

TAREA 5

Normas para la presentación de la Tarea

- La tarea puede realizarse **en parejas** de cualquier sección.
- La presente tarea puede realizarla a computador o a mano. Debe tener en cuenta que la presentación del informe puede influir en la calificación final.
- El informe debe ser presentado en hojas blancas, numeradas, impresión por ambos lados y en la parte superior de cada hoja se debe especificar el nombre y código de cada estudiante.
- **La primera hoja de su tarea debe contener el formato que se presenta en la siguiente página.** Las tareas que no tengan este formato tendrán una penalización de cero punto cinco (0.5) sobre la nota final de la tarea.
- Será responsabilidad de los integrantes del grupo verificar el contenido de la tarea antes de la entrega. Luego de entregado el documento, no se recibirán adiciones por motivos de problemas de impresión en fórmulas o ecuaciones.
- Debe respetar el horario y el lugar de entrega de las tareas. **Las tareas entregadas después de este plazo no serán recibidas y su calificación será de cero (0).**
- **Por ningún motivo la tarea será recibida por correo electrónico.**
- El incumplimiento de alguna de las anteriores instrucciones tendrá un impacto negativo en la nota de la tarea.
- Cualquier sospecha de fraude será tratada de acuerdo con el reglamento de la Universidad.
- Si usted encuentra algún GAZAPO¹ en la solución correspondiente a esta tarea por favor comuníquelo a **ma.galvis138**. Si su observación es válida, se verá recompensado con un incremento del 5% en la nota de la tarea.

Forma de entrega

- El informe de la tarea debe ser entregado en los casilleros de Ingeniería Industrial, en el séptimo piso del ML, antes de la fecha límite de entrega. El casillero será habilitado el día anterior a la entrega de la tarea.
- Adicionalmente, el informe de la tarea junto a sus archivos de soporte deberán ser colgados en el link habilitado en Sicua Plus, antes de la fecha límite de entrega. Por lo tanto, si usted realizó su tarea a mano, debe escanear el documento y subirlo al link correspondiente.

Fecha de entrega

La fecha límite de entrega es el **jueves 19 de noviembre de 2015, antes de las 5:00 p.m.**

¹ Yerro que por inadvertencia deja escapar quien escribe o habla. (Definición según La Real Academia de la Lengua Española)

Integrante 1: _____ Código: _____ Sección: _____
 Integrante 2: _____ Código: _____ Sección: _____

Numeral		Puntaje
1	a)	/2
	b)	/12
	c)	/4
	d)	/2
	e)	/5
2	a)	/2
	b)	/4
	c)	/4
	d)	/2
	e)	/2
	f)	/1
	g)	/16
	h)	/8
	i)	/4
3	a)	/2
	b)	/10
	c)	/4
	d)	/5
	e)	/4
	f)	/2
	g)	/8

TOTAL	/105
NOTA	/5

Punto 1. (25 puntos)

Diferentes organizaciones, entre las que se incluye *The American Journal of Psychiatry*, han presentado como una problemática creciente la sustitución de las interacciones sociales por las interacciones a través de medios digitales (redes sociales), lo que de acuerdo con los expertos está correlacionado con problemas psicológicos (estrés, ansiedad, adicción). Con el propósito de realizar un estudio formal en cierta universidad, se propuso como primera etapa determinar si el tiempo, en minutos, que un estudiante dedica a visitar las redes sociales se puede modelar como una variable aleatoria con una distribución de probabilidad conocida, particularmente, si se asemeja a una distribución normal. Para el desarrollo de esta primera etapa, se hizo seguimiento al tiempo que una muestra aleatoria de 150 estudiantes dedica en las redes sociales diariamente. Los resultados de este seguimiento se presentan en el archivo de excel *DatosRedesSociales.xlsx*.

- a. (2 puntos) Plantee la prueba de hipótesis nula y alterna de la prueba.
- b. (12 puntos) Identifique el estadístico de prueba con su distribución y calcúlelo. Utilice 10 clases equiprobables.
- c. (4 puntos) Utilizando un nivel de significancia del 5%, realice una gráfica de la distribución del estadístico de prueba, identifique el punto crítico y la región de rechazo. Adicionalmente, ubique el valor del estadístico de prueba.
- d. (2 puntos) Concluya en términos del problema.
- e. (5 puntos) Verifique los resultados obtenidos utilizando Crystal Ball, reporte las salidas que arroja Crystal Ball.

Punto 2. (45 puntos)

Diana es una comerciante exitosa que se caracteriza por utilizar métodos estadísticos como soporte a la toma de decisiones. En esta ocasión, ella realizó un estudio con el propósito de predecir el valor total de las ventas mensuales (variable aleatoria Y), en millones de pesos, utilizando los gastos en publicidad mensual, en millones de pesos, como la variable dependiente X. Para ello, recolectó información de 12 sucursales de empresas similares. A continuación se presenta el resumen de los datos obtenidos por el comerciante:

$$n = 12 \quad \bar{X} = 34.17 \quad \bar{Y} = 453.75$$

$$\sum_{i=1}^n y_i^2 = 2,512,925 \quad \sum_{i=1}^n x_i^2 = 15,650 \quad \sum_{i=1}^n x_i y_i = 191,325$$

$$SCE = 6338.438$$

- a. (2 puntos) Escriba el modelo de regresión lineal que le permite predecir el valor total de las ventas semanales en función de la publicidad semanal.
- b. (4 puntos) Utilizando el método de mínimos cuadrados ordinarios, encuentre de manera teórica la ecuación para estimar los coeficientes del modelo de regresión lineal planteado en el literal a.

- c. **(4 puntos)** Estime los parámetros del modelo de regresión lineal simple planteado en el literal a. Realice una interpretación de cada uno de los coeficientes estimados.
- d. **(2 puntos)** Estime la tabla ANOVA correspondiente al modelo de regresión lineal simple planteado en el literal a.
- e. **(2 puntos)** Plantee las hipótesis nula y alterna de la prueba de significancia global. Defina el estadístico de prueba y concluya en términos del contexto del problema; utilice el criterio del p-valor
- f. **(1 puntos)** Estime el valor del coeficiente de determinación, e interprételo.
- g. **(16 puntos)** Plantee una prueba de significancia individual para el intercepto y para la pendiente. Para cada una, plantee las hipótesis nula y alterna y defina el estadístico de prueba y concluya en términos del contexto del problema; utilice el criterio del p-valor.
- h. **(8 puntos)** Calcule un intervalo de confianza ($\alpha = 0.05$) para el intercepto y para el coeficiente asociado a la variable independiente. Interprete cada uno de los intervalos.
- i. **(4 puntos)** Diana decidió invertir 40 millones de pesos en publicidad, y las ventas mensuales fueron de 480 millones de pesos. ¿Cuál es el valor que Diana había pronosticado para las ventas? ¿Cuál es el residuo asociado a este pronóstico?

Punto 3. (35 puntos)

Descargue de Siciuplus el archivo de Excel con el nombre “*DatosRegresion.xls*”. Allí encontrará información relacionada con la contaminación atmosférica en 38 ciudades de Estados Unidos. Las variables que se presentan en el archivo son las siguientes:

- Contenido de SO_2 (Dióxido de Azufre) en el aire. Esta variable se mide en microorganismos por metro cúbico.
- Número de fábricas con más de 20 empleados.
- Número de habitantes (miles).
- Velocidad media del viento al año (millas por hora).
- Precipitación media anual (litros por pulgada).
- Número medio de días con lluvia al año.

Estime el modelo de regresión lineal múltiple para la variable *Contenido de SO_2 en el aire* usando todas las variables descritas. Para esto de respuesta a los siguientes literales:

- a. **(2 puntos)** Escriba a continuación el modelo de regresión lineal que le permite predecir el *Contenido de SO_2 en el aire* con base en las variables independientes.
- b. **(10 puntos)** Realice una gráfica de dispersión entre cada una de las variables independientes y la dependiente. Indique, para cada caso, si de manera visual se puede percibir una relación lineal.
- c. **(4 puntos)** Con ayuda de SPSS, presente a continuación la tabla de ANOVA, y la tabla de coeficientes.

- d. **(5 puntos)** Plantee las hipótesis nula y alterna de la prueba de significancia global, en términos de los betas. Defina el estadístico de prueba y concluya en términos del contexto del problema; utilice el criterio del p-valor.
- e. **(4 puntos)** Para aquellos coeficientes que resultan significativos, presente la interpretación correspondiente a cada uno.
- f. **(2 puntos)** Estime el valor del coeficiente de determinación, e interprételo.
- g. **(8 puntos)** Verifique los supuestos del modelo de regresión lineal múltiple. Para ello, utilizando SPSS, realice una prueba de bondad de ajuste sobre los errores estimados del modelo de regresión lineal múltiple.