

TAREA 4

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 **viernes 6 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)	/4
	b)	/4
	c)	/3
	d)	/2
	e)	/2
	f)	/3
	g)	/4
2	a)	/3
	b)	/3
	c)	/3
	d)	/3
3	a)	/6
	b)	/4
4	a)	/4
	b)	/4
	c)	/4
	d)	/4
5	a)	/4
	b)	/4
	c)	/4
	d)	/4
6	a)	/4
	b)	/4
	c)	/2
7	a)	/4
	b)	/4
	c)	/4
	d)	/4
	e)	/4

Numeral		Puntaje
8	a)	/4
	b)	/4
	c)	/4
9	a)	/2
	b)	/7
	c)	/2
	d)	/6

TOTAL	/135
--------------	------

NOTA	/5
-------------	----

Punto 1. (22 puntos)

Un grupo de inversionistas extranjeros está evaluando la posibilidad de invertir parte de su capital en activos que se transan en la Bolsa de Valores de Colombia (BVC). Con el fin de conocer un poco más sobre las diferentes opciones de inversión, se obtuvo información de las acciones de 35 empresas que transan actualmente en bolsa, la cual se encuentra disponible en el archivo de Excel “**InformaciónEmpresas.xls**”.

A continuación se muestra una breve descripción de las variables que fueron recolectadas para cada una de las empresas:

- **Precio:** corresponde al precio histórico promedio de la acción, en pesos colombianos, de cada compañía registrado entre los años 2010 y 2014.
- **Sector:** corresponde al sector empresarial en el que se desempeña la compañía.
- **Dividendo:** corresponde al valor total en pesos colombianos del dividendo que un accionista espera recibir durante el año 2016.
- **Merco:** corresponde a la posición que obtuvo la compañía en la encuesta *Merco Empresas 2014*, donde se listan las mejores 100 empresas para trabajar en Colombia (aquellas empresas que tienen un “-”, son empresas que no están en dicho listado de 100 empresas).

De acuerdo con la información presentada en el archivo, de solución a los siguientes literales:

- (4 puntos)** Construya un histograma de frecuencias de las 35 empresas de acuerdo con el sector al que pertenece cada una. Muestre la tabla de frecuencias y la gráfica del histograma. ¿Cuáles son los 3 sectores de mayor frecuencia? ¿Qué porcentaje de la muestra es representado por las empresas que pertenecen a estos sectores? ¿Cuáles son los sectores de menor frecuencia? ¿Qué porcentaje de la muestra es representado por las empresas que pertenecen a estos sectores?
- (4 puntos)** Construya un diagrama de caja que represente la distribución del precio promedio de la acción de las 35 empresas de la muestra. Comente sobre todos los elementos que representa el diagrama de caja.
- (3 puntos)** Se desea analizar una nueva variable correspondiente al cociente entre el precio de la acción y los dividendos que promete (P/D). De acuerdo con este indicador se clasifica cada empresa de la siguiente manera:

Rango	Categoría
$P/D < 35$	Bajo
$35 \leq P/D < 100$	Medio
$100 \leq P/D$	Alto

Genere la nueva variable y construya un histograma de frecuencias. Muestre la tabla de frecuencias y el histograma.

- (2 puntos)** Los inversionistas están interesados en conocer cuáles son los tres sectores que prometen en promedio mayores dividendos. Para ello construya y presente una tabla dinámica que le permita comparar cada sector en términos de sus dividendos promedio.
- (2 puntos)** Los inversionistas han decidido establecer la siguiente clasificación partir de la ubicación de las empresas en el ranking Merco:

Rango	Categoría
$M < 35$	A+

$35 \leq M < 100$	A
No Ranking.	B

Construya una tabla dinámica que le permita comprobar si existen compañías con un a clasificación A+ o A, que se cataloguen en una relación precio-dividendo de rango Bajo.

- f. **(3 puntos)** Obtenga las estadísticas descriptivas de la variable precio/dividendo. ¿Qué puede concluir sobre la simetría y la curtosis de la distribución de los datos para la variable precio/dividendo? ¿Cuál es la diferencia conceptual entre el error típico y la desviación estándar?
- g. **(4 puntos)** Presente la tabla de percentiles del precio promedio de las acciones.
- ¿A qué compañía pertenece el percentil 50?
 - ¿Cuál es el precio de la acción y los dividendos asociados con esta compañía?
 - ¿Cómo interpreta este resultado?

Punto 2. (12 puntos)

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n , en la que $n > 2$, de una población con media μ y varianza σ^2 . Se cuenta con tres estimadores para la media poblacional de los cuales se desea evaluar su conveniencia. A continuación se presentan los estimadores propuestos:

$$\hat{\mu}_1 = \bar{X}$$

$$\hat{\mu}_2 = \frac{X_1 + 2X_2 + 3X_3}{6}$$

$$\hat{\mu}_3 = \frac{X_1 + X_6 + X_{10}}{4}$$

- (3 puntos)** Calcule el sesgo para cada uno de los estimadores planteados. ¿Cuáles de estos estimadores son insesgados?
- (3 puntos)** Calcule el error cuadrático medio de los tres estimadores.
- (3 puntos)** Para cada uno de los estimadores, evalúe si es consistente o si no lo es.
- (3 puntos)** De acuerdo a las conclusiones obtenidas en los literales anteriores, ¿cuál de los tres estimadores sería el más adecuado? Justifique su respuesta.

Punto 3. (10 puntos)

La variable aleatoria X se distribuye normal con parámetros μ y σ^2 , y con la siguiente función de densidad de probabilidad:

$$f(x; \theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$$

- a. **(6 puntos)** Encuentre el estimador de máxima verosimilitud para dichos parámetros. Para esto siga los siguientes pasos:
- Plantee la función de verosimilitud:
 - Halle el logaritmo natural de la función de verosimilitud.
 - Derive el resultado anterior respecto al parámetro correspondiente.
 - Igualé la derivada a cero para encontrar el estimador de máxima verosimilitud para cada parámetro.
- b. **(4 puntos)** Determine si el estimador de máxima verosimilitud para el parámetro μ es eficiente.

Punto 4. (16 puntos)

ColSteel Ltda. es una compañía dedicada al procesamiento de piezas metálicas de distintas herramientas diseñadas para la perforación en la industria petrolera. Con el fin de evaluar la posible renovación de uno de los dos tornos verticales, utilizados para el procesamiento de piezas de mayor tamaño, se tomó una muestra aleatoria del tiempo de procesamiento en cada. A continuación se resumen los resultados de la muestra:

	Tiempo en procesar una pieza (horas)						
Torno A	0.936	1.116	2.353	3.228	0.359	1.075	0.244
	2.601	0.35	3.542	1.548	0.657	0.801	0.167
	1.203	1.716	0.358	0.117	2.62	3.187	1.502
Torno B	2.242	1.76	2.515	1.989	0.337	1.482	1.479
	1.467	1.244	1.521	1.469	2.851	1.828	3.681
	0.677	0.746	1.24	1.896	0.794	2.734	1.378

El tiempo en procesar una pieza cilíndrica en el Torno A es una variable aleatoria X la cual sigue una distribución normal $N(\mu_X, \sigma_X^2)$, el tiempo en procesar una pieza en el Torno B es una variable aleatoria Y la cual sigue una distribución normal $N(\mu_Y, \sigma_Y^2)$. Teniendo en cuenta esta información responda los siguientes literales:

- (4 puntos)** Calcule un intervalo de confianza del 95% para la media del tiempo de procesamiento de cada uno de los tornos. Interprete cada uno de los intervalos de confianza. Interprete este resultado.
- (4 puntos)** Calcule un intervalo de confianza del 90% para la varianza del tiempo de procesamiento de cada uno de los tornos. Interprete este resultado.
- (4 puntos)** Construya un intervalo de confianza del 90% que le permita evaluar si las varianzas del tiempo de procesamiento de cada uno de los tornos son estadísticamente diferentes. ¿Qué puede concluir?
- (4 puntos)** Teniendo en cuenta la conclusión obtenida en el literal anterior, construya un intervalo de confianza del 95% para la diferencia de las medias de procesamiento de ambos tornos $(\mu_X - \mu_Y)$. ¿Qué puede concluir?

Punto 5. (16 puntos)

El escándalo del software ilegal de Volkswagen, instalado en varios vehículos de motor diésel en EE.UU utilizado para alterar los resultados de emisiones contaminantes durante los controles técnicos, ha prendido las alarmas de las agencias ambientales en el mundo. En Colombia, El Ministerio de Ambiente ha encomendado a un grupo de investigación de la Universidad de los Andes para estudiar las emisiones de gases de vehículos diésel de la marca en el país, tanto en condiciones de revisión técnico mecánica, como en condiciones normales de uso. Luego de dos semanas de muestreo se ha determinado que, durante revisión la revisión técnico mecánica, 23 de 75 vehículos presentaron emisiones de óxido de nitrógeno por encima de la reglamentación vigente. Por otra parte, en condiciones normales de uso, 17 de 83 vehículos no cumplieron con las normas establecidas de emisiones de gases.

Con base en esta información, responda los siguientes literales:

- (4 puntos)** Calcule un intervalo de confianza del 98% para la proporción de automóviles diésel que no cumplen con los estándares de emisiones durante revisión técnico-mecánica.
- (4 puntos)** Calcule un intervalo de confianza del 98% para la proporción de automóviles diésel que cumplen con los estándares de emisiones en condiciones normales de uso.
- (4 puntos)** Interprete los dos intervalos de confianza construidos en los puntos anteriores e indique si estos son comparables. Justifique su respuesta.
- (4 puntos)** Calcule un intervalo de confianza del 95% que le permita establecer si existe diferencia entre la proporción de vehículos que cumplen con las emisiones mínimas de gases en condiciones normales de uso y la proporción de vehículos que cumplen durante revisión técnico-mecánica. ¿Qué puede concluir?

Punto 6. (10 puntos)

La compañía Cannondale desea evaluar un nuevo sistema de frenado de disco hidráulico para incorporar a su catálogo de bicicletas de ruta a partir del 2016. Este nuevo sistema de frenado se utilizará únicamente si la media de la distancia en la que se detiene una bicicleta de prueba, durante un frenado de emergencia y bajo condiciones contraladas de velocidad, es menor a 8 metros. Para realizar la evaluación, la empresa tomó una muestra aleatoria de 80 pruebas independientes. Por otra parte, las especificaciones técnicas del nuevo sistema de frenado definen que la desviación estándar de la distancia a la que se detiene una bicicleta de prueba es de $\sigma = 1.6$ metros.

Para evaluar si se debería considerar la implementación del sistema de frenado, la compañía definió las siguientes pruebas estadísticas:

- Prueba 1:** Rechazar H_0 si $\bar{X} \leq 7.77$ metros.
- Prueba 2:** Rechazar H_0 si $\bar{X} \geq 7.75$ metros.

Con base en esta información, responda los siguientes literales:

- (4 puntos)** Halle el Error Tipo I, el Error Tipo II y la potencia de la prueba 1. Utilice como hipótesis alterna $\mu = 7.8$ metros.
- (4 puntos)** Halle el Error Tipo I, el Error Tipo II y la potencia de la prueba 2. Utilice como hipótesis alterna $\mu = 7.8$ metros.

- c. **(2 puntos)** Si la compañía desea evaluar la hipótesis planteada con un nivel de significancia máximo del 0.1, ¿cuál de los dos pruebas establecidas, le recomendaría usar a la compañía?

Punto 7. (20 puntos)

Arfa S.A. es una empresa manufacturera que produce botellas de vidrio y de plástico para suplir la demanda de diferentes empresas embotelladoras. Se define la variable aleatoria X como el tiempo que se demora un producto terminado en salir de la línea producción de botellas de vidrio y la variable aleatoria Y como el tiempo que se demora un producto terminado en salir de la línea de producción de botellas de plástico. Se tomó una muestra aleatoria para caracterizar el tiempo de fabricación promedio en cada una de las líneas de producción, obteniendo los siguientes resultados:

$$\bar{X} = 22.43 \text{ segundos}$$

$$S_X = 2.35 \text{ segundos}$$

$$n_X = 33$$

$$\bar{Y} = 11.98 \text{ segundos}$$

$$S_Y = 3.21 \text{ segundos}$$

$$n_Y = 42$$

El jefe de producción de la compañía estima que existe una diferencia entre los tiempos de producción de ambas líneas, lo cual estaría afectando la etapa de distribución. Por disposiciones particulares de la planta, la diferencia estadística de tiempos de proceso entre la línea de producción de botellas de vidrio y la de botellas de plástico debe ser igual a 10 segundos por ítem producido. Si esta diferencia no se mantiene se generaría un cuello de botella en la etapa de distribución que afectaría el abastecimiento adecuado de la demanda. Con base en esta información, y asumiendo independencia entre las líneas de producción, responda los siguientes literales:

- a. **(4 puntos)** Construya una prueba de hipótesis para determinar si las varianzas poblacionales de los tiempos de producción de las dos líneas son diferentes. Plantee la hipótesis nula y alterna, el estadístico de prueba, la región de rechazo y concluya en términos del problema. Utilice un nivel de significancia del 5%.
- b. **(4 puntos)** Teniendo en cuenta el resultado anterior, plantee una prueba de hipótesis para comprobar si el jefe de producción tiene razón sobre la diferencia del tiempo entre las líneas de producción. Escriba las hipótesis nula y alterna, el estadístico de prueba y concluya en términos del problema. Utilice un nivel de significancia del 10%, y utilice el criterio del p-valor para evaluar la prueba.

Adicionalmente, el jefe de planta desea evaluar la línea de producción de botellas de vidrio únicamente, ya que considera que ésta línea presenta un tiempo de producción mayor a 22 segundos/botella, lo cual representaría una reducción considerable en la cantidad de botellas de vidrio producidas para suplir la demanda del próximo mes.

- c. **(4 puntos)** Suponga que se conoce la varianza poblacional del tiempo que se demora un producto en salir de la línea de producción de botellas de vidrio ($\sigma^2 = 2.4^2$). Plantee una prueba de hipótesis para comprobar si el jefe de planta tiene razón sobre el tiempo de producción en la línea de botellas de vidrio. Plantee las hipótesis nula y alterna, el estadístico de prueba, la región de rechazo y concluya en términos del problema. Utilice un nivel de significancia del 10%.
- d. **(4 puntos)** El jefe de planta tiene dudas sobre la varianza poblacional de la línea de producción de botellas de vidrio ya que considera que es mayor a la establecida en el punto anterior. Plantee las hipótesis nula y alterna, el estadístico de prueba, la región de rechazo y concluya en términos del problema. Utilice un nivel de significancia del 5%.

- e. **(4 puntos)** Asuma ahora que la información acerca de la varianza poblacional del tiempo que se demora un producto en salir de la línea de producción de botellas de vidrio es descartada. Teniendo esto en cuenta, plantee nuevamente una prueba de hipótesis para probar si el jefe de planta tiene razón sobre el tiempo de producción en dicha línea. Escriba las hipótesis nula y alterna, el estadístico de prueba, la región de rechazo y concluya en términos del problema. Utilice un nivel de significancia del 10%.

Punto 8. (12 puntos)

Un grupo de analistas desea realizar un estudio sobre la proporción de estudiantes de ingeniería de dos universidades diferentes que se gradúan con un promedio en la carrera superior a 4.0. Para tal fin, se tomó una muestra aleatoria de 105 estudiantes de la universidad A y 110 estudiantes de la universidad B, obteniendo los siguientes resultados:

Rango Promedio (X)	Universidad A (número de estudiantes)	Universidad B (número de estudiantes)
$3.25 \leq X < 3.50$	17	14
$3.50 \leq X < 3.75$	27	24
$3.75 \leq X < 4.00$	22	32
$4.00 \leq X < 4.25$	24	19
$4.25 \leq X < 4.50$	9	12
$4.50 \leq X < 4.75$	4	3
$4.75 \leq X < 5.00$	2	6

- a. **(4 puntos)** El grupo de analistas asegura que la proporción de estudiantes de la Universidad A con un promedio inferior a 4.0 es menor a 0.65. Construya una prueba de hipótesis que le permita evaluar si la afirmación es verdadera. Concluya a partir del cálculo del p-valor y utilice un nivel de significancia del 5%.
- b. **(4 puntos)** Así mismo, el grupo de analistas asegura que la proporción de estudiantes de la Universidad B con un promedio superior a 4.0 es mayor a 0.4. Construya una prueba de hipótesis que le permita evaluar si la afirmación es verdadera. Concluya a partir del cálculo del p-valor y utilice un nivel de significancia del 5%.
- c. **(4 puntos)** Históricamente la proporción de alumnos de la universidad A que se gradúan con un promedio superior a 4.0 es mayor que los alumnos que se gradúan de manera similar de la universidad B. Construya una prueba que le permita evaluar si esta promoción cumple con la afirmación planteada. Concluya a partir del cálculo del p-valor y utilice un nivel de significancia del 5%.

Punto 9. (17 puntos)

La compañía ColComp Ltda. es una empresa dedicada a la comercialización de productos electrónicos para el sector educativo. Actualmente la empresa está considerando ofrecer un paquete promocional para colegios que consta de 5 tabletas y 3 computadores, el cual espera vender en conjunto a un precio de \$5500 dólares. El gerente de ventas de la compañía considera que dicho precio establecido para cada paquete no es el adecuado. Según él, si se vende la misma cantidad de productos por separado, el precio es incluso menor a los \$5500 dólares. El gerente teme que si

el costo establecido para el paquete promocional es erróneo, esto pueda considerarse como publicidad engañosa y tener consecuencias negativas en un futuro. De esta forma, se tomó una muestra aleatoria del precio de venta (en dólares) de cada producto por separado en la cual se obtuvieron los siguientes resultados:

Producto	Tamaño Muestral	Media Muestral
Tabletas	23	\$490
Computadores	25	\$995

Adicional a esto, de investigaciones anteriores se sabe que el precio en dólares de las tabletas se puede representar por medio de una variable aleatoria $N(\mu_T, \sigma_T^2 = 1080)$, mientras que para los computadores se puede representar por medio de una variable aleatoria $N(\mu_C, \sigma_C^2 = 2020)$. Tenga en cuenta que estas variables son independientes entre sí. Con base en esta información responda los siguientes literales:

- (2 puntos)** Plantee la hipótesis nula y alterna que le permite al CEO de la compañía probar la afirmación sobre el precio de venta del paquete.
- (7 puntos)** Construya y calcule el estadístico de prueba apropiado que le permita evaluar la hipótesis definida en el numeral anterior. Defina paso a paso la construcción del estadístico y no olvide determinar su distribución.
- (2 puntos)** Especifique la región de rechazo y concluya usando un nivel de significancia del 5%.
- (6 puntos)** Construya y calcule un intervalo de confianza del 95% para el precio del paquete promocional a partir de la media de los precios de cada producto por separado.