EXAMEN DE COMPETENCIAS FUNDAMENTALES:

Ejercicios resueltos de Probabilidades y Estadística

Elaboración por Constanza Prado
Revisión por Daniela Hurtado

Problema 1. El estudio de los métodos de organización, resumen y presentación de datos estadísticos forma parte de la rama de la estadística que se denomina:

(a) Estadística descriptiva.

(b) Muestreo.

(c) Estadística Inferencial.

(d) Análisis de datos.

Solución:

$\boxed{\text{La respuesta es (a)}}$

Problema 2. En un estudio se observan los tiempos de permanencia hospitalaria (en días) de una muestra de 6 niños. Los datos registrados son: 2, 7, 5, 8 y 1. El sexto niño presentó complicaciones, por lo que pemaneció hospitalizado durante 60 días. ¿Cuál será el efecto del tiempo de permanencia hospitalaria de este niño sobre los indicadores de resumen?

(a) Un incremento en la mediana.

(b) Un incremento en la moda.

(c) Un incremento en la media aritmética.

(d) Un incremento en la media, moda y mediana.

Solución:

$\boxed{\text{La respuesta es (c)}}$

Problema 3. En una serie de mediciones del nivel de glucosa (mg/dl) de una muestra de individuos, el grado de dispersión de los valores alrededor del centro se pueden expresar en términos de

(a) Mediana.

(b) Rango.

(c) Media.

(d) Desviación estándar.

Solución:

$\boxed{\text{La respuesta es (d)}}$

Problema 4. Las siguientes medidas son todas de tendencia central, excepto

(a) La media.

(b) La mediana.

(c) El coeficiente de variación.

(d) El cuartil 2.

Solución:

$\boxed{\text{La respuesta es (c)}}$

Problema 5. ¿Cuál es la mediana, segundo cuartil y el rango de 9, 3, 44, 15, 6 ?

(a) La mediana es 9, el segundo cuartil es 9 y el rango es 41.

(b) La mediana es 9, el segundo cuartil 44 y el rango es 3.

(c) La mediana es 15.4, el segundo cuartil es 9 y el rango es 41.

(d) La mediana es 15.4, el segundo cuartil 15 y el rango 38.

Solución:

$\boxed{\text{La respuesta es (a)}}$

Problema 6. Los alumnos que tomaron un curso con el Profesor A tuvieron una desviación estándar de 2.4 en una prueba, mientras que los alumnos que tomaron el curso con el Profesor B tuvieron una desviación estándar de 1.2 en la misma prueba. ¿Qué se puede afirmar sobre las notas en ambos grupos de alumnos?.

(a) Los alumnos del Profesor A son más homogéneos que los del Profesor B.

(b) Los alumnos del Profesor B son menos heterogéneos que los del Profesor A.

(c) A los alumnos del Profesor B les fue peor en la prueba que a los del Profesor A.

(d) A los alumnos del Profesor A les fue el doble mejor que a los alumnos del Profesor B.

<u>Solución</u>:

$\boxed{\text{La respuesta es (b)}}$

Problema 7. ¿Qué dice el teorema del límite central?

(a) Dice que bajo un muestreo aleatorio simple desde una población con media μ y varianza finita σ^2, la distribución de la media muestral converge a una distribución normal con media μ y varianza σ^2.

(b) Dice que bajo un muestreo aleatorio simple desde una población con media μ y varianza finita σ^2, la distribución de la media muestral converge a una distribución normal con media μ/n y varianza σ^2/n^2, donde n es el tamaño de la muestra.

(c) Dice que bajo un muestreo aleatorio simple desde una población con media μ y varianza finita σ^2, la distribución de la media muestral converge a una distribución normal con media μ y varianza σ^2/n, donde n es el tamaño de la muestra.

(d) Dice que bajo un muestreo aleatorio simple desde una población con media μ y varianza finita σ^2, la distribución de la media muestral converge a una distribución normal con media μ/n y varianza σ^2/n, donde n es el tamaño de la muestra.

<u>Solución</u>:

$\boxed{\text{La respuesta es (c)}}$

Problema 8. Suponga que p representa la probabilidad de que en un embarazo nazca un varón. Una pareja desea tener exactamente dos niñas en su familia. La familia tendrá hijos hasta que se satisfaga la condición.

¿Cuántos varones se esperarían que tuviera esta familia sabiendo que $p = \frac{1}{2}$?.

(a) 1

(b) 2

(c) 3

(d) 4

Solución:

Sea X = número de varones que tiene una familia hasta que hayan 2 mujeres, luego

$$X \sim BinNeg\left(r = 2, p = \frac{1}{2}\right)$$

La esperanza de una distribución Binomial Negativa está dada por

$$\mathbb{E}(X) = \frac{r}{p}(1-p) = \frac{2}{\frac{1}{2}}\left(\frac{1}{2}\right) = 2$$

$\boxed{\text{La respuesta es (b)}}$

Problema 9. Un peaje cobra tarifas distintas para autobuses de pasajeros y otros vehículos particulares. Suponga que durante una hora, el 60 % de todos los vehículos son autobuses de pasajeros.
Si los autobuses de pasajeros llevan un promedio de 15 pasajeros y los particulares un promedio de 3 pasajeros. ¿Cuál es la probabilidad de que un vehículo que se detiene en el peaje lleve 4 pasajeros?

(a) 0,068

(b) 0,101

(c) 0,168

(d) 0,263

Solución:

Sea X =número de pasajeros que lleva un vehículo que pasa en un peaje durante un hora y sean los eventos:

A : El vehículo es un Autobus.
P : El vehículo es particular.

Luego,

$$X|A \sim Poi(\lambda = 15)$$
$$X|P \sim Poi(\lambda = 3)$$

Por ley de probabilidades totales,

$$\begin{aligned} P(X=4) &= P(X=4|A)P(A) + P(X=4|P)P(P) \\ &= \frac{15^4 e^{-15}}{4!} \times 0{,}6 + \frac{3^4 e^{-3}}{4!} \times 0{,}4 \\ &= 0{,}068 \end{aligned}$$

$\boxed{\text{La respuesta es (a)}}$

Problema 10. Se ha comprobado que el tiempo de vida de un marcapasos es una variable aleatoria con función de distribución acumulada

$$F_X(x) = 1 - e^{-\lambda x}, \quad x > 0$$

Donde $\lambda = \frac{1}{20}$ si el marcapasos es de tipo A y $\lambda = \frac{1}{16}$ si el marcapasos es de tipo B.
En un hospital no se tiene registro del tipo de marcapasos que se le ha implantado a un paciente, sólo se sabe que el 30 % de ellos usa marcapasos de tipo B.
Si se selecciona un paciente al azar. ¿Cuál es la probabilidad de que su marcapasos dure más de 30 años?.

(a) 0,174

(b) 0,825

(c) 0,798

(d) 0,202

Solución:

Por ley de probabilidades totales

$$\begin{aligned}P(X>30) &= P(X>30|A)P(A)+P(X>30|B)P(B)\\ &= e^{-30/20}\times 0,7+e^{-30/16}\times 0,3\\ &= 0,202\end{aligned}$$

$\boxed{\text{La respuesta es (d)}}$

Problema 11. Con respecto a la pregunta anterior. Si se examinan de manera sucesiva e independiente pacientes con marcapasos. ¿Cuál es la probabilidad de que se deba examinar 5 pacientes hasta encontrar 3 cuyo marcapasos tenga una duración superior a 30 años?.

(a) $3,2\times 10^{-9}$

(b) $7,68\times 10^{-5}$

(c) $4,61\times 10^{-5}$

(d) $0,004$

Solución:

Sea Y = número de pacientes examinados hasta tener 3 cuyo marcapasos tiene duración superior a 30 años, luego

$$Y\sim BinNeg(r=3,p=0,202)$$

Se pide calcular

$$\begin{aligned}P(Y=5) &= \binom{4}{2}0,02^3 0,98^2\\ &= 4,61\times 10^{-5}\end{aligned}$$

$\boxed{\text{La respuesta es (c)}}$

Problema 12. Para cualquier distribución de probabilidad, se define el Rango Intercuartil como la distancia que hay entre los percentiles 25 y 75. Calule el Rango Intercuartil para una distribución $N(\mu, \sigma^2)$.

(a) $0,25\sigma$

(b) $0,5\sigma$

(c) $0,67\sigma$

(d) 1,34σ

<u>Solución</u>:

Se quiere buscar (a, b) tal que $P(X < a) = 0,25$ y $P(X < b) = 0,75$, con $X \sim N(\mu, \sigma^2)$.

$$\begin{aligned}
P(X < b) &= 0{,}75 \\
P\left(Z < \frac{b-\mu}{\sigma}\right) &= 0{,}75 \\
\frac{b-\mu}{\sigma} &= 0{,}67 \\
b &= \mu + 0{,}67\sigma \\
P(X < a) &= 0{,}25 \\
P\left(Z < \frac{a-\mu}{\sigma}\right) &= 0{,}25 \\
\frac{a-\mu}{\sigma} &= -0{,}67 \\
a &= \mu - 0{,}67\sigma
\end{aligned}$$

Luego, el rango intercuartil está dado por

$$b - a = \mu + 0{,}67\sigma - \mu + 0{,}67\sigma = 1{,}34\sigma$$

$\boxed{\text{La respuesta es (d)}}$

Problema 13. Se efectuaron 10 mediciones independientes $X_1, X_2, ..., X_{10}$ en una población donde cada una de ellas sigue una distribución $N(\mu, \sigma^2)$. Suponga que estas mediciones satisfacen la propiedad que los eventos $(X_1 \geq c_1), (X_2 \geq c_2), ..., (X_{10} \geq c_{10})$ son independientes para cualquier $c_1, ..., c_{10} \in \mathbb{R}$. ¿Cuál es la probabilidad de que a lo menos 2 de ellas superan un valor $c = \mu + 2{,}2\sigma$?.

(a) 0,008

(b) 0,099

(c) 0,003

(d) 1

Solución:
Sea $Y =$ el número de observaciones que superan el valor $c = \mu + 2{,}2\sigma$, luego

$$Y \sim Binom(n = 10, p)$$

Con $p = P(X > \mu + 2,2\sigma) = 1 - P(Z < 2,2) = 1 - 0,9861 = 0,0139$.
Se pide calcular

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y = 0) - P(Y = 1) \\ &= 1 - \binom{10}{0} 0,0139^0 (1 - 0,0139)^1 0 - \binom{10}{1} 0,0139^1 (1 - 0,0139)^9 \\ &= 0,008 \end{aligned}$$

$\boxed{\text{La respuesta es (a)}}$

Problema 14. Un maestro de una universidad nunca termina su clase antes de finalizada la hora y siempre concluye su clase dentro de dos minutos después de la hora. Sea X el tiempo que transcurre entre el final de la hora y el fin de la clase, y suponga que su función de densidas está dada por

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq 2 \\ 0 & e.o.c \end{cases}$$

¿Cuál es la probabilidad de que la clase continúe más allá de la hora entre 60 y 90 segundos?.

(a) 0,297

(b) 0,375

(c) 0,422

(d) 0,125

Solución:

Primero, se debe calcular el valor de k,

$$\begin{aligned}\int_0^2 kx^2dx &= \frac{k8}{3}\\ &= 1\\ k &= \frac{3}{8}\end{aligned}$$

Segundo, se procede a calcular la probabilidad que se pide

$$\begin{aligned}P(1 < X < 1{,}5) &= \int_1^{1,5} \frac{3}{8}x^2dx\\ &= 0,297\end{aligned}$$

$\boxed{\text{La respuesta es (a)}}$

Problema 15. Suponga que el 25 % de los automovilistas se detienen por completo en un cruce con luces rojas intermitentes en todas las direcciones cuando no está visible ningún otro automóvil. ¿Cuál es la probabilidad de que de 20 automovilistas seleccionados al azar que llegan a la intersección, al menos seis se detengan por completo?.

(a) 0,214

(b) 0,383

(c) 0,786

(d) 0,617

Solución:

Sea X =número de vehículos que se detienen por completo en un cruce. Luego

$$X \sim Binom(n = 20, p = 0{,}25)$$

Se pide $P(X \geq 6) = 1 - P(X \leq 5) = 0{,}383$.

$\boxed{\text{La respuesta es (b)}}$

Problema 16. Considere X como una variable aleatoria que indica la cantidad de sellos obtenidos al lanzar una moneda cargada n veces (con probabilidad p de obtener sello).

¿Cuál es la distribución de la variable aleatoria X? ¿Cuál es la probabilidad de que se hayan obtenido n sellos?.

(a) $X \sim BinNeg(r = n, p)$ y $(1 - p)^n$

(b) $X \sim BinNeg(r = n, p)$ y p^n

(c) $X \sim Binom(n, p)$ y p^n

(d) $X \sim Binom(n, p)$ y $(1 - p)^n$

Solución:

$\boxed{\text{La respuesta es (c)}}$

Problema 17. Considere la variable aleatoria Y que denota la cantidad de lanzamientos obtenidos hasta obtenerse el primer sello (con probabilidad p de obtener sello en un lanzamiento). ¿Cuál es la probabilidad de que se haya lanzado la moneda y veces ?.

(a) $p^{y-1}(1 - p)$

(b) $p^y(1 - p)$

(c) $p(1 - p)^y$

(d) $p(1 - p)^{y-1}$

Solución:

Sea X =número de lanzamientos hasta obtener el primer sello. Luego

$$X \sim Geo(p)$$

Por lo tanto

$$P(X = y) = p(1 - p)^{y-1}$$

$\boxed{\text{La respuesta es (d)}}$

Problema 18. Suponga que X es una variable aleatoria con distribución Geométrica con parámetro p.
Sea $s, t > 0$, ¿Cuál de las alternativas es equivalente a la expresión?

$$P(X > s + t | X > t)$$

(a) $P(X \geq t)$

(b) $P(X > t)$

(c) $P(X > s)$

(d) $P(X \geq s)$

Solución:

Primero, se calcula $P(X > T)$, para un T cualquiera

$$\begin{aligned}
P(X > T) &= \sum_{x=T+1}^{\infty} p(1-p)^{x-1} \\
&= \sum_{u=0}^{\infty} p(1-p)^{u+T} \\
&= p(1-p)^T \sum_{u=0}^{\infty} (1-p)^u \\
&= p(1-p)^T \frac{1}{1-(1-p)} \\
&= (1-p)^T
\end{aligned}$$

Luego,

$$\begin{aligned}
P(X > t + s | X > t) &= \frac{P(X > s + t, X > t)}{P(X > t)} \\
&= \frac{P(X > s + t)}{P(X > t)} \\
&= \frac{(1-p)^{s+t}}{(1-p)^t} \\
&= (1-p)^s \\
&= P(X > s)
\end{aligned}$$

$\boxed{\text{La respuesta es (c)}}$

Problema 19. Estudios realizados recientemente por una AFP han revelado que el tiempo de duración de sus cuentas de ahorro voluntario de afiliados sigue un comportamiento normal con media de 24 meses y una desviación estándar de 10.2 meses.
Si se escogen 5 personas afiliadas con cuenta de ahorro voluntario, ¿cuál es la probabilidad de que a lo más 1 de ellas mantenga la cuenta por menos de 24 meses?

(a) 0,1875

(b) 0,8125

(c) 0,96875

(d) 0,03125

Solución:

Sea Y =número de personas afiliadas con cuenta de ahorro que tiene su cuenta por lo menos 24 meses. Luego

$$Y \sim Binom(n = 5, p)$$

Con p dado por

$$\begin{aligned} P(X > 24) &= P\left(Z > \frac{24 - 24}{10{,}2}\right) \\ &= P(Z > 0) \\ &= 0,5 \end{aligned}$$

Se pide calcular

$$P(Y \leq 1) = P(Y = 0) + P(Y = 1) = 0{,}1875$$

$\boxed{\text{La respuesta es (a)}}$

Problema 20. El número de solicitudes de auxilio que recibe un servicio de remolque de vehículos sigue una distribución Poisson con tasa $\lambda = 4$ por hora.
Si los operadores de las grúas toman un receso de 30 minutos para almorzar. ¿Cuál es la probabilidad de que no pierdan ninguna llamada de auxilio?.

(a) 0,406

(b) 0,018

(c) 0,981

(d) 0,135

Solución:

Sea X =número de llamadas recibidas en una hora. Luego

$$X \sim Poi(\lambda = 4)$$

Como el receso es de media hora, la tasa de tiempo cambia a $\lambda = 4 \times \frac{1}{2} = 2$. Se pide calcular

$$P(X = 0|\lambda = 2) = \frac{2^0 e^{-2}}{0!} = 0,135$$

La respuesta es (d)

Problema 21. Suponga que el número de horas X que funcionará una máquina antes de fallar es una variable aleatoria con distribución Normal de parámetros $\mu = 720$ y $\sigma^2 = 480^2$. Suponga que en el momento en que la máquina comienza a funcionar, usted debe decidir cuándo el inspector regresará a revisarla. Si vuelve antes de que la máquina falle, se ocasiona un costo de a dólares por haber desperdiciado una inspección. Si vuelve después de que la máquina ha fallado, se ocasiona un costo de b dólares por el no funcionamiento de la máquina.

Se observa este proceso durante 15 períodos y suponga que el inspector siempre decide volver en un tiempo de $t = 816$. ¿Cuál será la esperanza de la variable Y = Número de veces que el inspector llega tarde?

(a) $8,6895$

(b) $0,5793$

(c) $8,8695$

(d) $0,5973$

Solución:

Se puede concluir que $Y \sim Binom(n = 15, p)$, con p dado por

$$p = P(X < 816) = P(Z < 0,2) = 0,5793$$

Por lo tanto, la esperanza está dada por

$$\mathbb{E}(Y) = 15 \times 0,5793 = 8,6895$$

La respuesta es (a)

Problema 22. Sea $X_1, X_2, ..., X_n$ una muestra aleatoria del tiempo de falla (en horas) de n componentes electrónicos. Suponga que el tiempo de falla de este tipo de componentes sigue una distribucipon Weibull$(\lambda, 2)$, cuya función de densidad de probabilidad está dada por:

$$f(x) = \frac{2x}{\lambda} \exp\left(-\frac{x^2}{\lambda}\right), \quad x > 0, \quad \lambda > 0$$

¿Cuál será el estimador de máxima verosimilitud de λ?

(a) $\hat{\lambda} = \bar{X}$

(b) $\hat{\lambda} = \sum_{i=1}^{n} x_i^2$

(c) $\hat{\lambda} = \sum_{i=1}^{n} x_i$

(d) Ninguna se las anteriores

Solución:

No hay forma directa de despejar λ es la función de log-verosimilitud.

$\boxed{\text{La respuesta es (d)}}$

Problema 23. Suponga que se observó una muestra de tamaño $n = 50$ proveniente de una población con distribución $N(\mu, \sigma^2)$, donde $\sum_{i=1}^{n} X_i^2 = 1559$ y $\sum_{i=1}^{n} X_i = 32$. Entregue un intervalo de 95 % de confianza para μ.

(a) $(-0,653\ ,\ 1,933)$

(b) $(-0,897\ ,\ 2,177)$

(c) $(-0,977\ ,\ 2,257)$

(d) $(-0,953\ ,\ 2,233)$

Solución:

El intervalo de confianza para la media con varianza desconocida está dada por

$$\bar{x} - t_{(49,0,975)} \times \frac{S}{\sqrt{n}} \quad , \quad \bar{x} + t_{(49,0,975)} \times \frac{S}{\sqrt{n}}$$

Reemplazando, se tiene que

$$\bar{X} = \frac{32}{50} = 0,64$$

$$S^2 = \frac{1}{50-1}\left(1559 - 50\times(0{,}64)^2\right) = 31{,}39837$$

$$t_{0,975,49} = 2{,}01$$

Luego, el IC está dado por $(-0{,}953\ ,\ 2{,}233)$

$\boxed{\text{La respuesta es (b)}}$

Problema 24. Sea $X_1, X_2, ..., X_n$ una muestra de una población X cuya función de densidad es

$$f(x) = e^{-x+\theta},\ \ x \geq \theta$$

con $\mathbb{E}(X) = \theta + 1$. ¿Cuáles serán sus estimadores de momentos y de máxima verosimilitud respectivamente?

(a) $\hat{\theta}_M = \bar{X} + 1$ y $\hat{\theta}_{EMV} = \text{mín}\{X_i\}$, $i = 1, ..., n$.

(b) $\hat{\theta}_M = \bar{X} - 1$ y $\hat{\theta}_{EMV} = \text{mín}\{X_i\}$, $i = 1, ..., n$.

(c) $\hat{\theta}_M = \bar{X} - 1$ y $\hat{\theta}_{EMV} = \text{máx}\{X_i\}$, $i = 1, ..., n$.

(d) $\hat{\theta}_M = \bar{X} + 1$ y $\hat{\theta}_{EMV} = \text{máx}\{X_i\}$, $i = 1, ..., n$.

<u>Solución</u>:

El estimador de momentos, plantea que

$$\mathbb{E}(X) = \theta + 1 = \bar{X}$$

Despejando

$$\hat{\theta} = \bar{X} - 1$$

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} e^{-x_i+\theta} 1_{\{x_i \geq \theta\}} \\ &= e^{-\sum_{i=1}^{n} x_i} e^{n\theta} 1_{\{min\{x_i\} \geq \theta\}} \end{aligned}$$

La función de arriba se maximiza cuando θ toma el valor más grande de su recorrido, es decir, cuando toma el valor de $\text{mín}\{x_i\}$.

$\boxed{\text{La respuesta es (b)}}$

Problema 25. Suponga que una variable aleatoria discreta X tiene una función de probabilidad $f(x|\theta)$, $\theta \in \{1, 2, 3\}$ dada por

x	$f(x\|\theta = 1)$	$f(x\|\theta = 2)$	$f(x\|\theta = 3)$
0	$\frac{1}{3}$	$\frac{1}{4}$	0
1	$\frac{1}{3}$	$\frac{1}{4}$	0
2	0	$\frac{1}{4}$	$\frac{1}{4}$
3	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{2}$
4	$\frac{1}{6}$	0	$\frac{1}{4}$

Suponga que se observa $X = 3$. ¿Cuál es la estimación máximo verosímil de θ?

(a) $\hat{\theta} = 1$.

(b) $\hat{\theta} = 2$.

(c) $\hat{\theta} = 3$.

(d) $\hat{\theta} = 4$.

Solución:

Se comparan los resultados de las funciones de densidad cuando $X = 3$. El valor máximo de la función de densidad es $\frac{1}{2}$, que se obtiene cuando $\theta = 3$. Por lo tanto, $\hat{\theta} = 3$.

$\boxed{\text{La respuesta es (c)}}$

Problema 26. Se sabe que la probabilidad de que cualquier persona de una población tenga una enfermedad dada es $\theta \geq 0{,}5$. Con el objetivo de conocer exactamente el valor de esta probabilidad, se toma una muestra aleatoria de n personas y se regista

$$Y_i = \begin{cases} 1 & \text{Si la persona } i \text{ está enferma} \\ 0 & \text{Si la persona } i \text{ no está enferma} \end{cases}$$

Suponga que en una muestra de tamaño 5 se observan dos enfermos. ¿Cuál es la estimación de momentos para θ?

(a) 0,2

(b) 0,3

(c) 0,4

(d) 0,5

Solución:

El estimador de momentos está dado por $\bar{x}$. Por lo tanto, la estimación será $\frac{2}{5} = 0{,}4$.

$\boxed{\text{La respuesta es (c)}}$

Problema 27. Sea X_i el número de clientes que solicitan información a una empresa constructora durante el día i. Asuma que $X_1, X_2, ..., X_n \overset{iid}{\sim}$ Poisson(λ). Se quiere saber el npumero esperado de clientes que solicitan información en un día. Para esto, se tomó una muestra aleatoria dutante 50 días de la cantidad de clientes que llegaron por día, donde se obtuvo

Número de clientes por día	0	1	2	3	4
Cantidad de días observados	17	22	7	3	1

Encuentre la estimación máximo verosímil de la probabilidad de que el número de clientes durante un día sea 0.

(a) 0,397

(b) 0,283

(c) 0,98

(d) 0,375

Solución:

$P(X = 0) = e^{-\lambda}$. Como no se conoce λ, este debe ser estimado. Se sabe que $\hat{\lambda} = \bar{x}$. Para este caso,

$$\bar{x} = \frac{1}{50}(22 + 14 + 9 + 4) = \frac{49}{50}$$

Por lo tanto,

$$P(X = 0) = e^{-49/50} = 0,375$$

$\boxed{\text{La respuesta es (d)}}$

Problema 28. Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de una población con distribución Normal(μ, σ^2).
Suponga que se ha observado una muestra con los siguientes datos

23,0 24,3 24,5 25,1 25,5 26,9 28,3 30,0 31,4 32,0 34,4 34,8 38,5 39,4

¿Cuál será una estimación intervalar para μ utilizando un 98 % de confianza?

(a) $(27,37\ ,\ 33,18)$

(b) $(26,87\ ,\ 33,68)$

(c) $(26,30\ ,\ 34,25)$

(d) $(26,89\ ,\ 33,65)$

Solución:

Se tiene que

$$\bar{x} = 30{,}27692$$
$$S^2 = 28{,}85192$$
$$t_{(0{,}99{,}12)} = 2{,}68$$

Así, el IC está dado por

$$30{,}27692 - 2{,}68 \times \sqrt{28{,}58192/13}\ ,\ 30{,}27692 + 2{,}68 \times \sqrt{28{,}58192/13}$$

Que es igual a $(26,30\ ,\ 34,25)$

$\boxed{\text{La respuesta es (c)}}$

Problema 29. Se realizó un experimento considerando 61 pacientes varones de similares características que llegan a un servicio de urgencia con fuertes dolores producidos por cálculos renales. Se les suministró una dosis de 5ml. de un nuevo fármaco para calcular tales dolores, midiéndose el tiempo transcurrido hasta que el dolor desaparece completamente. Los resultados del experimento entregaron los siguientes resultados

$$\bar{X} = 20 \text{ minutos}, \qquad S = 5 \text{ minutos}$$

Además, 7 pacientes reaccionaron negativamente por la dosis.
Mediante un intervalo de confianza del 95 % encuentre los límites que permitan estimar el tiempo que tarda el medicamento en eliminar el dolor. (Asuma que el tiempo medio transcurrido hasta que el dolor desaparezca completamente tiene una distribución normal).

(a) $(18,74\ ,\ 21,25)$

(b) $(18,94\ ,\ 21,05)$

(c) $(18,72\ ,\ 21,28)$

(d) $(18,65\ ,\ 21,34)$

Solución:

$$20 - t_{(0{,}975{,}60)} \times \sqrt{25/61} = 18,72$$
$$20 + t_{(0{,}975{,}60)} \times \sqrt{25/61} = 21,28$$

$\boxed{\text{La respuesta es (c)}}$

Problema 30. Con respecto al problema anterior. Si se toma en consideración la información recopilada hasta el momento y se desea construir un intervalo con 90 % de confianza para la proporción de caos que reacciona negativamente, de tal manera de lograr un error de estimación del 3 % como máximo. ¿Cuál es la cantidad mínima de pacientes que debe construir el grupo experimental?

(a) 304

(b) 303

(c) 302

(d) 300

Solución:

El error de estimación está dado por

$$z_{1-\alpha/2} \times \sqrt{p(1-p)/n} \leq 0{,}03$$

Donde p es la probabilidad de que un paciente reaccione negativamente al tratamiento. La estimación de p es $7/61$. Luego,

$$z_{(0{,}95)} \times \sqrt{\left(\frac{\frac{7}{61}\frac{54}{61}}{n}\right)} \leq 0{,}03$$

$$1{,}64^2 \frac{7}{61}\frac{54}{61}\frac{1}{0{,}03^2} \leq n$$
$$303{,}5829 \leq n$$

Es decir, $n = 304$.

$\boxed{\text{La respuesta es (a)}}$

Problema 31. Se tiene una muestra aleatoria de 48 alumnos que rindieron una prueba que cuenta con los puntajes obtenidos respectivos. Se sabe que

$$\sum_{i=1}^{n} X_i = 387{,}8 \quad \sum_{i=1}^{n} X_i^2 = 4247{,}08$$

¿Cuál será un intervalo de confianza del 95 % para el puntaje medio de la prueba?

(a) (6,60 , 9,50)

(b) (6,70 , 9,45)

(c) (6,65 , 9,48)

(d) (6,67 , 9,49)

Solución:

$$(387{,}8/48 - t_{(0{,}975{,}47)}\sqrt{S^2/48}\ ,\ 387{,}8 + t_{(0{,}975{,}47)}\sqrt{S^2/48})$$

Donde

$$S^2 = \frac{1}{47}\left(4247{,}08 - 48(387/48)^2\right) = 23{,}70168$$

$$t_{(0{,}975{,}47)} = 2{,}01$$

Así, el IC está dado por

$$(6{,}67\ ,\ 9{,}49)$$

$\boxed{\text{La respuesta es (d)}}$

Problema 32. Una reconocida línea aérea ha desarrollado un plan para un grupo de clientes llamado Çlub Viajero Ejecutivo"sobre la premisa de que el 30 % de sus clientes cumplirán los requisitos para ser socios. Se toma una muestra aleatoria de n clientes, $X_1, X_2, ..., X_n$ provenientes de la distribución de la variable aleatoria X, donde

$$X = \begin{cases} 1 & \text{Si el cliente cumple los requisitos para ser socio del club} \\ 0 & \text{Si no} \end{cases}$$

Suponga que finalmente se toma una muestra aleatoria de 200 clientes, en la que se observó que 40 clientes sí cumplirán con los requisitos. Con estos datos, ¿Cuál será la esimación intervalar con un 95 % de confianza para el parámetro de la distribución de X? ¿Qué se puede decir acerca de la premisa de la línea aérea?

(a) El intervalo de confianza es (0,145 , 0,255) y estadísticamente, no se cumple la premisa

(b) El intervalo de confianza es (0,145 , 0,255) y estadísticamente, sí se cumple la premisa

(c) El intervalo de confianza es (0,155 , 0,355) y estadísticamente, sí se cumple la premisa

(d) El intervalo de confianza es (0,155 , 0,355) y estadísticamente, no se cumple la premisa

Solución:

El IC etsá dado por

$$\hat{p} - 1{,}96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0{,}2 - 1{,}96\sqrt{\frac{0{,}2 \times 0{,}8}{200}} = 0{,}145$$

$$\hat{p} + 1{,}96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0{,}2 - 1{,}96\sqrt{\frac{0{,}2 \times 0{,}8}{200}} = 0{,}255$$

Como el 0.3 no pertenece al intervalo, no se puede asegurar la premisa que plantea el problema.

$\boxed{\text{La respuesta es (a)}}$

Problema 33. Un fabricante de neumáticos afirma que, bajo condiciones normales, la vida promedio de sus neumáticos es mayor que 50.000 km. En una muestra aleatoria simple de 15 neumáticos sometidos a pruebas de duración, se obtiene un promedio de 54.000 km. con una desviación estándar de 10.000 km. Suponiendo que la duración de los neumáticos sigue una distribución normal, ¿cuál es el intervalo con un 99 % de confianza para la duración promedio de los neumáticos?.

(a) (46,306 , 61,694)

(b) (48,939 , 59,061)

(c) (47,235 , 60,764)

(d) (46,306 , 60,764)

Solución:

$$54 - t_{(0{,}995,14)}\sqrt{10^2/15} = 54 - 2{,}98 \times 2{,}581989 = 46{,}30567$$
$$54 + t_{(0{,}995,14)}\sqrt{10^2/15} = 54 + 2{,}98 \times 2{,}581989 = 61{,}69433$$

$\boxed{\text{La respuesta es (a)}}$

Problema 34. Con respecto a la pregunta anterior, ¿cuál será el resultado del test de hipótesis que contrasta la afirmación realizada por el fabricante con un 5 % de significancia?

(a) Se rechaza la hipótesis nula, ya que la media puede tomar valores entre 45 y 60.

(b) No se rechaza la hipótesis nula, ya que la media es mayor a 50.

(c) Se rechaza la hipótesis nula, ya que la media es mayor a 50.

(d) No se puede rechazar la hipótesis nula.

Solución:

El test plantea

$$H_0 : \mu = 50$$

$$H_0 : \mu > 50$$

Se rechaza la hipótesis nula si

$$T = \frac{54 - 50}{\sqrt{\frac{100}{15}}} > t_{(0{,}95{,}14)} = 1{,}76$$

(Al ser un test de una cola). Reemplazando, se tiene que

$$T = 1{,}55$$

Como no se cumple la condición, no se puede rechazar H_0.

$\boxed{\text{La respuesta es (d)}}$

Problema 35. Cuando en un test de hipótesis NO se rechaza la hipótesis nula, ¿qué es correcto?

(a) Se ha demostrado que la hipótesis nula es verdadera

(b) Se ha demostrado que la hipótesis nula es falsa

(c) Se ha demostrado que la hipótesis alternativa es verdadera

(d) No hay evidencia para demostrar nada

Solución:

$\boxed{\text{La respuesta es (d)}}$

Problema 36. Con el fin de conocer la incidencia sobre la contaminación de una determinada industria, se mide la concentración de monóxido de sulfuro en dos zonas diferentes de una ciudad. La tabla muestra las medidas obtenidas en unidades típicas de este contaminante

Zona												
A	86	82	75	61	89	64	81	68	65	80	90	63
B	87	74	63	55	76	70	69	57	53	72		

Suponiendo que las varianzas son conocidas e iguales a $\sigma^2 = 9$, ¿se puede rechazar la hipótesis de que la zona A está más contaminada que la zona B al nivel de significancia del 5 %?

(a) No se rechaza la hipótesis nula.

(b) Se rechaza la hipótesis nula.

(c) Se rechaza la hipótesis nula ya que el promedio muestral del contaminante A es mayor al de B.

(d) No se puede decir nada al respecto.

Solución:

$$H_0 : \mu_A - \mu_B = 0$$
$$H_1 : \mu_A - \mu_B > 0$$

Se rechaza H_0 si

$$Z_c = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{9}{12} + \frac{9}{10}}} > z_{(0,95)} = 1{,}65$$

$barX_A = 75{,}3$ y $\bar{X}_B = 67{,}3$. El resultado de $Z_c = 6{,}020392$. Luego, se rechaza la hipótesis nula .

$\boxed{\text{La respuesta es (b)}}$

Problema 37. Suponga que $X_1, X_2, ..., X_n \overset{idd}{\sim} N(\mu, 49)$ y que se quiere evaluar las hipótesis

$$H_0 : \mu = 80 \quad H_1 : \mu = 78$$

En base a este test, ¿qué decisión tomaría usted si se obtiene un promedio de 79.1 en una muestra de tamaño $n = 100$?

(a) Se rechaza la hipótesis nula con un 5 % de significancia ya que el promedio muestral es cercano a 78.

(b) No se rechaza la hipótesis nula con un 5 % de significancia ya que el promedio muestral es cercano a 80.

(c) Se rechaza la hipótesis nula con un 5 % de significancia ya que el promedio muestral es cercano a 80.

(d) No se puede rechazar la hipótesis nula con un 5 % de significancia.

Solución:

$$\begin{aligned}
\alpha &= P(\bar{x} < k|\mu = 80) \\
0{,}05 &= P(Z < \frac{k - 80}{\sqrt{(0{,}49)}}) \\
-1{,}65 &= \frac{k - 80}{\sqrt{0{,}49}} \\
k &= 78{,}845
\end{aligned}$$

Luego, la región de rechazo está dada por

$$R = \{x : \bar{x} < 78{,}845\}$$

Como $\bar{x} = 79{,}1$, no existe evidencia estadística para rechazar H_0.

$\boxed{\text{La respuesta es (d)}}$

Problema 38. Un programa dietético anuncia: "Pierda 20 kilos en 4 meses". En una muestra aleatoria de 25 clientes de este programa, se observó una pérdida media de 16 kilos y una desviación estándar de 6 kilos. ¿Qué puede decir respecto de la veracidad del anuncio?

(a) Es cierto, ya que se pierden menos de 20 kilos.

(b) Es falso, ya que la media es 16 kilos.

(c) Es estadísticamente falso con un 5 % de significancia.

(d) Es estadísticamente cierto con un 5 % de significancia.

Solución:

La hipótesis es

$$H_0 : \mu = 20$$
$$H_1 : \mu > 20$$

Se rechaza H_0 si

$$\frac{16-20}{\frac{6}{5}} = -3{,}333 > t_{(0{,}95{,}24)} = 1,71$$

Por lo tanto, no se puede rechazar H_0.

$\boxed{\text{La respuesta es (c)}}$

Problema 39. Una empresa de arriendo de vehículos dice que sus costos son menores en más de \$40 por km. que los de su competencia, la cual ofrece vehículos más cómodos. Para la próxima Cumbre del Mercosur a realizarse en Chile, el Estado piensa arrendar más de 100 vehículos, por lo cual decide realizar un experimento. Arrienda 10 vehículos a cada compañía, seleccionando al azar entre todos los disponibles. Los vehículos son asignados al azar a 20 choferes de su staff de choferes disponibles para la Cumbre. Los gastos de operación por km. son registrados y se muestran en la siguiente tabla (auma normalidad y varianzas conocidas).

Compañía	Promedio	Desviación Estándar
A: menor costo	250	16
B: mayor comodidad	300	14

El estado chileno decidirá por los vehículos más económicos si estadísticamente los gastos de operación por km. son inferiores a \$40. A través de un test de nivel $\alpha 5\,\%$ ¿qué le diría al estado chileno?

(a) Existe evidencia estadística para decidir por los vehículos de menor costo.

(b) Existe evidencia estadística para decidir por los vehículos de mayor comodidad.

(c) No existe evidencia estadística para decidir por los vehículos de menor costo.

(d) No se puede decir nada al respecto.

Solución:

Las hipótesis son

$$H_0 : \mu_A - \mu_B = 40$$
$$H_0 : \mu_A - \mu_B < 40$$

Se rechaza H_0 si

$$\frac{250 - 300}{\sqrt{\frac{16^2}{10} + \frac{14^2}{10}}} = -7{,}437051 < z_{(0{,}05)} = -1{,}65$$

Se rechaza H_0.

$\boxed{\text{La respuesta es (a)}}$

Problema 40. Se hizo un estudio para definirse si los ejercicios aeróbicos reducen el ritmo cardíaco de una persona durante el descanso, y al examinar 10 voluntarios antes y después de seguir un programa de ese tipo durante 6 meses, sus pulsaciones, en latidos por minuto, dieron los siguientes resultados

Voluntario	1	2	3	4	5	6	7	8	9	10
Antes	73	77	68	62	72	80	76	64	70	72
Después	68	72	64	60	71	77	74	60	64	68

¿Los ejercicios aeróbicos reducen el ritmo cardíaco durante el reposo? utilice 5 % de significancia.

(a) Estadísticamente, se puede decir que los ejercicios aeróbico reducen el ritmo cardíaco con un 5 % de significancia.

(b) Estadísticamente, no se puede decir que los ejercicios aeróbico reducen el ritmo cardíaco con un 5 % de significancia.

(c) Estadísticamente, se puede decir que los ejercicios aeróbico reducen el ritmo cardíaco con un 95 % de significancia.

(d) Estadísticamente, no se puede decir que los ejercicios aeróbico reducen el ritmo cardíaco con un 95 % de confianza.

Solución:

Para realizar esto, se debe considerar la variable X_d = Diferencia entre las pulsaciones de antes y después. Luego,

$$H_0 : \mu_d = 0$$

$$H_1 : \mu_d > 0$$

Los datos serían

Voluntario	1	2	3	4	5	6	7	8	9	10
Diferencia	5	5	4	2	1	3	2	4	6	4

Así

$$\bar{X}_d = 3{,}6$$

$$S_d^2 = 2{,}488889$$

Se rechaza H_0 si

$$\frac{3{,}6 - 0}{\sqrt{2{,}488889/10}} = 7{,}216053 > t_{(0{,}95{,}9)} = 1{,}83$$

Luego, se rechaza H_0.

$\boxed{\text{La respuesta es (a)}}$

Problema 41. Considere las siguientes observaciones muestrales acerca de la viscosidad estabilizada de especímenes de asfalto

$$2781, \quad 2900, \quad 3013, \quad 2856, \quad 2888$$

Suponga que para una aplicación particular, se requiere que la viscosidad promedio real sea 3000. ¿Al parecer se ha cumplido este requerimiento? utilice un 10 % de significancia.

(a) Estadísticamente, se puede decir que se cumple el requerimiento, es decir, se rechaza H_0.

(b) Estadísticamente, se puede decir que la viscosidad promedio es superior a 3000.

(c) Estadísticamente, se puede decir que no se cumple el requerimiento, pues, se rechaza H_0.

(d) Estadísticamente, se puede decir que se cumple el requerimiento, pues no se rechaza H_0.

Solución:

Se rechaza H_0 si

$$\left| \frac{\bar{X} - 3000}{\sqrt{S^2/5}} \right| = 2{,}991161 > t_{(0{,}95{,}4)} = 2{,}13$$

Donde

$$\bar{X} = 2887{,}6$$

$$S^2 = 7060{,}3$$

Luego, se rechaza H_0.

$\boxed{\text{La respuesta es (c)}}$

Problema 42. Suponga que se lanza una moneda honesta tres veces de manera sucesiva e independiente. Se definen las variables aleatorias X : Número de caras observadas en el primer lanzamiento e Y : Numero total de caras observadas en los tres lanzamientos.

Si usted recibe \$300 por cada cara obtenida más \$200 adicionales si una de ellas fue en el primer lanzamiento. ¿Cuánto dinero espera ganar?

(a) 625

(b) 600

(c) 300

(d) 350

Solución:

La función de ganancia está dada por

$$g(x, y) = 300Y + 200X$$

La función de probabilidad de cada variable aleatoria es

$$P(X = x) = \begin{cases} 0{,}5 & x = 0, 1 \\ 0 & e.o.c \end{cases}$$

$$P(Y = y) = \begin{cases} 1/8 & y = 0, 3 \\ 3/8 & y = 1, 2 \\ 0 & e.o.c \end{cases}$$

Así,

$$\mathbb{E}(X) = 0{,}5$$

$$\mathbb{E}(Y) = \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = \frac{14}{8} = 1{,}75$$

Por lo tanto, la ganacia esperada será

$$\mathbb{E}(g) = 300 \times 1{,}75 + 200 \times 0{,}5 = 625$$

$\boxed{\text{La respuesta es (a)}}$

Problema 43. Considere el estimador de la varianza $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ para una muestra de tamaño n proveniente de una población con distribución f de media μ y varianza σ^2. ¿Qué valor tiene $\mathbb{E}(S^2)$?

(a) σ^2

(b) $n\sigma^2$

(c) $\frac{n}{n-1}\sigma^2$

(d) $\frac{n-1}{n}\sigma^2$

Solución:

$$\begin{aligned}
\mathbb{E}(S^2) &= \frac{1}{n}\left(\sum_{i=1}^{n}\mathbb{E}(X_i^2) - n\mathbb{E}(\bar{X})\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}(\sigma^2 + \mu^2) - n(\sigma^2/n + \mu^2)\right) \\
&= \frac{1}{n}\left(n\sigma^2 - \sigma^2\right) \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}$$

$\boxed{\text{La respuesta es (d)}}$

Problema 44. Considere el siguiente modelo

$$y_i = \beta_0 + \beta_2 x_i + \varepsilon_i, \;\; i = 1, ..., n$$

¿Qué supuestos debe cumplir ε_i para que el modelo anterior sea un modelo lineal?

(a) $\mathbb{E}(\varepsilon_i) = 0$ para todo $i = 1, ..., n$

(b) $Var(\varepsilon_i) = \sigma^2$ y para todo $i = 1, ..., n$

(c) $Cov(\varepsilon_i, \varepsilon_j) = 0$ para todo $i \neq j$

(d) $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$, para todo $i = 1, ..., n$

Solución:

$\boxed{\text{La respuesta es (d)}}$

Problema 45. Considere el siguiente modelo lineal

$$\begin{aligned} Y_1 &= \alpha_1 + \varepsilon_1 \\ Y_2 &= 2\alpha_1 - \alpha_2 + \varepsilon_2 \\ Y_3 &= \alpha_1 + 2\alpha_2 + \varepsilon_3 \end{aligned}$$

¿Cuáles son los estimadores de máxima verosimilitud de α_1 y α_2?

(a) $\hat{\alpha_1} = \frac{y_1+2y_2+y_3}{6}$ y $\hat{\alpha_2} = \frac{2y_3-y_2}{5}$

(b) $\hat{\alpha_1} = \hat{\alpha_1} = \frac{y_1+2y_2+y_3}{5}$ y $\hat{\alpha_2} = \frac{2y_3-y_2}{6}$

(c) $\hat{\alpha_1} = \frac{2y_3-y_2}{5}$ y $\hat{\alpha_2} = \frac{y_1+2y_2+y_3}{6}$

(d) $\hat{\alpha_1} = \frac{y_1+y_2+y_3}{3}$ y $\hat{\alpha_2} = \frac{2y_3-y_2}{6}$

Solución:

La matriz de diseño está dada por

$$X = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}$$

Luego,

$$(X^T X)^{-1} = \begin{pmatrix} 1/6 & 0 \\ 0 & 1/5 \end{pmatrix}$$

Así,

$$\begin{pmatrix} \hat{\alpha_1} \\ \hat{\alpha_2} \end{pmatrix} = (X^T X)^{-1} X^T Y = \begin{pmatrix} \frac{y_1+2y_2+y_3}{6} \\ \frac{2y_3-y_2}{5} \end{pmatrix}$$

$\boxed{\text{La respuesta es (a)}}$

Problema 46. Suponga que se tiene modelo lineal

$$Y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i, \quad i = 1, ..., n$$

Considere la siguiente condición

$$\sum_{i=1}^{n} x_{i2} = \sum_{i=1}^{n} x_{i3} = \sum_{i=1}^{n} x_{i2}x_{i3} = 0$$

¿Qué se puede concluir a partir de lo anterior?

(a) $\hat{\beta}_1$ es independiente de $\hat{\beta}_2$

(b) $\hat{\beta}_2$ es independiente de $\hat{\beta}_3$

(c) $\hat{\beta}_1$ es independiente de $\hat{\beta}_3$

(d) $\hat{\beta}_1$ es independiente de $\hat{\beta}_2$ y de $\hat{\beta}_3$

Solución:

Recordar que

$$\hat{\beta} \sim N_3(\beta, (X^TX)^{-1}\sigma^2)$$

Por lo tanto, para ver independencia, basta con mirar

$$X^TX = \begin{pmatrix} 1 & 1 & ... & 1 \\ x_{12} & x_{22} & ... & x_{n2} \\ x_{13} & x_{23} & ... & x_{n3} \end{pmatrix} \begin{pmatrix} 1 & x_{12} & x_{13} \\ 1 & x_{22} & x_{23} \\ ... & ... & ... \\ 1 & x_{n2} & x_{n3} \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n x_{i2} & \sum_{i=1}^n x_{i3} \\ \sum_{i=1}^n x_{i2} & \sum_{i=1}^n x_{i2}^2 & \sum_{i=1}^n x_{i2}x_{i3} \\ \sum_{i=1}^n x_{i3} & \sum_{i=1}^n x_{i2}x_{i3} & \sum_{i=1}^n x_{i3}^2 \end{pmatrix}$$

Que reemplzando resulta

$$\begin{pmatrix} n & 0 & 0 \\ 0 & \sum_{i=1}^n x_{i2}^2 & 0 \\ 0 & 0 & \sum_{i=1}^n x_{i3}^2 \end{pmatrix}$$

$\boxed{\text{La respuesta es (d)}}$

Problema 47. El diámetro de un árbol a la altura del pecho está influenciado por muchos factores. Los datos relacionan el diámetro de árboles de eucalipto (Y) con: su edad (X_1), promedio de precipitación pluvial en su localidad (X_2) y elevación de la localidad (X_3). ¿Cómo sería el modelo? ¿Cuáles son los supuestos?

(a) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$, con $\mathbb{E}(\varepsilon_i) = 0$

(b) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$, con $\varepsilon_i \sim N(0, \sigma_i^2)$

(c) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$, con $Y_i \sim N(\beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3}, \sigma^2)$

(d) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$, con $Var(\varepsilon_i) = \sigma^2$

Solución:

$\boxed{\text{La respuesta es (c)}}$

Problema 48. Considere el modelo lineal

$$Y_i \sim N(\mu_i, \sigma^2)$$

donde $\mu_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$, $i = 1, 2, 3, 4$. Los datos se encuentran en la siguiente tabla

Observación	y	x_1	x_2
1	6	1	2
2	1	-1	2
3	11	0	-3
4	3	0	-1

¿Cuál es la estimación de β_0, β_1 y β_2?

(a) $\hat{\beta}_0 = 1$, $\hat{\beta}_1 = 3{,}8$ y $\hat{\beta}_2 = -5{,}55$

(b) $\hat{\beta}_0 = 5{,}25$, $\hat{\beta}_1 = 2$ y $\hat{\beta}_2 = 1$

(c) $\hat{\beta}_0 = -5{,}25$, $\hat{\beta}_1 = -2$ y $\hat{\beta}_2 = 1$

(d) $\hat{\beta}_0 = 5{,}25$, $\hat{\beta}_1 = 2{,}5$ y $\hat{\beta}_2 = -1{,}22$

Solución:

La estimación está dada por

$$(X^T X)^{-1} X^T y = \begin{pmatrix} 5{,}25 \\ 2{,}5 \\ -1{,}22 \end{pmatrix}$$

Donde

$$X = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 2 \\ 1 & 0 & -3 \\ 1 & 0 & -1 \end{pmatrix}$$

$\boxed{\text{La respuesta es (d)}}$

Problema 49. Se generó el siguiente modelo de regresión lineal que asocia la calidad del vino mediante algunas variables explicativas. Se obtuvieron los siguientes resultados

$$Calidad = 4{,}18 + 0{,}509 Aroma + 0{,}078 Consistencia + 1{,}14 Sabor$$

Predictor	Coef	D.E	T	Valor-p
Constante	4.185	1.230	3.40	0.002
Aroma	0.5085	0.2825	1.80	0.081
Consistencia	0.0779	0.3291	0.24	0.814
Sabor	1.1371	0.3260	3.49	0.001

$$S = 1{,}24576, \quad R^2 = 65{,}9\,\%, \quad R^2_{adj} = 62{,}9\,\%.$$

Con la información entregada, ¿qué podría afirmar?

(a) La variable Sabor es la única variable estadísticamente significativa en el modelo.

(b) Utilizando un 10 % de significancia, la variable Consistencia es estadísticamente significativa en el modelo.

(c) Utilizando un 10 % de significancia, las variables Aroma, Sabor y la constante son estadísticamente significativas en el modelo.

(d) Utilizando un 1 % de significancia, sólo la constante es estadísticamente significativa.

Solución:

$\boxed{\text{La respuesta es (c)}}$

Problema 50. Suponga que se tiene un modelo de regresión lineal simple que relaciona los niveles de Proteina C Reactiva (PCR) medida en mmol/L como variable dependiente con el número de articulaciones dolorosas o hipersensibles de un paciente con artritis idiomática (variable independiente). El modelo ajustado en una muestra de 100 pacientes en el siguiente

$$Y = 0{,}47 + 0{,}08X$$

En el modelo, ¿cómo se interpreta el coeficiente 0.08?

(a) Dado que es superior a 0.05, quiere decir que no existe asociación entre las variables dependiente e independiente.

(b) Es el cambio obtenido en la medida de PCR como consecuancia de oasar a tener una articulación adicional inflamada.

(c) Es el valor medio de la PCR para los sujetos con una sola articulación afecta.

(d) Es el valor medio basal de PCR para los individuos sanos (sin articulaciones afectas).

Solución:

$\boxed{\text{La respuesta es (b)}}$