

Answer on Question #57133 – Math – Statistics and Probability

The table below shows the results of a study on smoking and three illnesses. We are interested in determining if the proportions of smokers in the three categories are different from each other.

	Emphysema	Heart Problem	Cancer	Total
Smoker	145	140	300	585
Non-smoker	60	100	230	390
Total	205	240	530	975

p1 = proportions of smokers with emphysema

p2 = proportions of smokers with heart problem

p3 = proportions of smokers with cancer

- What represents the null and alternative hypotheses in this case?
- The expected frequency of smokers with emphysema is?
- The chi-square value is?
- With $\alpha = 0.1$, the critical χ^2 value is?
- $P_1 - P_2$ is? $P_2 - P_3$ is?
- Using $\alpha = 0.1$, the conclusion of the pairwise comparison is?

Solution

Observed data

	Emphysema	Heart Problem	Cancer	Total
Smoker	145	140	300	585
Non-smoker	60	100	230	390
Total	205	240	530	975

Expected data

	Emphysema	Heart Problem	Cancer
Smoker	$\frac{205 \cdot 585}{975} = 123$	$\frac{240 \cdot 585}{975} = 144$	$\frac{530 \cdot 585}{975} = 318$
Non-smoker	$\frac{205 \cdot 390}{975} = 82$	$\frac{240 \cdot 390}{975} = 96$	$\frac{530 \cdot 390}{975} = 212$

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

- What represents the null and alternative hypotheses in this case?

Null hypothesis: the proportions of smokers in the three categories are the same.

Alternative hypothesis: the proportions of smokers in the three categories are not the same.

b. The expected frequency of smokers with emphysema is

$$\frac{205 \cdot 585}{975} = 123$$

c. The chi-square value is

$$\chi^2 = \frac{(145 - 123)^2}{123} + \frac{(140 - 144)^2}{144} + \frac{(300 - 318)^2}{318} + \frac{(60 - 82)^2}{82} + \frac{(100 - 96)^2}{96} + \frac{(230 - 212)^2}{212} = 12.662$$

d. With $\alpha = 0.1$, the critical chi-square value is

$$\chi_{crit}^2 = \chi^2(2; 0.1) = 4.605$$

e. $P_1 - P_2$ is? $P_2 - P_3$ is?

$$P_1 - P_2 = \frac{145}{205} - \frac{140}{240} = 0.124.$$

$$P_2 - P_3 = \frac{140}{240} - \frac{300}{530} = 0.017.$$

f. Using $\alpha = 0.1$, the conclusion of the pairwise comparison is?

We used the Marascuillo procedure.

For an overall level of significance of 0.1, the critical value of the chi-square distribution having 3-1=2 degrees of freedom is $\chi^2(2; 0.1) = 4.605$.

$$r_{12} = \sqrt{\chi^2(2; 0.1)} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{4.605} \sqrt{\frac{\frac{145}{205}(1-\frac{145}{205})}{205} + \frac{\frac{140}{240}(1-\frac{140}{240})}{240}} = 0.0965.$$

$$r_{23} = \sqrt{\chi^2(2; 0.1)} \sqrt{\frac{p_3(1-p_3)}{n_3} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{4.605} \sqrt{\frac{\frac{300}{530}(1-\frac{300}{530})}{530} + \frac{\frac{140}{240}(1-\frac{140}{240})}{240}} = 0.0824.$$

contrast	value	critical range	significant
$ P_1 - P_2 $	0.124	0.0965	Yes
$ P_2 - P_3 $	0.017	0.0824	No

The conclusion of the pairwise comparison: the proportion of smokers with emphysema is different from two others; there is not enough data to conclude that the proportion of smokers with heart problem and the proportion of smokers with cancer are different.