dels contractes signats l'any 2016 als municipis/districtes de la província de Barcelona val:

```
t.test(dades$M2016[dades$PROV=="Barcelona"],conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: dades$M2016[dades$PROV == "Barcelona"]
## t = 28.373, df = 319, p-value < 2.2e-16
```

```
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 397.6002 456.8490
## sample estimates:
## mean of x
## 427.2246
```

Els altres intervals de la mateixa variable per les altres províncies serà:

```
t.test(dades$M2016[dades$PROV=="Girona"],conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: dades$M2016[dades$PROV == "Girona"]
## t = 16.45, df = 220, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 265.8846 338.2646
## sample estimates:
## mean of x
## 302.0746
```

```
t.test(dades$M2016[dades$PROV=="Lleida"],conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: dades$M2016[dades$PROV == "Lleida"]
## t = 9.3897, df = 230, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 79.23915 121.32587
## sample estimates:
## mean of x
## 100.2825
```

```
t.test(dades$M2016[dades$PROV=="Tarragona"],conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: dades$M2016[dades$PROV == "Tarragona"]
```

```
## t = 12.676, df = 183, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  157.3050 215.3031
## sample estimates:
## mean of x
##  186.304
```

Observem que la mitjana de lloguers més alts es produeixen a la província de Barcelona i els més baixos a la província de Lleida.

- 3) Primer calculem la proporció de municipis/districtes on durant l'any 2017 tenen més de 100 contractes signats:

```
(nombre.més.100 = length(dades$N2017[dades$N2017 >100]))
```

```
## [1] 181
```

```
(nombre.total = length(dades$N2017))
```

```
## [1] 956
```

```
(prop.mostral = nombre.més.100/nombre.total)
```

```
## [1] 0.1893305
```

Ens demanem l'interval de confiança al 95% de confiança per la proporció de municipis/districtes on durant l'any 2017 tenen més de 100 contractes signats.

```
alfa=0.05
```

```
(esquerra=prop.mostral-qnorm(1-alfa/2)*
  sqrt(prop.mostral*(1-prop.mostral)/nombre.total))
```

```
## [1] 0.1644963
```

```
(dreta=prop.mostral+qnorm(1-alfa/2)*
  sqrt(prop.mostral*(1-prop.mostral)/nombre.total))
```

```
## [1] 0.2141648
```

L'interval demanat serà (0.164, 0.214).