



MA34B – Estadística

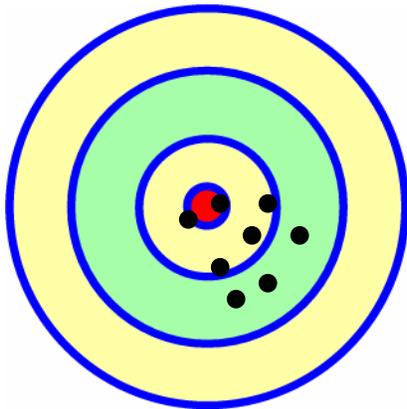
Propiedades de Estimadores

Prof. Rodrigo Abt B.

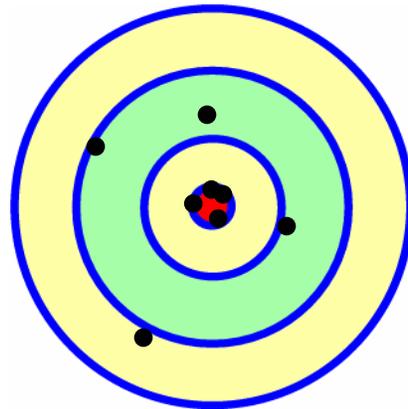
rabt@dim.uchile.cl

Comparaciones entre Estimadores (1)

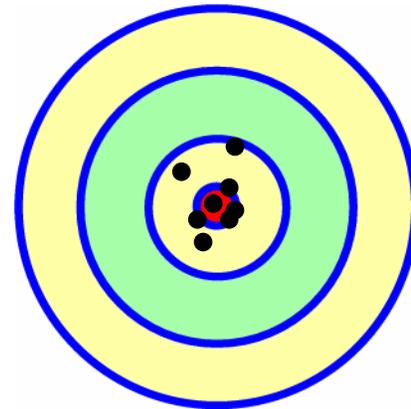
- Supongamos que un dirigente deportivo debe escoger un tirador de tiro al blanco para los próximos Juegos Olímpicos. Para ello el dirigente convoca a 3 participantes y los somete a pruebas de tiro, obteniendo los siguientes resultados:



Tirador 1



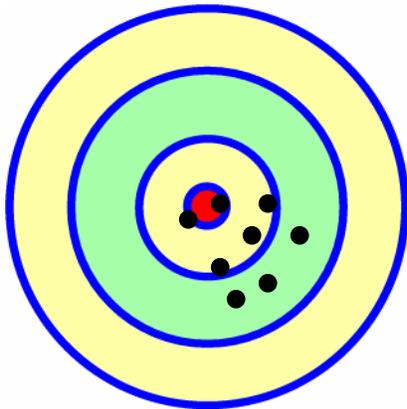
Tirador 2



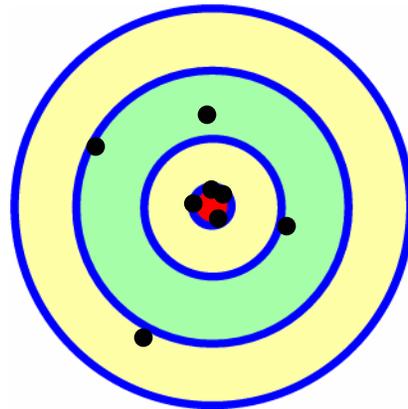
Tirador 3

Comparaciones entre Estimadores (2)

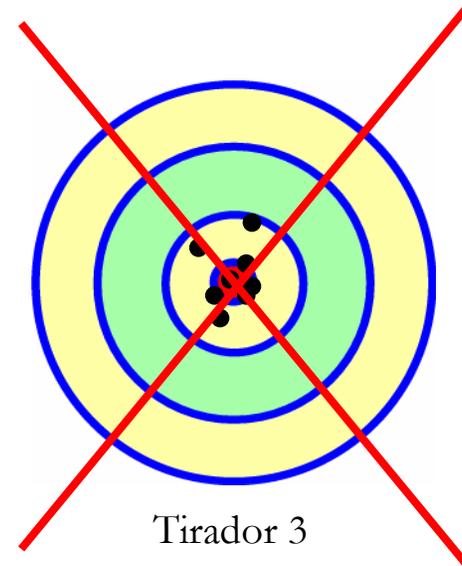
- Claramente el tirador 3 es nuestra mejor carta, ya que en *promedio* le da al blanco y la *dispersión* de sus tiros es baja.
- ¿Qué sucede si se enferma este último tirador?. ¿Es claro quién es mejor tirador entre el 1 y el 2?



Tirador 1



Tirador 2



Tirador 3

Comparaciones entre Estimadores (3)

- Veamos por tirador:
 - El tirador 1 llega menos al blanco (2 de 8 en el blanco), pero sus tiros son menos dispersos.
 - El tirador 2 llega en 4 de 8 veces al blanco, es decir, en promedio llega más al blanco, pero sus tiros son más dispersos.
- Este problema es análogo al siguiente:

Sean $\hat{\theta}$ y $\tilde{\theta}$ dos estimadores de θ , tales que $\hat{\theta}$ es sesgado pero de varianza pequeña, y $\tilde{\theta}$ insesgado pero de varianza grande.

¿Cuál escojo?

El Error Cuadrático Medio (2)

- Por lo tanto escogeremos aquel estimador que presente el menor ECM.
- Algunas observaciones:
 - El ECM es siempre mayor o igual a 0. Mientras más cercano a 0, mejor es el estimador.
 - Tener $\text{ECM} \rightarrow 0$ es equivalente a:

$$E(\hat{\theta}) \rightarrow \theta \quad \wedge \quad \text{Var}(\hat{\theta}) \rightarrow 0$$

- Si dos estimadores son insesgados, el ECM se reduce a una comparación por varianza.

Comportamiento Asintótico (1)

- Para estudiar las características de un estimador no solo basta con saber el sesgo y la varianza, sino que además es útil hacer un análisis de su comportamiento y estabilidad en el largo plazo, esto es, su comportamiento asintótico.
- Cuando hablamos de estabilidad en largo plazo, se viene a la mente el concepto de *convergencia*.
- Luego, podemos construir sucesiones de estimadores y estudiar el fenómeno de la convergencia.

Comportamiento Asintótico (2)

- En el caso de las variables aleatorias, existen diversos tipos de convergencia, dentro de las cuales podemos distinguir:
 - Convergencia en probabilidad (o débil)
 - Convergencia casi segura (o fuerte)
 - Convergencia en media cuadrática
 - Convergencia en distribución

Convergencia En Probabilidad

Una sucesión de variables aleatorias δ_n se dirá que *converge en probabilidad* a c si:

$$P(|\delta_n - c| > \varepsilon) \rightarrow 0 \quad \forall \varepsilon > 0, n \rightarrow \infty$$

Esta es la denominada ley débil de los grandes números.

Si $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ es un estimador de θ , $\hat{\theta}$ se dirá *consistente* si converge en probabilidad a θ . Lo cual notamos por:

$$\hat{\theta} \xrightarrow{p} \theta$$

Convergencia Casi Segura

Una sucesión de variables aleatorias δ_n se dirá que *casi seguramente* a c si:

$$P(\delta_n \rightarrow c) = 1 \quad n \rightarrow \infty$$

Esta es la denominada ley fuerte de los grandes números.

Si $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ es un estimador de θ , $\hat{\theta}$ que converge casi seguramente a θ , entonces se denotará por:

$$\hat{\theta} \xrightarrow{c.s.} \theta$$

Convergencia En Media Cuadrática

Un estimador $\hat{\theta}$ de θ se dirá que *converge en Media Cuadrática* si:

$$ECM(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] \rightarrow 0 \quad n \rightarrow \infty$$

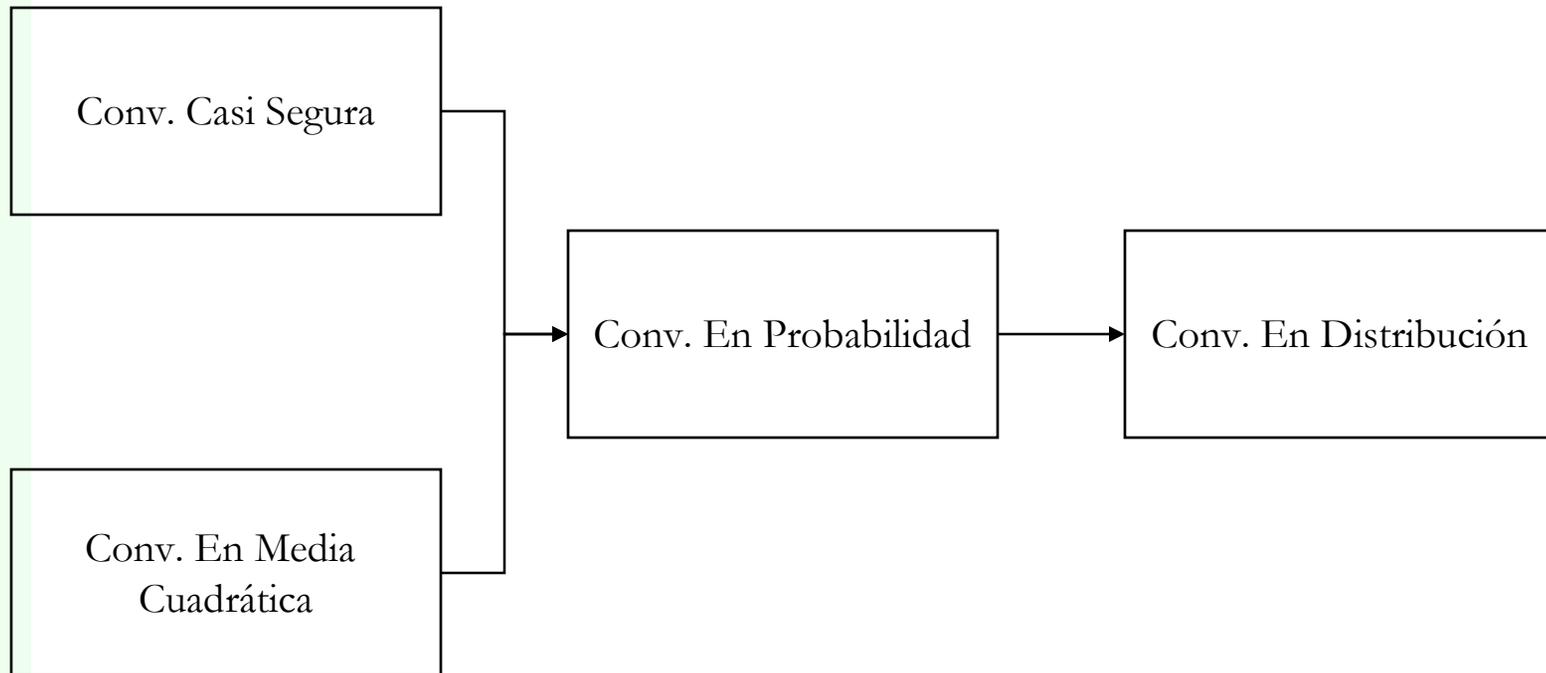
Esto equivale a que:

$$E(\hat{\theta}) \rightarrow \theta \quad \wedge \quad Var(\hat{\theta}) \rightarrow 0$$

Lo que se denota por:

$$\hat{\theta} \xrightarrow{m.c.} \theta$$

Relación Entre Los Tipos De Convergencia



Sobre La Consistencia

- Notemos que dado que la convergencia en Media Cuadrática implica convergencia en probabilidad, una forma muy recurrida de probar la consistencia de un estimador es a través de este último tipo de convergencia.

Concepto de Eficiencia

- Si nos restringimos al espacio de estimadores insesgados, ¿será posible encontrar un estimador que tenga mínima varianza?
- En general la respuesta es afirmativa, sobre todo, cuando las distribuciones involucradas pertenecen a la llamada “Familia de las Exponenciales”, que incluyen a la Normal, Binomial, Exponencial, Poisson , etcétera.
- Esta cota inferior para la varianza se puede determinar bajo ciertas condiciones de regularidad:
 - El rango de los datos no depende del parámetro estimado
 - Existen las segundas derivadas de la función de verosimilitud
 - Las funciones son lo suficientemente continuas como para permitir el intercambio de los operadores de derivación e integración
- Si un estimador alcanza dicha cota se dirá que es *eficiente*

La Información De Fisher (1)

- Para analizar la eficiencia, es útil analizar la sensibilidad de la función de la función de verosimilitud dado el parámetro a estimar en función de los datos.
- Sea X una v.a. con f.d.p. $f(x|\theta)$. Se define entonces la Función de Fisher para una observación x como:

$$I(\theta) = E \left[\left(\frac{\partial \ln f(x|\theta)}{\partial \theta} \right)^2 \right]$$

- La Información de Fisher representa la cantidad de información que trae la observación x respecto del parámetro θ

La Información de Fisher (2)

- Sea x_1, x_2, \dots, x_n una muestra observada de la v.a. X . Se define la Información de Fisher para la muestra como:

$$I_n(\theta) = E \left[\left(\frac{\partial \ln f_n(x|\theta)}{\partial \theta} \right)^2 \right]$$

- Se puede mostrar además que $I_n(\theta) = nI(\theta)$. Esto es, la Información de Fisher para la muestra es “n” veces más informativa que una sola observación.

Cota de Cramer-Rao

- Finalmente, se puede mostrar que si $T=T(X_1, X_2, \dots, X_n)$ es un estimador insesgado de θ , entonces:

$$\text{Var}(T) \geq \frac{1}{I_n(\theta)}$$

- En que $1/I_n(\theta)$ es la denominada Cota de Cramer-Rao para estimadores insesgados.
- Si un estimador alcanza la cota de Cramer-Rao, entonces dicho estimador es *eficiente*

Suficiencia (1)

- Hemos visto que en varios de los problemas de estimación hemos recurrido a estadísticos como la media muestral y la varianza muestral. Dichos estadísticos se caracterizan por ser expresiones que no dependen del orden particular de aparición de los valores.
- Si estimamos el parámetro “ p ” de una Bernoulli, sabemos que la media muestral es el EMV correspondiente. Y además, podemos notar que solo conociendo el promedio de los valores es suficiente para estimar “ p ”. No necesitaríamos conocer los valores individuales de las observaciones.

Suficiencia (2)

- Un estadístico $t=t(x_1, x_2, \dots, x_n)$ y con valores en Θ se dice suficiente para θ si la distribución conjunta de los valores x_1, x_2, \dots, x_n condicional a t no depende de θ .
- En la práctica esta última definición es poco práctica para probar la suficiencia de un estadístico. Por esto se recurre al Teorema de Factorización:

Si $t(x_1, x_2, \dots, x_n)$ es suficiente para θ , y $g(t|\theta)$ es la densidad de la v.a. $T(X_1, X_2, \dots, X_n)$, entonces:

$$f_n(x|\theta) = g(t(x_1, x_2, \dots, x_n)|\theta)h(x_1, x_2, \dots, x_n)$$

- Es decir, podemos escribir la función de verosimilitud de X en función del estadístico suficiente y una función que sólo depende los valores muestrales

Observaciones Finales

- El EMV cumple la propiedad de *invarianza*, esto es, el EMV de una transformación biyectiva $g(\theta)$ es la función evaluada en el EMV, i.e., $EMV(g(\theta))=g(EMV(\theta))$.
- El EMV es siempre función de un estadístico suficiente.
- El δ es el EMV de θ , entonces δ es asintóticamente normal con media θ y varianza $1/I_n(\theta)$.
- Los estimadores de Momentos son siempre consistentes por construcción.