VI. Inferencia estadística

Inferencia Estadística

- La inferencia estadística es primordialmente de naturaleza inductiva y llega a generalizar respecto de las características de una población valiéndose de observaciones empíricas de la muestra.
- Al utilizar estadísticas muestrales para estudiar un parámetro de la población es muy normal que ambos sean diferentes y la igualdad entre ambos sea mera coincidencia. La diferencia entre la estadística muestral y el correspondiente parámetro de la población se suele llamar error de estimación. Solo conoceríamos dicho error si se conociera el parámetro poblacional que por lo general se desconoce. La única forma de tener alguna certeza al respecto es hacer todas las observaciones posibles del total de la población; en la mayoría de las aplicaciones prácticas es imposible o impracticable.

Inferencia Estadística

- Las inferencias estadísticas se hacen por posibilidades o probabilidades.
- Por ejemplo de la media de una muestra se hacen inferencias sobre la media de la población. Exactamente no sabemos cuál es la diferencia entre ambas. Lo que si sabemos es que es pequeña la probabilidad de que esta diferencia sea mayor que, por ejemplo 3 o 2 errores estándares.

VI.1. Estimación Puntual

La inferencia estadística más sencilla es la Estimación Puntual o por punto, en la que se calcula un valor único (estadístico) con las datos muestrales para estimar un parámetro poblacional

VI.2. Estimador

Definición

- Un estimador es en si mismo una variable aleatoria y por lo mismo tiene una distribución (muestral) teórica.
- Un estimador de un parámetro θ es una función de los valores muestrales aleatorios X1, X2,....., Xn que proporciona una estimación puntual de θ.

3.2. Estimador

Ejemplo:

Sean los valores siguientes 120, 117.5, 115 tomados de una población finita, obtener la estimación resultante:

La estimación:

$$\hat{X} = \frac{X_1 + X_2 + X_3}{3}$$

Se interpreta como el proceso de "tomar una muestra de tres valores y promediarlos".

De la muestra que se da en particular:

$$x_1 = 120$$

$$x_2 = 117.5$$

$$x_3 = 115$$

Y se obtiene como una estimación de la media poblacional que está basada en la muestra especificada.

VI.2.1. Características de los Estimadores

VI.2.1.1.- Estimador Insesgado.

Un estimador que es una función de los datos muestrales $X_1, X_2, \ldots, X_n$ se conoce como estimador insesgado del parámetro poblacional θ si su valor esperado es igual a θ. Dicho de otra manera, es un estimador insesgado del parámetro θ si $E( ) = \theta$.

La condición de que el estimador es insesgado supone que el valor **promedio** de es exactamente correcto. No dice que un valor particular sea exactamente correcto.

VI.2.1. Características de los Estimadores

VI.2.1.2- Estimador Eficiente

Se dice que un estimador es el más eficiente, para un problema en particular , cuando tiene el error estándar más pequeño de todos los estimadores insesgados posibles.

Se utiliza la palabra eficiente porque, el estimador hace el mejor uso posible de los datos muestrales.

VI.2.1.3.- Estimador Consistente

Un estimador es consistente si se aproxima al parámetro poblacional con probabilidad uno a medida que el tamaño de la muestra tiende a infinito.

VI.3. Método de Máxima Verosimilitud

Si $x_1, x_2, \ldots\ldots, y\ x_n$ son los valores de una muestra tomada al azar de una población con el parámetro θ, la función de verosimilitud de la muestra está dada por:

$$L(\theta) = f\,(x_1, x_2\ ,\ldots\ldots, x_n\ ;\theta)$$

Para valores de θ contenidos en un dominio dado.

El método de máxima verosimilitud consiste en maximizar la función de verosimilitud con respecto a θ y nos referimos al valor de θ que maximiza la función de probabilidad como la estimación de máxima verosimilitud de θ.

VI.3.1. De la distribución de probabilidad binomial determinar el estimador de máxima verosimilitud del parámetro θ

$$L(\theta)=\binom{n}{x}\theta^{x}(1-\theta)^{n-x}$$

Obteniendo el logaritmo natural.

$$\ln(L(\theta))=\ln\binom{n}{x}+x\ln\theta+(n-x)\ln(1-\theta)$$

Derivando con respecto a θ.

$$\frac{d[\ln L(\theta)]}{d\theta}=\frac{x}{\theta}-\frac{(n-x)}{(1-\theta)}$$

VI.3.1. De la distribución de probabilidad binomial determinar el estimador de máxima verosimilitud del parámetro θ

- Igualando a cero:

$$\frac{x}{\theta}-\frac{(n-x)}{(1-\theta)}=0$$

$$\frac{x}{\theta}=\frac{(n-x)}{(1-\theta)}$$

$$x(1-\theta)=(n-x)\theta$$

$$x-x\theta=\theta n-x\theta$$

$$x-x\theta+x\theta=\theta n$$

$$x=\theta n$$

$$\therefore$$

$$\overline{\theta}=\frac{x}{n}$$

Es el estimador de máxima verosimilitud del parámetro θ de la distribución binomial.

VI.3.2. De la distribución de probabilidad Poisson determinar el estimador de máxima verosimilitud del parámetro λ en muestras aleatorias simples de tamaño n

$$L(X:\lambda) = e^{-n\lambda} \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!}$$

Obteniendo el logaritmo natural:

$$\ln L(X:\lambda) = -n\lambda + \ln \lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln x_i!$$

Derivando parcialmente con respecto a λ:

$$\frac{\partial \ln L(X:\lambda)}{\partial \lambda} = -n + \frac{\sum_{i=1}^{n} x_i}{\lambda}$$

VI.3.2. De la distribución de probabilidad Poisson determinar el estimador de máxima verosimilitud del parámetro λ en muestras aleatorias simples de tamaño n

Igualando a cero y resolviendo:

$$-n + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0$$

$$n = \frac{\sum_{i=1}^{n} x_i}{\lambda}$$

$$n\lambda = \sum_{i=1}^{n} x_i$$

$$\bar{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

El estimador es la media.

VI.3.3. Dada la función de Probabilidad normal con media μ y varianza σ², determinar los estimadores de máxima verosimilitud

- La función de probabilidad está dada por:

$$L(\mu,\sigma^2)=\prod_{i=1}^{n} n(x_i;\mu,\sigma)$$

$$L(\mu,\sigma^2)=\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)}$$

- Obteniendo el logaritmo natural:

$$\ln L(\mu,\sigma^2)=n\ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right)+\left(-\frac{1}{2\sigma^2}\right)\sum_{i=1}^{n}(x_i-\mu)^2\ln e$$

$$\ln L(\mu,\sigma^2)=n\left[\ln(1)-\ln\left(\sigma\sqrt{2\pi}\right)\right]-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\ln L(\mu,\sigma^2)=n\left[-\left(\ln\sigma+\ln\sqrt{2\pi}\right)\right]-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\ln L(\mu,\sigma^2)=-n\ln\sigma-n\ln\sqrt{2\pi}-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

VI.3.3. Dada la función de Probabilidad normal con media μ y varianza σ², determinar los estimadores de máxima verosimilitud

Derivando parcialmente con respecto a μ:

$$\frac{\partial \ln\left[L(\mu,\sigma)\right]}{\partial \mu} = -\frac{1}{2\sigma^2}(2)\sum_{i=1}^{n}(x_i-\mu)(-1)$$

$$\frac{\partial \ln\left[L(\mu,\sigma)\right]}{\partial \mu} = \frac{1}{2\sigma^2}(2)\sum_{i=1}^{n}(x_i-\mu)$$

Igualando a cero:

$$\frac{1}{2\sigma^2}(2)\sum_{i=1}^{n}(x_i-\mu)=0$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0$$

$$\sum_{i=1}^{n}(x_i-\mu)=0$$

VI.3.3. Dada la función de Probabilidad normal con media μ y varianza σ^2, determinar los estimadores de máxima verosimilitud

$$\sum_{i=1}^{n}\left(x_i\right)-n\mu=0$$

$$\sum_{i=1}^{n}\left(x_i\right)=n\mu$$

$$\overline{\mu}=\frac{\sum_{i=1}^{n}\left(x_i\right)}{n}=\overline{x}$$

El valor obtenido es la media.

VI.3.3. Dada la función de Probabilidad normal con media μ y varianza σ², determinar los estimadores de máxima verosimilitud

Ahora se deriva parcialmente con respecto a σ:

$$\frac{\partial \ln L\left(\mu,\sigma^2\right)}{\partial \sigma} = -\frac{n}{\sigma} - \frac{1}{2\sigma^3}(-2)\sum_{i=1}^{n}\left(x_i-\mu\right)^2$$

$$\frac{\partial \ln L\left(\mu,\sigma^2\right)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}\left(x_i-\mu\right)^2$$

VI.3.3. Dada la función de Probabilidad normal con media μ y varianza σ^2, determinar los estimadores de máxima verosimilitud

Igualando a cero:

$$-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}\left(x_i-\mu\right)^2=0$$

$$\frac{1}{\sigma^3}\sum_{i=1}^{n}\left(x_i-\mu\right)^2=\frac{n}{\sigma}$$

$$\sum_{i=1}^{n}\left(x_i-\mu\right)^2=\frac{n}{\sigma}\left(\sigma^3\right)$$

$$\sum_{i=1}^{n}\left(x_i-\mu\right)^2=n\sigma^2$$

$$\sigma^2=\frac{\sum_{i=1}^{n}\left(x_i-\mu\right)^2}{n}$$

VI.3.3. Dada la función de Probabilidad normal con media μ y varianza σ², determinar los estimadores de máxima verosimilitud

Sustituyendo el valor de se obtiene:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n}$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n}}$$

Es una estimación de máxima verosimilitud de σ.

VI.3.4. De la Distribución de probabilidad exponencial negativa obtener el estimador de máxima verosimilitud

$$f(x) = \theta e^{-\theta x}$$

o

$$f(x) = \lambda e^{-\lambda x}$$

$0 < x < \infty$

Obtener el estimador de máxima verosimilitud.

Una demostración es:

Como,

$$E(X) = \frac{1}{\lambda}$$

$$E(X)\lambda = 1$$

$$\lambda = \frac{1}{E(X)}$$

VI.3.4. De la Distribución de probabilidad exponencial negativa obtener el estimador de máxima verosimilitud

Donde:

$$E(X) = \frac{\sum_{i=1}^{n} x_i}{n}$$

Entonces:

$$\lambda = \frac{1}{\frac{\sum_{i=1}^{n} x_i}{n}}$$

$$\therefore$$

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

La otra demostración es por medio de los estimadores de máxima verosimilitud.

VI.4. Estimación por Intervalos

- Generalmente las estimaciones de punto rara vez son iguales a los parámetros que estiman, razón por la cual es deseable dar alguna libertad mediante la utilización de estimación por intervalo. Una estimación por intervalo de un parámetro θ es un intervalo de la forma

$$\hat{\theta}_1 < \theta < \hat{\theta}_2$$

donde $\hat{\theta}_1$ y $\hat{\theta}_2$ dependen del valor que tome el estimador $\hat{\theta}$ en una muestra dada y también en la distribución muestral de $\hat{\theta}$.

VI.4. Estimación por Intervalos

- Como generalmente diferentes muestras producen valores diferentes de $\hat{\theta}$ y, por consecuencia diferentes valores de $\hat{\theta}_1$ y de $\hat{\theta}_2$, esto puntos extremos del intervalo son valores de variables aleatorias correspondientes a $\hat{\theta}_1$ y $\hat{\theta}_2$.
- Por lo tanto con base en la distribución muestral de $\hat{\theta}$ se puede afirmar, con una probabilidad, si este intervalo en realidad contiene el parámetro que se supone estima.

VI.4. Estimación por Intervalos

- En otras palabras, podemos utilizar la distribución muestral de $\hat{\theta}$ para elegir $\hat{\theta}_1$ y $\hat{\theta}_2$ tal que para cualquier probabilidad especificada $1 - \alpha$, donde $0 < \alpha < 1$, entonces:

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$$

- El intervalo $(\hat{\theta}_1 < \theta < \hat{\theta}_2)$, determinado en relación con una muestra en particular, recibe el nombre de **intervalo de confianza** del $(1 - \alpha)100\%$, la fracción $1 - \alpha$ se conoce como **coeficiente de confianza o grado de confianza** y los extremos $\hat{\theta}_1$ y $\hat{\theta}_2$ reciben el nombre de **límites de confianza inferior** y **superior**.

VI.4. Estimación por Intervalos

- Los intervalos de confianza de parámetros dados no son únicos, existen numerosos intervalos de confianza de μ donde tienen el mismo grado de confianza. Igual que en el caso de la estimación puntual, los métodos de obtención de intervalos de confianza deben de juzgarse por sus propiedades estadísticas.

VI.4.1. Intervalos de Confianza para Medias

Teorema (Intervalo de confianza para μ, cuando σ es conocida)

- Si es el valor de la media de una muestra aleatoria de tamaño n tomada de una población normal con la varianza σ^2, un intervalo de confianza del (1 − α)100% para μ está dado por:

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

- Este resultado se puede utilizar, con base en el teorema central de límite, para muestras aleatorias tomadas de poblaciones no normales con la varianza conocida σ^2, siempre que n sea lo suficientemente grande; un valor que se considera suficientemente grande es con n ≥ 30.

VI.4.1.2. La distribución t

Teorema

- Si y s^2 son la media y la varianza de una muestra aleatoria de tamaño n tomada de una población normal con la media μ y la varianza σ^2, entonces:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

- Tiene la distribución t con n – 1 grados de libertad.

VI.4.1.2. La distribución t

Por lo tanto:

$$P(-t_{\frac{\alpha}{2},n-1} < t < t_{\frac{\alpha}{2},n-1}) = 1-\alpha$$

Al sustituir a t, la desigualdad queda como:

$$P(-t_{\frac{\alpha}{2},n-1} < \frac{\overline{x}-\mu}{\frac{s}{\sqrt{n}}} < t_{\frac{\alpha}{2},n-1}) = 1-\alpha$$

VI.4.1.2. La distribución t

- Lo que es equivalente a:

$$P(\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}) = 1 - \alpha$$

VI.4.1.2. La distribución t

Propiedades:

1. La distribución t, igual que la distribución normal, es simétrica con respecto a la media μ = 0.
2. La distribución t tiene una mayor variabilidad que la distribución z.
3. Hay muchas distribuciones t que son diferentes. Se determina una en particular al especificar sus grados de libertad, g.l. Si se toma una muestra aleatoria de una población normal, el estadístico:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Tiene una distribución t con g.l. = n – 1.

4. A medida que n se incrementa (o, lo que es lo mismo, g.l se incrementa), la distribución t se aproxima a la de z.

Lo anterior conduce al siguiente teorema para muestras pequeñas de **μ** (aunque es válida para muestras de cualquier tamaño).

VI.4.1.3. Teorema (Intervalo de confianza para μ, cuando σ es desconocida)

- Si y s son los valores de la media y la desviación estándar de una muestra aleatoria de tamaño n tomada de una población normal con la varianza desconocida , un intervalo de confianza del (1 –)100% para μ está dado por:

$$\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}$$

VI.4.2. Intervalo de Confianza para Proporciones

- Existen situaciones en las cuales se debe de obtener la proporción, probabilidad, porcentaje o índice (tasa), como la proporción de unidades defectuosas en un cargamento grande de televisores, la probabilidad de que un auto tenga los frenos en mal estado, la tasa de mortalidad que provoca una enfermedad, etc.
- En muchos de estos casos es razonable suponer que se muestrea una población binomial y, que el problema se reduce a calcular el parámetro binomial θ. Utilizando el hecho de que para n grande la distribución binomial se obtiene por una aproximación de la distribución normal, es decir, que la variable aleatoria:

$$z = \frac{x - np}{\sqrt{npq}}$$

- Se puede considerar como si tuviera la distribución normal estándar.

VI.4.2.1. Teorema (Intervalo de Confianza de muestra grande para p)

- Un intervalo de confianza del (1 –)100% para el parámetro binomial p está dado por:

$$\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Donde:

$$\hat{p} = \frac{x}{n}$$

VI.4.2.2. Intervalo de Confianza para Varianzas

- Dada una muestra aleatoria n tomada de una población normal, se puede obtener un intervalo de confianza del (1 – α)100% para σ^2 utilizando el siguiente teorema.

VI.4.2.2. Intervalo de Confianza para Varianzas

Teorema

- Si y s^2 son la media y la varianza de muestra aleatoria de tamaño n tomada de una población normal con la media µ y la varianza , entonces:
  1. $\bar{x}$ y s^2 son independientes.
  2. La variable aleatoria $\frac{(n-1)s^2}{\sigma^2}$ tiene una distribución ji-cuadrada con n – 1 grados de libertad.
- Según el cual:

$$\frac{(n-1)s^2}{\sigma^2}$$

- Es una variable aleatoria que tiene una distribución ji-cuadrada con n – 1 grados de libertad.

VI.4.2.2. Intervalo de Confianza para Varianzas

- Por lo tanto:

$$P(\chi^2_{1-\frac{\alpha}{2},n-1} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2},n-1}) = 1-\alpha$$

- Despejando para obtener el valor de:

$$P(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

- De lo anterior se deduce el siguiente teorema

VI.4.2.3. Teorema (Intervalo de Confianza para σ^2)

- Si s^2 es el valor de la varianza de una muestra aleatoria de tamaño n tomada de una población normal, un intervalo de confianza del $(1 - \alpha)100\%$ para σ^2 está dado por:

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

- Se pueden obtener límites de confianza del $(1 - \alpha)100\%$ correspondientes para σ sacando las raíces cuadradas de los límites de confianza para σ^2.

VI.4.2.3. Teorema (Intervalo de Confianza para σ^2)

- Lo que equivale a lo siguiente:

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}}$$

- Donde $\chi_a{}^2$ separa un área a en el extremo derecho de la distribución χ^2 con *n – l* grados de libertad.