

Chi square Test

Date: / /

[Class Assignment] - 8

[1] In the 2000 US Census, the ages of individuals in a small town were found to be following

<18	18-35	>35
20%	30%	50%

In 2010, ages of $n = 500$ individuals were sampled. Below are the results.

<18	18-35	>35
121	288	91

Using $\alpha = 0.05$, would you conclude that the population distribution of ages has changed in the last 10 years.

Sas

	<18	18-35	>35
Expected	20%	30%	50%

$n = 500$

	<18	18-35	>35
OBSERVED	121	288	91
EXPECTED	100	150	250

(1) H_0 : The data meets the expected data

H_1 : The data does not meet the expected distribution

(2) $\alpha = 0.05$ $CI = 95\%$ [Chi square
2 tailed test]

(3) Degree of freedom

$$df = k - 1$$

$$= 3 - 1$$

$$= 2$$

(4) Decision boundary

$$\text{Chi square value} = 5.991$$

If $\chi^2 > 5.991$ (Reject the Null hypothesis)

(5) Chi square test statistics

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

f_o = observed

f_e = expected

$$= \left[\frac{(121 - 100)^2}{100} + \frac{(288 - 150)^2}{150} + \frac{(91 - 250)^2}{250} \right]$$

$$= \left[\frac{(21)^2}{100} + \frac{(138)^2}{150} + \frac{(-159)^2}{250} \right]$$

$$= 4.41 + 126.96 + 101.124$$

$$= 232.494$$

⑥ Conclusion

$$\text{Since } \chi^2 > 5.99$$

Reject the Null hypothesis.

12

500 elementary school boys and girls were asked what is their favorite color - blue, green or pink. Results are shown below.

OBSERVED	BLUE	GREEN	PINK	
BOYS	100	150	20	270
GIRLS	20	30	180	230
	120	180	200	

Using $\alpha = 0.05$, would you conclude that there is a relationship between gender and favorite color?

Ans

(1) Null hypothesis H_0 : Gender and fav colour are related

H_1 : Gender and fav color are not related

(2) $\alpha = 0.05$ CI = 95%

(3) Degree of freedom

$$(rows - 1)(columns - 1)$$

$$= (2 - 1)(3 - 1)$$

$$= (1)(2)$$

$$= 2$$

(4) Decision boundary chi square value = 5.991

(5) Test Statistic

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$f_e = \frac{P_r P_c}{P_n}$$

Expected	Blue	Green	Pink	
Boys	64.8	97.8	108	270
Girls	55.2	82.8	92	230
	120	180	200	

$$(Boys, Blue) = \frac{120 \times 270}{500} = 64.8$$

$$(Boys, Green) = \frac{180 \times 270}{500} = 97.8$$

$$(Boys, Pink) = \frac{200 \times 270}{500} = 108$$

$$(Girls, Blue) = \frac{120 \times 230}{500} = 55.2$$

$$(Girls, Green) = \frac{180 \times 230}{500} = 82.8$$

$$(Girls, Pink) = \frac{200 \times 230}{500} = 92$$

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$= \left[\frac{(100 - 64.8)^2}{64.8} + \frac{(150 - 97.8)^2}{97.8} + \frac{(20 - 108)^2}{108} \right]$$

$$+ \left[\frac{(20 - 55.2)^2}{55.2} + \frac{(30 - 82.8)^2}{82.8} + \frac{(180 - 92)^2}{92} \right]$$

$$= 259.79$$

$$(5) \text{ Conclusion } \chi^2 = 259.79 > 5.99$$

→ Reject the Null Hypothesis

→ Gender and fav color are not related.