Estadística II
Ejercicios Tema 1

---

1. Una muestra aleatoria simple de diez coches del modelo X dio los siguientes datos de consumo de combustible (en millas por galón):

   27.2 27.2 26.8 26.9 25.3 26.0 26.4 25.7 28.1 25.7

   Otra muestra aleatoria simple independiente de la anterior para doce coches del modelo Y dio los resultados siguientes:

   24.2 24.3 25.3 24.8 25.1 25.0 24.9 23.9 26.0 26.1 26.0 26.3

   Empleando la siguiente información obtenida de Excel, conteste a las preguntas a continuación:

	A	B	C	D	E	F	G
1		X	Y	X*X	Y*Y	X>25.8	Y>25.8
2		27.2	24.2	739.84	585.64	1	0
3		27.2	24.3	739.84	590.49	1	0
4		26.8	25.3	718.24	640.09	1	0
5		26.9	24.8	723.61	615.04	1	0
6		25.3	25.1	640.09	630.01	0	0
7		26	25	676	625	1	0
8		26.4	24.9	696.96	620.01	1	0
9		25.7	23.9	660.49	571.21	0	0
10		28.1	26	789.61	676	1	1
11		25.7	26.1	660.49	681.21	0	1
12			26		676		1
13			26.3		691.69		1
14	sum	265.3	301.9	7045.17	7602.39	7	4
15							

   (a) Emplee un procedimiento de estimación insesgado para calcular estimaciones puntuales de la media de la población μ_X y la varianza de la población σ_X^2.

   (b) Emplee un procedimiento de estimación insesgado para calcular una estimación puntual de la proporción en la población de los coches del modelo X, p_X, cuyo consumo de combustible es superior a 25.8.

   (c) Emplee un procedimiento de estimación insesgado para calcular estimaciones puntuales de la media de la población μ_Y y la varianza de la población σ_Y^2.

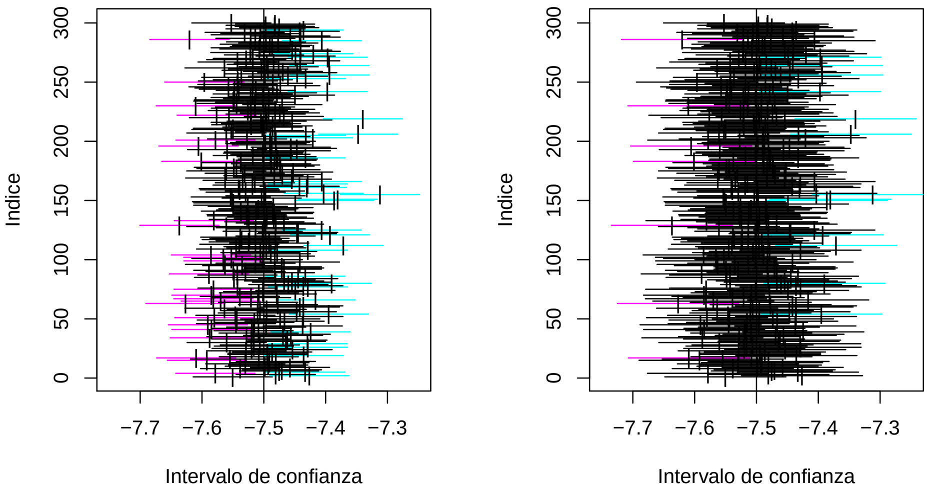
   (d) Emplee un procedimiento de estimación insesgado para calcular una estimación puntual de la proporción en la población de los coches del modelo Y, p_Y, cuyo consumo de combustible es superior a 25.8.

   (e) Emplee un procedimiento de estimación insesgado para calcular una estimación puntual de la diferencia de los consumos medios de combustible en las poblaciones entre los coches modelo X y modelo Y, esto es, de $\mu_X - \mu_Y$.

   (f) Emplee un procedimiento de estimación insesgado para calcular una estimación puntual de la diferencia entre la proporción en la población de coches del modelo X con consumos superiores a 25.8 la proporción en la población de coches del modelo Y con consumos superiores a 25.8, esto es, de $p_X - p_Y$.

2. Obtenemos una muestra aleatoria simple de tamaño $n = 400$ de una población X y construimos un intervalo de confianza al 95% para la media de la población (desconocida) μ_X (suponemos que la desviación típica de la población, $\sigma_X = 1$, es conocida). Empleando *la misma muestra* construimos un intervalo de confianza al 80% para la media de la población μ_X.

Repetimos este proceso 300 veces. Como resultado, obtenemos 300 intervalos de confianza al 95% y otros 300 intervalos de confianza asociados para un nivel de confianza del 80%. Ambos conjuntos de intervalos se representan en la siguiente imagen.



(a) ¿Qué gráfico, el izquierdo o el derecho, representa los intervalos de confianza al 95% para μ_X? ¿Cuál es el gráfico correspondiente a los intervalos al 80% para μ_X? Justifique su respuesta.

(b) Para cada gráfico, indique cuantos de los 300 intervalos (aproximadamente) contendrán el valor de μ_X y cuantos no lo harán.

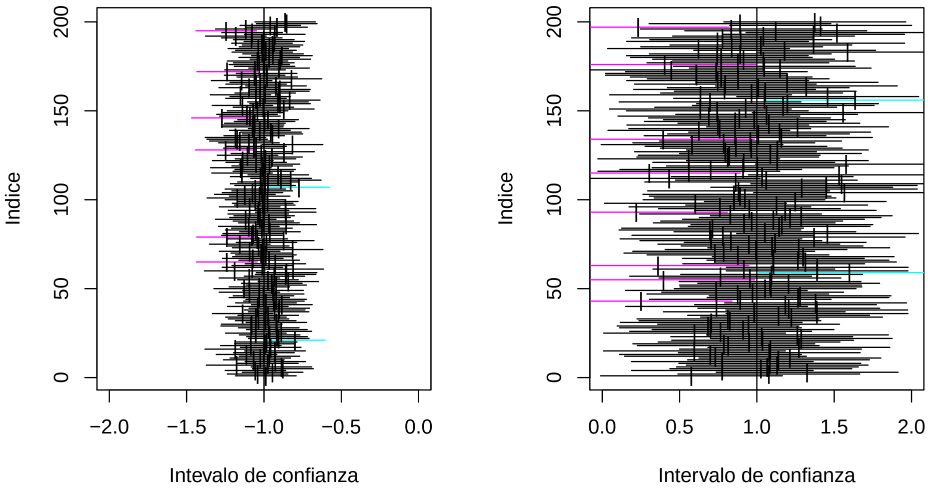
3. La edad media de todos los ministros del Gobierno de España en el momento de su nombramiento es 55 años. Tomamos una muestra aleatoria simple de 30 ministros. ¿Tendría sentido construir un intervalo de confianza al 95% para la edad media de la población?

4. Sea z_α el cuantil superior de la distribución normal estándar correspondiente a $\alpha \in (0, 1)$, esto es, z_α satisface $P(Z > z_\alpha) = \alpha$, donde $Z \sim N(0, 1)$. Si α aumenta, ¿aumenta o decrece el valor de z_α? Justifique su respuesta.

5. Considere dos poblaciones que siguen una distribución normal, $X \sim N(\mu_X, \sigma_X^2)$ y $Y \sim N(\mu_Y, \sigma_Y^2)$. Tomamos 200 muestras aleatorias simples de tamaño $n = 100$ de X y, suponiendo que la desviación típica de la población es conocida, construimos un intervalo de confianza al 95% para la media de la población a partir de cada muestra.

   Tomamos otras 200 muestras aleatorias simples de tamaño $n = 100$ de Y y, suponiendo que la desviación típica de la población es conocida, construimos un intervalo de confianza al 95% para la media de la población a partir de cada muestra.

   A continuación se representan los intervalos de confianza obtenidos:

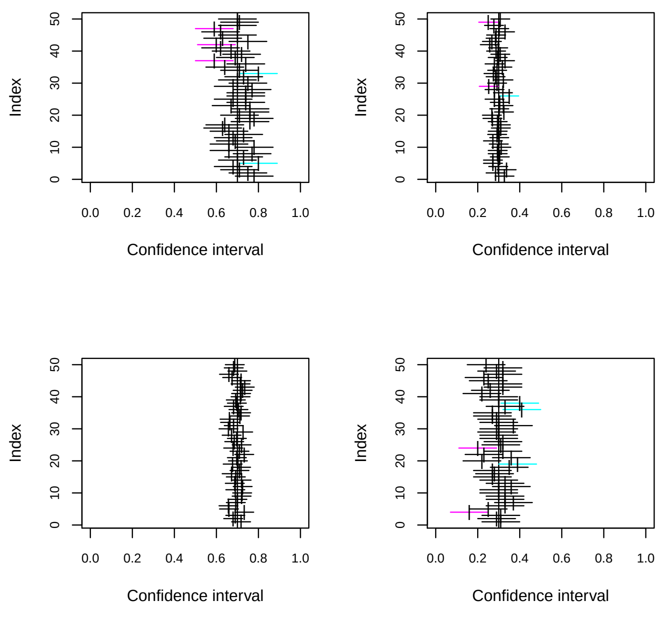


(a) Sabiendo que la desviación típica de la población σ_X es igual a 1 y que σ_Y es igual a 3, asocie cada gráfico a su población correspondiente (X o Y). Justifique su respuesta.

(b) Para cada gráfico, indique cuantos de los 200 intervalos (aproximadamente) contendrán bien μ_X o bien μ_Y y cuantos no.

6. Los siguientes cuatro gráficos muestran intervalos de confianza (para un nivel de confianza del 95%) para la proporción en la población (desconocida) p_X, obtenidos de 50 muestras aleatorias simples de tamaño n de X, bajo cuatro escenarios I-IV. Identifique cada escenario con su gráfico.

I.	$n = 400,$	$X \sim \text{Bernoulli}(0.3)$
II.	$n = 400,$	$X \sim \text{Bernoulli}(0.7)$
III.	$n = 100,$	$X \sim \text{Bernoulli}(0.3)$
IV.	$n = 100,$	$X \sim \text{Bernoulli}(0.7)$



7. El intervalo de confianza para la media de la población μ_X es simétrico en torno a la media muestral $\bar{x}$. ¿Es simétrico el intervalo de confianza para la varianza de la población σ_X^2 en torno a la cuasi-varianza de la muestra s_x^2? Justifique su respuesta.

8. Se ha obtenido el intervalo de confianza al 99%, $(75.7, 82.5)$, para la media de la población correspondiente al nivel de satisfacción de los estudiantes con sus coches. ¿Sería razonable decir que se tiene una probabilidad del 99% de que la media de la población para el nivel de satisfacción se encuentre entre 75.7 y 82.5? Si no fuese así, reemplace esta afirmación por otra correcta.

9. Un fabricante desea controlar los niveles de impurezas en los envíos de materia prima que recibe de un proveedor. En una muestra aleatoria simple de quince envíos se obtuvo una cuasi-desviación típica de 2.36% en el nivel de concentración de las impurezas. Se supone que la distribución de los niveles de impurezas sigue una distribución normal.

   (a) Calcule e interprete un intervalo de confianza al 95% para la varianza de la población.

   (b) Un intervalo de confianza al 99%, ¿sería menor o mayor que el obtenido en el apartado anterior? Justifique su respuesta sin realizar ningún cálculo.

   (c) Calcule e interprete un intervalo de confianza al 95% para la desviación típica de la población.

10. Las personas en una muestra aleatoria simple de cincuenta y cuatro representantes sindicales fueron preguntadas por la frecuencia con la que en sus negociaciones con la empresa habían aceptado una revisión salarial por debajo del aumento del coste de la vida. De los miembros de la muestra, catorce respondieron "nunca" a esa pregunta. Con esta información, un estudiante de Estadística calculó un intervalo de confianza entre 16% y 36% para el porcentaje de la población de representantes sindicales que nunca aceptaron una revisión inferior al aumento del coste de la vida.

    (a) Encuentre el nivel de confianza asociado a este intervalo.

    (b) Construya e interprete un intervalo de confianza al 80% para el porcentaje de la población que se está considerando.

11. Se supone que las notas obtenidas por un grupo numeroso de estudiantes que tomaron un examen siguen una distribución normal. Una muestra aleatoria de veinticinco notas nos proporciona los siguientes valores:

$$\sum_{i=1}^{25} x_i = 1508 \qquad \sum_{i=1}^{25} x_i^2 = 95628$$

    (a) Encuentre e interprete un intervalo de confianza al 90% para la media de la población.

    (b) Si llevar a cabo ningún cálculo, indique si un intervalo de confianza al 95% para la media de la población sería mayor o menor que el anterior.

12. La UC3M está interesada en analizar el tiempo que sus estudiantes dedican al estudio durante la semana. Una muestra aleatoria simple de dieciséis estudiantes tenía una media de tiempo dedicado al estudio de 18.36 h., con una cuasi desviación típica muestral de 3.92 h. Suponemos que los tiempo de estudio están distribuidos normalmente.

    (a) Encuentre e interprete un intervalo de confianza al 90% para el tiempo medio de estudio semanal de todos los estudiantes de la UC3M.

    (b) Indique, sin realizar cálculos adicionales, si el intervalo sería menor o mayor en cada uno de los casos siguientes:

        i. La muestra contiene treinta estudiantes (los demás valores no varían).

        ii. La cuasi desviación típica de la muestra es 4.15 h. (los demás valores no varían).

        iii. Se desea obtener un intervalo al 80% (los demás valores no varían).

    (c) Si la población no estuviese distribuida normalmente, ¿se podría construir un intervalo de confianza para la media de la población? Justifique su respuesta. ¿Que posible solución podría aplicarse en este caso?

    (d) Construya e interprete un intervalo de confianza al 90% para la desviación típica de la población.

13. Un proceso industrial para fabricar ladrillos genera productos cuyos pesos siguen una distribución normal con una desviación típica de 0.12 Kg. Una muestra aleatoria de dieciséis ladrillos de la producción de un día tiene un peso medio de 4.07 Kg.

    (a) Encuentre e interprete un intervalo de confianza al 95% para el peso medio de todos los ladrillos producidos en ese día.

    (b) La dirección piensa que el intervalo obtenido en la pregunta anterior es demasiado amplio y desea reducir su tamaño a la mitad. ¿Cuántas observaciones adicionales serían necesarias para conseguir un intervalo con esa propiedad?

    (c) Proporcione un valor para el error estándar de $\bar{X}$.

14. Una universidad hace un seguimiento de los sueldos de sus antiguos alumnos. A partir de una muestra de 20 antiguos alumnos de una titulación ha observado que la media de los sueldos mensuales transcurrido un año desde la graduación es de 1700 euros, y la cuasi desviación típica es de 350 euros. Suponemos que la población sigue una distribución normal.

    (a) Calcule un intervalo de confianza al 95% para el salario medio mensual de todos los graduados de la titulación, un año después de su graduación.

    (b) ¿Cuál sería el intervalo si el nivel de confianza se fija en el 99%?

    (c) ¿Qué tamaños de muestra darían intervalos de longitud menor de 100 euros, si los valores de la media y la cuasi desviación típica muestrales no varían, y el nivel de confianza es del 95%?

15. Una compañía lleva a cabo un control de calidad para los productos que recibe de uno de sus proveedores. Para controlar uno de los atributos de este producto se selecciona una muestra aleatoria de 20 y se mide el valor del mismo. Para esta muestra se obtiene una desviación típica muestral igual a 2.3. Suponemos que los valores del atributo siguen una distribución normal.

    (a) Calcule un intervalo de confianza al 90% para la varianza de la población.

    (b) Supongamos que para otra muestra de tamaño 40 se obtiene el mismo valor de la desviación típica muestral. ¿Cuál sería el valor del intervalo de confianza para la varianza de la población obtenido a partir de esta muestra, para el mismo nivel de confianza del 90%?

16. Un estudio de mercado ha llevado a cabo una encuesta a 64 personas. En ella se les ha preguntado si estarían interesadas en comprar dos productos, A y B. 29 personas han respondido afirmativamente a la pregunta en relación con el producto A, 32 han mostrado interés en el producto B y 26 estaban interesadas en ambos productos.

    Se pide que:

    (a) Emplee un procedimiento de estimación insesgado para estimar el porcentaje de clientes en la población interesado en comprar ambos productos.

    (b) Emplee un procedimiento de estimación insesgado para estimar el porcentaje de clientes interesados en adquirir el primer producto, pero no el segundo.

    (c) Indique la varianza del estimador empleado para responder a la pregunta anterior, y estime su valor a partir de la muestra.

    (d) Calcule un intervalo de confianza para la proporción de clientes interesados en la compra del primer producto, pero no del segundo (como en la segunda pregunta), para un nivel de confianza del 99%.

    (e) *Pregunta para discutir.* Calcule otro intervalo de confianza (aproximado) para esta proporción con el mismo nivel de confianza, pero con la propiedad de que el extremo izquierdo del intervalo sea igual a cero.

17. El equipo directivo de una gran superficie tiene información de que la edad de los clientes que acuden a su centro comercial sigue una distribución normal con desviación típica igual a 5. Se realiza una encuesta de forma aleatoria (m.a.s) a 100 clientes y se obtiene que la media de edad entre los encuestados es de 30 años. Se pide que:

    (a) Obtenga un intervalo de confianza al 96% para la media de edad de los clientes que acuden al centro.

    (b) ¿Cuál sería el tamaño de la muestra para obtener un intervalo de longitud 1 o inferior?

    (c) Calcule la probabilidad de que la cuasi-varianza muestral sea menor que 19.67.

    (d) En la muestra anterior 25 personas se han identificado como extranjeros. Calcule un intervalo de confianza al 95% para la proporción de extranjeros que compran en este centro.

    (e) La directora de ventas de la empresa dispone de información de otra muestra (independiente) sobre la proporción de extranjeros en sus centros. Para una muestra de tamaño 200 la proporción muestral es del 20%. Calcule un intervalo de confianza mejorado (a un nivel del 95%) para esta proporción utilizando toda la información disponible.

(f) Si el tamaño de la muestra es grande, se sabe que la mediana muestral, representada por $\hat{X}_m$, sigue una distribución normal con media μ y varianza $\frac{\pi}{2}\frac{\sigma^2}{n}$. Queremos comparar la eficiencia de la mediana muestral con respecto a la media muestral; ¿cúal de los dos estimadores es más eficiente para n grande?

18. El número de reservas de billetes en vuelos de una cierta ruta de una compañía aérea sigue una distribución normal. Se toma una muestra aleatoria de 81 observaciones de números de reservas en vuelos de esta ruta. Para dicha muestra, el número medio de reservas resulta ser de 112 y la cuasidesviación típica 36. De estos 81 vuelos, 30 llegaron a su destino con un retraso de más de quince minutos.

    (a) Calcule un intervalo de confianza del 95% para la varianza poblacional del número de reservas.

    (b) Calcule un intervalo de confianza del 95% para la proporción poblacional de vuelos que llegan con un retraso de más de 15 minutos.

    (c) En base a los resultados de los apartados anteriores, indique si las siguientes afirmaciones son verdaderas o falsas razonando la respuesta:

        i. Basándonos en la información muestral y suponiendo que el valor estimado de la proporción no varía, para construir un intervalo de confianza al 95% para la proporción poblacional de vuelos en la ruta que llegan con un retraso de más de quince minutos con una longitud total de 0.1 necesito tomar una muestra de al menos 200 observaciones.

        ii. El intervalo de confianza para la varianza poblacional del número de reservas para un vuelo contiene el valor 1000, para un nivel de confianza del 99%.

        iii. Para algunos valores muestrales, el intervalo de confianza para la proporción poblacional de vuelos que llegan con un retraso de más de quince minutos podría incluir valores negativos.

19. Una clínica ofrece tratamientos de reducción de peso. Se supone que la pérdida de peso observada tras un tratamiento de dos meses sigue una distribución normal. Se ha recogido una muestra aleatoria simple de 40 pacientes, y los valores observados pueden encontrarse en el fichero de Excel "datos_ej1.xlsx"

    Utilice Excel para responder a las preguntas siguientes:

    (a) Calcule un intervalo de confianza al 99% para la pérdida de peso de todos los pacientes que siguen el tratamiento.

    (b) ¿Cuál sería el valor del intervalo si el nivel de confianza deseado fuera del 90%?

    (c) Calcule un intervalo de confianza al 99% para la varianza de la pérdida de peso.

20. Se ha recogido información de 46 clientes importantes de una empresa en relación con sus expectativas de cambio en su demanda para el año próximo. Esta información está disponible en el fichero Excel "datos_ej1.xlsx" (Muestra 1). Suponemos que estas previsiones son una muestra aleatoria simple, y que la población correspondiente sigue una distribución normal.

    (a) Calcule un intervalo de confianza al 90% y al 99% para el cambio en la demanda en el año próximo.

    (b) Calcule un intervalo de confianza al 99% para la desviación típica de la población (el cambio estimado en la demanda).

    (c) Con posterioridad, se han recibido 18 nuevas respuestas sobre expectativas de cambio (Muestra 2). Se supone que estos valores corresponden a la misma población y son independientes de los anteriores. ¿Cuáles serían los valores revisados de los intervalos de confianza (para un nivel del 99%) si se utilizase toda la información disponible?

    (d) ¿Sería razonable esperar que la demanda no cambiase en el año próximo?

21. La longitud del intervalo para la proporción en la población, p_X, para un nivel de confianza de $(1-\alpha)$ viene dada por $w = 2z_{\alpha/2}\sqrt{\hat{p}_x(1-\hat{p}_x)/n}$. Para un nivel de confianza dado y un tamaño muestral n fijo, represente esta longitud en función del valor de la proporción observada en la muestra $\hat{p}_x$ utilizando Excel (por ejemplo, tome $\alpha = 0.05$ y $n = 100$).

    ¿Para que valor de $\hat{p}_x$ se obtiene la mayor longitud del intervalo?

22. Para los datos del Problema 1, calcule un intervalo de confianza al 95% para la media de la población μ_X (¿qué hipótesis son necesarias?) y un intervalo de confianza al 90% para la media de la población μ_Y (¿qué hipótesis son necesarias?) utilizando Excel.

23. Sea $\chi^2_{n;\alpha}$ un cuantil $\alpha \in (0,1)$ de una distribución χ^2_n, esto es, $\chi^2_{n;\alpha}$ cumple $P(\chi^2_n > \chi^2_{n;\alpha}) = \alpha$. Para un valor fijo de n (por ejemplo, $n = 10$) represente $\chi^2_{n;\alpha}$ en función de α en Excel.