Probabilidades y Estadística

Clase Auxiliar

Profesor: Fernando Lema
Auxiliares: Víctor Carmi - Abelino Jiménez G.

Universidad de Chile

Indice General

1 Contenidos
- Intervalos de Confianza

Indice General

1 Contenidos
- Intervalos de Confianza

Motivación

Hasta ahora, se tienen métodos para construir estimadores de parámetros de distribuciones específicas.

Sin embargo, no podemos esperar que el valor del estimador nos dé de manera precisa el valor del parámetro.

Queremos entonces que

$$P(|\hat{\theta} - \theta| > u) = \alpha$$

con α pequeño. (fijado a priori)

Esto equivale a

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$$

Motivación

Hasta ahora, se tienen métodos para construir estimadores de parámetros de distribuciones específicas.

Sin embargo, no podemos esperar que el valor del estimador nos dé de manera precisa el valor del parámetro.

Queremos entonces que

$$P(\mid \hat{\theta} - \theta \mid > u) = \alpha$$

con α pequeño. (fijado a priori)

Esto equivale a

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$$

Motivación

Hasta ahora, se tienen métodos para construir estimadores de parámetros de distribuciones específicas.

Sin embargo, no podemos esperar que el valor del estimador nos dé de manera precisa el valor del parámetro.

Queremos entonces que

$$P(\mid \hat{\theta} - \theta \mid > u) = \alpha$$

con α pequeño. (fijado a priori)

Esto equivale a

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$$

Motivación

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$$

Definición NIVEL DE CONFIANZA

Al valor $1 - \alpha$ se le denomina Nivel de Confianza

Si α es pequeño, entonces podemos decir que es altamente probable encontrar al parámetro θ en el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$

Definición INTERVALO DE CONFIANZA

Se dice que el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$ es el Intervalo de Confianza para θde nivel de confianza $(1 - \alpha)$

OBSERVACIONES:

- Un intervalo de confianza es aleatorio.

Motivación

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$$

Definición NIVEL DE CONFIANZA

Al valor $1-\alpha$ se le denomina Nivel de Confianza

Si α es pequeño, entonces podemos decir que es altamente probable encontrar al parámetro θ en el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$

Definición INTERVALO DE CONFIANZA

Se dice que el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$ es el Intervalo de Confianza para θde nivel de confianza $(1-\alpha)$

OBSERVACIONES:

- Un intervalo de confianza es aleatorio.

Motivación

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$$

Definición NIVEL DE CONFIANZA

Al valor $1 - \alpha$ se le denomina Nivel de Confianza

Si α es pequeño, entonces podemos decir que es altamente probable encontrar al parámetro θ en el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$

Definición INTERVALO DE CONFIANZA

Se dice que el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$ es el Intervalo de Confianza para θde nivel de confianza $(1 - \alpha)$

OBSERVACIONES:

- Un intervalo de confianza es aleatorio.

Motivación

$$P(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$$

Definición NIVEL DE CONFIANZA

Al valor $1 - \alpha$ se le denomina Nivel de Confianza

Si α es pequeño, entonces podemos decir que es altamente probable encontrar al parámetro θ en el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$

Definición INTERVALO DE CONFIANZA

Se dice que el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$ es el Intervalo de Confianza para θde nivel de confianza $(1 - \alpha)$

OBSERVACIONES:

- Un intervalo de confianza es aleatorio.

Intervalo para una Media

Sean $(X_1, ..., X_n)$ una muestra aleatoria de la v.a. $X \sim N(\mu, \sigma^2)$.

Queremos un intervalo de confianza para μ

Tenemos dos casos:

- Caso con Varianza Conocida
- Caso con Varianza Desconocida.

Intervalo para una Media

Sean $(X_1, ..., X_n)$ una muestra aleatoria de la v.a. $X \sim N(\mu, \sigma^2)$.

Queremos un intervalo de confianza para μ

Tenemos dos casos:

- Caso con Varianza Conocida
- Caso con Varianza Desconocida.

Intervalo para una Media

Caso con Varianza Conocida

Si $X \sim N(\mu, \sigma^2)$, entonces

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$\Rightarrow$

$$Z = \frac{(\overline{X} - \mu) \cdot \sqrt{n}}{\sigma} \sim N(0,1)$$

Queremos obtener el valor de a tal que

$$P(-a \leq Z \leq a) = 1 - \alpha$$

Este valor se determina por tablas!!!

- Para $\alpha = 0,05$, se tiene que $a = 1,96$
- Para $\alpha = 0,01$, se tiene que $a = 2,56$

Intervalo para una Media

Caso con Varianza Conocida

Si $X \sim N(\mu, \sigma^2)$, entonces

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$\Rightarrow$

$$Z = \frac{(\overline{X} - \mu) \cdot \sqrt{n}}{\sigma} \sim N(0,1)$$

Queremos obtener el valor de a tal que

$$P(-a \leq Z \leq a) = 1 - \alpha$$

Este valor se determina por tablas!!!

- Para $\alpha = 0,05$, se tiene que $a = 1,96$
- Para $\alpha = 0,01$, se tiene que $a = 2,56$

Intervalo para una Media

Caso con Varianza Conocida

Si $X \sim N(\mu, \sigma^2)$, entonces

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$\Rightarrow$

$$Z = \frac{(\overline{X} - \mu) \cdot \sqrt{n}}{\sigma} \sim N(0,1)$$

Queremos obtener el valor de a tal que

$$P(-a \leq Z \leq a) = 1 - \alpha$$

Este valor se determina por tablas!!!

- Para $\alpha = 0,05$, se tiene que $a = 1,96$
- Para $\alpha = 0,01$, se tiene que $a = 2,56$

Intervalo para una Media

Caso con Varianza Conocida

Si $X \sim N(\mu, \sigma^2)$, entonces

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$\Rightarrow$

$$Z = \frac{(\overline{X} - \mu) \cdot \sqrt{n}}{\sigma} \sim N(0,1)$$

Queremos obtener el valor de a tal que

$$P(-a \leq Z \leq a) = 1 - \alpha$$

Este valor se determina por tablas!!!

- Para $\alpha = 0,05$, se tiene que $a = 1,96$
- Para $\alpha = 0,01$, se tiene que $a = 2,56$

Intervalo para una Media

Luego, se tiene

$$P(-a \leq \frac{(\overline{X}-\mu)\cdot\sqrt{n}}{\sigma} \leq a) = 1-\alpha$$

$\Rightarrow$

$$P(\overline{X} - \frac{a\cdot\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{a\cdot\sigma}{\sqrt{n}}) = 1-\alpha$$

Es decir, I.C, con nivel de confianza $1-\alpha$, es $\left[\overline{X} - \frac{a\cdot\sigma}{\sqrt{n}} \; , \; \overline{X} + \frac{a\cdot\sigma}{\sqrt{n}}\right]$
Luego, si:

- $\alpha = 0,05$, el I.C. es $\left[\overline{X} - \frac{1,96\cdot\sigma}{\sqrt{n}} \; , \; \overline{X} + \frac{1,96\cdot\sigma}{\sqrt{n}}\right]$
- $\alpha = 0,01$, el I.C. es $\left[\overline{X} - \frac{2,56\cdot\sigma}{\sqrt{n}} \; , \; \overline{X} + \frac{2,56\cdot\sigma}{\sqrt{n}}\right]$

Intervalo para una Media

Luego, se tiene

$$P(-a \leq \frac{(\overline{X}-\mu)\cdot\sqrt{n}}{\sigma} \leq a) = 1-\alpha$$

$\Rightarrow$

$$P(\overline{X} - \frac{a\cdot\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{a\cdot\sigma}{\sqrt{n}}) = 1-\alpha$$

Es decir, I.C, con nivel de confianza $1-\alpha$, es $\left[\overline{X} - \frac{a\cdot\sigma}{\sqrt{n}}\ ,\ \overline{X} + \frac{a\cdot\sigma}{\sqrt{n}}\right]$
Luego, si:

- $\alpha = 0,05$, el I.C. es $\left[\overline{X} - \frac{1{,}96\cdot\sigma}{\sqrt{n}}\ ,\ \overline{X} + \frac{1{,}96\cdot\sigma}{\sqrt{n}}\right]$
- $\alpha = 0,01$, el I.C. es $\left[\overline{X} - \frac{2{,}56\cdot\sigma}{\sqrt{n}}\ ,\ \overline{X} + \frac{2{,}56\cdot\sigma}{\sqrt{n}}\right]$

Intervalo para una Media

Luego, se tiene

$$P(-a \leq \frac{(\overline{X}-\mu)\cdot\sqrt{n}}{\sigma} \leq a) = 1-\alpha$$

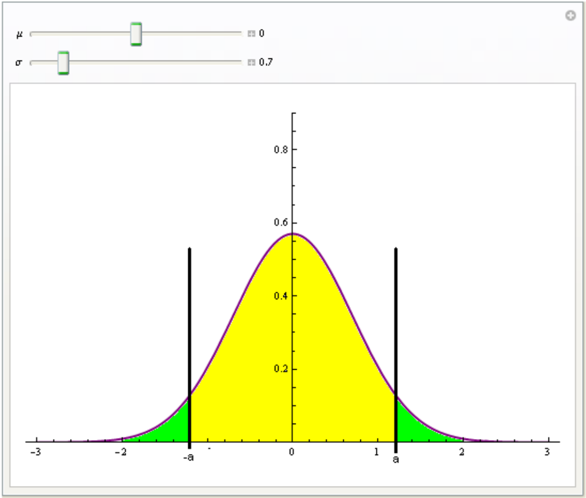
$\Rightarrow$

$$P(\overline{X} - \frac{a\cdot\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{a\cdot\sigma}{\sqrt{n}}) = 1-\alpha$$

Es decir, I.C, con nivel de confianza $1-\alpha$, es $\left[\overline{X} - \frac{a\cdot\sigma}{\sqrt{n}} \, , \, \overline{X} + \frac{a\cdot\sigma}{\sqrt{n}}\right]$
Luego, si:

- $\alpha = 0,05$, el I.C. es $\left[\overline{X} - \frac{1{,}96\cdot\sigma}{\sqrt{n}} \, , \, \overline{X} + \frac{1{,}96\cdot\sigma}{\sqrt{n}}\right]$
- $\alpha = 0,01$, el I.C. es $\left[\overline{X} - \frac{2{,}56\cdot\sigma}{\sqrt{n}} \, , \, \overline{X} + \frac{2{,}56\cdot\sigma}{\sqrt{n}}\right]$

Intervalo para una Media



Intervalo para una Media

Caso con Varianza Desconocida
La dificultad en este caso radica en buscar un estadístico cuya distribución no dependa de σ, cosa que no tenemos con el estadístico natural.

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Necesitamos la siguiente distribución.

Definición DISTRIBUCIÓN t-STUDENT

Sea $X \sim N(0,1)$ e $Y \sim \chi_n^2$ independientes. Se dice $T \sim t_n$ si es de la forma

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

Intervalo para una Media

Caso con Varianza Desconocida

La dificultad en este caso radica en buscar un estadístico cuya distribución no dependa de σ, cosa que no tenemos con el estadístico natural.

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Necesitamos la siguiente distribución.

Definición DISTRIBUCIÓN t-STUDENT

Sea $X \sim N(0,1)$ e $Y \sim \chi^2_n$ independientes. Se dice $T \sim t_n$ si es de la forma

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

t-Student

Si $T \sim t_n$, se tiene:

- $$f_T(x) = \frac{\Gamma(n+1/2)(1+\frac{x^2}{n})^{\frac{-(n+1)}{2}}}{\sqrt{n\pi}\cdot\Gamma(\frac{n}{2})}$$
- Caso $n=1$, T tiene una distribución de Cauchy.
- $E(T) = \begin{cases} 0 & si\ \ n>1 \\ \infty & si\ \ no \end{cases}$
- $Var(T) = \frac{n}{n-2} \qquad si\ \ n>2$

t-Student

Si $T \sim t_n$, se tiene:

- $$f_T(x) = \frac{\Gamma(n+1/2)(1+\frac{x^2}{n})^{\frac{-(n+1)}{2}}}{\sqrt{n\pi}\cdot\Gamma(\frac{n}{2})}$$
- Caso $n=1$, T tiene una distribución de Cauchy.
- $E(T) = \begin{cases} 0 & si\ n>1 \\ \infty & si\ no \end{cases}$
- $Var(T) = \frac{n}{n-2} \quad si\ n>2$

t-Student

Si $T \sim t_n$, se tiene:

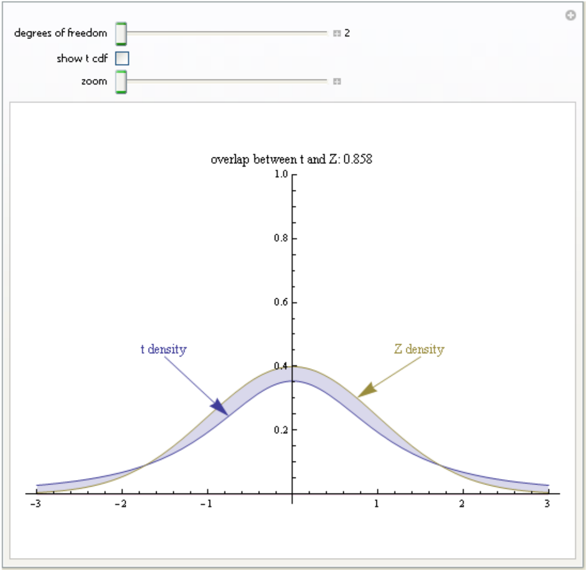
- $$f_T(x) = \frac{\Gamma(n+1/2)(1+\frac{x^2}{n})^{\frac{-(n+1)}{2}}}{\sqrt{n\pi}\cdot\Gamma(\frac{n}{2})}$$
- Caso $n = 1$, T tiene una distribución de Cauchy.
- $E(T) = \begin{cases} 0 & si \;\; n > 1 \\ \infty & si \;\; no \end{cases}$
- $Var(T) = \frac{n}{n-2} \qquad si \;\; n > 2$

t-Student

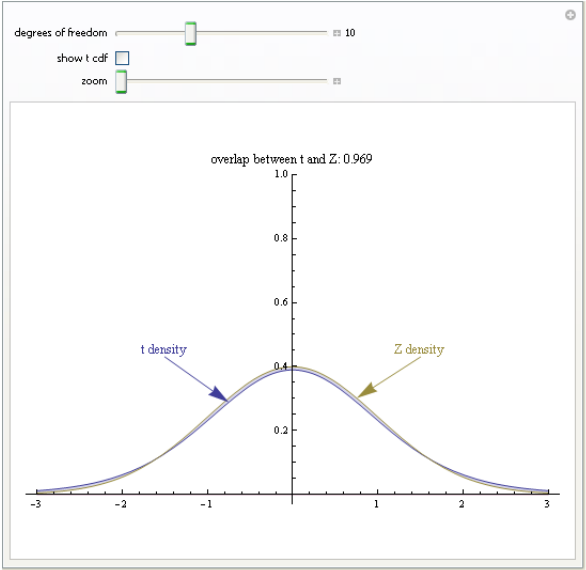
Si $T \sim t_n$, se tiene:

- $$f_T(x) = \frac{\Gamma(n+1/2)(1+\frac{x^2}{n})^{\frac{-(n+1)}{2}}}{\sqrt{n\pi}\cdot\Gamma(\frac{n}{2})}$$
- Caso $n=1$, T tiene una distribución de Cauchy.
- $E(T) = \begin{cases} 0 & si \;\; n > 1 \\ \infty & si \;\; no \end{cases}$
- $Var(T) = \frac{n}{n-2} \qquad si \;\; n > 2$

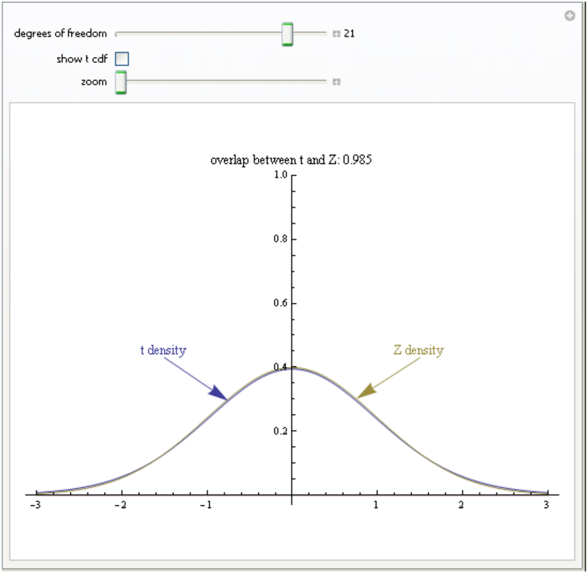
t-Student


degrees of freedom
2
show t cdf
zoom
overlap between t and Z: 0.858
1.0
0.8
0.6
0.4
0.2
t density
Z density
-3
-2
-1
1
2
3

t-Student


degrees of freedom
10
show t cdf
zoom
overlap between t and Z: 0.969
1.0
0.8
0.6
0.4
0.2
t density
Z density
-3
-2
-1
1
2
3

t-Student



degrees of freedom
21
show t cdf
zoom
overlap between t and Z: 0.985
1.0
0.8
0.6
0.4
0.2
t density
Z density
-3
-2
-1
1
2
3

Intervalo para una Media

¿De qué nos sirve la t-Student?

Recordemos que

$\overline{X}$ y S_n^2 son independientes si provienen de una distribución Normal

Además

$$Q = \frac{n}{\sigma^2} \cdot S_n^2 = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \sim \chi^2_{n-1}$$

Entonces, como $Z = \frac{\sqrt{n}}{\sigma}(\overline{X} - \mu) \sim N(0,1)$, se tiene que

$$\frac{Z}{\sqrt{\frac{Q}{n-1}}} \sim t_{n-1}$$

Intervalo para una Media

¿De qué nos sirve la t-Student?

Recordemos que

$\overline{X}$ y S_n^2 son independientes si provienen de una distribución Normal

Además

$$Q = \frac{n}{\sigma^2} \cdot S_n^2 = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \sim \chi^2_{n-1}$$

Entonces, como $Z = \frac{\sqrt{n}}{\sigma}(\overline{X} - \mu) \sim N(0,1)$, se tiene que

$$\frac{Z}{\sqrt{\frac{Q}{n-1}}} \sim t_{n-1}$$

Intervalo para una Media

¿De qué nos sirve la t-Student?

Recordemos que

$\overline{X}$ y S_n^2 son independientes si provienen de una distribución Normal

Además

$$Q = \frac{n}{\sigma^2} \cdot S_n^2 = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \sim \chi^2_{n-1}$$

Entonces, como $Z = \frac{\sqrt{n}}{\sigma}(\overline{X} - \mu) \sim N(0,1)$, se tiene que

$$\frac{Z}{\sqrt{\frac{Q}{n-1}}} \sim t_{n-1}$$

Intervalo para una Media

¿De qué nos sirve la t-Student?

Recordemos que

$\overline{X}$ y S_n^2 son independientes si provienen de una distribución Normal

Además

$$Q = \frac{n}{\sigma^2} \cdot S_n^2 = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \sim \chi_{n-1}^2$$

Entonces, como $Z = \frac{\sqrt{n}}{\sigma}(\overline{X} - \mu) \sim N(0,1)$, se tiene que

$$\frac{Z}{\sqrt{\frac{Q}{n-1}}} \sim t_{n-1}$$

Intervalo para una Media

Es decir

$$\frac{\frac{\sqrt{n}}{\sigma}(\overline{X}-\mu)}{\sqrt{\frac{\frac{1}{\sigma^2}\cdot\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}}} \sim t_{n-1}$$

Simplificando, queda

$$T = \frac{\sqrt{n-1}\cdot(\overline{X}-\mu)}{S_n} \sim t_{n-1}$$

Notemos que este estadístico no depende de σ.

Intervalo para una Media

Es decir

$$\frac{\frac{\sqrt{n}}{\sigma}(\overline{X}-\mu)}{\sqrt{\frac{\frac{1}{\sigma^2}\cdot\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}}} \sim t_{n-1}$$

Simplificando, queda

$$T = \frac{\sqrt{n-1}\cdot(\overline{X}-\mu)}{S_n} \sim t_{n-1}$$

Notemos que este estadístico no depende de σ.

Intervalo para una Media

Es decir

$$\frac{\frac{\sqrt{n}}{\sigma}(\overline{X}-\mu)}{\sqrt{\frac{\frac{1}{\sigma^2}\cdot\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}}} \sim t_{n-1}$$

Simplificando, queda

$$T = \frac{\sqrt{n-1}\cdot(\overline{X}-\mu)}{S_n} \sim t_{n-1}$$

Notemos que este estadístico no depende de σ.

Intervalo para una Media

Para encontrar el intervalo de confianza, procedemos igual que el caso anterior

Dado un nivel de confianza, buscamos (por medio de tablas) el valor de a que cumpla con:

$$P(-a \leq T \leq a) = 1 - \alpha$$

Una vez encontrado el valor de a, simplemente se despeja, creando un intervalo para μ, es decir

$$P(-a \leq \frac{\sqrt{n-1} \cdot (\overline{X} - \mu)}{S_n} \leq a) = 1 - \alpha$$

$$P(\overline{X} - \frac{a \cdot S_n}{\sqrt{n-1}} \leq \mu \leq \overline{X} + \frac{a \cdot S_n}{\sqrt{n-1}}) = 1 - \alpha$$

Intervalo para una Media

Para encontrar el intervalo de confianza, procedemos igual que el caso anterior

Dado un nivel de confianza, buscamos (por medio de tablas) el valor de a que cumpla con:

$$P(-a \leq T \leq a) = 1 - \alpha$$

Una vez encontrado el valor de a, simplemente se despeja, creando un intervalo para μ, es decir

$$P(-a \leq \frac{\sqrt{n-1} \cdot (\overline{X} - \mu)}{S_n} \leq a) = 1 - \alpha$$

$$P(\overline{X} - \frac{a \cdot S_n}{\sqrt{n-1}} \leq \mu \leq \overline{X} + \frac{a \cdot S_n}{\sqrt{n-1}}) = 1 - \alpha$$

Intervalo para una Media

Para encontrar el intervalo de confianza, procedemos igual que el caso anterior

Dado un nivel de confianza, buscamos (por medio de tablas) el valor de a que cumpla con:

$$P(-a \leq T \leq a) = 1 - \alpha$$

Una vez encontrado el valor de a, simplemente se despeja, creando un intervalo para μ, es decir

$$P(-a \leq \frac{\sqrt{n-1} \cdot (\overline{X} - \mu)}{S_n} \leq a) = 1 - \alpha$$

$$P(\overline{X} - \frac{a \cdot S_n}{\sqrt{n-1}} \leq \mu \leq \overline{X} + \frac{a \cdot S_n}{\sqrt{n-1}}) = 1 - \alpha$$

Intervalo para una Media

Para encontrar el intervalo de confianza, procedemos igual que el caso anterior

Dado un nivel de confianza, buscamos (por medio de tablas) el valor de a que cumpla con:

$$P(-a \leq T \leq a) = 1 - \alpha$$

Una vez encontrado el valor de a, simplemente se despeja, creando un intervalo para μ, es decir

$$P(-a \leq \frac{\sqrt{n-1} \cdot (\overline{X} - \mu)}{S_n} \leq a) = 1 - \alpha$$

$$P(\overline{X} - \frac{a \cdot S_n}{\sqrt{n-1}} \leq \mu \leq \overline{X} + \frac{a \cdot S_n}{\sqrt{n-1}}) = 1 - \alpha$$

En Síntesis,

Para el Caso Con Varianza Conocida
Intervalo de Confianza para μ, con nivel de confianza $1-\alpha$, es

$$\left[\overline{X} - \frac{a \cdot \sigma}{\sqrt{n}} \ , \ \overline{X} + \frac{a \cdot \sigma}{\sqrt{n}}\right]$$

donde

$$P(Z \geq a) = \frac{\alpha}{2} \quad con \ Z \sim N(0,1)$$

Para el Caso Con Varianza Desconocida
Intervalo de Confianza para μ, con nivel de confianza $1-\alpha$, es

$$\left[\overline{X} - \frac{a \cdot S_n}{\sqrt{n-1}} \ , \ \overline{X} + \frac{a \cdot S_n}{\sqrt{n-1}}\right]$$

donde

$$P(T \geq a) = \frac{\alpha}{2} \quad con \ T \sim t_{n-1}$$

En Síntesis,

Para el Caso Con Varianza Conocida

Intervalo de Confianza para μ, con nivel de confianza $1-\alpha$, es

$$\left[\overline{X} - \frac{a\cdot\sigma}{\sqrt{n}}\ ,\ \overline{X} + \frac{a\cdot\sigma}{\sqrt{n}}\right]$$

donde

$$P(Z \geq a) = \frac{\alpha}{2} \quad con\ Z \sim N(0,1)$$

Para el Caso Con Varianza Desconocida

Intervalo de Confianza para μ, con nivel de confianza $1-\alpha$, es

$$\left[\overline{X} - \frac{a\cdot S_n}{\sqrt{n-1}}\ ,\ \overline{X} + \frac{a\cdot S_n}{\sqrt{n-1}}\right]$$

donde

$$P(T \geq a) = \frac{\alpha}{2} \quad con\ T \sim t_{n-1}$$

Intervalo para la Varianza

Sean $(X_1, ..., X_n)$ una muestra aleatoria de la v.a. $X \sim N(\mu, \sigma^2)$.

Queremos ahora un intervalo de confianza para σ^2

Nuevamente, tenemos dos casos:

- Caso con Media Conocida
- Caso con Media Desconocida.

Intervalo para la Varianza

Sean $(X_1, ..., X_n)$ una muestra aleatoria de la v.a. $X \sim N(\mu, \sigma^2)$.

Queremos ahora un intervalo de confianza para σ^2

Nuevamente, tenemos dos casos:

- Caso con Media Conocida
- Caso con Media Desconocida.

Intervalo para la Varianza

Caso con Media Conocida

Si $X \sim N(\mu, \sigma^2)$, recordemos que

$$U = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \sim \chi_n^2$$

queremos encontrar a y b, tales que

$$P(a \leq U \leq b) = 1 - \alpha$$

despejando, tendríamos

$$P\left(\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{a} \right) = 1 - \alpha$$

Intervalo para la Varianza

Caso con Media Conocida
Si $X \sim N(\mu, \sigma^2)$, recordemos que

$$U = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$$

queremos encontrar a y b, tales que

$$P(a \leq U \leq b) = 1 - \alpha$$

despejando, tendríamos

$$P\left(\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{a}\right) = 1 - \alpha$$

Intervalo para la Varianza

Caso con Media Conocida

Si $X \sim N(\mu, \sigma^2)$, recordemos que

$$U = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \sim \chi_n^2$$

queremos encontrar a y b, tales que

$$P(a \leq U \leq b) = 1 - \alpha$$

despejando, tendríamos

$$P\left(\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{a}\right) = 1 - \alpha$$

Intervalo para la Varianza

¿Cómo determinamos a y b?

El criterio que se ocupa es:

$$P(U \leq a) = P(U \geq b) = \frac{\alpha}{2}$$

Intervalo para la Varianza

Caso con Media Desconocida

Si $X \sim N(\mu, \sigma^2)$, recordemos que

$$W = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

procedemos de manera similar, queremos encontrar a y b, tales que

$$P(a \leq W \leq b) = 1 - \alpha$$

despejando, tendríamos

$$P\left(\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{a} \right) = 1 - \alpha$$

a y b se determinan igual que en el caso anterior.

Intervalo para la Varianza

Caso con Media Desconocida

Si $X \sim N(\mu, \sigma^2)$, recordemos que

$$W = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

procedemos de manera similar, queremos encontrar a y b, tales que

$$P(a \leq W \leq b) = 1 - \alpha$$

despejando, tendríamos

$$P\left(\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{a} \right) = 1 - \alpha$$

a y b se determinan igual que en el caso anterior.

Intervalo para la Varianza

Caso con Media Desconocida

Si $X \sim N(\mu, \sigma^2)$, recordemos que

$$W = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{\sigma^2} \sim \chi^2_{n-1}$$

procedemos de manera similar, queremos encontrar a y b, tales que

$$P(a \leq W \leq b) = 1 - \alpha$$

despejando, tendríamos

$$P\left(\frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{a}\right) = 1 - \alpha$$

a y b se determinan igual que en el caso anterior.

Intervalo para la Varianza

Caso con Media Desconocida

Si $X \sim N(\mu, \sigma^2)$, recordemos que

$$W = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{\sigma^2} \sim \chi^2_{n-1}$$

procedemos de manera similar, queremos encontrar a y b, tales que

$$P(a \leq W \leq b) = 1 - \alpha$$

despejando, tendríamos

$$P\left(\frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{b} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{a}\right) = 1 - \alpha$$

a y b se determinan igual que en el caso anterior.

En Síntesis,

Para el Caso Con Media Conocida

Intervalo de Confianza para σ^2, con nivel de confianza $1-\alpha$, es

$$\left[\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{b}, \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{a} \right]$$

donde

$$P(U \leq a) = P(U \geq b) = \frac{\alpha}{2} \quad con \;\; U \sim \chi^2_n$$

Para el Caso Con Media Desconocida

Intervalo de Confianza para σ^2, con nivel de confianza $1-\alpha$, es

$$\left[\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{b}, \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{a} \right]$$

donde

$$P(W \leq a) = P(W \geq b) = \frac{\alpha}{2} \quad con \;\; W \sim \chi^2_{n-1}$$

En Síntesis,

Para el Caso Con Media Conocida

Intervalo de Confianza para σ^2, con nivel de confianza $1-\alpha$, es

$$\left[\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{b} , \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{a} \right]$$

donde

$$P(U \leq a) = P(U \geq b) = \frac{\alpha}{2} \quad con \;\; U \sim \chi^2_n$$

Para el Caso Con Media Desconocida

Intervalo de Confianza para σ^2, con nivel de confianza $1-\alpha$, es

$$\left[\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{b} , \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{a} \right]$$

donde

$$P(W \leq a) = P(W \geq b) = \frac{\alpha}{2} \quad con \;\; W \sim \chi^2_{n-1}$$

Intervalos de Confianza

Problema 1

Una empresa desea estimar el promedio de tiempo que necesita una secretaria para llegar a su trabajo.
Se toma una m.a.s. de **36 secretarias** y se encuentra un **promedio de 40 minutos**.
Suponiendo que el tiempo de trayecto proviene de una $N(\mu,\sigma^2)$, con $\sigma = 12$, dé un intervalo de confianza al 99% para la media.

Intervalos de Confianza

Problema 2

Se dispone de 10 muestras de sangre tomadas en las mismas condiciones a una misma persona. Se obtiene para cada una la dosis de Colesterol (en gramos) 245, 248, 250, 247, 249, 247, 247, 246, 246, 248. Cada medida puede considerarse como una realización particular de la variable "tasa de Colesterol" $X \sim N(\mu, \sigma^2)$.

a) Dé un intervalo de confianza para μ al 95 % suponiendo $\sigma^2 = 1,5$.

b) Dé un intervalo de confianza para μ al 95 % suponiendo σ^2 desconocido.

c) Construya un intervalo de confianza para σ^2 al 95 %, suponiendo μ desconocido.

Intervalo para la Diferencia de dos Medias

Problema 3

Se tienen dos muestras
$(X_1, \ldots, X_{n_1})$ una muestra aleatoria de la v.a. $X \sim N(\mu_1, \sigma_1^2)$.
$(Y_1, \ldots, Y_{n_2})$ una muestra aleatoria de la v.a. $Y \sim N(\mu_2, \sigma_2^2)$.

Queremos un intervalo de confianza para $\mu_1 - \mu_2$

Intervalo para la Diferencia de dos Medias

llamemos

$$\overline{X} = \frac{1}{n_1} \cdot \sum_{i=1}^{n_1} X_i \quad y \quad \overline{Y} = \frac{1}{n_2} \cdot \sum_{i=1}^{n_2} Y_i$$

Es claro que

$$\overline{X} \sim N(\mu_1, \frac{\sigma_1^2}{n_1}) \quad y \quad \overline{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$$

Luego, por independencia, se tiene

$$\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2 \ , \ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

Entonces, si σ_1 y σ_2 son conocidos, se procede igual que el primer caso estudiado.

Intervalo para la Diferencia de dos Medias

llamemos

$$\overline{X} = \frac{1}{n_1} \cdot \sum_{i=1}^{n_1} X_i \quad y \quad \overline{Y} = \frac{1}{n_2} \cdot \sum_{i=1}^{n_2} Y_i$$

Es claro que

$$\overline{X} \sim N(\mu_1, \frac{\sigma_1^2}{n_1}) \quad y \quad \overline{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$$

Luego, por independencia, se tiene

$$\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2 \ , \ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

Entonces, si σ_1 y σ_2 son conocidos, se procede igual que el primer caso estudiado.

Intervalo para la Diferencia de dos Medias

llamemos

$$\overline{X} = \frac{1}{n_1} \cdot \sum_{i=1}^{n_1} X_i \quad y \quad \overline{Y} = \frac{1}{n_2} \cdot \sum_{i=1}^{n_2} Y_i$$

Es claro que

$$\overline{X} \sim N(\mu_1, \frac{\sigma_1^2}{n_1}) \quad y \quad \overline{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$$

Luego, por independencia, se tiene

$$\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2 \ , \ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

Entonces, si σ_1 y σ_2 son conocidos, se procede igual que el primer caso estudiado.

Intervalo para la Diferencia de dos Medias

llamemos

$$\overline{X} = \frac{1}{n_1} \cdot \sum_{i=1}^{n_1} X_i \quad y \quad \overline{Y} = \frac{1}{n_2} \cdot \sum_{i=1}^{n_2} Y_i$$

Es claro que

$$\overline{X} \sim N(\mu_1, \frac{\sigma_1^2}{n_1}) \quad y \quad \overline{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$$

Luego, por independencia, se tiene

$$\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2 \; , \; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

Entonces, si σ_1 y σ_2 son conocidos, se procede igual que el primer caso estudiado.

Intervalo para la Diferencia de dos Medias

¿Qué pasa si σ_1 y σ_2 no se conocen?

Por un lado tenemos

$$W_1 = \frac{1}{\sigma_1^2} \cdot \sum_{i=1}^{n_1} \left(X_i - \overline{X}\right)^2 \sim \chi^2_{n_1-1}$$

Análogamente

$$W_2 = \frac{1}{\sigma_2^2} \cdot \sum_{i=1}^{n_2} \left(Y_i - \overline{Y}\right)^2 \sim \chi^2_{n_2-1}$$

Por independencia, tenemos

$$W_1 + W_2 \sim \chi^2_{n_1-1+n_2-1}$$

Intervalo para la Diferencia de dos Medias

¿Qué pasa si σ_1 y σ_2 no se conocen?

Por un lado tenemos

$$W_1 = \frac{1}{\sigma_1^2} \cdot \sum_{i=1}^{n_1} \left(X_i - \overline{X}\right)^2 \sim \chi^2_{n_1-1}$$

Análogamente

$$W_2 = \frac{1}{\sigma_2^2} \cdot \sum_{i=1}^{n_2} \left(Y_i - \overline{Y}\right)^2 \sim \chi^2_{n_2-1}$$

Por independencia, tenemos

$$W_1 + W_2 \sim \chi^2_{n_1-1+n_2-1}$$

Intervalo para la Diferencia de dos Medias

¿Qué pasa si σ_1 y σ_2 no se conocen?

Por un lado tenemos

$$W_1 = \frac{1}{\sigma_1^2} \cdot \sum_{i=1}^{n_1} \left(X_i - \overline{X}\right)^2 \sim \chi^2_{n_1 - 1}$$

Análogamente

$$W_2 = \frac{1}{\sigma_2^2} \cdot \sum_{i=1}^{n_2} \left(Y_i - \overline{Y}\right)^2 \sim \chi^2_{n_2 - 1}$$

Por independencia, tenemos

$$W_1 + W_2 \sim \chi^2_{n_1 - 1 + n_2 - 1}$$

Intervalo para la Diferencia de dos Medias

¿Qué pasa si σ_1 y σ_2 no se conocen?

Por un lado tenemos

$$W_1 = \frac{1}{\sigma_1^2} \cdot \sum_{i=1}^{n_1} \left(X_i - \overline{X}\right)^2 \sim \chi^2_{n_1-1}$$

Análogamente

$$W_2 = \frac{1}{\sigma_2^2} \cdot \sum_{i=1}^{n_2} \left(Y_i - \overline{Y}\right)^2 \sim \chi^2_{n_2-1}$$

Por independencia, tenemos

$$W_1 + W_2 \sim \chi^2_{n_1-1+n_2-1}$$

Intervalo para la Diferencia de dos Medias

Tenemos entonces

$$\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2 \ , \ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

$$W_1 + W_2 \sim \chi^2_{n_1+n_2-2}$$

Es decir

$$K = \frac{((\overline{X} - \overline{Y}) - (\mu_1 - \mu_2))}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Luego

$$\frac{K}{\sqrt{\frac{W_1+W_2}{n_1+n_2-2}}} \sim t_{n_1+n_2-2}$$

Intervalo para la Diferencia de dos Medias

Tenemos entonces

$$\overline{X}-\overline{Y}\sim N(\mu_1-\mu_2\ ,\ \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2})$$

$$W_1+W_2\sim\chi^2_{n_1+n_2-2}$$

Es decir

$$K=\frac{((\overline{X}-\overline{Y})-(\mu_1-\mu_2))}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

Luego

$$\frac{K}{\sqrt{\frac{W_1+W_2}{n_1+n_2-2}}}\sim t_{n_1+n_2-2}$$

Intervalo para la Diferencia de dos Medias

Tenemos entonces

$$\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2\ ,\ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

$$W_1 + W_2 \sim \chi^2_{n_1+n_2-2}$$

Es decir

$$K = \frac{((\overline{X} - \overline{Y}) - (\mu_1 - \mu_2))}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Luego

$$\frac{K}{\sqrt{\frac{W_1+W_2}{n_1+n_2-2}}} \sim t_{n_1+n_2-2}$$

Intervalo para la Diferencia de dos Medias

Quedando

$$\frac{\left((\overline{X}-\overline{Y})-(\mu_1-\mu_2)\right)/\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}{\sqrt{\left(\frac{n_1 S_{n_1}^2}{\sigma_1^2}+\frac{n_2 S_{n_2}^2}{\sigma_2^2}\right)/(n_1+n_2-2)}} \sim t_{n_1+n_2-2}$$

Aún sigue dependiendo de σ_1^2 y σ_2^2!!!

Intervalo para la Diferencia de dos Medias

Quedando

$$\frac{\left((\overline{X}-\overline{Y})-(\mu_1-\mu_2)\right)/\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}{\sqrt{\left(\frac{n_1 S_{n_1}^2}{\sigma_1^2}+\frac{n_2 S_{n_2}^2}{\sigma_2^2}\right)/(n_1+n_2-2)}} \sim t_{n_1+n_2-2}$$

Aún sigue dependiendo de σ_1^2 y σ_2^2!!!

Intervalo para la Diferencia de dos Medias

Si suponemos que $\sigma_1^2 = k^2 \cdot \sigma_2^2$, con k conocido, entonces se tiene que

$$\frac{\left(\left(\overline{X} - \overline{Y}\right) - (\mu_1 - \mu_2)\right) / \sqrt{\frac{1}{n_1} + \frac{k^2}{n_2}}}{\sqrt{\left(n_1 S_{n_1}^2 + \frac{n_2 S_{n_2}^2}{k^2}\right) / (n_1 + n_2 - 2)}} \sim t_{n_1+n_2-2}$$

Finalmente, se procede igual que en el segundo caso estudiado.

Intervalo para la Diferencia de dos Medias

Si suponemos que $\sigma_1^2 = k^2 \cdot \sigma_2^2$, con k conocido, entonces se tiene que

$$\frac{\left(\left(\overline{X} - \overline{Y}\right) - (\mu_1 - \mu_2)\right) / \sqrt{\frac{1}{n_1} + \frac{k^2}{n_2}}}{\sqrt{\left(n_1 S_{n_1}^2 + \frac{n_2 S_{n_2}^2}{k^2}\right) / (n_1 + n_2 - 2)}} \sim t_{n_1+n_2-2}$$

Finalmente, se procede igual que en el segundo caso estudiado.