Módulo 4

Probabilidad Multivariable

Breve anuncio de los temas — Probabilidad Multivariable

- → Densidad conjunta y marginales, esperanza, covarianza
- → Densidad condicional → para simular secuencialmente v.aleatorias
  (X,Y) $\quad Y|_{X=x}$ — Primero simulan X luego Y (basados en el valor de $X=x$.
- → Transformaciones de vectores → aprovechar para simular v.a. normales.
- → Noción de independencia → Algunas transformaciones de v.a. normales, resultan en vectores con v.a. independ.!
- → Multinomiales y concepto de convolución
- → Estadísticos de orden
- → Cópulas (simuladas con densidad condic)

→ Fundamentos otras materias

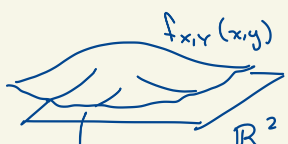
Nociones Básicas vectores de variables continuas

Para 'conocer' las variables en (X,Y) usamos la densidad conjunta

$$f_{X,Y}(x,y) \geq 0$$

con la propiedad de que

$$\iint_{\mathbb{R}^2} f_{X,Y}(x,y)\,dx\,dy = 1$$



volumen debajo de la sábana sea 1.

Si $X \sim Nor(0,1)$

$$\Downarrow$$

$$\mathbb{P}(X \in A) = \int_A f(x)\,dx$$

El Soporte de (X,Y) es

$$\{(x,y) : f_{X,Y}(x,y) > 0\}$$

$$\iint_{\mathbb{R}^2} f_{X,Y}(x,y)\,dx\,dy = 1$$

volumen debajo de la sábana sea 1.

En general, nos interesa obtener densidades 'individuales'

Marginal de X se obtiene de

$$\mathbb{P}(X \leq x) = \int_{-\infty}^{x} \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,dt}_{\text{densidad marginal}}\,ds = \int_{-\infty}^{x} f_X(s)\,ds$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(s,t)\,dt \qquad \text{marginal de } X$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(s,t)\,ds \qquad \text{marginal de } Y$$

Si tenemos la densidad $f_{X,Y,Z}$ de (X,Y,Z)

marginal de Y: $\quad f_Y(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(s,y,w)\,ds\,dw$

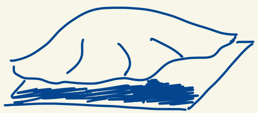
marginal de X,Y: $\quad f_{X,Y}(x,y) = \int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,w)\,dw$

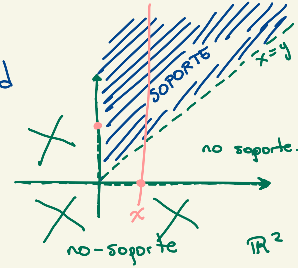
$$\mathbb{P}(Y \in (-1,1)) = \int_{-1}^{1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(s,t,w)\,ds\,dw\,dt$$

EJEMPLO:

Tenemos el vector (X,Y) con densidad

$$f_{X,Y}(x,y) = \begin{cases} e^{-y} & 0<x<y \\ 0 & \text{otro caso} \end{cases}$$



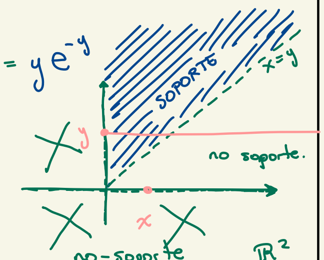


$x = \frac{1}{2}y$ pendiente 2

$y = 2x$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,t)\,dt = \int_x^{\infty} e^{-t}\,dt = -e^{-t}\Big|_{t=x}^{\infty} = e^{-x}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(s,y)\,ds = \int_0^y e^{-y}\,ds = y e^{-y}$$



Si definimos $Z = X + Y$

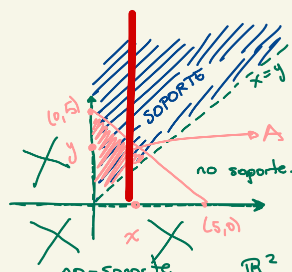
podemos calcular la distribución de Z

$$\mathbb{P}(Z \le 5) = \mathbb{P}((X,Y) \in A)$$

donde

$$A = \{(x,y) \in \text{soporte} : x+y \le 5\}$$

$$\mathbb{P}(Z \le 5) = \int_0^{5/2}\int_x^{5-x} e^{-t}\,dt\,ds$$



$$\mathbb{P}(Z \le 5) = \int_0^{5/2}\int_s^{5-s} e^{-t}\,dt\,ds = \int_0^{5/2} -e^{-t}\Big|_{t=s}^{5-s} ds = \int_0^{5/2} -e^{-(5-s)} + e^{-s}\,ds$$

$$= e^{-5} - 2e^{-5/2} + 1$$

$$\mathbb{P}(Z \le z) = e^{-z} - 2e^{-z/2} + 1$$

Nota: A partir de las densidades conjuntas e.g. de (X,Y) podemos recuperar densidades marginales y también podemos construir variables nuevas

e.g. $X+Y$, X/Y, $\frac{5X^2}{\sqrt{Y}}$
$X-Y$,

El jueves terminamos Def. Esperanza y Covarianza

$$\mathbb{E}[h(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\cdot f_{X,Y}(x,y)\,dx\,dy$$

$\mathrm{Cov}(X,Y)$ está definida para coincidir con $\mathrm{Var}(X)$
$\downarrow$
$\mathrm{cov}(X,X)$

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] \underset{\uparrow \text{ ejercicio}}{=} \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Término de estadística:

Correlación $\quad \mathrm{cor}(X,Y) = \rho_{X,Y} = \dfrac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}} \in [-1,1]$

Útil: Si $X \perp Y$ (independiente) entonces $\mathrm{cor}(X,Y) = 0$

CUIDADO: Hay variables con $\mathrm{cor}(X,Y)=0$ pero X, Y no son independ.

Que nos indica la correlación:

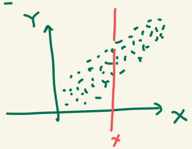
$$\mathrm{cor}(X,Y) = \frac{\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \qquad \mathrm{cov}(X,X) = \mathrm{Var}(X)$$

Nos da una indicación de dependencias lineales:

Dos casos específicos, si $a>0$ $b\in\mathbb{R}$

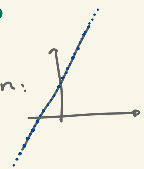
$$\left.\begin{array}{ll} Y = aX+b & \Rightarrow \mathrm{cor}(X,Y) = 1 \\ Y = -aX+b & \quad\ \ \mathrm{cor}(X,Y) = -1 \end{array}\right\} \text{ejercicio}$$

Ejemplo: Si tomamos 1000 muestras de (X,Y) y graficamos



entonces encontraremos que $\mathrm{cor}(X,Y) > 0$

Si $Y = 2X+5$ las muestras se observan:



Cuando $\mathrm{cor}(X,Y) = 0$ se puede ver:

Correlación no nula tampoco significa Causalidad

→ Un estudio encuentra que gran # corredores tienen problemas en la piel.

✗: Correr cause problemas en la piel!

✓: Hay una variable subyacente que afecta las observación:

Tiempo de luz diaria → anima a deportistas a correr
Tiempo de luz diaria → causa lesiones en la piel.

Coeficiente Multinomial: n experimientos, m resultados posibles.

Cuenta: ___ ___ _ _ _ _ _ _ donde cada resultado aparece cierto número de veces.

$$\binom{n}{n_1,n_2,\dots n_m} = \frac{n!}{n_1!\,n_2!\cdots n_m!} \qquad \text{donde } n_1+n_2+n_3\cdots+n_m = n$$

$$= \binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-n_2\cdots-n_{m-1}}{n_m} = \frac{n!}{n_1!\,\cancel{(n-n_1)!}}\cdot\frac{\cancel{(n-n_1)!}}{n_2!\,\cancel{(n-n_1-n_2)!}}\cdot\frac{\cancel{\cdots}}{n_3!}\cdots$$

Vector Multinomial

Es la descripción del resultado de n experimentos donde cada experimiento tiene k resultados posibles.

Si uno mira el espacio muestral más amplio tenemos que cada caso está codificado por

$(B_1, B_2, \dots B_n)$ donde $B_i = j$ es el resultado del i-ésimo experimiento

Si lanzamos 6 monedas entonces $(B_1,\dots B_6) = (A,S,S,A,S,A) = (1,2,2,1,2,1)$

Si lanzamos 4 dados entonces $(B_1,\dots B_4) = (3,3,5,1)$

Lo que nos interesa es saber cuántos resultados de cada tipo.

Vector Multinomial $(X_1, X_2, X_3, \dots X_k)$ se define como

X_i = # veces que el experimiento resultó en i

Vector Multinomial $(X_1, X_2, X_3, \dots X_k)$ se define como

$X_i = \#$ veces que el experimiento resultó en i

Nota: Tenemos $X_1 + X_2 + X_3 \dots + X_k = n$

En el caso de volados: $(X_1, X_2) \rightarrow$ $X_1 = \#$ Aguilas $X_2 = n - X_1$
$X_2 = \#$ Soles.

Si suponemos que $\mathbb{P}(B_i = j) = p_j$ $j \in \{1, 2, \dots, m\}$

$$\mathbb{P}(X_1 = n_1, X_2 = n_2 \dots X_k = n_k) = \binom{n}{n_1, n_2, \dots, n_k} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

cuenta todas las formas en que pueden 'alternarse' los resultados de $(B_1, B_2, \dots, B_n)$ de modo que hay n_i resultados tipo i, $i \in \{1, \dots k\}$

Ejemplo y pregunta:

Si (X_1, X_2, X_3, X_4) es multinomial $(n, .25, .15, .35, .25)$

Que distribución tiene (Y_1, Y_2, Y_3) dado por $Y_1 = X_1$, $Y_2 = X_2$
$Y_3 = X_3 + X_4$

$\rightarrow$ Multinomial $(n, .25, .15, .60)$ "para el nuevo tipo, su probabilidad de ocurrir es la suma de las prob. de 3, 4."

Para demostrar que sí es cierta esta conjetura, verificamos

$\mathbb{P}(Y_1 = k_1, Y_2 = k_2, Y_3 = k_3) =$ fórmula de multinomial cuando $k_1 + k_2 + k_3 = n$

$$\mathbb{P}(Y_1 = k_1, Y_2 = k_2, Y_3 = k_3) = \mathbb{P}(X_1 = k_1, X_2 = k_2, X_3 + X_4 = k_3)$$

$$= \sum_{l=0}^{k_3} \mathbb{P}(X_1 = k_1, X_2 = k_2, X_3 = l, X_4 = k_3 - l)$$

* $k_1 + k_2 + k_3 = n$

$$= \sum_{l=0}^{k_3} \binom{n}{k_1, k_2, l, k_3 - l} p_1^{k_1} p_2^{k_2} p_3^{l} p_4^{k_3 - l}$$

$$\mathbb{P}(Bin(n,p) = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= p_1^{k_1} p_2^{k_2} \sum_{l=0}^{k_3} \frac{n!}{k_1! \, k_2! \, l! \, k_3 - l!} p_3^{l} p_4^{k_3 - l}$$

$$= p_1^{k_1} \cdot p_2^{k_2} \frac{n!}{k_1! k_2! k_3!} \sum_{l=0}^{k_3} \frac{k_3!}{l! \, k_3 - l!} p_3^{l} p_4^{k_3 - l}$$

$$= p_1^{k_1} p_2^{k_2} \frac{n!}{k_1! k_2! k_3!} (p_3 + p_4)^{k_3}$$

$$= \binom{n}{k_1, k_2, k_3} p_1^{k_1} p_2^{k_2} (p_3 + p_4)^{k_3}$$

Comprobamos que $(Y_1, Y_2, Y_3) \sim$ Multinom $(n, p_1, p_2, p_3 + p_4)$

CUIDADO! Recuerden que todo el "experimento completo" se codifica en $\underbrace{(B_1, B_2, \dots, B_n)}$ experimentos independientes idénticos.

Pero como las X_i son conteos

entonces $X_1 + X_2 + \dots + X_k = n$ $\leftarrow$ Que es una dependencia

En general, la convolución se utiliza al manejar sumas de variables aleatorias

Si X, Y son discretas $(\in \mathbb{Z})$ entonces $Z = X+Y$ tiene función de prob. de masa

$$\mathbb{P}(Z = k) = \sum_{j\in\mathbb{Z}} \mathbb{P}(X = j, Y = k-j)$$

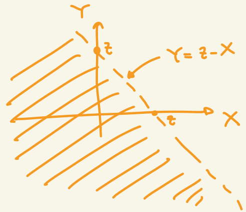
Si X, Y son continuas, con densidad conjunta $f_{X,Y}(x,y)$

entonces $Z = X+Y$ tiene densidad

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(s, z-s)\, ds$$

Demostración corta:

$$\mathbb{P}(X+Y \le z) = \int_{-\infty}^{z}\int_{-\infty}^{\infty} f_{X,Y}(s, t-s)\, ds\, dt$$



$$= \int_{-\infty}^{z} f_Z(u)\, du$$

Transformaciones de vectores aleatorios.

Suponemos que tenemos un vector (X,Y) con densidad $f_{X,Y}(x,y)$

Antes:

$$\mathbb{P}(\underbrace{|X-Y|}_{W} \le 5) = \int_A f_{X,Y}(x,y)\, dx\, dy \quad \text{donde } A = \{(x,y) : |x-y| \le 5\}$$

debería tener su propia densidad y distribución → Dependiendo de la transformación tenemos métodos automáticos.

Suponemos que el vector (U,V) se puede construir a partir de (X,Y) con una función biyectiva. Es decir, existe

$$\varphi : (X,Y) \longmapsto (U,V) \qquad U = \varphi_1(X,Y) \qquad V = \varphi_2(X,Y)$$

Entonces la densidad conjunta del nuevo vector es

$$f_{U,V}(u,v) = f_{X,Y}(\varphi^{-1}(u,v))\, |J(u,v)|$$

donde

$$J(u,v) = \begin{vmatrix} \frac{\partial}{\partial u}\varphi_1^{-1} & \frac{\partial}{\partial v}\varphi_1^{-1} \\ \frac{\partial}{\partial u}\varphi_2^{-1} & \frac{\partial}{\partial v}\varphi_2^{-1} \end{vmatrix}$$

Para n variables: $\varphi : (X_1, X_2, \dots X_n) \longmapsto (U_1, U_2, \dots\dots U_n)$

$\varphi_k(x, \dots)$

$\varphi_1(x_1 \dots x_n)$

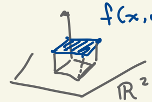
$J(u_1, u_2 \dots u_n)$ ← determinante de una matriz de $n \times n$.

$$f_{U,V}(u,v) = f_{X,Y}\left(\varphi^{-1}(u,v)\right)\,|J(u,v)|$$

Ejemplo:

Si (X,Y) son variables independientes y uniformes en $[0,1]$

$$f_{X,Y}(x,y) = \begin{cases} 1 & x,y \in (0,1) \\ 0 & \text{caso} \end{cases}$$



Definimos $U = X+Y$, $V = X-Y$ entonces $\varphi_1(x,y) = x+y$, $\varphi_2(x,y) = x-y$

$\varphi : (X,Y) \to (U,V)$ $\qquad \varphi^{-1} : (U,V) \to (X,Y)$

$U = x+y$

$V = x-y$

$$x = \frac{U+V}{2} = \varphi_1^{-1}(u,v)$$

$$y = \frac{U-V}{2} = \varphi_2^{-1}(u,v)$$

Lo siguiente es calcula el jacobiano (recuerden quedarse con el valor absoluto)

$$J(u,v) = \begin{vmatrix} \frac{\partial}{\partial u}\frac{u+v}{2} & \frac{\partial}{\partial v}\frac{u+v}{2} \\ \frac{\partial}{\partial u}\frac{u-v}{2} & \frac{\partial}{\partial v}\frac{u-v}{2} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

$$\varphi^{-1}(u,v) = \left(\varphi_1^{-1}(u,v),\ \varphi_2^{-1}(u,v)\right) = \left(\frac{u+v}{2},\ \frac{u-v}{2}\right)$$

Si X,Y fueran $N(0,1)$ independientes.

$$f_{X,Y}(x,y) = \frac{e^{-\frac{(x^2+y^2)}{2}}}{2\pi} \quad \forall x,y.$$

entonces

$$f_{U,V}(u,v) = \frac{1}{2}\cdot\frac{e^{-\frac{(u^2+v^2)}{4}}}{2\pi} \quad \forall u,v$$

$$f_{U,V}(u,v) = \begin{cases} 1/2 & \varphi^{-1}(u,v) \in [0,1]^2 \\ 0 & \text{en otro caso} \end{cases}$$

$$\varphi^{-1}(u,v) = \left(\overset{x}{\frac{u+v}{2}},\ \overset{y}{\frac{u-v}{2}}\right)$$

$$f_{U,V}(u,v) = f_{X,Y}\left(\varphi^{-1}(u,v)\right)\,|J(u,v)|$$

Para tener una expresión explícita resolvemos

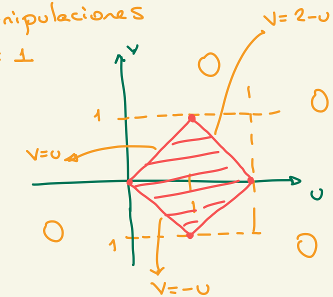
$$0 \le \frac{U+V}{2} \le 1$$

$$0 \le \frac{U-V}{2} \le 1$$

* sumando las dos expresiones $0 \le U \le 2$

* con otras manipulaciones $-1 \le V \le 1$

Resolver el sistema de desigualdades = desglosar cada una de las 4 desigualdades.



P: $0 \le \frac{U+V}{2} \longrightarrow V = -U$

O: $0 \le \frac{U-V}{2} \longrightarrow U = V$

F: $\frac{U+V}{2} \le 1 \longrightarrow V = 2-U$

G: $\frac{U-V}{2} \le 1 \longrightarrow V = -2+U$

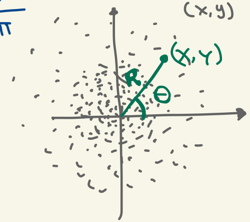
Hacia el algoritmo de Box-Müller.

Suponemos que (X,Y) son v.a. independientes normales $(0,1)$

Su densidad es

$$f_{X,Y}(x,y) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \cdot \frac{e^{-y^2/2}}{\sqrt{2\pi}} = \frac{e^{-\frac{(x^2+y^2)}{2}}}{2\pi}$$

Varias muestras de (x,y)



Utilizaremos la transformación: (R,Θ) como sus coordenadas polares

$$R = \sqrt{X^2+Y^2}$$
$$\Theta = \arctan\left(\frac{Y}{X}\right)$$

$$\varphi(x,y) = \left(\sqrt{x^2,y^2}, \arctan\left(\frac{y}{x}\right)\right)$$
$$\varphi^{-1}(r,\theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta)$$

Continuamos con

$$J(r,\theta) = \begin{vmatrix} \frac{\partial}{\partial r} r\cos\theta & \frac{\partial}{\partial \theta} r\cos\theta \\ \frac{\partial}{\partial r} r\,\mathrm{sen}\,\theta & \frac{\partial}{\partial \theta} r\,\mathrm{sen}\,\theta \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{vmatrix}$$
$$= r\,(\cos^2\theta + \mathrm{sen}^2\theta) = r$$

La función de densidad: De φ soporte es $R\in[0,\infty)$, $\Theta\in[0,2\pi]$

$$f_{R,\Theta}(r,\theta) = \frac{r\cdot e^{-r^2/2}}{2\pi} = r e^{-r^2/2}\cdot\frac{1}{2\pi} \qquad \forall\, r\in[0,\infty),\ \theta\in[0,2\pi]$$

El algoritmo de Box-Müller

Si tenemos un par (R,Θ) con densidad:

$$f_{R,\Theta}(r,\theta) = \frac{r\cdot e^{-r^2/2}}{2\pi} = r e^{-r^2/2}\cdot\frac{1}{2\pi} \qquad \forall\, r\in[0,\infty),\ \theta\in[0,2\pi]$$

Entonces el vector (X,Y) dado por

$$X = R\cos\Theta$$
$$Y = R\,\mathrm{sen}\,\Theta$$

tiene distribución binormal con entradas independientes y estándar.

Pasos del Algoritmo:

1. Genera $\Theta = \mathrm{Unif}[0,2\pi]$
2. Genera $W = \mathrm{Exp}(1/2)$ ← Método de transformada inv.
3. Definimos $R = W^{1/2}$
4. Devolvemos $X = R\cos\Theta$, $Y = R\,\mathrm{sen}\,\Theta$

Verificando paso 3 del algoritmo:

Cambio $r = s^{1/2}$, $dr = \frac{1}{2}s^{-1/2}ds$

$$\mathbb{P}(\underbrace{W^{1/2}}_{\text{"R"}} \le x) = \mathbb{P}(W\le x^2) = \int_0^{x^2} \frac{1}{2}\cdot e^{-\frac{s}{2}}\,ds = \int_0^{x^2} s^{1/2} e^{-s/2}\cdot \underbrace{\frac{1}{2}s^{-1/2}ds}_{dr}$$
$$= \int_0^x r e^{-r^2/2}\,dr = \mathbb{P}(R\le x)$$