MDS7101

Estadística: Teoría y Aplicaciones

ESCRIBAS
Naomí Cautivo B.
Máximo Flores Valenzuela

ÍNDICE

1. SEMANA 1: REPASO DE PROBABILIDADES.

- **¿Qué es una probabilidad?** Una probabilidad es una medida de incertidumbre.
- Tiene dos enfoques: frecuentista y bayesiano. Para el frecuentista, la probabilidad es algo inherente a la naturaleza, y su paradigma de cálculo es $\text{casos favorables}/\text{casos totales}$. Para el bayesiano, la probabilidad es un invento del ser humano, y ya no se usa la fórmula anterior.

1.1. NOTACIONES BÁSICAS

En el curso, usaremos Ω para denotar el espacio muestral, ω para los eventos, y $\mathbb{P}$ para la medida de probabilidad, que corresponde a una función que asigna una probabilidad a cualquier evento en $\mathcal{F}$, donde $\mathcal{F}$ es una colección de subconjuntos de Ω, no necesariamente una partición.

1.2. PROPIEDADES BÁSICAS DE $\mathbb{P}$

① La probabilidad del espacio muestral debe ser siempre 1, es decir, $\mathbb{P}(\Omega) = 1$.

② La probabilidad es no negativa, es decir, para cualquier evento $A \in \mathcal{F}$, $\mathbb{P}(A) \geq 0$.

③ La probabilidad de la unión de eventos disjuntos es la suma de sus probabilidades por separado, es decir, $\mathbb{P}\left(\bigcup_i A_i\right) = \sum_i \mathbb{P}(A_i)$ cuando $\forall i \neq j, A_i \cap A_j = \emptyset$.

1.3. VARIABLES ALEATORIAS

> **Nota**
>
> Por convención, en este curso usaremos letras mayúsculas para denotar las variables aleatorias (en adelante, abreviadas como v. a.).

Son funciones que toman elementos del espacio muestral, y les asigna a cada uno un número real. Podemos definir una v. a. X como $X : \Omega \to \mathbb{R}$. Por ejemplo, sea X el número de caras en el lanzamiento de una moneda no cargada 3 veces, entonces $X = \{0, 1, 2, 3\}$, porque son las distintas cantidades de caras que pueden salir.

1.3.1. VARIABLES ALEATORIAS DISCRETAS

Se dice que X es una v. a. discreta si toma valores de un conjunto finito, o infinito numerable, y además $\forall x$, $\mathbb{P}(X = x) \neq 0$.

1.3.2. VARIABLES ALEATORIAS CONTINUAS

Se dice que X es una v. a. continua si X toma cualquier valor real con probabilidad cero, es decir, $\forall x, \mathbb{P}(X = x) = 0$.

1.3.3. FUNCIONES DE DENSIDAD

Existen dos funciones de densidad que permiten ver el comportamiento de una variable aleatoria.

- PDF: *Probability Density Function* ($f(x)$). Describe cómo se distribuye la probabilidad a lo largo de los posibles valores de la v. a. En específico, $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x) \,\mathrm{d}x$.
- CDF: *Cummulative Density Function* ($F(x)$). Acumula la probabilidad desde $-\infty$ hasta un valor x en el dominio. En específico, $F(x) = \mathbb{P}(X \leq x)$.

Estas funciones están directamente relacionadas mediante la fórmula $F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$, lo que puede ser observado gráficamente en la Figura 1.

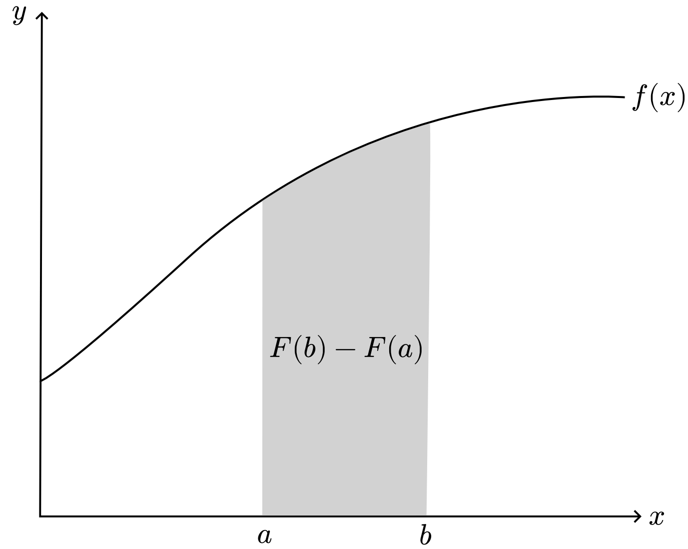

Figura 1: Funciones «PDF» ($f(x)$) y «CDF» ($F(x)$).

Si se conoce F, podemos conocer la probabilidad de un intervalo mediante la siguiente fórmula $\mathbb{P}(a \leq X \leq b) = F(b) - F(a)$.

1.3.4. ESPERANZA DE UNA VARIABLE ALEATORIA

Definimos la esperanza de una variable aleatoria para las v. a. discretas y continuas como:

- X discreta: $\mathbb{E}[X] = \sum_{\Omega} x \cdot \mathbb{P}(X = x)$.
- X continua: $\mathbb{E}[X] = \int_{\mathbb{R}_X} x \cdot f(x)\,\mathrm{d}x$.

También se puede definir como el primer momento de distribución. Los momentos de distribución se definen como $\mathbb{E}[X], \mathbb{E}[X^2], \mathbb{E}[X^3]$, etc.

1.3.5. VARIANZA DE UNA VARIABLE ALEATORIA

Definimos la varianza de una v. a. discreta y continua como:

- X discreta: $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$.
- X continua: $\mathbb{V}\mathrm{ar}(X) = \int_{\mathbb{R}_X} (X - \mathbb{E}(X))^2 \cdot f(x)\,\mathrm{d}x$.

Con esto mismo podemos definir la desviación estándar de una variable aleatoria, la cual viene a ser la raíz cuadrada de su varianza. Se le conoce también como σ o $\mathtt{STD}(X)$.

1.3.6. ESTANDARIZACIÓN DE UNA VARIABLE ALEATORIA

Sea X una variable aleatoria, se define la variable $Z = (X - \mu)/\sigma$ con $\mu = \mathbb{E}[X]$ y $\sigma = \sqrt{\mathbb{V}\mathrm{ar}(X)}$. Se dice que Z es la estandarización de X, pues cumple $\mathbb{E}[Z] = 0$ y $\mathbb{V}\mathrm{ar}(Z) = 1$.

> ⚠ **Advertencia**
>
> En algunas librerías de programación, la «estandarización» de una v. a. se considera como su «normalización», pero estos términos no son equivalentes.

1.4. DISTRIBUCIONES DISCRETAS

En el curso, veremos principalmente las siguientes distribuciones discretas:

① Bernoulli: $X :=$ lanzamiento de una moneda sólo una vez. Entonces $X \sim \mathrm{Bernoulli}(p)$. Sus valores se definen como:

$$X = \begin{cases} 1 \text{ en el caso de éxito} \\ 0 \text{ en el caso de fracaso} \end{cases}$$

Además, $\mathbb{P}(X = 1) = p$ (probabilidad de éxito) y $\mathbb{P}(X = 0) = 1 - p$ (probabilidad de fracaso). El éxito puede ser, por ejemplo, «obtener cara al lanzar la moneda».

② Binomial: si realizamos el experimento anterior n veces, entonces $X :=$ número de éxitos en n ensayos independientes. Luego, $X \sim \mathrm{Binomial}(p, n)$. La probabilidad asociada a k éxitos es la siguiente:

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

Además, $\mathbb{E}(X) = np$ y $\mathbb{V}\mathrm{ar}(X) = np \cdot (1 - p)$.

Si p es un vector multivariado $(p_1, p_2, ..., p_n)$, se transforma en una distribución multinomial, denominada $X \sim \mathrm{Multinomial}(p, n)$.

1.4.1. DISTRIBUCIONES CONTINUAS

① Normal: $X \sim \mathcal{N}(\mu, \sigma^2)$. Su función de densidad es:

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

- Normal estándar: si $X \sim \mathcal{N}(\mu, \sigma^2)$ y $Z = (X - \mu)/\sigma$, entonces $Z \sim \mathcal{N}(0, 1)$.

② «Chi cuadrado» (χ^2): si $Z \sim \mathcal{N}(0, 1)$ entonces:

$$Y = Z^2 \to Y \sim \chi^2_{[1]}$$

donde el subíndice $[1]$ denota los grados de libertad, que es algo que se tratará en las próximas secciones.

③ t-Student: si $Z \sim \mathcal{N}(0, 1)$ e $Y \sim \chi^2_{[n]}$. Entonces definimos t-Student como:

$$t = \frac{Z}{\sqrt{Y/n}} \sim t_{[n]}$$

④ Fischer (F): combinamos dos χ^2 independientes:

$$X_1 \sim \chi^2_{[n_1]} \wedge X_2 \sim \chi^2_{[n_2]} \text{ entonces } F = \frac{X_1/n_1}{X_2/n_2} \sim F_{n_1,n_2}$$

1.4.2. COVARIANZA DE DOS VARIABLES ALEATORIAS

Medida de cómo en promedio varían linealmente dos variables aleatorias entre sí.

$$\begin{aligned}\mathbb{C}\mathrm{ov}(X,Y) &= \mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))] \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\end{aligned}$$

Si estas variables X, Y son independientes, entonces su covarianza será cero, pues $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$ por la propiedad heredada de la esperanza.

> ⚠ **Advertencia**
>
> La implicancia $\mathbb{C}\mathrm{ov}(X,Y) = 0 \Rightarrow X, Y \text{ son independientes}$ es falsa, y es un error muy común asumir que es cierta.

1.4.3. CORRELACIÓN DE DOS VARIABLES ALEATORIAS.

$$\mathrm{Corr}(X,Y) = \frac{\mathbb{C}\mathrm{ov}(X,Y)}{\sqrt{\mathbb{V}\mathrm{ar}(X) \cdot \mathbb{V}\mathrm{ar}(Y)}} = \rho(X,Y)$$

Ojo: Correlación en cero no implica que serán variables aleatorias independientes, un caso clave para esto es de que estas estén relacionadas de forma no lineal (por verse en clases).