

EXAMEN FINAL ESTADÍSTICA GENERAL (Ejemplo 1)

- Este examen contiene 15 preguntas con 5 respuestas propuestas cada una. Identificar y marcar la única respuesta correcta en cada caso.
- Se aprueba con 9 repuestas correctamente identificadas.
- Tiempo disponible: 2 horas

Apellido y Nombre:

Problema 1

Suponiendo que el precio de la carne (\$/kg) en un comercio a tomar al azar en el Mercado Central es una variable aleatoria con distribución Normal con $\mu = 36$ y $\sigma = 4$.

1.1) ¿Cuál es el valor aproximado del tercer cuartil de la distribución de probabilidad de la variable mencionada?

- (a) 28,8 \$/kg
- (b) 36,8 \$/kg
- (c) 38,7 \$/kg
- (d) 43,3 \$/kg
- (e) mayor que 45 \$/kg

1.2) ¿Cuál es la probabilidad de que en un comercio a tomar al azar en el Mercado Central el precio de la carne esté entre 33 y 37 \$/kg?

- (a) 0,37
- (c) 0,44
- (b) 0,58
- (d) 0,63
- (e) 1,2

1.3) Si se eligieran al azar 8 comercios en el Mercado Central ¿cuál sería la probabilidad de encontrar al menos uno en el cual el precio de la carne no superara los 33 \$/kg?

- (a) 0,6413
- (b) 0,7490
- (c) 0,1587
- (d) 0,8720
- (e) 0,810

Problema 2

Un profesor universitario termina siempre sus clases varios minutos después de hora. El tiempo adicional que el profesor se toma para cada clase es una variable aleatoria T cuya función de densidad es la siguiente:

$$f(t) = -\frac{2}{9} \cdot t^2 + \frac{14}{9} \cdot t - \frac{20}{9}, \text{ definida en el intervalo } [2 \text{ min}; 5 \text{ min}].$$

Si se elige al azar una de las clases del profesor ¿cuál es la probabilidad de que el tiempo adicional exceda los 3 minutos?

- (a) 0,2592
- (b) 0,4444
- (c) 0,5555
- (d) 0,7407
- (e) 1,7407

Problema 3

En una población formada por 1600 pacientes del Hospital Tornú se han registrado las frecuencias absolutas de pacientes con las diferentes combinaciones de Nivel de Colesterol en la Sangre (Alto, Mediano o Bajo) y Presión Arterial (Alta, Normal o Baja) que figuran en la siguiente tabla de contingencia.

		Presión Arterial		
		Alta	Normal	Baja
Nivel de Colesterol en la Sangre	Alto	116	125	147
	Mediano	126	134	166
	Bajo	148	170	468

3.1) Se elije al azar un paciente de esta población, ¿cuál es la probabilidad de que el paciente tenga Nivel de Colesterol en Sangre Alto y que su Presión Arterial no sea Baja?

- (a) 0,8100 []
- (b) 0,5675 []
- (c) 0,2644 []
- (d) 0,1506 []
- (e) 0,0781 []

3.2) Se elije al azar un paciente con Nivel de Colesterol Alto. Bajo esta condición, ¿cuál es la probabilidad de que la Presión Arterial del paciente no sea Alta?

- (a) 0,6130 []
- (b) 0,7010 []
- (c) 0,3778 []
- (d) 0,2913 []
- (e) 0,7889 []

Problema 4

Un técnico conduce una investigación para comparar la concentración de ácido linoleico en la leche de ovejas de dos razas (Frisona y Sarda). Para ello, planea obtener una muestra aleatoria de ovejas de cada una de dichas razas y determinar el contenido de dicho ácido en la leche de cada una de ellas [g / kg grasa]. Con los datos que obtenga, pondrá a prueba la siguiente hipótesis nula: H_0) *el las medias poblacionales de los contenidos de ácido linoleico en la leche de ovejas tienen igual valor en ambas razas.*

4.1) El nivel de significación que elige para la prueba es $\alpha = 0,05$. Esto implica que:

- (a) Si H_0 fuera verdadera, la probabilidad de rechazarla sería 0,05 []
- (b) Si H_0 fuera verdadera, la probabilidad de no rechazarla sería 0,05 []
- (c) H_0 es falsa si la probabilidad de rechazarla es 0,05 []
- (b) H_0 es verdadera si la probabilidad de no rechazarla es 0,05 []
- (e) Hay una probabilidad de 0,95 de rechazar H_0 si es falsa []

Luego de generados los datos el técnico calculó los estadísticos correspondientes a las muestras de cada raza que figuran en la tabla:

Raza	Tamaño de muestra (n)	Media muestral (\bar{x})	Desvío estándar (s)
Frisona	6	25,52 g / kg grasa	2,416 g / kg grasa
Sarda	8	20,74 g / kg grasa	2,544 g / kg grasa

4.2) Si fuera razonable dar por cierto que las varianzas de las concentraciones de ácido linoleico son iguales entre razas, ¿cuál sería el valor del estimador insesgado de la varianza de los contenidos de ácido linoleico en las ovejas de ambas razas?

- (a) 6,1545 (g / kg grasa)² []
- (b) 6,2074 (g / kg grasa)² []
- (c) 3,4728 g / kg grasa []
- (d) 2,4914 g / kg grasa []
- (e) 5,2274 g / kg grasa []

- 4.3) ¿Por qué decimos que el estimador de la varianza de la concentración de ácido linoleico indicado en el punto anterior es insesgado?
- (a) Porque cuando aumenta el tamaño de la muestra se acerca al verdadero valor de la varianza []
 - (b) Porque su esperanza coincide con el verdadero valor de la varianza que intenta estimar []
 - (c) Porque coincide con el verdadero valor de la varianza que intenta estimar []
 - (d) Porque su esperanza es muy precisa []
 - (e) Porque su varianza es mínima []
- 4.4) A partir de los datos obtenidos, ¿en qué rango se encuentra el valor p para la prueba de la hipótesis propuesta?
- (a) $0,001 < \text{valor } p < 0,01$ []
 - (b) $0,025 < \text{valor } p < 0,05$ []
 - (c) $0,01 < \text{valor } p < 0,02$ []
 - (d) $0,1 < \text{valor } p < 0,2$ []
 - (e) $0,2 < \text{valor } p < 0,4$ []
- 4.5) ¿Qué significado tiene un intervalo del 95% de confianza construido para la diferencia entre las concentraciones promedio de ácido linoleico en la leche de la raza Frisona y Sarda?
- (a) Podemos tener un nivel de confianza de 0,95 en que el intervalo contiene al verdadero valor de la diferencia entre las concentraciones promedio de ácido linoleico en la leche de la raza Frisona y Sarda. []
 - (b) Podemos tener un nivel de confianza de 0,95 en que la media muestral de las diferencias está dentro del intervalo. []
 - (c) La probabilidad de que este intervalo contenga a la media muestral de las diferencias es de 0,95. []
 - (d) Es un intervalo que contiene al promedio de las concentraciones de ácido linoleico de la leche en todas las ovejas de ambas razas. []
 - (e) Ninguno de los anteriores. []

Problema 5

El chañar (*Geoffrea decorticans*) es una especie de árbol que invade campos de pastoreo y de cultivo en regiones semiáridas de la Argentina. Con el propósito de establecer la posible asociación entre la presencia o ausencia de esta especie y la modalidad de uso de la tierra en la provincia de La Pampa, se decide obtener una muestra aleatoria formada por 100 campos a elegir al azar dentro de dicha región y registrar en cada uno la presencia o ausencia de chañar y la modalidad de uso de la tierra (agrícola o ganadera). Con estos datos se realizará una prueba de independencia.

Si π_{ij} es la probabilidad conjunta de que un campo a tomar al azar corresponda a la i -ésima categoría de presencia o ausencia de chañar y a la j -ésima categoría de modalidad de uso de la tierra, $\pi_{i\cdot}$ es la probabilidad de que un campo a tomar al azar corresponda a la i -ésima categoría de presencia o ausencia de chañar y $\pi_{\cdot j}$ es la probabilidad de que un campo tomado al azar corresponda a la j -ésima modalidad de uso de la tierra.

5.1) ¿Cuál es la hipótesis nula que el investigador pondrá a prueba?

- (a) $\pi_{ij} \neq \pi_{i\cdot} \cdot \pi_{\cdot j}$, para $i, j = 1, 2$. []
- (b) $\pi_{ij} = \pi_{i\cdot} \cdot \pi_{\cdot j}$, para $i, j = 1, 2$. []
- (c) $\pi_{i\cdot} = \pi_{\cdot j}$, para $i, j = 1, 2$. []
- (d) $\pi_{11} = \pi_{12} = \pi_{21} = \pi_{22}$ []
- (e) Ninguna de las anteriores. []

Luego de obtenidos los datos se calcularon las frecuencias observadas que se presentan en la siguiente tabla:

Uso de la tierra	Chañar	
	Ausente	Presente
Agrícola	35	19
Ganadero	20	26

5.2) ¿Cuál de los siguientes valores de α es el mínimo nivel de significación con el cual se rechaza la hipótesis nula respectiva?

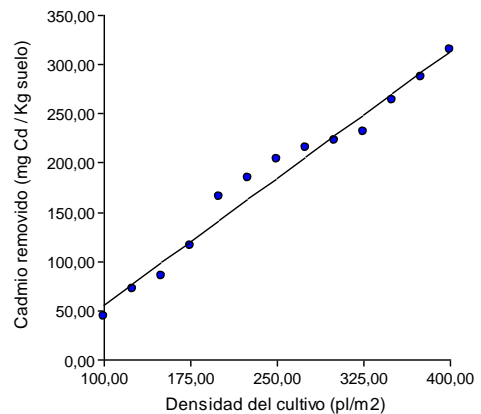
- (a) $\alpha = 0,2$ []
- (b) $\alpha = 0,1$ []
- (c) $\alpha = 0,05$ []
- (d) $\alpha = 0,025$ []
- (e) $\alpha = 0,01$ []

Problema 6

Una de las técnicas que se pueden aplicar para remediar la contaminación del suelo con cadmio es establecer un cultivo de geranios, ya que dichas plantas que absorben y retienen este contaminante. Con el propósito de evaluar la posible relación estadística entre la densidad de los cultivos de geranio y la cantidad de cadmio que extraen, se seleccionaron al azar 13 parcelas en un suelo homogéneamente contaminado con cadmio y en cada una se estableció un cultivo de geranio con una densidad preestablecida entre 100 y 400 plantas/m². Al cabo de 2 meses, se evaluó la cantidad de cadmio removido del suelo de cada parcela (mg de Cd / kg de suelo). Con los datos obtenidos se realizó un análisis de regresión lineal simple cuyos principales resultados se presentan a continuación

Variable	N	R ²
Cadmio removido	13	0,97

Coef	Est.	E.E.	T	valor p
b ₀	-29,84	11,34	-2,63	0,0233
b ₁	0,86	0,04	20,23	< 0,0001



6.1) ¿Qué se concluye cuando se rechaza la hipótesis nula que dice la pendiente de la recta de regresión es igual a cero?

- (a) El aumento promedio en la cantidad de cadmio que los cultivos de geranio extraen del suelo cuando la densidad del cultivo aumenta en 1 pl/m² es igual a cero. []
- (b) La cantidad promedio de cadmio removido del suelo por un cultivo de geranio varía con la densidad del cultivo (entre 100 y 400 pl/m²) []
- (c) La cantidad promedio de cadmio removido del suelo por un cultivo de geranio con densidad igual a 0 pl/m² es 0 g de Cd / kg de suelo. []
- (d) La cantidad promedio de cadmio removido del suelo por un cultivo de geranio no varía con la densidad del cultivo. []
- (e) El error estándar de la estimación de la pendiente es menor que el de la estimación de la ordenada al origen. []

6.2) ¿Qué porcentaje de la variación en la cantidad de cadmio extraída del suelo fue explicada por las diferencias en la densidad de los cultivos de geranio?

- (a) < 0,0001 []
- (b) 97 []
- (c) 2,33 []
- (d) 4 []
- (e) 13 []

RESPUESTAS CORRECTAS:

- 1.1 c, 1.2 a, 1.3 d,
- 2 d
- 3.1 d, 3.2 b
- 4.1 a, 4.2 b, 4.3 b, 4.4 a, 4.5 a
- 5.1 b, 5.2 c,
- 6.1 b, 6.2 b