UNIVERSIDAD NACIONAL DE COLOMBIA

Estadística I

Pruebas de Hipótesis para proporciones

Suponga que X es una v.a tal que $X \sim B(n, P)$, con P desconocida. Sea $n \geq 30$ y $p_0 \in (0,1)$ un valor de interés.

Paso 1: Hipótesis

1. $H_0 : p = p_0$ **vs** $H_1 : p > p_0$
2. $H_0 : p = p_0$ **vs** $H_1 : p < p_0$
3. $H_0 : p = p_0$ **vs** $H_1 : p \neq p_0$

Paso 2: Estadístico de prueba

$$Z_c = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim_{aprox} N(0,1) \text{ bajo } H_0 \text{ con } \hat{P} = \frac{X}{n}$$

Suponga que $X_1 \sim B(n_1, P_1)$ y $X_2 \sim B(n_2, P_2)$, con P_1 y P_2 desconocidas. Sean $n_1, n_2 \geq 30$.

Paso 1: Hipótesis

1. $H_0 : p_1 - p_2 = 0$ **vs** $H_1 : p_1 - p_2 > 0$
2. $H_0 : p_1 - p_2 = 0$ **vs** $H_1 : p_1 - p_2 < 0$
3. $H_0 : p_1 - p_2 = 0$ **vs** $H_1 : p_1 - p_2 \neq 0$

Paso 2: Estadístico de prueba

$$Z_c = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim_{aprox} N(0,1) \text{ bajo } H_0 \text{ donde } \hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

Paso 3: Regiones de rechazo

1. $R_c : \{Z_c \mid Z_c > Z_\alpha\}$
2. $R_c : \{Z_c \mid Z_c < -Z_\alpha\}$
3. $R_c : \{Z_c \mid Z_c < -Z_{\alpha/2} \vee Z_c > Z_{\alpha/2}\}$

Paso 4: Valor P

1. $VP = P(Z \geq Z_c)$
2. $VP = P(Z \leq Z_c)$
3. $VP = P(|Z| \geq |Z_c|)$

Pruebas de Hipótesis para proporciones

Ejemplo:

En una muestra aleatoria de 80 rodamientos de cigüeñales para automotores, 15 de los rodamientos tienen un terminado superficial cuya aspereza rebasa lo que permiten las especificaciones. A un nivel de significancia del 1%, ¿es posible apoyar la hipótesis de que más del 20% de los rodamientos de cigüeñales están fuera de las especificaciones?

Pruebas de Hipótesis para proporciones

Pruebas de Hipótesis para proporciones

Ejemplo:

Cierta empresa está realizando un nuevo shampoo sin sal ni parabenos el cual se espera que tenga un efecto positivo en más del 80% de sus consumidores para que pueda ser lanzado al mercado. Si 60 personas se ofrecen para usarlo y probar sus efectos ¿Cuál debe ser el número mínimo de personas que manifiesten efectos positivos en su cabello para que el shampoo sea lanzado? Use $\alpha = 0.05$.

Pruebas de Hipótesis para proporciones

Pruebas de Hipótesis para proporciones

UNIVERSIDAD NACIONAL DE COLOMBIA

Prueba de Bondad de Ajuste χ^2

Introducción

Una prueba de bondad de ajuste es un procedimiento de pruebas de hipótesis donde lo que se quiere es probar si un conjunto de datos provienen o no de una distribución conocida.

Una de las pruebas más usadas es la prueba da bondad de ajuste chi-cuadrado, debido a que el estadístico de prueba sigue esa distribución. Otras pruebas de bondad de ajuste muy utilizadas son: Kolmogorov-Smirnov, Shapiro-Wilk, Cramer Von Mises, Anderson Darling, entre otras.

UNIVERSIDAD NACIONAL DE COLOMBIA

Prueba de Bondad de Ajuste χ^2

Introducción

Se dispone de n observaciones independientes de una v.a X, las cuales se agrupan en C categorías o clases.

Estrato
1
2
3
4
5
6

Marca
A
B
C
D

Edades
0-10
11-20
21-30
31-40
41-50

Colores
Rojo
Verde
Azul
Rosado
Amarillo

Se agrupan especialmente cuando son continuas

Si es una variable "infinita" (de $(0,\infty)$) se definen las clases por ejemplo

Edades
0-10
11-20
21-30
>30

Y el número de observaciones en cada clase son presentadas en la siguiente tabla:

Clase	1	2	3	4	...	c
Frecuencia Observada	O_1	O_2	O_3	O_4		O_c

Prueba de Bondad de Ajuste χ^2

Introducción

En las pruebas de bondad de ajuste chi-cuadrado se trata de comparar los valores o frecuencias observadas con las frecuencias que habrían en cada grupo o clase (frecuencia esperada o valor esperado) si se cumple la hipótesis nula.

Las diferencias entre lo observado y lo esperado dan las discrepancias entre la teoría y la realidad. Si no hay diferencias, la realidad coincidiría perfectamente con la teoría y por el contrario, si las diferencias son grandes indica que la realidad y la teoría no se parecen.

Prueba de Bondad de Ajuste χ^2

Introducción

Como lo vimos anteriormente, se dispone de n observaciones independientes de una v.a X, las cuales se agrupan en C categorías o clases, O_i representa la cantidad de observaciones en la clase i (frecuencia observada) y sea p_i la probabilidad de que cada observación quede clasificada en la categoría i. Así la información se puede resumir como sigue:

Clase	Frec. Observada	Probabilidad	Frec. Esperada
1	O_1	p_1	$e_1 = np_1$
2	O_2	p_2	$e_2 = np_2$
3	O_3	p_3	$e_3 = np_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
c	O_c	p_c	$e_c = np_c$

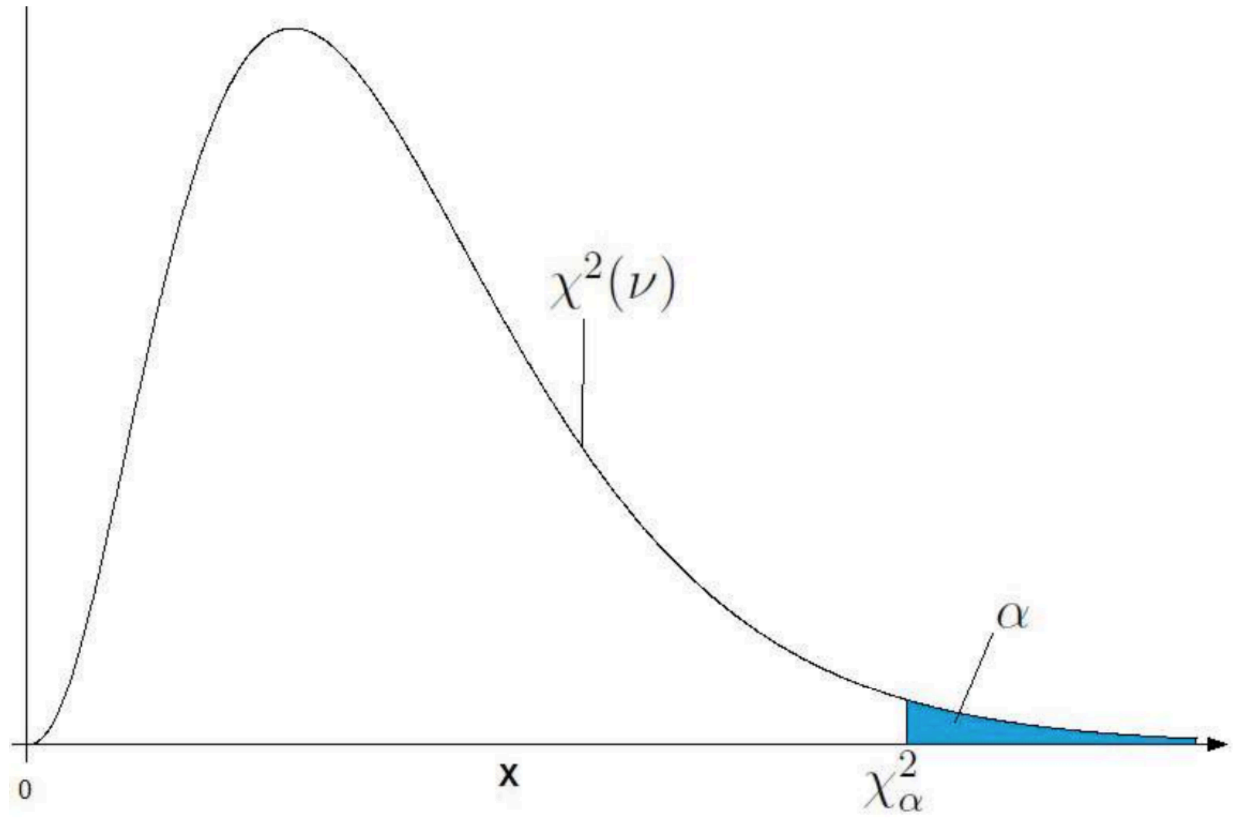
Prueba de Bondad de Ajuste χ^2

Paso 1: Hipótesis

- H_0 : **los datos provienen de una distrib. determinada vs** H_1 : **los datos no provienen de la distrib dada en** H_0
- $H_0 : F(x) = F_0(x)$ **vs** $H_1 : F_0(x)$ **no es la distrib asociada a la muestra.**
- $H_0 : p_i = p_{i0} \; i = 1,2,\ldots,c$ **vs** $H_1 : \exists j$ **tal que** $p_j \neq p_{j0}$

Paso 2: Estadístico de prueba

$$\chi_c = \sum_{i=1}^{c} \frac{(O_i - e_i)^2}{e_i} \sim \chi^2_{(c-k-1)}$$



Paso 3: Región de Rechazo

$$R_c = \{\chi_c \,|\, \chi_c > \chi^2_{\alpha,c-k-1}\}$$

Paso 4: Valor-P

$$VP = P\left(\chi^2_{c-k-1} > \chi_c\right)$$

Notas:

I. k es el número de parámetros estimados.

II. Debe cumplirse que $e_i = np_i \geq 5$. En caso de que no se cumpla en alguna categoría, éstas deben agruparse.

Clase	Frec.	Probabilidad	Frec.
1	O_1	p_1	$e_1 = np_1$
2	O_2	p_2	$e_2 = np_2$
3	O_3	p_3	$e_3 = np_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
c	O_c	p_c	$e_c = np_c$

Prueba de Bondad de Ajuste χ^2

Ejemplo:

El director de mercadeo de una empresa tiene la responsabilidad de controlar el nivel de existencias para cuatro tipos de botes vendidos por la empresa. En el pasado ha ordenado nuevos botes bajo la premisa de que los cuatro tipos son igualmente populares y la demanda por cada tipo es la misma. Recientemente las existencias han sido difíciles de controlar. El director quiere establecer si la premisa que ha venido manejando es coherente con la demanda actual. Para ello, toma como muestra las 58 ventas de los últimos meses. Los resultados obtenidos para cada tipo de bote (respecto a las ventas) son:

Tipo de bote	1	2	3	4
Frec. Ventas	18	12	15	13

¿Es coherente la premisa de mercadeo con lo observado en la muestra? Use $\alpha = 0.05$.

UNIVERSIDAD NACIONAL DE COLOMBIA

Prueba de Bondad de Ajuste χ^2

Prueba de Bondad de Ajuste χ^2

UNIVERSIDAD
NACIONAL
DE COLOMBIA

Prueba de Bondad de Ajuste χ^2

Prueba de Bondad de Ajuste χ^2

Ejemplo:

A cierta empresa de la ciudad de Medellín llegan todos los días una cantidad de camiones para descargar su materia prima. El encargado de la bodega registró el número de camiones por día durante 36 días y obtuvo los siguientes datos:

Nº de camiones (x)	0	1	2	3	≥ 4
Nº de días con x camiones	6	8	10	6	6

Pruebe si la muestra de datos se ajusta a una distribución Poisson. Use $\alpha = 0.02$.

Prueba de Bondad de Ajuste χ^2

Prueba de Bondad de Ajuste χ^2

Prueba de Bondad de Ajuste χ^2

Gracias

Clase preparada por: Verónica Guarín