

Probabilty & Statistics

Problem set №13. June 2020

1. The data represents fuel consumption in liters per 100 km. H_0 : median = 7.75, significance level $\alpha = 0.05$. Use the sign test.
 2. Two candidates take part in the second round of the election. Out of 500 people surveyed, 275 declared support for candidate A, 225 - for candidate B. Can the winner be identified at $\alpha = 0.05$?
 3. Data: a comparison of the test results of pupils from the E class (group) and the T class. Use the Wilcoxon test to check if there is a difference in the results of these classes. (E≡ experimental, T≡ traditional)
 4. Data is the weekly number of TV viewing for women and men who are in a some kind of partnership. Using the Wilcoxon rank test for data pairs, check whether there is a difference in TV viewing by men and women.
 5. We compare 4 types of diet. The columns contain weight loss as a result of diet. Use Kruskal-Wallis test to check if diets are effective to the same extent.
 6. In the first column there are assessments of food products by specialists, in the second - by so called "ordinary consumers". Calculate the Spearman correlation coefficient.
 7. The data includes the duration of action of 4 drugs. Use the Friedman test to check if the duration of these medications is the same.
 8. Several (seven) farmers have identified the most important treatments limiting the opportunities for farming. 1 means the most important factor, etc. Using the Friedman test, check if any treatment is significantly different.
-
9. **(E1)** The first column is the number of cigarettes smoked per day, the second - pressure. Calculate the Pearson correlation coefficient.
 10. **(E1)** Two referees (eg. in ski-jumping) rated 10 players. Use the Pearson's correlation coefficient to check if their ratings are similar.
 11. **(E2)** For $n = 10$ and $n = 50$ we calculate values of $B(n, p)$ distribution, where $x = 0, 1, \dots, n$ and $p = 0.4$. Then we approximate the Bernoulli distribution with the the distribution $N(0, 1)$. What is the maximum error of both approximations?
 12. **(E2)** Data is such as in the file `data1301.txt`. We test the hypothesis about the mean value $H_0 : \mu = 8.2$, the alternative hypothesis is $H_a : \mu > 8.2$, we do not know the variance. Consider three variants of data: data such as in a file, we repeat each value twice, we repeat each of the values 5 times. Therefore, \bar{X} will not change.
 - (a) How does the variance of the sample change? (give for example the multiplier)
 - (b) Give the value of three **p-values**. ($n_1 = 31$, $n_2 = 62$, $n_3 = 155$)