Cálculo de Probabilidades

Espacios de probabilidad

DFN 1 (Espacio de probabilidad). *Un espacio muestral es la terna* $(\Omega, \mathcal{F}, \mathbb{P})$ *donde*

- Ω *es el espacio muestral. Un conjunto no-vacío*
- $\mathcal{F}$ *es la* σ*-álgebra, una colección de subconjuntos de* Ω *que contiene a* $\Omega, \varnothing$ *y es cerrada bajo complementos, e intersecciones* *contables.*
- *Una medida* $\mathbb{P} : \mathcal{F} \to [0,1]$ *con* $\mathbb{P}(\varnothing) = 0$ *y* $\mathbb{P}(\Omega) = 1$*, tal que* $\mathbb{P}$ *es* σ*-aditiva.*

Teorema 1 (Propiedades de $\mathbb{P}$): Si $\mathbb{P}$ es una medida de probabilidad válida se cumple:

1. $\mathbb{P}(A) \geq 0 \quad \forall A \in \mathcal{F}$.
2. $\mathbb{P}(\Omega) = 1$.
3. Si $A_1, A_2, \dots \in \mathcal{F}$ con $A_i \cap A_j = \varnothing \ \forall i \neq j$, entonces $\mathbb{P}(\bigcup_i A_i) = \sum_i \mathbb{P}(A_i)$.

Las propiedades 1 a 3 se conocen como los axiomas de Kolmogorov.

4. Para cualquier $A \in \mathcal{F}$, $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$.
5. Si $A \subseteq B$ con $A, B \in \mathcal{F} \implies \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.
6. Si $A \subseteq B$ con $A, B \in \mathcal{F} \implies \mathbb{P}(A) \leq \mathbb{P}(B)$. Conocida como *monotonicidad*.
7. Para cualesquiera $A_1, A_2 \in \mathcal{F}$ se cumple $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$. Conocido como *inclusión exclusión*.
8. Para cualquier familia $\{A_i\}_{i \in I} \in \mathcal{F}$ se cumple que $\mathbb{P}(\bigcup_i A_i) \leq \sum_i \mathbb{P}(A_i)$. Conocida como la *desigualdad de Boole*.

Probabilidad condicional

El concepto de probabilidad condicional surge como una manera de medir probabilidades una vez que se ha reducido el espacio muestral.

DFN 2 (Probabilidad Condicional). *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad y sean* $A, B \in \mathcal{F}$ *tal que* $\mathbb{P}(B) > 0$*. Entonces definimos*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Teorema 2: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $B \in \mathcal{F}$ tal que $\mathbb{P}(B) > 0$. Sea $\mathbb{Q} : \mathcal{F} \to [0,1]$ definida como $\mathbb{Q}(A) := \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. Entonces $\mathbb{Q}$ es una medida válida de probabilidad sobre $(\Omega, \mathcal{F})$.

Teorema 3 (Probabilidad Total): Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $B_1, B_2, \dots, B_n \in \mathcal{F}$ una partición de Ω. Entonces, para cualquier $A \in \mathcal{F}$, se satisface

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B) \cdot \mathbb{P}(B).$$

Teorema 4 (Teorema de Bayes): Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $B_1, B_2, \dots, B_n \in \mathcal{F}$ una partición de Ω. Sea $A \in \mathcal{F}$ tal que $\mathbb{P}(A) > 0$. Entonces

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}.$$

DFN 3 (Eventos Independientes). *Decimos que dos eventos* $A, B \in \mathcal{F}$ *son independientes si y solamente si* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$*. En dado caso, denotamos que* A *y* B *son independientes como* $A \perp B$*.*

Obs: La independencia de eventos es un caso muy particular. No se puede asumir que es la norma.

Teorema 5: Supongamos $A \perp B$ $(A, B \in \mathcal{F})$. Entonces se cumplen

a) $A^c \perp B$.

b) $A^c \perp B^c$.

c) $A \perp B^c$.

Probabilidad condicional

Teorema 6 (Regla de la multiplicación): Sean $A_1, A_2, \dots, A_n \in F$ cualesquiera tales que $\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) > 0$. Entonces

$$\mathbb{P}(\cap_{i=1}^{n} A_i) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_n| \cap_{i=1}^{n-1} A_n).$$

Variables Aleatorias

DFN 4 (Variable Aleatoria). *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad. Se dice que* $X : \Omega \to \mathbb{R}$ *es una variable aleatoria (v.a) sobre el espacio si y solamente si*

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}.$$

La definición garantiza la permanencia en conjuntos medibles.

Las Variables Aleatorias se pueden clasificar en contínuas o discretas.

DFN 5 (Función de densidad de probabilidades). *Asociada a toda v.a existe una función* $f_X : \mathbb{R} \to [0,1]$ *llamada función de densidad de probabilidades (f.d.p) dada por*

$$f_X(x) := \mathbb{P}[\{\omega \in \Omega \mid X(\omega) = x\}].$$

Como azucar sintáctica podemos escribir lo anterior como

$$\mathbb{P}[\{\omega \in \Omega \mid X(\omega) = x\}] \overset{\text{not.}}{=\!=} \mathbb{P}[X = x].$$

Obs: Con el símbolo $\overset{\text{not.}}{=\!=}$ señalamos que es una notación convencional, no una igualdad en el sentido riguroso.

DFN 6 (Soporte de la f.d.p). *El conjunto* $S \subseteq \mathbb{R}$ *formado por los puntos con probabilidad no-cero es llamado el soporte de la función de densidad de probabilidades.*

Variables Aleatorias

Teorema 7 (Propiedades de f_X)**:** Sea f_X la f.d.p de X una variable aleatoria. Entonces

1. $f_X(x) \geq 0 \quad x \in \mathbb{R}$.
2. (a) En el caso discreto: $\sum_{x \in \mathbb{R}} f_X(x) = 1$
   (b) En el caso contínuo: $\int_{\mathbb{R}} f_X(x) = 1$.

DFN 7 (Función de distribución acumulada)**.** *Asociada a toda v.a se define una función $F_X : \mathbb{R} \to [0,1]$ llamada función de distribución acumulada (f.d.a) dada por*

$$F_X(x) \coloneqq \mathbb{P}[\{\omega \in \Omega \mid X(\omega) \leq x\}].$$

Una vez más, abreviamos como

$$F_X(x) \coloneqq \mathbb{P}[\{\omega \in \Omega \mid X(\omega) \leq x\}] \overset{\text{not.}}{=\!=} \mathbb{P}[X \leq x].$$

Obs: Vemos que

- $F_X(x) = \sum_{t \leq x} f_X(t)$. En el caso discreto.
- $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$. En el caso contínuo.

Teorema 8 (Propiedades de F_X)**:** Sea F_X la f.d.a de una variable aleatoria. Entonces

1. F_X es monótona no-decreciente.
2. F_X es contínua con la derecha.
3. $\lim_{x \to -\infty} F_X(x) = 0$ & $\lim_{x \to \infty} F_X(x) = 1$.

Características de una Variable Aleatoria

Estudiamos otras funciones asociadas a una v.a.

DFN 8 (Función de supervivencia)**.** *Se define la función de supervivencia de una v.a X como*

$$S_X(x) \coloneqq \mathbb{P}[\omega \in \Omega \mid X(\omega) > x].$$

Teorema 9 (Propiedades de S_X)**:** La función de supervivencia S_X cumple

1. S_X es monótona no-creciente.
2. S_X es contínua por la derecha.
3. $S_X(0) = 1$ & $\lim_{t \to \infty} S_X(t) = 0$.

Características de una Variable Aleatoria

DFN 9 (Función de riesgo)**.** *Si X es una variable aleatoria que modela el tiempo de falla de algún componente, definios la función o tasa de riesgo como*

$$h_X(x) = \frac{f_X(x)}{S_X(x)}.$$

Mostramos las definiciones únicamente para el caso contínuo por brevedad.

DFN 10 (Esperanza)**.** *Sea X una v.a con f.d.f $f_X(x)$. Supongamos que $\sum_x |x| \cdot f_X(x)$ está acotada (en el caso discreto, la integral). Entonces, se define la esperanza de X como*

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \cdot f_X(x)\,dx$$

Teorema 10 (Ley del estadístico inconsciente)**:** Sea X una variable aleatoria contínua, y $g : \mathbb{R} \to \mathbb{R}$ una función. Entonces

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) \cdot f_X(x)\,dx$$

DFN 11 (Varianza)**.** *Sea X una variable aleatoria con media μ. Entonces se define la varianza de X como*

$$\mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2].$$

Teorema 11: $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mu^2$.

DFN 12 (k-ésimo momento de X)**.** *Definimos el k-ésimo momento de la variable aleatoria X como*

$$\mathbb{E}\left[X^k\right] = \int_{\mathbb{R}} x^k \cdot f_X(x)\,dx$$

DFN 13 (k-ésimo centrado momento de X)**.** *Definimos el k-ésimo momento de la variable aleatoria X como*

$$\mathbb{E}[(X-\mu)^k] = \int_{\mathbb{R}} (x-\mu)^k f_X(x)$$

Teorema 12 (Linealidad)**:** Se cumple

1. $\mathbb{E}[aX+b] = a\mathbb{E}[X] + b$,
2. $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$.

Características de una Variable Aleatoria

DFN 14 (Funcion generadora de momentos)**.** *Definimos la función generadora de momentos de la variable aleatoria X como*

$$M_X(t) = \mathbb{E}\left[e^{tX}\right]$$

Teorema 13: Sea X una v.a y sea M_X su función generadora de momentos. Entonces

$$\left.\frac{d^k}{dt^k} M_X(T)\right|_{t=0} = \mathbb{E}\left[x^k\right]$$

DFN 15 (Coeficiente de Asimetría)**.** *Se define el coef. de asimetría $\nu_3(X)$ como:*

$$\nu_3(X) = \frac{\mathbb{E}[(X-\mu)^3]}{\mathbb{E}[(X-\mu)^2]^{3/2}}.$$

DFN 16 (Coeficiente de Kurtosis)**.** *Se define el coef. de kurtosis $\nu_3(X)$ como:*

$$\nu_4(X) = \frac{\mathbb{E}[(X-\mu)^4]}{\mathbb{E}[(X-\mu)^2]^2}.$$

Distribución Uniforme Discreta

Decimos que una v.a X se distribuye uniforme discreta ($X \sim \text{unif}\{x_1, \ldots, x_n\}$) si la probabilidad de que X tome cualquiera de los valores en $\{x_1, \ldots, x_n\}$ es constante e igual a $1/n$.
Su f.d.p es

$$f_X(x) = \frac{1}{n} \mathbb{1}_{\{x_1,\ldots,x_n\}}(x)$$

su f.d.a es

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1, \\ \frac{1}{n} & \text{si } 2 \leq x < 3, \\ \frac{2}{n} & \text{si } 3 \leq x < 4, \\ \vdots & \\ \frac{n}{n} & \text{si } x \geq n. \end{cases}$$

Características:

$$\mathbb{E}[X] = \frac{1}{n}\sum_{i=1}^{n} x_i = \mu$$
$$\text{Var}[X] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$
$$M_X(t) = \frac{e^t(1-e^{nt})}{n(1-e^t)}.$$

Distribución Bernoulli

Decimos que X se distribuye Bernoulli ($X \sim \text{Ber}(p)$) si modela un experimento con dos resultados: exito y fracaso con probabilidad p y $1-p$ respectivamente.
Su f.d.p es

$$f_X(x) = p^x(1-p)^{1-x}\,\mathbb{1}_{\{0,1\}}(x)$$

su f.d.a es

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1-p & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Características:

$$\mathbb{E}[X] = p$$
$$\text{Var}[X] = p(1-p)$$
$$M_X(t) = (1-p)pe^t.$$

Distribución binomial

Decimos que X se distribuye binomial ($X \sim \text{bin}(n,p)$) si modela una serie de n ensayos independientes Bernoulli.
Su f.d.p es

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}\,\mathbb{1}_{\{0,1,\ldots,n\}}(x)$$

su f.d.a no tiene una expresión reducida.
Características:

$$\mathbb{E}[X] = np$$
$$\text{Var}[X] = np(1-p)$$
$$M_X(t) = (pe^t + (1-p))^n.$$

Distribución Geométrica

Decimos que X se distribuye binomial ($X \sim \text{geo}(p)$) si modela una sucesión infinita ensayos independientes Bernoulli.
Su f.d.p es

$$f_X(x) = p(1-p)^{x-1}\,\mathbb{1}_{\{0,1,\ldots\}}(x)$$

su f.d.a es

$$F_X(x) = \sum_{u \leq x} f_X(u) \begin{cases} 0 & \text{si } x < 0, \\ 1-(1-p)^{k+1} & \text{si } k \leq x < k+1. \end{cases}$$

Características:

$$\mathbb{E}[X] = \frac{1-p}{p}$$
$$\text{Var}[X] = \frac{1-p}{p^2}$$
$$M_X(t) = \frac{p}{1-(1-p)e^t}.$$

Distribución Binomial Negativa

Decimos que X se distribuye binomial negativa ($X \sim \text{bin neg}(r,p)$) si modela el número de fracasos antes de obtener el r-ésimo éxito en una infinitos ensayos Bernoulli.

$$f_X(x) = \binom{r+x-1}{x} p^r (1-p)^x\,\mathbb{1}_{\{0,1,\ldots\}}(x)$$

Distribución Binomial Negativa

Su f.d.a no tiene expresión reducida.
Características:

$$\mathbb{E}[X] = \frac{r(1-p)}{p}$$
$$\text{Var}[X] = \frac{r(1-p)}{p^2}$$
$$M_X(t) = \frac{p^r}{[1-(1-p)e^t]^r}$$

Distribución Hipergeométrica

Esta distribución modela la probabilidad de obtener un número particular de objetos en una muestra de tamaño n de N totales, con K de interés. Escribimos $X \sim \text{hipergeo}(N,K,n)$.
Su f.d.p es

$$f_X(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}\,\mathbb{1}_{\{0,1,\ldots\}}(x)$$

Su f.d.a no tiene expresión reducida.
Sus primeros dos momentos son

$$\mathbb{E}[X] = n\frac{K}{N},$$
$$\text{Var}[X] = n\frac{K}{N}\frac{N-K}{N}\frac{N-n}{N-1}.$$

Distribución Poisson

Modela la el número de eventos que ocurren dentro de un intervalo de tiempo dado, conociendo la tasa media de ocurrencia λ. Escribimos $X \sim \text{Po}(\lambda)$
Su f.d.p es

$$f_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}\,\mathbb{1}_{\{0,1,\ldots\}}(x)$$

Características:

$$\mathbb{E}[X] = \lambda$$
$$\text{Var}[X] = \lambda$$
$$M_X(t) = e^{\lambda(e^t-1)}.$$

Obs: Se puede aproximar una v.a $X \sim \text{bin}(n,p)$ mediante otra $Y \sim \text{Po}(\lambda)$ con $\lambda = np$.

Distribución Uniforme Contínua

Esta distribución modela un espacio equiprobable en el intervalo (a,b). Escribimos $X \sim \text{unif}(a,b)$.
Su f.d.p es

$$f_X(x) = \frac{1}{b-a}\mathbb{1}_{(a,b)}(x)$$

Su f.d.a es

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq a, \\ \frac{x-a}{b-a} & \text{si } a < x < b, \\ 1 & \text{si } x \geq b. \end{cases}$$

Características:

$$\mathbb{E}[X] = \frac{a+b}{2}$$
$$\text{Var}[X] = \frac{(b-a)^2}{12}$$
$$M_X(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}.$$

Distribución Exponencial

Decimos $X \sim \exp(\lambda)$ con $\lambda > 0$. Puede pensarse como la versión contínua de la distribución Poisson.
Su f.d.p es

$$f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{(0,\infty)}(x)$$

Su f.d.a es

$$F_X(x) = 1 - e^{-\lambda x}\mathbb{1}_{(0,\infty)}(x)$$

Características:

$$\mathbb{E}[X] = \frac{1}{\lambda}$$
$$\text{Var}[X] = \frac{1}{\lambda^2}$$
$$M_X(t) = \frac{\lambda}{\lambda - t}.$$

Distribución Gamma

Decimos $X \sim \text{gamma}(\alpha,\lambda)$ con $\alpha > 0$, $\lambda > 0$.
Su f.d.p es

$$f_X(x) = \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}\mathbb{1}_{(0,\infty)}(x)$$

con

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$$

la función Gamma.
Su f.d.a no tiene expresión compacta.
Características:

$$\mathbb{E}[X] = \frac{\alpha}{\lambda}$$
$$\text{Var}[X] = \frac{\alpha}{\lambda^2}$$
$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha.$$

Distribución Beta

Decimos $X \sim \text{beta}(a,b)$.
Su f.d.p es

$$f_X(x) = \frac{1}{B(a,b)}x^{a-1}\mathbb{1}_{(0,1)}(x)$$

con

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,dt$$

la función Beta.
Su f.d.a no tiene expresión compacta.
Características:

$$\mathbb{E}[X] = \frac{a}{a+b}$$
$$\text{Var}[X] = \frac{ab}{(a+b+1)(a+b)^2}$$
$$M_X(t) = \sum_{k=0}^{\infty}\frac{t^k}{k!}\frac{B(\alpha+k,\beta)}{B(\alpha,\beta)}$$

Distribución Weibull

Decimos $X \sim \text{Weibull}(\alpha,\lambda)$
Su f.d.p es

$$f_X(x) = \lambda\alpha(\lambda x)^{\alpha-1}e^{-(\lambda\alpha)^a}\mathbb{1}_{(0,\infty)}(x)$$

Su f.d.a es

$$F_X(x) = 1 - e^{-(\lambda x)^a}\mathbb{1}_{(0,\infty)}(x)$$

Sus primeros dos momentos son

$$\mathbb{E}[X] = \frac{1}{\lambda}\Gamma\left(1+\frac{1}{\lambda}\right),$$
$$\text{Var}[X] = \frac{1}{\lambda^2}\left[\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma^2\left(1+\frac{1}{\alpha}\right)\right].$$

Distribución Normal

Decimos $X \sim \text{N}(\mu,\sigma^2)$
Su f.d.p es

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Su f.d.a no tiene una forma compacta, y la integral que la describe no tiene forma cerrada.
Características:

$$\mathbb{E}[X] = \mu$$
$$\text{Var}[X] = \sigma^2$$
$$M_X(t) = e^{\mu t}e^{\frac{1}{2}\sigma^2 t^2}.$$

Para la mayoría de los cálculos es necesario tener en cuenta lo siguiente:
Si $X \sim \text{N}(\mu,\sigma^2)$, entonces $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. Por lo tanto, $\mathbb{P}[X \leq x] = \Phi\left(\frac{x-\mu}{\sigma}\right)$.
Donde

$$\Phi(z) = \int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\,du.$$

Para una normal estándar se cumple:

$$\mathbb{E}[X] = 0$$
$$\text{Var}[X] = 1$$
$$M_X(t) = e^{\frac{1}{2}t^2}.$$