
NO TENEMOS DUDAS...
SON LAS MUTACIONES QUE HA IDO SUFRIENDO EL VIRUS DESDE QUE INGESÓ AL URUGUAY

Probabilidad - Clase 20
Variables aleatorias independientes

Ernesto Mordecki

Facultad de Ciencias, Universidad de la República. Montevideo, Uruguay

Curso de la Licenciatura en Matemática - 2020

Contenidos

Vectores aleatorios

- Sean $X_1, \ldots, X_n$ variables aleatorias definidas en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbf{P})$.
- El vector $X = (X_1, \ldots, X_n)$ se denomina *vector aleatorio*,
- Este vector toma valores en $\mathbb{R}^n$, es decur

$$X\colon \Omega \to \mathbb{R}^n.$$

- En el caso particular $n = 1$, que llamamos *unidimensional*, obtenemos una variable aleatoria.
- Como el conjunto $\{\omega\colon X_k(\omega) \leq x_k\}$ es un suceso (es decir, pertenece a $\mathcal{A}$) para cada $k = 1, \ldots, n$ y reales $x_1, \ldots, x_k$ arbitrarios, tenemos que

$$\bigcap_{i=1}^{n} \{\omega\colon X_k(\omega) \leq x_k\} \in \mathcal{A},$$

- Se puede definir la función real de n variables

$$F(x_1, \dots, x_n) = \mathbf{P}\Big(\bigcap_{i=1}^{n}\{\omega\colon X_k(\omega) \leq x_k\}\Big),$$

  que se denomina *función de distribución n–dimensional* del vector aleatorio X.

- La probabilidad recién considerada se designa también

$$\mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n).$$

- Luego,

$$F(x_1, \dots, x_n) = \mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n). \tag{1}$$

- Para $n = 1$ la definición en dada coincide con la definición de distribución de una variable aleatoria.

Distribución uniforme en el hiper-cubo

Ejemplo. Decimos que el vector aleatorio $X = (X_1, \ldots, X_n)$ tiene *distribución uniforme* en el hiper-cubo $[0, 1]^n$ del espacio $\mathbb{R}^n$, si tiene densidad dada por

$$p(x_1, \ldots, x_n) = \begin{cases} 1, & \text{si } 0 \leq x_i \leq 1 \text{ para } i = 1, \ldots, n, \\ 0, & \text{en caso contrario.} \end{cases}$$

Observación: El comando `runif(n)` nos da un punto con distribución uniforme en $[0, 1]^n$.

Distribución normal n dimensional

Decimos que el vector aleatorio $X = (X_1, \dots, X_n)$ tiene *distribución normal n–dimensional* si tiene densidad dada por

$$p(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(B)}} e^{-\frac{1}{2}(x-a)B^{-1}(x-a)'} \tag{2}$$

donde

- $x = (x_1, \dots, x_n)$ y $a = (a_1, \dots, a_n)$ son vectores fila de números reales;
- B es una matriz de dimensión $n \times n$, definida positiva[1],
- no singular y simétrica, B^{-1} es la matriz inversa de la matriz B,
- x' denota el vector traspuesto de x.

[1] Una matriz $B = (b_{ij})$ de dimensión $n \times n$ es *definida positiva*, si para todo vector fila $x = (x_1, \dots, x_n)$ no nulo, se verifica $xBx' = \sum_{i,j} x_i b_{ij} x_j > 0$.

Casos $n = 1, 2$

En el caso $n = 1$, esta densidad se reduce a la fórmula de la densidad normal, donde el "vector" a es el número a, y la matriz de dimensión 1×1 es $B = [\sigma^2]$.

Caso bidimensional. Consideremos un vector aleatorio (X, Y) con *distribución normal bidimensional* [2]. No es difícil de verificar que la matriz

$$B = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix}, \tag{3}$$

verifica las tres condiciones indicadas, si $\sigma_1 > 0$, $\sigma_2 > 0$, y $-1 < \rho < 1$.

[2] Decimos *bidimensional* en vez de 2–dimensional.

Veamos esto:

- Es claramente simétrica
- $$\det(B) = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = (1-\rho^2)\sigma_1^2\sigma_2^2 > 0.$$
- Los valores propios tienen el mismo signo:
$$\lambda_1\lambda_2 = \det(B) > 0$$
- Y su suma es positiva
$$\lambda_1 + \lambda_2 = \text{traza}(B) = \sigma_1^2 + \sigma_2^2 > 0$$

Luego ambos valores propios son positivos y la matriz es definida positiva.

Si $a = (a_1, a_2)$, y sustituimos la matriz B dada en (3) en la densidad, obtenemos,

$$p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$
$$\times \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-a_1}{\sigma_1}\right)^2 + \left(\frac{y-a_2}{\sigma_2}\right)^2 - 2\rho\frac{(x-a_1)}{\sigma_1}\frac{(y-a_2)}{\sigma_2}\right]\right\}. \tag{4}$$

Distribución marginal

Dado un vector X, cada coordenada X_i tiene una distribución que se llama *marginal*.
Demostremos que para el vector (X, Y) normal bidimensional, X tiene distribución normal con parámetros (a_1, σ_1): [3] Tenemos

$$\mathbf{P}(X \leq x) = \mathbf{P}(X \leq x, Y < \infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} p(u, v) du dv$$
$$= \int_{-\infty}^{x} \Big(\int_{-\infty}^{\infty} p(u, v) dv \Big) du. \tag{5}$$

Para calcular la integral con respecto de v, introducimos el cambio de variable

$$t = \big((v - a_2)/\sigma_2 - \rho(u - a_1)/\sigma_1\big)/\sqrt{1 - \rho^2}.$$

[3] Luego Y tiene distribución normal con parámetros (a_2, σ_2).

Calculando t^2 y sustituyendo en el exponente en la densidad, obtenemos

$$\int_{-\infty}^{\infty} p(u,v)dv = \frac{1}{2\pi\sigma_1}\int_{-\infty}^{\infty} e^{-\left(t^2/2+(u-a_1)^2/(2\sigma_1^2)\right)}dt$$
$$= \frac{1}{\sigma_1\sqrt{2\pi}}e^{-(u-a_1)^2/(2\sigma_1^2)},$$

donde utilizamos que $\int_{-\infty}^{\infty} e^{-t^2/2}dt = \sqrt{2\pi}$.

Sustituyendo esta expresión, resulta

$$\mathbf{P}(X \leq x) = \frac{1}{\sigma_1\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(u-a_1)^2/(2\sigma_1^2)} du,$$

es decir, la variable aleatoria X tiene distribución normal con parámetros (a_1, σ_1).
La afirmación relativa a Y se demuestra en forma análoga.

Vector normal estándar

Un caso importante es cuando

$$a = 0, \qquad B = Id_n.$$

En este caso

$$\begin{aligned} p(x) &= \frac{1}{(2\pi)^{n/2}\sqrt{\det(B)}} e^{-\frac{1}{2}(x-a)B^{-1}(x-a)'} \\ &= \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}xx'} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\|x\|^2} \\ &= \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}(x_1^2+\cdots+x_n^2)} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_i^2} \\ &= \prod_{i=1}^{n} \varphi(x_i). \end{aligned}$$

Variables independientes

- Sean $X_1, \dots, X_n$ v.a. en $(\Omega, \mathcal{A}, \mathbf{P})$,
- Sea $F_k(x)$ la distribución de la variable aleatoria X_k $(k = 1, \dots, n)$,
- Sea $F(x_1, \dots, x_n)$ a la distribución del vector aleatorio $(X_1, \dots, X_n)$.
- Decimos que $X_1, \dots, X_n$ son *variables aleatorias independientes* cuando

$$F(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n)$$

para reales $x_1, \dots, x_n$ arbitrarios.

Ejemplo: distribución uniforme en el hiper-cubo

Sea $X = (X_1, \dots, X_n)$ tiene uniforme en $[0,1]^n$. Tenemos

$$\begin{aligned} F(x_1, \dots, x_n) &= \mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n) \\ &= x_1 \dots x_n. \end{aligned}$$

(es el hiper-volumen). Por otra parte

$$F_k(x_k) = \mathbf{P}(X_k \leq x_k) = x_k.$$

Entonces

$$F(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n).$$

Es decir, las v.a. $X_1, \dots X_n$ son independientes.

Lema. Las variables aleatorias $X_1, \dots, X_n$ son independientes si y solo si, para reales $x_1, \dots x_n$ arbitrarios, son independientes los sucesos $\{\omega\colon X_1(\omega) \leq x_1\}, \dots, \{\omega\colon X_n(\omega) \leq x_n\}$

Demostración. En efecto, si los sucesos anteriores son independientes, tenemos

$$\begin{aligned} &\mathbf{P}(\{\omega\colon X_1(\omega) \leq x_1\}, \dots, \{\omega\colon X_n(\omega) \leq x_n\}) \\ &\qquad = \mathbf{P}(\{\omega\colon X_1(\omega) \leq x_1\}) \dots, \mathbf{P}(\{\omega\colon X_n(\omega) \leq x_n\}), \end{aligned}$$

es decir

$$F(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n).$$

(el recíproco se lee al revés).

Si consideramos variables aleatorias discretas o absolutamente continuas, podemos formular la condición de independencia

$$F(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n).$$

en términos de las probabilidades de los valores que toman las variables aleatorias discretas, o en términos de densidades en el caso absolutamente continuo. Para ésto, utilizamos el resultado siguiente.

Lema

Consideremos dos variables aleatorias X, Y independientes. Para $a < b$, $c < d$, se verifica

$$\mathbf{P}(a < X \leq b, c < Y \leq d) = \mathbf{P}(a < X \leq b)\,\mathbf{P}(c < Y \leq d).$$

Demostración Sea $F(x, y)$ la función de distribución del vector (X, Y). Tenemos

$$F(b, d) = \mathbf{P}(X \leq b, Y \leq d) = \mathbf{P}(X \leq a, Y \leq d) + \mathbf{P}(X \leq b, Y \leq c) \\ - \mathbf{P}(X \leq a, Y \leq c) + \mathbf{P}(a < X \leq b, c < Y \leq d).$$

Aplicando la fórmula de independencia, obtenemos

$$\begin{aligned}
\mathbf{P}(a < X \leq b, c < Y \leq d) \\
&= \mathbf{P}(X \leq b)\,\mathbf{P}(Y \leq d) - \mathbf{P}(X \leq a)\,\mathbf{P}(Y \leq d) \\
&\quad - \mathbf{P}(X \leq b)\,\mathbf{P}(Y \leq c) + \mathbf{P}(X \leq a)\,\mathbf{P}(Y \leq c) \\
&= \mathbf{P}(X \leq b)\big(\mathbf{P}(Y \leq d) - \mathbf{P}(Y \leq c)\big) \\
&\quad - \mathbf{P}(X \leq a)\big(\mathbf{P}(Y \leq d) - \mathbf{P}(Y \leq c)\big) \\
&= \mathbf{P}(a < X \leq b)\,\mathbf{P}(c < Y \leq d),
\end{aligned}$$

lo que concluye la demostración.

Proposición

Consideremos dos variables aleatorias X, Y con distribución discreta. Sean $x_1, x_2, \ldots$ los valores que toma la variable X; $y_1, y_2, \ldots$ los valores que toma la variable Y. Las variables aleatorias X e Y son independientes, si y solo si se verifica

$$\mathbf{P}(X = x_k, Y = y_j) = \mathbf{P}(X = x_k)\,\mathbf{P}(Y = y_j), \tag{6}$$

para todos $k, j = 1, 2, \ldots$.

Demostración Supongamos primero que las variables aleatorias X e Y son independientes. Para cada k y cada n naturales, consideramos el suceso

$$\mathbf{A}_{k,n} = \{\omega \colon x_k - 1/n < X(\omega) \leq x_k\}.$$

Tenemos

$$\mathbf{A}_{k,1} \supset \mathbf{A}_{k,2} \supset \cdots ,$$

y además

$$\bigcap_{n=1}^{\infty} \mathbf{A}_{k,n} = \{\omega \colon X(\omega) = x_k\}.$$

Por la propiedad de continuidad, obtenemos

$$\lim_{n\to\infty} \mathbf{P}(\mathbf{A}_{k,n}) = \mathbf{P}(X = x_k).$$

En forma análoga, obtenemos

$$\lim_{n\to\infty} \mathbf{P}(y_j - 1/n < Y \leq y_j) = \mathbf{P}(Y = y_j), \quad \text{para } j \text{ arbitrario.}$$

Con la misma argumentación, obtenemos que

$$\lim_{n\to\infty} \mathbf{P}(x_k - 1/n < X \leq x_k, y_j - 1/n < Y \leq y_j) = \mathbf{P}(Y = y_j, X = x_k) \tag{7}$$

para j, k arbitrarios.

Aplicando el lema 1, se tiene

$$\begin{aligned}\mathbf{P}(x_k - 1/n < X \leq x_k, y_j - 1/n < Y \leq y_j) \\ = \mathbf{P}(x_k - 1/n < X \leq x_k)\,\mathbf{P}(y_j - 1/n < Y \leq y_j),\end{aligned}$$

que sustituido en (7), da la igualdad (6).

Dem del recíproco.

Supongamos que las variables aleatorias X e Y verifican

$$\mathbf{P}(X = x_k, Y = y_j) = \mathbf{P}(X = x_k)\,\mathbf{P}(Y = y_j),$$

Consideremos el vector aleatorio (X, Y) y su función de distribución

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

Tenemos que ver

$$F(x, y) = \mathbf{P}(X \leq x)\,\mathbf{P}(Y \leq y) = F_X(x)F_Y(y).$$

para x e y arbitrarios.

Tenemos

$$
\begin{aligned}
F(x,y) &= \sum_{k,j:\ x_k\le x, y_j\le y} \mathbf{P}(X=x_k, Y=y_j) \\
&= \sum_{k,j:\ x_k\le x, y_j\le y} \mathbf{P}(X=x_k)\,\mathbf{P}(Y=y_j) \\
&= \sum_{k:\ x_k\le x}\ \sum_{j:\ y_j\le y} \mathbf{P}(X=x_k)\,\mathbf{P}(Y=y_j) \\
&= \sum_{k:\ x_k\le x} \mathbf{P}(X=x_k) \sum_{j:\ y_j\le y} \mathbf{P}(Y=y_j) \\
&= \mathbf{P}(X\le x)\,\mathbf{P}(Y\le y),
\end{aligned}
$$

LQQD

Utilizando un método similar, se obtiene la siguiente generalización de la proposición 1.

Proposición

Consideremos n variables aleatorias $X_1, \dots, X_n$ con distribución discreta, siendo $x_{k1}, x_{k2}, \dots$ los valores que toma cada variable aleatoria X_k, para $k = 1, \dots, n$. Las variables aleatorias dadas son independientes, si y solo si se verifica

$$\mathbf{P}(X_1 = x_{1m_1}, \dots, X_n = x_{nm_n}) = \prod_{k=1}^{n} \mathbf{P}(X_k = x_{km_k}),$$

para naturales $m_1, \dots, m_n$ no nulos y arbitrarios.

El resultado que sigue es el análogo de la proposición 2 para variables aleatorias absolutamente continuas.

Proposición

Consideremos un vector aleatorio $X = (X_1, \dots, X_n)$ *absolutamente continuo, y sea* $p_k(x)$ *la densidad de la variable aleatoria* $X_k (k = 1, \dots, n)$*. Las variables aleatorias* $X_1, \dots, X_n$ *son independientes, si y solo si se verifica*

$$p(x_1, \dots, x_n) = p_1(x_1) \cdots p_n(x_n)$$

Demostración Si las variables aleatorias $X_1, \dots, X_n$ son independientes, la función de distribución se escribe de dos formas: como producto de las distribuciones $F_k(x_k)$ $(k = 1, \dots, n)$, y como integral múltiple de la densidad $p(u_1, \dots, u_n)$, es decir

$$\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(u_1, \dots, u_n) du_1 \cdots du_n = F_1(x_1) \cdots F_n(x_n).$$

Derivando n veces en ambos miembros de esta identidad, primero respecto de x_1, luego respecto de x_2, ..., y finalmente respecto de x_n, obtenemos

$$p(x_1, \dots, x_n) = p_1(x_1) \cdots p_n(x_n)$$

Recíproco

Supongamos

$$p(x_1, \dots, x_n) = p_1(x_1) \cdots p_n(x_n)$$

Integramos n veces en ambos lados de esta igualdad, primero con respecto de x_1 en el intervalo $(-\infty, y_1)$, . . . , y por último con respecto de x_n en el intervalo $(-\infty, y_n)$. Obtenemos

$$F(y_1, \dots, y_n) = F_1(y_1) \cdots F_n(y_n).$$

y se verifica la definición de independencia.

Ejemplo

Sea nuevamente (X, Y) normal bidimensional, con densidad

$$p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$
$$\times\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-a_1}{\sigma_1}\right)^2+\left(\frac{y-a_2}{\sigma_2}\right)^2-2\rho\frac{(x-a_1)}{\sigma_1}\frac{(y-a_2)}{\sigma_2}\right]\right\}.$$

Sabemos que

$$X \sim (a_1, \sigma_1^2) \qquad Y \sim (a_2, \sigma_2).$$

Si $\rho = 0$, la densidad se reduce a

$$p(x, y) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-a_1)^2/(2\sigma_1^2)} \times \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(y-a_2)^2/2\sigma_2^2} = p_1(x)p_2(y),$$

donde

- $p_1(x)$ es la densidad de la variable aleatoria X,
- $p_2(y)$ la densidad de Y.

Como consecuencia en el caso $\rho = 0$, las variables aleatorias X e Y son independientes.

Ejemplo Consideremos un vector aleatorio $(X_1, \dots, X_n)$ con distribución normal n–dimensional, con densidad $p(x)$ dada en la fórmula (2). Supongamos que la matriz B es diagonal, y está dada por

$$B = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}. \tag{8}$$

Si $\sigma_k > 0$ ($k = 1, \ldots, n$) la matriz dada verifica las tres condiciones indicadas en la definición (ser definida positiva, simétrica, y no singular). Sustituyendo la matriz B en la fórmula (2), obtenemos la densidad del vector considerado, que es

$$p(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\sigma_1 \cdots \sigma_n} e^{-\sum_{k=1}^{n}(x_k - a_k)^2/(2\sigma_k^2)}. \qquad (9)$$

Como ocurre en el caso $n = 2$, se verifica que cada variable aleatoria X_k tiene distribución normal, con parámetros (a_k, σ_k), es decir, tiene densidad $p_k(x) = (\sigma_k\sqrt{2\pi})^{-1} e^{-(x-a_k)^2/(2\sigma_k^2)}$, para $k = 1, \ldots, n$. Mas aún, factorizando la fórmula (9), tenemos

$$p(x_1, \ldots, x_n) = \prod_{k=1}^{n} \frac{1}{\sigma_k\sqrt{2\pi}} e^{-(x_k - a_k)^2/(2\sigma_k^2)} = \prod_{k=1}^{n} p_k(x_k). \qquad (10)$$

Aplicando la proposición 3 se obtiene la independencia mutua de las variables aleatorias $X_1, \ldots, X_n$.

Reciprocamente, si las variables aleatorias $X_1, \dots, X_n$ son independientes, y cada una de las variables X_k tiene distribución normal con parámetros (a_k, σ_k) $(k = 1, \dots, n)$, aplicando la proposición 3, obtenemos que la densidad $p(x_1, \dots, x_n)$ del vector $X = (X_1, \dots, X_n)$ verifica la fórmula (10), y en consecuencia, que el vector X tiene distribución normal n–dimensional, con densidad dada en (2), y matriz B dada en (8).

Hoy: ¡salpicón!

- ¿Qué es un terapeuta? 1024 gigapeutas.
- ¿Por qué un fotón no puede hacer una pizza? Porque no tiene masa.
- ¿Que es un niño complejo?: el que tiene madre real y padre imaginario.