Part VII

Estimación puntual

Introducción a la inferencia estadística

Introducción

- La inferencia estadística puede definirse como el conjunto de métodos mediante los cuales podemos extraer información sobre distintas características de interés de cierta distribución de probabilidad de la cual se ha observado una serie de datos.
- La inferencia estadística comprende las fases de recogida y depuración de los datos, estimación, contrastes de simplificación, diagnosis y validación del modelo.
- Según el objeto de estudio, la inferencia estadística se clasifica en inferencia paramétrica y no paramétrica.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos ...

Introducción

- La inferencia paramétrica se ocupa de aquellos casos en los que la distribución de probabilidad de la población objeto de estudio se supone conocida salvo los valores que toman ciertos coeficientes, llamados parámetros. El objetivo es estimar, dar intervalos de confianza o contrastar hipótesis sobre dichos parámetros
- La inferencia no paramétrica trata problemas similares cuando se tiene una distribución poblacional totalmente desconocida, sobre la cual tan sólo se realizan suposiciones muy generales (como, por ejemplo, que es una distribución continua, que tiene una única moda, etc.).

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Introducción

- El enfoque clásico trata los parámetros poblacionales desconocidos como valores fijos o constantes.
- El enfoque bayesiano considera que los parámetros desconocidos del modelo son variables aleatorias, para las cuales debe fijarse una distribución inicial, llamada distribución a priori. Utilizando la información muestral junto con la mencionada distribución a priori, los métodos bayesianos hacen uso de la regla de Bayes para ofrecer una distribución a posteriori sobre los parámetros.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Conceptos generales

- Se denomina población a un conjunto homogéneo de individuos sobre los que se estudian una o varias características que son, de alguna forma, observables.
- Una muestra es un subconjunto de la población. El tamaño muestral es el número de elementos de la muestra.
- Un método de muestreo no es más que el procedimiento empleado para la obtención de la muestra.
- Un parámetro es cualquier característica medible (normalmente numérica) de la población.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Muestreo en poblaciones finitas

- Muestreo aleatorio simple: Cada muestra posible tiene la misma probabilidad de ser elegida.
- Muestreo sistemático: Se usa muy frecuentemente cuando los individuos de la población están ordenados en listas.
- Muestreo estratificado: Si somos capaces de dividir la población en estratos, se actua dentro de cada estrato según un muestreo aleatorio simple.
- Muestreo por conglomerados: Si la población puede dividirse en conglomerados, que son homogéneos entre sí, se eligen aleatoriamente algunos conglomerados y, en cada uno de ellos, una muestra representativa.
- Muestreo polietápico: Utilizar distintos tipos de muestreo en sucesivas etapas.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Introducción

- Matemáticamente, podemos pensar que estamos interesados en el estudio de una variable aleatoria X, cuya distribución, F, es en mayor o menor grado desconocida.
- Supondremos que la familia $\mathcal{F}$ de distribuciones a la que pertence F es de la forma: $\mathcal{F} = \{F_\theta \mid \theta \in \Theta\}$, siendo $\Theta \subset \mathbb{R}^k$.
- Para tratar de disminuir el desconocimiento de la distribución teórica F de la variable aleatoria X en estudio, tomamos una muestra.

Estimación puntual

Muestra aleatoria simple

- Una muestra aleatoria simple, de tamaño n, de una variable aleatoria X, con distribución teórica F, son n variables aleatorias, $X_1, X_2, \ldots, X_n$, independientes e igualmente distribuidas, con distribución común F.
- Una realización de la muestra son los valores particulares, $x_1, x_2, \ldots, x_n$, observados para las variables $X_1, X_2, \ldots, X_n$.
- Como consecuencia, la función de distribución conjunta de $X_1, X_2, \ldots, X_n$ es:

$$F\left(x_1, x_2, \ldots, x_n\right) = F(x_1) \cdot F(x_2) \cdots F(x_n)$$

Estimación puntual

Otros conceptos

- Llamaremos espacio muestral al conjunto de muestras posibles, de tamaño determinado, que pueden obtenerse al seleccionar una muestra aleatoria de una población.
- Llamaremos estadístico a cualquier función T de la muestra.
- Un estadístico $T(X_1, X_2, \ldots, X_n)$, como función de variables aleatorias, es también una variable aleatoria. Tiene, por tanto, una distribución de probabilidad, llamada distribución en el muestreo de T.
- Los estadísticos independientes del parámetro (que son función únicamente de la muestra) se denominan estimadores ($\widehat{\theta} = T(X_1, X_2, \ldots, X_n)$).

Estimación puntual

Ejemplos de estadísticos

- $T(X_1, X_2, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i$.
- $T(X_1, X_2, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$.
- $T(X_1, X_2, \ldots, X_n) = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$.
- $T(X_1, X_2, \ldots, X_n) = \min(X_1, X_2, \ldots, X_n)$.
- $T(X_1, X_2, \ldots, X_n) = \max(X_1, X_2, \ldots, X_n)$.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Insesgadez

- Se denomina sesgo del estimador $\widehat{\theta}$ como estimador de θ a

$$Sesgo\left(\widehat{\theta}\right) = E\left(\widehat{\theta}\right) - \theta$$

- Si $E\left(\widehat{\theta}\right) = \theta$, para cada $\theta \in \Theta$, el estimador $\widehat{\theta}$ se dice centrado o insesgado para θ.
- Pueden existir muchos estimadores centrados para un parámetro. Por ejemplo, para estimar la media μ de una distribución cualquiera, todos los estimadores del tipo $\widehat{\mu} = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ con $\sum_{i=1}^n a_i = 1$ son centrados.

Propiedades deseables de los estimadores

Insesgadez asintótica

Un estimador, $\widehat{\theta}_n$, se dice asintóticamente insesgado para θ si:

$$\lim_{n \to \infty} E\left(\widehat{\theta}_n\right) = \theta, \text{ para cada } \theta \in \Theta$$

Propiedades deseables de los estimadores

Error cuadrático medio

- Se define el error cuadrático medio de un estimador $\widehat{\theta}$ de un parámetro θ como el número

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

- De la definición anterior, se deduce fácilmente que

$$ECM\left(\widehat{\theta}\right) = \left[Sesgo\left(\widehat{\theta}\right)\right]^2 + Var\left(\widehat{\theta}\right)$$

Propiedades deseables de los estimadores

Eficiencia

- Llamaremos precisión o eficiencia de $\widehat{\theta}$, como estimador de θ, al inverso de su error cuadrático medio:

$$efic\left(\widehat{\theta}\right) = \frac{1}{ECM\left(\widehat{\theta}\right)}$$

- Diremos que $\widehat{\theta}_2$ es más eficaz o más preciso que $\widehat{\theta}_1$ si para cualquier tamaño muestral:

$$efic\left(\widehat{\theta}_2\right) \geq efic\left(\widehat{\theta}_1\right)$$

Es decir, si $ECM\left(\widehat{\theta}_2\right) \leq ECM\left(\widehat{\theta}_1\right)$.

Propiedades deseables de los estimadores

Eficiencia relativa

- Llamaremos eficiencia relativa de $\widehat{\theta}_2$ respecto a $\widehat{\theta}_1$ al cociente entre sus eficiencias:

$$ER\left(\widehat{\theta}_2/\widehat{\theta}_1\right) = \frac{efic\left(\widehat{\theta}_2\right)}{efic\left(\widehat{\theta}_1\right)} = \frac{ECM\left(\widehat{\theta}_1\right)}{ECM\left(\widehat{\theta}_2\right)}$$

- Si $\widehat{\theta}$ es centrado, su error cuadrático medio coincide con su varianza, por lo que entre estimadores insesgados, en términos de eficiencia, es preferible el que tenga menor varianza.

Propiedades deseables de los estimadores

Consistencia

- Un estimador $\widehat{\theta}_n$ se dice consistente en media cuadrática para estimar un parámetro θ si

$$\lim_{n\to\infty} ECM\left(\widehat{\theta}_n\right) = 0$$

- En adelante, cuando hablemos de consistencia de un estimador nos referiremos a la consistencia en media cuadrática.
- Una condición necesaria y suficiente para que $\widehat{\theta}_n$ sea consistente es que sea asintóticamente insesgado y que $\lim_{n\to\infty} Var\left(\widehat{\theta}_n\right) = 0$.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Media muestral

- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria X con $E(X) = \mu$ y $Var(X) = \sigma^2$. Un estimador razonable del parámetro μ es la media muestral.

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

- Es un estimador insesgado de μ.

$$E\left(\overline{X}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

Estimación de la media de una población

Media muestral

- Es un estimador consistente de μ. Sólo hay que probar que su varianza tiende a cero.

$$Var\left(\overline{X}\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var\left(X_i\right) = \frac{\sigma^2}{n}$$

- La distribución exacta de $\overline{X}$ depende de la distribución de la población. Si $X \in N(\mu, \sigma)$, entonces $\overline{X} \in N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$. Por el TCL, para n grande $(n \geq 30)$, la distribución de $\overline{X}$ puede aproximarse por $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

Estimación de la varianza de una población

Varianza muestral

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria X con $E(X) = \mu$ y $Var(X) = \sigma^2$. Se define la varianza muestral como:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Varianza muestral

Es un estimador asintóticamente insesgado de σ^2. Teniendo en cuenta que

$$\sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 + n\left(\mu - \overline{X}\right)^2$$

obtenemos que

$$E\left(S^2\right) = \sigma^2 - \frac{\sigma^2}{n} = \sigma^2 \frac{n-1}{n}$$

y, por tanto,

$$\lim_{n\to\infty} E\left(S^2\right) = \sigma^2$$

Estimación de la varianza de una población

Varianza muestral

- La varianza de S^2 puede calcularse también, pero es bastante complicada (depende de los coeficientes de asimetría y de curtosis de la población).
- Es un estimador consistente de σ^2.
- La distribución de S^2 es, en general, asimétrica y su forma depende del tamaño muestral y de la población base. Usando el TCL, puede aproximarse la distribución de S^2 por la distribución normal, pero la aproximación es muy lenta y sólo se manifiesta en tamaños muestrales muy grandes.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos

Varianza muestral

- Si $X \in N(\mu, \sigma)$, se puede probar que $\frac{nS^2}{\sigma^2} \in \chi^2_{n-1}$.
- Dado que $E\left(\chi^2_n\right) = n$, se tiene que $E\left(\frac{nS^2}{\sigma^2}\right) = n - 1$ y, por tanto,

$$E\left(S^2\right) = \frac{n-1}{n}\sigma^2$$

- Como $Var\left(\chi^2_n\right) = 2n$, se tiene que $Var\left(\frac{nS^2}{\sigma^2}\right) = 2(n-1)$ y, por tanto,

$$Var\left(S^2\right) = \frac{2(n-1)}{n^2}\sigma^4$$

Estimación de la varianza de una población

Varianza muestral corregida o cuasivarianza

- Con el objeto de corregir el sesgo de la varianza muestral, se define la varianza muestral corregida o cuasivarianza como

$$\widehat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{n}{n-1}S^2$$

- Es un estimador insesgado de σ^2.
- Es un estimador consistente de σ^2.
- Si $X \in N(\mu, \sigma)$, las variables aleatorias $\overline{X}$ y $\widehat{S}^2$ son independientes y $\dfrac{(n-1)\widehat{S}^2}{\sigma^2} \in \chi^2_{n-1}$. Entonces $E\left(\widehat{S}^2\right) = \sigma^2$ y $Var\left(\widehat{S}^2\right) = \dfrac{2}{n-1}\sigma^4$.

Estimación de la varianza de una población

Comentarios

- Si se conoce la media poblacional, μ, utilizaremos el estimador que resulta de reemplazar $\overline{X}$ por μ en la expresión de S^2. El estimador resultante es insesgado y tiene menor varianza que S^2 y $\widehat{S}^2$.
- En los programas informáticos de estadística suele utilizarse la cuasivarianza muestral, debido a su carácter insesgado.

Estimación de una proporción

Proporción muestral

- Se desea estimar la proporción, p, de individuos de una población que poseen determinada característica. Tomamos una m.a.s. de n elementos anotando un 1 si el individuo elegido posee la característica y un 0 si carece de ella. Un estimador razonable de p es:

$$\widehat{p} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

- $\widehat{p}$ es la media muestral de n variables i.i.d. $B(p)$.
- $E(\widehat{p}) = p,\ Var(\widehat{p}) = \dfrac{p(1-p)}{n}$ y su distribución en el muestreo, para $(n > 30)$ es aprox. $N\left(p, \sqrt{(p(1-p))/n}\right)$.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos ...

Método de los momentos

- Se tiene una m.a.s $X_1, X_2, \ldots, X_n$ de una de X, que se caracteriza por k parámetros $\theta_1, \theta_2, \ldots, \theta_k$ que deseamos estimar.
- El método de los momentos consiste en igualar los primeros momentos poblacionales, que no sean constantes, a los correspondientes momentos muestrales y resolver las ecuaciones

$$A_j = g_j(\theta_1, \theta_2, \ldots, \theta_k) \ , \ j = 1, 2, \ldots, k,$$

donde $E(X^j) = g_j(\theta_1, \theta_2, \ldots, \theta_k)$ y $A_j = \frac{1}{n}\sum_{i=1}^{n} X_i^j$, $j = 1, 2, \ldots, k$.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos ...

Método de máxima verosimilitud

- La idea de la estimación máximo-verosímil de los parámetros $\theta_1, \theta_2, \ldots, \theta_k$ que caracterizan una variable aleatoria, X, es elegir los valores de los parámetros que hacen que la muestra observada, $x_1, x_2, \ldots, x_n$, sea la más verosímil.
- Si en una bolsa hay 6 caramelos de fresa y menta, no todos del mismo sabor, pero se desconoce cuántos hay de cada uno, la proporción de caramelos de fresa puede ser $p = \frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}$ o $\frac{5}{6}$. Si se extraen dos caramelos, con reemplazamiento, y se obtiene uno de cada sabor, el valor de p que otorga mayor probabilidad al suceso observado es $p = 1/2$. En este caso la estimación de máxima verosimilitud para p es $\widehat{p} = 1/2$.

Estadística I
Mario Francisco
Introducción a la inferencia estadística
Estimación puntual
Propiedades deseables de los estimadores
Estimación de la media de una población
Estimación de la varianza de una población
Estimación de una proporción
Procedimientos ...

Método de máxima verosimilitud

- Sea $X_1, X_2, \ldots, X_n$ una m.a.s. de una población cuya distribución depende de un parámetro, θ.
- En el caso discreto, se denomina función de verosimilitud a

$$l(\theta) = P_\theta\left(X_1 = x_1, \ldots X_n = x_n\right) = P_\theta\left(X = x_1\right) \cdots P_\theta\left(X = x_n\right)$$

- Si la población es continua la verosimilitud se define por

$$l(\theta) = f_\theta(x_1, x_2, \ldots, x_n) = f_\theta\left(x_1\right) \cdot f_\theta\left(x_2\right) \cdots f_\theta\left(x_n\right)$$

donde f_θ es la función de densidad de X.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos ...

Método de máxima verosimilitud

- El estimador de máxima verosimilitud de θ es aquel que hace máxima la función de verosimilitud.
- Si l es diferenciable, una condición necesaria es $\frac{\partial l(\theta)}{\partial \theta} = 0$
- Si la distribución de X depende de varios parámetros, $\theta_1, \theta_2, \ldots, \theta_k$, y l es derivable,

$$\frac{\partial l(\theta_1, \theta_2, \ldots, \theta_k)}{\partial \theta_j} = 0, \quad j = 1, 2, \ldots, k$$

- En muchos casos será más conveniente trabajar con $L = \ln l$, llamada función soporte.

Estadística I

Mario Francisco

Introducción a la inferencia estadística

Estimación puntual

Propiedades deseables de los estimadores

Estimación de la media de una población

Estimación de la varianza de una población

Estimación de una proporción

Procedimientos para la construcción de estimadores

Método de máxima verosimilitud. Propiedades

Para distribuciones cuyo rango de valores posibles es conocido a priori y no depende de ningún parámetro:

- Asintóticamente insesgados.
- Con distribución asintóticamente normal.
- Asintóticamente de varianza mínima.
- Si $\widehat{\theta}$ es el estimador de máxima verosimilitud de θ y g es una función biyectiva, $g\left(\widehat{\theta}\right)$ es el estimador de máxima verosimilitud de $g\left(\theta\right)$.