Instituto Superior de Engenharia do Porto

Estatística

2018/2019

LEEC

C6 – Teste de ajuste do Qui-quadrado

Teste do Qui-quadrado

O teste do qui-quadrado para a avaliação da qualidade de ajuste baseia-se na comparação da distribuição dos dados amostrais com a distribuição teórica à qual se supõe pertencer uma amostra. Pretende assim testar a hipótese dos dados da amostra serem consistentes com o facto população ter uma dada distribuição.

A metodologia que se adota neste teste, traduz-se num teste não paramétrico unilateral à direita.

1. Formulação das hipóteses.

H_0: A população tem uma distribuição teórica proposta

vs.

H_1: A população não segue a distribuição proposta

Considera-se um nível de significância α

Teste do Qui-quadrado

2. Construção do quadro comparativo.

Este método requer a recolha e classificação de uma amostra aleatória de tamanho **n**, em **k** classes C_i: i=1,..,k, de variação dos valores amostra contendo a seguinte informação:

Classe	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{C_3}$	**…**	$\mathbf{C_k}$
Frequência absoluta observada - n_i	n_1	n_2	n_3		n_k
Frequência esperada – e_i (De acordo com H_o)	$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{e_3}$		$\mathbf{e_k}$

De acordo com H0 (a população tem uma determinada distribuição) calculam-se as frequência teóricas esperadas:

$$e_i = n \times P(X|_{Ho} \in C_i) = n \times p_i : i = 1,2, \ldots, k.$$

onde,

$$n = \sum_{i=1}^{k} n_k = \sum_{i=1}^{k} e_k$$

(Devem unir-se classes adjacentes se e_i<5)

Teste do Qui-quadrado

3. Identificação da estatística teste

Pode mostrar-se que a **estatística teste**

$$Q = \sum_{i=0}^{k} \frac{(N_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

é uma v.a. que se aproxima de uma distribuição Qui-quadrado com *k-m-1* graus de liberdade. Considera-se que a qualidade da aproximação é suficientemente boa se $n \geq 30$ e $e_i \geq 5$.

k – número total de classes ou valores individuais considerados na amostra (após reformulação do quadro comparativo).

m – número de parâmetros que foi necessário estimar para definir H_0

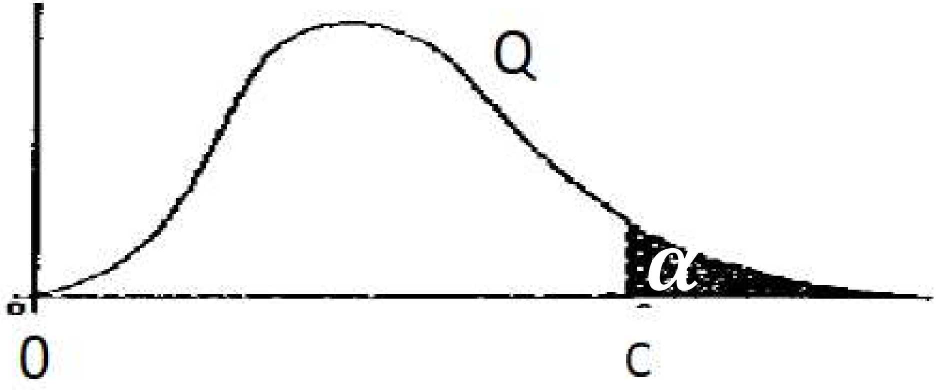
Teste do Qui-quadrado

4. Região crítica

$$RC_{\chi} = [c, +\infty[,$$

onde

$$P(Q_{H0} > c) = \alpha,$$



$$c = inv.chi(\alpha; k1 - m - 1) \quad : Excel\ versão\ PT$$

5. Decisão

Calcular $q_0 = \sum_{i=0}^{k1} \frac{(n_i - e_i)^2}{e_i}$

Se $q_0 < c$ H0 não deve ser rejeitada, aceitando-se o ajuste ao nível de α*100%.

Se $q_0 \geq c$ H0 é rejeitada - a distribuição teórica proposta é rejeitada.

Teste do Qui-quadrado

Procedimento geral:

1 - Formulação das Hipóteses

H0: X tem uma distribuição proposta

H1: X não segue a distribuição proposta

2 - Construção do quadro comparativo das frequências observadas versus as frequências esperadas. Modificação do quadro (enquanto algum $e_i<5$).

3 – Determinação dos graus de liberdade a partir do passo 2. e identificação da estatística teste.

4 - Cálculo da região crítica, $RC = [c, +\infty[$

5 – Calcular o qui-quadrado observado, q_0, e decidir comparando esse valor com a região crítica:

- Se $q_0 < c$, aceita-se o ajuste (H0) ao nível de significância $\alpha \times 100\%$
- Se $q_0 \geq c$ a distribuição proposta é rejeitada.

Exemplo de aplicação I

O registo do número de aviões que aterraram num aeródromo em 60 dias escolhidos aleatoriamente produziu o quadro seguinte:

Número de aviões	0	1	2	3	4	5
Nº de dias	10	18	17	10	4	1

Ajuste uma distribuição teórica aos dados observados verificando a qualidade do ajuste ao nível de 5%

Exemplo de aplicação II

Numa inspeção alimentar foi analisada a quantidade de corantes presente por cada 100 ml de uma dada bebida. Os dados registados para 55 bebidas foram os seguintes:

Quantidade (ml)	[7,2-7,6[	[7,6 -8,0[	[8,0-8,4[	[8,4-8,8[	[8,8-9,2[	[9,2-9,6]
Nº de bebidas	2	7	18	15	9	4

Ajuste uma distribuição teórica aos dados observados verificando a qualidade do ajuste ao nível de 5%