

EXAMEN FINAL DE INFERENCIA ESTADÍSTICA – período 091
Cali, mayo 21 de 2009.

1. La cantidad de tiempo que tarda un pagador con cada cliente tiene una distribución normal con una media de 3 minutos y una desviación de 30 segundos. Si se seleccionan 16 clientes ¿Cuál es la probabilidad de que el pagador se demore con estos 16 clientes un tiempo promedio que difiera de la media de la población en más de 15 segundos?

(5%)

2. Una empresa farmacéutica quiere que la concentración de impurezas de sus píldoras no supere el 3%. Se sabe que la concentración de impurezas de un lote tiene una distribución normal que tiene una desviación estándar de 0.4%. Se toma una muestra de 64 píldoras de un lote y se encuentra una concentración promedio de 3.07%.
- Utilizando un nivel de significancia del 5%, confirman los datos que el contenido de impurezas de las píldoras es superior al 3%?
 - Construya un intervalo de confianza del 90% para el contenido promedio de impurezas de las píldoras. Compare el resultado con la decisión anterior.
 - Calcule la probabilidad de cometer error tipo II y la potencia de la prueba, si la verdadera media de concentración de impurezas fuera 3.1%.
 - Calcule el tamaño de muestra que se debe evaluar, si se desea tener una confianza del 98% y un margen de error de 0.03%.

(10%)

3. Una empresa desea comparar el nivel de calidad de dos proveedores de materia prima, la característica a comparar es la dureza del material cuya característica es "mientras mayor mejor" el proveedor A es su proveedor actual y la compañía lo conservará mientras la dureza promedio NO sea superada por el proveedor B. Para cada proveedor se tomó una muestra aleatoria independiente midiendo la respectiva dureza del material (en unidades de dureza) y obteniendo los siguientes resultados (MUESTRALES):

	PROVEEDOR A	PROVEEDOR B
n	50	50
promedio	14	16
desv. Estandar	3.1	2.9

- a. ¿Considera usted que se debe cambiar de proveedor? Justifique su respuesta.

(10%)

4. Una cadena de restaurantes de comida rápida desea determinar si existe diferencia entre tres medios de comunicación: Revistas, Televisión y Radio; en términos de que el consumidor recuerde el anuncio. Los resultados de un estudio de publicidad son los siguientes:

Habilidad para recordar	MEDIO DE PUBLICIDAD			Total
	REVISTA	T.V.	RADIO	

Número de personas que lo recuerdan	55	10	7	72
Número de personas que no lo recuerdan	73	93	108	274
Total	128	103	115	346

- Para un nivel de significancia del 0.05, determine si existen indicios de un efecto significativo del medio en cuanto a la proporción de individuos que pueden recordar el anuncio.
- Si existen diferencias significativas entre las proporciones de individuos que recuerdan los anuncios, determine cuál de los medios de comunicación sería el más recomendable para continuar realizando la publicidad de la cadena de negocios de comida rápida. Sustente.

(25%)

5. Un profesor de estadística evalúa su materia a partir de tres parciales y un final, el profesor desea saber si el rendimiento de los alumnos, ha ido aumentando. Para ello seleccionó una muestra de 5 alumnos. Obteniendo las siguientes calificaciones:

Parcial I	Parcial II	Parcial III	Final
2,5	3,0	3,5	4,7
3,2	3,5	3,0	5,0
2,3	2,5	3,1	3,9
1,7	2,3	2,8	4,0
2,8	3,2	3,6	4,5

Después de procesar la información, encontró los siguientes resultados:

Anova de Levene

Fuente de Variación	Grados de libertad	Suma de cuadrados	Cuadrados Medios	$F_{calculada}$	valor-p	$F_{critico}$
Entre grupos	3		0,0193		0,853	3,239
Dentro de los grupos		1,192				
Total		1,25				

Anova para los promedios

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados Medios	$F_{calculada}$	valor-p	$F_{critico}$
Entre grupos	3		3,427		0,000057	3,24
Dentro de los grupos		3,568				
Total		13,850				

Procedimiento de Tukey-Kramer

Group	Sample Mean	Sample Size	Comparison	Absolute Difference
1	2,5	5	Group 1 to Group 2	0,4
2	2,9	5	Group 1 to Group 3	0,7
3	3,2	5	Group 1 to Group 4	1,92
4	4,42	5	Group 2 to Group 3	0,3
			Group 2 to Group 4	1,52
			Group 3 to Group 4	1,22

Other Data	
Level of significance	0,05
Numerator d.f.	4
Denominator d.f.	16
MSW	0,223
Q Statistic	4,05

a. ¿Puede afirmarse que los alumnos obtienen mejores notas en la medida que el curso avanza?

b. ¿Los alumnos obtienen mejores notas en el último examen?

Diligencie los espacios en blanco de las tablas, planteé las hipótesis pertinentes, justifique si el procedimiento de Tukey-Kramer se debía aplicar, tome las decisiones con un $\alpha = 0,05$ e interprete los resultados.

(15%)

6. La gerencia de una aerolínea considera que hay una relación entre los gastos publicitarios y el número de pasajeros que la escogen. Para determinar si esta relación existe se tomaron 15 observaciones con la siguiente información:

Publicidad (en US\$1000's)	10	12	8	17	10	15	10	14	19	10	11	13	16	10	12
Pasajeros (000)	15	17	13	23	16	21	14	20	24	17	16	18	23	15	16

La información se procesó en Excel, y entre otros resultados, se encontró lo siguiente:

ANOVA de la regresión

Fuente de Variación	Grados de libertad	Suma de cuadrados	Cuadrados Medios	F calculada	valor-p
Regresión		161,04			3,238E-09
Residuos			0,82		
Total		171,73			

$$\sum_{i=2}^{15} (e_i - e_{i-1})^2 = 26.5588$$

$$\sum_{i=1}^{15} e_i^2 = 10.6892$$

Diferencias entre la distribución empírica de los residuales y la correspondiente distribución teórica normal:

Dmax: 0,1235

Dmin: - 0,1059

El coeficiente de determinación R^2 de la regresión auxiliar entre los cuadrados de los residuales y los valores estimados de la variable de respuesta: 0,0013

Con los resultados anteriores:

- Encuentre el modelo apropiado para las 2 variables e interprete los coeficientes que encuentre.
- Diligencie la tabla de ANOVA de la regresión y con ella, plantee y resuelva la hipótesis sobre la significancia global del modelo.
- Encuentre el coeficiente de determinación R^2 e interprete su significado.
- Plantee las hipótesis y resuelva para los 4 supuestos del modelo de regresión indicando los valores críticos y calculados de las pruebas y su decisión en cada hipótesis y concluya si el modelo encontrado es apropiado para pronosticar.
- Por último, cuando $X=10$ usted encontró los siguientes intervalos: $(14.54 - 15.85)$ y $(13.13 - 17.26)$ Identifique cuál de ellos es el intervalo de confianza y cuál el de predicción e interprete cada uno de los intervalos, dependiendo de su conclusión en el literal anterior.

(35%)

Scatter Diagram

