Teoría de Probabilidad. 1º de Carrera. Tema 2: Variables Aleatorias

Índice

1 Introducción

El objetivo de la propiedad, como hemos visto en el apartado anterior, busca estudiar los fenómenos aleatorios que resultan de un experimento. Sin embargo, muchas veces no nos interesan los restultados en si, sino una cantidad que nos permita extraer información de estos valores. Pongamos que estamos jugando a un juego de azar. No nos interesan tanto los resultados del juego como el beneficio o la pérdida que se pueda tener.

$$\Omega = \{\text{Resultado del juego}\}$$

Nos interesa:

$$\Omega(resultado) \rightarrow \mathbb{R}(ganador)$$

Por ejemplo, pongamos que jugamos con un dado trucado. Lo lanzamos muchas veces, pero el que anota el resultado solo comprueba si el resultado es par o impar. Los resultados posibles de este experimento serán $\Omega = \{1, 2, 3, 4, 5, 6\}$. Digamos que el suceso $A =$ "par".

- Obtenemos como resultado:
$$\mathbb{P}(A) \approx \frac{7}{10}$$
- Información: $\sigma - algebra$
$$\mathcal{A} = \{\emptyset, \Omega, A, A^c\}$$
- Observación: no podemos calcular $\mathbb{P}(1)$
- P queda determinada por $\mathbb{P}(A) \approx \frac{7}{10}$
- Apuesta: Se gana 10 euros si sale un **5 o 6**. Se pierden 2 euros si sale cualquier otro número.
- Aplicación:
$$X : \Omega \rightarrow \mathbb{R}$$
$$X(5) = X(6) = 10 \text{ €}; X(1) = X(2) = X(3) = X(4) = -2 \text{ €}$$

En nuestro modelo $(\Omega, (A), \mathbb{P})$, **NO** podemos calcular la probabilidad de ganar 10€:

$$\mathbb{P}(\text{ganar 10 €}) = \mathbb{P}(\{w \in \Omega, X(\omega) = 10\}) = \mathbb{P}(\{5, 6\}) \ \rightarrow \text{No pertenece a } \mathcal{A}$$

DEFINICIÓN: dado $(\Omega, (A), \mathbb{P})$, diremos que una aplicación

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

es una **variable aleatoria** si, para todo intervalo $B \in \mathbb{R}$ (o semirecta) tenemos que:

$$\{\omega \in \Omega, X(\omega) \in B\} \in \mathcal{A}$$

En el caso mencionado anteriormente, X no es una variable aleatoria, ya que el resultado sería $\{5, 6\}$, que no pertenece a $\mathcal{A}$
Observaciones:

1. En la definición de variable aleatoria no interviene $\mathbb{P}$, pese a que la motivación es calcular:
$$\mathbb{P}(\underbrace{\{\omega \in \Omega, X(\omega) \in B\}}_{X^{-1}(B)}) = \mathbb{P}(X \in B)$$
2. Notación: Si $B = \{a\}$, entonces:
$$\mathbb{P}(X \in B) = \mathbb{P}(X \in \{a\}) = \mathbb{P}(X = a)$$
o también:
$$a \leq b : B = [a, b] : \mathbb{P}(X \in [a, b]) = \mathbb{P}(a \leq X \leq b)$$
3. Si $\mathcal{A} = P(\Omega)$, entonces toda aplicación $X : \Omega \rightarrow \mathbb{R}$ es variable aleatoria.

2 Como hallar variables aleatorias

2.1 Operaciones

Sean X, Y variables aleatorias y $a \in \mathbb{R}$ entonces son variables aleatorias:

- $X + Y$
- $X \cdot Y$
- $a \cdot X$
- $absX$
- X^a
- Si $X(\omega) \neq 0; \forall \omega \in \Omega$, entonces $\dfrac{1}{X}$ también es una variable aleatoria.

2.2 Sucesiones

Tomemos $\{X_n, n \geq 1\}$ como una sucesión de variables aleatorias tales que $\forall \omega \in \Omega$, la sucesión

$$\{X_n(\omega) \in \mathbb{R}, n \geq 1\}$$

es convergente. En ese caso, definiremos

$$X(\omega) \coloneqq \lim_{n \to \infty} X_n(\omega) \in \mathbb{R}$$

Entonces, X es una variable aleatoria.

3 Función de distribución de una variable aleatoria

DEFINICIÓN: Sea X una variable aleatoria. La **función de distribución** de X es:

$$F : \mathbb{R} \to [0, 1]$$

$$x \to F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\{\omega \in \Omega, X(\omega) \leq x\} = \mathbb{P}(\underbrace{\{\omega \in \Omega, X(\omega) \in B\}}_{\in \mathcal{A}}) \text{ donde } B = (-\infty, x]$$

3.1 Propiedades

1. F es no decreciente: $x < y \Rightarrow F(x) \leq F(y)$
2. F es continua por la derecha y tiene límite por la izquierda en todos sus puntos (càdlàg *continu à droit, limit à gauche*).
3. $$\lim_{x \to -\infty} F(x) = 0 \; ; \; \lim_{x \to +\infty} F(x) = 1$$
4. F tiene como máximo un numero numerable de puntos de discontinuidad. Esto se debe a que la recta real se puede expresar como la unión de $\mathbb{Z}$ intervalos.
5. $$\mathbb{P}(s \leq X \leq t) = F(t) - F(s)$$
6. $$\mathbb{P}(X < x) = F(x^-) = \lim_{X \to x^-} F(X)$$
7. $$\mathbb{P}(X = x) = F(x) - F(x^-)$$

   si F es discontinua en x:

   $$\mathbb{P}(X = x) > 0$$

4 Variables aleatorias discretas

Una variable aleatoria X es **discreta** si:

$$\exists S \subset R \text{ finito o numerable tal que } \mathbb{P}(X \in S) = 1$$

Siendo S el soporte de X. Suponemos que todo $x \in S$ cumple que $\mathbb{P}(X = a) > 0$. También:

$$S = \{x_i, i \in I\}; I \subseteq \mathbb{N} = \{1, 2, 3, ...\}$$

4.1 Función de probabilidad

$$p : S \to [0, 1]$$
$$x_i \to p(x_i) = \mathbb{P}(X = x_i)$$

En la notación de esta función normalmente se escribirá p_i para referirse a $p(x_i)$.
Observaciones:

A
$$\sum_{i \in I} p_i = \sum_{i \in I} \mathbb{P}(X = x_i) = \mathbb{P}(X \in S) = 1$$

B $\forall B \subset \mathbb{R}$ intervalo/semirecta
$$\mathbb{P}(X \in B) = \sum_{i \in I, x \in B} p_i$$

C La función de probabilidad de X permite calcular todas las probabilidades relacionadas con la variable aleatoria X y determina la **función de distribución**:

$$F(x) = \mathbb{P}(X \leq x) = \sum_{i \in I, x_i \leq x} pi$$

4.2 Tipos de variable discreta

4.2.1 Caso Degenerado

$$X = a \Rightarrow S = \{a\} \text{ y } \mathbb{P}(X = a) = 1$$

4.2.2 Caso Bermoulli

$$X = \{1, \text{ con probabilidad } p \in (0, 1) \text{ y } 0, \text{ con probabilidad } 1 - p \in (0, 1)\}$$
$$S = \{a\}; \mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$$

X tiene distribución Bermouilli de parámetro p:

$$X \sim B(p)$$

4.2.3 Uniforme discreta sobre n puntos

Toman valores en $\{x_1, x_2, x_3, ..., x_n\}$ con la misma probabilidad. Tenemos que:

$$\mathbb{P}(X = x_i) = \frac{1}{n}$$

Denotamos:
$$X \sim U\{x_1, ..., x_n\};\ S = \{x_1, ..., x_n\};\ p(x_i) = \frac{1}{n}, \forall i = 1, ..., n$$

4.2.4 Binomial

- Repetimos n veces el mismo experimento
- Tenemos un suceso A con $p = \mathbb{P}(A)$. Si A se realiza es un éxito, sino es un fracaso.
- $$X = \text{Numero de veces que ocurre } A \text{ en las } n \text{ repeticiones; } S = \{0, 1, 2, 3, 4, ..., n\}$$
$$\mathbb{P}(X = k) = \binom{n}{k} p^k \cdot (1-p)^{n-k}, k = 0, 1, ..., n$$
- Denotamos como:
$$X \sim B(n, p)$$

Observación: Pongamos que tenemos una serie de variables aleatorias $X_1, X_2, ..., X_n$ donde

$$X_i = \begin{cases} 1, \text{Si en la i-ésima repetición se realiza } A \\ 0, \text{En caso contrario} \end{cases} \tag{1}$$

Entonces:

$$X_i = B(p)$$

Tenemos que:

$$X = \sum_{i=1}^{n} X_i \sim B(n, p) \text{ (repeticiones independientes)}$$

4.2.5 Geométrica

Pongamos que tenemos un experimento aleatorio con una probabilidad:

$$p = \mathbb{P}(\text{éxito})$$

$$X = \text{"número de pruebas hasta obtener el primer éxito (incluida la prueba con el éxito)"}$$

$$S = \{1, 2, 3, ...\}$$

Función de probabilidad:

$$P(X = k) = (1-p)^{k-1} \cdot p$$

Notación:

$$X_n \sim Geom(p)$$

Propiedades:

$$\mathbb{P}(X > l + k | X > l) = \mathbb{P}(X > k), \ \forall k, l \in \mathbb{N}$$

Esta propiedad es conocida como falta de memoria. Para entenderlo puede pensarse de la siguiente forma. Supongamos que estás lanzando una moneda desde las 9 de la mañana hasta las 11, y no sale ninguna cara. Si alguien llegara a ĺas 11 y los dos empezarais a lanzar monedas en ese instante, quien de los dos tendría ¿más posibilidades de sacar una cara antes? Los dos tendrían la misma.

4.2.6 Binomial negativa

$$X = \text{número de pruebas hasta que obtenga } m \text{ éxitos}$$

Observación: Si $m = 1 \Rightarrow X_n \ Geom(p)$
En general:

$$S = \{m, m+1, m+2, ...\} \text{ con } \mathbb{P}(X = k) = \binom{k-1}{m-1} (1-p)^{k-m} \cdot p^m$$

Notación:

$$X_n \sim BN(m, p)$$

4.2.7 Hipergeométrica

Pongamos que tenemos una urna con N bolas. B bolas blancas y $N - B$ rojas. De esta urna sacamos $n \leq N$ bolas.

$$X = \text{"número de bolas blancas obtenidas"}$$

Si estamos trabajando con reemplazamiento:

$$X \sim B(n, \frac{B}{N})$$

Si trabajamos sin reemplazamiento:

$$\mathbb{P}(X = j) = \frac{\frac{n!}{(n-j)!}}{\binom{N}{n}}$$

Es más, podemos deducir que:

$$max(n - N + B, 0) \leq j \leq min(B, n)$$

Notación:

$$H \sim H(N, B, n)$$

Pongamos un ejemplo. En una mesa de 10 personas hay 4 menores de edad. Todas las bebidas que se beben son alcoholicas y el camarero pide 5 DNI al azar.

- ¿$\mathbb{P}$(Descubrir los 4 menores de edad)?:

$$N = 10 \ ; \ B = 4 \ ; \ N - B = 6 \ ; \ n = 5$$

$$X = \text{"Número de menores de edad"} \Rightarrow X \sim H(10, 4, 5)$$

$$\mathbb{P}(X = 4) = \frac{\frac{5!}{(5-4)!}}{\binom{10}{4}} = 0.024$$

- ¿Cuantos DNI se deben pedir para que la probabilidad de detectar a los 4 menores sea minimo 0.25?

$$n = 7 \Rightarrow \mathbb{P}(X = 4) = \frac{\frac{7!}{(7-4)!}}{\binom{10}{7}} = 0.16667$$

$$n = 8 \Rightarrow \mathbb{P}(X = 4) = \frac{\frac{8!}{(8-4)!}}{\binom{10}{8}} = 0.3333 > 0.25$$

4.2.8 Poisson

$\lambda \in \mathbb{R}$. Se dice que X tiene ley de **Poisson de parámetros** λ si se cumple que $\mathbb{N} \cup \{u\}$ y

$$\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \ k = 0, 1, 2...$$

Se escribe:

$$X \sim Poiss(\lambda)$$

Esta variable aleatoria aparece cuando se busca calcular el número de veces que ocurre un suceso aleatorio en un parámetro de tiempo fijo. Más adelante veremos que:

$$\lambda = \text{Valor medio de los sucesos que contamos } (\textit{por unidad de tiempo/espacio})$$

Aproximación de la Binomial por la Poisson: Pongamos que tenemos una sucesión:

$$\{p_n, n \geq 1\} \text{ tal que } 0 < p_n < 1 \ y \ \lim_{n \to \infty} np_n = \lambda > 0 \Rightarrow \lim_{n \to \infty} p_n = 0$$

También se deduce que:

$$\lambda > 0 \Rightarrow p_n = \frac{\lambda}{n}$$

$$X_n \sim B(n, p_n)$$

$$\mathbb{P}(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k},\ k = 0, 1, ..., n$$

Si $k = 0$:

$$\lim_{n\to\infty} \mathbb{P}(X_n = 0) = \lim_{n\to\infty} (1-p)^n = \lim_{n\to\infty} (1 + \frac{1}{-\frac{1}{p_n}})^{n(-p_n)\cdot(\frac{1}{-p_n})}$$

$$= \lim_{n\to\infty} [(1 + \frac{1}{-\frac{1}{p_n}})^{-\frac{1}{p_n}}]^{-np_n} = e^{\lambda}$$

Si $k \geq 1$ para n suficientemente grande:

$$\mathbb{P}(X_n = k) = \frac{n!}{(n-k)!k!} p_n^k (1-p_n)^{n-k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{n^k} \frac{(np_n)^k}{k!} (1-p_n)^n (1-p_n)^{-k}$$

$$\xrightarrow{n\to\infty} prob(X = k) = 1 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1 = \!\Rightarrow X \sim Poiss(\lambda)$$

Ejemplo: Tenemos un lote de 200 unidades.

$$X = \text{"Unidades defectuosas"}$$

$$X \sim B(200, p)\ \{p_1 = 0.001;\ p_2 = 0.05\}$$

$$\mathbb{P}(\text{Aceptar el lote}) = \mathbb{P}(X \leq 2)$$

Si $p = 0.001$:

$$\mathbb{P}(\text{Aceptar el lote}) = \mathbb{P}(X \leq 2) \sim \mathbb{P}(Y \leq 2)$$

Siendo $Y \sim Poiss(0.001 \times 200)$:

$$= \mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 2) =$$

5 Variables aleatorias continuas

Una variable aleatoria X es **continua** si $\exists f : \mathbb{R} \to \mathbb{R}$:

- La función es no negativa.

$$f(x) \geq 0, \forall x \in \mathbb{R}$$

- f es integrable en $\mathbb{R}$ y

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

- La probabilidad de que la variable tome valores entre a y b es la integral entre ambos valores.

$$\forall -\infty \leq a \leq b \leq +\infty\ \mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$$

- El valor f se denomina **densidad de** X.
- La probabilidad de que la variable tome un valor exacto es 0

$$\mathbb{P}(X = a) = \mathbb{P}(a \leq X \leq a) = \int_a^a f(x)dx = 0,\ \forall a \in \mathbb{R}$$

5.1 Función de distribución

Para calcular la función de distribución de una variable aleatoria continua tomamos la siguiente función:

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(-\infty < X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Esto implica que no solo $F(x)$ es continua sino que además es absolutamente continua.

5.2 Tipos de variable continua

5.2.1 Uniforme en (0,1)

NOTACIÓN:

$$X \sim U(0,1)$$

Esta variable aleatoria representa el resultado de escoger un número al azar del intervalo $(0,1)$ (uniformemente)

Densidad:

$$f(x) = \begin{cases} 1, \ x \in (0,1) \\ 0, \ \text{en caso contrario} \end{cases}$$

Para calcular la probabilidad entre un intervalo dado a, b empleariamos la integral:

$$\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f(x)dx = \int_{a}^{b} 1dx = b - a$$

Función de distribución:

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, \ si \ x < 0 \\ \int_0^x 1dt = x, \ si \ 0 \leq x \leq 1 \\ 1, \ si \ x \geq 1 \end{cases} = \begin{cases} 0, \ si \ x < 0 \\ x, \ si \ 0 \leq x < 1 \\ 1, \ si \ x \geq 1 \end{cases}$$