Universidad Panamericana
Maestría en Ciencia de Datos
Econometría

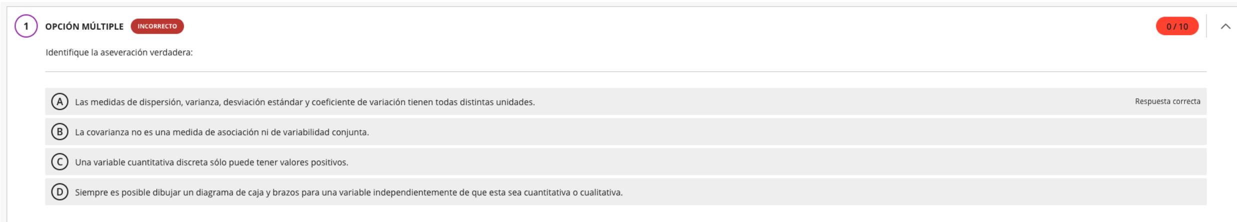
Solución de la Evaluación Diagnóstica de Estadística

Enrique Ulises Báez Gómez Tagle

2 de agosto de 2025

Índice

1. Pregunta 1


1 OPCIÓN MÚLTIPLE INCORRECTO 0 / 10

Identifique la aseveración verdadera:

(A) Las medidas de dispersión, varianza, desviación estándar y coeficiente de variación tienen todas distintas unidades. Respuesta correcta

(B) La covarianza no es una medida de asociación ni de variabilidad conjunta.

(C) Una variable cuantitativa discreta sólo puede tener valores positivos.

(D) Siempre es posible dibujar un diagrama de caja y brazos para una variable independientemente de que esta sea cuantitativa o cualitativa.

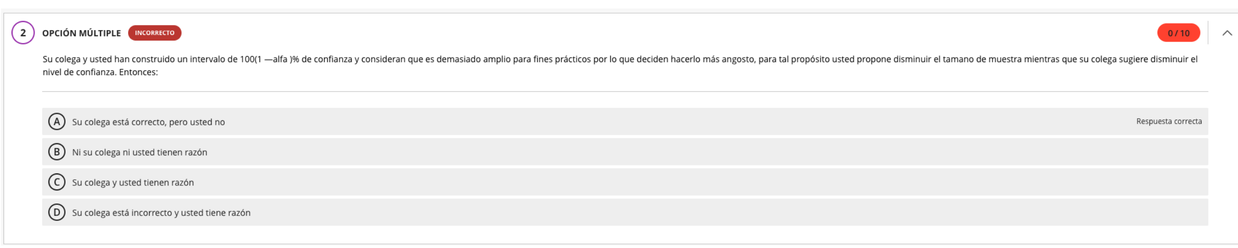
Respuesta correcta: Las medidas de dispersión, varianza, desviación estándar y coeficiente de variación tienen todas distintas unidades.
Cada una de estas medidas de dispersión tiene unidades diferentes:

- *Varianza:* mide dispersión en unidades cuadradas (cuadrado).
- *Desviación estándar:* raíz cuadrada de la varianza (base).
- *Coeficiente de variación:* desviación estándar normalizada por la media (magnitud adimensional).

Se descartan:

- *"La covarianza no es una medida de asociación ni de variabilidad conjunta."*: La covarianza mide justamente como varían 2 variables conjuntamente, es una medida de asociación lineal.
- *Üna variable cuantitativa discreta sólo puede tener valores positivos."*: Una variable discreta puede tomar tanto valores positivos como negativos.
- *"Siempre es posible dibujar un diagrama de caja y brazos para una variable independientemente de que esta sea cuantitativa o cualitativa."*: Los boxplots requieren una variable cuantitativa por la distancia/cuartiles que usan.

2. Pregunta 2


2 OPCIÓN MÚLTIPLE INCORRECTO 0 / 10

Su colega y usted han construido un intervalo de 100(1 —alfa)% de confianza y consideran que es demasiado amplio para fines prácticos por lo que deciden hacerlo más angosto, para tal propósito usted propone disminuir el tamano de muestra mientras que su colega sugiere disminuir el nivel de confianza. Entonces:

(A) Su colega está correcto, pero usted no Respuesta correcta

(B) Ni su colega ni usted tienen razón

(C) Su colega y usted tienen razón

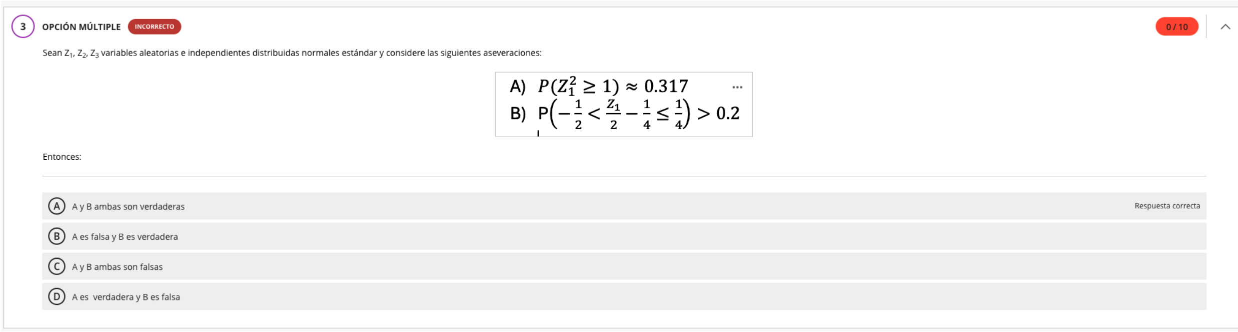
(D) Su colega está incorrecto y usted tiene razón

Respuesta correcta: Su colega está correcto, pero usted no.
Para reducir un intervalo de confianza de $100(1-\alpha)\,\%$ existen 2 opciones:

- *Disminuir el nivel de confianza* (aumentar α), lo que reduce el valor crítico $z_{1-\alpha/2}$ y por ende el margen de error.
- *Aumentar el tamaño de muestra* n, lo que disminuye el error estándar $\sigma/\sqrt{n}$.

Disminuir el tamaño de muestra *aumentaría* el error estándar y *ampliaría* el intervalo, por lo que estamos equivocados. En cambio, nuestro amigo propone **disminuir el nivel de confianza**, que *reduce* el valor crítico y *estrecha* el intervalo, por lo que él está en lo correcto.

3. Pregunta 3


3 OPCIÓN MÚLTIPLE INCORRECTO 0 / 10

Sean Z_1, Z_2, Z_3 variables aleatorias e independientes distribuidas normales estándar y considere las siguientes aseveraciones:

A) $P(Z_1^2 \geq 1) \approx 0.317$

B) $P\left(-\frac{1}{2} < \frac{Z_1}{2} - \frac{1}{4} \leq \frac{1}{4}\right) > 0.2$

Entonces:

- (A) A y B ambas son verdaderas — Respuesta correcta
- (B) A es falsa y B es verdadera
- (C) A y B ambas son falsas
- (D) A es verdadera y B es falsa

Respuesta correcta: A y B ambas son verdaderas.

- *A*:

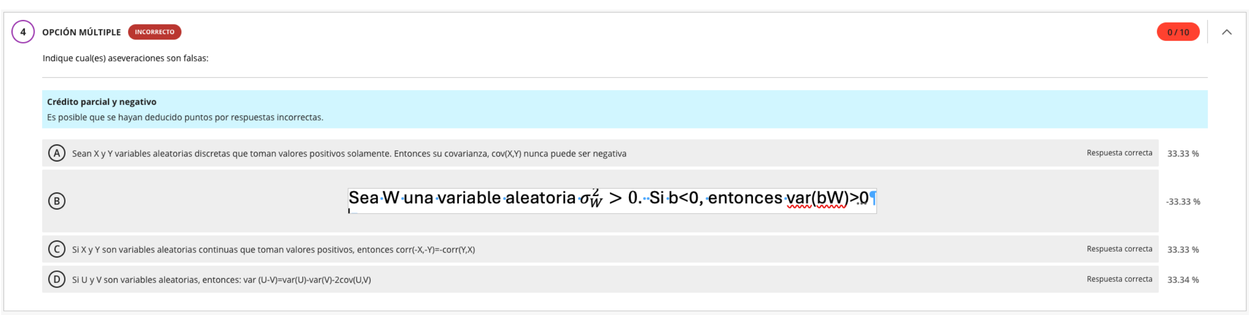
$$P(Z_1^2 \geq 1) \;=\; P(|Z_1| \geq 1) \;=\; 2\big[1 - \Phi(1)\big] \;\approx\; 2(0.1587) \;\approx\; 0.3174.$$

- *B*:

$$T = \frac{Z_1}{2} - \frac{1}{4} \;\sim\; N\big(\mu = -\tfrac{1}{4},\; \sigma^2 = \tfrac{1}{4}\big),$$

$$\begin{aligned} P\Big(-\tfrac{1}{2} \leq T \leq \tfrac{1}{4}\Big) &= P\Big(\frac{-\frac{1}{2}+\frac{1}{4}}{0.5} \leq U \leq \frac{\frac{1}{4}+\frac{1}{4}}{0.5}\Big) \\ &= P\big(-0.5 \leq U \leq 1\big) = \Phi(1) - \Phi(-0.5) \;\approx\; 0.8413 - 0.3085 \;\approx\; 0.5328 > 0.2. \end{aligned}$$

4. Pregunta 4


4 OPCIÓN MÚLTIPLE INCORRECTO 0 / 10

Indique cual(es) aseveraciones son falsas:

Crédito parcial y negativo

Es posible que se hayan deducido puntos por respuestas incorrectas.

- (A) Sean X y Y variables aleatorias discretas que toman valores positivos solamente. Entonces su covarianza, cov(X,Y) nunca puede ser negativa — Respuesta correcta — 33.33 %
- (B) Sea W una variable aleatoria $\sigma_W^2 > 0$. Si b<0, entonces var(bW)>0 — -33.33 %
- (C) Si X y Y son variables aleatorias continuas que toman valores positivos, entonces corr(-X,-Y)=-corr(Y,X) — Respuesta correcta — 33.33 %
- (D) Si U y V son variables aleatorias, entonces: var (U-V)=var(U)-var(V)-2cov(U,V) — Respuesta correcta — 33.34 %

Respuesta correcta: A, C y D son falsas.

Las aseveraciones A, C y D son falsas por:

- *"Sean X y Y variables aleatorias discretas que toman valores positivos solamente. Entonces su covarianza, cov(X,Y) nunca puede ser negativa"*: Falso, la covarianza puede ser negativa si valores altos de X se asocian con valores bajos de Y.

- *"Si X y Y son variables aleatorias continuas que toman valores positivos, entonces corr(-X,Y)=corr(Y,X)"*: Falso

$$\operatorname{corr}(-X, Y) = \frac{\operatorname{Cov}(-X, Y)}{\sigma_X\, \sigma_Y} = -\frac{\operatorname{Cov}(X, Y)}{\sigma_X\, \sigma_Y} = -\operatorname{corr}(X, Y) \neq \operatorname{corr}(Y, X)\,.$$

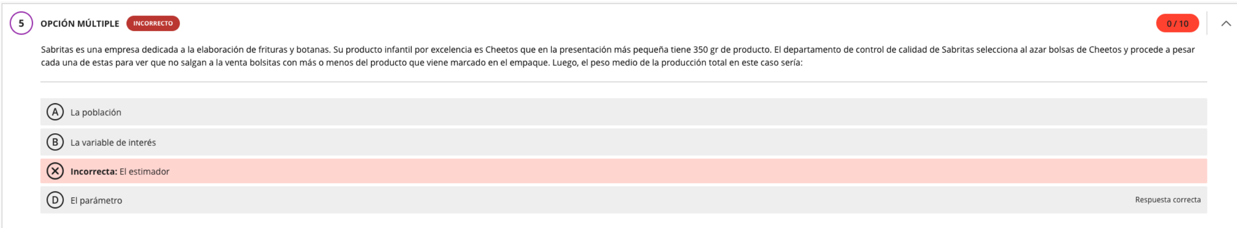
- *"Si U y V son variables aleatorias, entonces $var(U - V) = var(U) - var(V) - 2\ \operatorname{Cov}(U, V)$"*: Falso, la fórmula sería:

$$\operatorname{Var}(U - V) = \operatorname{Var}(U) + \operatorname{Var}(V) - 2\ \operatorname{Cov}(U, V)\,.$$

Queda como verdadera

- *"Sea W una variable aleatoria $\sigma_W^2 > 0$. Si $b < 0$, entonces* $\mathrm{Var}(bW) > 0$*"*:
  Verdadero, porque $\mathrm{Var}(bW) = b^2\,\mathrm{Var}(W) > 0$ siempre que $\mathrm{Var}(W) > 0$ y $b \neq 0$.

5. Pregunta 5

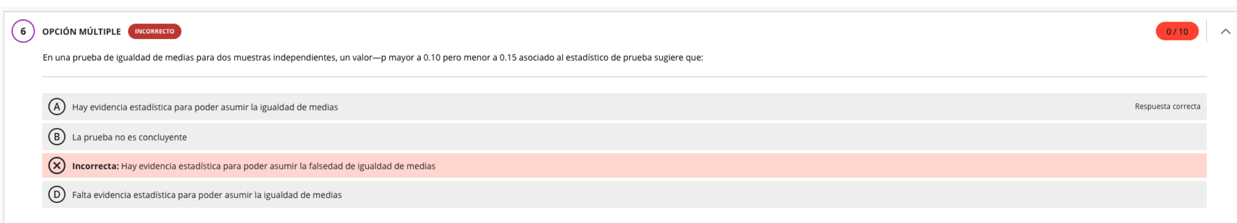


Respuesta correcta: El parámetro. El peso medio de todas las bolsas de Cheetos producidas es una característica *fija* de la población completa, es decir, un **parámetro** poblacional.

Se descartan:

- *"La población"*: conjunto de todas las bolsas, no la *medida* de su peso medio.
- *"La variable de interés"*: variable de interés, el *peso* de cada bolsa, no la *media* de esa variable.
- *El Estimador*: el estimador sería la *media muestral* calculada a partir de una muestra.

6. Pregunta 6

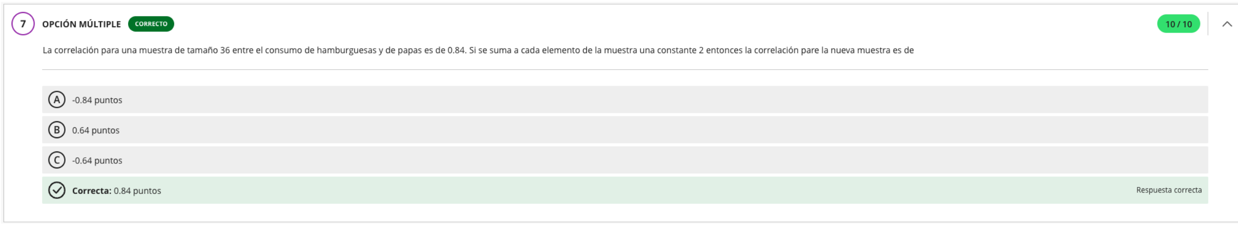


Respuesta correcta: Hay evidencia estadística para poder asumir la igualdad de medias. Con un nivel de significancia $\alpha = 0.10$, si el valor-p asociado al estadístico de prueba es mayor que 0.10 (aunque menor que 0.15), no rechazamos H_0. Al no encontrar evidencia contra la hipótesis nula, podemos *asumir* que las medias son iguales en ese nivel de confianza.

Se descartan:

- *"La prueba no es concluyente"*: Aunque el p-valor esté en la "zona gris" entre 0.10 y 0.15, sigue siendo mayor que el umbral de 0.10, por lo que no se rechaza H_0.
- *"Hay evidencia estadística para poder asumir la falsedad de igualdad de medias"*: Sería falso, pues para rechazar H_0 necesitaríamos $p \leq 0.10$.
- *"Falta evidencia estadística para poder asumir la igualdad de medias"*: Equivocado, porque al no rechazar H_0 ya tenemos evidencia suficiente para asumir la igualdad en ese nivel de confianza.

7. Pregunta 7



7 OPCIÓN MÚLTIPLE CORRECTO 10 / 10

La correlación para una muestra de tamaño 36 entre el consumo de hamburguesas y de papas es de 0.84. Si se suma a cada elemento de la muestra una constante 2 entonces la correlación pare la nueva muestra es de

(A) -0.84 puntos

(B) 0.64 puntos

(C) -0.64 puntos

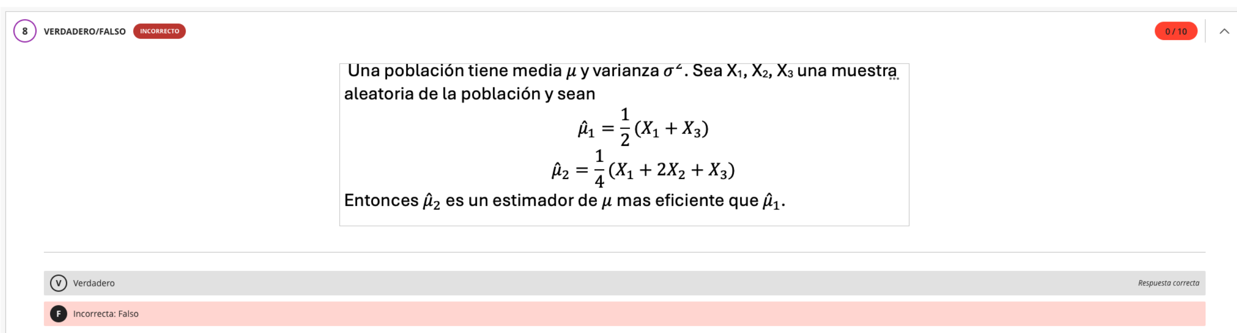
Correcta: 0.84 puntos — Respuesta correcta

Respuesta correcta: 0.84 puntos.
La correlación muestral es invariante ante traslaciones de las variables y ante escalados positivos. Cuando sumamos 2 unidades a cada observación, la covarianza y las desviaciones estándar cambian, pero su cociente (la correlación) se mantiene igual.

Se descartan:

- *0.84 puntos"*: sumar constantes no invierte el signo de la correlación, solo un factor negativo lo afecta.
- *"0.64 puntos"*: no hay cambio de magnitud, ya que no se multiplica por un factor distinto de 1.
- *0.64 puntos"*: no aplica ni inversión de signo, ni cambio de magnitud.

8. Pregunta 8



8 VERDADERO/FALSO INCORRECTO 0 / 10

Una población tiene media μ y varianza σ^2. Sea X_1, X_2, X_3 una muestra aleatoria de la población y sean

$$\hat{\mu}_1 = \frac{1}{2}(X_1 + X_3)$$

$$\hat{\mu}_2 = \frac{1}{4}(X_1 + 2X_2 + X_3)$$

Entonces $\hat{\mu}_2$ es un estimador de μ mas eficiente que $\hat{\mu}_1$.

(V) Verdadero — *Respuesta correcta*

(F) Incorrecta: Falso

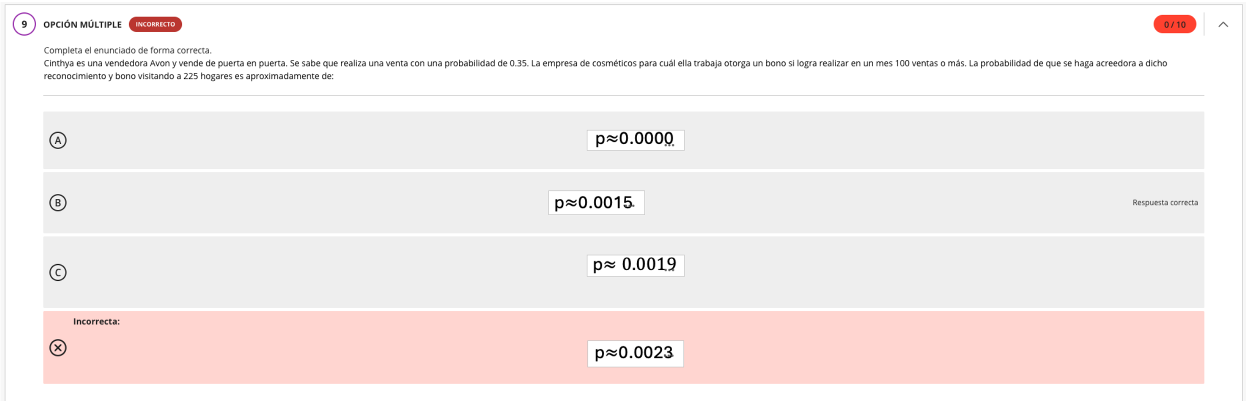
Respuesta correcta: Verdadero.
Ambos $\hat{\mu}_1$ y $\hat{\mu}_2$ son *no sesgados* para μ. La eficiencia se compara a través de sus varianzas:

$$\mathrm{Var}(\hat{\mu}_1) = \mathrm{Var}\Big(\tfrac{1}{2}(X_1 + X_3)\Big) = \tfrac{1}{4}\big[\mathrm{Var}(X_1) + \mathrm{Var}(X_3)\big] = \tfrac{1}{4}(2\sigma^2) = \tfrac{1}{2}\,\sigma^2.$$

$$\mathrm{Var}(\hat{\mu}_2) = \mathrm{Var}\Big(\tfrac{1}{4}(X_1 + 2X_2 + X_3)\Big) = \tfrac{1}{16}\big[\mathrm{Var}(X_1) + 4\,\mathrm{Var}(X_2) + \mathrm{Var}(X_3)\big] = \tfrac{1}{16}(6\sigma^2) = \tfrac{3}{8}\,\sigma^2.$$

Como $\frac{3}{8}\sigma^2 < \frac{1}{2}\sigma^2$, $\hat{\mu}_2$ tiene menor varianza y por tanto es más eficiente que $\hat{\mu}_1$.

9. Pregunta 9


9 OPCIÓN MÚLTIPLE INCORRECTO 0 / 10

Completa el enunciado de forma correcta.
Cinthya es una vendedora Avon y vende de puerta en puerta. Se sabe que realiza una venta con una probabilidad de 0.35. La empresa de cosméticos para cuál ella trabaja otorga un bono si logra realizar en un mes 100 ventas o más. La probabilidad de que se haga acreedora a dicho reconocimiento y bono visitando a 225 hogares es aproximadamente de:

(A) p≈0.0000

(B) p≈0.0015 Respuesta correcta

(C) p≈ 0.0019

Incorrecta: (D) p≈0.0023

Respuesta correcta: $p \approx 0.0015$**.**
Sea $X \sim \text{Bin}(n = 225,\ p = 0.35)$. Buscamos

$$P(X \geq 100).$$

Usando la aproximación normal sin corrección de continuidad:

$$\mu = np = 225 \times 0.35 = 78.75, \qquad \sigma = \sqrt{np(1-p)} = \sqrt{78.75 \times 0.65} \approx 7.15.$$

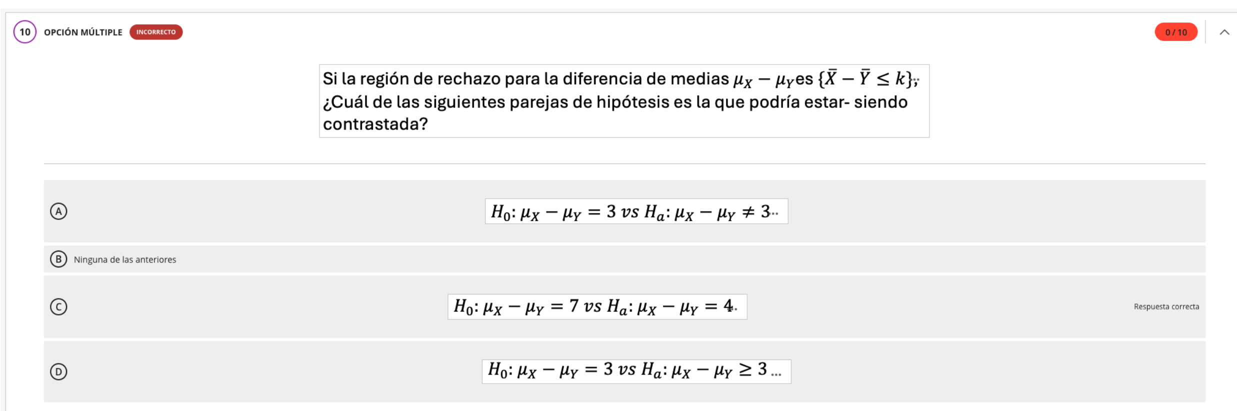
Entonces

$$Z = \frac{100 - 78.75}{7.15} \approx 2.97,$$

y

$$P(X \geq 100) \approx 1 - \Phi(2.97) \approx 0.0015.$$

10. Pregunta 10


10 OPCIÓN MÚLTIPLE INCORRECTO 0 / 10

Si la región de rechazo para la diferencia de medias $\mu_X - \mu_Y$ es $\{\bar{X} - \bar{Y} \leq k\}$, ¿Cuál de las siguientes parejas de hipótesis es la que podría estar- siendo contrastada?

(A) H_0: $\mu_X - \mu_Y = 3$ vs H_a: $\mu_X - \mu_Y \neq 3$..

(B) Ninguna de las anteriores

(C) H_0: $\mu_X - \mu_Y = 7$ vs H_a: $\mu_X - \mu_Y = 4$. Respuesta correcta

(D) H_0: $\mu_X - \mu_Y = 3$ vs H_a: $\mu_X - \mu_Y \geq 3$ ...

Respuesta correcta: $H_0 : \mu_X - \mu_Y = 7$ **vs.** $H_a : \mu_X - \mu_Y < 4$**.**
La región de rechazo $\{\bar{X} - \bar{Y} \leq k\}$ resulta de un test de una cola a la izquierda, donde se rechaza la igualdad de medias si la diferencia muestral es suficientemente pequeña.

Se descartan:

- "$H_0 : \mu_X - \mu_Y = 3$ *vs.* $H_a : \mu_X - \mu_Y \neq 3$": es un test bilateral, no cuadra con una sola cola.
- "$H_0 : \mu_X - \mu_Y = 3$ *vs.* $H_a : \mu_X - \mu_Y \geq 3$": ese sería un test de cola derecha (rechaza por valores grandes), al contrario del enunciado.