INFERENCIA ESTADÍSTICA

INTERVALOS DE CONFIANZA

IC I

Generalidades

Cuando se estima un parámetro en forma puntual, generalmente, pero no siempre, el resultado de la estimación es simplemente un número, no teniéndose ninguna idea de la fiabilidad de la estimación.
La estimación por intervalos de confianza soluciona este problema ya que la respuesta que da este método es un intervalo dentro del cual se "decreta" que está el parámetro (es decir una aproximación del mismo), junto con una medida de la confiabilidad del "decreto".
A continuación se analizarán diversos casos en los que se usa la estimación de parámetros por el método de los intervalos de confianza.

IC II

Intervalo de confianza para el valor medio de una variable aleatoria correspondiente a una distribución cualquiera de probabilidad en el caso de disponerse de una muestra grande

IC II.1

a. El hecho de analizar una muestra de tamaño *n* de una población homogénea equivale a efectuar *n* repeticiones independientes de un mismo experimento aleatorio, suponiéndose que existe permanencia estadística a lo largo de todos ellos.
Supóngase que a los resultados a obtener en las repeticiones 1, ..., *n* se les asocien respectivamente las variables aleatorias $X_1, \ldots, X_n$.
Evidentemente, todas estas variables serán independientes y tendrán un mismo valor medio *m* y una misma desviación típica σ.
Supóngase que *m* sea desconocido y que, por el momento σ sea conocido.
En el caso de que *n* sea grande, por el teorema de Lindeberg se tiene (ver **[5]** de BNP VII) que:

$$\overline{X} = \frac{X_1 + \ldots + X_n}{n} \qquad \textbf{[1]}$$

es aproximadamente normal (m , $\sigma/\sqrt{n}$).
Por lo tanto:

$$\frac{\overline{X} - m}{\sigma/\sqrt{n}} = \frac{\sqrt{n}}{\sigma}(\overline{X} - m) \qquad \textbf{[2]}$$

será aproximadamente normal (0 , 1), y en lo que sigue se la considerará como tal.

b. Usando la tabla de la F. de D. normal (0 , 1) búsquese un número α tal que:

$$P(-\alpha < \frac{\sqrt{n}}{\sigma}(\overline{X} - m) < \alpha) = 0{,}95 \qquad \textbf{[3]}$$

y entonces resultara que:

$$P\left(-\frac{\alpha\sigma}{\sqrt{n}} \le \overline{X} - m \le \frac{\alpha\sigma}{\sqrt{n}}\right) = 0{,}95$$

Es decir que:

$$P\left(\overline{X} - \alpha\frac{\sigma}{\sqrt{n}} \le m \le \overline{X} + \alpha\frac{\sigma}{\sqrt{n}}\right) = 0{,}95 \qquad \textbf{[4]}$$

En palabras, esta expresión dice lo siguiente:

<u>Existe un probabilidad igual a 0,95 de que el valor que asuma $\overline{X}$ al efectuarse *n* repeticiones del experimento sea tal que el valor de m esté dentro del intervalo</u>:

$$\left(\overline{X} - \alpha\frac{\sigma}{\sqrt{n}} \qquad \overline{X} + \alpha\frac{\sigma}{\sqrt{n}}\right)$$

Entonces se define que:
Si se obtiene un valor observado $\overline{x}$, una estimación de *m* consiste en el intervalo de confianza:

$$\left(\overline{x} - \alpha\frac{\sigma}{\sqrt{n}} \qquad \overline{x} + \alpha\frac{\sigma}{\sqrt{n}}\right) \qquad \textbf{[5]}$$

con un grado de confiabilidad igual a 0,95 (ó 95%).

c. Observación 1ª (muy importante):
Notar que <u>no</u> puede decirse que exista una probabilidad 0,95 de que el valor de *m* esté dentro del intervalo indicado en **[5]**. Los límites de <u>este</u> intervalo son números fijos y *m* también es un número fijo y, entonces, *m* está o no está dentro de dicho intervalo, no pudiendo hablarse de una probabilidad.
La interpretación de un intervalo de confianza es la siguiente:
<u>La probabilidad de que la afirmación " El valor de *m* está dentro del intervalo de confianza considerado" sea cierta es igual a 0,95</u>.
Supóngase que se efectúen 10^6 tandas de *n* realizaciones del experimento cada una. Cada tanda dará un valor observado $\overline{x}$ propio, obteniéndose así 10^6 valores observados $\overline{x}_1,...,\overline{x}_{10^6}$ que, seguramente, serán todos distintos entre sí. Como cada uno de esos valores observados determina un intervalo de confianza, se obtendrán entonces 10^6 intervalos de confianza todos distintos entre sí. El valor de *m* estará, aproximadamente, dentro del 95% de dichos intervalos.

d. Observación 2ª:
La longitud del intervalo de confianza indicado en **[5]** es:

$$l = \left(\overline{x} + \alpha \frac{\sigma}{\sqrt{n}} \right) - \left(\overline{x} - \alpha \frac{\sigma}{\sqrt{n}} \right) = \frac{2\,\sigma\,\alpha}{\sqrt{n}} \qquad \textbf{[6]}$$

Al aumentar *n* esta longitud disminuye, aumentando así la precisión de la estimación.

e. Observación 3ª:
En el problema se supuso que se conocía la desviación típica σ_X de la F. de D. de la variable X. Es bastante raro que en la práctica ocurra que se conozca la desviación típica de una variable y no se conozca su valor medio. Lo usual es que no se conozca ninguno de los dos parámetros.
Si se quiere conservar la pureza de lo visto hasta ahora, se tendrá que el método de estimación por intervalos de confianza no pasa de la categoría de especulación teórica.
A fin de dar importancia práctica al método, para poder estimar por intervalos al valor medio, el único camino que queda abierto es estimar previamente a σ. En vista a lo indicado en **[7]** de IE IV parece lógico tomar:

$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2} \qquad \textbf{[7]}$$

y usando esta estimación en lugar del verdadero σ, se podrá salvar el método, si bien a costa de algo de su "pureza". Por lo tanto el intervalo de confianza resultante habrá que tomarlo ligeramente a beneficio de inventario. Esto no es precisamente una situación ideal, pero desgraciadamente no hay alternativa.

IC II.2

Aplicación

a. En 60000 tiros de dado salieron 5400 ases. Se pide hallar un intervalo de confianza para la probabilidad de sacar un as en un único tiro de dado. Se pide un grado de confiabilidad igual a 0,95.

b. Sea X la variable aleatoria asociada a un tiro de dado. Se dirá que X = 1 si sale un as y que X = 0 si no sale un as.
Llamando *p* a la probabilidad (desconocida) de sacar un as en un tiro de dado, se tiene que:

$$m_X = p$$

Con lo que resulta que el problema de estimar la probabilidad se reduce al problema de estimar el valor medio m_X .

c. Según arriba indicado, se efectuarán 60000 pruebas, a las cuales se les asocian las variables $X_1, \ldots, X_{60000}$. Como la cantidad de pruebas es grande, se tiene que la variable:

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

tendrá una distribución aproximadamente normal.
Cabe entonces usar el procedimiento indicado en el párrafo IC II.1.

d. Dado que el grado de confiabilidad es de 0,95, se buscará un valor α tal que para una cierta variable normal (0 , 1), Y, se tenga que:

$$P(-\alpha \leq Y \leq \alpha) = 0{,}95$$

Esto implica que:

$P(Y < -\alpha \cup Y > \alpha) = 1 - 0{,}95 = 0{,}05$

$P(Y < -\alpha) + P(Y > \alpha) = 0{,}05$

Como Y tiene una F. de D. normal (0 , 1), se tiene que $P(Y < -\alpha) = P(Y > \alpha)$, y por lo tanto:

$2\,P(Y > \alpha) = 0{,}05$
$P(Y > \alpha) = 0{,}025$
$P(Y < \alpha) = 1 - 0{,}025 = 0{,}975$

Yendo a la tabla de la F. de D. normal (0 , 1) se obtiene entonces que:

$$\alpha = 1{,}96$$

(Notar que Y hizo las veces de variable aleatoria "muda").

e. Como, según los datos del problema, de las 60000 variables X, 5400 asumieron el valor 1, y los restantes asumieron el valor 0, es decir que 5400 valores x_i son iguales a 1 y los restantes son iguales a 0, se tienen los siguientes valores observados:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{60000}5400 = 0{,}09$$

$$\hat{\sigma}_X = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} = \sqrt{\frac{1}{60000-1}\left[5400(1-0{,}09)^2 + 54600(0-0{,}09)^2\right]} = 0{,}286$$

f. Según indicado en **[5]** un intervalo de confianza del valor m_X con un grado de confiabilidad 0,95 será:

$$\left(\overline{x} - \alpha\frac{\sigma}{\sqrt{n}} \qquad \overline{x} + \alpha\frac{\sigma}{\sqrt{n}}\right)$$

En el problema en estudio, se conocen los valores de $\overline{x}$, α y n (respectivamente 0,09; 1,963 y 60000), pero no se conoce el verdadero valor de σ_X.
Si en vez de dicho valor se usa el valor observado $\hat{\sigma}_X = 0{,}286$ se obtendrá un intervalo de confianza:

$$\left(0{,}09 - 1{,}96\frac{0{,}286}{\sqrt{60000}} \qquad 0{,}09 + 1{,}96\frac{0{,}286}{\sqrt{60000}}\right)$$

Efectuando las operaciones indicadas, se llega a que para la estimación de p se obtiene un intervalo de confianza:

$$(0{,}08771 \quad ; \quad 0{,}09229)$$

con un grado de confiabilidad que sería de 0,95 si el valor estimado de la desviación típica coincidiera con el valor verdadero. El hecho de no poder garantizar que esto ocurra disminuye algo de confiabilidad del resultado pero, desgraciadamente, no se sabe en que medida.

IC III

Intervalo de confianza para el valor medio de una variable aleatoria correspondiente a una distribución normal

IC III.1

a. Teóricamente la muestra puede ser grande o chica, es decir que las repeticiones del experimento pueden ser muchas o pocas, pero siempre las variables aleatorias correspondientes deben ser normales, independientes y con un mismo valor medio m y una misma varianza σ^2.

b. Sean X_1, ..., X_n las variables aleatorias correspondientes a las repeticiones 1ª, ..., nª del experimento.
Sean:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad \textbf{[1]}$$

Entonces, por las condiciones impuestas en **a.** y por lo indicado en **[5]** de JtF IV se tiene que:

$T = \sqrt{n-1}\,\dfrac{\overline{X} - m}{S}$ tiene una distribución t de Student con $n - 1$ grados de libertad. **[2]**

c. Supóngase que se desee hallar un intervalo de confianza con una confiabilidad igual a 0,95 del valor medio mencionado en **a.** mediante la ejecución de n experimentos independientes.

Usando la tabla de la distribución t para $\alpha = \dfrac{1-0{,}95}{2} = 0{,}025$ y $n - 1$, hállese el valor $t_{0{,}025\,;\,n-1}$ correspondiente.

Se tiene entonces que:

$$P(T > t_{0{,}025 \quad n-1}) = 0{,}025 \Rightarrow P(T \leq t_{0{,}025 \quad n-1}) = 0{,}975 \qquad \textbf{[3]}$$

y,

$$P(\mathrm{T} < -\mathrm{t}_{0,025 \quad n-1}) = P(\mathrm{T} > \mathrm{t}_{0,025 \quad n-1}) = 0{,}025 \qquad \textbf{[4]}$$

Ver **[4]** de JtF IV

y entonces por **[3]** y **[4]** resulta que:

$$P(-\mathrm{t}_{0,025 \quad n-1} < \mathrm{T} \leq \mathrm{t}_{0,025 \quad n-1}) = 0{,}975 - 0{,}025 = 0{,}95$$

y por **[2]**:

$$P(-\mathrm{t}_{0,025 \quad n-1} < \sqrt{n-1}\,\frac{\overline{X}-m}{S} < \mathrm{t}_{0,025 \quad n-1}) = 0{,}95$$

De donde resulta que:

$$P\left(\overline{\mathrm{X}} - \frac{\mathrm{S.}\ \mathrm{t}_{0,025 \quad n-1}}{\sqrt{n-1}} < m < \overline{\mathrm{X}} + \frac{\mathrm{S.}\ \mathrm{t}_{0,025 \quad n-1}}{\sqrt{n-1}}\right) = 0{,}95$$

Es decir que existe una probabilidad igual a 0,95 de que el valor que asuma $\overline{\mathrm{X}}$ al efectuarse *n* repeticiones del experimento esté dentro del intervalo:

$$\left(\overline{\mathrm{X}} - \frac{\mathrm{S.}\ \mathrm{t}_{0,025 \quad n-1}}{\sqrt{n-1}} \qquad \overline{\mathrm{X}} + \frac{\mathrm{S.}\ \mathrm{t}_{0,025 \quad n-1}}{\sqrt{n-1}}\right) \qquad \textbf{[5]}$$

Si se obtienen valores observados $\overline{x}$ y s se tiene que <u>un</u> intervalo de confianza (entre infinitos posibles) para la estimación de *m* con una confiabilidad e 0,95 es:

$$\left(\overline{\mathrm{x}} - \frac{\mathrm{s.}\ \mathrm{t}_{0,025 \quad n-1}}{\sqrt{n-1}} \qquad \overline{\mathrm{x}} + \frac{\mathrm{s.}\ \mathrm{t}_{0,025 \quad n-1}}{\sqrt{n-1}}\right) \qquad \textbf{[6]}$$

d. Observación 1ª.
Análoga a la observación hecha en **c.** de IC II.1.

e. Observación 2ª.
La longitud del intervalo de confianza indicado en **[6]** es:

$$l = 2.\frac{\mathrm{s}\cdot \mathrm{t}_{0,05 \quad n-1}}{\sqrt{n-1}}$$

Como para un cierto valor de α el valor de $\mathrm{t}_{\alpha,n}$ disminuye al aumentar *n* (ver tabla de figura JtF II.b) y por otra parte $\sqrt{n-1}$ aumenta, se tiene que al aumentar *n* la longitud disminuye, es decir que la precisión aumenta.

f. Muy a menudo el método de estimación de un valor medio recién desarrollado se lo supone relacionado únicamente con el caso de "muestras chicas".
Eso es enteramente falso. Este método sirve tanto para el caso de muestras chicas como de muestras grandes, estando su uso limitado únicamente por el alcance de la tabulación de la distribución t de que se disponga.
Pero por otra parte, en el caso de muestras chicas éste es el único método disponible, <u>a condición de que se cumplan los requisitos indicados en **a.**</u>

g. Y ya que se habla de muestras chicas, vale tener en cuenta lo expresado por M. J. Moroney en su libro “Hechos y Estadísticas”: “En la vida diaria muy a menudo tomamos decisiones en base a la evidencia obtenida a partir de muestras pequeñas.”
“La práctica de hacer juicios ligeros sobre una primera evidencia es un hábito insidioso que a menudo trasladamos a cuestiones donde sí es importante que estemos equivocados o no. La evidencia de las muestras pequeñas puede ser muy traicionera.”

IC III.2

Aplicación

Sea una fábrica de alambre para resistencias eléctricas. Se sabe que la resistencia por metro sigue una ley normal (m , σ), siendo m y σ desconocidas.
Se hicieron 10 mediciones sobre otras tantas muestras de 1 m tomadas al azar obteniéndose valores $\bar{x} =$ 10,48 y s = 1,29.
Se pide hallar un intervalo de confianza para el valor medio m con una confiabilidad igual a 0,90.
Para $\alpha = \frac{1-0{,}9}{2} = 0{,}05$ y $n - 1 = 9$, de la tabla de la función t se obtiene $t_{0,05\,;\,9} = 1{,}833$, y entonces según **[6]** se obtiene el intervalo de confianza:

$$\left(10{,}48 - \frac{1{,}29 \cdot 1{,}833}{\sqrt{9}} \qquad 10{,}48 + \frac{1{,}29 \cdot 1{,}833}{\sqrt{9}}\right) = (9{,}69 \quad ; \quad 11{,}27)$$

IC IV

Intervalo de confianza para la diferencia de los valores medios correspondientes a dos distribuciones normales distintas pero que tienen la misma varianza

a. Sean dos experimentos aleatorios distintos a los cuales corresponden respectivamente las variables X e Y, ambas normales pero con la misma varianza σ^2.
Supóngase que se efectúen n_x repeticiones del primer experimento y n_y del segundo.

b. Según indicado en **[8]** de JtF IV se tiene que la variable:

$$T = \sqrt{\frac{(n_x + n_y - 2)}{(n_x S_x^2 + n_y S_y^2)\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}\left[(\bar{X} - \bar{Y}) - (m_X - m_Y)\right]$$

tiene una distribución t de Student con $n_x + n_y - 2$ grados de libertad.
Entonces, por un procedimiento análogo al indicado en **c.** IC III.1 se llega a que un intervalo de confianza para $m_X - m_Y$ con una confiabilidad igual a 0,95 está dado por:

$$\left((\bar{x}-\bar{y}) - \sqrt{\frac{(n_x s_x^2 + n_y s_y^2)\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}{(n_x+n_y-2)}}\, t_{0,025;(n_x+n_y-2)} \quad ; \quad (\bar{x}-\bar{y}) + \sqrt{\frac{(n_x s_x^2 + n_y s_y^2)\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}{(n_x+n_y-2)}}\, t_{0,025;(n_x+n_y-2)} \right)$$

siendo $\bar{x}$, $\bar{y}$, s_x y s_y los valores observados.

IC V

Intervalo de confianza para la varianza de una distribución normal

a. Sea un experimento aleatorio al cual corresponde una variable aleatoria X que sea normal (m , σ), siendo m y σ por el momento desconocidas.
Supóngase que se efectúen n repeticiones de dicho experimento, y que en base a los resultados obtenidos se pretende hallar un intervalo de confianza para la varianza σ^2 con un grado de confiabilidad de, digamos, 0,95.

b. Interesa pues hallar dos números α_1 y α_2 tales que:

$$P(\alpha_1 \leq \sigma^2 \leq \alpha_2) = 0{,}95 \qquad \textbf{[1]}$$

Es decir que debe ser:

$$P\left(\frac{\alpha_1}{nS^2} \leq \frac{\sigma^2}{nS^2} \leq \frac{\alpha_2}{nS^2}\right) = 0{,}95 \Rightarrow P\left(\frac{nS^2}{\alpha_1} \geq \frac{nS^2}{\sigma^2} \geq \frac{nS^2}{\alpha_2}\right) = 0{,}95 \Rightarrow$$
$$\Rightarrow P\left(\frac{nS^2}{\alpha_2} \leq \frac{nS^2}{\sigma^2} \leq \frac{nS^2}{\alpha_1}\right) = 0{,}95$$

Esto se cumplirá si, por ejemplo:

$$1^{\text{o}})\ P\left(\frac{nS^2}{\sigma^2} > \frac{nS^2}{\alpha_1}\right) = 0{,}025 \qquad \textbf{[2]}$$

y además

$$2^{\text{o}})\ P\left(\frac{nS^2}{\sigma^2} < \frac{nS^2}{\alpha_2}\right) = 0{,}025 \Rightarrow P\left(\frac{nS^2}{\sigma^2} \geq \frac{nS^2}{\alpha_2}\right) = 0{,}975 \qquad \textbf{[3]}$$

Ahora bien, como según se vio en JtF III $\frac{nS^2}{\sigma^2}$ tiene una distribución χ^2 con $n-1$ grados de libertad, por **[2]** y **[3]** resulta que:

$$\frac{nS^2}{\alpha_1} = \chi^2_{0,025\ ;\ n-1} \qquad \text{y} \qquad \frac{nS^2}{\alpha_2} = \chi^2_{0,975\ ;\ n-1}$$

y por lo tanto será:

$$\alpha_1 = \frac{nS^2}{\chi^2_{0,025;n-1}} \qquad \text{y} \qquad \alpha_2 = \frac{nS^2}{\chi^2_{0,975;n-1}}$$

Entonces, por **[1]**:

$$P\left(\frac{nS^2}{\chi^2_{0,025;n-1}} \leq \sigma^2 \leq \frac{nS^2}{\chi^2_{0,975;n-1}}\right) = 0,95 \qquad \textbf{[4]}$$

Es decir que existe una probabilidad igual a 0,95 de que el valor que asuma S^2 al efectuarse *n* repeticiones del experimento sea tal que el valor σ^2 esté dentro del intervalo:

$$\left[\frac{nS^2}{\chi^2_{0,025;n-1}} \ ; \ \frac{nS^2}{\chi^2_{0,975;n-1}}\right]$$

Entonces se define que:
Si se obtiene un valor observado s^2, la estimación de σ^2 consiste en el intervalo de confianza:

$$\left[\frac{ns^2}{\chi^2_{0,025;n-1}} \ ; \ \frac{ns^2}{\chi^2_{0,975;n-1}}\right] \qquad \textbf{[5]}$$

con una confiabilidad de 0,95.

c. En el caso de intervalos de confianza de la varianza, muy a menudo interesa un intervalo del tipo [0 , *l*].
Evidentemente, en este caso el intervalo será:

$$\left[0 \ ; \ \frac{ns^2}{\chi^2_{0,95;n-1}}\right] \qquad \textbf{[6]}$$

con una confiabilidad de 0,95.

d. Observación 1ª:
Análoga a la observación **c** de IC II.1.

e. Observación 2ª:
Si bien todo lo antedicho es rigurosamente válido solo para el caso de una distribución normal, en el caso de una distribución cualquiera y una cantidad muy grande de distribuciones del experimento, el método recién indicado será aproximadamente válido.

IC V.2

Aplicación

Sea el mismo caso indicado en IC III.2 y supóngase que se desee hallar un intervalo de confianza para la varianza σ^2.
Como es s = 1,29 y n = 10, por **[5]** se tiene que un intervalo de confianza para σ^2 con una confiabilidad de 0,95 será:

$$\left[\frac{10(1{,}29)^2}{\chi^2_{0,025;9}} \quad ; \quad \frac{10(1{,}29)^2}{\chi^2_{0,975;9}}\right] = \left[\frac{16{,}641}{19{,}02} \quad ; \quad \frac{16{,}641}{2{,}70}\right] = [0{,}875 \ ; \ 6{,}16]$$

IC VI

Intervalo de confianza para el cuociente de varianzas de dos distribuciones normales

a. Sean dos experimentos aleatorios distintos a los cuales corresponden respectivamente las variables X e Y ambas normales. Supóngase que dichas variables tengan varianzas σ^2_X y σ^2_Y respectivamente, las cuales son por el momento desconocidas.
Supóngase que se efectúen n_X repeticiones del primer experimento y n_Y repeticiones del segundo, y que en base a los resultados obtenidos interese hallar un intervalo de confianza para $\dfrac{\sigma^2_Y}{\sigma^2_X}$ con un grado de confiabilidad igual a 0,9.

b. Interesa pues hallar dos números α_1 y α_2 tales que:

$$P\left(\alpha_1 \leq \frac{\sigma^2_Y}{\sigma^2_X} \leq \alpha_2\right) = 0{,}9 \qquad \textbf{[1]}$$

Como:

$\dfrac{n_X S^2_X}{\sigma^2_X}$ tiene una distribución χ^2 con $n_X - 1$ grados de libertad.

$\dfrac{n_Y S^2_Y}{\sigma^2_Y}$ tiene una distribución χ^2 con $n_Y - 1$ grados de libertad.

Por lo visto en JtF V se tiene que:

$$\frac{\dfrac{n_X S_X^2}{\sigma_X^2 (n_X - 1)}}{\dfrac{n_Y S_Y^2}{\sigma_Y^2 (n_Y - 1)}} \text{ tiene una distribución } F_{(n_X-1)(n_Y-1)}$$

y como:

$$\frac{\dfrac{n_X S_X^2}{\sigma_X^2 (n_X - 1)}}{\dfrac{n_Y S_Y^2}{\sigma_Y^2 (n_Y - 1)}} = \overbrace{\boxed{\frac{n_X S_X^2 (n_Y - 1)}{n_Y S_Y^2 (n_X - 1)}}}^{\gamma} \frac{\sigma_Y^2}{\sigma_X^2} = \gamma \frac{\sigma_Y^2}{\sigma_X^2}$$

resulta que:

$$\gamma \frac{\sigma_Y^2}{\sigma_X^2} \text{ tiene una distribución } F_{(n_X-1)(n_Y-1)} \text{ siendo } \frac{n_X S_X^2 (n_Y - 1)}{n_Y S_Y^2 (n_X - 1)} = \gamma \qquad \textbf{[2]}$$

c. Por **[1]** se tiene que:

$$P\left(\gamma\, \alpha_1 \le \gamma \frac{\sigma_Y^2}{\sigma_X^2} \le \gamma\, \alpha_2 \right) = 0{,}9 \qquad \textbf{[3]}$$

Esta expresión se cumple cuando:

$$1^{\text{o}})\ P\left(\gamma \frac{\sigma_Y^2}{\sigma_X^2} > \gamma\, \alpha_2 \right) = 0{,}05 \Rightarrow \gamma\alpha_2 = f_{0{,}05(n_X-1)(n_Y-1)}$$

Por tener $\gamma \frac{\sigma_Y^2}{\sigma_X^2}$ una distribución $F_{(n_X-1)(n_Y-1)}$

$$2^{\text{o}})\ P\left(\gamma \frac{\sigma_Y^2}{\sigma_X^2} < \gamma\, \alpha 1 \right) = 0{,}05 \Rightarrow$$

$$\Rightarrow P\left(\gamma \frac{\sigma_Y^2}{\sigma_X^2} > \gamma\, \alpha 1 \right) = 0{,}95 \Rightarrow \gamma\alpha_1 = f_{0{,}95(n_X-1)(n_Y-1)} \qquad \textbf{[4]}$$

Entonces; por **[1]**:

$$P\left(\alpha_1 \le \frac{\sigma_Y^2}{\sigma_X^2} \le \alpha_2\right) = 0{,}9 \Rightarrow P\left(\frac{1}{\gamma}\gamma\alpha_1 \le \frac{\sigma_Y^2}{\sigma_X^2} \le \frac{1}{\gamma}\gamma\alpha_2\right) = 0{,}9 \overset{\text{Por } \mathbf{[2]} \text{ y } \mathbf{[4]}}{\Rightarrow}$$

$$\Rightarrow P\left(\frac{n_Y(n_X-1)S_Y^2}{n_X(n_Y-1)S_X^2} f_{0,95(n_X-1)(n_Y-1)} \le \frac{\sigma_Y^2}{\sigma_X^2} \le \frac{n_Y(n_X-1)S_Y^2}{n_X(n_Y-1)S_X^2} f_{0,05(n_X-1)(n_Y-1)}\right) = 0{,}9 \qquad \mathbf{[5]}$$

Es decir que existe una probabilidad igual a 0,9 de que los valores que asuman S_X^2 y S_Y^2 al efectuarse n_X veces el primer experimento y n_Y veces el segundo sean tales que el valor de $\frac{\sigma_Y^2}{\sigma_X^2}$ esté dentro del intervalo:

$$\left(\frac{n_Y(n_X-1)S_Y^2}{n_X(n_Y-1)S_X^2} f_{0,95(n_X-1)(n_Y-1)} \;;\; \frac{n_Y(n_X-1)S_Y^2}{n_X(n_Y-1)S_X^2} f_{0,05(n_X-1)(n_Y-1)}\right) \qquad \mathbf{[6]}$$

Si se obtuvieron valores observados s_X^2 y s_Y^2, un intervalo de confianza para $\frac{\sigma_Y^2}{\sigma_X^2}$ con confiabilidad 0,9 será:

$$\left(\frac{n_Y(n_X-1)\,s_Y^2}{n_X(n_Y-1)\,s_X^2} f_{0,95(n_X-1)(n_Y-1)} \;;\; \frac{n_Y(n_X-1)\,s_Y^2}{n_X(n_Y-1)\,s_X^2} f_{0,05(n_X-1)(n_Y-1)}\right) \qquad \mathbf{[7]}$$

d. Aplicación.

Sean $n_X = 10$, $n_Y = 12$ y que se obtuvo $s_X^2 = 5{,}1$ y $s_Y^2 = 4{,}7$.

De la tabla de la figura JtF III.b:

$f_{0,05\,;\,10-1\,;\,12-1} = 2{,}90$

$$f_{0,95\,;\,10-1\,;\,12-1} \underset{\text{Por } \mathbf{[4]} \text{ de JtF V}}{=} \frac{1}{f_{(1-0,95);10-1;12-1}} = \frac{1}{f_{0,05;10-1;12-1}} = \frac{1}{2{,}90} = 0{,}345$$

y por **[7]** resulta un intervalo de confianza para $\frac{\sigma_Y^2}{\sigma_X^2}$ con confiabilidad 0,9.

$$\left[\frac{12(10-1)}{10(12-1)}\frac{4{,}7}{5{,}1}0{,}345 \;;\; \frac{12(10-1)}{10(12-1)}\frac{4{,}7}{5{,}1}2{,}90\right] = [0{,}312 \;;\; 2{,}62]$$

Problemas sobre Intervalos de Confianza

IC 1 Suponga que la variable aleatoria X correspondiente al nivel de contaminación del aire en un ambiente industrial tiene una F. de D. normal. En 16 días tomados al azar se obtuvo una media observada $\bar{x} = 9$.

Se pide:

1º. Hallar un intervalo de confianza para m_X con una confiabilidad de 0,95 si se supiera con certeza que $\sigma_X = 4$.

2º. Hallar un intervalo de confianza para m_X con una confiabilidad de 0,95 si "a priori" se ignorara el valor de σ_X pero se tuviera que $s^2 = 25$.

3º. Indicar durante cuantos días se debería medir el nivel de contaminación si se desea que el error cometido en la estimación del valor medio de la contaminación sea menor que 0,25 con una confiabilidad de 0,95. Suponer que $\sigma_X = 4$.

IC 2 En una hilandería hay 50000 husos. Se quiere estimar el porcentaje de husos inactivos con una precisión del 2%, y con ese objeto se toma una muestra al azar. Indicar el tamaño de dicha muestra si se desea asegurar que la confiabilidad de la estimación sea de 0,95.

IC 3 En dos ocasiones distintas se efectuaron medidas sistemáticas de la estatura de los candidatos a ingresar al profesorado de cultura física. Los resultados fueron:

	N	$\bar{x}$	**s**
Año 1998	400	168 cm	6
Año 1999	900	172 cm	12

Se pide hallar un intervalo de confianza para el aumento de estatura media, con una confiabilidad de 0,95.

IC 4 Una fábrica ha llevado la cuenta de la venta diaria de un nuevo producto a lo largo de un mes. Si X es la variable correspondiente a dicha venta diaria, supóngase que se obtuvo $\bar{x} = 10$; $s = 1$.

En base a estos datos, hallar un intervalo de confianza con confiabilidad de 0,95 para la venta total correspondiente al próximo semestre (180 días).

IC 5 A la cantidad de fisiones de átomos radioactivos que ocurren en un cierto período corresponde una distribución de Poisson. Si considerando 1000 períodos se obtuvo $\bar{x} = 10$, se pide obtener un intervalo de confianza con confiabilidad 0,95 para el parámetro λ de la distribución de Poisson antedicha.

IC 6 Suponga que a la longitud de las barras de acero de una cierta producción corresponde una variable aleatoria X cuya F. de D. sea normal. Supóngase que se extraigan al azar 9 barras de dicha producción y que se obtenga un valor $s^2 = 84{,}5$.

Se pide indicar un intervalo de confianza con una confiabilidad de 0,95 de la varianza de dicha producción.

IC 7 A la longitud de los cables de acero de una cierta producción se le asocia una variable aleatoria X que se supone normal.

Se toma una muestra aleatoria de 12 cables, obteniéndose los siguientes resultados:

9,2 ; 9,7 ; 9,8 ; 10,2 ; 10,4 ; 10 ; 9,4 ; 9,5 ; 9,5 ; 10,3 ; 9,9 ; 9,7

Se pide indicar intervalos de confianza con una confiabilidad de 0,95 para el valor medio y para la varianza de X.