



Unit 11: Non Parametric Tests

FLIGHT The duration of a flight between two cities (X , in minutes) is randomly distributed. A random simple sample of 12 flights shows an average of 100 minutes with standard deviation 15.

- If Jarque-Bera test is performed over the sample, obtaining $d_{JB}^* = 4.1$, which should be the conclusion of the test?
- According to the flight attendant: *"the expected time of flight does not exceed the hour and a half."* Test this statement with the information available.
- If the test for independence between the duration of the flight (considering the categories normal and high) and climatology (normal or adverse) shows $d_{IND}^* = 10.5$, which should be the conclusion of the test?

FISH Weekly fish consumption (in grams) is analyzed over a simple random sample of 96 children.

- The sample shows the following coefficients of skewness, and kurtosis, $g_1 = 0.307$ and $g_2 = -0.53$, respectively. What should be the conclusion of the Jarque-Bera test?
- Some experts believe that the variance of consumption is at least 22,500. Knowing that the previous sample shows $s^2 = 20790.69$, what could be said about experts' belief?
- The level of fish consumption is said to be independent from gender. Given the following sample information, test the assumption of independence.

Level of consumption Gender	Low	Medium	High
Female	12	17	16
Male	14	22	15

CLIMBER The authors of a guide need to estimate the expected time (in hours) required to reach the highest point of a mountain. A sample of 30 ascensions is considered showing an average ascension time of 7.07 hours.

- If out of the 30 sampled ascensions, 15 are classified as fast and define a total of 13 runs. What could be said about the randomness of the sample?
- Recently a new route has been opened. A random sample of 22 ascensions over the new route shows an average of 6.63 hours. Assuming the normality of ascension times in both routes with variances 0.69 and 0.83, respectively, could it be claimed that the new route is faster?

WORKERS A sample of 200 unskilled manufacturing workers provides the following information on wages and experience:

Wage / Experience	Low	Medium	High
Low	62	10	2
Medium	14	38	12
High	2	9	51

- Discuss the possible independence of the two characteristics at the 1% significance level.
- Based on sample information, could it be said that the percentage of low-wage workers in this sector is at least 50?



PROPOSED

CAR A car manufacturer is carrying out a market study to analyze the possible relationship between the *type of vehicle purchased* (sports, family and compact) and *gender of the buyer*. The next table shows Data on a simple random sample of 100 customers who have purchased any of their cars in the last month:

Sex \ Type of vehicle	Sports	Family	compact
Man	8	24	30
Woman	14	12	12

- The marketing department states that the type of vehicle purchased is independent of the gender of the buyer. What could be said about such statement?
- The company is also analyzing car longevity (in thousands of km). A simple random sample of 40 cars shows the following coefficients of skewness and kurtosis, respectively, $g_1 = 0.1637$ and $g_2 = -0.8027$.
 - Test the normality of the variable car longevity.
 - The company states that the variance of car longevity is greater than 500. What could be said about this claim, knowing that $s^2 = 142$?

SOLUTION

a) $p < 0.005$; b1) $p < 0.9$; b2) $p > 0.999$

HEALTH CARE The time a doctor spends with each patient is a random variable (X, in minutes per patient). A simple random sample of 15 patients shows:

$$\bar{x} = 9,866 \quad s_x = 1,505$$

- Test the assumption of normality for the variable time, knowing that the coefficients of skewness and kurtosis are $g_1 = 0.224$ and $g_2 = -0.187$, respectively, for sample data.
- The media claim that the expected time spent per patient does not exceed 9 minutes. What could be said about such claim?

SOLUTION

a) $p > 0.9$; b) $p < 0.025$

NGOs An NGO is analyzing the donations (in hundreds of euros) received from its members in certain region A, which are a random variable that follows a normal distribution with $\sigma = 0.75$. A random simple sample of 30 members shows an average donation of 2.5.

- A runs test is performed over the sample, obtaining $d_R^* = -0.89$. State the null hypothesis and the conclusion of the test.
- The donations received by the NGO in another region B are also normally distributed with $\sigma = 0.8$. A simple random sample of 32 members in region B shows an average contribution of 2.45. Given that both samples are independent, could it be claimed that the expected donations are the same in both regions?

SOLUTION

a) $p = 0.373$; b) $p = 0.1286$



PRODUCTION A simple random sample of 20 months on the production of a factory (X, in thousands of tons) shows the following results: $\bar{X} = 24$ and $S = 4$.

- a) If Jarque-Bera test is performed over the sample, $d^* = 1.3$ with $p = 0.522$, state the null hypothesis of the test and the conclusion.
- b) An economic analyst argues that the expected monthly production is at least 25 thousand tons. Is his claim admissible?

SOLUTION

- a) We may assume that the population is normally distributed; b) $p > 0.1$.