



Examen Final - Inferencia Estadística - período 2013 1  
Santiago de Cali, Mayo 9 de 2013 - Prof. Ma. Amparo Giraldo Q.

- (20%) Para estudiar el problema del uso del celular mientras se maneja, se realizó una encuesta a 500 conductores que utilizan celular mientras manejan, encontrándose que el 46% de los encuestados dijeron haber tenido que virar bruscamente cuando manejando hablaba por celular. Construya un intervalo de confianza del 96% para el porcentaje de conductores real que ha tenido que virar bruscamente cuando conduciendo habla por celular. Interprete su respuesta.
- (20%) En una fábrica de comida para mascotas hay dos líneas de llenado en latas de 8 onzas, se tomaron muestras de dos líneas de llenado encontrándose los siguientes resultados:

	Línea A	Línea B
Promedios de la muestra	8.005	7.997
Desviaciones estándar muestrales	0.012	0.005
Tamaños de muestra	11	16

$\alpha = 0,02$

Suponiendo que las varianzas son iguales, con un nivel de significancia de 0.025:

- Existe evidencia de la diferencia entre la media del peso de las latas llenadas en las dos líneas de llenado? (plantee las hipótesis, grafique y concluya)
- Aproxime del valor p

- (20%) Se realiza un estudio sobre Edad Vs empleo a 1000 personas de cierta Ciudad, encontrándose los siguientes resultados.

Género	Empleado	Desempleado	Total
Menor i iguala 25 años	200	100	300
De 26 a 40 años	400	50	450
De 41 o más años	100	150	250
Total	700	300	1000

- Plantee las hipótesis adecuadas para probar si existe evidencia que el desempleo tiene alguna relación con la Edad
- Pruebe la hipótesis con un nivel de significancia del 0.01.

- (20%) Los siguientes datos corresponden al número de cirugías diarias realizadas en 3 hospitales, en cinco días de la semana anterior.

ANÁLISIS DE VARIANZA

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Día semana -Bloques	78,27777778	5	15,65555556		0,649471897	
Hospital -Tratamientos	1083,444444	2	541,7222222		0,000166303	4,102821015
Error	230,5555556	10	23,05555556			
Total	1392,277778	17				

- Complete la tabla de ANOVA
- Plantee y grafique las hipótesis para adecuadas
- Con un nivel de significancia del 0.05, puede usted concluir que existe alguna diferencia significativa en el número de cirugías diarias por Hospital o por día de la semana?

- (20%) Los siguientes resultados corresponden al análisis de regresión, busca establecer si existe alguna relación lineal entre las variables número de faltas a clase de Inferencia Estadística y Nota definitiva obtenida por una muestra de estudiantes en el año 2012;  $r = -0.78$

ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	1	5,887340322	5,887340322	27,1191735	7,11364E-05
Residuos	17	3,690554415	0,217091436		
Total	18	9,577894737			

	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%
Intercepción	4,691581109	0,16890669	27,77617097	1,31975E-15	4,335219146	5,047943072
Faltas	-0,23963039	0,04601545	-5,207607272	7,11364E-05	-0,336714502	-0,142546279

- Identifique las variables independiente y dependiente y especifique el modelo en términos del ejercicio
- Interprete el coeficiente de determinación
- Interprete los parámetros de Modelo, en términos del ejercicio.
- De acuerdo a los resultados considera que el modelo es adecuado? Explique

**FÓRMULAS:**

**3.1. ESTIMACIÓN DEL INTERVALO DE CONFIANZA PARA LA MEDIA -  $\mu$**

- Con varianza  $\sigma$  conocida:

$$\bar{X} - Z \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z \frac{\sigma}{\sqrt{n}}$$

- Con varianza  $\sigma$  desconocida:

$$\bar{X} - t_{n-1} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{n-1} \frac{S}{\sqrt{n}}$$

**3.1. ESTIMACIÓN INTERVALO DE CONFIANZA PARA LA PROPORCIÓN -  $\pi$**

$$P - Z \sqrt{\frac{p(1-p)}{n}} \leq \pi \leq P + Z \sqrt{\frac{p(1-p)}{n}}$$

**PRUEBAS DE HIPÓTESIS PARA UNA MUESTRA:**

- Para el Promedio o Media – con varianza conocida

Estadístico de la prueba:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

**PRUEBAS DE HIPÓTESIS PARA UNA MUESTRA:**

- Para el Promedio o Media – con varianza conocida

Estadístico de la prueba:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- Para el Promedio o Media – con varianza desconocida

Estadístico de la prueba:

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

- Para la Proporción

Estadístico de la prueba:

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

Donde:

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \quad \text{Es el error estándar para la proporción}$$

**PRUEBAS DE HIPÓTESIS PARA DOS MUESTRAS INDEPENDIENTES:**

- Para la Diferencia de promedios – con varianzas desconocidas e iguales

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$G.L = n_1 + n_2 - 2$$

c. Para la Diferencia de promedios – con varianzas desconocidas desiguales

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$gl = \frac{\left[ \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \right]^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}}$$

**ANÁLISIS DE VARIANZA (ANOVA):**

Fuente de Variación	Suma de Cuadrados	Grados de Libertad	Media Cuadrática	F
Tratamientos	SST	k-1	MST=SST/k-1	MST/MSE
Error	SSE	n-k	MSE=SSE/n-k	
Total	SSTotal	n-1		

Valor Crítico para el estadístico F:  $F_{\alpha, k-1, n-k}$

**ANÁLISIS DE VARIANZA (ANOVA EN DOS VÍAS):**

Fuente de Variación	Suma de Cuadrados	Grados de Libertad	Media Cuadrática	Fcal
Tratamientos	SST	k-1	MST=SST/k-1	MST/MSE
Bloques	SSB	b-1	MSB=SSB/b-1	MSB/MSE
Error	SSE	(k-1)*(b-1)	MSE=SSE/((k-1)*(b-1))	
Total	SSTotal	n-1		

Valor Crítico para el estadístico F tratamientos:  $F_{\alpha, k-1, (k-1)*(b-1)}$

Valor Crítico para el estadístico F bloques:  $F_{\alpha, b-1, (k-1)*(b-1)}$

**PRUEBA  $\chi^2$  – CUADRADA PARA TABLAS DE CONTINGENCIA Y BONDAD DE AJUSTE:**

$$\chi^2 - \text{cuadrado cal} = \sum \left| \frac{(f_{os} - f_e)^2}{f_e} \right|$$

g.l.=(filas - 1)x(columnas - 1) - Tablas de contingencia

g.l.=(celdas - 1) – bondad de ajuste