

# UN CATÁLOGO DE MODELOS DE PROBABILIDAD

Julián de la Horra

Departamento de Matemáticas U.A.M.

## 1 Introducción

En este capítulo vamos a dar un catálogo de algunos de los modelos de probabilidad más utilizados, tanto en los desarrollos teóricos como en las aplicaciones.

En los capítulos anteriores hemos visto que lo esencial en las variables aleatorias y en los vectores aleatorios, no son los modelos de probabilidad de partida ni las propias funciones que definen la variable o el vector aleatorio. Lo más importante son los modelos de probabilidad que se generan sobre  $\mathfrak{R}$  o sobre  $\mathfrak{R}^n$ . Por eso, en este capítulo vamos a dar, en primer lugar, modelos de probabilidad discretos sobre  $\mathfrak{R}$ , continuaremos con modelos de probabilidad continuos sobre  $\mathfrak{R}$ , y terminaremos con el modelo de probabilidad Normal bivariante sobre  $\mathfrak{R}^2$ .

## 2 Modelos de probabilidad discretos sobre $\mathfrak{R}$

Los primeros modelos de probabilidad que vamos a definir están basados en las llamadas pruebas de Bernoulli.

**Definición.-** Una *prueba de Bernoulli* consiste en considerar una partición de todos los posibles resultados de un experimento aleatorio en dos conjuntos que llamaremos éxito ( $E$ ) y fracaso ( $F$ ), con  $P(E) = p$  y  $P(F) = 1 - p$ .

Ejemplos de pruebas de Bernoulli:

(a) Podemos considerar el lanzamiento de una moneda equilibrada, y tomar  $E = \{Cara\}$  y  $F = \{Cruz\}$ . En este caso,  $P(E) = 1/2$  y  $P(F) = 1/2$ .

(b) Podemos considerar el lanzamiento de un dado equilibrado, y tomar  $E = \{1, 2\}$  y  $F = \{3, 4, 5, 6\}$ . En este caso,  $P(E) = 1/3$  y  $P(F) = 2/3$ .

(c) Podemos considerar el nivel de colesterol de una persona elegida al azar en una población, y tomar  $E = \text{"Nivel de colesterol superior a 200"}$  y  $F = \text{"Nivel de colesterol inferior a 200"}$ . En este caso, los valores de  $P(E) = p$

y  $P(F) = 1 - p$  no son previamente conocidos. Se pueden utilizar métodos de inferencia estadística para estimarlos.

El primer modelo de probabilidad que vamos a ver consiste, simplemente, en una codificación sencilla de los posibles resultados de una prueba de Bernoulli:

**Definición.-** Realizamos una prueba de Bernoulli con  $P(E) = p$ . El *modelo o distribución de Bernoulli*( $p$ ) es el modelo de probabilidad generado por la variable aleatoria que resulta de codificar el resultado de la prueba de Bernoulli de la siguiente forma:

$$X = \begin{cases} 1 & \text{si obtenemos } E \\ 0 & \text{si obtenemos } F \end{cases}$$

La función de masa generada por esta variable aleatoria es:

$$P(X = 0) = 1 - p \quad P(X = 1) = p$$

que también podemos escribir de una forma más condensada:

$$P(X = x) = p^x (1 - p)^{1-x} \quad \text{para } x = 0, 1.$$

La esperanza del modelo de Bernoulli( $p$ ) es:

$$E[X] = \sum_n x_n P(X = x_n) = (0)(1 - p) + (1)(p) = p$$

La varianza del modelo de Bernoulli( $p$ ) es:

$$\begin{aligned} V(X) &= \sum_n (x_n - E[X])^2 P(X = x_n) = \sum_n x_n^2 P(X = x_n) - (E[X])^2 \\ &= (0)^2(1 - p) + (1)^2(p) - p^2 = p - p^2 = p(1 - p) \quad \bullet \end{aligned}$$

El modelo de probabilidad más importante generado a partir de pruebas de Bernoulli es el modelo binomial:

**Definición.-** Realizamos  $n$  pruebas de Bernoulli independientes con  $P(E) = p$  en cada prueba. El *modelo o distribución Binomial*,  $Bin(n; p)$ , es el modelo de probabilidad generado por la variable aleatoria  $X = \text{“Número de éxitos obtenidos en las } n \text{ pruebas”}$ . La función de masa generada por esta variable aleatoria es sencilla de obtener:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad \text{para } x = 0, 1, \dots, n$$

Para obtener la esperanza y la varianza de la distribución Binomial podemos aplicar directamente las definiciones, pero ese camino es enormemente tedioso. Hay un camino indirecto que es mucho más cómodo. Para esto, definimos para  $i = 1, \dots, n$ :

$$X_i = \begin{cases} 1 & \text{si obtenemos } E \text{ en la prueba } i\text{-ésima} \\ 0 & \text{si obtenemos } F \text{ en la prueba } i\text{-ésima} \end{cases}$$

De esta forma, tenemos que  $X_1, \dots, X_n$  son variables aleatorias independientes con distribución de Bernoulli( $p$ ) y, además,  $X = X_1 + \dots + X_n$ . Por tanto:

$$E[X] = E[X_1 + \dots + X_n] = E[X_1] + \dots + E[X_n] = p + \dots + p = np$$

$$V(X) = V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n) = p(1-p) + \dots + p(1-p) = np(1-p) \quad \bullet$$

Calcular probabilidades correspondientes a la distribución Binomial, si  $n$  no es muy grande, no es nada complicado mediante una calculadora, y también pueden utilizarse las tablas disponibles en muchos libros. Cuando  $n$  es grande, veremos qué aproximaciones pueden utilizarse.

La forma más natural de presentar el siguiente modelo es como límite del modelo Binomial:

**Definición.-** El *modelo o distribución de Poisson*,  $Poisson(\lambda)$ , surge como límite del modelo Binomial,  $Bin(n; p)$ , cuando  $n \rightarrow \infty$ ,  $p \rightarrow 0$ , and  $np \rightarrow \lambda$ . Por lo tanto, su función de masa, para  $x = 0, 1, \dots$ , es de la forma:

$$\begin{aligned} P(X = x) &= \lim \binom{n}{x} p^x (1-p)^{n-x} \\ &= \lim \frac{n(n-1) \dots (n-x+1)}{x!} p^x (1-p)^{n-x} \\ &= \lim \frac{n(n-1) \dots (n-x+1)}{x! n^x} (np)^x (1-p)^{n-x} \\ &= \lim \frac{1}{x!} \left( \frac{n}{n} \frac{n-1}{n} \dots \frac{n-x+1}{n} \right) (np)^x \frac{(1-p)^n}{(1-p)^x} \\ &= \frac{\lambda^x}{x!} \lim (1-p)^n = \frac{\lambda^x}{x!} e^{\lim(-pn)} = \frac{e^{-\lambda} \lambda^x}{x!} \end{aligned}$$

La esperanza del modelo de Poisson ( $\lambda$ ) es:

$$\begin{aligned}
E[X] &= \sum_n x_n P(X = x_n) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=1}^{\infty} x \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} = e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{(y+1)}}{y!} \quad (y = x - 1) \\
&= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda e^{-\lambda} e^{\lambda} = \lambda
\end{aligned}$$

La varianza del modelo de Poisson( $\lambda$ ) se obtiene mediante unos cálculos similares:

$$\begin{aligned}
V(X) &= \sum_n (x_n - E[X])^2 P(X = x_n) = \sum_n x_n^2 P(X = x_n) - (E[X])^2 \\
&= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda} \lambda^x}{x!} - \lambda^2 = \dots = (\lambda^2 + \lambda) - \lambda^2 = \lambda \quad \bullet
\end{aligned}$$

Teniendo en cuenta que el modelo de Poisson( $\lambda$ ) surge como límite del modelo Bin( $n$ ;  $p$ ) podemos también obtener la esperanza y la varianza de una forma rápida e intuitiva, aunque menos rigurosa:

$$\begin{aligned}
E[X] &= \lim np = \lambda \\
V(X) &= \lim np(1 - p) = [\lim np][\lim(1 - p)] = \lambda
\end{aligned}$$

El modelo de Poisson ( $\lambda$ ) se utiliza fundamentalmente como aproximación del modelo Binomial,  $Bin(n; p)$ , cuando  $n \rightarrow \infty$  (es decir, cuando  $n$  es grande) y  $p \rightarrow 0$  (es decir, cuando  $p$  es próximo a cero).

Por fijar ideas, podríamos decir que aproximamos el modelo  $Bin(n; p)$  mediante un modelo de Poisson ( $\lambda$ ), cuando  $n \geq 30$ ,  $p \leq 0,10$ , y tomamos  $\lambda$  como  $\lambda = np$ .

Ejemplos de modelos de Poisson:

(a) Podemos utilizar el modelo de Poisson para modelizar la variable aleatoria “Número de erratas por página en un libro”, ya que cada vez que se imprime un carácter podemos considerar que se está llevando a cabo una prueba de Bernoulli donde  $E$ =“Errata”. Evidentemente, el número de caracteres impresos por página es grande, y la probabilidad de errata es pequeña.

(b) Podemos utilizar el modelo de Poisson para modelizar la variable aleatoria “Número de siniestros anuales de cierto tipo a los que tiene que hacer frente una compañía de seguros”, ya que cada póliza de seguro se

puede considerar como una prueba de Bernoulli donde  $E$  = “Siniestro”. Evidentemente, el número de pólizas es grande, y la probabilidad de siniestro es pequeña.

Si en vez de fijar el número de pruebas de Bernoulli que vamos a llevar a cabo, lo que fijamos es el número de éxitos antes de detener la experimentación, obtenemos otros modelos o distribuciones de probabilidad. Veremos a continuación la distribución Geométrica:

**Definición.-** Realizamos pruebas de Bernoulli independientes, con  $P(E) = p$  en cada prueba, hasta obtener el primer éxito. El *modelo o distribución Geométrica* (con parámetro  $p$ ) es el modelo de probabilidad generado por la variable aleatoria  $X$  = “Número de fracasos obtenidos antes de obtener el primer éxito”. La función de masa generada por esta variable aleatoria es sencilla de obtener:

$$P(X = x) = (1 - p)^x p \quad \text{para } x = 0, 1, \dots$$

La esperanza del modelo Geométrico se calcula de la siguiente forma:

$$\begin{aligned} E[X] &= \sum_n x_n P(X = x_n) \\ &= \sum_{x=0}^{\infty} x(1-p)^x p = \sum_{x=1}^{\infty} x(1-p)^x p \\ &= p \begin{pmatrix} (1-p) & & \\ (1-p)^2 & + (1-p)^2 & \\ (1-p)^3 & + (1-p)^3 & + (1-p)^3 \\ \dots & \dots & \dots \end{pmatrix} \\ &= p \left[ \frac{(1-p)}{1 - (1-p)} + \frac{(1-p)^2}{1 - (1-p)} + \dots \right] \\ &= (1-p) + (1-p)^2 + (1-p)^3 + \dots \\ &= \frac{(1-p)}{1 - (1-p)} = \frac{1-p}{p} \quad \bullet \end{aligned}$$

Desafortunadamente, no hay unanimidad en los libros a la hora de definir la distribución Geométrica. A continuación, se da una definición ligeramente distinta, y vemos cómo es la relación con la definición anterior:

**Definición (alternativa).-** Realizamos pruebas de Bernoulli independientes, con  $P(E) = p$  en cada prueba, hasta obtener el primer éxito. El *modelo o distribución Geométrica (alternativa)* (con parámetro  $p$ ) es el modelo de

probabilidad generado por la variable aleatoria  $Y =$  “Número de pruebas realizadas hasta obtener el primer éxito”.

La relación entre las dos variables aleatorias es obvia:

$$Y = X + 1$$

Esta relación nos lleva fácilmente a la función de masa y la esperanza de la variable aleatoria  $Y$ :

$$P(Y = y) = P(X + 1 = y) = P(X = y - 1) = (1 - p)^{y-1}p \quad \text{para } y = 1, 2, \dots$$

$$E[Y] = E[X + 1] = E[X] + 1 = \frac{1 - p}{p} + 1 = \frac{1}{p} \quad \bullet$$

Pasamos ahora a la distribución Binomial Negativa:

**Definición.-** Realizamos pruebas de Bernoulli independientes, con  $P(E) = p$  en cada prueba, hasta obtener  $r$  éxitos. El *modelo o distribución Binomial Negativa* (con parámetros  $p$  y  $r$ ) es el modelo de probabilidad generado por la variable aleatoria  $X =$  “Número de fracasos obtenidos antes de obtener el  $r$ -ésimo éxito”. La función de masa generada por esta variable aleatoria es sencilla de obtener:

$$P(X = x) = \binom{x + r - 1}{x} (1 - p)^x p^r \quad \text{para } x = 0, 1, \dots$$

Para calcular, de forma sencilla, la esperanza de la distribución Binomial Negativa, vamos a utilizar un procedimiento similar al utilizado para la distribución Binomial. Definimos para  $i = 1, \dots, r$ :

$X_i =$  “Número de fracasos entre el  $(i - 1)$ -ésimo éxito y el  $i$ -ésimo éxito”

De esta forma, tenemos que  $X_1, \dots, X_r$  son variables aleatorias independientes con distribución Geométrica (con parámetro  $p$ ) y, además,  $X = X_1 + \dots + X_r$ . Por tanto:

$$E[X] = E[X_1 + \dots + X_r] = E[X_1] + \dots + E[X_r] = \frac{1 - p}{p} + \dots + \frac{1 - p}{p} = \frac{r(1 - p)}{p} \quad \bullet$$

Desafortunadamente, con la distribución Binomial Negativa ocurre lo mismo que con la distribución Geométrica: no hay unanimidad en los libros a la hora de definirla. A continuación, se da una definición ligeramente distinta, y vemos cómo es la relación con la definición anterior:

**Definición (alternativa).**- Realizamos pruebas de Bernoulli independientes, con  $P(E) = p$  en cada prueba, hasta obtener  $r$  éxitos. El *modelo o distribución Binomial Negativa (alternativa)* (con parámetros  $p$  y  $r$ ) es el modelo de probabilidad generado por la variable aleatoria  $Y =$  “Número de pruebas realizadas hasta obtener el  $r$ -ésimo éxito”.

La relación entre las dos variables aleatorias es obvia:

$$Y = X + r$$

Esta relación nos lleva fácilmente a la función de masa y la esperanza de la variable aleatoria  $Y$ :

$$P(Y = y) = P(X + r = y) = P(X = y - r) = \binom{y-1}{y-r} (1-p)^{y-r} p^r \quad \text{para } y = r, r+1, \dots$$

$$E[Y] = E[X + r] = E[X] + r = \frac{r(1-p)}{p} + r = \frac{r}{p} \quad \bullet$$

Para acabar con este catálogo de modelos para variables aleatorias discretas, vamos a abordar la distribución Hipergeométrica. Todas las distribuciones anteriores se obtienen a partir de pruebas de Bernoulli independientes, que podríamos visualizar como extracciones con reemplazamiento de una urna en la que hay una proporción  $p$  de éxitos y, por tanto, una proporción  $1 - p$  de fracasos. Naturalmente, si las extracciones se efectúan sin reemplazamiento, la proporción de éxitos y fracasos va cambiando a medida que hacemos extracciones y, por lo tanto, el proceso sería diferente. Las extracciones sin reemplazamiento son la base de la distribución Hipergeométrica:

**Definición.**- Consideramos una población con  $N$  elementos, de los cuales  $D$  son éxitos y  $N - D$  fracasos. El *modelo o distribución Hipergeométrica* es el modelo de probabilidad generado por la variable aleatoria  $X =$  “Número de éxitos obtenidos en  $n$  observaciones sin reemplazamiento de la población”. Su función de masa es de la forma:

$$P(X = x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}$$

Observaciones sobre la distribución Hipergeométrica:

(a) Para obtener la anterior función de masa es conveniente darse cuenta de que la obtención de  $n$  elementos de uno en uno sin reemplazamiento es equivalente a la obtención simultánea de  $n$  elementos.

(b) Normalmente, los valores que puede tomar una variable aleatoria con distribución hipergeométrica son  $x = 0, 1, \dots, n$ . Pero esto no es cierto si el número de éxitos o el número de fracasos en la población es menor que el número  $n$  de observaciones. La expresión general válida para todas las situaciones es:

$$x = \max\{0, n - (N - D)\}, \dots, \min\{n, D\}$$

(c) Si el número de éxitos ( $D$ ) y de fracasos ( $N - D$ ) en toda la población es muy grande en comparación con el número  $n$  de observaciones, la diferencia entre tomar las observaciones con o sin reemplazamiento es insignificante. Por tanto, en esa situación, la distribución Hipergeométrica se comporta prácticamente igual que la distribución  $B(n; p)$ , donde  $p = D/N$ .

Para obtener la esperanza de la distribución Hipergeométrica vamos a seguir un camino similar al utilizado para el modelo Binomial. Para esto, definimos para  $i = 1, \dots, n$ :

$$X_i = \begin{cases} 1 & \text{si obtenemos éxito en la } i\text{-ésima observación} \\ 0 & \text{si obtenemos fracaso en la } i\text{-ésima observación} \end{cases}$$

De esta forma, tenemos que  $X_1, \dots, X_n$  son variables aleatorias no independientes, pero (es fácil probar que) todas ellas tienen distribución de Bernoulli, con  $p = D/N$ . Además,  $X = X_1 + \dots + X_n$ . Por tanto:

$$E[X] = E[X_1 + \dots + X_n] = E[X_1] + \dots + E[X_n] = \frac{D}{N} + \dots + \frac{D}{N} = \frac{nD}{N} \quad \bullet$$

### 3 Modelos de probabilidad continuos sobre $\mathfrak{R}$

**Definición.-** El *modelo o distribución Uniforme* de parámetros  $a$  y  $b$  (con  $a < b$ ),  $U(a, b)$ , es un modelo para variables aleatorias continuas con función de densidad:

$$f(x) = \frac{1}{b - a} \quad \text{para } a < x < b.$$

La idea intuitiva es que este modelo reparte uniformemente la probabilidad en el intervalo  $(a, b)$ .



Es fácil comprobar que es función de densidad:

$$\int_{\mathfrak{R}} f(x) dx = \int_a^b \frac{1}{b-a} dx = 1$$

Además, tenemos:

$$E[X] = \int_{\mathfrak{R}} x f(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

$$\begin{aligned} V(X) &= E[X^2] - (E[X])^2 = \int_{\mathfrak{R}} x^2 f(x) dx - (E[X])^2 \\ &= \int_a^b x^2 \frac{1}{b-a} dx - \left( \frac{a+b}{2} \right)^2 = \frac{1}{b-a} \left[ \frac{x^3}{3} \right]_a^b - \left( \frac{a+b}{2} \right)^2 \\ &= \frac{b^3 - a^3}{3(b-a)} - \left( \frac{a+b}{2} \right)^2 = \frac{b^2 + ab + a^2}{3} - \frac{a^2 + 2ab + b^2}{4} \\ &= \frac{b^2 - 2ab + a^2}{12} = \frac{(b-a)^2}{12} \quad \bullet \end{aligned}$$

La distribución Normal es, con gran diferencia, la distribución más utilizada para modelizar multitud de variables: características morfológicas en una población (pesos, envergaduras, ...) mediciones obtenidas en experimentos físicos, químicos, ...

**Definición.-** El modelo o distribución *Normal* de parámetros  $\mu$  y  $\sigma$  (con  $\mu \in \mathfrak{R}$  y  $\sigma > 0$ ),  $N(\mu; \sigma^2)$ , es un modelo para variables aleatorias continuas con función de densidad:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{para } x \in \mathfrak{R}$$

En muchos libros, se representa por  $N(\mu; \sigma)$ . Podemos comprobar que, efectivamente,  $f(x)$  es una función de densidad:

$$\begin{aligned} \int_{\mathfrak{R}} f(x) dx &= \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathfrak{R}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathfrak{R}} e^{-\frac{y^2}{2}} dy \quad \left( \text{definiendo } y = \frac{x-\mu}{\sigma} \right) \\ &= \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{y^2}{2}} dy \quad (\text{por ser una función par}) \\ &= \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-h} h^{-1/2} dh \quad \left( \text{definiendo } h = \frac{y^2}{2} \right) \\ &= \frac{1}{\sqrt{\pi}} \Gamma(1/2) = \frac{1}{\sqrt{\pi}} \sqrt{\pi} = 1 \end{aligned}$$

Podemos calcular también la esperanza y la varianza del modelo Normal:

$E[X] = \mu$  (por ser  $f(x)$  una función de densidad simétrica con respecto a  $\mu$ ).

$V(X) = \sigma^2$  (tras unos laboriosos cálculos). •

Vemos a continuación dos propiedades muy útiles de la distribución Normal: su tipificación o paso a la  $N(0; 1)$ , y el hecho de que cualquier combinación lineal de variables aleatorias Normales independientes, sigue teniendo una distribución Normal, con una estructura muy sencilla:

### Propiedades.-

(a) Sea  $X$  una variable aleatoria con distribución  $N(\mu; \sigma^2)$ . Entonces, la variable aleatoria  $Z = \frac{X-\mu}{\sigma}$  tiene distribución  $N(0; 1)$ .

(b) Sean  $X_1 \sim N(\mu_1, \sigma_1^2)$ , ... ,  $X_n \sim N(\mu_n, \sigma_n^2)$  variables aleatorias independientes. Entonces:

$$\sum_{i=1}^n a_i X_i + b \sim N\left(\sum_{i=1}^n a_i \mu_i + b; \sum_{i=1}^n a_i^2 \sigma_i^2\right) \quad \bullet$$

### Demostración.-

(a) Tenemos el siguiente cambio de variable:

$$z = \frac{x - \mu}{\sigma} \quad \Rightarrow \quad x = \sigma z + \mu \quad \Rightarrow \quad \frac{dx}{dz} = \sigma$$

Por tanto, la densidad de la nueva variable aleatoria  $Z$  será:

$$g(z) = f(x) \left| \frac{dx}{dz} \right| = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

que, evidentemente, corresponde a la  $N(0; 1)$ .

(b) Esta propiedad de la distribución Normal será muy sencilla de obtener a partir de la función característica, que es una herramienta muy útil en Probabilidad. •

**Definición.-** El modelo o distribución *Gamma* de parámetros  $\alpha > 0$  y  $\beta > 0$ ,  $Gamma(\alpha; \beta)$ , es un modelo para variables aleatorias continuas con función de densidad:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \quad \text{para } x > 0$$

Podemos comprobar que, efectivamente,  $f(x)$  es una función de densidad:

$$\begin{aligned}\int_{\mathbb{R}} f(x)dx &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/\beta} dx \\ &= \frac{1}{\Gamma(\alpha)} \int_0^\infty y^{\alpha-1} e^{-y} dy \quad \left( \text{definiendo } y = \frac{x}{\beta} \right) \\ &= \frac{1}{\Gamma(\alpha)} \Gamma(\alpha) = 1\end{aligned}$$

Podemos calcular también la esperanza y la varianza del modelo  $Gamma(\alpha; \beta)$ :

$$\begin{aligned}E[X] &= \int_{\mathbb{R}} xf(x)dx = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^\alpha e^{-x/\beta} dx \\ &= \frac{\beta}{\Gamma(\alpha)} \int_0^\infty y^\alpha e^{-y} dy \quad \left( \text{definiendo } y = \frac{x}{\beta} \right) \\ &= \frac{\beta}{\Gamma(\alpha)} \Gamma(\alpha + 1) = \frac{\beta}{\Gamma(\alpha)} \alpha \Gamma(\alpha) = \alpha\beta\end{aligned}$$

Mediante cálculos similares:

$$V(X) = E[X^2] - (E[X])^2 = \int_{\mathbb{R}} x^2 f(x)dx - (\alpha\beta)^2 = \dots = \alpha\beta^2 \quad \bullet$$

**Definición.-** El modelo o distribución *Exponencial* de parámetro  $\beta > 0$ ,  $Exp(\beta)$ , es un modelo para variables aleatorias continuas con función de densidad:

$$f(x) = \frac{1}{\beta} e^{-x/\beta} \quad \text{para } x > 0$$

Es un caso particular, pero especialmente importante, del modelo  $Gamma(\alpha; \beta)$ , donde  $\alpha = 1$ . Es muy utilizado para modelizar tiempos de vida.

Obviamente, para el modelo  $Exp(\beta)$ , tenemos:

$$E[X] = \beta \quad ; \quad V(X) = \beta^2 \quad \bullet$$

**Definición.-** El modelo o distribución *Beta* de parámetros  $\alpha$  y  $\beta > 0$ ,  $Beta(\alpha; \beta)$ , es un modelo para variables aleatorias continuas con función de densidad:

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad \text{para } 0 < x < 1$$

Evidentemente, se trata de una función de densidad, y es fácil calcular su esperanza y su varianza:

$$\begin{aligned}
E[X] &= \int_{\mathbb{R}} xf(x)dx = \frac{1}{B(\alpha, \beta)} \int_0^1 x^\alpha (1-x)^{\beta-1} dx \\
&= \frac{1}{B(\alpha, \beta)} B(\alpha+1, \beta) \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)} \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}$$

Mediante cálculos similares:

$$V(X) = E[X^2] - (E[X])^2 = \int_{\mathbb{R}} x^2 f(x) dx - \left( \frac{\alpha}{\alpha+\beta} \right)^2 = \dots = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad \bullet$$

## 4 Modelos de probabilidad sobre $\mathfrak{R}^n$

El principal inconveniente de los modelos de probabilidad sobre  $\mathfrak{R}^n$  es que la notación se complica inevitablemente. Por este motivo, nos vamos a conformar con definir el modelo de probabilidad Normal bivariante sobre  $\mathfrak{R}^2$ , y a comentar algunos de sus aspectos más interesantes. La distribución Normal bivariante se utiliza muy a menudo para modelizar pares de características morfológicas: peso y estatura de personas, altura y envergadura de aves, ...

**Definición.-** El modelo o distribución *Normal bivariante de parámetros*  $\mu_1 \in \mathfrak{R}$ ,  $\mu_2 \in \mathfrak{R}$ ,  $\sigma_1 > 0$ ,  $\sigma_2 > 0$  y  $\rho \in (-1, 1)$  es un modelo para vectores aleatorios continuos bivalentes con función de densidad para  $(x, y) \in \mathfrak{R}^2$ :

$$\begin{aligned}
f(x, y) &= \frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_1\sigma_2} \\
&\exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \left( \frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left( \frac{x-\mu_1}{\sigma_1} \right) \left( \frac{y-\mu_2}{\sigma_2} \right) + \left( \frac{y-\mu_2}{\sigma_2} \right)^2 \right] \right\}
\end{aligned}$$

Las propiedades más destacables de este modelo son las siguientes:

$$E[X] = \mu_1 \quad E[Y] = \mu_2 \quad V(X) = \sigma_1^2 \quad V(Y) = \sigma_2^2 \quad Cov(X, Y) = \rho\sigma_1\sigma_2 \quad \bullet$$

Adicionalmente, podemos observar que si  $\rho = 0$ , la función de densidad queda de la siguiente forma:

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2} \exp \left\{ -\frac{1}{2} \left[ \left( \frac{x - \mu_1}{\sigma_1} \right)^2 + \left( \frac{y - \mu_2}{\sigma_2} \right)^2 \right] \right\} \\ &= \frac{1}{\sigma_1\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{x - \mu_1}{\sigma_1} \right)^2 \right\} \frac{1}{\sigma_2\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{y - \mu_2}{\sigma_2} \right)^2 \right\} \end{aligned}$$

lo cual, obviamente, corresponde al caso en el que las variables aleatorias  $X$  e  $Y$  tienen distribuciones  $N(\mu_1, \sigma_1^2)$  y  $N(\mu_2, \sigma_2^2)$ , respectivamente, y son independientes.