

Introducció a R

R (<http://www.r-project.org/>) és una eina de lliure distribució orientat a la realització de càlculs estadístics.

Lliurament

Creeu un document on es vegi el resultat de cadascun dels següents exercicis. En acabar la sessió, pugeu el document a través de l'activitat que s'ha creat al campus virtual. Aquesta activitat té un pes del 5% sobre la nota final de l'assignatura. El nom del fitxer ha de ser **Cognom1Nom1_Cognom2Nom2_ActivitatR.pdf**. Qualsevol fitxer que es lliuri sense aquesta nomenclatura no serà corregit.

Primer exercici

1. Inicieu el programa R i situeu-vos en una consola.
2. Anem a generar un vector amb dades (anomenat 'dades'). Introduïu la comanda:
dades<-c(7.9, 6.2, 8.3, 6.3, 8.1, 9.1, 7.0, 7.6, 6.8, 3.5, 4.0, 6.3, 7.9, 10.0)
3. Per visualitzar el contingut del vector, executeu:
dades
4. Calculeu, utilitzant una calculadora, la mitjana d'aquest conjunt i anoteu-ne el valor.
5. R facilita molt el treball amb vectors de dades. Per exemple, per calcular la suma de tots els seus valors, podem executar:
sum(dades)
6. La longitud d'un vector s'obté mitjançant:
length(dades)
7. Finalment, podem calcular el valor mitjà del vector com
sum(dades)/length(dades)
8. Per fer-ho més fàcil, R ens proporciona una comanda que fa el càlcul directament:
mean(dades)
9. Podem calcular la variància de les nostres dades mitjançant:
sum((dades-mean(dades))^2)/length(dades)
10. Compareu el resultat amb allò que obteniu en executar la comanda:
var(dades)
11. Haureu pogut comprovar que els dos resultats no coincideixen. Recordeu que molt programari implementa l'anomenada 'variància ajustada'. Modifiqueu la fórmula de l'apartat 9 per a calcular aquest valor.
12. Si voleu veure el vostre conjunt ordenat:
sort(dades)
13. Quina és la mediana del vostre conjunt?
14. La mediana es pot obtenir mitjançant:
median(dades)
15. Calculeu manualment el valor de la primera i la tercera quartil·la del vostre conjunt.
16. Executeu la comanda
quantile(dades,c(0.25))

17. Utilitzeu R per calcular la tercera quantil·la. Tingueu en compte que R implementa fins 9 criteris diferents per al càlcul de les quantil·les. Per tant, és molt probable que el criteri utilitzat per defecte difereixi del vostre.
18. Indiqueu a R que dibuixi un diagrama de capsa del vostre conjunt de dades:
`boxplot(dades)`
19. Genereu un nou vector que contingui 10 cops el valor 3. Tot seguit calculeu-ne la mitjana i la variància. Comproveu que obteniu els valors esperats.
`dades2<-seq(length=10,from=3,to=3)`

Segon exercici

Si llencem 10 daus, el nombre de cops que obtindrem el número '3' pot oscil·lar entre 0 i 10. Aquest nombre segueix una distribució de probabilitat binomial, Bin(10, 1/6).

1. Sabent que la funció **choose(n,m)** calcula el número combinatori $\binom{n}{m}$, calculeu, amb R, utilitzant la fórmula adequada, la probabilitat amb que una variable aleatòria que segueix una distribució binomial Bin(10, 1/6) retorna cadascun dels valors del seu espai mostral.
2. Compareu els resultats que heu obtingut a l'apartat anterior amb aquells retornats per la funció **dbinom(x, 10, 1/6)**, per cadascun dels valors x. Aquesta funció retorna la probabilitat p[Bin(10, 1/6) = x].
3. Els càlculs anteriors es poden automatitzar de la següent manera:
`nums<-seq(0,10,by=1)`
`probabilitats<-dbinom(nums,10,1/6)`
4. Utilitzant la funció **sum**, comproveu que la suma de les probabilitats dels valors l'espai mostral dona 1.
5. Podeu visualitzar com es distribueix aquesta probabilitat mitjançant:
`plot(seq(0,10,by=1),probabilitats)`
6. I també:
`lines(seq(0,10,by=1),probabilitats)`

Tercer exercici

Si llencem un dau repetidament fins obtenir el primer '5' i comptem el nombre de llançaments que hem hagut de fer, aquest nombre segueix una distribució geomètrica, Geom(1/6).

1. Utilitzant R, calculeu la probabilitat de que el nombre de llançaments sigui 1, 2, 3, ..., fins 8.
2. La funció **dgeom(x, 1/6)**, retorna la probabilitat de que el primer '5' aparegui **després d'haver realitzat 'x' llançaments fallits**. Utilitzeu aquesta funció per comprovar els resultats que heu obtingut a l'apartat anterior.
3. Genereu una gràfica que mostri com es distribueixen aquestes probabilitats.