

Examen Final de Inferencia Estadística
Cali, Mayo 25 de 2006

NO SE RESPONDERÁN PREGUNTAS DURANTE EL EXAMEN.

1. Un comerciante al menudeo lleva a cabo un estudio para determinar la relación entre los gastos de publicidad y las ventas (en períodos semanales). Se registró la siguiente información para una muestra de tamaño $n=15$. Se ajustó un modelo de regresión lineal para explicar las ventas en función de los gastos de publicidad y a continuación se presentan los resultados obtenidos en Microsoft Excel.

TABLA DE ANÁLISIS DE VARIANZA

| Fuente de Variación | Grados de Libertad | Suma de Cuadrados | Cuadrados Medios | Estadístico F_{calc} | Valor-p |
|---------------------|--------------------|-------------------|------------------|------------------------|----------|
| Regresión | | | | | 0.000000 |
| Residuales | | 10.6893 | | | |
| Total | | 171.7333 | | | |

TABLA DE CÁLCULO DE LOS COEFICIENTES DEL MODELO

| | Coefficientes | Error típico | Estadístico t | Probabilidad | Inferior 95% | Superior 95% |
|----------------|---------------|--------------|---------------|--------------|--------------|--------------|
| Intercepto | 4,386254 | 0,991282 | 4,424829 | 0,000686 | 2,244718 | 6,527789 |
| publicidad(\$) | | 0,077265 | 13,994918 | 0,000000 | 0,914396 | 1,248237 |

Estadístico Durbin-Watson: 2,49 ($d_L=1,08$; $d_U=1,36$)

- Complete los espacios en blanco de las tablas arriba descritas e interprete el resultado de la tabla del análisis de varianza.
 - Escriba el modelo de regresión lineal ajustado a esta serie de datos e interprete la pendiente del modelo.
 - Plantee, resuelva e interprete el resultado de las pruebas de hipótesis inherentes a β_0 y β_1 con un $\alpha=0,05$ escogiendo el valor t de la tabla adecuado entre los colocados en los anexos.
 - Calcule e interprete el coeficiente de determinación del modelo.
 - Interprete el resultado del estadístico Durbin-Watson, suponiendo que la normalidad de los residuales ya está probada con anterioridad.
- Valor del punto: 40%

2. Los clientes de un concesionario de autos de lujo se clasifican en tres niveles de ingreso (Bajo, Medio, Alto); se tomó una muestra aleatoria de clientes para determinar si existe alguna relación entre el nivel de ingresos y la importancia que dan al precio de los autos (Poca, Moderada, Grande); las frecuencias obtenidas de la muestra son:

| | INGRESO | | | |
|-------------|---------|-------|------|-------|
| IMPORTANCIA | Bajo | Medio | Alto | TOTAL |
| Grande | 83 | 62 | 37 | 182 |
| Moderada | 52 | 71 | 49 | 172 |
| Poca | 63 | 58 | 63 | 184 |
| TOTAL | 198 | 191 | 149 | 538 |

Plantee, resuelva e interprete la hipótesis respectiva para establecer si existe relación entre las variables nivel de ingresos y nivel de importancia otorgada al precio con un $\alpha=0,05$ escogiendo para la comparación el valor teórico de la tabla adecuado entre los colocados en los anexos que le aplica a esta prueba de hipótesis.

Valor del punto: 20%

En los puntos 3, 4 y 5 encierre en un círculo la respuesta que usted piense es correcta.

3. Los resultados de la encuesta antes del día de elecciones en cuanto a las personas a favor de un determinado candidato presidencial, afirman que existe mayor preferencia en el estrato 4 que en el 2 por dicho candidato. Usted es contratado para evaluar dicha afirmación. ¿Qué prueba de hipótesis plantearía?

- a. Diferencia de Medias.
- b. Diferencia de Medias de datos apareados.
- c. Igualdad de Varianzas.
- d. Diferencia de Proporciones de una sola cola.
- e. Diferencia de Proporciones de dos colas.

Valor del punto: 5%

4. En una empresa se obtuvo, mediante una encuesta a 36 de sus empleados de la parte operativa y un intervalo de confianza del 95%, que el salario promedio estaba entre (\$450 mil, \$600 mil). Si se desea cambiar el nivel de confianza a 99%, el intervalo sería:

- a. Más estrecho e involucraría un riesgo más grande de ser incorrecto.
- b. Más ancho e involucraría un riesgo más pequeño de ser incorrecto.
- c. Más estrecho e involucraría un riesgo más pequeño de ser incorrecto.
- d. Más ancho e involucraría un riesgo más grande de ser incorrecto.
- e. Más ancho pero no puede determinarse si el riesgo de ser incorrecto sería más grande o más pequeño.

Valor del punto: 5%

5. Si usted aumenta la región de rechazo en una prueba de hipótesis, lograría que:

- a. Se disminuyan los errores tipo I y II.
- b. Se disminuya solo uno de los errores.
- c. Se disminuya el tamaño de la muestra.
- d. Se disminuya la variabilidad.
- e. Ninguna de las anteriores opciones.

Valor del punto: 5%

6. Un director de recursos humanos, decide investigar la percepción de los empleados sobre la equidad de dos métodos de evaluación del desempeño. Para probar la diferencia entre los dos métodos, 160 empleados se asignaron al azar para ser evaluados con uno de los dos métodos: 78 se asignaron al método 1, donde los individuos proporcionan retroalimentación al supervisor como parte del proceso de evaluación; 82 se asignaron al método 2, donde los individuos aportan la auto evaluación de su desempeño. Después de las evaluaciones, se preguntó a los empleados si consideraban justa o injusta la evaluación del desempeño. De los 78 empleados en el método 1, 63 dijeron que era justa; 49 de los 82 empleados en el método 2 dijeron que era justa.

- a. Con un $\alpha=0.05$ de significancia, ¿Existen indicios de una diferencia significativa entre los dos métodos?
- b. Encuentre el valor-p e interprete su significado.

Valor del punto: 25%

ANEXOS:

Estadístico de la diferencia de proporciones:

$$Z \cong \frac{(p_{s_1} - p_{s_2}) - (p_1 - p_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}, \text{ donde: } p_{s_i} = \frac{X_i}{n_i} \quad \text{y} \quad \bar{p} = \frac{\sum X_i}{\sum n_i}$$

Estadístico X^2 para la prueba de independencia:

$$X^2 = \sum_{\text{todas las celdas}} \frac{(f_o - f_e)^2}{f_e}, \text{ que se distribuye } X^2 \text{ con } (r-1)*(c-1) \text{ grados de libertad}$$

donde r y c son el número de filas y columnas respectivamente que tiene la tabla de contingencia.

$$X^2_{0.05;4 \text{ gl}} = 9,488$$

$$X^2_{0.05;6 \text{ gl}} = 12,592$$

$$X^2_{0.025;4 \text{ gl}} = 11,143$$

$$X^2_{0.025;6 \text{ gl}} = 14,449$$

Coefficiente de determinación:

$r^2 = \frac{SCR}{SCT}$, donde SCR es la suma de cuadrados de la regresión y SCT es la suma de cuadrados totales.

Estadístico t para la pendiente de regresión:

$t = \frac{b_1}{S_{b_1}}$, donde S_{b_1} es el error típico del coeficiente, b_1 es el estimador puntual del coeficiente, y t se distribuye t -student con $n-2$ grados de libertad.

$$t_{13; 0.025} = \pm 2,1604$$

$$t_{13; 0.05} = \pm 1,7709$$

$$t_{14; 0.05} = \pm 1,7613$$

$$t_{14; 0.025} = \pm 2,1448$$