

5- VARIABLES ALEATORIAS BIDIMENSIONALES

5.1 – Generalidades

Hasta ahora hemos considerado el caso de variables aleatorias unidimensionales. Esto es, el resultado del experimento de interés se registra como un único número real.

En muchos casos, sin embargo, nos puede interesar asociar a cada resultado de un experimento aleatorio, dos o más características numéricas. Por ejemplo, de los remaches que salen de una línea de producción nos puede interesar el diámetro X y la longitud Y . Teniendo en cuenta la inevitable variabilidad en las dimensiones de los remaches debido a las numerosas causas presentes en el proceso de fabricación, los podemos representar asociándoles dos variables aleatorias X e Y que pueden pensarse como una **variable aleatoria bidimensional**: (X, Y) .

Sea ε un experimento aleatorio y S un espacio muestral asociado a él. Sean $X : S \rightarrow R, Y : S \rightarrow R$, que a cada resultado $s \in S$ le asignan el par de números reales (x, y)

Llamaremos a (X, Y) **variable aleatoria bidimensional**.

Si en lugar de dos variables aleatorias, tenemos n variables aleatorias X_1, X_2, \dots, X_n , llamaremos a (X_1, X_2, \dots, X_n) **variable aleatoria n-dimensional**

En lo que sigue nos referiremos en particular a variables aleatorias n -dimensionales con $n=2$, es decir nos concentraremos en **variables aleatorias bidimensionales** por cuanto son las más simples de describir, fundamentalmente en relación a la notación. Pero debemos tener presente que las propiedades que estudiemos para ellas se pueden extender sin demasiada dificultad al caso general.

Al conjunto de valores que toma la variable aleatoria bidimensional (X, Y) lo llamaremos **recorrido** de la v.a. (X, Y) y lo indicaremos R_{XY} . En otras palabras $R_{XY} = \left\{ (x, y) : x = X(s) \text{ e } y = Y(s) \text{ con } s \in S \right\}$, es decir, es la imagen por (X, Y) del espacio muestral S .

Notar que el recorrido de (X, Y) es un subconjunto del espacio Euclidiano: $R_{XY} \subseteq R^2$. Como antes, puede considerarse al recorrido R_{XY} como un espacio muestral cuyos elementos son ahora pares de números reales.

Como con cualquier espacio muestral, según el número de elementos que lo constituyen, podemos clasificar a los recorridos R_{XY} en numerables (finitos o infinitos) y no-numerables.

Los recorridos numerables son, en general, de la forma

$$R_{XY} = \left\{ (x_i, y_j) \text{ con } i = 1, 2, \dots, n \text{ y } j = 1, 2, \dots, m \right\} = \{(x_1, y_1), (x_1, y_2), \dots, (x_n, y_m)\} \quad (\text{finito})$$

$$R_{XY} = \left\{ (x_i, y_j) \text{ con } i = 1, 2, \dots \text{ y } j = 1, 2, \dots \right\} = \{(x_1, y_1), (x_1, y_2), \dots\} \quad (\text{infinito numerable})$$

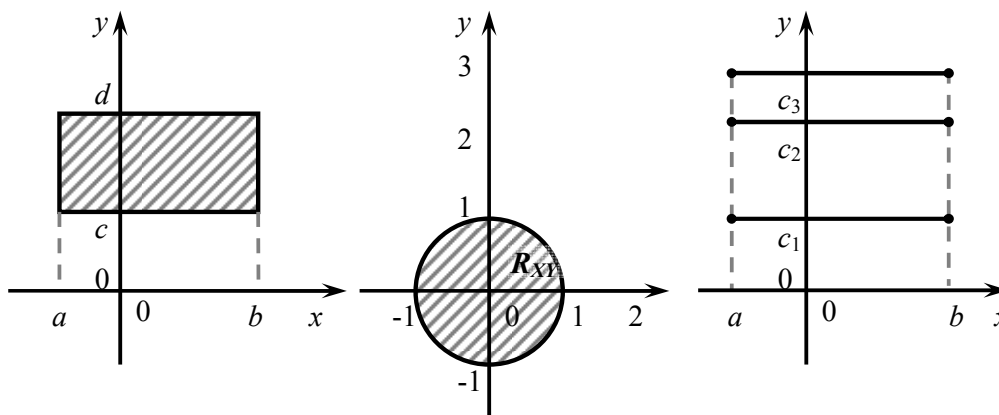
Los recorridos no numerables son regiones o subconjuntos no numerables del plano Euclidiano. Por ejemplo:

$$R_{XY} = \left\{ (x, y) : a \leq x \leq b; \quad c \leq y \leq d \right\} \quad (\text{no numerable})$$

$$R_{XY} = \{(x, y) : x^2 + y^2 \leq 1\} \quad (\text{no numerable})$$

$$R_{XY} = \{(x, y_j) : a \leq x \leq b, y_j = c_1, c_2, c_3\} \quad (\text{no numerable "mixto"})$$

cuyas gráficas se pueden apreciar en la figura siguiente. Notar en el último recorrido, X es v.a. continua e Y discreta.



Clasificaremos a las variables aleatorias bidimensionales de la siguiente manera:

(X, Y) es v.a. **bidimensional discreta** si X e Y son discretas

(X, Y) es v.a. **bidimensional continua** si X e Y son continuas

El caso X continua, Y discreta (o viceversa) no lo consideramos.

Sea (X, Y) una variable aleatoria bidimensional discreta y sea R_{XY} su recorrido (numerable). Sea $p : R_{XY} \rightarrow \mathbb{R}$ una función que a cada elemento (x_i, y_j) le asigna un número real $p(x_i, y_j)$ tal que

$$P(X = x_i, Y = y_j) = p(x_i, y_j) \quad \forall (x_i, y_j) \in R_{XY} \text{ y que verifica.}$$

a) $p(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{XY}$

b) $\sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = \sum_i \sum_j p(x_i, y_j) = 1$

A esta función la llamaremos **función de probabilidad puntual conjunta** de la variable aleatoria bidimensional (X, Y) . En forma abreviada la designaremos **fdp conjunta**.

Ejemplos:

1-Dos líneas de producción, señaladas I y II, manufacturan cierto tipo de artículo a pequeña escala. Supóngase que la capacidad máxima de producción de la línea I es cinco artículos por día, mientras que para la línea II es 3 artículos/día. Debido a los innumerables factores presentes en todo proceso de producción, el número de artículos realmente producido por cada línea puede pensarse como una variable aleatoria. En conjunto podemos pensar en una variable aleatoria bidimensional (X, Y) discreta, donde la primera componente X corresponde a la producción de la línea I y la segunda componente Y a los artículos que salen de la línea II. La **fdp** conjunta correspondiente a variables aleatorias bidimensionales suele presentarse, por comodidad, como una tabla. Supongamos que la para la v.a. (X, Y) que nos interesa aquí la tabla correspondiente a $p(x_i, y_j)$ es

Y/X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

¿Cuál es la probabilidad de que salgan más artículos de la línea I que de la línea II?

Antes de calcular la probabilidad que nos pide el problema, hagamos algunas consideraciones sobre la tabla que representa a $p(x_i, y_j)$.

Se trata de una tabla a doble entrada donde en la primera fila se indican los valores que puede tomar la v.a. X (en este caso $X=0,1,2,3,4,5$) y la primera columna indica los valores que puede tomar la variable Y ($0,1,2,3$). Para determinar el valor de la $p(x_i, y_j)$ cuando la v.a. (X, Y) toma el valor (x_i, y_j) consideramos el número que se encuentra en la columna correspondiente a $X = x_i$ y la fila correspondiente a $Y = y_j$. Por ejemplo: $p(4,2) = P(X = 4, Y = 2) = 0.05$.

Podemos verificar fácilmente que la *fdp* conjunta definida por esta bien definida. En efecto verifica las condiciones a) $p(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{XY}$ y b) $\sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = 1$.

Para contestar la pregunta del enunciado, consideremos el suceso $B \subset R_{XY}$ definido

B : “es el suceso que ocurre cuando la línea I produce más artículos que la línea II” o, $B = \{X > Y\}$. Luego:

$$P(B) = P(X > Y) = \sum_{y_j=0}^3 \sum_{x_i > y_j} p(x_i, y_j) = 0.01+0.03+0.05+0.07+0.09+0.04+0.05+0.06+0.08+ \\ +0.05+0.05+0.06+0.06+0.05=0.75.$$

2- Hay tres cajas registradoras a la salida de un supermercado. Dos clientes llegan a las cajas en diferentes momentos cuando no hay otros clientes ante aquellas. Cada cliente escoge una caja al azar e independientemente del otro.

Sean las variables aleatorias X : “nº de clientes que escogen la caja 1” e Y : “nº de clientes que escogen la caja 2”. Hallar la *fdp* conjunta de (X, Y)

Podemos suponer que el espacio muestral original S es el conjunto de pares ordenados $S = \{(1,1); (1,2); (1,3); (2,1); (2,2); (2,3); (3,1); (3,2); (3,3)\}$ donde la primera componente del par indica la caja elegida por el cliente 1 y la segunda componente del par indica la caja elegida por el cliente 2.

Además notar que X como Y pueden tomar los valores 0, 1, 2

El punto muestral (3,3) es el único punto muestral que corresponde al evento $\{X = 0, Y = 0\}$

Entonces

$P(X = 0, Y = 0) = \frac{1}{9}$; pensando de forma análoga los otros casos:

$P(X = 1, Y = 0) = \frac{2}{9}$; $P(X = 2, Y = 0) = \frac{1}{9}$; $P(X = 0, Y = 1) = \frac{2}{9}$, $P(X = 1, Y = 1) = \frac{2}{9}$,

$P(X = 0, Y = 2) = \frac{1}{9}$; $P(X = 1, Y = 2) = P(X = 2, Y = 2) = 0$

Disponemos estas probabilidades en una tabla de la siguiente forma

Y \ X	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

5.2 - Funciones de distribución marginales de una v.a. (X,Y) discreta

En el ejemplo 1, supongamos que queremos saber cuál es la probabilidad de que el número de artículos producidos por la línea I sea 2, o sea $P(X = 2)$

Como el evento $\{X = 2\}$ es igual a $\{X = 2\} \cap (\{Y = 0\} \cup \{Y = 1\} \cup \{Y = 2\} \cup \{Y = 3\})$, y a su vez $\{X = 2\} \cap (\{Y = 0\} \cup \{Y = 1\} \cup \{Y = 2\} \cup \{Y = 3\}) = (\{X = 2\} \cap \{Y = 0\}) \cup (\{X = 2\} \cap \{Y = 1\}) \cup (\{X = 2\} \cap \{Y = 2\}) \cup (\{X = 2\} \cap \{Y = 3\})$

Entonces

$$\begin{aligned}
 P(X = 2) &= \\
 &= P(\{X = 2\} \cap \{Y = 0\}) + P(\{X = 2\} \cap \{Y = 1\}) + P(\{X = 2\} \cap \{Y = 2\}) + P(\{X = 2\} \cap \{Y = 3\}) = \\
 &= P(X = 2, Y = 0) + P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 2, Y = 3) = \sum_{j=0}^3 P(X = 2, Y = j)
 \end{aligned}$$

Razonando de la misma forma podemos escribir

$$P(X = i) = \sum_{j=0}^3 P(X = i, Y = j) \quad i = 0, 1, \dots, 5$$

Es decir obtenemos la **función de distribución de probabilidad de X**

Análogamente obtenemos

$$P(Y = j) = \sum_{i=0}^5 P(X = i, Y = j) \quad j = 0, 1, 2, 3$$

Que es la **función de distribución de probabilidad de Y**

En general se las denomina **distribuciones marginales de X e Y**, y su definición sería la siguiente

Sea (X, Y) discreta y sea $p(x_i, y_j)$ ($i=1, 2, \dots, n, j=1, 2, \dots, m$) su función de probabilidad conjunta (Eventualmente n y/o m pueden ser ∞).

La función de probabilidad marginal de X es

$$p(x_i) = P(X = x_i) = \sum_{j=1}^m p(x_i, y_j) \quad (i=1, 2, \dots, n)$$

La función de probabilidad marginal de Y es

$$q(y_j) = P(Y = y_j) = \sum_{i=1}^n p(x_i, y_j) \quad (j=1, 2, \dots, m)$$

Observación: Remarcamos que la función de probabilidad marginal de X , es decir $p(x_i)$ calculada a partir de $p(x_i, y_j)$ en la forma indicada, coincide con la función de probabilidad de la variable aleatoria unidimensional X considerada en forma aislada. Análogamente la función de probabilidad marginal de

Y , es decir $q(y_j)$ calculada a partir de $p(x_i, y_j)$ en la forma indicada, coincide con la función de probabilidad de variable aleatoria unidimensional Y considerada en forma aislada.

Ejemplo:

Siguiendo con el ejemplo 1,

$$p(5) = P(X = 5) = p(5,0) + p(5,1) + p(5,2) + p(5,3) = 0.09 + 0.08 + 0.06 + 0.05 = 0.28$$

$$q(1) = P(Y = 1) = p(0,1) + p(1,1) + p(2,1) + p(3,1) + p(4,1) + p(5,1) = 0.01 + 0.02 + 0.04 + 0.05 + 0.06 = 0.26$$

Observemos que se verifica la condición de normalización para cada una de las marginales:

$$\sum_{x_i=0}^5 p(x_i) = 0.03 + 0.08 + 0.16 + 0.21 + 0.24 + 0.28 = 1$$

$$\sum_{y_j=0}^3 q(y_j) = 0.25 + 0.26 + 0.25 + 0.24 = 1$$

5.3 - Funciones de probabilidades condicionales

Consideremos nuevamente el ejemplo de las dos líneas I y II que producen cierto artículo a pequeña escala. Definimos la v.a. (X, Y) cuya función de probabilidad conjunta $p(x_i, y_j)$ está dada por la tabla anterior que repetimos

Y/X	0	1	2	3	4	5	$q(y_j)$
0	0	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
$p(x_i)$	0.03	0.08	0.16	0.21	0.24	0.28	1

Supongamos que deseamos conocer la probabilidad de que la línea I produzca tres artículos sabiendo que la línea II ha fabricado dos. Tenemos que calcular una probabilidad condicional. Entonces

$$P(X = 3|Y = 2) = \frac{P(X = 3, Y = 2)}{P(Y = 2)} = \frac{p(3,2)}{q(2)} = \frac{0.05}{0.25} = 0.2$$

En general definimos la **función de probabilidad puntual de X condicional a Y** como sigue:

$p(x_i|y_j) = P(X = x_i|Y = y_j) = \frac{p(x_i, y_j)}{q(y_j)}$, es decir como el cociente de la función de probabilidad conjunta de (X, Y) y la función de probabilidad puntual marginal de Y .

Análogamente, definimos la **función de probabilidad puntual de Y condicional a X** :

$q(y_j|x_i) = P(Y = y_j|X = x_i) = \frac{p(x_i, y_j)}{p(x_i)}$, es decir como el cociente de la función de probabilidad puntual conjunta de (X, Y) y la función de probabilidad puntual marginal de X .

5.4– Variables aleatorias independientes

Ya se discutió el concepto de independencia entre dos eventos A y B . Esas mismas ideas podemos trasladarlas en relación a dos variables aleatorias X y Y que, eventualmente, podemos considerarlas como las componentes de una variable aleatoria bidimensional (X, Y) .

De acuerdo con esto, intuitivamente decimos que dos variables, X e Y , son independientes si el valor que toma una de ellas no influye de ninguna manera sobre el valor que toma la otra. Esto lo establecemos más formalmente:

Sea (X, Y) una variable aleatoria bidimensional discreta. Sea $p(x_i, y_j)$ su *fdp* conjunta y $p(x_i)$ y $q(y_j)$ las correspondientes *fdp* marginales de X e Y . Decimos que X e Y son **variables aleatorias independientes** si y sólo si

$$p(x_i, y_j) = p(x_i)q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$$

Observación: Notar que para poder afirmar la independencia de X e Y debe cumplirse la factorización de la *fdp* conjunta como producto de las *fdp* marginales para **todos** los pares de valores de la v.a. (X, Y) . Por lo tanto, para verificar la independencia es necesario demostrar la validez de la factorización para todos los pares. En cambio, es suficiente encontrar un solo par que no la verifica, para afirmar, de acuerdo con la definición, que las variables X e Y son no independientes, es decir, que son dependientes. Esto es, para demostrar la dependencia es suficiente con encontrar un solo par que no verifique la factorización señalada.

Vimos que dos sucesos A y B son independientes si y sólo si $P(A|B) = P(A)$ y $P(B|A) = P(B)$ (donde por supuesto debía ser $P(A) \neq 0$ y $P(B) \neq 0$). En términos de variables aleatorias, esta forma de ver la independencia se manifiesta en la igualdad entre las *fdp* condicionales y las correspondientes *fdp* marginales, como demostramos en este

Teorema

Sea (X, Y) una variable aleatoria bidimensional discreta cuyas *fdp* conjunta, condicionales y marginales son, respectivamente, $p(x_i, y_j)$; $p(x_i|y_j)$, $q(y_j|x_i)$ y $p(x_i)$, $q(y_j)$.

Entonces, X e Y son variables aleatorias independientes si y sólo si

1) $p(x_i|y_j) = p(x_i) \quad \forall (x_i, y_j) \in R_{XY}$, o

2) $q(y_j|x_i) = q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$, que es equivalente a lo anterior

Dem.)

Demostraremos solamente 1). La equivalencia entre 1) y 2) la dejamos como ejercicio.

Para demostrar 1) verificaremos la doble equivalencia entre ésta y la definición de v.a. independientes.

⇒)

Sean X e Y variables aleatorias independientes. Entonces $\forall (x_i, y_j) \in R_{XY}$

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{q(y_j)} = \frac{p(x_i)q(y_j)}{q(y_j)} = p(x_i)$$

Aquí la primera igualdad es la definición de *fdp* condicional y la segunda sale de la definición de independencia al suponer que X e Y son independientes.

⇐)

Supongamos que se verifica 1). Entonces $\forall (x_i, y_j) \in R_{XY}$

$$p(x_i|y_j) = p(x_i) \rightarrow \frac{p(x_i, y_j)}{q(y_j)} = p(x_i) \rightarrow p(x_i, y_j) = p(x_i)q(y_j) \rightarrow X \text{ e } Y \text{ independientes}$$

Aquí, la primera implicación se debe a la definición de *fdp* condicional y la tercera a la definición de v.a. independientes.

Ejemplo:

1- Supongamos que una máquina se usa para un trabajo específico a la mañana y para uno diferente en la tarde. Representemos por X e Y el número de veces que la máquina falla en la mañana y en la tarde respectivamente. Supongamos que la tabla siguiente da la función de probabilidad conjunta $p(x_i, y_j)$ de la variable aleatoria bidimensional discreta (X, Y) .

Y/X	0	1	2	$q(y_j)$
0	0.1	0.2	0.2	0.5
1	0.04	0.08	0.08	0.2
2	0.06	0.12	0.12	0.3
$P(x_i)$	0.2	0.4	0.4	1

Deseamos saber si las variables aleatorias X e Y son independientes o dependientes.

Para demostrar que son independientes debemos probar que se verifica $\forall (x_i, y_j) \in R_{XY}$

$$p(x_i, y_j) = p(x_i)q(y_j) \text{ Verificamos directamente que}$$

$$\begin{aligned}
 p(0,0) &= 0.1 = p(0)q(0) = 0.2 \times 0.5 \\
 p(0,1) &= 0.04 = p(0)q(1) = 0.2 \times 0.2 \\
 p(0,2) &= 0.06 = p(0)q(2) = 0.2 \times 0.3 \\
 p(1,0) &= 0.2 = p(1)q(0) = 0.4 \times 0.5 \\
 p(1,1) &= 0.08 = p(1)q(1) = 0.4 \times 0.2 \\
 p(1,2) &= 0.12 = p(1)q(2) = 0.4 \times 0.3 \\
 p(2,0) &= 0.2 = p(2)q(0) = 0.4 \times 0.5 \\
 p(2,1) &= 0.08 = p(2)q(1) = 0.4 \times 0.2 \\
 p(2,2) &= 0.12 = p(2)q(2) = 0.4 \times 0.3
 \end{aligned}$$

Luego X e Y son independientes.

Podríamos haber usado las condiciones 1) $p(x_i|y_j) = p(x_i) \quad \forall (x_i, y_j) \in R_{XY}$, o su equivalente

2) $q(y_j|x_i) = q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$. Veamos, como muestra para un solo valor, que se verifica

$p(2|1) = \frac{p(2,1)}{q(1)} = \frac{0.08}{0.2} = 0.4 = p(2)$. Para demostrar la independencia por este camino habría que demostrar que se cumple la condición para el resto de los pares de valores. Se deja este cálculo como ejercicio optativo.

Observaciones

1- De la definición de las *fdp* marginales, vemos que tanto en el caso discreto como en el continuo, la *fdp* conjunta determina unívocamente las *fdp* marginales. Es decir, si (X, Y) es discreta del conocimiento de la función de probabilidad conjunta $p(x_i, y_j)$ podemos determinar unívocamente las funciones de probabilidad $p(x_i)$ y $q(y_j)$. Sin embargo la inversa no se cumple en general. Es decir del conocimiento de $p(x_i)$ y $q(y_j)$ no se puede, en general, reconstruir $p(x_i, y_j)$ a menos que X e Y sean variables independientes en cuyo caso es $p(x_i, y_j) = p(x_i)q(y_j)$.

2- El concepto de independencia entre dos variables aleatorias se puede generalizar a n variables aleatorias X_1, X_2, \dots, X_n

5.5 - Función de una variable aleatoria bidimensional

Existen muchas situaciones en las que dado una variable aleatoria bidimensional nos interesa considerar otra variable aleatoria que es función de aquella. Por ejemplo, supongamos que las variables aleatorias X e Y denotan la longitud y el ancho, respectivamente, de una pieza, entonces $Z = 2X + 2Y$ es una v.a. que representa el perímetro de la pieza, o la v.a. $W = X.Y$ representa el área de la pieza. Tanto Z como W **son variables aleatorias**.

En general, sea S un espacio muestral asociado a un experimento probabilístico ε , sean $X : S \rightarrow R$ e $Y : S \rightarrow R$ dos variables aleatorias que definen una variable aleatoria bidimensional (X, Y) cuyo recorrido es R_{XY} , y sea una función de dos variables reales $H : R_{XY} \rightarrow R$ que a cada elemento (x, y) del recorrido R_{XY} le hace corresponder un número real $z = H(x, y)$, entonces la función compuesta $Z = H(X, Y) : S \rightarrow R$ es una variable aleatoria, puesto que a cada elemento $s \in S$ le hace corresponder un número real $z = H[X(s), Y(s)]$. Diremos que la variable aleatoria Z es **función de la variable aleatoria bidimensional (X, Y)** .

Algunas variables aleatorias que son función de variables aleatorias bidimensionales son $Z = X + Y$, $Z = X.Y$, $Z = X / Y$, $Z = \min(X, Y)$, $Z = \max(X, Y)$, etc.

Lo anterior se puede generalizar si en lugar de dos variables aleatorias tenemos n variables aleatorias X_1, X_2, \dots, X_n , y $z = H(x_1, x_2, \dots, x_n)$ es una función de n variables a valores reales.

Ejemplos:

1- Sea $Z \sim B(n, p)$

Podemos escribir a Z como suma de variables aleatorias de la siguiente forma.

Recordar que Z cuenta el número de éxitos en n repeticiones o ensayos del experimento ε

Si definimos

$$X_i = \begin{cases} 1 & \text{si en la } i\text{-ésima repetición de } \varepsilon \text{ ocurre éxito} \\ 0 & \text{caso contrario} \end{cases} \quad i = 1, 2, \dots, n$$

Notar que a cada X_i se la puede considerar $B(1, p)$, y además X_1, X_2, \dots, X_n son independientes

Podemos escribir $Z = X_1 + X_2 + \dots + X_n$

2- Sea Z v.a. binomial negativa con parámetros r y p , es decir $Z \sim BN(r, p)$

Si definimos

X_1 : “número de repeticiones del experimento requeridos hasta el 1° éxito”

X_2 : “número de repeticiones del experimento adicionales requeridos hasta el 2° éxito”

X_3 : “número de repeticiones del experimento adicionales requeridos hasta el 3° éxito”

Y en general

X_i : “número de repeticiones del experimento adicionales después del $(i-1)$ -ésimo éxito requeridos hasta el i -ésimo éxito”

Entonces cada variable tiene **distribución geométrica con parámetro p** y $Z = X_1 + X_2 + \dots + X_r$

Notar además que X_1, X_2, \dots, X_r son independientes

Esperanza de una v.a. que es función de una v.a. bidimensional

Sea una variable aleatoria bidimensional (X, Y) cuya *fdp* conjunta es la función de probabilidad conjunta $p(x_i, y_j)$ si es discreta o la función de densidad de probabilidad conjunta $f(x, y)$ si es continua y sea una función real de dos variables $z = H(x, y)$ de manera que podemos definir una variable aleatoria Z que es función de la variable aleatoria bidimensional (X, Y) de la forma $Z = H(X, Y)$. Si la *fdp* de Z es $q(z_i)$, siendo Z discreta, entonces la esperanza matemática de Z es, de acuerdo con la definición general,

$$E(Z) = \sum_{z_i \in R_Z} z_i \cdot q(z_i) \quad (Z \text{ discreta})$$

Nuevamente lo interesante es considerar la posibilidad de evaluar $E(Z)$ sin tener que calcular previamente la *fdp* de Z . El siguiente teorema nos muestra cómo hacerlo.

Teorema Sea (X, Y) una variable aleatoria bidimensional y sea $Z = H(X, Y)$ una variable aleatoria que es función de (X, Y) .

Si Z es variable aleatoria discreta que proviene de la variable aleatoria bidimensional discreta (X, Y) cuyo recorrido es R_{XY} y su *fdp* conjunta es $p(x_i, y_j)$, entonces:

$$E(Z) = E[H(X, Y)] = \sum_{(x_i, y_j) \in R_{XY}} H(x_i, y_j) p(x_i, y_j)$$

Dem.) sin demostración

Esperanza de una suma de variables aleatorias

Sean X e Y dos variables aleatorias arbitrarias. Entonces $E(X + Y) = E(X) + E(Y)$.

Dem.) en el teorema anterior consideramos $H(x, y) = x + y$

Si (X, Y) es discreta

$$E(Z) = E[H(X, Y)] = \sum_{(x_i, y_j) \in R_{XY}} H(x_i, y_j) p(x_i, y_j) = \sum_{(x_i, y_j) \in R_{XY}} (x_i + y_j) p(x_i, y_j) =$$

Aplicando la propiedad distributiva y separando en dos sumas

$$\begin{aligned} E(Z) &= \sum_{(x_i, y_j) \in R_{XY}} (x_i + y_j) p(x_i, y_j) = \sum_{(x_i, y_j) \in R_{XY}} x_i p(x_i, y_j) + \sum_{(x_i, y_j) \in R_{XY}} y_j p(x_i, y_j) = \\ &= \sum_i \sum_j x_i p(x_i, y_j) + \sum_i \sum_j y_j p(x_i, y_j) = \sum_i x_i \sum_j p(x_i, y_j) + \sum_j y_j \sum_i p(x_i, y_j) = \end{aligned}$$

Pero $\sum_j p(x_i, y_j) = p(x_i)$ y $\sum_i p(x_i, y_j) = q(y_j)$, por lo tanto

$$= \sum_i x_i p(x_i) + \sum_j y_j q(y_j) = E(X) + E(Y)$$

Para el caso (X, Y) continua sigue siendo válida esta propiedad.

Podemos generalizar la propiedad anterior a un número finito cualquiera de variables aleatorias:

Sean X_1, X_2, \dots, X_n n variables aleatorias arbitrarias. Entonces:

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) \text{ o, en notación más concentrada, :}$$

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

(leeremos: “la esperanza de la suma es la suma de las esperanzas”)

Dem.) Se deduce por inducción completa sobre el número n de variables aleatorias.

Observación: se deduce que la esperanza verifica la **propiedad lineal**:

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

Ejemplos:

1- Vamos a aplicar algunas de las propiedades anteriores para calcular de una manera alternativa la esperanza matemática de una variable aleatoria X distribuida binomialmente.

Sea entonces una v.a. $X \sim B(n, p)$.

Ya vimos que podemos escribir $X = X_1 + X_2 + \dots + X_n$ donde cada X_i se la puede considerar $B(1, p)$, y además X_1, X_2, \dots, X_n son independientes

Entonces

$$E(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0) = P(X_i = 1) = p \text{ para cualquier } i$$

Por lo tanto

$$E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = \underbrace{p + p + \dots + p}_{n \text{ veces}} = np$$

Observación: muchas veces es conveniente descomponer una variable aleatoria como suma de otras más simples para facilitar los cálculos

2- Esperanza de una v.a. binomial negativa

Cuando se trató la v.a. binomial negativa se dijo cuál era su esperanza. Ahora damos una demostración Sea X v.a. binomial negativa con parámetros r y p , es decir $X \sim BN(r, p)$

Si definimos

X_1 : “número de repeticiones del experimento requeridos hasta el 1° éxito”

X_2 : “número de repeticiones del experimento adicionales requeridos hasta el 2° éxito”

X_3 : “número de repeticiones del experimento adicionales requeridos hasta el 3° éxito”

Y en general

X_i : “número de repeticiones del experimento adicionales después del $(i-1)$ -ésimo éxito requeridos hasta el i -ésimo éxito”

Entonces cada variable tiene **distribución geométrica con parámetro p** y $X = X_1 + X_2 + \dots + X_r$

Por lo tanto

$$E(X) = E(X_1 + X_2 + \dots + X_r) = E(X_1) + E(X_2) + \dots + E(X_r) = \underbrace{\frac{1}{p} + \frac{1}{p} + \dots + \frac{1}{p}}_{r \text{ veces}} = r \frac{1}{p} = \frac{r}{p}$$

3- Esperanza de una v.a. hipergeométrica

Si $X \sim H(n, M, N)$ entonces $E(X) = \frac{nM}{N}$

Para facilitar la demostración supongamos que tenemos N bolillas en una urna de las cuales M son rojas y $N-M$ son blancas. Queremos hallar el número esperado de bolillas rojas extraídas

Definimos las variables

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ésima bolilla roja es extraída} \\ 0 & \text{caso contrario} \end{cases}$$

Las variables X_1, X_2, \dots, X_M **no son independientes**

Se puede escribir $X = X_1 + X_2 + \dots + X_M$, además

$$E(X_i) = P(X_i = 1) = \frac{\binom{1}{1} \binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}$$

Por lo tanto

$$E(X) = E(X_1 + X_2 + \dots + X_M) = E(X_1) + E(X_2) + \dots + E(X_M) = \underbrace{\frac{n}{N} + \frac{n}{N} + \dots + \frac{n}{N}}_{M \text{ veces}} = M \frac{n}{N} = \frac{nM}{N}$$

Ejemplo

El espesor X de una cuña de madera (en milímetros) tiene una función de densidad de probabilidad

$$f(x) = \begin{cases} \frac{3}{4} - \frac{3(x-5)^2}{4} & 4 \leq x \leq 6 \\ 0 & \text{en otro lado} \end{cases}$$

- a) Determine $E(X)$
- b) Si Y denota el espesor de una cuña en pulgadas (1mm = 0.0394 pulgadas), determine $E(Y)$
- c) Si se seleccionan tres cuñas de manera independiente y las apilamos una encima de otra, encuentre la media y la varianza del espesor total.

a) Verifique el lector que

$$E(X) = \int_4^6 x \left(\frac{3}{4} - \frac{3}{4}(x-5)^2 \right) dx = 5$$

- b) $Y = 0.0394X$ entonces $E(Y) = E(0.0394X) = 0.0394E(X) = 0.197$
- c) Notar que si X_i : “espesor de cuña i ”, $i = 1, 2, 3$ entonces $X = X_1 + X_2 + X_3$ es el espesor total
 Por lo tanto $E(X) = E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 5 + 5 + 5 = 15$

En general **la esperanza de un producto de variables aleatorias no es igual al producto de las esperanzas**

Si (X, Y) es una variable aleatoria bidimensional tal que X e Y son variables aleatorias **independientes**, entonces: $E(X.Y) = E(X).E(Y)$

(leeremos:” la esperanza del producto es el producto de las esperanzas”).

Dem.) análoga a la demostración de la propiedad anterior.

Para el caso (X, Y) continua sigue siendo válida esta propiedad.

Ejemplo:

Supongamos que debido a innumerables causas incontrolables la corriente i y la resistencia r de un circuito varían aleatoriamente de forma tal que pueden considerarse como variables aleatorias I y R independientes. Supongamos que las correspondientes *fdp* son:

$$g(i) = \begin{cases} 2i & 0 \leq i \leq 1 \\ 0 & \text{demás valores} \end{cases} \qquad h(r) = \begin{cases} \frac{r^2}{9} & 0 \leq r \leq 3 \\ 0 & \text{demás valores} \end{cases}$$

Nos interesa considerar el voltaje $v = i.r$ de manera que podemos definir la variable aleatoria $V = I.R$. Hallar el valor esperado o esperanza matemática del voltaje: $E(V)$.

Como I y R son independientes, usando la propiedad anterior

$$E(V) = E(I)E(R)$$

$$E(I) = \int_0^1 i(2i)di = 2 \frac{i^3}{3} \Big|_0^1 = \frac{2}{3} \quad E(R) = \int_0^3 r \left(\frac{r^2}{9} \right) dr = \frac{1}{9} \frac{r^4}{4} \Big|_0^3 = \frac{1}{9} \times \frac{3^4}{4} = 1$$

$$\therefore E(V) = \frac{2}{3} \times 1 = \frac{2}{3}$$

Varianza de una suma de variables aleatorias

$$V(X + Y) = V(X) + V(Y) + 2\sigma_{XY} \quad \text{con} \quad \sigma_{XY} = E(X.Y) - E(X).E(Y)$$

Dem.) Escribimos la varianza en su forma alternativa

$V(X + Y) = E([X + Y]^2) - [E(X + Y)]^2$. Desarrollamos los cuadrados y aplicamos la propiedad lineal de la esperanza:

$$V(X + Y) = E(X^2 + 2.X.Y + Y^2) - [E(X) + E(Y)]^2 \\ = E(X^2) + 2E(X.Y) + E(Y^2) - \{[E(X)]^2 + 2E(X)E(Y) + [E(Y)]^2\}$$

Agrupando convenientemente:

$$V(X + Y) = \{E(X^2) - [E(X)]^2\} + \{E(Y^2) - [E(Y)]^2\} + 2\{E(X.Y) - E(X)E(Y)\} \quad , \text{ es decir} \\ = V(X) + V(Y) + 2\{E(X.Y) - E(X)E(Y)\}$$

$$V(X + Y) = V(X) + V(Y) + 2\sigma_{XY}$$

$$\sigma_{XY} = E(X.Y) - E(X).E(Y) \text{ se la llama la } \textit{covarianza de X e Y}.$$

Observaciones:

1- Teniendo presente la definición de la desviación estándar de una v.a. X : $\sigma_X = \sqrt{V(X)}$, vemos que a la propiedad anterior la podemos escribir:

$$V(X + Y) = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}$$

2- Análogamente se prueba que $V(X - Y) = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}$

3- X e Y son independientes, entonces $V(X + Y) = V(X - Y) = V(X) + V(Y)$

Esto es porque si las variables aleatorias X e Y son independientes, entonces $E(X.Y) = E(X).E(Y)$.

Por lo tanto la covarianza vale cero : $\sigma_{XY} = E(X.Y) - E(X).E(Y) = 0$.

4- Podemos generalizar, usando el principio de inducción completa, al caso de n variables aleatorias independientes:

Si X_1, X_2, \dots, X_n son n variables aleatorias independientes entonces:

$$V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n) \text{ o, en forma más compacta, } V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i).$$

5- Vemos que la esperanza de la suma de dos variables aleatorias X e Y es igual a la suma de las esperanzas $E(X + Y) = E(X) + E(Y)$ cualesquiera sean X e Y . En cambio la varianza de la suma de las variables aleatorias X e Y es, en general, igual a la suma de las varianzas, $V(X + Y) = V(X) + V(Y)$, sólo si X e Y son variables independientes.

Ejemplos:

1- Podemos ejemplificar la aplicación de las propiedades de la varianza, calculando nuevamente la varianza de una v.a. X distribuida binomialmente con parámetros n y p .

Sea entonces una v.a. $X \sim B(n, p)$. Vimos que se puede escribir:

$X = X_1 + X_2 + \dots + X_n$, donde las n variables aleatorias son independientes entre sí y tienen todas la misma distribución:

$$X_i \sim B(1, p) \quad \forall i = 1, 2, \dots, n$$

Entonces, tratándose de n variables aleatorias independientes

$V(X) = V(X_1) + V(X_2) + \dots + V(X_n)$ todas la varianzas son iguales y podemos escribir la suma como n veces una cualquiera de ellas:

$$V(X) = nV(X_i). \text{ Pero}$$

$$V(X_i) = E(X_i^2) - [E(X_i)]^2.$$

$$\text{Ya vimos que } E(X_i) = 1 \cdot p + 0(1 - p) = p$$

$$\text{Además es: } E(X_i^2) = 1^2 \cdot p + 0^2(1 - p) = p$$

$$\text{Entonces: } V(X_i) = E(X_i^2) - [E(X_i)]^2 = p - p^2 = p(1 - p).$$

Luego:

$$V(X) = nV(X_i) = np(1 - p)$$

que es el resultado que habíamos obtenido a partir de la definición y llevando las sumas involucradas a la forma del desarrollo de un binomio de Newton.

2- Varianza de una v.a. binomial negativa

Ya vimos que podemos escribir $X = X_1 + X_2 + \dots + X_r$, donde cada variable X_i tiene **distribución geométrica con parámetro p**

Por lo tanto

$$V(X) = V(X_1) + V(X_2) + \dots + V(X_r) = r \frac{1 - p}{p^2}$$

5.6 - Covarianza

Sean X e Y dos variables aleatorias. La **covarianza de X e Y** se define:

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\} = E(X \cdot Y) - E(X) \cdot E(Y)$$

Notación: la notación usual para la covarianza de X e Y es σ_{XY} o $Cov(X, Y)$

La última igualdad surge de desarrollar el producto y aplicar las propiedades de la esperanza:

$$E\{[X - E(X)][Y - E(Y)]\} = E\{X.Y - X.E(Y) - E(X).Y + E(X)E(Y)\}$$

Teniendo presente que $E(X)$ y $E(Y)$ son constantes:

$$E\{[X - E(X)][Y - E(Y)]\} = E(X.Y) - E(X).E(Y) - E(X)E(Y) + E(X)E(Y) = E(X.Y) - E(X).E(Y).$$

Si X e Y son variables aleatorias independientes, entonces $Cov(X, Y) = 0$.

Dem.)

Según vimos, si X e Y son variables aleatorias independientes, entonces $E(X.Y) = E(X).E(Y)$, de donde se sigue la propiedad.

Propiedades de la covarianza

Las siguientes propiedades son útiles y su verificación se deja como ejercicio

- 1- $Cov(a + bX, c + dY) = bdCov(X, Y)$
- 2- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$
- 3- $Cov\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m Cov(X_i, Y_j)$
- 4- $Cov(X, X) = V(X)$

Ejemplos:

1) ***Varianza de una v.a. hipergeométrica***

Si $X \sim H(n, M, N)$ entonces $V(X) = n \frac{M}{N} \left(\frac{N-M}{N} \right) \left(\frac{N-n}{N-1} \right)$

Para facilitar la demostración supongamos que tenemos N bolillas en una urna de las cuales M son rojas y $N-M$ son blancas. Queremos hallar la varianza del número de bolillas blancas extraídas

Como antes definimos las variables

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ésima bolilla roja es extraída} \\ 0 & \text{caso contrario} \end{cases}$$

Las variables X_1, X_2, \dots, X_M **no son independientes**

Se puede escribir $X = X_1 + X_2 + \dots + X_M$, además

$$E(X_i) = P(X_i = 1) = \frac{\binom{1}{1} \binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N} \quad \text{y}$$

$$V(X_i) = E(X_i^2) - (E(X_i))^2 = \frac{n}{N} - \left(\frac{n}{N}\right)^2 = \frac{n}{N} \left(1 - \frac{n}{N}\right)$$

Por lo tanto $V(X) = V(X_1 + X_2 + \dots + X_M) = \sum_{i=1}^M V(X_i) + 2 \sum_{i < j \leq M} Cov(X_i, X_j)$

Por otro lado $Cov(X_i; X_j) = E(X_i X_j) - E(X_i)E(X_j)$

Y $E(X_i X_j) = \frac{n(n-1)}{N(N-1)}$, entonces $Cov(X_i; X_j) = \frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2$

Aplicando algunos pasos algebraicos se llega a $Cov(X_i; X_j) = \frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2 = -\frac{n}{N} \left(\frac{1}{N-1}\right) \left(1 - \frac{n}{N}\right)$

Reemplazando

$$V(X) = \sum_{i=1}^M V(X_i) + 2 \sum_{i < j \leq M} Cov(X_i, X_j) = M \frac{n}{N} \left(1 - \frac{n}{N}\right) + 2 \left[\binom{M}{2} \left(-\frac{1}{N-1}\right) \frac{n}{N} \left(1 - \frac{n}{N}\right) \right]$$

Nuevamente, luego de algunos cálculos algebraicos se llega a

$$V(X) = n \frac{M}{N} \left(\frac{N-M}{N}\right) \left(\frac{N-n}{N-1}\right)$$

2) De una caja con frutas que contiene 3 naranjas, 2 manzanas y 3 plátanos se selecciona una muestra de 4 frutas.

Sean las variables aleatorias

X: “nº de naranjas extraídas”

Y: “nº de manzanas extraídas”

Notar que la *f.d.p.* conjunta de (X, Y) es

$$P(X = x, Y = y) = \frac{\binom{3}{x} \binom{2}{y} \binom{3}{4-x-y}}{\binom{8}{4}} \quad x = 0, 1, 2, 3 ; y = 0, 1, 2 ; 1 \leq x + y \leq 4$$

Es un ejemplo de *v.a. hipergeométrica bidimensional*.

También se podría haber presentado la *f.d.p.* conjunta en una tabla, donde también figuran las distribuciones marginales de X e Y.

a) ¿Cuales son E(X), V(X), E(Y) y V(Y)?

$$E(X) = 0 \times \frac{0}{70} + 1 \times \frac{3}{70} + 2 \times \frac{9}{70} + 3 \times \frac{3}{70} = \frac{105}{70} = \frac{3}{2}$$

$$E(Y) = 0 \times \frac{15}{70} + 1 \times \frac{40}{70} + 2 \times \frac{15}{70} = 1$$

Verifique el lector que $V(X) = \frac{15}{28}$ y $V(Y) = \frac{3}{7}$

X/Y	0	1	2	
0	0	2/70	3/70	5/70
1	3/70	18/70	9/70	30/70
2	9/70	18/70	3/70	30/70
3	3/70	2/70	0	5/70
	15/70	40/70	15/70	

b) ¿Son X e Y independientes?

$$P(X = 0, Y = 0) = 0 \quad \text{pero} \quad P(X = 0)P(Y = 0) = \frac{5}{70} \frac{15}{70} \neq 0$$

Por lo tanto X e Y son dependientes, lo que implica que $Cov(X, Y) \neq 0$

c) ¿Cuál es la $Cov(X, Y)$?

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

$$\begin{aligned} E(XY) = & 0 \times 0 \times 0 + 0 \times 1 \times \frac{2}{70} + 0 \times 2 \times \frac{3}{70} + \\ & + 1 \times 0 \times \frac{3}{70} + 1 \times 1 \times \frac{18}{70} + 1 \times 2 \times \frac{9}{70} + \\ & + 2 \times 0 \times \frac{9}{70} + 2 \times 1 \times \frac{18}{70} + 2 \times 2 \times \frac{3}{70} + \\ & + 3 \times 0 \times \frac{3}{70} + 3 \times 1 \times \frac{2}{70} + 3 \times 2 \times 0 = \frac{90}{70} \end{aligned}$$

$$\text{Entonces } Cov(X, Y) = E(XY) - E(X)E(Y) = \frac{9}{7} - \frac{3}{2} \times 1 = -\frac{3}{14}$$

d) $Z = X+Y$ simboliza el total de naranjas y manzanas extraídas

¿Cuál es la $E(Z)$ y $V(Z)$?

$$E(Z) = E(X + Y) = E(X) + E(Y) = \frac{3}{2} + 1 = 2.5$$

$$V(Z) = V(X) + V(Y) + 2Cov(X, Y) = \frac{15}{18} + \frac{3}{7} + 2 \times \left(-\frac{3}{14} \right) = \frac{15}{28}$$

e) Supongamos que cada naranja cuesta 2\$ y cada manzana cuesta 1.5\$ entonces

$W = 2X + 1.5Y$ es el costo del total de frutas extraídas. Hallar $E(W)$ y $V(W)$

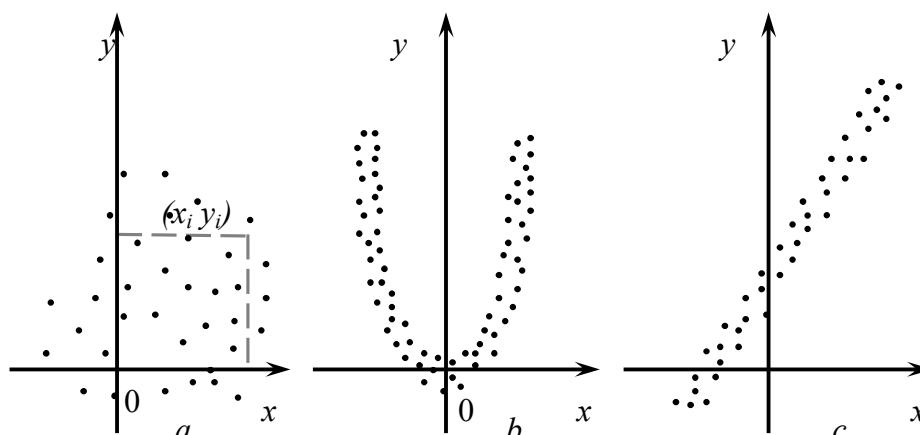
$$E(W) = E(2X + 1.5Y) = 2E(X) + 1.5E(Y) = 4.5\$$$

$$V(W) = V(2X + 1.5Y) = 2^2V(X) + 1.5^2V(Y) + 2 \times 2 \times 1.5Cov(X, Y) = \frac{51}{28}$$

5.7 - Coeficiente de correlación lineal.

En realidad más que la covarianza aquí nos interesa considerar una cantidad relacionada con σ_{XY} y que según veremos nos dará información sobre el grado de asociación que existe entre X e Y . Más concretamente nos contará si existe algún grado de relación lineal entre X e Y . Esa cantidad es el coeficiente de correlación lineal.

En el mismo sentido en que podemos tener una idea aproximada sobre la probabilidad de un suceso A si repetimos el experimento y consideramos las ocurrencias de A en las n repeticiones, así podemos tener también una primera idea sobre la existencia de una relación funcional, específicamente una relación lineal, entre X e Y si consideramos un **diagrama de dispersión**. Consiste en dibujar pares de valores (x_i, y_j) medidos de la variable aleatoria (X, Y) en un sistema de coordenadas. En la figura mostramos diversas situaciones posibles.



De la figura *a* se deduciría que entre X e Y no hay ningún tipo de relación funcional. La figura *b* sugiere la posibilidad de que exista una relación funcional que corresponde a una parábola. La figura *c*, por su parte, sugiere una relación lineal entre X e Y . Este último es el comportamiento que nos interesa caracterizar. Con ese fin definimos el coeficiente de correlación lineal como sigue:

Sea (X, Y) una variable aleatoria bidimensional. Definimos el **coeficiente de correlación lineal entre X e Y** como
$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

En consecuencia:

$$\rho_{XY} = \frac{E\{[X - E(X)][Y - E(Y)]\}}{\sqrt{V(X)V(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}}$$

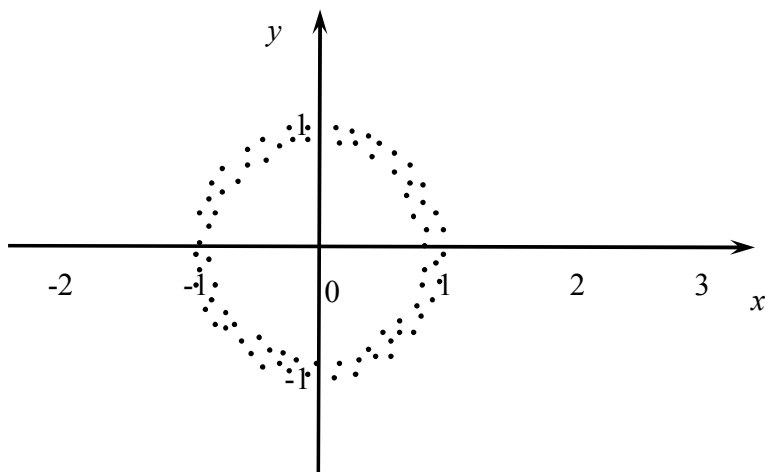
Daremos una serie de propiedades de ρ_{XY} que nos permitirán establecer más concretamente su significado.

Propiedad 1

Si X e Y son variables aleatorias independientes entonces $\rho_{XY} = 0$.

Dem.) inmediata a partir del hecho que si X e Y son independientes entonces $E(XY) = E(X)E(Y)$

Observación: La inversa no es necesariamente cierta. Puede ser que $\rho_{XY} = 0$ y sin embargo X e Y no sean variables aleatorias independientes. En efecto si tenemos una v.a. bidimensional (X, Y) que da lugar a un diagrama de dispersión como el que se muestra en la figura, veremos que correspondería a un coeficiente de correlación lineal $\rho_{XY} = 0$ y sin embargo la figura sugiere que entre X e Y existe la relación funcional $X^2 + Y^2 = 1$, es decir X e Y son v.a. dependientes. En realidad, como veremos, ρ_{XY} es una medida de la existencia de una relación lineal entre X e Y y una circunferencia se aleja mucho de una línea recta.



Propiedad 2 :

$$-1 \leq \rho_{XY} \leq 1$$

Dem.)

Si consideramos la v.a. $\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}$ entonces

$$0 \leq V\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = \frac{V(X)}{\sigma_X^2} + \frac{V(Y)}{\sigma_Y^2} + \frac{2Cov(X,Y)}{\sigma_X\sigma_Y} = 2(1 + \rho_{XY})$$

Implicando que $-1 \leq \rho_{XY}$

Por otro lado: $0 \leq V\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = \frac{V(X)}{\sigma_X^2} + \frac{V(Y)}{\sigma_Y^2} - \frac{2Cov(X,Y)}{\sigma_X\sigma_Y} = 2(1 - \rho_{XY})$

Implicando que $\rho_{XY} \leq 1$

$$\therefore -1 \leq \rho_{XY} \leq 1$$

Propiedad 3 :

Si $\rho_{XY}^2 = 1$, entonces con probabilidad 1 es $Y = A.X + B$ donde A y B son constantes.

Dem.) Si $\rho_{XY}^2 = 1$ entonces $\rho_{XY} = 1$ o $\rho_{XY} = -1$

Si $\rho_{XY} = -1$ entonces de la demostración anterior se deduce que

$V\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 2(1 + \rho_{XY}) = 0$, lo que implica que la v.a. $Z = \frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}$ tiene varianza cero. Según la interpretación de varianza podemos deducir (en forma intuitiva) que la v.a. **no tiene dispersión con respecto a su esperanza**, es decir la v.a. **Z es una constante con probabilidad 1**

Por lo tanto esto implica que $Y = A.X + B$ con $A = -\frac{\sigma_Y}{\sigma_X} < 0$

Análogamente $\rho_{XY} = 1$ implica que $Y = A.X + B$ con $A = \frac{\sigma_Y}{\sigma_X} > 0$

Propiedad 4 :

Si X e Y son dos variables aleatorias tales que $Y = A.X + B$, donde A y B son constantes, entonces $\rho_{XY}^2 = 1$. Si $A > 0$ es $\rho_{XY} = 1$ y si $A < 0$ es $\rho_{XY} = -1$.

Dem.) se deja como ejercicio

Observación: Claramente las propiedades anteriores establecen que el coeficiente de correlación lineal es una medida del grado de linealidad entre X e Y .

Ejemplo

En el ejemplo anterior

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{-\frac{3}{14}}{\sqrt{\frac{15}{28} \cdot \frac{3}{7}}} = -\frac{\sqrt{5}}{5} = -0.44721$$