Capítulo 3

Principales distribuciones. Vectores aleatorios.

El objetivo es familiarizarse con las distribuciones de probabilidad más usuales, tanto con las de tipo discreto como con las continuas. En la segunda parte del capítulo nos porponemos introducir la teoría básica relativa a variables aleatorias bidimensionales.

3.1. Principales distribuciones discretas

3.1.1. Degenerada

Sea la v.a.

$$\begin{array}{rccc} \xi : & (\Omega, \mathcal{A}, P) & \rightarrow & (\mathbf{R}, \mathcal{B}) \\ & \omega & \rightarrow & \xi(\omega) = a. \end{array}$$

Se trata de un fenómeno determinista pues identificamos a todas las posibles concreciones que componen Ω con un mismo número real a, no hay aleatoriedad. Se comprueba fácilmente que

$$F_{\xi}(x) = \begin{cases} 0 & \text{si } x < a \\ 1 & \text{si } x \geq a, \end{cases}$$

y también que

$$\begin{aligned} \varphi_{\xi}(t) &= e^{ita}, \\ E[\xi] &= a, \text{ y} \\ V[\xi] &= 0. \end{aligned}$$

3.1.2. Distribución binomial

Sea un fenómeno con sólo dos posibles concreciones, éxito (E) o fracaso (F). Suponemos que

$$P(E) = p \in (0,1) \text{ y } P(F) = 1 - p.$$

Realizamos n veces - de manera independiente - el fenómeno y definimos

$$\begin{array}{llll} \xi : & (\Omega, \mathcal{A}, P) & \rightarrow & (\mathbb{R}, \mathcal{B}) \\ & \omega & \rightarrow & \xi(\omega) = \text{n}^0 \text{ de éxitos al cabo de n realizaciones.} \end{array}$$

Notamos que Ω estará compuesto por $\omega' s$ de la forma

$$E, E, F, \overset{n)}{F,} E, F, ..., E,$$

así que si ω consta de x éxitos y $n - x$ fracasos entonces $\xi(\omega) = x$ y $P(\omega) = p^x q^{n-x}$ (siendo $q = 1 - p$).
Vemos como ξ es de tipo discreto con $\operatorname{Im} \xi = \{0, 1, 2, ..., n\} \subset \mathbb{R}$, y la probabilidad P^* inducida por la v.a. ξ, P_ξ^*, en estos puntos de $\operatorname{Im} \xi$ es

$$P_\xi^*(x) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0, 1, 2, ..., n.$$

Cálculo de $\varphi_\xi(t)$:

$$\begin{aligned} \varphi_\xi(t) &= \sum_{x=0}^{n} e^{itx} P_\xi^*(x) = \\ &= \sum_{x=0}^{n} e^{itx} \binom{n}{x} p^x q^{n-x} = \\ &= \sum_{x=0}^{n} \binom{n}{x} \left(pe^{it}\right)^x q^{n-x} \\ &= \left(q + pe^{it}\right)^n. \end{aligned}$$

Se comprueba que

$$E[\xi] = np \text{ y } V[\xi] = npq.$$

Notación: $\xi \sim B(n, p)$.

3.1.3. Distribución geométrica

Seguimos refiriéndonos a un fenómeno con dos posibles concreciones, (E) o (F) con $P(E) = p \in (0,1)$. Realizamos el fenómeno varias veces -de manera independiente- hasta que aparece el primer éxito. Consideramos Ω como el conjunto cuyos elementos son de la forma

$$\{E\}, \{F, E\}, \{F, F, E\}, ..., \{F, F, F, ..., E\}, ...;$$

observamos que si el número de realizaciones de un ω hasta obtener el primer éxito es x entonces

$$P(\omega) = pq^{x-1}.$$

Definimos

$$\xi = \text{ número de realizaciones hasta obtener } (E);$$

para esta v.a. se tiene que, $\operatorname{Im} \xi = \{1, 2, 3, ...\}$ y

$$P_\xi^*(x) = pq^{x-1}, \qquad x = 1, 2, 3...$$

Los cálculos en este caso se desarrollan con facilidad, llegándose a que

$$\begin{aligned}
\varphi_\xi(t) &= \frac{p}{q}\left(\frac{qe^{it}}{1-qe^{it}}\right), \\
E[\xi] &= \frac{1}{p}, \text{ y} \\
V[\xi] &= \frac{q}{p^2}.
\end{aligned}$$

Notación: $\xi \sim G(p)$.

3.1.4. Binomial negativa

Sea como antes el esquema fenómeno:(E) o (F) donde $P(E) = p \in (0,1)$. Se llevan a cabo realizaciones del experimento hasta obtener n éxitos. Conformamos Ω como el conjunto de los ω de la forma

$$\omega = \underset{n-1 \text{ éxitos}}{E, F, E}, F, F, F, ..., \underset{n-\text{ésimo éxito}}{E}.$$

Así, si ω está compuesto por n éxitos y x fracasos

$$P(\omega) = q^x p^n.$$

Definimos

$$\xi = \text{n}^0 \text{ de fracasos hasta obtener } n \text{ éxitos;}$$

para esta v.a. tenemos que su imagen es $\{0, 1, 2, ...\}$ y

$$P_\xi^* (x) = \left(\begin{array}{c} n - 1 + x \\ x \end{array} \right) q^x p^n.$$

Se comprueba por inducción que

$$\varphi_\xi (t) = \left(\frac{p}{1 - qe^{it}} \right)^n ,$$

y a partir de esta función deducimos

$$E [\xi] = \frac{np}{q} \text{ y } V [\xi] = \frac{nq}{p^2}.$$

Notación: $\xi \sim BN (n, p)$

3.1.5. Distribución de Poisson

Este tipo de distribución puede ser obtenida a partir de una binomial $B (n, p)$ cuando n es suficientemente grande y p es pequeño. Siendo esto así se consigue una distribución de Poisson de parámetro $\lambda = np$ (en el límite $\left(\begin{array}{c} n \\ x \end{array} \right) p^x q^{n-x} \approx e^{-\lambda} \frac{\lambda^x}{x!}$). la aproximación es aconsejable si $n \geq 30$ y $p \leq 0{,}1$.
Una v.a. discreta ξ cuya función de densidad de masa asociada P_ξ^* sea de la forma

$$P_\xi^* (x) = e^{-\lambda} \frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, ...$$

se dirá que tiene asociada una distribución de Poisson de parámetro λ; se denotará $\xi \sim P (\lambda)$.
Para esta situación tenemos

$$\begin{array}{rcl} \varphi_\xi (t) & = & \exp \left(\lambda \left(e^{it} - 1 \right) \right) , \\ E [\xi] & = & V [\xi] = \lambda. \end{array}$$

3.1.6. Distribución hipergeométrica

La idea a generalizar es la del siguiente modelo: sea una población de N elementos de los cuales D son (E) y $N - D$ son (F) . Se toman muestras de tamaño n (sin reemplazamiento) y contamos el número de éxitos; construimos así el espacio de probabilidad $(\Omega, \mathcal{A}, P)$ donde

$$\Omega = \{\omega : \omega \text{ muestra de tamaño } n\}, \qquad \mathcal{A} = \mathcal{P}(\Omega)$$

y

$$P(\omega) = \frac{1}{\begin{pmatrix} N \\ n \end{pmatrix}}.$$

Definimos

$$\xi = \text{n}^0 \text{ de éxitos };$$

entonces

$$P_\xi^*(x) = \begin{pmatrix} D \\ x \end{pmatrix} \begin{pmatrix} N-D \\ n-x \end{pmatrix} \frac{1}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

[1] y

$$E[\xi] = \frac{Dn}{N}.$$

Notación: $\xi \sim HG(N, n, p)$.

3.1.7. Ejemplos

EJEMPLO 3.1 *Se sabe que la probabilidad de que un individuo reaccione desfavorablemente tras la inyección de una vacuna es de 0.002. Determinar la probabilidad de que en un grupo de 2000 personas vacunadas haya como mucho tres que reaccionen desfavorablemente.*

$\lambda = np = 0{,}002 \cdot 2000 = 4$; *así la v.a.* $\xi =$ *número de individuos que reaccionan desfavorablemente tiene asociada una distribución de Poisson de parámetro* $\lambda = 4$, *i.e.* $\xi \sim P(4)$. *Por tanto*

$$P_\xi^*(x) = P(\xi = x) = e^{-4}\frac{4^x}{x!}, \qquad x = 0, 1, 2, ...$$

[1]Para que tenga sentido se ha de cumplir

$$\text{máx}\{0, n-(N-D)\} \leq x \leq \text{mín}\{n, D\}$$

y la probabilidad pedida es

$$\begin{aligned} P\left(\xi \leq 3\right) &= P(\xi = 0) + P(\xi = 1) + \\ +P(\xi &= 2) + P(\xi = 3) \approx 0{,}433. \end{aligned}$$

Ejemplo 3.2 *Supongamos que la probabilidad de que en la elaboración de un producto alimenticio - perteneciente a cierto supermercado - éste contenga sal es de 0.03. Al coger el carro en el supermercado nos proponemos ir eligiendo de forma aleatoria alimentos hasta encontrar uno con sal; supondremos que depositando este último en el carro concluiremos la compra. Si cada uno de los productos cuesta 1000 pesetas, ¿ cuál es la probabilidad de que nos gastemos 100000 pesetas ?. Calcular el precio esperado de la compra.*

Identificamos éxito con la elección de un producto con sal, así $P(E) = 0{,}03$. *Definimos la v.a.*

$$\xi = \; n^0 \textit{ de artículos hasta encontrar uno con sal;}$$

vemos que $\xi \sim G(0{,}03)$ *y que por tanto la probabilidad pedida es* $P(\xi = 100)$, *i.e.*

$$P(\xi = 100) = (0{,}03)\,(1 - 0{,}03)^{99}\,.$$

El precio esperado de la compra coincide con la esperanza multiplicada por 1000 (que es lo que vale cada artículo), es

$$10^3 E\left[\xi\right] = 10^3 \frac{1}{0{,}03} = 33{,}300 \textit{ pesetas.}$$

Ejemplo 3.3 *Una gama de 10 productos usa el conservante EZ en 3 de éstos. Si compramos en el mercado 5 productos de dicha gama, ¿ cuál será la probabilidad de encontrar EZ en 2 de los artículos ?.*

Se trata de una población con $N = 10$, $D = 3$ *de éstos contienen EZ y tomamos una muestra de tamaño* $n = 5$. *La ley de probabilidad de la v.a.*

$$\xi \left(\textit{muestra}\right) = \; n^0 \textit{ de productos con EZ en la muestra}$$

es una hipergeométrica $HG(N = 10, n = 5, p = 0{,}3)$. *La probabilidad pedida es*

$$P(\xi = 2) = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \begin{pmatrix} 10 - 3 \\ 5 - 2 \end{pmatrix} \frac{1}{\begin{pmatrix} 10 \\ 5 \end{pmatrix}}.$$

3.2. Principales distribuciones continuas

3.2.1. Uniforme

Se dice que la v.a. ξ está distribuida según una distribución uniforme en el intervalo (a,b) si

$$f_\xi(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in (a,b) \\ 0 & \text{en el resto.} \end{cases}$$

Se comprueba que

$$E[\xi] = \frac{a+b}{2}$$

y

$$V[\xi] = \frac{(b-a)^2}{12}.$$

Por otro lado el cálculo de su función característica es

$$\begin{aligned} \varphi_\xi(t) &= \int_{\mathbb{R}} e^{itx} dF_\xi(x) = \\ &= \int_{\mathbb{R}} e^{itx} f_\xi(x)\, dx = \\ &= \int_a^b \frac{e^{itx}dx}{b-a} = \left(\frac{1}{it}\right) \frac{e^{itb}-e^{ita}}{b-a}. \end{aligned}$$

Queda como ejercicio comprobar que

$$F_\xi(x) = \begin{cases} 0 & \text{si } x \le a \\ \dfrac{x-a}{b-a} & \text{si } x \in (a,b) \\ 1 & \text{si } x \ge b. \end{cases}$$

3.2.2. Distribución de Cauchy.

Cuando la función de densidad de la v.a. ξ sea

$$f_\xi(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad x \in \mathbb{R}$$

diremos que ξ se distribuye según una Cauchy. Queda como ejercicio comprobar que f_ξ es una función de densidad de probabilidad.

Observación: ξ no tiene momentos y $\varphi_\xi(t) = e^{-|t|}$.

3.2.3. Distribución de Pareto.

Una v.a. ξ está distribuida según una Pareto si su densidad de probabilidad es tal que

$$\begin{aligned} P(\xi > x) &= 1 - F_{\xi}(x) = \\ &= \begin{cases} \left(\frac{a}{x}\right)^{\alpha} & \text{si } x > a \\ 1 & \text{si } x \leq a, \end{cases} \end{aligned}$$

con α y a constantes positivas. Para este tipo de distribuciones se tiene

$$E[\xi] = \frac{\alpha a}{\alpha - 1} \text{ y } E\left[\xi^2\right] = \frac{\alpha a^2}{\alpha - 2}.$$

3.2.4. Distribución normal

Las variables aleatorias normales son las más frecuentes en el estudio de fenómenos, bien porque en la práctica ocurren o bien porque la ley de probabilidad que estudiamos se puede aproximar por una normal, esto último sucede con mucha frecuencia.

La densidad asociada a una variable aleatoria normal (o a una distribución de probabilidad normal) tiene forma de campana, lo cual dice que la mayoría de los valores que alcanza la v.a. están cercanos al centro de dicha campana (el centro será μ, la esperanza de la v.a.), y según nos alejamos de este centro la probabilidad va disminuyendo. Existirá un coeficiente - σ - en función del cual veremos cuánto disminuirá (o aumentará) la probabilidad a medida que nos alejemos (o nos acerquemos) del centro μ.

DEFINICIÓN 3.1 *Una v.a. ξ sigue una distribución normal de parámetros (μ, σ) si su función de densidad de probabilidad es*

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$

donde $\mu \in \mathbf{R}$ y $\sigma > 0$. La notación empleada es $\xi \sim N(\mu, \sigma)$.
El gráfico de la función de densidad está representado en la Figura 1.

El aspecto es el de una campana que será más picuda cuanto más pequeño sea el valor de σ - ver Figura adjunta -.

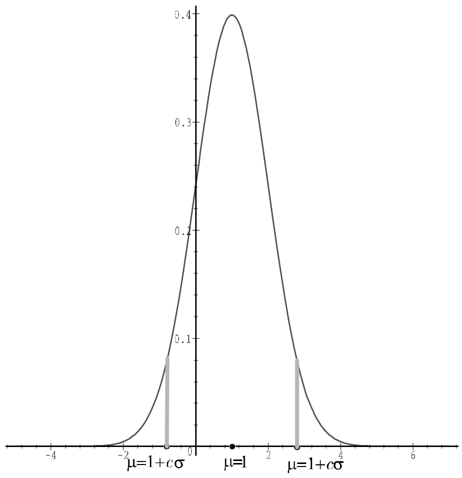

Figura 3.1: Densidad de una distribución $N(1, 1)$.

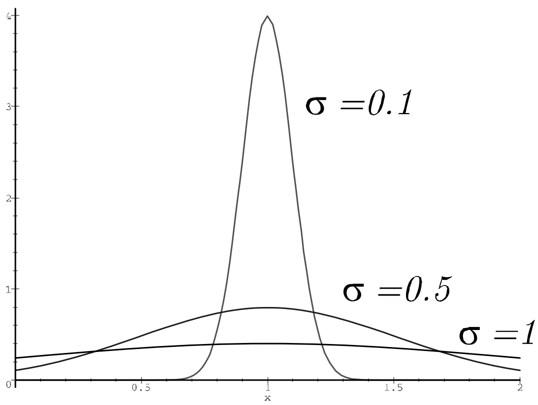

Figura 3.2:

Recordamos que con σ pequeño logramos que la probabilidad se concentre en torno al valor de μ (la probabilidad que en esta situación queda en las colas es menor). Observamos que se trata de una función simétrica con puntos de inflexión en $\mu-\sigma$ y $\mu+\sigma$.

Por otro lado sabemos que según la desigualdad de Tchevichev $P\left(|\xi-\mu|\geq c\sigma\right)\leq 1/c^2$, lo que se interpreta en el gráfica de la manera siguiente: el área (probabilidad) que queda en las dos 'colas' a partir de los segmentos de color gris - ver Figura 1- se hace pequeña a medida que c se toma grande.

Comprobar que $f_\xi\left(x\right)\geq 0$ es evidente pues se trata de una función exponencial, pero comprobar

$$\int_{\mathbb{R}} f_\xi\left(x\right)dx=1$$

es algo complicado. Por ello tendremos que recurrir a la siguiente fórmula:

$$\int_{-\infty}^{+\infty} e^{-x^2/2}dx=\sqrt{2\pi}. \tag{3.1}$$

Con esta fórmula y un cambio de variables comprobamos que la densidad $f_\xi\left(x\right)$ tiene integral uno sobre $\mathbb{R}$.

Definición 3.2 *Diremos que la v.a.* ξ *se distribuye como una normal* tipificada *si* $\xi\sim N(0,1)$.

Normal tipificada

Si $\xi\sim N\left(\mu,\sigma\right)$ entonces

$$F_\xi\left(x\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(t-\mu)^2}{2\sigma^2}}dt,$$

pero la dificultad reside en que evaluar esta integral para todos los posibles valores de μ y σ es imposible en la práctica. Para resolver esto nos ocuparemos de evaluar dicha integral cuando $\mu=0$ y $\sigma=1$ - esto es, en el caso tipificado -. Disponemos $F_\eta\left(x\right)$ para $\eta\sim N(0,1)$ - ver tablas de la normal $N(0,1)$ - y con esto podremos evaluar el resto de los casos. En efecto, sabiendo el valor de $F_\eta(x)$ podemos saber cuanto es $F_\xi(x)$ siendo $\xi \sim N(\mu, \sigma)$:

$$\begin{aligned} F_\xi(x) &= P(\xi \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt \\ &\overset{z=\frac{t-\mu}{\sigma}}{=} \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz \\ &= F_\eta\left(\frac{x-\mu}{\sigma}\right) \end{aligned}$$

i.e.

$$F_\xi(x) = F_\eta\left(\frac{x-\mu}{\sigma}\right).$$

En la práctica cuando $\xi \sim N(\mu, \sigma)$

$$\begin{aligned} F_\xi(x) &= P(\xi \leq x) = P\left(\frac{\xi-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\ &= F_\eta\left(\frac{x-\mu}{\sigma}\right) = P\left(\eta \leq \frac{x-\mu}{\sigma}\right) \end{aligned}$$

por lo tanto la v.a.

$$\phi = \frac{\xi-\mu}{\sigma}$$

está distribuida según una $N(0,1)$, se denomina *la tipificada de* ξ.[2] Resumiendo

TEOREMA 3.1 *Si* $\xi \sim N(\mu, \sigma)$ *entonces*

$$F_\xi(x) = F_\eta\left(\frac{x-\mu}{\sigma}\right)$$

con $\eta \sim N(0,1)$ *y*

$$\frac{\xi-\mu}{\sigma} \sim N(0,1).$$

---

[2]Una manera de probar que $\phi \sim N(0,1)$:

$$\begin{aligned} F_\phi(y) &= P(\phi \leq y) \\ &= P(\xi \leq \sigma y + \mu) \\ &= F_\eta\left(\frac{(\sigma y+\mu)-\mu}{\sigma}\right) \\ &= F_\eta(y). \end{aligned}$$

Ejemplo 3.4 *Evaluar* $P(\xi > 55)$ *si* $\xi \sim N(50,5)$.

$$\begin{aligned} P(\xi > 55) &= 1 - P(\xi \leq 55) \\ &= 1 - F_\xi(55) \\ &= 1 - F_\eta\left(\frac{55-50}{5}\right) \\ &= 1 - F_\eta(1) \end{aligned}$$

donde según tablas $F_\eta(1) \approx 0{,}8143$.

Ejemplo 3.5 *Con* $\xi \sim N(50,5)$ *determinar un número* k *tal que* $P(\xi > k) = 0{,}25$.
Para ello utilizamos las tablas y tenemos en cuenta que

$$P(\xi > k) = 1 - F_\eta\left(\frac{55-k}{5}\right),$$

de donde se sigue que $F_\eta\left(\frac{55-k}{5}\right) = 0{,}75$. *Aquel valor de* x *tal que* $F_\eta(x) = 0{,}75$ *es - aproximadamente - según tablas* $x = 0{,}675$*, luego* $x = \frac{55-k}{5}$ *y por ende* $k = 55 - 5\,(0{,}675) = 51{,}625$.

Función característica.

Nos ocupamos en deducir cuál es la función característica de una v.a. normal: si $\xi \sim N(\mu,\sigma)$ entonces

$$\begin{aligned} \varphi_\xi(t) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{\mathbb{R}} e^{itx} e^{\frac{-(x-\mu)^2}{2\sigma^2}}dx = \\ &\overset{z=\frac{x-\mu}{\sigma}}{=} \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{1}{2}(z-it\sigma)^2}\, e^{-\frac{\sigma^2}{2}t^2+it\mu}dz \end{aligned}$$

ya que

$$\begin{aligned} itx - \frac{(x-\mu)^2}{2\sigma^2} &= -\frac{z^2}{2} + it(\sigma z + \mu) \\ &= -\frac{1}{2}(z-it\sigma)^2 + \frac{i^2\sigma^2 t^2}{2} + it\mu, \end{aligned}$$

haciendo ahora $v = z - it\sigma$, $dv = dz$, y teniendo en cuenta la fórmula (3.1) tenemos que

$$\begin{aligned} \varphi_\xi(t) &= \left(\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{v^2}{2}}\,dz\right) e^{-\frac{\sigma^2}{2}t^2+it\mu} \\ &= e^{-\frac{\sigma^2}{2}t^2+it\mu}. \end{aligned}$$

Y con este cálculo estamos en condiciones de evaluar los momentos con respecto al origen de ξ : el primer momento es

$$E[\xi] = \left. \frac{\left(e^{-\frac{\sigma^2}{2}t^2+it\mu}\right)'}{i} \right|_{t=0} = \left. \frac{(i\mu - 2\sigma t)\, e^{-\frac{\sigma^2}{2}t^2+it\mu}}{i} \right|_{t=0} = \mu.$$

Se deja como ejercicio comprobar que

$$V[\xi] = \sigma^2.$$

Así pues los parámetros de una normal son la esperanza y la desviación típica, i.e. si $\xi \sim N(\mu, \sigma)$

$$E[\xi] = \mu$$

y

$$V[\xi] = \sigma^2.$$

3.3. Vectores aleatorios

Al estudiar un fenómeno en el que intervienen dos o más fuentes de aleatoriedad, el vector aleatorio es el elemento preciso para la descripción del mismo ya que permite un análisis fluido y facilita el cálculo de probabilidades.

Cuando haya dos fuentes de aleatoriedad consideraremos vectores aleatorios bidimensionales, digamos que sta situación va a ser en la que nos detendremos con detalle ya que el caso n-dimensional es análogo a éste.

Nos vamos a ocupar de problemas en los que ocurren simultáneamente dos variables aleatorias, donde cada una de ellas por separado la sabemos manejar, en cambio conjuntamente originan una ley de probabilidad que tiene una estructura concreta, más compleja y que manifiesta la vinculación que desde el punto estadístico tienen los fenómenos aleatorios que concurren en el estudio.

3.3.1. Propiedades fundamentales

Definición 3.3 *Sea $(\Omega, \mathcal{A}, P)$ espacio de probabilidad, y ξ_1 y ξ_2 dos v.a.; definimos el vector aleatorio $\xi = (\xi_1, \xi_2)$ bidimensional como*

$$\begin{array}{rccl} (\xi_1, \xi_2): & \Omega & \rightarrow & \mathbb{R}^2 \\ & \omega & \rightarrow & (\xi_1(\omega), \xi_2(\omega)). \end{array}$$

Este vector inducirá cierta probabilidad P_ξ^* en $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$ donde

$$\mathcal{B}_{\mathbb{R}^2} = \left\{(a,b] \times (c,d] \subset \mathbb{R}^2 : a,b,c,d \in \mathbb{R}\right\},$$

y tal probabilidad inducida se define como

$$P_\xi^* \{(a,b] \times (c,d]\} = P\{\omega \in \Omega : \xi_1(\omega) \in (a,b] \ \ y \ \xi_2(\omega) \in (c,d]\}.$$

DEFINICIÓN 3.4 *Se define la función de distribución del vector aleatorio $\xi = (\xi_1, \xi_2)$ como*

$$F_{(\xi_1,\xi_2)}(x,y) = P\{\omega \in \Omega : \xi_1(\omega) \leq x, \ \xi_2(\omega) \leq y\}.$$

Análogo al caso de las variables aleatorias, la forma de la distribución de un vector se expresa a través de probabilidades acumuladas:

$$F_{(\xi_1,\xi_2)}(x,y) = P_\xi^*((-\infty,x] \times (-\infty,y]).$$

Esto es, $F_\xi(x,y)$ se puede leer como la medida de la probabilidad que hay en el rectángulo infinito $(-\infty,x] \times (-\infty,y]$.

PROPOSICIÓN 3.2

1. $F(+\infty,+\infty) = 1$
2. $F(-\infty,y) = F(x,-\infty) = 0 \quad \forall x,y.$
3. $F(x,y) \leq F(x+h,y) \quad \forall x,y$ *y* $\forall h > 0.$
4. $F(x,y) \leq F(x,y+k) \quad \forall x,y$ *y* $\forall k > 0.$
5. $F(x,y) \leq F(x+h,y+k) \quad \forall x,y, \forall k > 0$ *y* $\forall k \geq 0.$
6. $F(-\infty,-\infty) = 0.$

EJEMPLO 3.6 *Se tira un dado dos veces. Se define la v.a. bidimensional $\xi = (\xi_1, \xi_2)$ donde $\xi_1 = n^0$ puntos obtenidos en el primer lanzamiento y $\xi_2 = n^0$ puntos obtenidos en el segundo lanzamiento. observamos que* $\mathrm{Im}\,\xi = \{(1,1),(1,2),...,(6,6)\} \subset \mathbb{R}^2$. *Este ejemplo es un caso típico de vector de tipo discreto.*

EJEMPLO 3.7 *Se selecciona al azar un estudiante de la UCLM. Definimos* ξ_1 = *el peso del estudiante y* ξ_1 = *la altura del estudiante, y definimos el vector aleatorio bidimensional* $\xi = (\xi_1, \xi_2)$ *cuya imagen podemos suponer incluida en el conjunto* $[30, 200] \times [0'5, 3]$. ξ *definido así es de tipo continuo.*

DEFINICIÓN 3.5 (TIPOS DE VECTORES ALEATORIOS) *Sea* $\xi = (\xi_1, \xi_2)$ *un v.a.*

1. *Diremos que* ξ *es de tipo discreto si la masa o probabilidad se concentra sobre un conjunto finito o infinito numerable de puntos de* $\mathbf{R}^2$*, es decir, si* $\operatorname{Im}\xi$ *es un conjunto finito o infinito numerable, esto es*

$$\operatorname{Im}\xi = \{a_{11}, a_{12}, ..., a_{ij}, ...\} \subset \mathbb{R}^2,$$

y para cada uno de estos puntos podemos definir

$$P\{\omega \in \Omega : \xi(\omega) = a_{ij}\} = P_\xi^*(a_{ij}) = p_{ij}$$

siendo los $p_{ij} > 0$ *y* $\sum_{i,j} p_{ij} = 1$. *A* p_{ij} *lo llamamos peso o probabilidad en* a_{ij}. *La función de distribución de probabilidad asociada al vector* ξ *queda definida como*

$$F_{(\xi_1,\xi_2)}(x, y) = \sum_{x_i \le x} \sum_{y_j \le y} P_\xi^*(x_i, y_j).$$

A $P_\xi^*(x_i, y_j)$ *la llamaremos función de densidad de masa de probabilidad asociada al vector. Nótese que* P_ξ^* *es cero si* $(x_i, y_j) \notin \{a_{11}, a_{12}, ..., a_{ij}, ...\}$. *El conjunto* $\{a_{11}, a_{12}, ..., a_{ij}, ...\}$ *es la imagen del v.a.* ξ *y recibe el nombre de soporte de* P_ξ^*.
Téngase en cuenta que las dos propiedades que caracterizan a cualquier densidad de masa son $p_{ij} > 0$ *y* $\sum_{i,j} p_{ij} = 1$.

2. *Diremos que el v.a.* ξ *es de tipo continuo si su función de distribución* $F_{(\xi_1,\xi_2)}(x, y)$ *es continua en* $\mathbb{R}^2$. *Si* ξ *es de tipo continuo entonces* $\operatorname{Im}\xi$ *es un subconjunto de* $\mathbb{R}^2$ *de área positiva.*

Para esta situación

$$P_\xi^*(B) = \int_B f(x_1, x_2)\, dx_1 dx_2,$$

con $f(x_1, x_2)$ *cierta función que llamaremos de densidad y* B *en* $\mathbb{R}^2$. *Precisando: supongamos que* $B = (x, x+h] \times (y, y+k]$ *es un rectángulo, entonces*

$$\begin{aligned} P^*_\xi(B) &= P\{x < \xi_1 \leq x+h,\ y < \xi_2 \leq y+k\} = \\ &= \left[F_{(\xi_1,\xi_2)}(x+h, y+k) - F_{(\xi_1,\xi_2)}(x+h, y)\right] - \\ &\quad - \left[F_{(\xi_1,\xi_2)}(x, y+k) - F_{(\xi_1,\xi_2)}(x, y)\right] \\ &\overset{TVM}{=} kh\ F''_{xy}(x+\theta' h, y+\theta k), \end{aligned}$$

con lo cual la densidad media será

$$\frac{P^*_\xi(B)}{\acute{a}rea(B)} = \frac{kh\ F''_{xy}(x+\theta' h, y+\theta k)}{kh} = F''_{xy}(x+\theta' h, y+\theta k).$$

Por lo tanto la probabilidad que hay en un rectángulo de lados de longitud infinitesimal dx *y* dy *se evalúa haciendo* $k \to 0$ *y* $h \to 0$, *tal probabilidad es* $F''_{xy}(x, y)dxdy$.

Definimos la densidad de probabilidad asociada al vector ξ *como*

$$f_\xi(x, y) = F''_{xy}(x, y) = \frac{\partial^2 F}{\partial x \partial y}(x, y).$$

Por todo lo dicho podemos establecer que

$$P^*_\xi(B) = \int_B f_\xi(u, v)dvdu = \int_x^{x+h}\int_y^{y+k} f_\xi(u, v)dvdu$$

En general, si $B = (a, b] \times (c, d]$, *entonces*

$$P^*_\xi(B) = \int_a^b\int_c^d f_\xi(u, v)dvdu.$$

En particular la distribución se escribe como

$$F_\xi(x, y) = P^*_\xi((-\infty, x] \times (-\infty, y]) = \int_{-\infty}^x\int_{-\infty}^y f_\xi(u, v)dvdu.$$

Las dos propiedades que definen a una función de densidad son, $f_\xi(x, y) \geq 0$ *y* $\int_{-\infty}^\infty\int_{-\infty}^\infty f_\xi(u, v)dudv = 1$.

EJEMPLO 3.8 *Se realizan dos extracciones sin reemplazamiento en una urna con cuatro bolas rojas y cinco blancas. Encontrar la distribución de probabilidad del vector* $\xi = (\xi_1, \xi_2)$ *si*

$$\xi_i = \begin{cases} \text{1 si sale blanca en i-ésima extracción} \\ \text{0 si sale roja.} \end{cases} .$$

Vemos que $P(B) = 5/9$, $P(R) = 4/9$ *y la densidad de masa de probabilidad se calcula así:*

$$\begin{aligned} P(B,B) &= P(\xi_1 = 1, \xi_2 = 1) = P_\xi^*(1,1) = \frac{5}{9}\frac{4}{8} = \frac{5}{18}, \\ P(B,N) &= P(\xi_1 = 1, \xi_2 = 0) = P_\xi^*(1,0) = \frac{5}{9}\frac{4}{8} = \frac{5}{18}, \\ P(N,N) &= P(\xi_1 = 0, \xi_2 = 0) = P_\xi^*(0,0) = \frac{4}{9}\frac{3}{8} = \frac{3}{18}, \\ P(N,B) &= P(\xi_1 = 0, \xi_2 = 1) = P_\xi^*(0,1) = \frac{4}{9}\frac{5}{8} = \frac{5}{18}. \end{aligned}$$

Determinemos su función de distribución:

$$F_\xi(x,y) = \begin{cases} 0 & si\ x < 0 \text{ ó } y < 0 \\ 3/18 & si\ 0 \le x < 1,\ y < 1 \\ 8/18 & si\ 0 \le x < 1,\ y \ge 1 \\ 8/18 & si\ x \ge 1,\ y < 1 \\ 1 & si\ x \ge 1,\ y \ge 1 \end{cases} .$$

EJEMPLO 3.9 *Si* ξ *es un vector aleatorio de tipo continuo con densidad de probabilidad*

$$f_\xi(x,y) = \begin{cases} 1/9 & si\ x \in (0,3),\ y \in (1,4) \\ 0 & resto \end{cases} ,$$

evaluar $P\{\xi_1 \le 2,\ \xi_2 \le 2\}$ *y determinar la fucnción de distribución* $F_\xi(x,y)$.

3.3.2. Momentos.

DEFINICIÓN 3.6

1. *Se define el momento de orden* (k_1, k_2) *con respecto al origen del vector aleatorio* $\xi = (\xi_1, \xi_2)$ *como*

$$\alpha_{k_1,k_2} = E\left[\xi_1^{k_1}\xi_2^{k_2}\right] = \int_{\mathbb{R}^2} x_1^k x_2^k dF_{(\xi_1,\xi_2)}(x_1,x_2).$$

2. *A*

$$\mu_{k_1,k_2} = E\left[(\xi_1 - \mu_1)^{k_1}(\xi_2 - \mu_2)^{k_2}\right] = \int_{\mathbb{R}^2} (x_1 - \mu_1)^{k_1}(x_2 - \mu_2)^{k_2}\, dF_{(\xi_1,\xi_2)}(x_1,x_2)$$

se le denomina momento centrado del vector aleatorio $\xi = (\xi_1, \xi_2)$ *donde*

$$\mu_i = \int_{\mathbb{R}} x_i dF_{\xi_i}(x_i).$$

Se definen a su vez

$$\begin{aligned} \mu_{2,0} &= \sigma_1^2 = \int_{\mathbb{R}} (x_1 - \mu_1)^2\, dF_{\xi_1}(x_1), \\ \mu_{0,2} &= \sigma_2^2 = \int_{\mathbb{R}} (x_2 - \mu_2)^2\, dF_{\xi_2}(x_2). \end{aligned}$$

3. *Se define la covarianza entre las componentes del vector* $\xi = (\xi_1, \xi_2)$ *como*

$$\operatorname{Cov}(\xi_1, \xi_2) = E\left[(\xi_1 - \mu_1)(\xi_2 - \mu_2)\right] = \int\int_{\mathbb{R}^2} (x_1 - \mu_1)(x_2 - \mu_2)\, dF_{(\xi_1,\xi_2)}(x_1, x_2).$$

La notación habitual para la covarianza es $\operatorname{Cov}(\xi_1, \xi_2) = \sigma_{12}$.

4. *Llamaremos matriz de covarianzas del vector* $\xi = (\xi_1, \xi_2)$ *a la matriz*

$$\mathcal{M} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}.$$

NOTA 3.1

$$\operatorname{Cov}(\xi_1, \xi_2) = E[\xi_1\xi_2] - E[\xi_1]\, E[\xi_2].$$

Supongamos que la variable aleatoria $\psi = g(\xi_1, \xi_2)$ está definida como la composición de una función g con el vector aleatorio $\xi = (\xi_1, \xi_2)$,

$$\Omega \xrightarrow{\xi} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}.$$

Definimos la esperanza y la varianza de ψ como

$$E[\psi] = \int_{\mathbb{R}^2} g(x_1, x_2)\, dF_{(\xi_1,\xi_2)}(x_1, x_2)$$

y

$$V[\psi] = \int_{\mathbb{R}^2} (g(x_1, x_2) - E[\psi])^2\, dF_{(\xi_1,\xi_2)}(x_1, x_2).$$

PROPOSICIÓN 3.3 *Sean a_1 y a_2 números reales y $\xi = (\xi_1, \xi_2)$ v.a.. Entonces*

1.
$$E[a_1\xi_1 + a_2\xi_2] = a_1 E[\xi_1] + a_2 E[\xi_2]$$

2.
$$\begin{aligned} E[a_1\xi_1 + a_2\xi_2] &= \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\ &= a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2\sigma_{12} \end{aligned}$$

NOTA 3.2 *Para el cálculo de momentos hemos de tener en cuenta si el vector $\xi = (\xi_1, \xi_2)$ es discreto o continuo:*

1. Si $\xi = (\xi_1, \xi_2)$ es de tipo continuo entonces

$$\begin{aligned} \alpha_{k_1,k_2} &= \int_{\mathbb{R}^2} x_1^{k_i} x_2^{k_2} f_\xi(x_1, x_2)\, dx_1 dx_2, \\ \mu_{k_1,k_2} &= \int_{\mathbb{R}^2} (x_1 - \mu_1)^{k_1} (x_2 - \mu_2)^{k_2} f_\xi(x_1, x_2)\, dx_1 dx_2 \end{aligned}$$

2. Si el vector ξ es de tipo discreto entonces

$$\begin{aligned} \alpha_{k_1,k_2} &= \sum_i \sum_j a_i^{k_i} a_j^{k_2} P_\xi^*(x_i, y_j), \\ \mu_{k_1,k_2} &= \sum_i \sum_j (x_i - \mu_1)^{k_1} (y_j - \mu_2)^{k_2} P_\xi^*(x_i, y_j). \end{aligned}$$

NOTA 3.3 *Nótese que al evaluar por ejemplo $\mu_1 = \int_{\mathbb{R}} x_1 dF_{\xi_1}(x_1)$ observamos que*

$$\int_{\mathbb{R}} x_1 dF_{\xi_1}(x_{1,}) = \int_{\mathbb{R}^2} x_1 dF_{\xi_1,\xi_2}(x_1, x_2),$$

$F_{\xi_1}(x_1)$ está asociada a la distribución de la variable aleatoria ξ_1, es lo denominamos distribución marginal de ξ_1 del vector $\xi = (\xi_1, \xi_2)$:

3.3.3. Distribuciones marginales.

DEFINICIÓN 3.7 *Sea $\xi = (\xi_1, \xi_2)$ v.a.; se define la distribución marginal de ξ_1 como*

$$F_{\xi_1}(x_1) = F_{(\xi_1,\xi_2)}(x_1, +\infty).$$

Vemos claramente que

$$F_{\xi_1}(x_1) = P\left\{\omega \in \Omega : \xi_1(\omega) \leq x_1\right\} = F_{(\xi_1,\xi_2)}(x_1, +\infty).$$

De manera idéntica se define

$$F_{\xi_2}(x_2) = F_{(\xi_1,\xi_2)}(+\infty, x_2).$$

Veamos cómo son los cálculos de las distribuciones marginales:

1. Caso continuo: para $\xi = (\xi_1, \xi_2)$ v.a. continuo sabemos que

$$F_{(\xi_1,\xi_2)}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{\xi_1,\xi_2}(u, v)\, dvdu,$$

luego

$$\begin{aligned} F_{\xi_1}(x_1) &= F_{(\xi_1,\xi_2)}(x_1, +\infty) = \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} f_{\xi_1,\xi_2}(u, v)\, dvdu \\ &= \int_{-\infty}^{x_1} f_{\xi_1}(u)\, du, \end{aligned}$$

donde

$$f_{\xi_1}(u) = \int_{-\infty}^{\infty} f_{\xi_1,\xi_2}(u, v)\, dv$$

función de densidad de probabilidad asociada a la marginal ξ_1. Análogamente, la densidad de la marginal ξ_2 es

$$f_{\xi_2}(v) = \int_{-\infty}^{\infty} f_{\xi_1,\xi_2}(u, v)\, du.$$

2. Caso discreto: siendo $\xi = (\xi_1, \xi_2)$ v.a. discreto dispondremos de

$$p_{ij} = P^*_{(\xi_1,\xi_2)}(x_i, y_j)$$

(i.e. los puntos de masa son los puntos (x_i, y_j) de $\mathbb{R}^2$) y el cálculo de la masa marginal de ξ_1, en un punto x será

$$P^*_{\xi_1}(x) = \lim_{h \to 0^+} F_\xi(x, \infty) - F_\xi(x - h, \infty),$$

lo que da lugar a

$$P^*_{\xi_1}(x) = \sum_j P^*_{(\xi_1,\xi_2)}(x, y_j),$$

densidad de masa de la marginal ξ_1. Nótese que como los puntos de masa son (x_i, y_j) entonces $P^*_{\xi_1}(x)$ será distinta de cero sólo cuando x coincida con x_i. De la misma forma se define la densidad de masa de la marginal ξ_2

$$P^*_{\xi_2}(y) = \sum_i P^*_{(\xi_1,\xi_2)}(x_i, y).$$

Notación: $P^*_{\xi_1} = P_{\cdot i}$ y $P^*_{\xi_2} = P_{j\cdot}$.
El cálculo de las distribuciones marginales es sencillo:

$$\begin{aligned} F_{\xi_1}(x) &= \text{ suma de los } P^*_{\xi_1}(x_i) \text{ con } x_i \leq x \\ &= \sum_{x_i \leq x} \sum_j P^*_{(\xi_1,\xi_2)}(x_i, y_j); \end{aligned}$$

y

$$F_{\xi_2}(y) = \sum_{y_j \leq y} \sum_i P^*_{(\xi_1,\xi_2)}(x_i, y_j).$$

3.3.4. Distribuciones condicionadas

Dado $\xi = (\xi_1, \xi_2)$ estudiaremos qué quiere decir la distribución de ξ_2 dado que $\xi_1 = x$, i.e. la variable aleatoria condicionada $\xi_2|_{\xi_1=x}$ cuya función de distribución asociada denotaremos por $F(x_2/x_1)$.

Definición 3.8 *Dado $\xi = (\xi_1, \xi_2)$ v.a. definimos la función de distribución de la v.a. ξ_2 condicionada por $\xi_1 = x$ como*

$$F(y/x) = \lim_{h \to 0^+} \frac{F_\xi(x, y) - F_\xi(x - h, y)}{F_{\xi_1}(x) - F_{\xi_1}(x - h)}.$$

Distingamos los resultado del límite anterior según sea el vector de tipo discreto o continuo:
(i) Si $\xi = (\xi_1, \xi_2)$ v.a. de tipo discreto entonces

$$F(y/x) = \sum_{y_j \leq y} \frac{P^*_{(\xi_1,\xi_2)}(x, y_j)}{P^*_{\xi_1}(x)}$$

lo que afirma que la densidad de masa asociada a $F(y/x)$, o a la variable aleatoria $\xi_2|_{\xi_1=x}$ es

$$P^*(y/x) = \frac{P^*_{(\xi_1,\xi_2)}(x,y)}{P^*_{\xi_1}(x)}.$$

De manera similar se define la distribución asociada a la variable condicionada $\xi_1|_{\xi_2=y}$ y su densidad de masa será

$$P^*(x/y) = \frac{P^*_{(\xi_1,\xi_2)}(x,y)}{P^*_{\xi_2}(y)}$$

(ii) Si $\xi = (\xi_1, \xi_2)$ v.a. de tipo continuo entonces

$$F(y/x) = \int_{-\infty}^{y} \frac{f_\xi(x,v)}{f_{\xi_1}(x)} dv,$$

por la tanto la densidad asociada es

$$f(y/x) = \frac{f_\xi(x,y)}{f_{\xi_1}(x)}.$$

Análogamente

$$f(x/y) = \frac{f_\xi(x,y)}{f_{\xi_2}(y)}.$$

Obsérvese que si consideramos la v.a. $\psi = g(\xi_1, \xi_2)$ entonces

$$E[\psi] = \int_{\mathbb{R}^2} g(x,y)\, dF_{\xi_1,\xi_2}(x,y);$$

distingamos los casos: si el vector $\xi = (\xi_1, \xi_2)$ es discreto entonces

$$\begin{aligned} E[\psi] &= \sum_i \sum_j g(x_i, x_j)\, P^*(x_i, y_j) \\ &= \sum_i \sum_j g(x_i, y_j)\, P^*(x_i/y_j)\, P^*_{\xi_2}(y_j). \end{aligned}$$

La situación en la que el vector es de tipo continuo se leva a cabo de manera similar:

$$\begin{aligned} E[\psi] &= \int_{\mathbb{R}^2} g(x,y)\, dF_\xi(x,y)\, dxdy \\ &= \int_{\mathbb{R}^2} g(x,y)\, dF(x/y)\, dF_{\xi_1}(y) \\ &= \int_{\mathbb{R}^2} g(x,y)\, f(x/y)\, f_{\xi_2}(y) dxdy. \end{aligned}$$

3.3.5. Independencia

Dadas dos variables aleatorias ξ_1 y ξ_2 podemos construir un vector aleatorio $\xi = (\xi_1, \xi_2)$ el cual tendrá cierta distribución conjunta $F_\xi(x, y)$. Sabemos además que $f_\xi(x,y) = f(x/y) f_{\xi_2}(y)$ - relación entre densidades de probabilidad - o bien $P^*(x,y) = P^*(x/y) P^*_{\xi_2}(y)$ - relación entre las densidades de masa -. A partir de estas relaciones nos vamos a ayudar para decidir si las variables aleatorias ξ_1 y ξ_2 son independientes:

Definición 3.9 *ξ_1 y ξ_2 son independientes si $\forall B_1 \in \mathcal{B}$ y $\forall B_2 \in \mathcal{B}$ se tiene que*

$$P\{\omega \in \Omega : \xi_1(\omega) \in B_1,\ \xi_2(\omega) \in B_2\} = P\{\xi_1(\omega) \in B_1\} P\{\xi_2(\omega) \in B_2\}.$$

Tal y como ha sido definida la independencia vemos que ésta es difícil de comprobar, no obstante disponemos de la

Proposición 3.4

1. *ξ_1 y ξ_2 son independientes si y sólo si*
$$F_\xi(x,y) = F_{\xi_1}(x) F_{\xi_2}(y) \quad \forall x, y.$$
2. *Si $\xi = (\xi_1, \xi_2)$ es discreto entonces, ξ_1 y ξ_2 son independientes si y sólo si*
$$P^*(x,y) = P^*_{\xi_1}(x) P^*_{\xi_2}(y) \quad \forall x, y.$$
3. *Si $\xi = (\xi_1, \xi_2)$ es continuo entonces, ξ_1 y ξ_2 son independientes si y sólo si*
$$f_\xi(x,y) = f_{\xi_1}(x) f_{\xi_2}(y) \quad \forall x, y.$$

Proposición 3.5 *Sean ξ_1 y ξ_2 variables aleatorias independientes, entonces*

1. $E[\xi_1 \xi_2] = E[\xi_1] E[\xi_2]$.
2. $E[g(\xi_1) h(\xi_2)] = E[g(\xi_1)] E[h(\xi_2)]$.
3. $V[a_1\xi_1 + a_2\xi_2] = a_1^2 V[\xi_1] + a_2^2 V[\xi_2]$.
4. $\varphi_{\xi_1+\xi_2}(t) = \varphi_{\xi_1}(t)\varphi_{\xi_2}(t)$.

DEFINICIÓN 3.10 *Se define la función característica de un vector aleatorio* $\xi = (\xi_1, \xi_2)$ *como*

$$\varphi_{(\xi_1,\xi_2)}(t_1,t_2) = \int_{\mathbb{R}^2} e^{i(t_1x+t_2y)} dF_{(\xi_1,\xi_2)}(x,y).$$

PROPOSICIÓN 3.6 *Si* ξ_1 *y* ξ_2 *son variables aleatorias independientes, entonces*

$$\varphi_{(\xi_1,\xi_2)}(t_1,t_2) = \varphi_{\xi_1}(t_1)\,\varphi_{\xi_2}(t_2).$$

De manera análoga se define la independencia entre n *variables aleatorias* $\xi_1, \xi_2, ..., \xi_n$*, llegándose a*

1. $\xi_1, \xi_2, ..., \xi_n$ *independientes si y sólo si*

$$F_{(\xi_1,\xi_2,...,\xi_n)}(x_1,x_2,...,x_n) = F_{\xi_1}(x_1)\,F_{\xi_2}(x_2)\,...F_{\xi_n}(x_n).$$

2. $(\xi_1, \xi_2, ..., \xi_n)$ *de tipo continuo son independientes si y sólo si*

$$f_{(\xi_1,\xi_2,...,\xi_n)}(x_1,x_2,...,x_n) = f_{\xi_1}(x_1)\,f_{\xi_2}(x_2)\,...f_{\xi_n}(x_n).$$

3. $(\xi_1, \xi_2, ..., \xi_n)$ *de tipo discreto son independientes si y sólo si*

$$P^*_{(\xi_1,\xi_2,...,\xi_n)}(x_1,x_2,...,x_n) = P^*_{\xi_1}(x_1)\,P^*_{\xi_2}(x_2)\,...P^*_{\xi_n}(x_n).$$

4. *Si* $\xi_1, \xi_2, ..., \xi_n$ *son independientes, entonces*

$$\varphi_{\xi_1+\xi_2+...+\xi_n}(t) = \varphi_{\xi_1}(t)\,\varphi_{\xi_2}(t)...\varphi_{\xi_n}(t).$$

EJEMPLO 3.10 *Sea el vector aleatorio* $\xi = (\xi_1, \xi_2)$ *distribución de probabilidad es*

$$F_\xi(x,y) = \begin{cases} (1-e^{-2x})\,(1-e^{-3y}) & si\ x > 0\ e\ y > 0 \\ 0 & en\ el\ resto \end{cases}.$$

Se pide:

1. *Comprobar que en efecto* F_ξ *es distribución de probabilidad.*
2. *Calcular la densidad conjunta de* ξ.
3. *Evaluar* $P(\xi_1 \leq 1, \xi_2 \leq 2)$

4. *Determinar las densidades marginales.*
5. *Estudiar si ξ_1 y ξ_2 son independientes.*

1. *Para verificar que es una distribución de probabilidad basta con ver si su densidad asociada es de probabilidad.*
$$\begin{aligned} f_\xi(x,y) &= \frac{\partial^2 F_\xi(x,y)}{\partial x \partial y} \\ &= \begin{cases} 6e^{-2x}e^{-3y} & \text{si } x>0 \text{ e } y>0 \\ 0 & \text{en el resto} \end{cases} \end{aligned}$$
vemos que $f_\xi(x,y) \geq 0$ y además
$$\begin{aligned} \int_{\mathbb{R}^2} f_\xi(x,y)dxdy &= \int_0^\infty \int_0^\infty 6e^{-2x-3y}dxdy \\ &= \int_0^\infty 6e^{-3y}\left[\int_0^\infty e^{-2x}dx\right]dy \\ &= \int_0^\infty 6e^{-3y}\left[\frac{-1}{2}e^{-2x}\right[_0^\infty dy \\ &= \int_0^\infty 3e^{-3y}dy = 1. \end{aligned}$$
2. *Ya está hecho.*
3. $P(\xi_1 \leq 1, \xi_2 \leq 2) = F(1,2) = (1-e^{-2})(1-e^{-6}),$
4. *Calculamos la densidad marginal de ξ_1 :*
$$\begin{aligned} f_{\xi_1}(x) &= \int_{\mathbb{R}} f_\xi(x,y)dy = \begin{cases} \int_0^\infty 6e^{-2x-3y}dy & \text{si } x>0 \\ 0 & \text{si } x\leq 0 \end{cases} \\ &= \begin{cases} 6e^{-2x}\int_0^\infty e^{-3y}dy & \text{si } x>0 \\ 0 & \text{si } x \leq 0 \end{cases} \\ &= \begin{cases} 2e^{-2x} & \text{si } x>0 \\ 0 & \text{si } x\leq 0 \end{cases} . \end{aligned}$$
De la misma forma vemos que
$$f_{\xi_2}(y) = \begin{cases} 3e^{-3y} & \text{si } y>0 \\ 0 & \text{si } y \leq 0 \end{cases} .$$

5. *Es sencillo comprobar que ξ_1 y ξ_2 son independientes ya que*

$$f_\xi(x,y) = \left[\begin{cases} 2e^{-2x} & si\ x > 0 \\ 0 & si\ x \leq 0 \end{cases} \right] \cdot \left[\begin{cases} 3e^{-3y} & si\ y > 0 \\ 0 & si\ y \leq 0 \end{cases} \right].$$

Ejemplo 3.11 *(continuación del Ejemplo 3). Se pide determinar la distribución marginal de ξ_1. Esta se calcula evaluando*

$$F_{\xi_1}(x) = F_\xi(x,\infty) = \begin{cases} 0 & si\ x < 0 \\ 8/18 & si\ x \in [0,1) \\ 1 & si\ x \geq 1 \end{cases} .$$

A partir de $F_{\xi_1}(x)$ vemos que

$$\begin{aligned} P^*_{\xi_1}(0) &= 8/18 \\ P^*_{\xi_1}(1) &= 10/18, \end{aligned}$$

los únicos puntos de masa de la distribución marginal ξ_1 son 0 y 1.

3.4. Regresión y Correlación

Supongamos que nos dan una tabla de valores de dos variables aleatorias ξ_1 y ξ_2. Consideramos la correspondiente nube de puntos en un sistema de coordenadas $\xi_1\xi_2$. Nuestro objetivo es ver si existe alguna función H tal que $H(\xi_1) = \xi_2$ para todo valor que tomen ξ_1 y ξ_2. Esto en la práctica es imposible, en cambio podemos encontrar una función H de forma que

$$E\left[(\xi_2 - H(\xi_1))^2\right] \tag{3.2}$$

sea mínima entre todas las posibles funciones H.

Que (3.2) sea pequeño dice que la distancia al cuadrado entre los valores de ξ_2 y $H(\xi_1)$ son en media pequeños, lo cual dice que ξ_2 y $H(\xi_1)$ se parecerán más cuanto más pequeña se la esperanza (3.2), de tal forma que si esta esperanza es cero entonces podremos concluir afirmando que $H(\xi_1) = \xi_2$ con probabilidad 1.

Procedamos con la determinación de H : observamos que

$$
\begin{aligned}
E\left[(\xi_2 - H(\xi_1))^2\right] &= \int_{\mathbb{R}^2} (y - H(x))^2 \, dF_{(\xi_1,\xi_2)}(x,y) \\
&= \int_{\mathbb{R}^2} (y - H(x))^2 \, dF(y/x) dF_{\xi_1}(x) \\
&= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} (y - H(x))^2 \, dF(y/x) \right] dF_{\xi_1}(x),
\end{aligned}
$$

con lo cual, $E\left[(\xi_2 - H(\xi_1))^2\right]$ se hará mínimo cuando y sólo cuando

$$
L(x) \doteq \int_{\mathbb{R}} (y - H(x))^2 \, dF(y/x)
$$

se haga mínimo. Observamos que $L(x)$ es una desviación media cuadrática, luego si aplicamos el hecho conocido de que esta desviación es mínima cuando se hace con respecto a la esperanza, concluimos con que L es mínimo al elegir

$$
H(x) = \int_{\mathbb{R}} y \, dF(y/x). \tag{3.3}
$$

A (3.3) lo llamamos regresión de ξ_2 sobre ξ_1 o también esperanza condicionada de ξ_2 cuando $\xi_1 = x$ (o lo que es lo mismo, la esperanza de la variable aleatoria condicionada $\xi_2|_{\xi_1=x}$). Si (3.3) se expresa como $H(x) = E\left[\xi_2|_{\xi_1=x}\right]$, entonces

$$
\min_H E\left[(\xi_2 - H(\xi_1))^2\right] = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \left(y - E\left[\xi_2|_{\xi_1=x}\right] \right)^2 dF(y/x) \right] dF_{\xi_1}(x).
$$

De manera similar se puede buscar una función G tal que haga mínimo el valor de $E\left[(\xi_1 - G(\xi_2))^2\right]$, el análisis es el mismo llegándose a que

$$
G(y) = \int_{\mathbf{R}} x \, dF(x/y)
$$

(regresión de ξ_1 sobre ξ_2) y

$$
\min_H E\left[(\xi_1 - G(\xi_2))^2\right] = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \left(y - E\left[\xi_1|_{\xi_2=y}\right] \right)^2 dF(x/y) \right] dF_{\xi_2}(y).
$$

3.4.1. Regresión lineal

Supóngase que buscamos la regresión $H(x)$ con la restricción de que ésta sea una función lineal, es decir imponiendo que $H(x) = ax + b$. Así pues habremos de determinar los valores de a y b para los cuales se hace mínimo el valor de la integral

$$\int_{\mathbf{R}^2} (y-(ax+b))^2 \, dF_{(\xi_1,\xi_2)}(x,y). \tag{3.4}$$

Para ello considermos la función $\mathcal{F}:\ \mathbf{R}^2 \to \mathbf{R}$ tal que

$$\mathcal{F}(a,b) = \int_{\mathbf{R}^2} (y-(ax+b))^2 \, dF_{(\xi_1,\xi_2)}(x,y).$$

Desarrollamos la integral y queda

$$\mathcal{F}(a,b) = E\left[\xi_2^2\right] + a^2 E\left[\xi_1^2\right] - 2aE\left[\xi_1\xi_2\right] - 2bE\left[\xi_2\right] + 2abE\left[\xi_1\right] + b^2.$$

Para calcular los puntos a y b en los que se alcanza el mínimo resolvemos el sistema de ecuaciones

$$\begin{aligned} \frac{\partial F(a,b)}{\partial a} &= 0, \\ \frac{\partial F(a,b)}{\partial b} &= 0. \end{aligned}$$

Tal sistema es

$$\begin{aligned} aE\left[\xi_1^2\right] + bE\left[\xi_1\right] &= E\left[\xi_1\xi_2\right], \\ aE\left[\xi_1\right] + b &= E\left[\xi_2\right], \end{aligned}$$

con lo cual

$$a = \frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_1]} \quad \text{y} \quad b = E[\xi_2] - E[\xi_1]\frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_1]}.$$

Por consiguiente la regresión buscada es

$$\begin{aligned} H(x) &= \left[\frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_1]}\right]x + \left[E[\xi_2] - E[\xi_1]\frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_1]}\right] = \\ &= E[\xi_2] + \frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_1]}(x - E[\xi_1]) \end{aligned} \tag{3.5}$$

cuyo nombre es la *recta de regresión lineal de* ξ_2 *sobre* ξ_1.

De forma análoga se calcula la *recta de regresión lineal de* ξ_1 *sobre* ξ_2,

$$G(y) = E[\xi_1] + \frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_2]}(y - E[\xi_2]).$$

DEFINICIÓN 3.11 *Al número adimensional*

$$\rho = \frac{\operatorname{Cov}(\xi_1, \xi_2)}{\sqrt[2]{V[\xi_1]\, V[\xi_2]}}$$

se le denomina coeficiente de regresión lineal.

Observamos que ρ^2 coincide con el producto de las dos pendientes de las rectas de regresión lineal; además se cumplen las siguientes propiedades:

PROPOSICIÓN 3.7 *1.* $-1 \leq \rho \leq 1.$

2. *Las rectas de regresión se cortan en el punto* $(E[\xi_1], E[\xi_2])$.

3. *Si* $\rho \neq 0$ *entonces* ξ_1 *y* ξ_2 *no están incorreladas, es decir,* $\operatorname{Cov}(\xi_1, \xi_2) \neq 0$; *en tal caso las rectas de regresión tienen pendientes con el mismo signo y se puede establecer que hay dependencia aleatoria entre las dos variables.*

4. *Si* $\rho = 0$ *entonces* ξ_1 *y* ξ_2 *están incorreladas (no hay relación entre las dos v.a.) y en este caso las rectas son*

$$\begin{aligned} H(x) &= E[\xi_2], \\ G(y) &= E[\xi_1]. \end{aligned}$$

5. *Calculadas las regresiones lineales se obtienen los mínimos (por ejemplo, sustituyendo (3.5) en (3.4))*

$$r_y^2 = E\left[(\xi_2 - H(\xi_1))^2\right] = V[\xi_2]\left(1 - \rho^2\right),$$

varianza residual de la regresión lineal de ξ_2 *sobre* ξ_1*, y*

$$r_x^2 = E\left[(\xi_1 - G(\xi_2))^2\right] = V[\xi_1]\left(1 - \rho^2\right),$$

varianza residual de la regresión lineal de ξ_1 *sobre* ξ_2*.*

6. *Si* $\rho^2 = 1$ *entonces las dos rectas de regresión coinciden y podemos decir que existe relación funcional entre las dos variables.*

EJEMPLO 3.12 *Sea $\xi = (\xi_1, \xi_2)$ un vector aleatorio cuya densidad de masa de probabilidad es*

$$P^*_\xi(0,1) = \frac{1}{10}, \quad P^*_\xi(0,2) = \frac{3}{10}, \quad P^*_\xi(1,1) = \frac{1}{10}, \quad P^*_\xi(1,2) = \frac{5}{10}.$$

hallar la regresión de ξ_2 sobre ξ_1.
Se trata de evaluar

$$\begin{aligned} E\left[\xi_2|_{\xi_1=x}\right] &= \int_{\mathbb{R}} y\, dF(y/x) \\ &= \sum_j y_j P^*(y_j/x). \end{aligned}$$

Vemos que

$$\begin{aligned} P^*_{\xi_1}(0) &= P^*_\xi(0,1) + P^*_\xi(0,2) = 4/10, \\ P^*_{\xi_1}(1) &= P^*_\xi(1,1) + P^*_\xi(1,2) = 6/10; \end{aligned}$$

por tanto

$$P^*(1/x) = \begin{cases} \frac{P^*_\xi(0,1)}{P^*_{\xi_1}(0)} & si\ x=0 \\ \frac{P^*_\xi(1,1)}{P^*_{\xi_1}(1)} & si\ x=1 \end{cases} = \begin{cases} \frac{\frac{1}{10}}{\frac{4}{10}} & si\ x=0 \\ \frac{\frac{1}{10}}{\frac{6}{10}} & si\ x=1 \end{cases}$$

y

$$P^*(2/x) = \begin{cases} \frac{P^*_\xi(0,2)}{P^*_{\xi_1}(0)} & si\ x=0 \\ \frac{P^*_\xi(1,2)}{P^*_{\xi_1}(1)} & si\ x=1 \end{cases} = \begin{cases} \frac{\frac{3}{10}}{\frac{4}{10}} & si\ x=0 \\ \frac{\frac{5}{10}}{\frac{6}{10}} & si\ x=1 \end{cases}.$$

Entonces

$$\begin{aligned} E\left[\xi_2|_{\xi_1=x}\right] &= \sum_j y_j P^*(y_j/x) \\ &= \begin{cases} 1P^*(1/0) + 2P^*(2/0) & si\ x=0 \\ 1P^*(1/1) + 2P^*(2/1) & si\ x=1 \end{cases} \\ &= \begin{cases} 1/4 + 2(3/4) & si\ x=0 \\ 1/6 + 2(5/6) & si\ x=1 \end{cases} \\ &= \begin{cases} 7/4 & si\ x=0 \\ 11/6 & si\ x=1 \end{cases}. \end{aligned}$$

EJEMPLO 3.13 *Sea $\xi = (\xi_1, \xi_2)$ un vector aleatorio cuya densidad de probabilidad es*

$$f_\xi(x,y) = \begin{cases} \frac{1}{4}\left(1 + xy(x^2 - y^2)\right) & \text{si } -1 \leq x \leq 1 \text{ y } -1 \leq y \leq 1 \\ 0 & \text{en el resto.} \end{cases} .$$

Se pide:

1. *Determinar las densidades marginales de ξ_1 y ξ_2.*
2. *La densidad de $\xi_2|_{\xi_1=x}$.*
3. *¿Son ξ_1 y ξ_2 independientes?*
4. *Regresión de ξ_2 sobre ξ_1.*
5. *Recta de refresión lineal de ξ_2 sobre ξ_1.*

1.

$$\begin{aligned} f_{\xi_1}(x) &= \int_{\mathbb{R}} f_\xi\,(x,y)\,dy = \\ &= \begin{cases} 0 & \text{si } x \notin [-1,1] \\ \frac{1}{4}\displaystyle\int_{\mathbb{R}} \left[1 + xy(x^2 - y^2)\right] dy & \text{si } x \in [-1,1] \end{cases} \\ &= \begin{cases} 0 & \text{si } x \notin [-1,1] \\ \frac{1}{2} & \text{si } x \in [-1,1] \end{cases} . \end{aligned}$$

De la misma forma se tiene que

$$f_{\xi_2}(y) = \begin{cases} 0 & \text{si } y \notin [-1,1] \\ \frac{1}{2} & \text{si } y \in [-1,1] \end{cases} .$$

2. $f\,(y/x) = \dfrac{f_\xi\,(x,y)}{f_{\xi_1}(x)} = \frac{1}{2}\displaystyle\int_R \left[1 + xy(x^2 - y^2)\right] dy$ *si* $x \in [-1,1]$.

3. *Puesto que $f_{\xi_1}(x) f_{\xi_2}(y) \neq f_\xi\,(x,y)$ no hay independencia.*

4.

$$\begin{aligned} H(x) &= E\left[\xi_2|_{\xi_1=x}\right] = \int_{\mathbb{R}} y\,dF(y/x) \\ &= \int_{\mathbb{R}} \frac{y}{2}\left(1 + xy(x^2 - y^2)\right) dF(y/x) \text{ si } x \in [-1,1] \\ &= \frac{x^3}{3} - \frac{x}{5} \text{ si } x \in [-1,1] \end{aligned}$$

5. *Sabemos que* $H(x) = ax + b$, *siendo* $a = \frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_1]}$ *y* $b = E[\xi_2] - E[\xi_1]\frac{\mathrm{Cov}(\xi_1,\xi_2)}{V[\xi_1]}$. *Resta pues el evaluar tales coeficientes:*
   $\mathrm{Cov}(\xi_1, \xi_2) = E[\xi_1\xi_2] - E[\xi_1]E[\xi_2]$ *y,*

$$\begin{aligned} E[\xi_1\xi_2] &= \int_{\mathbb{R}^2} xy\, f_\xi(x,y) \\ &= \frac{1}{4}\int_{-1}^{1}\int_{-1}^{-1}(xy + x^4y^2 - x^2y^4)dxdy \\ &= \frac{1}{4}\int_{-1}^{1}\left(\frac{2}{5}y^2 - \frac{2}{3}y^4\right)dy = 0, \end{aligned}$$

   y

$$\begin{aligned} E[\xi_1] &= \int_{\mathbb{R}} x f_{\xi_1}(x)\, dx = 0, \\ E[\xi_2] &= \int_{\mathbb{R}} y f_{\xi_2}(y)\, dy = 0, \end{aligned}$$

   y $V[\xi_1] = 1/3$. *Como consecuencia* $H(x) = 0$.

3.4.2. Regresión descriptiva

En este epígrafe hacemos un breve apartado con objeto de conocer la relación entre dos variables discretas ξ_1 y ξ_2. Damos a conocer los elementos básicos para poder encontrar una fórmula que exprese una variable en función de la otra. El caso que aquí se trata es una situación particular del general que hemos tratado en el apartado anterior.

El punto de partida es una muestra de tamaño n para el vector aleatorio discreto (ξ_1, ξ_2), digamos que es

$$(x_1, y_1), ..., (x_n, y_n)$$

la cual suponemos que representa a la población sobre la que estamos realizando el estudio. De este modo la asignación de las probabilidades tiene sentido sólo en los puntos de la muestra, y su probabiliadad coincide con la frecuencia relativa.

El objetivo es el siguiente: con indepndencia de que se pueda realizar un estudio descriptivo sobre las variables marginales ξ_1 y ξ_2, queremos ver si existe alguna relación funcional que ligue los resultados de una con los de la otra. Para ello, el primer paso será realizar una representación de la nube de puntos que forma la muestra en $\mathbb{R}^2$. Hecho esto se puede en principio, vislumbrar algún tipo de función que pueda relacionar a las variables dadas. En la etapa siguiente nos proponemos estudiar si la relación funcional es lineal[3]: buscamos la recta de regresión lineal. La medida en que tal dependencia lineal es válida la da el siguiente criterio: consideramos la suma de cuadrados

$$E = \frac{1}{n}\sum_{i=1}^{n}(y_i - a - bx_i)^2$$

E es una función que depnde de dos variables, a y b. Al objeto validar una relación lineal de tipo $y = ax + b$ (es decir, a fin de explicar ξ_2 como una función del tipo $a\xi_1 + b$) minimizamos simultánemaente los cuadrados $(y_i - a - bx_i)^2$, lo cual equivale minimizar E[4].

El resultado que se obtiene, la recta de regresión de ξ_2 sobre ξ_1 es

$$y = ax + b = \left(\frac{\mathrm{Cov}\,(\xi_1,\xi_2)}{V\,[\xi_1]}\right)x + \left(E\,[\xi_2] - E\,[\xi_1]\,\frac{\mathrm{Cov}\,(\xi_1,\xi_2)}{V\,[\xi_1]}\right)$$

donde

$$\begin{aligned} \mathrm{Cov}\,(\xi_1,\xi_2) &= \frac{1}{n}\sum_{i=1}^{n}(y_i - E\,[\xi_2])(x_i - E\,[\xi_1]), \\ E\,[\xi_1] &= \frac{1}{n}\sum_{i=1}^{n}x_i, \\ E\,[\xi_2] &= \frac{1}{n}\sum_{i=1}^{n}y_i. \end{aligned}$$

Es importante hacer notar que el formato de la muestra no es general como en (3.5) sino que tiene el formato

$$\{(x_i, y_j)\}$$

[3]Se pueden suponer otro tipo de relaciones no lineales que según las situaciones puedan ajustar mejor los datos y proporcionar un error cuadrático menor.

[4]Como podemos ver el proceso es el mismo que antes, dado que el vector aleatorio es de tipo discreto la integral (3.4) se expresa como

$$\begin{aligned} \int_{\mathbf{R}^2}(y-(ax+b))^2\,dF_{(\xi_1,\xi_2)}(x,y) &= \sum_{i=1}^{n}(y_i-(ax_i+b))^2\,P^*_{(\xi_1,\xi_2)}(x_i,y_i) \\ &= \sum_{i=1}^{n}(y_i-(ax_i+b))^2\,\frac{1}{n} \end{aligned}$$

donde $i = 1, ..., n$, y $j = 1, ..., m$, de suerte que la muestra dada es de tamaño $N = mn$. Asociada a esta muestra se ha de considerar la siguiente tabla de frecuencias absolutas

$\xi_2 \backslash \xi_1$	x_1	...	x_n	
y_1	f_{11}	...	f_{n1}	$f_{\cdot 1}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
y_m	f_{1m}	...	f_{nm}	$f_{\cdot m}$
	$f_{1\cdot}$	...	$f_{n\cdot}$	N

donde f_{ij} es el número de veces que aparece el dato (x_i, y_j) y por ende $P^*_{(\xi_1,\xi_2)}(x_i, y_j) = \frac{f_{ij}}{N}$, $P^*_{\xi_1}(x_i) = \dfrac{f_{i\cdot}}{N}$ y $P^*_{\xi_1}(y_j) = \dfrac{f_{\cdot j}}{N}$.

EJEMPLO 3.14 *El número de bacterias por unidad de volumen presentes en un cultivo después de x horas, viene representado en la tabla siguiente:*

Número de horas	0	1	2	3	4	5
Número de bacterias	12	19	23	34	56	62

Se pide:

1. *Hallar las rectas de regresión y deducir cuantas horas tienen que pasar hasta obtener 100 bacterias*

2. *Determinar el coeficiente de corelación lineal e interpretar su valor.*

Definimos las variables aleatorias del siguiente modo. ξ_1 = *número de horas y* ξ_2 = *número de bácterias.*

1. *Para la primera variable*

$$\begin{aligned} E[\xi_1] &= \frac{15}{6} = 2{,}5, \\ V[\xi_1] &= \frac{\sum_{i=1}^{5}(x_i - 2{,}5)^2}{6} = \frac{\sum_{i=1}^{5} {x_i}^2}{6} - (2{,}5)^2 = 2{,}9 \end{aligned}$$

y para la segunda

$$\begin{aligned} E[\xi_2] &= \frac{206}{6} = 34{,}33, \\ V[\xi_2] &= \frac{\sum_{j=1}^{5}(y_j - 34{,}33)^2}{6} = \frac{\sum_{j=1}^{5} {y_j}^2}{6} - (34{,}33)^2 = 349{,}5 \end{aligned}$$

Además

$$
\begin{aligned}
\operatorname{Cov}(\xi_1, \xi_2) &= \frac{1}{6}\sum_{i=1}^{6}(y_i - E[\xi_2])(x_i - E[\xi_1]) \\
&= \frac{1}{6}\sum_{i=1}^{6}(x_1 y_i) - E[\xi_1]\,E[\xi_2]) \\
&= \frac{701}{6} - (2{,}5)\,(34{,}33) = 31
\end{aligned}
$$

Las rectas de regresión son:

$$
\begin{aligned}
x - 2{,}5 &= \frac{31}{349{,}5}(y - 34{,}33), \\
y - 34{,}33 &= \frac{31}{2{,}9}(x - 2{,}5).
\end{aligned}
$$

Usando por ejemplo, la primera se estima que para $y = 100$ bacterias el número de horas que le correspondería es solución de la ecuación

$$x - 2{,}5 = \frac{31}{349{,}5}(100 - 34{,}33)$$

i.e. $x = 8.\,3248$

2. *Sabemos que* $\rho = \frac{\operatorname{Cov}(\xi_1,\xi_2)}{\sqrt[2]{V[\xi_1]V[\xi_2]}}$, *y esto en nuestro caso da*

$$\rho = \frac{31}{\sqrt[2]{(2{,}9)\,(349{,}5)}} = 0{,}973\,73$$

Esto permite establecer establecer una gran dependencia aleatoria (cuanto más próximo a 1 ó -1 mayor es la correlación o dependencia aleatoria, y en el caso de $\rho = 1$ ó $\rho = -1$ se da la dependencia funcional).

EJERCICIO 3.1 *Se realiza un estudio sobre la presencia de un determinado agente químico en la atmósfera, en relacion con la temperatura en el momento y lugar de la medida. El estudio determina una función de densidad de masa de probabilidad conjunta para el modelo. Las probabilidades se dan en la tabla* *adjunta*

$\xi_1 \backslash \xi_2$	2	4	6	8	10	$P^*_{\xi_1}(x)$
5	0,04	0,02	0,02	0	0	0,08
10	0,02	0,04	0,08	0	0	0,14
15	0	0,06	0,02	0,02	0	0,1
20	0	0,06	0,04	0,02	0	0,12
25	0,02	0,04	0,08	0,02	0	0,16
30	0,04	0,04	0,04	0	0,02	0,14
35	0,02	0,04	0,06	0,06	0,02	0,2
40	0	0,04	0,02	0	0	0,06
$P^*_{\xi_2}(y)$	0,14	0,34	0,36	0,12	0,04	1

Se pide

EJEMPLO 3.15 *1. Distribuciones marginales de las variables ξ_1 y ξ_2.*

2. Distribuciones condicionadas

3. Covarianza y coeficiente de correlación lineal

4. Regresión de ξ_2 sobre ξ_1.

5. Recta de regresión lineal de ξ_2 sobre ξ_1.

3.4.3. Normal multivariante

Un ejemplo de los más importantes de vector aleatorio es el caso de la distribución normal multivariante, que al remitirnos al caso dos dimensional será llamada normal bivariante.

DEFINICIÓN 3.12 *Decimos que el vector $\xi = (\xi_1, \xi_2)$ está distribuido según una normal bivariante de parámetros $\mu = (\mu_1, \mu_2) \in \mathbb{R}^2$ y $R \in \mathcal{M}_{2\times 2}(\mathbb{R})$, matriz real 2×2 definda positiva[5], si la densidad de probabilidad asociada a ξ es*

$$f_\xi(x,y) = \frac{\sqrt{\det R}}{\left(\sqrt{2\pi}\right)^2} \exp\left(\frac{-1}{2}(x-\mu_1, y-\mu_2) R \begin{pmatrix} x-\mu_1 \\ x-\mu_2 \end{pmatrix}\right)$$

para $(x,y) \in \mathbb{R}^2$.

[5]Recuérdese que una matriz real R es definida positiva si $(x,y) R \begin{pmatrix} x \\ y \end{pmatrix} \geq 0$ y es $= 0$ si y sólo si $(x,y) = (0,0)$.

Se demuestra con cierta facilidad que $f_\xi(x,y)$ es una densidad de probabilidad, es decir, que $f_\xi(x,y) \geq 0$ y que $\int_{\mathbb{R}^2} f_\xi(x,y)\,dxdy = 1$. La notación que se emplea para un vector así es $\xi \sim N(\mu, R^{-1})$ donde $\mu = (\mu_1, \mu_2)$ y R^{-1} es la matriz inversa de R.

TEOREMA 3.8 *Supongamos que* $\xi = (\xi_1, \xi_2) \sim N((\mu_1, \mu_2), R^{-1})$ *con*

$$R^{-1} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix},$$

las distribuciones marginales de ξ_i *, para* $1 = 1, 2$, *son variables aleatorias con distribución* $N\left((\mu_i, \sqrt[2]{\sigma_{ii}}\right)$ *, esto es, las densidades marginales de probabilidad son*

$$f_{\xi_i}(x_i) = \frac{1}{\sqrt{2\pi}\sigma_{ii}} \exp\left(\frac{-(x_i - \mu_i)^2}{2\sigma_{ii}}\right).$$

Además, la distribución condicionada de $\xi_2|_{\xi_1 = x}$ *es*

$$N(\mu_2 + \frac{\sigma_{12}}{\sigma_{11}}(x - \mu_1), \sqrt[2]{\sigma_{22}}\sqrt[2]{1 - \rho^2})$$

donde ρ *es el coeficiente de correlación lineal definido anteriormente:*

$$\rho = \frac{\sigma_{12}}{\sqrt[2]{\sigma_{11}\sigma_{22}}} = \frac{\operatorname{Cov}(\xi_1, \xi_2)}{\sqrt[2]{V[\xi_1]\,V[\xi_2]}}$$

TEOREMA 3.9 *Supongamos que* $\xi = (\xi_1, \xi_2) \sim N((\mu_1, \mu_2), R^{-1})$, *entonces*

1. $E[\xi_i] = \mu_i$.

2. R^{-1} *es la matriz de covarianzas del vector* $\xi = (\xi_1, \xi_2)$, *i.e.*

$$R^{-1} = \mathcal{M} = (\operatorname{Cov}(\xi_i, \xi_j)).$$

3.
$$f_\xi(x,y) = \frac{\exp\left(\frac{-1}{2}(\overrightarrow{x} - \overrightarrow{\mu})^t (\operatorname{Cov}(\xi_i, \xi_j))^{-1} (\overrightarrow{x} - \overrightarrow{\mu})\right)}{2\pi\sqrt{\det(\operatorname{Cov}(\xi_i, \xi_j))}},$$

siendo $(\overrightarrow{x} - \overrightarrow{\mu}) = \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix}$ *y* $(\overrightarrow{x} - \overrightarrow{\mu})^t = (x - \mu_1, y - \mu_2)$.

TEOREMA 3.10 *Supongamos que* $\xi = (\xi_1, \xi_2) \sim N((\mu_1, \mu_2), R^{-1})$, *entonces* ξ_1 *y* ξ_2 *son independientes si y sólo si* $\operatorname{Cov}(\xi_1, \xi_2) = 0$.

3.4.4. Ejemplos

3.5. Problemas

1. La aplicación de determinado producto químico en el tratamiento de aguas residuales produce cierta mejoría en el 70 % de los casos. Se aplica tal producto en diez plantas de tratamiento de aguas. Se pide:

   a) La probabilidad de que mejore el agua en 4 de las plantas.

   b) La probabilidad de que al menos mejore en 3 de las plantas.

2. Una partícula se mueve en un fluido de manera rectilínea, digamos que lo hace a lo largo del eje coordenado OX del plano $\mathbb{R}^2$. Sobre este movimiento se sabe lo siguiente: al cabo de un segundo la partícula se mueve a la derecha 0,001 mm con una probabilidad de $\frac{1}{60}$ y $-0{,}001$ mm a la izquierda con una probabilidad $1-\frac{1}{60}$. Se supone además que cada movimiento en un segundo es independiente del movimiento realizado en segundos anteriores. Se pide:

   a) Calcular la probabilidad de que la partícula realice 15 movimientos a la derecha al cabo de 360 segundos.

   b) Calcular la probabilidad de que la partícula, al cabo de 360 segundos, esté en el punto $(-0{,}3, 0)$ si se supone que en el instante inicial (justamente antes de empezar a contabilizar los 360 segundos) se encontraba en el origen de coordenadas $(0,0)$.

   c) Encontrar la posición esperada de la partícula al cabo de los 360 segundos.

3. La probabilidad de que se produzca un choque entre una partícula de un fluido y las paredes del recinto en el que esta confinado es 0.002. Sea ψ la v.a. que representa el número de choques en un grupo de 1200 partículas. Se pide:

   a) $P(\psi \leq 5)$

   b) $P(\psi = 7)$

4. Una jaula de laboratorio contine 25 ratones, de los que 8 son blancos y el resto pardos. Un ayudante de laboratorio con los ojos vendados extrae sin reemplazaminto 4 ratones de la jaula. Calcular la probabilidad que sólamente uno de los ratones sea blanco.

5. Si X es una v.a. con distribución $N(2,3)$,calcúlese:

   a) $P(X \leq 6{,}32)$ y $P(6{,}15 \leq X \leq 7{,}35)$.

   b) $P(X^2 \geq 3{,}15)$.

6. Si X es una variable aleatoria con distribución $N(3, 0{,}5)$, calcúlese:

   a) $P(X \leq 3{,}32)$ y $P(2{,}15 \leq X \leq 3{,}35)$.

   b) $P(X^2 \geq 9)$.

7. Supongamos que el número de llamadas que recibe una centralita en 5 minutos viene dado por una v.a. ψ tal que $\psi \sim \mathcal{P}(5)$. Se piden las siguientes probabilidades:

   a) De tener 6 llamadas en 5 miniutos

   b) De tener 3 en 10 minutos.

8. Las velocidades de dos partículas dentro de un fluido vienen dadas por lasvariables aleatorias ξ_1y ξ_2 respectivamente. Se supone que la función de densidad de probabilidad para ξ_1 es

$$f_{\xi_1}(x) = \begin{cases} 2 & \text{si } x \in (0, 0{,}5) \\ 0 & \text{en el resto,} \end{cases}$$

   y que para todo $x \in (0, 0{,}5)$ la densidad de probabilidad de ξ_2 condicionada por la primera es

$$f\left(y|\, x\right) = \begin{cases} \frac{1}{2} + x\,(y-1) & \text{si } y \in (0,2) \\ 0 & \text{en el resto.} \end{cases}$$

   Se pide:

   a) Determinar la función de densidad de probabilidad conjunta del vector aleatorio (ξ_1, ξ_2).

b) Hallar las densidades marginales del vector aleatorio (ξ_1, ξ_2).

c) ¿Qué velocidad se espera para la segunda partícula cuando la velocidad de la primera es de 0.75 unidades?

9. Dado el vector aleatorio $\xi = (\xi_1, \xi_2)$ cuya ley de probabilidades viene dada a través de la función de densidad de probabilidad

$$f_\xi(x,y) = \begin{cases} x+y & \text{si } x \in (0,1) \text{ e } y \in (0,1) \\ 0 & \text{en el resto,} \end{cases}$$

se pide:

a) Calcular la probabilidad del suceso $\{\xi_1 \leq 0{,}5, \xi_2 \leq 0{,}2\}$.

b) ¿Son ξ_1 y ξ_2 independientes?

c) Calcular $E\left[\xi_1|_{\xi_2=0{,}5}\right]$.

10. Sean, ξ_1 la v.a. que representa a los gastos por impagados en millones de euros de una determinada empresa, y ξ_2 la v.a. de los ingresos netos. Se sabe que la función de densidad del vector (ξ_1, ξ_2) es

$$f_{(\xi_1,\xi_2)}(x,y) = \begin{cases} 5(1+x)e^{20(2-y)} & \text{si } x \in (0,2) \text{ e } y > 2 \\ 0 & \text{en el resto} \end{cases}$$

a) Calcular la probabilidad de que los gastos por impagados sean menores a un millón de euros.

b) ¿Son independientes los gastos por impagados y los ingresos netos?

c) ¿Qué cantidad de ingresos netos se esperan obtener si los impagados ascendieron hasta 1 millón de euros?

11. Una comapañía aérea sabe por experiencia que el 12 % de las reservas telefónicas de plazas no se llevan a efecto, de modo que reserva más plazas de las que dispone. Si en un vuelo hay 150 plazas, ¿cuántas reservas puede hacer la compañía para que la probabilidad de cubrir al menos 145 plazas sea del 99 %?
Si la compañía reserva 160 plazas, ¿cuál es la probabilidad de que, al menos un pasajero no tenga plaza disponible a la hora de embarcar?

12. Una fábrica produce fusibles eléctricos, resultando defectuosos el 3 %. Los fusibles se empaquetan en cajas de 24 unidades. Se pide:

a) Calcular la probabilidad de que una caja elegida al azar contenga al menos un fusible defectuoso.

b) Si selecionamos 5 cajas al azar, ¿cuál es la probabilidad de que exactamente en 2 de ellas no haya ningún fusible defectuoso?

c) Vamos selecionando cajas al azar y comprobando si sus fusibles son defectuosos o no. ¿Cuál es el número medio de cajas que tendremos que inspeccionar hasta encontar algún fusible defectuoso?

13. Se sabe que la distribución ξ de los coeficientes intelectuales de los alumnos de un colegio sigue la ley nomal. Se sabe que $P\left(\xi \geq 1{,}4\right) = 0{,}1056$ y que $P(\xi > 1) = 0{,}4013$. Calcular los parámetros μ y σ de la citada distribución.

14. La duración media sin rotura de las sábanas de un Sanatorio de 300 camas es de 50 días con una desviación típica de 8. Si suponemos que la duración de las sábanas sigue una distribución normal, calcular:

a) ¿Cuántas sábanas se habrán tenido que reponer antes de los 35 días?

b) ¿Cuántas sábanas habrá que reponer pasados 60 días?