

Apuntes

Nicolas Muñoz

Estadística
Licenciatura en Matemática
Pontificia Universidad Católica - Chile

2 de septiembre de 2025

Índice

1. Clase 1: Razonamiento Estadístico.	2
1.1. La Historia que dio Origen al Diseño Experimental	2
1.2. El Experimento de Fisher	2
1.3. Análisis Probabilístico y Toma de Decisión	2
1.4. Pasos Fundamentales del Razonamiento Estadístico	3
2. Clase 2: Conceptos Básicos	3
3. Clase 3: Definiciones de Probabilidad	4
4. Clase 4: Probabilidad Condicional e Independencia	4
5. Clase 5: Variables Aleatorias	5
6. Clase 6: Variables Aleatorias Discretas	6
7. Clase 7: Distribuciones Discretas	6

1. Clase 1: Razonamiento Estadístico.

1.1. La Historia que dio Origen al Diseño Experimental

La primera clase introduce el concepto de **razonamiento estadístico** a través de una historia real que ocurrió en Cambridge, Inglaterra, a finales de la década de 1920. En una reunión social, una dama afirmó que el sabor del té era diferente dependiendo de si se había vertido primero la leche en la taza o el té.

El eminente estadístico, Sir Ronald Aylmer Fisher, en lugar de desestimar la afirmación, propuso evaluarla a través de un experimento diseñado para obtener evidencia objetiva. Este evento es considerado el origen del Diseño Experimental moderno.

1.2. El Experimento de Fisher

Para verificar la afirmación de la dama, Fisher diseñó un experimento con las siguientes características:

- **Preparación:** Se prepararon 8 tazas de té en total.
- **Grupos:** En 4 de las tazas se sirvió el té antes que la leche, y en las 4 restantes, la leche antes que el té.
- **Aleatorización:** Las 8 tazas se le presentaron a la dama en un orden completamente aleatorio para que ella las probara y clasificara.
- **Información previa:** A la dama se le informó que había exactamente 4 tazas de cada tipo de preparación.

La pregunta central del experimento era: ¿qué resultados nos harían concluir que la dama realmente tiene la habilidad que dice tener y no está simplemente adivinando?.

1.3. Análisis Probabilístico y Toma de Decisión

El núcleo del razonamiento de Fisher fue el uso de la probabilidad para determinar qué tan sorprendente sería un resultado si la dama estuviera adivinando. La idea es evaluar la probabilidad de que un resultado ocurra **por puro azar**.

- **Hipótesis Nula (H_0):** La dama no tiene ninguna habilidad para distinguir la preparación. Cualquier acierto es producto del azar.
- **Hipótesis Alternativa (H_A):** La dama sí tiene la habilidad de distinguir la preparación.

Fisher calculó la probabilidad de los posibles resultados bajo el supuesto de que la hipótesis nula era cierta:

- La probabilidad de acertar **todas las tazas** (las 4 de un tipo) por azar es de solo **1.4 %**. Este es un evento muy improbable.

- La probabilidad de acertar **exactamente 3** de las 4 tazas de un tipo por azar es de aproximadamente **23 %**. Este evento es considerablemente más probable.

Conclusión del experimento: Se determinó que acertar las 4 tazas sería evidencia suficiente para convencer a los presentes de su habilidad, ya que la probabilidad de que esto ocurriera por azar era muy baja (1.4 %). Un acierto de solo 3 tazas no se consideraría evidencia suficiente.

1.4. Pasos Fundamentales del Razonamiento Estadístico

La historia de la dama catadora de té ilustra los pilares de la inferencia estadística y las pruebas de hipótesis:

1. **Definición de hipótesis:** Se formula una hipótesis nula y una alternativa.
2. **Diseño de un experimento:** Se crea un método para recolectar datos que permita discriminar entre las hipótesis.
3. **Establecimiento de la significancia:** Se fija un umbral de probabilidad (la máxima probabilidad de error que se permitirá) por debajo del cual un resultado se considera raro.º “estadísticamente significativo” si la hipótesis nula fuera cierta.
4. **Determinación de la región de rechazo:** Se definen los resultados del experimento que se considerarán evidencia suficiente para rechazar la hipótesis nula.

2. Clase 2: Conceptos Básicos

Definición 2.1 (Experimento Aleatorio). Un experimento aleatorio corresponde a una situación cuyo resultado no se puede predecir con total certeza.

Definición 2.2 (Espacio Muestral). El espacio muestral de un experimento corresponde al conjunto de todos sus resultados posibles. Usualmente se denota por Ω .

Definición 2.3 (Evento Aleatorio). Un evento aleatorio corresponde a cualquier subconjunto del espacio muestral Ω .

Definición 2.4 (σ -álgebra). Un σ -álgebra corresponde a una colección de subconjuntos de Ω , denotada por \mathcal{F} , que cumple con las siguientes condiciones:

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
3. $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Definición 2.5 (Conjunto Potencia). El conjunto de todos los subconjuntos posibles de un espacio muestral Ω corresponde a un σ -álgebra, se denomina conjunto potencia de Ω y se denota por 2^Ω .

Definición 2.6 (Espacio Medible). El par (Ω, \mathcal{F}) , donde Ω es no vacío y \mathcal{F} es una σ -álgebra de subconjuntos de Ω , se denomina espacio medible.

Definición 2.7 (Medida de Probabilidad). Sea (Ω, \mathcal{F}) un espacio medible. Una medida de probabilidad, definida sobre él, es una función $P : \mathcal{F} \rightarrow [0, 1]$, que cumple con:

- **Axioma I:** $P(\Omega) = 1$
- **Axioma II:** $P(A) \geq 0, \forall A \in \mathcal{F}$
- **Axioma III:** (σ -aditividad) Si A_1, A_2, \dots son disjuntos dos a dos, entonces:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Definición 2.8 (Espacio de Probabilidad). El trío (Ω, \mathcal{F}, P) , donde Ω es no vacío, \mathcal{F} es una σ -álgebra de subconjuntos de Ω , y P es una medida de probabilidad, se denomina espacio de probabilidad.

3. Clase 3: Definiciones de Probabilidad

Definición 3.1 (Probabilidad Clásica). Sea (Ω, \mathcal{F}) un espacio medible, con Ω numerable y finito. Para un evento $A \in \mathcal{F}$, la medida de probabilidad clásica se define como:

$$P(A) = \frac{\#A}{\#\Omega}$$

Definición 3.2 (Ocurrencia de un Evento). Sea (Ω, \mathcal{F}) un espacio medible asociado a un experimento aleatorio. Se dice que un evento $A \in \mathcal{F}$ ocurrió si, una vez realizado el experimento, su resultado está contenido en A .

Definición 3.3 (Probabilidad Frecuentista). Sea (Ω, \mathcal{F}) un espacio medible asociado a un experimento aleatorio que se repite infinitas veces. Sea $f_n(A)$ la frecuencia relativa de veces en que el evento $A \in \mathcal{F}$ ha ocurrido en las primeras n repeticiones. La probabilidad frecuentista se define como:

$$P(A) = \lim_{n \rightarrow \infty} f_n(A)$$

Definición 3.4 (Probabilidad Subjetiva o Bayesiana). La probabilidad es una medida racional de la incertidumbre, condicional al estado de conocimiento. Este paradigma asigna probabilidades a objetos determinísticos para describir la incertidumbre sobre sus valores, la cual debe ser actualizada a la luz de nuevos datos.

4. Clase 4: Probabilidad Condicional e Independencia

Definición 4.1 (Probabilidad Condicional). Sean (Ω, \mathcal{F}, P) un espacio de probabilidad, y sean A y B elementos de \mathcal{F} . La probabilidad condicional de un evento A , dado un evento B tal que $P(B) > 0$, se define como:

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

Definición 4.2 (Independencia entre dos Eventos). Dos eventos, A y B, se dicen independientes si se cumple cualquiera de las siguientes condiciones equivalentes:

$$P(A|B) = P(A) \Leftrightarrow P(B|A) = P(B) \Leftrightarrow P(A \cap B) = P(A)P(B)$$

Definición 4.3 (Independencia Mutua). Los eventos E_1, \dots, E_n se dicen mutuamente independientes si y solo si, para toda subcolección de tamaño m , con $1 \leq m \leq n$, se cumple que:

$$P(E_{j_1} \cap E_{j_2} \cap \dots \cap E_{j_m}) = \prod_{k=j_1}^{j_m} P(E_k)$$

Definición 4.4 (Partición de un Conjunto). Sea E_1, E_2, \dots una colección de subconjuntos de Ω tales que:

- (i) $E_i \cap E_j = \emptyset$, para todo $i \neq j$
- (ii) $\bigcup_{i=1}^{\infty} E_i = \Omega$

Decimos que esta colección corresponde a una partición del conjunto Ω .

Teorema 4.5 (Teorema de Probabilidad Total). Sea E_1, E_2, \dots una partición de Ω . Entonces, para todo evento A, se cumple:

$$P(A) = \sum_{i=1}^{\infty} P(A|E_i)P(E_i)$$

Teorema 4.6 (Teorema de Bayes). Sea la colección de eventos E_1, E_2, \dots una partición de Ω . Entonces, para todo evento A y para cualquier $j \in \mathbb{N}$, se cumple:

$$P(E_j|A) = \frac{P(A|E_j)P(E_j)}{\sum_{i=1}^{\infty} P(A|E_i)P(E_i)}$$

5. Clase 5: Variables Aleatorias

Definición 5.1 (Variable Aleatoria). Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Una variable aleatoria es una función $X : \Omega \rightarrow \mathbb{R}$ que satisface la condición:

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}, \quad \text{para todo } x \in \mathbb{R}$$

Definición 5.2 (Función de Distribución). Sea X una variable aleatoria definida sobre un espacio de probabilidad (Ω, \mathcal{F}, P) . Se denomina función de distribución de X a la función $F_X : \mathbb{R} \rightarrow [0, 1]$ definida como:

$$F_X(x) = P(X \leq x) \equiv P(\{\omega \in \Omega : X(\omega) \leq x\}), \quad x \in \mathbb{R}$$

6. Clase 6: Variables Aleatorias Discretas

Definición 6.1 (Soporte de una Variable Aleatoria). El conjunto $\mathcal{B} \subseteq \mathbb{R}$ más pequeño que cumple con $P(X \in \mathcal{B}) = 1$ se denomina soporte de la variable aleatoria X .

Definición 6.2 (Variable Aleatoria Discreta). Una variable aleatoria X se dice discreta si y solo si su soporte es un conjunto numerable.

Definición 6.3 (Función de Probabilidad). Sea X una variable aleatoria discreta con valores en $\{x_1, x_2, \dots\}$. Se denomina función de probabilidad de X a la función p_X definida como:

$$p_X(x_i) = P(X = x_i), \quad i = 1, 2, \dots$$

Definición 6.4 (Media o Valor Esperado). Sea X una variable aleatoria discreta. La media, esperanza, o valor esperado de X se define como:

$$\mu_X = E(X) = \sum_x x \cdot p_X(x)$$

Definición 6.5 (Varianza). Sea X una variable aleatoria discreta con media μ_X . La varianza de X se define como:

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

Definición 6.6 (Desviación Estándar). La desviación estándar de una variable aleatoria X corresponde a la raíz cuadrada positiva de la varianza de X .

7. Clase 7: Distribuciones Discretas

Definición 7.1 (Distribución Bernoulli). Una variable aleatoria X tiene distribución Bernoulli si solo toma los valores 0 y 1. Su función de probabilidad, con parámetro $p \in [0, 1]$, es:

$$p_X(x) = p^x(1-p)^{1-x}I_{\{0,1\}}(x)$$

Definición 7.2 (Proceso de Bernoulli). Una secuencia infinita X_1, X_2, \dots de variables aleatorias independientes e idénticamente distribuidas Bernoulli(p) se denomina proceso de Bernoulli.

Definición 7.3 (Distribución Binomial). Sea X_1, \dots, X_n las primeras n realizaciones de un proceso de Bernoulli. La variable aleatoria $X = \sum_{i=1}^n X_i$, que cuenta el número de éxitos, sigue una distribución Binomial con parámetros $n \in \mathbb{N}$ y $p \in [0, 1]$.

Definición 7.4 (Distribución Geométrica). Considere un proceso de Bernoulli con parámetro $p \in [0, 1]$. La variable aleatoria X que representa el número de ensayos hasta que se observa el primer éxito sigue una distribución Geométrica de parámetro p . Su función de probabilidad es:

$$p_X(x) = p(1-p)^{x-1}I_{\mathbb{N}}(x)$$

Definición 7.5 (Distribución Binomial Negativa). Considere un proceso de Bernoulli con parámetro $p \in [0, 1]$. La variable aleatoria X que representa el número de ensayos hasta que ocurre el r -ésimo éxito sigue una distribución Binomial Negativa con parámetros (r, p) . Su función de probabilidad es:

$$p_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} I_{\{r, r+1, \dots\}}(x)$$

Definición 7.6 (Proceso de Poisson). Sea $N(t)$ el número de eventos que ocurren en el intervalo de tiempo $[0, t]$. Estos eventos constituyen un Proceso de Poisson de tasa $\lambda > 0$ si se cumplen las siguientes condiciones:

1. $N(0) = 0$
2. Los eventos ocurren de manera independiente.
3. La distribución del número de eventos en un intervalo depende únicamente de la longitud del intervalo.
4. Existe $\lambda > 0$ tal que $\lim_{h \rightarrow 0} \frac{P(N(h)=1)}{h} = \lambda$.
5. $\lim_{h \rightarrow 0} \frac{P(N(h) \geq 2)}{h} = 0$.

Definición 7.7 (Distribución de Poisson). Una variable aleatoria X sigue una distribución de Poisson de parámetro $\lambda > 0$ si su función de probabilidad está dada por:

$$p_X(x) = \frac{e^{-\lambda} \lambda^x}{x!} I_{\mathbb{N} \cup \{0\}}(x)$$

Definición 7.8 (Distribución Hipergeométrica). Una variable aleatoria X tiene distribución Hipergeométrica con parámetros N, n, r (donde $N, n, r \in \mathbb{N}$, $n \leq N$, $r \leq N$) si su función de probabilidad es de la forma:

$$p_X(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}} I_{\{0, 1, \dots, \min(r, n)\}}(x)$$