XS-2110. Métodos Estadísticos.

**Clase. Pruebas de Bondad de Ajuste: X2, G2 y contraste de Kolmogorov-Smirnov: Bondad de ajuste y media para una población.**

**Pruebas de Bondad de Ajuste.**

El objetivo de las pruebas o contrastes de bondad de ajuste es inferir si una muestra proviene de una población con cierto tipo de parámetros. De esta forma, las hipótesis nula y alternativa serían:

H0: Muestra fue seleccionada de una población con distribución f(x | θ): X~f(x | θ)

H1: Muestra NO fue seleccionada de una población con distribución f(x | θ): X~f(x | θ)

Comúnmente también se dice que las pruebas de bondad de ajuste se utilizan para definir si una muestra tiene una cierta distribución (normal, Poisson, exponencial, gamma).

H0: X se distribuye con distribución f(x | θ): X~f(x | θ)

H1: X no se distribuye con distribución f(x | θ): X~f(x | θ)

Es interesante empezar con contrastes de hipótesis de bondad de ajuste, porque en muchas de sus aplicaciones, buscamos no rechazar H0.

Como dice Gutiérrez (p.36), estas pruebas se usan para:

1. analizar cumplimiento de supuestos de otras pruebas estadísticas.
2. analizar el ajuste de los datos a leyes probabilísticas (como las leyes mendelianas de la genética) o si un experimento estadístico a base de simulaciones genera la distribución esperada.

Para esta clase veremos 2 pruebas generales de bondad de ajuste la prueba X2 (conocida también como chi-cuadrado) de Pearson y la prueba de razón de verosimilitud para bondad de ajuste (G2). Estudiaremos también una prueba específica para contrastar bondad de ajuste a la distribución normal: Prueba de Shapiro.

Además, la bondad de ajuste también se puede estudiar a partir de gráficos. No se puede hacer una conclusión probabilística, pero sí realizar una inspección visual a partir de los gráficos. El histograma (que estudiaron en Estadística Introductoria I) y los gráficos QQ (QQ plots) son usados para esta inspección visual.

Contraste X2 de Pearson:

Requisitos o supuestos:

1. Muestra aleatoria (supone simple al azar, pero se puede estimar con otros diseños muestrales).
2. Muestra de mediana a grande: Se evalúa analizando si más del 75% de los valores esperados son mayores a 5.
3. Se puede analizar la bondad de ajuste para cualquier escala de medición: variables cualitativas y cuantitativas.
4. Para variables continuas, los datos muestrales tienen que estar agrupados en clases: Como las distribuciones de frecuencias que se estudian en Estadística Introductoria I.
5. Los grados de libertad de la prueba son equivalentes a k-p-1, donde:
   1. k = # de clases
   2. p = # de parámetros de la distribución de ajuste.
6. La prueba es de una sola cola.

La fórmula de X2 es:

La teoría para que este estadístico se distribuya como una χ2 es la siguiente:

1. Se supone que la variable de conteo (los fi) tienen una distribución Poisson con media igual a Fi;
2. Pero si los Fi son suficientemente grandes, una variable con distribución Poisson se aproxima a una distribución normal.
3. El término  es equivalente a: o sea, a una variable normal estandarizada.
4. Una suma de variables con distribución normal estándar se distribuye como una chi-cuadrado con grados de libertad iguales a la suma de términos no redundantes.

Prueba de razón de verosimilitud para bondad de ajuste.

Mientras que la prueba X2 se basa en inferencia clásica, la prueba de razón de verosimilitudes se basa en inferencia vía máxima verosimilitud. Este curso no entrará en detalles de lo que significa la verosimilitud, pues lo estudiarán más a profundidad en Modelos Continuos, Teoría Estadística y Diseño de Experimentos 2.

La fórmula del estadístico de prueba es:

Requisitos y supuestos.

1. El 75% de las Fi deberían ser mayores a 5.
2. Para usar G2 se debería de tener una muestra razonablemente grande (digamos, más que 100), más grande que la que se puede usar para X2 (digamos que 40 ó más). Sin embargo, en la práctica, después de n=50, las distribuciones muestrales de X2 y G2 son prácticamente iguales.
3. El estadístico G2 se distribuye también como una χ2 con grados de libertad equivalentes a k-p-1.
4. Si se analiza una variable continua, los datos tienen que estar agrupados en clases.

Cuando hagan los cálculos, notarán que los valores de X2 y G2 son muy parecidos, y serán más parecidos cuanto más grande sea la muestra.

Contraste de normalidad de Shapiro

La prueba de Shapiro (también conocida como prueba de Shapiro-Wilk) se usa para contrastar si una muestra particular proviene de una distribución normal. La fórmula del estadístico de prueba es:

donde ai es una serie de pesos que aparecen en el artículo original de Shapiro y Wilks. Estos pesos se calcularon a partir de los cuantilos de una distribución teórica estadística. El denominador de W es el numerador de la fórmula de variancia. Los xi\* tienen que estar ordenados de mayor a menor.

Supuestos y requisitos:

1. Muestras relativamente pequeñas (menor a 1000).
2. Es estrictamente para contrastar el ajuste a una distribución normal.
3. Se rechaza si el estadístico W se aleja de uno, según tabla proporcionada por el profesor.

Gráfico QQ-Plot.

En este gráfico, se grafica la frecuencia acumulada esperada contra la frecuencia acumulada observada de los datos ordenados. Si los datos se alejan demasiado de una recta de 45°, entonces se considera que el ajuste es malo. Las desviaciones más graves de la normalidad se dan en la mitad de la recta.

Contraste de Kolmogorov-Smirnov para una muestra

La prueba de Kolmogorov-Smirnov es un contraste no paramétrico frecuentemente usado para analizar bondad de ajuste. Al ser no paramétrica, es una prueba más conservadora que las otras opciones: X2 y G2. Esto quiere decir que al mismo nivel de α, es más difícil rechazar H0.

Aunque se utiliza como prueba de bondad de ajuste, también se puede usar como alternativa a la prueba z ó la prueba t para una media, dado que uno de los requisitos indispensables del contraste de Kolmogorov-Smirnov (ó KS) es:

**Que se tiene que conocer o plantear de antemano los valores de los parámetros que definen la distribución.**

Además, la prueba KS tiene la ventaja de que se puede usar con variables cuantitativas y también para comparación de distribuciones de variables cualitativas ordinales.

En cualquier caso, la prueba de KS compara la distribución acumulada observada de frecuencias ( Fo(Y) )contra la distribución esperada ( Fh(Y) ). Por eso, la hipótesis nula se puede plantear de varias formas:

H0: Fo(Y)=Fh(Y)

H1: Fo(Y)≠ Fh(Y).

ó bien:

H0: Muestra fue seleccionada de una población con distribución de probabilidad acumulada F(Y | θ)

H1: Muestra NO fue seleccionada de una población con distribución de probabilidad acumulada F(Y | θ)

ó bien:

H0: θ=θk

H1: θ≠θk

donde θk puede ser el valor del parámetro en la distribución de interés.

El procedimiento por utilizar es el siguiente:

1. Establecer la distribución relativa acumulada de frecuencias de la variable de interés.
2. Con los parámetros establecidos, calcular la distribución de probabilidad acumulada esperada según la hipótesis nula.
3. Se calcula: Δ=|F0(Yi)-Fh(Yi)| para cada una de las clases, y se determina Δmax.
4. Se calcula: Ω=|F0(Yi-1)-Fh(Yi)| para cada una de las clases, y se determina Ωmax.
5. Se busca Dmax que es igual a Dmax=max(Δmax,Ωmax).
6. El valor Dmax se compara con un Dc.(Tabla E, página 152)
7. Si Dmax es mayor que el Dc, entonces se rechaza H0.
8. Si la prueba fuera de una cola, entonces Δmax yΩmax se calculan solo con las diferencias positivas (si θ>θk) o solo con las negativas (si θ<θk), antes de pasarlas a valor absoluto.