

# Modelos y Simulación: Parcial 1

Ferré Valderrama, Eduardo

14 de abril de 2025

Modelos y Simulación

## Índice

<b>1. Probabilidad</b>	<b>3</b>
1.1. Espacio de probabilidad . . . . .	3
1.1.1. Axiomas de probabilidad . . . . .	3
1.1.2. Probabilidad condicional . . . . .	3
1.2. Variables aleatorias . . . . .	4
1.2.1. Distribución conjunta . . . . .	5
1.2.2. Distribución condicional . . . . .	6
1.2.3. Convolucion de distribuciones . . . . .	7
1.2.4. Valor esperado . . . . .	7
1.2.5. Varianza . . . . .	8
1.2.6. Desigualdad de Chebyshev . . . . .	9
1.2.7. Leyes de los grandes números . . . . .	9
1.3. Distribuciones de probabilidad . . . . .	10
1.3.1. Variables aleatorias discretas . . . . .	10
1.3.2. Variables aleatorias continuas . . . . .	12
<b>2. Procesos de Poisson</b>	<b>15</b>
2.1. El proceso de poisson homogéneo . . . . .	15
2.1.1. Distribución del número de llegadas $N(t)$ . . . . .	15
2.1.2. Distribución del tiempo entre arribos . . . . .	15
2.1.3. Distribución del tiempo de arribo . . . . .	16
2.1.4. Superposición de procesos de Poisson homogéneos . . . . .	16
2.1.5. Refinamiento de procesos de Poisson homogéneos . . . . .	16
2.2. El proceso de poisson no homogéneo . . . . .	17
<b>3. Generación de Números Aleatorios</b>	<b>18</b>
3.1. Generadores congruenciales . . . . .	19
3.1.1. Generadores congruenciales lineales . . . . .	19
3.1.2. Generadores mixtos . . . . .	20
3.1.3. Generadores multiplicativos . . . . .	20
3.1.4. El problema de los hiperplanos . . . . .	20
3.1.5. Generadores congruenciales lineales combinados . . . . .	21

<b>4. Método de Monte Carlo</b>	<b>21</b>
4.1. Fundamentos del Método de Monte Carlo . . . . .	21
4.2. Estimación de integrales definidas . . . . .	22
4.2.1. Integración sobre $(0, 1)$ . . . . .	22
4.2.2. Integración sobre $[a, b]$ . . . . .	22
4.2.3. Integración sobre $(0, \infty)$ . . . . .	23
4.2.4. Integración sobre $-(\infty), 0$ . . . . .	23
4.2.5. Integración sobre $(a, \infty)$ con transformación racional . . . . .	24
4.2.6. Integración sobre $(-\infty, b]$ . . . . .	24
4.2.7. Integración sobre $(-\infty, \infty)$ . . . . .	25
4.3. Estimación de integrales múltiples . . . . .	26

# 1. Probabilidad

## 1.1. Espacio de probabilidad

Un espacio de probabilidad se define como una terna  $(S, \mathcal{F}, \mathbb{P})$  donde:

- $S$  es el espacio muestral, que es el conjunto de todos los resultados posibles de un experimento aleatorio.
- $\mathcal{F}$  es el conjunto de eventos.
- $P$  es una función de probabilidad

En el presente texto omitiremos a  $\mathcal{F}$  en la notación y solo denotaremos  $(S, \mathbb{P})$ .

### 1.1.1. Axiomas de probabilidad

Consideramos un espacio muestral  $S$ , y suponemos que existe una función  $P$  definida sobre el conjunto de eventos  $S$  que satisface los siguientes axiomas:

- $0 \leq P(A) \leq 1$  para todo evento  $A$ .
- $P(S) = 1$ .
- Si  $A_1, A_2, \dots$  son eventos mutuamente excluyentes dos a dos, entonces

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i), \quad n = 1, 2, \dots$$

Una función  $P$  se la denomina **probabilidad**, y  $P(A)$  es la probabilidad del evento  $A$ . De los axiomas podemos concluir además que:

- $P(A^c) = 1 - P(A)$ , para todo evento  $A$ .
- $P(\emptyset) = 0$ .
- $P(A \cup B) = P(A) + P(B) - P(AB)$ , para cualquier par de eventos  $A$  y  $B$ .

Un **espacio de probabilidad** es un par  $(S, P)$  donde  $S$  es un espacio muestral y  $P$  es una probabilidad sobre  $S$ . Si es necesario especificar la familia  $\mathcal{F}$  de eventos, se utiliza la notación  $(S, \mathcal{F}, P)$ .

### 1.1.2. Probabilidad condicional

Dados dos eventos  $A$  y  $F$ , y una probabilidad  $P$ , la **probabilidad condicional** de que ocurra  $A$  dado  $F$  se define como:

$$P(A|F) = \frac{P(AF)}{P(F)}, \quad P(F) > 0$$

Dos eventos  $A$  y  $B$  son **independientes** si  $P(A|B) = P(A)$ . En este caso se cumple:

$$P(AB) = P(A) \cdot P(B)$$

Notemos que para cualquier par de eventos  $A$  y  $F$ , se tiene que  $A = AF \cup AF^c$ . Dado que  $F$  y  $F^c$  son mutuamente excluyentes, entonces también lo son  $AF$  y  $AF^c$ . Por lo tanto

$$P(A) = P(AF) + P(AF^c).$$

Esto nos permite calcular la probabilidad de  $A$  como una suma ponderada de probabilidades condicionales:

$$\begin{aligned} P(A) &= P(A | F)P(F) + P(A | F^c)P(F^c) \\ &= P(A | F)P(F) + P(A | F^c)(1 - P(F)). \end{aligned} \quad (1.1)$$

En particular, también tenemos que  $P(AF) = P(A | F)P(F)$  y  $P(AF) = P(F | A)P(A)$ . Entonces:

$$P(F | A) = \frac{P(AF)}{P(A)} = \frac{P(A | F)P(F)}{P(A)}. \quad (1.2)$$

La igualdad (1.2) se conoce como **Fórmula de Bayes**, y permite calcular la probabilidad condicional  $P(F | A)$  (a posteriori) en términos de  $P(F)$  (a priori). Puede generalizarse a un número finito de eventos. Esto es: si  $F_1, F_2, \dots, F_n$  son eventos mutuamente excluyentes tales que  $F_1 \cup F_2 \cup \dots \cup F_n = S$ , entonces para cualquier evento  $A$  se tiene que:

$$\begin{aligned} P(A) &= P(AF_1) + P(AF_2) + \dots + P(AF_n) \\ &= P(A | F_1)P(F_1) + P(A | F_2)P(F_2) + \dots + P(A | F_n)P(F_n) \\ &= \sum_{i=1}^n P(A | F_i)P(F_i) \end{aligned}$$

Luego, la probabilidad condicional de que ocurra el evento  $F_j$  dado  $A$  es igual a:

$$P(F_j | A) = \frac{P(A | F_j)P(F_j)}{P(A)} = \frac{P(A | F_j)P(F_j)}{\sum_{i=1}^n P(A | F_i)P(F_i)}. \quad (1.3)$$

## 1.2. Variables aleatorias

Dado un espacio muestral  $(S, P)$ , una **variable aleatoria** es una función  $X : S \rightarrow \mathbb{R}$  tal que el conjunto  $\{s \in S | X(s) \leq x\}$  es un evento sobre el que está definido  $P$ , para todo  $x \in \mathbb{R}$ .

Dada una variable aleatoria  $X$  en un espacio de probabilidad  $S$ , se define la **función de distribución acumulada** como:

$$F(X) = P(X \leq x) = P(\{s \in S | X(s) \leq x\}), \quad (1.4)$$

La función  $F$  tiene dominio en los números reales y toma valores en el intervalo  $[0, 1]$ . En particular cumple las siguientes propiedades:

- $F$  es no decreciente.
- para todo  $x \in \mathbb{R}$ , se cumple que  $0 \leq F(x) \leq 1$ .
- $F$  es continua a derecha.

## Variable aleatoria discreta

Diremos que una variable aleatoria es **discreta** si toma sólo un número finito o numerable de valores. En este caso, se define la **función de probabilidad de masa** como:

$$p_X(x) = P(X = x).$$

Si la variable  $X$  toma valores en un conjunto finito  $\{x_1, x_2, \dots, x_n\}$ , entonces se cumple que:

$$\sum_{i=1}^n p_X(x_i) = 1.$$

y si toma una cantidad infinita numerable de valores  $\{x_1, x_2, \dots\}$ , entonces se cumple que:

$$\sum_{i=1}^{\infty} p_X(x_i) = 1.$$

## Variable aleatoria continua

Una variable aleatoria se dice que es una variable **absolutamente** continua si existe una función  $f$  tal que para todo subconjunto  $C \subseteq \mathbb{R}$  se cumple que:

$$P(X \in C) = \int_C f(x) dx.$$

La función se denomina **función de densidad de probabilidad** de la variable aleatoria  $X$ .

Notemos además que si  $X$  es continua, entonces  $P(X = a) = \int_a^a f(t) dt$ , y por lo tanto  $P(X = a) = 0$ . En particular,  $f$  y la función de distribución acumulada  $F$  están relacionadas por la siguiente expresión:

$$F(a) = P(X \leq a) = \int_{-\infty}^a f(t) dt.$$

Derivando con respecto a  $a$  tenemos que  $F'(a) = f(a)$ , para todo  $a \in \mathbb{R}$  tal que  $F$  es derivable en  $a$ .

### 1.2.1. Distribución conjunta

Si  $X$  e  $Y$  son variables aleatorias sobre un espacio de probabilidad  $(S, P)$ , se llama **función de distribución acumulada conjunta** de  $X$  e  $Y$  a la función  $F : \mathbb{R} \times \mathbb{R} \mapsto [0, 1]$  dada por

$$F(a, b) = P(X \leq a, Y \leq b).$$

En particular, si  $X$  e  $Y$  son variables aleatorias discretas se define la **función de masa de probabilidad conjunta** de  $X$  e  $Y$  como:

$$p(a, b) = P(X = a, Y = b).$$

Dos variables aleatorias  $X$  e  $Y$  se dicen **conjuntamente continuas** si existe una función  $f$  llamada **función de densidad conjunta** tal que

$$P(X \in C, Y \in D) = \int \int_{x \in C, y \in D} f(x, y) dx dy.$$

Si  $F$  es función de distribución conjunta de  $X$  e  $Y$ , entonces pueden calcularse las distribuciones de probabilidad de  $X$  e  $Y$  a partir de  $F$ , también llamadas **distribuciones marginales**  $F_X$  y  $F_Y$ . Esto es:

$$F_X(a) = P(X \leq a) = F(a, \infty), \quad F_Y(b) = P(Y \leq b) = F(\infty, b).$$

Si  $X$  e  $Y$  son variables aleatorias discretas con función de masa conjunta  $p$ , entonces las funciones de probabilidad de masa marginales de  $X$  e  $Y$  están dadas respectivamente por:

$$p_X(a) = \sum_b p(a, b), \quad p_Y(b) = \sum_a p(a, b).$$

Aquí los subíndices de la sumatoria  $b$  y  $a$  toman todos los valores posibles en el rango de  $Y$  y  $X$ , respectivamente.

Si  $X$  e  $Y$  son conjuntamente continuas con función de densidad conjunta  $f$ , las distribuciones marginales están dadas por:

$$F_X(a) = \int_{-\infty}^a \int_{-\infty}^{\infty} f(x, y) dy dx = \int_{-\infty}^a f_X(x) dx.$$

$$F_Y(b) = \int_{-\infty}^{\infty} \int_{-\infty}^b f(x, y) dx dy = \int_{-\infty}^b f_Y(y) dy.$$

Notemos que las correspondientes densidades marginales  $f_X$  y  $f_Y$  están dadas por:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{y} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

Así, a partir de la función de distribución conjunta es posible obtener las distribuciones marginales de  $X$  e  $Y$ . La recíproca no es cierta en general.

### 1.2.2. Distribución condicional

Si  $X$  e  $Y$  son variables aleatorias discretas, entonces la **probabilidad de masa condicional**  $p_{X|Y}$  se define como

$$p_{X|Y}(x | y) = P(X = x | Y = y).$$

En particular, tenemos que:

$$p_{X|Y}(x | y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x, y)}{p_Y(y)}.$$

Si  $X$  e  $Y$  son conjuntamente continuas, se define la **función de densidad condicional** como:

$$f_{X|Y}(x | y) = \frac{f(x, y)}{f_Y(y)}.$$

En particular, la probabilidad condicional  $P(X \leq x | Y = y)$  está dada por

$$P(X \leq a | Y = y) = \int_{-\infty}^a \frac{f(x, y)}{f_Y(y)} dx = \int_{-\infty}^a \frac{f(x, y)}{\int_{-\infty}^{\infty} f(s, y) ds} dx.$$

Tenemos además una versión de la fórmula de Bayes que relaciona las distribuciones condicionales y marginales para dos variables aleatorias  $X$  e  $Y$ :

$$p_{X|Y}(x | y) = \frac{p_{Y|X}(y | x)p_X(x)}{p_Y(y)}, \quad f_{X|Y}(x | y) = \frac{f_{Y|X}(y | x)f_X(x)}{f_Y(y)}.$$

### 1.2.3. Convolution de distribuciones

Dadas dos variables aleatorias  $X$  e  $Y$  sobre un espacio de probabilidad  $(S, P)$ , consideremos la distribución de la suma de estas variables,  $X + Y$ . Así, si  $X$  e  $Y$  son variables discretas, entonces  $X + Y$  también es discreta y se tiene que  $X + Y$  toma un valor  $a$  cada vez que  $X$  e  $Y$  toman valores  $x$  y  $a - x$ , y recíprocamente. Por lo tanto:

$$P(X + Y = a) = \sum_x P(X = x, Y = a - x) = \sum_x p(x, a - x),$$

donde  $p(x, y)$  denota la probabilidad conjunta de  $X$  e  $Y$ . Si  $X$  e  $Y$  son conjuntamente continuas, entonces  $X + Y \leq a$  cada vez que  $X$  toma un valor  $x$  e  $Y$  es menor o igual a  $a - x$ . Luego,

$$P(X + Y \leq a) = \int_{-\infty}^{\infty} \int_{-\infty}^{a-x} f_{X,Y}(x, y) dy dx = \int_{-\infty}^{\infty} \int_{-\infty}^a f_{X,Y}(x, y - x) dy dx,$$

con  $f_{X,Y}$  la densidad conjunta de  $X$  e  $Y$ .

Si las variables son independientes, entonces para el caso discreto resulta que la probabilidad de masa de la variable aleatoria  $X + Y$  está dada por:

$$P(X + Y = a) = \sum_x p_X(x) p_Y(a - x).$$

Para el caso continuo tenemos que

$$P(X + Y \leq a) = \int_{-\infty}^{\infty} \int_{-\infty}^a f_X(x) f_Y(y - x) dy dx = \int_{-\infty}^a \int_{-\infty}^{\infty} f_X(x) f_Y(y - x) dx dy,$$

de donde vemos que la densidad de  $X + Y$  está dada por  $\int_{-\infty}^{\infty} f_X(x) f_Y(y - x) dx$ . Aquí  $p_X$ ,  $p_Y$ ,  $f_X$  y  $f_Y$  son las probabilidades de masa y densidades marginales respectivamente.

Esto da lugar al concepto de **convolución** de funciones de probabilidad. En particular, para las probabilidades de masa la convolución  $p_X * p_Y(a)$  se define por:

$$p_X * p_Y(a) = \sum_x p_X(x) p_Y(a - x).$$

Para las funciones de densidad, la convolución  $f_X * f_Y(a)$  se define por:

$$f_X * f_Y(a) = \int_{-\infty}^{\infty} f_X(x) f_Y(a - x) dx.$$

Así, si  $X$  e  $Y$  son variables aleatorias discretas independientes, entonces la probabilidad de masa de  $X + Y$  está dada por la convolución  $p_X * p_Y$ . Si son continuas e independientes, la densidad de  $X + Y$  está dada por la convolución de las densidades marginales,  $f_X * f_Y$ .

### 1.2.4. Valor esperado

Dada una variable aleatoria discreta  $X$  que toma valores  $x_i$ ,  $i = 1, 2, \dots, n, \dots$ , se llama **valor esperado o esperanza matemática** a la cantidad (si existe)

$$E[X] = \sum_i x_i P(X = x_i).$$

Si  $X$  es una variable aleatoria continua, su valor esperado se define por el valor de la integral (si existe):

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

Es importante notar que el valor esperado no es necesariamente un valor posible de  $X$ .

Sean  $X$  e  $Y$  dos variables aleatorias sobre un mismo espacio de probabilidad  $S$ . Entonces se cumplen las siguientes propiedades:

a) Si  $g : \mathbb{R} \mapsto \mathbb{R}$ , entonces  $g(X)$  es una variable aleatoria y

$$E[g(X)] = \sum_i g(x_i) p(x_i), \quad (\text{si } X \text{ es discreta}),$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx, \quad (\text{si } X \text{ es continua}).$$

En particular, tomando  $g(x) = ax + b$  se deduce que

$$E[aX + b] = aE[X] + b.$$

b) El valor esperado es un operador lineal. Esto es,

$$E[X + Y] = E[X] + E[Y]$$

### 1.2.5. Varianza

La **varianza** es una medida de la dispersión de  $X$  en torno a su valor esperado  $E[X] = \mu$ , y está definida por:

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

Notemos que

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

La varianza es siempre un número positivo a menos que la variable aleatoria sea siempre constante. Es importante notar que si  $a$  y  $b$  son números reales, entonces

$$\text{Var}(aX + b) = a^2 \text{Var}(X).$$

La varianza no verifica la condición de linealidad. En efecto, si  $E[X] = \mu$ ,  $E[Y] = \theta$ , entonces

$$\begin{aligned} \text{Var}(X + Y) &= E[((X + Y) - (\mu + \theta))^2] \\ &= E[(X - \mu)^2 + (Y - \theta)^2 + 2(X - \mu)(Y - \theta)] \\ &= E[(X - \mu)^2] + E[(Y - \theta)^2] + 2E[(X - \mu)(Y - \theta)] \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{cov}(X, Y). \end{aligned}$$

Si  $X$  e  $Y$  son variables aleatorias, se define la **covarianza** de  $X$  e  $Y$  por:

$$\text{cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)].$$

De esta manera,

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{cov}(X, Y).$$



Si  $X$  e  $Y$  son independientes entonces  $\text{cov}(X, Y) = 0$  (la recíproca en general no es cierta). Esto puede verse observando que  $\text{cov}(X, Y) = E[XY] - E[X]E[Y]$ . Luego, si  $X$  e  $Y$  son independientes y conjuntamente continuas con densidad conjunta  $f$ , entonces  $f(x, y) = f_X(x)f_Y(y)$  y por lo tanto:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_X(x) y f_Y(y) dx dy = E[X]E[Y].$$

Si  $X$  e  $Y$  son discretas con probabilidad de masa conjunta  $p$ , entonces  $p(x, y) = p_X(x)p_Y(y)$  y

$$E[XY] = \sum_i \sum_j x_i y_j p(x_i, y_j) = \sum_i \sum_j x_i p_X(x_i) y_j p_Y(y_j) = E[X]E[Y].$$

Otra medida de dispersión de una variable aleatoria es la **desviación estándar**  $\sigma(X)$ , definida por

$$\sigma(X) = \sqrt{\text{Var}(X)}.$$

Tiene la ventaja de mantener las mismas unidades (distancia, longitud, tiempo, etc.) que  $X$  y  $E[X]$ . A su vez, se define la **correlación** de dos variables aleatorias  $X$  e  $Y$  por:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X) \cdot \sigma(Y)}.$$

Se cumple que  $\rho(X, Y)$  es un número entre  $-1$  y  $1$ , y da una medida normalizada de la covarianza entre dos variables aleatorias.

### 1.2.6. Desigualdad de Chebyshev

Si  $X$  toma sólo valores no negativos y  $a > 0$ , entonces se cumple la desigualdad

$$P(X \geq a) \leq \frac{E[X]}{a}. \quad (1.5)$$

**Teorema 1** (Desigualdad de Chebyshev). Si  $X$  es una variable aleatoria con media  $\mu$  y varianza  $\sigma^2$ , entonces para  $k > 0$  se cumple:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

### 1.2.7. Leyes de los grandes números

Dos variables aleatorias se dicen **idénticamente distribuidas** si tienen una misma función de distribución acumulada.

Si  $X_1, X_2, \dots, X_n$  son variables aleatorias independientes e idénticamente distribuidas, con media  $\mu$ , se cumplen las siguientes propiedades:

- **Ley débil de los grandes números:**

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| > \epsilon\right) \rightarrow 0 \quad \text{cuando } n \rightarrow \infty.$$

■ **Ley fuerte de los grandes números:**

Con probabilidad 1 se cumple que:

$$\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu.$$

### 1.3. Distribuciones de probabilidad

#### 1.3.1. Variables aleatorias discretas

## Distribuciones Discretas

### Distribución Uniforme Discreta: $U\{1, n\}$

Se dice que una variable aleatoria tiene **distribución uniforme** si todos sus valores son equiprobables. Con  $U\{1, n\}$  denotaremos a la variable aleatoria que toma valores en el conjunto  $\{1, 2, \dots, n\}$ , todos con la misma probabilidad  $\frac{1}{n}$ .

$$p(i) = \frac{1}{n}, \quad i = 1, 2, \dots, n.$$

El valor esperado y la varianza están dados por:

$$E[X] = \frac{n+1}{2}, \quad \text{Var}(X) = \frac{n^2-1}{12}.$$

### Distribución de Bernoulli: $B(p)$

Una variable aleatoria que toma dos valores con probabilidad  $p$  (éxito) y  $1-p$  (fracaso), se dice de **Bernoulli**. La distribución de Bernoulli teórica se corresponde con la variable aleatoria que toma los valores 1 y 0:

$$X = \begin{cases} 1 & \text{con prob. } p \\ 0 & \text{con prob. } 1-p \end{cases}$$

Su valor esperado y varianza están dados por:

$$E[X] = p, \quad \text{Var}(X) = p(1-p).$$

### Distribución Binomial: $B(n, p)$

Si consideramos un experimento que consiste en  $n$  ensayos independientes, cada uno con probabilidad  $p$  de éxito, entonces el número de éxitos tiene una distribución binomial de parámetros  $n$  y  $p$ . El rango de la variable aleatoria es el conjunto  $\{0, 1, 2, \dots, n\}$ , y la función de masa de probabilidad está dada por:

$$p(i) = P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad 0 \leq i \leq n.$$

Más adelante resultará útil la siguiente fórmula recursiva para las probabilidades de masa:

$$p(0) = (1-p)^n, \quad p(i+1) = \frac{n-i}{i+1} \frac{p}{1-p} p(i), \quad 0 \leq i \leq n-1.$$

El valor esperado y la varianza están dados por:

$$E[X] = np, \quad \text{Var}(X) = np(1 - p).$$

## Distribución de Poisson: $\mathcal{P}(\lambda)$

Una variable aleatoria se dice que es de **Poisson con parámetro**  $\lambda$  si toma valores en  $\mathbb{N} \cup \{0\}$  con probabilidad de masa

$$p(i) = P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i \geq 0.$$

Una fórmula recursiva para estas probabilidades está dada por:

$$p(0) = e^{-\lambda}, \quad p(i+1) = \frac{\lambda}{i+1} p(i), \quad i \geq 0.$$

El valor esperado y la varianza están dados por:

$$E[X] = \lambda, \quad \text{Var}(X) = \lambda.$$

## Distribución Geométrica: $\text{Geom}(p)$

Dada una sucesión de ensayos independientes con probabilidad  $p$  de éxito, la variable aleatoria geométrica cuenta el número de ensayos independientes hasta obtener el primer éxito. Su rango es el conjunto de números naturales,  $\mathbb{N}$ .

La función de probabilidad de masa está dada por:

$$p(n) = P(X = n) = p(1 - p)^{n-1}, \quad n \geq 1.$$

El valor esperado y la varianza están dados por:

$$E[X] = \frac{1}{p}, \quad \text{Var}(X) = \frac{1 - p}{p^2}.$$

## Distribución Binomial Negativa o Pascal: $\text{Bn}(r, p)$

Esta distribución se corresponde con la variable aleatoria que mide el número de ensayos independientes con probabilidad de éxito  $p$ , hasta obtener  $r$  éxitos. La variable toma valores en el intervalo  $\{r, r+1, r+2, \dots\} = \{n \in \mathbb{N} \mid n \geq r\}$ . La función de probabilidad de masa está dada por:

$$P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \quad n \geq r.$$

El valor esperado y la varianza están dados por:

$$E[X] = \frac{r}{p}, \quad \text{Var}(X) = \frac{r(1-p)}{p^2}.$$

## Distribución Hipergeométrica: $H(n, N, M)$

Esta distribución se corresponde con la variable aleatoria que mide el número de éxitos en una muestra de tamaño  $n$  extraída de un conjunto de  $N + M$  elementos, donde un éxito equivale a extraer un elemento del subconjunto de cardinal  $N$ .

El rango de esta distribución es  $\{0, 1, 2, \dots, \min(n, N)\}$ . La función de probabilidad de masa está dada por:

$$p(i) = P(X = i) = \frac{\binom{N}{i} \binom{M}{n-i}}{\binom{N+M}{n}}.$$

La probabilidad  $p(i)$  es 0 si  $i > N$  o  $n - i > M$ . El valor esperado y la varianza están dados por:

$$E[X] = \frac{nN}{N+M}, \quad \text{Var}(X) = \frac{nNM}{(N+M)^2} \left(1 - \frac{n-1}{N+M-1}\right).$$

### 1.3.2. Variables aleatorias continuas

## Función Indicadora

Denotaremos con  $I_A$  a la **función indicadora del conjunto**  $A$ , dada por

$$I_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

## Distribución Uniforme Continua: $U(a, b)$

Una variable aleatoria continua  $X$  se dice uniformemente distribuida en  $(a, b)$  si su función de densidad está dada por

$$f(x) = \frac{1}{b-a} \mathbb{I}_{(a,b)}(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{c.c.} \end{cases}$$

Su función de distribución acumulada está dada por:

$$F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$$

y la varianza y su valor esperado son:

$$E[X] = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{1}{12}(b-a)^2.$$

## Distribución Normal: $N(\mu, \sigma^2)$

Una variable aleatoria continua  $X$  se dice normalmente distribuida con media  $\mu$  y varianza  $\sigma^2$  si su función de densidad de probabilidad está dada por

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

En tal caso usamos la notación  $X \sim N(\mu, \sigma^2)$ . Si  $Z \sim N(0, 1)$  se dice que su distribución es **normal estándar**. La función de densidad es entonces:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

Si  $X \sim N(\mu, \sigma^2)$ , entonces  $\frac{X-\mu}{\sigma}$  tiene distribución normal estándar. Denotemos con  $\Phi$  a la función de distribución acumulada de una variable aleatoria normal estándar,  $Z \sim N(0, 1)$ :

$$\Phi(x) = P(Z \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt. \quad (1.6)$$

La integral  $\Phi(x)$  no tiene una fórmula cerrada, por lo cual es usual utilizar valores tabulados para el cálculo o interpolación de valores de  $\Phi$ . De la expresión para la función de densidad, se observa que es una función par, esto es,  $f(x) = f(-x)$ . Luego se cumple que

$$\Phi(x) = 1 - \Phi(-x), \quad \text{para todo } x \in \mathbb{R}.$$

Si  $X$  es normal con media  $\mu$  y varianza  $\sigma^2$ , entonces su función de distribución acumulada puede expresarse en términos de  $\Phi$ :

$$F(x) = P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

Algunos valores importantes a recordar son las probabilidades de que los valores de una variable aleatoria normal se distribuyan alrededor de la media, a una distancia menor de  $k\sigma$ , para ciertos valores de  $k$ . Para  $k = 1, 2, 3$  tenemos:

$$\begin{aligned} P(|X - \mu| < \sigma) &\approx 68\%, \\ P(|X - \mu| < 2\sigma) &\approx 95\%, \\ P(|X - \mu| < 3\sigma) &\approx 99.7\%. \end{aligned}$$

Si  $Z \sim N(0, 1)$  y  $\alpha$  es un número entre 0 y 1, se suele denotar  $z_\alpha$  al número real tal que

**Teorema 2** (Teorema Central del Límite). Sean  $X_1, X_2, \dots, X_n$  variables aleatorias independientes e idénticamente distribuidas con media  $\mu$  y varianza  $\sigma^2$ . Entonces

$$\lim_{n \rightarrow \infty} P\left(\frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}} < x\right) = \Phi(x).$$

Este teorema resulta útil para estimaciones estadísticas de intervalos de confianza a partir de una muestra de tamaño  $n$ .

## Distribución Exponencial: $\mathcal{E}(\lambda)$

Una variable aleatoria  $X$  tiene distribución exponencial con parámetro  $\lambda$ ,  $\lambda > 0$ , si su función de densidad está dada por:

$$f_\lambda(x) = \lambda e^{-\lambda x} \mathbb{I}_{(0, \infty)}(x).$$

Su valor esperado y varianza están dados por:

$$E[X] = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

Si  $X \sim \mathcal{E}(\lambda)$ , entonces  $cX \sim \mathcal{E}\left(\frac{\lambda}{c}\right)$ .

Se dice que una variable aleatoria  $X$  tiene **falta de memoria** si:

$$P(X > s + t \mid X > s) = P(X > t).$$

### Distribución Lognormal: $LN(\mu, \sigma)$

Una variable aleatoria continua  $X$  se dice **lognormal** si su logaritmo  $\ln(X)$  tiene distribución normal. Denotaremos  $X \sim LN(\mu, \sigma)$  si el logaritmo de  $X$  tiene media  $\mu$  y desviación estándar  $\sigma$ . En tal caso su función de densidad está dada por:

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-(\ln(x)-\mu)^2/(2\sigma^2)} \cdot \mathbb{I}_{(0,\infty)}(x).$$

Su valor esperado y varianza están dados por:

$$E[X] = e^{\mu+\sigma^2/2}, \quad \text{Var}(X) = e^{2\mu+2\sigma^2}(e^{\sigma^2} - 1).$$

### Distribución Gamma: $\Gamma(\alpha, \beta)$

Una variable aleatoria Gamma con parámetros  $\alpha$  y  $\beta$  tiene una función de densidad dada por:

$$f(x) = \frac{1}{\Gamma(\alpha)} \beta^{-\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}, \quad x > 0.$$

Su valor esperado y varianza están dados por:

$$E[X] = \alpha\beta, \quad \text{Var}(X) = \alpha\beta^2.$$

### Distribución Weibull: $W(\alpha, \beta)$

Una variable aleatoria continua Weibull con parámetros  $\alpha$  y  $\beta$ , con  $\alpha > 0$  y  $\beta > 0$ , tiene función de densidad dada por:

$$f(x) = \alpha\beta^{-\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} \mathbb{I}_{(0,\infty)}(x).$$

Su valor esperado y varianza son:

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

## 2. Procesos de Poisson

Un proceso de Poisson puede pensarse como el proceso de contar el número de arribos o llegadas ocurridos hasta el tiempo  $t$ , sabiendo que la tasa de llegada por unidad de tiempo es  $\lambda$ .

### 2.1. El proceso de poisson homogéneo

Entonces por lo mencionado anteriormente:

- a) Al momento  $t = 0$  no se contabiliza ningún arribo.
- b) **Incrementos independientes:** Si se consideran dos o más intervalos de tiempo no solapados entre sí, el número de arribos que ocurre en uno y otro intervalo son variables aleatorias independientes.
- c) **Incrementos estacionarios:** La distribución del número de llegadas que ocurre en un período de tiempo depende sólo del tiempo transcurrido y no de la ubicación en el tiempo de este período. Esta propiedad es la que determina que el proceso sea homogéneo.
- d) En un intervalo pequeño de tiempo:
  - La probabilidad de que ocurra una llegada es proporcional a la longitud del intervalo, con constante de **intensidad** igual a la tasa de arribos  $\lambda$ .
  - La probabilidad de que lleguen dos o más eventos simultáneamente tiende a ser nula cuando el intervalo de tiempo es reducido.

#### 2.1.1. Distribución del número de llegadas $N(t)$

$$P(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

Donde  $k$  la cantidad de llegadas (o eventos) que ocurren en el intervalo de tiempo  $t$ . Entonces  $N(t)$  tiene una distribución de Poisson con parámetro  $\lambda t$ .

$$N(t) \sim \mathcal{P}(\lambda t).$$

#### 2.1.2. Distribución del tiempo entre arribos

- $N(t)$ , tasa  $\lambda$
- $X_1$  : Tiempo hasta el primer evento
- $X_k$  : Tiempo transcurrido desde el evento  $k - 1$  hasta el evento  $k$ .

$X_k$  tiene una distribución exponencial con parámetro  $\lambda$ .

$$X_k \sim \mathcal{E}(\lambda).$$

### 2.1.3. Distribución del tiempo de arribo

- $N(t)$ , tasa  $\lambda$
- $X_1, X_2, \dots, X_n$  : Tiempo hasta el evento  $n$ .

Llamemos  $S_n$  el tiempo en el que ocurre el  $n$ -ésimo evento.

$$S_n = X_1 + X_2 + \dots + X_n.$$

Entonces  $S_n$  tiene una distribución gamma con parámetros  $n$  y  $\frac{1}{\lambda}$ .

$$S_n \sim \Gamma(n, \frac{1}{\lambda}).$$

### 2.1.4. Superposición de procesos de Poisson homogéneos

- $N_1(t)$ , tasa  $\lambda_1$
- $N_2(t)$ , tasa  $\lambda_2$
- $t \geq 0$
- $N_1(t)$  y  $N_2(t)$  independientes entre sí.

$$M(t) = N_1(t) + N_2(t).$$

$$M(t) \sim \mathcal{P}(\lambda_1 t + \lambda_2 t).$$

$\{M(t), t \geq 0\}$  es un proceso de Poisson homogéneo con tasa  $\lambda_1 + \lambda_2$ .

Se obtiene que si sumamos:

$$M(t) = N_1(t) + N_2(t) + \dots + N_n(t).$$

$N_1(t), N_2(t), \dots, N_n(t)$  son independientes entre sí y cada uno de ellos es un proceso de Poisson homogéneo con tasas  $\lambda_1, \lambda_2, \dots, \lambda_n$  respectivamente.

Entonces  $M(t)$  es un proceso de Poisson homogéneo con tasa  $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_n$ .

$$P(\{\text{Primer evento sea del proceso } N_k\}) = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_n}.$$

### 2.1.5. Refinamiento de procesos de Poisson homogéneos

- $N(t)$ , tasa  $\lambda$
- $t \geq 0$
- $\lambda > 0$
- En lugar de contar todos los eventos, vamos a contar los eventos y marcarlos con una cierta probabilidad  $p$  de un determinado tipo  $T$ .



Entonces tenemos:

$$\begin{aligned} 0 < p_1, p_2, \dots, p_n < 1 \\ p_1 + p_2 + \dots + p_n &= 1 \end{aligned}$$

Esto produce:

$$\begin{aligned} N_1(t) &: \text{Eventos de tipo } T_1 \\ N_2(t) &: \text{Eventos de tipo } T_2 \\ &\vdots \\ N_n(t) &: \text{Eventos de tipo } T_n \end{aligned}$$

Donde  $N_1(t), N_2(t), \dots, N_n(t)$  son procesos de Poisson homogéneos e independientes

$N_j(t)$  es un proceso de Poisson homogéneo con tasa  $\lambda_j = \lambda p_j$ .

Observación:  $\lambda p_1 + \lambda p_2 + \dots + \lambda p_n$

## 2.2. El proceso de poisson no homogéneo

En este caso la tasa de arribos  $\lambda(t)$  no es constante, sino que depende del tiempo. En este caso, la función de intensidad  $\lambda(t)$  es una función integrable en el intervalo  $[0, t]$ .

La función **valor medio del proceso** mide la intensidad media del número de llegadas en un intervalo. Está dada por:

$$m(t) = \int_0^t \lambda(s) ds$$

Notemos que en este caso los incrementos son independientes pero no estacionarios, ya que la distribución de  $N(t+s) - N(t)$  dependerá de la función de intensidad  $\lambda$  en el período  $(t, t+s]$ . Por otra parte si  $\lambda(t) = \lambda$ , constante, entonces  $m(t) = \lambda \cdot t$  y es el caso del proceso de Poisson homogéneo.

En particular, se tiene que para cada  $t \geq 0$  y  $s > 0$ , el número de llegadas en el intervalo  $(t, t+s]$  es una variable aleatoria Poisson con media  $m(t, t+s) = m(t+s) - m(t)$ :

$$m(t, t+s) = m(t+s) - m(t) = \int_t^{t+s} \lambda(x) dx.$$

Es decir:

$$P(N(t+s) - N(t) = j) = e^{-m(t,t+s)} \cdot \frac{(m(t, t+s))^j}{j!}.$$

**Proposición 1.** Sea  $N(t)$  el número de eventos ocurridos hasta el tiempo  $t$  en un proceso de Poisson homogéneo con intensidad  $\lambda$ . Supongamos que en tiempo  $t$  un evento es contado con probabilidad  $p(t)$ , independientemente de lo ocurrido hasta ese instante. Entonces el proceso de conteo de estos eventos  $M(t)$  es un proceso de Poisson no homogéneo con intensidad  $\lambda(t) = \lambda \cdot p(t)$ .

En particular, si  $N(t)$  es un proceso de Poisson no homogéneo con función de intensidad  $\lambda(t)$  y  $\lambda \in \mathbb{R}$  es una constante tal que

$$\lambda(t) \leq \lambda,$$

para todo  $t$ , entonces  $N(t)$  puede verse como el proceso de contar eventos de un proceso de Poisson homogéneo con intensidad  $\lambda$  donde los eventos son contados con probabilidad

$$p(t) = \frac{\lambda(t)}{\lambda}.$$

### 3. Generación de Números Aleatorios

#### Método del Cuadrado Medio de von Neumann

Uno de los primeros trabajos que sugieren un método bien definido de generación de una secuencia determinística intentando imitar una secuencia aleatoria, fue de von Neumann. Este método conocido como *mid square* puede ser escrito de la siguiente manera:

1. Sea  $X_0$  un número entero de 4 dígitos decimales (puede ser que los dígitos más a la izquierda sean 0). Hacer  $i = 0$ .
2. Calcular  $X_i^2$ . El resultado es un número de 8 dígitos:  $X_i^2 = d_8d_7d_6d_5d_4d_3d_2d_1$ , con  $d_j$  en el conjunto  $\{0, 1, \dots, 9\}$ ,  $0 \leq j \leq 7$ .
3. Definir  $X_{i+1} = d_6d_5d_4d_3$ , esto es, los cuatro dígitos centrales o "middle".
4. Hacer  $i = i + 1$  y seguir en el paso 2.

```
def vonNeumann(u):
    u = (u**2 // 100) % 10000
    return u
```

Este método es un ejemplo de lo que llamaríamos un **mal generador de números aleatorios**. Veamos por qué.

Aunque ciertas secuencias obtenidas sean bien aleatorias y se repitan solo después de un número grande de términos, sucede que no se conocen bien las propiedades del generador. En particular, pareciera que las secuencias que se obtienen dependen fuertemente del valor inicial  $X_0$ , o semilla. Los siguientes ejemplos muestran algunas secuencias obtenidas a partir de este algoritmo, donde el primer valor es la semilla dada:

- a) **4010**: 801, 6416, 1650, 7225, 2006, 240, 576, 3317, 24, 5, 0, 0, ...
- b) **2100**: 4100, 8100, 6100, 2100, 4100, 8100, 6100, 2100, 4100, 8100, ...
- c) **3792**: 3792, 3792, 3792, 3792, 3792, 3792, ...
- d) **1234**: 5227, 3215, 3362, 3030, 1809, 2724, 4201, 6484, 422, 1780, ...

En (a) puede verse que la secuencia degenera en el valor 0 luego de algunas iteraciones, en (b) la secuencia alterna entre solo cuatro valores y en (c) los valores se repiten paso tras paso.

### 3.1. Generadores congruenciales

Es claro que con una computadora no es posible generar cualquier número real, en particular si posee infinitas cifras decimales. Los generadores que analizaremos a continuación producen en realidad una secuencia de números enteros:

$$y_1, y_2, \dots, y_N, \quad y_j \in \{0, 1, \dots, M-1\},$$

para un cierto entero positivo  $M$  "grande", y a partir de esta muestra se considera la sucesión de números en  $[0, 1)$  como:

$$u_1 = \frac{y_1}{M}, \quad u_2 = \frac{y_2}{M}, \quad \dots, \quad u_N = \frac{y_N}{M}, \quad \dots$$

Nos ocuparemos entonces de generar secuencias de números enteros mediante la relación de recurrencia:

$$y_n = f(y_{n-1}) \text{ mód } M,$$

para algún  $M$  entero positivo, comenzando de un valor inicial (semilla)  $y_0$ .

#### 3.1.1. Generadores congruenciales lineales

**Definición 1** (Generador Congruencial Lineal). Sea  $M$  un entero positivo,  $M \geq 2$ . Una sucesión  $y_1, y_2, \dots, y_n, \dots$  con valores en  $\{0, 1, \dots, M-1\}$  se dice **generada por el generador congruencial lineal** con parámetros  $a$ ,  $c$  y  $M$  y semilla  $y_0$  si

$$y_n = (ay_{n-1} + c) \text{ mód } M, \quad n \geq 1,$$

donde  $a$ ,  $c$  e  $y_0$  son enteros del conjunto  $\{0, \dots, M-1\}$ .

En la terminología usual:

- $a$  se dice **multiplicador**
- $c$  es el **incremento**
- $M$  es el **módulo**

Si  $c \neq 0$  el generador se dice **mixto** y si  $c = 0$  se dice **multiplicativo**.

```
def randMixto(a, c, M, u):  
    return (a * u + c) % M
```

```
def randMulti(a, M, u):  
    return (a * u) % M
```

Una secuencia generada por un generador congruencial lineal puede tener a lo sumo  $M$  números diferentes. Además, si un número de la secuencia se repite entonces también se repite la secuencia que sigue a ese número. Entonces, si  $K$  es el menor número tal que

$$y_n = y_{n+K}, \quad \text{para todo } n \geq N_0,$$

para algún  $N_0$ , diremos que  $K$  es el **período** de la secuencia  $y_0, y_1, \dots$ .

En principio podemos afirmar que  $K \leq M$ , es decir, el período no puede superar a la cantidad de números que es posible generar.

### 3.1.2. Generadores mixtos

**Teorema 3** (Condiciones para período máximo). Consideremos una secuencia dada por el generador:

$$y_{i+1} = ay_i + c \text{ mód } M, \quad c \neq 0.$$

Entonces la secuencia tiene período  $M$  si y sólo si se cumplen todas las siguientes condiciones:

1. El máximo común divisor entre  $c$  y  $M$  es 1:  $(c, M) = 1$ .
2.  $a \equiv 1 \text{ mód } p$ , para cualquier factor primo  $p$  de  $M$ .
3. Si 4 divide a  $M$ , entonces  $a \equiv 1 \text{ mód } 4$ .

**Corolario 1.** Si  $M$  es primo, el período máximo ocurre sólo si  $a = 1$ .

### 3.1.3. Generadores multiplicativos

**Definición 2** (Raíz primitiva). Sea  $M$  un número natural. Se dice que  $a$  es una **raíz primitiva** de  $M$  si

$$a^{(M-1)/p} \not\equiv 1 \pmod{M}$$

para cualquier factor primo  $p$  de  $M - 1$ .

**Teorema 4** (Propiedades del período en generadores multiplicativos). Para un generador multiplicativo

$$y_{i+1} = ay_i \text{ mód } M,$$

el período  $K$  de la secuencia verifica las siguientes tres propiedades:

1. Si  $K = M - 1$  entonces  $M$  es primo.
2. Si  $M$  es primo, entonces  $K$  divide a  $M - 1$ .
3.  $K = M - 1$  si y sólo si  $a$  es raíz primitiva de  $M$  y  $M$  es primo.

### 3.1.4. El problema de los hiperplanos

Para el caso de los generadores congruenciales lineales, ocurre lo siguiente. Si la secuencia producida es:

$$y_0, y_1, y_2, \dots, y_n, \dots,$$

está demostrado que los puntos  $(y_j, y_{j+1}, \dots, y_{j+k-1})$ ,  $j = 0, 1, 2, \dots$  están ubicados en no más de

$$(k! M)^{1/k} = (k!)^{1/k} \sqrt[k]{M}$$

hiperplanos paralelos en  $\mathbb{R}^k$ . ¿Qué nos dice esto? Que los puntos determinados por estas  $k$ -uplas quedan ubicados en una cantidad finita de hiperplanos, y por lo tanto existen "franjas" en las que no cae ninguna  $k$ -upla.

Un mal generador y ejemplo de este problema es **RANDU**. Este es un generador del tipo congruencial lineal multiplicativo ( $c = 0$ ), con  $M = 2^{31}$  y  $a = 2^{16} + 3$ :

$$y_n = 65539 \cdot y_{n-1} \text{ mód } 2147483648.$$

El problema fundamental es que estos generadores exhiben un *fenómeno de no-aleatoriedad*.

### 3.1.5. Generadores congruenciales lineales combinados

**Teorema 5** (Suma de variables uniformes discretas). Sean  $W_1, W_2, \dots, W_n$  variables aleatorias discretas independientes, tales que  $W_1 \sim U(\{0, d-1\})$  para cierto  $d \geq 1$ . Entonces

$$W = \left( \sum_{j=1}^n W_j \right) \text{ mód } d$$

es una variable aleatoria uniforme discreta en  $\{0, d-1\}$ .

## 4. Método de Monte Carlo

### 4.1. Fundamentos del Método de Monte Carlo

El método de Monte Carlo se basa en dos resultados fundamentales de la Teoría de la Probabilidad:

#### 1. Ley Fuerte de los Grandes Números

Si  $X_1, X_2, \dots, X_n, \dots$  son variables aleatorias independientes, idénticamente distribuidas, con media  $\theta$ , entonces:

$$P \left( \lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} = \theta \right) = 1,$$

o equivalentemente con probabilidad 1:

$$\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} = \theta.$$

#### 2. Valor Esperado de Funciones de Variables Aleatorias

Si  $X$  es una variable aleatoria absolutamente continua, con función de densidad  $f$ , y  $g : \mathbb{R} \mapsto \mathbb{R}$  es una función, entonces  $g \circ X$  es una variable aleatoria y su valor esperado está dado por:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x) dx.$$

### El método de Monte Carlo

- Se desea calcular  $\theta$ .
- $\theta = E[g(X)]$ . Que es difícil de calcular.
- Consideremos  $X$  una variable aleatoria con distribución conocida.

Tomamos una muestra  $X_1, X_2, \dots, X_N$  de variables independientes e idénticamente distribuidas (i.i.d.) con la misma distribución que  $X$ .

Entonces:

$$\frac{g(x_1) + g(x_2) + \dots + g(X_N)}{N} \approx E[g(X_i)] = E[g(X)] = \theta.$$

$\theta$  se estima con

$$[g(x_1) + g(x_2) + \dots + g(x_n)] \cdot \frac{1}{N}$$

## 4.2. Estimación de integrales definidas

### 4.2.1. Integración sobre $(0, 1)$

Queremos estimar:

$$\begin{aligned}\int_0^1 g(x) dx &= \theta \\ \theta &= \int_0^1 g(x) \cdot \frac{1}{\mathbb{I}_{(0,1)}(x)} dx \\ f(x) &= \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{c.c.} \end{cases} = \mathbb{I}_{[0,1]}(x)\end{aligned}$$

$f$  es la función de densidad de  $U \sim \mathcal{U}(0, 1)$ .

$$\theta = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx = E[g(u)]$$

Por Monte Carlo, tomamos una muestra  $U_1, U_2, \dots, U_N$  de variables aleatorias independientes, donde  $U_i \sim \mathcal{U}(0, 1)$  y entonces:

$$\frac{1}{N} \sum_{i=1}^N g(u_i) = \frac{g(u_1) + g(u_2) + \dots + g(u_N)}{N} \approx \theta = \int_0^1 g(x) dx.$$

### 4.2.2. Integración sobre $[a, b]$

Queremos calcular:

$$\theta = \int_a^b g(x) dx$$

Hacemos el cambio de variable:

$$\begin{aligned}y &= \frac{x-a}{b-a} \Rightarrow x = y(b-a) + a \\ dy &= \frac{1}{b-a} dx \Rightarrow dx = (b-a) dy\end{aligned}$$

Transformando la integral:

$$\int_a^b g(x) dx = \int_0^1 \underbrace{g((b-a)y + a)(b-a)}_{h(y)} dy = \int_0^1 h(y) dy$$

Luego con Monte Carlo generamos  $N$  variables aleatorias uniformes:

$$U_1, U_2, \dots, U_N \sim \mathcal{U}(0, 1)$$

Entonces tenemos:

$$\begin{aligned}\frac{h(u_1) + h(u_2) + \dots + h(u_N)}{N} &= \frac{1}{N} \sum_{i=1}^N g((b-a)u_i + a)(b-a) \\ &= \frac{b-a}{N} \sum_{i=1}^N g((b-a)u_i + a) \\ &\approx \theta \\ &= \int_a^b g(x) dx\end{aligned}$$

### 4.2.3. Integración sobre $(0, \infty)$

Consideremos la integral:

$$\theta = \int_0^{\infty} g(x) dx$$

Realizamos un cambio de variable biyectivo:

$$(0, \infty) \xrightarrow{\text{biyectiva}} (0, 1)$$

Definimos la transformación:

$$y = \frac{1}{x+1} \Rightarrow x = \frac{1}{y} - 1$$

El diferencial se transforma como:

$$dy = -\frac{1}{(x+1)^2} dx \Rightarrow dx = -\frac{1}{y^2} dy$$

Transformando la integral:

$$\int_0^{\infty} g(x) dx = \int_1^0 g\left(\frac{1}{y} - 1\right) \cdot -\frac{1}{y^2} dy = \int_0^1 \underbrace{g\left(\frac{1}{y} - 1\right) \cdot \frac{1}{y^2}}_{h(y)} dy = \theta$$

Notar que usamos el hecho de que si  $a < b$  y tenemos  $\int_b^a f(x) dx$  entonces:

$$\int_b^a f(x) dx = -\int_a^b f(x) dx$$

Luego generamos  $U_1, U_2, \dots, U_N$  variables aleatorias independientes, donde  $U_i \sim \mathcal{U}(0, 1)$  y entonces:

$$\frac{h(u_1) + h(u_2) + \dots + h(u_N)}{N} = \frac{1}{N} \sum_{i=1}^N g\left(\frac{1}{u_i} - 1\right) \cdot \frac{1}{u_i^2} \approx \theta = \int_0^{\infty} g(x) dx.$$

### 4.2.4. Integración sobre $-(\infty), 0$

Queremos calcular la integral impropia:

$$\theta = \int_{-\infty}^0 g(x) dx$$

Para ello hacemos un cambio de variable:

$$y = -x \Rightarrow x = -y$$

$$dy = -dx \Rightarrow dx = -dy$$

Entonces transformamos la integral:

$$\int_{-\infty}^0 g(x) dx = \int_{\infty}^0 g(-y) \cdot (-dy) = \int_0^{\infty} g(-y) dy = \theta$$

Luego generamos  $U_1, U_2, \dots, U_N$  variables aleatorias independientes, donde  $U_i \sim \mathcal{U}(0, 1)$  y entonces:

$$\theta \approx \frac{1}{N} \sum_{i=1}^N g(-U_i) \cdot (-1) = -\frac{1}{N} \sum_{i=1}^N g(-U_i) \approx \int_{-\infty}^0 g(x) dx$$

#### 4.2.5. Integración sobre $(a, \infty)$ con transformación racional

Consideremos la integral impropia:

$$\theta = \int_a^\infty g(x) dx \quad \text{con } a > 0$$

Realizamos un cambio de variable racional:

$$y = \frac{a}{x} \quad \Rightarrow \quad x = \frac{a}{y}$$

El diferencial se transforma como:

$$dy = -\frac{a}{x^2} dx \quad \Rightarrow \quad dx = -\frac{a}{y^2} dy$$

Transformando la integral:

$$\int_a^\infty g(x) dx = \int_1^0 g\left(\frac{a}{y}\right) \cdot \left(-\frac{a}{y^2}\right) dy = \int_0^1 \underbrace{g\left(\frac{a}{y}\right) \cdot \frac{a}{y^2}}_{h(y)} dy = \theta$$

Donde usamos que:

$$\int_b^a f(x) dx = - \int_a^b f(x) dx \quad \text{si } a < b$$

**Método de Monte Carlo.** Sea  $U_i \sim \mathcal{U}(0, 1)$  para  $i = 1, \dots, N$ . Entonces, una estimación de  $\theta$  es:

$$\theta \approx \frac{1}{N} \sum_{i=1}^N g\left(\frac{a}{U_i}\right) \cdot \frac{a}{U_i^2}$$

**Resumen de las Transformaciones Clave:**

Límite inferior	Cambio de variable	Jacobiano	Nuevos límites
$a = 1$	$y = \frac{1}{x}$	$\frac{1}{y^2}$	$y \in (0, 1]$
$a > 0$ (general)	$y = \frac{a}{x}$	$\frac{a}{y^2}$	$y \in (0, 1]$

#### 4.2.6. Integración sobre $(-\infty, b]$

Queremos calcular la integral impropia:

$$\theta = \int_{-\infty}^b g(x) dx$$

Realizamos un cambio de variable racional:

$$y = \frac{1}{b-x} \quad \Rightarrow \quad x = b - \frac{1}{y}$$

El diferencial se transforma como:

$$dy = \frac{1}{(b-x)^2} dx \quad \Rightarrow \quad dx = \frac{1}{y^2} dy$$

Transformando la integral:

$$\int_{-\infty}^b g(x) dx = \int_0^1 g\left(b - \frac{1}{y}\right) \cdot \frac{1}{y^2} dy = \theta$$



**Método de Monte Carlo.** Sea  $U_i \sim \mathcal{U}(0, 1)$  para  $i = 1, \dots, N$ . Entonces:

$$\theta \approx \frac{1}{N} \sum_{i=1}^N g\left(b - \frac{1}{U_i}\right) \cdot \frac{1}{U_i^2}$$

#### 4.2.7. Integración sobre $(-\infty, \infty)$

Queremos evaluar:

$$\theta = \int_{-\infty}^{\infty} g(x) dx$$

Dividimos la integral en dos partes:

$$\theta = \int_{-\infty}^0 g(x) dx + \int_0^{\infty} g(x) dx$$

Para cada parte aplicamos un cambio de variable racional similar:

- Para  $\int_0^{\infty} g(x) dx$ : usamos  $y = \frac{1}{x+1} \Rightarrow x = \frac{1}{y} - 1$ , con  $dx = -\frac{1}{y^2} dy$

$$\int_0^{\infty} g(x) dx = \int_0^1 g\left(\frac{1}{y} - 1\right) \cdot \frac{1}{y^2} dy$$

- Para  $\int_{-\infty}^0 g(x) dx$ : usamos  $y = \frac{1}{1-x} \Rightarrow x = 1 - \frac{1}{y}$ , con  $dx = \frac{1}{y^2} dy$

$$\int_{-\infty}^0 g(x) dx = \int_0^1 g\left(1 - \frac{1}{y}\right) \cdot \frac{1}{y^2} dy$$

Entonces:

$$\theta = \int_0^1 \left[ g\left(\frac{1}{y} - 1\right) + g\left(1 - \frac{1}{y}\right) \right] \cdot \frac{1}{y^2} dy$$

**Método de Monte Carlo.** Sea  $U_i \sim \mathcal{U}(0, 1)$  para  $i = 1, \dots, N$ . Entonces:

$$\theta \approx \frac{1}{N} \sum_{i=1}^N \left[ g\left(\frac{1}{U_i} - 1\right) + g\left(1 - \frac{1}{U_i}\right) \right] \cdot \frac{1}{U_i^2}$$

#### Aclaración sobre el cambio de variables

Cuando realizamos un cambio de variable, la idea es buscar una función que mapee los límites de integración, a los límites de integración 0 y 1.

Es decir el "y", tiene que tender a 0 y 1, cuando el  $x$  tiende a los límites de integración.

Por ejemplo:

$$\int_1^{\infty} g(x) dx$$

Hacemos el cambio de variable:

$$y = \frac{1}{x} \quad \Rightarrow \quad x = \frac{1}{y}$$

$$dy = -\frac{1}{x^2}dx \quad \Rightarrow \quad dx = -\frac{1}{y^2}dy$$

Entonces  $y \rightarrow 0$  cuando  $x \rightarrow \infty$  y  $y \rightarrow 1$  cuando  $x \rightarrow 1$ . Luego tenemos:

$$\int_1^\infty g(x) dx = \int_1^0 g\left(\frac{1}{y}\right) \cdot \left(-\frac{1}{y^2}\right) dy = \int_0^1 g\left(\frac{1}{y}\right) \cdot \frac{1}{y^2} dy$$

Otro caso por ejemplo es si tenemos:

$$\int_{-1}^\infty g(x) dx$$

Aquí tenemos dos opciones, o bien hacemos el cambio de variable:

$$y = x + 1 \quad \Rightarrow \quad x = y - 1$$

$$dy = dx \quad \Rightarrow \quad dx = dy$$

Entonces  $y \rightarrow 0$  cuando  $x \rightarrow -1$  y  $y \rightarrow \infty$  cuando  $x \rightarrow \infty$ . Luego tenemos:

$$\int_{-1}^\infty g(x) dx = \int_0^\infty g(y-1) dy$$

La otra opción es separar la integral en dos:

$$\int_{-1}^0 g(x) dx + \int_0^\infty g(x) dx$$

Y estas integrales sabemos aproximarlas por Monte Carlo.

### 4.3. Estimación de integrales múltiples

Para calcular la cantidad:

$$\theta = \int_0^1 \cdots \int_0^1 g(x_1, \dots, x_l) dx_1 \dots dx_l$$

utilizamos el hecho que:

$$\theta = E[g(U_1, \dots, U_l)]$$

con  $U_1, \dots, U_l$  independientes y uniformes en  $(0, 1)$ . Esto es así porque su distribución conjunta está dada por:

$$f(x_1, x_2, \dots, x_l) = \mathbb{I}_{(0,1)^l}(x_1, x_2, \dots, x_l),$$

donde  $\mathbb{I}_{(0,1)^l}$  es la función indicadora del hipercubo unitario  $l$ -dimensional.

Entonces:

$$\theta = \int_{\mathbb{R}^l} g(x_1, \dots, x_l) f(x_1, \dots, x_l) dx_1 \dots dx_l$$

Si se tienen  $N$  muestras independientes de estas  $l$  variables:

$$(U_1^{(1)}, \dots, U_l^{(1)}), \quad (U_1^{(2)}, \dots, U_l^{(2)}), \quad \dots, \quad (U_1^{(N)}, \dots, U_l^{(N)})$$

podemos estimar el valor de  $\theta$  como:

$$\theta \approx \frac{1}{N} \sum_{i=1}^N g(U_1^{(i)}, \dots, U_l^{(i)})$$