AI1110 : Probability and Random Variables

Assignment 9

Mannem Charan(AI21BTECH11019)

June 13, 2022

Outline

Question

Question Example 8.25: In a certain university, 60% of all first-year students are male and 75% of all entering students graduate. We select at random the records of 299 males and 101 females and we find that 168 males and 68 females graduated. Test the hypothesis that the events $B = \{male\}$ and $C = \{graduate\}$ are independent with $\alpha = 0.05$.

Solution

Solution : We shall test the hypothesis that two events B and C are independent;
H_0 :

$$\Pr(BC) = \Pr(B)\Pr(C) \tag{1}$$

against, H_1 :

$$\Pr(BC) \neq \Pr(B)\Pr(C) \tag{2}$$

Given that the probabilities,

$$\Pr(B) = 0.6 \tag{3}$$
$$\Pr(C) = 0.75 \tag{4}$$

Now we will apply chi-square test to the partition consisting of the four events,

$$A_1 = BC \tag{5}$$
$$A_2 = BC' \tag{6}$$
$$A_3 = B'C \tag{7}$$
$$A_4 = B'C' \tag{8}$$

Under the hypothesis H_0, the components of each of the events A_i are independent. Hence,

$$p_{01} = \Pr(B)\Pr(C) \tag{9}$$
$$= 0.45 \tag{10}$$
$$p_{02} = \Pr(B)(1 - \Pr(C)) \tag{11}$$
$$= 0.15 \tag{12}$$
$$\tag{13}$$

$$p_{03} = (1 - \Pr(B)) \Pr(C) \tag{14}$$
$$= 0.3 \tag{15}$$
$$p_{04} = (1 - \Pr(B)) (1 - \Pr(C)) \tag{16}$$
$$= 0.1 \tag{17}$$

where,p_{0i} is probability of event A_i.
Now according to the chi-square test we will accept H_0 iff,

$$\sum_{i=1}^{4} \frac{(k_i - np_{0i})^2}{np_{0i}} < \chi^2_{1-\alpha}(3) \tag{18}$$

The left side of the inequality is known as test statistic(q),

$$\mathbf{q} = \sum_{i=1}^{m} \frac{(\mathbf{k}_i - np_{0i})^2}{np_{0i}} \tag{19}$$

where, k_i is the number of occurences of the event A_i, m is the no. of classes of the partition and $\chi^2(k)$ is chi-square distribution.

In this case,

$$k_1 = 168 \tag{20}$$
$$k_2 = 131 \tag{21}$$
$$k_3 = 68 \tag{22}$$
$$k_4 = 33 \tag{23}$$

Now after substituting we will get,

$$q \approx 4.1 \tag{24}$$

And since,

$$\chi^2_{0.95}(3) \approx 7.81 > 4.1 \tag{25}$$

We can accept the independent hypothesis.

```
no. of observations:400
No. of events in the partition:4
enter the percentile of chi-square distribution:7.81
enter the number of occurences of each event:
168
131
68
33
enter the theoritical probailities of each event:
0.45
0.15
0.3
0.1
The value of q(test statistic): 4.092625
The two events are independent
```

Figure 1: Output of python code