

# EXAMEN FINAL ESTADÍSTICA GENERAL (Ejemplo 5)

## Apellido y nombre:

- 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

### Problema 1

Un ingeniero agrónomo debe decidir si dos variedades de trigo (A y B) tienen igual rendimiento esperado bajo las condiciones de cultivo habituales en el partido de Carlos Pellegrini (Buenos Aires). Para ello, obtiene dentro del partido muestras aleatorias de lotes cultivados con cada una de estas variedades y registra el rendimiento de cada lote (en quintales/ha). Los datos obtenidos aparecen resumidos en la siguiente tabla:

Variedad de trigo	A	B
Número de parcelas	15	17
Rendimiento promedio [qq/ha]	32,0	38,0
Desvío estándar [qq/ha]	3,0	4,0

1.1) ¿Cuál de las siguientes es la hipótesis nula que debería plantear el ingeniero?

- (a)  $\mu_A = \mu_B$  [ ]
- (b)  $\mu_A \leq \mu_B$  [ ]
- (c)  $\mu_A \geq \mu_B$  [ ]
- (d)  $\bar{X}_A = \bar{X}_B$  [ ]
- (e)  $\bar{X}_A \leq \bar{X}_B$  [ ]

1.2) Si se considera razonable dar por cierto que los rendimientos en lotes a tomar al azar son variables aleatorias con distribución Normal y que las varianzas poblacionales correspondientes a ambas variedades son iguales y con los datos disponibles se realiza la prueba de hipótesis con un nivel de significación  $\alpha = 0,05$ , el resultado es:

- (a)  $0,05 \leq \text{valor } p < 0,1$ . Se rechaza la hipótesis nula. [ ]
- (b)  $0,01 \leq \text{valor } p < 0,05$ . Se rechaza la hipótesis nula. [ ]
- (c)  $0,001 \leq \text{valor } p < 0,01$ . Se rechaza la hipótesis nula. [ ]
- (d)  $\text{valor } p < 0,001$ . Se rechaza la hipótesis nula. [ ]
- (e) Se acepta la hipótesis nula. [ ]

1.3) El nivel de significación  $\alpha = 0,05$  adoptado implica que:

- (a) La probabilidad de concluir que las dos variedades tenían distinto rendimiento esperado en el caso en que efectivamente lo tuvieran era del 5% [ ]
- (b) La probabilidad de concluir que las dos variedades tenían igual rendimiento esperado en el caso en que no lo tuvieran era del 5% [ ]
- (c) La probabilidad de rechazar la hipótesis nula en el caso en que las dos variedades tuvieran diferente rendimiento esperado era del 5% [ ]
- (d) La probabilidad de concluir que las dos variedades tenían distinto rendimiento esperado cuando en realidad tuvieran igual rendimiento esperado era del 5% [ ]
- (e) La probabilidad de cometer error de tipo II era del 95% [ ]

### Problema 2

En un lote de semillas de una especie vegetal en peligro de extinción que han sido coleccionadas para cultivar en un jardín botánico, sólo un 5% de las semillas están en condiciones de germinar. El técnico a cargo prepara bandejas donde pone a germinar grupos de 20 de semillas elegidas al azar de dicho lote.

2.1) ¿Cuál es la probabilidad de que en una bandeja cualquiera germinen por lo menos 2 semillas?

- (a) 0.1887
- (b) 0.2642
- (c) 0.7358
- (d) 0.9245
- (e) 1,0000

2.2) ¿Qué valores tienen la esperanza y el desvío estándar del número de semillas que germinan por bandeja?

- (a) 1 y 0,9747
- (b) 1 y 0,9500
- (c) 1 y 0,0475
- (d) 0,05 y 0,0475
- (e) Ninguno de los anteriores

### Problema 3

Para evaluar la contaminación con nitratos del agua de los pozos de una región agrícola, se planea tomar una muestra aleatoria de 30 pozos de la región, extraer agua de cada uno de los pozos y medir su contenido de nitratos  $X_i$  [en ppm = partes por millón]. Con los datos a obtener, se calcularán los siguientes estadísticos:

$$\bar{X} = \frac{1}{30} \sum_{i=1}^{30} X_i \qquad S^2 = \frac{1}{29} \sum_{i=1}^{30} (X_i - \bar{X})^2$$

3.1) ¿Qué es  $S^2$ ?

- (a) La suma de los cuadrados de las diferencias entre los contenidos de nitratos observados y la media muestral.
- (b) El estimador insesgado de la suma de los cuadrados de los contenidos de nitratos de todos los pozos de la Pampa Ondulada.
- (c) El estimador insesgado de la varianza de los contenidos de nitratos de todos los pozos de la Pampa Ondulada.
- (d) El verdadero valor de la varianza de los contenidos de nitratos de todos los pozos de la Pampa Ondulada.
- (e) El valor esperado del cuadrado de la diferencia entre los contenidos de nitratos en el agua de los pozos de la Pampa Ondulada y su media poblacional.

3.2) Suponiendo que el promedio de los contenidos de nitratos de todos los pozos de la región fuera 35,0 ppm y la varianza de dichos contenidos fuera 42 ppm<sup>2</sup>, ¿aproximadamente cuánto valdría la probabilidad de que  $\bar{X}$  tomara un valor mayor que 36 ppm?

- (a) 1,0000
- (b) 0,8010
- (c) 0,2375
- (d) 0,1990
- (e) 0,0000

3.3) Suponiendo que el promedio de los contenidos de nitratos de todos los pozos de la región fuera 35,0 ppm y la varianza de dichos contenidos fuera 42 ppm<sup>2</sup>, ¿cuánto valdría la probabilidad de que  $\bar{X} = 35$  ppm?

- (a) 1,0000 [ ]
- (b) 0,5000 [ ]
- (c) 0,7374 [ ]
- (d) 0,0000 [ ]
- (e) No se puede calcular [ ]

**Problema 4**

Para evaluar la relación estadística entre la ganancia de carbono las hojas de girasol y la intensidad de la luz que reciben, se midió la ganancia de carbono [mg CO<sub>2</sub>/dm<sup>2</sup>/hora] de hojas sometidas a intensidades de luz preestablecidas ntre 0 Watt/m<sup>2</sup> (oscuridad) y 500 Watt/m<sup>2</sup>. A continuación se presenta parte de los resultados de un análisis de regresión lineal simple realizado con los datos obtenidos.

Variable	N	R <sup>2</sup>		Est.	EE	valor p
Ganancia neta de carbono	16	0,97	b <sub>0</sub>	-4,59	1,23	0,002
			b <sub>1</sub>	0,36	0,02	<0,001

4.1) Según estos resultados, la hipótesis nula que dice que *la media poblacional de la ganancia neta de carbono en la oscuridad es igual a cero*:

- (a) se acepta [ ]
- (b) no se rechaza con  $\alpha = 0,001$  porque valor  $p > 0,001$  [ ]
- (c) no se rechaza con  $\alpha = 0,001$  porque valor  $p < 0,001$  [ ]
- (d) se rechaza con  $\alpha = 0,001$  porque valor  $p > 0,001$  [ ]
- (e) se rechaza con  $\alpha = 0,001$  porque valor  $p < 0,001$  [ ]

4.2) La varianza estimada del estimador de la pendiente de la recta de regresión es:

- (a) 0,0004 [ ]
- (b) 0,0200 [ ]
- (c) 0,9700 [ ]
- (d) 1,2300 [ ]
- (e) 1,5129 [ ]

**Problema 5**

Un genetista que estudia la herencia del color de las flores de *Vicia spp.* está interesado en poner a prueba la hipótesis nula que dice que el cruzamiento entre dos plantas con flores rosadas produce una descendencia formada por 25% de plantas con flores blancas, 50% de plantas con flores rosadas y 25% de plantas con flores rojas. Para ello, examina 50 plantas provenientes de ese tipo de cruzamiento y encuentra que 14 tienen flores blancas, 18 tienen flores rosadas y 19 tienen flores rojas. La hipótesis nula que pondrá a prueba con estos datos es:

- (a) H<sub>0</sub>: la proporción en la descendencia es 50:50 [ ]
- (b) H<sub>0</sub>: las proporciones en la descendencia son 28:36:38 [ ]
- (c) H<sub>0</sub>: las proporciones en la descendencia son 1/3:1/3:1/3 [ ]
- (d) H<sub>0</sub>: las proporciones en la descendencia son 25:50:25 [ ]
- (e) H<sub>0</sub>: no hay buen ajuste a las proporciones esperadas [ ]

### Problema 6

Si  $X$  es una variable aleatoria continua que puede tomar valores entre 0 y 1 y la función de densidad de probabilidad de  $X$  es  $f(x) = 1,5(1 - x^2)$ .

6.1) ¿Cuál es la probabilidad de que  $X$  tome un valor comprendido en el intervalo  $[0,2 ; 0,7]$ ?

- (a) -0,0050
- (b) 0,2960
- (c) 0,5825
- (d) 0,8785
- (e) 3,9825

6.2) ¿Cuanto vale la esperanza de  $X$ ?

- (a) 0,375
- (b) 0,500
- (c) 0,750
- (d) 1,000
- (e) 1,500

### Problema 7

En un estudio sobre la susceptibilidad de plantas de duraznero a dos cepas diferentes de un virus, se tomaron de un vivero 5 plantas al azar; en cada planta se seleccionaron 2 hojas y cada una fue inoculada con una de las dos cepas virales. Al cabo de una semana, se midieron las áreas de las lesiones producidas por los virus en las hojas inoculadas (en  $\text{mm}^2$ ). Los datos obtenidos figuran en la tabla:

Planta	1	2	3	4	5
Lesión de cepa viral A [ $\text{mm}^2$ ]	40	47	52	50	49
Lesión de cepa viral B [ $\text{mm}^2$ ]	42	42	46	43	47

7.1) A partir de estos datos, el intervalo del 95% de confianza para la diferencia media entre el área de las lesiones producidas por la cepa A menos la de las lesiones producidas por la cepa B es:

- (a) (-1,739 ; 8,938)
- (b) (-0,928 ; 8,128)
- (c) (-0,089 ; 7,289)
- (d) (1,739 ; 8,938)
- (e) (0,928 ; 8,128)

7.2) En relación al intervalo del 95% de confianza calculado en el punto anterior se puede decir:

- (a) Podemos tener un nivel de confianza igual a 0,95 en que ese intervalo contiene al verdadero valor de la diferencia media entre las áreas de las lesiones producidas por ambas cepas
- (b) El verdadero valor de la diferencia media entre las áreas de las lesiones producidas por ambas cepas está dentro del intervalo construido
- (c) La probabilidad de que el verdadero valor de la diferencia media entre las áreas de las lesiones producidas por ambas cepas esté en el intervalo construido es de 0,95
- (d) La probabilidad de que la diferencia entre las medias muestrales de los tamaños de las lesiones producidas por ambas cepas esté en ese intervalo es de 0,95.
- (e) Es más amplio (menos preciso) que el intervalo del 99% de confianza calculado con los mismos datos.

Respuestas correctas:

1.1 a, 1.2 d, 1.3 d, 2.1 b, 2.2 a, 3.1 c, 3.2 d, 3.3 d, 4.1 b, 4.2 a, 5 d, 6.1 c, 6.2 a, 7.1 b, 7.2 a