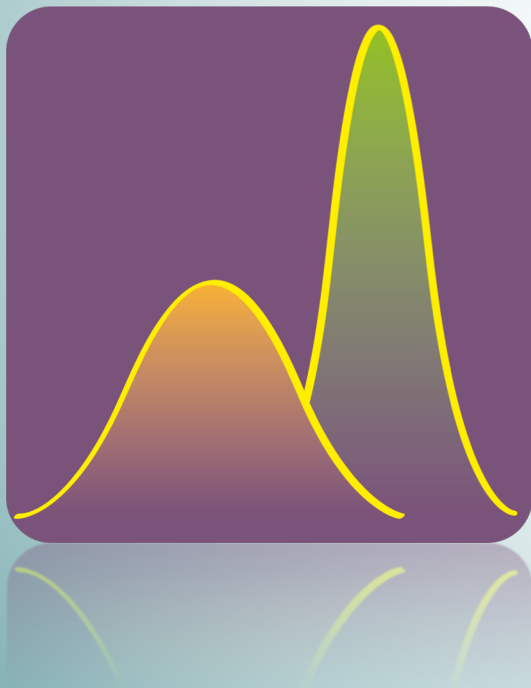


Estadística II

TEMA 4. MUESTREO Y ESTIMACIÓN PUNTUAL



Marta Guijarro Garvi
David Gutiérrez Sobrao

DEPARTAMENTO DE ECONOMÍA

Este material se publica bajo licencia:

[Creative Commons BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)



Tema 4: Muestreo y estimación puntual

Índice de contenidos

- Principales conceptos: población y muestra.
- Estadísticos en el muestreo de una población: media, varianza y cuasivarianza muestral.
- Muestreo de una población normal.
- Muestreo de una población de Bernoulli: proporción muestral.
- Estimación puntual: propiedades de los estimadores y método de la máxima verosimilitud.

Tema 4: Muestreo y estimación puntual

Resultados de aprendizaje

- Comprender los conceptos de población, muestra y estadístico.
- Entender la naturaleza aleatoria de la muestra y de los estadísticos.
- Conocer los principales estadísticos en el muestreo: media, varianza, cuasivarianza y proporción.
- Establecer las diferencias entre parámetro, estimador y estimación.
- Conocer las propiedades deseables de un estimador.
- Saber aplicar el método de la máxima verosimilitud para la obtención de estimadores de un parámetro.

Principales conceptos: población

La **inferencia estadística** tiene como finalidad asignar a un fenómeno aleatorio el modelo probabilístico que mejor se ajuste a los datos obtenidos de su observación.

Este modelo probabilístico se expresa mediante la distribución de probabilidades de una variable aleatoria, X , denominada **población**.

El propósito de la **inferencia paramétrica** es realizar una “aproximación razonable” a ciertos valores desconocidos de la distribución de probabilidades de la población. Estos valores se denomina **parámetros**.

Principales conceptos: población

La **distribución de probabilidades de la población** se concreta con $f(x; \theta)$, siendo θ el parámetro, el cual toma valores en un subconjunto de la recta real denominado **espacio paramétrico**, Θ . Cuando la población es discreta, $f(x; \theta)$ es su función de probabilidad y, cuando es continua, $f(x; \theta)$ es su función de densidad.

La media de la distribución de probabilidades, media de la población o **media poblacional** es $E(X) = \mu$. La varianza de la distribución de probabilidades, varianza de la población o **varianza poblacional** es $Var(X) = \sigma^2$.

Ejemplo 4.1. La distribución de probabilidades de una población de Bernoulli de parámetro p queda determinada por la función de probabilidad:

$$f(x; p) = p^x \cdot (1 - p)^{1-x} \quad x = 0, 1.$$

Ejemplo 4.2. La distribución de probabilidades de una población exponencial negativa de parámetro a queda determinada por la función de densidad:

$$f(x; a) = a \cdot e^{-a \cdot x} \quad x > 0.$$

Principales conceptos: muestra

Una **muestra aleatoria simple** de tamaño n de una población X , $(X_1, \dots, X_n)$, es una variable aleatoria n -dimensional formada por n variables aleatorias independientes y con la misma distribución que X .

La observación de la muestra es la **realización de la muestra**, $(x_1, \dots, x_n)$.

La **distribución de probabilidades de la muestra** queda determinada por su función de probabilidad, cuando la población es discreta, y por su función de densidad, cuando la población es continua:

$$f(x_1, \dots, x_n; \theta) = f(x_1; \theta) \cdot \dots \cdot f(x_n; \theta).$$

Estadísticos en el muestreo de una población: media y varianza muestral

Un **estadístico** es una variable aleatoria función de la muestra:

$$T(X_1, \dots, X_n; \theta).$$

Para una realización de la muestra, $(x_1, \dots, x_n)$, se obtiene el **valor del estadístico**:

$$T(x_1, \dots, x_n; \theta).$$

Estadísticos en el muestreo de una población: media y varianza muestral

Sea X una población con $E(X) = \mu$ y $Var(X) = \sigma^2$ y sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de tamaño n de X . Los principales estadísticos en el muestreo y sus correspondientes expresiones figuran en la tabla siguiente.

Denominación	Media muestral	Varianza muestral	Cuasivarianza muestral
Expresión	$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^n X_i$	$S^2 = \frac{1}{n} \cdot \sum_{i=1}^n (X_i - \bar{X})^2$	$S_C^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (X_i - \bar{X})^2$
Propiedad: $n \cdot S^2 = (n-1) \cdot S_C^2$			

Estadísticos en el muestreo de una población: media muestral

En la tabla siguiente se recogen la media, la varianza y la desviación típica de la variable media muestral, donde $E(X) = \mu$ y $Var(X) = \sigma^2$ son la media y la varianza de la población, respectivamente.

Media	Varianza	Desviación típica o error típico
$E(\bar{X}) = \mu$	$Var(\bar{X}) = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

Ejemplo 4.3. Sea X una población con distribución binomial negativa de parámetros r y p y sea una muestra aleatoria simple de tamaño n de X , $(X_1, \dots, X_n)$. Como la media y la varianza de la población son, respectivamente, $E(X) = r \cdot (1-p)/p$ y $Var(X) = r \cdot (1-p)/p^2$, entonces, la media de la media muestral es $E(\bar{X}) = r \cdot (1-p)/p$, su varianza es $\sigma_{\bar{X}}^2 = r \cdot (1-p)/(n \cdot p^2)$ y su error típico o desviación típica es $\sigma_{\bar{X}} = \sqrt{r \cdot (1-p)/(p \cdot \sqrt{n})}$.

Tema 4. Muestreo y estimación puntual

Muestreo de una población normal

Sea X una población con distribución normal de parámetros μ y σ , y sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de tamaño n de X . La variable media muestral, $\bar{X}$, sigue también una distribución normal.

Además, como la media y la varianza de una población normal coinciden con los parámetros de su distribución, aplicando el resultado general, la media y la varianza de la media muestral son, en este caso, $E(\bar{X}) = \mu$ y $\sigma_{\bar{X}}^2 = \sigma^2/n$, respectivamente.

En definitiva, la media muestral de una población normal sigue una distribución normal de parámetros μ y $\sigma/\sqrt{n}$.

Muestreo de una población de Bernoulli: proporción muestral

Sea X una población con distribución de Bernoulli de parámetro p y sea una muestra aleatoria simple de tamaño n de X , $(X_1, \dots, X_n)$.

En este caso, la variable media muestral es igual a la proporción de éxitos en la muestra o **proporción muestral**:

$$\hat{p} = \hat{p}(X_1, \dots, X_n) = \bar{X} = \frac{1}{n} \cdot \sum_{i=1}^n X_i.$$

Además, como la media y la varianza de una población de Bernoulli son, respectivamente, $E(X) = p$ y $Var(X) = p \cdot (1 - p)$, aplicando el resultado general, se cumple que la media y la varianza de la proporción muestral son $E(\bar{X}) = p$ y $\sigma_{\hat{p}}^2 = p \cdot (1 - p) / n$, respectivamente.

Asimismo, si $n > 30$, la distribución de la proporción muestral es *aproximadamente*¹ una normal de parámetros p y $p \cdot (1 - p) / n$.

¹Este resultado se obtiene a partir del Teorema del Límite Central.

Estimación puntual: concepto de estimador

Sea X una población con distribución determinada por $f(x; \theta)$ ($\theta \in \Theta$) y sea una muestra aleatoria simple de tamaño n de la población, $(X_1, \dots, X_n)$. Un **estimador** del parámetro es un estadístico, $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$, que toma valores en el espacio paramétrico.

En consecuencia, para cualquier realización de la muestra, $(x_1, \dots, x_n)$, se cumple:

$$\hat{\theta} = \hat{\theta}(x_1, \dots, x_n) \in \Theta,$$

valor que se denomina **estimación**.

Ejemplo 4.4. Sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de una población de Bernoulli de parámetro p . El estadístico suma, $T_1(X_1, \dots, X_n; p) = \sum_{i=1}^n X_i$, no es un estimador porque no toma valores en el espacio paramétrico, $[0, 1]$; el estadístico proporción muestral, $T_2(X_1, \dots, X_n; p) = \hat{p}(X_1, \dots, X_n) = \sum_{i=1}^n X_i / n$, sí es un estimador del parámetro pues toma valores en el espacio paramétrico.

Estimación puntual: propiedades de un estimador

Sea X una población con distribución determinada por $f(x; \theta)$. Sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de tamaño n de X y $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ un estimador del parámetro. Las siguientes propiedades son deseables para un estimador:

Insesgadez. Se dice que un estimador es insesgado cuando “se mueve” en torno al parámetro, esto es, cuando su media coincide con el parámetro: $E(\hat{\theta}) = \theta$. Cuando esta propiedad se cumple en la medida en que aumenta el tamaño de la muestra, se dice que el estimador es **asintóticamente insesgado**.

Eficiencia relativa. Sean dos estimadores insesgados, $\hat{\theta}_1$ y $\hat{\theta}_2$. Se dice que $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$ si $Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$.

Suficiencia. Un estimador es suficiente cuando mantiene toda la información sobre el parámetro que contiene la muestra.

Consistencia. Un estimador es consistente si, para todo número real positivo, ε , se cumple: $\lim_{n \rightarrow \infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1$.

Estimación puntual: Método de la máxima verosimilitud

Sea X una población con distribución determinada por $f(x; \theta)$. Sea una muestra aleatoria simple de tamaño n de X , $(X_1, \dots, X_n)$.

Para una realización de la muestra, $(x_1, \dots, x_n)$, la **función de verosimilitud** del parámetro es

$$L(x_1, \dots, x_n; \theta) = f(x_1; \theta) \cdot \dots \cdot f(x_n; \theta).$$

La **estimación máximo verosímil** del parámetro es el valor de θ que maximiza la función de verosimilitud.

Ejemplo 4.5. Sea X una población con distribución exponencial negativa de parámetro a y sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de X . Dada una realización de la muestra, $(x_1, \dots, x_n)$, la función de verosimilitud del parámetro es

$$L(x_1, \dots, x_n; a) = f(x_1; a) \cdot \dots \cdot f(x_n; a) = a^n \cdot e^{-a \cdot \sum_{i=1}^n x_i}.$$

Maximizando el logaritmo neperiano² de la función anterior, se obtiene la estimación máximo verosímil del parámetro, $\hat{a}(x_1, \dots, x_n) = 1/\bar{x}$, siendo el estimador máximo verosímil del parámetro, $\hat{a}(X_1, \dots, X_n) = 1/\bar{X}$.

²Puesto que el logaritmo neperiano es una función creciente, maximizar la función de verosimilitud es equivalente a maximizar su logaritmo. Este procedimiento resulta más sencillo en la práctica.

Tema 4. Muestreo y estimación puntual