ESTADÍSTICA – EXAMEN PEI2 – CURSO 2024-25 (10 puntos)

APELLIDOS: _____ NOMBRE: _____

DNI: _____ CÓDIGO NUMÉRICO DE LAS HOJAS DE AYUDA: _____

INSTRUCCIONES:

Al finalizar el examen hay que devolver las hojas de ayuda, en las que no se puede escribir nada.

Hay que entregar esta hoja del examen y todas las hojas para los cálculos, en las que hay que poner el nombre. Al terminar hay que entregar todas las hojas, aunque alguna esté en blanco. Si hay alguna hoja que no se quiere que sea leída por el profesor, debe tacharse su contenido.

Sólo se puede utilizar una calculadora; no se puede usar ningún dispositivo móvil.

ENUNCIADO:

En una empresa de desarrollo de software trabajan dos programadores llamados Ana y Juan. Se selecciona aleatoriamente una muestra de 4 programas de los muchos desarrollados por Ana desde que trabaja en la empresa, y se ha detectado el siguiente número de errores en cada uno de ellos: 3, 4, 8 y 9 errores. Se selecciona una muestra de 4 programas de los muchos desarrollados por Juan, y se ha detectado que tienen los siguientes errores: 1, 2, 8 y 9 errores. Juan ha tardado las siguientes horas en desarrollar esos 4 programas: 10, 20, 40 y 50 horas respectivamente. NOTA: Se asume una distribución Normal de todos los datos.

1) (2p) Con un nivel de confianza del 90%, ¿entre qué valores se puede afirmar que se encuentra la media de errores por programa de Ana desde que trabaja en la empresa?

Con un nivel de confianza del 90%, ¿se puede afirmar que Ana es mejor programadora que Juan, es decir comete menos errores al programar?
NOTA: Suponer varianzas poblacionales desconocidas pero iguales.

2) (1p) Resolverlo calculando el intervalo de confianza de la diferencia de medias poblacionales de errores por programa.
3) (1.5p) Resolverlo con un contraste de hipótesis paramétrico sobre la media poblacional de errores por programa, aplicando el método de la región de aceptación.
4) (1.5p) Resolverlo con el mismo contraste del apartado anterior, aplicando el método del p-valor.
5) (2p) Resolverlo con un contraste de hipótesis no paramétrico sobre la mediana poblacional de errores por programa.

6) (2p) Si Juan acaba de terminar un nuevo programa al que ha dedicado 30 horas; con un nivel de confianza del 90%, ¿cuál será el mínimo y el máximo número de errores que se prevé que tendrá ese nuevo programa? NOTA: Redondear por defecto el valor mínimo (ej. 100’8 sería 100) y por exceso el valor máximo (ej. 200’3 sería 201).

RESPUESTAS: No se aceptan respuestas que no estén claramente avaladas por los cálculos en las hojas adjuntas. **NOTA: Redondear los resultados a 2 cifras decimales.**

RESPUESTAS

1) Valor mínimo = _____ errores por programa

Valor máximo = _____ errores por programa

El resultado obtenido significa que, si calculásemos 100 intervalos de confianza para la media poblacional de errores cometidos por Ana desde que trabaja en la empresa, haciendo 100 muestreos diferentes, entonces _____

_____

Dibujar el diagrama de densidad de probabilidad del estimador, indicando el centro y los límites del intervalo de confianza calculado.



2) Intervalo de la diferencia de medias = [_____, _____]

La respuesta a ¿Ana es mejor programadora que Juan? es (SI/NO) _____, porque _____

_____

_____

3) Contraste de hipótesis formulado: H_0: _____ vs H_A: _____

La respuesta a ¿Ana es mejor programadora que Juan? es (SI/NO) _____ porque _____ está _____ de la región de aceptación, cuyos límites son (_____, _____)

4) La respuesta a ¿Ana es mejor programadora que Juan? es (SI/NO) _____ porque el p-valor = _____

es _____ que _____

5) Contraste de hipótesis formulado: H_0: _____ vs H_A: _____

La respuesta a ¿Ana es mejor programadora que Juan? es (SI/NO) _____ porque el p-valor = _____

es _____ que _____

6) Coeficiente de regresión = _____

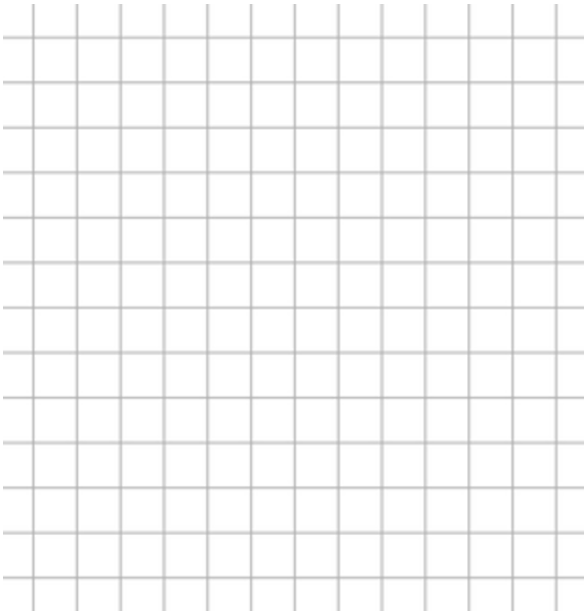
Fórmula recta de regresión: _____

Valor mínimo = _____ errores (núm. entero)

Valor máximo = _____ errores (núm. entero)

(Se ha redondeado por defecto el valor mínimo y por exceso el valor máximo)

Dibujo de la recta de regresión y del intervalo, sobre un diagrama de dispersión con las parejas de valores de las variables.

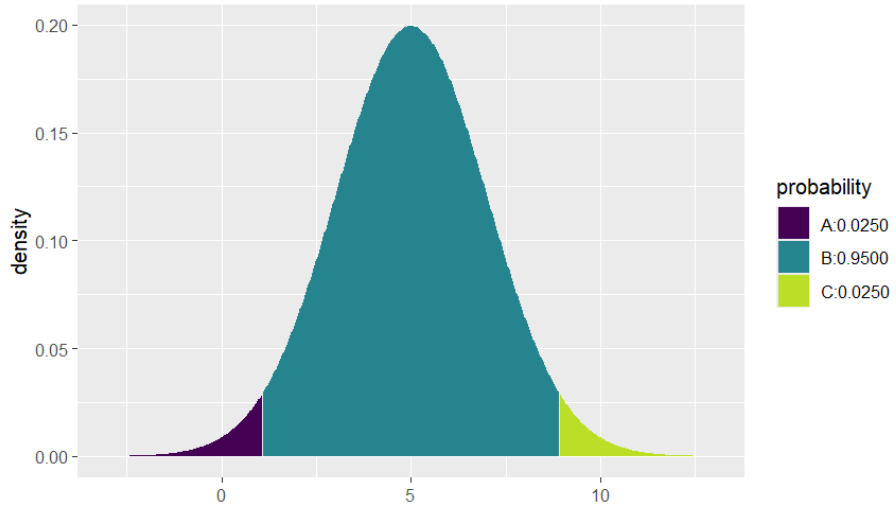


OTRAS POSIBLES PREGUNTAS NO INCLUIDAS EN EL EXAMEN

7) Si en lugar de una muestra pequeña de 4 programas desarrollados por Ana, se selecciona una muestra de grande de 30 programas, y la media muestral en ese caso es de 6 errores por programa, y la desviación estándar muestral es de 2,94 errores por programa, calcular, un intervalo de confianza para la media de errores por programa de Ana desde que trabaja en la empresa, con un nivel de confianza del 90%. Dibujar el diagrama de densidad de probabilidad del estimador, indicando el centro y los límites del intervalo de confianza calculado.

8) En la muestra grande de 30 programas desarrollados por Ana, se ha encontrado que hay 7 programas con menos de 4 errores. Con un nivel de confianza del 90%, ¿entre qué valores se puede afirmar que se encuentra la proporción de programas desarrollados por Ana desde que trabaja en la empresa que tengan menos de 4 errores? Dibujar el diagrama de densidad de probabilidad del estimador utilizado, indicando los límites del intervalo de confianza calculado.

9) El siguiente diagrama representa la función de densidad de probabilidad de un estimador para una media poblacional. Indicar los límites del intervalo de confianza para esa media poblacional mostrado en la imagen y el nivel de confianza para el intervalo.
NOTA: El tamaño de la muestra es 100, la media muestral es 5, y la varianza muestral es 4.



10) Utilizando los datos de las muestras de 4 programas de Ana y 4 de Juan, con un nivel de confianza del 90%, ¿se puede afirmar que Ana es mejor programadora que Juan, es decir comete menos errores al programar? Resolverlo calculando el intervalo de confianza de la diferencia de medias poblacionales de errores por programa. Suponer varianzas poblacionales desconocidas y diferentes.
NOTA: Es igual que la pregunta 2), pero suponiendo varianzas diferentes.

11) En lugar de dos muestras pequeñas de 4 programas de Ana y Juan, se seleccionan 30 programas de Ana y 40 de Juan. En el caso de Ana, la media muestral es de 6 errores por programa, y la desviación estándar muestral es de 2,94 errores por programa. En el caso de Juan, la media muestral es de 5 errores por programa, y la desviación estándar muestral es de 4,08 errores por programa. Con un nivel de confianza del 90%, ¿se puede afirmar que Ana es mejor programadora que Juan, es decir comete menos errores al programar? Resolverlo calculando el intervalo de confianza de la diferencia de medias poblacionales de errores por programa. NOTA: Son muestras grandes.

12) Con un nivel de confianza del 90%, ¿se puede afirmar que la media de errores por programa de Ana desde que trabaja en la empresa es superior a 4 errores por programa? Resolverlo por el método de la Región de Aceptación y por el método del p-valor.

13) Con un nivel de confianza del 90%, ¿se puede afirmar que la media de errores por programa de Ana desde que trabaja en la empresa es inferior a 9 errores por programa? Resolverlo por el método de la Región de Aceptación y por el método del p-valor.

14) Con un nivel de confianza del 90%, ¿se puede afirmar que la media de errores por programa de Ana desde que trabaja en la empresa es diferente a 5 errores por programa? Resolverlo por el método de la Región de Aceptación y por el método del p-valor.

15) En una muestra de 30 programas desarrollados por Ana, se ha encontrado que hay 7 programas con menos de 4 errores. En una muestra de 40 programas desarrollados por Juan, se ha encontrado que hay 7 programas con menos de 4 errores. Con un nivel de confianza del 90%, ¿se puede afirmar que la proporción de programas desarrollados por Ana desde que trabaja en la empresa que tengan menos de 4 errores es mayor que la de Juan? Resolverlo por el método de la Región de Aceptación y por el método del p-valor.