

## ARTIFICIAL INTELLIGENCE

### Lab5. Testing hypotheses – single population

CAUTION! Most of the data which appear in the tasks below is available in [data\\_hip.csv](#)

1. An engineer wish to build of a new wind power plant in a given region. To verify, if such investment is reasonable (the average strength of the wind should be above 4 m/sec), every month of a year the strength of a wind in the region was measured (in m/sec):

5.9; 4.4; 5.4; 3.8; 4.0; 4.2; 3.0; 3.6; 4.6; 6.5; 5.6; 4.8.

Assuming that the strength of a wind is normally distributed random variable, verify engineer's claim. Assume significance level  $\alpha = 0.05$ . Write your own function and compare with the results obtained from respective R function.

$t=2.4186$ ,  $p\text{-val}=0.017$ , reject  $H_0$

2. A manufacturer of heat pumps declares that the averaged efficiency index (COP) of its pumps is at least 3.5 (it means that at least 70% of a heat comes from a natural source, whilst the remaining comes from the compressor). A consumer who bought a pump claims that COP of his equipment is much smaller than 3.5 (and thus he wonders if he should complain with the pump). Therefore, during some time, the efficiency index of his pump was measured, giving the results:

3.5; 3.2; 3.6; 3.0; 3.3; 3.8; 2.5; 3.0; 3.7; 3.9.

Assuming normality of a random variable describing COP verify if the data confirms consumer's suspicion. Assume significance level  $\alpha = 0.01$ .

$t=-1.09$ ,  $p\text{-val}=0.15$ , no reason to reject  $H_0$

3. A manufacturer of sprinkler systems used for fire protection in office buildings declares that the true average system-activation temperature is 54°C. A sample of 9 systems, when tested, gives the results:

53, 56, 55, 55, 56, 54, 57, 55, 54.

If the distribution of activation times is normal with standard deviation 2°C, does the data confirm, that the activation time might be different than declared? Assume significance level 0.01. Write your own function and compare the decision with the result of respective R function.

$z=1.5$ ,  $p\text{-val}=0.13$ , no reason to reject  $H_0$

4. It is known that the distribution of the depth of the sea in some area is normal with standard deviation 5 m. In view of climate changes, the depth of the sea in this area was measured 5 times giving the following results (in meters):

862, 870, 876, 866, 871.

Assuming significance level  $\alpha = 0.05$  does the data confirm, that the climate changes, that is that the average depth of the sea in this area is significantly different than 870 m, measured 50 years ago?

$z=-0.447$ ,  $p\text{-val}=0.65$ , no reason to reject  $H_0$

5. A machine produces metal plates of a given size with a nominal thickness of 0.04 mm. An engineer suspects that the machine should be calibrated, as in his opinion produced metal plates are thinner than nominal level. Thus, a thickness of 40 plates produced by this machine was measured, with the results available in [data\\_hip.csv](#) file (variable *thickness*). Assuming significance level  $\alpha = 0.05$  verify engineer's suspicion.

$z=-2.0055$ ,  $p\text{-val}=0.022$ , reject  $H_0$

6. Cuckoos used to put their eggs into the nests of other birds, also small wrens. It is known that the average length of wren's eggs is 16 mm and the standard deviation of the length is 3 mm. Some naturalists think that the size of a cuckoos' egg can become like the size of egg of „adopting” birds. Thus, the length of 21 of eggs (put to the wren's nests) was measured (in mm) and the results are as follows:

17.93; 18.52; 19.66; 14.30; 17.52; 20.76; 20.26;  
 19.82; 21.40; 16.54; 18.64; 17.62; 20.79; 19.14;  
 16.74; 14.93; 18.56; 15.43; 15.19; 21.05; 20.79.

Assuming normality of the length of cuckoos' eggs and significance level 0.05, verify naturalists' claim about:

(a) mean length of cuckoos' eggs;

$t=4.9268$ ,  $p\text{-val}=0$ , reject  $H_0$

(b) variance of length of cuckoos' eggs;

$\chi^2=10.721$ ,  $p\text{-val}=0.09$ , no reason to reject  $H_0$

Estimate with 95% of confidence the true mean length of cuckoos' eggs. What is the relation between the confidence interval and decision from the hypothesis test? (17.36;19.37)

7. In a sample of milk 10 measurements of fat content (in %) were obtained as

1.5; 1.8; 1.5; 1.7; 1.6; 1.6; 1.8; 1.6; 1.7; 1.6.

Assuming normality of the fat content in milk and significance level  $\alpha = 0.05$  verify, whether

(a) the mean fat content in milk is different than 1.7 %?

$t=-1.765$ ,  $p\text{-val}=0.11$ , no reason to reject  $H_0$

(b) the variance of fat content in milk is smaller than 0.02 (%)<sup>2</sup>.

$\chi^2=5.2$ ,  $p\text{-val}=0.18$ , no reason to reject  $H_0$

8. Environment Protection Agency fixed the admissible average of pollution in industrial area as 55 milligrams per m<sup>3</sup> (in the radius of 2 km from the factory), with the standard deviation 18 (milligrams per m<sup>3</sup>). An ecologist measured the pollution level in an industrial area 100 times, in various days and nights, and then he computed the mean and standard deviation of measurements: 60 and 20 milligrams per m<sup>3</sup>, respectively. Assuming significance level 0.01 does the data confirm, that the factory is against the law (for verification of hypothesis related to variance assume normality of pollution level).

mean:  $z=2.5$ ,  $p\text{-val}=0.006$ , reject  $H_0$ ; variance:  $\chi^2=122.22$ ,  $0.05 < p\text{-val} < 0.06$ , no reason to reject  $H_0$

9. There is a common opinion that more than 90% of polish population does not read any book last year. Thus, 1100 Poles were asked about reading books and only 100 of them declared, that they read books regularly. At the significance level 0.05 verify, whether the data confirms the opinion. Use the formulas from the lecture and compare the decision with the result of binom.test and prop.test

$z=1.005$ , no reason to reject  $H_0$  /  $p\text{-val}=0.17$ , no reason to reject  $H_0$  /  $p\text{-val}=0.17$ , no reason to reject  $H_0$

10. A random sample of 150 recent donations at a certain blood bank reveals that 82 were type A blood. Does this suggest that the actual percentage of type A donations differs from 40%, the percentage of the population having type A blood? Assume significance level 0.01. Use the formulas from the lecture and compare the decision with the result of binom.test and prop.test.

$z=3.67$ , reject  $H_0$  /  $p\text{-val}=0$ , reject  $H_0$  /  $p\text{-val}=0$ , reject  $H_0$

11. Natural cork in wine bottles is subject to deterioration, and as a result wine in such bottles may experience contamination. The experiment checking 91 bottles of chardonnays showed, that 16 bottles were considered spoiled to some extent by cork-associated characteristics. Does this data provide strong evidence for concluding that more than 15% of all such bottles are contaminated in this way? Assume significance level 0.1.

no reason to reject  $H_0$