(1)

ESTAD – T201. ~~ESTIMACION PUNTUAL I.~~ 0. Introducción.

2. ~~Absorbido~~ PROPIEDADES de los estimadores puntuales.

3. ERROR CUADRÁTICO MEDIO.

4. ESTIMADORES INSESGADOS, CONSISTENTES Y SUFICIENTES.

---

0. INFERENCIA. INTRODUCCIÓN

~~1. ESTIMACION PUNTUAL I.~~

Recordemos brevemente que la Inferencia Estadística consiste en ~~generalizar~~ (sacar conclusiones) sobre la poblac. que nos interesa estudiar a partir ~~de~~ ~~la~~ la información que nos proporciona una muestra aleatoria basándonos en la Teoría de la Probabilidad.

Si estamos interesados en ~~en~~ estudiar el valor de una característica poblacional θ, la Inferencia que podemos utilizar es:

- Estimación,
- Contrastación

La Estimación consiste en dar un valor aproximado del parám. poblacional a partir de la información ~~proporcionada~~ muestral.

La Contrastación consiste es formular una conjetura (hipótesis) sobre el valor del parámetro poblacional y utilizar la inform. muestral para aceptar ó rechazar dicha hipótesis.

En el caso de la estimación, se puede aproximar:

- Estim. puntual $\rightarrow$ valor concreto
- Estim. por intervalos de confianza $\rightarrow$ intervalo

En ambos métodos se utiliza un estadístico (función real de la muestra) para estimar.

La ventaja de la estimación puntual es que da un valor concreto del estimador que puede sustituirse directamente en el parámetro, pero sólo cuenta con las propiedades como garantía.

La ventaja de la estimación por intervalos es que el intervalo va acompañado de un grado de confianza de que el verdadero valor del parámetro se encuentre dentro del intervalo, pero su inconveniente es que ofrece infinitas soluciones.

1 _ ESTIMACIÓN PUNTUAL I.

La estimación puntual consiste en obtener un único número, calculado a partir de las observaciones muestrales, y que es utilizado como estimación del parámetro poblacional θ. Se le llama estimación puntual porque a ese nº, que se utiliza como estimación puntual de θ, se le puede asignar un punto de la recta real.

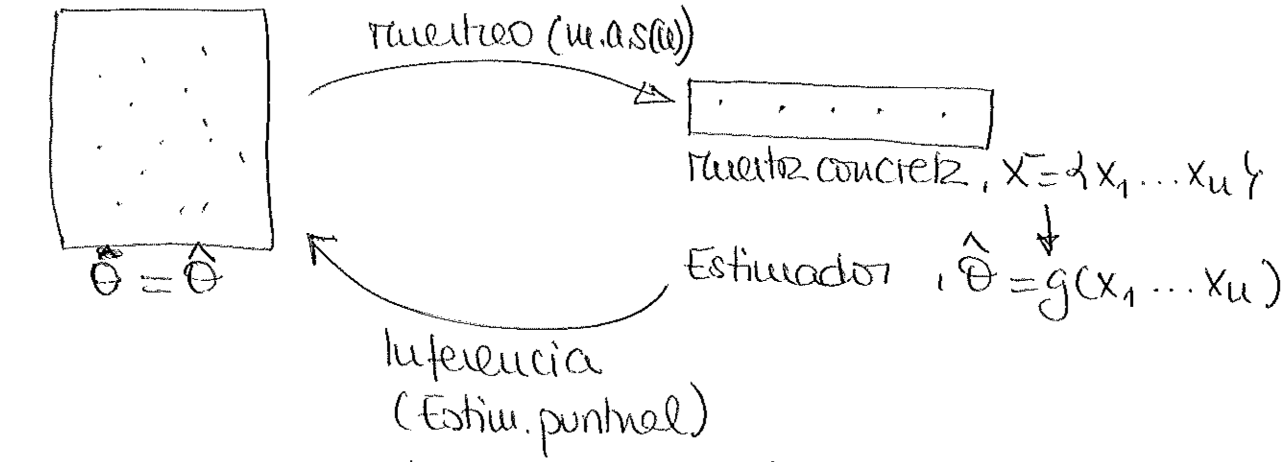
Esquema de la estimac. puntual:

Población

$\xi \rightarrow F(x,\theta)$ θ desconocido

Muestra

Espacio muestral, $\mathbb{R}^n = \{X_1 \dots X_n\}$



Para una población ξ (realización de una v.a.), representada a partir de su función de distribución $F(x,\theta)$, que depende de un parámetro θ cuyo valor concreto se desconoce, se toma una muestra aleatoria (normalmente en poblac. infinitas se utiliza m.a.s.) con n elementos que, al ser el muestreo aleatorio tb será una v.a. A partir de la muestra se construye un estadístico (~~et~~ depende del parámetro) que seguirá siendo una v.a. que se utiliza para estimar el parámetro, por eso se llama estimador.

Para una muestra concreta se obtendrá un valor concreto del ~~estimación~~ estimador, que recibe el nombre de estimación puntual del parámetro poblacional.

Antes de continuar, merece la pena detenerse en los tres conceptos mencionados anteriormente para no confundirlos.

Parámetro → Constante con valor desconocido. ~~Relacionado~~ Característica poblacional, no depende de la muestra y es único (≠ muestras, 1 parám).

Estimador → Estadístico (→ función de la muestra) que se utiliza para ofrecer una aproximación del parámetro desconocido. Es una variable aleatoria, no una cte → tiene distrib. de probab. que depende de los valores muestrales.

Estimación → Valor de la v.a. estimador para una muestra concreta → es una cte.
≠ muestras, = estimador ⇒ ≠ estimaciones

Para seleccionar el estadístico que utilizaremos como estimador del parámetro poblacional tendremos en cuenta las propiedades de la distribución muestral del estadístico, por lo que:

1º. Estudiar distribución de los estad. en el muestreo.
2º. Conocer las propiedades deseables de los estimadores puntuales, para estudiar su bondad.
3º. Métodos de obtención

Afortunadamente, las propiedades y los métodos nos sugieren los estimadores que se nos ocurren de manera natural.

Para la media poblacional μ, la media muestral $\bar{x}$

Para la varianza poblacional σ^2, una corrección de la varianza muestral, la cuasivarianza muestral (en términos infinitos es igual).

Para la proporción poblacional P, la proporción muestral p.

2_ PROPIEDADES de los estim. puntuales

Dada una población ξ, con función de distribución $F(x,\theta)$ donde θ es un parámetro poblacional desconocido que pretendemos estimar con ayuda de la muestra aleatoria simple de tamaño n $(X_1, X_2 \dots X_n)$, a partir del estimador $\hat{\theta} = g(X_1, X_2, \dots X_n)$. ($X_i \to$ v.a. , $x_i \to$ realización muestral)

$\hat{\theta}$ es un estadístico $\Rightarrow$ v.a. $\Rightarrow$ distribución muestral, su media y su varianza.

Nos interesa encontrar la función real de la muestra que nos proporcione el mejor estimador de θ, para lo que tendremos que utilizar alguna medida que nos permita dar algún criterio para seleccionar el mejor estimador. Esta medida será el error cuadrático medio del estimador.

A partir del ECM se deducen algunas propiedades deseables del estimador $\to$ insesgadez
eficiencia
consistencia

Pero tb. hay otras no tan evidentes, pero muy importantes $\to$ suficiencia
invarianza
robustez.

3_ ERROR CUADRÁTICO MEDIO.

Para establecer la bondad de un estimador partiremos del hecho de ser deseable conocer si la estimación se encuentra lejos o cerca del valor verdadero, siempre desconocido $\Rightarrow$ - Práctica: imposible de comprobar
- Teoría: Seguir el planteamiento.

Parámetro, θ
Estimador, $\hat{\theta}$

$$\left.\begin{matrix}\text{Parámetro}, \theta \\ \text{Estimador}, \hat{\theta}\end{matrix}\right\} \begin{matrix}\text{Error} = \hat{\theta} - \theta \text{ ó } \theta - \hat{\theta} \rightarrow \text{no importa signo} \\ \text{Error}^2 = (\hat{\theta} - \theta)^2\end{matrix}$$

Al ser $\hat{\theta}$ una v.a. no tiene un valor concreto, luego el error será distinto para cada muestra $\Rightarrow$ estudiamos el error en términos de su esperanza, del valor esperado de la v.a. (medida del error global), que es justamente el error cuadrático medio, desviación cuadrática media ó acuracidad(*) del estimador.

Un valor pequeño del ~~er~~ ECM indicará que, en media, el estimador no se encuentra lejos del parámetro θ, y cuanto mayor sea el ECM más lejos estará en media el estimador del verdadero valor del parámetro.

$$ECM(\hat{\theta}) = E[(\hat{\theta}-\theta)^2] = E[\hat{\theta}-\theta]^2$$

Acuracidad $\equiv$ concentración de las estimaciones respecto al ~~del~~ verdadero valor del parámetro.

Precisión $\equiv$ concentración de las estimaciones respecto al ~~del~~ valor medio del estimador.

Desarrollando la expresión tendremos:

tomando esperanzas

$$ECM(\hat{\theta}) = E[(\hat{\theta}-\theta)^2] = E[\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2] =$$

$$= E[\hat{\theta}^2] - 2\theta E[\hat{\theta}] + \theta^2 =$$

Sumando y restando $E[\hat{\theta}]^2$

$$= \left(E[\hat{\theta}^2] - E[\hat{\theta}]^2\right) + \left(E[\hat{\theta}]^2 - 2\theta E[\hat{\theta}] + \theta^2\right) =$$

$$= Var(\hat{\theta}) + (E[\hat{\theta}] - \theta)^2 =$$

$$ECM(\hat{\theta}) = Var(\hat{\theta}) + [sesgo(\hat{\theta})]^2$$

Por lo que el ECM del estimador se puede descomponer en la suma de dos cantidades no negativas:

- Varianza del estimador, $V(\hat{\theta}) = E[\hat{\theta}^2] - (E[\hat{\theta}])^2$
- Cuadrado del sesgo del estimador, $sesgo^2(\theta) = (E(\hat{\theta}) - \theta)^2$

- Lo deseable es que el ECM sea pequeño, que equivale a desear que estas dos cantidades sean pequeñas.

$$\min ECM(\hat{\theta}) \Rightarrow \begin{cases} \min V(\hat{\theta}) \longrightarrow \hat{\theta} \text{ var. mínima.} \\ sesgo(\hat{\theta}) = 0 \longrightarrow \hat{\theta} \text{ insesgado.} \end{cases}$$

- Pero no siempre es posible encontrar un $\hat{\theta}$ que minimice el ECM para todos los posibles valores de θ. Puede ser que dependiendo del valor de θ cambie el estimador que minimice el ECM.

- Por tanto, la utilización del ECM para la elección de un buen estimador es insuficiente, siendo necesario dar otros criterios $\rightarrow$ dependerá de otras propiedades.

4 – ESTIMADORES INSESGADOS, CONSISTENTES Y SUFICIENTES

4.1. INSESGADEZ:

Hemos definido el sesgo del estimador como la diferencia del valor esperado del estimador y el verdadero valor del parámetro.

$$\text{Sesgo}(\hat{\theta}) = E[\hat{\theta}] - \theta$$

Para que el $ECM(\hat{\theta})$ sea mínimo lo han de ser el sesgo y la varianza del estimador.

El sesgo se hace mínimo cuando vale 0, es decir, cuando la esperanza matemática del estimador coincide con el verdadero valor del parámetro. Esta es la propiedad de insesgadez.

Definición: $\hat{\theta} = g(X_1, \dots X_n)$ es un parámetro insesgado o centrado del parámetro θ si la esperanza matemática del estimador $\hat{\theta}$ es igual al parámetro θ,

$$E[\hat{\theta}] = \theta$$

para todos los valores de θ, y entonces su sesgo es 0.

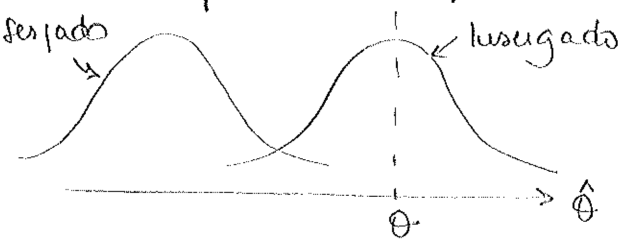
$$b(\hat{\theta}) = \text{Sesgo}(\hat{\theta}) = E[\hat{\theta}] - \theta = 0$$

En caso contrario, diremos que el estimador es sesgado ó no centrado.

En general, la esperanza de un estimador se puede descomponer en la suma del parámetro y el sesgo.

$$E[\hat{\theta}] = \theta + b(\hat{\theta})$$

El sesgo del estimador equivale al error sistemático, no aleatorio, positivo o negativo. El signo del sesgo tiene una interpretación importante: si es positivo, el estimador sobreestima el valor del parámetro y si es negativo lo infraestima.



La insesgadez es una propiedad de la var. aleatoria estimador, no de un valor concreto del estimador, y ha de entenderse como propiedad "media", en el sentido de que si tomamos todas las posibles m.a.s. de un tamaño concreto, se calcula para cada una de ellas el valor del estimador (estimación puntual) y se halla la media de todas las estimaciones, el resultado es igual al valor del parámetro si el estimador es insesgado.

La definición de insesgadez indica la manera de verificar si un estimador es insesgado, no hay más que calcular su esperanza matemática.

En Ectnia no es una prop. tan importante

Fotocopiar tabla 2.2 de Casas (pág 97)

~~Pto~~

Propiedades de los estimadores insesgados:

P1 _ La combinación lineal convexa de dos estimadores insesgados, tb es un estimador insesgado.

$$\left.\begin{array}{l} E[\hat{\theta}_1]=\theta \\ E[\hat{\theta}_2]=\theta \end{array}\right\} \hat{\theta} = c\hat{\theta}_1 + (1-c)\hat{\theta}_2 \ / \ E[\hat{\theta}]=\theta .$$

P2 _ Si existen dos estimadores con el mismo sesgo, entonces existen infinitos estimadores de esa clase. Tómese la c.l. convexa

P3 _ La media muestral es un estimador insesgado de la media poblacional.

$$E[\bar{x}] = E\left[\frac{x_1+\dots+x_n}{n}\right] = \frac{1}{n}\left(E(x_1)+\dots+E(x_n)\right) \overset{E(x_i)=\mu}{=} \frac{n\mu}{n} = \mu .$$

P4 _ La varianza muestral es un estimador sesgado de la varianza poblacional, pero la cuasivarianza muestral es insesgado de la varianza poblacional.

$$S^2 = \frac{\sum(x_i-\bar{x})^2}{n} \rightarrow E[S^2] = \frac{n-1}{n}\sigma^2 = \left(1-\frac{1}{n}\right)\sigma^2$$

Si tomamos la cuasivarianza.

$$S_1^2 = \frac{n}{n-1}S^2 \ , \ E[S_1^2] = \frac{n}{n-1}E[S^2] = \sigma^2$$

P5 _ Los momentos muestrales respecto al origen son insesgados respecto a los momentos poblacionales respecto al origen: $E[a_r] = \alpha_r$.

En los momentos muestrales respecto a la media el resultado es distinto.

P6 _ Un estimador es asintóticamente insesgado si es sesgado, pero su sesgo tiende a 0 cuando el tamaño muestral tiende a ∞.

$$E[\hat{\theta}] = \theta + b(\hat{\theta}) \xrightarrow[n\to\infty]{} \theta$$

4.2. CONSISTENCIA

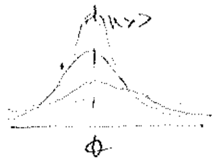
Parece lógico esperar que un estimador será tanto mejor cuanto mayor sea el tamaño de la muestra. Al aumentar el tamaño de la muestra, la información que proporciona es más completa y la varianza del estimador suele ser menor y su distribución muestral tenderá a encontrarse más concentrada alrededor del parámetro que queremos estimar.

Además, teniendo en cuenta el tma. de Glivenko-Cantelli, cuando el tamaño muestral es suficientemente grande, la muestra puede llegar a proporcionar una información casi exacta de la población, por lo que el valor del estimador tiende a coincidir con el valor del parámetro.

Consistencia en probabilidad $\hat{\theta}_n \xrightarrow{P} \theta$ → Convergencia débil.

Definición: Una sucesión de estimadores $\{\hat{\theta}_n\}$ es consistente en probabilidad, si la sucesión converge en probabilidad hacia el parámetro θ.

$\forall \varepsilon > 0, \lim_{n \to \infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1 \quad , \forall \theta$.



y cada elemento de la sucesión se llama estimador consistente.

Tma de consist. pág 25

Partiendo del planteamiento del ECM, el error cometido por un estimador es $\hat{\theta} - \theta$, y pretendemos que sea muy difícil encontrarnos con valores elevados de ese error, es decir, que en el límite, la probabilidad del suceso $\{|\hat{\theta} - \theta| \geq \varepsilon\}$ sea 0, que resulte muy raro que el error cometido supere una cantidad tan reducida como deseemos ε, consistencia en probabilidad,

$\{\hat{\theta}_n\}$ sucesión formada por $\hat{\theta}_1 = g(X_1)$, $\hat{\theta}_2 = g(X_1, X_2)$, ... $\hat{\theta}_n = g(X_1 \dots X_n)$ ...

o bien consiguiendo que los errores $|\hat{\Theta}-\Theta|$ sean muy pequeños si el ECM = 0 al límite, convergencia en ~~medi~~ error cuadrático medio. Tb. se puede considerar la consistencia casi segura.

Después de la definición de ~~convergencia~~ consistencia en probabilidad:

Aplicando el teorema de Chebychev:

$$P(|\hat{\Theta}_n-\Theta| \geq \varepsilon) \leq \frac{E(\hat{\Theta}_n-\Theta)^2}{\varepsilon^2}$$

$$E(\hat{\Theta}_n-\Theta)^2 = ECM(\hat{\Theta}_n) = V(\hat{\Theta}_n) + b(\hat{\Theta}_n)$$

Tomando límites en la desig. Chebichev:

$$\lim_{n\to\infty} P(|\hat{\Theta}_n-\Theta| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{E(\hat{\Theta}_n-\Theta)^2}{\varepsilon^2} = \lim_{n\to\infty} V(\hat{\Theta}_n) + \lim_{n\to\infty} b^2(\hat{\Theta}_n)$$

lo que significa que para que el estimador sea consistente en probabilidad ($\lim = 0$), ~~se~~ se tiene que verificar

$\lim_{n\to\infty} V(\hat{\Theta}_n) = 0$ → distrib. degenerada

y

$\lim_{n\to\infty} b^2(\hat{\Theta}_n) = 0$ → estimador asintóticamente insesgado

Luego si el sesgo y la varianza de un estimador tienden a 0 cuando $n\to\infty$, el estimador es consistente.

<u>Propiedades</u> de los estimadores <u>consistentes en probabilidad</u>

P1_ Los momentos muestrales respecto al origen son estimadores consistentes de los correspondientes poblacionales.

P2_ Los momentos muestrales respecto a la media son estimadores consistentes de los correspondientes poblacionales.

Consistencia en error cuadrático medio ó en media cuadrática

$\{\hat{\theta}_n\}$ es consistente en media cuadrática (ECM) par el parámetro θ si se verifica:

$$\lim_{n\to\infty} E[\hat{\theta}_n-\theta]^2=0$$

y cada elemento de la sucesión se dirá que es un estimador consistente en media cuadrático (ECM).

Recordando la definición de ECM y tomando límites, deberá verificarse que tanto la varianza como el sesgo han de tender a 0.

Tma: Consistencia en ECM $\Longrightarrow$ Consistencia en Probab
$\not\Longleftarrow$

Consistencia en ECM $\to \lim\limits_{n\to\infty} E[\hat{\theta}_n-\theta]^2=0$

Consistencia en Probab $\to P(|\hat{\theta}_n-\theta|>\varepsilon) \xrightarrow[n\to\infty]{} 0, \ \forall\varepsilon>0, \forall\theta$

Por la desig. Chebichev: $P(|\hat{\theta}_n-\theta|^2>\varepsilon^2) \leq \dfrac{E[\hat{\theta}_n-\theta]^2}{\varepsilon^2}$??

tomando límites: $\lim$ $\leq \lim = 0$

0

si $\hat{\theta}$ consist. en ECM

Consistencia casi segura

$\{\hat{\theta}_n\}$ es consistente casi seguro para θ si se verifica:

$$P\left(\lim_{n\to\infty} \hat{\theta}_n=\theta\right)=1$$

y cada elemento de la sucesión se dirá que es un estimador consistente casi seguro.

Tma: Consistencia CASI SEGURO $\Longrightarrow$ Consistencia en Probab.

4.3. SUFICIENCIA

Un estadístico es suficiente cuando resume el conjunto de información relevante contenida en la muestra, y ningún otro estadístico puede proporcionar información adicional acerca del parámetro desconocido de la población.

Al ser el estimador una función de los valores muestrales, y por tanto un estadístico, parece razonable extender a los estimadores la propiedad de suficiencia.

Definición: Un estimador $\hat{\theta} = T(X_1 \ldots X_n)$ es suficiente para el parámetro θ si la distribución condicionada de la muestra dado el valor del estimador, no depende del parámetro θ. $P(X/\hat{\theta}=t)$ no depende de θ

Esta definición nos permite comprobar si un estadístico dado es suficiente, pero si queremos obtenerlo utilizaremos el tma. de factorización de Fisher-Neyman:

Un estadístico es suficiente si y sólo si podemos descomponer la función de probabilidad (cuantía o densidad) en producto de dos factores no negativos, uno que depende solamente del parámetro y de la muestra a través del estadístico y el otro factor que no depende de θ.

AMPLIAR (Canavos)

~~Sufic~~
Eficiente $\Rightarrow$ Suficiente

Suficiente $\not\Rightarrow$ Eficiente