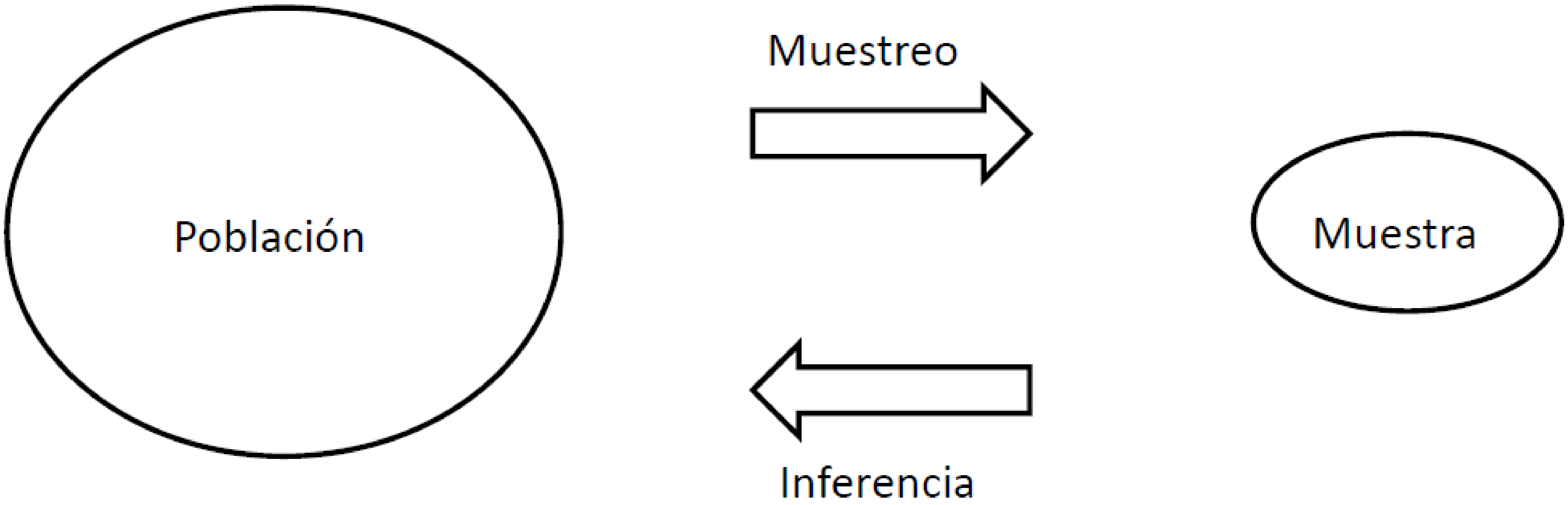
Intervalos de confianza

Docente: Johanna Trochez González

La inferencia estadística consiste en aquellos métodos por medio de los cuales se realizan inferencias o generalizaciones acerca de una población.



Métodos de inferencia

Métodos

- Clásico ✓
- Bayesiano

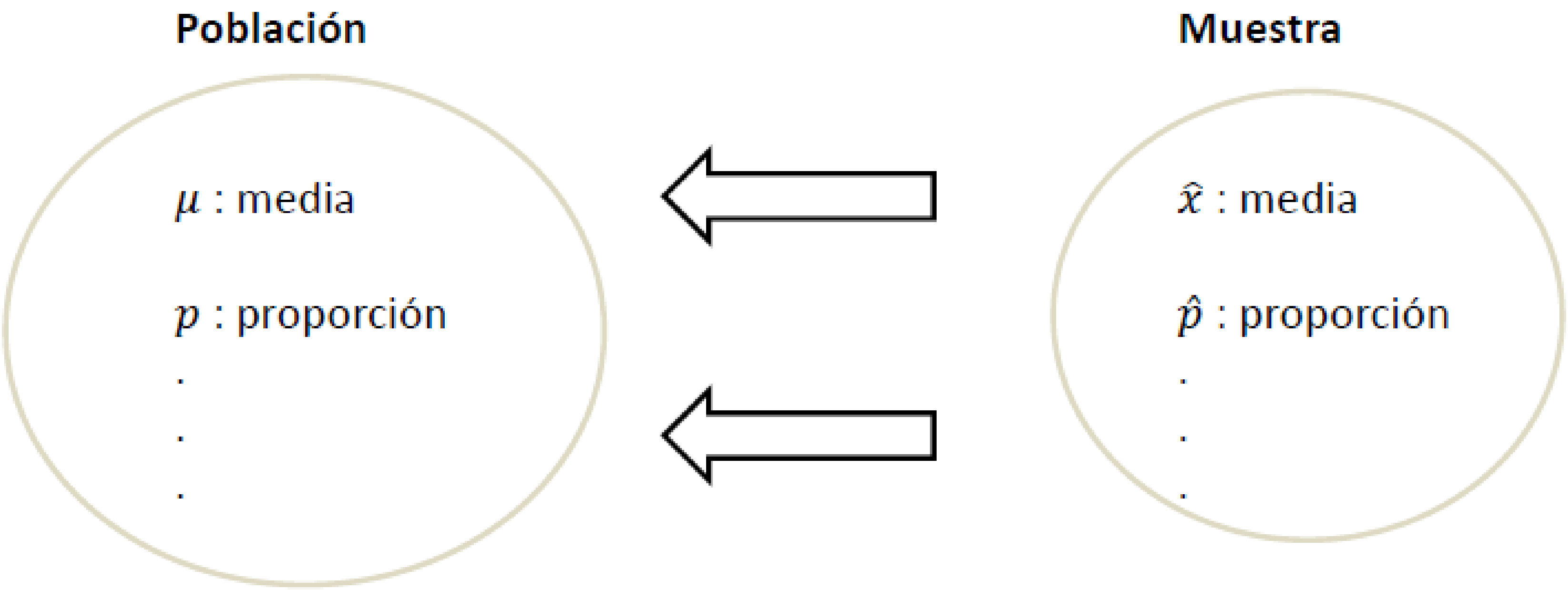
Problemas de inferencia

- Problemas
  - Estimación
    - Puntual
    - Por intervalo
  - Prueba de hipótesis

Estimación puntual

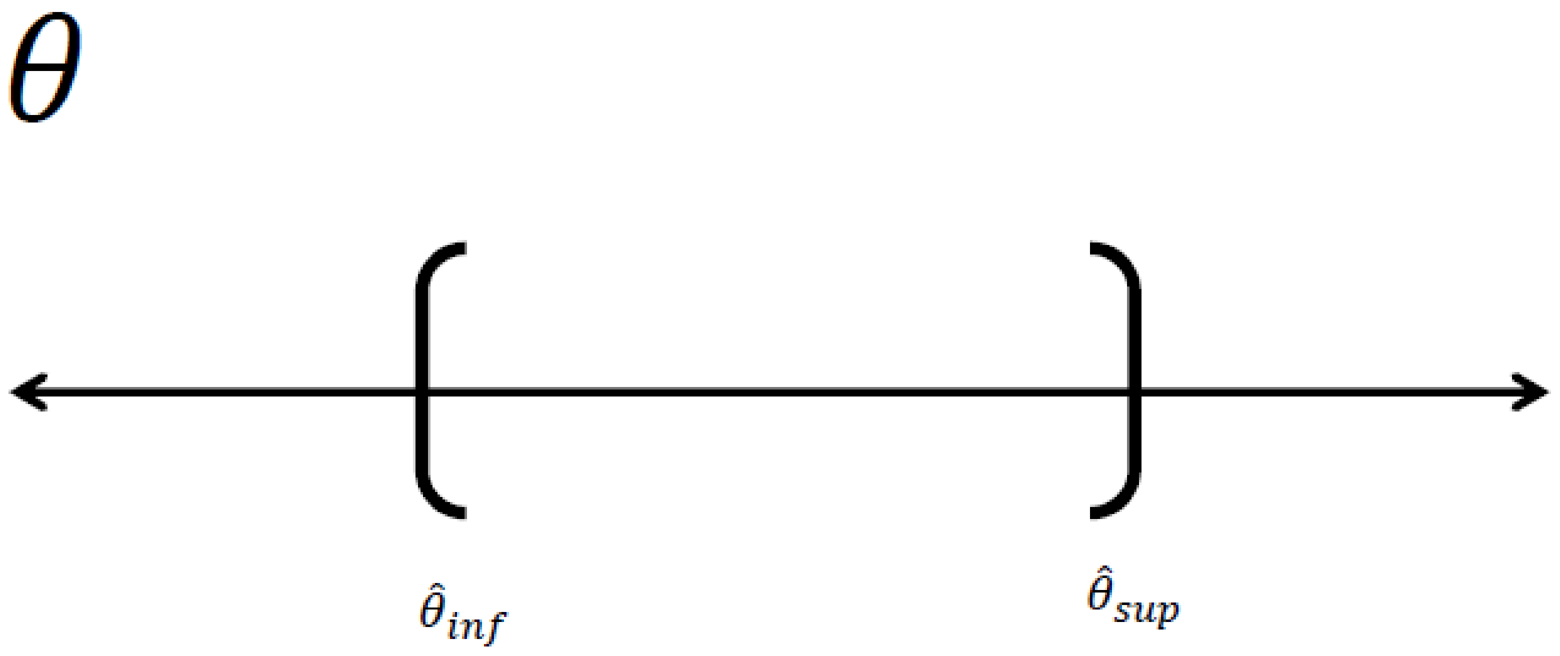
Una estimación puntual de algún parámetro θ de la población es un valor $\hat{\theta}$ de una estadística $\widehat{\Theta}$.

Ejemplo:



Estimación por intervalo

Una estimación por intervalo de un parámetro θ de la población es un intervalo de la forma $\hat{\theta}_{inf} < \theta < \hat{\theta}_{sup}$ donde $\hat{\theta}_{inf}$ y $\hat{\theta}_{sup}$ dependen de la estadística $\widehat{\Theta}$ y de la distribución de muestreo de $\widehat{\Theta}$.



Nivel de confianza y significancia de una estimación por intervalo

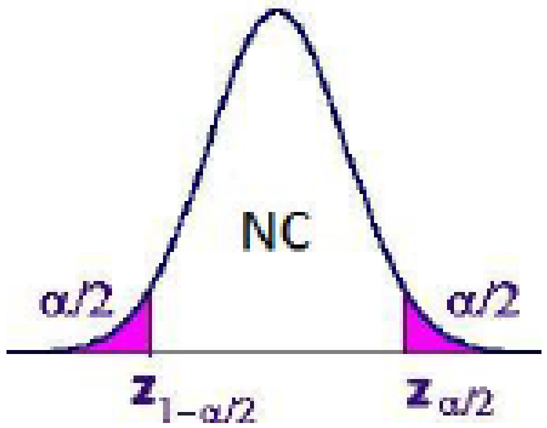
$$NC + \alpha = 1$$

$$\Downarrow \qquad\qquad \Downarrow$$

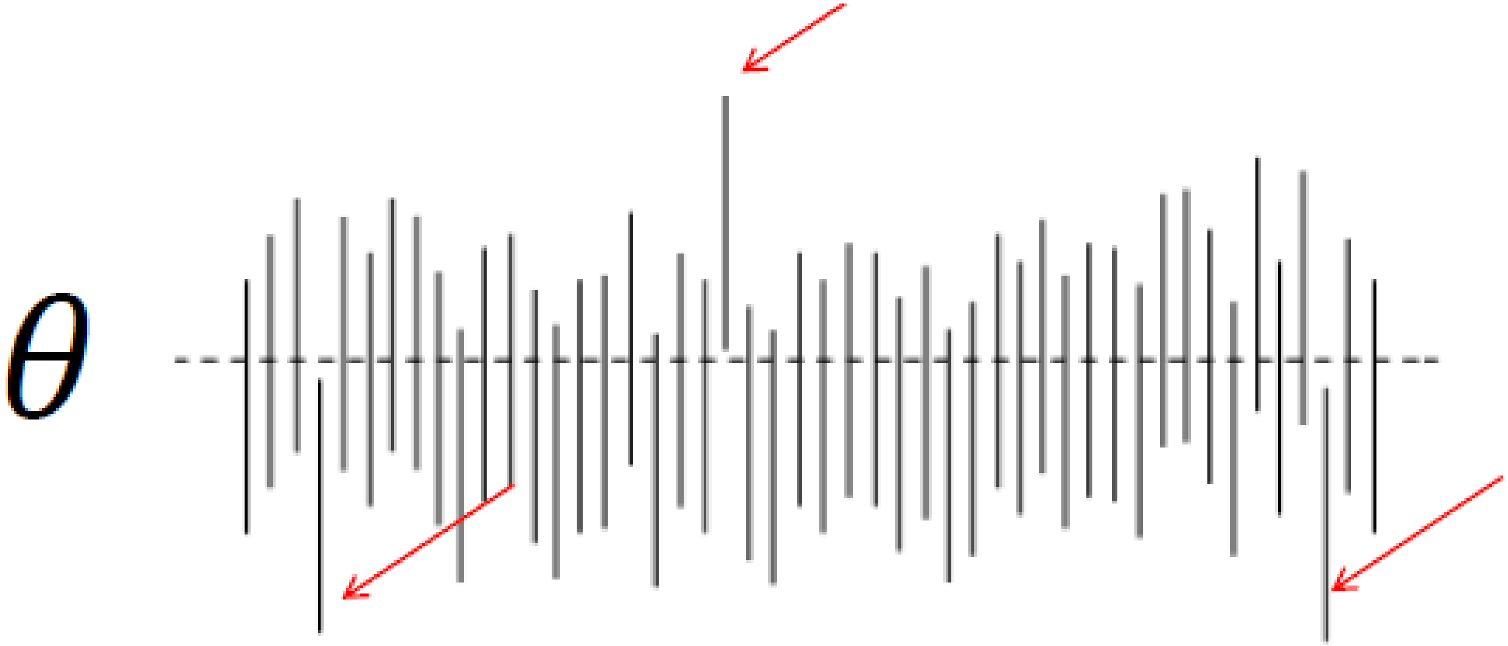
NC	α
0.90	0.10
0.95	0.05
0.99	0.01

Valores usuales de $Z_{\alpha/2}$ para intervalos de confianza

Nivel de confianza	α	$z_{\alpha/2}$
90%	0.10	1.645
95%	0.05	1.960
99%	0.01	2.576



Representación del nivel de confianza



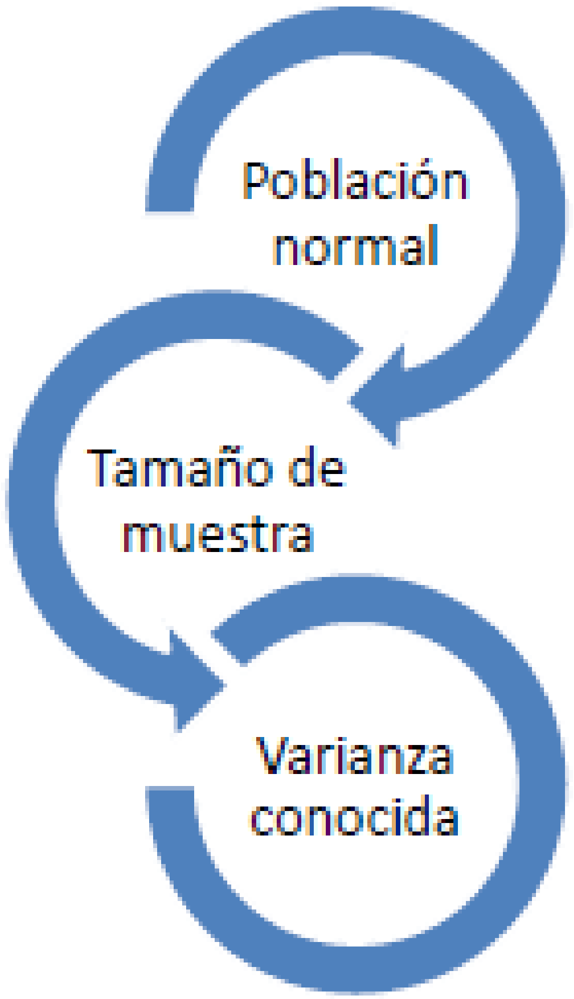
Ejemplo: si se construyen 100 intervalos de la forma $\hat{\theta}_{inf} < \theta < \hat{\theta}_{sup}$ con un nivel de confianza $NC = 0.97$ entonces se espera que 97 de los 100 intervalos contengan al parámetro θ.

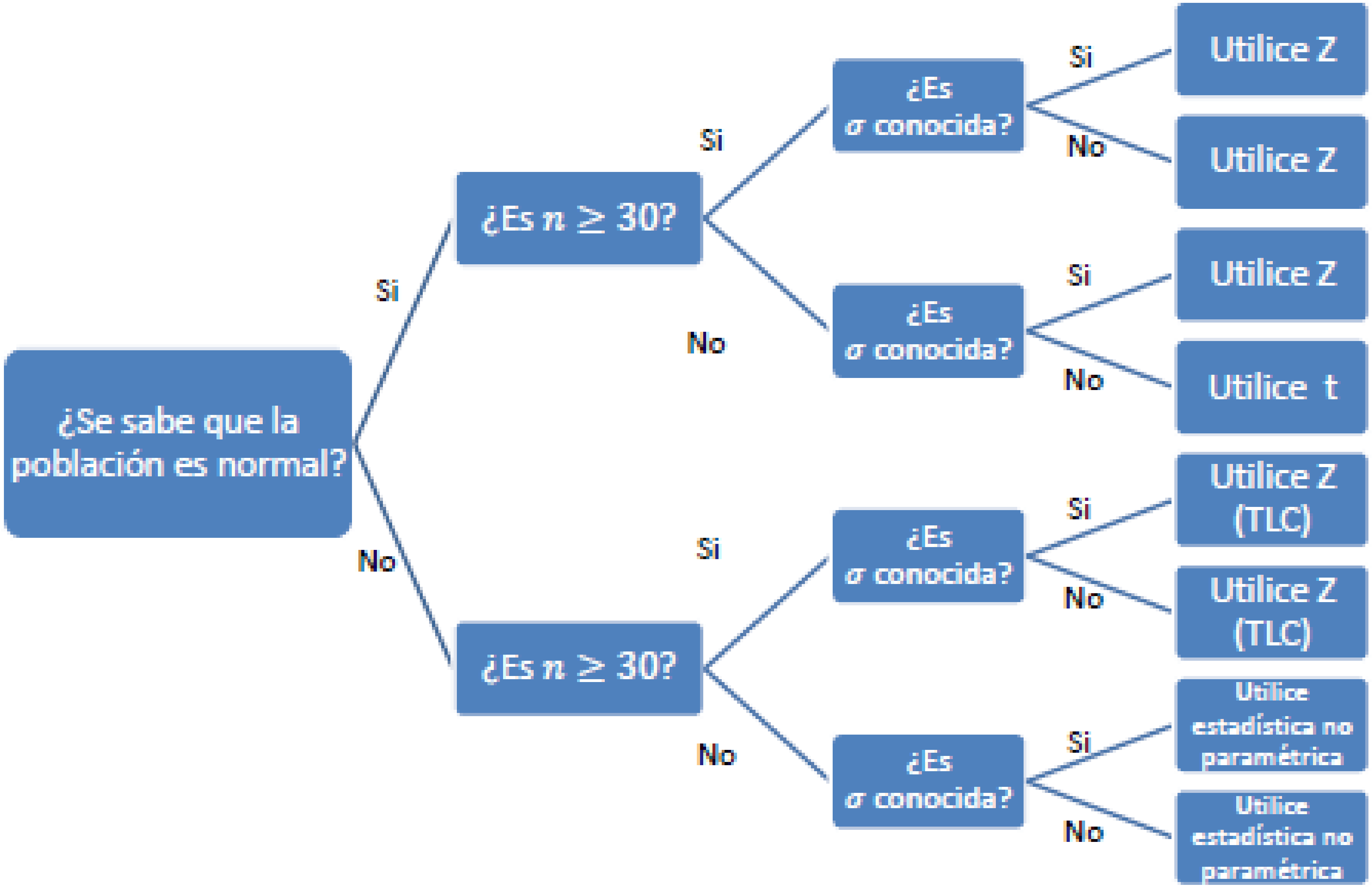
Intervalos de confianza más usados

- Intervalos de confianza
  - Para una muestra
    - μ
    - p
    - σ^2
  - Para dos muestras
    - $\mu_1 - \mu_2$
    - $\mu_1 - \mu_2$ pareadas
    - $p_1 - p_2$
    - σ_1^2 / σ_2^2

Intervalos de confianza para una muestra

Aspectos relevantes para construir intervalos de confianza para μ




¿Se sabe que la población es normal?
Si
No
¿Es $n \geq 30$?
¿Es $n \geq 30$?
Si
No
Si
No
¿Es σ conocida?
¿Es σ conocida?
¿Es σ conocida?
¿Es σ conocida?
Si
No
Si
No
Si
No
Si
No
Utilice Z
Utilice Z
Utilice Z
Utilice t
Utilice Z (TLC)
Utilice Z (TLC)
Utilice estadística no paramétrica
Utilice estadística no paramétrica

Intervalo de confianza para µ con σ conocida

Si $\bar{x}$ es la media de una muestra aleatoria de tamaño n de una población normal con varianza σ^2conocida, un intervalo de confianza del $1-\alpha\ 100\%$ para μ está dado por

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

donde $Z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha.

EJEMPLO

Las siguiente mediciones se registraron para el tiempo de secado, en horas, de cierta marca de pintura de látex:

3.4	2.5	4.8	2.9	3.6
2.8	3.3	5.6	3.7	2.8
4.4	4.0	5.2	3.0	4.8

Suponga que las mediciones representan una muestra aleatoria de una población normal con una desviación de 1.2 minutos. Encuentre un intervalo de confianza para el tiempo medio de secado con una confianza del 95%.

Solución:

La muestra proviene de una distribución normal.

$\alpha = 0.05$

$n = 15$

$\bar{x} = 3.79$ minutos

$\sigma = 1.2$ minutos

Intervalo de confianza para μ con σ desconocida

Si $\bar{x}$ es la media de una muestra aleatoria de tamaño n de una población normal con varianza σ^2 desconocida, un intervalo de confianza del $(1-\alpha)\,100\%$ para μ está dado por

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

donde $t_{\frac{\alpha}{2},n-1}$ es el valor t que deja un área de $\alpha/2$ a la derecha, con n-1 grados de libertad

EJEMPLO

Una máquina produce piezas metálicas de forma cilíndrica. Se toma una muestra de las piezas y los diámetros son 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01, y 1.03 centímetros. Encuentre un intervalo de confianza del 99% para el diámetro medio de las piezas de esta máquina, suponga una distribución aproximadamente normal

Solución:

La muestra proviene de una distribución normal.

$\alpha = 0.01$

$n = 9$

$\bar{x} = 1.0056$ cm

$s = 0.0246$ cm

Con un nivel de confianza del 99% se espera que el diámetro medio (μ) esté entre 0.9781 cm y 1.0331 cm.

Intervalo de confianza para la proporción $\boldsymbol{p}$

Si $\hat{p}$ es la proporción de éxitos en una muestra aleatoria de tamaño n, un intervalo de confianza aproximado de $(1 - \alpha)$ 100% para p está dado por

$$\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

donde $Z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha.

Nota: usar solo cuando $np \geq 50$ y $n(1 - p) \geq 50$

EJEMPLO

El gerente de una estación de televisión debe determinar en la ciudad qué porcentaje de casas tienen más de un televisor. Una muestra aleatoria de 500 casas revela que 275 tienen dos o más televisores. ¿Cuál es el intervalo de confianza del 90% para estimar la proporción de todas las casas que tienen dos o más televisores?

Solución:

$$\alpha = 0.10$$
$$n = 500$$
$$\hat{p} = \frac{275}{500} = 0.55$$

Con un nivel de confianza del 90% se espera que la proporción (p) de casas con dos o más televisores esté entre 0.5134 y 0.5866.

Distribución muestral para la varianza

Estimar la **variabilidad** en los diámetros de tuercas de cierto tipo.

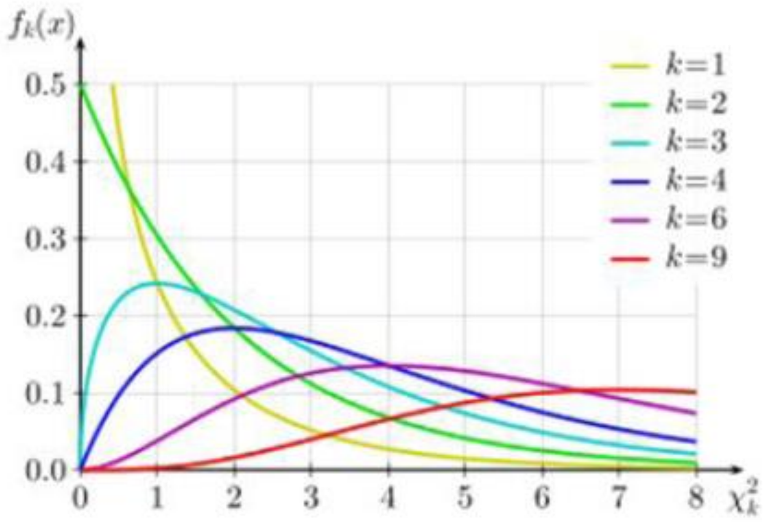
La distribución de muestreo es:

$$\chi^2{}_{(n-1)} \sim \frac{(n-1)S^2}{\sigma^2}$$

Distribución chi-cuadrado

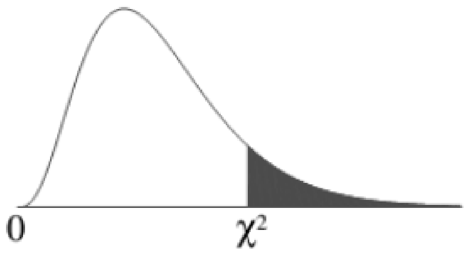
$$f(x;\,k) = \begin{cases} \dfrac{x^{(k/2)-1}e^{-x/2}}{2^{k/2}\Gamma\left(\frac{k}{2}\right)}, & x \geq 0; \\ 0, & \text{otherwise.} \end{cases}$$

with x ∈ $[0, \infty)$ and $k > 0$



The parameter k is called degrees of freedom (df)

Tabla de la distribución chi-cuadrado



Areas under curve

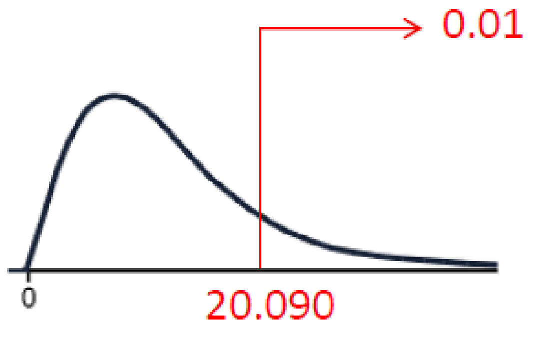
df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$
1	0.000	0.000	0.001	0.004	0.016	2.706
2	0.010	0.020	0.051	0.103	0.211	4.605
3	0.072	0.115	0.216	0.352	0.584	6.251
4	0.207	0.297	0.484	0.711	1.064	7.779
5	0.412	0.554	0.831	1.145	1.610	9.236

Degrees of freedom

Quantiles

Obtain α if $x^2_{\alpha,\ 8} = 20$

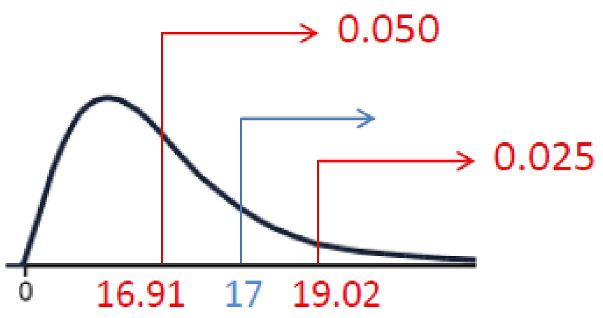
df	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	3.841	5.024	6.635	7.879
2	5.991	7.378	9.210	10.597
3	7.815	9.348	11.345	12.838
4	9.488	11.143	13.277	14.860
5	11.070	12.833	15.086	16.750
6	12.592	14.449	16.812	18.548
7	14.067	16.013	18.475	20.278
8	15.507	17.535	20.090	21.955
9	16.919	19.023	21.666	23.589
10	18.307	20.483	23.209	25.188



Answer: $\alpha = 0.01$

Obtain α if $x^2_{\alpha,\,9} = 17$

df	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	3.841	5.024	6.635	7.879
2	5.991	7.378	9.210	10.597
3	7.815	9.348	11.345	12.838
4	9.488	11.143	13.277	14.860
5	11.070	12.833	15.086	16.750
6	12.592	14.449	16.812	18.548
7	14.067	16.013	18.475	20.278
8	15.507	17.535	20.090	21.955
9	16.919	19.023	21.666	23.589
10	18.307	20.483	23.209	25.188



Answer: $\alpha \in (0.025, 0.050)$

Intervalo de confianza para la varianza σ^2

Si s^2 es la varianza de una muestra aleatoria de tamaño n de una población normal, un intervalo de confianza aproximado de (1 − α) 100% para σ^2 está dado por

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

donde los denominadores son obtenidos de una chi-cuadrada.

Nota: un intervalo de confianza para σ se puede obtener tomando a los límites del intervalo anterior.

EJEMPLO

La duración en horas de un conjunto de bombillas se muestra a continuación:

1470 1510 1690 1740 1900 2000 2030 2010 2190
2200 2290 2380 2390 2480 2500 2580 2700

Construir un intervalo de confianza del 95% para la varianza poblacional σ^2.

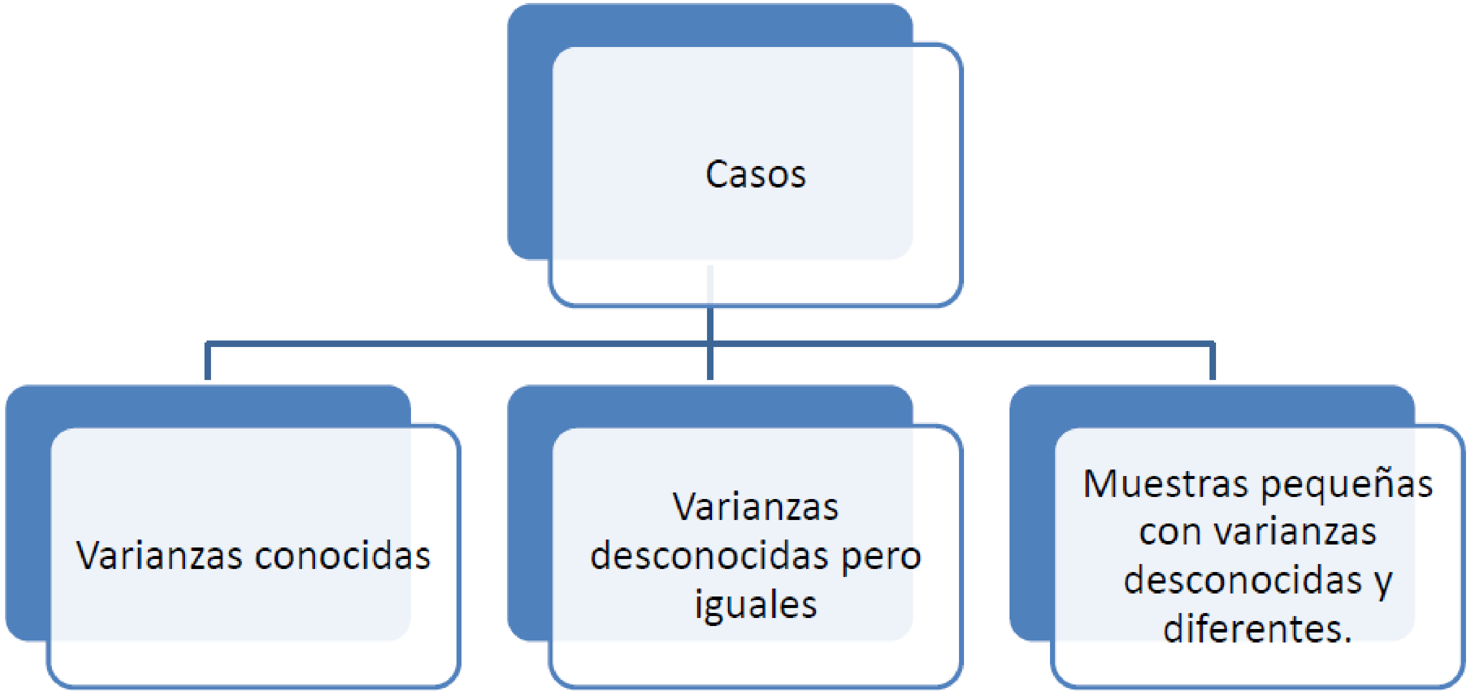
Solución:

$$\alpha = 0.05$$
$$n = 17$$
$$s^2 = 138098.5$$

Con un nivel de confianza del 95% se espera que la varianza poblacional esté entre 76601.7 y 316558.2

Intervalos de confianza para dos muestras

Casos para intervalos de confianza de $\mu_1 - \mu_2$



Distribución muestral para la diferencia de medias

- ***Medicina*.-** Poner aprueba los efectos de una dieta mediante las medidas del peso en la misma persona antes y después de aplicar una dieta.

- ***Agricultura.***- Poner a prueba los efectos de dos fertilizantes en la producción granos comparando la producción de parcelas similares en las mismas condiciones.

- ***Industria:*** Poner a prueba dos marcas de llantas en cuanto al desgaste del piso.

Distribución muestral para la diferencia de medias con σ conocida

Sean dos muestras aleatorias independientes de tamaños $n_1 \ y \ n_2$, de dos poblaciones discretas o continuas, con medias $\mu_1 \ y \ \mu_2$ y varianzas $\sigma_1^2 \ y \ \sigma_2^2$ respectivamente.

La distribución muestral para la diferencias de las medias $\bar{X}_1 - \bar{X}_2$ está distribuida de manera aproximadamente normal con media y varianzas dadas por:

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_1 - \mu_2$$

$$s^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1}\right) + \left(\frac{\sigma_2^2}{n_2}\right)}}$$

$\mu_1 \ y \ \mu_2$ medias poblacionales

σ_1^2 y σ_2^2 varianzas poblacionales conocidas

$\bar{X}_1 \ y \ \bar{X}_2$ medias muestrales

Intervalo de confianza para la diferencia entre dos medias

Si $\overline{x_1}$ y $\overline{x_2}$ son las medias muestrales de muestras aleatorias independientes de tamaño n1 y n2 de poblaciones normales con varianzas conocidas $\sigma_1^2\ y\ \sigma_2^2$, respectivamente, un intervalo de confianza de (1 – α) 100% para μ1 – μ2 está dado por

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\sqrt{\left(\frac{\sigma_1^2}{n_1}\right) + \left(\frac{\sigma_2^2}{n_2}\right)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\sqrt{\left(\frac{\sigma_1^2}{n_1}\right) + \left(\frac{\sigma_2^2}{n_2}\right)}$$

EJEMPLO

Se lleva a cabo un experimento en que se comparan dos tipos de motores, A y B. Se mide el rendimiento en kilómetros por galón de gasolina. Se realizan 50 experimentos con el motor A y 75 con el motor B. La gasolina que se utiliza y demás condiciones se mantienen constantes. El rendimiento promedio de gasolina para el motor A es de 36 Km/galón y el promedio para el motor B es de 42 Km/galón. Encuentre un intervalo de confianza de 96% sobre $\mu_B - \mu_A$, donde μ_B y μ_A son los rendimientos de gasolina poblacional para los motores B y A respectivamente. Suponga desviaciones estándar de seis y ocho para los motores A y B respectivamente.

Solución:

$$\alpha = 0.04$$

$$n_1 = 50 \qquad \bar{x}_1 = 36 \qquad \sigma_1 = 6$$

$$n_2 = 75 \qquad \bar{x}_2 = 42 \qquad \sigma_2 = 8$$

$$3.43 < \mu_B - \mu_A < 8.57$$

Con un nivel de confianza del 96% se espera que $\mu_B - \mu_A$ esté entre 3.43 Km/galón y 8.57 Km/galón. **Como el 0 NO está en el intervalo hay evidencias de que μ_B es mayor que μ_A.**

Distribución muestral para la diferencia de medias con σ desconocida

Para que la distribución este asociada a una distribución t de student n < 30 y $\sigma_1^{\ 2} = \sigma_2^{\ 2}$

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_1 - \mu_2$$

$$s^2{}_p = \left[\frac{(n_1-1)\ S_1^{\ 2} + (n_2-1)\ S_2^{\ 2}}{n_1 + n_1 - 2}\right]$$

$$t_{(n_1+n_2-2)} = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{s^2{}_{\bar{X}_1-\bar{X}_2}}}$$

Intervalo de confianza para la diferencia entre dos medias con varianzas poblacionales desconocidas

Si $\overline{x_1}$ y $\overline{x_2}$ son las medias muestrales de muestras aleatorias independientes de tamaño $n1$ y $n2$ de poblaciones normales con varianzas iguales pero desconocidas, un intervalo de confianza de $(1 - \alpha)$ 100% para $\mu1 - \mu2$ está dado por

$$(\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2},v} S_p \sqrt{\left(\frac{1}{n_1}\right) + \left(\frac{1}{n_2}\right)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2},v} S_p \sqrt{\left(\frac{1}{n_1}\right) + \left(\frac{1}{n_2}\right)}$$

donde $v = n1 + n2 - 2$

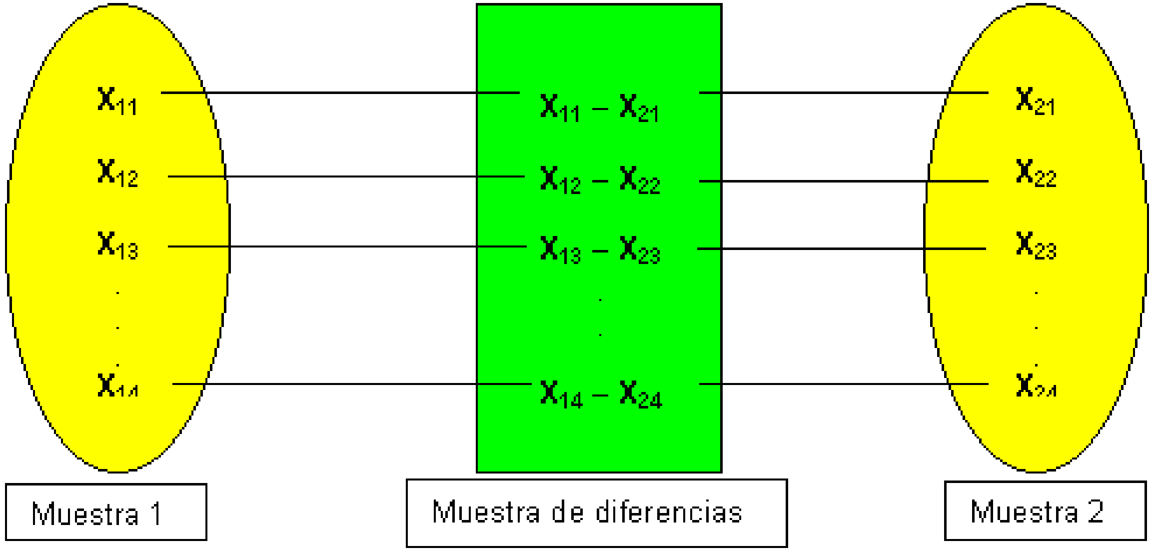
Intervalo de confianza para la diferencia entre dos medias con varianzas poblacionales desconocidas y diferentes

Si $\overline{x_1}$ y $\overline{x_2}$ son las medias muestrales de muestras aleatorias independientes de tamaño $n1$ y $n2$ de poblaciones normales con varianzas desconocidas y diferentes, un intervalo de confianza de (1 − α) 100% para $\mu1$ − $\mu2$ está dado por

$$(\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2},v}\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2},v}\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}$$

$$v = \frac{\left(\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_1}\right)^2}{n_2 - 1}}$$

Distribución muestral para las diferencias pareadas



Distribución muestral para las diferencias pareadas

Considera dependencia entre pares, el estadístico es:

$$t_{(n-1)} = \frac{\bar{d} - \mu_d}{sd/\sqrt{n}}$$

Donde:

$\bar{d}$: Promedio de las diferencias pareadas

μ_d: Promedio esperado para las diferencias pareadas

sd: Desviación estándar muestral

n: Tamaño de la muestra

Intervalo de confianza para la diferencia de medias pareadas

Si d y sd son la media y la desviación estándar de las diferencias distribuidas normalmente de n pares aleatorios de mediciones, entonces un intervalo de confianza de 1 − α 100% para μd = μ1 − μ2 está dado por

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

donde v = n − 1.

Ejemplo

Se tienen los datos del rendimiento (millas/galón) en carretera y en ciudad para 12 vehículos.

Vehículo/ Rendimiento	Carretera	Ciudad	Diferencia pareada
1	31	25	6
2	25	18	7
3	26	20	6
4	26	19	7
5	30	22	8
6	31	22	9
7	28	19	9
8	25	16	9
9	27	19	8
10	25	16	9
11	25	16	9
12	36	25	11
Media	**27.92**	**19.75**	**8.17**
Desviación muestral	**3.45**	**3.19**	**1.47**

Distribución muestral para la diferencia de proporciones

APLICACIONES:

- **Educación:** ¿Es mayor la proporción de los estudiantes que aprueban matemáticas que las de los que aprueban inglés?

- **Medicina:** ¿Es menor el porcentaje de los usuarios del medicamento A que presentan una reacción adversa que el de los usuarios del fármaco B que también presentan una reacción de ese tipo?

- **Administración:** ¿Hay diferencia entre los porcentajes de hombres y mujeres en posiciones gerenciales?.

- **Ingeniería:** ¿Existe diferencia entre la proporción de artículos defectuosos que genera la máquina A a los que genera la máquina B?

Distribución muestral para la diferencia de proporciones

Sean las variables aleatorias X y Y tal que:

$$\left.\begin{array}{l} X \rightarrow B(n_x, p_x) \\ Y \rightarrow B(n_Y, p_Y) \end{array}\right\} \text{Independientes}$$

- Cuando n_x y n_Y son grandes:

$$X \rightarrow B(n_x p_x, n_x p_x q_x)$$
$$Y \rightarrow B(n_Y p_Y, n_Y p_Y q_Y)$$

Definimos las proporciones muestrales como:

$$\boxed{\begin{aligned} \hat{p}_X &= \frac{X}{n_X} \\ \hat{p}_Y &= \frac{Y}{n_Y} \end{aligned}}$$

El estadístico corresponde a:

$$Z = \frac{\left(\widehat{p_x} - \widehat{p_y}\right) - (p_x - p_y)}{\sqrt{\left(\frac{p_x q_x}{n_x}\right) + \left(\frac{p_y q_y}{n_y}\right)}}$$

Intervalo de confianza para la diferencia de proporciones

$\widehat{p_x}\;\; y \;\; \widehat{p_y}$ son las proporciones de éxito de dos muestras aleatorias independientes de tamaño $n1$ y $n2$, entonces un intervalo de confianza de (1 – α) 100% para $p1$ – $p2$ está dado por

$$\left(\widehat{p_x} - \widehat{p_y}\right) - Z_{\frac{\alpha}{2}}\sqrt{\left(\frac{p_x q_x}{n_x}\right) + \left(\frac{p_y q_y}{n_y}\right)} < p_x - p_y < \left(\widehat{p_x} - \widehat{p_y}\right) + Z_{\frac{\alpha}{2}}\sqrt{\left(\frac{p_x q_x}{n_x}\right) + \left(\frac{p_y q_y}{n_y}\right)}$$

Ejemplo

Se sabe que en una población el 28% de las mujeres y el 25% de los hombres son fumadores. Se extraen muestras de 42 mujeres y 40 hombres.

Determinar la probabilidad de que las mujeres fumadoras superen a los hombres fumadores en al menos el 4%.

Se definen las variables:

$$P_X = Proporción\ de\ mujeres\ fumadoras\ en\ la\ población = 0.28$$
$$P_y = Proporción\ de\ hombres\ fumadores\ en\ la\ población = 0.25$$
$$\widehat{P}_x = Proporción\ de\ mujeres\ \ fumadoras\ en\ la\ \ muestra$$
$$\widehat{P}_y = Proporción\ de\ hombres\ fumadores\ en\ la\ muestra$$

Ejemplo 9

Se quiere determinar si un cambio en el método de fabricación de una piezas ha sido efectivo o no. Para esta comparación se tomaron 2 muestras, una antes y otra después del cambio en el proceso y los resultados obtenidos son los siguientes.

	Antes del cambio	Después del cambio
N° de piezas defectuosas	75	80
N° de piezas analizadas	1500	2000

Construir un intervalo de confianza del 90% para decidir si el cambio tuvo efecto positivo o no.

Solución:

$$\hat{p}_1 = 75/1500 = 0.05 \qquad \hat{p}_2 = 80/2000 = 0.04$$

Distribución muestral para la razón de varianzas

- Sean las variables aleatorias X e Y tales que

$$\left.\begin{array}{l} X \longrightarrow N(\mu_X, \sigma_X) \\ Y \longrightarrow N(\mu_Y, \sigma_Y) \end{array}\right\} \text{ Independientes}$$

Consideramos:

$$\left.\begin{array}{c} \text{m.a.s. de tamaño } n_X \text{ de } X \\ X_1, X_2, \ldots, X_{n_X} \end{array}\right\} \quad \overline{X},\ S_X^2$$

$$\left.\begin{array}{c} \text{m.a.s. de tamaño } n_Y \text{ de } Y \\ Y_1, Y_2, \ldots, Y_{n_Y} \end{array}\right\} \quad \overline{Y},\ S_Y^2$$

$$\overline{X} = \frac{1}{n_X} \sum_{i=1}^{n_X} X_i \qquad S_X^2 = \frac{1}{n_X - 1} \sum_{i=1}^{n_X} (X_i - \overline{X})^2$$

$$\overline{Y} = \frac{1}{n_Y} \sum_{j=1}^{n_Y} Y_j \qquad S_Y^2 = \frac{1}{n_Y - 1} \sum_{j=1}^{n_Y} (Y_j - \overline{Y})^2$$

El estadístico es:

$$F = \frac{S_1^2 * \sigma_2^2}{S_2^2 * \sigma_1^2}$$

tiene una distribución F de Snedecor con $n_X - 1, n_Y - 1$, grados de libertad

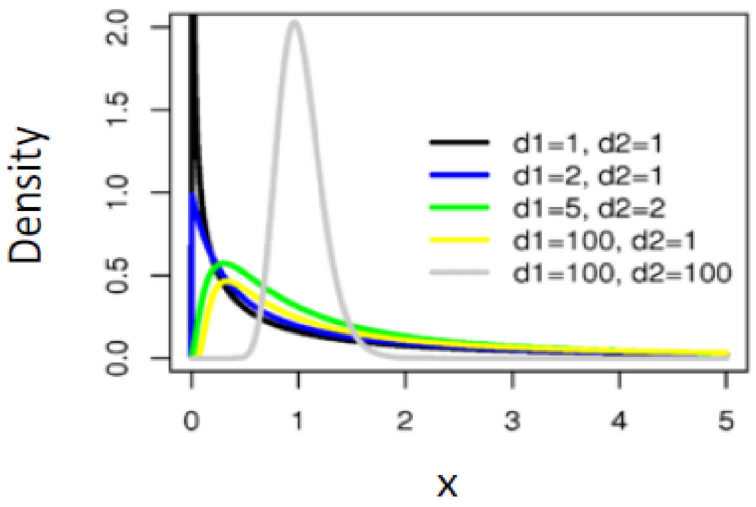
F distribution

This distribution is used in analysis of variance.

The probability density function (fdp) is given by

$$f(x; d_1, d_2) = \frac{\sqrt{\frac{(d_1\, x)^{d_1}\, d_2^{d_2}}{(d_1\, x + d_2)^{d_1 + d_2}}}}{x\, \mathrm{B}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}$$

with $x \in [0, \infty)$ and $d_1, d_2 > 0$

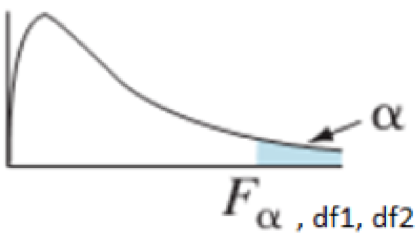


The parameter d_1, d_2 are called degrees of freedom (df)

Where $\mathrm{B}(x, y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x + y)}$

F table

Areas under curve



df_2	α	df_1 1	2	3	4
1	.25	5.83	7.50	8.20	8.58
	.10	39.86	49.50	53.59	55.83
	.05	161.4	199.5	215.7	224.6
	.025	647.8	799.5	864.2	899.6
	.01	4052	5000	5403	5625
2	.25	2.57	3.00	3.15	3.23
	.10	8.53	9.00	9.16	9.24
	.05	18.51	19.00	19.16	19.25
	.025	38.51	39.00	39.17	39.25
	.01	98.50	99.00	99.17	99.25
	.005	198.5	199.0	199.2	199.2
	.001	998.5	999.0	999.2	999.2

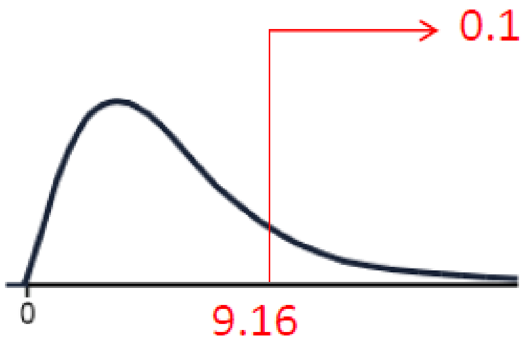
Degrees of freedom df1

Quantiles

Degrees of freedom df2

Obtain $f_{3,\,2,\,0.1}$

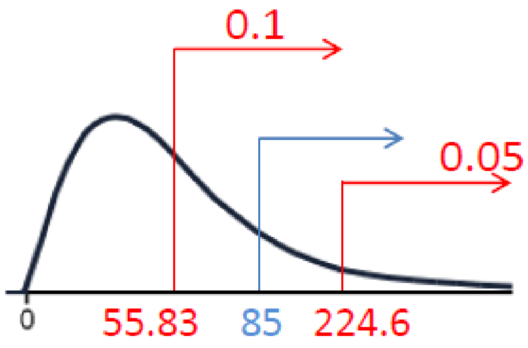
		df₁			
df_2	α	1	2	3	4
1	.25	5.83	7.50	8.20	8.58
	.10	39.86	49.50	53.59	55.83
	.05	161.4	199.5	215.7	224.6
	.025	647.8	799.5	864.2	899.6
	.01	4052	5000	5403	5625
2	.25	2.57	3.00	3.15	3.23
	.10	8.53	9.00	9.16	9.24
	.05	18.51	19.00	19.16	19.25
	.025	38.51	39.00	39.17	39.25
	.01	98.50	99.00	99.17	99.25
	.005	198.5	199.0	199.2	199.2
	.001	998.5	999.0	999.2	999.2



Answer: $f_{3,\,2,\,0.1} = 9.16$

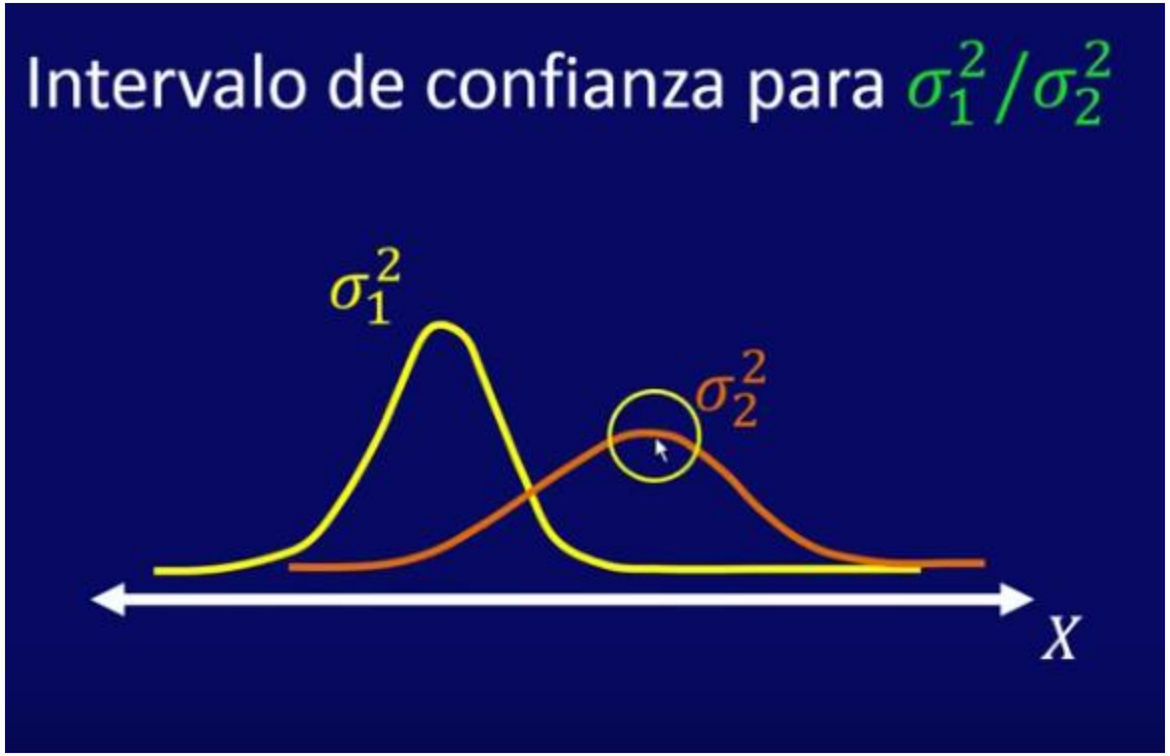
Obtain α if $f_{4,\,1,\,\alpha} = 85$

df_2	α	df_1: 1	2	3	4
1	.25	5.83	7.50	8.20	8.58
	.10	39.86	49.50	53.59	55.83
	.05	161.4	199.5	215.7	224.6
	.025	647.8	799.5	864.2	899.6
	.01	4052	5000	5403	5625
2	.25	2.57	3.00	3.15	3.23
	.10	8.53	9.00	9.16	9.24
	.05	18.51	19.00	19.16	19.25
	.025	38.51	39.00	39.17	39.25
	.01	98.50	99.00	99.17	99.25
	.005	198.5	199.0	199.2	199.2
	.001	998.5	999.0	999.2	999.2



Answer: $\alpha \in (0.05\ 0.10)$

Intervalo de confianza para la razón de varianzas σ_1^2/σ_2^2



La pregunta de interés consiste en saber si son iguales las varianzas poblacionales?

El intervalo de confianza corresponde a:

$$\frac{{S_1}^2}{{S_2}^2} * \frac{1}{f_{\frac{\alpha}{2},n_1-1,n_2-1}} < \frac{\boldsymbol{\sigma_1}^{\mathbf{2}}}{\boldsymbol{\sigma_2}^{\mathbf{2}}} < \frac{{S_1}^2}{{S_2}^2} * f_{\frac{\alpha}{2},n_2-1,n_1-1}$$

Se observa si el intervalo de confianza contiene al 1:

- Si lo contiene se dice que hay evidencia de que las varianzas poblacionales son iguales

- Si el intervalo es inferior a 1 la varianza de la segunda población es superior a la varianza de la población 1

- Si el intervalo es superior a 1 la varianza de la primera población es superior a la varianza de la segunda población.

EJEMPLO

- Se requiere determinar si dos tipos de motores, uno con gas y el otro con gasolina, tiene la misma variabilidad en millas por galón de gasolina o su equivalente en gas. Se tomaron un conjunto de autos de cada tipo y midió el rendimiento del combustible en un recorrido estándar, dando los siguientes resultados:

Gasolina	34	36	39	31	33	26	45	34	39	38	37	36		
Gas	33	41	39	32	29	28	33	34	25	28	36	33	35	35

- Construya un intervalo de confianza del 90% para la razón de varianzas

Solución

$$NC = 0.9 \qquad \alpha = 0.1 \qquad \frac{\alpha}{2} = 0.05$$

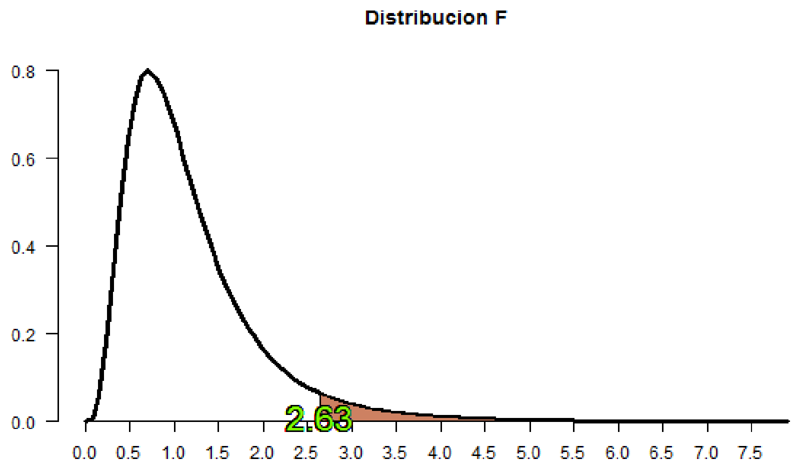
Limite inferior: $f_{\frac{\alpha}{2},n_1-1,n_2-1}$=$f_{0.05,11,13=2.634}$

Limite superior: $f_{\frac{\alpha}{2},n_2-1,n_1-1}$ =$f_{0.05,13,11=2.7614}$

	Gasolina 1	**Gas 2**
n	12	14
$\boldsymbol{S^2}$	22.24	19.15

Limite inferior:

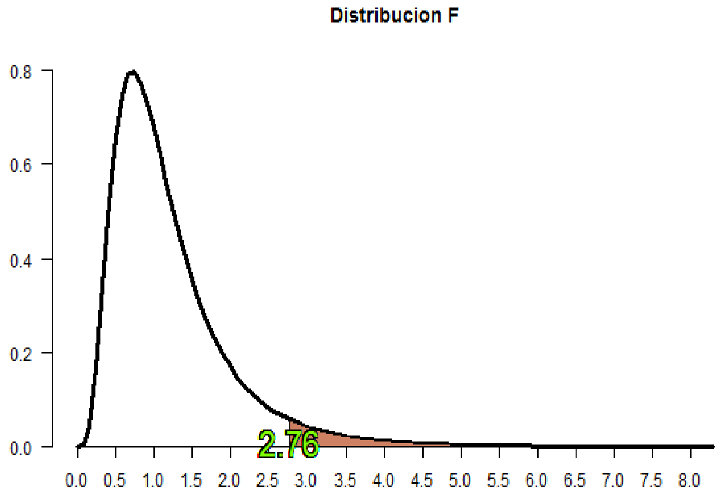
$$f_{\frac{\alpha}{2}, n_1-1, n_2-1} = f_{0.05,11,13=2.634}$$



$P(F > 2.63465) = 0.05$

Limite superior:

$$f_{\frac{\alpha}{2}, n_2-1, n_1-1} = f_{0.05,13,11=2.7614}$$



$P(F > 2.761417) = 0.05$

Reemplazando en la fórmula:

$$\frac{S_1{}^2}{S_2{}^2} * \frac{1}{f_{\frac{\alpha}{2},n_1-1,n_2-1}} < \frac{\boldsymbol{\sigma}_1{}^{\mathbf{2}}}{\boldsymbol{\sigma}_{\mathbf{2}}{}^{\mathbf{2}}} < \frac{S_1{}^2}{S_2{}^2} * f_{\frac{\alpha}{2},n_2-1,n_1-1}$$

$$\frac{22.24}{19.15} * \frac{1}{2.634} < \frac{\boldsymbol{\sigma}_1{}^{\mathbf{2}}}{\boldsymbol{\sigma}_{\mathbf{2}}{}^{\mathbf{2}}} < \frac{22.24}{19.15} * 2.76$$

$$0.45 < \frac{\boldsymbol{\sigma}_1{}^{\mathbf{2}}}{\boldsymbol{\sigma}_{\mathbf{2}}{}^{\mathbf{2}}} < 3.2$$

El intervalo contiene al 1 por lo tanto existe evidencia para afirmar que hay igualdad entre las varianzas.

Otro ejemplo

Un fabricante de automóviles pone a prueba dos nuevos métodos de ensamblaje de motores respecto al tiempo en minutos. Los resultados se muestran en la siguiente tabla

Método	**1**	**2**
n	31	25
$\boldsymbol{S^2}$	50	24

Construya un intervalo de confianza del 90% para la razón de varianzas