



PES University, Bangalore

(Established under Karnataka Act No. 16 of 2013)

UE19CS203 – STATISTICS FOR DATA SCIENCE

Unit-4 - Hypothesis and Inference

QUESTION BANK

Distribution Free Tests:

Exercises for section 6.9: [Text Book Exercise 6.9– Pg. No. [457 – 458]]

1. The article “Wastewater Treatment Sludge as a Raw Material for the Production of *Bacillus thuringiensis* Based Biopesticides” (M. Tirado Mon- tiel, R. Tyagi, and J. Valero, *Water Research*, 2001:3807–3816) presents measurements of total solids, in g/L, for seven sludge specimens. The results (rounded to the nearest gram) are 20, 5, 25, 43, 24, 21, and 32. Assume the distribution of total solids is approximately symmetric.
 - a. Can you conclude that the mean concentration of total solids is greater than 14 g/L? Compute the appropriate test statistic and find the P -value.
 - b. Can you conclude that the mean concentration of total solids is less than 30 g/L? Compute the appropriate test statistic and find the P -value.
 - c. An environmental engineer claims that the mean concentration of total solids is equal to 18 g/L. Can you conclude that the claim is false?
2. The thicknesses of eight pads designed for use in air- craft engine mounts are measured. The results, in mm, are 41.83, 41.01, 42.68, 41.37, 41.83, 40.50, 41.70, and 41.42. Assume that the thicknesses are a sample from an approximately symmetric distribution.
 - a. Can you conclude that the mean thickness is greater than 41 mm? Compute the appropriate test statistic and find the P -value.

- b. Can you conclude that the mean thickness is less than 41.8 mm? Compute the appropriate test statistic and find the P -value.
- c. The target thickness is 42 mm. Can you conclude that the mean thickness differs from the target value? Compute the appropriate test statistic and find the P -value.

3. The article “Reaction Modeling and Optimization Using Neural Networks and Genetic Algorithms: Case Study Involving TS-1-Catalyzed Hydroxylation of Benzene” (S. Nandi, P. Mukherjee, et al., *Industrial and Engineering Chemistry Research*, 2002:2159–2169) presents benzene conversions (in mole percent) for 24 different benzene hydroxylation reactions. The results are

52.3 41.1 28.8 67.8 78.6 72.3 9.1 19.0
 30.3 41.0 63.0 80.8 26.8 37.3 38.1 33.6
 14.3 30.1 33.4 36.2 34.6 40.0 81.2 59.4.

- a. Can you conclude that the mean conversion is less than 45? Compute the appropriate test statistic and find the P -value.
- b. Can you conclude that the mean conversion is greater than 30? Compute the appropriate test statistic and find the P -value.
- c. Can you conclude that the mean conversion differs from 55? Compute the appropriate test statistic and find the P -value.

4. The article “Abyssal Peridotites > 3,800 Ma from Southern West Greenland: Field Relationships, Petrography, Geochronology, Whole-Rock and Mineral Chemistry of Dunite and Harzburgite Inclusions in the Itsaq Gneiss Complex” (C. Friend, V. Bennett, and A. Nutman, *Contributions to Mineral Petrology*, 2002:71–92) presents silicon dioxide (SiO_2) concentrations (in weight percent) for 10 dunites. The results are

40.57 41.48 40.76 39.68 43.68 43.53
 43.76 44.86 43.06 46.14.

- a. Can you conclude that the mean concentration is greater than 41? Compute the appropriate test statistic and find the P -value.
 - b. Can you conclude that the mean concentration is less than 43? Compute the appropriate test statistic and find the P -value.
 - c. Can you conclude that the mean concentration differs from 44? Compute the appropriate test statistic and find the P -value.
5. This exercise shows that the signed-rank test can be used with paired data. Two gauges that measure tire tread depth are being compared. Ten different locations on a tire are measured once by each gauge. The results, in mm, are presented in the following table.

Location	Gauge 1	Gauge 2	Difference
1	3.95	3.80	0.15
2	3.23	3.30	−0.07
3	3.60	3.59	0.01
4	3.48	3.61	−0.13
5	3.89	3.88	0.01
6	3.76	3.73	0.03
7	3.45	3.56	−0.11
8	3.01	3.02	−0.01
9	3.82	3.77	0.05
10	3.44	3.49	−0.05

Assume the differences are a sample from an approximately symmetric population with mean μ . Use the Wilcoxon signed-rank test to test $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$.

6. The article “ n -Nonane Hydroconversion on Ni and Pt Containing HMFI, HMOR and HBEA” (G. Kinger and H. Vinek, *Applied Catalysis A: General*, 2002:139–149) presents hydroconversion rates (in $\mu\text{mol/g} \cdot \text{s}$) of n -nonane over both HMFI and HBEA catalysts. The results are as follows:

HMFI: 0.43 0.93 1.91 2.56 3.72 6.19 11.00

HBEA: 0.73 1.12 1.24 2.93

Can you conclude that the mean rate differs between the two catalysts?

7. A new postsurgical treatment is being compared with a standard treatment. Seven subjects receive the new treatment, while seven others (the controls) receive the standard treatment. The recovery times, in days, are as follows:

Treatment (X): 12 13 15 19 20 21 27

Control (Y): 18 23 24 30 32 35 40

Can you conclude that the mean rate differs between the treatment and control?

8. In an experiment to determine the effect of curing time on compressive strength of concrete blocks, two samples of 15 blocks each were prepared identically except for curing time. The blocks in one sample were cured for two days, while the blocks in the other were cured for six days. The compressive strengths of the blocks, in MPa, are as follows:

Cured 2 days (X): 1326 1302 1314 1270
 1287 1328 1318 1296
 1306 1329 1255 1310
 1255 1291 1280

Cured 6 days (Y): 1387 1301 1376 1397
 1399 1378 1343 1349
 1321 1364 1332 1396
 1372 1341 1374

Can you conclude that the mean strength is greater for blocks cured for six days?

9. In a comparison of the effectiveness of distance learning with traditional classroom instruction, 12 students took a business administration course online, while 14 students took it in a classroom. The final exam scores were as follows.

Online: 66 75 85 64 88 77 74

 91 72 69 77 83

Classroom: 80 83 64 81 75 80 86

 81 51 64 59 85 74 77

Can you conclude that the mean score differs between the two types of course?

- 10.** A woman who has moved into a new house is trying to determine which of two routes to work has the shorter average driving time. Times in minutes for six trips on route A and five trips on route B are follows:

A: 16.0 15.7 16.4 15.9 16.2 16.3

B: 17.2 16.9 16.1 19.8 16.7

Can you conclude that the mean time is less for route A?