Fonaments de Probabilitat i Estadística.
Notes de classe

Pedro Delicado

6 d'abril de 2018

Índex

Tema 3

Vectors aleatoris

3.1 Variables aleatòries multivariants

Molts experiments aleatoris produeixen més d'una única quantitat rellevant, és a dir, més d'una variable aleatoria. Aquí teniu alguns exemples.

1. En estudis d'economia aplicada podem seleccionar una llar a l'atzar i anotar els seus ingressos X i les seves despeses Y al llarg d'un mes.

2. En estudis mèdics s'estudien conjuntament:

   - el temps que passa entre el primer símptoma d'una malaltia i la seva diagnosi,
   - l'extensió de la malaltia en el moment del diagnòstic,
   - el temps entre el diagnòstic i el pas del malalt a un nou estadi (curació, mort, etc.). En el cas de la mort, aquest temps es coneix com a *temps de supervivència*.

3.1.1 Variables aleatòries bivariants

Definició 3.1.1. *Sigui* $(\Omega, \mathcal{A}, \Pr)$ *un espai de probabilitat. Una funció*

$$\begin{array}{rccl} (X,Y): & \Omega & \longrightarrow & \mathbb{R}^2 \\ & w & \longmapsto & (X,Y)(w) = (X(w), Y(w)) \end{array}$$

tal que

$$(X,Y)^{-1}\left((-\infty, x] \times (-\infty, y]\right) = \{w \in \Omega : X(w) \leq x \ i \ Y(w) \leq y\} \in \mathcal{A}$$

per a tot $x \in \mathbb{R}$ *i tot* $y \in \mathbb{R}$ *és una* VARIABLE ALEATÒRIA BIVARIANT.

Nota: De manera similar al que passa al cas univariant, és suficient que $(X,Y)^{-1}\left((-\infty,x]\times(-\infty,y]\right)\in\mathcal{A}$ per a tot $(x,y)\in\mathbb{R}^2$ per a garantir que

$$(X,Y)^{-1}(B)\in\mathcal{A} \text{ per a tot } B\in\mathcal{B}(\mathbb{R}^2).$$

Per tant, qualsevol variable aleatoria bivariant (X,Y) indueix una mesura de probabilitat en $\mathbb{R}^2$ que permet assignar probabilitats a qualsevol subconjunt *Borelià* B de $\mathbb{R}^2$:

$$P_{(X,Y)}(B)=\Pr\left((X,Y)^{-1}(B)\right)=\Pr\left(\{w\in\Omega:(X(w),Y(w))\in B\}\right).$$

Aquesta mesura de probabilitat induïda s'anomena DISTRIBUCIÓ DE PROBABILITAT CONJUNTA de (X,Y).

Definició 3.1.2. *Donades dues variables aleatòries X,Y definim la* FUNCIÓ DE DISTRIBUCIÓ CONJUNTA F *per a tot $x\in\mathbb{R}$ i $y\in\mathbb{R}$ com*

$$F(x,y)=\Pr\{X\leq x, Y\leq y\}.$$

Nota: La funció de distribució conjunta de (X,Y) identifica completament la distribució conjunta de (X,Y).

Exemple 3.1.1.

En aquest exemple calcularem la probabilitat de que un punt qualsevol de $\mathbb{R}^2$ pertanyi a un rectangle $R=[a,b]\times[c,d]$, usant una funció de distribució conjunta F.

$$\begin{aligned}\Pr\{(X,Y)\in R\} &= \Pr\{a\leq X\leq b, c\leq Y\leq d\}\\ &= \Pr\{a\leq X\leq b, Y\leq d\}-\Pr\{a\leq X\leq b, Y\leq c\}\\ &= \Pr\{X\leq b, Y\leq d\}-\Pr\{X\leq a, Y\leq d\}\\ &\quad -\Pr\{X\leq b, Y\leq c\}+\Pr\{X\leq a, Y\leq c\}\\ &= F(b,d)-F(a,d)-F(b,c)+F(a,c).\end{aligned}$$

□

3.2 Variables aleatòries bivariants discretes i contínues

Variable aleatòria bivariant discreta

Definició 3.2.1. *La variable aleatòria (X,Y) té* DISTRIBUCIÓ CONJUNTA DISCRETA *si conjuntament poden prendre només un nombre finit o numerable* *de valors* (x^i, y^i), $i = 1, 2, \ldots$.

La Figura 3.2 il·lustra la definició de variable aleatoria bivariant discreta. En aquest cas el conjunt imatge de (X, Y) només té 4 elements: (x^1, y^1), (x^2, y^2), (x^3, y^3), (x^4, y^4). Usualment es fa servir una notació diferent. Anomenarem x_1 x_2, x_3, les abscisses dels punts (x^i, y^i), $i = 1, \ldots, 4$, i y_1 y_2, y_3, les seves ordenades. Aleshores

$$(x^1, y^1) = (x_1, y_3),\ (x^2, y^2) = (x_1, y_2),\ (x^3, y^3) = (x_2, y_1),\ (x^4, y^4) = (x_2, y_3).$$

Direm que la variable aleatoria (X, Y) pren els valors (x_i, y_j), $i = 1, 2, 3$, $j = 1, 2, 3$, tot i que hagi alguns valors (x_i, y_j) que no pertanyen al conjunt imatge de (X, Y) (per exemple, (x_3, y_1)). En aquests casos simplement escriurem que la seva probabilitat és 0:

$$\Pr((X, Y) = (x_3, y_1)) = 0.$$

Definició 3.2.2. *Sigui* (X, Y) *una variable aleatòria discreta tal que* x_i, $i = 1, 2, \ldots$ *i* y_j, $j = 1, 2, \ldots$ *són, respectivament les abscisses i les ordenades dels punts del conjunt imatge de* (X, Y). *Es defineix la* FUNCIÓ DE PROBABILITAT *de* (X, Y) *com*

$$p(x_i, y_j) = \Pr\{(X, Y) = (x_i, y_j)\} = \Pr\{X = x_i, Y = y_j\}.$$

Observeu que

$$\Pr\{(X, Y) \in A\} = \sum_{i,j:(x_i, y_j) \in A} p(x_i, y_j).$$

La funció de probabilitat p satisfà dues condicions:

- $p(x_i, y_j) \geq 0$ per a tot x_i, y_j,
- $\sum_i \sum_j p(x_i, y_j) = 1$.

Exemple 3.2.1.

A qualsevol mecanisme aleatori que doni com a resultat una distribució uniforme discreta en $\{1, \ldots, m\}$, $m \geq 1$, l'anomenarem DAU DE m CARES. Siguin (D_1, D_2) els resultats de llançar consecutivament i de forma independent un dau de quatre cares i un altre de tres. És a dir, $D_1 \sim \mathrm{UD}\{1, 2, 3, 4\}$, $D_2 \sim \mathrm{UD}\{1, 2, 3\}$ i D_1 i D_2 són independents. La seva distribució de probabilitat conjunta és

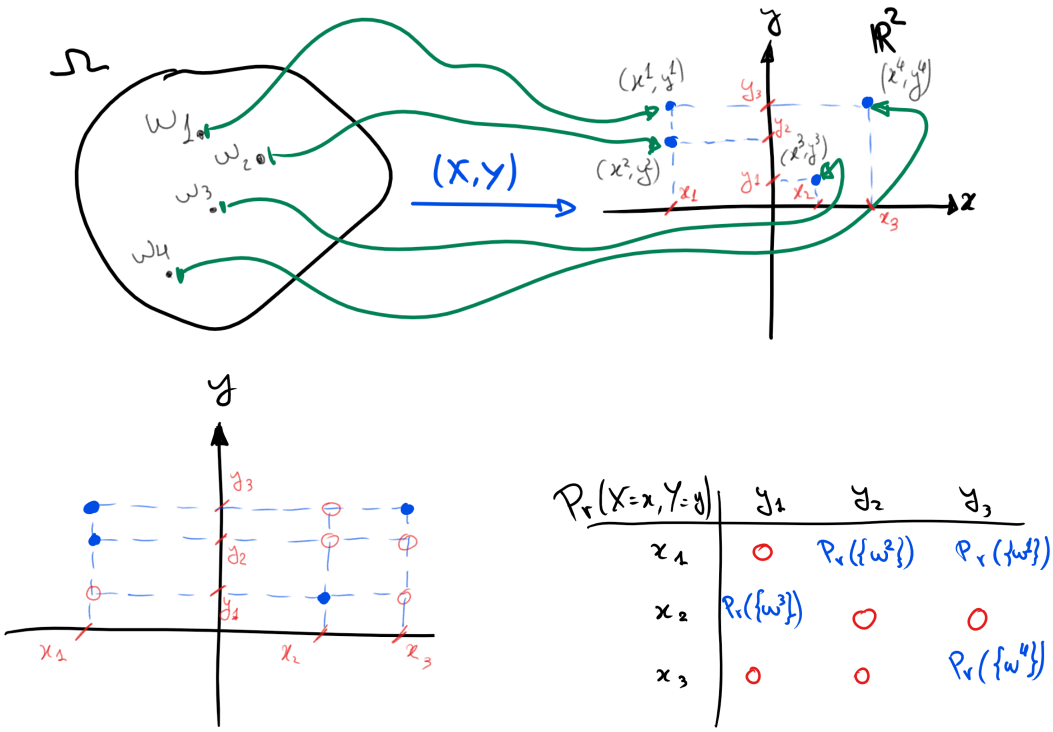


Figura 3.1: Variable aleatòria vibariant discreta. Es suposa que només són 4 els punts de $\mathbb{R}^2$ que són imatge de punts de Ω per la funció (X, Y): $(x^1, y^1) = (x_1, y_3)$, $(x^2, y^2) = (x_1, y_2)$, $(x^3, y^3) = (x_2, y_1)$, $(x^4, y^4) = (x_2, y_3)$.

$\Pr(D_1 = d_1, D_2 = d_2) =$ $\Pr(D_1 = d_1)\Pr(D_2 = d_2)$		d_2		
		1	2	3
	1	1/12	1/12	1/12
d_1	2	1/12	1/12	1/12
	3	1/12	1/12	1/12
	4	1/12	1/12	1/12

Considerem la distribució conjunta de (X, Y) on

$$X = D_1 + D_2, \quad Y = |D_1 - D_2|.$$

La següent taula mostra la distribució conjunta de (X, Y).

$\Pr(X=x, Y=y)$		y = 0	1	2	3
	2	1/12	0	0	0
	3	0	2/12	0	0
x	4	1/12	0	2/12	0
	5	0	2/12	0	1/12
	6	1/12	0	1/12	0
	7	0	1/12	0	0

Considerem ara la distribució conjunta de (Z, Y) on

$$Z = \max\{D_1, D_2\}, \quad Y = |D_1 - D_2|.$$

La següent taula mostra la distribució conjunta de (Z, Y).

$\Pr(Z=z, Y=y)$		y = 0	1	2	3
	1	1/12	0	0	0
	2	1/12	2/12	0	0
z	3	1/12	2/12	2/12	0
	4	0	1/12	1/12	1/12

□

Integrals dobles

Definició 3.2.3. *Sigui $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ una funció no negativa. Sigui D un subconjunt de $\mathbb{R}^2$. El volum sota el gràfic de la funció f restringida a aquest conjunt D (veure Figure 3.2) es denota per*

$$\int\int_D f(x,y)\,dydx \text{ o per } \int\int_D f(x,y)\,dxdy$$

i s'anomena INTEGRAL DOBLE *de f sobre D.*

El càlcul de les integrals dobles és un càlcul iteratiu que implica fer dues integrals simples, una darrera l'altra. És a dir, primer es calcula la integral interna i després l'externa.

> **Nota:** Podeu aprendre més sobre integrals dobles a Nykamp DQ, "Introduction to double integrals." From Math Insight. http://mathinsight.org/double_integral_introduction

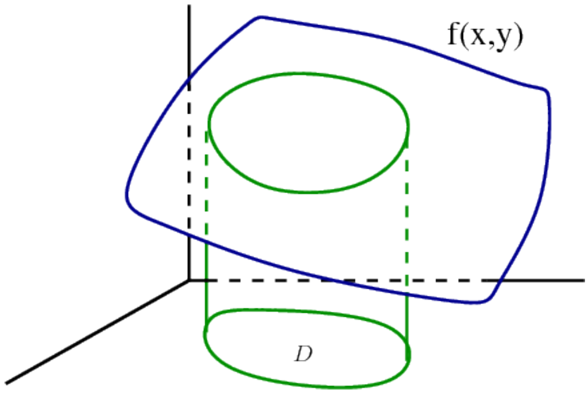


Figura 3.2: Integral doble com a volum. Font: Nykamp DQ, "Double integral as volume under a surface." From Math Insight.
http://mathinsight.org/image/double_integral_volume_under_surface

Exemple 3.2.2.

Calculem el volum que la funció $f(x,y) = xy$ deixa sobre el rectangle $D = [2,4] \times [2,3]$. Per fer-ho calcularem la integral doble

$$\int\int_D xy\ dydx = \int_2^4 \int_2^3 xy\ dydx.$$

Al rectangle $D = [2,4] \times [2,3]$ l'anomenem *recinte d'integració.*

Com acabem de dir, per calcular la integral doble farem dues integrals simples consecutivament:

$$\int_2^4 \int_2^3 xy\ dydx = \int_2^4 \left(\int_2^3 xy\ dy \right)\ dx.$$

Treballem primer la integral interior:

$$\int_2^3 xy\ dy = \left[x\frac{y^2}{2} \right]_2^3 = x\left(\frac{3^2}{2} - \frac{2^2}{2} \right) = \frac{5}{2}x.$$

Ara tenim que

$$\int_2^4 \int_2^3 xy\ dydx = \int_2^4 \left(\int_2^3 xy\ dy \right)\ dx = \int_2^4 \frac{5}{2}x\ dx.$$

Aquesta és una nova integral simple que resolem de la forma habitual:

$$\int_2^4 \int_2^3 xy\ dydx = \int_2^4 \frac{5}{2}x\ dx = \frac{5}{2}\left[\frac{x^2}{2}\right]_2^4 = \frac{5}{2}\left(\frac{4^2}{2} - \frac{2^2}{2}\right) = \frac{5}{2}\frac{12}{2} = \frac{60}{4} = 15.$$

També és possible calcular la integral doble canviant l'ordre d'integració:

$$\int\int_D xy\ dxdy = \int_2^3 \int_2^4 xy\ dxdy = \int_2^3 \left(\int_2^4 xy\ dx\right) dy =$$

$$\int_2^3 y \left[\frac{x^2}{2}\right]_2^4 dy = \int_2^3 y\frac{12}{2}dy = \frac{12}{2}\left[\frac{y^2}{2}\right]_2^3 = \frac{12}{2}\frac{5}{2} = \frac{60}{4} = 15.$$

□

Exemple 3.2.3.

Considerem ara un cas en que el recinte d'integració no és un rectangle. Calculem la integral doble

$$\int\int_A xy\ dydx$$

on el recinte d'integració és el triangle

$$A = \{(x, y) \in \mathbb{R}^2 : 2 \leq x \leq 4, 2 \leq y \leq x\}$$

representat a la Figura 3.3. Farem aquesta integral doble integrant primer en y:

$$\int\int_A xy\ dydx = \int_2^4 \int_2^x xy\ dydx = \int_2^4 \left(\int_2^x xy\ dy\right) dx.$$

Fixeu-vos que la integral interna té el límit superior d'integració en funció de x. Això no planteja cap problema perquè en aquesta integral interna x juga el paper de constant. Així,

$$\int_2^4 \int_2^x xy\ dydx = \int_2^4 \left(\int_2^x xy\ dy\right) dx = \int_2^4 x\left[\frac{y^2}{2}\right]_2^x dx =$$

$$\int_2^4 x\left(\frac{x^2}{2} - \frac{4}{2}\right) dx = \frac{1}{2}\left[\frac{x^4}{4}\right]_2^4 - \left[x^2\right]_2^4 = 18.$$

Observeu en la Figura 3.3 que el conjunt A es pot recórrer de dues maneres i cadascuna d'elles és apropiada per a fer l'integral doble en un ordre d'integració o en l'altre. Quan volem fer la integral doble integrant primer

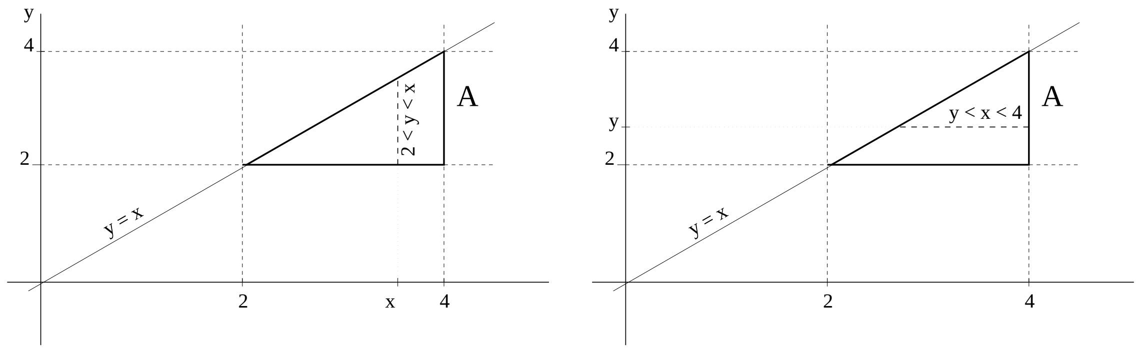


Figura 3.3: Dues maneres de recórrer el recinte d'integració A.

en y i després en x (és a dir, com $\int\int \cdots dydx$), es fixa x i es troba el rang de variació de y dins del recinte d'integració A en funció de x. En el cas contrari, quan integrem primer en x i després en y (escrivim $\int\int \cdots dxdy$), es pren y fix i es troba la variació de x en funció de y:

$$A = \{(x, y) \in \mathbb{R}^2 : 2 \leq y \leq 4, y \leq x \leq 4\}.$$

Aleshores

$$\int\int_A xy\ dxdy = \int_2^4 \int_y^4 xy\ dxdy = \int_2^4 y \left[\frac{x^2}{2}\right]_y^4 dx =$$

$$\int_2^4 y \left(\frac{4^2}{2} - \frac{y}{2}\right) dx = \left[4y^2\right]_2^4 - \frac{1}{2}\left[\frac{y^4}{4}\right]_2^4 = 18.$$

Veiem que el resultat de la integral doble (el volum de la funció xy sobre el rectangle A) no depèn de l'ordre d'integració:

$$\int\int_A xy\ dydx = \int_2^4 \int_2^x xy\ dydx = \int\int_A xy\ dxdy = \int_2^4 \int_y^4 xy\ dxdy = 18.$$

$\square$

Considerem ara un recinte d'integració rectangular, amb costats de longitud infinitesimal: $[x, x + dx] \times [y, y + dy]$. Si la funció $f(\cdot, \cdot)$ és contínua, el volum que deixa sobre aquest rectangle (Figura 3.2) és aproximadament

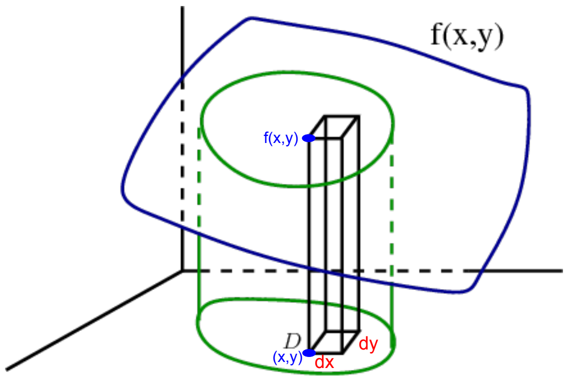


Figura 3.4: Integral doble com a volum, amb la il·lustració del volum sobre un rectangle d'àrea infinitesimal: $\int_x^{x+dx}\int_y^{y+dy} f(u,v)dvdu \approx f(x,y)dxdy$
Font: Nykamp DQ, "Double integral as volume under a surface, with box illustrating Riemann sum." From Math Insight.
http://mathinsight.org/image/double_integral_volume_under_surface_box

igual al volum del prisma rectangular de base $[x, x+dx] \times [y, y+dy]$ i alçada $f(x,y)$:

$$\int_x^{x+dx}\int_y^{y+dy} f(u,v)\, dvdu \approx f(x,y)dxdy,$$

perquè, per la seva continuïtat, la funció $f(\cdot,\cdot)$ és pràcticament constant al llarg d'aquest rectangle $[x, x+dx] \times [y, y+dy]$ d'àrea infinitesimal.

Variable aleatòria bivariant contínua

Definició 3.2.4. *Una variable aleatòria bivariant* (X,Y) *té* DISTRIBUCIÓ CONJUNTA CONTÍNUA *si existeix una funció* f*, definida sobre* $\mathbb{R}^2$*, tal que per a tot subconjunt* A *borelià de* $\mathbb{R}^2$ *es compleix*

$$\Pr\{(X,Y) \in A\} = \int\int_A f(x,y)\, dx\, dy.$$

En aquest cas f *satisfà dues condicions:*

- $f(x,y) \geq 0$ *per a tot* $x, y \in \mathbb{R}$,
- $\int\int_{\mathbb{R}^2} f(x,y)\, dx\, dy = 1.$

La funció f *s'anomena* FUNCIÓ DE DENSITAT CONJUNTA *de* (X,Y)*.*

Figura 3.5: Funció de densitat de la variable aleatòria bivariant (X, Y).

De les propietats de les integrals dobles es deueix que la probabilitat conjunta que (X, Y) assigna al conjunt A és el volum sota el gràfic de la funció de densitat restringida a aquest conjunt A. La Figura 3.5 representa aquesta propietat per un rectangle $A = [a, b] \times [c, d]$.

Algunes conseqüències derivades de la definició de variable aleatòria bivariant contínua són les següents:

1. En notació diferencial tenim:

   $$\Pr\{x \leq X \leq x+dx, y \leq Y \leq y+dy\} = \Pr\{(X,Y) \in (x, x+dx] \times (y, y+dy]\} =$$
   $$\int_x^{x+dx} \int_y^{y+dy} f(u,v)dvdu \approx f(x,y)dxdy.$$

   Aleshores,

   $$f(x,y) \approx \frac{\Pr\{x \leq X \leq x+dx, y \leq Y \leq y+dy\}}{dxdy} \equiv \frac{\text{Probabilitat}}{\text{Àrea}}$$

   i en deduïm que la unitat de mesura de la densitat de probabilitat és probabilitat per unitat d'àrea.

2. Qualsevol punt o successió infinita de punts té probabilitat 0, perquè a sobra d'un conjunt d'àrea cero no es pot encerclar un volum positiu.

3. Qualsevol corba unidimensional té probabilitat 0. En particular $\Pr\{Y = a + bX\} = 0$ i $\Pr\{X^2 + Y^2 = r^2\} = 0$.

Exemple 3.2.4.

Calculeu el valor de k que fa que

$$f(x,y) = \begin{cases} kxy & \text{si } x \in [1,4],\ y \in [2,3], \\ 0 & \text{en altre cas.} \end{cases}$$

sigui una funció de densitat.
També podem escriure

$$f(x,y) = kxyI_{[1,4]\times[2,3]}(x,y) = kxyI_{[1,4]}(x)I_{[2,3]}(y),$$

on fem servir la funció indicadora d'un conjunt:

$$\text{Sigui } A \subseteq \Omega, \text{ per a } w \in \Omega,\ I_A(w) = \begin{cases} 1 & \text{si } w \in A, \\ 0 & \text{si } w \notin A. \end{cases}$$

Com que $f(x,y) \geq 0$ per tot $(x,y) \in \mathbb{R}^2$, f serà una funció de densitat si i només si

$$1 = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)dydx = \int_{\mathbb{R}}\int_{\mathbb{R}} kxyI_{[1,4]}(x)I_{[2,3]}(y)dydx = k\int_1^4\int_2^3 xy\ dydx.$$

Calculem

$$\int_1^4\int_2^3 xy\ dydx.$$

La integral interna és $\int_2^3 xy\ dy$ on la variable és y (fixeu-vos en el dy) i x juga el paper d'una constant. Així,

$$\int_2^3 xy\ dy = x\left[\frac{y^2}{2}\right]_2^3 = x\left[\frac{9}{2} - \frac{4}{2}\right] = \frac{5}{2}x.$$

Ara en la integral exterior x és la variable i tenim que

$$\int_1^4\int_2^3 xy\ dydx = \int_1^4 \frac{5}{2}x\ dx = \frac{5}{2}\left[\frac{x^2}{2}\right]_1^4 = \frac{5}{2}\frac{15}{2} = \frac{75}{4}.$$

Nota: Si la integral doble és finita, el càlcul es pot fer canviant l'ordre d'integració: primer integrem respecte x i després respecte y. És a dir,

$$\int_a^b\int_c^d g(x,y)dydx = \int_c^d\int_a^b g(x,y)dxdy.$$

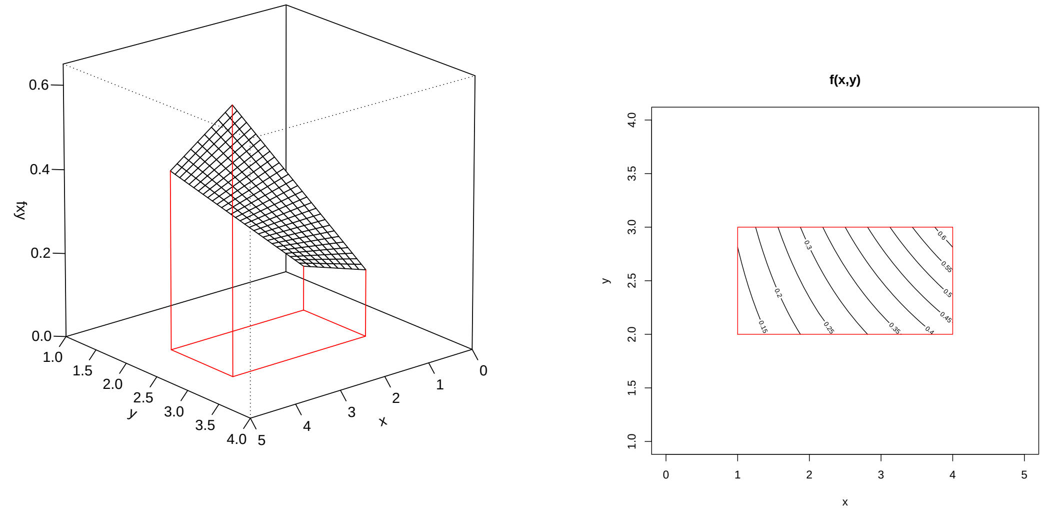


Figura 3.6: Gràfica de la densitat $f(x,y) = \frac{4}{75}xyI_{[1,4]}(x)I_{[2,3]}(y)$. A la dreta, les seves corbes de nivell $f(x,y)$.

Ara recuperem l'equació inicial, on k era l'incògnita:

$$1 = k\int_1^4\int_2^3 xy\,dydx = k\frac{75}{4} \Longrightarrow k = \frac{4}{75}$$

i la funció de densitat és

$$f(x,y) = \frac{4}{75}xyI_{[1,4]}(x)I_{[2,3]}(y).$$

La Figura 3.6 (a l'esquerra) mostra la gràfica d'aquesta funció de densitat. $\square$

Exemple 3.2.5.

Siguin X, Y dues variables aleatòries amb funció de densitat

$$f(x,y) = \begin{cases} c(x^2+y) & \text{si } 0 \le y \le 1-x^2 \\ 0 & \text{altrament} \end{cases}$$

Volem calcular la probabilitat $\Pr(Y \le X+1)$.

Observeu que el conjunt d'$\mathbb{R}^2$ on la densitat pot ser diferent de 0 és

$$A = \{(x,y) \in \mathbb{R}^2 : -1 \le x \le 1,\ 0 \le y \le 1-x^2\},$$

que també es pot expressar com

$$A = \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq 1, -\sqrt{1-y} \leq x \leq \sqrt{1-y}\}.$$

Abans de calcular la probabilitat pròpiament demanada necessitem calcular la constant c. Per fer-ho usarem la propietat que ens diu que la integral d'una funció de densitat sobre el seu domini és 1.

Tenim $-1 \leq X \leq 1$ i $0 \leq Y \leq 1 - X^2 \leq 1$.

$$\begin{aligned}
1 &= \int_{-1}^{1} \int_{0}^{1-x^2} c(x^2 + y)\, dy\, dx = c \int_{-1}^{1} \left[x^2 y + \frac{y^2}{2} \right]_0^{1-x^2} dx \\
&= c \int_{-1}^{1} \left(x^2(1-x^2) + \frac{(1-x^2)^2}{2} \right) dx = c \int_{-1}^{1} \left(\frac{1}{2} - \frac{x^4}{2} \right) dx \\
&= c \left[\frac{x}{2} - \frac{1}{2} \frac{x^5}{5} \right]_{-1}^{1} = c \left(\frac{1}{2} - \frac{1}{10} + \frac{1}{2} - \frac{1}{10} \right) = \frac{8}{10} c
\end{aligned}$$

Per tant $c = \frac{5}{4}$.

Ara ja podem calcular el que volíem:

$$\Pr(Y \leq X + 1) = \Pr((X, Y) \in B) = \int \int_B f(x, y)\, dx\, dy,$$

on $B = \{(x, y) \in A : y \leq x + 1\}$. Observeu que

$$\begin{aligned}
\int \int_B f(x, y)\, dx\, dy &= \int_{-1}^{0} \int_{0}^{x+1} f(x, y)\, dy\, dx + \int_{0}^{1} \int_{0}^{1-x^2} f(x, y)\, dy\, dx \\
&= \int_{-1}^{0} c \left[x^2 y + \frac{y^2}{2} \right]_0^{x+1} dx + \int_0^1 c \left(\frac{1}{2} - \frac{x^4}{2} \right) dx = \\
&= c \int_{-1}^{0} (x^3 + 3/2x^2 + x + 1/2)\, dx + c \int_0^1 \left(\frac{1}{2} - \frac{x^4}{2} \right) dx \\
&= c \left[\frac{x^4}{4} + \frac{x^3}{2} + \frac{x^2}{2} + \frac{x}{2} \right]_{-1}^{0} + c \left[\frac{x}{2} - \frac{x^5}{10} \right]_0^1 \\
&= c \left(-\frac{1}{4} + \frac{1}{2} - \frac{1}{2} + \frac{1}{2} + \frac{1}{2} - \frac{1}{10} \right) = \frac{5}{4} \cdot \frac{13}{20} = \frac{13}{16}.
\end{aligned}$$

Observeu que també podem calcular la probabilitat demanada d'una altra manera:

$$\Pr(Y \leq X + 1) = 1 - \Pr(Y > X + 1) = 1 - \int_{-1}^{0} \int_{x+1}^{1-x^2} f(x, y)\, dy\, dx.$$

D'aquesta manera només hem de calcular una integral doble. □

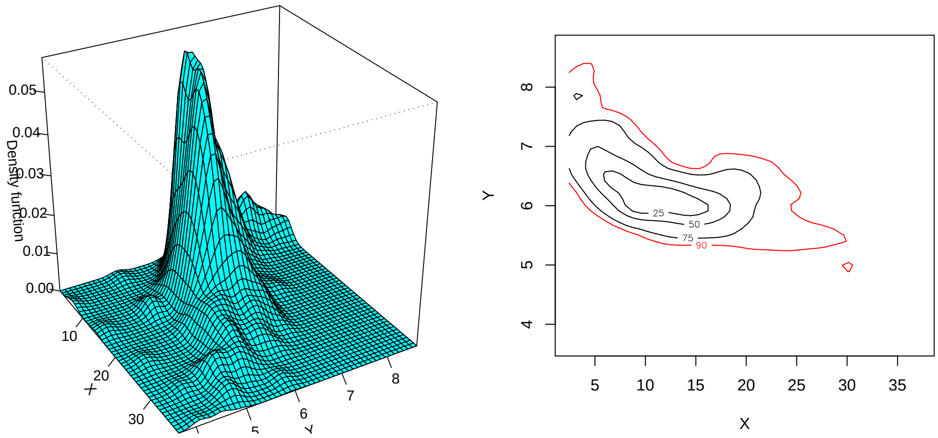


Figura 3.7: Densitat conjunta de dues variables (X, Y) i les sever corbes de nivell.

Definició 3.2.5. *Els conjunts de punts de $\mathbb{R}^2$ pels que la funció de densitat pren el mateix valor s'anomenen* CORBES DE NIVELL DE LA FUNCIÓ DE DENSITAT*:*

$$\Lambda_C = \{(x, y) \in \mathbb{R}^2 : f(x, y) = C\}$$

Hi ha una corba de nivell per a cada $C \in [\min_{(x,y)} f(x, y), \max_{(x,y)} f(x, y)]$.

A vegades el dibuix d'unes quantes corbes de nivell d'una funció de densitat és la millor manera de representar la distribució de probabilitat associada. Tingueu en compte que les corbes de nivell es poden dibuixar en el pla i que per dibuixar la funció de densitat hem de fer-ho a $\mathbb{R}^3$.

Exemple 3.2.6.
La Figura 3.7 mostra una funció de densitat conjunta de dues variables (X, Y) i les seves corbes de nivell. Els nombres 25, 50, 75 i 90 indiquen que la corba de nivell corresponent encercla, respectivament, el 25%, 50%, 75% i 90% de la probabilitat total.
$\square$

Exemple 3.2.4, pàgina 87. Continuació. La Figura 3.6 (a la dreta) mostra les corbes de nivell de la funció de densitat $f(x, y) = (4/75)xyI_{[1,4]}(x)I_{[2,3]}(y)$. Observeu que

$$f(x, y) = C \iff \frac{4}{75}xy = C \iff y = \frac{75C}{4}\frac{1}{x}.$$

Aleshores les corbes de nivell són fragments de les gràfiques de les funcions

$$y = g_C(x) = \frac{75C}{4}\frac{1}{x}.$$

El següent resultat relaciona les funcions de densitat $f(x, y)$ i de distribució $F(x, y)$ d'una variable aleatoria bivariant contínua. Es tracta d'una conseqüència del Teorema Fonamental del Càlcul per funcions de varies variables.

Proposició 3.2.6. *Sigui (X, Y) una variable aleatoria bivariant contínua amb funció de densitat conjunta $f(x, y)$ i funció de distribució $F(x, y)$, llavors*

$$F(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(u, v)\, du\, dv, \qquad f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

A continuació presentem una de les distribucions bivariants usuals: la distribució uniforme.

Definició 3.2.7. *Direm que el parell de variables aleatòries (X, Y) és* UNIFORME *o està* DISTRIBUÏT UNIFORMEMENT *sobre la regió $R \subset \mathbb{R}^2$ si per qualsevol $A \subseteq R$*

$$\Pr\{(X, Y) \in A\} = \frac{\grave{a}rea(A)}{\grave{a}rea(R)}.$$

Observeu que si $A \subseteq R$ aleshores

$$\Pr((X, Y) \in A) = \frac{1}{\text{àrea}(R)} \int\int_A 1 dydx = \int\int_A \frac{1}{\text{àrea}(R)} dydx.$$

Per altra banda, si A no està totalment contingut en R, la seva probabilitat és

$$\Pr((X, Y) \in A) = \Pr((X, Y) \in A \cap R) =$$

$$\int\int_{A \cap R} \frac{1}{\text{àrea}(R)} dydx = \int\int_A \frac{1}{\text{àrea}(R)} I_R(x, y) dydx.$$

Per tant la funció de densitat conjunta de $(X, Y) \sim U(R)$ és

$$f(x, y) = \frac{1}{\text{àrea}(R)} I_R(x, y).$$

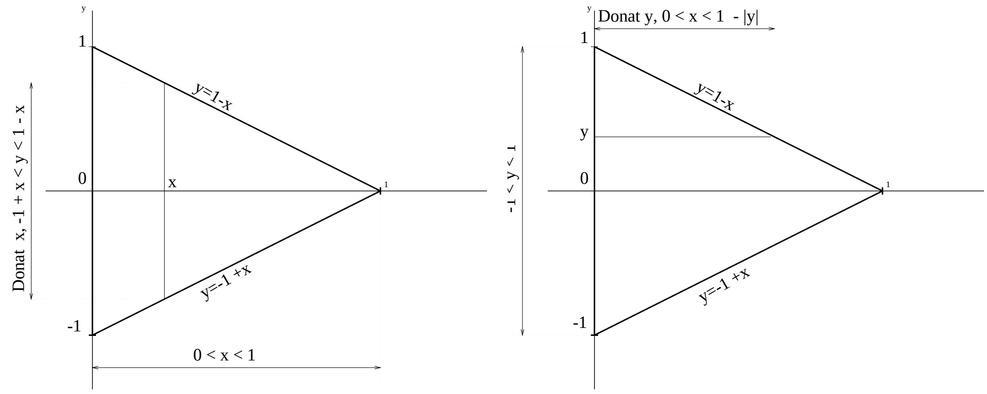


Figura 3.8: Triangle A.

Exemple 3.2.7.
Sigui la variable aleatoria bidimensional (X, Y) amb distribució uniforme en el triangle de la Figura 3.8, $A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, -1+x \leq y \leq 1-x\}$.

La funció de densitat serà

$$f(x,y) = \begin{cases} k & \text{si } (x,y) \in A \\ 0 & \text{en cas contrari} \end{cases}$$

Volem trobar el valor de k:

$$1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)dydx = \int \int_{\{(x,y)\in A\}} k\ dydx = \int_0^1 \int_{-1+x}^{1-x} k\ dydx =$$

$$= \int_0^1 [ky]_{-1+x}^{1-x} dx = \int_0^1 2k(1-x)dx = 2k[x-x^2/2]_0^1 = 2k(1-1/2) = k \Longrightarrow k = 1.$$

□

Donada la variable aleatòria bivariant (X, Y) i una funció (*mesurable*) $g\colon \mathbb{R}^2 \to \mathbb{R}$, es té que $Z = g(X, Y)$ és també una variable aleatoria univariant. El següent resultat (que no provarem) estableix que l'esperança de Z (quan existeix) es pot calcular sense haver de calcular prèviament la funció de probabilitat (o de densitat) de Z.

Proposició 3.2.8. *L'esperança d'una funció de* (X, Y)*,* $Z = g(X, Y)$*, que* *sigui variable aleatoria univariant es pot calcular com*

$$\mathbb{E}(g(X,Y)) = \begin{cases} \sum_i \sum_j g(x_i, y_j) p_{ij} & \text{si}(X,Y) \ \text{és discreta} \\ & \quad \text{amb funció de probabilitat} \\ & \quad \Pr(X = x_i, Y = y_j) = p_{ij} \\ \\ \int_{\mathbb{R}} \int_{\mathbb{R}} g(x,y) f(x,y) & \text{si}(X,Y) \ \text{és contínua} \\ & \quad \text{amb densitat } f(x,y) \end{cases}$$

Exemple 3.2.8.

Una barra d'un metre es trenca aleatòriament per dos punts en tres peces. Quina és la llargada mitjana de la peça del mig?

Per estudiar el problema considerem dues variables aleatòries $(U_1, U_2) \sim \mathrm{U}([0,1] \times [0,1])$.

La longitud de la peça central ve donada per la variable aleatòria $X = |\mathrm{U}_2 - \mathrm{U}_1|$. Llavors

$$\mathbb{E}(|\mathrm{U}_1 - \mathrm{U}_2|) = \int_0^1 \int_0^1 |u_1 - u_2| \, du_1 \, du_2$$

$$= \int_0^1 \int_0^1 1_{(u_1 \geq u_2)} (u_1 - u_2) \, du_1 \, du_2 + \int_0^1 \int_0^1 1_{(u_1 \leq u_2)} (u_2 - u_1) \, du_1 \, du_2$$

$$= \int_0^1 (\int_0^{u_1} (u_1 - u_2) \, du_2) \, du_1 + \int_0^1 (\int_0^{u_2} (u_2 - u_1) \, du_1) \, du_2$$

$$= \int_0^1 [u_1 u_2 - \frac{u_2^2}{2}]_0^{u_1} \, du_1 + \int_0^1 [u_2 u_1 - \frac{u_1^2}{2}]_0^{u_2} \, du_2 = \int_0^1 (u_1^2 - \frac{u_1^2}{2}) \, du_1 + \int_0^1 (u_2^2 - \frac{u_2^2}{2}) \, du_2$$

$$= \int_0^1 (\frac{u_1^2}{2}) \, du_1 + \int_0^1 (\frac{u_2^2}{2}) \, du_2 = 2 \int_0^1 \frac{u_1^2}{2} \, du_1 = \frac{1}{3}.$$

De fet observeu que

$$|\mathrm{U}_1 - \mathrm{U}_2| = \max(\mathrm{U}_1, \mathrm{U}_2) - \min(\mathrm{U}_1, \mathrm{U}_2).$$

Es pot provar que

$$\mathbb{E}(\min(\mathrm{U}_1, \mathrm{U}_2)) = 1/3, \ \mathbb{E}(\max(\mathrm{U}_1, \mathrm{U}_2)) = 2/3$$

i per tant (aplicant la linealitat de l'esparança, que porvarem a la Proposició 3.3.1)

$$\mathbb{E}(|\mathrm{U}_1 - \mathrm{U}_2|) = \mathbb{E}(\max(\mathrm{U}_1, \mathrm{U}_2)) - \mathbb{E}(\min(\mathrm{U}_1, \mathrm{U}_2)) = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

Veiem que, en efecte, $\mathbb{E}(\min(\mathrm{U}_1, \mathrm{U}_2)) = 1/3$ i $\mathbb{E}(\max(\mathrm{U}_1, \mathrm{U}_2)) = 2/3$. Calculem les funcions de distribució de $U_{(1)} = \min(\mathrm{U}_1, \mathrm{U}_2)$ i $U_{(2)} = \max(\mathrm{U}_1, \mathrm{U}_2)$ per $x \in [0, 1]$:

$$F_{U_{(1)}}(x) = 1 - (1 - F_{U_1}(x))^2 = 1 - (1 - x)^2 = 2x - x^2,$$

$$F_{U_{(2)}}(x) = (F_{U_1}(x))^2 = x^2.$$

Per tant les seves densitats són

$$f_{U_{(1)}}(x) = (2 - 2x)I_{[0,1]}(x), \quad f_{U_{(2)}}(x) = 2xI_{[0,1]}(x),$$

i les seves esperances

$$\mathbb{E}(U_{(1)}) = \int_0^1 x f_{U_{(1)}}(x)dx = \int_0^1 (2x - 2x^2)dx = \left(x^2 - \frac{2x^3}{3}\right)\Big|_0^1 = 1 - \frac{2}{3} = \frac{1}{3},$$

$$\mathbb{E}(U_{(2)}) = \int_0^1 x f_{U_{(2)}}(x)dx = \int_0^1 2x^2 dx = \frac{2x^3}{3}\Big|_0^1 = \frac{2}{3}.$$

□

3.3 Marginals i condicionades. Independència de variables aleatorias

Un cop coneixem la funció de distribució conjunta F de dues variables aleatòries (X, Y), ens pot interessar conèixer les seves funcions de distribució per separat. Són les DISTRIBUCIONS MARGINALS.

Distribucions marginals. Cas discret. Sigui (X, Y) una variable aleatòria bivariant discreta amb funció de probabilitat conjunta $\Pr(X = x_i, Y = y_j)$, $i = 1, 2, \ldots$, $j = 1, 2, \ldots$ Aleshores les FUNCIONS DE PROBABILITAT MARGINALS són

$$\Pr(X = x_i) = \sum_j \Pr(X = x_i, Y = y_j), \quad \Pr(Y = y_j) = \sum_i \Pr(X = x_i, Y = y_j).$$

Exemple 3.2.1, pàgina 79. Continuació. Considerem una altra vegada dos daus, un de 4 cares i l'altre de 3. Siguin (D_1, D_2) els resultats de llançar-los consecutivament. Considerem ara la distribució conjunta de (X, Y) on

$$X = D_1 + D_2, \quad Y = |D_1 - D_2|.$$

La següent taula mostra la distribució conjunta de (X, Y). Als *marges* de la taula es mostren les distribucions *marginals* (d'aquí ve el nom), obtingudes sumant per files i per columnes les probabilitats conjuntes.

		y			
$\Pr(X = x, Y = y)$	0	1	2	3	$\Pr(X = x)$
2	1/12	0	0	0	1/12
3	0	2/12	0	0	2/12
x 4	1/12	0	2/12	0	3/12
5	0	2/12	0	1/12	3/12
6	1/12	0	1/12	0	2/12
7	0	1/12	0	0	1/12
$\Pr(Y = y)$	3/12	5/12	3/12	1/12	1

Considerem ara la distribució conjunta de (Z, Y) on

$$Z = \max\{D_1, D_2\}, \quad Y = |D_1 - D_2|.$$

La següent taula mostra la distribució conjunta de (Z, Y) i les marginals.

		y			
$\Pr(Z = z, Y = y)$	0	1	2	3	$\Pr(Z = z)$
1	1/12	0	0	0	1/12
2	1/12	2/12	0	0	3/12
z 3	1/12	2/12	2/12	0	5/12
4	0	1/12	1/12	1/12	3/12
$\Pr(Y = y)$	3/12	5/12	3/12	1/12	1

Distribucions marginals. Cas continu.

Pel cas continu farem servir la següent notació:

$$F_X(x) = Prob(X \leq x) = \lim_{y \longrightarrow \infty} F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(s, y) dy ds,$$

$$F_Y(y) = Prob(Y \leq y) = \lim_{x \longrightarrow \infty} F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{\infty} f(x, t) dx dt.$$

Aleshores tenim que les densitats marginals són

$$f_X(x) = F_X'(x) = \int_{-\infty}^{\infty} f(x, y)\, dy,$$

$$f_Y(y) = F_Y'(y) = \int_{-\infty}^{\infty} f(x, y)\, dx.$$

Exemple 3.3.1.
Considerem la funció de densitat conjunta de dues variables aleatòries:

$$f(x,y) = cx^2y, \text{ amb } x^2 \le y \le 1.$$

Volem calcular les funcions de densitat marginals. Abans però, calculem la constant c per a poder determinar la funció de densitat donada:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx = \int_{-1}^{1}\int_{x^2}^{1} cx^2y\,dy\,dx = \frac{4}{21}c.$$

Per tant $c = \frac{21}{4}$.

Si $0 \le y \le 1$ llavors

$$f_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} cx^2y\,dx = \frac{7}{2}y^{\frac{5}{2}}.$$

Si $-1 \le x \le 1$ llavors

$$f_X(x) = \int_{x^2}^{1} cx^2y\,dy = \frac{21}{8}x^2(1-x^4).$$

Aquesta funció de densitat és simètrica respecte al 0, és bimodal i té les seves modes a $x = \pm 3^{-1/4} \approx \pm 0.76$ (basta buscar els ceros de la seva derivada). $\square$

Exemple 3.3.2.
Seguidament calculem les distribucions marginals d'alguns parells (X,Y) de variables aleatòries a partir de la seva densitat conjunta.

- Sigui $f(x,y) = c(x^2+y), 0 \le y \le 1-x^2$. A l'exemple 3.2.5 havíem trobat que $c = 5/4$.

$$\begin{aligned} f_Y(y) &= \int_{-\sqrt{1-y}}^{\sqrt{1-y}} f(x,y)\,dx \\ &= c\left[\frac{x^3}{3} + yx\right]_{-\sqrt{1-y}}^{\sqrt{1-y}} \\ &= ((2/3)(1-y)^{3/2} + 2y(1-y)^{1/2})c, \text{ per a } 0 \le y \le 1. \end{aligned}$$

$$\begin{aligned} f_X(x) &= \int_0^{1-x^2} f(x,y)\,dy \\ &= c\left[x^2y + \frac{y^2}{2}\right]_0^{1-x^2} \\ &= \left(x^2(1-x^2) + \frac{(1-x^2)^2}{2}\right)c, \text{ per a } -1 \le x \le 1. \end{aligned}$$

- Sigui $f(x,y) = c, (x,y) \in R = \{(x,y) | -1 \le x \le 1, 0 \le y \le 1-x^2\}$. Calculem c:

$$1 = \int_{-1}^{1}\int_0^{1-x^2} c\,dy\,dx = \int_{-1}^{1}(1-x^2)\,dx$$

$$= c\left[x - \frac{x^3}{3}\right]_{-1}^{1} = c(1-\frac{1}{3}) - c\left((-1) - \frac{(-1)^3}{3}\right) = 4/3c.$$

Per tant $c = 3/4$.

$$f_X(x) = \frac{3}{4}(1-x^2), \text{ per a } -1 \le x \le 1,$$

$$f_Y(y) = \frac{3}{2}\sqrt{1-y}, \text{ per a } 0 \le y \le 1.$$

□

Si coneixem la llei conjunta podem obtenir les marginals (com en els casos dels exemples anteriors), però el recíproc no sempre és cert. Per exemple, les dos taules de probabilitats conjuntes següents comparteixen les distribucions marginals:

$\Pr(X=x,Y=y)$		$y=1$	$y=2$	$\Pr(X=x)$
x	1	1/2	0	1/2
	2	0	1/2	1/2
$\Pr(Y=y)$		1/2	1/2	1

$\Pr(X=x,Y=y)$		$y=1$	$y=2$	$\Pr(X=x)$
x	1	1/4	1/4	1/2
	2	1/4	1/4	1/2
$\Pr(Y=y)$		1/2	1/2	1

Ara ja podem fer la demostració d'un resultat que segurament ja hem fet servir abans: l'esperança de la suma és la suma de les esperances. Ho provarem pel cas continu (el cas discret és anàleg) i per a la suma de dues variables aleatorias (el cas de la suma de n variables aleatorias es demostra per inducció en n). Aquest resultat també és cert quan la distribució conjunta de (X,Y) no és i discreta ni contínua (la demostració, per això, és diferent).

Proposició 3.3.1. *Sigui* (X,Y) *una variable aleatoria bivariant amb* $\mathbb{E}(|X|)<\infty$ *i* $\mathbb{E}(|Y|)<\infty$*, aleshores* $\mathbb{E}(X+Y)=\mathbb{E}(X)+\mathbb{E}(Y)$*.*

Demostració:

$$\mathbb{E}(X+Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\,dydx$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\ f(x,y)\,dydx+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y\ f(x,y)\,dxdy$$

$$=\int_{-\infty}^{\infty}x\left(\int_{-\infty}^{\infty}f(x,y)\,dy\right)dx+\int_{-\infty}^{\infty}y\left(\int_{-\infty}^{\infty}f(x,y)\,dx\right)dy$$

$$=\int_{-\infty}^{\infty}x\ f_X(x)\,dx+\int_{-\infty}^{\infty}y\ f_Y(y)\,dy=\mathbb{E}(X)+\mathbb{E}(Y).$$

□

3.3.1 Independència de variables aleatòries.

Definició 3.3.2. *Dues variables aleatòries* X,Y *són* INDEPENDENTS *si*

$$\Pr(X\in A,Y\in B)=\Pr(X\in A)\cdot\Pr(Y\in B),$$

per qualssevol A*,* B*, subconjunts borelians de* $\mathbb{R}$*.*

Teorema 3.3.3. *Donades dues variables aleatòries* X,Y *aleshores les següents afirmacions són equivalents:*

(a) X *i* Y *són independents.*

(b) $F(x,y)=F_X(x)F_Y(y)$ *per qualssevol* x*,* y*.*

(c) $f(x,y) = f_X(x)f_Y(y)$ *per qualssevol* x, y, *on* f, f_X *i* f_Y *són funcions de probabilitat o de densitat, segons si la distribució de* (X,Y) *és discreta o contínua, respectivament.*

Demostració: La farem només pel cas continu.
(a) $\Longrightarrow$ *(b):* Prenent $A = (-\infty, x]$ i $B = (-\infty, y]$.
(b) $\Longrightarrow$ *(c):* Derivant respecte de x i de y.
(c) $\Longrightarrow$ *(a):* $\Pr(X \in A, Y \in B) = \int_A \int_B f_X(x)f_Y(y)dydx =$
$\int_A f_X(x)dx \int_B f_Y(y)dy = \Pr(X \in A)\Pr(Y \in B)$. □

Una conseqüència important d'aquest teorema és que quan hi ha independència entre X i Y podem calcular les funcions de probabilitat, o de densitat, conjuntes a partir de les marginals.

El següent resultat estableix que, si dues variables aleatorias són independents, l'esperança del seu producte és el producte de les seves esperances.

Proposició 3.3.4. *Sigui* X *i* Y *són variables aleatorias independents amb* $\mathbb{E}(|X|) < \infty$, $\mathbb{E}(|Y|) < \infty$ *i* $\mathbb{E}(|XY|) < \infty$, *aleshores* $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Demostració: Farem la demostració pel cas en que X i Y tinguin funció de densitat conjunta.

$$\mathbb{E}(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dydx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\,dxdy$$

$$= \int_{-\infty}^{\infty} x\ f_X(x)\,dx \int_{-\infty}^{\infty} y\ f_Y(y)\,dy = \mathbb{E}(X)\mathbb{E}(Y).$$

□

3.3.2 Distribucions condicionades

Siguin X i Y dues variables aleatòries amb densitat conjunta $f(x,y)$ (o funció de probabilitat conjunta $p(x_i,y_j)$) i marginals $f_X(x)$ i $f_Y(y)$ (o $p_X(x_i)$ i $p_Y(y_j)$).

Suposem primer que (X,Y) és discreta i que hem observat $Y = y$, llavors necessàriament $p_Y(y) > 0$. Coneixent aquesta informació volem especificar les probabilitats pels possibles valors de X.

Definició 3.3.5. *Siguin* X *i* Y *dues variables aleatòries amb distribució conjunta discreta, amb funció de probabilitat conjunta* $p(x_i,y_j) = \Pr(X = x_i, Y = y_j)$ *i marginals* $p_X(x_i) = \Pr(X = x_i)$ *i* $p_Y(y_j) = \Pr(Y = y_j)$.

Suposem que hem observat $Y = y_j$. Llavors es defineix la FUNCIÓ DE PROBABILITAT CONDICIONAL *de X donat $Y = y_j$ com*

$$p_{X|Y=y_j}(x_i) = \Pr(X = x_i | Y = y_j) = \frac{\Pr(X = x_i, Y = y_j)}{\Pr(Y = y_j)} = \frac{p(x_i, y_j)}{p_Y(y_j)}.$$

Observeu que, fixat y_j, $p_{X|Y=y_j}$ és una funció de probabilitat,

$$\sum_i p_{X|Y=y_j}(x_i) = \sum_i \frac{p(x_i, y_j)}{p_Y(y_j)} = \frac{\sum_i p(x_i, y_j)}{p_Y(y_j)} =$$

$$\frac{\sum_i \Pr(X = x_i, Y = y_j)}{\Pr(Y = y_j)} = \frac{\Pr(Y = y_j)}{\Pr(Y = y_j)} = 1.$$

Exemple 3.2.1, pàgina 79. Continuació. Considerem una altra vegada dos daus, un de 4 cares i l'altre de 3. Siguin (D_1, D_2) els resultats de llançar-los consecutivament. Considerem ara la distribució conjunta de (X, Y) on

$$X = D_1 + D_2, \quad Y = |D_1 - D_2|.$$

La seva distribució conjunta de (X, Y) i les marginals estan recollides a la següent taula:

		y			
$\Pr(X = x, Y = y)$	0	1	2	3	$\Pr(X = x)$
2	1/12	0	0	0	1/12
3	0	2/12	0	0	2/12
x 4	1/12	0	2/12	0	3/12
5	0	2/12	0	1/12	3/12
6	1/12	0	1/12	0	2/12
7	0	1/12	0	0	1/12
$\Pr(Y = y)$	3/12	5/12	3/12	1/12	1

Calculem les distribucions condicionades de X pels diferents valors de Y:

x_i	$\Pr(X = x_i \mid Y = 0)$
2	1/3
3	0
4	1/3
5	0
6	1/3
7	0

x_i	$\Pr(X = x_i \mid Y = 1)$
2	0
3	2/5
4	0
5	2/5
6	0
7	1/5

x_i	$\Pr(X = x_i \mid Y = 2)$
2	0
3	0
4	2/3
5	0
6	1/3
7	0

x_i	$\Pr(X = x_i \mid Y = 3)$
2	0
3	0
4	0
5	1
6	0
7	0

Calculem les distribucions condicionades de Y pels diferents valors de X:

y_j	0	1	2	3
$\Pr(Y = y_j \mid X = 2)$	1	0	0	0

y_j	0	1	2	3
$\Pr(Y = y_j \mid X = 3)$	0	1	0	0

y_j	0	1	2	3
$\Pr(Y = y_j \mid X = 4)$	1/3	0	2/3	0

y_j	0	1	2	3
$\Pr(Y = y_j \mid X = 5)$	0	2/3	0	1/3

y_j	0	1	2	3
$\Pr(Y = y_j \mid X = 6)$	1/2	0	1/2	0

y_j	0	1	2	3
$\Pr(Y = y_j \mid X = 7)$	0	1	0	0

Recordeu que vam definir $Z = \max\{D_1, D_2\}$. Calculeu ara, com a exercici, les distribucions condicionades següents:
Z donat que $X = x_i$, per $x_i = 2, \ldots, 7$,
Z donat que $Y = y_j$, per $y_j = 0, \ldots, 3$,
X donat que $Z = z_k$, per $z_k = 1, \ldots, 4$,
Y donat que $Z = z_k$, per $z_k = 1, \ldots, 4$.

Considerem ara el cas en que (X, Y) té distribució conjunta contínua.

Definició 3.3.6. *Siguin X i Y dues variables aleatòries amb densitat conjunta f i marginals f_X i f_Y. Suposem que $f_Y(y) > 0$. Llavors es defineix la* DENSITAT CONDICIONAL *de X donat $Y = y$ com*

$$f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)}, \text{ per a } -\infty \le x \le \infty.$$

Fixem-nos que, fixat y, $f_{X|Y=y}$ és una densitat, és a dir, integra 1 (donat que és no negativa):

$$\int_{-\infty}^{\infty} f_{X|Y=y}(x)\, dx = \int_{-\infty}^{\infty} \frac{f(x,y)}{f_Y(y)} dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f(x,y) dx = \frac{1}{f_Y(y)} f_Y(y) = 1.$$

Suposem que interpretem que $f(x, y)$ representa la versemblança relativa pel punt (x, y). Si és sabut que $Y = y_0$ llavors el punt (x, y) està sobre la recta $Y = y_0$ i la versemblança relativa de qualsevol punt en aquesta recta és $f(x, y_0)$. Per tant $f_{X|Y=y_0}(x)$ ha de ser proporcional a $f(x, y_0)$. Tenim doncs que essencialment $f(x, y_0)$ i $f_{X|Y=y_0}(x)$ representen el mateix però $f_{X|Y=y_0}(x)$ inclou el factor $1/f_Y(y_0)$, que és necessari per a què sigui densitat.

Una altra justificació de l'expressió de la densitat condicionada és la següent. Tot i que no podem condicionar per un esdeveniment amb probabilitat 0, sí que podem pensar que la densitat condicional de X donat que $Y = y$ hauria de verificar que

$$f_{X|Y=y}(x)dx \approx \Pr(X \in [x, x+dx]|Y = y) \approx \Pr(X \in [x, x+dx]|Y \in [y, y+dy]) =$$

$$\frac{\Pr(X \in [x, x+dx], Y \in [y, y+dy])}{\Pr(Y \in [y, y+dy])} \approx \frac{f(x,y)dxdy}{f_Y(y)dy} = \frac{f(x,y)}{f_Y(y)}dx.$$

Per tant, la densitat condicional ha de ser

$$f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)}.$$

Anàlogament es defineix

$$f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)}, \text{ per a } -\infty \leq y \leq \infty.$$

En resum, tenim les relacions:

$$f_X(x)f_{Y|X=x}(y) = f(x,y) = f_Y(y)f_{X|Y=y}(x).$$

Exemple 3.3.3.

Siguin X, Y dues variables aleatòries amb funció de densitat conjunta:

$$f(x,y) = \frac{3}{16}(4 - 2x - y), \text{ si } x, y > 0 \text{ i } 2x + y < 4.$$

Volem determinar la probabilitat $\Pr(Y \geq 2|X = 0.5)$.

Determinem primer la distribució marginal de X:

$$f_X(x) = \int_0^{4-2x} \frac{3}{16}(4 - 2x - y)\, dy = \frac{3}{16}\left[4y - 2xy - \frac{y^2}{2}\right]_0^{4-2x}$$

$$= \frac{3}{16}\left((4 - 2x)^2 - \frac{(4-2x)^2}{2}\right) = \frac{3}{16}\frac{(4-2x)^2}{2} = \frac{3(2-x)^2}{8}.$$

Calculem $f_{Y|X=x}(y)$ per a $0 < y < 4 - 2x$:

$$f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)} = \frac{\frac{3}{16}(4-2x-y)}{\frac{3}{8}(2-x)^2}.$$

Per tant:

$$f_{Y|X=x}(y) = \begin{cases} \frac{1}{2}\frac{(4-2x-y)}{(2-x)^2} & \text{si } 0 \le y \le 4-2x \\ 0 & \text{altrament} \end{cases}$$

Ara ja podem calcular el que volíem:

$$\Pr(Y \ge 2|X = 0.5) = \int_2^3 \frac{1}{2}\frac{(4-2(0.5)-y)}{(2-0.5)^2}\,dy = \int_2^3 \frac{(3-y)}{2(1.5)^2}\,dy$$

$$= \frac{2}{9}\int_2^3 (3-y)\,dy = \frac{2}{9}\left[3y - \frac{y^2}{2}\right]_2^3 = \frac{2}{9}(9 - \frac{9}{2} - 6 + \frac{4}{2}) = 1/9.$$

□

Construcció d'una distribució conjunta

Volem construir una distribució conjunta coneixent la condicionada. Definirem la distribució a partir de la relació

$$f(x,y) = \begin{cases} f_{X|Y=y}(x) f_Y(y) & \text{si } f_Y(y) > 0 \\ 0 & \text{si } f_Y(y) = 0 \end{cases}$$

Però si $f_Y(y_0) = 0$ per algun y_0, pot suposar-se que $f(x, y_0) = 0$ per tot x i per tant la relació anterior la tenim per tot x i per tot y.

Exemple 3.3.4.

Escollim un punt X a l'atzar de l'interval $(0,1)$ i un cop hem observat x escollim un punt Y a l'atzar a $(x,1)$. Ens preguntem quina és la densitat marginal de Y.

$$f_X(x) = 1, \text{ per a } 0 < x < 1.$$

$$f_{Y|X=x}(y) = \frac{1}{1-x}, \text{ per a } x < y < 1.$$

$$f(x,y) = \frac{1}{1-x}, \text{ per a } 0 < x < 1, x < y < 1.$$

$$f_Y(y) = \int_0^y \frac{1}{1-x}\,dx = -\log(1-y), \text{ per a } 0 < y < 1.$$

□

Exemple 3.3.5.
Es llença un dau de 4 cares. Sigui X el resultat obtingut. Ara, si $X = k$ agafem un dau de k cares, el llencem i anomenem Y al seu resultat. Aquesta situació produeix les següents distribucions condicionades:

y_j	1	2	3	4
$\Pr(Y = y_j \mid X = 1)$	1	0	0	0

y_j	1	2	3	4
$\Pr(Y = y_j \mid X = 2)$	1/2	1/2	0	0

y_j	1	2	3	4
$\Pr(Y = y_j \mid X = 3)$	1/3	1/3	1/3	0

y_j	1	2	3	4
$\Pr(Y = y_j \mid X = 4)$	1/4	1/4	1/4	1/4

Recordem que X és uniforme discreta a $\{1, 2, 3, 4\}$. A partir d'aquí podem calcular la distribució conjunta i les marginals:

		y_j			
$\Pr(X = x_i, Y = y_j)$	1	2	3	4	$\Pr(X = x_i)$
1	1/4	0	0	0	1/4
x_i 2	1/8	1/8	0	0	1/4
3	1/12	1/12	1/12	0	1/4
4	1/16	1/16	1/16	1/16	1/4
$\Pr(Y = y_j)$	25/48	13/48	7/48	3/48	1

□

Distribucions multivariades

Definició 3.3.7. *Sigui* $(\Omega, \mathcal{A}, \Pr)$ *un espai de probabilitat. Una funció p-dimensional* $\mathbf{X} = (X_1, \ldots, X_p)$,

$$\begin{array}{rccl} \mathbf{X} = (X_1, \ldots, X_p): & \Omega & \longrightarrow & \mathbb{R}^p \\ & w & \longmapsto & \mathbf{X}(w) = (X_1(w), \ldots, X_p(w)) \end{array}$$

tal que $\{w \in \Omega : X_i(w) \leq x_i,\ i = 1, \ldots, p\} \in \mathcal{A}$ *per a tot* $\mathbf{x} = (x_1, \ldots, x_p) \in \mathbb{R}^p$*, és una* VARIABLE ALEATÒRIA MULTIVARIANT *de dimensió p, també anomenat* VECTOR ALEATORI.

Definició 3.3.8. *Un vector aleatori* $\mathbf{X} = (X_1, \ldots, X_p)$ *té una* DISTRIBUCIÓ CONJUNTA DISCRETA *si només pot prendre un nombre finit o numerable de valors* $\mathbf{x} = (x_1, \ldots, x_p) \in \mathbb{R}^p$.

Definició 3.3.9. *Un vector aleatori* $\mathbf{X} = (X_1, \ldots, X_p)$ *té* DISTRIBUCIÓ CONTÍNUA *si existeix una densitat conjunta* $f : \mathbb{R}^p \longrightarrow \mathbb{R}$, *funció no negativa, tal que per tot* $A \subseteq \mathbb{R}^p$

$$\Pr\{(X_1, \ldots, X_p) \in A\} = \int \cdots \int_A f(x_1, \ldots, x_p)\, dx_1, \ldots dx_p$$

En aquest cas es diu que f *és la funció de densitat de* $\mathbf{X}$.

Definició 3.3.10. *Sigui* $\mathbf{X}$ *un vector aleatori amb distribució contínua. Aleshores es defineix la* FUNCIÓ DE DISTRIBUCIÓ CONJUNTA $F : \mathbb{R}^p \longrightarrow [0,1]$ *com*

$$F(x_1, \ldots, x_p) = \Pr\{X_1 \leq x_1, \ldots, X_p \leq x_p\}.$$

De forma anàloga a com ho vam fer per dimensió 2, es pot prova que si la derivada p-èsima de F existeix aleshores

$$f(x_1, \ldots, x_p) = \frac{\partial^p F(x_1, \ldots, x_p)}{\partial x_1 \cdots \partial x_p}.$$

Les MARGINALS DE QUALSEVOL ORDRE es defineixen anàlogament al cas bivariant. Així

$$f_i(x) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-1} f(x_1 \ldots x_{i-1}, x, x_{i+1} \ldots x_p)\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_p.$$

$$f_{ij}(x,y) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-2} f(x_1 \ldots x \ldots y \ldots x_p)\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_{j-1}\, dx_{j+1} \ldots dx_p.$$

La funció de distribució marginal es defineix de forma anàloga:

Definició 3.3.11. *Sigui* $\mathbf{X}$ *un vector aleatori amb funció de distribució conjunta* $F(x_1, \ldots, x_n)$, *aleshores es defineix la* FUNCIÓ DE DISTRIBUCIÓ MARGINAL *com*

$$F_i(x_i) = \lim_{\substack{x_j \longrightarrow \infty \\ j \neq i}} F(x_1, x_2, \ldots, x_p).$$

Exemple 3.3.6.

Considerem la funció de densitat conjunta

$$f(x_1, x_2, x_3) = \begin{cases} c(x_1 + 2x_2 + 3x_3) & \text{si } 0 \leq x_i \leq 1 \text{ per } i = 1,2,3 \\ 0 & \text{altrament.} \end{cases}$$

Calcularem algunes distribucions conjuntes i marginals a partir d'aquesta.

(a) Calculem la constant c:

$$\frac{1}{c} = \int_0^1 \int_0^1 \int_0^1 \frac{1}{c} f(x_1, x_2, x_3)\, dx_1\, dx_2\, dx_3$$

$$= \int_0^1 \int_0^1 \int_0^1 (x_1+2x_2+3x_3)\, dx_1\, dx_2\, dx_3 = \int_0^1 \int_0^1 \left[\frac{x_1^2}{2} + 2x_2x_1 + 3x_3x_1\right]_0^1 dx_2\, dx_3$$

$$= \int_0^1 \int_0^1 (1/2 + 2x_2 + 3x_3)\, dx_2\, dx_3 = \int_0^1 \left[(x_2/2 + x_2^2 + 3x_3x_2)\right]_0^1 dx_3$$

$$= \int_0^1 (1/2 + 1 + 3x_3)\, dx_3 = \left(\frac{3}{2}x_3 + \frac{3x_3^2}{2}\right)\bigg|_0^1 = 3.$$

Per tant $c = 1/3$.

(b)

$$f_{12}(x_1, x_2) = \begin{cases} \int_0^1 \frac{1}{3}(x_1 + 2x_2 + 3x_3)\, dx_3 = \frac{1}{3}(x_1 + 2x_2 + \frac{3}{2}) & \text{si } 0 \le x_i \le 1 \\ & \text{per } i = 1, 2, 3 \\ 0 & \text{altrament.} \end{cases}$$

(c)

$$F_{12}(x_1, x_2) = \Pr(X_1 \le x_1, X_2 \le x_2)$$

$$= \int_0^{x_1} \int_0^{x_2} 1/3(s + 2t + 3/2)\, ds\, dt = \int_0^{x_1} \left[1/3(st + 2\frac{t^2}{2} + \frac{3}{2}t)\right]_0^{x_2} ds$$

$$= \int_0^{x_1} 1/3 \left(sx_2 + x_2^2 + \frac{3}{2}x_2\right) ds = 1/3 \left(\frac{x_1^2 x_2}{2} + x_1x_2^2 + \frac{3}{2}x_1x_2\right).$$

(d)

$$F_1(x_1) = \lim_{x_2 \longrightarrow \infty} F_{12}(x_1, x_2) = F_{12}(x_1, 1) = 1/3 \left(\frac{x_1^2}{2} + x_1 + \frac{3}{2}x_1\right).$$

(e)

$$f_X(x_1) = \frac{\partial}{\partial x_1} F_1(x_1) = 1/3(x_1 + 5/2), \text{ per a } 0 \le x_1 \le 1.$$

Fixem-nos que $\int_0^1 f_X(x_1)\, dx_1 = 1/3 \left[\frac{x_1^2}{2} + \frac{5x_1}{2}\right]_0^1 = (\frac{1}{3})3 = 1$. També

$$f_X(x_1) = \int_0^1 f_{12}(x_1, x_2)\, dx_2 = 1/3 \left[x_1x_2 + 2\frac{x_2^2}{2} + \frac{3}{2}x_2\right]_0^1 = 1/3(x_1+5/2).$$

$\square$

Definició 3.3.12. *Direm que $X_1, X_2, \ldots, X_p$ són variables aleatòries* INDEPENDENTS *si per a tot $A_1, A_2, \ldots, A_p \subseteq \mathbb{R}$ (que siguin borelians) tenim que*

$$\Pr(X_1 \in A_1, \ldots, X_p \in A_p) = \prod_{i=1}^{p} \Pr(X_i \in A_i).$$

Teorema 3.3.13. *Siguin $X_1, X_2, \ldots, X_p$ variables aleatòries. Aleshores les següents afirmacions són equivalents:*

- $X_1, \ldots, X_p$ *són independents.*
- $F(x_1, \ldots, x_p) = \prod_{i=1}^{p} F_i(x_i)$.
- $f(x_1, \ldots, x_p) = \prod_{i=1}^{p} f_i(x_i)$.

Definició 3.3.14. *Direm que $X_1, \ldots, X_n$ són variables aleatòries* INDEPENDENTS I IDÈNTICAMENT DISTRIBUÏDES *(i.i.d.) o que formen una* MOSTRA ALEATÒRIA *d'una població, dessignada per X i amb funció de densitat (o de probabilitat) f, si $X_1, \ldots, X_n$ són variables aleatòries independents i amb la mateixa llei marginal. En aquest cas la funció de densitat (o de probabilitat) conjunta és*

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i).$$

Ara demostrarem que l'esperança del producte de variables aleatorias independents és el producte de les esperances (un resultat que ja hem provat per a 2 variables).

Proposició 3.3.15. *Si $X_1, \ldots, X_n$ són independents, aleshores $\mathbb{E}(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} \mathbb{E}(X_i)$.*

Demostració: Farem la demostració pel cas de distribució conjunta contínua.

$$\begin{aligned}
\mathbb{E}(\prod_{i=1}^n X_i) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\prod_{i=1}^n x_i) f(x_1 \dots x_n)\, dx_1 \dots dx_n \\
&= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\prod_{i=1}^n x_i)(\prod_{i=1}^n f_i(x_i))\, dx_1 \dots dx_n \\
&= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\prod_{i=1}^n x_i\ f_i(x_i))\, dx_1 \dots dx_n \\
&= \prod_{i=1}^n \int_{-\infty}^{\infty} x_i\ f_i(x_i)\, dx_i \\
&= \prod_{i=1}^n \mathbb{E}(X_i).
\end{aligned}$$

□

Distribucions condicionals multivariades

Siguin $X_1, \dots, X_p$ variables aleatòries amb funció de densitat conjunta $f(x_1, \dots, x_p)$. Denotem per $f_{(i)}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$ la funció de densitat marginal de $X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_p$.

Exemple 3.3.7.

- $f_{(1)}(x_2, \dots, x_p)$ és la funció de densitat marginal de $X_2, \dots, X_p$.
- $f_{(2)}(x_1, x_3, \dots, x_p)$ és la f.d. marginal de $X_1, X_3, \dots, X_p$.

⋮

- $f_{(p)}(x_1, \dots, x_{p-1})$ és la f.d. marginal de $X_1, X_2, \dots, X_{p-1}$.

□

Fixem un subíndex i qualsevol, per exemple $i = 1$. Llavors tenim la següent definició:

Definició 3.3.16. *Per qualsevol conjunt de valors $(x_2, \dots, x_p)$ tal que $f_{(1)}(x_2, \dots, x_p) > 0$, la funció de densitat condicional de X_1 a $X_2 = x_2, \dots, X_p = x_p$, ve donada per*

$$f_{X_1|X_2=x_2,\dots,X_p=x_p}(x_1) = \frac{f(x_1, x_2, \dots, x_p)}{f_{(1)}(x_2, \dots, x_p)}.$$

Si generalitzem la definició, en notació vectorial, tenim:

Sigui $\mathbf{X} = (X_1, X_2, \ldots, X_p) = (\mathbf{Y}, \mathbf{Z})$, amb $\mathbf{Y} = (X_1, \ldots, X_k)$ i $\mathbf{Z} = (X_{k+1}, \ldots, X_p)$. Llavors, per qualsevol punt $\mathbf{z} \in \mathbb{R}^{p-k}$ tal que $f_{\mathbf{Z}}(\mathbf{z}) > 0$, definim

$$f_{\mathbf{Y}|\mathbf{Z}=\mathbf{z}}(\mathbf{y}) = \frac{f(\mathbf{y}, \mathbf{z})}{f_{\mathbf{Z}}(\mathbf{z})}, \mathbf{y} \in \mathbb{R}^k.$$

Direm que $\mathbf{Y}$ i $\mathbf{Z}$ són vectors aleatoris independents si

$$f_{(\mathbf{Y},\mathbf{Z})}(\mathbf{y}, \mathbf{z}) = f_{\mathbf{Y}}(\mathbf{y}) f_{\mathbf{Z}}(\mathbf{z})$$

per a qualssevol $\mathbf{y} \in \mathbb{R}^k$ i $\mathbf{z} \in \mathbb{R}^{p-k}$.

Definicions anàlogues a aquestes es donen també per a vectors aleatoris discrets.

Exemple 3.3.8.

Sigui X_1 una variable aleatòria exponencial de paràmetre $\alpha = 1$.

Donat $X_1 = x_1$, definim dues variables aleatòries independents X_2, X_3 amb funció de densitat condicional $\gamma(\alpha = 1, \beta = x_1)$, és a dir:

- $f_{X_2|X_1=x_1}(x_2) = x_1 e^{-x_1 x_2}$, si $x_2 > 0$. Per tant, $(X_2|X_1 = x_1) \sim \mathrm{Exp}(x_1)$.
- $f_{X_3|X_1=x_1}(x_3) = x_1 e^{-x_1 x_3}$, si $x_3 > 0$. Per tant, $(X_3|X_1 = x_1) \sim \mathrm{Exp}(x_1)$.

Volem calcular:

(a) La densitat marginal conjunta de (X_2, X_3).
Fem primer uns càlculs previs.

(a.1)
$$f_{(X_2,X_3)|X_1=x_1}(x_2, x_3) = x_1 e^{-x_1 x_2} x_1 e^{-x_1 x_3} = x_1^2 e^{-x_1(x_2+x_3)},$$
amb $x_2, x_3 > 0$, on hem usat que $X_2|X_1 = x_1$ és independent de $X_3|X_1 = x_1$.

(a.2)
$$f(x_1, x_2, x_3) = f_{(X_2,X_3)|X_1=x_1}(x_2, x_3) f_1(x_1) =$$
$$x_1^2 e^{-x_1(x_2+x_3)} e^{-x_1} = x_1^2 e^{-x_1(1+x_2+x_3)},$$
amb $x_1, x_2, x_3 > 0$.

(a.3) Ara ja podem calcular la densitat marginal conjunta:

$$\begin{aligned}
f_{(X_2,X_3)}(x_2,x_3) &= \int_0^\infty x_1^2 e^{-x_1(1+x_2+x_3)}\,dx_1 \\
&\left\{\begin{array}{ll} u = x_1^2 & du = 2x_1 dx_1 \\ dv = e^{-x_1(1+x_2+x_3)} & v = \frac{-1}{1+x_2+x_3} e^{-x_1(1+x_2+x_3)} \end{array}\right\} \\
&= \underbrace{x_1^2 \frac{-e^{-x_1(1+x_2+x_3)}}{1+x_2+x_3}\Big|_0^\infty}_{0,\text{ fent límits}} - \int_0^\infty \frac{-e^{-x_1(1+x_2+x_3)}}{1+x_2+x_3} 2x_1\,dx_1 \\
&= \int_0^\infty \frac{2x_1}{1+x_2+x_3} e^{-x_1(1+x_2+x_3)}\,dx_1 \\
&= \frac{2}{(1+x_2+x_3)^3} \underbrace{\int_0^\infty \frac{(1+x_2+x_3)^2}{1} x_1 e^{-x_1(1+x_2+x_3)}\,dx_1}_{1} \\
&= \frac{2}{(1+x_2+x_3)^3},
\end{aligned}$$

i l'última integral val 1 ja que es tracta de l'expressió de la densitat d'una $\gamma(2, 1+x_2+x_3)$.

(b) La probabilitat $\Pr(X_2 + X_3 < 4)$.

$$\Pr(X_2+X_3<4) = \int_0^4 \int_0^{4-x_3} \frac{2}{(1+x_2+x_3)^3}\,dx_2\,dx_3 = \int_0^4 \left[\frac{-1}{(1+x_2+x_3)^2}\right]_0^{4-x_3} dx_3$$

$$= \int_0^4 \left(\frac{-1}{5^2} + \frac{1}{(1+x_3)^2}\right) dx_3 = \left[\frac{-1}{1+x_3}\right]_0^4 - \left[\frac{x_3}{25}\right]_0^4 = 16/25.$$

(c) La probabilitat $\Pr(X_1 \le 1 | X_2 = 1, X_3 = 4)$.

Tenim que $\Pr(X_1 \le 1 | X_2 = 1, X_3 = 4) = \int_0^1 f_{X_1|X_2=1,X_3=4}(x_1)\,dx_1$.

$$f_{X_1|X_2=x_2,X_3=x_3}(x_1) = \frac{f(x_1,x_2,x_3)}{f_{(X_2,X_3)}(x_2,x_3)}$$

$$= \frac{x_1^2 e^{-x_1(1+x_2+x_3)}}{\frac{2}{(1+x_2+x_3)^3}} = \frac{1}{2} x_1^2 (1+x_2+x_3)^3 e^{-x_1(1+x_2+x_3)}, \text{ amb } x_1 > 0.$$

Fixem-nos que $(X_1 | X_2 = x_2, X_3 = x_3) \sim \gamma(3, 1+x_2+x_3)$.

Ara ja podem acabar el càlcul:

$$\Pr(X_1 \le 1 | X_2 = 1, X_3 = 4) = \int_0^1 \frac{1}{2} x_1^2 (1+1+4)^3 e^{-x_1(1+1+4)}\,dx_1$$

$$= \int_0^1 \frac{6^3 x_1^2}{2} e^{-6x_1}\, dx_1 = 0.938,$$

on l'última integral l'hem resolt integrant per parts.

$\square$

Estadístics extrems

Siguin $X_1, X_2, \ldots, X_n$ variables aleatòries independents i idènticament distribuïdes amb funció de distribució F i de densitat f.

Considerem $U = \text{màx}\{X_1, \ldots, X_n\} = X_{(n)}$ i $V = \text{mín}\{X_1, \ldots, X_n\} = X_{(1)}$. En general $X_{(i)}$ indica la i-èsima variable aleatòria ordenada.

- **Llei de $X_{(n)}$:**

$$F_{X_{(n)}}(u) = \Pr(X_{(n)} \le u) = \Pr(X_1 \le u, \ldots, X_n \le u) = [F(u)]^n.$$

$$f_{X_{(n)}}(u) = nf(u)[F(u)]^{n-1}.$$

- **Llei de $X_{(1)}$:**

$$F_{X_{(1)}}(v) = 1-\Pr(X_{(1)} \le v) = 1-\Pr(X_1 \le v, \ldots, X_n \le v) = 1-[1-F(v)]^n.$$

$$f_{X_{(1)}}(v) = nf(v)[1-F(v)]^{n-1}.$$

Exemple 3.3.9.

Siguin $T_1, \ldots, T_n$ variables aleatòries independents i idènticament distribuïdes segons una $\text{Exp}(\lambda)$ que mesuren el temps de vida de n components d'un sistema. Distingirem dos casos:

(a) El sistema conté aquestes n components en sèrie. La funció de densitat del temps de funcionament del sistema equival a la funció de densitat del mínim de les $T_1, \ldots, T_n$, llavors és

$$f_{X_{(1)}}(v) = nf(v)[1-F(v)]^{n-1} = n\lambda e^{-\lambda v}(e^{-\lambda v})^{n-1} = n\lambda e^{-n\lambda v}.$$

Per tant, el temps de funcionament del sistema és $\text{Exp}(n\lambda)$.

(b) El sistema conté les n components en paral·lel. La funció de densitat del temps de funcionament del sistema equival a la funció de densitat del màxim de les $T_1, \ldots, T_n$, és a dir:

$$f_{X_{(n)}}(u) = nf(u)[F(u)]^{n-1} = n\lambda e^{-\lambda u}(1-e^{-\lambda u})^{n-1}.$$

$\square$

Els estadístics d'ordre són $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$. Si n és senar, és a dir: $n = 2m + 1$, aleshores $X_{(m)}$ és la mediana de les X_i, amb $i = 1, \ldots, n$.

Generalitzant el que hem fet fins ara tenim el següent resultat:

Teorema 3.3.17. *La funció de densitat del k-èsim estadístic d'ordre és*

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} f(x)[F(x)]^{k-1}[1-F(x)]^{n-k}.$$

3.4 Matriu de variàncies i covariàncies. Correlació. Desigualtat de Cauchy-Schwarz

Definició 3.4.1. *Donada la variable aleatòria bivariant (X, Y), es defineix el seu* VECTOR D'ESPERANCES *com un vector de dimensió 2, amb coordenades iguals a les esperances de les coordenades de (X, Y):*

$$\mathbb{E}(X, Y) = \mathbb{E}\left[\begin{pmatrix} X \\ Y \end{pmatrix}\right] = \begin{pmatrix} \mathbb{E}(X) \\ \mathbb{E}(Y) \end{pmatrix}.$$

Sigui (X, Y) amb

$$\mathbb{E}(X) = \mu_x, \mathrm{Var}\,(X) = \sigma_X^2, \quad \mathbb{E}(Y) = \mu_y, \mathrm{Var}\,(Y) = \sigma_Y^2.$$

La següent és una mesures de la dependència lineal entre X i Y.

Definició 3.4.2. *Definim la* COVARIÀNCIA *de X i Y com*

$$Cov(X, Y) = \mathbb{E}\{(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\}.$$

Si $\sigma_X^2 < \infty$ i $\sigma_Y^2 < \infty$ aleshores la covariància existeix i és finita.

La Figura 3.9 il·lustra per què la covariància entre dues variables aleatòries és una bona mesura de la seva relació lineal. Pel cas de (X, Y) continua, aquesta figura mostra la contribució de cada quadrant a la integral

$$\int\int_{\mathbb{R}^2} (x - \mu_X)(y - \mu_Y) f(x, y)\, dy\, dx$$

que defineix la covariància entre X i Y. Les rectes verticals i horitzontals passen per $\mu_X = \mathbb{E}(X)$ i per $\mu_Y = \mathbb{E}(Y)$, respectivament. Es contemplen tres possibles tipus de distribucions conjuntes de (X, Y): relació lineal positiva (esquerra), independència (centre) i relació lineal negativa (dreta). En aquests tres casos la covariància entre X i Y serà positiva, nul·la o negativa, respectivament.

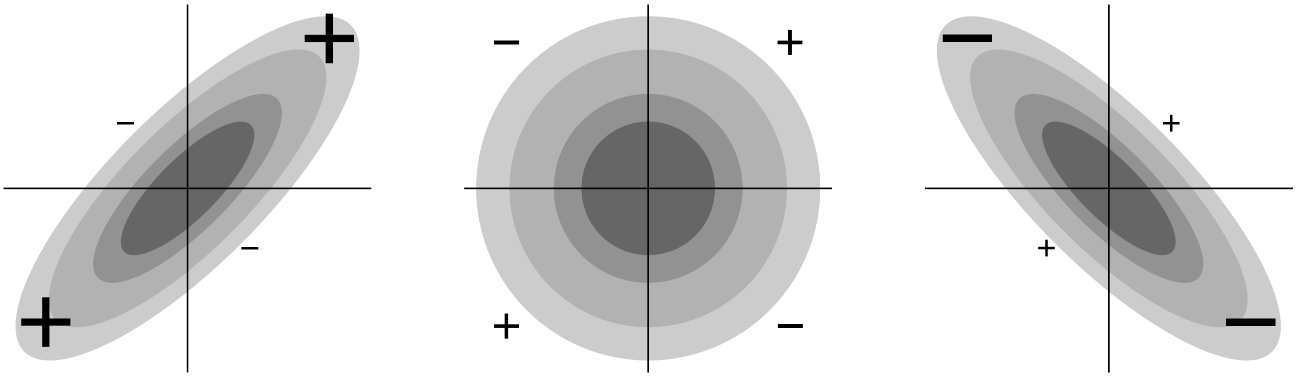


Figura 3.9: Contribució de cada quadrant a la integral $\int\int_{\mathbb{R}^2}(x-\mu_X)(y-\mu_Y)f(x,y)\,dy\,dx$ que defineix la covariància entre X i Y. Tres possibles distribucions conjuntes de (X,Y): relació lineal positiva (esquerra), independència (centre) i relació lineal negativa (dreta).

Proposició 3.4.3. *Siguin X i Y variables aleatòries tals que $\sigma_X^2 < \infty$ i $\sigma_Y^2 < \infty$, llavors*

$$Cov(X,Y) = \mathbb{E}(X\cdot Y) - \mathbb{E}(X)\cdot\mathbb{E}(Y).$$

Demostració:

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}\{(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\} \\
&= \mathbb{E}\{(X-\mathbb{E}(X))Y-(X-\mathbb{E}(X))\mathbb{E}(Y)\} \\
&= \mathbb{E}\{(X-\mathbb{E}(X))Y\}-\mathbb{E}\{(X-\mathbb{E}(X))\mathbb{E}(Y)\} \\
&= \mathbb{E}(XY)-\mathbb{E}\{\mathbb{E}(X)Y\}-\mathbb{E}\{X\mathbb{E}(Y)\}+\mathbb{E}(X)\mathbb{E}(Y)\} \\
&= \mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)-\mathbb{E}(X)\mathbb{E}(Y)+\mathbb{E}(X)\mathbb{E}(Y)\} \\
&= \mathbb{E}(X\cdot Y)-\mathbb{E}(X)\cdot\mathbb{E}(Y).
\end{aligned}$$

□

Proposició 3.4.4. *Siguin (X,Y,Z) variables aleatòries amb variància finita i $a \in R$.*

1. *La covariància és un operador simètric: $Cov(X,Y) = Cov(Y,X)$.*

2. *La variància és un cas particular de covariància: $Var(X) = Cov(X,X)$.*

3. *La covariància amb una constant és 0:* $Cov(X, a) = 0$.

4. *La covariància és un operador bilineal, és a dir, és lineal en els seus dos arguments:*

$$Cov(aX, Y) = Cov(X, aY) = a\,Cov(X, Y).$$

$$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z),$$
$$Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z).$$

5. $Var\,(X + Y) = Var\,(X) + Var\,(Y) + 2Cov(X, Y)$.

6. $Var\,(X - Y) = Var\,(X) + Var\,(Y) - 2Cov(X, Y)$.

Demostració: La demostració dels tres primers apartats és immediata. Provem la linealitat en el primer argument (això, juntament amb la simetria, prova la linealitat en el segon argument):

$$\begin{aligned} \mathrm{Cov}(aX, Y) &= \mathbb{E}[aXY] - \mathbb{E}(aX)\mathbb{E}(Y) \\ &= a\mathbb{E}(XY) - a\mathbb{E}(X)\mathbb{E}(Y) \\ &= a\left(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\right) \\ &= a\mathrm{Cov}(X, Y). \end{aligned}$$

$$\begin{aligned} \mathrm{Cov}(X + Y, Z) &= \mathbb{E}[(X + Y)Z] - \mathbb{E}(X + Y)\mathbb{E}(Z) \\ &= \mathbb{E}(XZ) + \mathbb{E}(YZ) - \{\mathbb{E}(X) + \mathbb{E}(Y)\}\mathbb{E}(Z) \\ &= \mathbb{E}(XZ) + \mathbb{E}(YZ) - \mathbb{E}(X)\mathbb{E}(Z) - \mathbb{E}(Y)\mathbb{E}(Z) \\ &= \mathbb{E}(XZ) - \mathbb{E}(X)\mathbb{E}(Z) + \mathbb{E}(YZ) - \mathbb{E}(Y)\mathbb{E}(Z) \\ &= \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z). \end{aligned}$$

La demostració dels dos últims apartats segueix del tercer, escrivint $\mathrm{Var}\,(X + Y) = \mathrm{Cov}(X + Y, X + Y)$ i $\mathrm{Var}\,(X - Y) = \mathrm{Cov}(X - Y, X - Y)$. □

Una propietat no desitjable de la covariància com a mesura de la força de la relació lineal entre dues variables aleatòries, és el fet que la covariància depèn de las unitats en que mesurem les variables: si canviem les unitats de mesura, canvia el valor de la covariància. Per exemple, siguin les següents variables aleatòries:

- X, altura d'un nen de 10 anys, mesurada en metres.
- X^*, altura d'un nen de 10 anys, mesurada en centímetres: $X^* = 100X$.

- Y, pes d'un nen de 10 anys, mesurada en quilos.

Observeu que

$$\text{Cov}(X^*, Y) = \text{Cov}(100X, Y) = 100\text{Cov}(X, Y).$$

Per corregir aquest defecte s'introdueix a continuació el coeficient de correlació entre dues variables aleatòries.

Definició 3.4.5. *Si* $0 < Var(X) = \sigma_X^2 < \infty$, $0 < Var(Y) = \sigma_Y^2 < \infty$ *i* $Cov(X, Y) = \sigma_{XY}$, *es defineix la* CORRELACIÓ *(o el* coeficient de correlació*) entre* X *i* Y *com*

$$\rho_{XY} = \rho(X, Y) = Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)\, Var(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Observeu que, si $a > 0$ i $b > 0$ són dues constants positives,

$$\text{Corr}(aX, bY) = \frac{\text{Cov}(aX, bY)}{\sqrt{\text{Var}\,(aX)\text{Var}\,(bY)}} = \frac{ab\text{Cov}(X, Y)}{\sqrt{a^2\text{Var}\,(X) b^2 \text{Var}\,(Y)}}$$

$$= \frac{ab\text{Cov}(X, Y)}{ab\sqrt{\text{Var}\,(X)\text{Var}\,(Y)}} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}\,(X)\text{Var}\,(Y)}} = \text{Corr}(X, Y).$$

Per tant el coeficient de correlació no depèn de l'escala en que estan mesurades les variables.

Definició 3.4.6. *Sigui la variable aleatoria bivariant* (X, Y). *Siguin* $\sigma_X^2 = Var(X)$, $\sigma_Y^2 = Var(Y)$, $\rho = Corr(X, Y)$ *i* $Cov(X, Y) = \sigma_{XY} = \rho\sigma_X\sigma_Y$. *La matriu*

$$Var(X, Y) = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$$

s'anomena MATRIU DE VARIÀNCIES-COVARIÀNCIES *de* (X, Y). *La matriu*

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

s'anomena MATRIU DE CORRELACIONS *de* (X, Y).

Desigualtat de Cauchy-Schwarz

Teorema 3.4.7 (Desigualtat de Cauchy-Schwarz)**.** *Siguin* U, V *variables aleatòries amb* $0 < \mathbb{E}(U^2) < \infty$ *i* $0 < \mathbb{E}(V^2) < \infty$. *Llavors*

$$[\mathbb{E}(UV)]^2 \leq \mathbb{E}(U^2)\mathbb{E}(V^2).$$

Es dona la igualtat si i només si

$$\Pr(V = \lambda U) = 1, \ \ \textit{on}\ \lambda = \frac{\mathbb{E}(UV)}{\mathbb{E}(U^2)} \neq 0,$$

o, de forma equivalent, si

$$\Pr(U = \gamma V) = 1, \ \ \textit{on}\ \gamma = \frac{\mathbb{E}(UV)}{\mathbb{E}(V^2)} \neq 0.$$

Demostració: Sigui λ un nombre real qualsevol i sigui $X = V - \lambda U$. Aleshores, per a tot $\lambda \in \mathbb{R}$,

$$0 \leq \mathbb{E}(X^2) = \mathbb{E}\left((V - \lambda U)^2\right) = \mathbb{E}(U^2 - 2\lambda UV + \lambda^2 V^2) = \mathbb{E}(U^2) - 2\lambda\mathbb{E}(UV) + \lambda^2\mathbb{E}(V^2).$$

Per tant, l'equació de segon grau en λ

$$\mathbb{E}(U^2) - 2\lambda\mathbb{E}(UV) + \lambda^2\mathbb{E}(V^2) = 0$$

té, com a molt, una solució real. Donat que les solucions es poden calcular com

$$\lambda = \frac{2\mathbb{E}(UV) \pm \sqrt{4\mathbb{E}(UV)^2 - 4\mathbb{E}(U^2)\mathbb{E}(V^2)}}{2\mathbb{E}(U^2)},$$

com que no poden haver 2 solucions reals, es segueix que el *discriminant* no pot ser positiu:

$$4\mathbb{E}(UV)^2 - 4\mathbb{E}(U^2)\mathbb{E}(V^2) \leq 0 \Leftrightarrow \mathbb{E}(UV)^2 \leq \mathbb{E}(U^2)\mathbb{E}(V^2).$$

Observeu que es dona la igualtat si i només si existeix una única sol·lució de l'equació de segon grau, que val

$$\lambda = \frac{\mathbb{E}(UV)}{\mathbb{E}(U^2)},$$

que no pot ser 0 perquè, en aquest cas, l'equació de segon grau implicaria que $\mathbb{E}(U^2) = 0$, i això contradiria la hipòtesi de que $0 < \mathbb{E}(U^2)$.

En cas d'igualtat, a més a més,

$$0 = \mathbb{E}(X^2) \Leftrightarrow \Pr(X^2 = 0) \Leftrightarrow \Pr(X = 0) \overset{X=V-\lambda U}{\Leftrightarrow} \Pr(V = \lambda U) = 1,$$

on, com hem dit abans, $\lambda = \frac{\mathbb{E}(UV)}{\mathbb{E}(U^2)}$.

Com que $\lambda \neq 0$, podem definir $\gamma = 1/\lambda$, i tenir que

$$\mathbb{E}(UV)^2 = \mathbb{E}(U^2)\mathbb{E}(V^2) \Leftrightarrow \mathbb{E}(X^2) = 0 \Leftrightarrow \Pr(V = \lambda U) = 1 \Leftrightarrow \Pr(U = \gamma V) = 1,$$

on

$$\gamma = \frac{1}{\lambda} = \frac{\mathbb{E}(U^2)}{\mathbb{E}(UV)} = \frac{\mathbb{E}(U^2)\mathbb{E}(UV)}{\mathbb{E}(UV)^2} = \frac{\mathbb{E}(U^2)\mathbb{E}(UV)}{\mathbb{E}(U^2)\mathbb{E}(V^2)} = \frac{\mathbb{E}(UV)}{\mathbb{E}(V^2)}.$$

□

Es pot provar que la desigualtat de Cauchy-Schwarz continua sent vàlida si $\mathbb{E}(U^2) = 0$, o $\mathbb{E}(U^2) = \infty$, o $\mathbb{E}(V^2) = 0$, o $\mathbb{E}(V^2) = \infty$.

Proposició 3.4.8. *Sigui* (X, Y) *variable aleatoria bivariant amb* $\mathbb{E}(X) = \mu_X$, $\mathbb{E}(Y) = \mu_Y$, $0 < Var(X) = \sigma_X^2 < \infty$, $0 < Var(Y) = \sigma_Y^2 < \infty$ *i* $Cov(X, Y) = \sigma_{XY}$. *Sigui*

$$Corr(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

el coeficient de correlació entre X *i* Y. *Aleshores*

$$-1 \leq Corr(X, Y) \leq 1,$$

amb igualtat si i només si

$$Y = a + bX, \text{ per a constants } a \text{ i } b, \text{ amb } b \neq 0,$$

o de forma equivalent, si

$$X = \alpha + \beta Y, \text{ per a constants } \alpha = -a/b \text{ i } \beta = 1/b, \text{ amb } \beta \neq 0.$$

En cas d'igualtat,

$$b > 0 \Leftrightarrow \beta > 0 \Leftrightarrow Corr(X, Y) = 1,$$

$$b < 0 \Leftrightarrow \beta < 0 \Leftrightarrow Corr(X, Y) = -1.$$

Demostració: Prenem $U = X - \mu_X$ i $V = Y - \mu_Y$. La desigualtat de Cauchy-Schwarz ($\mathbb{E}(UV)^2 \leq \mathbb{E}(U^2)\mathbb{E}(V^2)$) ens diu que

$$\mathrm{Cov}^2(X, Y) = \mathbb{E}\left((X - \mu_X)(Y - \mu_Y)\right)^2 \leq \mathbb{E}\left((X - \mu_X)^2\right)\mathbb{E}\left((Y - \mu_Y)^2\right) = \sigma_X^2 \sigma_Y^2.$$

Aleshores

$$\frac{\mathrm{Cov}^2(X, Y)}{\sigma_X^2 \sigma_Y^2} \leq 1,$$

i per tant

$$-1 \leq \mathrm{Corr}(X, Y) \leq 1.$$

Es donarà la igualtat $|\mathrm{Corr}(X, Y)| = 1$ si i només si $\mathbb{E}(UV)^2 = \mathbb{E}(U^2)\mathbb{E}(V^2)$, si i només si

$$\Pr(V = \lambda U) = 1, \text{ on } \lambda = \frac{\mathbb{E}(UV)}{\mathbb{E}(U^2)},$$

si i només si

$$\Pr(Y - \mu_Y = \lambda(X - \mu_X)) = 1, \text{ on } \lambda = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} \neq 0,$$

si i només si

$$\Pr(Y = (\mu_Y - \lambda\mu_X) + \lambda X) = 1.$$

Tenim així que es compleix $\Pr(Y = a + bX) = 1$, on $a = \mu_Y - \lambda\mu_X$ i $b = \lambda \neq 0$.

Veurem ara que l'última afirmació és certa. Observeu que si $Y = a + bX$ amb probabilitat 1,

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))),$$

$$\text{on } Y - \mathbb{E}(Y) = a + bX - (a + b\mathbb{E}(X)) = b(X - \mathbb{E}(X)).$$

Per tant $\mathrm{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}(X))b(X - \mathbb{E}(X))] = b\sigma_X^2$. D'altra banda $\mathrm{Var}(Y) = \mathbb{E}((Y - \mathbb{E}(Y))^2) = b^2\mathrm{Var}(X) = b^2\sigma_X^2$. Finalment:

$$\mathrm{Corr}(X,Y) = \frac{b\sigma_X^2}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

$$= \frac{b\sigma_X^2}{\sigma_X|b|\sigma_X} = \begin{cases} +1 & \text{si i només si } b > 0 \\ -1 & \text{si i només si } b < 0 \end{cases}$$

□

Definició 3.4.9. *Direm que X, Y estan*

- POSITIVAMENT CORRELACIONADES *si Corr$(X,Y) > 0$ (Cov$(X,Y) > 0$).*
- NEGATIVAMENT CORRELACIONADES *si Corr$(X,Y) < 0$.*
- NO CORRELACIONADES *si Corr$(X,Y) = 0$.*

Proposició 3.4.10. *Siguin X i Y variables aleatòries independents amb $0 < \sigma_X^2 < \infty$ i $0 < \sigma_Y^2 < \infty$, llavors*

$$\mathit{Cov}(X,Y) = \mathit{Corr}(X,Y) = 0.$$

Demostració: Si X i Y són independents aleshores $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Això implica que

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y) = 0.$$

□

El recíproc d'aquesta última propietat no sempre és cert perquè el coeficient de correlació detecta bé les relacions lineals entre variables, però no les relacions no lineals. Vegem un exemple de variables no independents amb coeficient de correlació igual a 0.

Exemple 3.4.1.

Siguin X i Y dues variables aleatòries discretes tals que

$$X = \begin{cases} -1 & \text{amb probabilitat } 1/3 \\ 0 & \text{amb probabilitat } 1/3 \\ 1 & \text{amb probabilitat } 1/3 \end{cases}$$

$$\text{i } Y = X^2.$$

Llavors

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X^3) = (-1)^3(1/3) + 0^3(1/3) + 1^3(1/3) = 0,$$

$$\mathbb{E}(X) = 0.$$

Per tant $\mathrm{Cov}(X, Y) = 0$ i X i Y no estan correlacionades.

D'altra banda X i Y són òbviament dependents ja que el valor de Y està totalment determinat pel de X.

□

Proposició 3.4.11. *Siguin X, Y variables aleatòries amb $Var(X) < \infty$ i $Var(Y) < \infty$, i a, b, c nombres reals, llavors*

$$Var(aX + bY + c) = a^2\, Var(X) + b^2\, Var(Y) + 2ab Cov(X, Y).$$

Demostració:

$$\begin{aligned} \mathrm{Var}\,(aX + bY + c) &= \mathrm{Cov}(aX + bY + c, aX + bY + c) \\ &\quad \text{(linealitat de la covariància)} \\ &= a\,\mathrm{Cov}(X, aX + bY + c) + b\,\mathrm{Cov}(Y, aX + bY + c) \\ &\quad +\mathrm{Cov}(c, aX + bY + c) \\ &\quad \text{(linealitat de la covariància i covariància 0 amb constants)} \\ &= a^2\mathrm{Cov}(X, X) + a\,b\,\mathrm{Cov}(X, Y) + b\,a\,\mathrm{Cov}(Y, X) + b^2\mathrm{Cov}(Y, Y) \\ &\quad \text{(simetria de la covariància)} \\ &= a^2\mathrm{Var}\,(X) + b^2\mathrm{Var}\,(Y) + 2ab\mathrm{Cov}(X, Y). \end{aligned}$$

□

Proposició 3.4.12. *Siguin* $X_1, \ldots, X_n$ *variables aleatòries amb* $Var(X_i) < \infty$ *per* $i = 1, \ldots, n$*, llavors*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n}\sum_{i<j} Cov(X_i, X_j).$$

Demostració:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{i=1}^{n} \mathrm{Cov}\left(X_i, \sum_{j=1}^{n} X_j\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}\left(X_i, X_j\right) = \sum_{i=1}^{n} \mathrm{Cov}\left(X_i, X_i\right) + \sum_{i=1}^{n}\sum_{j\neq i} \mathrm{Cov}\left(X_i, X_j\right)$$

$$= \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right) + 2\sum_{i=1}^{n}\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

□

Un cas particular del resultat anterior és el següent.

Proposició 3.4.13. *Siguin* $X_1, \ldots, X_n$ *variables aleatòries amb* $Var(X_i) < \infty$ *per* $i = 1, \ldots, n$*. Si són no correlacionades, llavors*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i).$$

3.4.1 Càlcul de variàncies en format matricial

Les següents definicions i els resultats que les segueixen, permeten operar de forma matricial amb vectors d'esperances i matrius de covariàncies de vectors aleatoris p-dimensionals.

Definició 3.4.14. *Siguin* W_{ij}*,* $i = 1, \ldots, n$*,* $j = 1, \ldots, m$*, variables aleatorias. Es considera la matriu* $n \times m$ $\mathbf{W}$ *que té per element* (i, j) *la variable aleatoria* W_{ij}*. Es diu que* $\mathbf{W}$ *és una* MATRIU ALEATÒRIA*. Es defineix l'esperança de* $\mathbf{W}$ *com la matriu* $n \times m$ $\mathbb{E}(\mathbf{W})$ *que té per element* (i, j) *la* $\mathbb{E}(W_{ij})$*.*

Definició 3.4.15. *Sigui* $\mathbf{X}$ *una variable aleatoria* p*-dimensional amb* $\mathbb{E}(\mathbf{X}) = \mu \in \mathbb{R}^p$*. Es definieix la seva matriu de variàncies (o de variàncies i covariancies) com*

$$Var(\mathbf{X}) = \mathbb{E}((\mathbf{X} - \mu)(\mathbf{X} - \mu)^t)$$

que és una matriu $p \times p$*.*

És fàcil comprovar que pel cas $p = 2$ la definició anterior coincideix amb la Definicó 3.4.6.

Proposició 3.4.16. *Sigui* $\mathbf{X}$ *una variable aleatoria p-dimensional amb* $\mathbb{E}(\mathbf{X}) = \mu \in \mathbb{R}^p$ *i* $Var(\mathbf{X}) = \Sigma \in \mathbb{R}^{p\times p}$. *Sigui la matriu* $A \in \mathbb{R}^{q\times p}$. *Aleshores*

$$\mathbb{E}(A\mathbf{X}) = A\mathbb{E}(\mathbf{X}) = A\mu, \;\; \mathbb{E}(\mathbf{X}A^t) = \mathbb{E}(\mathbf{X})A^t = \mu A^t,$$

$$Var(A\mathbf{X}) = A\,Var(\mathbf{X})A^t = A\Sigma A^t.$$

Demostració: Siguin $a_1, \ldots, a_q$ els vectors fila de A. Aleshores

$$\begin{aligned}
\mathbb{E}(A\mathbf{X}) &= \mathbb{E}\left[\begin{pmatrix} a_1 \\ \vdots \\ a_q \end{pmatrix}\mathbf{X}\right] = \mathbb{E}\left[\begin{pmatrix} a_1\mathbf{X} \\ \vdots \\ a_q\mathbf{X} \end{pmatrix}\right] = \begin{pmatrix} \mathbb{E}(a_1\mathbf{X}) \\ \vdots \\ \mathbb{E}(a_q\mathbf{X}) \end{pmatrix} \\
&= \begin{pmatrix} a_1\mathbb{E}(\mathbf{X}) \\ \vdots \\ a_q\mathbb{E}(\mathbf{X}) \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_q \end{pmatrix}\mathbb{E}(\mathbf{X}) = A\mathbb{E}(\mathbf{X}) = A\mu.
\end{aligned}$$

De la mateixa forma es prova que $\mathbb{E}(\mathbf{X}A^t) = \mathbb{E}(\mathbf{X})A^t = \mu A^t$.

Ara farem servir aquestes expresins matricials de la linealitat de l'esperança per provar la fòrmula per la variància:

$$\begin{aligned}
\mathrm{Var}\,(A\mathbf{X}) &= \mathbb{E}\left[(A\mathbf{X} - \mathbb{E}(A\mathbf{X}))(A\mathbf{X} - \mathbb{E}(A\mathbf{X}))^t\right] \\
&= \mathbb{E}\left[(A\mathbf{X} - A\mu))(A\mathbf{X} - A\mu))^t\right] \\
&= \mathbb{E}\left[A(\mathbf{X} - \mu)(A(\mathbf{X} - \mu))^t\right] \\
&= \mathbb{E}\left[A(\mathbf{X} - \mu)(\mathbf{X} - \mu)^t A^t\right] \\
&= A\mathbb{E}\left[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^t A^t\right] \\
&= A\mathbb{E}\left[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^t\right] A^t \\
&= A\mathrm{Var}\,(\mathbf{X})A^t = A\Sigma A^t.
\end{aligned}$$

□

3.5 L'esperança condicionada com a variable aleatòria

Siguin X, Y dues variables aleatòries amb funció de densitat conjunta $f(x, y)$. Denotem per $f_X(x)$ i $f_Y(x)$ les funcions de densitat (o de probabilitat) marginals corresponents a X i a Y.

Considerem per a tot x tal que $f_X(x) > 0$ la funció de densitat (o de probabilitat) condicionada $f_{Y|X=x}(y)$ de la variable aleatòria Y condicionada a $X = x$.

L'esperança condicionada d'Y donat que $X = x$, es calcula com

$$\mathbb{E}(Y|X = x) = \begin{cases} \int_{-\infty}^{\infty} y\, f_{Y|X=x}(y)\, dy & \text{en el cas continu,} \\ \sum_y y\, f_{Y|X=x}(y) & \text{en el cas discret.} \end{cases}$$

Quan el valor x de X al que condicionem varia, l'esperança condicionada d'Y també canvia.

Això ens permet definir la següent funció,

$$m_Y(x) = \mathbb{E}(Y|X = x),$$

que està definida per a tot x tal que $f_X(x) > 0$ i verifiqui que la distribució de probabilitat condicionada, determinada per la densitat (o probabilitat) condicionada $f_{Y|X=x}(y)$, té esperança finita.

Definició 3.5.1. *La funció $m_Y(x) = \mathbb{E}(Y|X = x)$ s'anomena* FUNCIÓ ESPERANÇA CONDICIONADA *de Y donat X o també* FUNCIÓ DE REGRESSIÓ *de Y sobre X.*

Exemple 3.2.1, pàgina 79. Continuació. Tornem a l'exemple on llencem un dau de 4 cares i un altre de 3, i definim X i Y com la suma de punts obtinguts i el valor absolut de la diferència de punts, respectivament. Calculem tots els possibles valors de les esperances condicionades de Y, donat que $X = x_i$.

$$\begin{array}{r|cccc} y_j & 0 & 1 & 2 & 3 \\ \hline \Pr(Y = y_j|X = 2) & 1 & 0 & 0 & 0 \end{array} \qquad \Longrightarrow \mathbb{E}(Y|X = 2) = 0.$$

$$\begin{array}{r|cccc} y_j & 0 & 1 & 2 & 3 \\ \hline \Pr(Y = y_j|X = 3) & 0 & 1 & 0 & 0 \end{array} \qquad \Longrightarrow \mathbb{E}(Y|X = 3) = 1.$$

$$\begin{array}{r|cccc} y_j & 0 & 1 & 2 & 3 \\ \hline \Pr(Y = y_j|X = 4) & 1/3 & 0 & 2/3 & 0 \end{array} \qquad \Longrightarrow \mathbb{E}(Y|X = 4) = \tfrac{4}{3}.$$

$$\begin{array}{r|cccc} y_j & 0 & 1 & 2 & 3 \\ \hline \Pr(Y = y_j|X = 5) & 0 & 2/3 & 0 & 1/3 \end{array} \qquad \Longrightarrow \mathbb{E}(Y|X = 5) = \tfrac{5}{3}.$$

$$\begin{array}{r|cccc} y_j & 0 & 1 & 2 & 3 \\ \hline \Pr(Y = y_j|X = 6) & 1/2 & 0 & 1/2 & 0 \end{array} \qquad \Longrightarrow \mathbb{E}(Y|X = 6) = 1.$$

$$\begin{array}{r|cccc} y_j & 0 & 1 & 2 & 3 \\ \hline \Pr(Y = y_j|X = 7) & 0 & 1 & 0 & 0 \end{array} \qquad \Longrightarrow \mathbb{E}(Y|X = 7) = 1.$$

Aleshores, la funció esperança condicionada ve donada per la següent taula:

x_i	$m_Y(x_i) = \mathbb{E}(Y\|X = x_i)$
2	0
3	1
4	4/3
5	5/3
6	1
7	1

Recordeu que vam definir $Z = \max\{D_1, D_2\}$. Calculeu ara, com a exercici, les següents funcions esperances condicionades:
$\mathbb{E}(Z|X = x_i)$, per $x_i = 2, \dots, 7$,
$\mathbb{E}(Z|Y = y_j)$, per $y_j = 0, \dots, 3$,
$\mathbb{E}(X|Z = z_k)$, per $z_k = 1, \dots, 4$,
$\mathbb{E}(Y|Z = z_k)$, per $z_k = 1, \dots, 4$.

Exemple 3.5.1.

Sigui X, Y variables aleatòries tals que $(X, Y) \sim \mathrm{U}(C)$ on $C = \{(x, y)|x^2 + y^2 \leq 1,\ x \geq 0\}$ és el semicercle de radi 1 i que té abscisses positives. Volem calcular $\mathbb{E}(X|Y = y)$.

La funció de densitat conjunta és

$$f(x, y) = \begin{cases} \frac{1}{\pi/2} & \text{si } (x, y) \in C \\ 0 & \text{altrament} \end{cases}$$

Fixat $y \in [-1, 1]$ tenim $\mathbb{E}(X|Y = y) = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y}(x)\, dx$.

Calculem $f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)}$. En primer lloc,

$$f_Y(y) = \int_0^{\sqrt{1-y^2}} \frac{2}{\pi}\, dx = \frac{2}{\pi}\sqrt{1 - y^2}, \text{ si } -1 \leq y \leq 1.$$

Per tant, fixat $y \in [-1, 1]$, tenim

$$f_{X|Y=y}(x) = \frac{\frac{2}{\pi}}{\frac{2}{\pi}\sqrt{1 - y^2}} = \frac{1}{\sqrt{1 - y^2}}, \text{ per } 0 \leq x \leq \sqrt{1 - y^2},$$

és a dir, $(X|Y = y) \sim U([0, \sqrt{1 - y^2}])$. Per tant

$$m_X(y) = \mathbb{E}(X|Y = y) = \frac{0 + \sqrt{1 - y^2}}{2} = \frac{1}{2}\sqrt{1 - y^2}.$$

$\square$

Definició 3.5.2. *L'expressió $\mathbb{E}(Y|X)$ és la variable aleatòria definida com a la transformació de la variable aleatòria X mitjançant la funció $m_Y(x)$:*

$$\mathbb{E}(Y|X) = m_Y(X).$$

A $\mathbb{E}(Y|X)$ l'anomenarem ESPERANÇA CONDICIONADA *de Y donat X.*

La funció de densitat (o de probabilitat) de la variable aleatòria $\mathbb{E}(Y|X)$ es calcularà a partir de l'expressió de la funció esperança condicionada i de la densitat (o probabilitat) de X.

Exemple 3.2.1, pàgina 79. Continuació. Estudiem quins valors pren la variable aleatòria $\mathbb{E}(Y|X)$ i amb quina probabilitat els pren. Per això, farem servir la taula de la funció $m_Y(x_i)$ calculada abans:

$$\mathbb{E}(Y|X) = \begin{cases} 0 & \text{si } X = 2, \\ 4/3 & \text{si } X = 4, \\ 5/3 & \text{si } X = 5, \\ 1 & \text{si } X = 3, 6, 7. \end{cases}$$

Per tant

$$\mathbb{E}(Y|X) = \begin{cases} 0 & \text{amb probabilitat } \Pr(X = 2) = 1/12, \\ 4/3 & \text{amb probabilitat } \Pr(X = 4) = 3/12, \\ 5/3 & \text{amb probabilitat } \Pr(X = 5) = 3/12, \\ 1 & \text{amb probabilitat } \Pr(X = 3) + \Pr(X = 6) + \Pr(X = 7) = 5/12. \end{cases}$$

La distribució de probabilitat de la variable aleatoria $\mathbb{E}(Y|X)$ la podem donar amb una taula:

y	0	4/3	5/3	1
$\Pr(\mathbb{E}(Y\|X) = y)$	1/12	3/12	3/12	5/12

Veiem doncs que efectivament $\mathbb{E}(Y|X)$ és una variable aleatoria.

Recordeu que vam definir $Z = \max\{D_1, D_2\}$. Calculeu ara, com a exercici, la distribució de probabilitat de les següents esperances condicionades, vistes com a variables aleatorias: $\mathbb{E}(Z|X)$, $\mathbb{E}(Z|Y)$, $\mathbb{E}(X|Z)$, $\mathbb{E}(Y|Z)$.

Exemple 3.5.1, pàgina 123. Continuació. $(X, Y) \sim \mathrm{U}(C)$ on $C =$

$\{(x,y)|x^2+y^2 \leq 1,\ x \geq 0\}$ és el semicercle de radi 1 i que té abscisses positives. Volem ara caracteritzar la variable aleatòria $\mathbb{E}(X|Y)$.

Hem vist que

$$m_X(y) = \mathbb{E}(X|Y=y) = \frac{1}{2}\sqrt{1-y^2}.$$

Per tant, la variable aleatòria que volem calcular és

$$\mathbb{E}(X|Y) = m_X(Y) = \frac{1}{2}\sqrt{1-Y^2}.$$

En aquest cas $\mathbb{E}(X|Y)$ és una variable aleatoria contínua, de la qual podríem calcular la seva funció de densitat

$$f_{\mathbb{E}(X|Y)}(z).$$

D'una forma anàloga a com hem definit la variable aleatòria esperança de Y condicionada a X tenim la variància condicionada:

Definició 3.5.3. *L'expressió $Var(Y|X)$ és una variable aleatòria definida com a funció de la variable aleatòria X, $v_Y(X)$, on la funció $v_Y(x)$ és*

$$v_Y(x) = Var(Y|X=x) = \mathbb{E}(Y^2|X=x) - (\mathbb{E}(Y|X=x))^2.$$

Aquesta definició ens diu que $\mathrm{Var}(Y|X) = \mathbb{E}(Y^2|X) - (\mathbb{E}(Y|X))^2$. Recordeu que els tres termes que apareixen en aquesta expressió són variables aleatorias.

Anem a veure algunes propietats de les noves variables aleatòries que hem definit

Proposició 3.5.4. *Siguin X, Y variables aleatòries. Aleshores*

1. Llei de l'esperança iterada *(o* Llei de l'esperança total*):*

$$\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y).$$

2. Llei de la variància iterada *(o* Llei de la variància total*):* *Si $Var(X) < \infty$ i $Var(Y) < \infty$ llavors*

$$Var(Y) = \mathbb{E}(Var(Y|X)) + Var(\mathbb{E}(Y|X)).$$

Demostració:

1. En efecte,
$$E(E(Y|X)) = \int_{-\infty}^{\infty} E(Y|X=x) f_X(x)\,dx$$
$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y f_{Y|X=x}(y)\,dy \right) f_X(x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \underbrace{f_{Y|X=x}(y) f_X(x)}_{f(x,y)}\,dy\,dx$$
$$= \int_{-\infty}^{\infty} y \underbrace{\left(\int_{-\infty}^{\infty} f(x,y)\,dx \right)}_{f_Y(y)} dy = \int_{-\infty}^{\infty} y f_Y(y)\,dy = \mathbb{E}(Y).$$
La demostració pel cas discret és anàloga.

2. D'una banda tenim
$$\mathbb{E}(\mathrm{Var}\,(Y|X)) = \mathbb{E}(\mathbb{E}(Y^2|X)) - \mathbb{E}((\mathbb{E}(Y|X))^2) = \mathbb{E}(Y^2) - \mathbb{E}((\mathbb{E}(Y|X))^2).$$
D'altra banda
$$\mathrm{Var}\,(\mathbb{E}(Y|X)) = \mathbb{E}((\mathbb{E}(Y|X))^2) - (\mathbb{E}(\mathbb{E}(Y|X)))^2 = \mathbb{E}((\mathbb{E}(Y|X))^2) - (\mathbb{E}(Y))^2.$$
Per tant
$$\mathbb{E}(\mathrm{Var}\,(Y|X)) + \mathrm{Var}\,(\mathbb{E}(Y|X)) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = \mathrm{Var}\,(Y).$$

□

Exemple 3.2.1, pàgina 79. Continuació. A l'exemple on llencem un dau de 4 cares i un altre de 3, hem definit X i Y com la suma de punts obtinguts i el valor absolut de la diferència de punts, respectivament. A la pàgina 124 hem vist que la distribució de probabilitat de la variable aleatoria $\mathbb{E}(Y|X)$ ve donada per aquesta taula:

y	0	4/3	5/3	1
$\Pr(\mathbb{E}(Y\|X) = y)$	1/12	3/12	3/12	5/12

Calculem $\mathbb{E}(\mathbb{E}(Y|X))$:
$$\mathbb{E}(\mathbb{E}(Y|X)) = 0 \cdot \frac{1}{12} + \frac{4}{3} \cdot \frac{3}{12} + \frac{5}{3} \cdot \frac{3}{12} + 1 \cdot \frac{5}{12} = \frac{14}{12} = \frac{7}{6}.$$

Per altra banda, a la pàgina 94 hem calculat la distribució marginal de Y:

y	0	1	2	3
$\Pr(Y=y)$	3/12	5/12	3/12	1/12

Aleshores,

$$\mathbb{E}(Y) = 0\cdot\frac{3}{12} + 1\cdot\frac{5}{12} + 2\cdot\frac{3}{12} + 3\cdot\frac{1}{12} = \frac{14}{12} = \frac{7}{6}.$$

Comprovem doncs que $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X))$, tal com diu la Llei de l'esperança iterada.

Com a exercici comproveu que també es compleix la Llei de la variància iterada: $\mathrm{Var}\,(Y) = \mathbb{E}(\mathrm{Var}\,(Y|X)) + \mathrm{Var}\,(\mathbb{E}(Y|X))$.

Exemple 3.5.1, pàgina 123. Continuació. Calculem $\mathbb{E}(X)$ aplicant la Llei de l'Esperança Iterada:

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}\left(\frac{1}{2}\sqrt{1-Y^2}\right) = \int_{-1}^{1}\frac{1}{2}\sqrt{1-y^2}\frac{2}{\pi}\sqrt{1-y^2}dy =$$

$$\frac{1}{\pi}\int_{-1}^{1}(1-y^2)dy = \frac{2}{\pi} - \frac{1}{\pi}\frac{y^3}{3}\bigg|_{-1}^{1} = \frac{2}{\pi} - \frac{2}{3\pi} = \frac{4}{3\pi}.$$

Exemple 3.5.2.

En un peatge s'instal·la un comptador automàtic de cotxes que amb probabilitat p detecta cada cotxe que hi passa (suposem també independència entre la detecció o no de diferents cotxes). Sigui N la variable aleatoria que compta el nombre de cotxes que passen pel peatge en un minut i sigui Y el nombre de cotxes registrats pel comptador automàtic. Si suposem que $N \sim \mathrm{Poisson}(\lambda)$, quina és l'esperança de Y?

Observeu que

$$Y = \sum_{i=1}^{N} X_i,$$

on $X_1, \ldots$ són variables aleatorias independents amb distribució de Bernoulli de paràmetre p. Aleshores

$$(Y|N=n) \sim B(n,p) \Longrightarrow m(n) = \mathbb{E}(Y|N=n) = np$$

Aplicant la Llei de l'Esperança Iterada,

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|N)) = \mathbb{E}(m(N)) = \mathbb{E}(Np) = \mathbb{E}(N)p = \lambda p.$$

$\square$

Exemple 3.5.3.
Siguin X, Y variables aleatòries contínues amb densitat conjunta

$$f(x,y) = \begin{cases} e^{-y} & \text{si } 0 \le x \le y \\ 0 & \text{altrament} \end{cases}$$

(a)
$$f_X(x) = \int_x^\infty f(x,y)\,dy = \int_x^\infty e^{-y}\,dy = \left[-e^{-y}\right]_x^\infty = e^{-x}.$$
Per tant $X \sim \mathrm{Exp}(1)$.

(b)
$$f_Y(y) = \int_0^y f(x,y)\,dx = \int_0^y e^{-y}\,dx = e^{-y}\,[x]_0^y = ye^{-y},$$
Per tant $Y \sim \gamma(\alpha = 2, \beta = 1)$.

(c)
$$f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)} = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}.$$
Per tant $(X|Y = y) \sim \mathrm{U}[0, y]$.

(d)
$$f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, \text{ si } y \ge x.$$
Per tant $(Y|X = x)$ segueix una exponencial desplaçada.

(e)
$$E(X|Y = y) = \int_0^y x\frac{1}{y}\,dx = \frac{1}{y}\left[\frac{x^2}{2}\right]_0^y = \frac{1}{y}\frac{y^2}{2} = \frac{y}{2}.$$
De fet ho podíem veure directament ja que $(X|Y = y) \sim \mathrm{U}[0, y]$. Usant això tenim
$$\mathrm{Var}\,(X|Y = y) = \frac{1}{12}y^2.$$
Per tant
$E(X|Y) = \frac{Y}{2}$ i $\mathrm{Var}\,(X|Y) = \frac{Y^2}{12}$.

(f)

$$E(Y|X=x) = \int_x^\infty y e^{-(y-x)}\,dy = \left\{ \begin{array}{cc} u = y & du = 1 \\ dv = e^{-(y-x)} & v = -e^{-(y-x)} \end{array} \right\}$$

$$= e^x \left[-ye^{-y}\right]_x^\infty - \int_x^\infty e^{-y}\,dy = e^x \left[xe^{-x} + e^{-x}\right] = x+1.$$

$$E(Y^2|X=x) = \int_x^\infty y^2 e^{-(y-x)}\,dy = \left\{ \begin{array}{c} z = y-x \\ dz = dy \end{array} \right\}$$

$$= \int_0^\infty (z+x)^2 e^{-z}\,dz = \int_0^\infty z^2 e^{-z}\,dz + x^2 \int_0^\infty e^{-z}\,dz + 2x \int_0^\infty z e^{-z}\,dz$$

$$= \frac{\Gamma(2)}{1^2} \underbrace{\int_0^\infty z \frac{1^2}{\Gamma(2)} z e^{-z}\,dz}_{\mathbb{E}(\gamma(\alpha=2,\beta=1))} + x^2 + 2x = 2 + x^2 + 2x = (x+1)^2 + 1,$$

on hem usat que si $X \sim \gamma(\alpha, \beta)$ llavors
$E(X) = \frac{\alpha}{\beta}$,
$E(X^2) = \frac{\alpha(\alpha+1)}{\beta^2}$ i $\mathrm{Var}\,(X) = \frac{\alpha}{\beta^2}$.
Amb aquests càlculs podem trobar la variable aleatòria $\mathrm{Var}\,(Y|X)$:

$$\mathrm{Var}\,(Y|X=x) = 1 + (x+1)^2 - (x+1)^2 = 1.$$

Per tant
$E(Y|X) = X+1$ i $\mathrm{Var}\,(Y|X) = 1$.

(g)

$$\mathrm{Var}\,(Y) = \mathrm{Var}\,(\gamma(\alpha=2, \beta=1)) = \frac{2}{1^2} = 2.$$

Ara podem comprovar la segona part de la proposició anterior:

$$\mathrm{Var}\,(Y) = E(\mathrm{Var}\,(Y|X)) + \mathrm{Var}\,(E(Y|X)) = E(1) + \mathrm{Var}\,(X+1)$$

$$= 1 + \mathrm{Var}\,(X) = 1 + 1 = 2.$$

(h)

$$1 = \mathrm{Var}\,(X) = E(\mathrm{Var}\,(X|Y)) + \mathrm{Var}\,(E(X|Y)) = E(Y^2/12) + \mathrm{Var}\,(Y/2)$$

$$= \frac{1}{12}(\mathrm{Var}\,(Y) + E(Y)^2) + \frac{1}{4}\mathrm{Var}\,(Y) = \frac{1}{12}(2+2^2) + \frac{1}{4}2 = 1.$$

$\square$

Exemple 3.5.4.
Suposem que X, Y són dues variables aleatòries tals que $\mathbb{E}(Y|X) = aX + b$. Suposem que mesuren, per exemple, la relació entre el pes i l'alçada d'un individu. Ens interessa calcular $\mathbb{E}(XY)$.

$$\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(XY|X)) = \mathbb{E}(X\mathbb{E}(Y|X))$$

$$= \mathbb{E}(X(aX + b)) = \mathbb{E}(aX^2 + bX) = a\mathbb{E}(X^2) + b\mathbb{E}(X).$$

$\square$

3.5.1 El problema de la predicció

Donades dues variables aleatòries (X, Y) amb funció de densitat conjunta $f(x, y)$ volem predir el valor de Y a partir de l'observació del valor de X. Cerquem per tant una funció de X, $r(X)$, que es pugui considerar una bona aproximació de Y.

Per traduir matemàticament la idea d'una bona aproximació m el que farem és minimitzar l'ERROR QUADRÀTIC MIG (EQM): $\mathbb{E}((Y - m)^2)$. En el cas de condicionar al valor de X, cerquem $r(X)$ per a cada valor de X de forma que minimitzem $\mathbb{E}((Y - r(X))^2|X)$.

Per simplificar els càlculs ens centrarem en minimitzar $\mathbb{E}((Y - m)^2)$, on m és una constant.

$$\mathbb{E}((Y - m)^2) = \mathbb{E}(Y^2) - 2m\mathbb{E}(Y) + m^2,$$

llavors

$$\frac{d}{dm}\mathbb{E}((Y - m)^2) = -2\mathbb{E}(Y) + 2m = 0 \Rightarrow m = \mathbb{E}(Y).$$

Obtenim doncs que $r(X) = \mathbb{E}(Y|X)$ és la millor predicció de Y, en tant que minimitza l'error quadràtic mig.

Si observem $X = x$ prediem Y mitjançant $r(x) = \mathbb{E}(Y|X = x)$. La funció r així definida s'anomena FUNCIÓ DE REGRESSIÓ DE Y SOBRE X. L'error quadràtic mig d'aquesta predicció és $\mathrm{Var}\,(Y|X = x)$ i l'error quadràtic mig global (sense especificar el valor x) és $\mathbb{E}(\mathrm{Var}\,(Y|X))$.

Fixem-nos que, pel que hem vist en la segona part de la proposició, la reducció de l'error quadràtic mig gràcies a la predicció és

$$\mathrm{Var}\,(Y) - \mathbb{E}(\mathrm{Var}\,(Y|X)) = \mathrm{Var}\,(\mathbb{E}(Y|X)).$$

Hi han dues definicions interessants al voltant de l'error quadràtic mig:

Definició 3.5.5. *Anomenem* ERROR QUADRÀTIC MIG GLOBAL *(EQMG) a* $\mathbb{E}(\mathit{Var}\,(Y|X))$*.*

Anomenem ERROR QUADRÀTIC MIG DE LA PREDICCIÓ *quan* $X = x$ *a* $\mathit{Var}\,(Y|X = x)$*.*

Observeu que la igualtat derivada de la Llei de la variància itrerada,

$$\mathrm{Var}\,(Y) = \mathbb{E}(\mathrm{Var}\,(Y|X)) + \mathrm{Var}\,(\mathbb{E}(Y|X))$$

es pot expressar en termes d'errors quadratics mitjans de predicció:

$$\mathrm{EQM}_{\mathbb{E}(Y)} = \mathrm{EQMG}_{\mathbb{E}(Y|X)} + \mathrm{Var}\,(\mathbb{E}(Y|X)).$$

Exemple 3.5.5.

La variable aleatòria Y s'escull amb idèntica probabilitat entre dues poblacions P_1 i P_2.

Si treballem amb P_1 la funció de densitat de Y és la d'una $\mathrm{U}[0,1]$. Si treballem amb P_2 la funció de densitat de Y és

$$f_{Y|P_2}(y) = \begin{cases} 2y & \text{si } 0 \le y \le 1 \\ 0 & \text{altrament} \end{cases}$$

1. Volem calcular la densitat de la variable aleatòria Y, el valor predit per Y i l'error quadràtic mig per Y.

2. Definim la variable aleatòria

$$X = \begin{cases} 1 & \text{amb probabilitat } 1/2 \\ 2 & \text{amb probabilitat } 1/2 \end{cases}$$

   Si $X = 1$ la variable aleatòria Y s'escull de P_1 i si $X = 2$ s'escull de P_2.

   Volem calcular el valor predit per Y i l'error quadràtic mig si Y s'escull de P_1 i també si s'escull de P_2.

Solucions:

1. P_1 té probabilitat 1/2 i funció de densitat

$$f_{Y|P_1}(y) = \begin{cases} 1 & 0 \le y \le 1 \\ 0 & \text{altrament} \end{cases}.$$

   P_2 té probabilitat 1/2 i funció de densitat

$$f_{Y|P_2}(y) = \begin{cases} 2y & 0 \le y \le 1 \\ 0 & \text{altrament} \end{cases}.$$

   La funció de distribució de Y és

$$F_Y(y) = \Pr(Y \le y) = \Pr(Y \le y|X = 1)\Pr(X = 1) + \Pr(Y \le y|X = 2)\Pr(X = 2) =$$

$$\frac{1}{2}F_1(y) + \frac{1}{2}F_2(y).$$

   Per tant la densitat de Y és $g(y) = \frac{1}{2}(f_{Y|P_1}(y) + f_{Y|P_2}(y))$ per $0 \le y \le 1$. Llavors

$$\mathbb{E}(Y) = \int_0^1 y\, g(y)\, dy = 1/2 \int_0^1 (y + 2y^2)\, dy = \frac{7}{12}.$$

$$\mathbb{E}(Y^2) = \int_0^1 y^2 g(y)\, dy = \frac{5}{12}.$$

   Per tant $\mathrm{Var}\,(Y) = \frac{5}{12} - (\frac{7}{12})^2 = \frac{11}{144}$.

   Tenim doncs que el valor predit per Y és $7/12$ amb error quadràtic mig $\frac{11}{144} = 0.076$.

2. El valor predit per Y si s'escull de la població P_i és $\mathbb{E}(Y|X = x_i)$ i l'error d'aquesta predicció és $\mathrm{Var}\,(Y|X = x_i)$.

   Població 1: Tenim $Y|X = 1 \sim \mathrm{U}[0, 1]$. Llavors

$$\mathbb{E}(Y|X = 1) = 1/2, \mathrm{Var}\,(Y|X = 1) = 1/12.$$

   Població 2: Tenim $\mathbb{E}(Y|X = 2) = 2/3$ i $\mathrm{Var}\,(Y|X = 2) = 1/18$.

   Per tant, el valor predit és

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y|X = 1)\Pr(X = 1) + \mathbb{E}(Y|X = 2)\Pr(X = 2) =$$

$$\frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{2}{3} = \frac{7}{12}.$$

Per la seva banda, l'error quadràtic mig global de predir Y a partir de X és

$$\begin{aligned}\mathbb{E}(\mathrm{Var}\,(Y|X)) &= \mathrm{Var}\,(Y|X=1)\,\Pr(X=1) + \mathrm{Var}\,(Y|X=2)\,\Pr(X=2)\\ &= \frac{1}{2}\left(\frac{1}{12}+\frac{1}{18}\right) = \frac{10}{144} = 0.069.\end{aligned}$$

Per tant, podem concloure que, si podem saber de quina població s'ha escollit Y l'error quadràtic mig es redueix de $\frac{11}{144}$ a $\frac{10}{144}$.

Observeu que

$$\mathbb{E}(Y|X) = \begin{cases} 6/12 = 3/6 \text{ amb probabilitat } 1/2\\ 8/12 = 4/6 \text{ amb probabilitat } 1/2\end{cases}$$

Aleshores

$$\mathbb{E}(Y|X) = \frac{1}{6}(3+B),$$

on $B \sim \mathrm{Bern}(1/2)$. Aleshores

$$\mathrm{Var}\,(\mathbb{E}(Y|X)) = \frac{1}{36}\mathrm{Var}\,(B) = \frac{1}{36}\frac{1}{4} = \frac{1}{144}.$$

Observeu que

$$\mathrm{Var}\,(\mathbb{E}(Y|X)) + \mathbb{E}(\mathrm{Var}\,(Y|X)) = \frac{1}{144} + \frac{10}{144} = \frac{11}{144} = \mathrm{Var}\,(Y).$$

□

3.6 Transformacions de variables aleatòries multivariants

Sigui $\mathbf{X} = (X_1, \ldots, X_n)$ amb funció de densitat conjunta $f(x_1, \ldots, x_n)$ i amb suport S, és a dir: $\Pr\{(X_1, \ldots, X_n) \in S\} = 1$.

Donada $r : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ amb $m \leq n$, sigui

$$\mathbf{Y} = (Y_1, \ldots, Y_m) = r(X_1, \ldots, X_n) \text{ i}$$

$$r_i(X_1, \ldots, X_n) = Y_i, \text{ amb } i = 1, \ldots, m.$$

Considerem també $T = r(S)$, és a dir, T és el suport de $\mathbf{Y}$.

El nostre objectiu és determinar la llei conjunta del vector $\mathbf{Y}$.

Mètode general

Calcularem la funció de distribució $G_{\mathbf{Y}}$, això és:

$$G_{\mathbf{Y}}(\mathbf{y}) = \Pr(\mathbf{Y} \leq \mathbf{y}) = \Pr(Y_1 \leq y_1, \dots, Y_m \leq y_m)$$
$$= \Pr(r(X_1, \dots, X_n) \leq \mathbf{y}) = \int \cdots \int_{A_{\mathbf{y}}} f(x_1, \dots, x_n)\, dx_1 \cdots dx_n,$$

on $A_{\mathbf{y}} = \{(x_1, \dots, x_n) \in \mathbb{R}^n | r(x_1, \dots, x_n) \leq \mathbf{y}\}$.

Quan $G_{\mathbf{Y}}(\mathbf{y})$ sigui diferenciable respecte cadascuna de les variables y_i, definirem la densitat conjunta de $\mathbf{Y}$ com

$$g(y_1, \dots, y_m) = \frac{\partial^m}{\partial y_1 \cdots \partial y_m} G_{\mathbf{Y}}(\mathbf{y}).$$

3.6.1 Distribució de la suma

Sigui (X, Y) variable aleatoria bivariant. Sigui $Z = X+Y$. Volem determinar la distribució de probabilitat de Z, a partir de la distribució conjunta de (X, Y).

Cas (X, Y) variable aleatoria bivariant discreta

Sigui $p(x_i, y_j)$ la funció de probabilitat conjunta de (X, Y):

$$p(x_i, y_j) = \Pr(X = x_i, Y = y_j),\ i = 1, 2, \dots,\ j = 1, 2, \dots$$

La variable aleatoria suma, $Z = X + Y$, és una variable aleatoria univariant discreta perquè, com a molt, pot prendre tants valors com (X, Y). Els valors que pren Z són de la forma

$$x_i + y_j,\ i = 1, 2, \dots,\ j = 1, 2, \dots$$

alguns dels quals poden estar repetits. Sigui z_ℓ algun dels valors que pot prendre $Z = X + Y$. Aleshores la funció de probabilitat de Z és

$$\begin{aligned}\Pr(Z = z_\ell) &= \Pr(X + Y = z_\ell)\\ &= \sum_{i,j:x:i+y_j=z_\ell} \Pr(X = x_i, Y = y_j)\\ &= \sum_i \Pr(X = x_i, Y = z_\ell - x_i)\\ &= \sum_j \Pr(X = z_\ell - y_j, Y = y_j).\end{aligned}$$

Considerem ara el cas important que es presenta quan X i y són independents. En aquest cas l'expressió anterior es transforma en la següent:

$$\begin{aligned}\Pr(Z = z_\ell) &= \sum_i \Pr(X = x_i)\Pr(Y = z_\ell - x_i) \\ &= \sum_j \Pr(X = z_\ell - y_j)\Pr(Y = y_j).\end{aligned}$$

Aquestes expressions es coneixen com a CONVOLUCIÓ de les funcions de probabilitat de X i Y.

Diem doncs que la funció de probabilitat de la suma de dues variables aleatòries discretes independents és la convolució de les seves funcions de probabilitat.

Exemple 3.6.1.
La suma de Poissons independents és Poisson.
Recordeu que la variable aleatòria X segueix la LLEI POISSON, de paràmetre $\lambda > 0$, si pren valors $0, 1, \ldots$ i la seva funció de probabilitat és

$$\Pr(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \ldots$$

Recordeu també la fórmula del Binomi de Newton:

$$(a + b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^k b^{n-k}.$$

Siguin $X \sim \text{Poisson}(\lambda)$ i $Y \sim \text{Poisson}(\mu)$ independents. Calculem la funció de probabilitat de $Z = X + Y$. En primer lloc, Z pot prendre qualsevol valor sencer no negatiu. Sigui $z \in \{0,1,2,\ldots\}$,

$$\begin{aligned}
\Pr(Z=z) &= \sum_{k=0}^{\infty} \Pr(X=k)\Pr(Y=z-k) \\
&\quad \text{(ha de ser } z-k \geq 0\text{)} \\
&= \sum_{k=0}^{z} \Pr(X=k)\Pr(Y=z-k) \\
&= \sum_{k=0}^{z} e^{-\lambda}\frac{\lambda^k}{k!}e^{-\mu}\frac{\mu^{z-k}}{(z-k)!} \\
&= e^{-(\lambda+\mu)}\frac{1}{z!}\sum_{k=0}^{z}\frac{z!}{k!(z-k)!}\lambda^k\mu^{z-k} \\
&\quad \text{(Binomi de Newton)} \\
&= e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^z}{z!}.
\end{aligned}$$

Observem doncs que la funció de probabilitat de Z coincideix amb la de la Poisson de paràmetre $\lambda+\mu$. Per tant concloem que si $X \sim \text{Poisson}(\lambda)$ i $Y \sim \text{Poisson}(\mu)$ independents, aleshores

$$X+Y \sim \text{Poisson}(\lambda+\mu).$$

$\square$

Cas (X,Y) variable aleatoria bivariant contínua

Sigui $f(x,y)$ la funció de densitat conjunta de (X,Y). Veurem que la variable aleatoria suma, $Z=X+Y$, és una variable aleatoria univariant contínua i calcularem la seva funció de densitat.

Començarem calculant la funció de distribució de Z, avaluada a $z \in \mathbb{R}$:

$$
\begin{aligned}
F_Z(z) = & \Pr(Z \leq z) = \Pr(X+Y \leq z) = \Pr((X,Y) \in A_z) \\
& (\text{on } A_z = \{(x,y) \in \mathbb{R}^2 : x+y \leq z\}) \\
= & \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x,y)\, dy\, dx \\
& (s = x+y \Leftrightarrow y = s-x,\ ds = dy) \\
= & \int_{-\infty}^{\infty} \int_{-\infty}^{z} f(x,s-x)\, ds\, dx \\
& (\text{canviem l'ordre d'integració}) \\
= & \int_{-\infty}^{z} \int_{-\infty}^{\infty} f(x,s-x)\, dx\, ds \\
= & \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} f(x,s-x)\, dx \right) ds \\
& (g(s) = \int_{-\infty}^{\infty} f(x,s-x)\, dx) \\
= & \int_{-\infty}^{z} g(s)\, ds
\end{aligned}
$$

Pel Teorema Fonamental del Càlcul, F_Z és derivable i la seva derivada és

$$F'_Z(z) = g(z) = \int_{-\infty}^{\infty} f(x,z-x)\, dx.$$

Del fet que $F_Z(z)$ sigui derivable es segueix que $Z = X+Y$ és variable aleatoria contínua i que la seva densitat és $F'_Z(z)$:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x,z-x)\, dx.$$

Considerem ara el cas important que es presenta quan X i y són independents. En aquest cas l'expressió anterior es transforma en la següent:

$$
\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx \\
&= \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\, dy.
\end{aligned}
$$

Aquestes expressions es coneixen com a CONVOLUCIÓ de les funcions de densitat de X i Y.

Diem doncs que la funció de densitat de la suma de dues variables aleatòries contínues independents és la convolució de les seves funcions de densitat.

Exemple 3.6.2.
Proveu que si $X_1 \sim \text{Gamma}(\alpha_1, \beta)$ i $X_2 \sim \text{Gamma}(\alpha_2, \beta)$ són independents, aleshores $X_1 + X_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \beta)$.
Efectivament, aplicant la fórmula de la convolució, la funció de densitat de $X_1 + X_2$ serà

$$\begin{aligned} f_{X_1+X_2}(y) = & \int_0^y \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}(y-x)^{\alpha_1-1}x^{\alpha_2-1}e^{-\beta(y-x)-\beta x}dx \\ & \text{(fem el canvi de variable } t = x/y,\ dt = (1/y)dx) \\ = & \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}y^{\alpha_1+\alpha_2-1}e^{-\beta y}\int_0^1 (1-t)^{\alpha_1-1}t^{\alpha_2-1}dt \\ = & \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}y^{\alpha_1+\alpha_2-1}e^{-\beta y}\frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)} \\ = & \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}y^{\alpha_1+\alpha_2-1}e^{-\beta y} \end{aligned}$$

i es conclou que $X_1 + X_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \beta)$.

Hem fet servir la igualtat

$$\int_0^1 (1-t)^{\alpha_1-1}t^{\alpha_2-1}dt = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}.$$

que està provada a l'Annex **??**, quan es defineix la distribució Beta. □

Densitat de la suma de dos normals independents (convolució).

Proposició 3.6.1. *Siguin $X \sim N(0,1)$ i $Y \sim N(0,\sigma^2)$ independents. Aleshores $X + Y \sim N(0, 1+\sigma^2)$*

Demostració. Les funcions de densitat de X i Y són

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2},\ f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}y^2}.$$

Aplicant la fòrmula de la concolució de dos densitats, tenim que

$$\begin{aligned}
f_{X+Y}(s) &= \int_{-\infty}^{\infty} f_X(s-y)f_Y(y)dy \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(s-y)^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}y^2} dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}\left(y^2+\sigma^2(s-y)^2\right)} dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}\left(y^2(1+\sigma^2)-2\sigma^2 sy+\sigma^2 s^2\right)} dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1+\sigma^2}{2\sigma^2}\left(y^2-2\frac{\sigma^2}{1+\sigma^2}sy+\frac{\sigma^2 s^2}{1+\sigma^2}\right)} dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1+\sigma^2}{2\sigma^2}\left(\left\{y^2-2\frac{\sigma^2}{1+\sigma^2}sy+\frac{\sigma^4}{(1+\sigma^2)^2}s^2\right\}+\left\{-\frac{\sigma^4}{(1+\sigma^2)^2}s^2+\frac{\sigma^2 s^2}{1+\sigma^2}\right\}\right)} dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1+\sigma^2}{2\sigma^2}\left(y-\frac{\sigma^2}{1+\sigma^2}s\right)^2} e^{-\frac{1+\sigma^2}{2\sigma^2}\left(\frac{-\sigma^4 s^2+(1+\sigma^2)\sigma^2 s^2}{(1+\sigma^2)^2}\right)} dy \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1+\sigma^2}{2\sigma^2}\frac{\sigma^2 s^2}{(1+\sigma^2)^2}} \frac{1}{\sqrt{1+\sigma^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{1+\sigma^2}}} e^{-\frac{1}{2\frac{\sigma^2}{1+\sigma^2}}\left(y-\frac{\sigma^2}{1+\sigma^2}s\right)^2} dy \\
&= \frac{1}{\sqrt{2\pi}\sqrt{1+\sigma^2}} e^{-\frac{1}{2(1+\sigma^2)}s^2},
\end{aligned}$$

que és la funció de densitat d'una $N(0, 1+\sigma^2)$. □

Proposició 3.6.2. *Siguin $X \sim N(0,\sigma_1^2)$ i $Y \sim N(0,\sigma_2^2)$ independents. Aleshores $X+Y \sim N(0,\sigma_1^2+\sigma_2^2)$*

Demostració. Observeu que

$$\frac{X}{\sigma_1} \sim N(0,1),\ \frac{Y}{\sigma_1} \sim N(0,(\sigma_2/\sigma_1)^2).$$

Aplicant la proposició anterior tenim que

$$\frac{X}{\sigma_1} + \frac{Y}{\sigma_1} \sim N\left(0, 1+(\sigma_2/\sigma_1)^2\right).$$

Per tant

$$X+Y = \sigma_1\left(\frac{X}{\sigma_1}+\frac{Y}{\sigma_1}\right) \sim N(0,\sigma_1^2(1+(\sigma_2/\sigma_1)^2) = \sigma_1^2+\sigma_2^2).$$

□

Proposició 3.6.3. *Siguin* $X \sim N(\mu_1, \sigma_1^2)$ *i* $Y \sim N(\mu_2, \sigma_2^2)$ *independents. Aleshores* $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Demostració. Per la Proposició anterior tenim que

$$(X + Y) - (\mu_1 + \mu_2) = (X - \mu_1) + (Y - \mu_2) \sim N(0, \sigma_1^2 + \sigma_2^2).$$

Per tant

$$X + Y = \{(X + Y) - (\mu_1 + \mu_2)\} + (\mu_1 + \mu_2) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

□

3.6.2 Transformacions bijectives de vectors aleatoris continus

Recordem com es duu a terme un canvi de variable en una integral doble (les fórmules s'estenen a qualsevol dimensió). Sigui $y_1 = y_1(x_1, x_2)$, $y_2 = y_2(x_1, x_2)$ una transformació bijectiva

$$\begin{aligned} r \ : \ & S \subseteq \mathbb{R}^2 \to T \subset \mathbb{R}^2 \\ & (x_1, x_2) \mapsto (y_1, y_2) \end{aligned}$$

La transformació es pot invertir: $x_1 = x_1(y_1, y_2)$, $x_2 = x_2(y_1, y_2)$. Si existeixen les derivades parcials $\frac{\partial x_i}{\partial y_j}$ per a tot i, j i per a tot $(y_1, y_2) \in R$, es defineix el determinant Jacobià de la transformació inversa de r com el determinant de la matriu Jacobiana:

$$J_{r^{-1}}(y_1, y_2) = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1}.$$

Sigui $|J_{r^{-1}}(y_1, y_2)|$ el seu valor absolut. Suposem que totes les derivades parcials són contínues.

Teorema 3.6.4 (Canvi de variable en integrals dobles)**.** *Sigui* $g : \mathbb{R}^2 \to \mathbb{R}$. *Suposem que* r *transforma el conjunt* $A \subseteq S$ *en* $B \subseteq T$. *Aleshores*

$$\int\int_A g(x_1, x_2) dx_1 dx_2 = \int\int_B g(x_1(y_1, y_2), x_2(y_1, y_2)) |J_{r^{-1}}(y_1, y_2)| dy_1 dy_2.$$

Sigui $\mathbf{X} = (X_1, \ldots, X_n)$ amb funció de densitat conjunta $f(x_1, \ldots, x_n)$ i amb suport S, és a dir: $\Pr\{(X_1, \ldots, X_n) \in S\} = 1$.

Donada $r : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ bijectiva (això implica que les dimensions de $\mathbf{X}$ i de $r(\mathbf{X})$ han de coincidir), sigui $\mathbf{Y} = r(\mathbf{X})$. Volem determinar la distribució conjunta del vector $\mathbf{Y}$.

Sigui $T = r(S)$ el suport de $\mathbf{Y}$. Si r és bijectiva existeix la funció inversa:

$$\begin{array}{rccl} s: & T \subseteq \mathbb{R}^n & \longrightarrow & S \subseteq \mathbb{R}^n \\ & (y_1, \ldots, y_n) & \longmapsto & s(y_1, \ldots, y_n) = (x_1, \ldots, x_n) \end{array}$$

i les components $s_i(\mathbf{y}) = x_i$, per $i = 1, 2, \ldots, n$.

En aquests casos passarem de

$$\left.\begin{array}{l} Y_1 = r_1(X_1, \ldots, X_n) \\ \vdots \\ Y_n = r_n(X_1, \ldots, X_n) \end{array}\right\} \text{ a } \left.\begin{array}{l} X_1 = s_1(Y_1, \ldots, Y_n) \\ \vdots \\ X_n = s_n(Y_1, \ldots, Y_n) \end{array}\right\}$$

Teorema 3.6.5. *Sigui $r : T \subseteq \mathbb{R}^n \longrightarrow S \subseteq \mathbb{R}^n$ bijectiva amb inversa $r^{-1} = s : S \longrightarrow T$. Si existeixen totes les derivades parcials $\frac{\partial s_i}{\partial y_j}$ per a tot i, j i per a tot $\mathbf{y} \in T$, aleshores existirà la matriu Jacobiana de s a $\mathbf{y}$, $Jac_s(\mathbf{y})$, que tindrà la forma següent:*

$$Jac_s(\mathbf{y}) = \begin{pmatrix} \frac{\partial s_1}{\partial y_1} & \cdots & \frac{\partial s_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial s_n}{\partial y_1} & \cdots & \frac{\partial s_n}{\partial y_n} \end{pmatrix}.$$

Sigui $J_s(\mathbf{y})$ el determinant de $Jac_s(\mathbf{y})$ (l'anomenarem determinant Jacobià, o simplement Jacobià), i sigui $|J_s(\mathbf{y})|$ el seu valor absolut. Aleshores la funció de densitat conjunta de $\mathbf{Y}$ ve donada per

$$g_{\mathbf{Y}}(y_1, \ldots, y_n) = f(s_1(\mathbf{y}), \ldots, s_n(\mathbf{y}))|J_s(\mathbf{y})|,$$

per $\mathbf{y} \in T$.

Demostració: La demostració és una aplicació immediata del canvi de variable en integració múltiple. Sigui $A \subseteq \mathbb{R}^n$, calculem

$$\begin{aligned} \Pr(\mathbf{Y} \in A) &= \Pr(s(\mathbf{Y}) \in s(A)) \\ &= \Pr(\mathbf{X} \in s(A)) \\ &= \int \cdots \int_{s(A)} f(x_1, \ldots, x_n) dx_1 \ldots dx_n \\ &\quad \text{(apliquem el canvi de variable } \mathbf{y} = r(\mathbf{x}) \Longleftrightarrow \mathbf{x} = s(\mathbf{y})\text{)} \\ &= \int \cdots \int_A f(s_1(\mathbf{y}), \ldots, s_n(\mathbf{y}))|J_s(\mathbf{y})| dy_1 \ldots dy_n. \end{aligned}$$

Per tant, $\Pr(\mathbf{Y} \in A)$ es pot calcular integrant sobre el conjunt A la funció, $f(s_1(\mathbf{y}), \ldots, s_n(\mathbf{y}))|J_s(\mathbf{y})|$, i aquesta és precisament la definició de funció de densitat de la variable aleatoria $\mathbf{Y}$. $\square$

Exemple 3.6.3.
Transformació producte i quocient
Sigui (X_1, X_2) vector aleatori amb densitat conjunta

$$f(x_1, x_2) = 4x_1x_2, \text{ si } 0 < x_i < 1 \text{ amb } i = 1, 2.$$

Volem determinar la llei conjunta de (Y_1, Y_2) on $Y_1 = \frac{X_1}{X_2}$ i $Y_2 = X_1X_2$.
Tenim les transformacions:

$$\left\{ \begin{array}{l} Y_1 = r_1(X_1, X_2) = \frac{X_1}{X_2} \\ Y_2 = r_2(X_1, X_2) = X_1X_2 \end{array} \right. ,$$

que és equivalent a

$$\left\{ \begin{array}{l} X_1 = s_1(Y_1, Y_2) = (Y_1Y_2)^{1/2} \\ X_2 = s_2(Y_1, Y_2) = (\frac{Y_2}{Y_1})^{1/2} \end{array} \right. .$$

Els respectius suports són

$$S = \{(x_1, x_2) \in \mathbb{R}^2 \text{ tals que } 0 < x_i < 1, i = 1, 2\},$$

$$T = \{(y_1, y_2) \in \mathbb{R}^2 \text{ tals que } 0 < y_1, 0 < y_2 < 1, 0 < y_1y_2 < 1, 0 < \frac{y_2}{y_1} < 1\}.$$

$$J_s(y_1, y_2) = det \begin{pmatrix} \frac{y_2}{2(y_1y_2)^{1/2}} & \frac{y_1}{2(y_1y_2)^{1/2}} \\ y_2^{1/2}(-\frac{1}{2})y_1^{-3/2} & \frac{1}{2(y_1y_2)^{1/2}} \end{pmatrix}$$

$$= \frac{y_2}{4(y_1y_2)} + \frac{y_1}{2(y_1y_2)^{1/2}} y_2^{1/2} \frac{1}{2y^{3/2}} = \frac{1}{4y_1} + \frac{1}{4y_1} = \frac{1}{2y_1}.$$

Finalment podem calcular la llei conjunta demanada:

$$g(y_1, y_2) = f\left((y_1y_2)^{1/2}, \left(\frac{y_2}{y_1}\right)^{1/2}\right) \frac{1}{2y_1}$$

$$= 4(y_1y_2)^{1/2} \left(\frac{y_2}{y_1}\right)^{1/2} \frac{1}{2y_1} = 2\frac{y_2}{y_1}, \text{ amb } (y_1, y_2) \in T.$$

$\square$

Exemple 3.6.4.
Donades X_1, X_2 variables aleatòries amb funció de densitat conjunta

$$f(x_1,x_2) = \left\{ \begin{array}{cc} x_1 + x_2 & \text{si } 0 < x_i < 1, \text{ per } i = 1,2 \\ 0 & \text{altrament} \end{array} \right. .$$

Volem trobar la funció de densitat de $Y = X_1 \cdot X_2$.

Definim el canvi bijectiu:

$$\left. \begin{array}{c} Y_1 = X_1 X_2 \\ Y_2 = X_2 \end{array} \right\}$$

$$\left. \begin{array}{c} X_1 = \frac{Y_1}{Y_2} \\ X_2 = Y_2 \end{array} \right\}$$

Tenim:

$$S = [0,1]^2, T = \{(y_1,y_2) : 0 < y_1 < y_2 < 1\}.$$

$$J_s(y_1,y_2) = det \begin{pmatrix} \frac{1}{y_2} & \frac{-y_1}{y_2^2} \\ 0 & 1 \end{pmatrix} = \frac{1}{y_2}.$$

Llavors:

$$g(y_1,y_2) = f\left(\frac{y_1}{y_2}, y_2\right)\frac{1}{y_2} = ((y_1/y_2) + y_2)\frac{1}{y_2}, \text{ amb } (y_1,y_2) \in T.$$

Calculem una de les densitats marginals:

$$g_{Y_1}(y_1) = \int_{y_1}^{1} g(y_1,y_2)\,dy_2 = \int_{y_1}^{1} ((y_1/y_2) + y_2)\frac{1}{y_2}\,dy_2$$

$$= \int_{y_1}^{1} ((y_1/y_2) + 1)\,dy_2 = [(-y_1/y_2) + y_2]_{y_1}^{1} = 2(1 - y_1).$$

□

Exemple 3.6.5.
Siguin X, Y variables aleatòries independents i idènticament distribuïdes segons la llei Exp(1). Volem determinar la llei de les variables aleatòries $\frac{X}{X+Y}$ i $X + Y$ i comprovar si són o no independents.

Tenim que

$$f(x,y) = \begin{cases} e^{-(x+y)} & \text{si } x, y > 0, \\ 0 & \text{altrament,} \end{cases}$$

amb $S = (\mathbb{R}^+)^2$.

Les transformacions són:

$$\left. \begin{array}{c} U = \frac{X}{X+Y} \\ V = X + Y \end{array} \right\}$$

$$\left. \begin{array}{c} UV = X \\ \mathrm{Var}\,(1-U) = Y \end{array} \right\}$$

amb $T = [0,1] \times \mathbb{R}^+$.

Llavors

$$J_s(u,v) = det \begin{pmatrix} v & u \\ -v & 1-u \end{pmatrix} = v.$$

Calculem ara la densitat conjunta de (U, V) i les marginals de U i V:

$$g_{(U,V)}(u,v) = f_{(X,Y)}(uv, v(1-u)) \cdot v$$

$$= e^{-(uv+v(1-u))} \cdot v = ve^{-v}, \text{ per a } 0 < u < 1 \text{ i } v > 0.$$

$$g_U(u) = 1, \text{ per a } 0 < u < 1, U \sim \text{Uniforme}(0,1),$$

$$g_V(v) = ve^{-v}, V \sim \gamma(2,1).$$

Podem concloure que U i V són independents. □

Exemple 3.6.6.

Transformacions lineals. Convolució.

Sigui $(X_1, \ldots, X_n)$ amb funció de densitat conjunta $f(x_1, \ldots, x_n)$.

Donada una matriu A no singular d'ordre n, definim

$$\mathbf{Y} = A\mathbf{X}, \mathbf{Y} = (Y_1, \ldots, Y_n).$$

Aleshores la funció de densitat conjunta de $\mathbf{Y}$ ve donada per:

$$g(\mathbf{y}) = g(y_1, \ldots, y_n) = \frac{1}{|\mathrm{det}A|} f(A^{-1}\mathbf{y}), \text{ per } y \in \mathbb{R}^n.$$

Per exemple, considerem $Y_1 = X_1 + X_2$ i $Y_2 = X_2$. Llavors:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

amb det(A)=1, i

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Llavors

$$g(y_1, y_2) = f(y_1 - y_2, y_2),$$
$$g(y_1) = \int_{-\infty}^{\infty} f(y_1 - y_2, y_2)\, dy_2.$$

En particular, si X_1 i X_2 són independents tenim que $f(x_1, x_2) = f_1(x_1)f_2(x_2)$, aleshores

$$g(y_1) = \int_{-\infty}^{\infty} f_1(y_1 - y_2) f_2(y_2)\, dy_2,$$

que és la CONVOLUCIÓ de f_1 i f_2. □

3.7 Distribució multinomial

Definició 3.7.1. *Considerem un experiment aleatori que pot acabar en k possibles resultats mutuament excloents (no es poden presentar dos diferents simultàneament). Suposem que el resultat i-èsim es presenta amb probabilitat p_i, $i = 1, \ldots, k$. Per tant*

$$\sum_{i=1}^{k} p_i = 1.$$

Repetim n vegades l'experiment de forma que els n experiments siguin independents i anomenem X_i al número de vegades que hem obtingut el resultat de tipus i, $i = 1, \ldots, k$.

Llavors diem que $\mathbf{X} = (X_1, \ldots, X_k)$ presenta una DISTRIBUCIÓ MULTINOMIAL *de dimensió k amb paràmetres n i $\mathbf{p} = (p_1, \ldots, p_k)$,*

$$(X_1, \ldots, X_k) \sim M_k(n; p_1, \ldots, p_k).$$

La seva funció de probabilitat és

$$\Pr(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k}$$

si $\sum_{i=1}^{k} x_i = n$.

Exemple 3.7.1.
En n experiments independents de Bernoulli, amb probabilitat d'èxit p, anomenem X_1 al nombre d'èxits obtinguts i $X_2 = n - X_1$ al nombre de fracasos. Aleshores

$$(X_1, X_2) \sim M_2(n; p, 1-p).$$

$\square$

Exemple 3.7.2.
Considerem una població que conté objectes de k tipus diferents ($k \geq 2$), amb proporcions p_i (amb $i = 1, 2, \ldots, k$), de manera que $\sum_{i=1}^{k} p_i = 1$.

Seleccionem n objectes amb reemplaçament i anomenem X_i al número d'objectes que hem obtingut del tipus i.

Llavors $(X_1, \ldots, X_k) \sim M_k(n; p_1, \ldots, p_k)$.

$\square$

Alguns valors característics són:

- Cada X_i és $B(n, p_i)$, i per tant $\mathbb{E}(X_i) = np_i$ i $\mathrm{Var}\,(X_i) = np_i(1-p_i)$.
- $\mathrm{Var}\,(X_i + X_j) = n(p_i + p_j)(1 - p_i - p_j)$, perquè $X_i + X_j \sim B(n, p_i + p_j)$.
- $\mathrm{Cov}(X_i, X_j) = -np_ip_j$. Efectivament,

$$n(p_i+p_j)(1-p_i-p_j) = \mathrm{Var}\,(X_i+X_j) = \mathrm{Var}\,(X_i)+\mathrm{Var}\,(X_j)+2\mathrm{Cov}(X_i, X_j)$$

$$= np_i(1-p_i) + np_j(1-p_j) + 2\mathrm{Cov}(X_i, X_j).$$

Aleshores

$$\begin{aligned} \mathrm{Cov}(X_i, X_j) &= \frac{1}{2}(n(p_i + p_j)(1 - p_i - p_j) - np_i(1-p_i) + np_j(1-p_j)) \\ &= -np_ip_j. \end{aligned}$$

3.8 Normal bivariant

Definició 3.8.1. *Sigui* $V_1 \sim N(0,1)$, $V_2 \sim N(0,1)$ *amb* V_1 *i* V_2 *independents. Llavors*

$$f_{(V_1,V_2)}(v_1,v_2) = f_{V_1}(v_1) f_{V_2}(v_2)$$

$$= \frac{1}{\sqrt{2\pi}} e^{\frac{-v_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-v_2^2}{2}} = \frac{1}{2\pi} e^{-(1/2)(v_1^2+v_2^2)} = \frac{1}{2\pi} e^{-(1/2)\mathbf{v}^t\cdot\mathbf{v}},$$

amb $\mathbf{v} = (v_1, v_2)^t \in \mathbb{R}^2$. *La llei conjunta de dues variables aleatòries independents* $N(0,1)$ *s'anomena* Llei Normal bivariant estàndard amb marginals independents.

Exemple 3.8.1.

Aquest exemple requereix un canvi de variables en integrals dobles.
Anem a comprovar que l'expressió de la funció $\varphi(x)$ d'una normal estàndard és efectivament una densitat.

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{\frac{-x^2}{2}\}\, dx = 1 \Leftrightarrow I = \int_{-\infty}^{\infty} \exp\{\frac{-x^2}{2}\}\, dx = \sqrt{2\pi} \Leftrightarrow I^2 = 2\pi.$$

Veurem que l'última igualtat és certa:

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}(x^2+y^2)\}\, dy\, dx.$$

Considerem el següent canvi a coordenades polars:

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}.$$

Tenim $x^2 + y^2 = r^2$ i

$$\begin{vmatrix} \frac{dx}{dr} & \frac{dx}{d\theta} \\ \frac{dy}{dr} & \frac{dy}{d\theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

Llavors

$$
\begin{aligned}
I^2 &= \int_0^{2\pi} \int_0^{\infty} \exp(-\frac{1}{2}r^2) r \, dr \, d\theta \\
&= \int_0^{2\pi} d\theta \int_0^{\infty} \exp(-\frac{1}{2}r^2) r \, dr \\
&= \theta|_0^{2\pi} \int_0^{\infty} \exp(-\frac{1}{2}r^2) r \, dr \\
&= 2\pi \int_0^{\infty} \exp(-\frac{1}{2}r^2) r \, dr \\
& \quad \{u = r^2/2, \ du = r \, dr\} \\
&= 2\pi \int_0^{\infty} e^{-u} \, du \\
&= 2\pi \left(-e^{-u}\right)\Big|_0^{\infty} \\
&= 2\pi.
\end{aligned}
$$

$\square$

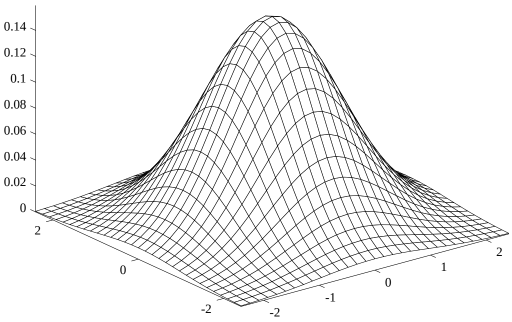
Alguns fets destacables de la distribució normal bivariant estàndard amb marginals independents són els següents:

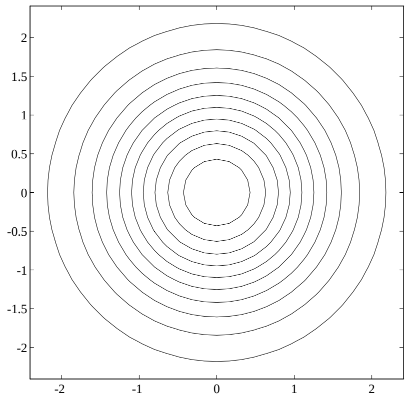
- La densitat de (Z_1, Z_2) és més gran a mesura que ens acostem a $(0, 0)$.

- Les corbes de nivell de (Z_1, Z_2) són circumferències centrades al punt $(0, 0)$:

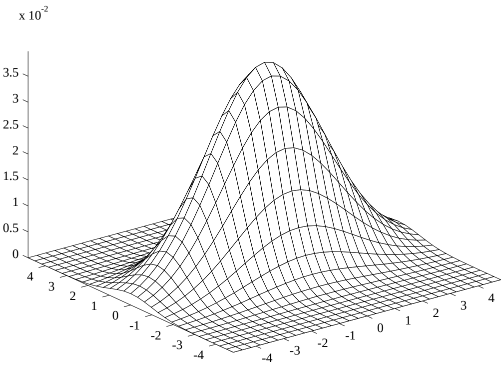
$$f_{(Z_1,Z_2)}(z_1, z_2) = k \iff z_1^2 + z_2^2 = k',$$

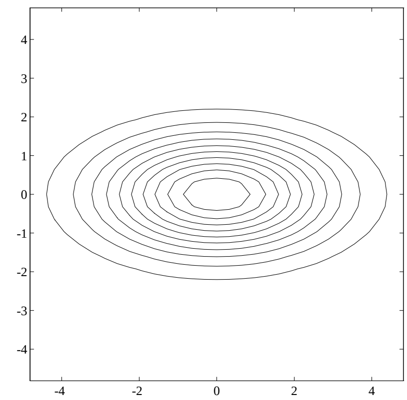
on $k' = -2\log(2\pi k)$.



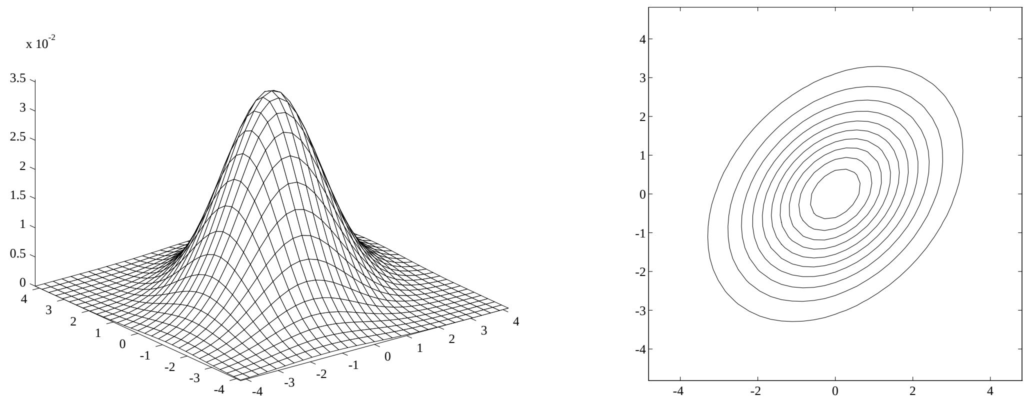


- Sigui $(X, Y) = (aZ_1, bZ_2)$ per $a, b \in \mathbb{R}$. La distribució de (X, Y) ja no és la mateixa d'abans, però les característiques fonamentals de (X, Y) no varien: és la distribució conjunta de dues distribucions normals independents. La nova distribució també és normal bivariant. Observeu





- Imagineu que podem agafar la distribució (X, Y) i que li apliquem una rotació que deixi fix el $(0, 0)$. Les corbes de nivell de la distribució resultant són encara el·lipses, però ara ja no tenen els seus eixos sobre les eixos de coordenades, però l'essencial de la distribució normal de sortida es manté. Ara les marginals de la nova distribució ja no són independents.



- Si traslladem la distribució des del $(0,0)$ a qualsevol altre punt del pla, la distribució resultant és bàsicament la mateixa: seguirà essent normal bivariant.

Essencialment, totes les distribucions normals bivariants es podem obtenir a partir de la nostra normal bivariant original (Z_1, Z_2) (amb marginals independents $N(0,1)$) mitjançant transformacions senzilles com les descrites abans: canvis d'escala, rotacions i translacions. A continuació es presenta la definició formal de la distribució normal bivariant general. El punt de partida serà la Llei Normal bivariant estàndard amb marginals independents definida abans.

Definició 3.8.2. *Siguin V_1, V_2 v.a. independents $N(0,1)$ i $\rho \in (-1,1)$. Siguin*

$$\left.\begin{array}{rcl} Z_1 & = & V_1, \\ Z_2 & = & \rho V_1 + \sqrt{1-\rho^2} V_2. \end{array}\right\}$$

Direm que (Z_1, Z_2) segueix una distribució normal bivariant estàndard amb marginals correlades de paràmetre ρ

Proposició 3.8.3. *Sigui (Z_1, Z_2) v.a. normal bivariant estàndard amb marginals correlades de paràmetre $\rho \in (-1,1)$. Aleshores la seva funció de densitat conjunta és*

$$f(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\right)$$

on $\rho \in (-1,1)$.

Demostració: Per definició,

$$(Z_1, Z_2) \sim (V_1, \rho V_1 + \sqrt{1-\rho^2} V_2),$$

on V_1 i V_2 són variables aleatorias independents $N(0,1)$. Per tant, $Z_1 \sim N(0,1)$ i la condicionada de Z_2 donada Z_1,

$$\begin{aligned}(Z_2|Z_1 = z_1) \sim & (\rho V_1 + \sqrt{1-\rho^2} V_2 | V_1 = z_1) \\ \sim & (\rho z_1 + \sqrt{1-\rho^2} V_2 | V_1 = z_1) \\ & \text{(V_1 i V_2 són independents)} \\ \sim & \rho z_1 + \sqrt{1-\rho^2} V_2 \\ \sim & N(\rho z_1, 1-\rho^2).\end{aligned}$$

Per tant

$$\begin{aligned} f(z_1, z_2) = & f_{Z_2|Z_1=z_1}(z_2) f_{Z_1}(z_1) \\ = & \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(z_2 - \rho z_1)^2\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_1^2\right) \\ = & \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(z_2^2 - 2\rho z_1 z_2 + \rho^2 z_1^2 + (1-\rho^2) z_1^2)\right) \\ = & \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\right). \end{aligned}$$

□

Observeu que les corbes de nivell de la funció de densitat conjunta de (Z_1, Z_2) són el·lipses centrades al $(0,0)$:

$$f(z_1, z_2) = C \Leftrightarrow z_1^2 - 2\rho z_1 z_2 + z_2^2 = k.$$

Proposició 3.8.4. *Sigui (Z_1, Z_2) normal bivariant estàndard amb marginals correlades de paràmetre ρ.*

1. $\mathbb{E}(Z_1) = \mathbb{E}(Z_2) = 0$.
2. $V(Z_1) = V(Z_2) = 1$.
3. $Cov(Z_1, Z_2) = Cor(Z_1, Z_2) = \rho$.
4. *Les marginals Z_1 i Z_2 són normals univariants estàndards.*
5. *Z_1 i Z_2 són independents si i només si $\rho = 0$.*

Demostració: Per definició,

$$(Z_1, Z_2) \sim (V_1, \rho V_1 + \sqrt{1-\rho^2} V_2),$$

on V_1 i V_2 són variables aleatorias independents $N(0,1)$. Per tant $Z_1 \sim V_1 \sim N(0,1)$, aleshores $\mathbb{E}(Z_1) = 0$, $\mathrm{Var}(Z_1) = 1$.

Per altra banda, $Z_2 \sim \rho V_1 + \sqrt{1-\rho^2} V_2$. Aleshores, Z_2 és suma de dos normals univariants independents i, en conseqüència, és normal (vegeu la Proposició 3.6.3). A més a més,

$$\begin{aligned}
\mathbb{E}(Z_2) =& \rho \mathbb{E}(V_1) + \sqrt{1-\rho^2}\mathbb{E}(V_2) = 0, \\
\mathrm{Var}\,(Z_2) =& \mathrm{Var}\left(\rho V_1 + \sqrt{1-\rho^2} V_2\right) \\
& (V_1 \text{ i } V_2 \text{ són independents, aleshores } \mathrm{Cov}(V_1,V_2){=}0) \\
=& \rho^2 \mathrm{Var}\,(V_1) + (1-\rho^2)\mathrm{Var}\,(V_2) \\
=& \rho^2 + (1-\rho^2) \\
=& 1. \\
\mathrm{Cor}(Z_1, Z_2) =& \mathrm{Cov}(Z_1, Z_2) \\
=& \mathrm{Cov}(V_1, \rho V_1 + \sqrt{1-\rho^2} V_2) \\
=& \rho \, \mathrm{Cov}(V_1, V_1) + \sqrt{1-\rho^2}\, \mathrm{Cov}(V_1, V_2) \\
=& \rho.
\end{aligned}$$

Per tant, $\rho = \mathrm{Cov}(Z_1, Z_2) = 0$ si Z_1 i Z_2 són independents. Recíprocament, si $\rho = 0$ aleshores

$$(Z_1, Z_2) \sim (V_1, V_2)$$

i, per tant, Z_1 i Z_2 són independents. □

Definició 3.8.5 (Distribució normal bivariant)**.** *Siguin Z_1 i Z_2 normals univariants estàndards amb marginals correlades amb correlació $\rho \in (-1, 1)$. Per qualsevol $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}^+$ definim*

$$\left.\begin{array}{rcl} X_1 &=& \mu_1 + \sigma_1 Z_1, \\ X_2 &=& \mu_2 + \sigma_2 Z_2 \end{array}\right\}$$

Aleshores direm que (X_1, X_2) té distribució NORMAL BIVARIANT *amb vector d'esperances (μ_1, μ_2) i matriu de variàncies i covariàncies*

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Escriurem també

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right).$$

Proposició 3.8.6. *Siguin V_1 i V_2 normals univariants estàndards independents. Per qualsevol $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}^+$ i $-1 < \rho < 1$, definim*

$$\left.\begin{array}{rcl} X_1 & = & \mu_1 + \sigma_1 V_1, \\ X_2 & = & \mu_2 + \sigma_2\rho V_1 + \sigma_2\sqrt{1-\rho^2}\, V_2. \end{array}\right\}$$

Aleshores (X_1, X_2) té distribució normal bivariant amb vector d'esperances (μ_1, μ_2) i matriu de variàncies i covariàncies

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

Es demostra fàcilment, definint $Z_1 = V_1$ i $Z_2 = \rho V_1 + \sqrt{1-\rho^2}\, V_2$.

Proposició 3.8.7. *Sigui (X_1, X_2) normal bivariant.*

1. $X_1 \sim N(\mu_1, \sigma_1^2)$ *i* $X_2 \sim N(\mu_2, \sigma_2^2)$.
2. *Cov*$(X_1, X_2) = \rho\sigma_1\sigma_2$, *Cor*$(X_1, X_2) = \rho$.
3. $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$.
4. X_1 *i* X_2 *són independents si i només si* $\rho = 0$.

Demostració: Com que, per a $i = 1, 2$, $Z_i \sim N(0, 1)$ i $X_i = \mu_i + \sigma_i Z_i$, aleshores $X_i \sim N(\mu_i, \sigma_i^2)$.

Per altra banda, si V_1 i V_2 són normals univariants estàndards independents,

$$\begin{aligned} \mathrm{Cov}(X_1, X_2) =& \mathrm{Cov}(\mu_1 + \sigma_1 V_1, \mu_2 + \sigma_2\rho V_1 + \sigma_2\sqrt{1-\rho^2}\, V_2) \\ =& \sigma_1\sigma_2\rho\mathrm{Cov}(V_1, V_1) + \sigma_1\sigma_2\sqrt{1-\rho^2}\mathrm{Cov}(V_1, V_2) \\ =& \rho\sigma_1\sigma_2. \end{aligned}$$

Això implica que $\mathrm{Cor}(X_1, X_2) = \rho$.

Observeu que

$$X_1 + X_2 \sim (\mu_1 + \sigma_1 V_1) + \left(\mu_2 + \sigma_2\rho V_1 + \sigma_2\sqrt{1-\rho^2}\, V_2\right)$$

$$\sim (\mu_1 + \mu_2) + (\sigma_1 + \sigma_2\rho)V_1 + \sigma_2\sqrt{1-\rho^2}V_2$$

que és normal per ser $(\sigma_1+\sigma_2\rho)V_1$ i $\sigma_2\sqrt{1-\rho^2}V_2$ normals independents (vegeu la Proposició 3.6.3).

Si X_1 i X_2 són independents, aleshores la seva correlació ρ és 0. Recíprocament, si $\rho=0$, $X_1=\mu_1+\sigma_1 V_1$, i

$$X_2=\mu_2+\sigma_2\rho V_1+\sigma_2\sqrt{1-\rho^2}\,V_2=\mu_2+\sigma_2 V_2.$$

Per tant X_1 i X_2 són independents, perquè V_1 i V_2 ho són. □

Proposició 3.8.8. *Sigui (X_1,X_2) amb distribució normal bivariant amb vector d'esperances (μ_1,μ_2) i matriu de variàncies i covariàncies*

$$\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}.$$

Aleshores la seva densitat conjunta és

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}Q(x_1,x_2)\right),$$

on Q és la forma quadràtica

$$Q(x_1,x_2)=\frac{1}{(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right].$$

Demostració: Es pot demostrar fent un canvi de variable bivariant. Ho farem, però, d'una altra manera. Calcularem primer la funció de distribució conjunta de (X_1,X_2): sigui $(x_1,x_2)\in\mathbb{R}^2$,

$$\begin{aligned}F_{(X_1,X_2)}(x_1,x_2)&=\Pr(X_1\le x_1,\ X_2\le x_2)\\&=\Pr(\mu_1+\sigma_1 Z_1\le x_1,\ \mu_2+\sigma_2 Z_2\le x_2)\\&=\Pr\left(Z_1\le\frac{x_1-\mu_1}{\sigma_1},Z_2\le\frac{x_2-\mu_2}{\sigma_2}\right)\\&=F_{(Z_1,Z_2)}\left(\frac{x_1-\mu_1}{\sigma_1},\frac{x_2-\mu_2}{\sigma_2}\right)\end{aligned}$$

Prenem ara derivades parcials respecte de x_1 i x_2:

$$\begin{aligned}
f_{(X_1,X_2)}(x_1,x_2) =& \frac{\partial^2}{\partial x_1 \partial x_2} F_{(X_1,X_2)}(x_1,x_2) \\
=& \frac{\partial^2}{\partial x_1 \partial x_2} F_{(Z_1,Z_2)}\left(\frac{x_1-\mu_1}{\sigma_1}, \frac{x_2-\mu_2}{\sigma_2}\right) \\
=& \frac{1}{\sigma_1}\frac{1}{\sigma_2} f_{(Z_1,Z_2)}\left(\frac{x_1-\mu_1}{\sigma_1}, \frac{x_2-\mu_2}{\sigma_2}\right) \\
& \left(f_{(Z_1,Z_2)}(z_1,z_2) = \frac{1}{2\pi\sqrt{1-\rho 2}} \exp\left(-\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\right)\right) \\
=& \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}Q(x_1,x_2)\right).
\end{aligned}$$

□

3.9 Normal multivariant

Proposició 3.9.1. *Si* $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \Sigma)$ *la seva funció de densitat és*

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})},$$

on $\mathbf{x} = (x_1, x_2)^t$, $p = 2$ *és la dimensió de l'espai* $\mathbb{R}^2$ *on està definida la variable aleatoria* $\mathbf{X}$ *i* $|\Sigma|$ *és el determinant de* Σ.

Només cal identificar termes per demostra-ho.

Aquesta expressió en forma vectorial i matricial de la densitat de la normal bivariant, s'estén a normals multivariants de dimensió p qualsevol. Treballarem amb vectors columna $\mathbf{x} \in \mathbb{R}^p$, i $\mathbf{x}'$ indica el vector transposat de $\mathbf{x}$.

Definició 3.9.2. *El vector aleatori* $\mathbf{X} = (X_1, \ldots, X_p)'$ *té* DISTRIBUCIÓ NORMAL MULTIVARIANT, *i escriurem* $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$, *si la seva densitat conjunta és*

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})},\ \mathbf{x} \in \mathbb{R}^p,$$

on $\boldsymbol{\mu} \in \mathbb{R}^p$ *i* Σ *és una matriu simètrica* $p \times p$ *definida positiva (és a dir, per a tot* $\mathbf{a} \neq \mathbf{0}$ *en* $\mathbb{R}^p$, $\mathbf{a}'\Sigma\mathbf{a} > 0$*).*

Donat que Σ és una matriu simètrica $p \times p$ definida positiva, Σ admet una descomposició en funció dels seus valors i vectors propis:

$$\Sigma = \mathbf{C}\Lambda\mathbf{C}'$$

on $\mathbf{C}$ és la matriu $p \times p$ ortonormal que té per columnes els vectors propis de Σ i Λ és la matriu diagonal que té per elements els valors propis $\lambda_1 \geq \cdots \geq \lambda_p > 0$ de Σ: $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$. Aleshores, és fàcil veure que

$$\Sigma^{-1} = \mathbf{C}\Lambda^{-1}\mathbf{C}',\ \Sigma^{1/2} = \mathbf{C}\Lambda^{1/2}\mathbf{C}',\ \Sigma^{-1/2} = \mathbf{C}\Lambda^{-1/2}\mathbf{C}',$$

amb $\Lambda^\alpha = \text{diag}(\lambda_1^\alpha, \ldots, \lambda_p^\alpha)$ per $\alpha = -1, -1/2, 1/2$.

Teorema 3.9.3. *Siguin $\boldsymbol{\mu} \in \mathbb{R}^p$ i Σ una matriu simètrica $p \times p$ definida positiva.*

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma) \Leftrightarrow \Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}_p, I_p) \Leftrightarrow \mathbf{X} = \boldsymbol{\mu} + \Sigma^{1/2}\mathbf{Z},$$

amb $\mathbf{Z} \sim N(\mathbf{0}_p, I_p)$.

La demostració és conseqüència directa de les fórmules del canvi de variable per transformacions lineals bijectives.

Teorema 3.9.4. *Si $\mathbf{X}$ és $N(\boldsymbol{\mu}, \Sigma)$ aleshores*

(a) $\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}$, *això és* $\mathbb{E}(X_i) = \mu_i$ *per tot* i.

(b) $\Sigma = (\sigma_{ij})_{ij}$ *és la matriu de variàncies i covariàncies de* $\mathbf{X}$, *és a dir,* $\sigma_{ij} = Cov(X_i, X_j)$.

La demostració és immediata a partir del Teorema anterior i dels càlculs d'esperances i variàncies en notació matricial.

Teorema 3.9.5. *Sigui $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ i $\mathbf{D}$ una matriu $m \times p$ de rang $m \leq p$. sigui $\mathbf{Y} = \mathbf{D}\mathbf{X}$. Aleshores*

$$\mathbf{Y} \sim N_m(\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\Sigma\mathbf{D}').$$

La demostració és conseqüència de les fórmules del canvi de variable per transformacions lineals bijectives. Abans de poder aplicar-lo, però, s'ha de definir un vector auxiliar $\mathbf{V} = \mathbf{F}\mathbf{X}$, on $\mathbf{F}$ sigui una matriu $(p - m) \times p$ de rang $(p - m)$ tal que la matriu $p \times p$ $\mathbf{M} = (\mathbf{D}', \mathbf{F}')'$ tingui rang complet p i que $\mathbf{D}\Sigma\mathbf{F}' = \mathbf{0}_{m\times(p-m)}$. Torbar una matriu $\mathbf{F}$ que compleixi aquestes propietats sempre és possible. Aleshores, les fórmules del canvi de variable per transformacions lineals bijectives garanteixen que $\mathbf{M}\mathbf{X}$ és normal p-variant, amb sub-vectors $\mathbf{Y}$ i $\mathbf{V}$ independents normals de dimensions m i $(p - m)$, respectivament.

Corol·lari 3.9.6. *Les marginals de qualsevol dimensió d'una normal multivariant també són normals multivariants.*

Corol·lari 3.9.7. *Sigui* $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$ *i siguin* $a_1, \ldots, a_p$ *nombres reals tals que* $\sum_{i=1}^p a_i^2 > 0$. *Aleshores*

$$\sum_{i=1}^p a_i X_i \sim N\left(\sum_{i=1}^p a_i \mu_i, \sum_{i=1}^p a_i^2 \sigma_i^2 + 2\sum_{i<j} a_i a_j \sigma_{ij}\right).$$

Teorema 3.9.8. *Sigui el vector normal multivariant de dimensió* $(p+q)$

$$\begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{pmatrix} \sim N_{p+q}\left(\begin{pmatrix} \boldsymbol{\mu_1} \\ \boldsymbol{\mu_2} \end{pmatrix}, \begin{pmatrix} \Sigma_1^2 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2^2 \end{pmatrix},\right)$$

on $\mathbf{X_1}$ *i* $\boldsymbol{\mu_1}$ *són de dimensió* p, $\mathbf{X_2}$ *i* $\boldsymbol{\mu_2}$ *són de dimensió* q, *i les dimensions de les matrius* Σ_{11}, Σ_{12}, Σ_{21} *i* Σ_{22} *són* $p\times p$, $p\times q$, $q\times p$ *i* $q\times q$, *respectivament.*

Aleshores,

1. $(\mathbf{X_1}|\mathbf{X_2} = \mathbf{x_2}) \sim N_p\left(\boldsymbol{\mu_1} + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x_2} - \boldsymbol{\mu_2}), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$.
2. $\mathbf{X_1}$ *i* $\mathbf{X_2}$ *són independents si i només si* $\Sigma_{12} = \mathbf{0}_{p\times q}$.

La demostració no és immediata. La veurem pel cas bivariant ($p = q = 1$) a la secció següent.

Observeu que $\mathbb{E}(\mathbf{X_1}|\mathbf{X_2} = \mathbf{x_2})$ és lineal en $\mathbf{x_2}$ i que $\mathrm{Var}\,(\mathbf{X_1}|\mathbf{X_2} = \mathbf{x_2})$ no depèn de $\mathbf{x_2}$.

3.10 Predicció en el cas de la normal bivariant

Una de les propietats de la normal bivariant (i també en el cas multivariant) és que totes les seves distribucions condicionades també són normals.

Proposició 3.10.1. *Sigui* (X_1, X_2) *un vector aleatori que segueix una distribució normal bivariant,*

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right).$$

Aleshores la distribució condicional de X_2, *donat* $X_1 = x_1$, *és la distribució normal de mitjana* $\mu_* = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)$ *i de variància* $\sigma_*^2 = \sigma_2^2(1-\rho^2)$. *És a dir,*

$$(X_2|_{X_1=x_1}) \sim N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1-\rho^2)\right).$$

Anàlogament,

$$(X_1|_{X_2=x_2}) \sim N\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2), \sigma_1^2(1-\rho^2)\right).$$

La demostració es fa usualment aplicant la definició de densitat condicionada (quocient de la densitat conjunta entre la densitat marginal de la variable que condiciona) al cas concret de la normal bivariant i manipulant l'expressió resultant. Aquesta manipulació és feixuga. És per això que farem aquí una demostració alternativa.

Demostració: Per la Proposició 3.8.6, sabem que

$$(X_1, X_2) \sim \left(\mu_1 + \sigma_1 V_1,\ \mu_2 + \sigma_2\rho V_1 + \sigma_2\sqrt{1-\rho^2}\, V_2\right),$$

on V_1 i V_2 són normals univariants estàndards independents. Aleshores,

$$\begin{aligned}
(X_2|X_1 = x_1) &\sim \left(\mu_2 + \sigma_2\rho V_1 + \sigma_2\sqrt{1-\rho^2}\, V_2 \middle| \mu_1 + \sigma_1 V_1 = x_1\right) \\
&\quad \text{(es pot provar que això és cert, per ser } \mu_1 + \sigma_1 V_1 \\
&\quad \text{transformació bijectiva derivable de } V_1) \\
&\sim \left(\mu_2 + \sigma_2\rho V_1 + \sigma_2\sqrt{1-\rho^2}\, V_2 \middle| V_1 = \frac{x_1-\mu_1}{\sigma_1}\right) \\
&\sim \left(\mu_2 + \sigma_2\rho\frac{x_1-\mu_1}{\sigma_1} + \sigma_2\sqrt{1-\rho^2}\, V_2 \middle| V_1 = \frac{x_1-\mu_1}{\sigma_1}\right) \\
&\quad (V_1 \text{ i } V_2 \text{ són independents}) \\
&\sim \mu_2 + \sigma_2\rho\frac{x_1-\mu_1}{\sigma_1} + \sigma_2\sqrt{1-\rho^2}\, V_2 \\
&\sim N\left(\mu_2 + \sigma_2\rho\frac{x_1-\mu_1}{\sigma_1}, \sigma_2^2(1-\rho^2)\right).
\end{aligned}$$

□

Emprarem ara la notació (X, Y), en comptes de (X_1, X_2). Ambdues es fan servir indistintament a la pràctica. Sigui (X, Y) normal bivariant,

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}\right).$$

Sigui $\sigma_{XY} = \rho\sigma_X\sigma_Y$. Observeu que

$$\begin{aligned}
r_1(x) = E(Y|X = x) &= \mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x-\mu_X) \\
&= \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2}(x-\mu_X) \\
&= (\mu_Y - \mu_X\frac{\sigma_{XY}}{\sigma_X^2}) + x\frac{\sigma_{XY}}{\sigma_X^2} \\
&= \beta_0 + \beta_1 x,
\end{aligned}$$

on $\beta_1 = \frac{\sigma_{XY}}{\sigma_X^2}$ i $\beta_0 = \mu_Y - \mu_X \beta_1$. Per tant, en el cas de la normal bivariant la FUNCIÓ DE REGRESSIÓ és lineal en x: és una recta que té pendent β_1 i passa pel punt (μ_X, μ_Y):

$$r_1(\mu_X) = E(Y|X = \mu_X) = \mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(\mu_X - \mu_X) = \mu_Y.$$

Per altra banda, la variància condicionada $\mathrm{Var}\,(Y|X = x)$ no depèn del valor x pel qual es condiciona:

$$\mathrm{Var}\,(Y|X = x) = \sigma_2^2(1 - \rho^2).$$

Vegeu Figura 3.10.

Fixarem ara la nostra atenció a la variable aleatòria esperança condicionada de Y donada X:

$$\mathbb{E}(Y|X) = r_1(X) = \mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(X - \mu_X) \sim N\left(\mu_Y, \rho^2\sigma_Y^2\right).$$

Hem fet servir la propietat que diu que $\mathbb{E}(Y|X)$ és normal univariant, per ser una transformació lineal de X, que sabem que és normal. Hem calculat l'esperança i variància de $\mathbb{E}(Y|X)$ fent servir la seva expressió com a transformació lineal de X.

Observeu que es verifica la Llei de l'Esperança Iterada,

$$\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y),$$

i també la Llei de la Variància Iterada,

$$\mathrm{Var}\,(Y) = \sigma_Y^2 = \sigma^2(1 - \rho^2) + \rho^2\sigma^2 = \mathbb{E}(\mathrm{Var}\,(Y|X)) + \mathrm{Var}\,(\mathbb{E}(Y|X)).$$

El segon sumand, $\mathrm{Var}\,(\mathbb{E}(Y|X))$, és la part de la variància de Y que vé explicada per la variable X, mentre que el primer sumand, $\mathbb{E}(\mathrm{Var}\,(Y|X))$, és la part no explicada per X. Podem escriure

$$\text{Var. total de } Y = \text{Var. de } Y \text{ no explicada per } X + \text{Var. de } Y \text{ explicada per } X,$$

o també,

$$\mathrm{VT} = \mathrm{VNE} + \mathrm{VE}.$$

La proporció de variància de Y explicada per X és ρ^2:

$$\frac{\mathrm{VE}}{\mathrm{VT}} = \frac{\sigma_Y^2\rho^2}{\sigma_Y^2} = \rho^2.$$

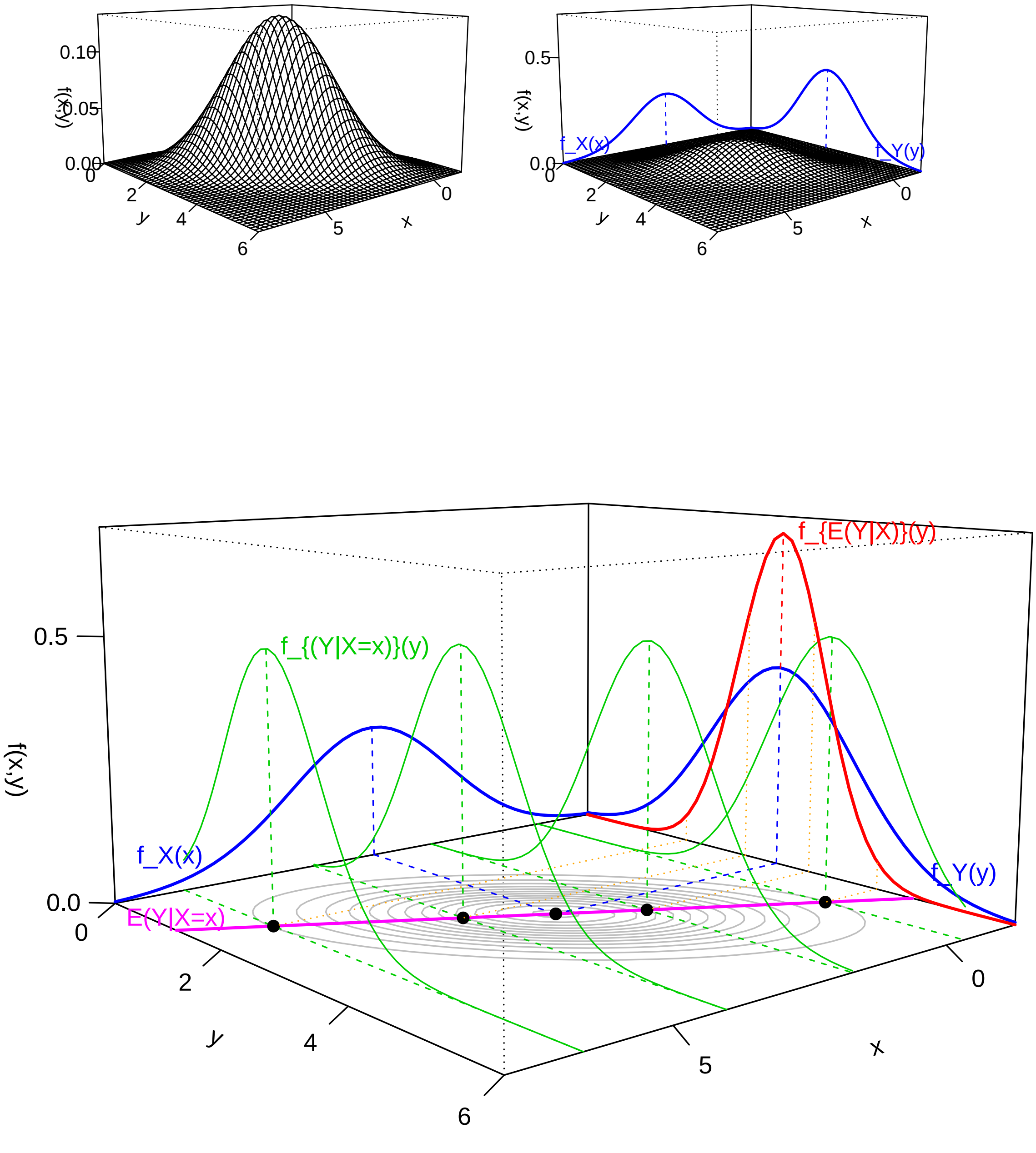


Figura 3.10: (X, Y) normal bivariant. A dalt, esquerra: densitat conjunta. A dalt, dreta: densitat conjunta i densitats marginals. Abaix, marginals, condicionades i densitat de $\mathbb{E}(Y|X)$.

Aquesta és la versió poblacional (estem parlant d'una variable aleatòria bivariant (X, Y)) del coeficient de determinació R^2 que es defineix quan s'ajusta una recta de regressió a un conjunt de dades bivariants (que suposadament són una mostra de (X, Y)). Es pot provar que de fet R^2 és el quadrat del coeficient de correlació mostral de les dades.

Exemple 3.10.1.

Suposem que X_1 representa l'alçada i X_2 el pes d'una determinada persona. Per experiències prèvies sabem que la llei conjunta de (X_1, X_2) és normal bivariant. Estem interessats en predir el pes X_2 d'una persona en dues situacions diferents:

1. *Sense tenir més informació:* prediem un pes μ_2 amb variància σ_2^2.

2. *Si coneixem l'alçada d'aquesta persona $X_1 = x_1$:* llavors el pes predit serà $E(X_2|X_1 = x_1) = \mu_2 + \sigma_2\rho\frac{x_1-\mu_1}{\sigma_1}$ amb variància $\sigma_2^2(1-\rho^2) < \sigma_2^2$ (fixem-nos que la variància no depèn de x_1, per tant, es comet el mateix error predient el pes de persones baixes i altes).

Resoldrem el problema suposant que (X_1, X_2) segueixen una llei normal bivariant amb els següents valors:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 159 \\ 55 \end{pmatrix}, \begin{pmatrix} 3 & 1.3856 \\ 1.3856 & 4 \end{pmatrix} \right).$$

Fixem-nos que

$$\sigma_1^2 = 3,\ \sigma_2^2 = 4 \text{ i } \rho\, \sigma_1\, \sigma_2 = 1.3856.$$

Per tant:

$$\sigma_1 = 1.73,\ \sigma_2 = 2 \text{ i } \rho = 0.4.$$

Resolem ara les qüestions inicials:

1. El pes predit és 55 kg. amb variància 4.

2. Suposant que sabem que l'alçada de la persona és 165 cm. llavors el pes predit és $55+2\ (0.4)\ \frac{165-159}{\sqrt{3}} = 58.08$ amb variància $4\ (1-(0.4)^2) = 3.36$.

$\square$

3.11 Distribucions relacionades amb la normal

Definició 3.11.1. *Siguin $X_1, \dots, X_n$ v.a.i.i.d. $N(\mu, \sigma^2)$. Es defineix la* MITJANA MOSTRAL

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

i la VARIÀNCIA MOSTRAL

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

Aquests ESTADÍSTICS seran ESTIMADORS de μ i σ^2, respectivament.

Proposició 3.11.2.

$$\bar{X} \sim N(\mu, \sigma^2/n), \quad \mathbb{E}(\bar{X}) = \mu, \quad \mathit{Var}(\bar{X}) = \frac{\sigma^2}{n}, \quad \mathbb{E}(S^2) = \sigma^2.$$

La demostració de tots els punts, excepte potser l'últim, és fàcil. Provem l'últim punt:

$$\begin{aligned}
\mathbb{E}(S^2) =& \frac{1}{n-1} \mathbb{E}\left(\sum_{i=1}^{n} (X_i - \bar{X})^2 \right) \\
=& \frac{n}{n-1} \mathbb{E}\left((X_1 - \bar{X})^2 \right) \\
& (\mathbb{E}(X_1 - \bar{X}) = 0) \\
=& \frac{n}{n-1} \mathrm{Var}\left((X_1 - \bar{X})^2 \right) \\
=& \frac{n}{n-1} \left(\mathrm{Var}(X_1) + \mathrm{Var}(\bar{X}) - 2\mathrm{Cov}(X_1, \bar{X}) \right) \\
& \left(\mathrm{Cov}(X_1, \bar{X}) = \frac{1}{n} \mathrm{Cov}(X_1, \sum_{j=1}^{n} X_j) = \frac{1}{n} \mathrm{Cov}(X_1, X_1) = \frac{\sigma^2}{n} \right) \\
=& \frac{n}{n-1} \left(\sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n} \right) \\
=& \frac{n}{n-1} \frac{(n-1)\sigma^2}{n} \\
=& \sigma^2.
\end{aligned}$$

La primera de les distribucions relacionades amb la normal que definirem és la χ^2 (llegim *txi quadrat*).

Definició 3.11.3. *Siguin $X \sim N(0,1)$ i sigui $Y = X^2$. Direm que Y segueix una distribució χ_1^2, que llegirem* TXI QUADRAT AMB UN GRAU DE LLIBERTAT.

Siguin $X_1, \ldots, X_n$ v.a.i.i.d. $N(0,1)$ i sigui $Y_n = \sum_{i=1}^n X_i^2$. Direm que Y_n segueix una distribució χ_n^2, que llegirem TXI QUADRAT AMB n GRAUS DE LLIBERTAT.

Veurem que la distribució χ_1^2 és un cas particular de la llei Gamma. Recordeu que una variable aleatòria X segueix una LLEI GAMMA de paràmetres $\alpha > 0$ i $\beta > 0$ si X té una funció de densitat definida per

$$f(x|\alpha,\beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{si } x > 0, \\ 0 & \text{altrament,} \end{cases}$$

on α i β són els anomenats paràmetres de forma i d'escala respectivament. Escrivim que $X \sim \gamma(\alpha,\beta)$.

En determinats llibres podeu trobar una altra parametrització d'aquesta distribució, que s'obté definint $b = 1/\beta$:

$$f(x|\alpha,b) = \begin{cases} \frac{1}{\Gamma(\alpha) b^\alpha} x^{\alpha-1} e^{-\frac{x}{b}} & \text{si } x > 0, \\ 0 & \text{altrament.} \end{cases}$$

Proposició 3.11.4. $Y \sim \chi_1^2 \Longleftrightarrow Y \sim \gamma(1/2, 1/2)$.

Demostració: Suposem $X \sim \mathrm{N}(0,1)$ i definim $Y = X^2 \sim \chi_1^2$. Llavors

$$F_Y(y) = P(Y \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y}),$$

on Φ és la funció de distribució de la $N(0,1)$. Aleshores, la funció de densitat de Y serà

$$f_Y(y) = F_Y'(y) = \varphi(\sqrt{y})\frac{1}{2\sqrt{y}} + \varphi(-\sqrt{y})\frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{y}}\varphi(\sqrt{y}) = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}},$$

on φ és la funció de densitat de la $N(0,1)$.

Aleshores, prenent $\alpha = \beta = 1/2$, hem provat que la llei de Y és una $\gamma(1/2, 1/2)$ □

Observeu que, a la demostració anterior, igualant les constants tenim que

$$\frac{\beta^\alpha}{\Gamma(\alpha)} = \frac{1}{\sqrt{2\pi}} \Rightarrow \Gamma(1/2) = \sqrt{\pi}.$$

Proposició 3.11.5. $Y_n \sim \chi_n^2 \Longleftrightarrow Y \sim \gamma(n/2, 1/2)$.

Demostració: Sabem que $Y_n = \sum_{i=1}^n X_i^2$, amb $X_1, \ldots, X_n$ v.a.i.i.d. $N(0,1)$. Per tant Y_n és la suma de n v.a.i.i.d distribuïdes com una $\chi_1^2 \equiv \gamma(1/2, 1/2)$. Fent servir la propietat que diu que la suma de Gammas independents (amb el mateix paràmetre d'escala β) és també gamma (amb paràmetre de forma α igual a la suma dels paràmetres de forma; hem provat aquest resultat a l'exemple 3.6.2, i en donarem una prova alternativa al Tema **??**, fent servir la funció generadora de moments), es té

$$Y_n = \sum_{i=1}^n X_i^2 \sim \gamma(n/2, 1/2) \equiv \chi_n^2.$$

□

Les següents proposicions garanteixen que certes formes quadràtiques definides a partir de normals multivariants tenen una distribució txi quadrat.

Proposició 3.11.6. *Sigui* $\mathbf{Z} = (Z_1, \ldots, Z_p)^t$ *variable aleatoria p-dimensional amb distribució normal estàndard,*

$$\mathbf{Z} \sim N_p(\mathbf{0}, I_p).$$

Aleshores

$$\mathbf{Z}^t\mathbf{Z} = Z_1^2 + \cdots + Z_p^2 \sim \chi_p^2.$$

La demostració és segueix directament de la definició de la χ_p^2.

Proposició 3.11.7. *Sigui* $\mathbf{X}$ *variable aleatoria normal p-dimensional,*

$$\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma).$$

Aleshores

$$(\mathbf{X} - \boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_p^2.$$

Demostració: Sigui

$$\mathbf{Z} = \Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu}).$$

Aleshores $\mathbf{Z} \sim N_p(\mathbf{0}, I_p)$. Per altra banda

$$(\mathbf{X} - \boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) = \left((\mathbf{X} - \boldsymbol{\mu})^t \Sigma^{-1/2}\right) \left(\Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu})\right) = \mathbf{Z}^t \boldsymbol{Z}$$

i aplicant la proposició anterior tenim el resultat provat. □

Ara podem enunciar el Teorema de Fisher:

Teorema 3.11.8 (Teorema de Fisher)**.** . *Si* $X_1, \ldots, X_n$ *són v.a.i.i.d.* $N(\mu, \sigma^2)$ *aleshores* $\bar{X}$ *i* S^2 *són independents. A més*

$$\bar{X} \sim N(\mu, \sigma^2/n), \ \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Demostració:
Que $\bar{X} \sim N(\mu, \sigma^2/n)$ es segueix de les propietats de les combinacions lineals de normals independents.

Veurem ara que $\bar{X}$ i S^2 són independents. Observeu que

$$\sum_{i=1}^{n}(X_i - \bar{X}) = 0 \Rightarrow X_n - \bar{X} = -\sum_{i=1}^{n-1}(X_i - \bar{X}),$$

i per tant

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n-1}(X_i - \bar{X})^2 + \frac{1}{n-1}\left(\sum_{i=1}^{n-1}(X_i - \bar{X})\right)^2,$$

d'on es segueix que S^2 és funció de $(X_1 - \bar{X}, \ldots, X_{n-1} - \bar{X})$. Aleshores, per demostrar que $\bar{X}$ i S^2 són independents el que farem serà provar que $\bar{X}$ i $(X_1 - \bar{X}, \ldots, X_{n-1} - \bar{X})$ són independents. Com que sabem que la seva correlació és zero ($\mathrm{Cov}(\bar{X}, X_i - \bar{X}) = 0$, $i = 1, \ldots, n-1$) el que farem serà provar que conjuntament tenen distribució normal, ja que sota normalitat conjunta incorrelació és equivalent a independència.

Observeu que

$$\begin{pmatrix} \bar{X} \\ X_1 - \bar{X} \\ \vdots \\ X_{n-1} - \bar{X} \end{pmatrix} = \begin{pmatrix} 1/n & 1/n & \cdots & 1/n & 1/n \\ 1-1/n & -1/n & \cdots & -1/n & -1/n \\ 1/n & 1-1/n & \cdots & -1/n & -1/n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -1/n & -1/n & \cdots & 1-1/n & -1/n \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-1} \\ X_n \end{pmatrix}$$

$$= \left(\begin{array}{c:c} \frac{1}{n}\mathbf{1}'_{n-1} & \frac{1}{n} \\ \hdashline \mathbf{I}_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1} & -\frac{1}{n}\mathbf{1}_{n-1} \end{array}\right) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \mathbf{AX},$$

on $\mathbf{1}_k$ és el vector columna de dimensió k amb tots els elements iguals a 1, $\mathbf{I}_k$ és la matriu identitat $k \times k$,

$$\mathbf{A} = \left(\begin{array}{c:c} \frac{1}{n}\mathbf{1}'_{n-1} & \frac{1}{n} \\ \hdashline \mathbf{I}_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1} & -\frac{1}{n}\mathbf{1}_{n-1} \end{array}\right)$$

i $\mathbf{X} = (X_1, \ldots, X_n)'$. Vegem que $\mathbf{A}$, matriu $n \times n$, és de rang complet:

$$\mathrm{rg}(\mathbf{A}) = \mathrm{rg}\left(\begin{array}{cccc|c} 1 & 1 & \cdots & 1 & 1 \\ \hline 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right) = \mathrm{rg}\left(\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \\ \hline 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right) = \mathrm{rg}\,(\mathbf{I}_n) = n.$$

A la primera igualtat, primer hem sumat la primera fila a les altres i després hem multiplicat per n. A la segona igualtat hem restat l'última columna a les altres. A la tercera igualtat hem mogut l'última columna al davant.

Donat que $\mathbf{A}$ és de rang complet, $(\bar{X}, X_1 - \bar{X}, \ldots, X_{n-1} - \bar{X})' = \mathbf{AX}$ és normal multivariant, com volíem provar.

Finalment hem de provar que $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. Observeu que $(X_1 - \bar{X}, \ldots, X_{n-1} - \bar{X})')$ és normal multivariant de dimensió $n-1$, per ser un subconjunt de coordenades d'una normal n-dimensional. Serà doncs

$$\begin{pmatrix} X_1 - \bar{X} \\ \vdots \\ X_{n-1} - \bar{X} \end{pmatrix} = \left(\begin{array}{c|c} \mathbf{I}_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1} & -\frac{1}{n}\mathbf{1}_{n-1} \end{array}\right) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$= \mathbf{BX} \sim N_{n-1}\left(\mathbf{0}, \sigma^2\mathbf{BB}'\right).$$

Per altra banda,

$$\frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n-1}(X_i - \bar{X})^2 + \left(\sum_{i=1}^{n-1}(X_i - \bar{X})\right)^2\right)$$

$$= \frac{1}{\sigma^2}\left((X_1 - \bar{X}, \ldots, X_{n-1} - \bar{X}) \begin{pmatrix} X_1 - \bar{X} \\ \vdots \\ X_{n-1} - \bar{X} \end{pmatrix}\right.$$

$$\left. + (X_1 - \bar{X}, \ldots, X_{n-1} - \bar{X})\mathbf{1}_{n-1}\mathbf{1}'_{n-1} \begin{pmatrix} X_1 - \bar{X} \\ \vdots \\ X_{n-1} - \bar{X} \end{pmatrix}\right)$$

$$= \frac{1}{\sigma^2}(X_1 - \bar{X}, \ldots, X_{n-1} - \bar{X})\left(\mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}'_{n-1}\right) \begin{pmatrix} X_1 - \bar{X} \\ \vdots \\ X_{n-1} - \bar{X} \end{pmatrix}$$

Sigui $\mathbf{C} = \mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}'_{n-1}$. Per la Proposició 3.11.7, basta provar que $(1/\sigma^2)\mathbf{C}$ és la matriu inversa de $\sigma^2\mathbf{BB}'$, o de forma equivalent, que

$$\mathbf{CBB}' = \mathbf{I}_{n-1}, \; = \mathbf{BB}'\mathbf{C} = \mathbf{I}_{n-1}.$$

Com que $\mathbf{C}$ és simètrica, és suficient provar que $\mathbf{CBB}' = \mathbf{I}_{n-1}$. Calculem primer

$$\mathbf{BB}' = \left(\begin{array}{c|c} \mathbf{I}_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1} & -\frac{1}{n}\mathbf{1}_{n-1} \end{array} \right) \left(\begin{array}{c} \mathbf{I}_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1} \\ \hline -\frac{1}{n}\mathbf{1}'_{n-1} \end{array} \right)$$

$$\left(\mathbf{I}_{n-1} - \frac{2}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1} + \frac{n-1}{n^2}\mathbf{1}_{n-1}\mathbf{1}'_{n-1}\right) + \frac{1}{n^2}\mathbf{1}_{n-1}\mathbf{1}'_{n-1}$$

$$= \mathbf{I}_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1},$$

i ara

$$\mathbf{CBB}' = \left(\mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}'_{n-1}\right)\left(\mathbf{I}_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1}\right)$$

$$= \mathbf{I}_{n-1} + \mathbf{1}_{n-1}\mathbf{1}'_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1} - \frac{n-1}{n}\mathbf{1}_{n-1}\mathbf{1}'_{n-1}$$

$$= \mathbf{I}_{n-1} + \left(1 - \frac{1}{n} - \frac{n-1}{n}\right)\mathbf{1}_{n-1}\mathbf{1}'_{n-1} = \mathbf{I}_{n-1}.$$

□

Definició 3.11.9. *Siguin $X \sim N(0,1)$ i $Y \sim \chi^2_r$ independents. Sigui*

$$T = \frac{X}{\sqrt{Y/r}}.$$

Es diu que T segueix una distribució t de Student amb r gruas de llibertat i s'escriu $T \sim t(r)$ o $T \sim t_r$.

Proposició 3.11.10. *Si T és una t de Student amb r gruas de llibertat la seva funció de densitat és*

$$f_T(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)}\left(1 + \frac{t^2}{r}\right)^{-(r+1)/2}, \; -\infty < t < \infty.$$

Demostració: La demostració es fa aplicant un canvi de variable i calculant la densitat d'una marginal. Siguin $X \sim N(0,1)$ i $Y \sim \chi_r^2$ independents.

Fem el canvi de variable bidimensional de (X,Y) a (T,V), on

$$T = \frac{X}{\sqrt{Y/r}}, \; V = Y.$$

La transformació directa és $t = x/\sqrt{y/r}$, $v = y$, i la transformació inversa és

$$\left.\begin{array}{l} x = t\frac{v^{1/2}}{r^{1/2}} \\ y = v \end{array}\right\}$$

El determinant Jacobià del canvi invers és

$$J(t,v) = \begin{vmatrix} \frac{v^{1/2}}{r^{1/2}} & \frac{t}{r^{1/2}}\frac{1}{2}v^{-1/2} \\ 0 & 1 \end{vmatrix} = \frac{v^{1/2}}{r^{1/2}}.$$

Apliquem ara la fórmula del canvi de variable: per a $t \in \mathbb{R}$ i $v > 0$,

$$f_{T,V}(t,v) = f_{X,Y}\left(t\frac{v^{1/2}}{r^{1/2}}, v\right)\frac{v^{1/2}}{r^{1/2}} = f_X\left(t\frac{v^{1/2}}{r^{1/2}}\right) f_Y(v)\frac{v^{1/2}}{r^{1/2}}$$

$$= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\frac{v}{r}t^2}\frac{(1/2)^{r/2}}{\Gamma(r/2)}v^{\frac{r}{2}-1}e^{-\frac{1}{2}v}\frac{v^{1/2}}{r^{1/2}} = \frac{1}{\sqrt{\pi r}}\frac{(1/2)^{(r+1)/2}}{\Gamma(r/2)}v^{\frac{r+1}{2}-1}e^{-\frac{1}{2}\left(1+\frac{t^2}{r}\right)v}.$$

Ara calculem la marginal de T. Observeu que, com a funció de v, l'expressió de $f_{T,V}(t,v)$ recorda a la de la funció de densitat d'una gamma amb paràmetres $(r+1)/2$ i $(1/2)(1+(t^2/r))$, només falta ajustar les constants:

$$f_T(t) = \int_0^\infty f_{T,V}(t,v)dv$$

$$= \frac{1}{\sqrt{\pi r}}\frac{\Gamma((r+1)/2)}{\Gamma(r/2)}\left(1+\frac{t^2}{r}\right)^{-\frac{r+1}{2}}\int_0^\infty \frac{(1/2)^{\frac{r+1}{2}}\left(1+\frac{t^2}{r}\right)^{\frac{r+1}{2}}}{\Gamma((r+1)/2)}v^{\frac{r+1}{2}-1}e^{-\frac{1}{2}\left(1+\frac{t^2}{r}\right)v}dv$$

$$= \frac{1}{\sqrt{\pi r}}\frac{\Gamma((r+1)/2)}{\Gamma(r/2)}\left(1+\frac{t^2}{r}\right)^{-\frac{r+1}{2}}.$$

□

Proposició 3.11.11. *Sigui T una t de Student amb r graus de llibertat. Aleshores*

$$\mathbb{E}(T) = 0 \text{ si } r > 1$$

(per $r \in (0,1]$ la esperança no existeix) i

$$Var(T) = \frac{r}{r-2} \text{ si } r > 2$$

(per $r \in (1,2]$ la variància és infinita i per $r \in (0,1]$ no existeix).

La demostració es fa integrant. No és immediata.

Corol·lari 3.11.12. *Siguin* $X_1, \ldots, X_n$ *v.a.i.i.d.* $N(\mu, \sigma^2)$ *i* $\bar{X}$ *i* S^2 *definits més amunt. Aleshores*

$$\frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1).$$

Definició 3.11.13. *Siguin* $X \sim \chi_r^2$ *i* $Y \sim \chi_s^2$ *independents. Sigui*

$$F = \frac{X/r}{Y/s}.$$

Es diu que F *segueix una distribució* F *amb* r *i* s *gruas de llibertat, i s'escriu* $F \sim F(r,s)$ *o* $F \sim F_{r,s}$.

Proposició 3.11.14. *Sigui* F *amb distribució* F *amb* r *i* s *gruas de llibertat. La seva funció de densitat és*

$$f(x) = \frac{r\Gamma((r+s)/2)}{s\Gamma(r/2)\Gamma(s/2)} \frac{(rx/s)^{(r/2)-1}}{[1+(rx/s)]^{(r+s)/2}}, \; x > 0.$$

La demostració es fa aplicant un canvi de variable i canculant la densitat d'una marginal. No és immediata.

Proposició 3.11.15. $F^{-1} \sim F(s,r)$ *i* $T^2 \sim F(1,r)$ *si* $T \sim t(r)$.

Es demostra fàcilment a partir de la definició de la distribució F.

Proposició 3.11.16. *Siguin* $X_1, \ldots, X_n$ *v.a.i.i.d.* $N(\mu_1, \sigma_1^2)$, *i* $Y_1, \ldots, Y_m$ *v.a.i.i.d.* $N(\mu_2, \sigma_2^2)$, *dues mostres independents de dues distribucions normals. Siguin* S_X^2 *i* S_Y^2 *les variàncies mostrals calculades a partir d'aquestes mostres. Aleshores*

$$\frac{S_X^2/\sigma_1^2}{S_Y^2/\sigma_2^2} \sim F(n-1, m-1).$$

Es demostra fàcilment a partir de la definició de la distribució F.

Proposició 3.11.17. *Sigui* F *amb distribució* F *amb* r *i* s *graus de llibertat. Aleshores*

$$\mathbb{E}(F) = \frac{s}{s-2} \text{ si } s > 2$$

(per $s \in (0,2]$ *la esperança és infinita) i*

$$Var(F) = \frac{2s^2(r+s-2)}{r(s-2)^2(s-4)} \text{ si } s > 4$$

(per $s \in (2,4]$ *la variància és infinita i per* $s \in (0,2]$ *no existeix).*

La demostració es fa integrant. No és immediata.