INSTITUTO TECNOLÓGICO DE COSTA RICA
ESCUELA DE MATEMÁTICA
MA 3405 ESTADÍSTICA
II SEMESTRE 2025
Profesora: Mag. Nathalia Saborío Cordero

TEC | Tecnológico de Costa Rica

ESTIMACIÓN CON DOS POBLACIONES

Estudiaremos a continuación herramientas para comparar dos poblaciones. Dicha comparación puede darse para estimar por intervalo la diferencia entre dos medias poblacionales o bien, para probar hipótesis acerca de si dos medias poblacionales son iguales.

Cuando se comparan dos poblaciones es posible dar respuesta a preguntas como las siguientes:

- ¿Cuál es la diferencia (si la hay) entre la durabilidad promedio de los zapatos de una marca con respecto a los zapatos de otra marca?
- ¿Hay alguna diferencia entre la proporción de unidades defectuosas producidas por un método y las producidas por otro método alternativo?
- ¿Los trabajadores de una planta producen en promedio más que los trabajadores de otra planta?
- Muchas otras…

Varias de las pruebas estadísticas parten de la suposición de que las varianzas poblacionales son iguales. Para poder demostrar esa suposición es necesario utilizar la **distribución F**, llamada así en honor a Ronald A. Fisher.

1. RAZÓN ENTRE DOS VARIANZAS

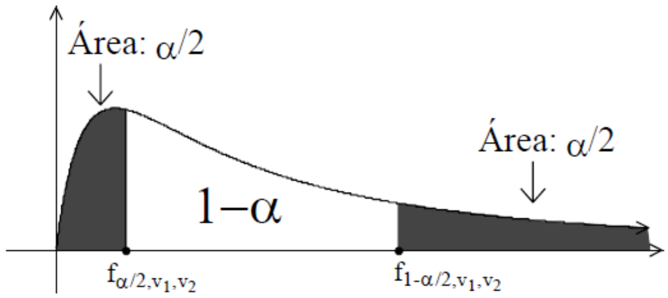
Teorema:

Considere las poblaciones dadas por las variables aleatorias X, Y que siguen una distribución normal con varianzas poblacionales de σ_1^2 y σ_2^2 respectivamente. Sean S_1^2 y S_2^2 varianzas de muestras aleatorias independientes de tamaño n_1 y n_2, tomadas de cada población respectivamente. Se tiene que

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

sigue una distribución F con $v_1 = n_1 - 1$ y $v_2 = n_2 - 1$ grados de libertad.

Veamos la gráfica de esta distribución:



Se cumple que $f_{\frac{\alpha}{2},n_1-1,n_2-1}$ deja un área de $\frac{\alpha}{2}$ a la izquierda y además $f_{1-\frac{\alpha}{2},n_1-1,n_2-1}$ deja un área de $\frac{\alpha}{2}$ a la derecha.

Teorema:

Considere las poblaciones 1 y 2 que siguen una distribución normal con varianzas poblacionales de σ_1^2 y σ_2^2 respectivamente. Sean s_1^2 y s_2^2 las varianzas de muestras aleatorias independientes de tamaño n_1 y n_2 tomadas de cada población respectivamente. Un intervalo de confianza del $100(1-\alpha)\%$ para $\frac{\sigma_2^2}{\sigma_1^2}$ está dado por

$$\left]\frac{s_2^2}{s_1^2 f_{1-\frac{\alpha}{2},n_2-1,n_1-1}}, \frac{s_2^2\, f_{1-\frac{\alpha}{2},n_1-1,n_2-1}}{s_1^2}\right[$$

Ejercicios:

1. Un curso es impartido tradicionalmente por 2 profesores A y B. Se tiene que 15 de los estudiantes del profesor A tienen una nota final promedio de 67.1 con una desviación estándar de 15, y 21 estudiantes del profesor B tienen un promedio de 62.8 con una desviación estándar de 18. Suponga que ambas notas siguen una distribución normal.
   a) Encuentre un intervalo de confianza de 90% para el cociente de las desviaciones estándar de las notas obtenidas por ambos profesores.
   b) ¿Puede suponerse que las varianzas son iguales? ¿Por qué?

2. Según un estudio acerca de la cantidad de tiempo que pasan conectados a internet, por mes, adultos y jóvenes, en el año 2000, se concluyó que, en promedio, los adultos pasan más tiempo conectados a internet que los jóvenes. Para confirmar esto, se realiza otro estudio para el que se toma una muestra de 26 adultos y otra de 30 jóvenes. Las desviaciones estándar de las cantidades de tiempo que pasan conectados a internet son 94 y 58 minutos, respectivamente. Use una confiabilidad del 95% y calcule un intervalo de confianza para el cociente de varianzas.

3. El Instituto del Consumidor desea comparar la variabilidad en la cantidad de cierto componente químico de un medicamento elaborado por las compañías A y B. Ambos medicamentos se distribuyen en forma de tabletas y contiene en promedio 250 mg del componente químico. En una muestra de 26 tabletas de cada compañía, se encontraron las desviaciones estándares de 1.25 y 1.18 en la cantidad del componente químico para cada compañía respectivamente. (Suponga que cada población se distribuye normalmente).

   A) Calcule un intervalo de confianza de 95% para $\frac{\sigma_A^2}{\sigma_B^2}$ $\quad$ *R/]0.503, 2.5025[*

PRUEBA DE HIPÓTESIS CON DOS VARIANZAS

Tomar en consideración la siguiente notación:

Población	Varianza poblacional	Tamaño de la muestra	Varianza muestral
X_1	σ_1^2	n_1	S_1^2
X_2	σ_2^2	n_2	S_2^2

Suponga además que las poblaciones X_1 y X_2 siguen una distribución normal y los estadísticos S_1^2 y S_2^2 son independientes. Entonces

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

Recuerde también que F tiene una distribución con $\nu_1 = n_1 - 1$ y $\nu_2 = n_2 - 1$ grados de libertad. Para pruebas de dos varianzas, la hipótesis nula es de la forma

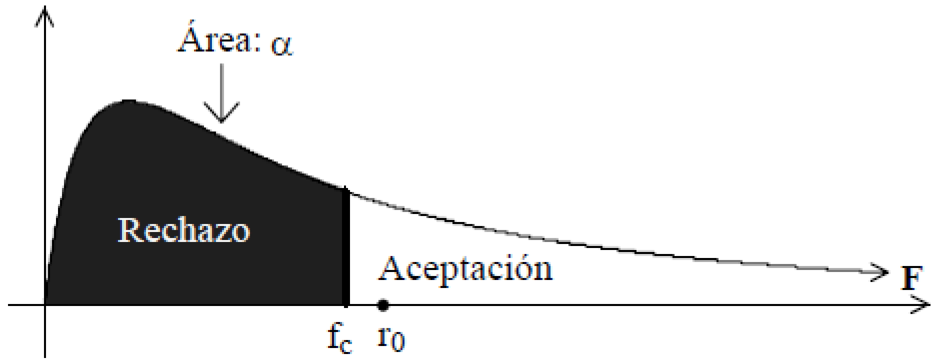
$$H_0\colon \frac{\sigma_1^2}{\sigma_2^2} = r_0$$

Donde r_0 se conoce como razón nula. Entonces, bajo H_0, se tiene que

$$F = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} = \frac{S_1^2}{r_0 S_2^2} \sim f\,(n_1 - 1, n_2 - 1)$$

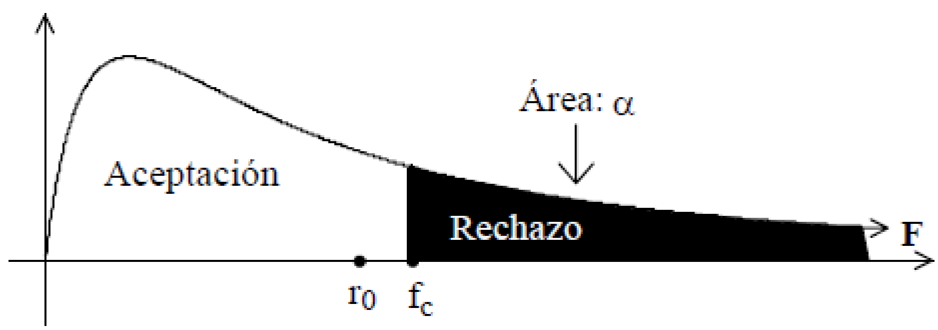
Analicemos los distintos casos para la hipótesis alternativa (H_1):

1. Si $H_1 : \dfrac{\sigma_1^2}{\sigma_2^2} < r_0$ entonces



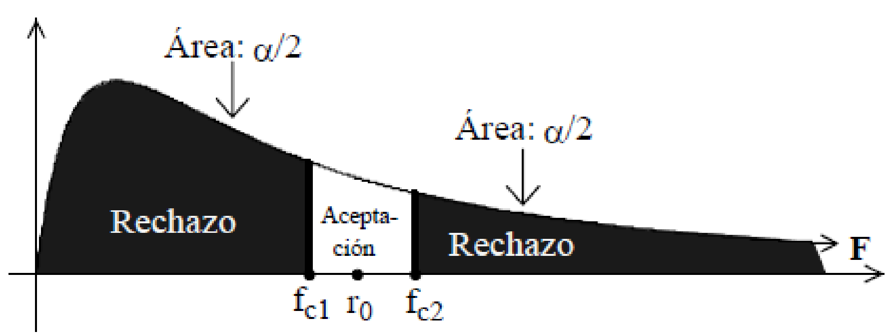
donde $f_c = f_{\alpha, n_1 - 1, n_2 - 1}$ y el Valor P es $P\left(F < f_{obs}\right)$ con $f_{obs} = \dfrac{s_1^2}{r_0 s_2^2}$.

2. Si $H_1 : \dfrac{\sigma_1^2}{\sigma_2^2} > r_0$ entonces



donde $f_c = f_{1-\alpha, n_1 - 1, n_2 - 1}$ y el Valor P es $P\left(F > f_{obs}\right)$.

3. Si $H_1 : \dfrac{\sigma_1^2}{\sigma_2^2} \neq r_0$ entonces



donde $f_{c1} = f_{\alpha/2, n_1 - 1, n_2 - 1}$, $f_{c2} = f_{1-\alpha/2, n_1 - 1, n_2 - 1}$. Dado que el Valor P sería la probabilidad de que F tome el valor f_{obs} o un valor extremo de forma que $F \neq r_0$, así si

(a) $f_{obs} > r_0$ entonces el valor P es aproximadamente[8] $2P\left(F > f_{obs}\right)$.

(b) $f_{obs} < r_0$ entonces el valor P es aproximadamente $2P\left(F < f_{obs}\right)$.

EJERCICIOS:

1. Un vendedor de accesorios de computadoras afirma que los discos duro modelo A tienen una vida útil promedio que supera a la vida útil promedio de los modelos B en más de 1 año. Para investigar se obtuvo la siguiente información donde $\bar{x}$ es la duración promedio muestral (en años) y s la desviación estándar muestral (en años):

Disco duro	Tamaño de muestra	$\bar{x}$	s
Modelo A	15	6.5	2.3
Modelo B	17	5.4	1.6

   a) Pruebe la hipótesis de que las varianzas de las duraciones de ambos modelos son iguales a un nivel de significancia del 5%.
   b) ¿Cuál es el valor P en la prueba anterior?

2. Un investigador desea comparar la influencia que tiene el uso de la computadora en un curso de estadística elemental para estudiantes de colegio; para ello se tomaron dos grupos. El grupo 1 tomó el curso asistido por computadora, mientras que el grupo 2 recibió el curso sin utilización de la computación. Al final del curso se aplicó un test a ambos grupos, obteniendo los siguientes resultados:

	Tamaño de muestra	$\bar{x}$	s
Grupo 1 (con computación)	11	60.3	3.8
Grupo 2 (sin computación)	15	67.2	2.1

   a) Pruebe la hipótesis de que las varianzas son iguales al nivel de significancia de 0.05.

   b) De acuerdo con el ejercicio "a", ¿indican los datos que el resultado de quienes tuvieron experiencia computacional fue significativamente menor que el de aquellos sin tal experiencia?

R/Valor P=0.044, se rechaza H_0

3. La oficina del Consumidor desea determinar si hay diferencia entre el peso promedio de dos tipos de bolsas de arroz A y B, que se venden con peso nominal de dos kilos. Para ello, realizó una inspección, obteniendo los siguientes resultados:

Tipo	Tamaño de muestra	Media muestral	Desviación muestral
A	49	2.05 kg	0.2 kg
B	21	1.97 kg	0.12 kg

   a) Pruebe la hipótesis de que las variancias son iguales al nivel de significancia de 0.05.

Texto base utilizado: Sanabria Brenes, Giovanni, 2011. **Comprendiendo la Estadística Inferencial**.