Índice

4.2. Distribución de Probabilidad

Supongamos una población constituida por 50 millones de individuos. Como estudiamos en el capítulo anterior, la selección aleatoria de los individuos de esta población puede formalizarse, a Nivel I, mediante un espacio probabilístico $(\Omega, \mathcal{A}, P)$ en el que el espacio muestral esté constituido por los individuos de la población

$$\Omega = \{\omega_1 = \text{Abad Abad}, ..., \omega_{50,000,000} = \text{Zurdo Zamora}\}$$

y tal que sobre el conjunto $\mathcal{A}$ de los sucesos esté definida una probabilidad P, de forma que todos los sucesos elementales sean equiprobables: Modelo Uniforme.

Habitualmente estaremos interesados en alguna característica de la población más que en la población misma. Así, es habitual desear conocer el peso medio de la población o la estatura media, etcétera, interesándonos, por tanto, no los individuos ω_i, sino una función suya $X(\omega_i)$ como por ejemplo su peso.

Es decir, habitualmente estaremos interesados no en el espacio probabilístico, sino en una transformación suya (el Nivel II), tal que no sólo nos dé los valores de la característica en estudio para los individuos de la población,

$$X : \Omega \longrightarrow \mathbb{R}$$

sino que conserve la probabilidad P, aglutinando la nueva P_X las probabilidades de los sucesos elementales ω_i a los que corresponda el mismo valor mediante X,

$$P_X(A) = P\{\omega \in \Omega : X(\omega) \in A\} = P\{X^{-1}(A)\}.$$

Así por ejemplo, si en la población existiesen 20 millones de individuos con un peso entre 60 y 75 kilos, la transformación X debe ser tal que

$$P_X\{[60\,,\,75]\} = P\{\omega \in \Omega : 60 \leq X(\omega) \leq 75\} = \frac{2}{5}.$$

La función X recibe el nombre de *variable aleatoria* y P_X el de su *distribución de probabilidad*.

Evidentemente, sobre un espacio probabilístico es posible definir muchas variables aleatorias. Cuando se consideran a la vez varias de ellas, $X_1, ..., X_p$, de forma que en los individuos de la población se observan varios caracteres, queda constituido lo que se denomina una *variable aleatoria multidimensional*, o *vector aleatorio* $X = (X_1, ..., X_p)$.

Nada impide que los sucesos elementales del espacio muestral Ω sean números reales, por lo que, en ese caso, la aplicación identidad es la variable aleatoria natural a considerar.

En otras ocasiones, aunque los sucesos elementales no sean numéricos, la variable aleatoria a estudiar resulta obligada. Tanto es así que en ocasiones se identifican a los sucesos elementales con los valores de ésta.

Ejemplo 4.1

Consideremos el experimento aleatorio del lanzamiento de un dado. El espacio muestral es

$$\Omega = \{ \boxed{\cdot}, \boxed{\cdot\,\cdot}, \boxed{\cdot\,\cdot}, \boxed{\cdot\,\cdot}, \boxed{\cdot\,\cdot}, \boxed{\cdot\,\cdot} \}$$

y la probabilidad igual a 1/6 sobre cada uno de ellos.
La variable aleatoria a considerar de forma natural es X = *número de puntos de la cara superior del dado*. Es tan evidente la consideración de tal variable que en el Ejemplo 3.2 incluso reemplazamos los sucesos elementales por los valores de dicha variable aleatoria.
La distribución de probabilidad de X es, para $x = 1, ..., 6$,

$$P_X(\{x\}) = P\{\omega : X(\omega) = x\} = \frac{1}{6}$$

Asociada a toda variable aleatoria existe una función $F(x)$, denominada *función de distribución* de X, la cual va midiendo la probabilidad *acumulada* por X hasta el punto x. Es decir

$$F(x) = P\{\omega \in \Omega : X(\omega) \leq x\}.$$

Esta función tiene la propiedad de caracterizar la distribución de probabilidad de X, P_X. Es decir, a partir de una de ellas se obtiene la otra, siendo habitualmente más cómodo trabajar con la función de distribución.

Ejemplo 4.1 (continuación)

La función de distribución de X será una función en escalera que salta 1/6 en los valores de la variable,

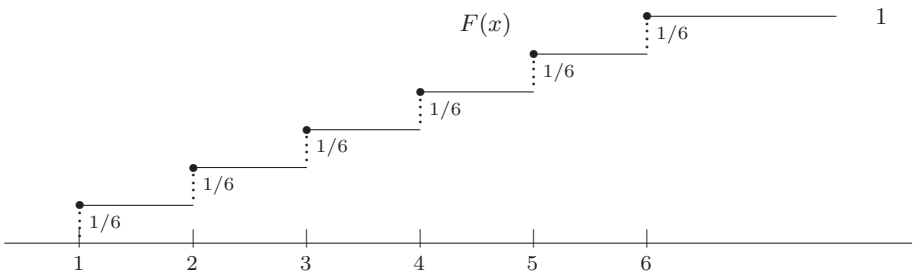

Figura 4.1

Si una variable aleatoria toma valores aislados, como ocurría en el ejemplo anterior, se denomina *discreta*. Si por el contrario puede tomar cualquier valor

de un intervalo, como por ejemplo ocurre con el peso, o la talla, la variable aleatoria recibe el nombre de *continua*. Estos calificativos se aplican también a su distribución, hablando de *distribuciones discretas* o *continuas*.

De la misma definición se deduce que la función de distribución de una variable aleatoria discreta es una función en escalera como la de la Figura 4.1, mientras que la correspondiente a una variable continua es una función continua como la de la Figura 4.2.

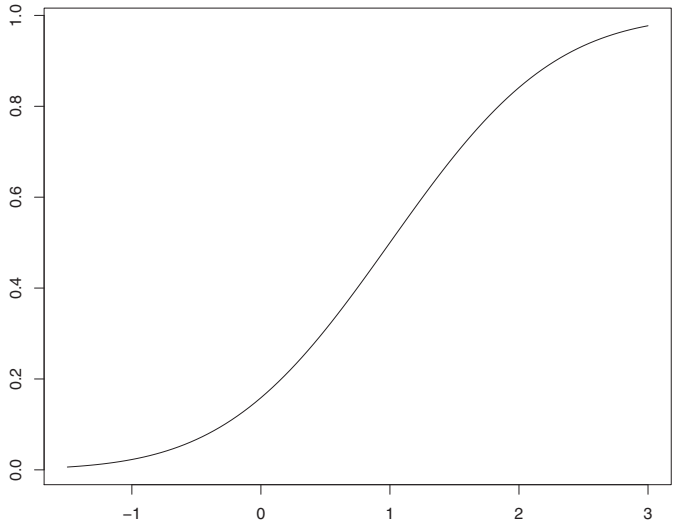

Figura 4.2

A partir de las propiedades de las probabilidades se puede deducir que las funciones de distribución son

1. No decrecientes.
2. Continuas por la derecha.
3. $\lim_{x\to-\infty} F(x) = 0$ y $\lim_{x\to\infty} F(x) = 1$.

Las variables aleatorias discretas X, las cuales hemos visto tienen una función de distribución en escalera, tienen asociadas una función, denominada *función de masa*, $p_X(x)$, la cual da la probabilidad de los valores de dicha variable aleatoria; es decir,

$$p_X(x) = P_X(\{x\}) = P\{\omega : X(\omega) = x\}.$$

Por la definición de función de distribución, las funciones de masa y de distribución de una variable aleatoria discreta están relacionadas por las expresiones:

$$p_X(x) = F(x) - F(x-)$$

en donde $F(x-)$ es el límite por la izquierda de F en x. Se ve, por tanto, que la función de masa recoge el valor del salto de la función de distribución, e inversamente,

$$F(x) = \sum_{y \leq x} p_X(y).$$

De manera análoga, las variables aleatorias continuas X tienen asociada una función, denominada *función de densidad*, $f_X(x)$, la cual indica la *velocidad* a la crece su función de distribución, siendo

$$f_X(x) = \frac{d}{dx} F(x)$$

e inversamente,

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

lo que implica, por la Propiedad 3 de las funciones de distribución, que sea $\int_{-\infty}^{\infty} f_X(y)\, dy = F(\infty) = 1$.
(Las denominadas *integrales impropias*, como estas en las que aparecen los símbolos $-\infty$ o $+\infty$, pueden interpretarse como $\int_{-\infty}^{x} f_X(y)\, dy = \lim_{b \to -\infty} \int_{b}^{x} f_X(y)\, dy$. Se han incluido en el texto para completar el concepto que se estudia pero no tendrá que calcular el lector ninguna de ellas.)

Así pues, la distribución de una variable aleatoria se puede caracterizar por su distribución de probabilidad, por su función de distribución, o por su función de masa o densidad (esta última según sea discreta o continua):

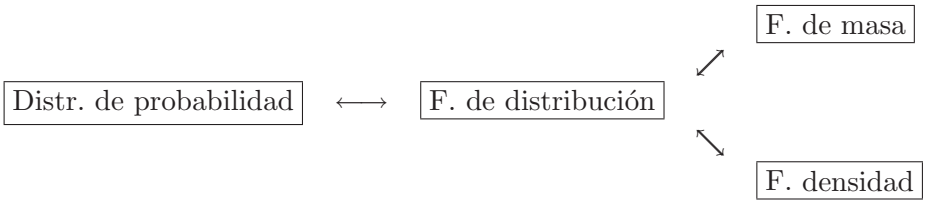

Figura 4.3

Estas funciones de masa y de densidad tienen una interpretación clara a partir, respectivamente, del diagrama de barras y del histograma: son *modelos teóricos* de donde proceden los datos que el investigador examina. De

ahí que se hable de que la variable en estudio tiene una determinada *Distribución de Probabilidad*, o mejor aún, un determinado *Modelo Probabilístico*. Éste habrá que suponerlo con objeto de hacer inferencias sobre X y, como veremos más adelante, si nuestros datos presentan —por ejemplo en el caso continuo— un histograma tal que cuando las bases de los rectángulos que lo forman tienden a cero a medida que la frecuencia total aumenta, la curva resultante se ajusta bien al modelo supuesto, las inferencias que hagamos serán aceptables. En caso contrario deberemos cambiar el modelo.

En las Secciones 4.4 y 4.5 estudiaremos algunos de los modelos más importantes.

Características de una distribución de probabilidad

Dado que los modelos de probabilidad representan, básicamente, un *ideal* de las distribuciones de frecuencias estudiadas en el Capítulo 2, tendrán, al igual que éstas, una medidas de posición, de dispersión, etc. Aquí sólo nos centraremos en una de posición y dos de dispersión, definiéndolas primero para el caso discreto y luego para el continuo.

Dada una variable aleatoria discreta X, con función de masa p_X, llamaremos *media* o *esperanza* de X a la suma de los valores que toma por las probabilidades con que los toma

$$\mu_X = E[X] = \sum_x x\, p_X(x)$$

y *varianza* de X a

$$\sigma_X^2 = V(X) = \sum_x (x - \mu_X)^2\ p_X(x).$$

Dada una variable aleatoria continua X, con función de densidad f_X, llamaremos *media* o *esperanza* de X a la integral

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

y *varianza* de X a

$$\sigma_X^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2\ f_X(x)\, dx.$$

En ambos casos, llamaremos *desviación típica* de X a la raíz cuadrada de la varianza:

$$\sigma_X = D(X) = \sqrt{\sigma_X^2}$$

teniendo estas medidas las misma interpretación que tenían en el Capítulo segundo.

Ejemplo 4.1 (continuación)

Esta distribución de probabilidad es de tipo discreto puesto que toma valores aislados. Como dijimos más arriba, su media será igual a los valores que toma por las probabilidades con que los toma,

$$E[X] = \sum_{x=1}^{6} x\, p_X(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3'5.$$

La varianza será igual a los valores que toma la variable, menos la media acabada de calcular al cuadrado, por las probabilidades con que los toma,

$$V(X) = \sum_{x} (x - \mu_X)^2\, p_X(x) = (1-3'5)^2 \cdot \frac{1}{6} + (2-3'5)^2 \cdot \frac{1}{6} + (3-3'5)^2 \cdot \frac{1}{6} + (4-3'5)^2 \cdot \frac{1}{6} +$$

$$+(5-3'5)^2 \cdot \frac{1}{6} + (6-3'5)^2 \cdot \frac{1}{6} = 6'25 \cdot \frac{1}{6} + 2'25 \cdot \frac{1}{6} + 0'25 \cdot \frac{1}{6} + 0'25 \cdot \frac{1}{6} + 2'25 \cdot \frac{1}{6} + 6'25 \cdot \frac{1}{6} = \frac{17'5}{6} = 2'92.$$

Una forma alternativa de calcular la varianza es utilizando la expresión

$$V(X) = E[X^2] - (E[X])^2$$

la cual es válida no sólo para distribuciones discretas sino también para continuas.
La *media de los cuadrados* será, en el caso de distribuciones discretas

$$E[X^2] = \sum_{x} x^2\, p_X(x)$$

es decir, valores que toma la variable, al cuadrado, por las probabilidades con que los toma, y en el caso de distribuciones de tipo continuo con función de densidad f_X, la integral

$$E[X^2] = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx.$$

En el ejemplo que nos ocupa sería

$$E[X^2] = \sum_{x} x^2\, p_X(x) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} = \frac{91}{6} = 15'17$$

con lo que la varianza sería, utilizando la fórmula anterior,

$$V(X) = 15'17 - 3'5^2 = 2'92.$$

La desviación típica sería la raíz cuadrada de la varianza, es decir,

$$\sigma_X = D(X) = \sqrt{2'92} = 1'71.$$

No debe perderse de vista que estas características de la distribución de X no son más que *idealizaciones* de las medidas de posición y dispersión de la distribución de frecuencias de las observaciones (ver Capítulo 2), cuyo histograma/diagrama de barras, de frecuencias relativas sugiere cuál debe ser la distribución de probabilidad de la variable en estudio.