UAB

Econometria I

Llistat 1: Exercicis de repàs d'estadística

Una variable: repàs

1. Considera una variable aleatòria Z que només pot prendre cinc valors, cadascun amb la mateixa probabilitat:

$$Z = \{z_1, z_2, z_3, z_4, z_5\}.$$

La variable Z és una variable mesurada en milers d'euros.

(a) Dóna l'expressió que defineix el valor esperat de Z, $\mathbb{E}[Z]$, la variància, $\mathrm{Var}(Z)$, i la desviació estandard, $\mathrm{Sd}(Z)$. Presenta les expressions utilitzant sumatoris.

(b) Quina informació sobre el comportament de Z proporciona cadascun d'aquests paràmetres?

(c) Considera ara que tenim els cinc valors que la variable Z pot prendre són:

$$Z = \{1, 2, 3, 4, 5\}.$$

Amb l'ajuda d'un full de càlcul (Excel o similar) calcula $\mathbb{E}[Z]$, $\mathrm{Var}(Z)$ i $\mathrm{Sd}(Z)$, definint les fórmules corresponents. *(És a dir, no utilitzis les funcions pre-definides pel full de calcul).*

(d) Considera ara que redefinim Z en euros. És a dir, volem canviar les unitats de mesura de Z. Etiquetem a la variable Z definida en euros com Z^*. Així, ara els cinc valors que Z^* pot prendre són:

$$Z^* = \{1000, 2000, 3000, 4000, 5000\}.$$

Modifica el full de càlcul que has creat abans i calcula $\mathbb{E}[Z^*]$, $\mathrm{Var}(Z^*)$ i $\mathrm{Sd}(Z^*)$.

(e) Donat els resultats anteriors, diries que el valor esperat d'una variable depèn de les unitats de mesura o és independent a aquests canvis? I la variància? I la desviació estàndard? Justifica.

(f) Entra les observacions de Z a *Gretl*. Crea un guió d'instruccions i, utilitzant les instruccions corresponents, calcula $\mathbb{E}[Z]$, $\mathrm{Var}(Z)$, $\mathrm{Sd}(Z)$ y $\mathbb{E}[Z^*]$, $\mathrm{Var}(Z^*)$, $\mathrm{Sd}(Z^*)$. Coincideixen aquests valors amb els teus càlculs anteriors? Comenta.

2. Defineix la variable aleatòria Z igual al resultat de tirar un dau no trucat.

(a) És Z una variable aleatòria continua o discreta?

(b) Troba la funció de probabilitat de Z. Dibuixa-la. Com etiquetaries aquest tipus de distribució?

3. Considera que $Z \sim \mathcal{N}(0,1)$. Amb l'ajuda de les taules que proporciona *Gretl*, troba el valor de c en cada cas:

   (a) $\Pr\{Z > c\} = 0.025$

   (b) $\Pr\{|Z| > c\} = 0.10$

   (c) $\Pr\{Z > 1.96\} = c$

   (d) $\Pr\{|Z| > 1.96\} = c$

   (e) $\Pr\{-1.96 < Z < 1.96\} = c$

4. Si sabem que $Z_1 \sim \mathcal{N}(0,1)$, què podries dir sobre la distribució (tipus, esperança i variància) de les següents variables:

   (a) $Z_2 = Z_1 \cdot 4$

   (b) $Z_3 = 2 + 4Z_1$

   (c) $Z_4 = Z_1^2$

5. Coneixem que la població d'una variable aleatòria Z ve donada per:

$$Z = \{6, 7, 6, 8, 5, 7, 6, 9, 10, 6\}.$$

   (a) Calcula $\mathbb{E}[Z]$ i $\mathrm{Var}(Z)$.

   (b) Considera que en lloc de tota la població, només tenim accés a les observacions incloses en la següent mostra, extreta aleatòriament, d'aquesta població: $\{6, 8, 10\}$. Calcula la mitjana mostral, $\bar{Z}$, i la variància mostral de Z, $\widehat{\mathrm{Var}(Z)}$.

   (c) Considera ara que en lloc de la mostra anterior tenim la següent mostra procedent també d'aquesta població: $\{6, 5, 9\}$. Calcula la mitjana mostral, $\bar{Z}$, i la variància mostral de Z associada a aquesta segona mostra.

   (d) Creus que la mitjana mostral, $\bar{Z}$, és una variable aleatória? I la variància mostral? Raona la resposta.

Una variable: simulació

6. Crea un fitxer d'instruccions de *Gretl* que generi 100 observacions d'una variable aleatòria $Z \sim \mathcal{N}(0,1)$. Abans de la comanda que genera la variable, inclou la comanda "`set seed 1234`". Executa el guió. Un cop generades aquestes observacions, utilitzant els menús de la finestra principal:

   (a) Troba el histograma de densitat amb les observacions de Z de la mostra. Comenta.

   (b) Troba els estadístics descriptius bàsics per aquesta variable. Comenta.

7. Repeteix l'exercici anterior però generant ara 10.000 observacions de Z_1 en lloc de 100. Comenta els resultats: fixa't amb la mitjana mostral, la variància mostral i l'histograma obtinguts en aquest exercici en comparació a l'obtingut en l'exercici anterior. Han variat en la direcció que esperaves?

8. Crea un fitxer d'instruccions de *Gretl* que generi 1.000 observacions d'una variable aleatòria Z_2 si $Z_2 = Z_1$ i $Z_1 \sim \mathcal{N}(0,1)$. Abans de la commnda que genera la variable, inclou la comanda "`set seed 1234`". Executa el guió. Un cop generades aquestes observacions, utilitzant els menús de la finestra principal:

    (a) Troba l'histograma de densitat amb les observacions de Z_2 de la mostra. Comenta.

    (b) Troba els estadístics descriptius bàsics per aquesta variable. Comenta.

Una variable: aplicació

9. El fitxer `Spain2013.gdt` conté informació sobre la renda anual (`hy022`), en euros, de 12.053 llars espanyoles per 2013.

    (a) Amb l'ajuda de *Gretl*, estima la distribució de la renda a Espanya utilitzant un histograma. Comenta.

    (b) Calcula el estadístics descriptius bàsics per `hy022`. Comenta.

    (c) Creus que seria raonable suposar que la renda es distribueix normalment? Justifica la resposta.

Dues variables: repàs

10. Sabent que la distribució conjunta de les variables aleatòries W and Z és la donada per la taula següent:

	W=0	W=1	W=2
Z=0	0	0.2	0
Z=1	0.3	0.3	0
Z=2	0.1	0	0.1

Calcula $\mathbb{E}[W]$, $\mathbb{E}[Z]$, $\mathbb{E}[Z|W=1]$, $\mathbb{E}[W|Z=1]$.

11. Considera dues variables Z i W, ambdues mesurades en milers d'euros, que només podem prendre 5 parells de valors, tots ells amb la mateixa probabilitat:

Z	W
Z_1	W_1
Z_2	W_2
Z_3	W_3
Z_4	W_4
Z_5	W_5

    (a) Dóna l'expressió que defineix la covariància entre Z i W, $\text{Cov}(Z,W)$, i també l'expressió que defineix el coeficient de correlació entre Z i W, $\text{Corr}(Z,W)$. Presenta les expressions utilitzant sumatoris.

    (b) Quina informació sobre el comportament de Z i W proporciona cadascun d'aquests paràmetres?

(c) Considera que tenim totes les observacions dels parells (Z, W):

Z	W
1	2
2	4
3	6
4	8
5	10

Amb l'ajuda d'un full de càlcul (Excel o similar) calcula $\mathrm{Cov}(Z, W)$ i $\mathrm{Corr}(Z, W)$, definint les fórmules corresponents. *(És a dir, no utilitzis les funciona pre-definides pel full de calcul).*

(d) Considera que canviem les unitats de mesura de la variable Z de milers d'euros a euros:

Z^*	W
1000	2
2000	4
3000	6
4000	8
5000	10

Modifica el full de càlcul abans creat i calcula $\mathrm{Cov}(Z^*, W)$ i $\mathrm{Corr}(Z^*, W)$.

(e) Donat els resultats anteriors, diries que la covariància depèn de les unitats de mesura o és independent a aquests canvis? I el coeficient de correlació simple? Justifica.

(f) Entra les observacions de (Z, W) a *Gretl*. Crea un guió d'instruccions i utilizant les comandes adients, calcula $\mathrm{Cov}(Z, W)$, $\mathrm{Corr}(Z, W)$, $\mathrm{Cov}(Z^*, W)$ i $\mathrm{Corr}(Z^*, W)$. Coincideixen els valors emb els teus càlculs anteriors? Comenta.

12. Demostra que si entre dues variables, X i Y, existeix una relació linial exacta, $Y = a + bX$ amb $b \neq 0$, el coeficient de correlació entre X i Y és igual a 1 si $b > 0$ i igual a -1 si $b < 0$.

13. Sigui $z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ un vector aleatori. Sabent que $\mathbb{E}[Z_1] = 4$, $\mathbb{E}[Z_2] = 8$, $\mathrm{Var}(Z_1) = 100$, $\mathrm{Var}(Z_2) = 49$ i $\mathrm{Cov}(Z_1, Z_2) = -2$, troba $\mathbb{E}[Z]$ i $\mathrm{Var}(z)$.

14. Si definim:

$$z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \right)$$

(a) Què podem dir de $\mathbb{E}[Z]$, $\mathrm{Var}(z)$, $\mathbb{E}[Z_1]$, $\mathrm{Var}(Z_1)$, $\mathbb{E}[Z_2]$, $\mathrm{Var}(Z_2)$, $\mathrm{Cov}(Z_1, Z_2)$, $\mathrm{Cov}(Z_2, Z_1)$, $\mathrm{Corr}(Z_1, Z_2)$?

(b) Què podem dir de la distribució de Z_1? I de Z_2?

15. Definim:

$$z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

com un vector que inclou les variables aleatòries Z_1 i Z_2. Per cadascún dels següents casos, (i) calcula el coeficient de correlació simple entre Z_1 i Z_2, (ii) comenta quin aspecte creus que probablement tindria una mostra d'observacions de (Z_1, Z_2) en un gràfic de punts 2-D.

(a) $\text{Var}(z) = \begin{bmatrix} 9 & 0 \\ 0 & 100 \end{bmatrix}$

(b) $\text{Var}(z) = \begin{bmatrix} 9 & 30 \\ 30 & 100 \end{bmatrix}$

(c) $\text{Var}(z) = \begin{bmatrix} 9 & -18 \\ -18 & 100 \end{bmatrix}$

16. Considera la informació donada en la pregunta (15) sobre les variables Z_1 i Z_2. Per cadascún dels casos considerats:

    (a) Calcula $\text{Var}(Z_1 + Z_2)$.

    (b) Calcula $\text{Var}(Z_1 - Z_2)$.

17. Considera que:

    (a) $\mathbb{E}[Z_1] = 8$, $\mathbb{E}[Z_2] = -3$ i $\mathbb{E}[Z_3] = 5$, $\text{Var}(Z_1) = 8$, $\text{Var}(Z_2) = 4$ i $\text{Var}(Z_3) = 16$, $\text{Cov}(Z_1, Z_2) = 16$, $\text{Cov}(Z_1, Z_3) = -0.2$ i $\text{Cov}(Z_2, Z_3) = 9$,

    Detalla $\mathbb{E}[Z]$ i $\text{Var}(z)$ si $z = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix}$.

18. Considera dues variables aleatòries, Z_1 and Z_2, amb $\mathbb{E}[Z_1] = 4$, $\mathbb{E}[Z_2] = 10$, $\text{Var}(Z_1) = 2$, $\text{Var}(Z_2) = 4$ i $\text{Cov}(Z_1, Z_2) = 0$. Definim dues variables aleatòries: $W_1 = 4Z_1$ i $W_2 = Z_1 + 3Z_2$.

    (a) Aplicant les propietats del valor esperat, variància i covariància troba: (i) $\mathbb{E}[W_1]$, $\mathbb{E}[W_2]$ (ii) $\text{Var}(W_1)$, $\text{Var}(W_2)$ (iii)$\text{Cov}(Z_1, Z_2)$ (iv)tipus de distribució de W_1 i de W_2.

    Volem repetir els càlculs en 18a però utilizant vectors. Definim el vector columna z:

    $$z \equiv \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

    i introduim les següents 2 propietat:

    $$(P1) \quad E(Az) = A\,\mathbb{E}[Z] \qquad (P2) \quad \text{Var}(Az) = Avar(z)A'$$

    on A és una matriu de constants.

    (b) Si $\mathbb{E}[Z] = \mu$ i $\text{Var}(z) = \Sigma$, detalla els elements de μ i Σ.

    (c) Defineix el vector columna w com:

    $$w \equiv \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}.$$

    Després de definir adientment la matriu A, aplica les dues propietats ($P1$) i ($P2$) per repetir els càlculs en 18a.

(d) Aplicant les mateixes dues propietats, troba els valors esperats, variàncies i covariàncies de les següents tres variables:

$$\begin{aligned} W_1 &= Z_1 + 2Z_2 \\ W_2 &= Z_1 - Z_2 \\ W_3 &= Z_1 + Z_2. \end{aligned}$$

Dues variables: Simulació

19. Considera la variable aleatòria $Z_1 \sim N(0,4)$ i definim $Z_2 = Z_1 + 3$.

    (a) Què pots dir de la correlació entre Z_1 i Z_2?

    (b) Escriu un guió d'instruccions de *Gretl* que generi 100 observacions d'aquestes dues variables. Executa el guió. Inclou la instrucció "`set seed 1234`" abans de generar les observacions. Fes un plot amb les observacions generades, en el pla (Z_1, Z_2). Utilitza el menu de *Gretl* per trobar el coeficient de correlació simple d'observacions de Z_1 i Z_2 incloses en la mostra. Comenta.

20. Sabem que $Z_1 \sim N(3,4)$ i $Z_2 = 2 \cdot Z_1 + v$ where $v \sim N(0,9)$.

    (a) Què pots dir de la correlació entre Z_1 i Z_2?

    (b) Amb l'ajuda d'un guió d'instruccions de *Gretl* genera 100 observacions de Z_1 i Z_2. Inclou la instrucció "`set seed 1234`" abans de generar les observacions. Fes un plot amb les observacions generades, en el pla (Z_2, Z_1). Troba el coeficient de correlació simple mostral entre Z_1 i Z_2.

    (c) Si volguessis disminuir la correlació entre les dues variables, quin element en la generació de Z_2 podríem canviar?

21. Considera $Z_1 \sim \mathcal{N}(0,4)$. Defineix una altra variable aleatòria normal Z_2 de forma independent. Amb l'ajuda de *Gretl*, genera 1.000 observacions de Z_1 i 1.000 observacions de Z_2. Inclou la instrucció "`set seed 10101`" abans de generar les observacions.

    (a) Calcula el coeficient de correlació simple mostral entre les dues. Ha de sortit com esperaves?

    (b) Fes un plot de les 1000 observacions en el plà (Z_1, Z_2). Comenta.

Dues variables: aplicació

22. L'article de Sachs & Wagner (2001) a la *European Economic Review* és un dels primers en contrastar empíricament l'anomenada "maledicció dels recursos naturals", que sosté que els països amb abundancia de recursos naturals tendeixen a tenir un creixement econòmic més baix. El fitxer `nr97m.gdt` conté les dades de 71 països utilitzades per aquests autors per contrastar aquesta hipòtesi. Els autors utilitzen el pes de les exportacions del recursos natural sobre el PNB per 1980 (`sxp80`) com a mesura de l'abundància dels recursos naturals d'un país. Com a indicador del creixement econòmic, utilitzen el promig de les taxes de creixement econòmic anuals entre 1970 i 1990 (`gea7090`).

(a) Amb l'ajuda de *Gretl*, calcula el coeficient de correlació simple entre `sxp80` i `gea7090`. Creus què el valor trobat dóna suport a la hipòtesi de la maledicció dels recursos naturals? Comenta.

(b) Fes un gràfic (scatter) de les observacions dels 71 països posant `gea7090` en l'eix vertical i `sxp80` en l'eix horitzontal. Creus què el gràfic dóna suport a la hipòtesi de la maledicció dels recursos naturals? Comenta.