Taller de Análisis y diseño de experimentos

Tamaños de muestra

1. Para cada una de las siguientes encuestas, describa la población objetivo, el marco de muestreo, la unidad de muestreo y la unidad de observación.

a) Un estudiante desea estimar el porcentaje de fondos cuyas acciones aumentaron de precio la semana pasada. Usando una lista que salió en el periódico Portafolio, el alumno eligió uno de cada 10 fondos y con estos cálculo el porcentaje de aquellos donde el precio de la acción había aumentado.

b) Se extrae una muestra de 8 arquitectos de una ciudad con 45 arquitectos. Para formar la muestra de la encuesta, cada arquitecto fue contactado por teléfono por orden de aparición en el directorio telefónico. Los primeros ocho que acordaron ser entrevistados conformaron la muestra.

c) Para estimar cuantos libros de la biblioteca deben ser encuadernados de nuevo, un bibliotecario utiliza una tabla de números aleatorios para elegir al azar 100 posiciones en los estantes de la biblioteca. Luego, camina hasta cada posición, busca el libro que se encuentra en ese punto y registra si el libro debe encuadernarse o no.

d) Se realiza un estudio para determinar el peso promedio de las vacas en una región del país. Usando una lista de las granjas de la región, se eligen 50 de ellas. Luego se registra el peso de cada vaca de las 50 granjas elegidas.

e) En un juicio sobre marcas registradas, un demandante que afirme que otra compañía infringe sus marcas debe mostrar con frecuencia que las marcas tienen un significado secundario en el mercado; es decir, los usuarios potenciales del producto asocian las marcas registradas con el demandante aun cuando no esté presente el nombre de la compañía. Estos estudios son denominados de confusión de marca. Un estudio de confusión de marca entre los empaques de Something Special y Haig Supreme aplicado con una encuesta realizada a 500 personas consumidoras de whisky, mostró que cuando se mostraba la botella de Haig Supreme (Sin etiqueta) más del 50 % de los consumidores lo asociaron con Something Special.

Ejemplos:

- Un biólogo quiere estimar el peso promedio de los ciervos cazados en el estado de Maryland. Un estudio anterior de diez ciervos cazados mostró que la desviación estándar de sus pesos es de 12.2 libras. ¿Qué tan grande debe ser una muestra para que el biólogo tenga el 95% de confianza de que el error de estimación es a lo más de 4 libras?

Solución:

$$n = \left(\frac{z\,\sigma}{\varepsilon}\right)^2 = \left(\frac{(1.96)(12.2)}{4}\right)^2 = 35.736$$

En consecuencia, si el tamaño de la muestra es 36, se puede tener un 95% de confianza en que μ difiere en menos de 4 libras de $\bar{x}$.

- Una empresa eléctrica fabrica focos que tienen una duración aproximadamente normal con una desviación estándar de 40 horas. ¿De qué tamaño se necesita una muestra si se desea tener 96% de confianza que la media real esté dentro de 10 horas de la media real?

$$n = \left(\frac{z\sigma}{\varepsilon}\right)^2 = \left(\frac{(2.053)(40)}{10}\right)^2 = 67.43$$

Se necesita una muestra de 68 focos para estimar la media de la población y tener un error máximo de 10 horas.

- ¿Qué pasaría si en lugar de tener un error de estimación de 10 horas sólo se requiere un error de 5 horas?

$$n = \left(\frac{z\sigma}{\varepsilon}\right)^2 = \left(\frac{(2.053)(40)}{5}\right)^2 = 269.74$$

Se puede observar como el tamaño de la muestra aumenta, pero esto tiene como beneficio una estimación más exacta.

- Suponga que en el ejercicio anterior se tiene una población de 300 focos, y se desea saber de qué tamaño debe de ser la muestra. El muestreo se realizará sin reemplazo.

Solución:

Como se tiene una población finita y un muestreo sin reemplazo es necesario utilizar la formula con el factor de corrección.

$$n = \frac{z^2\sigma^2 N}{\varepsilon^2(N-1)+z^2\sigma^2} = \frac{(2.053)^2(40)^2(300)}{(10)^2(300-1)+(2.053)^2(40)^2} = 55.21$$

Si se tiene una población finita de 300 focos sólo se tiene que extraer de la población una muestra sin reemplazo de 56 focos para poder estimar la duración media de los focos restantes con un error máximo de 10 horas.

Cálculo del Tamaño de la Muestra para Estimar una Proporción

- En una muestra aleatoria de 500 familias que tienen televisores en la ciudad de Hamilton, Canadá, se encuentra que 340 están suscritas a HBO. ¿Qué tan grande se requiere que sea una muestra si se quiere tener 95% de confianza de que la estimación de P esté dentro de 0.02?

  Solución:

Se tratarán a las 500 familias como una muestra preliminar que proporciona una estimación de p=340/500=0.68.

$$n = \frac{z^2 pq}{e^2} = \frac{(1.96)^2 (0.68)(0.32)}{(0.02)^2} = 2090$$

Por lo tanto si basamos nuestra estimación de P sobre una muestra aleatoria de tamaño 2090, se puede tener una confianza de 95% de que nuestra proporción muestral no diferirá de la proporción real por más de 0.02.

- Una legisladora estatal desea encuestar a los residentes de su distrito para conocer qué proporción del electorado conoce la opinión de ella, respecto al uso de fondos estatales para pagar abortos. ¿Qué tamaño de muestra se necesita si se requiere un confianza del 95% y un error máximo de estimación de 0.10?

Solución:

En este problema, se desconoce totalmente la proporción de residentes que conoce la opinión de la legisladora, por lo que se utilizará un valor de 0.5 para *p.*

$$n = \frac{z^2 pq}{e^2} = \frac{(1.96)^2 (0.50)(0.50)}{(0.10)^2} = 96.04$$

- Una compañía de productos alimenticios contrató a una empresa de investigación de mercadotecnia , para muestrear dos mercados, I y II, a fin de comparar las proporciones de consumidores que prefieren la comida congelada de la compañía con los productos de sus competidores. No hay información previa acerca de la magnitud de las proporciones P_1 y P_2. Si la empresa de productos alimenticios quiere estimar la diferencia dentro de 0.04, con una probabilidad de 0.95, ¿ cuántos consumidores habrá que muestrear en cada mercado?

$$n = \frac{z^2 (p_1 q_1 + p_2 q_2)}{e^2} = \frac{(1.96^2)[(0.5)(0.5) + (0.5)(0.5)]}{0.04^2} = 1200.5$$

Se tendrá que realizar encuestas a 1201 consumidores de cada mercado para tener una estimación con una confianza del 95% y un error máximo de 0.04.

Problemas propuestos

1. Se quiere estudiar la tasa de combustión de dos propelentes sólidos utilizados en los sistemas de escape de emergencia de aeroplanos. Se sabe que la tasa de combustión de los dos propelentes tiene aproximadamente la misma desviación estándar; esto es $\sigma_1 = \sigma_2 =$ 3 cm/s. ¿Qué tamaño de muestra debe utilizarse en cada población si se desea que el error en la estimación de la diferencia entre las medias de las tasas de combustión sea menor que 4 cm/s con una confianza del 99%?. n= 8

2. Una tienda de donas se interesa es estimar su volumen de ventas diarias. Supóngase que el valor de la desviación estándar es de $50. Si el volumen de ventas se encuentra aproximado por una distribución normal, a) ¿Cuál debe ser el tamaño de la muestra para que con una probabilidad de 0.95 la media muestral se encuentre a no más de $20 del verdadero volumen de ventas promedio? Respuesta [25]

b) Si no es posible suponer que la distribución es normal, obtener el tamaño necesario de la muestra para la pregunta a. Respuesta [125]

3. Se planea realizar un estudio de tiempos para estimar el tiempo medio de un trabajo, exacto dentro de 4 segundos y con una probabilidad de 0.90, para terminar un trabajo de montaje. Si la experiencia previa sugiere que σ = 16 seg. mide la variación en el tiempo de montaje entre un trabajador y otro al realizar una sola operación de montaje, ¿cuántos operarios habrá que incluir en la muestra? Rta/ n= 44

4. Se desea estimar el número medio de horas de uso continuo antes de que cierto tipo de computadora requiera una reparación inicial. Si podemos suponer que u=20 días, ¿De qué tamaño debe ser una muestra a fin de suponer con una confianza del 90% que la media muestral difiera a lo más 5 días? Respuesta [44]

5. Un ingeniero de control de calidad quiere estimar la fracción de elementos defectuosos en un gran lote de lámparas. Por la experiencia, cree que la fracción real de defectuosos tendría que andar alrededor de 0.2. ¿Qué tan grande tendría que seleccionar la muestra si se quiere estimar la fracción real, exacta dentro de 0.01, utilizando un nivel de confianza fe 95%? Rta/ n= 6147

6. Se tienen que seleccionar muestras aleatorias independientes de n_1=n_2=n observaciones de cada una de dos poblaciones binomiales, 1 y 2. Si se desea estimar la diferencia entre los dos parámetros binomiales, exacta dentro de 0.05, con una probabilidad de 0.98. ¿qué tan grande tendría que ser n?. No se tiene información anterior acerca de los valores P_1 y P_2, pero se quiere estar seguro de tener un número adecuado de observaciones en la muestra.

   Rta/ n= 1086

7. Mayr et al.(1994) tomaron una muestra mediante un diseño MAS de tamaño n = 240 niños con edad entre 2 y 6 años, ellos encontraron la siguiente distribución de frecuencias para la variable: edad en que los niños empezaron a caminar sin ayuda.

Edad (meses)	9	10	11	12	13	14	15	16	17	18	19	20
Número de niños	13	35	44	69	36	24	7	3	2	5	1	1

a) Construya un histograma de la distribución de la edad al comenzar a caminar. ¿Cree Ud. que la variable sigue una distribución normal? La distribución de la edad promedio en que los niños comienzan a caminar sigue una distribución normal? Explique.

b) Determine la media, el coeficiente de variación y un intervalo de confianza del 95% para la edad promedio en que los niños comienzan a caminar solos.

c) Suponga que en el 2009 se desea hacer otro estudio en una región diferente y se desea que el intervalo de confianza del 95% para la edad promedio en que los niños comienzan a caminar solos tenga un margen de error de 5 %. Basándose en la información auxiliar proporcionada ¿qué tamaño de muestra se necesitará usando un diseño MAS?

8. El director administrativo de un colegio desea usar la media de una muestra aleatoria para estimar la cantidad promedio de tiempo que tardan los alumnos en ir de una clase a la siguiente, y además quiere poder asegurar con una confianza del 99% que el error es a lo más de 0.25 minutos. Si se supone por experiencia ? =. 1.4 minutos, ¿ Qué tamaño debe tener la muestra? Respuesta [208]

Ejercicios de intervalos de confianza

1. Se llevan a cabo pruebas de resistencia a la tensión sobre dos diferentes clases de largueros de aluminio utilizados en la fabricación de alas de aeroplanos comerciales. De la experiencia pasada con el proceso de fabricación se supone que las desviaciones estándar de las resistencias a la tensión son conocidas. La desviación estándar del larguero 1 es de 1.0 Kg/mm^2 y la del larguero 2 es de 1.5 Kg/mm^2. Se sabe que el comportamiento de las resistencias a la tensión de las dos clases de largueros son aproximadamente normal. Se toma una muestra de 10 largueros del tipo 1 obteniéndose una media de 87.6 Kg/mm^2, y otra de tamaño 12 para el larguero 2 obteniéndose una media de 74.5 Kg/mm^2. Estime un intervalo de confianza del 90% para la diferencia en la resistencia a la tensión promedio.

   Rta/ $12.22 \leq \mu_1 - \mu_2 \leq 13.98$

2. Se seleccionaron dos muestras de 400 tubos electrónicos, de cada una de dos líneas de producción, A y B. De la línea A se obtuvieron 40 tubos defectuosos y de la B 80. Estime la diferencia real en las fracciones de defectuosos para las dos líneas, con un coeficiente de confianza de 0.90 e interprete los resultados. Rta/ $0.059 \leq P_B - P_A \leq 0.141$

3. Se probó una muestra aleatoria de 400 cinescopios de televisor y se encontraron 40 defectuosos. Estime el intervalo que contiene, con un coeficiente de confianza de 0.90, a la verdadera fracción de elementos defectuosos. Rta/ $0.07532 \leq P \leq 0.1246$

4. Una carta publicada en una revista indicaba lo siguiente: "He observado que en los últimos números que no hay ganadores del sur en los concursos. Ustedes siempre dicen que los ganadores se eligen al azar. significa esto que ustedes venden menos en el sur? ". En respuesta, los editores realizaron una muestra con un diseño MAS de 1000 datos de los últimos 10000 concursos y encontraron que 175 provenían del sur.

a. Determine un intervalo de confianza para los datos provenientes del sur.

b. De acuerdo con el reporte de censos de los editores, el 30.9% de los suscriptores vive en estados considerados del sur. >Hay alguna evidencia en el intervalo de confianza de que el porcentaje de premios entregados en el sur difiere del porcentaje de suscriptores que viven en esa región del país?

5. Suponga que se quiere estimar la producción media por hora, en un proceso que produce antibiótico. Se observa el proceso durante 100 períodos de una hora, seleccionados al azar y se obtiene una media de 34 onzas por hora con una desviación estándar de 3 onzas por hora. Estime la producción media por hora para el proceso, utilizando un nivel de confianza del 95%.

Rta/ $33.412 \leq \mu \leq 34.588$

6. El decano registró debidamente el porcentaje de calificaciones D y F otorgadas a los estudiantes por dos profesores universitarios de matemáticas. El profesor I alcanzó un 32%, contra un 21% para el profesor II, con 200 y 180 estudiantes, respectivamente. Estime la diferencia entre los porcentajes de calificaciones D y F otorgadas por los dos profesores. Utilice un nivel de confianza del 95% e interprete los resultados.

Rta/$0.0222 \leq P_1 - P_2 \leq 0.1978$

7. Se midió la resistencia a la ruptura por torcimiento de un cierto tipo de tela en un lote con los siguientes resultados (en psi): 182, 172, 176, 178. La desviación estándar basada en la experiencia previa es de 5 psi, Encuentre un intervalo de confianza del 99% para la resistencia promedio de la ruptura por torcimiento del lote.

Respuestas P[170.5625<□<183.4375]=0.99

8. Un fabricante de fibras sintéticas que desea estimar la tensión de ruptura media de una fibra. Diseña un experimento para observar las tensiones de ruptura en libras, de 1 6 hilos del proceso seleccionados azar. Las tensiones son 20.8, 20.6, 21.0, 20.9, 19.9, 20.2, 19.8, 19.6, 20.9,.21.1, 20.4, 20.6, 19.7, 19.6, 20.3 v 20.7. Supóngase que la tensión de ruptura de una fibra se encuentra modelada por una distribución normal con desviación estándar de 0.45 Libras. Construir un intervalo de confianza estimado del 98% para la tensión de ruptura promedio de la fibra. Respuesta [20.1196,20.6427]

9. Los siguientes datos representan medidas de porosidad en una muestra de un cargamento de coque. Encuentre un intervalo de confianza del 95% para la media verdadera. Suponga que □ = 0.25,

2.16, 2.07, 2.34, 1.97, 1.97, 1.90, 2.19, 2.23, 2.15, 2.47, 2.31, 1.94, 2.31, 1.86, 2.25, 2.14, 2.15, 2.161 2.30, 2.48, 2.11, 2.15, 2.24, 2.04, 2.21, 1.91, 2.01, 2.09, 2.07, 2.25

Respuesta [2.0581, 2.2370]

10. El crecimiento del tronco principal para una muestra de 17 pinos rojos de 4 años, tiene una media de 11.3 pulgadas y una desviación estándar de 3.4 pulgadas. Obtenga un intervalo de confianza de 90% para la media del crecimiento del tronco principal para una población de pinos rojos de 4 años sujeta a condiciones ambientales similares. Supóngase que el crecimiento tiene una distribución normal.

Respuesta (9.8602,12.7397)

11. En un proceso químico se han producido, en promedio 800 toneladas de cierto producto por día. Las producciones diarias para la semana pasada fueron 785, 805, 790, 793 y 802 toneladas. Estimar a partir de los datos la media de la producción diaria con un coeficiente de confianza de 90 %. Respuesta (787.0528, 802.947)

12. Debido a la variabilidad en los descuentos por los automóviles entregados a cambio, la ganancia por auto nuevo vendido por un distribuidor de automóviles varia de uno a otro. Las ganancias por ventas (en cientos de dólares); registradas la semana pasada, fueron 2.1, 3.0, 1.2, 6.2, 4.5 y 5.1. Obtener un intervalo de confianza de 95% para la ganancia media por venta. Respuesta (1.684, 5.6825)

13. Se registro el tiempo transcurrido entre la facturación y el pago recibido, para una muestra aleatoria de 91 clientes de una empresa de contadores públicos. La media y la desviación estándar de dicha muestra fueron 39.1 días y 17.3 días, respectivamente. Obtener un intervalo de confianza de 90% para el tiempo medio que transcurre entre la facturación y el pago recibido para todas las cuentas de las firmas de contadores públicos. Interpreta el resultado.

Respuesta (36.089, 42.114)

Interpretación nueve de cada diez veces, el tiempo transcurrido entre la facturación y el pago será aproximadamente entre 36 y 43 días.

14. La cámara de comercio de una ciudad se interesa en estimar la cantidad promedio de dinero que gasta la gente que asiste a convenciones, calculando comidas, alojamiento y entretenimiento por día. De las distintas convenciones que se llevan a cabo en la ciudad, se seleccionaron 16 personas de las que se obtuvo la siguiente información en dólares: 150, 175, 163, 148, 142, 189, 135, 174, 168, 152, 158, 184, 134, 146, 155, 163. Si, se supone que la cantidad de dinero gastada en un día es una variable aleatoria distribuida normal, obtener los intervalos de confianza estimados del 90, 95 y 98% para la cantidad promedio real.

Respuestas (151.3055, 165.6944), (149.753,167,2469), (147.8197, 169.182)

15. Los pesos de la fruta contenida en 21 latas de duraznos seleccionadas al azar fueron (en onzas): 11.0, 11.6, 10.9, 12.0, 11.5, 12.0, 11.2, 10.5, 12.2, 11.8,12. 1, 11.6, 11.7, 11.6, 11.2, 12.0, 11.4, 10.8, 11.8, 10.9 y 11.4. Determine el intervalo de confianza de 98%,para estimar el peso promedio por lata de los duraznos.

Respuesta (11.2238,11.7475)

16. Se aplicó un examen de matemáticas a un grupo de 50 alumnos seleccionados al azar de la secundaria A y a un grupo de 45 de estudiantes seleccionados al azar de la secundaria B. El grupo de la secundaria, A obtuvo una media de 75 puntos con una desviación estándar de 10 puntos. El grupo de la secundaria B logró una media 72 puntos con una desviación de 8 puntos. Construir un intervalo de confianza del 95% para la diferencia en los resultados medios.

Respuesta Sp= 9.1086, (-0.6684, 66684)

17. Una comparación de los tiempos de reacción a dos estímulos diferentes en un experimento psicológico de asociación de palabras aplicado a una muestra aleatoria de 16 personas, produjo los resultados (en segundos) que se muestran en la siguiente tabla. Obtener un intervalo de confianza de 90% para $(\mu_1 - \mu_2)$.

Estimulo	Tiempo de reacción (en segundos							
1	1	2	3	1	2	3	1	2
2	4	1	3	3	2	2	3	3

Respuesta Sp 0.8767, (-1.5216.0.0216)

18. Estime un intervalo de confianza del 95 % la diferencia del coeficiente de inteligencia (IQ) entre los miembros mas viejos y más jóvenes (hombres y mujeres) de una familia tomando como base la siguiente muestra aleatoria de sus IQ.

Mas viejos	145	133	116	128	85	100	105	150	97	110	120	130
Mas jóvenes	131	119	103	93	108	100	111	130	135	113	108	125

Respuesta Sp=16.8993, (10.7181.17.8981)

19. Se aplicaron 2 métodos para enseñar la lectura a dos grupos de niños de una escuela primaria y se compararon los resultados mediante una prueba de lectura y comprensión. El método 1 se aplicó a 11 niños para los cuales se obtuvo una media y una desviación estándar de 64 y 52 puntos respectivamente. El método 2 se probó en 14 niños que al final de la prueba obtuvieron una media de 69 y una desviación estándar de 71 puntos. Obtener un intervalo de confianza de 95% para $(\mu_1 - \mu_2)$.

Respuesta Sp=7.9208, (-1.6020,11.6020)

20. Se administraron dos nuevos medicamentos a pacientes con cierto padecimiento cardiaco. El primer medicamento bajo la presión sanguínea de 16 pacientes en un promedio de 11 puntos, con una desviación estándar de 6 puntos. El segundo fármaco disminuyó la presión sanguínea de 20 pacientes en un promedio de 12 puntos, con una desviación estándar de 8 puntos. Desarrollar un intervalo de confianza del 95% para la diferencia en la reducción media de la presión sanguínea, bajo el supuesto que las mediciones se distribuyen normales con varianzas iguales.

Respuesta Sp=7.1866, (-5.89,3.89)

21. Se midieron las presiones sanguíneas diastólicas de 15 pacientes utilizando dos técnicas: el método estándar utilizado por personal médico y otro método que utiliza un aparato electrónico con indicador digital. Los resultados fueron los siguientes:

Método	Paciente														
Estándar	72	80	88	80	80	75	92	77	80	65	69	96	77	75	60
Indicador digital	70	76	87	77	81	75	90	75	82	64	72	95	80	70	61

Determine el intervalo de confianza del 90% para la diferencia media de las dos lecturas

Respuesta Sp= 9.7119, (5.3026,6.7626)

22. En un proceso químico se han producido, en promedio, 800 toneladas de cierto producto por día las producciones diarias para la semana pasada fueron 785, 805, 790, 793 y 802 toneladas. Determinar un intervalo de confianza de 90% para la varianza σ^2 de la producción diaria.

Respuesta P[29.3013<σ^2<391.1668]=0.90

23. Se indica que las anormalidades congénitas ocurren mayormente entre niños varones engendrados por padres de mayor edad promedio. Se obtuvieron histories clínicas de este tipo de anormalidades correspondientes a 20 infantes varones cuyas madres tuvieron las edades siguientes: 31, 21, 29, 28, 34, 45, 21, 41, 27, 37, 43, 21, 39, 38, 32, 28, 37, 28, 16 y 39. Determine el intervalo de confianza de 90% para la desviación estándar de la edad en madres de hijos con anormalidades congénitas.

Respuesta P[40.85<σ^2<121.70]=0.90

24. Si 32 mediciones del punto de ebullición del azufre tienen una desviación estándar de 0.83 grados Celsius, constrúyase un intervalo con un nivel de confianza del 98% para la desviación estándar real de tales mediciones.

P[0.38<σ^2<0.70]=0.98

25. Se espera tener una cierta variación aleatoria nominal en el espesor de las láminas de plástico que una máquina produce. Para determinar cuando la variación en el espesor se encuentra dentro de ciertos límites, cada día se seleccionan en forma aleatoria 12 láminas de plástico y se mide en milímetros su espesor. Los datos que se obtuvieron son los siguientes: 12.6, 11.9, 12.3, 12.8 11.8, 11.7, 12.4, 12.1, 12.3, 12.0, 12.5, 12.9. Si se supone que el espesor es una variable aleatoria distribuida normal, obtener el intervalo de confianza estimado del 99% para la varianza desconocida del espesor. Si no es aceptable una varianza mayor de 0.90 mm, ¿existe alguna razón para preocuparse con base en esta evidencia?

P[0.0614<σ^2<0.6309]=0.99 No. La muestra no proporciona evidencia de que ocurra una varianza de 0.9 mm^2, con una confianza del 99%

26. Se tiene interés en la variabilidad de los puntajes obtenidos en un examen TOEFL (de Test of english as a Foreign Languaje). Se obtiene una muestra aleatoria de puntajes correspondientes a estudiantes extranjeros con los siguientes resultados: 495, 525, 580, 605, 552, 490, 590, 505, 551, 600. Obtenga un intervalo de confianza de 95 % para la desviación estándar delas calificaciones del examen TOEFL.

P[30.16<σ^2<80.04]=0.95

27. Mientras realizan una tarea extenuante, el ritmo cardiaco de 25 trabajadores se incrementa en un promedio de 18.4 pulsaciones por minuto, con una desviación estándar de 4.9 pulsaciones por minuto. Calcular un intervalo con un nivel de confianza del 95% para la correspondiente desviación estándar de la población.

$P[3.83<\sigma^2<6.82]=0.95$

Pruebas de hipótesis

1. Los salarios diarios en cierta, rama de la industria en particular presentan una distribución normal con media de $13.20 y una desviación estándar de 2.50. Una compañía X que emplea a 40 trabajadores paga en promedio $12.20, ¿puede acusarse a esta compañía de pagar sueldos bajos?. Emplear = 0.01.

Respuestas: Estadístico de Prueba =-2.5298

Regla de decisión: Rechazar Ho si $Z_c< Z_{0.01}=2.325$

Conclusión: Rechazar Ho en favor de Ha. Es decir con un nivel de significancia = 0.01, existe suficiente evidencia muestral para decir que la compañía X paga salarios inferiores a los de la rama de la industrias a la que pertenece.

2. Las casas cercanas a una universidad tienen un valor medio igual a $58,000. Se supone que aquellas que están situadas en la vecindad de la universidad tienen un valor superior. Se tomo una muestra aleatoria de 12 casas en el área universitaria para contrastar esta teoría. Su avaluó promedio es de $62,460, siendo su desviación estándar de $5,200. Realice un contraste de hipótesis utilizando =0.01

Conclusión: Rechazar Ho en favor de Ha con un nivel significancia de 0.01. Es decir la muestra aporta suficiente evidencia para manifestar que las casas cercanas a una universidad tienen un valor medio mayor a $58,000.

3. Un grupo de estudiantes sostiene que el alumno promedio invierte al menos 25 minutos diarios para llegar a la universidad. El departamento de servicios escolares obtuvo una muestra del tiempo empleado (un solo sentido) por 36 estudiantes cuya media y desviación estándar fue 22 y 7.3 minutos, respectivamente. ¿ Tiene el departamento evidencia para rechazar la afirmación de los estudiantes? Utilice = 0.01

Estadístico de Prueba =-2.4657

Regla de decisión: Rechazar Ho si t_c $-t_{0.01,35}=-2.4377$

Conclusión: Rechazar Ho en favor de Ha con un nivel de significancia de 0.01. Es decir existe suficiente evidencia muestral para rechazar la afirmación de los estudiantes.

4. En las etiquetas de una marca de leche evaporada se afirma que esta contiene "no menos de 850 U.I. (Unidades internacionales) de vitamina D por litro". Se realizan 15

determinaciones del contenido (por litro) de vitamina D y se obtienen los siguientes resultados:

Conclusión: Con un nivel de significancia = 0.025 no existe suficiente evidencia muestral para rechazar Ho, por lo que se concluye que la leche evaporada contiene al no menos de 850 U.I. de vitamina D por litro.

5. En una muestra aleatoria de seis varillas de acero se obtuvo una resistencia media a la comprensión de 58,392 psi (libras por pulgada cuadrada) con una desviación estándar de 648 psi. Emplear esta información y un nivel de significancia de = 0.05 para probar si la media de la resistencia real a la comprensión del acero del cual proviene esta muestra es de 58,000 psi.

Conclusión: No existe suficiente evidencia muestral para rechazar Ho con un nivel de significancia de 0.05. Es decir la media, de la resistencia a la compresión del acero del cual proviene esta muestra es de 58,000 psi.

6. Una muestra aleatoria de los archivos de una compañía que contiene información detallada indica que las órdenes para cierta pieza de máquina fueron entregados en 10, 12, 19, 14, 15, 18, 11 y 13 días. Usar un nivel de significancia = 0.01 para probar

La afirmación que el tiempo medio de entrega es de 10.5 días. Elegir la Hipótesis alterna de manera que el rechazo de la hipótesis nula = 10.5 implique que la entrega de las órdenes toma más tiempo del indicado.

Conclusión: Rechazar Ho en favor de Ha con un nivel de significancia de 0.01. Es decir existe evidencia muestral que indica un tiempo de entrega mayor de 10.5 días

7. Cinco mediciones del contenido de alquitrán de cierta marca de cigarros producen los siguientes resultados: 14.5, 14.2, 14.4, 14.3 y 14.5 mg por cigarro. Probar que la diferencia entre el promedio muestral y la media del contenido de alquitrán que indica el fabricante = 14.0 es significativa, con = 0.05.

Conclusión: Rechazar Ho en favor de Ha con un nivel de significancia de 0.05. Es decir existe suficiente evidencia muestral para rechazar la afirmación del fabricante

8. Supóngase que en el ejercicio anterior la primera medición es anotada incorrectamente como 16.0 en lugar de 14.5. Verificar si ahora la diferencia entre la media muestral y el contenido de alquitrán que indica el fabricante = 14.0 no es significativa con = 0.05. Explicar la aparente paradoja de que, a pesar de que la diferencia entre y ha aumentado, no hay significancia estadística.

Conclusión: No rechazar Ho con un nivel de significancia de 0.05. Es decir no existe suficiente evidencia muestral para rechazar la afirmación del fabricante. Note el incremento desproporcionado de la varianza con respecto al de la media lo que causa la aparente paradoja.

9. Un fabricante de bombas de pozo profundo asegura que a lo sumo el 30 % de sus bombas requieren reparación en los primeros 5 años de operación. Si una muestra aleatoria de 120 bombas incluye 47 que requieren reparación en los primeros 5 años se puede afirmar que esto contradice la afirmación del fabricante. Use = 0.05.

Conclusión: Rechazar Ho en favor de Ha con un nivel de significancia de 0.05. Es decir existe evidencia muestral para contradecir la proporción (30%) de las bombas que requieren reparación los primeros 5 años de operación.

10. La experiencia de un comerciante en aparatos y accesorios mostró que 10% de sus clientes que compran a plazos liquidan sus cuentas antes del vencimiento de la última mensualidad (la vigésimo cuarta). Al sospechar un incremento en este porcentaje el comerciante selecciona al azar 200 compradores a crédito para saber sus intenciones, treinta y tres ellos afirmaron tener planeado pagar adeudos antes de la última mensualidad. ¿Son los datos suficientes para indicar que el porcentaje de compradores a plazos que pagarán sus deudas antes de la última mensualidad, excede de 10%? . Usar un nivel de significancia $\alpha = 0.05$.

Conclusión: Rechazar Ho en favor de Ha con un nivel de significancia de 0.05. Es decir la muestra aporta evidencia suficiente para indicar que el porcentaje de compradores a plazos que pagaran sus deudas antes de la última mensualidad, excede el 10%.

11. . El rendimiento de una computadora se observa en un periodo de dos años para verificar la afirmación de que la probabilidad del tiempo perdido por fallas exceda a 5 horas en una semana cualquiera es de 0.2. ¿ Qué se puede concluir con un nivel de significancia $\alpha=0.05$, si hubo solo 11 semanas en las cuales el tiempo perdido de la computadora excedió las 5 horas?. (Recuerde que un año tiene 52 semanas).

Conclusión: Rechazar Ho en favor de Ha con un nivel de significancia de 0.05. Es decir existe suficiente evidencia muestral para refutar la afirmación del productor acerca del tiempo perdido por fallas en una semana.

12. Un fabricante modificó una línea de producción para reducir el promedio de la fracción de defectuosos. Para determinar si la modificación fue efectiva, el fabricante sacó una muestra aleatoria de 400 artículos antes de la modificación de la línea de producción y otra muestra aleatoria de 400 artículos después de tal cambio. Los porcentajes de defectuosas en las muestras eran

Antes	Después
5.25 %	3.5%

Pruebe la Hipótesis de que la modificación disminuye la proporción de artículos defectuosos con un nivel de significancia $\alpha= 0.05$.

Conclusión: No rechazar Ho con un nivel de significancia $\alpha = 0.01$. Es decir no existe evidencia muestral para afirmar que la modificación reduce significativamente el número de artículos defectuosos.

13. Un genetista está interesado en la proporción de machos y hembras de una población que tiene cierta enfermedad menor en la sangre. En una muestra aleatoria de 100 machos se encuentran 31 afectados mientras que solamente 24 de 100 hembras presentan la enfermedad. Se puede concluir, con un nivel de significancia de $\alpha = 0.01$, que la proporción de machos afectados por esta enfermedad de la sangre es mayor que la proporción de hembras también afectadas? (Walpole V Myers, 1989)

Conclusión: No rechazar Ho con un nivel de significancia α= 0.01. Es decir existe evidencia muestral para afirmar que la enfermedad afecta por igual a ambos sexos.

14. Dos empresas que fabrican artículos equivalentes afirman tener la misma proporción de preferencia hacia sus productos entre los consumidores. Una muestra aleatoria indica que 102 de 300 y 152 do 400 consumidores prefrieren los productos A V B respectivamente. ¿Indica esta evidencia una diferencia significativa entre las proporciones?. Utilizar α= 0.02.

Conclusión: No rechazar Ho con un nivel de significancia α=0.02. Es decir no existe evidencia muestral que sugiera que el consumidor prefiere un producto en especial.

15. Supóngase que deseamos investigar si en promedio el sueldo del hombre excede en más de $ 20 por semana al de la mujer en cierta industria. Si los datos revelan que 60 hombres ganan en promedio $292.50 a la semana con una desviación estándar de $ 1 5.60, mientras que 60 mujeres perciben en promedio $ 266.10 por semana con una desviación estándar de $18.20. ¿Qué puede concluirse con un nivel de significancia de 0.017

Conclusión: No rechazar H_o con un nivel de significancia α = 0.01. Es decir no existe suficiente evidencia muestral para decir que el promedio del sueldo para hombres excede en mas de $ 20 por semana al de la mujer en cierta industria.

16. Un fabricante de motores eléctricos comparó la productividad de trabajadores de ensamblaje para dos tipos de horarios semanales de trabajo de 40 horas. Uno cuatro días de 10 horas (horaria 1) y cl horario estándar de 5 días de 8 horas (horario 2). Se asignaron 20 trabajadores a cada horario de trabajo y se registro el número de unidades armadas durante una semana las medias (en cientos de unidades) y las varianzas muestrales se indican a continuación

Estadística	Horario	
	1	2
Media Muestral	43.10	44.60
Varianza muestral	4.28	3.89

Proporcionan los datos evidencia suficiente para indicar una diferencia en la productividad media para los dos horarios de trabajo?. Haga la prueba con un nivel de significancia α = 0.05.

Conclusión: Rechazar H_o en favor H_a con un nivel de significancia α = 0.05. Es decir la evidencia muestral indica una diferencia de productividad entre horarios.

17. Se aplicó un examen relacionado con los aspectos fundamentales del sida a dos grupos uno de estudiantes universitarios de licenciatura y el otro de egresados del bachillerato. A continuación se presenta un resumen de los resultados de el examen.

Graduados	n	media	Desviación estándar
Universitarios	75	77.5	6.2
Bachilleres	75	60.4	7.4

¿Indican estos datos que los graduados de universidad tuvieron en promedio un resultado significativamente mayor de 13 puntos en el examen?. Utilizar α = 0.001.

Conclusión: Rechazar H_o en favor H_a con un nivel de significancia α = 0.001. Es decir existe suficiente evidencia muestral para decir que los estudiantes universitarios tienen un puntaje, trece puntos significativamente superior que los estudiantes de bachillerato en aspectos relacionados con el sida.

18. Con el fin de reducir los costos en la alimentación de cerdos, se generó una dieta con ingredientes no convencionales y de bajo costo. Para el experimento se contó con 24 cerdos de la misma raza, edad y peso inicial similar. Doce cerdos fueron alimentados con la dieta no convencional y otros doce con un alimento comercial. Se midió la ganancia de peso al final del experimento, los resultados obtenidos se presentan a continuación.

Dieta	Media	Desviación Estándar
Comercial	49.2	3.9
No convencional	40.0	2.5

Probar la hipótesis de que ambas dietas producen igual ganancia de peso. Utilizar un nivel de significancia de α = 0.001.

Conclusión: No rechazar H_o con un nivel de significancia α = 0.001. Es decir existe suficiente evidencia muestral para decir que ambas dietas producen la misma ganancia de peso.

19. Dos grupos de 10 ratones de laboratorio fueron alimentados con una dieta preestablecida. Al finalizar tres semanas se registró el peso ganado por cada animal. ¿Justifican los datos de la tabla siguiente la conclusión de que el peso medio ganado con la dieta B fue mayor que con la dieta A, al nivel de significación α = 0.05?

Dieta A	5	14	7	9	11	13	14	12	8	7
Dieta B	5	21	4	9	16	23	16	13	19	21

Conclusión: Rechazar H_o en favor H_a con un nivel de significancia α = 0.05. ES decir existe suficiente evidencia muestral para decir que el peso medio ganado con la dieta B fue mayor que el de la dieta A.

20. Se ha desarrollado una nueva cura para cemento Pórtland. Se efectúan ensayos para determinar si la nueva cura tiene un efecto (positivo o negativo) en la resistencia. Se ha producido un lote sometido a ambas curas, la estándar y la experimental. Las resistencias a la compresión (psi) son las siguientes

Cura estándar X	4.125	4.225	4.35	3.575	3.875	3.825	3.975	3.80	3.775	3.850
Cura experimental Y	4.25	3.95	3.9	4.075	4.55	4.45	4.15	4.55	3.70	4.25

Pruebe el efecto en la resistencia del cambio de cura a un nivel de significancia de α= 0.05.

Conclusión: Rechazar H_o en favor H_a con un nivel de significancia α = 0.05. Es decir existe suficiente evidencia muestral para decir que la nueva cura tiene efecto (positivo o negativo) en la resistencia del cemento.

21. Datos de archivo indican que la varianza de las mediciones efectuadas sobre lámina metálica grabada, las cuales fueron obtenidas por inspectores expertos en control de calidad es de 0.18 pulgadas cuadradas. Las mediciones realizadas por un inspector sin experiencia podrían tener una varianza mayor (debido quizás a su poca destreza para leer los instrumentos) o también una varianza muy pequeña (quizás porque las mediciones excesivamente altas o bajas se han descartado). Si un nuevo inspector mide 101 láminas grabadas con una varianza de 0.13 pulgadas cuadradas, pruébese con un nivel de significancia de 0.05 si el inspector realiza mediciones satisfactorias.

Conclusión: Rechazar H_o en favor de H_a con un nivel de significancia α = 0.05. Es decir existe evidencia muestral suficiente para indicar que el nuevo inspector no toma satisfactoriamente sus mediciones.

22. El gerente de una planta sospecha que el número de piezas que produce un trabajador en particular por día, fluctúa más allá del valor normal esperado. El gerente decide observar el número de piezas que produce este trabajador durante 10 días, seleccionados estos al azar. Los resultados son 15, 12, 8, 13, 12, 15, 16, 9, 8 y 14. Si se sabe que la desviación estándar para todos los trabajadores es de 2 unidades y si el número de estas que se produce diariamente se encuentra modelada en forma adecuada por una distribución normal, aun nivel de α = 0.05, ¿ Tiene apoyo la sospecha del gerente?

Conclusión: Rechazar H_o en favor de H_a con un nivel de significancia α = 0.05. Es decir existe evidencia muestral para avalar la sospecha del gerente.

23. En un proceso de llenado, la tolerancia para el peso de los recipientes es 8 gramos para reunir este requisito la desviación estándar en el peso puede ser de 2 gramos. Los pesos de 25 recipientes seleccionados al azar dieron como resultado una desviación estándar de 2.8 gramos. Si los pesos se distribuyen normales, determinar si la varianza de estos es diferente del valor necesario. Emplear α = 0.01

Conclusión: Rechazar H_o en favor de H_a con un nivel de significancia α = 0.01. Es decir la evidencia muestral indica que el proceso no tiene la tolerancia requerida.

24. Un fabricante de máquinas empacadoras de jabón en polvo afirma, que su producto podría llenar las cajas con un peso dado con una amplitud de no más 2/5 de onza. La media y la varianza da una muestra de 8 cajas de 3 onzas resultaron ser iguales a 3.1 y 0.01 8 onzas, respectivamente. Pruebe la hipótesis de que la varianza de la población de mediciones del peso es $\square^2 = 0.01$ contra la alternativa de $\square^2 > 0.01$. Emplear un nivel de significancia $\square = 0.05$.

Conclusión: No rechazar H_o con un nivel de significancia $\square = 0.05$. Es decir la muestra no proporciona evidencia suficiente para decir que $\square^2 > 0.01$.

25. Un agricultor labra todo su terreno en la misma época con un solo cultivo. En consecuencia, desea sembrar una variedad de fríjol cuya maduración sea uniforme (que sea pequeña la desviación estándar entre los momentos de madurez de las plantas). Una productora de semillas ha desarrollado un nuevo híbrido que considera idóneo para el agricultor. El tiempo de maduración de la variedad estándar tiene una media igual a 50 días con una desviación estándar de 2.1 días. Una muestra aleatoria de 30 plantas del nuevo híbrido señala una desviación estándar de 1.65 días. ¿Indica esta muestra una disminución significativa de la desviación estándar al nivel de significancia $\square = 0.05$?.

Conclusión: No rechazar Ho con un nivel significancia de $\square = 0.05$. Es decir la muestra de la nueva variedad no proporciona evidencia significativa de tener una desviación estándar menor de 2.1.

26. Un grupo de ecologistas manifiesta que la temperatura durante el verano en cierta región es más variable actualmente como consecuencia de la contaminación. Si la temperatura máxima histórica (20 años) es de 34° con una desviación estándar de 4°. Se tomó una muestra de tamaño 21 de las temperaturas máximas obtenidas durante los últimos 3 años en dicha región y se obtuvo una desviación estándar de 7.5°. Probar la hipótesis de los ecologistas con un nivel de significancia de 0.05

Conclusión: Rechazar H_o en favor de H_a con un nivel de significancia de $\square = 0.001$. Es decir existe evidencia muestral para validar un incremento en la variabilidad de la temperatura.

27. Un director de personal que proyectaba utilizar una prueba t de student para comparar el promedio del número de inasistencias mensuales para dos categorías de empleados se encontró con una posible dificultad. La variación en el número de tales inasistencias parecía ser diferente para los dos grupos. Para verificar esto, seleccionó aleatoriamente 5 meses y contó el número de faltas de asistencia para cada grupo. Los datos se muestran en la siguiente tabla.

Categoría A	20	14	19	22	25
Categoría B	37	29	51	40	26

a) ¿Cuál fue la suposición necesaria para poder usar la prueba t que preocupa al director de personal? Los datos provienen de poblaciones que se distribuyen normales con varianzas poblacionales desconocidas pero iguales. En este problema aparentemente las varianzas no son iguales. b) Proporcionan los datos evidencia suficiente para indicar que las varianzas difieren para las poblaciones de las inasistencias para las dos categorías de empleados?. Emplear $\square = 0.10$ e interpretar los resultados.

Conclusión: No rechazar H_o con un nivel de significancia α= 0.10. Es decir no existe evidencia muestral, para decir que las varianzas poblacionales de las dos muestras son diferentes, por lo tanto es posible que efectuar la prueba t proyectada por el director de personal.

28. La cantidad de cera superficial en cada lado de bolsas de papel encerado es una variable aleatoria. Hay razones para sospechar que hay una mayor variación en la cantidad de cera en el interior de la bolsa que era el exterior. Se ha obtenido una muestra de 61 observaciones de la cantidad de cera de cada lado de estas bolsas con los siguientes resultados.

Cera en libras por unidad de área muestreada		
Estadísticas	Superficie exterior	Superficie interior
Media	0.9480	0.6520
Varianza	0.3189	0.7043

Conduzca una prueba para determinar si la variabilidad de la cantidad de cera de la superficie interior es mayor que la contenida en la superficie exterior, Usar un nivel de significancia α = 0.01.

Conclusión: Rechazar H_o en favor de H_a con un nivel de significancia = 0.01. ES decir existe suficiente evidencia muestral, para indicar que la superficie interior de las bolsas contiene más cera que la exterior.

29. Una panadería está considerando la compra de uno de dos hornos. Se requiere que la temperatura permanezca constante durante la operación de horneado. Se hizo un estudio para medir la varianza en temperatura de los dos hornos en funcionamiento. Antes de que el termostato restableciera la flama, la varianza en la temperatura del horno A fue igual a 2.4, resultante de 16 mediciones. La varianza del horno B fue de 3.2 resultante de 12 mediciones. ¿ Proporciona esta información suficiente evidencia para concluir que existe una diferencia entre las varianzas para los dos hornos?. Utilizar un nivel de significancia de α=0.02.

Conclusión: No rechazar H_o con un nivel de significancia α = 0.02. Es decir no existe suficiente evidencia muestral, para concluir que hay diferencia entre las varianzas para los dos hornos.

30. Se realizó un estudio para decidir si hay o no la misma variabilidad en la presión sanguínea sistólica entre hombres y mujeres. Se utilizaron muestras aleatorias de 16 hombres y 13 mujeres para contrastar la afirmación de los investigadores en el sentido de que las varianzas eran diferentes. Realice el contraste de hipótesis, con α = 0.05, utilizando los datos siguientes:

Hombres	120 120 118 112 120 114 130 114 124 125 130 100 120 108 112 122
Mujeres	122 102 118 126 108 130 104 116 102 122 120 118 130

Conclusión: No rechazar H_o con un nivel de significancia α = 0.05. La muestra no aporta evidencia para decir que la variabilidad de la presión sanguínea sistólica depende del individuo.

31. En un experimento acerca de la contaminación del aire, se comparan dos tipos de instrumentos para medir fa cantidad de monóxido de sulfuro en la atmósfera. Se desea determinar si los dos tipos de instrumentos producen mediciones que tienen la misma variabilidad variabilidad. Se registraron las siguientes lecturas para los dos instrumentos.

Monóxido de sulfuro		
Instrumento A		Instrumento B
0.86		0.87
0.82		0.74
0.75		0.63
0.61		0.55
0.89		0.76
0.64		0.70
0.81		0.69
0.68		0.57
0.65		0.53
Media	0.7455	0.673111
Desviación estándar	0.10405	0.11174

Suponiendo que las distribuciones de las poblaciones están distribuidas aproximadamente en forma normal, probar la hipótesis planteada con un nivel de significancia de α=0.02.

Conclusión: No rechazar H_o con un nivel de significancia α= 0.02. Es decir no existe suficiente evidencia muestral que manifieste diferencias en la variabilidad de los instrumentos.

32. El Instituto del consumidor desea comparar la variabilidad en la eficacia de un medicamento elaborado por las compañías X e Y. Ambos medicamentos se distribuyen en forma de tabletas de 250 mg. Se determinó la eficacia en 25 tabletas en cada compañía encontrándose S^2_1 = 2.09 y S^2_2 =1.06. Realizar una prueba para contrastar la variabilidad de ambos medicamentos con α =0.10.

Conclusión: Rechazar H_o a favor de H_a con un nivel de significancia α =0.10. Es decir la muestra proporciona evidencia sobre la diferencia en la variabilidad de ambos productos.