

# EYP1027 Modelos Probabilísticos

## Clase 5-1

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística  
Pontificia Universidad Católica de Chile

Primer Semestre 2020



# Contenido I

## 1 Función de distribución

- Definición
- Propiedades básicas
- Ejemplos
- Igualdad en distribución

## 2 Variables Aleatorias Discretas y Continuas

- Definición
- Función de masa de probabilidad (fmp)
- Función de densidad de probabilidad (fdp)

# Variables Aleatorias

## Definición

Dada una variable aleatoria  $X$  en  $(\Omega, \mathcal{A}, P)$ , su distribución de probabilidad,  $P_X(B) = P(X \in B) = P(\{\omega \in \Omega : X(\omega) \in B\})$ ,  $B \in \mathcal{B}$ , queda completamente determinada por las probabilidades de la forma  $P_X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \in B\})$ ,  $x \in \mathbb{R}$ .

### Definición 1.1

**Función de distribución acumulada:** Dada una variable aleatoria  $X$  en  $(\Omega, \mathcal{A}, P)$ , la función definida por

$$F_X(x) = P_X((-\infty, x]) = P(X \leq x), \quad x \in \mathbb{R}.$$

se llama **función de distribución acumulada (fda)** de la variable aleatoria  $X$ .

# Función de distribución

## Propiedades básicas

### Teorema 1.1

Sea  $X$  una variable aleatoria definida sobre  $(\Omega, \mathcal{A}, P)$ . La fda  $F_X$  satisface las siguientes propiedades,

**F1)** Si  $x \leq y$ , entonces  $F_X(x) \leq F_X(y)$   
( $F_X$  es no-decreciente)

**F2)** i)  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  ; ii)  $\lim_{x \rightarrow \infty} F_X(x) = 1$   
( $0 \leq F_X(x) \leq 1 \quad \forall x \in \mathbb{R}$ )

**F3)**  $\lim_{h \rightarrow 0} F_X(x + h) = F_X(x) \quad \forall x \in \mathbb{R}$   
( $F_X$  es continua por la derecha)

**Nota:** La reciproca también se cumple; es decir, cualquier función  $F(x)$ ,  $x \in \mathbb{R}$ , que verifique las propiedades F1), F2) y F3), es la fda de alguna variable aleatoria  $X$ . En otras palabras, una función  $F(x)$ ,  $x \in \mathbb{R}$ , es una fda si y sólo si se cumplen las condiciones F1), F2) y F3).

# Función de distribución

## Demostración 1.1

F1) Si  $x \leq y$ , entonces  $\{X \leq x\} \subseteq \{X \leq y\}$ . Luego, por la definición de  $F_X$  y la propiedad P4) de una medida de probabilidad, se tiene que

$$F_X(x) = P(X \leq x) \leq P(X \leq y) = F_X(y).$$

F2i) Basta probar que  $\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P(X \leq x_n) = 0$ , para cualquier secuencia  $x_n \downarrow -\infty$ . Sea  $x_n = -n$ ; entonces  $\{X \leq -n\} \downarrow \emptyset$ , es decir,  $\lim_{n \rightarrow \infty} \{X \leq -n\} = \bigcap_{n=1}^{\infty} \{X \leq -n\} = \emptyset$ . Luego, por la propiedad P6b) de una medida de probabilidad, se tiene que

$$\lim_{n \rightarrow \infty} P(X \leq -n) = P\left(\lim_{n \rightarrow \infty} \{X_n \leq -n\}\right) = P(\emptyset) = 0.$$

## Función de distribución

F2ii) Similarmente, como  $\{X \leq n\} \uparrow \Omega$  cuando  $n \rightarrow \infty$ , es decir,  $\lim_{n \rightarrow \infty} \{X \leq n\} = \bigcup_{n=1}^{\infty} \{X_n \leq n\} = \Omega$ , entonces la propiedad P6a) de una medida de probabilidad, implica que

$$\lim_{n \rightarrow \infty} P(X \leq n) = P(\lim_{n \rightarrow \infty} \{X_n \leq n\}) = P(\Omega) = 1.$$

F3) Análogamente, como  $x_n = x + \frac{1}{n} \downarrow x$  cuando  $n \rightarrow \infty$ , entonces  $\{X \leq x + \frac{1}{n}\} \downarrow \{X \leq x\}$ . Aplicando nuevamente P6b), se concluye que

$$\begin{aligned} \lim_{n \rightarrow \infty} F_X(x + \frac{1}{n}) &= \lim_{n \rightarrow \infty} P(X \leq x + \frac{1}{n}) \\ &= P(\lim_{n \rightarrow \infty} \{X \leq x + \frac{1}{n}\}) \\ &= P(X \leq x) = F_X(x) \quad \forall x \in \mathbb{R}. \end{aligned}$$

**Nota:** Recuerde que  $\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A} \quad \forall B \in \mathcal{B}$ .

# Función de distribución

## Corolario 1.1

Sea  $X$  una variable aleatoria definida sobre  $(\Omega, \mathcal{A}, P)$ ,  $F_X$  su fda y  $a, b \in \mathbb{R}$  con  $a < b$ ; entonces,

- 1)  $F_X(a^-) = \lim_{h \rightarrow 0^+} F_X(a - h) = P_X(X < a)$ ,
- 2)  $P(X = a) = F_X(a) - F_X(a^-)$  (*salto* de  $F_X$  en el punto  $a$ ),
- 3)  $P(X > a) = 1 - F_X(a)$ ,  $P(X \geq a) = 1 - F_X(a^-)$
- 4)  $P(a < X < b) = F_X(b^-) - F_X(a)$ .
- 5)  $P(a < X \leq b) = F_X(b) - F_X(a)$ ,
- 6)  $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$ ,
- 7)  $P(a \leq X < b) = F_X(b^-) - F_X(a^-)$ ,

**Nota:** De 1) y 2) se concluye que  $F_X(x)$  es continua en un punto  $x$  ssi  $P(X = x) = 0$  ssi  $F_X(x) = F_X(x^-)$ . Luego,  $F_X(x)$  es una función continua en  $\mathbb{R}$  ssi  $P(X = x) = 0 \forall x \in \mathbb{R}$ .

# Función de distribución

## Ejemplos

### Ejemplo 1.1

Sea  $X$  = número de caras obtenidas en tres lanzamientos de una moneda justa. Entonces, el recorrido de  $X$  es  $\mathcal{X} = \{0, 1, 2, 3\}$ , y su fda  $X$  es

$$F_X(x) = \begin{cases} 0 & \text{si } -\infty < x < 0, \\ \frac{1}{8} & \text{si } 0 \leq x < 1, \\ \frac{1}{2} & \text{si } 1 \leq x < 2, \\ \frac{7}{8} & \text{si } 2 \leq x < 3, \\ 1 & \text{si } 3 \leq x < \infty. \end{cases}$$

La función escalera  $F_X(x)$  está graficada en la Figura 1.  $F_X$  se define para todos los valores de  $x \in \mathbb{R}$ , no solo para  $x \in \mathcal{X} = \{0, 1, 2, 3\}$ . Así, por ejemplo,

$$F_X(2.5) = P(X \leq 2.5) = P(X = 0, \text{ o } 1, \text{ o } 2) = \frac{7}{8}.$$

$F_X$  tiene un salto para cada  $x \in \mathcal{X}$ , y el tamaño del salto en  $x$  está dado por  $P_X(X = x) = F_X(x) - F_X(x^-)$ .



# Función de distribución

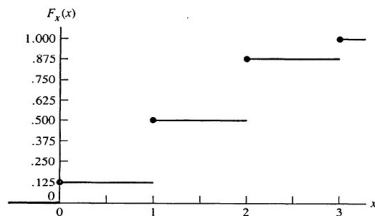


Figura 1: Fda Ejemplo 1.1. (Fuente: Casella & Berger (2002))

# Función de distribución

## Ejemplo 1.2

Experimento: Lanzar una moneda de forma independiente hasta que aparezca una cara. Sea  $p$  ( $0 < p < 1$ ) la probabilidad de que salga cara en cualquier lanzamiento, y defina la variable aleatoria  $X$  = número de lanzamientos necesarios para obtener una cara. Entonces, la probabilidad de requerir  $x$  lanzamientos (independientes) es  $P(X = x) = P(\{\underbrace{(s, \dots, s)}_{x-1 \text{ sellos}}, c\}) = (1 - p)^{x-1}p$ ,  $x = 1, 2, \dots$

Es claro que  $F_X(x) = 0$  si  $x < 1$ . Además,  $F_X(x)$  es plana entre pares consecutivos de enteros positivos. Ahora, para cada  $x = 1, 2, \dots$ ,

$$F_X(x) = P(X \leq x) = \sum_{k=1}^x P(X = k) = \sum_{i=1}^x (1 - p)^{k-1}p = 1 - (1 - p)^x.$$

Es fácil demostrar que  $F_X(x)$  satisface las condiciones F1) a F3) de una fda.

## Función de distribución

F1) Para verificar F1), simplemente observe que la suma anterior contiene más términos positivos a medida que aumenta  $x$ .

F2)

i)  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  ya que  $F_X(x) = 0$  para todo  $x < 1$ ,

ii)  $\lim_{x \rightarrow \infty} F_X(x) = \lim_{x \rightarrow \infty} \{1 - (1 - p)^x\} = 1 \quad (0 < p < 1).$

F3) Finalmente, para verificar F3), note que, para cualquier  $x$ ,  $F_X(x + h) = F_X(x)$  si  $h > 0$  es suficientemente pequeño. Por lo tanto,

$$\lim_{h \downarrow 0} F_X(x + h) = F_X(x), \quad (\text{es decir, es continua por la derecha})$$

**Nota:**  $F_X(x)$  es la fda de una **distribución geométrica**; y se muestra en la Figura 2.

# Función de distribución

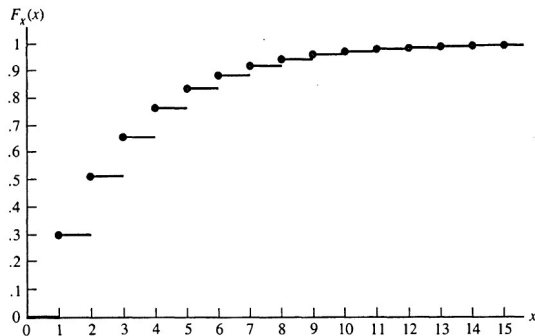


Figura 2: Fda, distribución geométrica, con  $p = 0.3$ . (Fuente: [Casella & Berger \(2002\)](#))

# Función de distribución

## Ejemplo 1.3

Un ejemplo de una fda continua es  $F_X(x) = \frac{1}{1+e^{-x}}$ ,  $x \in \mathbb{R}$ , la cual verifica las condiciones F1) a F3) de una fda:

- i)  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  ya que  $\lim_{x \rightarrow -\infty} e^{-x} = \infty$
- ii)  $\lim_{x \rightarrow \infty} F_X(x) = 1$  ya que  $\lim_{x \rightarrow \infty} e^{-x} = 0$ .

Al derivar  $F_X(x)$  con respecto a  $x$ , se tiene que

$$\frac{d}{dx} F_X(x) = \frac{e^{-x}}{(1 + e^{-x})^2} > 0,$$

de modo que  $F_X(x)$  es una función creciente. Por último, es claro que  $F_X$  es una función continua y, por tanto, es continua a la derecha.

**Nota:** Esta fda es un caso especial de la **distribución logística**

# Función distribución

## Igualdad en distribución

Es importante notar también que dos o más variables aleatorias (definidas sobre el mismo modelo de probabilidad) pueden tener la misma distribución o fda.

### Definición 1.2

Se dice que las variables aleatorias  $X$  y  $Y$  tienen la misma distribución (o son igualmente distribuídas), y se denota como  $X \stackrel{d}{=} Y$ , si para cada conjunto  $B \in \mathcal{B}$ , se tiene que  $P_X(B) = P_Y(B)$ .

**Nota:** El hecho de que dos variables aleatorias tengan la misma distribución, no implica necesariamente que ellas sean iguales.

# Variables Aleatorias Discretas y Continuas

## Ejemplo 1.4

Al lanzar una moneda justa tres veces, defina las variables aleatorias  $X$  = número de caras observadas, e  $Y$  = número de sellos observados. La distribución de  $X$  está dada en el Ejemplo 1.1 y se puede verificar que la distribución de  $Y$  es exactamente la misma. Es decir, para cada  $x = 0, 1, 2, 3$ , tenemos  $P(X = x) = P(Y = x)$ . Así,  $X$  e  $Y$  son idénticamente distribuídas. Sin embargo, para ningún punto del espacio muestral, se tiene que  $X(\omega) = Y(\omega)$ ; de hecho, es fácil ver que  $Y(\omega) = 3 - X(\omega) \forall \omega \in \{c, s\}^3$ , implicando que  $X \stackrel{d}{=} 3 - X$

# Variables Aleatorias Discretas y Continuas

## Teorema 1.2

Las siguientes dos afirmaciones son equivalentes,

- a) Las variables aleatorias  $X$  e  $Y$  son idénticamente distribuidas.
- b)  $F_X(x) = F_Y(x)$  para cada  $x$ .

## Demostración 1.2

Se debe demostrar que cada afirmación implica la otra. Primero, se probará que  $a) \Rightarrow b)$ . Debido a que  $X$  e  $Y$  son idénticamente distribuidas, para cualquier evento  $B \in \mathcal{B}$ ,  $P_X(B) = P_Y(B)$ . En particular, para cada  $x$ , el intervalo  $(-\infty, x]$  está en  $\mathcal{B}$ , y

$$F_X(x) = P_X((-\infty, x]) = P_Y((-\infty, x]) = F_Y(x).$$

Para  $b) \Rightarrow a)$ , ver [Casella & Berger \(2002\)](#).



# Variables Aleatorias Discretas y Continuas

## Definición

Cada variable aleatoria tiene asociada una distribución o fda, y viceversa. De este modo, el tipo de variable aleatoria determina el tipo de distribución y viceversa. Distinguimos dos tipos principales de variables aleatorias o distribuciones, **las discretas** y **las continuas**; aunque se debe mencionar que pueden existir situaciones que deben representarse mediante variables aleatorias o distribuciones mixtas.

La asociación entre una variable aleatoria y su fda es tan estrecha que una forma simple para distinguir entre variable aleatoria discreta y variable aleatoria continua, queda establecida en la siguiente definición:

### Definición 2.1

Una variable aleatoria  $X$  es **discreta** si  $F_X(x)$  es una función escalera de  $x$ . Una variable aleatoria  $X$  es **continua** si  $F_X(x)$  es una función continua de  $x$ .

# Variables Aleatorias Discretas y Continuas

## Función de masa de probabilidad (fmp)

Asociada con una variable aleatoria  $X$  y su fda  $F_X$  hay otra función, llamada función de masa de probabilidad (fmp) en el caso discreto, y función de densidad de probabilidad (fdp) en el caso continuo.

### Definición 2.2

**Función de masa de probabilidad (fmp):** La **fmp** de una variable aleatoria discreta  $X$  está dada por,

$$f_X(x) = P(X = x) \quad \text{para todo } x.$$

# Variables Aleatorias Discretas y Continuas

Claramente, cualquier fmp  $f_X(x) = P(X = x) \geq 0$  para todo  $x$ . Además, si  $\mathcal{X} = \{x_1, x_2, \dots\}$  es el recorrido de  $X$ , entonces:

- a)  $f_X(x_i) > 0$  para todo  $i = 1, 2, \dots$ ,
- b)  $\sum_{i=1}^{\infty} f_X(x_i) = 1$ ,
- c)  $P_X(B) = P(X \in B) = \sum_{\{i: x_i \in B\}} f_X(x_i) \quad \forall B \in \mathcal{B}$

En particular, al tomar  $B = (-\infty, x]$ ,  $x \in \mathbb{R}$ , se obtiene la fda de una variable aleatoria discreta a partir de su fmp como

$$F_X(x) = \sum_{\{i: x_i \leq x\}} f_X(x_i) \quad \forall x \in \mathbb{R}.$$

# Variables Aleatorias Discretas y Continuas

## Ejemplo 2.1

**Distribución Geométrica:** La fmp de la distribución geométrica es,

$$f_X(x) = P_X(X = x) = \begin{cases} (1-p)^{x-1}p & \text{para } x = 1, 2, \dots \\ 0 & \text{en otros casos.} \end{cases}$$

Podemos usar la fmp para calcular probabilidades. Como ahora podemos medir la probabilidad de un solo punto, solo necesitamos sumar la fmp sobre todos los puntos del evento deseado. Por ejemplo, para dos enteros positivos  $a$  y  $b$ , con  $a \leq b$ , tenemos,

$$P(a \leq X \leq b) = \sum_{k=a}^b f_X(x) = \sum_{k=a}^b (1-p)^{x-1}p.$$

# Variables Aleatorias Discretas y Continuas

## Función de densidad de probabilidad (fdp)

El procedimiento análogo en el caso continuo, consiste en sustituir sumas por integrales, de modo que

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t)dt.$$

Usando el teorema fundamental del cálculo, si  $F_X(x)$  es continua, tenemos que,

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

### Definición 2.3

La **función de densidad de probabilidad** o **fdp** de una variable aleatoria continua  $X$ , es una función no-negativa  $f_X(x)$  que satisface,

$$F_X(x) = \int_{-\infty}^x f_X(t)dt \quad \forall x \in \mathbb{R}.$$

## Variables Aleatorias Discretas y Continuas

Note que para el caso continuo,  $P(X = x) = 0$ . En efecto, ya que, para cualquier  $h > 0$ ,  $\{X = x\} \subseteq \{x - h < X \leq x\}$ , entonces,

$$P(X = x) \leq P(x - h < X \leq x) = F_X(x) - F_X(x - h)$$

Por lo tanto,

$$0 \leq P(X = x) \leq \lim_{h \downarrow 0} [F_X(x) - F_X(x - h)] = 0$$

por la continuidad de  $F_X$ . Luego, si  $X$  es una variable aleatoria continua, entonces,

$$P(a < X < b) = P(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b).$$

En general, para cualquier conjunto  $B \in \mathcal{B}$ , se tiene que

$$P_X(B) = P(X \in B) = \int_B f_X(x) dX.$$

# Variables Aleatorias Discretas y Continuas

## Ejemplo 2.2

Para la distribución logística tenemos,  $F_X(x) = \frac{1}{1+e^{-x}}$ , y por lo tanto,

$$f_X(x) = \frac{d}{dx}F_X(x) = \frac{e^{-x}}{(1 + e^{-x})^2}.$$

El área bajo la curva  $f_X(x)$  nos da la probabilidad del intervalo,

$$\begin{aligned}P(a < X < b) &= F_X(b) - F_X(a) \\&= \int_{-\infty}^b f_X(x)dx - \int_{-\infty}^a f_X(x)dx \\&= \int_a^b f_X(x)dx.\end{aligned}$$

# Variables Aleatorias Discretas y Continuas

## Teorema 2.1

Una función  $f_X(x)$  es una fdp (o fmp) de una variable aleatoria  $X$  si y sólo si,

- a)  $f_X(x) \geq 0$  para todo  $x$ ,
- b)  $\int_{-\infty}^{\infty} f_X(x)dx = 1$  (fdp)    ( $\sum_x f_X(x) = 1$  (fmp))

## Demostración 2.1

Si  $f_X(x)$  es una fdp (o fmp), entonces las propiedades a) y b) son directas de la definición. En particular, para una fdp tenemos,

$$1 = \lim_{x \rightarrow \infty} F_X(x) = \int_{-\infty}^{\infty} f_X(t)dt.$$

La implicación inversa no es difícil de probar. Una vez que se tenga  $f_X(x)$ , se puede obtener  $F_X(x)$  y usar el Teorema 1.1.



# Variables Aleatorias Discretas y Continuas

## Ejemplo 2.3

Calculando probabilidades. Suponga que la fdp de  $X$  es de la forma,

$$f_X(x) = \begin{cases} \frac{x}{8} & \text{para } 0 < x < 4 \\ 0 & \text{eoc} \end{cases}$$

Claramente,  $f_X(x) \geq 0$  para todo  $x$ . Además, es fácil ver que

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_0^4 \frac{x}{8} dx = 1.$$

Finalmente, las probabilidades  $P(1 \leq X \leq 2)$  y  $P(X > 2)$ , por ejemplo, pueden calcularse como

$$P(1 \leq X \leq 2) = \int_1^2 \frac{x}{8} dx = \frac{3}{16} \quad \text{y} \quad P(X > 2) = \int_2^4 \frac{x}{8} dx = \frac{3}{4}.$$

# Variables Aleatorias Discretas y Continuas

## Ejercicio 2.1

Sea  $X$  una variable aleatoria con fda dada por

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

- a) Haga un gráfico para esta fda.
- b) Verifique que  $F_X(x)$  es efectivamente una fda.
- c) Calcule  $P(X < \frac{1}{2})$  y  $P(X = \frac{1}{2})$ .
- d) Identifique la naturaleza de esta fda (continua, discreta u otra?).

# Variables Aleatorias Discretas y Continuas

## Ejercicio 2.2

Se extrae una muestra aleatoria (m.a.) de tamaño  $n$  de un lote de  $N$  artículos, de los cuales una fracción  $p$  son defectuosos. Sea  $X$  = número de artículos defectuosos en la m.a. Determine la fmp

$f_X(x) = P(X = x)$  si

- a) La m.a. es ordenada con devolución.
- b) La m.a. es ordenada sin devolución.

Suponga que  $N = 100$ ,  $n = 10$  y  $p = 0.10$ , y que el lote se rechaza si más de la mitad de los artículos inspeccionados son defectuosos.

Calcule la probabilidad de rechazar el lote en ambos casos, a) y b).

# References

- Blanco, L., Arunachalam, V. y Dharmaraja, S. (2012). *Introduction to Probability and Stochastic Processes*. John Wiley, New Jersey.
- Casella, G. y Berger, R.L. (2002). *Statistical Inference*. Second Edition. Duxbury, California.