Mathematica Aeterna, Vol. 6, 2016, no. 4, 557 - 559

An Application of Mathematical Statistics --------Relationship between population and air pollution

Hanyu Zhang

College of Mathematics and Information Science,
Hebei University
Baoding,071002,P.R.China

Jian Shi
College of Mathematics and Information Science,
Hebei University
Baoding,071002,P.R.China

Abstract: In this paper,we use chi square test of independence to search the influence of population on the air pollution.The results shows that the air pollution is not related to the population.
Mathematics Subject Classification:62A01
Keywords: r×c contingency table, test of independence, air pollution

1 Introducton

With the increasing development of urbanization and industrialization, air pollution in the city draws many people's attention. Some people blame the air pollution for the quantity of people is too large. The more people, the more waste emissions. However, Shenzhen, a city in China, with more than 12 million people, has an excellent air environment. Is air pollution related to the number of people? In this paper, we use mathematical statistics to solve this problem.

Theoretically, chi square test of independence[1] is referred in this problem. This method has been used by S. Mao used to test some studies in Biology. We also use the r×c contingency table[1] to determine whether the two attributes of the same sample is dependent or not.

In this paper, we set the capital letter P stands for the probability of events, the letter n stands for the frequency of events. If events A and B is independent, that is $P(AB)=P(A)P(B)$.

In the r×c contingency table,we can make

original hypothesis $H_0 : P_{ij} = P_{i\cdot}P_{\cdot j}$ $i = 1,\cdots,r$ $j = 1,\cdots,c$

alternative hypothesis H_1 : the opposite of H_0.

The test statistics is $\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(n_{ij} - n\hat{P}_{ij})}{n\hat{P}_{ij}}$, $\hat{P}_{ij} = \hat{P}_{i\cdot}\hat{P}_{\cdot j} = \frac{n_{i\cdot}}{n}\cdot\frac{n_{j\cdot}}{n}$

For a given significance level α ,the reject filed is

$$W = \{\chi^2 = \chi^2_{1-\alpha}((r-1)(c-1))\}$$

To determine the relation between the air pollution and the population, we investigate 158 cities of China. The information of the cities' air pollution situation and the population is in the year 2015.We collect data on the website[2][3].

2 Result

We divide the air quality level into three categories: excellent, good, poor. We set A for the concentration pollution of PM2.5,the value is $\mu g/m^3$ .If $0 < A \le 50$,we regard the air quality is excellent. If $50 < A \le 100$,we regard the air quality is good. If $A > 100$,we regard the air quality is poor. We set B for the population of a city, the value is a billion. There are three types of cities: medium city($0 < B \le 100$),large city($100 < B \le 500$),mega-city($B > 500$).We put each sample into the 3×3 contingency table according to the standard above.

table 1

population population air quality	excellent	good	poor	total
medium city	0	6	0	6
large city	22	57	2	81
mega-city	18	52	1	71
total	40	115	3	158

We make original hypothesis $H_0 : P_{ij} = P_{i\cdot}P_{\cdot j}$ $i = 1,2,3$ $j = 1,2,3$

Based on the table 1,the respectively MLE:

$$n\hat{P}_{11} = 158 \times \frac{40}{158} \times \frac{6}{158} = 1.5202$$

$$\mathrm{n}\hat{P}_{12} = 158 \times \frac{115}{158} \times \frac{6}{158} = 4.3697$$

$$\vdots$$

$$\mathrm{n}\hat{P}_{33} = 158 \times \frac{3}{158} \times \frac{71}{158} = 1.3491$$

$$\chi^2 = \frac{(0-1.5202)^2}{1.5202} + \frac{(6-4.3697)^2}{4.3697} + \cdots + \frac{(1-1.3497)^2}{1.3497} = 2.7485$$

$$\chi^2_{0.9}(4) = 7.7794$$

$\chi^2 < \chi^2_{0.9}(4)$,that means there are a high possibility that the original hypothesis is right. We can get the air pollution is not related to population.

Acknowledgement:

The author Jian Shi (corresponding author) is supported by Hebei Education Department (No. ZC2016009, QN2016030), Hebei University Funds for Distinguished Young Scientists and Natural Science Foundation of Shandong Province (No. BS2015SF006).

References

[1] Shisong Mao et al, Probability and Statistics, Higher Education Press, 2014.7, page.
[2] www.greenpeace.org.cn
[3] www.stats.gov.cn

Received: July, 2016