

Segona entrega

Instruccions

El treball s'ha d'entregar abans de dilluns **21 de novembre** a les **23:59h**.

Heu d'entregar dos fitxers: un pdf amb nom *E2_Grup Δ .pdf* i un R Script amb nom *E2_Grup Δ .R*, on en lloc de Δ heu de posar la lletra que corresponen al vostre grup. Només ha d'entregar una persona per cada grup.

Les bases de dades que heu d'utilitzar les trobeu al campus virtual.

Responen a cada exercici de forma completa i ordenada i especifiqueu a quin apartat esteu contestant. Les respostes que no siguin justificades poden tenir puntuació nul·la.

Le figures que adjunteu han de tenir títol o descripció i els eixos s'han d'etiquetar.

Podeu incloure al pdf els trossos de codi que us semblin rellevants.

L'avaluació del treball es farà majoritàriament del fitxer pdf. El script es mirarà en el cas que alguna cosa no quedi clara en el pdf, tot i que s'avaluarà que el codi entregat sigui net i ordenat. En el script poseu els comentaris que considereu oportuns i especifiqueu sempre els exercicis i apartats que esteu resolent.

Exercicis

1. Les dades del fitxer **enquesta.csv** corresponen a una enquesta feta a 120 persones escollides a l'atzar en un centre cultural. Les variables són:

EDAT	Edat en anys.
SEXE	Sexe (<i>h</i> = home, <i>d</i> = dona).
PES	Pes en kilograms.
ALTURA	Alçada en centímetres.
TABAC	Si té hàbits de fumador (0 = no, 1 = sí).
OCI	Activitat d'oci a la qual dediquen majoritàriament el seu temps lliure (1 = televisió, 2 = ordinador, 3 = música o lectura, 4 = esports).

Responen a les següents preguntes explicant els càlculs que feu i quines fórmules utilitzeu.

- a) Si agafem una persona a l'atzar entre les enquestades, quina és la probabilitat que fumi? Aquesta probabilitat és més alta entre els que fan esport o els que no?
- b) Si agafem una dona a l'atzar entre les enquestades, quina és la probabilitat que sigui alta més de 1.60 m? I la probabilitat que l'activitat d'oci a la qual majoritàriament es dediqui la dona sigui la televisió o l'ordinador?

- c) Si agafem un home que practica esport a l'atzar entre els enquestats, quina és la probabilitat que fumi? I si agafem una persona que fuma a l'atzar entre els enquestats, quina és la probabilitat que sigui home i practiqui esport?
 - d) Entre les persones enquestades, n'hi més que fan esport i no fumen o que dediquen el seu temps lliure a l'ordinador o a la música (o lectura) i fumen?
 - e) Entre els enquestats, és més probable agafar una persona que pesi més de 60 kg entre les que tenen almenys 20 anys o una persona que pesi menys de 70 kg entre les que tenen com a molt 50 anys?
2. Suposem que tenim un dau una mica particular: té les cares numerades de l'1 al 6, però no és regular. Sigui X la variable aleatòria que representa el resultat d'aquest dau. La seva funció de massa de probabilitat és:

$$P(X = k) = \frac{k-1}{5k} \quad \text{si } k \text{ és parell,}$$

$$P(X = k) = \frac{1+2k}{36} \quad \text{si } k \text{ és senar.}$$

- a) Escriviu en R una funció `dau()` que, donat un valor $k \in \mathbb{R}$, us torni la funció de massa de probabilitat de X per aquest valor. Copieu al pdf el codi que heu escrit.
 - b) Representeu la funció de massa de probabilitat i la corresponent funció de distribució de X . Representeu en un únic gràfic la funció de distribució de X i la d'un dau regular i compareu-les.
 - c) Suposeu de tirar el dau un cop. Per quin valor apostaríeu? Apostaríeu abans per un resultat parell o per un senar? Apostaríeu abans per un resultat més gran que 4 o per un més petit que 5? Justifiqueu les respostes.
 - d) Calculeu l'esperança i la variància teòriques de X especificant els càlculs que feu. Genereu una mostra de $n = 300$ tirades del dau i calculeu-ne la mitjana i la variància empíriques. Heu obtingut una bona aproximació? Justifiqueu la resposta.
 - e) Suposeu de tirar el dau dos cops i sumeu els dos resultats. Si sabem que la suma obtinguda és més gran o igual que 5, quina és la probabilitat que sigui menor que 11?
3. Sigui X una variable aleatòria absolutament contínua amb funció de densitat:

$$f(x) = \frac{1}{4}(x-1)^3, \quad 1 < x < 3.$$

- a) Representeu la funció de densitat. Calculeu la funció de distribució de X i representeu-la.
- b) Simuleu $n = 80000$ valors de X i dibuixeu l'histograma. Superposeu-hi el dibuix de la funció de densitat.
- c) Calculeu teòricament l'esperança i la variància de X . Després calculeu la mitjana i variància empíriques de la mostra que heu generat a l'apartat anterior. Heu obtingut una bona aproximació? Justifiqueu la resposta.

- d) Utilitzeu la mostra generada a l'apartat b) per comprovar el teorema central del límit. Expliqueu i justifiqueu cada pas que feu.
- e) Calculeu el valor t tal que

$$P(X < t) = 0.8.$$

Quantes de les simulacions que heu fet tenen un valor menor de t ? És el que us esperàveu? Justifiqueu la resposta.

4. Una empresa utilitza una màquina per omplir les ampolles d'un refresc. Es suposa que les ampolles contenen 300 ml. En realitat, el contingut de les ampolles varia segons una distribució normal de mitjana $\mu = 298$ ml i desviació típica $\sigma = 3$ ml.
- a) Quina és la probabilitat que una ampolla contingui menys de 295 ml? I entre 290 ml i 310 ml?
- b) Quina és la quantitat mínima de refresc que pot contenir una ampolla per estar en el 5% de les més plenes?
- c) L'empresa empaqueta el refresc en caixes de 6 ampolles cadascuna. En compro 100 caixes. Em retornen els diners d'una caixa si hi ha almenys 2 ampolles que contenen menys de 295 ml. Denotem amb X el nombre de caixes de les quals em retornen els diners. De quina variable es tracta? Calculeu quina és la probabilitat (exacta) que em retornin els diners d'almenys 30 caixes.
- d) La variable X es pot aproximar per una distribució Normal. Quins són els paràmetres d'aquesta distribució? Justifiqueu la resposta.
- e) Utilitzeu l'aproximació de l'apartat anterior per calcular la probabilitat (aproximada) que em retornin els diners d'almenys 30 caixes. Considereu que heu obtingut una bona aproximació? Justifiqueu la resposta.