Tema 1: Introducción a la inferencia estadística

Estimación I

Grado en Estadística Aplicada

Curso 2019-2020

En el **cálculo de probabilidades** se estudian varios aspectos de las distribuciones de probabilidad, asumiendo que la distribución considerada es **conocida**.

No obstante, en la práctica, la **distribución de una variable** de interés puede no ser conocida.

DEFINICIÓN: INFERENCIA ESTADÍSTICA

La inferencia estadística (o estadística matemática) es un área de la estadística compuesta por una serie de técnicas (basados en el cálculo de probabilidades) que permiten **obtener información** acerca de la ley de probabilidad de un fenómeno aleatorio, denominado población, **mediante la observación** del mismo.

Para obtener información sobre dicho fenómeno aleatorio, se llevan a cabo **repeticiones** del mismo o se seleccionan **individuos** de la población. El conjunto de dichas repeticiones/individuos recibe el nombre de **muestra**.

DEFINICIÓN: MUESTREO ALEATORIO SIMPLE

El muestreo aleatorio simple es un procedimiento para seleccionar muestras en el que todos los individuos de la población tienen la **misma probabilidad** de ser elegidos.

Matemáticamente, si X es la variable aleatoria población, una muestra aleatoria simple (m.a.s.) de tamaño n son n **variables aleatorias independientes e identicamente distribuidas** (v.a.i.i.d.) $X_1, X_2, \ldots, X_n$ con la misma distribución que X.

X_i representa la **aleatoriedad** del i-ésimo individuo elegible en la muestra.

Sea $f_X()$ la función de masa o de densidad (según corresponda) de la v.a. X. Entonces, la **función de masa/densidad conjunta** de la m.a.s. viene dada por:

$$f(x_1, \ldots, x_n) = f_X(x_1) \cdots f_X(x_n) = \prod^{n} f_X(x_i),$$

TIPOS DE INFERENCIA

- **Inferencia paramétrica**.- En este caso, se asume que la distribución de probabilidad del fenómeno de interés pertenece a una **familia paramétrica**, como pueden ser la distribución binomial o la normal, pero se **desconoce** el (o los) parámetros que rigen dicha distribución.
  Así el objetivo de la inferencia paramétrica es obtener información sobre el valor de dicho(s) **parámetro(s)**.
- **Inferencia no paramétrica**.- En este caso se **desconoce la distribución** de probabilidad del fenómeno de interés, por lo que se intentará obtener información sobre la misma u **otras cuestiones** como, por ejemplo, la independencia entre variables aleatorias.

INFERENCIA PARAMÉTRICA: TÉCNICAS

Dentro de la inferencia paramétrica, podemos encontrar 3 tipos de técnicas:

- **Estimación puntual**.- Consiste en hacer **un pronóstico** (técnicamente una estimación) sobre el valor del parámetro desconocido. Se busca que el valor proporcionado sea lo más cercano posible al verdadero valor del parámetro.
- **Estimación por intervalos**.- En este caso, en lugar de ofrecer un único valor, se proporciona un **rango de valores** en el que existe una probabilidad alta (normalmente el 95 %) de encontrar el valor del parámetro desconocido.
- **Constrastes de hipótesis**.- Consiste en proporcionar una **regla de decisión** para elegir entre posibles conjuntos de valores para el parámetro desconocido (por ejemplo, si es igual a 0 o no).

En esta asignatura estudiaremos **inferencia paramétrica**, resolviendo los problemas de estimación puntual y por intervalos.

Por ello, vamos a asumir que la v.a. población X se distribuye según una **familia parámetrica** de distribuciones conocida $\mathcal{F}$, pero cuyo parámetro (o vector de parámetros) θ es **desconocido**: $\mathcal{F} = \{f_{X,\theta}(x), \theta \in \Theta\}$.

Θ representa el conjunto de posibles valores de θ y recibe el nombre de **espacio paramétrico**.

Para reforzar la idea de que los cálculos **dependen del valor del parámetro**, la función de verosimilitud se suele denotar de la siguiente manera:

$$f_\theta(x_1, \ldots, x_n) = f_{X,\theta}(x_1) \cdots f_{X,\theta}(x_n) = \prod_{i=1}^{n} f_{X,\theta}(x_i),$$

Obtén la función de verosimilitud muestral si $X_1, X_2, \ldots, X_n$ es una m.a.s. de X, que se distribuye según: a) $B(1, p)$; b) $U(a, b)$ y c) $\Gamma(a, p)$.

EJERCICIO 1

Obtén la función de verosimilitud muestral si $X_1, X_2, \ldots, X_n$ es una m.a.s. de X, que se distribuye según: a) $B(m,p)$; b) $P(\lambda)$ y c) $N(\mu,\sigma)$.

Como acabamos de ver, es relativamente sencillo obtener información sobre la muestra completa a través de su función de verosimilitud.

No obstante, generalmente no necesitamos información sobre **toda la muestra** si no solo sobre algún aspecto concreto de la misma.

DEFINICIÓN: ESTADÍSTICO MUESTRAL

Sea $X_1, \dots, X_n$ una m.a.s. de X. Llamamos **estadístico muestral** (o estimador) a cualquier función de la muestra T, que no dependa de parámetros desconocidos:

$$T : \Omega^n \to \mathbb{R}^k$$

Generalmente, $k = 1$, en cuyo caso podemos encontrar, por ejemplo, la media, el mínimo o el máximo.

Nótese que, dado que $X_1, X_2, \dots, X_n$ son v.a., $T(X_1, X_2, \dots, X_n)$ es también una **v.a.** y, por tanto, tendrá una distribución.

Algunos de los estadísticos muestrales más utilizados son:

- Media muestral: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Varianza muestral: $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$
- Cuasivarianza muestral: $S_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n}{n-1}S^2$
- Momento muestral respecto al origen de orden r: $A_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$
- Desviación típica muestral: $S = \sqrt{S^2}$
- Cuasidesviación típica muestral: $S_c = \sqrt{S_c^2}$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = A_2 - A_1^2$$

Generalmente, la distribución en el muestreo de los estadísticos muestrales depende de la **distribución concreta de X**. No obstante, en algunos casos concretos es posible obtener expresiones para su **esperanza** y su **varianza**.

PROPIEDADES DE LA MEDIA Y LA VARIANZA MUESTRAL

Sea $X_1, \ldots, X_n$ una m.a.s. de X, tal que $E[X] = \mu$ y $Var[X] = \sigma^2$. Entonces,

- $E[\bar{X}] = \mu$ (la esperanza de la media muestral es la media poblacional).
- $Var[\bar{X}] = \frac{\sigma^2}{n}$ (la varianza de la media muestral es la varianza poblacional entre el tamaño muestral por lo que disminuye cuando este último aumenta).
- $E[S^2] = \frac{n-1}{n}\sigma^2$.
- $E[S_c^2] = \sigma^2$ (la esperanza de la cuasivarianza muestral es la varianza poblacional).

EJERCICIO 2

Sea $X_1, \ldots, X_n$ una m.a.s. de X, con $\alpha_r = E[X^r]$. Demuestra que, en ese caso:

$$E[A_r] = \alpha_r$$

$$Var[A_r] = \frac{\alpha_{2r} - \alpha_r^2}{n}$$

Como consecuencia de los resultados anteriores, tenemos que

- Si $X_1, \ldots, X_n$ es m.a.s. de $B(1,p)$,
  - $E[\bar{X}] = p \qquad Var[\bar{X}] = \frac{p(1-p)}{n}$
- Si $X_1, \ldots, X_n$ es m.a.s. de $B(m,p)$,
  - $E[\bar{X}] = mp \qquad Var[\bar{X}] = \frac{mp(1-p)}{n}$
- Si $X_1, \ldots, X_n$ es m.a.s. de $P(\lambda)$,
  - $E[\bar{X}] = \lambda \qquad Var[\bar{X}] = \frac{\lambda}{n}$
- Si $X_1, \ldots, X_n$ es m.a.s. de $U(a,b)$,
  - $E[\bar{X}] = \frac{a+b}{2} \qquad Var[\bar{X}] = \frac{(b-a)^2}{12n}$
- Si $X_1, \ldots, X_n$ es m.a.s. de $\Gamma(\alpha,\lambda)$,
  - $E[\bar{X}] = \frac{\alpha}{\lambda} \qquad Var[\bar{X}] = \frac{\alpha}{n\lambda^2}$

De una población con distribución de Poisson de parámetro λ, se obtiene una m.a.s. $(X_1, \ldots, X_n)$. Determina la distribución en el muestreo de la media muestral y comprueba que su esperanza coincide con la esperanza poblacional.

EJERCICIO 3

Calcular la distribución en el muestreo del estadístico $\sum_{i=1}^{n} X_i$ obtenido a partir de una m.a.s. de tamaño n de una población:

a) $B(m, p)$ b) $Exp(\lambda)$ c) $P(\lambda)$

d) $N(\mu, \sigma)$ e) $\Gamma(\alpha, \lambda)$

Como acabamos de ver, existen ciertos estadísticos cuya distribución en el muestreo es relativamente sencilla de obtener, si la **distribución poblacional** cumple ciertas condiciones (como la reproductividad).

No obstante, este no es siempre así por lo que la distribución en el muestreo puede ser **compleja de obtener**. En esos casos, resulta de utilidad el resultado siguiente.

DISTRIBUCIÓN ASINTÓTICA DE LOS ESTADÍSTICOS MUESTRALES

Sea $X_1, \ldots, X_n$ una m.a.s. de X, tal que $E[X] = \mu$ y $Var[X] = \sigma^2$. Entonces, por la **ley débil de los grandes números**,

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu$$

Así mismo, como la continuidad respeta la convergencia en probabilidad,

$$S^2 \xrightarrow{p} \sigma^2 \qquad S_c^2 \xrightarrow{p} \sigma^2$$

Por otro lado, por el **Teorema Central del Límite**, El resultado anterior implica que, en la práctica, si el tamaño muestral n es **suficientemente grande** (superior a 30), independientemente de la distribución de X, se tiene que:

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

DEFINICIÓN: ESTADÍSTICO DE ORDEN

Sea $X_1, \dots, X_n$ una m.a.s. de X. Llamaremos **estadístico de orden** al estadístico que ordena la muestra de **menor a mayor** valor y lo denotaremos como:

$$(X_{(1)}, X_{(2)}, \dots, X_{(n)})$$

Nótese que, aunque las $X_1, \dots, X_n$ son v.a.i.i.d., las variables $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ no son independientes, ni identicamente distribuidas. De hecho, su función de densidad/masa **conjunta** se obtiene como:

$$f_\theta(x_{(1)}, \dots, x_{(n)}) = n! \prod_{i=1}^{n} f_{X,\theta}(x_{(i)}), \quad x_{(1)} < x_{(2)} < \cdots < x_{(n)}$$

$$X_{(1)} = \min\{X_1, \dots, X_n\} \quad \rightarrow \quad F_{X_{(1)}}(y) = 1 - (1 - F_{X,\theta}(y))^n$$

$$X_{(n)} = \max\{X_1, \dots, X_n\} \quad \rightarrow \quad F_{X_{(n)}}(y) = (F_{X,\theta}(y))^n$$

EJERCICIO 4

Sea $X_1, \ldots, X_n$ una m.a.s. de $U(0, \theta)$. Calcula la esperanza y la varianza del máximo muestral.