

# Inferencia estadística

## Propiedades asintóticas

San Andrés

May 31, 2022

# Convergencia en probabilidad

Tanto el Teorema Central del Límite (TCL) como la Ley de los grandes números (LGN) nos dicen algo sobre el comportamiento de la media muestral  $\bar{X}_n$  cuando el tamaño de muestra  $n$  tiende a infinito.

## Theorem (LGN)

*Sea  $X_1, \dots, X_n$  una muestra aleatoria con esperanza igual a  $\mu$ . Entonces, para todo  $\varepsilon > 0$ :*

$$P(|\bar{X}_n - \mu| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0.$$

La LGN nos dice que  $\bar{X}_n$  converge, en cierto sentido, a

$$\mu = E(X)$$

A este tipo de convergencia de sucesiones de variables aleatorias la llamaremos **convergencia en probabilidad**.

# Convergencia en probabilidad

## Definition (Convergencia en probabilidad)

Sea  $V_1, \dots, V_n, \dots$  una sucesión de variables aleatorias y sea  $V$  una variable aleatoria. Decimos que  $V_n$  **converge en probabilidad** a  $c$  si para todo  $\varepsilon > 0$

$$P(|V_n - \overset{\text{cte}}{\cancel{V}}| \geq \varepsilon) \xrightarrow[n \rightarrow \infty]{} 0.$$

En este caso, usaremos la notación

$$V_n \xrightarrow{P} c.$$

# Consistencia y convergencia en probabilidad

## LGN

Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que con esperanza igual a  $\mu$ .  
Entonces

$$\overline{X}_n \xrightarrow{P} \mu.$$

## Definition (Consistencia)

Sea  $X_1, \dots, X_n$  una muestra aleatoria con una distribución que depende de cierto parámetro  $\theta$ . Sea  $\hat{\theta}_n$  un estimador del parámetro  $\theta$ . Decimos que  $\hat{\theta}_n$  es **consistente** si

$$\hat{\theta}_n \xrightarrow{P} \theta$$

## Idea

$\hat{\theta}_n$  es consistente si para tamaños de muestra grande, es “muy” probable que  $\hat{\theta}_n$  esté “cerca” de  $\theta$ . Recordar que  $\hat{\theta}_n$  es una variable aleatoria y que  $\theta$  está fijo! **Pedirle a un estimador que sea consistente ¡es lo mínimo indispensable!**

# Consistencia y convergencia en probabilidad

## LGN

Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que con esperanza igual a  $\mu$ .  
Entonces

$$\overline{X}_n \xrightarrow{P} \mu.$$

## Definition (Consistencia)

Sea  $X_1, \dots, X_n$  una muestra aleatoria con una distribución que depende de cierto parámetro  $\theta$ . Sea  $\hat{\theta}_n$  un estimador del parámetro  $\theta$ . Decimos que  $\hat{\theta}_n$  es **consistente** si

$$\hat{\theta}_n \xrightarrow{P} \theta$$

## Idea

$\hat{\theta}_n$  es consistente si para tamaños de muestra grande, es “muy” probable que  $\hat{\theta}_n$  esté “cerca” de  $\theta$ . Recordar que  $\hat{\theta}_n$  es una variable aleatoria y que  $\theta$  está fijo! **Pedirle a un estimador que sea consistente ¡es lo mínimo indispensable!**

# Consistencia y convergencia en probabilidad

## LGN

Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que con esperanza igual a  $\mu$ . Entonces

$$\overline{X}_n \xrightarrow{P} \mu.$$

## Definition (Consistencia)

Sea  $X_1, \dots, X_n$  una muestra aleatoria con una distribución que depende de cierto parámetro  $\theta$ . Sea  $\hat{\theta}_n$  un estimador del parámetro  $\theta$ . Decimos que  $\hat{\theta}_n$  es **consistente** si

$$\hat{\theta}_n \xrightarrow{P} \theta$$

## Idea

$\hat{\theta}_n$  es consistente si para tamaños de muestra grande, es “muy” probable que  $\hat{\theta}_n$  esté “cerca” de  $\theta$ . Recordar que  $\hat{\theta}_n$  es una variable aleatoria y que  $\theta$  está fijo! **Pedirle a un estimador que sea consistente ¡es lo mínimo indispensable!**

# Convergencia en probabilidad

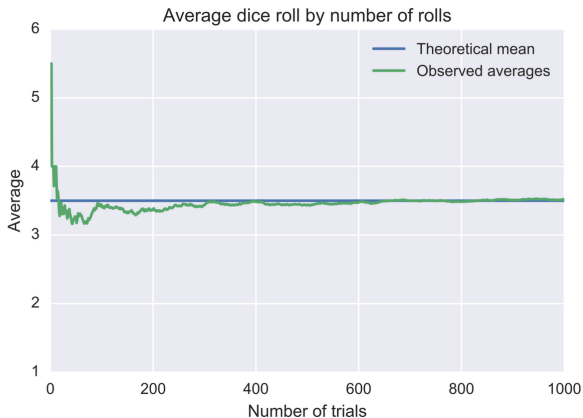


Figure: Evolución del promedio de un dado. Fuente: [Wikipedia](#)

# Consistencia y convergencia en probabilidad

## Estimador de la varianza

Nos interesa estudiar el estado de salud de la población Argentina, en particular su peso. Supongamos que el peso de una persona elegida al azar es una variable aleatoria con esperanza  $\mu$  y varianza  $\sigma^2$  desconocidos.

Dada una muestra aleatoria de pesos  $X_1, \dots, X_n$ , sabemos que  $\bar{X}_n$  es un estimador consistente de  $\mu$ . ¿Qué sucede con el estimador de la varianza?

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2.$$



# Consistencia y convergencia en probabilidad

## Estimador de la varianza

Sabemos que  $\bar{X}_n$  es un estimador consistente de  $E(X_1)$ , por lo tanto *esperamos* que

$$(\bar{X}_n)^2 \xrightarrow{P} E^2(X_1).$$

Por otro lado, la ley de los grandes números nos dice

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} E(X_1^2).$$

# Consistencia y convergencia en probabilidad

## Estimación de la varianza

*Esperamos* entonces que

$$\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2 \xrightarrow{P} \sigma^2.$$

Para ver que esto efectivamente es así, tenemos que entender cómo se comporta la convergencia en probabilidad al hacer operaciones entre variable aleatorias: sumar, resta, multiplicar, elevar al cuadrado, etc.

# Propiedades de la convergencia en probabilidad

## Proposición ( ALGEBRA DE LÍMITES )

Sea  $a_n$  es una sucesión de números reales tal que  $a_n \xrightarrow[n \rightarrow \infty]{} a$ . Sean

$V_n$  y  $W_n$  tales que  $V_n \xrightarrow{P} V$  y  $W_n \xrightarrow{P} W$ . Entonces

Siendo  $V$  y  $W$  CONSTANTES

①  $a_n + V_n \xrightarrow{P} a + V.$

②  $a_n V_n \xrightarrow{P} aV.$

③  $V_n + W_n \xrightarrow{P} V + W.$

④  $V_n W_n \xrightarrow{P} VW.$

⑤ Si  $P(W = 0) = 0$ ,  $V_n/W_n \xrightarrow{P} V/W.$

⑥ Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  es continua  $f(V_n) \xrightarrow{P} f(V).$

⑦ Más generalmente, si  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  es continua  
 $g(V_n, W_n) \xrightarrow{P} g(V, W).$

# Estimadores consistentes

## Álgebra de límites

Con el álgebra de límite podremos mostrar que  $\widehat{\sigma^2}$  es consistente

$$\widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2 \xrightarrow{P} \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

También que el estimador  $S^2$  es consistente

$$S_n^2 = \frac{n}{n-1} \widehat{\sigma_n^2} \xrightarrow{P} 1 \cdot \sigma^2$$

# Estimadores consistentes

## Ejercicios

- Proponer un estimador consistente para el desvío estándar
- Proponer un estimador consistente para el coeficiente de variación

$$CV(X) = \frac{\mu}{\sigma}.$$

## Observación

Gracias al álgebra de límites si nuestro estimador depende de promedios podemos mostrar la consistencia de manera sencilla.

# Estimadores consistentes

## Ejercicios

- Proponer un estimador consistente para el desvío estándar
- Proponer un estimador consistente para el coeficiente de variación

$$CV(X) = \frac{\mu}{\sigma}.$$

## Observación

Gracias al álgebra de límites si nuestro estimador depende de promedios podemos mostrar la consistencia de manera sencilla.

# Consistencia

Muchas veces probaremos la consistencia de estimadores usando LGN y las propiedades vistas antes. En algunos casos la consistencia de un estimador se puede probar directamente a partir del siguiente resultado.

## Proposición

Sea  $V_1, \dots, V_n, \dots$  una sucesión de variables aleatorias. Sea  $c \in \mathbb{R}$  una constante. Si  $E(V_n) \xrightarrow{n \rightarrow \infty} c$  y  $\text{Var}(V_n) \xrightarrow{n \rightarrow \infty} 0$  entonces

$$V_n \xrightarrow{P} c$$

## Idea

Si el centro del histograma (densidad) de  $V_n$  se está acercando a  $c$  y la variabilidad de  $V_n$  se está acercando a cero, para  $n$  grande es muy probable que  $V_n$  esté “cerca” de  $c$ .

# Propiedades de la convergencia en probabilidad

Proof.

Lo hacemos suponiendo que  $E(V_n) = c$  para todo  $n$ .

Fijemos  $\varepsilon > 0$ . Por Tchebysheff

$$P(|V_n - c| \geq \varepsilon) \leq \frac{\text{Var}(V_n)}{\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0.$$

El caso general queda de ejercicio (opcional).





# Propiedades de la convergencia en probabilidad

## Máximo de uniformes

Sea  $X_1, \dots, X_n \sim \mathcal{U}(0, \theta)$  vimos que

$$E(\hat{\theta}_n) = \frac{n}{n+1}\theta$$

$$\text{Var}(\hat{\theta}_n) = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2.$$

# Propiedades de la convergencia en probabilidad

## Máximo de uniformes

Como

$$E(\hat{\theta}_n) \xrightarrow{n \rightarrow \infty} \theta$$

y

$$Var(\hat{\theta}_n) \xrightarrow{n \rightarrow \infty} 0,$$

por la proposición que vimos antes

$$\hat{\theta}_n \xrightarrow{P} \theta,$$

es decir,  $\hat{\theta}_n$  es consistente para  $\theta$ .

E

sta propiedad dice que un estimador es consistente si para todo  $\theta$

$$ECM_{\theta}(\hat{\theta}) = \text{Sesgo}(\hat{\theta})^2 + Var(\hat{\theta}) \rightarrow 0$$

# Convergencia en distribución

## Theorem (TCL)

Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que  $E(X) = \mu$  y  $\text{Var}(X) = \sigma^2$ , entonces, para todo  $z \in \mathbb{R}$

$$P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq z\right) \xrightarrow{n \rightarrow \infty} \Phi(z),$$

donde  $\Phi(z) = P(Z \leq z)$  para  $Z \sim \mathcal{N}(0, 1)$ .

# Convergencia en distribución

## Idea

Para tamaños de muestra grande, la distribución de  $\bar{X}_n$  es aproximadamente normal:

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

y

$$\bar{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right).$$

El TCL nos dice que la distribución de una versión debidamente re-escalada de  $\bar{X}_n$  converge en cierto sentido a una normal  $\mathcal{N}(0, 1)$ .

# Convergencia en distribución

## Definition (Convergencia en distribución)

Sea  $V_1, \dots, V_n, \dots$  una sucesión de variables aleatorias y sea  $V$  una variable aleatoria. Decimos que  $V_n$  converge en distribución a  $V$  si, para todo  $v$  tal que la función de distribución acumulada de  $V$ ,  $F_V$  es continua en  $v$ , vale que

$$P(V_n \leq v) \xrightarrow{n \rightarrow \infty} P(V \leq v).$$

## Notación

Usaremos la notación  $V_n \xrightarrow{d} V$ . También haremos el siguiente abuso de notación. Por ejemplo, si  $V \sim \mathcal{N}(0, 1)$ , escribiremos

$$V_n \xrightarrow{d} \mathcal{N}(0, 1),$$

o, si  $V \sim \text{Pois}(\lambda)$ , escribiremos

$$V_n \xrightarrow{d} \text{Pois}(\lambda), \text{ etc.}$$

# Convergencia en distribución

## Idea

Si  $V_n \xrightarrow{d} V$ , la función de distribución acumulada de  $V_n$  está “cerca” de la distribución de  $V$  para valores de  $n$  grandes.

**Esto no quiere decir que  $V_n$  esté cerca de  $V$ !**

De hecho,  $V_1, \dots, V_n$  y  $V$  pueden estar definidas todas sobre espacios muestrales distintos!

La convergencia en distribución, más que una convergencia de variables aleatorias, es una convergencia de funciones de distribución. De hecho, si llamamos  $F_{V_n}$  a la función de distribución de  $V_n$  y  $F_V$  a la función de distribución de  $V$ ,  $V_n \xrightarrow{d} V$  si

$$F_{V_n}(v) \xrightarrow{n \rightarrow \infty} F_V(v)$$

para todo  $v$  donde  $F_V$  es continua.

# Convergencia en distribución

## TCL

Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que  $E(X) = \mu$  y  $\text{Var}(X) = \sigma^2$ , entonces, para todo  $z \in \mathbb{R}$

$$P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq z\right) \xrightarrow{n \rightarrow \infty} \Phi(z),$$

donde  $\Phi(z) = P(Z \leq z)$  para  $Z \sim \mathcal{N}(0, 1)$ . Luego

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

Notemos que la función de distribución acumulada de una normal (o de cualquier variable aleatoria continua) es continua en todo punto.

# Convergencia en distribución

## Proposición

*Si  $V_n \xrightarrow{P} V$ , entonces  $V_n \xrightarrow{d} V$ .*

## ¡Ojo!

No es cierto que si  $V_n \xrightarrow{d} V$  entonces  $V_n \xrightarrow{P} V$ .

La convergencia en distribución es más **débil** que la convergencia en probabilidad.



# Convergencia en distribución

Volvamos al ejemplo de la estimación del peso de la población argentina. Usando LGN y las propiedades de convergencia en probabilidad, vimos que  $\bar{X}_n$  es un estimador consistente de  $\mu$  y que

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2$$

es un estimador consistente de  $\sigma^2$ .

Por TCL sabemos que, si  $n$  es grande, la distribución de

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$$

es aproximadamente normal standard. Luego

$$\bar{X}_n - \mu \approx N\left(0, \frac{\sigma^2}{n}\right).$$

Si conociésemos  $\sigma^2$ , tendríamos una aproximación de cómo están distribuidos los *errores* de nuestra estimación de  $\mu$  por  $\bar{X}_n$ .

# Convergencia en distribución

Pero no conocemos  $\sigma^2$ . ¿Qué pasará si reemplazamos  $\sigma^2$  por  $\hat{\sigma}^2$  en la fórmula anterior? ¿Será cierto lo siguiente?

$$\bar{X}_n - \mu \approx N\left(0, \frac{\hat{\sigma}^2}{n}\right)$$

Para poder responder a esta pregunta, necesitamos el **Lema de Slutsky**.

# Propiedades de la convergencia en distribución

## Proposición (Lema de Slutsky)

Sea  $c \in \mathbb{R}$  una constante. Si  $W_n \xrightarrow{P} c$  y  $V_n \xrightarrow{d} V$  entonces

$$W_n V_n \xrightarrow{d} cV$$

y

$$V_n + W_n \xrightarrow{d} V + c$$

# Convergencia en distribución

Volvamos al ejemplo de antes. Por TCL

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

Habíamos visto que

$$\hat{\sigma} \xrightarrow{P} \sigma$$

lo que implica, por las propiedades de convergencia en probabilidad que vimos antes, que

$$\frac{\sigma}{\hat{\sigma}} \xrightarrow{P} 1.$$

Ahora

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}} = \underbrace{\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}}_{\xrightarrow{d} \mathcal{N}(0,1)} \underbrace{\frac{\sigma}{\hat{\sigma}}}_{\xrightarrow{P} 1} \xrightarrow{d} \mathcal{N}(0, 1),$$

por el lema de Slutsky.

# Convergencia en distribución

Luego

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

y por lo tanto (informalmente)

$$\bar{X}_n - \mu \approx \mathcal{N}\left(0, \frac{\hat{\sigma}^2}{n}\right).$$

# Convergencia en distribución

## Ejemplo

Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que  $E(X_1) = 0$  y  $E(X_1^2) = 2$ . Calcular el límite en distribución de

$$Y_n = \frac{\sqrt{n} \sum_{i=1}^n X_i}{\sum_{i=1}^n X_i^2}.$$

# Consistencia de máxima verosimilitud

## Theorem (Consistencia de máxima verosimilitud)

*Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que  $X_i$  tiene densidad  $f(x; \theta_0)$  para un  $\theta_0$  desconocido. Bajos condiciones de regularidad sobre  $f(x; \theta_0)$ ,*

$$\hat{\theta}_{MV} \xrightarrow{P} \theta_0.$$

Las condiciones de regularidad sobre  $f(x; \theta)$  que se necesitan piden, esencialmente, que  $f(x; \theta)$  sea un función 'suave' de  $\theta$ , en el sentido de que tenga suficientes derivadas.

# Consistencia de momentos

## Theorem (Consistencia de momentos)

*Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que  $X_i$  tiene densidad  $f(x; \theta_0)$  para un  $\theta_0$  desconocido. Bajos condiciones de regularidad sobre  $f(x; \theta_0)$ ,*

$$\hat{\theta}_{MM} \xrightarrow{P} \theta_0.$$



# Distribución asintótica

## Definition (Normalidad asintótica)

Decimos que un estimador  $\hat{\theta}_n$  de  $\theta$  es asintóticamente normal con media cero si

$$n^{1/2}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, V(\theta)),$$

es decir, si para tamaños de muestra grande,

$$\sqrt{n}(\hat{\theta}_n - \theta) \approx N(0, V(\theta)).$$

Llamaremos a  $V(\theta)$  la varianza asintótica de  $\hat{\theta}_n$ . Llamaremos a  $\sqrt{V(\theta)}$  el error standard asintótico o el desvío standard asintótico de  $\hat{\theta}_n$ . Diremos que  $\hat{\theta}_n$  es asintóticamente insesgado.

# Distribución asintótica

## Observación

Si tenemos entonces un estimador de  $V(\theta)$ , por ejemplo  $V(\hat{\theta}_n)$ , podemos medir la variabilidad de  $\hat{\theta}_n$  como

$$\frac{V(\hat{\theta}_n)}{n}.$$

Para que esto funcione, tendríamos que probar que el estimador de  $V(\theta)$  que usamos, que podría ser  $V(\hat{\theta}_n)$ , es consistente.

Veremos más adelante la utilidad de conocer la distribución asintótica de estimadores puntuales para poder hacer estimación por intervalos.

# Distribución asintótica

## Example

Consideremos una muestra aleatoria  $X_1, \dots, X_n$  y llamemos  $\mu = E(X_1)$  y  $\sigma^2 = \text{Var}(X_1)$ . Por el TCL,

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2).$$

Luego  $\bar{X}_n$  es asintóticamente normal con media cero y su varianza asintótica es  $\sigma^2$ .

Cómo podemos estimar la varianza (asintótica) de  $\bar{X}_n$ ?

# Distribución asintótica

## Example (Exponencial)

Consideremos una muestra aleatoria de una exponencial de parámetro  $\lambda$ . Sabemos que el estimador de máxima verosimilitud es  $1/\bar{X}_n$ . Cómo vemos si  $1/\bar{X}_n$  es asintóticamente normal?

Vamos a probar que

$$\sqrt{n} \left( \frac{1}{\bar{X}_n} - \lambda \right) \xrightarrow{d} N(0, \lambda^2).$$

La idea es que, si hacemos una expansión de Taylor de primer orden de la función  $g(x) = 1/x$  alrededor de  $1/\lambda$ , tenemos que

$$\frac{1}{x} \approx \lambda - \lambda^2 \left( x - \frac{1}{\lambda} \right),$$

siempre que  $x$  esté cerca de  $1/\lambda$ .

# Distribución asintótica

## Example (Exponencial)

Como ya sabemos que con probabilidad alta, para  $n$  grande,  $\bar{X}_n$  está cerca de  $1/\lambda$ , tenemos que

$$\frac{1}{\bar{X}_n} - \lambda \approx -\lambda^2 \left( \bar{X}_n - \frac{1}{\lambda} \right)$$

y por lo tanto usando el TCL

$$\sqrt{n} \left( \frac{1}{\bar{X}_n} - \lambda \right) \approx -\lambda^2 \sqrt{n} \left( \bar{X}_n - \frac{1}{\lambda} \right) \xrightarrow{d} N \left( 0, \lambda^4 \cdot \frac{1}{\lambda^2} \right).$$

Luego el estimador de máxima verosimilitud es asintóticamente normal con media cero.

¿Cómo podemos estimar la varianza asintótica de  $1/\bar{X}_n$ ?

# Método delta

La técnica que usamos para calcular la distribución asintótica en el ejemplo anterior se conoce como el método delta.

## Theorem (Método delta)

*Supongamos que  $\hat{\theta}_n$  cumple que*

$$\sqrt{n} \left( \hat{\theta}_n - \theta \right) \xrightarrow{d} N(0, V(\theta)).$$

*Supongamos que  $g(x)$  es una función con derivada continua y que  $g'(\theta) \neq 0$ . Entonces*

$$\sqrt{n} \left( g(\hat{\theta}_n) - g(\theta) \right) \xrightarrow{d} N(0, V(\theta)(g'(\theta))^2).$$

# Distribución asintótica

Vamos a ver a continuación que bajo condiciones de regularidad, los estimadores de máxima verosimilitud son asintóticamente normales con media cero y que su varianza asintótica es en cierto sentido, la mínima posible.

# Distribución asintótica de máxima verosimilitud

## Theorem (Distribución asintótica de máxima verosimilitud)

Sea  $X_1, \dots, X_n$  una muestra aleatoria tal que  $X_i$  tiene densidad  $f(x; \theta_0)$  para un  $\theta_0$  desconocido. Bajos condiciones de regularidad sobre  $f(x; \theta_0)$ ,

$$\sqrt{n} \left( \hat{\theta}_{MV} - \theta_0 \right) \xrightarrow{d} N(0, I_1(\theta_0)^{-1}),$$

donde

$$I_1(\theta_0) = E \left\{ \left[ \frac{d \log(f(X; \theta))}{d\theta}(\theta_0) \right]^2 \right\}$$

es la información de Fisher.



# Distribución asintótica de máxima verosimilitud

## Observación

Este resultado nos dice que, bajo ciertas condiciones, si usamos máxima verosimilitud tenemos garantizado que nuestro estimador sea consistente y asintóticamente normal con media cero.

# Distribución asintótica de máxima verosimilitud

Proof.

Ver All of Statistics.



# Optimalidad asintótica de máxima verosimilitud

La inversa de la información de Fisher es la menor varianza asintótica que puede tener un estimador asintóticamente insesgado de  $\theta$ .

Por esto se suele decir que el estimador de máxima verosimilitud es asintóticamente eficiente o también asintóticamente óptimo.

# Optimalidad asintótica de máxima verosimilitud

## Theorem

*Bajo condiciones de regularidad, si  $\tilde{\theta}_n$  es un estimador de  $\theta$  tal que*

$$\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{d} N(0, \tilde{V}(\theta))$$

*entonces*

$$I_1^{-1}(\theta) \leq \tilde{V}(\theta).$$

## Idea

La varianza asintótica de cualquier estimador asintóticamente normal e insesgado es mayor o igual que la varianza asintótica del estimador de máxima verosimilitud