Introduction to Statistics - Homework #7

- Use Appendix C: Distribution tables (e.g. Z-table) from page 408 to 417 in the textbook if necessary.

Exercise 7.7 - Sleep habits of New Yorkers

New York is known as "the city that never sleeps." A random sample of $n = 25$ New Yorkers were asked how much sleep they get per night, yielding the following summary statistics:

※ $\sum_{i=1}^{25} x_i = 193.25, \ \sum_{i=1}^{25}(x_i - \bar{x})^2 = 14.2296$

The point estimate suggests New Yorkers sleep less than 8 hours a night on average. Is the result statistically significant? We conduct **a hypothesis test for a population mean** using a significance level of $\alpha = 0.05$. (Assume that normality condition is met.)

(a) State the null and alternative hypotheses.

Solution:

$$H_0 : \mu = 8, \qquad H_A : \mu < 8$$

(b) Find the null distribution of the test statistic.

Solution:

Under H_0, the test statistic follows:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\bar{X} - 8}{S/\sqrt{25}} \overset{H_0}{\sim} t(24)$$

(c) Compute the observed test statistic.

Solution:

$$\bar{x} = \frac{1}{25}\sum_{i=1}^{25} x_i = \frac{193.25}{25} = 7.73, \qquad s = \sqrt{\frac{1}{25-1}\sum_{i=1}^{25}(x_i - \bar{x})^2} = \sqrt{\frac{14.2296}{24}} = \sqrt{0.5929} = 0.77.$$

The observed test statistic is

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{7.73 - 8}{0.77/\sqrt{25}} = \frac{-0.27}{0.154} \approx \boxed{-1.7532}.$$

(d) Compute the p-value and complete the hypothesis test. State the conclusion in the context of the data.

Solution:

$$p\text{-value} = P(T < -1.7532) \approx \boxed{0.0462}$$

Since the p-value is less than $\alpha = 0.05$, we reject the null hypothesis.

Conclusion: There is statistically significant evidence that New Yorkers sleep less than 8 hours per night on average.

Exercise: Light Bulb Lifetime with a New Filament Design

A lighting manufacturer has historically produced light bulbs whose lifetime (in hours) is normally distributed with mean $\mu_0 = 1000$ and known standard deviation $\sigma = 60$. The R&D team has developed a new filament design and suspects that the new design has **increased** the mean bulb lifetime. A random sample of $n = 36$ bulbs from the new design will be tested. We conduct **a hypothesis test for a population mean** using a significance level of $\alpha = 0.05$. (Assume the population standard deviation σ remains unchanged.)

(a) State the null and alternative hypotheses.

Solution:

$$H_0 : \mu = 1000, \qquad H_A : \mu > 1000$$

(b) Find the null distribution of the test statistic.

Solution: Since σ is known, under H_0,

$$\bar{X} \overset{H_0}{\sim} N\big(1000,\ 10^2\big), \quad \text{equivalently,} \quad Z = \frac{\bar{X} - 1000}{10} \overset{H_0}{\sim} N(0,1).$$

(c) Find the rejection region.

Solution 1 (using $\bar{x}$): Reject H_0 when $\bar{x} > c$, where c satisfies $P(\bar{X} > c) = 0.05$ under H_0. Standardizing gives $(c - 1000)/10 = z_{0.05} = 1.645$, so $c = 1000 + 1.645 \times 10 = 1016.45$. Rejection region:

$$\boxed{\bar{x} > 1016.45}.$$

Solution 2 (using z): The critical value at $\alpha = 0.05$ is $z_{0.05} = 1.645$, so the rejection region is

$$\boxed{z > 1.645}.$$

(d) Given that $\bar{x} = 1020$, complete the hypothesis test. State the conclusion in the context of the data.

Solution 1 (using $\bar{x}$): Since $\bar{x} = 1020 > 1016.45$, we reject H_0.

Solution 2 (using z): The observed test statistic is $z = (1020 - 1000)/10 = 2.00$. Since $2.00 > 1.645$, we reject H_0.

Conclusion: There is statistically significant evidence at the $\alpha = 0.05$ level that the mean lifetime of bulbs with the new filament design is greater than 1000 hours.

(e) Assuming that the true population mean is $\mu = 1030$, compute the power of the test.

Solution: Under H_A with true mean $\mu = 1030$, $\bar{X} \overset{H_A}{\sim} N(1030,\ 10^2)$. Using the rejection region $\bar{X} > 1016.45$ from (c),

$$\text{Power} = P(\bar{X} > 1016.45 \mid \mu = 1030) = P\left(\frac{\bar{X} - 1030}{10} > \frac{1016.45 - 1030}{10}\right) = P(Z > -1.355) \approx \boxed{0.9123}.$$

Exercise 7.20 - High School and Beyond, Part I

The National Center of Education Statistics conducted a survey of high school seniors, collecting test data on reading, writing, and several other subjects. Here we examine a simple random sample of 200 students from this survey. We want to know whether the difference between reading and writing scores is statistically significant using **paired t-test**.

(a) State the null and alternative hypotheses.

Solution:

Let μ_{diff} denote the true mean difference between reading and writing scores among high school seniors.

$$H_0 : \mu_{\text{diff}} = 0, \qquad H_A : \mu_{\text{diff}} \neq 0$$

(b) Find the null distribution of the test statistic.

Solution:

Let $n_{\text{diff}} = 200$ be the number of paired differences. Under H_0, the test statistic follows:

$$T = \frac{\bar{X}_{\text{diff}} - 0}{S_{\text{diff}}/\sqrt{n_{\text{diff}}}} = \frac{\bar{X}_{\text{diff}}}{S_{\text{diff}}/\sqrt{200}} \overset{H_0}{\sim} t(199)$$

(c) The average observed difference in scores is $\bar{x}_{\text{diff}} = -0.545$, and the standard deviation of the differences is $s_{\text{diff}} = 8.887$. Compute the observed test statistic.

Solution:

$$t = \frac{-0.545 - 0}{8.887/\sqrt{200}} = \frac{-0.545}{0.6284} \approx \boxed{-0.8673}$$

(d) Compute the p-value and complete the hypothesis test. State the conclusion in context of the data.

Solution:

$$p\text{-value} = 2 \cdot P(T < -0.8673) \approx \boxed{0.3866}$$

Since the p-value is larger than $\alpha = 0.05$, we fail to reject the null hypothesis.

Conclusion: There is not enough statistical evidence to suggest a difference between reading and writing scores among high school seniors.

Exercise 7.19 - Global warming, Part I

We analyzed temperature data from 197 NOAA stations where records were available for both 1948 and 2018, comparing the number of days exceeding 90°F. The average difference (2018 - 1948) was 1.5 days with a standard deviation of 17.2 days. We seek evidence that there were more days in 2018 that exceeded 90°F from NOAA's weather stations using **paired t-test** with a significance level of $\alpha = 0.05$.

(a) State the null and alternative hypothesis. (Use a **one-sided test.**)

Solution:

Let μ_{diff} denote the true mean difference (2018 - 1948) in the number of days exceeding 90°F.

$$H_0 : \mu_{\text{diff}} = 0, \qquad H_A : \mu_{\text{diff}} > 0$$

(b) Find the null distribution of the test statistic.

Solution:

Let $n_{\text{diff}} = 197$ be the number of paired differences. Under H_0, the test statistic follows:

$$T = \frac{\bar{X}_{\text{diff}} - 0}{S_{\text{diff}}/\sqrt{n_{\text{diff}}}} = \frac{\bar{X}_{\text{diff}}}{S_{\text{diff}}/\sqrt{197}} \overset{H_0}{\sim} t(196).$$

(c) Find the rejection region.

Solution: The critical value at $\alpha = 0.05$ is $t_{0.05}(196) \approx 1.6526$, so the rejection region is

$$\boxed{t > 1.6526}.$$

(d) Given that the sample mean of differences $\bar{x}_{\text{diff}} = 1.5$ and the sample standard deviation of differences $s_{\text{diff}} = 17.2$, complete the hypothesis test. State the conclusion in the context of the data.

Solution: The observed test statistic is

$$t = \frac{1.5}{17.2/\sqrt{197}} = \frac{1.5}{1.2240} \approx 1.2255.$$

Since $1.2255 < 1.6526$, we fail to reject H_0.

Conclusion: There is not enough statistical evidence at the $\alpha = 0.05$ level to conclude that there were more days exceeding 90°F in 2018 compared to 1948.