

UNIDAD V: ESTIMACIÓN DE PARÁMETROS Y PRUEBAS DE HIPOTESIS.

5.1 INTERVALOS DE CONFIANZA.

CONTENIDOS:

- 5.1.1. Intervalo de confianza y tamaño de muestra para una media.
- 5.1.2 Intervalo de confianza y tamaño de muestra para una proporción.
- 5.1.3 Intervalo de confianza para la varianza.
- 5.1.4 Intervalo de confianza para una diferencia de medias.
- 5.1.5 Intervalo de confianza para una diferencia de proporciones.
- 5.1.6 Intervalo de confianza para el cociente de varianzas.

OBJETIVOS:

- Estimar diferentes parámetros.
- Determinar un intervalo de confianza de nivel $1-\alpha$ para diferentes parámetros.
- Interpretar un intervalo de confianza expresando el significado del nivel de confianza.
- Redactar alguna conclusión con los resultados obtenidos del intervalo de confianza.
- Determinar el tamaño de muestra adecuado para estimar algún parámetro.

5.1.1 INTERVALO DE CONFIANZA Y TAMAÑO DE MUESTRA PARA UNA MEDIA.

CONCEPTOS CLAVES: Parámetro. Estimador. Error estándar de estimación. Error de estimación. Nivel de confianza. Tamaño de muestra.

RESUMEN DE CONCEPTOS Y PROPIEDADES:

1. Intervalo de confianza de nivel $1-\alpha$ para estimar la media en una población normal

$$\left[\bar{X} - d, \bar{X} + d \right] \quad ; \quad d = t_{\left(1-\frac{\alpha}{2}, n-1\right)} \sqrt{\frac{S_x^2}{n}}$$

2. Tamaño de muestra para estimar la media en la población con una confianza $1-\alpha$

$$n \geq \left(\frac{z_{1-\frac{\alpha}{2}}}{d_0} \right)^2 S_x^2 \quad ; \quad d_0 : \text{error de estimación máximo permitido}$$

EJERCICIO RESUELTO, PASO A PASO:

Ejercicio 1: (Aplicación en Ingeniería)

Se tomaron 12 muestras de mineral de una veta y se envían a un laboratorio para determinar el porcentaje de cobre soluble en cada una de ellas, obteniendo los siguientes resultados: 2.84 2.85 2.72 2.76 2.69 2.72 2.71 2.69 2.50 2.67 2.77 2.72

Pregunta 1: Determine un intervalo de 98% de confianza para estimar el porcentaje medio real de cobre soluble en la veta.

Pregunta 2: ¿Cuántas muestras se debe analizar para estimar el porcentaje medio de cobre soluble si se desea una confianza de 95% y un error máximo permitido de 0.03?

Esquema de solución

Paso 1: Leer cuidadosamente el enunciado del problema.

Paso 2: Identificar la variable en estudio y los parámetros involucrados.

Sea X = porcentaje de cobre soluble en muestras de mineral. En este caso se puede suponer que $X \sim N(\mu, \sigma^2)$ donde los parámetros involucrados son μ y σ^2 donde μ es el porcentaje medio de cobre y σ es la desviación estándar del porcentaje de cobre.

Paso 3: Estimar los parámetros. Se ingresan los datos a la calculadora fx – 350 MS o similar de la siguiente manera:

- Poner la calculadora en el modo SD
- Limpiar la memoria: SHIFT MODE 1 =
- Ingresar los datos de la siguiente manera:
2.84 M+ 2.85 M+2.72 M+
- Obtener resultados: SHIFT 2 1 = $\bar{X} = 2.72$ estima a μ
SHIFT 2 3 = $S_x = 0.0895$ estima a σ

Paso 4: Leer la pregunta 1 y revisar cual de los conceptos se debe usar para obtener lo pedido. En la pregunta 1 se pide un intervalo de 98% de confianza para el porcentaje medio de cobre soluble, luego el intervalo pedido es

$$[\bar{X} - d, \bar{X} + d] \quad ; \quad d = t_{\left(1-\frac{\alpha}{2}, n-1\right)} \sqrt{\frac{S_x^2}{n}}, \text{ en este caso la confianza es } 1-\alpha = 0.98,$$

luego $\alpha = 0.02 \Rightarrow \frac{\alpha}{2} = 0.01 \Rightarrow 1 - \frac{\alpha}{2} = 0.99$; en la tabla t de Student se encuentra

$$t_{\left(1-\frac{\alpha}{2}, n-1\right)} = t_{(0.99, 11)} = 2.72 \text{ y reemplazando se obtiene } d = 2.72 \frac{0.0895}{\sqrt{12}} = 0.07028$$

Luego un intervalo de 98% de confianza para el porcentaje medio de cobre soluble es:

$$[\bar{X} - d, \bar{X} + d] = [2.72 - 0.07028, 2.72 + 0.07028] = [2.6497, 2.7903]$$

Paso 5: Redactar una respuesta a la pregunta 1: Con 98% de confianza el verdadero porcentaje medio de cobre soluble en las muestras de mineral de la veta está contenido en el intervalo [2.6497, 2.7903]

Paso 6: Leer la pregunta 2 y revisar los conceptos a usar: En la pregunta 2 se pide un tamaño de muestra para estimar el porcentaje medio de cobre soluble con una confianza de 95% y un error máximo permitido de 0.03, luego el tamaño de muestra

$$\text{es } n \geq \left(\frac{z_{1-\frac{\alpha}{2}}}{d_0} \right)^2 S_X^2 \quad (*)$$

como la confianza es $1 - \alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975$;

luego en la tabla normal se obtiene $z_{0.975} = 1.96$ y reemplazando en (*) se obtiene

$$n \geq \left(\frac{1.96}{0.03} \right)^2 (0.0895)^2 = 34.1913 \approx 35$$

Paso 7: Redactar una respuesta a la pregunta 2: Es necesario analizar a lo menos 35 muestras de mineral para estimar el porcentaje medio de cobre soluble con una confianza de 95% y un error máximo permitido de 0.03%

EJERCICIOS PROPUESTOS:

EJERCICIOS PROPUESTOS:

1. **(Aplicación en Ciencias de la Salud)** Con el propósito de verificar la efectividad de un tratamiento basado en ejercicios para el aumento de la talla en niños de 10 años (en cms.), se realizó un experimento aplicando el tratamiento a 13 niños y considerando 16 como controles, el que arrojó los siguientes resultados:

TRATADOS $n = 13$, $\bar{X} = 138.6$, $S_x^2 = 29.16$

CONTROLES $m = 16$, $\bar{Y} = 125.8$, $S_y^2 = 20.7$

- Determine entre qué valores se encontrará la verdadera talla promedio de los niños tratados, si desea una confianza de 99%. Con este resultado ¿Podría asegurar que la talla promedio de los niños tratados en la población es de 145 cms.?
 - Determine entre qué valores se encontrará la verdadera talla promedio de los niños controles, si desea una confianza de 95%.
 - Si en un estudio posterior se desea estimar la talla promedio de los niños a los cuales se les aplica el tratamiento con una confianza de 95% y un error no superior a 2 cms.. ¿Cuántas observaciones se deberán hacer?
2. **(Aplicación en Ciencias de la Salud)** Los siguientes datos representan el aumento de peso (en Kilos.) de niños mediante dos dietas alternativas en la región de Antofagasta:

Dieta 1	0.98	0.62	0.53	0.76	0.43	0.36	0.51	0.24	0.24
Dieta 2	0.40	0.14	0.16	0.21	0.09	0.02	0.02	0.01	

- Determine un intervalo de 99% de confianza para el aumento de peso promedio en los niños con la dieta 1. Con esta información ¿Podría usted asegurar que el aumento de peso promedio con esta dieta es de 0.80 Kilos.?
- Determine un intervalo de 95% de confianza para el aumento de peso promedio en los niños con la dieta 2.
- Si se desea estimar el aumento de peso promedio en los niños con la dieta 1 con una confiabilidad de 95% y un error no mayor a 0.1 Kilos.. ¿Cuántas observaciones se deben hacer?

5.1.2 INTERVALO DE CONFIANZA Y TAMAÑO DE MUESTRA PARA UNA PROPORCIÓN.

CONCEPTOS CLAVES: Parámetro. Estimador. Error estándar de estimación. Error de estimación. Nivel de confianza. Tamaño de muestra.

RESUMEN DE CONCEPTOS Y PROPIEDADES:

1. Si A es una característica de interés y se desea estimar la probabilidad de A denotada como $p = P(A)$, entonces mediante una muestra de tamaño n se estima p como $\hat{p} = \frac{a}{n}$ donde a es el número de elementos de la muestra con la característica A , entonces un intervalo de confianza de nivel $1 - \alpha$ para estimar una proporción en una población binomial está dado por:

$$[\hat{p} - d, \hat{p} + d] \quad \text{donde} \quad d = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

2. Tamaño de muestra para estimar una proporción en la población

$$n \geq \left(\frac{z_{1-\frac{\alpha}{2}}}{d_0} \right)^2 \hat{p}(1-\hat{p}) \quad ; \quad d_0 : \text{Error de estimación máximo permitido}$$

EJERCICIO RESUELTO, PASO A PASO:

Ejercicio 2: (Aplicación en Ciencias de la salud)

En una universidad se desea estimar la proporción de estudiantes que fuma por lo menos dos cigarros al día. Se elige una muestra de 1500 estudiantes y se obtiene que 900 de ellos fuma al menos dos cigarros al día.

Pregunta 1: Determine un intervalo de 95% de confianza para estimar la verdadera proporción de estudiantes que fuma al menos dos cigarros al día.

Pregunta 2: ¿Cuántos estudiantes se debe encuestar para estimar la proporción de estudiantes que fuma al menos dos cigarros al día si se desea una confianza de 95% y un error de estimación máximo permitido de 0.05?

Esquema de solución

Paso 1: Leer cuidadosamente el enunciado del problema.

Paso 2: Identificar la característica de interés y los parámetros involucrados.

Sea $A = \{\text{Estudiantes de la universidad que fuma}\}$. En este caso el parámetro de interés es $p = P(A) =$ proporción de estudiantes que fuma al menos dos cigarros al día.

Paso 3: Estimar los parámetros.

En la muestra de $n = 1500$ estudiantes se tiene que $a = 900$, luego $\hat{p} = \frac{900}{1500} = 0.60$

Paso 4: Leer la pregunta 1 y revisar cual de los conceptos se debe usar para obtener lo pedido. En la pregunta 1 se pide un intervalo de 95% de confianza para la proporción de estudiantes que fuma al menos dos cigarros al día, luego el intervalo pedido

$$\text{es } [\hat{p} - d, \hat{p} + d] \quad ; \quad \hat{p} = \frac{\#A}{n} \quad ; \quad d = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

la confianza es $1 - \alpha = 0.95$, luego $\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975$.

$$\text{Reemplazando se obtiene } d = z_{0.975} \sqrt{\frac{0.60(0.40)}{1500}} = 1.96 \sqrt{0.00016} = 0.025.$$

Luego un intervalo de 95% de confianza para la proporción de estudiantes que fuma al menos dos cigarros al día en la universidad es:

$$[\hat{p} - d, \hat{p} + d] = [0.60 - 0.25, 0.60 + 0.25] = [0.575, 0.625]$$

Paso 5: Redactar una respuesta a la pregunta 1: Con 95% de confianza la verdadera proporción de estudiantes que fuma al menos dos cigarros al día en la universidad está contenido en el intervalo $[0.575, 0.625]$

Paso 6: Leer la pregunta 2 y revisar los conceptos a usar: En la pregunta 2 se pide un tamaño de muestra para estimar la proporción de estudiantes que fuma al menos dos cigarros con una confianza de 95% y un error de estimación máximo permitido de 0.05,

$$\text{luego el tamaño de muestra es } n \geq \left(\frac{z_{1-\frac{\alpha}{2}}}{d_0} \right)^2 \hat{p}(1-\hat{p}) \quad (*)$$

como la confianza es $1 - \alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975$;

luego en la tabla normal se obtiene $z_{0.975} = 1.96$ y reemplazando en (*) se obtiene

$$n \geq \left(\frac{1.96}{0.05} \right)^2 (0.60)(0.40) = 368.79 \approx 369$$

Paso 7: Redactar una respuesta a la pregunta 2: Se debe encuestar 369 estudiantes o más para estimar la proporción de estudiantes que fuma en la universidad.

EJERCICIOS PROPUESTOS:

1. **(Aplicación en Ciencias de la Salud)** Con el propósito de verificar la efectividad de un tratamiento basado en ejercicios para bajar de peso en niños con obesidad, se realizó un experimento aplicando el tratamiento a 300 niños. Después de un tiempo se observó que 100 bajaron de peso.
 - i. Determine entre qué valores se encontrará la verdadera proporción de niños que bajan de peso con el tratamiento, si desea una confianza de 95%. Con este resultado ¿Podría asegurar que la proporción de efectividad del tratamiento en la población es de 0.3?
 - ii. Si en un estudio posterior se desea estimar la proporción de efectividad del tratamiento con una confianza de 95% y un error no superior a 0.05. ¿Cuántos niños se deberá seleccionar para aplicarles el tratamiento?
2. **(Aplicación en Ciencias del Mar)** Se desea comparar la efectividad de dos análisis de laboratorio para detectar la presencia de bacterias en equinodermos (erizo rojo), para ello se selecciona dos muestras independientes de *Loxechinus albus* (erizo rojo comestible) en un sector de caleta coloso y los resultados de los análisis fueron los siguientes:

Análisis	Detecta la Bacteria		Total
	Si	No	
1	11	39	50
2	8	42	50
Total	19	81	100

- a) ¿Entre qué valores se encuentra la verdadera proporción de efectividad en detectar la bacteria del análisis 1 con una confiabilidad del 90%?
- b) Si se desea estimar la proporción de efectividad en detectar la bacteria del análisis 2 con una confiabilidad de 95% y un error en la estimación no superior a 0.05. ¿Cuántos especímenes de *Loxechinus albus* se deben seleccionar para hacerles el análisis?

5.1.3 INTERVALO DE CONFIANZA PARA LA VARIANZA DE UNA POBLACIÓN NORMAL.

CONCEPTOS CLAVES: Parámetro. Estimador. Nivel de confianza. Intervalo de confianza

RESUMEN DE CONCEPTOS Y PROPIEDADES:

1. Estimador puntual de la varianza σ^2

$$\hat{\sigma}^2 = S_X^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

2. Intervalo de confianza de nivel $1-\alpha$ para estimar la varianza en una población normal

$$\left[\frac{(n-1)S_X^2}{b}, \frac{(n-1)S_X^2}{a} \right] ; \quad a = \chi^2_{\left(\frac{\alpha}{2}, n-1\right)} ; \quad b = \chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}$$

EJERCICIO RESUELTO, PASO A PASO:

Ejercicio 3: (Aplicación en Ingeniería)

El tiempo necesario (en minutos) en realizar una tarea específica fue observada en una muestra de 10 trabajadores de una fábrica, obteniendo los siguientes resultados:

15.2 15.5 14.2 15.6 14.8 15.2 15.1 14.1 14.7 14.6

Pregunta: Determine un intervalo de 95% de confianza para estimar la varianza del tiempo en realizar la tarea en la fábrica.

Esquema de solución

Paso 1: Leer cuidadosamente el enunciado del problema.

Paso 2: Identificar la variable en estudio y los parámetros involucrados.

Sea X = tiempo necesario en realizar la tarea. En este caso se puede suponer que $X \sim N(\mu, \sigma^2)$ donde los parámetros involucrados son μ y σ^2 donde μ es el tiempo promedio en realizarla tarea y σ es la desviación estándar del tiempo en hacer la tarea.

Paso 3: Estimar los parámetros. Se ingresan los datos a la calculadora fx – 350 MS o similar de la siguiente manera:

- Poner la calculadora en el modo SD
- Limpiar la memoria: SHIFT MODE 1 =
- Ingresar los datos de la siguiente manera:
15.2 M+ 15.5M+14.6 M+
- Obtener resultados SHIFT 2 3 = $S_x = 0.5099$ estima a σ

Paso 4: Leer la pregunta y revisar cual de los conceptos se debe usar para obtener lo pedido. En la pregunta se pide un intervalo de 95% de confianza para la varianza del porcentaje de cobre soluble, luego el intervalo pedido es

$$\left[\frac{(n-1)S_x^2}{b}, \frac{(n-1)S_x^2}{a} \right] ; \quad a = \chi^2_{\left(\frac{\alpha}{2}, n-1\right)} ; \quad b = \chi^2_{\left(1-\frac{\alpha}{2}, n-1\right)}, \quad \text{en este caso la}$$

confianza es $1 - \alpha = 0.95$, luego $\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975$;

Luego en la tabla Chi - cuadrado se obtiene

$$a = \chi^2_{(0.025, 9)} = 2.70 ; \quad b = \chi^2_{(0.975, 9)} = 19.0$$

Reemplazando se obtiene que un intervalo de 95% de confianza para la varianza del tiempo en realizar la tarea por los trabajadores de la fábrica es:

$$\left[\frac{(n-1)S_x^2}{b}, \frac{(n-1)S_x^2}{a} \right] = \left[\frac{9(0.5099)^2}{19.0}, \frac{9(0.5099)^2}{2.70} \right] = [0.1232, 0.8667]$$

Paso 5: Redactar una respuesta a la pregunta: Con 95% de confianza la varianza del tiempo en realizar la tarea por los trabajadores de la fábrica está contenido en el intervalo $[0.1232, 0.8667]$

EJERCICIOS PROPUESTOS:

1. **(Aplicación en Ciencias de la Salud)** Con el propósito de verificar la efectividad de un tratamiento basado en ejercicios para el aumento de la talla en niños de 10 años (en cms.), se realizó un experimento aplicando el tratamiento a 13 niños y considerando 16 como controles, el que arrojó los siguientes resultados:

$$\text{TRATADOS} \quad n_1 = 13, \bar{x}_1 = 138.6, S_1^2 = 29.16$$

$$\text{CONTROLES} \quad n_2 = 16, \bar{x}_2 = 125.8, S_2^2 = 20.7$$

- Determine entre qué valores se encontrará la verdadera varianza de la talla de los niños tratados, si desea una confianza de 99%. Con este resultado ¿Podría asegurar que la varianza de la talla de los niños tratados en la población es igual a 35?
- Determine entre qué valores se encontrará la verdadera varianza de la talla de los niños controles, si desea una confianza de 95%.

3. **(Aplicación en Ciencias de la Salud)** Los siguientes datos representan el aumento de peso (en Kilos.) de niños mediante dos dietas alternativas en la región de Antofagasta:

Dieta 1	0.98	0.62	0.53	0.76	0.43	0.36	0.51	0.24	0.24
Dieta 2	0.40	0.14	0.16	0.21	0.09	0.02	0.02	0.01	

- Determine un intervalo de 99% de confianza para la varianza del aumento de peso en los niños con la dieta 1.
- Determine un intervalo de 95% de confianza para la varianza del aumento de peso en los niños con la dieta 2.

5.1.4 INTERVALO DE CONFIANZA PARA ESTIMAR LA DIFERENCIA DE MEDIAS.

CONCEPTOS CLAVES: Parámetro. Estimador. Error estándar de estimación. Error de estimación. Nivel de confianza.

RESUMEN DE CONCEPTOS Y PROPIEDADES:

1. Sea $X \sim N(\mu_1, \sigma^2)$; $Y \sim N(\mu_2, \sigma^2)$ dos poblaciones independientes.
2. En dos muestra de tamaños n y m de X e Y respectivamente se obtiene $\bar{X}, \bar{Y}, S_X$ y S_Y .
3. Parámetros a estimar: $\mu_1 - \mu_2$ y σ^2
4. Estimadores: $\hat{\mu}_1 - \hat{\mu}_2 = \bar{X} - \bar{Y}$ y $\hat{\sigma}^2 = S_C^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$
5. Intervalo de confianza de nivel $1 - \alpha$ para estimar la diferencia de medias en la población

$$\left[(\bar{X} - \bar{Y}) - d, (\bar{X} - \bar{Y}) + d \right] \quad ; \quad d = t_{\left(1-\frac{\alpha}{2}, n+m-2\right)} \sqrt{S_C^2 \left(\frac{1}{n} + \frac{1}{m} \right)}$$

EJERCICIO RESUELTO, PASO A PASO:

Ejercicio 4: (Aplicación en Ciencias de la salud)

Se desea analizar el contenido de vitamina A en la sangre en trabajadores a nivel del mar y en altura obteniendo los siguientes datos:

Nivel del mar: 25.2 30.4 46.9 51 46.4 48.5 39.3 55.9 34.3 31.2 40.7 29.8 35.7 40.1

En altura : 43.7 62.6 61.6 74.8 36.8 68.6 69.3 67 44 49 56.8 48.4 42.4 47.1

Pregunta: Determine un intervalo de 95% de confianza para estimar la diferencia de los contenidos medios de vitamina A en la sangre entre trabajadores a nivel del mar y en altura.

Esquema de solución

Paso 1: Leer cuidadosamente el enunciado del problema.

Paso 2: Identificar la variable en estudio y los parámetros involucrados.

Sea X = contenido de vitamina A en la sangre en trabajadores a nivel del mar.

Y = contenido de vitamina A en la sangre en trabajadores en altura. Suponer que

$$X \sim N(\mu_1, \sigma^2) \quad ; \quad Y \sim N(\mu_2, \sigma^2)$$

Paso 3: Estimar los parámetros. Se ingresan los datos a la calculadora fx – 350 MS (o similar) de igual manera que en el ejercicio 1 y obtener:

$$\bar{X} = 39.6714 \quad y \quad S_x = 0.5099$$

$$\bar{Y} = 55.15 \quad y \quad S_y = 12.1224$$

$$\hat{\mu}_1 - \hat{\mu}_2 = \bar{X} - \bar{Y} = 39.6714 - 55.15 = -15.4786$$

Estimadores:

$$\hat{\sigma}^2 = S_c^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2} = 114.9294$$

Paso 4: Leer la pregunta y revisar cual de los conceptos se debe usar para obtener lo pedido. En la pregunta se pide un intervalo de 95% de confianza para la diferencia de los contenidos medios entre trabajadores a nivel del mar y altura, luego el intervalo pedido es

$$\left[(\bar{X} - \bar{Y}) - d, (\bar{X} - \bar{Y}) + d \right] \quad \text{con} \quad d = t_{(1-\frac{\alpha}{2}, n+m-2)} \sqrt{S_c^2 \left(\frac{1}{n} + \frac{1}{m} \right)},$$

En este caso la confianza es $1 - \alpha = 0.95$, luego

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975;$$

Luego en la tabla t de Student se obtiene $t_{1-\frac{\alpha}{2}} = t_{0.975}(26) = 2.06$

$$\text{luego } d = t_{(1-\frac{\alpha}{2}, n+m-2)} \sqrt{S_c^2 \left(\frac{1}{n} + \frac{1}{m} \right)} = 2.06 \sqrt{114.9294 \left(\frac{1}{14} + \frac{1}{14} \right)} = 8.3288$$

Reemplazando se obtiene que un intervalo de 95% de confianza para la diferencia de los contenidos medios entre trabajadores a nivel del mar y altura es:

$$[-15.4786 - 8.3288, -15.4786 + 8.3288] = [-23.8074, -7.1497]$$

Paso 5: Redactar una respuesta a la pregunta: Con 95% de confianza la diferencia de los contenidos medios entre trabajadores a nivel del mar y altura está contenido en el intervalo $[-23.8074, -7.1497]$

EJERCICIOS PROPUESTOS:

1. **(Aplicación en Ciencias de la Ingeniería)** Los siguientes datos corresponden a la pureza del ácido en porcentaje observada en dos plantas de producción de ácido sulfúrico.

PLANTA A 96.8 95.8 96.2 96.5 96.4 96.7 96.3 96 97.1 96.5

PLANTA B 95.6 96 97.8 98.4 97.6 98.2 96.9 96.8

Determine entre qué valores se encontrará la verdadera diferencia media de los porcentajes promedios de pureza del ácido entre la planta A y la planta B, si desea una confianza de 99%. Con este resultado ¿Podría asegurar que no hay diferencia en el porcentaje promedio de pureza del ácido entre las dos plantas? ¿Por qué?

2. **(Aplicación en Ciencias del mar)** Los siguientes datos representan el crecimiento (en cms.) de *Protothaca Thaca* (Almejas) en dos sectores costeros de la región de Antofagasta:

BOLSICO	0.98	0.62	0.53	0.76	0.43	0.36	0.51	0.24	0.24
COLOSO	0.40	0.14	0.16	0.21	0.09	0.02	0.02	0.01	

Determine un intervalo de 95% de confianza para la diferencia de los crecimientos promedio entre Bolsico y Coloso.

3. **(Aplicación en Ciencias de la Salud)** Se tomaron dos muestras de presión sistólica (en mm Hg.) a sujetos normales y sujetos hospitalizados, obteniendo la siguiente información:

NORMALES (X) : 146 142 135 140 154 163 138 168

HOSPITALIZADOS (Y) : 164 176 165 172 169 171

Determine entre qué valores se encuentra la diferencia de la presión sistólica media entre sujetos normales y hospitalizados con una confiabilidad de 95%.

5.1.5 INTERVALO DE CONFIANZA PARA ESTIMAR LA DIFERENCIA DE PROPORCIONES.

CONCEPTOS CLAVES: Parámetro. Estimador. Error estándar de estimación. Error de estimación. Nivel de confianza.

RESUMEN DE CONCEPTOS Y PROPIEDADES:

Si A es una característica de interés y se desea estimar la probabilidad de A en dos poblaciones independientes Pob 1 y Pob 2 denotadas como $p_1 = P(A / \text{Pob 1})$ y $p_2 = P(A / \text{Pob 2})$,

Mediante una muestra de tamaño n de la población 1 se estima p_1 como $\hat{p}_1 = \frac{a_1}{n}$ donde a_1 es el número de elementos de la muestra de la población 1 con la característica A.

Mediante una muestra de tamaño m de la población 2 se estima p_2 como $\hat{p}_2 = \frac{a_2}{m}$ donde a_2 es el número de elementos de la muestra de la población 2 con la característica A.

Luego un intervalo de confianza de nivel $1 - \alpha$ para estimar la diferencia de proporciones en la población está dado por:

$$[\hat{p}_1 - \hat{p}_2 - d, \hat{p}_1 - \hat{p}_2 + d] \quad \text{con} \quad d = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{m}}$$

EJERCICIO RESUELTO, PASO A PASO:

Ejercicio 5: (Aplicación en Ciencias de la salud)

En un estudio para investigar los efectos secundarios de dos medicamentos, a 50 pacientes se les dio un medicamento X y a otros 50 un medicamento Y. De los que recibieron el medicamento X, 11 mostraron efectos secundarios, mientras que de los que recibieron el medicamento Y, 8 mostraron efectos secundarios.

Pregunta: Determine un intervalo de 95% de confianza para estimar la diferencia de proporciones de efectos secundarios entre los medicamentos X y Y..

Esquema de solución

Paso 1: Leer cuidadosamente el enunciado del problema.

Paso 2: Identificar la característica de interés y los parámetros involucrados.

Sean $A = \{\text{Pacientes con efectos secundarios}\}$.

$\text{Pob 1} = \{\text{Pacientes con medicamento X}\}$ y $\text{Pob 2} = \{\text{Pacientes con medicamento Y}\}$

Parámetros de interés:

$p_1 = P(A / Pob 1)$ = proporción de pacientes con efectos secundarios en el medicamento X .

$p_2 = P(A / Pob 2)$ = proporción de pacientes con efectos secundarios en el medicamento Y.

Paso 3: Estimar los parámetros.

En la muestra de $n = 50$ pacientes con medicamento X se tiene que $\#A = 11$, luego

$\hat{p}_1 = \frac{11}{50} = 0.22$ y en la muestra de $m = 50$ pacientes con medicamento Y se tiene que

$\#A = 8$, luego $\hat{p}_2 = \frac{8}{50} = 0.16$

Paso 4: Leer la pregunta y revisar cual de los conceptos se debe usar para obtener lo pedido. En la pregunta se pide un intervalo de 95% de confianza para la diferencia de proporciones de pacientes con efectos secundarios entre los medicamentos X e Y, luego el

intervalo pedido es, $[\hat{p}_1 - \hat{p}_2 - d, \hat{p}_1 - \hat{p}_2 + d]$ con $d = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{m}}$

la confianza es $1 - \alpha = 0.95$, luego $\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975$.

Luego $d = z_{0.975} \sqrt{\frac{0.22(0.78)}{50} + \frac{0.16(0.84)}{50}} = 1.96 \sqrt{0.00612} = 0.1533$.

Luego un intervalo de 95% de confianza para la diferencia de proporciones de pacientes con efectos secundarios entre los medicamentos X e Y es:

$$[0.22 - 0.16 - 0.1533, 0.22 - 0.16 + 0.1533] = [-0.0933, 0.2133]$$

Paso 5: Redactar una respuesta a la pregunta: Con 95% de confianza la diferencia de proporciones de pacientes con efectos secundarios entre los medicamentos X e Y está contenido en el intervalo $[-0.0933, 0.2133]$

EJERCICIOS PROPUESTOS:

1. **(Aplicación en Ciencias del Mar)** La siguiente tabla muestra la distribución de una muestra aleatoria de 400 truchas cafés de un gran río., según la longitud y el sector donde fueron extraídas.

SECTOR DEL RIO

LONGITUD	Alto	Centro	Bajo
Bajo el promedio	67	64	25
Promedio	42	76	56
Sobre el promedio	10	23	37

Verifique mediante un intervalo de 95% de confianza si la proporción de truchas de longitud bajo el promedio es mayor en el sector alto que en el sector bajo.

2. **(Aplicación en Ingeniería)** Una empresa se provee de componentes electrónicos provenientes de dos fábricas. El interés del dueño está centrado en el tiempo de vida útil de tales componentes. Al recibir una gran partida de cada fábrica, decide inspeccionar una muestra aleatoria de cada partida, las cuales somete a experimentación con el propósito de recavar los tiempos de vida útil. Los resultados obtenidos se muestran en la siguiente tabla:

F A B R I C A

VIDA UTIL (en horas)	1	2	TOTAL
0 - 300	4	6	
300 - 500	12	16	
500 - 1000	23	19	
1000 - 1500	6	4	
TOTAL			

¿Podría afirmar mediante un intervalo de 95% de confianza, que la proporción de componentes cuya vida útil es mayor que 500 horas es menor en la FABRICA 2 que en la FABRICA 1?

5.1.6 INTERVALO DE CONFIANZA PARA EL CUOCIENTE DE VARIANZAS.

CONCEPTOS CLAVES: Parámetro. Estimador. Error estándar de estimación. Nivel de confianza.

RESUMEN DE CONCEPTOS Y PROPIEDADES:

1. Sea $X \sim N(\mu_1, \sigma_1^2)$; $Y \sim N(\mu_2, \sigma_2^2)$ dos poblaciones independientes.
2. En dos muestra de tamaños n y m de X e Y respectivamente se obtiene $\bar{X}, \bar{Y}, S_X$ y S_Y .
3. Parámetro a estimar: $\frac{\sigma_2^2}{\sigma_1^2}$
4. Estimadores: $\hat{\sigma}_1^2 = S_X^2$; $\hat{\sigma}_2^2 = S_Y^2$
5. Intervalo de confianza de nivel $1 - \alpha$ para estimar el cuociente de varianzas

$$\left[a \frac{S_Y^2}{S_X^2}, b \frac{S_Y^2}{S_X^2} \right] \quad \text{donde} \quad a = F_{\left[\frac{\alpha}{2}, (n-1, m-1) \right]} \quad \text{y} \quad b = F_{\left[1 - \frac{\alpha}{2}, (n-1, m-1) \right]}$$

EJERCICIO RESUELTO, PASO A PASO:

Ejercicio 6: (Aplicación en Ciencias de la Ingeniería)

Un nuevo dispositivo de filtrado se instala en una planta química. Antes y después de su instalación una muestra aleatoria entrega la siguiente información del porcentaje de impurezas:

Antes de instalación : $n = 8$ $\bar{X} = 12.5$ $S_X^2 = 101.17$

Después de instalación: $m = 9$ $\bar{Y} = 12.5$ $S_Y^2 = 94.73$

Pregunta: Determine un intervalo de 95% de confianza para estimar el cuociente de varianzas del porcentaje de impurezas después y antes de la instalación del nuevo dispositivo. ¿Podría concluir que las varianzas son iguales?

Esquema de solución

Paso 1: Leer cuidadosamente el enunciado del problema.

Paso 2: Identificar la variable en estudio y los parámetros involucrados.

Sea X = porcentaje de impurezas antes de instalación del nuevo dispositivo.

Y = porcentaje de impurezas después de instalación del nuevo dispositivo.

Suponer que $X \sim N(\mu_1, \sigma_1^2)$; $Y \sim N(\mu_2, \sigma_2^2)$

Paso 3: Estimar los parámetros. De la información entregada por las muestras se obtiene que:

$$\hat{\sigma}_1^2 = S_X^2 = 101.17 \text{ y } \hat{\sigma}_2^2 = S_Y^2 = 94.73$$

Paso 4: Leer la pregunta y revisar cual de los conceptos se debe usar para obtener lo pedido. En la pregunta se pide un intervalo de 95% de confianza para el cuociente de varianzas del porcentaje de impurezas después y antes de la instalación del nuevo dispositivo, luego el intervalo pedido es

$$\left[a \frac{S_Y^2}{S_X^2}, b \frac{S_Y^2}{S_X^2} \right] \text{ donde } a = F_{\left[\frac{\alpha}{2}, (n-1, m-1) \right]} \text{ y } b = F_{\left[1-\frac{\alpha}{2}, (n-1, m-1) \right]}$$

En este caso la confianza es $1 - \alpha = 0.95$, luego

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975;$$

Luego en la tabla F de Fischer se obtiene:

$$a = F_{\left[\frac{\alpha}{2}, (n-1, m-1) \right]} = F_{[0.025, (7,8)]} = \frac{1}{F_{[0.975, (8,7)]}} = \frac{1}{4.90} = 0.204$$

$$b = F_{\left[1-\frac{\alpha}{2}, (n-1, m-1) \right]} = F_{[0.975, (7,8)]} = 4.52$$

Reemplazando se obtiene que el intervalo de confianza pedido es:

$$\left[0.204 \frac{94.73}{101.17}, 4.52 \frac{94.73}{101.17} \right] = [0.1911, 4.2403]$$

Paso 5: Redactar una respuesta a la pregunta: Con 95% de confianza el cuociente de varianzas del porcentaje de impurezas después y antes de la instalación del nuevo dispositivo de filtrado está contenido en el intervalo [0.1911, 4.2403]. Como el 1 pertenece al intervalo podemos concluir con un 95% de confianza que las varianzas poblacionales son iguales.

EJERCICIOS PROPUESTOS:

1. **(Aplicación en Ciencias de la Ingeniería)** Los siguientes datos corresponden a la pureza del ácido en porcentaje observada en dos plantas de producción de ácido sulfúrico.

PLANTA A 96.8 95.8 96.2 96.5 96.4 96.7 96.3 96 97.1 96.5

PLANTA B 95.6 96 97.8 98.4 97.6 98.2 96.9 96.8

Determine un intervalo de 95% de confianza para el cuociente de las varianzas de los porcentajes de pureza del ácido entre la planta B y la planta A. Con este resultado ¿Podría asegurar que las varianzas del porcentaje de pureza del ácido entre las dos plantas es la misma? ¿Por qué?

2. **(Aplicación en Ciencias del mar)** Los siguientes datos representan el crecimiento (en cms.) de *Protothaca Thaca* (Almejas) en dos sectores costeros de la región de Antofagasta:

BOLSICO	0.98	0.62	0.53	0.76	0.43	0.36	0.51	0.24	0.24
COLOSO	0.40	0.14	0.16	0.21	0.09	0.02	0.02	0.01	

Determine un intervalo de 95% de confianza para el cuociente de las varianzas del crecimiento de las almejas entre Coloso y Bolsico.

3. **(Aplicación en Ciencias de la Salud)** Se tomaron dos muestras de presión sistólica (en mm Hg.) a sujetos normales y sujetos hospitalizados, obteniendo la siguiente información:

NORMALES (X) : 146 142 135 140 154 163 138 168

HOSPITALIZADOS (Y) : 164 176 165 172 169 171

Determine un intervalo de 95% de confianza para el cuociente de las varianzas de la presión sistólica entre sujetos normales y hospitalizados. Con este resultado ¿Podría asegurar que las varianzas de la presión sistólica entre los dos grupos de pacientes es la misma? ¿Por qué?