APUNTES PROBABILIDAD

Tema 2: Vectores Aleatorios

Doble Grado en Informática y Estadística (INdat)

Autor
Juan Horrillo Crespo
Universidad de Valladolid

October 28, 2024

Contents

1 Distribución conjunta de un vector aleatorio

1.1 Modelos para vectores aleatorios

Cuando estudiamos experimentos aleatorios podemos tener interes en la variación conjunta de dos o más características o variables aleatorias del mismo.
Estas agrupaciones de varias características se conocen como vectores aleatorios. Por tanto el conjunto de vectores aleatorios posibles será ahora un subconjunto de $\mathbb{R}^n$, siendo n la dimensión del vector aleatorio, osea el número de caracteristicas estudiadas.
Ademas de distribuciones discretas, continuas y mixtas, es posible encontrar algunas que no entre en ninguno de los anteriores grupos, pero nuestro estudio se centrará en las distribuciones discretas y continuas.

Llamaremos *distribución conjunta del vector* al modelo probabilistico asociado a un vector aleatorio. Un vector aleatorio se describe como $X = (X_1, ..., X_n)$ siendo el resultado de la consideración conjunta de variables aleatorioas. Llamaremos *distribución marginal* X_i a la distribución asociada a la caracteristica descrita por dicha variable por separado.
Es posible crear un vector mediante otros vectores de la forma $X = (X_a, X_b)$; en estos casos, para referirnos a la distribución de X_a tambíen hablatemos de la *Distribución marginal* de X_a.

1.2 Vectores aleatorios discretos

Los modelos multivariantes discretos (modelos para vectores aleatorios) reproducen las características de los modelos univariantes:
El conjunto de los valores posibles de $X = (X_1, ..., X_n)$ es un conjunto discreto de elementos de $\mathbb{R}^n$ y la probabilidad se concentra en esos elementos. Los sucesos no elementales se calculan mediante las sumas de los elementales.
Emplearemos la notación $X_1 = x_1, ..., X_n = x_n$ para referirnos al suceso consistente en el que el vector aleatorio $X = (X_1, ..., X_n)$ toma el valor $x = (x_1, ..., x_n)$. De forma analoga, $(X_1 \in A_1, ..., X_n \in A_n)$ denota el suceso consistente en que el valor de X es un elemento de $A_1 \times ... \times A_n$. Es inmediato entonces

$$(X_1 \in A_1, ..., X_n \in A_n) = (X_1 \in A_1) \cap ... \cap (X_n \in A_n)$$

A la función

$$p(x_1, ..., x_n) = P(X_1 = x_1, ..., X_n = x_n), \qquad (x_1, ..., x_n) \in \mathbb{R}^n$$

la llamaremos *función de masa de probabilidad conjunta* del vector X.

Con esta notación los modelos probabilísticos para vectores aleatorios discretos pueden resumirse de la siguiente forma:

$X = (X_1, ..., X_n), \quad \Omega = \{x^1, x^2, ..., x^i, ...\}, \quad x^i = (x^i_1, ..., x^i_n) \in \mathbb{R}^n$
$P(X_1 = x_1, ..., X_n = x_n) = p(x_1, ..., x_n);$
$p(x_1, ..., x_n) \geq 0, \qquad \sum_{(x_1,...,x_n)\in\Omega} p(x_1, ..., x_n) = 1$

$$P((X_1, ..., X_n) \in A) = P\left(\bigcup_{(x_1,...,x_m)\in A} (X_1 = x_1, ..., X_n = x_n)\right) = \sum_{(x_1,...,x_n)\in A} p(x_1, ..., x_n)$$

1.3 Distribución conjunta de un vector aleatorio bidimensional

Dadas X e Y dos variables aleatorias definidas sobre el mismo espacio muestral Ω, la aplicación $(X, Y) : \Omega \to \mathbb{R}^2$ es un vector aleatorio bidimensional.
La distribución que describe simultáneamente el comportamiento de X e Y se llama *distribución de probabilidad conjunta*.

Dado un vector aleatorio (X, Y) y dados $x, y \in \mathbb{R}$, la *función de distribución conjunta* de (X, Y) evaluada en (x, y) se define como:

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = P((X \leq x) \cap (Y \leq y))$$

Esta función es útil principalmente en los casos continuos y mixtos.

1.4 Vectores aleatorios discretos

Dadas X e Y dos variables aleatorias discretas, el vector (X, Y) será discreto y tiene *función de probabilidad conjunta*

$$p(x, y) = P(X = x, Y = y)$$

Por tanto tenemos las siguientes propiedades

1. $p(x, y) \geq 0$
2. $\sum_x \sum_y p(x, y) = 1$

y una *función de distribución conjunta*

$$F(x_0, y_0) = P(X \leq x_0, Y \leq y_0) = \sum_{x\leq x_0} \sum_{y\leq y_0} p(x, y)$$

Ejemplo 1.1. Consideremos el experimento aleatorio consistente en seleccionar al azar y sin reemplazamiento dos números del conjunto $\{1, 2, 3, 4, 5, 6, 7, 8\}$. Centremos nuestra atención en el vector aleatorio (X, Y) en el que X es el mayor número elegido e Y es el menor. Los posibles valores de (X, Y) son los pares de números enteros (i, j) que verifican $1 \leq j < i \leq 8$. Si (i, j) verifica esas condiciones entonces

$$P(X = i, Y = j) = \frac{1}{\binom{8}{2}} = \frac{1}{28}$$

(el experimento consiste, esencialmente, en escoger una subpoblación de tamaño 2 de una población con 8 elementos).

A menudo es conveniente recoger en una tabla de doble entrada la función de masa de probabilidad de un vector aleatorio bidimensional. La tabla correspondiente a este ejemplo es la siguiente:

X \ Y	1	2	3	4	5	6	7
2	$\frac{1}{28}$	0	0	0	0	0	0
3	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0	0	0
4	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0	0
5	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0
6	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0
7	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0
7	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$

A partir de la función masa podemos calcular la probabilidad de cualquier otro suceso relacionado con (X, Y). Por ejemplo, si A es el suceso ”la diferencia entre los dos números elegidos es menor que 5”, entonces

$$P((X, Y) \in A) = P(X - Y < 5) = \sum_{(i,j) \in A} p(i, j)$$

es decir, que $P(X - Y < 5$ se calcula sumando los valores de la función de masa de probabilidad sombreados en la siguiente tabla

X \ Y	1	2	3	4	5	6	7
2	$\frac{1}{28}$	0	0	0	0	0	0
3	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0	0	0
4	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0	0
5	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0
6	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0
7	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0
7	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$

Por lo tanto $P(X - Y < 5) = \frac{22}{28} \approx 0,7857$

1.5 Vectores aleatorios continuos

Diremos que el vector aleatorio $X = (X_1, ..., X_n)$ es continuo y con función de densidad $f(x_1, ..., x_n)$ si las probabilidades de los sucesos de interés dentro del modelo se calculan segun la regla

$$P(X \in A) = \int_A f(x_1, ..., x_n) dx_1 \cdots dx_n$$

Para que la anterior expresión defina un modelo probabilistico correcto la función f debe verificar las siguientes condiciones

1. $f(x_1, ..., x_n) \geq 0, \qquad (x_1, ..., x_n) \in \Omega \subset \mathbb{R}^n$
2. $\int_\Omega f(x_1, ..., x_n) dx_1 \cdots dx_n = 1$

donde Ω, el conjunto de valores posibles de X, es subconjunto de $\mathbb{R}^n$.

1.5.1 Vector aleatorio continuo bidimensional

Dados X e Y dos variables aleatorias continuas, el vector aleatorio (X, Y) será contiuo y tiene *función de densidad conjunta* $f(x, y)$ que satisface

1. $f(x, y) \geq 0$;
2. $\int \int f(x, y) dx dy = 1$

y una *función de distribución conjunta* descrita como

$$F(x_0, y_0)) P(X \leq x_0, Y \leq y_0) = \int_{-\infty}^{x_0} \int_{-\infty}^{y_0} f(x, y) dy dx$$

Para calcular la probabilidad de que un suceso esté entre dos valores (a, c) y (b, d) tenemos que

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y) dy dx$$

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

y ademas

$$f(x, y) = \frac{\delta^2}{\delta x \delta y} F(x, y)$$

Aquí se pueden ver algunos ejemplos de funciones de densidad variables. El analisis de la probabilidad consiste en calcular el volumen encerrado bajo la superficie.

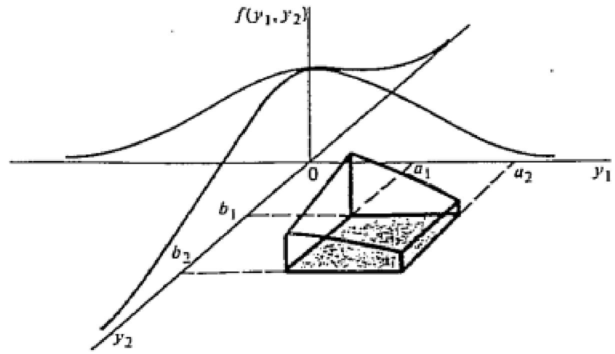

Figure 1: Una función de densidad bivariable

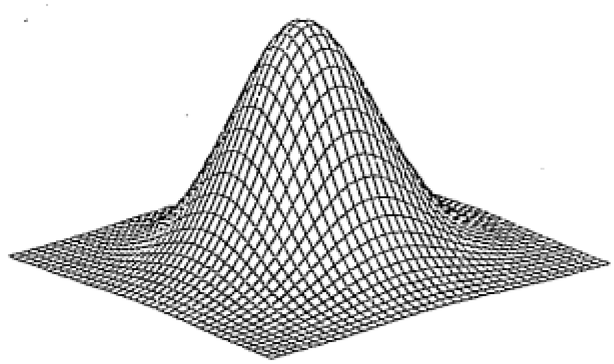

Figure 2: Densidad normal bivariable

Ejemplo 1.2. Consideremos el experimento consistente en seleccionar un punto al azar en un círculo de radio unidad (suponemos centrado en $(0,0)$). Llamaremos X e Y a la abcisa y a la ordenada del punto elegido. La elección puramente al azar se puede modelizar con al función de densidad

$$f(x,y) = \left\{ \begin{array}{ccc} \frac{1}{\pi} & si & x^2+y^2 \leq 1 \\ 0 & si & x^2+y^2 > 1 \end{array} \right\}$$

Esta función es una función de densidad. Fácilmente comprobamos que su integral es uno:

$$\int_{\{(x,y)\big/ x^2+y^2 \leq 1\}} f(x,y)dxdy = \int_{-1}^{1} \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy \right) dx =$$

$$= \frac{1}{\pi} \int_{-1}^{1} 2\sqrt{1-x^2}dx = \frac{\pi}{\pi} = 1$$

El cálculo de probabilidades dentro de este modelo se efectúa integrando la función de densidad en la región adecuada. Consideremos los sucesos $A =$ *"el punto elegido dista del origen menos de 0.5 unidades"* y $B =$ *"la abcisa del punto elegido está dentro del intervalo* $(0.3, 0.8)$*"*.

$$P((X,Y) \in A) = P(X^2+Y^2 < 0.5^2) = \int_{\{(x,y)\big/ x^2+y^2 < 0.25\}} f(x,y)dxdy =$$

$$= \int_{-0.5}^{0.5} \left(\int_{-\sqrt{0.25-x^2}}^{\sqrt{0.25-x^2}} \frac{1}{\pi} dy \right) dx = \frac{1}{\pi} \int_{-0.5}^{0.5} 2\sqrt{0.25-x^2}dx = \frac{\pi}{4\pi} = 0.25$$

$$P((X,Y) \in B) = P(0.3 < X < 0.8) = \int_{\{(x,y)\big/ 0.3 < x < 0.8\}} f(x,y)dxdy =$$

$$= \int_{0.3}^{0.8} \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy \right) dx = \frac{1}{\pi} \int_{0.3}^{0.8} 2\sqrt{1-x^2}dx =$$

$$\frac{1}{\pi} \left[\arcsin x + x\sqrt{1-x^2} \right]_{x=0.3}^{x=0.8} \approx 0.2599$$

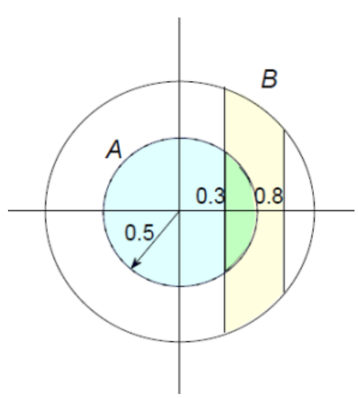

Figure 3: Circunferencia de radio unidad y los conjuntos A y B

2 Distribuciones marginales y condicionadas

2.1 Distribuciones marginales

A la distrubución de cada variable de las que componen un vector aleatorio se le denomina *distribución marginal*. Dado el vector aleatorio (X, Y) podemos hablar de la *distribución marginal de* X y de la *distribución marginal de* Y.

2.1.1 Variables discretas

Dadas X e Y variables discretas con función de probabilidad conjunta $p(x, y)$, las funciones de probabilidad marginales de ambas variables son:

$$p_X(x) = P(X = x) = \sum_y P(X = x, Y = y) = \sum_y p(x, y)$$

$$p_Y(y) = P(Y = y) = \sum_y P(X = x, Y = y) = \sum_x p(x, y)$$

Ejemplo 2.1. Podemos obtener la distribución marginal de X a partir de la tabla con la función de masa de probabilidad conjunta sumando los valores por filas. Análogamente, sumando los valores de las columnas obtendremos la de Y:

X \ Y	1	2	3	4	5	6	7	P_X
2	$\frac{1}{28}$	0	0	0	0	0	0	$\frac{1}{28}$
3	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0	0	0	$\frac{2}{28}$
4	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0	0	$\frac{3}{28}$
5	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0	0	$\frac{4}{28}$
6	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	0	$\frac{5}{28}$
7	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	0	$\frac{6}{28}$
7	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{1}{28}$	$\frac{7}{28}$
P_Y	$\frac{7}{28}$	$\frac{6}{28}$	$\frac{5}{28}$	$\frac{4}{28}$	$\frac{3}{28}$	$\frac{2}{28}$	$\frac{1}{28}$	

Por tanto

$$P(X = 3) = p_X(3) = \frac{3}{28}$$

$$P(Y = 2) = p_Y(2) = \frac{6}{28}$$

2.2 Variables continuas

Dadas X e Y variables continuas con función de densidad conjunta $f(x,y)$, las funciones de densidad marginales de ambas variables son

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

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

Ejemplo 2.2. Continuando con el vector (X,Y) del ejemplo continuo anterior cuya densidad conjunta uniforme sobre el círculo unidad era

$$f(x,y) = \left\{ \begin{array}{ccc} \frac{1}{\pi} & si & x^2+y^2 \leq 1 \\ 0 & si & x^2+y^2 > 1 \end{array} \right\}$$

La densidad marginal para X es:

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} f(x,y)dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{dy}{\pi} = \frac{2}{\pi}\sqrt{1-x^2}, \quad -1 < x < 1$$

2.3 Distribuciones condicionadas

La idea de una probabilidad condicionada se puede incorporar a los modelos de vectores aleatorios. Supongamos que el experimento de interés se recoge en el vector aleatorio $X = (X_a, X_b)$ del que sabemos que el subvector X_a toma el valor x_a.
El grado de creencia en los sucesos relativos a la otra parte del experimento, X_b, se verá modificado. A la distribución de X_b modificada para incorporar esta información la llamaremos *distribución de X_b condicionada a X_a*. En esta sección se estudiará la manera en la que se actualiza la distrubución de un vector aleatorio en estas circunstancias, considerando por separado los casos discreto y continuo.

2.3.1 Caso discreto

Las distribuciones condicionadas se obtienen a partir de la definición de probabilidad condicionada de sucesos. Supongamos por ejemplo que ha ocurrido $(X_a = x_a)$. La probabilidad de que $(X_b = x_b)$ será

$$P\left(X_b = x_b \middle/ X_a = x_a\right) = \frac{P(X_a = x_a, X_b = x_b)}{P(X_a = x_a)} = \frac{p(x_a, x_b)}{p_{x_a}(x_a)}$$

A la función de masa de probabilidad que incorpora la información $(X_a = x_a)$ se le llama *función de masa de probabilidad condicionada por x_a* y se denota $p(x_b|x_a)$, es decir

$$p\left(x_b \middle/ x_a\right) = P\left(X_b = x_b \middle/ X_a = x_a\right)$$

Esta función está bien definida para

$$P(X_a = x_a) > 0 \qquad p\left(x_b \Big/ x_a\right) = \frac{p(x_a, x_b)}{p_{x_a}(x_a)}$$

Esta última igualdad reescrita de la forma siguiente

$$p(x_a, x_b) = p_{x_a}(x_a)p\left(x_b \Big/ x_a\right)$$

o, equivalente, en la forma

$$P(X_a = x_a, X_b = x_b) = P(X_a = x_a)P\left(X_b = x_b \Big/ X_a = x_a\right)$$

se conoce como *"Regla de la multiplicación"*. Esta regla nos permite obtener la distribución conjunta a partir de las distribuciones marginales y condicionadas.

Ejemplo 2.3. Se lanza un dado. Lllamaremos X al resultado obtenido. A continuación se lanzan X monedas. Representaremos por Y el número de caras obtenidas. Calcula la distribución de X, la distribución de Y condicionada por X, la distribución conjunta de (X, Y) y la distribución marginal de Y.
Obviamente la dsitribución de X es $P(X = i) = \frac{1}{6}, \quad i = 1, 2, \ldots, 6$. Si se lanzan i monedas, el número de resultados diferentes que podemos obtener es 2^i (Combinatoria de i elementos tomados de 2 en 2).
De esas 2^i posibilidades, en $\binom{i}{j}$ se registran j caras. Asumiendo que las monedas no estan sesgadas podemos aplicar la *Regla de Laplace* para obtener

$$P\left(Y = j \Big/ X = i\right) = \binom{i}{j}\frac{1}{2^i}$$

para $0 \leq j \leq i$. La distribución conjunta es, por tanto,

$$P(X = i, Y = j) = P(X = i)P\left(Y = j \Big/ X = i\right) = \frac{1}{6}\binom{i}{j}\frac{1}{2^i}$$

para $i = 1, 2, \ldots, 6$ y $0 \leq j \leq i$. Podemos expresar la distribución conjunta en la siguiente tabla:

X \ Y	0	1	2	3	4	5	6
1	$\frac{1}{6}\frac{1}{2}$	$\frac{1}{6}\frac{1}{2}$	0	0	0	0	0
2	$\frac{1}{6}\frac{1}{4}$	$\frac{1}{6}\frac{5}{4}$	$\frac{1}{6}\frac{1}{4}$	0	0	0	0
3	$\frac{1}{6}\frac{1}{8}$	$\frac{1}{6}\frac{3}{8}$	$\frac{1}{6}\frac{3}{8}$	$\frac{1}{6}\frac{1}{8}$	0	0	0
4	$\frac{1}{6}\frac{1}{16}$	$\frac{1}{6}\frac{4}{16}$	$\frac{1}{6}\frac{6}{16}$	$\frac{1}{6}\frac{4}{16}$	$\frac{1}{6}\frac{1}{16}$	0	0
5	$\frac{1}{6}\frac{1}{32}$	$\frac{1}{6}\frac{5}{32}$	$\frac{1}{6}\frac{10}{32}$	$\frac{1}{6}\frac{10}{32}$	$\frac{1}{6}\frac{5}{32}$	$\frac{1}{6}\frac{1}{32}$	0
6	$\frac{1}{6}\frac{1}{64}$	$\frac{1}{6}\frac{6}{64}$	$\frac{1}{6}\frac{15}{64}$	$\frac{1}{6}\frac{20}{64}$	$\frac{1}{6}\frac{15}{64}$	$\frac{1}{6}\frac{6}{64}$	$\frac{1}{6}\frac{1}{64}$
$P(Y)$	0.1641	0.3125	0.2578	0.1667	0.0755	0.0208	0.0026

Sumando por columnas obtendremos la distribución marginal de Y, recogida en el margen inferior de la tabla

2.3.2 Caso continuo

En el caso continuo la distribución condicionada es también continua. A la función de densidad correspondiente se le llama función de densidad condicionada y se denota mediante cualquiera de los dos simbolos siguientes:

$$f\left(x_a \Big/ x_b\right) \quad \text{o} \quad f_{X_b/X_a=x_a}(x_b)$$

Cualquiera de los dos representa la densidad el vector aleatorio X_b cuando se incorpora al modelo de información consistente en que X_a toma el valor x_a.
Se puede comprobar que la densidad condicionada se calcula de forma análoga a la función de densidad condicionada del caso discreto:

$$f\left(x_a \Big/ x_b\right) = \frac{f(x_a, x_b)}{f_{X_a}}(x_a)$$

es decir, que la densidad condicionada se obtiene dividiendo la densidad conjunta por la densidad marginal de la variable (o vector) por la que condicionamos.
Igual que en el caso discreto, reescribir convenientemente la igualdad anterior nos conduce a una nueva *"Regla de la multiplicación"*:

$$f(x_a, x_b) = f_{X_a}(x_a) f\left(x_a \Big/ x_b\right)$$

Ejemplo 2.4. Consideremos el experimento aleatorio consistente en elegir un punto al azar, x, en el intervalo $(0, 1)$. Una vez elegido x escogemos otro punto y al azar en el intervalo $(0, x)$. Llamemos (X, Y) al vector aleatorio cuyas coordenadas son los puntos elegidos por el procedimiento anterior. Calcular

a) la densidad conjunta de (X, Y)

b) la densidad marginal de Y

c) la probabilidad de que Y tome un valor menor que 0.8

// Falta la resolución de a)
Teniendo $f(x, y) = \frac{1}{x}$ función de densidad conjunta, podemos calcular la densidad marginal de Y:

$$f_Y(y) = \int_y^1 f(x, y)dx = \int_y^1 \frac{1}{x}dx = -\log y, \quad 0 < y < 1$$

La función de distribución asociada a Y es

$$F_Y(y) = \int_0^y f_y(t)dt = y(1 - \log_y), \quad 0 < y < 1$$

La probabilidad que se pide en el apartado c) es fácil de calcular a partir de F_Y:

$$P(Y < 0.8) = F_Y(0.8) = 0.8(1 - \log 0.8) \approx 0.9785$$

3 Distribución multinomial

Dado un experimento aleatorio con k resultados posibles, de tal modo que la probabilidad de cada resultado $p_1, p_2, \ldots, p_k$ se mantiene constante, un vector aleatorio $(X_1, X_2, \ldots, X_k)$ sigue distribución multinomial si cada X_i representa el número de veces que ocurre el resultado i-esimo en n repeticiones independientes del experimento.
Dados $0 \leq x_1, x_2, \ldots, x_k \leq n$ y $0 \leq p_1, p_2, \ldots, p_k \leq 1$ con $\sum_{i=1}^{k} x_i = n$ y $\sum_{i=1}^{k} p_i = 1$, si $(X_1, X_2, \ldots, X_k)$ sigue distribución multinomial, tenemos

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

Propiedades:

- Todas las distribuciones marginales de un vector multinomial son multinomiales (binomiales en su caso)
- Todas las distribuciones condicionales de un vector multinomial son multinomiales (binomiales en su caso)

Por ejemplo, podemos decir que, si $(X_1, \ldots, X_k) \sim \mathcal{M}(n; p_1, \ldots, p_k)$ entonces:

- $X_i \sim b(n, p_i), \quad i = 1, \ldots, k$
- $X_1 + \cdots + X_i \sim b(n, p_1 + \cdots + p_k), \quad i = 1, \ldots, k$

4 Vectores mixtos

Sea (X, Y) vector aleatorio tal que conocemos la distribución de Y condicionada a $X = x_0$ y la distribución de X, en los siguentes casos, ¿cómo podemos obtener la ley de Y?

- X e y son v.a. discretas
- X e Y son v.a. continuas
- X es discreta e Y es continua
- X es continua e Y es discreta

5 Independencia entre variables aleatorias

Dos variable aleatorias X e Y se dicen *independientes* si el conocimiento del valor que toma una, no nos aporta información sobre el valor que tomará la otra.
Para cualesquiera $A, B \subset \mathbb{R}$,

$$P((X \in A) \cap (Y \in B)) = P(X \in A) P(Y \in B)$$

Equivalentemente, para cualesquiera $x, y \in \mathbb{R}$

$$P((X \leq x) \cap (Y \leq y)) = P(X \leq x) P(Y \leq y)$$

Propiedad: Si X e Y son independientes, entonces

$$Var[X+Y] = Var[X] + Var[Y]$$

- Variables aleatorias **discretas** son Indep. si
  1. $p(y|x) = p_Y(y)$ para cualesquiera x, y o bien
  2. $p(x|y) = p_X(x)$ para cualesquiera x, y o bien
  3. $p(x,y) = p(x|y)p_Y(y) = p_X(x)p_Y(y)$
- Variables aleatorias **continuas** son Indep. si
  1. $f(y|x) = f_Y(y)$ para cualesquiera x, y o bien
  2. $f(x|y) = f_X(x)$ para cualesquiera x, y o bien
  3. $f(x,y) = f(x|y)f_Y(y) = f_X(x)f_Y(y)$

6 Características de un vector aleatorio

6.1 Esperanza

El vector de medias de X es el vector cuyas componentes son las esperanzas de cada componente de X

$$E[X] = \begin{pmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_n] \end{pmatrix}$$

6.2 Covarianza

Cuando tenemos que estudiar la variación conjunta de dos variables aleatorias que forman un vector aleatorio, la descripción individual de cada una de las variables por medio de los correspondientes parámetros poblacionales resulta insuficiente como resumen numérico del vector aleatorio pues no informa de un aspecto tan importante como es la asociación entre las dos variables.
Las versiones poblacionales de las medidas de asociación que se manejan en Estadística Descriptiva nos servirán ahora para cuantificar la asociación entre variables aleatorias estudiadas sobre la misma población.

Covarianza: La covarianza de dos variables aleatorias X e Y se denota $Cov(X,Y)$ ó $\sigma_{X,Y}$ y se define como el valor medio de los productos cruzados de las distancias de los valores de cada una de las variables respecto a su media, es decir

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

Para el *Caso discreto* tenemos

- $\Omega = \{(x_i, y_j)/i, j = 1, 2, \dots\} \subset \mathbb{R}^2$
- f.m.p: $p(x_i, y_j)$

$$\sigma_{X,Y} = \sum_{i,j} (x_i - \mu_X)(y_j - \mu_Y) p(x_i, y_j)$$

En cambio para el *Caso continuo* tenemos

- $\Omega = \subset \mathbb{R}^2$
- f.dens: $f(x, y)$

$$\sigma_{X,Y} = \int (x - E[X])(y - E[Y]) f(x, y) dx$$

Dada la definición de covarianza de un vector podemos realizar las siguientes observaciones:

- La varianza indica el sentido de la asociación a través del signo:
  1. $Cov(X, Y) > 0 \Rightarrow$ asociación lineal creciente
  2. $Cov(X, Y) < 0 \Rightarrow$ asociación lineal decreciente
  3. $Cov(X, Y) = 0 \Rightarrow$ indica la ausencia de asociación lineal. En particular, si (X, Y) son variables aleatorias independientes se prueba facilmente que $E[XY] = E[X]E[Y]$, con lo que $Cov(X, Y) = 0$
- La covarianza es invariante por cambios de localización de las variables, pero no por cambios de escala:

$$(X, Y)\text{vector aleatorio}, \quad (X', Y') = (aX + b, cY + d), a, c > 0 \implies$$
$$\implies Cov(X', Y') = acCov(X, Y)$$

- La covarianza no se puede usar directamente como medida de asociación lineal porque su valor depende de las unidades de medida de las variables.

La covarianza tiene las siguientes **propiedades**:

- $Cov(X, Y) = E[XY] - E[X]E[Y]$
- si X e Y son independientes, entonces $Cov(X, Y) = 0$
- $Cov(X, Y) = 0$ no garantiza que X e Y sean independientes
- $Cov(X, X) = Var(X)$
- $Cov(aX + b, cY + d) = acCov(X, Y), \quad a, c > 0$
- $Cov(\sum_i X_i, \sum_j Y_j) = \sum_i \sum_j Cov(X_i, Y_j)$
- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$
- $Var(\sum_i X_i) = \sum_i Var(X_i) + \sum_i \sum_k Cov(X_i, X_k)$

6.3 Correlación

El **Coeficiente de Correlación** es una medida de asociación lineal entre variables aleatorias que se define a partir de la covarianza y evitando los inconvenientes de esta:

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

Dada la definición de correlación de un vector podemos realizar las siguientes observaciones:

- El coeficiente de correlación es invariante por cambios de locaclización y escala en las variables:

$$(X,Y)\text{vector aleatorio}, (X',Y') = (aX+b, cY+d), a,c > 0 \implies$$
$$\implies \rho(X',Y') = \rho(X,Y)$$

- El coeficiente de correlación es adimensional, no tiene unidades. De hecho, el coeficiente de correlación es igual a la covarianza de las variables tipificadas:

$$(X,Y)\text{vector aleatorio}, \quad Z_X = \frac{X-\mu_X}{\sigma_X}, Z_Y = \frac{Y-\mu_Y}{\sigma_Y} \implies$$
$$\implies \rho(X,Y) = Cov(Z_X, Z_Y)$$

- El valor del coeficiente de correlación está acotado:

$$|\rho(X,Y)| \leq 1$$

La correlación tiene las siguientes **propiedades**:

- Si X e Y son independientes, entonces $\rho(X,Y) = 0$
- $\rho(X,Y) = 0$ no garantiza que X e Y sean independientes
- $|\rho(X,Y)| \leq 1$
- Si $a,c > 0$, entonces $\rho(aX+b, cY+d) = \rho(X,Y)$
- $|\rho(X,X)| = 1$
- Si $a > 0$ entonces $|\rho(X, aX+b)| = 1$

6.4 Matriz de covarianzas

La matriz de varianzas y covarianzas, o simplemente matriz de covarianzas, es simetrica y semidefinida positiva. Se define como

$$\Sigma_X = \begin{pmatrix} Var[X_1] & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var[X_2] & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & Var[X_n] \end{pmatrix}$$

6.5 Matriz de correlaciones

La matriz de correlaciones es simetrica y semidefinida positiva. Se define como

$$R_X = \begin{pmatrix} 1 & \rho(X_1, X_2) & \cdots & \rho(X_1, X_n) \\ \rho(X_2, X_1) & 1 & \cdots & \rho(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(X_n, X_1) & \rho(X_n, X_2) & \cdots & 1 \end{pmatrix}$$

6.6 Características numéricas de la distribución multinomial

- Medias: $\mu_{X_i} = E[X_i] = np_i$
- Varianzas: $\sigma^2_{X_i} = Var[X_i] = np_i(1 - p_i)$
- Covarianzas: $\sigma_{X_i, X_j} = Cov(X_i, X_j) = -np_ip_j$

7 Transformaciones de vectores aleatorios

Dado el vector aleatorio $X = (X_1, X_2, \ldots, X_n)$ con función de densidad conjunta $f_X(x_1, x_2, \ldots, x_n)$, lo tranformamos en otro vector aleatorio $Y = (Y_1, Y_2, \ldots, Y_n)$ con la misma dimensión

$$\begin{pmatrix} Y_1 = g_1(X_1, X_2, \ldots, X_n) \\ Y_2 = g_2(X_1, X_2, \ldots, X_n) \\ \vdots \\ Y_n = g_n(X_1, X_2, \ldots, X_n) \end{pmatrix}$$

de tal modo que existan transformadas inversas. La función de densidad continua será

$$f_Y(y_1, y_2, \ldots, y_n) = f_X(g^{-1}(y_1, y_2, \ldots, y_n))|J| \quad \text{donde} \quad J = \begin{vmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} & \cdots & \frac{dx_1}{dy_n} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} & \cdots & \frac{dx_2}{dy_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{dx_n}{dy_1} & \frac{dx_n}{dy_2} & \cdots & \frac{dx_n}{dy_n} \end{vmatrix}$$

7.1 Convolución

Si X e Y son variables aleatorias idependientes con funciones de densidad $f_X(x)$ y $f_Y(y)$, respectivamente, entonces la función de densidad de $Z = X + Y$ es

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - x) f_Y(x) dx$$

7.2 Transformaciones lineales

Dado X un vector aleatorio n-dimensional y u un vector en $\mathbb{R}^n$, la variable aleatoria $Y = u^t X$ cumple

$$E[Y] = E[u^t X] = u^t E[X]$$

$$Var[Y] = Var[u^t X] = u^t \Sigma_X u$$

Dado X un vector aleatorio n-dimensional y A una matriz $m \times n$-dimensional, el vector aleatorio $Y = AX$(m-dimensional) cumple

$$E[Y] = E[AX] = AE[X]$$

$$\Sigma_Y = E[(AX - AE[X])(AX - AE[X])^t] = A \Sigma_X A^t$$

8 Distribución normal bivariante

Decimos que un vector aleatorio X sigue distrubución normal bivariante con vector de medidas (μ_X, μ_Y), matriz de varianzas y covarianzas Σ si tiene función de densidad conjunta

$$f(x_1, x_2) = \frac{1}{2\pi\,|\Sigma|^{\frac{1}{2}}} exp\left\{\frac{-1}{2}(x_1-\mu_1, x_2-\mu_2)\Sigma^{-1}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}\right\};$$

$$\text{si} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad \text{entonces} \qquad f(x_1, x_2) =$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left\{\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}$$

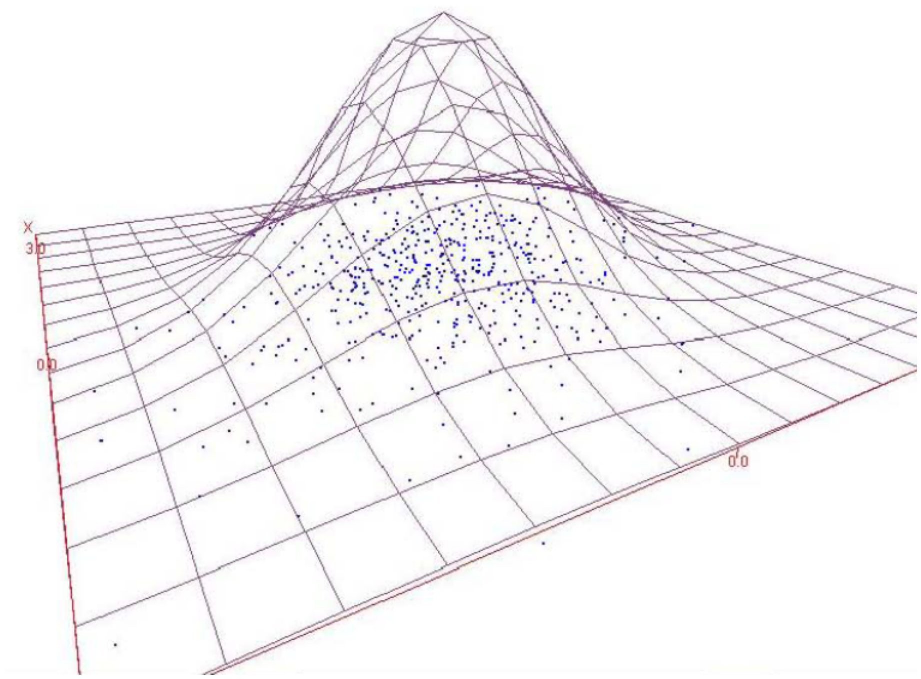

Figure 4: Ejemplo de distribución normal bivariante

Las distribuciones marginales de una distribución Normal Bivariante son normales unidimensionales,

$$X_1 \sim N(\mu_1, \sigma_1^2) \quad \text{y} \quad X_2 \sim N(\mu_2, \sigma_2^2)$$

donde la correlación ρ controla el grado de dependencia lineal entre ellas.

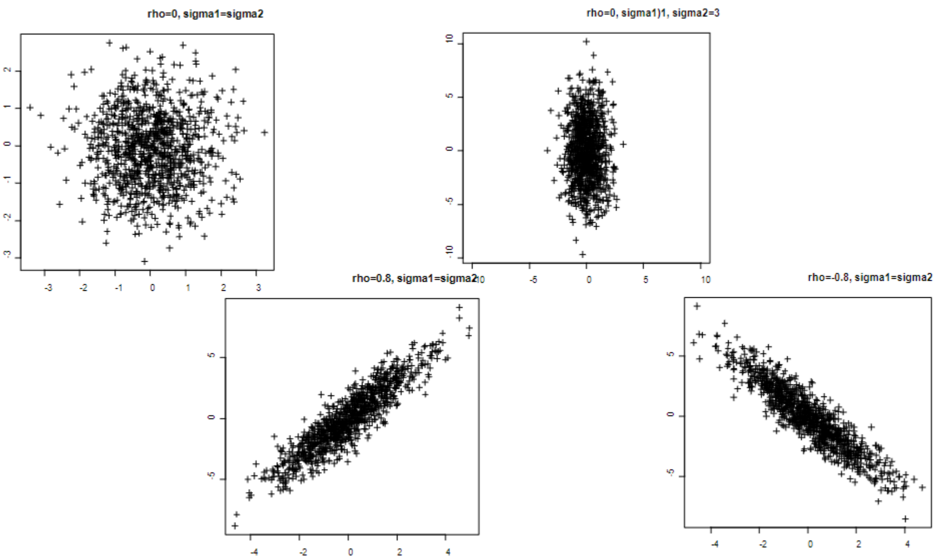

Figure 5: Ejemplo de distribución normal bivariante.

Las distribuciones normales bivariantes tienen las siguientes **propiedades**: Dado (X_1, X_2) un vector aleatorio normal con vector de medidas (μ_1, μ_2) y matriz de varianzas y covarianzas Σ,

- Si $\rho = 0$ entonces X_1 e X_2 son independientes
- Dados $a_1, a_2, a_3 \in \mathbb{R}$, $a_1X_1 + a_2X_2 + a_3$ es normal
- $X_1 \Big/ X_2 = x_2$ es normal y $X_2 \Big/ X_1 = x_1$ es normal
- Sea $A \in \mathcal{M}_{2\times 2}$, X vector aleatorio de dimensión 2, $X \sim N(\mu_X, \Sigma_X)$ entonces $Y = AX \sim N(\mu_Y, \Sigma_Y)$ donde $\mu_Y = A\mu_X$ y $\Sigma_Y = A\Sigma_X A^t$

Dado (X, Y) un vector aleatorio normal con medidas (μ_X, μ_Y) y matriz de varianzas y covarianzas Σ

$$Y \Big/ X = x \sim N(\mu_2 + \rho(\frac{\sigma_2}{\sigma_1})(x - \mu_1), \sigma_2^2(1 - \rho^2))$$

9 La Esperanza Condicionada como v.a.

Al hablar de distribuciones condicionadas, tiene sentido plantearse cómo se incorpora a la esperanza de una variable información sobre el comportamiento de otra o vector aleatorio asociado a ella. Por simplificar trataremos el caso bivariante.
Supongamos que (X, Y) es un vector aleatorio bidimensional. La esperanza de la variable Y es una caracteristica propia de la distribución de Y, en principio, P_Y. Si conocemos que $X = x$, el modelo asociado a Y pasará a estar dado por la distribución condicionada $P_{Y/X=x}$.

Al valor esperado de Y en este modelo se le llama esperanza de Y condicionada por $X = x$ y se denota

$$E\left(Y \middle/ X = x\right)$$

Recordando lo estudiado sobre las distribuciones condicionadas, es fácil darse cuenta que, en el caso discreto, la forma de calcular la esperanza condicionada está dada por la siguiente expresión:

$$E\left(Y \middle/ X = x\right) = \sum_y yP\left(Y = y \middle/ X = x\right) = \sum_y y\frac{p(x,y)}{p_X(x)}$$

mientras que en el caso continuo la regla de cálculo para la esperanza condicionada está dada por la siguiente expresión

$$E\left(Y \middle/ X = x\right) = \int yf_{Y/X=x}(y)dy = \int y\frac{f(x,y)}{f_X(x)}dy$$

Sea (X,Y) vector aleatorio. La función $\phi(x) \equiv E\left(Y \middle/ X = x\right)$ está bien definida en el conjunto de posibles valores de X y, por tanto, tiene sentido calcular $\phi(X)$. A esta variable aleatoria se le suele denotar como

$$E\left(Y \middle/ X\right)$$

y habitualmente nos referimos a ella como "esperanza de Y ". // $\phi(X)$ es una nueva variable aleatoria de la que se puede calcular su distribución, esperanza, varianza, etc. Por ejemplo, en el caso discreto, teniendo en cuenta la expresión correspondiente para $E\left(Y \middle/ X = x\right)$,

$$E(\phi(X)) = \sum_x \left(\sum_y y\frac{p(x,y)}{p_X(x)}\right)p_X(x) = \sum_{x,y} yp(x,y) = E(Y)$$

La esperanza condicionada tiene las siguientes **propiedades**:

- Si $a, b \in \mathbf{R}$ entonces

$$E\left(aY_1 \middle/ X\right) = aE\left(Y_1 \middle/ X\right) + bE\left(Y_2 \middle/ X\right)$$

- $E\left(E\left(Y \middle/ \right)\right) = E(Y)$

- Para cualquier función g

$$E\left(g(X)Y \middle/ X\right) = g(X)E\left(Y \middle/ X\right)$$

9.1 Curva de regresión de Y sobre X

Sea (X, Y) un vector aleatorio nos planteamos el problema de encontrar una función f que minimice el valor

$$E(Y - f(X))^2$$

Si tomamos la notación anterior para $\phi(x)$, entonces, de la tercera propiedad de la esperanza condicionada se deduce que

$$E(Y - f(X))^2 = E(Y - \phi(X))^2 + E(\phi(X) - f(X))^2$$

lo que significa que el valor mínimo de $E(Y - f(X))^2$ se alcanza cuando $f = \phi$. Claro que $\phi(X) = E\left(Y \middle/ X\right)$. Esto significa que $\phi(X)$ es la variable aleatoria función de X que mejor aproxima a Y en el sentido de los mínimos cuadrados. Por esta razón, nos referimos en ocasiones a $\phi(X)$ como la *curva de regresión de Y sobre X*.

9.2 Recta de regresión de Y sobre X

Sea (X, Y) un vector aleatorio nos planteamos el problema de encontrar $a, b \in \mathbb{R}$ que minimicen el valor

$$E(Y - aX - b)^2$$

Se puede ver que este error, como función de b se minimiza cuando $b = E(Y) - aE(X)$. Entonces la función a minimizar es

$$Var(Y - aX) = Var(Y) + a^2 Var(X) - 2aCov(X, Y)$$

y el mínimo se alcanza en $a = \frac{Cov(X,Y)}{Var(X)} = \rho_{X,Y} \frac{\sigma_Y}{\sigma_X}$. Por lo tanto la función a minimizar será

$$E(Y) + \rho_{X,Y} \frac{\sigma_Y}{\sigma_X} (X - E(X))$$

que es la variable aletoria función lineal de X que mejor representa a Y en el sentido del error cuadrático. A esta variable se la llama *rectra de regresión de Y sobre X*.

9.3 Ejemplos

Ejemplos **discretos**:

- Vector multinomial. Las condicionadas son binomiales y la curva de regresión es una recta
- Cálculos a partir de una tabla cruzada

Ejemplos **continuos**:

- Vector Normal. Las condicionadas son normales y la curva de regresión es una recta.
- Cálculos a partir de $f(x, y)$

9.4 La varianza condicionada como v.a.

Sea (X,Y) un vector aleatorio, definimos la v.a. en función de X:

$$Var\left(Y \middle/ X\right) = E\left(\left(Y - E\left(Y \middle/ X\right)\right)^2 \middle/ X\right)$$

tambien podemos calcularla como:

$$Var\left(Y \middle/ X\right) = E\left(Y^2 \middle/ X\right) - \left(E\left(Y \middle/ X\right)\right)^2$$

ademas se verifica la siguiente igualdad:

$$Var(Y) = E\left(Var\left(Y \middle/ X\right)\right) + Var\left(E\left(Y \middle/ X\right)\right)$$

10 Distribuciones asociadas a un vector ordenado

Sean $X_1, X_2, \ldots, X_n$ v.a. cuando las ordenamos las denotamos por

$$X_{(1)}, X_{(2)}, \ldots, X_{(n)}$$

El vector $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ es un estadístico n-dimensional puesto que es función de la muestra, aunque en este caso el estadístico no reduce la dimensión de ella. Este vector recibe el nombre de **estadístico ordenado**.
A su coordenada $X_{(i)}$ la llamaremos **estadístico ordenado i-esimo**.

La **distribución conjunta** del estadístico ordenado en el caso continuo:
Sean $X_1, X_2, \ldots, X_n$ v.a.i.i.d. con distribución F continua y sea f su función de densidad, la función de densidad que define la distribución n-dimensional del vector $T = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ es el siguiente producto

$$f_T = (x_1, x_2, \ldots, x_n) = n! f(x_1) f(x_2) \cdots f(x_n) \qquad \text{si} \quad x_1 < x_2 < \cdots < x_n$$

10.1 Distribución del máximo

Sean $X_1, X_2, \ldots, X_n$ v.a.i.i.d. con distribución F, la función de distribución de $X_{(n)}$ es

$$F_{X_{(n)}}(x) = F(x)^n$$

y su función de densidad de $X_{(n)}$ es

$$f_{X_{(n)}}(x) = n f(x) F(x)^{n-1}$$

10.2 Distribución del mínimo

Sean $X_1, X_2, \dots, X_n$ v.a.i.i.d. con distribución F, la función de distribución de $X_{(1)}$ es

$$F_{X_{(1)}}(x) = 1 - (1 - F(x))^n$$

y su función de densidad de $X_{(1)}$ es

$$f_{X_{(1)}}(x) = nf(x)(1 - F(x))^{n-1}$$

11 Familias de distribuciones asociadas al muestreo de la normal

Sean $X_1, X_2, \dots, X_n$ variables aleatorias i.i.d. $N(\mu, \sigma^2)$ se verifica que:

1. Las distribuciones exactas de $\overline{X}$ y S^2 son Normal y Gamma respectivamente:

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$n\frac{S^2}{\sigma^2} \sim \gamma\left(\frac{n-1}{2}, \frac{1}{2}\right)$$

2. $\overline{X}$ y S^2 son independientes

11.1 Chi-cuadrado

Sea $n = 1, 2, 3 \dots$ diremos que una v.a X tiene una distribución Chi-cuadrado con n grados de libertad cuando su distribución es $\gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ y lo denotaremos

$$X \sim \chi_n^2$$

Por tanto es inmediato comprobar que la variabilidad del muestreo sigue una distribución Chi-cuadrado de $n-1$ grados de libertad

$$n\frac{S^2}{\sigma^2} \sim \gamma\left(\frac{n-1}{2}, \frac{1}{2}\right) = \chi_{n-1}^2$$

Si $X_1, X_2, \dots, X_n$ son n variables aleatorias i.i.d. con distribución $N(0, 1)$, entonces $Y = X_1^2 + X_2^2 + \cdots + X_n^2$ es una variable aleatoria con distribución Chi-cuadrado con n grados de libertad

$$Y \sim \chi_n^2$$

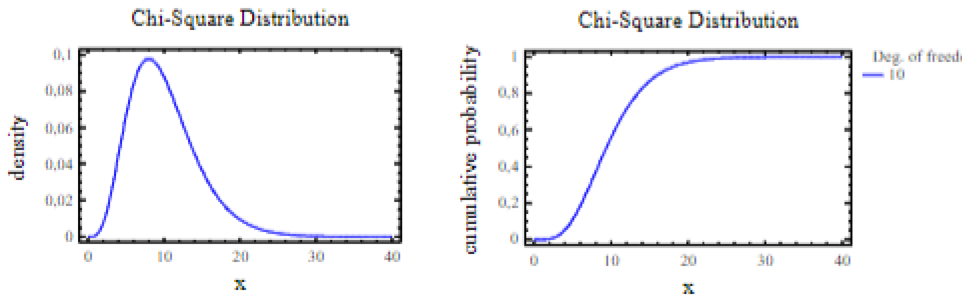

Figure 6: Chi-cuadrado con 10 grados de libertad

Si $X \sim \chi^2_n$ entonces

$$EX = n \qquad Var(X) = 2n$$

La densidad es asimétrica a la derecha y positiva en $(0, \infty)$
Sea $0 < \alpha < 1$ denotamos por $\chi^2_{n,\alpha}$ al $(1 - alpha) - cuantil$ de la distribución χ^2_n. Este valor siempre es positivo y está tabulado.

11.2 F de Fisher-Snedecor

Sean $n = 1, 2, 3, \ldots$ y $m = 1, 2, 3, \ldots$, y sean X e Y variables aleatorias independientes con distribuciones χ^2_n y χ^2_m respectivamente.
La distrubición del cociente de las variables se denomina F de Fisher-Snedecor con n y m grados de libertad.

$$\frac{\frac{X}{n}}{\frac{Y}{m}} \sim F_{n,m}$$

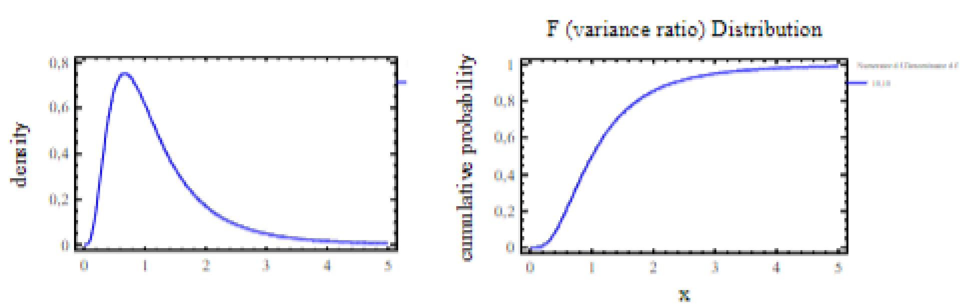

Figure 7: F de Fisher-Snedecor con n = 10, m = 10 grados de libertad

Si $X \sim F_{n,m}$ entonces

$$EX = \frac{m}{m-2} \quad m > 2 \qquad Var(X) = \frac{2m^2(m+n-2)}{n(m-2)^2(m-4)}$$

Sea $0 < \alpha < 1$ denotamos por $F_{n,m,\alpha}$ al $(1-alpha)-cuantil$ de la distribución $F_{n,m}$. Con esta notación, para usar las tablas de la distribución debemos tener en cuenta la siguiente propiedad

$$F_{n,m,\alpha} = \frac{1}{F_{m,n,1-\alpha}}$$

11.3 t de Student

Sea $n = 1, 2, 3 \ldots$ y sean Z y X v.a. independientes con distribuciones $N(0,1)$ y χ^2_n respectivamente. La distribución de $\frac{Z}{\sqrt{\frac{X}{n}}}$ se denomina t de Student con n grados de libertad y lo denotaremos

$$\frac{Z}{\sqrt{\frac{X}{n}}} \sim t_n$$

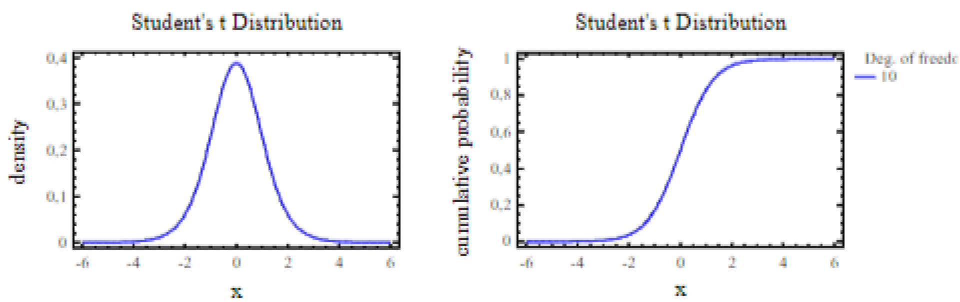

Figure 8: t de Student con 10 grados de libertad

Si $X \sim t_n$ entonces

$$EX = 0 \quad n > 1 \qquad Var(X) = \frac{n}{n-2}$$

Ademas, la distribución t_1 es una $Cauchy(0,1)$

Su densidad es acampanada y simétrica con respecto al origen. Sea $0 < \alpha < 1$ denotamos por $t_{n,\alpha}$ al $(1 - alpha) - cuantil$ de la distribución t_n. Con esta notación, para usar las tablas de la distribución debemos tener en cuenta la siguiente propiedad

$$t_{n,1-\alpha} = -t_{n,\alpha}$$

References

[1] Pilar Rodríguez de Tío,
Vectores aleatorios y distribuciones multivariantes.
Universidad de Valladolid, 2023.

[2] Pilar Rodríguez de Tío,
Distribuciones asociadas a variables independientes igualmente distribuidas.
Universidad de Valladolid, 2023.