Assignment Questions

**CONFIDENCE INTERVALS**

**EXERCISE 1.**What is the normal body temperature for healthy humans? A random sample of 130 healthy human body temperatures provided by Allen Shoemaker yielded 98.25 degrees and standard deviation 0.73 degrees.

Give a 99% confidence interval for the average body temperature of healthy people.

**EXERCISE 2.**The administrators for a hospital wished to estimate the average number of days required for inpatient treatment of patients between the ages of 25 and 34. A random sample of 500 hospital patients between these ages produced a mean and standard deviation equal to 5.4 and 3.1 days, respectively.

Construct a 95% confidence interval for the mean length of stay for the population of patients from which the sample was drawn.

**HYPOTHESIS TESTING**

**EXERCISE 3.**The hourly wages in a particular industry are normally distributed with mean $13.20 and standard deviation $2.50. A company in this industry employs 40 workers, paying them an average of $12.20 per hour. Can this company be accused of paying substandard wages? Use an α = .01 level test. *(Wackerly, Ex.10.18)*

**EXERCISE 4.** Shear strength measurements derived from unconfined compression tests for two types of soils gave the results shown in the following document (measurements in tons per square foot). Do the soils appear to differ with respect to average shear strength, at the 1% significance level?

**EXERCISE 5.**The following dataset is based on data provided by the World Bank (https://datacatalog.worldbank.org/dataset/education-statistics). World Bank Edstats.  [***2015 PISA Test Dataset***](https://docs.google.com/spreadsheets/d/14rVnIUfEm3CuK9bSvS5253RHWzQhXOuNc0I-cCkgpR8/edit?usp=sharing)

1. Get descriptive statistics (the central tendency, dispersion and shape of a dataset’s distribution) for each continent group (AS, EU, AF, NA, SA, OC).
2. Determine whether there is any difference (on the average) for the math scores among European (EU) and Asian (AS) countries (assume normality and equal variances). Draw side-by-side box plot

My Answers

*Lower bound CI = x\_bar - z(sigma/sqrt(n))*

*Upper bound CI = x\_bar + z(sigma/sqrt(n))*

**EXERCISE 1.**

x\_bar = 98.25 F degrees

sigma = 0.73

n = 130

z = 2.576

CI = [98.085, 98.415] F degrees

Conclusion: **Statistically, healthy human body temperature is between 98.085F and 98.415F degrees in the 99% confidence interval.**

**EXERCISE 2.**

x\_bar = 5.4

s = 3.1

n = 500

z = 1.96

CI = [5.128, 5.671]

Conclusion: **Statistically, average number of days required for inpatient treatment of patients between the ages of 25 and 34 is 5 days.**

**EXERCISE 3.**

H0 : mu = $13.20

H1 : mu < $13.20

One-tailed hypothesis

Significance  level = 0.01

n  = 40

sigma = $2.50

Since sigma is known and n >= 30, we can use z test.

z critical = -2.326    |    z test = -2.529    |    z critical > z test      |  Rejection region

p critical = 0.01       |    p value = 0.005  |    p critical > p value   |  Rejection region

Conclusion: **At the 1% significance level, there is statistically sufficient evidence to support the claim that the company paying substandard wages to their workers.**

**EXERCISE 4.**

H0 : mu1 = mu2

H1 : mu1 != mu2

Two-tailed hypothesis

Significance level = 0.01

Since we have 2 samples, independent from each other, and have equal variance according to levene test, we can use pooled t test.

p critical = 0.005   |    p value = 2.593e-06   |   p critical > p value   |   Rejection region

Conclusion:**At the 1% significance level, there is statistically sufficient evidence to support the claim that the soils appear to differ with respect to average shear strength.**

**EXERCISE 5.**

<https://github.com/ihsankoo/Clarusway-IT-School/blob/main/Statistics/Assignments/Statistics_Assignment_2_Exercise_5.ipynb>