TEMA 3: INTERVALOS DE CONFIANZA

Con la estimación puntual de un parámetro desconocido θ, asignamos a θ un valor que toma el parámetro en una determinada muestra. Tiene el inconveniente de que la estimación varia de una muestra a otra. No hay una estimación mejor que otra puesto que todas las muestras que podamos tomar son válidas para la estimación. Por ejemplo imaginemos la situación de que en una muestra de tamaño 1000 estimamos el valor del parámetro $\hat{\theta} = 50$. Con otra muestra de tamaño 5832 obtenemos una estimación de $\hat{\theta} = 78$. ¿Cuál es mejor estimación? Para evitar este problema se utiliza la estimacion por intervalos de confianza. Con la información que nos proporciona la muestra se halla un intervalo en el que, con una probabilidad (confianza) determinada a priori, se encuentra el valor desconocido del parámetro θ. La situación del ejemplo anterior ahora sería que con la primera muestra el valor de θ se encuentra en el intervalo $(45, 79)$ y con la segunda muestra este valor estaria en el intervalo $(40, 95)$. Esto es más aceptable.

Definición de intervalo aleatorio

Sean $T_1(x_1, x_2, \dots x_n)$ y $T_2(x_1, x_2, \dots x_n)$ dos estadísticos de la v.a. X. Se dice que (T_1, T_2) es un intervalo aleatorio para $h(\theta)$ con probabilidad c si:

$$P[T_1 < h(\theta) < T_2] = c \quad \forall \theta$$

Definición de intervalo de confianza

Un intervalo de confianza es la realización muestral de un intervalo aleatorio. Sea (T_1, T_2) un intervalo aleatorio de probabilidad c. Sea $(x_1, x_2, \dots, x_n)$ una realización muestral, $T_1(x_1, x_2, \dots, x_n) = t_1$, $T_2(x_1, x_2, \dots, x_n) = t_2$. Entonces (t_1, t_2) es un intervalo de confianza de nivel c para $h(\theta)$ y para la realización muestral $(x_1, x_2, \dots x_n)$

si verifica que:

$$P[t_1 < h(\theta) < t_2] = c$$

¿Qué diferencia hay entre un intervalo aleatorio de probabilidad c y un intervalo de confianza de nivel c?

Sea el estadístico $T_1(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i$,

$T_2(x_1, x_2, \ldots x_n) = \sum_{i=1}^{n} x_i^2$

Está claro que $T_1 < T_2$. Sea $h(\theta) = \theta^2 - 3\theta + 2$. Decimos que (T_1, T_2) es un intervalo aleatorio para $h(\theta)$ con probabilidad $\frac{3}{4}$ si se cumple que:

$$P\left[\sum_{i=1}^{n} x_i < \theta^2 - 3\theta + 2 < \sum_{i=1}^{n} x_i^2\right] = \frac{3}{4}$$

El correspondiente intervalo de confianza para $h(\theta)$ de nivel $\frac{3}{4}$ para la realización muestral 3, 7, 5 es $(15, 83)$ ya que:

$$\sum_{i=1}^{n} x_i = 3+7+5 = 15 = t_1 \qquad \sum_{i=1}^{n} x_i^2 = 3^2+7^2+5^2 = 83$$

$$P[15 < \theta^2 - 3\theta + 2 < 83] = \frac{3}{4}$$

Coeficiente de confianza, nivel de confianza y precisión

El nivel de confianza c nos indica la proporción del intervalo (t_1, t_2) que contiene al parámetro $h(\theta)$. El nivel de confianza puede ser del 90%, del 95%, de 60%, etc.

El coeficiente de confianza o nivel de significación $1-c$ nos indica la proporción del intervalo que no contiene al parámetro $h(\theta)$. La amplitud del intervalo es $t_2 - t_1$

Cuanto mayor es la precisión menor es la amplitud del intervalo y el nivel de confianza

Precisión	longitud del intervalo $t_2 - t_1$	Nivel de confianza c
$\uparrow$	$\downarrow$	$\downarrow$
$\downarrow$	$\uparrow$	$\uparrow$

Métodos de construcción de intervalos de confianza.

a) Método pivotal

Este método para la construcción de un intervalo de confianza sigue cuatro etapas:

1ª Etapa: hallar una función pivote $v_\theta(x_1, x_2, \dots x_n)$ que depende de la muestra y del parámetro sobre el que se quiere construir el intervalo de confianza. La distribución de $v_\theta(x_1, x_2, \dots x_n)$ está libre de parámetros desconocidos y en particular está libre de θ.

2ª Etapa: Seleccionar dos percentiles q_1 y q_2 de v_θ tales que $P[q_1 < v_\theta < q_2] = c$, q_1 y q_2 son constantes conocidas que no dependen de θ.

3ª Etapa: Pivotar (hacer transformaciones) sobre la desigualdad $q_1 < v_\theta < q_2$ para llegar a una expresión equivalente en la forma $T_1(x_1, x_2, \dots x_n) < h(\theta) < T_2(x_1, x_2 \dots x_n)$ que tenga la forma de un intervalo aleatorio ~~para~~ $h(\theta)$. El proceso debe ser tal que T_1 y T_2 no dependan de θ y que las desigualdades que nos quedan al final sean iguales que las que teníamos inicialmente, es decir:

$$P[T_1 < h(\theta) < T_2] = P[q_1 < v_\theta < q_2] = c \text{, por tanto}$$

(T_1, T_2) es un intervalo aleatorio de nivel c.

4ª Etapa: tomamos una realización muestral y obtenemos un intervalo de confianza (IC) para $h(\theta)$ de nivel c, que es:

$$\left(T_1(x_1, x_2, \dots x_n), T_2(x_1, x_2, \dots, x_n)\right)$$ siendo $(x_1, x_2 \dots x_n)$ una realización muestral

Para que este proceso sea factible han de ser v_θ y $h(\theta)$ monótonas (crecientes o decrecientes).

¿Cómo buscamos la función pivote?

Podemos tomar un estimador suficiente de θ o bien el estimador de máxima verosimilitud, haciendo en ellos transformaciones para que la función pivotal tenga una distribución independiente de θ.

¿Es posible usar el nivel de confianza que queramos?
Si la distribución es continua sí lo es
En caso de que la variable se distribuya según una discreta no está garantizado y en este caso se toma como norma escoger q_1 y q_2 de modo que den el nivel de confianza más próximo al deseado.

Unicidad de q_1 y q_2
q_1 y q_2 no son únicos, hay varias posibilidades. Aplicamos el criterio de la precisión que consiste en minimizar la amplitud del intervalo

$$\text{Min}: T_2(X_1, X_2, \dots X_n) - T_1(X_1, X_2, \dots X_n)$$

$$\text{s.a}: P(q_1 < V_\theta < q_2) = c$$

Veamos un ejemplo.

Ejemplo
Sea la v.a. $X \rightsquigarrow Exp\left(\frac{1}{\theta}\right)$. Construir a partir de una muestra de tamaño 1, un intervalo de confianza para θ con confianza $c = 0'9$.

Vamos a hallar una función pivote a partir de un estimador suficiente de θ. La distribución exponencial pertenece a la familia exponencial ya que su función de densidad se puede expresar como
$f(x) = a(\theta) \cdot b(x) \cdot e^{c(\theta) \cdot d(x)}$

Si $X \rightsquigarrow Exp\left(\frac{1}{\theta}\right)$ la función de densidad de X es

$$f_X(x) = \begin{cases} \frac{1}{\theta} \cdot e^{-\frac{1}{\theta}x} & \text{si } x > 0 \\ 0 & \text{en el resto} \end{cases}$$

$a(\theta) = \frac{1}{\theta}$ $\quad b(x) = 1$
$c(\theta) = -\frac{1}{\theta}$ $\quad d(x) = x$.

ETAPA 1:

El estadístico suficiente para una muestra de tamaño n es: $T(X_1, X_2, \dots X_n) = \sum_{i=1}^{n} d(x_i) = \sum_{i=1}^{n} x_i$. Si la muestra es de tamaño 1, el estadístico suficientes es $T(X_1)=X_1$

Al ser la función de densidad de $T(X_1)$

$$f_T(x_1) = \frac{1}{\theta} \cdot e^{-\frac{x_1}{\theta}} \quad x_1 > 0$$

ya que $T \rightsquigarrow Exp\left(\frac{1}{\theta}\right)$, no nos sirve $T(X_1) = X_1$ como función pivotal. Tenemos que hacer una transformación para conseguir una función pivotal. Hacemos el cambio $T = \frac{X_1}{\theta}$. Veamos cual es la función de densidad de T.

Hallamos la función de distribución de T y derivando obtendremos la función de densidad de T.

$$F_T(t) = P[T \le t] = P\left[\frac{X_1}{\theta} < t\right] = P[X_1 < t \cdot \theta] =$$

$$= \int_{-\infty}^{t\cdot\theta} f_X(x)\, dx = \int_0^{t\cdot\theta} \frac{1}{\theta} \cdot e^{-\frac{x}{\theta}} dx = \frac{1}{\theta} \cdot \left[-\theta \cdot e^{-\frac{x}{\theta}}\right]_0^{t\cdot\theta}$$

$= 1 - e^{-t} \Rightarrow f_T'(t) = e^{t}$, ahora la función de densidad no depende de θ, podemos tomar como función pivotal $v_\theta(X_1) = \frac{X_1}{\theta}$. Además como

$$\frac{\partial v_\theta(X_1)}{\partial \theta} = -\frac{X_1}{\theta^2} < 0 \Rightarrow v_\theta(X_1) \text{ es decreciente.}$$

Al ser $h(\theta) = \theta \Rightarrow \frac{\partial h(\theta)}{\partial \theta} = 1 > 0$ creciente.

ETAPA 2:

Elegimos q_1 y q_2 de modo que $P[q_1 < v_\theta(X_1) < q_2] = 0'95$

ETAPA: 3

Tenemos que hacer transformaciones para llegar a una expresión en la forma $P(T_1 < h(\theta) < T_2) = 0'95$

$$q_1 < v_\theta(X_1) < q_2 \Rightarrow q_1 < \frac{X_1}{\theta} < q_2$$

$$q_1 < \frac{x_1}{\theta} \Rightarrow \theta < \frac{x_1}{q_1}$$

$$\frac{x_1}{\theta} < q_2 \Rightarrow x_1 < \theta q_2 \Rightarrow \frac{x_1}{q_2} < \theta$$

$$\Rightarrow \frac{x_1}{q_2} < \theta < \frac{x_1}{q_1}$$

El intervalo aleatorio es $\left(\frac{x_1}{q_2}, \frac{x_1}{q_1} \right)$

Tenemos que calcular q_1 y q_2. Sabemos que:

$P[q_1 < v_\theta < q_2] = 0'95$. Por tanto se verifica que

$P[v_\theta < q_1] = 0'05$; $P[q_2 < v_\theta] = 0'05 \Rightarrow 1 - P[v_\theta < q_2] = 0'05$

$$\int_0^{q_1} e^{-t} dt = 0'05 \qquad 1 - \int_0^{q_2} e^{-t} dt = 0'05$$

Luego:

$$\left[-e^{-t}\right]_0^{q_1} = 0'05 \Rightarrow 1 - e^{-q_1} = 0'05 \Rightarrow e^{-q_1} = 0'95 \Rightarrow q_1 = -\ln 0'95$$

$$1 - \left[-e^{-t}\right]_0^{q_2} = 0'05 \Rightarrow 1 - (-e^{-q_2} + 1) = 0'05 \Rightarrow e^{-q_2} = 0'05$$

$$\Rightarrow q_2 = -\ln 0'05$$

Luego el intervalo aleatorio de la muestra aleatoria de tamaño 1 es:

$$IC = \left(-\frac{x_1}{\ln 0'05}, -\frac{x_1}{\ln 0'95} \right)$$

INTERVALOS DE CONFIANZA ASINTOTICOS

Para muestras grandes utilizaremos resultados asintóticos. Para construir un intervalo de confianza asintótico necesitamos:

- El método pivotal
- El T.C.L: Si $X_1, X_2, \ldots, X_n$ es una m.a.s. de una población de media μ y desviación típica σ, entonces $\bar{X} \rightsquigarrow N(\mu, \frac{\sigma}{\sqrt{n}})$
- El nivel de confianza

Intervalo de confianza para la media de una población normal con varianza conocida.

Sea X una v.a. que se distribuye según una población normal $N(\mu, \sigma)$, con σ conocida. Si el nivel de significación es $\alpha \Rightarrow C = 1-\alpha$ es el nivel de confianza, queremos hallar un intervalo de confianza para μ. Vamos a emplear el método pivotal.

Si $X_1, X_2, \ldots, X_n$ es una m.a.s., según el T.C.L.

$\bar{X} \rightsquigarrow N(\mu, \sigma)$. La función $\bar{X} = \frac{\sum_1^n X_i}{n} \rightsquigarrow N(\mu, \sigma)$

no nos sirve como función pivotal ya que su función de densidad depende de un parámetro desconocido que es μ. Si tipificamos obtenemos que:

$$\frac{X-\mu}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow N(0,1).$$ La función $$v_\mu = \frac{X-\mu}{\frac{\sigma}{\sqrt{n}}}$$ si

puede ser función pivotal ya que su distribución no depende de parámetro alguno. Queremos encontrar un intervalo (λ_1, λ_2) en donde se verifique

$P[\lambda_1 < v_\mu < \lambda_2] = 1-\alpha$. Vamos a pivotar la desigualdad $\lambda_1 < v_\mu < \lambda_2$ para acotar μ:

$$\lambda_1 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \lambda_2 \Rightarrow \lambda_1 \cdot \frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < \lambda_2 \cdot \frac{\sigma}{\sqrt{n}}$$

si cambiamos el signo:

$$-d_2 \cdot \frac{\sigma}{\sqrt{n}} < \mu - \bar{x} < d_1 \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow \bar{x} - d_2 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - d_1 \cdot \frac{\sigma}{\sqrt{n}}$$

El intervalo de confianza $(1-\alpha)$ en donde se encuentra μ es:

$$IC_{(1-\alpha)} = \left(\bar{x} - d_2 \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - d_1 \cdot \frac{\sigma}{\sqrt{n}} \right)$$

Elegimos el intervalo con menor amplitud: la amplitud viene dada por:

$$A(d_1, d_2) = \bar{x} - d_1 \cdot \frac{\sigma}{\sqrt{n}} - \bar{x} + d_2 \frac{\sigma}{\sqrt{n}} = (d_2 - d_1) \cdot \frac{\sigma}{\sqrt{n}}$$

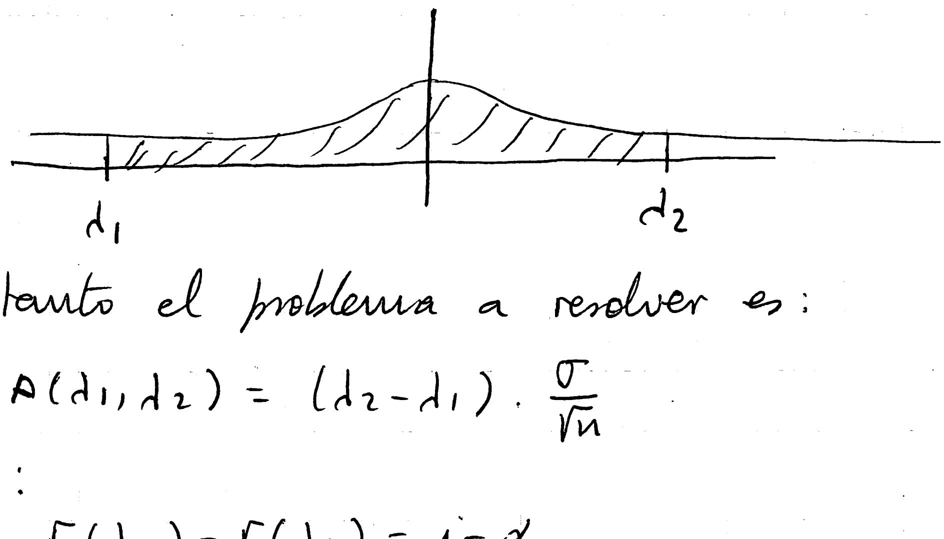
El problema que tenemos que resolver ahora es:

$$\text{Min } A(d_1, d_2) = (d_2 - d_1) \cdot \frac{\sigma}{\sqrt{n}}$$

$$\text{s.a: } P[d_1 < v(\mu) < d_2] = 1 - \alpha$$

Pero si $F(x)$ es la distribución de la normal $N(0,1)$ entonces:

$$P[d_1 < v(\mu) < d_2] = F(d_2) - F(d_1) = 1 - \alpha$$



Por tanto el problema a resolver es:

$$\text{Min } A(d_1, d_2) = (d_2 - d_1) \cdot \frac{\sigma}{\sqrt{n}}$$

s.a:

$$F(d_2) - F(d_1) = 1 - \alpha$$

Podemos resolver este problema por el método de los multiplicadores de Lagrange.

La función de Lagrange (o lagrangiana) es:

$$L(d_1, d_2, \lambda) = (d_2 - d_1) \cdot \frac{\sigma}{\sqrt{n}} - \lambda \cdot (F(d_2) - F(d_1) + \alpha - 1)$$

Ha de verificarse en un punto extremo (máximo o mínimo) que:

$$\frac{\partial L}{\partial d_1} = -\frac{\sigma}{\sqrt{n}} + \lambda \cdot F'(d_1) = 0 \Rightarrow -\frac{\sigma}{\sqrt{n}} + \lambda \cdot f(d_1) = 0$$

siendo f la función de densidad de $N(0,1)$

$$\frac{\partial L}{\partial d_2} = +\frac{\sigma}{\sqrt{n}} - \lambda \cdot F'(d_2) = 0 \Rightarrow +\frac{\sigma}{\sqrt{n}} - \lambda f(d_2) = 0$$

$$\frac{\partial L}{\partial \lambda} = F(d_2) - F(d_1) + \alpha - 1 = 0$$

El sistema obtenido es:

$$\left.\begin{array}{l} \lambda f(d_1) = \frac{\sigma}{\sqrt{n}} \\ \lambda f(d_2) = \frac{\sigma}{\sqrt{n}} \end{array}\right\} \Rightarrow f(d_1) = f(d_2) = \frac{\sigma}{\lambda \sqrt{n}}$$

$$F(d_2) - F(d_1) = 1 - \alpha$$

La función de densidad de la distribución normal $N(0,1)$ es:

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$$ es simétrica.

Al ser $f(d_1) = f(d_2) \Rightarrow \left\{\begin{array}{l} d_1 = d_2 \text{ (absurdo ya que suponemos } d_1 < d_2) \\ \text{ó} \\ d_1 = -d_2 \end{array}\right.$

Ha de ser $d_1 = -d_2$.

Como $F(d_2) - F(d_1) = 1 - \alpha \Rightarrow F(d_2) - F(-d_2) = 1 - \alpha$

Como $F(-d_2) = 1 - F(d_2)$ ya que $P[v(\mu) < -d_2] =$
$= P[d_2 < v(\mu)] = 1 - P(v(\mu) < d_2) = 1 - F(d_2)$:

$F(\lambda_2) - (1 - F(\lambda_2)) = 1 - \alpha$

$2F(\lambda_2) = 2 - \alpha \Rightarrow F(\lambda_2) = 1 - \frac{\alpha}{2}$

Luego λ_2 es el valor en donde:

$P[\,\upsilon(\mu) < \lambda_2\,] = 1 - \frac{\alpha}{2}$ y a este valor se le representa por $z_{1-\frac{\alpha}{2}}$, es decir $\lambda_2 = z_{1-\frac{\alpha}{2}} \Rightarrow$

$\Rightarrow \lambda_1 = -z_{1-\frac{\alpha}{2}}$. Por tanto el intervalo de confianza para la media μ de una normal $N(\mu, \sigma)$ de varianza σ conocida es:

$$IC_{(1-\alpha)} = \left(\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

Veamos un ejemplo.

Ejemplo: consideremos una población $X \rightsquigarrow N(\mu, 5)$ de la que se toma una muestra de tamaño $n = 100$. Si el nivel de significación es $\alpha = 0'05$, hallar el intervalo de confianza para la media poblacional

$1 - \alpha = 1 - 0'05 = 0'95$.

El intervalo de confianza 95% es:

$$IC_{0'95}(\mu) = \left(\bar{X} - z_{1-\frac{0'05}{2}} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{1-\frac{0'05}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

Como $z_{1-\frac{0'05}{2}} = z_{0'975}$, si buscamos en la tabla de la normal: $z_{0'975} = 1'96$. Entonces

$$IC_{0'95}(\mu) = \left(\bar{X} - 1'96 \cdot \frac{5}{\sqrt{100}},\ \bar{X} + 1'96 \cdot \frac{5}{\sqrt{100}}\right)$$

$$IC_{0'95}(\mu) = (\bar{X} - 0'98,\ \bar{X} + 0'98)$$

Si nos dan el dato adicional de que en la muestra de tamaño 100 la media es $\bar{X} = 15$, entonces:

$$IC_{0'95}(\mu) = (0'02,\ 15'98)$$

INTERVALO DE CONFIANZA PARA LA MEDIA DE UNA NORMAL CON VARIANZA DESCONOCIDA

Si $X \rightsquigarrow N(\mu, \sigma)$ según el lema de Fisher se tiene que $\bar{X} \rightsquigarrow N(\mu, \frac{\sigma}{\sqrt{n}}) \Rightarrow \dfrac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow N(0,1)$.

$$v(\mu) = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$$ no puede ser función pivotal,

si bien su distribución $N(0,1)$ no contiene parámetros desconocidos, esta función depende del parámetro a estimar μ, de la muestra y de un parámetro desconocido que es σ.

Según el lema de Fisher $\dfrac{nS^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$, siendo además $\bar{X}$ y S^2 independientes, donde

$S^2 = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$, $X_1, X_2, \ldots X_n$ m.a.s.

Podemos poner que:

$$\frac{\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{\frac{nS^2}{\sigma^2}}{n-1}}} \rightsquigarrow \frac{N(0,1)}{\sqrt{\frac{\chi^2_{n-1}}{n-1}}} = t_{n-1}$$

Por tanto $$\frac{\bar{x}-\mu}{\frac{\cancel{\sigma}}{\cancel{\sqrt{n}}} \cdot \frac{S\cdot\cancel{\sqrt{n}}}{\cancel{\sigma}\cdot\sqrt{n-1}}} = \frac{x-\mu}{\frac{S}{\sqrt{n-1}}} \rightsquigarrow t_{n-1}$$

$$v(\mu) = \frac{x-\mu}{\frac{S}{\sqrt{n-1}}}$$ sí puede ser función pivotal,

depende tan sólo de μ (parámetro a estimar) y de la muestra, su distribución es una t-Student con $(n-1)$ grados de libertad (simétrica) que no tiene parámetros desconocidos.

Tenemos que hallar un intervalo (λ_1, λ_2) tal que:

$$P[\lambda_1 < v(\mu) < \lambda_2] = 1-\alpha$$

Vamos a pivotar la expresión $\lambda_1 < v(\mu) < \lambda_2$ para acotar el parámetro μ.

$$\lambda_1 < v(\mu) < \lambda_2 \Rightarrow \lambda_1 < \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n-1}}} < \lambda_2 \Rightarrow$$

$$\Rightarrow \lambda_1 \cdot \frac{S}{\sqrt{n-1}} < \bar{x}-\mu < \lambda_2 \cdot \frac{S}{\sqrt{n-1}}$$ si cambiamos el signo:

$$-\lambda_2 \cdot \frac{S}{\sqrt{n-1}} < \mu - \bar{x} < -\lambda_1 \frac{S}{\sqrt{n-1}} \Rightarrow$$

$$\Rightarrow \bar{x} - \lambda_2 \cdot \frac{S}{\sqrt{n-1}} < \mu < \bar{x} - \lambda_1 \cdot \frac{S}{\sqrt{n-1}}$$

Los intervalos aleatorio en los que se encuentra μ con una confianza de $1-\alpha$ son de la forma

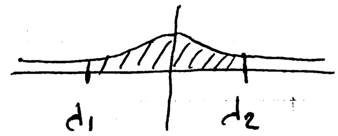
$\left(\bar{x} - \lambda_2 \frac{S}{\sqrt{n-1}}, \bar{x} - \lambda_1 \cdot \frac{S}{\sqrt{n-1}}\right)$, para decidir cual elegir de todos ellos seguiremos el criterio de máxima precisión que es lo mismo que decir menor amplitud sujeto a que la probabilidad en el intervalo sea $1-\alpha$. La amplitud del intervalo es:

$$A(\lambda_1, \lambda_2) = \bar{x} - \lambda_1 \cdot \frac{S}{\sqrt{n-1}} - \left(\bar{x} - \lambda_2 \cdot \frac{S}{\sqrt{n-1}}\right)$$

$$A(\lambda_1, \lambda_2) = (\lambda_2 - \lambda_1) \cdot \frac{S}{\sqrt{n-1}}$$

El problema que hay que resolver es:

$$\text{Min } A(\lambda_1, \lambda_2) = (\lambda_2 - \lambda_1) \cdot \frac{S}{\sqrt{n-1}}$$



s.a:

$$P(\lambda_1 < v(\mu) < \lambda_2) = 1 - \alpha \Rightarrow F(\lambda_2) - F(\lambda_1) = 1 - \alpha$$

siendo $F(\lambda_2) = P[v(\mu) < \lambda_2]$; $F[\lambda_1] = P[v(\mu) < \lambda_1]$

La función de Lagrange es:

$$L(\lambda_1, \lambda_2, \lambda) = (\lambda_2 - \lambda_1) \cdot \frac{S}{\sqrt{n-1}} - \lambda \left[F(\lambda_2) - F(\lambda_1) - 1 + \alpha \right]$$

Aplicamos las condiciones de 1er orden para obtener un extremo (en este caso mínimo):

$$\frac{\partial L}{\partial \lambda_1} = - \frac{S}{\sqrt{n-1}} + \lambda \cdot f(\lambda_1) = 0 \Rightarrow f(\lambda_1) = \frac{S}{\lambda \sqrt{n-1}}$$

$$\frac{\partial L}{\partial \lambda_2} = \frac{S}{\sqrt{n-1}} - \lambda \cdot f(\lambda_2) = 0 \Rightarrow f(\lambda_2) = \frac{S}{\lambda \sqrt{n-1}}$$

$$\frac{\partial L}{\partial \lambda} = F(\lambda_2) - F(\lambda_1) - 1 + \alpha = 0$$

siendo $f(\lambda_1) = F'(\lambda_1)$, $F'(\lambda_2) = f(\lambda_2)$. (la derivada de la función de distribución es la función de densidad)

De las ecuaciones anteriores se desprende que la función de densidad de la t-Student coincide en los valores λ_1 y λ_2, es decir $f(\lambda_1) = f(\lambda_2)$ y como es simétrica:

$$f(\lambda_1) = f(\lambda_2) \Rightarrow \begin{cases} \lambda_1 = \lambda_2 \text{ (absurdo ya que } \lambda_1 < \lambda_2) \\ \lambda_2 = -\lambda_1 \end{cases}$$

$$F(\lambda_2) - F(\lambda_1) - 1 + \alpha = 0 \Rightarrow F(\lambda_2) - F(-\lambda_2) = 1 - \alpha \Rightarrow$$

$$\Rightarrow F(\lambda_2) - (1 - F(\lambda_2)) = 1 - \alpha \Rightarrow 2F(\lambda_2) = 2 - \alpha \Rightarrow$$

$\Rightarrow F(\lambda_2) = 1 - \frac{\alpha}{2}$. Al valor λ_2 de la distribución t-Student tal que $P(X < \lambda_2) = 1 - \alpha$ lo representamos por $t_{n-1, 1-\frac{\alpha}{2}}$

Entonces el intervalo de confianza que buscamos es:

$$IC_{(1-\alpha)}(\mu) = \left(\bar{x} - t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n-1}} , \bar{x} + t_{n-1, 1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}} \right)$$

Veamos un ejemplo

Ejemplo:

Supongamos que una determinada población se distribuye según una normal $N(\mu, \sigma)$ con media y varianza desconocidas. Tomamos una muestra de tamaño $n = 9$ cuya media es $\bar{x} = 40$, siendo la varianza muestral $S^2 = 5$. Si el nivel de significación es del 2%, hallar un intervalo de confianza para la media poblacional μ.

Solución:

Según hemos visto:

$$IC_{1-\alpha}(\mu) = \left(\bar{x} - t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n-1}} , \bar{x} + t_{n-1, 1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}} \right)$$

$\alpha = 0'02 \Rightarrow 1 - \frac{\alpha}{2} = 0'99 \Rightarrow t_{n-1, 1-\frac{\alpha}{2}} = t_{8, 0'99} = 2'896$

(hay que mirar en la tabla t-Student)

$$IC_{98\%}(\mu) = \left(40 - 2'896 \frac{\sqrt{5}}{\sqrt{8}} , 40 + 2'896 \cdot \frac{\sqrt{5}}{\sqrt{8}} \right) = (37'709, 42'29)$$

INTERVALO DE CONFIANZA PARA LA VARIANZA DE UNA POBLACION NORMAL DE MEDIA CONOCIDA

Sea $X \rightsquigarrow N(\mu, \sigma)$. Tomamos una m.a.s. $X_1, X_2, \dots X_n$

Entonces $X_i \rightsquigarrow N(\mu, \sigma)$ $\forall i = 1 \dots n$. Si tipificamos:

$$\frac{x_i - \mu}{\sigma} \rightsquigarrow N(0,1) \Rightarrow \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^2 \rightsquigarrow \chi^2_n \Rightarrow$$

$$\Rightarrow \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} (x_i - \mu)^2 = v(\sigma) \rightsquigarrow \chi^2_n .$$

Podemos tomar $v(\sigma)$ como función pivotal para el parámetro σ ya que v depende tan solo de la muestra y del parametro a estimar σ^2, además su distribución χ^2_n no depende de parametros desconocidos. Si el nivel de significación es α, pretendemos hallar un intervalo (λ_1, λ_2) tal que:

$$P[\lambda_1 < v(\sigma) < \lambda_2] = 1-\alpha$$

Vamos a pivotar la desigualdad $\lambda_1 < v(\sigma) < \lambda_2$ para acotar σ^2:

$$\lambda_1 < \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2 < \lambda_2$$

$$\frac{1}{\lambda_2} < \frac{1}{\sum_{i=1}^{n} (x_i-\mu)^2} \sigma^2 < \lambda_2 \Rightarrow \frac{1}{\lambda_2} \sum_{i=1}^{n} (x_i-\mu)^2 < \sigma^2 < \frac{1}{\lambda_1} \sum_{i=1}^{n} (x_i-\mu)^2$$

El intervalo en el que está σ^2 es de la forma:

$$\left(\frac{1}{\lambda^2} \sum_{i=1}^{n} (x_i-\mu)^2 , \frac{1}{\lambda_1} \cdot \sum_{i=1}^{n} (x_i-\mu)^2 \right)$$

Hemos de imponer la condición de que la amplitud sea mínima:

$$A(\lambda_1, \lambda_2) = \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) \sum_{i=1}^{n} (x_i-\mu)^2$$

El problema que hemos de resolver es:

Minimizar $A(\lambda_1, \lambda_2)$

s.a. $P[\lambda_1 < v(\sigma) < \lambda_2] = 1-\alpha$

Si F es la función de distribución para χ^2_n entonces

$$P[\lambda_1 < v(\sigma) < \lambda_2] = 1-\alpha \Leftrightarrow F(\lambda_2) - F(\lambda_1) = 1-\alpha$$

De forma más clara, el problema a resolver es:

$$\text{Min } A(\lambda_1, \lambda_2) = \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) \sum_{i=1}^{n} (x_i-\mu)^2$$

s.a: $F(\lambda_2) - F(\lambda_1) = 1-\alpha$

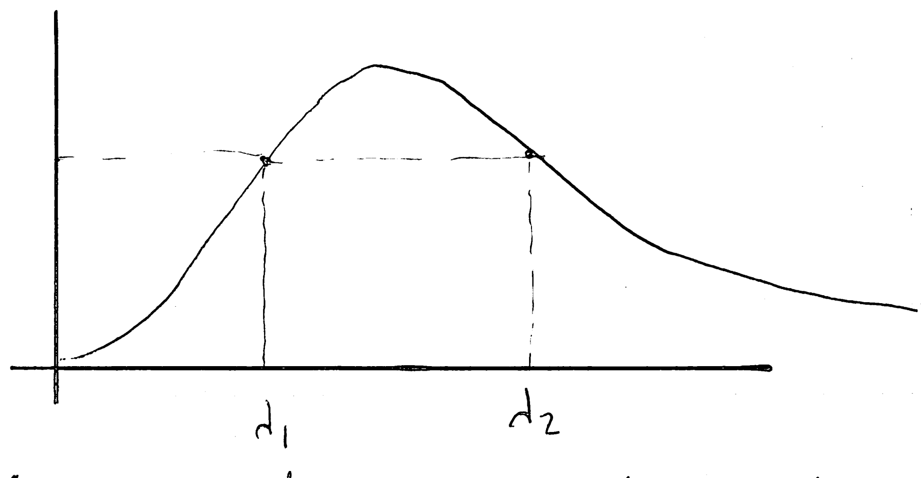
La función de Lagrange en este caso vale:

$$L(d_1, d_2, \lambda) = \left(\frac{1}{d_1} - \frac{1}{d_2}\right) \sum_{i=1}^{n} (x_i - \mu)^2 - \lambda \left[F(d_2) - F(d_1) + \alpha - 1\right]$$

Las condiciones de 1er orden para la existencia de mínimo son:

$$\left.\begin{aligned}
\frac{\partial L}{\partial d_1} = 0 &\Rightarrow -\frac{1}{d_1^2} \sum_{i=1}^{n} (x_i - \mu)^2 + \lambda \cdot f(d_1) = 0 \\
\frac{\partial L}{\partial d_2} = 0 &\Rightarrow +\frac{1}{d_2^2} \sum_{i=1}^{n} (x_i - \mu)^2 - \lambda f(d_2) = 0 \\
\frac{\partial L}{\partial \lambda} = 0 &\Rightarrow F(d_2) - F(d_1) = 1 - \alpha
\end{aligned}\right\}$$

En las ecuaciones anteriores $f(d_1) = F'(d_1)$ es la función de densidad de χ^2_n.
Obtenemos la igualdad $d_1^2 f(d_1) = d_2^2 f(d_2)$ pero esta igualdad es poco manejable y al ser la distribución χ^2_n no simétrica nos sirve de poco. ¿Qué hacer entonces?



Lo menos chapucero es hacer lo siguiente:
Asignamos a la probabilidad que deja en la derecha d_2 el valor $\frac{\alpha}{2}$, lo mismo que la probabilidad que deja en la izquierda d_1, tambien $\frac{\alpha}{2}$.
Es decir

$$P(X < d_1) = \frac{\alpha}{2} \qquad P[X > d_2) = \frac{\alpha}{2} \Rightarrow P[X < d_2) = 1 - \frac{\alpha}{2}$$

De esta manera $F(d_1) = \frac{\alpha}{2}$, $F(d_2) = 1 - \frac{\alpha}{2}$ y se verifica $F(d_2) - F(d_1) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$, de este modo podemos calcular d_1 y d_2. Veamos un ejemplo.

Ejemplo: En una población $N(12, \sigma^2)$ con varianza desconocida, tomamos una muestra de tamaño $n=6$ con un nivel de significación $\alpha = 0'02$ y teniendo en cuenta que $\sum_{i=1}^{6} (x_i - 12)^2 = 10$, hallar un intervalo de confianza para la varianza poblacional.

$\frac{\alpha}{2} = 0'01$. Si miramos en la tabla Ji-cuadrado con 6 grados de libertad

$$F(d_1) = 0'01 \Rightarrow d_1 = 0'872$$
$$F(d_2) = 1 - 0'01 = 0'99 \Rightarrow d_2 = 16'8$$

Entonces el intervalo de confianza al 98% de la varianza σ^2 es:

$$IC_{98\%}(\sigma^2) = \left(\frac{1}{16'8} \cdot 10 , \frac{1}{0'872} \cdot 10 \right) = (0'59 , 11'47)$$

INTERVALO DE CONFIANZA PARA LA VARIANZA DE UNA POBLACIÓN NORMAL CON MEDIA DESCONOCIDA.

Sea $X \rightsquigarrow N(\mu, \sigma)$, con μ desconocida. Sea $x_1, x_2, \dots x_n$ una m.a.s. Según el lema de Fisher-Cochran se tiene que $\frac{n s^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$ siendo:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n} \quad ; \quad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Podemos tomar como función pivotal la función $v(\sigma) = \frac{n s^2}{\sigma^2}$ ya que tan sólo depende de la muestra y del parámetro a estimar σ^2. Además su distribución χ^2_{n-1} no depende de parámetros desconocidos.

Si el nivel de significación es α, queremos hallar un intervalo (d_1, d_2) tal que:

$P[d_1 < v(\sigma) < d_2] = 1-\alpha$. Esto es equivalente a poner que $F(d_2) - F(d_1) = 1-\alpha$ siendo F la función de distribución de χ^2_{n-1}.

Pivotamos sobre la desigualdad $d_1 < v(\sigma) < d_2$ para acotar σ^2:

$$d_1 < \frac{ns^2}{\sigma^2} < d_2 \Rightarrow \frac{1}{d_2} < \frac{\sigma^2}{ns^2} < \frac{1}{d_1} \Rightarrow$$

$\Rightarrow \dfrac{ns^2}{d_2} < \sigma^2 < \dfrac{ns^2}{d_1}$. Por tanto el intervalo en donde se encuentra σ^2 es de la forma: $\left(\dfrac{ns^2}{d_2}, \dfrac{ns^2}{d_1}\right)$

La amplitud de este intervalo es:

$$A(d_1, d_2) = \frac{ns^2}{d_1} - \frac{ns^2}{d_2} = \left(\frac{1}{d_1} - \frac{1}{d_2}\right) ns^2$$

El problema que tenemos que resolver es:

$$\text{Min } A(d_1, d_2) = \left(\frac{1}{d_1} - \frac{1}{d_2}\right) ns^2$$

$$\text{s.a: } F(d_2) - F(d_1) = 1-\alpha$$

La función de Lagrange es:

$$L(d_1, d_2, d) = \left(\frac{1}{d_1} - \frac{1}{d_2}\right) ns^2 - d\left[F(d_2) - F(d_1) + \alpha - 1\right]$$

Las condiciones de primer orden para la existencia de extremo (en este caso mínimo) son:

$$\frac{\partial L}{\partial d_1} = -\frac{1}{d_1^2} \cdot ns^2 + d \cdot f(d_1) = 0 \Rightarrow d_1^2 f(d_1) = \frac{ns^2}{d}$$

$$\frac{\partial L}{\partial d_2} = \frac{1}{d_2^2} \cdot ns^2 - d \cdot f(d_2) = 0 \Rightarrow d_2^2 f(d_2) = \frac{ns^2}{d}$$

$$\frac{\partial L}{\partial d} = F(d_2) - F(d_1) + \alpha - 1 = 0 \Rightarrow F(d_2) - F(d_1) = 1-\alpha$$

siendo $f(d_i) = F'(d_i)$ la función de densidad de χ^2_{n-1} para el valor d_i $i=1,2$.

Lo que desprendemos de las ecuaciones anteriores es que $d_1^2 f(d_1) = d_2^2 \cdot f(d_2)$, esta ecuación es muy perra como ya hemos dicho antes y por eso tomamos el criterio que la probabilidad que deja a su izquierda d_1 sea igual a la probabilidad que deja a su derecha d_2 es decir:

$$P[X < d_1] = \frac{\alpha}{2} \qquad P[X > d_2] = 1 - P[X < d_2] = \frac{\alpha}{2} \Rightarrow$$

$\Rightarrow P[X < d_2] = 1 - \frac{\alpha}{2}$, con esto y con la tabla de χ^2_{n-1} podemos determinar d_1 y d_2. Nos preguntamos ahora, ¿en qué se diferencia entonces el intervalo de confianza para σ^2 si se conoce μ o no se conoce μ? En la precisión, tenemos más precisión si se conoce μ ya que el cálculo de d_1 y d_2 se obtiene de la tabla χ^2_n, sin embargo si no conocemos μ, d_1 y d_2 se obtienen en la tabla χ^2_{n-1} y a más grados de libertad mayor precisión.

Ejemplo: Sea $X \rightsquigarrow N(\mu, \sigma)$. Tomamos una m.a.s. de tamaño $n=7$ cuya varianza es $S^2 = 30$, siendo el nivel de significación $\alpha = 2\%$. Hallar un intervalo de confianza para la varianza poblacional

Sabemos que el intervalo de confianza es:

$$IC_{98\%}(\sigma^2) = \left(\frac{1}{d_2} nS^2 , \frac{1}{d_1} nS^2 \right)$$

$$P(X < d_1) = \frac{\alpha}{2} = 0'01$$

$$P(X < d_2) = 1 - P[X > d_2] = 1 - \frac{\alpha}{2} = 0'99$$

Como los grados de libertad son 7, si miramos en la tabla χ^2_6 obtenemos $d_1 = 0'872$ $d_2 = 16'8$. Por tanto:

$$IC_{98\%}(\sigma^2) = \left(\frac{7 \cdot 30}{16'8} , \frac{7 \cdot 30}{0'872} \right) = (12'5 , 240'8)$$

INTERVALO DE CONFIANZA PARA LA DIFERENCIA ENTRE LAS MEDIAS DE DOS POBLACIONES NORMALES CON VARIANZAS CONOCIDAS.

Sean las dos poblaciones normales

$X \rightsquigarrow N(\mu_1, \sigma_1)$, $Y \rightsquigarrow N(\mu_2, \sigma_2)$. En cada una de estas poblaciones se toman muestras de tamaño n y m respectivamente:

$X_1, X_2, \dots, X_n$ m.a.s. en $N(\mu_1, \sigma_1)$

$Y_1, Y_2, \dots Y_m$ m.a.s. en $N(\mu_2, \sigma_2)$

Entonces $\bar{X} \rightsquigarrow N\left(\mu_1, \frac{\sigma_1}{\sqrt{n}}\right)$, $\bar{Y} \rightsquigarrow N\left(\mu_2, \frac{\sigma_2}{\sqrt{n}}\right)$

Por lo tanto:

$$\bar{X} - \bar{Y} \rightsquigarrow N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right) \Rightarrow \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \rightsquigarrow N(0,1)$$

Podemos tomar como función pivotal

$$v(\mu_1 - \mu_2) = \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$ ya que depende tan

sólo de la muestra y del parámetro a estimar $\mu_1 - \mu_2$ ya que tanto σ_1 como σ_2 son conocidas.

Si el nivel de significación es α, tenemos que encontrar un intervalo (λ_1, λ_2) tal que:

$$P[\lambda_1 < v(\mu_1 - \mu_2) < \lambda_2] = 1 - \alpha$$

Vamos a pivotar la desigualdad $\lambda_1 < v(\mu_1 - \mu_2) < \lambda_2$ para acotar la diferencia $\mu_1 - \mu_2$.

$$\lambda_1 < \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} < \lambda_2 \Rightarrow \lambda_1 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < (\bar{X}-\bar{Y}) - (\mu_1 - \mu_2) < \lambda_2 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$$

$$-\lambda_2 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < (\mu_1 - \mu_2) - (\bar{X}-\bar{Y}) < -\lambda_1 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}.$$

$$(\bar{X}-\bar{Y}) - \lambda_2 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < (\bar{X}-\bar{Y}) - \lambda_1 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$$

Los intervalos en donde está $\mu_1-\mu_2$ son de la forma

$$I = \left((\bar{x}-\bar{y}) - d_2 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \ , \ (\bar{x}-\bar{y}) - d_1 \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right)$$

La amplitud de este intervalo es:

$$A(d_1, d_2) = (d_2 - d_1) \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$$ y tenemos que minimizar esta amplitud con la condición de que $P(d_1 < U(\mu_1-\mu_2) < d_2) = 1-\alpha$ lo que es equivalente a poner que $F(d_2) - F(d_1) = 1-\alpha$ siendo F la función de distribución de la normal $N(0,1)$.

El problema a resolver es:

$$\text{Min } A(d_1, d_2) = (d_2 - d_1) \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$$

$$\text{s.a: } \quad F(d_2) - F(d_1) = 1-\alpha$$

La función de Lagrange es:

$$L(d_1, d_2, \lambda) = (d_2 - d_1) \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} - \lambda \left[F(d_2) - F(d_1) + \alpha - 1 \right]$$

Las condiciones de primer orden para la existencia de extremo (mínimo en este caso) son:

$$\frac{\partial L}{\partial d_1} = -\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} + \lambda \cdot f(d_1) = 0$$

$$\frac{\partial L}{\partial d_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} - \lambda f(d_2) = 0$$

$f(d_1)$ es la función de densidad para $x = d_1$ y lo mismo ocurre con $f(d_2)$

$$\frac{\partial L}{\partial \lambda} = F(d_2) - F(d_1) + \alpha - 1 = 0$$

De estas ecuaciones se deduce que $f(d_1) = f(d_2)$.

$$\text{Si } f(d_1) = f(d_2) \Rightarrow \begin{cases} d_1 = d_2 & \text{(absurdo ya que } d_1 < d_2) \\ d_1 = -d_2 & \text{(ya que } N(0,1) \text{ es simétrica)} \end{cases}$$

Si $d_1 = -d_2 \Rightarrow F(d_2) - F(-d_2) + \alpha - 1 = 0 \Rightarrow$

$$F(d_2) - 1 + F(d_2) = 1-\alpha \Rightarrow 2F(d_2) = 2-\alpha \Rightarrow F(d_2) = 1 - \frac{\alpha}{2}$$

Luego λ_2 es el punto en donde $P(X < \lambda_2) = 1-\frac{\alpha}{2}$
y esto se representa por $z_{1-\frac{\alpha}{2}}$

$\lambda_2 = z_{1-\frac{\alpha}{2}} \Rightarrow \lambda_1 = -z_{1-\frac{\alpha}{2}}$. El intervalo es por tanto

$$IC_{1-\alpha}(\mu_1-\mu_2) = \left((\bar{x}-\bar{y}) - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\ ,\ (\bar{x}-\bar{y}) + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\right)$$

Ejemplo: Se consideran dos poblaciones normales $X \rightsquigarrow N(\mu_1, 2)$ $Y \rightsquigarrow N(\mu_2, 3)$. El nivel de significación es del 10%
En X tomamos una muestra de tamaño 400 y media $\bar{x} = 30$. En Y tomamos una muestra de tamaño 100 y media $\bar{y} = 29$. Hallar el intervalo de confianza para la diferencia de medias poblacionales $\mu_1-\mu_2$.

Según hemos visto

$$IC_{90\%}(\mu_1-\mu_2) = \left((\bar{x}-\bar{y}) - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\ ,\ (\bar{x}-\bar{y}) + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\right)$$

Se tiene en este caso:

$\bar{x}-\bar{y} = 30-29 = 1$; $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0'1}{2}} = z_{0'95} = 1'65$

$$\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}} = \sqrt{\frac{4}{400}+\frac{3}{100}} = \sqrt{\frac{4}{100}} = 0'2$$

$$IC_{90\%}(\mu_1-\mu_2) = (1 - 1'65 \cdot 0'2,\ 1 + 1'65 \cdot 0'2) = (0'67, 1'33)$$

Intervalo de confianza para la diferencia de medias entre dos poblaciones normales con la misma varianza desconocida.

Sean $X \rightsquigarrow N(\mu_1, \sigma)$, $Y \rightsquigarrow N(\mu_2, \sigma)$ dos poblaciones normales con la misma varianza σ^2 desconocida. Sea $X_1, X_2, \ldots X_n$ una m.a.s de X, $Y_1, Y_2 \ldots Y_m$ una m.a.s. de Y. Entonces por el lema de Fisher-Cochran:

$$\bar{x} \rightsquigarrow N\left(\mu_1, \frac{\sigma}{\sqrt{n}}\right) \qquad \bar{y} \rightsquigarrow N\left(\mu_2, \frac{\sigma}{\sqrt{m}}\right)$$

$$\frac{n S_x^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1} \qquad \frac{m S_y^2}{\sigma^2} \rightsquigarrow \chi^2_{m-1}$$

donde:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad S^2_x = \frac{\sum_{1}^{n} (x_i - \bar{x})^2}{n}$$

$$\bar{y} = \frac{\sum_{i=1}^{m} y_i}{m} \qquad S^2_y = \frac{\sum_{i=1}^{m} (y_i - \bar{y})^2}{m}$$

Entonces se verifica que:

$$\bar{x} - \bar{y} \rightsquigarrow N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}\right)$$

$$\bar{x} - \bar{y} \rightsquigarrow N\left(\mu_1 - \mu_2, \sigma \cdot \sqrt{\frac{n+m}{n \cdot m}}\right).$$ Por tanto

$$\frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{n+m}{n \cdot m}}} \rightsquigarrow N(0;1)$$

Al ser $\frac{n \cdot S^2_x}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$ $\quad \frac{m \cdot S^2_y}{\sigma^2} \rightsquigarrow \chi^2_{m-1} \Rightarrow$

$\Rightarrow \frac{n S^2_x + n S^2_y}{\sigma^2} \rightsquigarrow \chi^2_{m+n-2}$. Por tanto:

$$\frac{\dfrac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\cancel{\sigma} \sqrt{\frac{n+m}{n \cdot m}}}}{\sqrt{\dfrac{n S^2_x + n S^2_y}{(m+n-2) \cancel{\sigma^2}}}} \rightsquigarrow t_{m+n-2}$$

$$\frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{n S^2_x + m S^2_y}{(m+n-2)} \cdot \dfrac{n+m}{n \cdot m}}} \rightsquigarrow t_{m+n-2}$$

Si llamamos $S^* = \sqrt{\dfrac{n S^2_x + m S^2_y}{(m+n-2)} \cdot \dfrac{n+m}{n \cdot m}}$

Podemos tomar como función pivotal

$$v(\mu_1-\mu_2) = \frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{S^*}$$ ya que esta función

depende solo de la muestra y del parametro $\mu_1-\mu_2$ que queremos estimar, además su distribución χ^2_{n+m-2} no depende de ningún parámetro desconocido. Si el nivel de significación es α, pretendemos hallar un intervalo (λ_1, λ_2) tal que:

$$P[\lambda_1 < v(\mu_1-\mu_2) < \lambda_2] = 1-\alpha$$

Vamos a pivotar la expresión $\lambda_1 < v(\mu_1-\mu_2) < \lambda_2$ para acotar $\mu_1-\mu_2$:

$$\lambda_1 < \frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{S^*} < \lambda_2$$

$$\lambda_1 S^* < (\bar{x}-\bar{y})-(\mu_1-\mu_2) < \lambda_2 S^*$$

$$-\lambda_2 S^* < (\mu_1-\mu_2)-(\bar{x}-\bar{y}) < -\lambda_1 S^*$$

$$(\bar{x}-\bar{y})-\lambda_2 S^* < \mu_1-\mu_2 < (\bar{x}-\bar{y})-\lambda_1 S^*$$

El intervalo que queremos hallar tiene la forma

$(\bar{x}-\bar{y}-\lambda_2 S^*,\ \bar{x}-\bar{y}-\lambda_1 S^*)$, su amplitud es:

$$A(\lambda_1,\lambda_2) = (\lambda_2-\lambda_1)\cdot S^*$$

La condición $P(\lambda_1 < v(\mu_1-\mu_2) < \lambda_2) = 1-\alpha$ la podemos poner de forma equivalente como:

$F(\lambda_2) - F(\lambda_1) = 1-\alpha$ siendo F la función de distribución de t_{m+n-2}.

El problema que tenemos que resolver es:

$$\text{Min } A(\lambda_1,\lambda_2) = (\lambda_2-\lambda_1)\cdot S^*$$
$$\text{S.a: } F(\lambda_2)-F(\lambda_1) = 1-\alpha$$

La función de Lagrange es:

$$L(\lambda_1,\lambda_2,\lambda) = (\lambda_2-\lambda_1)S^* - \lambda[F(\lambda_2)-F(\lambda_1)+\alpha-1]$$

Las condiciones de primer orden para la existencia de extremo (mínimo) son:

$$\left.\begin{array}{l}\dfrac{\partial L}{\partial d_1} = -S^* + d \cdot f(d_1) = 0 \\[2ex] \dfrac{\partial L}{\partial d_2} = S^* - d\, f(d_2) = 0\end{array}\right\} \text{ siendo } f \text{ la función de densidad de } t_{m+n-2}$$

$$\frac{\partial L}{\partial d} = F(d_2) - F(d_1) + \alpha - 1 = 0$$

De las ecuaciones anteriores se deduce que $f(d_1) = f(d_2)$ y por tanto:

$$f(d_1) = f(d_2) \Rightarrow \begin{cases} d_1 = d_2 \text{ (absurdo ya que } d_1 < d_2) \\ d_1 = -d_2 \text{ (por la simetría de } t_{m+n-2}) \end{cases}$$

$$F(d_2) - F(d_1) = 1-\alpha \Rightarrow F(d_2) - F(-d_2) = 1-\alpha \Rightarrow$$

$$\Rightarrow F(d_2) - (1 - F(d_2)) = 1-\alpha \Rightarrow 2F(d_2) = 2-\alpha \Rightarrow$$

$\Rightarrow F(d_2) = 1 - \frac{\alpha}{2} \Rightarrow P[X \leq d_2) = 1 - \frac{\alpha}{2}$. Representamos por $d_2 = t_{m+n-2,1-\frac{\alpha}{2}}$ al punto de la distribución t-Student en donde $P(X < d_2) = 1 - \frac{\alpha}{2}$

Por tanto $d_1 = -t_{m+n-2,1-\frac{\alpha}{2}}$ y el intervalo de confianza pedido es:

$$IC_{(1-\alpha)}(\mu_1 - \mu_2) = \left((\bar{x}-\bar{y}) - t_{m+n-2,1-\frac{\alpha}{2}} \cdot S^*, (\bar{x}-\bar{y}) + t_{m+n-2,1-\frac{\alpha}{2}} \cdot S^*\right)$$

Veamos un ejemplo.

Ejemplo:

En las poblaciones $X \rightsquigarrow N(\mu_1, \sigma)$, $Y \rightsquigarrow N(\mu_2, \sigma)$ se toman muestras de tamaño 4 y 5 respectivamente. Si el nivel de significación es del 5% y las medias y varianzas muestrales son $\bar{X} = 12$, $\bar{Y} = 11$, $S_X^2 = 7$, $S_Y^2 = 6$, hallar un intervalo de confianza para la diferencia de medias $\mu_1 - \mu_2$.

$$t_{m+n-2,\,1-\frac{\alpha}{2}} = t_{4+5-2,\,1-\frac{0'05}{2}} = t_{7,\,0'975} = 2'365$$

$$S^* = \sqrt{\frac{nS_x^2 + m\cdot S_y^2}{n+m-2} \cdot \frac{n+m}{n\cdot m}} = \sqrt{\frac{4\cdot 7 + 5\cdot 6}{7} \cdot \frac{9}{20}} = 3'72$$

$$\bar{x} - \bar{y} = 12 - 11 = 1.$$

$$IC_{95\%}(\mu_1 - \mu_2) = (1 - 2'365 \cdot 3'72,\ 1 + 2'365\cdot 3'72) = (-7'8, 9'8)$$

En consecuencia con una confianza del 95% se acepta la hipótesis de que los valores que toma la diferencia de medias poblacionales $\mu_1 - \mu_2$ se encuentran en el intervalo $(-7'8, 9'8)$ o lo que es lo mismo, por ejemplo no se puede rechazar, con una confianza del 95%, que las medias de las dos poblaciones sean iguales $\mu_1 = \mu_2$ ya que $\mu_1 - \mu_2 = 0 \in (-7'8, 9'8)$.

Intervalo de confianza para la diferencia de medias entre dos poblaciones normales con varianzas desconocidas y distintas.

Sean las poblaciones $X \rightsquigarrow N(\mu_1, \sigma_1)$, $Y \rightsquigarrow N(\mu_2, \sigma_2)$ cuyas varianzas se desconocen.

La relación entre la varianza muestral y la poblacional viene dada por las expresiones:

$$S_x^2 = \frac{n-1}{n} \cdot \sigma_1^2 \qquad S_y^2 = \frac{m-1}{m} \sigma_2^2 .$$ Entonces

$$\sigma_1^2 = \frac{n}{n-1} S_x^2 \qquad \sigma_2^2 = \frac{m}{m-1} \cdot S_y^2 .$$ Entonces podemos considerar, si las muestras son grandes, que:

$$X \rightsquigarrow N\left(\mu_1, \sqrt{\frac{n}{n-1}} \cdot S_x\right) \qquad Y \rightsquigarrow N\left(\mu_2, \sqrt{\frac{m}{m-1}} \cdot S_y\right)$$

y queremos hallar el intervalo de confianza para la diferencia de la media poblacional $\mu_1 - \mu_2$ con varianza conocida. Este intervalo es:

$$IC_{1-\alpha}(\mu_1-\mu_2)=\left((\bar{x}-\bar{y})-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\frac{n s_x^2}{n-1}}{n}+\frac{\frac{m s_y^2}{m-1}}{m}}\,,\,(\bar{x}-\bar{y})+z_{1-\frac{\alpha}{2}}\sqrt{\frac{\frac{n s_x^2}{n-1}}{n}+\frac{\frac{m s_y^2}{m-1}}{m}}\right)$$

$$IC_{1-\alpha}(\mu_1-\mu_2)=\left((\bar{x}-\bar{y})-z_{1-\frac{\alpha}{2}}\sqrt{\frac{s_x^2}{n-1}+\frac{s_y^2}{m-1}}\,,\,(\bar{x}-\bar{y})+z_{1-\frac{\alpha}{2}}\sqrt{\frac{s_x^2}{n-1}+\frac{s_y^2}{m-1}}\right)$$

Intervalo de confianza para el cociente de las varianzas de dos poblaciones normales con medias conocidas

Sean las dos poblaciones normales $X \rightsquigarrow N(\mu_1, \sigma_1)$ $Y \rightsquigarrow N(\mu_2, \sigma_2)$ donde se conocen μ_1 y μ_2.
En X tomamos una m.a.s. de tamaño n, $x_1, x_2, \dots x_n$. En Y tomamos una m.a.s. de tamaño m, $Y_1, Y_2, \dots Y_m$.
Al ser $x_i \rightsquigarrow N(\mu_1, \sigma_1)$ $\forall i = 1, 2, \dots n \Rightarrow$

$$\Rightarrow \frac{x_i-\mu_1}{\sigma_1} \rightsquigarrow N(0,1) \Rightarrow \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma_1}\right)^2 \rightsquigarrow \chi^2_n \Rightarrow$$

$$\Rightarrow \frac{1}{\sigma_1^2}\cdot\sum_{i=1}^{n}(x_i-\mu_1)^2 \rightsquigarrow \chi^2_n$$

De igual manera se tiene que $\frac{1}{\sigma_2^2}\sum_{j=1}^{m}(y_j-\mu_2)^2 \rightsquigarrow \chi^2_m$

Entonces:

$$\frac{\frac{1}{n\sigma_1^2}\sum_{i=1}^{n}(x_i-\mu_1)^2}{\frac{1}{m\sigma_2^2}\sum_{j=1}^{m}(y_j-\mu_2)^2} = \frac{m\sigma_2^2}{n\sigma_1^2}\cdot\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{\sum_{j=1}^{m}(y_j-\mu_2)^2} \rightsquigarrow F_{n,m}$$

Podemos tomar como función pivotal si μ_1 y μ_2 son conocidas

$$v\left(\frac{\sigma_2^2}{\sigma_1^2}\right) = \frac{m}{n}\cdot\frac{\sigma_2^2}{\sigma_1^2}\cdot\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{\sum_{j=1}^{m}(y_j-\mu_2)^2}$$

ya que $v\left(\frac{\sigma_2^2}{\sigma_1^2}\right)$ depende tan sólo de $\frac{\sigma_2^2}{\sigma_1^2}$ y de la muestra y su distribución $F_{m,n}$ no contiene parámetros desconocidos. Queremos hallar un intervalo (d_1, d_2) tal que

$P\left[d_1 < v\left(\frac{\sigma_2^2}{\sigma_1^2}\right) < d_2\right] = 1-\alpha$ si el nivel de significación es α. Pero esto es equivalente a poner que $F(d_2) - F(d_1) = 1-\alpha$ siendo F la función de distribución de $F_{n,m}$

Pivotamos la desigualdad $d_1 < v\left(\frac{\sigma_2^2}{\sigma_1^2}\right) < d_2$ para acotar $\frac{\sigma_2^2}{\sigma_1^2}$

$$d_1 < \frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{m}{n} \cdot \frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{\sum_{j=1}^{m}(y_j-\mu_2)^2} < d_2$$

$$\frac{n\, d_1}{m} \cdot \frac{\sum_{j=1}^{m}(y_j-\mu_2)^2}{\sum_{i=1}^{n}(x_i-\mu_1)^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{n\, d_2}{m} \cdot \frac{\sum_{j=1}^{m}(y_j-\mu_2)^2}{\sum_{i=1}^{n}(x_i-\mu_1)^2}$$

El intervalo que queremos hallar tiene la forma:

$$\left(\frac{n}{m} \cdot \frac{\sum_{j=1}^{m}(y_j-\mu_2)^2}{\sum_{i=1}^{n}(x_i-\mu_1)^2}\, d_1 \; ; \; \frac{n}{m} \cdot \frac{\sum_{j=1}^{m}(y_j-\mu_2)^2}{\sum_{i=1}^{n}(x_i-\mu_1)^2} \cdot d_2\right)$$

La amplitud de este intervalo es:

$$A(d_1, d_2) = \frac{n}{m} \cdot \frac{\sum_{j=1}^{m}(y_j-\mu_2)^2}{\sum_{i=1}^{n}(x_i-\mu_1)^2} \cdot (d_2 - d_1)$$

El problema que tenemos que resolver es:

Minimizar
$$A(d_1, d_2) = \frac{n}{m} \cdot \frac{\sum_{j=1}^{m}(y_j-\mu_2)^2}{\sum_{i=1}^{n}(x_i-\mu_1)^2} (d_2 - d_1)$$

s.a: $F(d_2) - F(d_1) = 1-\alpha$

La función de Lagrange es:

$$L(d_1, d_2, \lambda) = \frac{n}{m} \cdot \frac{\sum_{j=1}^{m} (y_j - \mu_2)^2}{\sum_{i=1}^{n} (x_i - \mu_1)^2} (d_2 - d_1) - \lambda \left(F(d_2) - F(d_1) + \alpha - 1 \right)$$

Las condiciones de primer orden para la existencia de mínimo son:

$$\frac{\partial L}{\partial d_1} = -\frac{n}{m} \cdot \frac{\sum_{j=1}^{m} (y_j - \mu_2)^2}{\sum_{i=1}^{n} (x_i - \mu_1)^2} + \lambda \cdot f(d_1) = 0$$

$$\frac{\partial L}{\partial d_2} = \frac{n}{m} \cdot \frac{\sum_{j=1}^{m} (y_j - \mu_2)^2}{\sum_{i=1}^{n} (x_i - \mu_1)^2} - \lambda \cdot f(d_2) = 0$$

$$\frac{\partial L}{\partial \lambda} = F(d_2) - F(d_1) + \alpha - 1 = 0$$

donde f es la función de densidad de la distribución F de Snedecor.

Se obtiene que

$$f(d_1) = f(d_2) = \frac{n}{m\lambda} \cdot \frac{\sum_{j=1}^{m} (y_j - \mu_2)^2}{\sum_{i=1}^{n} (x_i - \mu_1)^2}$$

. Pero tenemos el inconveniente de que la distribución F no es simétrica. Entonces no podemos desprender ninguna consecuencia del hecho de que $f(d_1) = f(d_2)$. Imponemos la condición de siempre:

$F(d_1) = P[X < d_1] = \frac{\alpha}{2}$ (probabilidad que deja a la izquierda d_1)

$P(X > d_2) = \frac{\alpha}{2} \Rightarrow F(d_2) = P[X < d_2] = 1 - \frac{\alpha}{2}$

De las ecuaciones

$P[X < \lambda_1] = \frac{\alpha}{2} = F(\lambda_1)$

$P[X > \lambda_2] = 1 - P[X < \lambda_2) = 1 - F(\lambda_2) = \frac{\alpha}{2}$

$F(\lambda_1) = \frac{\alpha}{2}$

$F(\lambda_2) = 1 - \frac{\alpha}{2}$ obtenemos, de la tabla F-Snedecor, los valores de λ_1 y λ_2.

Intervalo de confianza para el cociente de las varianzas de dos poblaciones normales con medias desconocidas.

Sean dos poblaciones $X \rightsquigarrow N(\mu_1, \sigma_1)$ $Y \rightsquigarrow N(\mu_2, \sigma_2)$ con μ_1 y μ_2 desconocidos, en las que tomamos sendas m.a.s. de tamaños n y m respectivamente:

$X_1, X_2, \ldots, X_n$ e $Y_1, Y_2, \ldots Y_m$

Entonces por el lema de Fisher-Cochran:

$\frac{n\, S^2_x}{\sigma_1^2} \rightsquigarrow \chi^2_{n-1}$ $\frac{m \cdot S^2_y}{\sigma_2^2} \rightsquigarrow \chi^2_{m-1}$ siendo:

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}, \quad S^2_x = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}, \quad \bar{Y} = \frac{\sum_{j=1}^{m} y_j}{m}, \quad S^2_y = \frac{\sum_{j=1}^{m} (y_j - \bar{y})^2}{m}$$

Por lo tanto:

$$\frac{\frac{n \cdot S^2_x}{n\, \sigma_1^2}}{\frac{m \cdot S^2_y}{m\, \sigma_2^2}} = \frac{S^2_x}{S^2_y} \cdot \frac{\sigma_2^2}{\sigma_1^2} \rightsquigarrow F_{n-1, m-1}$$

Podemos tomar como función pivotal $U\left(\frac{\sigma_2^2}{\sigma_1^2}\right) = \frac{S^2_x}{S^2_y} \cdot \frac{\sigma_2^2}{\sigma_1^2}$ ya que esta depende de la muestra y del parámetro que queremos estimar $\frac{\sigma_2^2}{\sigma_1^2}$, además su distribución, que es $F_{n,m}$ no depende de parámetros desconocidos.

Queremos hallar un intervalo (λ_1, λ_2) en el que se verifique que:

$P\left[\lambda_1 < \vartheta\left(\frac{\sigma_2^2}{\sigma_1^2}\right) < \lambda_2 \right] = 1-\alpha$, siendo α el nivel de significación. Esto es equivalente a poner $F(\lambda_2) - F(\lambda_1) = 1-\alpha$ donde F es la función de distribución de $F_{n-1,m-1}$

Pivotamos la desigualdad:

$$\lambda_1 < \vartheta\left(\frac{\sigma_2^2}{\sigma_1^2}\right) < \lambda_2 \text{ para acotar } \frac{\sigma_2^2}{\sigma_1^2}$$

$$\lambda_1 < \frac{S^2_x}{S^2_y} \cdot \frac{\sigma_2^2}{\sigma_1^2} < \lambda_2 \Rightarrow \lambda_1 \cdot \frac{S^2_y}{S^2_x} < \frac{\sigma_2^2}{\sigma_1^2} < \lambda_2 \frac{S^2_y}{S^2_x}$$

El intervalo que buscamos es de la forma:

$$\left(\lambda_1 \cdot \frac{S^2_y}{S^2_x}, \lambda_2 \cdot \frac{S^2_y}{S^2_x} \right).$$

La amplitud de este intervalo es:

$$A(\lambda_1, \lambda_2) = \frac{S^2_y}{S^2_x}(\lambda_2 - \lambda_1).$$

El problema que tenemos que resolver es:

Minimizar $A(\lambda_1, \lambda_2) = \dfrac{S^2_y}{S^2_x}(\lambda_2 - \lambda_1)$

s.a: $F(\lambda_2) - F(\lambda_1) = 1-\alpha$

La función de Lagrange es:

$$L(\lambda_1, \lambda_2, \lambda) = \frac{S^2_y}{S^2_x}(\lambda_2 - \lambda_1) - \lambda\left[F(\lambda_2) - F(\lambda_1) + \alpha - 1\right]$$

Las condiciones de primer orden para la existencia de mínimo son:

$$\frac{\partial L}{\partial \lambda_1} = -\frac{S^2_y}{S^2_x} + \lambda \cdot f(\lambda_1) = 0$$

$$\frac{\partial L}{\partial \lambda_2} = \frac{S^2_y}{S^2_x} - \lambda \cdot f(\lambda_2) = 0$$

$$\frac{\partial L}{\partial \lambda} = F(\lambda_2) - F(\lambda_1) + \alpha - 1 = 0$$

De estas ecuaciones se deduce que $f(\lambda_1) = f(\lambda_2)$ y como $F_{n,m}$ no es simétrica no podemos sacar ninguna conclusión de este hecho. Aplicamos el criterio de siempre: la probabilidad que deja a la izquierda λ_1 = probabilidad que deja a la derecha $\lambda_2 = \frac{\alpha}{2}$

$$P[X < \lambda_1] = \frac{\alpha}{2} \Rightarrow F(\lambda_1) = \frac{\alpha}{2}$$

$$P[X > \lambda_2] = \frac{\alpha}{2} \Rightarrow 1 - P[X < \lambda_2] = \frac{\alpha}{2} \Rightarrow P[X < \lambda_2] = 1 - \frac{\alpha}{2}$$

$$\Rightarrow F(\lambda_2) = 1 - \frac{\alpha}{2}$$

De las expresiones:

$$F(\lambda_1) = \frac{\alpha}{2}$$

$$F(\lambda_2) = 1 - \frac{\alpha}{2}$$

y de la tabla de la F de Snedecor $F_{n-1, m-1}$ obtenemos los valores de λ_1 y λ_2.

Observemos que el cálculo del intervalo de confianza para el cociente de las varianzas de dos poblaciones normales es muy parecido tanto si se conocen las medias como si no se conocen. Si conocemos las medias obtenemos λ_1 y λ_2 de la tabla $F_{n,m}$. y si no las conocemos obtenemos λ_1 y λ_2 de la tabla $F_{n-1, m-1}$. Como a mayor grado de libertad hay más precisión, el intervalo que se obtiene conociendo las medias es más preciso que el que se obtiene sin conocerlas.

INTERVALO DE CONFIANZA CON MUESTRAS GRANDES

Si X es una población cuya distribución de probabilidad depende de un parámetro θ, si tomamos mm. aa. ss. de tamaño grande, podemos aproximar la distribución del estimador de máxima verosimilitud de θ a una normal:

$$\hat{\theta}_{EMV} \rightsquigarrow N\left(\theta, \sqrt{\text{cota } FCR(\theta)}\right) \Rightarrow$$

$$\Rightarrow \frac{\hat{\theta}_{EMV} - \theta}{\sqrt{\text{cota } FCR(\theta)}} \rightsquigarrow N(0,1)$$

El intervalo de confianza de θ lo determinamos por el método pivotal.