Sesión 2: Estimación e Incertidumbre

Escuela de Verano – CIMAT

Marco Antonio Aquino López [1]

[1]CIMAT
Probabilidad y Estadística.

5 July 2025

CIMAT

¿Qué tan buena es nuestra estimación?

- Queremos estimar parámetros desconocidos como proporciones o medias.
- ¿Qué tan confiables son esas estimaciones?
- ¿Cómo expresar la incertidumbre asociada?

Ley de los Grandes Números

Teorema (Ley de los Grandes Números (versión débil))

Sea $X_1, X_2, \ldots$ una secuencia i.i.d. con esperanza $\mu = \mathbb{E}[X_i]$ y varianza $\sigma^2 < \infty$. Entonces:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu \quad \text{cuando } n \to \infty.$$

Demostración de la LGN (boceto)

Demostración (boceto)

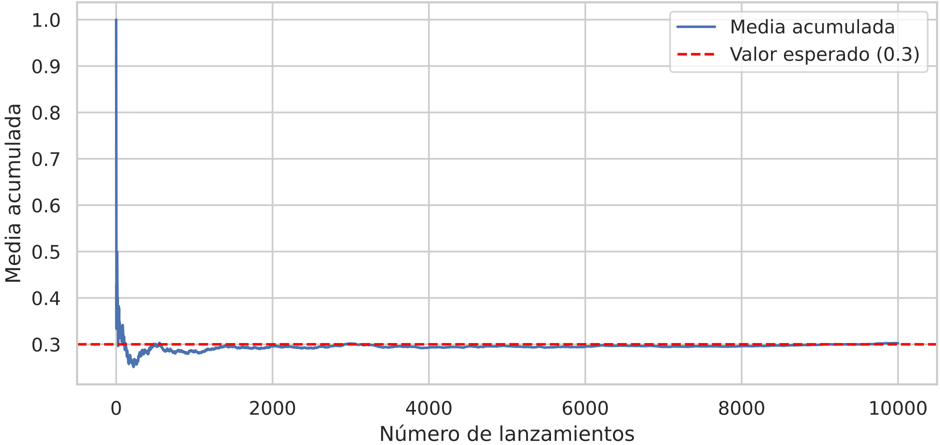
- *Sabemos que* $\mathbb{E}[\bar{X}_n] = \mu$ *y* $\mathrm{Var}(\bar{X}_n) = \sigma^2/n$.
- *Aplicamos la desigualdad de Chebyshev:*

$$\mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

- *Como* $n \to \infty$, *la probabilidad de desviarse de* μ *se vuelve arbitrariamente pequeña.*

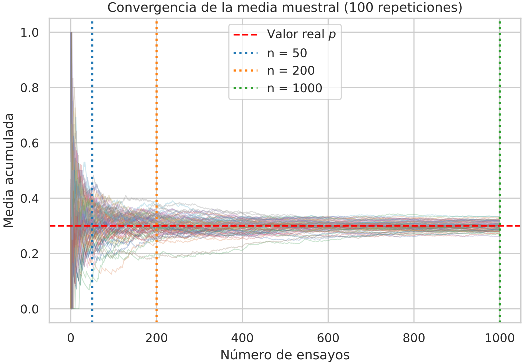
Esto implica que $\bar{X}_n \xrightarrow{P} \mu$.

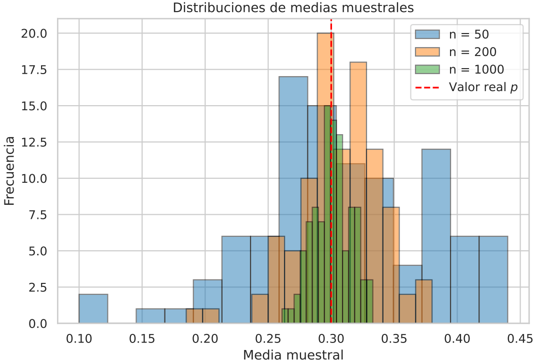
Visualización: media acumulada



La media muestral converge hacia $p = 0.3$ conforme aumentan las repeticiones.

Visualización: distribución de medias





Teorema Central del Límite (TCL)

Teorema (Teorema Central del Límite)

Sea $X_1, X_2, \ldots, X_n$ una secuencia i.i.d. con esperanza μ y varianza σ^2. Entonces, cuando n es grande:

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$$

Teorema Central del Límite (TCL)

Teorema (Teorema Central del Límite)

Sea $X_1, X_2, \ldots, X_n$ una secuencia i.i.d. con esperanza μ y varianza σ^2. Entonces, cuando n es grande:

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$$

Implicación: $\bar{X}_n$ se aproxima a una distribución normal con media μ y varianza σ^2/n.

Aplicación a proporciones

Si $X_i \sim \text{Bernoulli}(p)$:

- $\mathbb{E}[X_i] = p, \quad \mathrm{Var}(X_i) = p(1-p)$
- La proporción muestral $\hat{p} = \bar{X}_n$ se comporta como:

$$\hat{p} \approx \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

Esto nos permite construir intervalos de confianza para p.

Desde el TCL al intervalo de confianza (1/2)

Recordemos el resultado del Teorema Central del Límite aplicado a proporciones:

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

Desde el TCL al intervalo de confianza (1/2)

Recordemos el resultado del Teorema Central del Límite aplicado a proporciones:

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

Como queremos hacer inferencia sobre p, podemos estandarizar:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \approx \mathcal{N}(0, 1)$$

Desde el TCL al intervalo de confianza (1/2)

Recordemos el resultado del Teorema Central del Límite aplicado a proporciones:

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

Como queremos hacer inferencia sobre p, podemos estandarizar:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \approx \mathcal{N}(0, 1)$$

Ahora usamos la probabilidad simétrica en la normal:

$$\mathbb{P}\left(-z_{0.975} < Z < z_{0.975}\right) \approx 0.95$$

con $z_{0.975} \approx 1.96$

Desde el TCL al intervalo de confianza (2/2)

Sustituimos Z:

$$\mathbb{P}\left(-1.96 < \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} < 1.96\right) \approx 0.95$$

Desde el TCL al intervalo de confianza (2/2)

Sustituimos Z:

$$\mathbb{P}\left(-1.96 < \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} < 1.96\right) \approx 0.95$$

Multiplicamos y despejamos:

$$\mathbb{P}\left(\hat{p} - 1.96\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + 1.96\sqrt{\frac{p(1-p)}{n}}\right) \approx 0.95$$

Desde el TCL al intervalo de confianza (2/2)

Sustituimos Z:

$$\mathbb{P}\left(-1.96 < \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} < 1.96\right) \approx 0.95$$

Multiplicamos y despejamos:

$$\mathbb{P}\left(\hat{p} - 1.96\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + 1.96\sqrt{\frac{p(1-p)}{n}}\right) \approx 0.95$$

Como no conocemos p, lo aproximamos con $\hat{p}$ en la varianza:

$$IC_{95\%} = \left[\hat{p} \pm 1.96 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Desde el TCL al intervalo de confianza (2/2)

Sustituimos Z:

$$\mathbb{P}\left(-1.96 < \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} < 1.96\right) \approx 0.95$$

Multiplicamos y despejamos:

$$\mathbb{P}\left(\hat{p} - 1.96\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + 1.96\sqrt{\frac{p(1-p)}{n}}\right) \approx 0.95$$

Como no conocemos p, lo aproximamos con $\hat{p}$ en la varianza:

$$IC_{95\%} = \left[\hat{p} \pm 1.96 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Este es el intervalo de confianza clásico para una proporción.

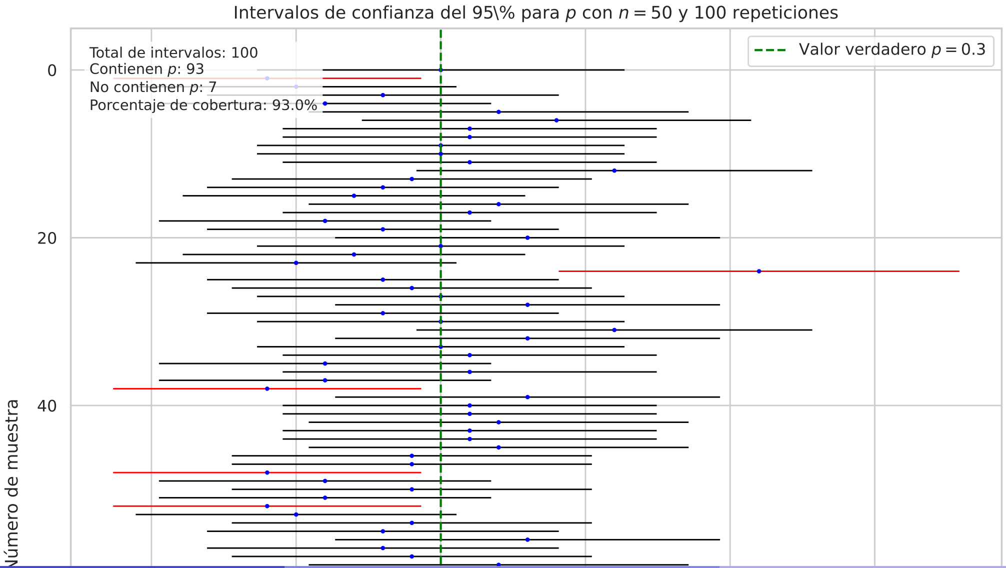
Intervalos de confianza (95%)

Fórmula general para proporciones

$$IC_{95\%} = \left[\hat{p} \pm z_{0.975} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

- $z_{0.975} \approx 1.96$
- Interpretable como: "en el 95% de los experimentos, el intervalo incluirá al valor verdadero de p"

Visualización: intervalos de confianza



Comparación: estimación puntual vs intervalo

- La media muestral sola no basta.
- El intervalo captura la variabilidad entre muestras.
- Nos da una herramienta para comunicar incertidumbre con base estadística.

Una pregunta distinta

En lugar de preguntar:

> *"¿Qué tan variable es $\hat{p}$ si repito el muestreo?" (frecuentista)*

preguntamos:

> **"¿Qué tan probable es que p esté en cierto intervalo, dado los datos observados?"** *(bayesiano)*

Teorema de Bayes

Teorema (Teorema de Bayes)

Sean eventos A y B con mathbbP(B) ¿ 0, entonces :$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$

Teorema de Bayes

Teorema (Teorema de Bayes)

Sean eventos A y B con mathbbP(B) ¿ 0, entonces :$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$

Demostración (boceto)

Por definición de probabilidad condicional:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad \mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

Despejando $\mathbb{P}(A \cap B)$ *y sustituyendo:*

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

Teorema de Bayes para inferencia

Sea p un parámetro desconocido, y x los datos observados. Entonces:

$$\underbrace{p(p \mid x)}_{\text{Posterior}} = \frac{\underbrace{p(x \mid p)}_{\text{Verosimilitud}} \cdot \underbrace{p(p)}_{\text{Prior}}}{\underbrace{p(x)}_{\text{Evidencia}}}$$

Teorema de Bayes para inferencia

Sea p un parámetro desconocido, y x los datos observados. Entonces:

$$\underbrace{p(p \mid x)}_{\text{Posterior}} = \frac{\underbrace{p(x \mid p)}_{\text{Verosimilitud}} \cdot \underbrace{p(p)}_{\text{Prior}}}{\underbrace{p(x)}_{\text{Evidencia}}}$$

- $p(p)$: creencia inicial sobre p (antes de observar datos)
- $p(x \mid p)$: cuán compatibles son los datos con distintos valores de p
- $p(p \mid x)$: creencia actualizada sobre p después de ver los datos

Ejemplo: proporción p con prior Beta

- Supongamos $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$
- Observamos $s = \sum X_i$ éxitos en n ensayos
- Elegimos una prior: $p \sim \text{Beta}(\alpha_0, \beta_0)$

Ejemplo: proporción p con prior Beta

- Supongamos $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$
- Observamos $s = \sum X_i$ éxitos en n ensayos
- Elegimos una prior: $p \sim \text{Beta}(\alpha_0, \beta_0)$

La verosimilitud es:

$$L(p) \propto p^s (1-p)^{n-s}$$

Ejemplo: proporción p con prior Beta

- Supongamos $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$
- Observamos $s = \sum X_i$ éxitos en n ensayos
- Elegimos una prior: $p \sim \text{Beta}(\alpha_0, \beta_0)$

La verosimilitud es:

$$L(p) \propto p^s (1-p)^{n-s}$$

Entonces, la posterior es proporcional a:

$$p(p \mid \text{datos}) \propto p^{\alpha_0 - 1 + s} (1-p)^{\beta_0 - 1 + n - s}$$

Posterior Beta y credibilidad

Distribución posterior

$$p \mid \text{datos} \sim \text{Beta}(\alpha_0 + s, \beta_0 + n - s)$$

Posterior Beta y credibilidad

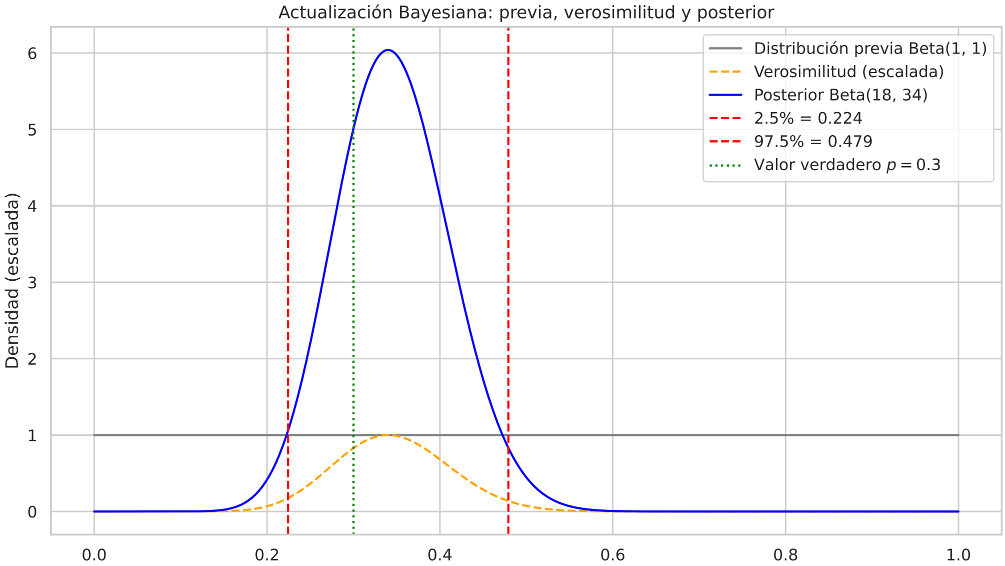
Distribución posterior

$$p \mid \text{datos} \sim \text{Beta}(\alpha_0 + s, \beta_0 + n - s)$$

Intervalo de credibilidad del 95%: los cuantiles 2.5% y 97.5% de esta distribución:

$$\text{Intervalo } [a, b] \text{ tal que } \mathbb{P}(p \in [a, b] \mid \text{datos}) = 0.95$$

Visualización: actualización bayesiana



Frecuentista vs Bayesiano

	Frecuentista	**Bayesiano**
¿Qué es el parámetro?	Fijo, desconocido	Variable aleatoria
¿Qué se modela?	Variabilidad de la muestra	Creencias sobre el parámetro
Inferencia	Basada en replicar el muestreo	Basada en probabilidad condicional
Intervalo	De confianza: $\hat{p}\pm$ margen	De credibilidad: cuantil posterior
Interpretación	“Cobertura en repetición”	“Probabilidad de p en un rango”

Visualmente: ¿qué cambia?

- En el enfoque clásico, lo que varía es la muestra.
- En el enfoque bayesiano, lo que varía es el parámetro dado lo observado.
- Ambos generan intervalos, pero su interpretación es diferente.

Visualmente: ¿qué cambia?

- En el enfoque clásico, lo que varía es la muestra.
- En el enfoque bayesiano, lo que varía es el parámetro dado lo observado.
- Ambos generan intervalos, pero su interpretación es diferente.

Idea clave

En la práctica, ambos enfoques pueden conducir a resultados similares, pero sus fundamentos y objetivos son distintos.

Exploración con código: sugerencias

Usando el script, explora lo siguiente:

- ¿Qué ocurre al cambiar el tamaño de muestra n? ¿Y el valor de p?
- ¿Qué pasa si usamos `Beta(2, 5)` como prior informativa?
- ¿Cómo cambia la cobertura si usamos pocos datos?
- ¿Cómo se ve la verosimilitud cuando hay más o menos éxitos?
- ¿El intervalo de credibilidad es más o menos conservador que el de confianza?

Actividad final: juega con los parámetros

Objetivo

Implementa una función que reciba como entrada:

- tamaño de muestra n,   número de éxitos s
- parámetros de la prior α_0, β_0

y grafique la posterior Beta con su intervalo de credibilidad.

Actividad final: juega con los parámetros

Objetivo

Implementa una función que reciba como entrada:

- tamaño de muestra *n*,   número de éxitos *s*
- parámetros de la prior α_0, β_0

y grafique la posterior Beta con su intervalo de credibilidad.

Luego, repite para distintos valores y reflexiona:

¿Cuánto influye la prior cuando tenemos pocos datos? ¿Cuándo deja de importar?

Dinámica de cierre: mini-proyectos

- Formaremos **tres equipos de trabajo**.
- Cada equipo elige o se le asigna un mini-proyecto.
- Tienen tiempo para explorarlo, discutirlo y preparar una breve presentación para mañana.
- El objetivo es aplicar lo aprendido y formular preguntas.

Proyecto 1: ¿Qué pasa con muestras pequeñas?

- Simula muchas muestras Bernoulli con tamaños pequeños ($n = 5$, $n = 10$, $n = 20$).
- Compara la cobertura real del intervalo de confianza clásico.
- ¿Qué tan bien se cumple el 95% de cobertura?
- ¿Mejora la cobertura con el intervalo bayesiano? Usa distintas priors.

Proyecto 2: El efecto de la prior

- Simula un experimento con pocos datos ($n = 20$) y una proporción baja ($p = 0.1$).
- Compara la posterior resultante usando:
  - Una prior uniforme `Beta(1,1)`
  - Una prior informativa `Beta(2,20)`
- ¿Cómo cambian la distribución posterior y el intervalo de credibilidad?
- ¿Cuándo deberíamos desconfiar de una prior fuerte?

Proyecto 3: Comparando coberturas

- Diseña un experimento para comparar la **frecuencia de cobertura** de:
  - El intervalo clásico
  - El intervalo bayesiano
- Usa simulaciones con $n = 50$ y repítelo muchas veces.
- ¿Cuál de los dos logra cubrir el valor verdadero de p más veces?
- ¿Hay alguna condición donde uno supere claramente al otro?

Proyecto 4: Verosimilitud y evidencia visual

- Explora cómo se comporta la función de verosimilitud cuando:
  - Aumenta el número de éxitos
  - Aumenta el tamaño de muestra
- Superpón la verosimilitud con la posterior resultante para distintas priors.
- Interpreta los casos donde la posterior se aleja o coincide con la verosimilitud.

Proyecto 5: Intervalos como comunicación

- Simula un mismo experimento con distintos valores de p (ej. 0.1, 0.5, 0.9).
- Calcula y grafica intervalos clásicos y bayesianos.
- Discute:
  - ¿Cuál comunica mejor la incertidumbre?
  - ¿Cuál es más útil para un tomador de decisiones?
- Presenta tus hallazgos con ejemplos visuales.