

# Distribuciones de probabilidad

Fernando González

April 25, 2025

## 1 Teoría

### 1.1 Variable aleatoria

Una **variable aleatoria** es una función que asigna valores numéricos a los resultados de un experimento aleatorio. Formalmente, es una aplicación  $X : \Omega \rightarrow \mathbb{R}$  donde  $\Omega$  es el espacio muestral. La importancia de este concepto radica en que nos permite cuantificar resultados cualitativos y aplicar herramientas matemáticas para su análisis. Se clasifican en:

- **Discretas:** Aquellas cuyo conjunto de valores posibles es finito o infinito numerable. Por ejemplo, el número de caras al lanzar una moneda 10 veces (valores: 0,1,...,10).
- **Continuas:** Variables que pueden tomar cualquier valor en un intervalo real. Por ejemplo, el tiempo de vida de un componente electrónico. En este curso nos enfocaremos principalmente en las discretas.

### 1.2 Distribución de probabilidad

La **función de distribución** (o función de masa de probabilidad para variables discretas) describe completamente el comportamiento probabilístico de una variable aleatoria. Para una v.a. discreta  $X$ , se define como:

$$P(X = x) = p(x)$$

Esta función debe satisfacer dos condiciones fundamentales:

1. **No-negatividad:**  $0 \leq p(x) \leq 1$  para todo valor posible  $x$  de  $X$ . Esto garantiza que cada probabilidad individual sea válida.
2. **Normalización:**  $\sum_x p(x) = 1$ . La suma de todas las probabilidades debe ser 1, lo que refleja que algún valor debe ocurrir con certeza.

### 1.3 Notación

Es crucial entender la notación probabilística:

- $P(X = m)$ : Probabilidad puntual, la chance exacta de que  $X$  tome el valor específico  $m$ .
- $P(X < m)$ : Probabilidad acumulada hasta (pero sin incluir)  $m$ . Para discretas, equivale a  $\sum_{k < m} P(X = k)$ .
- $P(a \leq X \leq b)$ : Probabilidad de que  $X$  caiga en el intervalo cerrado  $[a, b]$ . Para discretas, se calcula sumando  $P(X = k)$  para todos los  $k$  enteros entre  $a$  y  $b$ .

## 1.4 Función de distribución acumulada (CDF)

La **función de distribución acumulada**  $F(x) = P(X \leq x)$  proporciona la probabilidad de que  $X$  sea menor o igual a un valor dado  $x$ . Para v.a. discretas:

$$F(x) = \sum_{k \leq x} P(X = k)$$

Propiedades clave:

- Es no decreciente: si  $x_1 < x_2$  entonces  $F(x_1) \leq F(x_2)$
- Comportamiento en extremos:  $\lim_{x \rightarrow -\infty} F(x) = 0$  y  $\lim_{x \rightarrow \infty} F(x) = 1$
- Para discretas, es una función escalonada con saltos en los valores posibles de  $X$

## 1.5 Esperanza

La **esperanza matemática** o valor esperado  $E[X]$  representa el promedio teórico de  $X$  a largo plazo. Para v.a. discretas:

$$E[X] = \sum_x x \cdot p(x)$$

Interpretación intuitiva: Si repitiéramos el experimento infinitas veces,  $E[X]$  sería el promedio de los valores observados. Es un operador lineal que "baricentra" la distribución de probabilidad.

## 1.6 Propiedades de la esperanza

- **Linealidad:**  $E[aX + b] = aE[X] + b$  para constantes  $a, b$ . Esto permite manipular fácilmente transformaciones lineales.
- **Aditividad:**  $E[X + Y] = E[X] + E[Y]$ . El valor esperado de una suma es la suma de valores esperados, sin requerir independencia.
- **Independencia:** Si  $X$  e  $Y$  son independientes, entonces  $E[XY] = E[X]E[Y]$ . Esta propiedad falla cuando hay dependencia entre variables.

## 1.7 Varianza y desviación estándar

La **varianza** mide la dispersión de  $X$  alrededor de su media:

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

La **desviación estándar**  $\sigma_X = \sqrt{\text{Var}(X)}$  tiene las mismas unidades que  $X$  y proporciona una medida de variabilidad directamente interpretable.

## 1.8 Propiedades de la varianza

- **Escalado:**  $\text{Var}(aX + b) = a^2 \text{Var}(X)$ . Notar que las traslaciones (suma de  $b$ ) no afectan la varianza.
- **Independencia:** Si  $X$  e  $Y$  son independientes,  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ . La varianza es aditiva para variables independientes.

## 1.9 Distribución de Bernoulli

Modela experimentos con exactamente dos resultados posibles (éxito/fracaso). Su función de distribución es:

$$P(X = k) = \begin{cases} p & \text{si } k = 1 \text{ (éxito)} \\ 1 - p & \text{si } k = 0 \text{ (fracaso)} \end{cases}$$

Características:

- Media:  $E[X] = p$  (la probabilidad de éxito)
- Varianza:  $\text{Var}(X) = p(1 - p)$  (máxima cuando  $p = 0.5$ )

## 1.10 Distribución binomial

Describe el número de éxitos en  $n$  ensayos Bernoulli independientes con igual probabilidad  $p$ . Su función de distribución es:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n$$

Propiedades:

- Media:  $E[X] = np$  (intuitivo: número esperado de éxitos)
- Varianza:  $\text{Var}(X) = np(1 - p)$
- Surge como suma de  $n$  Bernoullis independientes

## 1.11 Distribución geométrica

Modela el número de ensayos necesarios para obtener el primer éxito en una secuencia de Bernoullis independientes. Su función de distribución es:

$$P(X = k) = (1 - p)^{k-1} p, \quad k \in \mathbb{N}$$

Propiedades notables:

- Media:  $E[X] = \frac{1}{p}$  (inversa de la probabilidad de éxito)
- Varianza:  $\text{Var}(X) = \frac{1-p}{p^2}$
- Tiene la propiedad de "pérdida de memoria":  $P(X > m + n | X > m) = P(X > n)$

## 1.12 Distribución hipergeométrica

La distribución hipergeométrica modela el número de éxitos en una secuencia de  $n$  extracciones **sin reemplazo** de una población finita de tamaño  $N$  que contiene  $K$  éxitos y  $N - K$  fracasos. A diferencia de la distribución binomial, donde los ensayos son independientes, en la distribución hipergeométrica cada extracción afecta la probabilidad de los resultados posteriores.

### 1.12.1 Definición formal

Sea  $X$  la variable aleatoria que representa el número de éxitos obtenidos en  $n$  extracciones sin reemplazo de una población de  $N$  elementos, de los cuales  $K$  son considerados éxitos y  $N - K$  son considerados fracasos. Entonces, la probabilidad de obtener exactamente  $x$  éxitos viene dada por la siguiente fórmula:

$$P(X = x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}$$

donde:

- $N$ : tamaño de la población
- $K$ : número de éxitos en la población
- $n$ : número de extracciones (tamaño de la muestra)
- $x$ : número de éxitos observados en la muestra ( $0 \leq x \leq \min(n, K)$ )
- $\binom{a}{b}$ : coeficiente binomial, que representa el número de formas de elegir  $b$  elementos de un conjunto de  $a$  elementos.

Además, se tienen los siguientes datos:

- $E[X] = n \frac{K}{N}$
- $\text{Var}(X) = n \frac{K}{N} \left(1 - \frac{K}{N}\right) \frac{N-n}{N-1}$

### 1.12.2 Deducción de la fórmula

Podemos entender la fórmula de la siguiente manera:

- $\binom{K}{x}$ : representa el número de formas de seleccionar  $x$  éxitos de los  $K$  éxitos disponibles en la población.
- $\binom{N-K}{n-x}$ : representa el número de formas de seleccionar  $n-x$  fracasos de los  $N-K$  fracasos disponibles en la población.
- El producto  $\binom{K}{x} \binom{N-K}{n-x}$  da el número total de formas de obtener exactamente  $x$  éxitos y  $n-x$  fracasos en la muestra de tamaño  $n$ .
- $\binom{N}{n}$ : representa el número total de formas posibles de seleccionar una muestra de tamaño  $n$  de la población de tamaño  $N$ .

Por lo tanto, la probabilidad de obtener exactamente  $x$  éxitos es la razón entre el número de resultados favorables (seleccionar  $x$  éxitos y  $n-x$  fracasos) y el número total de resultados posibles (seleccionar cualquier muestra de tamaño  $n$ ).

### 1.12.3 Ejemplo práctico

#### 1.12.4 Relación con la distribución binomial

- **Aproximación cuando  $N \gg n$ :** Cuando el tamaño de la población ( $N$ ) es mucho mayor que el tamaño de la muestra ( $n$ ), la diferencia entre muestrear con y sin reemplazo se vuelve insignificante. En este caso, la distribución hipergeométrica puede aproximarse mediante una distribución binomial con probabilidad de éxito  $p = K/N$ . Intuitivamente, si la muestra es pequeña en comparación con la población, la eliminación de unos pocos elementos no altera significativamente la proporción de éxitos en la población restante.
- **Naturaleza de los ensayos:** Ambas distribuciones modelan el conteo de éxitos en una serie de ensayos. Sin embargo, la clave de la distinción radica en la independencia de los ensayos. La distribución binomial asume que cada ensayo es independiente, con la misma probabilidad de éxito en cada uno. En contraste, la distribución hipergeométrica modela situaciones donde los ensayos son dependientes debido al muestreo sin reemplazo, lo que hace que la probabilidad de éxito cambie en cada extracción.

## 2 Ejemplos

1. **Binomial:** Se lanza un dado justo 10 veces. La probabilidad de obtener exactamente 3 seises es:

$$P(X = 3) = \binom{10}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^7 \approx 0.155$$

2. **Geométrica:** En una línea de producción, un artículo defectuoso aparece con probabilidad 0.01. La probabilidad de que el primer defectuoso sea el décimo artículo es:

$$P(X = 10) = (0.99)^9(0.01) \approx 0.0091$$

3. **Hipergeométrica:** En un lote de 50 productos, de los cuales 10 son defectuosos, se inspeccionan 15 productos al azar sin reemplazo. ¿Cuál es la probabilidad de encontrar exactamente 3 productos defectuosos?

Aquí, tenemos:

- $N = 50$  (tamaño del lote)
- $K = 10$  (número de productos defectuosos)
- $n = 15$  (tamaño de la muestra)
- $x = 3$  (número de productos defectuosos que queremos encontrar)

Aplicando la fórmula de la distribución hipergeométrica:

$$P(X = 3) = \frac{\binom{10}{3} \binom{50-10}{15-3}}{\binom{50}{15}} = \frac{\binom{10}{3} \binom{40}{12}}{\binom{50}{15}}$$

Calculando los coeficientes binomiales:

$$\binom{10}{3} = \frac{10!}{3!(10-3)!} = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120$$

$$\binom{40}{12} = \frac{40!}{12!(40-12)!} = \frac{40!}{12!28!} = 5,586,853,480$$

$$\binom{50}{15} = \frac{50!}{15!(50-15)!} = \frac{50!}{15!35!} = 30,050,029,200$$

Sustituyendo estos valores en la fórmula:

$$P(X=3) = \frac{120 \times 5,586,853,480}{30,050,029,200} \approx \frac{670,422,417,600}{30,050,029,200} \approx 0.0223$$

Por lo tanto, la probabilidad de encontrar exactamente 3 productos defectuosos en la muestra es aproximadamente 0.0223 o 2.23%.

4. **Esperanza y varianza:** Si  $X \sim \text{Binomial}(n=5, p=0.4)$ , entonces:

$$E[X] = 5 \times 0.4 = 2, \quad \text{Var}(X) = 5 \times 0.4 \times 0.6 = 1.2$$

### 3 Ejercicios

1. Sea  $X \sim \text{Binomial}(n=8, p=0.3)$ . Calcula  $P(2 \leq X < 5)$ .
2. En una distribución geométrica con  $p=0.2$ , encuentra el mínimo  $k$  tal que  $P(X \leq k) \geq 0.75$ .
3. Demuestra que para  $X \sim \text{Bernoulli}(p)$ ,  $\text{Var}(X) = p(1-p)$  usando la definición  $\text{Var}(X) = E[X^2] - (E[X])^2$ .
4. Si  $X$  e  $Y$  son independientes con  $E[X] = 3$ ,  $\text{Var}(X) = 2$ ,  $E[Y] = -1$ ,  $\text{Var}(Y) = 4$ , calcula  $E[2X - Y + 1]$  y  $\text{Var}(2X - Y + 1)$ .
5. Una v.a. continua tiene PDF  $f(x) = kx^2$  para  $0 \leq x \leq 2$  y 0 en otro caso. Determina  $k$  y calcula  $P(0.5 \leq X \leq 1.5)$ .

## 4 Profundización: repaso de sumas y series

### 4.1 Propiedades fundamentales de sumatorias

#### 4.1.1 Linealidad y descomposición

Las sumatorias cumplen propiedades algebraicas similares a las integrales. Dos propiedades esenciales son:

- **Linealidad:** La suma es compatible con la suma de términos y con la multiplicación por constantes.

$$\sum_{k=1}^n (a_k + b_k) = \sum_{k=1}^n a_k + \sum_{k=1}^n b_k \quad \text{y} \quad \sum_{k=1}^n c \cdot a_k = c \cdot \sum_{k=1}^n a_k$$

Esto nos permite "romper" sumas complejas en partes más simples.

- **Descomposición:** Podemos dividir una suma en segmentos:

$$\sum_{k=m}^n a_k = \sum_{k=1}^n a_k - \sum_{k=1}^{m-1} a_k \quad \text{para } m > 1$$

Útil cuando queremos calcular sumas parciales.

## 4.2 Suma de Gauss (suma de los primeros $n$ naturales)

La famosa fórmula descubierta por Carl Friedrich Gauss de niño:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}$$

### 4.2.1 Demostración visual

1. Escribe la suma dos veces, una en orden normal y otra en inverso:

$$S = 1 + 2 + \cdots + (n-1) + n$$

$$S = n + (n-1) + \cdots + 2 + 1$$

2. Suma verticalmente:

$$2S = \underbrace{(n+1) + (n+1) + \cdots + (n+1)}_{n \text{ veces}} = n(n+1)$$

3. Despeja  $S$  para obtener el resultado.

## 4.3 Series geométricas

### 4.3.1 Suma finita

Para  $r \neq 1$ , la suma de los primeros  $n+1$  términos es:

$$\sum_{k=0}^n r^k = \frac{1 - r^{n+1}}{1 - r}$$

### 4.3.2 Demostración

1. Sea  $S_n = 1 + r + r^2 + \cdots + r^n$ .

2. Multiplica ambos lados por  $r$ :

$$rS_n = r + r^2 + \cdots + r^{n+1}$$

3. Resta esta ecuación de la original:

$$S_n - rS_n = 1 - r^{n+1} \quad \Rightarrow \quad S_n(1 - r) = 1 - r^{n+1}$$

4. Despeja  $S_n$  (observa que  $1 - r \neq 0$ ):

$$S_n = \frac{1 - r^{n+1}}{1 - r}$$

## 4.4 Serie geométrica infinita

Cuando  $|r| < 1$ , la serie converge a:

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}$$

### 4.4.1 Explicación

Al tomar  $n \rightarrow \infty$  en la suma finita,  $r^{n+1} \rightarrow 0$  porque  $|r| < 1$ , dejando solo  $\frac{1}{1-r}$ .

## 4.5 Series telescópicas

$$\sum_{k=1}^n (a_{k+1} - a_k) = a_{n+1} - a_1$$

### 4.5.1 Ejemplo ilustrativo

Considera:

$$\sum_{k=1}^n \frac{1}{k(k+1)} = \sum_{k=1}^n \left( \underbrace{\frac{1}{k}}_{a_k} - \underbrace{\frac{1}{k+1}}_{a_{k+1}} \right) = 1 - \frac{1}{n+1}$$

La mayoría de términos se cancelan ("colapsan como un telescopio"), dejando solo el primero y el último.

## 4.6 Binomio de Newton

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$$

donde  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$  son coeficientes binomiales.

### 4.6.1 Interpretación

Cada término representa una forma de elegir  $k$  factores  $b$  y  $n-k$  factores  $a$  al expandir el producto.

## 4.7 Series de Taylor y Maclaurin

### 4.7.1 Origen de la fórmula

La serie de Taylor generaliza el polinomio de Taylor. Surge de querer aproximar funciones arbitrarias mediante polinomios cuyas derivadas coincidan con las de la función en un punto.

1. Supongamos que  $f(x)$  puede escribirse como:

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

2. Calcula derivadas en  $x=a$ :

$$\begin{aligned} f(a) &= c_0 \\ f'(a) &= c_1 \\ f''(a) &= 2!c_2 \\ &\vdots \\ f^{(n)}(a) &= n!c_n \end{aligned}$$

3. Despeja los coeficientes  $c_n = \frac{f^{(n)}(a)}{n!}$ .

## 4.8 Serie de Maclaurin (caso $a=0$ )

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$



### 4.8.1 Ejemplo: $e^x$

Como  $\frac{d^n}{dx^n}e^x = e^x$  y  $e^0 = 1$ :

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

## 4.9 Otras series esenciales

- **Cuadrados:**

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

- **Cubos** (notablemente, el cuadrado de la suma de Gauss):

$$\sum_{k=1}^n k^3 = \left( \frac{n(n+1)}{2} \right)^2$$

- **Armónica** (divergente pero lentamente):

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

- **Armónica alternante** (convergente):

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln(2)$$

## 5 Profundización: De Bernoulli a nuevas distribuciones

Recordemos que en la **binomial** ( $X \sim \text{Bin}(n, p)$ ) contábamos éxitos en  $n$  ensayos *independientes* con probabilidad  $p$ . Pero, ¿qué pasa cuando la independencia no se cumple?

### 5.1 Distribución hipergeométrica: Binomial sin reemplazo

#### 5.1.1 Contexto

Imagina que tienes:

- Una urna con  $N$  bolas ( $K$  éxitos,  $N - K$  fracasos)
- Extraes  $n$  bolas **sin reemplazo** (la probabilidad cambia en cada extracción)

$X$ : Número de éxitos obtenidos. Entonces  $X \sim \text{HGeo}(N, K, n)$ .

#### 5.1.2 Ejemplo práctico

En un lote de 50 productos (10 defectuosos), inspeccionas 15 al azar. La probabilidad de encontrar exactamente 3 defectuosos es:

$$P(X = 3) = \frac{\binom{10}{3} \binom{40}{12}}{\binom{50}{15}}$$

### 5.1.3 Relación con la binomial

- Si  $N \gg n$ , la hipergeométrica se aproxima a binomial (pues el no reemplazo afecta poco)
- Ambas modelan conteos de éxitos, pero la hipergeométrica considera dependencia entre ensayos

## 5.2 Distribución de Poisson: Binomial con eventos raros

### 5.2.1 Contexto

Surge al contar eventos en un intervalo continuo (tiempo, área) cuando:

- Los eventos son independientes
- Ocurren con tasa promedio  $\lambda$  conocida
- La probabilidad por intervalo pequeño es proporcional a su tamaño

$Y$ : Número de eventos en el intervalo. Entonces  $Y \sim \text{Poisson}(\lambda)$ .

### 5.2.2 Ejemplo práctico

Si un call center recibe en promedio 5 llamadas por hora, la probabilidad de recibir exactamente 2 llamadas en una hora es:

$$P(Y = 2) = \frac{e^{-5} \cdot 5^2}{2!}$$

### 5.2.3 Relación con la binomial

- Es límite de  $\text{Bin}(n, p)$  cuando  $n \rightarrow \infty$ ,  $p \rightarrow 0$  con  $np = \lambda$  constante
- Útil cuando  $n \geq 100$  y  $p \leq 0.01$  (regla práctica)

## 5.3 ¿Cuándo usar cada una?

Distribución	Contexto clave	Dependencia
Binomial	Ensayos independientes con reemplazo	$p$ constante
Hipergeométrica	Ensayos <b>dependientes</b> sin reemplazo	$p$ variable
Poisson	Eventos raros en intervalo continuo	Tasa $\lambda$ fija

## Anexo: pregunta generadora

[[Yo]]: Hola, soy estudiante; no tengo problema con explicaciones técnicas y pruebas, al igual que explicaciones intuitivas.

Genera un resumen del contenido de mi clase de Probabilidades en un archivo .tex.

Formato: letterpaper, márgenes de 1.27 cm

Título: Distribuciones de probabilidad

Autor: Fernando González

Sección: Teoría

Explicar los siguientes temas sin ejemplos. Que sea el intermedio entre breve y detallado. Busca que la fórmula tenga algún desarrollo breve para llegar a ella, y de esta u otra

manera poder a acercarme a cómo y por qué funciona esa fórmula.

- Definición de variable aleatoria (discreta y continua), con ejemplos
- Definición de una distribución de probabilidad
- Condiciones necesarias que cumple una distribución de probabilidad
- Notación  $P(X = m)$ ,  $P(x < m)$ , etc.
- Distribución acumulada de probabilidad, con ejemplos
- Definición de esperanza
- Propiedades de la esperanza
- Definición de varianza y desviación estándar
- Propiedades de la varianza
- Distribución de Bernoulli: cuándo sucede, su media y su varianza
- Distribución binomial: cuándo sucede, su media y su varianza
- Distribución geométrica: cuándo sucede, su media y su varianza

Sección: Ejemplos

- 3 ejemplos de complejidad media que abarquen los temas de arriba. Explicados (no necesitan ser paso a paso) y con respuesta

Sección: Ejercicios

- 3 a 5 ejercicios de complejidad media-alta, sin respuesta.

[[IA]]: Respuesta 1

[[Yo]]: Estoy estudiando Probabilidades. Llegué al tema de introducción a las distribuciones de probabilidad, por lo que necesitaré un repaso de sumatorias y series importantes. Genérame un snippet de LaTeX que contenga estas diferentes series que debería saber, junto con posiblemente sus fórmulas explícitas, y tal vez pruebas de estas fórmulas (las que ayuden a comprender cómo se llega a esa forma). Incluye mínimo las siguientes:

- Propiedades de la sumas (suma y resta de sumatorias, constante por sumatoria, etc.)
- Suma de Gauss
- Propiedades de las series
- Suma y serie geométrica
- Suma y serie geométrica
- Suma y serie telescópica
- Binomio de Newton
- Serie de Taylor y Maclaurin
- Caso particular de  $e^x$
- Algunas otras necesarias

[[IA]]: Respuesta 2

[[Yo]]: Ahora quiero un snippet diferente. Ya vi las distribuciones de experimentos de Bernoulli, junto con la distribución binomial y la geométrica. Entonces hazme un acercamiento a las siguientes distribuciones que veremos: la hipergeométrica y la de Poisson. Posiblemente usando los conceptos aprendidos de las distribuciones anteriores, con un experimento que permita ver la aplicación de esa distribución. No necesitas ahondar mucho, porque pronto me lo explicarán: es un acercamiento.

[[IA]]: Respuesta 3