Contraste de Bondad de Ajuste

Clase del 3/12/2024

Contenidos

Definición del problema

- **Objetivo**: Determinar si los datos observados provienen de una distribución teórica específica.
- **Distribución teórica**: Especificar la distribución de referencia (e.g., $\mathcal{N}(\mu, \sigma^2)$, $U(a, b)$, $B(n, p)$).
- **Hipótesis**:
  - H_0: Los datos siguen la distribución teórica.
  - H_1: Los datos no siguen la distribución teórica.

Elección del estadístico de prueba

- **Chi-cuadrado (χ^2)**: Comparar las frecuencias observadas y esperadas en diferentes categorías.
  - Usado cuando los datos son discretos o agrupados.
- **Kolmogorov-Smirnov (K-S)**: Evaluar la diferencia máxima entre la función de distribución empírica y la teórica.
  - Adecuado para datos continuos con una distribución específica.
- **Pruebas alternativas**: Anderson-Darling, Shapiro-Wilk, Jarque-Bera, etc.

Preparación de los datos

- **Frecuencias observadas** (O_i): Contar las observaciones en cada categoría (discretas o intervalos para datos continuos).
- **Frecuencias esperadas** (E_i):
  - Calcular las frecuencias bajo la hipótesis nula utilizando la distribución teórica.
  - **Nota**: $E_i > 5$ en cada categoría para aplicar χ^2 (reagrupar si es necesario).

Cálculo del estadístico de prueba

- **Para** χ^2:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

- **Para K-S**:

$$D = \text{máx} \left| F_{\text{empírica}}(x) - F_{\text{teórica}}(x) \right|$$

- Otros estadísticos dependen de la prueba elegida.

Determinación del valor crítico o *p*-valor

- **Nivel de significación** (α): Usualmente $\alpha = 0.05$ o $\alpha = 0.01$.
- Comparar el estadístico con el valor crítico de la distribución correspondiente.
  - χ^2: Tabla de distribución χ^2 con $k - m - 1$ grados de libertad (k: número de categorías, m: parámetros estimados).
  - K-S: Valores críticos de la tabla de K-S.

Conclusión del contraste

- **Rechazar** H_0: Si el estadístico calculado es mayor al valor crítico o el p-valor es menor que α.
- **No rechazar** H_0: Si el estadístico calculado es menor al valor crítico o el p-valor es mayor que α.

Sobre el lanzamiento de un dado

Lanzamos un dado 30 veces y obtenemos:

$$4, 5, 3, 5, 5, 2, 3, 3, 3, 5, 4, 3, 6, 5, 2, 4, 6, 6, 2, 4, 5, 1, 4, 2, 6, 5, 4, 1, 1, 3$$

Las frecuencias observadas (O_i) son: $O = [3, 4, 7, 6, 6, 4]$

- Cara 1: 3 veces.
- Cara 2: 4 veces.
- Cara 3: 7 veces.
- Cara 4: 6 veces.
- Cara 5: 6 veces.
- Cara 6: 4 veces.

Sobre el lanzamiento de un dado

Bajo la hipótesis nula (H_0), el dado es justo, por lo que cada cara tiene igual probabilidad de aparecer.
El estadístico de Chi-cuadrado se calcula con la fórmula:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

Sustituyendo los valores:

$$\chi^2 = \frac{(3-5)^2}{5} + \frac{(4-5)^2}{5} + \frac{(7-5)^2}{5} + \frac{(6-5)^2}{5} + \frac{(6-5)^2}{5} + \frac{(4-5)^2}{5} = 2.4$$

Sobre el lanzamiento de un dado

Con $k = 6$ categorías y sin estimación de parámetros, los grados de libertad son:

$$df = k - 1 = 5$$

Para $\chi^2 = 2.4$ y $df = 5$, $\chi^2_{0.05,5} \approx 11.07$,
$p - valor = P(\chi^2 > 2.4) = 0.791$
Por tanto, como no se cumple que $2.4 > 11.07$ o equivalentemente, como que $p = 0.791 > 0.05$,

Conclusión

NO hay evidencias para rechazar la hipótesis nula. Por lo tanto, no hay evidencia suficiente para concluir que la muestra proviene de un dado no uniforme.

Sobre una normal standard

Se aplicó el test de bondad de ajuste de Chi-cuadrado para verificar si una muestra de tamaño 30 proviene de una distribución normal estándar ($\mathcal{N}(0,1)$).

$$\begin{bmatrix} 0.4967 & -0.1383 & 0.6477 & 1.5230 & -0.2342 & -0.2341 \\ 1.5792 & 0.7674 & -0.4695 & 0.5426 & -0.4634 & -0.4657 \\ 0.2420 & -1.9133 & -1.7249 & -0.5623 & -1.0128 & 0.3142 \\ -0.9080 & -1.4123 & 1.4656 & -0.2258 & 0.0675 & -1.4247 \\ -0.5444 & 0.1109 & -1.1510 & 0.3757 & -0.6006 & -0.2917 \end{bmatrix}$$

Sobre una normal standard

Consideramos los siguientes intervalos:

$$\text{Intervalos}: (-\infty, -1), [-1, 0), [0, 1), [1, \infty)$$

Las frecuencias observadas (O_i) para cada intervalo son:

$$O = [6, 12, 9, 3]$$

Sobre una normal standard

Bajo la hipótesis nula (H_0), se calculan las probabilidades teóricas para cada intervalo usando la función de densidad acumulativa de una normal estándar:

$$\begin{aligned}
P(Z < -1) &= \Phi(-1) \approx 0.1587, \\
P(-1 \leq Z < 0) &= \Phi(0) - \Phi(-1) \approx 0.3413, \\
P(0 \leq Z < 1) &= \Phi(1) - \Phi(0) \approx 0.3413, \\
P(Z \geq 1) &= 1 - \Phi(1) \approx 0.1587.
\end{aligned}$$

Multiplicando estas probabilidades por el tamaño muestral ($n = 30$), las frecuencias esperadas (E_i) son:

$$E = [4.76, 10.24, 10.24, 4.76]$$

Sobre una normal standard

Con $k = 4$ categorías y sin estimación de parámetros, los grados de libertad son:

$$df = k - 1 = 3$$

Para $\chi^2 = 1.43$ y $df = 3$, $\chi^2_{0.05,3} \approx 7.81$, el $p - valor \approx 0.699$

Conclusión

Dado que $p = 0.699 > 0.05$, no se rechaza la hipótesis nula. Por lo tanto, no hay evidencia suficiente para concluir que la muestra no proviene de una distribución normal estándar.

Ajuste de una Poisson

Se han contado los coches que hay en un semáforo a los 30 segundos de ponerse en rojo, recogiendo los siguientes resultados.

$$6, 6, 4, 5, 4, 6, 4, 4, 9, 3, 5, 5, 6, 5, 10, 2, 5, 4, 11, 5, 4, 9, 4, 6, 8, 3, 4, 6$$

Se puede decir que sigue una distribución Poisson con $\lambda = 6$?

Ajuste de una Poisson

Se calculan las frecuencias observadas (O_i) y las esperadas (E_i) para cada valor posible, agrupando categorías con $E_i < 5$. Los resultados ajustados fueron:

Intervalo	Frecuencia Observada	Frecuencia Esperada
$[0, 4]$	11	10.16
$[5, 6]$	12	9.2
$[7, 8]$	1	5.77
$[9, \infty[$	4	2.87

Ajuste de una Poisson

El estadístico de Chi-cuadrado se calcula con la fórmula:

$$\chi^2_{3,\alpha} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(11 - 10.16)^2}{10.16} + \frac{(12 - 9.2)^2}{9.2} + \cdots$$

El estadístico Chi-cuadrado obtenido y el teórico son:

$$\chi^2_{3,0.05} \approx 5.2964 > (?)7.8147$$

Debemos concluir que a un nivel de significación del 5 % que NO hay evidencias para rechazar la hipótesis nula, o sea que no ha evidencia para rechazar que sea una Poisson con $\lambda = 6$.

* Observa que hay un intervalo donde el valor esperado no es superior a 5, por lo que realmente no reuniríamos los requisitos. Podríamos replantear el test en los siguientes términos.

Ajuste de una Poisson

Se calculan las frecuencias observadas (O_i) y las esperadas (E_i) para cada valor posible, agrupando categorías con $E_i < 5$. Los resultados ajustados fueron:

Intervalo	Frecuencia Observada	Frecuencia Esperada
$[0, 4]$	11	10.16
$[5, 6]$	12	9.2
$[7, 8]$	5	8.63

En esta ocasión, el estadístico Chi-cuadrado obtenido y el teórico son:

$$\chi^2 \approx 2.4509 > (?)5.9915$$

Concluyendo de igual forma que antes que a un nivel de significación del 5 % que NO hay evidencias para rechazar la hipótesis nula, o sea que no ha evidencia para rechazar que sea una Poisson con $\lambda = 6$.