Estimación por intervalos

Estadística II

Universidad de Salamanca

Curso 2011/2012

Outline

Introducción

En el tema anterior hemos visto métodos de estimación que asignan al valor desconocido del parámetro un valor concreto

En este tema, nuestro objetivo es obtener un intervalo en el que se encuentren los valores que podrán ser considerados "razonables" para ese parámetro. Lo que denominamos **intervalos de confianza**

Estimación por intervalos

Definición

Sea X una v. a. cuya función de densidad es $f_X(x, \theta)$, que depende de una parámetro θ (desconocido)

Los estadísticos $T_I = T_I(X_1, \ldots, X_n)$ y $T_S = T_S(X_1, \ldots, X_n)$ (no dependen de θ) forman un **intervalo de confianza** para θ a nivel de confianza $1 - \alpha \Rightarrow$

$$P_\theta[T_I \leq \theta \leq T_S] = 1 - \alpha$$

Nota

Para $(x_1, \ldots, x_n)$ una muestra concreta, T_I y T_S son valores concretos

Estimación por intervalos

Propiedades

- T_I y T_S son estimadores por defecto y por exceso de θ
- El nivel de confianza, la fiabilidad de la estimación, $1-\alpha$ se da en %
- Longitud del intervalo $L = T_S - T_I$
- Precisión $\frac{L}{2}$. A mayor L menor precisión. Para hayar el I.C. fijado el nivel de confianza, busco el intervalo de longitud mínima

Método de construcción de intervalos de confianza: Método del pivote

El método del pivote

Sea $X \rightsquigarrow f_X(x, \theta)$ una v. a. y una m.a.s de tamaño n.
Un **pivote** es un estadístico $T = T(X_1, \ldots, X_n, \theta)$ que verifica:

- Depende de la muestra y del parámetro
- Su distribución no depende de θ y es conocida
- Para una muestra concreta, T es estrictamente monótono en θ
- La solución de θ en la ecuación $T(x_1, \ldots, x_n, \theta) = t$ es única

$$P[a \leq Pivote \leq b] = 1 - \alpha$$

Outline

Para la media poblacional μ, con σ conocida

Pivote

$$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow N(0,1)$$

El intervalo de mínima longitud para una $N(0,1)$ es simétrico respecto a 0 dejando colas de $\frac{\alpha}{2}$ cada una

$$P\left[-z_{\frac{\alpha}{2}} \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

Para la media poblacional μ, con σ conocida

Intervalo de confianza

$$\left[\overline{X} - z_{\frac{\alpha}{2}} \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}, \overline{X} + z_{\frac{\alpha}{2}} \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right]$$

Cálculo de $z_{\frac{\alpha}{2}}$

$$P\left[Z \leq z_{\frac{\alpha}{2}}\right] = 1 - \frac{\alpha}{2}$$

Buscamos en las tablas de la $N(0, 1)$

Para la media poblacional μ, con σ desconocida

Pivote

$$\frac{\overline{X}-\mu}{\frac{S_c}{\sqrt{n}}} \rightsquigarrow t_{n-1}$$

El intervalo de mínima longitud para una t es simétrico respecto a 0 dejando colas de $\frac{\alpha}{2}$ cada una

$$P\left[-t_{\frac{\alpha}{2}} \leq \frac{\overline{X}-\mu}{\frac{S_c}{\sqrt{n}}} \leq t_{\frac{\alpha}{2}}\right] = 1-\alpha$$

Para la media poblacional μ, con σ desconocida

Intervalo de confianza

$$\left[\overline{X} - t_{\frac{\alpha}{2}} \frac{S_c}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2}} \frac{S_c}{\sqrt{n}}\right]$$

Cálculo de $t_{\frac{\alpha}{2}}$

$$P\left[t_{n-1} \leq t_{\frac{\alpha}{2}}\right] = 1 - \frac{\alpha}{2}$$

Buscamos en las tablas de la t de Student

Para la media poblacional μ, con σ desconocida

Nota

Sabemos que:

$$n.S_X^2 = (n-1)S_c^2 \Rightarrow \frac{S_X}{\sqrt{n-1}} = \frac{S_c}{\sqrt{n}}$$

Entonces, podemos considerar como Pivote

$$\frac{\overline{X}-\mu}{\frac{S_X}{\sqrt{n-1}}} \rightsquigarrow t_{n-1}$$

Outline

Para la varianza poblacional con μ desconocido

Pivote

$$\frac{n.S_X^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$$

El intervalo de mínima longitud para una χ^2 es el que deja colas de $\frac{\alpha}{2}$ cada una

$$P\left[\frac{n.S_X^2}{b} \leq \sigma^2 \leq \frac{n.S_X^2}{a}\right] = 1 - \alpha$$

Para la varianza poblacional con μ desconocido

Intervalo de confianza

$$\left[\frac{n.S_X^2}{b}, \frac{n.S_X^2}{a}\right]$$

donde:

$$a = \chi^2_{n-1}\left(\frac{\alpha}{2}\right)$$

$$b = \chi^2_{n-1}\left(1 - \frac{\alpha}{2}\right)$$

Outline

Intervalos de confianza para la diferencia de medias poblacionales con σ_1 y σ_2 conocidas

Sean $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$

- $x_{11}, \ldots, x_{1n_1}$ una m.a.s de X_1
- $x_{21}, \ldots, x_{2n_2}$ una m.a.s de X_2

Intervalos de confianza para la diferencia de medias poblacionales con σ_1 y σ_2 conocidas

Pivote

$$\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightsquigarrow N(0,1)$$

Intervalo de confianza

$$\left[(\overline{X_1} - \overline{X_2}) - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\overline{X_1} - \overline{X_2}) + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right]$$

Intervalos de confianza para la diferencia de medias poblacionales con $\sigma_1 = \sigma_2 = \sigma$ desconocidas

Sean $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$

- $x_{11}, \ldots, x_{1n_1}$ una m.a.s de X_1
- $x_{21}, \ldots, x_{2n_2}$ una m.a.s de X_2

Intervalos de confianza para la diferencia de medias poblacionales con $\sigma_1 = \sigma_2 = \sigma$ desconocidas

Pivote

$$\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{S.\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow t_{n_1+n_2-2}$$

$$S = \sqrt{\frac{n_1.S^2_{X_1} + n_2.S^2_{X_2}}{n_1 + n_2 - 2}}$$

Intervalos de confianza para la diferencia de medias poblacionales con $\sigma_1 = \sigma_2 = \sigma$ desconocidas

Intervalo de confianza

$$\left[(\overline{X_1} - \overline{X_2}) - t_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\overline{X_1} - \overline{X_2}) + t_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

Intervalos de confianza para la diferencia de medias poblacionales con σ_1 y σ_2 desconocidas

Sean $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$

- $x_{11}, \ldots, x_{1n_1}$ una m.a.s de X_1
- $x_{21}, \ldots, x_{2n_2}$ una m.a.s de X_2

Intervalos de confianza para la diferencia de medias poblacionales con σ_1 y σ_2 desconocidas

Pivote

$$\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_{c_1}^2}{n_1} + \frac{S_{c_2}^2}{n_2}}} \rightsquigarrow t_\phi$$

$$\phi = \frac{\left(\frac{S_{c_1}^2}{n_1} + \frac{S_{c_2}^2}{n_2}\right)^2}{\frac{\left(\frac{S_{c_1}^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_{c_2}^2}{n_2}\right)^2}{n_2 - 1}}$$

Intervalos de confianza para la diferencia de medias poblacionales con σ_1 y σ_2 desconocidas

Intervalo de confianza

$$\left[(\overline{X_1}-\overline{X_2})-t_{\frac{\alpha}{2}}.\sqrt{\frac{S_{c_1}^2}{n_1}+\frac{S_{c_2}^2}{n_2}},(\overline{X_1}-\overline{X_2})+t_{\frac{\alpha}{2}}\sqrt{\frac{S_{c_1}^2}{n_1}+\frac{S_{c_2}^2}{n_2}}\right]$$

Outline

Intervalos de confianza para el cociente de varianzas $\frac{\sigma_2^2}{\sigma_1^2}$

Sean $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$

- $x_{11}, \ldots, x_{1n_1}$ una m.a.s de X_1
- $x_{21}, \ldots, x_{2n_2}$ una m.a.s de X_2

Intervalos de confianza para el cociente de varianzas $\frac{\sigma_2^2}{\sigma_1^2}$

Pivote

$$\frac{S_{c_1}^2}{S_{c_2}^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \rightsquigarrow F_{n_1-1, n_2-1}$$

Intervalos de confianza para el cociente de varianzas $\frac{\sigma_2^2}{\sigma_1^2}$

Intervalo de confianza

$$\left[a.\frac{S_{c_2}^2}{S_{c_1}^2}, b.\frac{S_{c_2}^2}{S_{c_1}^2}\right]$$

$$a = \frac{1}{F_{n_2-1,n_1-1}(1-\frac{\alpha}{2})}$$

$$b = F_{n_1-1,n_2-1}(1-\frac{\alpha}{2})$$

Outline

Para una muestra con $n > 20$

Sea $X \rightsquigarrow b(p)$

- $X_1, \ldots, X_n$ una m.a.s de X
- El estimador máximo verosímil de p es $\overline{X}$
- $\sum_{i=1}^{n} X_i \rightsquigarrow B(n, p)$
- $n.\overline{X} \rightsquigarrow B(n, p)$
- Cuando $n > 20$ tenemos

$$\frac{\overline{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \rightsquigarrow N(0, 1)$$

Para una muestra con $n > 20$

Pivote

$$\frac{\overline{X} - p}{\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}} \rightsquigarrow N(0,1)$$

Para una muestra con $n > 20$

Intervalo de confianza

$$\left[\overline{X} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}, \overline{X} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right]$$

Outline

Para dos muestras: $p_1 - p_2$

Sean $X_1 \rightsquigarrow b(p_1)$ y $X_2 \rightsquigarrow b(p_2)$

- $X_{11}, \ldots, X_{1n_1}$ una m.a.s de X_1
- $X_{21}, \ldots, X_{2n_2}$ una m.a.s de X_2
- El estimador máximo verosímil de p es $\overline{X}$
- $\sum_{i=1}^{n} X_i \rightsquigarrow B(n, p)$
- $n.\overline{X} \rightsquigarrow B(n, p)$
- Cuando $n > 20$ tenemos

$$\frac{\overline{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \rightsquigarrow N(0, 1)$$

Para dos muestras: $p_1 - p_2$

Pivote

$$\frac{(\overline{X_1} - \overline{X_2}) - (p_1 - p_2)}{\sqrt{\frac{\overline{X_1}(1-\overline{X_1})}{n_1} + \frac{\overline{X_2}(1-\overline{X_2})}{n_2}}} \rightsquigarrow N(0,1)$$

Para dos muestras: $p_1 - p_2$

Intervalo de confianza

$$\left[(\overline{X_1} - \overline{X_2}) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\overline{X_1}(1 - \overline{X_1})}{n_1} + \frac{\overline{X_2}(1 - \overline{X_2})}{n_2}} \right]$$