

## 5.4 PRUEBAS CHI-CUADRADO

### CONTENIDOS:

- 5.4.1. Prueba de bondad de ajuste.
- 5.4.2 Prueba de independencia.
- 5.3.3 Prueba de homogeneidad.

### OBJETIVOS:

- Plantear hipótesis para diferentes propósitos.
- Determinar los pasos a seguir al realizar una prueba chi-cuadrado.
- Interpretar el nivel de significación de la prueba de hipótesis.
- Redactar una conclusión con los resultados obtenidos de la prueba de hipótesis realizada.
- Realizar pruebas chi-cuadrado en problemas prácticos

### 5.4.1 PRUEBA DE BONDAD DE AJUSTE.

**CONCEPTOS CLAVES:** Parámetro. Estimador. Hipótesis. Frecuencia esperada. Estadístico de prueba. Nivel de significación. Región de rechazo. Conclusión.

#### RESUMEN DE CONCEPTOS Y PROPIEDADES:

Sea  $X \sim \mathfrak{I}_0(\theta)$ ,  $\mathfrak{I}_0$  es una distribución teórica conocida que depende de un parámetro  $\theta$  y se tiene una muestra aleatoria de tamaño  $n$  de  $X$  agrupada en  $m$  categorías  $A_1, A_2, \dots, A_m$  con frecuencias observadas  $n_1, n_2, \dots, n_m$

Pasos a seguir al realizar la prueba de hipótesis:

P1: Plantear hipótesis.

Hipótesis nula  $H_0$ : Los datos se ajustan a la distribución teórica  $\mathfrak{I}_0$

v/s Hipótesis alternativa  $H_A$ : Los datos no se ajustan a la distribución teórica

P2: Estadístico de prueba:  $J_0 = \sum_{i=1}^m \frac{(n_i - e_i)^2}{e_i} \sim \chi^2(m - k - 1)$

Donde  $n_i$ : Frecuencia observada de la categoría  $A_i$

$e_i = nP(A_i)$ : Frecuencia esperada de la categoría  $A_i$

$k$ : número de parámetros estimados en la distribución teórica.

$m$ : número de categorías en que se agrupan los datos.

P3: Establecer un nivel de significación:  $\alpha = P(\text{Rechazar } H_0 / H_0 \text{ es verdadero})$

P4: Región de rechazo de  $H_0$

Para  $H_0$  v/s  $H_A \Rightarrow R = \{x / x > \chi^2_{(1-\alpha, m-k-1)}\}$

P5: Decisión: Si  $J_0 \in R \Rightarrow$  se rechaza  $H_0$  al nivel de significación  $\alpha$

P6: Conclusión: Se debe interpretar la decisión tomada en Paso 5.

## EJERCICIO RESUELTO, PASO A PASO:

### Ejercicio 1: (Aplicación en Ciencias de la salud)

El número de alumnos por semana que sufren algún tipo de accidente en un colegio durante 36 semanas del periodo escolar es la siguiente:

|                                          |   |   |    |   |         |
|------------------------------------------|---|---|----|---|---------|
| Nº alumnos accidentados (X)              | 0 | 1 | 2  | 3 | 4 o más |
| Nº de semanas con X accidentes ( $n_i$ ) | 6 | 8 | 10 | 6 | 6       |

Probar si la muestra de datos se ajusta a una distribución de Poisson con intensidad  $\lambda$ , con un nivel de significación de 5%

### Esquema de solución

#### Paso 1: Leer cuidadosamente el enunciado del problema.

#### Paso 2: Identificar la variable en estudio y los parámetros involucrados.

Sea  $X$  = Número de alumnos accidentados. En este caso se debe suponer que  $X \sim \mathcal{P}(\lambda)$ ; es decir,  $P(X = j) = \frac{\lambda^j e^{-\lambda}}{j!}$  y el parámetro involucrado es la intensidad  $\lambda$ , donde  $\lambda$  es el número promedio de alumnos accidentados por semana en la población.

**Paso 3: Estimar los parámetros.** En este caso se tiene que el estimador de la intensidad es la media muestral, luego de la tabla de frecuencias obtenemos que

$$\hat{\lambda} = \bar{X} = \frac{\sum_{i=1}^5 x_i n_i}{36} = \frac{0*6 + 1*8 + 2*10 + 3*6 + 4*6}{36} = \frac{70}{36} = 1.94$$

#### Paso 4: Leer la pregunta 1 y revisar cual de los conceptos se debe usar para obtener lo pedido.

Para responder la pregunta se debe realizar una prueba de bondad de ajuste donde las hipótesis deben ser:  $H_0$ : Los datos se ajustan a la distribución de Poisson

v/s  $H_A$ : Los datos no se ajustan a la distribución de Poisson.

#### Paso 5: Realizar la prueba siguiendo los seis pasos.

**P1:** Plantear hipótesis.

Hipótesis nula  $H_0$ : Los datos se ajustan a la distribución de Poisson

v/s Hipótesis alternativa  $H_A$ : Los datos no se ajustan a la distribución de Poisson

**P2:** Estadístico de prueba:  $J_0 = \sum_{i=1}^5 \frac{(n_i - e_i)^2}{e_i} \sim \chi^2(5-1-1)$

Donde  $e_i = P(A_i) * 36$  y  $A_i = (X = i-1)$  para  $i = 1, 2, 3, 4, 5$

Luego

$$P(A_1) = P(X = 0) = \frac{(1.94)^0 e^{-1.94}}{0!} = 0.1437 \Rightarrow e_1 = 36 * 0.1437 = 5.1732$$

$$P(A_2) = P(X = 1) = \frac{(1.94)^1 e^{-1.94}}{1!} = 0.2788 \Rightarrow e_2 = 36 * 0.2788 = 10.0368$$

$$P(A_3) = P(X = 2) = \frac{(1.94)^2 e^{-1.94}}{2!} = 0.2704 \Rightarrow e_3 = 36 * 0.2704 = 9.7344$$

$$P(A_4) = P(X = 3) = \frac{(1.94)^3 e^{-1.94}}{3!} = 0.1749 \Rightarrow e_4 = 36 * 0.1749 = 6.2964$$

$$P(A_5) = P(X \geq 4) = 1 - P(X < 4) = 0.1322 \Rightarrow e_5 = 36 * 0.1322 = 4.7592$$

Luego calculamos el valor del estadístico de prueba en la tabla siguiente:

| Valores de X | $n_i$ | $e_i$                                                   | $\frac{(n_i - e_i)^2}{e_i}$ |
|--------------|-------|---------------------------------------------------------|-----------------------------|
| 0            | 6     | 5.1732                                                  | 0.1321                      |
| 1            | 8     | 10.0368                                                 | 0.4133                      |
| 2            | 10    | 9.7344                                                  | 0.00725                     |
| 3            | 6     | 6.2964                                                  | 0.01395                     |
| 4 o más      | 6     | 4.7592                                                  | 0.3235                      |
|              |       | $J_0 = \sum_{i=1}^5 \frac{(n_i - e_i)^2}{e_i} = 0.8901$ |                             |

**P3:** Nivel significación:  $\alpha = 0.05$

**P4:** Región de rechazo de  $H_0$  v/s  $H_A$

$$R = \left\{ x / x > \chi^2_{(1-\alpha, m-k-1)} \right\} = \left\{ x / x > \chi^2_{(0.95, 3)} \right\} = \left\{ x / x > 7.8 \right\}$$

**P5:** Decisión. Como  $J_0 = 0.8901 < 7.81 \Rightarrow J_0 \notin R \Rightarrow$  No se rechaza  $H_0$  al nivel de significación 0.05

**P6:** Conclusión. Con 95% de confianza la muestra de datos correspondiente al número de accidentes escolares por semana se ajusta a la distribución de Poisson

### EJERCICIOS PROPUESTOS:

1. (**Aplicación en Ciencias de la Salud**) En una ciudad la distribución habitual del grupo sanguíneo es de un 35%, 10%, 6% y un 49% para los grupos A, B, AB y O respectivamente. En Antofagasta, se realizó el estudio en una muestra de 200 individuos obteniéndose una distribución de 100, 60, 36, y 20 individuos para los grupos A, B, AB y O respectivamente. Verifique si la muestra de datos de la ciudad de Antofagasta se ajusta a la distribución habitual del grupo sanguíneo en dicha ciudad usando un nivel de significación de 0.01.
2. (**Aplicación en Ciencias de la Salud**) Con el fin de conocer si un cierto tipo de bacterias se distribuyen al azar en un determinado cultivo o si, por el contrario, lo hacen con algún tipo de preferencia (el centro, los extremos, etc...), se divide un cultivo en 576 áreas iguales y se cuenta el número de bacterias en cada área. Los resultados son los siguientes:

|                    |     |     |    |    |          |
|--------------------|-----|-----|----|----|----------|
| $n^o$ de bacterias | 0   | 1   | 2  | 3  | $\geq 4$ |
| $n^o$ de áreas     | 229 | 211 | 93 | 35 | 8        |

¿Se ajustan los datos a una distribución de Poisson de intensidad  $\lambda$ ? Use un nivel de significación de 0.05

3. (**Aplicación en Ciencias de la Salud**) La siguiente tabla recoge la distribución de los triglicéridos en suero, expresados en mg/dl en 90 niños de 6 años:

| Nivel de triglicéridos | Frecuencias |
|------------------------|-------------|
| Menos de 20            | 5           |
| 20 - 30                | 11          |
| 30 - 40                | 15          |
| 40 - 50                | 24          |
| 50 - 60                | 18          |
| 60 - 70                | 12          |
| 70 o más               | 5           |

Contrastar la hipótesis de que los datos se ajustan a una distribución Normal de media  $\mu$  y varianza  $\sigma^2$  con un nivel de significación de 0.05

## 5.4.2 PRUEBA DE INDEPENDENCIA.

**CONCEPTOS CLAVES:** Muestra aleatoria. Estadístico de prueba. Nivel de significación. Región de rechazo. Conclusión. Coeficiente de contingencia.

### RESUMEN DE CONCEPTOS Y PROPIEDADES:

Supongamos que se tiene una muestra de  $n$  datos bidimensionales de las variables  $X$  e  $Y$  y que se clasifican en  $m$  categorías  $A_1, A_2, \dots, A_m$  para  $X$  y  $k$  categorías  $B_1, B_2, \dots, B_k$  para  $Y$ , en la siguiente tabla de frecuencias conjunta:

| $X \setminus Y$ | $B_1$           | $B_2$           |  | $B_k$           | Total          |
|-----------------|-----------------|-----------------|--|-----------------|----------------|
| $A_1$           | $n_{11}$        | $n_{12}$        |  | $n_{1k}$        | $n_{1\bullet}$ |
| $A_2$           | $n_{21}$        | $n_{22}$        |  | $n_{2k}$        | $n_{2\bullet}$ |
|                 |                 |                 |  |                 |                |
|                 |                 |                 |  |                 |                |
| $A_m$           | $n_{m1}$        | $n_{m2}$        |  | $n_{mk}$        | $n_{m\bullet}$ |
| Total           | $n_{\bullet 1}$ | $n_{\bullet 2}$ |  | $n_{\bullet k}$ | $n$            |

P1: Plantear hipótesis.

Hipótesis nula  $H_0$ : Las variables  $X$  e  $Y$  son independientes

v/s Hipótesis alternativa  $H_A$ : Existe alguna relación entre  $X$  e  $Y$

La hipótesis  $H_0$  es equivalente a probar que  $P(A_i \cap B_j) = P(A_i)P(B_j)$  es decir que los sucesos  $A_i$  y  $B_j$  son independientes  $\forall i$  y  $\forall j$

P2: Estadístico de prueba:  $J_0 = \sum_{i=1}^m \sum_{j=1}^k \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 [(m-1)(k-1)]$

Donde  $m$ : número de categorías para la variable  $X$ .

$k$ : número de categorías para la variable  $Y$

$n_{ij}$ : Frecuencia observada de la categoría  $A_i \cap B_j$

$e_{ij} = nP(A_i \cap B_j)$ : Frecuencia esperada de la categoría  $A_i \cap B_j$

Para obtener las frecuencias esperadas se supone que  $H_0$  es verdadera, es decir

$$e_{ij} = nP(A_i \cap B_j) = nP(A_i)P(B_j) = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}.$$

P3: Establecer un nivel de significación:  $\alpha = P(\text{Rechazar } H_0 / H_0 \text{ es verdadero})$

P4: Región de rechazo de  $H_0$

$$\text{Para } H_0 \text{ v/s } H_A \Rightarrow R = \left\{ x / x > \chi^2_{(1-\alpha, (m-1)(k-1))} \right\}$$

P5: Decisión: Si  $J_0 \in R \Rightarrow$  se rechaza  $H_0$  al nivel de significación  $\alpha$

P6: Conclusión: Se debe interpretar la decisión tomada en Paso 5.

**Grado de relación:** Para medir el grado de relación entre las variables cualitativas se usa como indicador el coeficiente de contingencia (CC), definido por:

$$CC = \left( \sqrt{\frac{J_0}{J_0 + n}} \right) * 100$$

## EJERCICIO RESUELTO, PASO A PASO:

### Ejercicio 1: (Aplicación en Ciencias de la ingeniería)

Una empresa minera hizo un estudio para verificar si el tipo de trabajo se relaciona con el grado de silicosis de los trabajadores. Para lo cual se elige una muestra aleatoria de 300 trabajadores y se clasifican en la tabla siguiente:

| Tipo de Trabajo | Grado Silicosis |     |     | Total |
|-----------------|-----------------|-----|-----|-------|
|                 | I               | II  | III |       |
| Oficina         | 42              | 24  | 30  | 96    |
| Terreno         | 54              | 78  | 72  | 204   |
| Total           | 96              | 102 | 102 | 300   |

- Probar la hipótesis de que el tipo de trabajo afecta el grado de silicosis del trabajador con un nivel de significación de 5%
- Determine el grado de relación

### Esquema de solución

**Paso 1: Leer cuidadosamente el enunciado del problema.**

**Paso 2: Identificar las variables en estudio.**

Sea  $X$ =Tipo de trabajo     $Y$ =Grado de silicosis

**Paso 4: Leer la pregunta y revisar cual de los conceptos se debe usar para obtener lo pedido.**

Para responder la pregunta se debe realizar una prueba de hipótesis de independencia para variables cualitativas

**Paso 5: Realizar la prueba siguiendo los seis pasos.**

**P1:** Plantear hipótesis.

$H_0$  : El grado de silicosis es independiente del tipo de trabajo

v/s  $H_A$  : Existe alguna relación entre grado de silicosis y tipo de trabajo

**P2:** Estadístico de prueba;

| $n_{ij}$ | $e_{ij}$ | $\frac{(n_{ij} - e_{ij})^2}{e_{ij}}$ |
|----------|----------|--------------------------------------|
| 42       | 30,72    | 4,1419                               |
| 24       | 32,64    | 2,2871                               |
| 30       | 32,64    | 0,2135                               |
| 54       | 65,28    | 1,9491                               |
| 78       | 69,36    | 1,0763                               |
| 72       | 69,36    | 0,1005                               |
|          |          |                                      |
|          | $J_0 =$  | 9,7683                               |

**P3:** Nivel significación;  $\alpha = 0.05$

**P4:** Región de rechazo de  $H_0$  v/s  $H_A$

$$R = \left\{ x / x > \chi^2_{(1-\alpha, (m-1)(k-1))} \right\} = \left\{ x / x > \chi^2_{(0.95, 2)} \right\} = \left\{ x / x > 5.99 \right\}$$

**P5:** Decisión. Como  $J_0 = 9.7683 > 5.99 \Rightarrow J_0 \in R \Rightarrow$  Se rechaza  $H_0$  al nivel de significación 0.05

**P6:** Conclusión. Con 95% de confianza Existe alguna relación entre grado de silicosis y tipo de trabajo

## EJERCICIOS PROPUESTOS:

### 1. (Aplicación en Ciencias de la Salud)

Una institución ambiental hizo un estudio para determinar si el sector habitacional se relaciona con el grado de contaminación por plomo. Una muestra aleatoria de 300 personas a las que se les examinó entregó los siguientes resultados:

| Sector habitacional | Grado de Contaminación |       |      |
|---------------------|------------------------|-------|------|
|                     | Alto                   | Medio | Bajo |
| Gran Vía            | 18                     | 6     | 12   |
| Coviefi             | 42                     | 24    | 30   |
| Playa Blanca        | 36                     | 72    | 60   |

- ¿Cree usted que el sector habitacional se relaciona con el grado de contaminación por plomo? Use una confiabilidad de 99% en su conclusión.
  - Determine el grado de relación entre las variables analizadas e interprete el resultado.
2. (Aplicación en Ciencias del Mar) Se desea comparar la efectividad de dos análisis de laboratorio para detectar la presencia de bacterias en equinodermos (erizo rojo), para ello se selecciona dos muestras independientes de *Loxechinus albus* (erizo rojo comestible) en un sector de caleta coloso y los resultados de los análisis fueron los siguientes:

| Análisis | Detección de la Bacteria |    | Total |
|----------|--------------------------|----|-------|
|          | Si                       | No |       |
| 1        | 11                       | 39 | 50    |
| 2        | 8                        | 42 | 50    |
| Total    | 19                       | 81 | 100   |

- Pruebe la hipótesis de que la detección de la bacteria es independiente del tipo de análisis usado, con un nivel de significación de 0.05
  - Determine el grado de relación entre las variables
3. (Aplicación en Ciencias del Mar) La siguiente tabla muestra la distribución de una muestra aleatoria de 400 truchas cafés de un gran río., según la longitud y el sector donde fueron extraídas.

#### SECTOR DEL RIO

| LONGITUD          | Alto | Centro | Bajo |
|-------------------|------|--------|------|
| Bajo el promedio  | 67   | 64     | 25   |
| Promedio          | 42   | 76     | 56   |
| Sobre el promedio | 10   | 23     | 37   |

- Pruebe la hipótesis de que existe alguna relación entre la longitud de las truchas y el sector del río donde fueron extraídas, usando un nivel de significación de 0.05.
- ¿Cuál es el grado de relación entre las variables?.

### 5.4.3 PRUEBA DE HOMOGENEIDAD.

**CONCEPTOS CLAVES:** Hipótesis nula. Estadístico de prueba. Frecuencia esperada. Nivel de significación. Región de rechazo. Conclusión.

#### RESUMEN DE CONCEPTOS Y PROPIEDADES:

Consideremos  $k$  poblaciones independientes, cada una particionada en las clases  $A_1, A_2, \dots, A_m$ .

Para cada  $A_i$  se definen las probabilidades

$$p_{ij} = P(A_i / \text{Población } j) \quad i = 1, 2, \dots, m \quad y \quad j = 1, 2, \dots, k$$

El objetivo es probar si cada clase  $A_i$  tiene la misma probabilidad en todas las poblaciones. Para realizar la prueba se toma una muestra de tamaño  $n_j$  de la población  $j$  y se clasifican según las clases  $A_i$  en la siguiente tabla

| Clases   | Población |          |  |          | Total          |
|----------|-----------|----------|--|----------|----------------|
|          | 1         | 2        |  | k        |                |
| $A_1$    | $n_{11}$  | $n_{12}$ |  | $n_{1k}$ | $n_{1\bullet}$ |
| $A_2$    | $n_{21}$  | $n_{22}$ |  | $n_{2k}$ | $n_{2\bullet}$ |
| $\vdots$ |           |          |  |          |                |
| $A_m$    | $n_{m1}$  | $n_{m2}$ |  | $n_{mk}$ | $n_{m\bullet}$ |
| Total    | $n_1$     | $n_2$    |  | $n_k$    | $n$            |

P1: Plantear hipótesis.

$$\text{Hipótesis nula } H_0 : \begin{pmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{m1} \end{pmatrix} = \begin{pmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{m2} \end{pmatrix} = \dots = \begin{pmatrix} p_{1k} \\ p_{2k} \\ \vdots \\ p_{mk} \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix}$$

v/s Hipótesis alternativa  $H_A$  : Alguna diferencia existe

$$\text{P2: Estadístico de prueba: } J_0 = \sum_{i=1}^m \sum_{j=1}^k \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 [(m-1)(k-1)]$$

Donde  $m$  : número de clases.

$k$  : número de poblaciones

$n_{ij}$  : Frecuencia observada de la clase  $A_i$  en la población  $j$

$e_{ij}$  : Frecuencia esperada de la clase  $A_i$  en la población  $j$

Para obtener las frecuencias esperadas se supone que  $H_0$  es verdadera, es decir basta

estimar  $p_i$  por  $\hat{p}_i = \frac{n_{i\bullet}}{n}$  luego  $e_{ij} = n_j \hat{p}_i$ .

P3: Establecer un nivel de significación:  $\alpha = P(\text{Rechazar } H_0 / H_0 \text{ es verdadero})$

P4: Región de rechazo de  $H_0$

$$\text{Para } H_0 \text{ v/s } H_A \Rightarrow R = \left\{ x / x > \chi^2_{(1-\alpha, (m-1)(k-1))} \right\}$$

P5: Decisión: Si  $J_0 \in R \Rightarrow$  se rechaza  $H_0$  al nivel de significación  $\alpha$

P6: Conclusión: Se debe interpretar la decisión tomada en Paso 5.

## EJERCICIO RESUELTO, PASO A PASO:

### EJERCICIO : (APLICACIÓN EN CIENCIAS DE LA INGENIERÍA)

En un proceso de fabricación de tornillos, el fabricante quería determinar si la proporción de tornillos defectuosos producidos por tres máquinas variaba de una máquina a otra. Para verificar esto se seleccionaron muestras de 400 tornillos de la producción de cada máquina y se contó el número de tornillos defectuosos en cada una, obteniendo la siguiente tabla de frecuencias.

| Calidad        | Máquina |     |     | Total |
|----------------|---------|-----|-----|-------|
|                | 1       | 2   | 3   |       |
| Defectuosos    | 16      | 24  | 9   | 49    |
| No Defectuosos | 384     | 376 | 391 | 1151  |
| Total          | 400     | 400 | 400 | 1200  |

Pregunta: Realizando la prueba de hipótesis adecuada, verifique si la proporción de tornillos defectuosos no varía entre las diferentes máquinas. Use un nivel de significación de 0.05.

#### Esquema de solución

**Paso 1: Leer cuidadosamente el enunciado del problema.**

**Paso 2: Identificar las clases y las poblaciones.**

En este caso tenemos tres poblaciones (las máquinas) particionadas en dos clases (defectuosos y no defectuosos).

**Paso 3: Leer la pregunta y revisar cual de los conceptos se debe usar para obtener lo pedido.**

Para responder la pregunta se debe realizar una prueba de homogeneidad  
Donde  $p_{1j} = P(\text{Defectuoso} / \text{Máquina } j)$  y  $p_{2j} = P(\text{No Defectuoso} / \text{Máquina } j)$

**Paso 4: Realizar la prueba siguiendo los seis pasos.**

**P1:** Plantear hipótesis.

$$H_0 : \begin{pmatrix} p_{11} \\ p_{21} \end{pmatrix} = \begin{pmatrix} p_{12} \\ p_{22} \end{pmatrix} = \begin{pmatrix} p_{13} \\ p_{23} \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \text{ v/s } H_A : \text{Alguna diferencia existe}$$

**P2:** Estadístico de prueba;

| $n_{ij}$ | $e_{ij}$                        | $\frac{(n_{ij} - e_{ij})^2}{e_{ij}}$ |
|----------|---------------------------------|--------------------------------------|
| 16       | $400 \frac{49}{1200} = 16.3$    | 4,1419                               |
| 384      | $400 \frac{1151}{1200} = 383.7$ | 2,2871                               |
| 24       | 16.3                            | 0,2135                               |
| 376      | 383.7                           | 1,9491                               |
| 9        | 16.3                            | 1,0763                               |
| 391      | 383.7                           | 0,1005                               |
|          | $J_0 =$                         | 9,7683                               |

**P3:** Nivel significación;  $\alpha = 0.05$

**P4:** Región de rechazo de  $H_0$  v/s  $H_A$

$$R = \left\{ x / x > \chi^2_{(1-\alpha, (m-1)(k-1))} \right\} = \left\{ x / x > \chi^2_{(0.95, 2)} \right\} = \left\{ x / x > 5.99 \right\}$$

**P5:** Decisión. Como  $J_0 = 9.7683 > 5.99 \Rightarrow J_0 \in R_1 \Rightarrow$  Se rechaza  $H_0$  al nivel de significación 0.05

**P6:** Conclusión. Con 95% de confianza alguna diferencia existe en la proporción

### EJERCICIOS PROPUESTOS:

1. (**Aplicación en Ciencias de Ingeniería**) Se desea probar si la distribución proporcional del rendimiento en una gran compañía variaba según el coeficiente intelectual de los trabajadores, para lo cual se toma una muestra para cada nivel del CI y se clasifica según el rendimiento en la compañía en la siguiente tabla.

**R E N D I M I E N T O**

| COEFICIENTE INTELECTUAL | Deficiente | Regular | Bueno |
|-------------------------|------------|---------|-------|
| Bajo el promedio        | 67         | 64      | 25    |
| Promedio                | 42         | 76      | 56    |
| Sobre el Promedio       | 10         | 23      | 37    |

Verifique la hipótesis que la distribución proporcional del rendimiento en la compañía es el mismo para cada nivel de CI. Use un nivel de significación de 5%

2. (**Aplicación en Ciencias de Ingeniería**) Con el propósito de decidir si la opinión sobre el nivel de las carreras de ingeniería, que tienen los académicos y los alumnos, son concordantes, se procede a encuestar a 100 académicos y a 200 alumnos, obteniéndose los siguientes resultados:

|           | Deficiente | Regular | Bueno | Excelente |
|-----------|------------|---------|-------|-----------|
| Académico | 36         | 55      | 4     | 5         |
| Alumnos   | 80         | 93      | 14    | 13        |

Pruebe la hipótesis que la distribución proporcional de la opinión entre los académicos y los alumnos es la misma, usando una significación de 0.05

3. (**Aplicación en Ciencias de la salud**) El Servicio nacional de salud desea verificar si la la distribución proporcional del estado nutricional de los niños no varía en tres ciudades de la región, para lo cual toma una muestra de niños de las tres ciudades y los clasifica según estado nutricional obteniendo la siguiente tabla:

|                    |           | Ciudad      |        |           |       |
|--------------------|-----------|-------------|--------|-----------|-------|
|                    |           | Antofagasta | Calama | Tocopilla | Total |
| Estado nutricional | Obeso     | 82          | 70     | 62        | 214   |
|                    | Sobrepeso | 93          | 62     | 67        | 222   |
|                    | Normal    | 25          | 18     | 21        | 64    |
|                    | Flaco     | 16          | 15     | 18        | 49    |
|                    | Total     | 216         | 165    | 168       | 549   |

Realice el análisis estadístico correspondiente destacando las hipótesis que usted deberá probar con una significación de 0.05 y los pasos que ha seguido para lograr su conclusión.