|  |  |  |
| --- | --- | --- |
| A picture containing circle  Description automatically generated | **Universidad Tecnológica de Panamá**  **Facultad de Ingeniería de Sistemas Computacionales** | Logo  Description automatically generated |

**Carrera:** Ing. en Sistemas de Información

**Cursado:** Simulación de Sistemas

**Profesor:** Modaldo Tuñón, PhD.

**L10 – Prueba de Chi-Cuadrado en Excel**

**Estudiante**

Johel Heraclio Batista Cárdenas {8-914-587}

**Grupo**: 1IF-131

**Fecha de Entrega**: 11 de Octubre de 2022

**Prueba de Bondad Ajuste Chi Cuadrado**

La prueba de Bondad de Ajuste Chi Cuadrado es la prueba de bondad de ajuste más utilizado. En general las pruebas de bondad de ajuste se utilizan para discriminar si un grupo de datos o muestra se ajusta a una distribución teórica.

La prueba de chi cuadrado puede utilizarse para trabajar tanto con distribuciones discretas (Ejemplo, distribución de Poisson), como distribuciones continuas (Ejemplo, Distribución Normal y Exponencial). Esta es una ventaja sobre las pruebas Kolmogórov-Smirnov y Anderson-Darling las cuales solo se pueden utilizar para distribuciones continuas.

La prueba de bondad de ajuste de chi cuadrado requiere que los datos estén agrupados en categorías o clases por lo que se deben generar u obtener tablas de frecuencias o histogramas.

Se hace importante mencionar que hay dos pruebas de chi cuadrado las cuales son diferentes y llegan a ser confundidas frecuentemente:

* **Prueba de bondad de ajuste chi cuadrado:** ajusta una variable categórica a una distribución teórica.
* **Prueba chi cuadrado para independencia:** compara dos conjuntos de datos para ver si existe relación entre ellas.

**Estadístico de Chi Cuadrado**

La fórmula de cálculo del estadístico de chi cuadrado utilizado en la prueba de bondad corresponde a:

Donde:

**C: Grados de Libertad de la Información (Datos)**

**Oi: Valor Observado para la i-ésima clase.**

**Ei: Valor Esperado para la i-ésima clase.**

**Hipótesis Nula y Alternativa**

En el caso de la prueba de bondad de ajuste de chi cuadrado las hipótesis nula y alternativa son:

* *H0:* Los datos se ajustan a una distribución dada.
* *Ha:* Los datos no se ajustan a una distribución dada.

Donde la distribución dada corresponde a una distribución teórica específica a la cual queremos ajustar nuestros datos categorizados.

Si por ejemplo decimos que nuestra hipótesis nula es que los datos se ajustan a una distribución normal. Dicha hipótesis será evaluada para un nivel de significancia (α) definido por el evaluador (comúnmente entre 1%, 5% y 10%).

Adicionalmente esta prueba nos permite obtener un valor p el cual dependerá de los grados de libertad y el nivel de significancia. Si el valor p es menor al nivel de significación se puede rechazar la hipótesis nula.

**Explicación con Ejemplo Práctico de la Prueba de Bondad de Ajuste Chi Cuadrado o x2**

**Consigna:** Un total de 256 artistas visuales fueron encuestados de modo de identificar su signo del zodiaco. Los resultados obtenidos fueron: Aries (29), Tauro (24), Géminis (22), Cáncer (19), Leo (21), Virgo (18), Libra (19), Escorpión (20), Sagitario (23), Capricornio (18), Acuario (20), Piscis (23). Se desea probar la hipótesis que los signos del zodiaco están distribuidos uniformemente entre los artistas visuales.

**Paso 1**: Preparar una tabla con las siguientes columnas:

Table

Description automatically generated with medium confidence

**Paso 2**: Llenamos las categorías junto con sus valores observados.

Table

Description automatically generated

**Paso 3**: Calcular el valor esperado de cada categoría. Como el enunciado dice que la distribución es uniforme cada categoría debería tener la misma distribución de 256/12.

Table, Excel

Description automatically generated

**Paso 4**: Calculamos la diferencia entre el valor observado y el valor esperado para cada categoría.

Table

Description automatically generated

**Paso 5**: Elevamos al cuadrado el resultado obtenido en el paso anterior (Oi - Ei).

Table

Description automatically generated

**Paso 6**: Finalmente para cada uno de los resultados del paso anterior lo dividimos por el valor esperado de cada categoría. Y luego sumamos todos estos resultados.

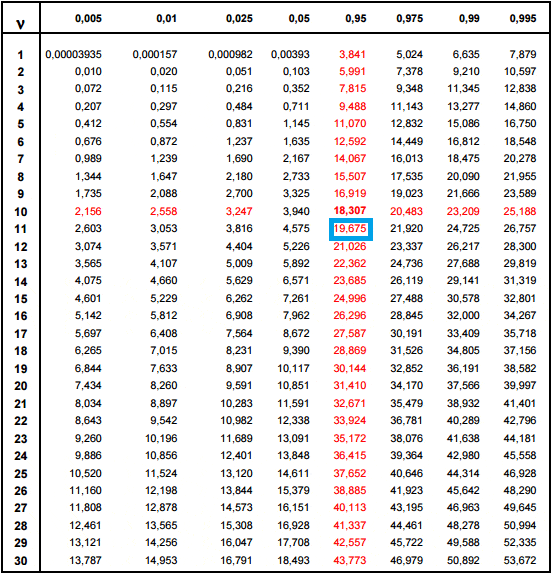
Table

Description automatically generated

Con esto ya tenemos la tabla para hacer nuestra prueba de bondad.

**Paso 7**: Para nuestro ejemplo queremos evaluar nuestra hipótesis nula (que los signos del zodiaco están distribuidos uniformemente entre los artistas visuales) para un ***nivel de significancia* de 5% (α = 0.05),** indicando un **nivel de confianza de 95%**. Además, tenemos que los ***grados de libertad* son 11** (cantidad de categorías menos 1).

Junto con la tabla de distribución chi cuadrado buscamos el valor para el nivel de confianza y grados de libertad.



Nuestro valor es 19.675 es mayor al valor obtenido en la tabla de chi cuadrado (5.09375) y en consecuencia no hay evidencia estadística que permita rechazar la hipótesis nula.

**Cálculo del valor p**

Esta es una alternativa al paso 7 donde utilizando Excel podemos obtener el valor p asociado a la prueba de bondad de ajuste chi cuadrado.

La fórmula que se utiliza es CHISQ.TEST([CONJUNTO\_Oi], [CONJUNTO\_Ei]), donde el primer set de valores son las distribuciones observadas y el segundo set las distribuciones esperadas.

Table

Description automatically generated

Como el valor p (0.9265) no es menor al nivel de significancia (0.05) no se rechaza la hipótesis nula.