STAT/MATH 415 HW#7

November 3, 2017

EXERCISES

8.1.1 Assume that IQ scores for a certain population are approximately $N(\mu, 100)$. To test $H_0 : \mu = 110$ against the alternative $H_1 : \mu > 110$, we take a random sample of size $n = 16$ from this population and observe $\bar{x} = 113.5$

a) Do we accept or reject H_0 at 5% significance level?

Answer: $Z_{0.05} = 1.645$ and we can calculate the test statistic as

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{113.5 - 110}{10/4} = 1.4$$

Since $Z = 1.4$ is not in the critical region $[1.645, +\infty)$, so we do not reject H_0

b) Do we accept or reject H_0 at 10% significance level?

Answer: $Z_{0.10} = 1.28$ and the critical region is $[1.28, +\infty)$. Now $Z = 1.4$ is in the critical region, so we reject H_0.

c) What is the p-value of this test?

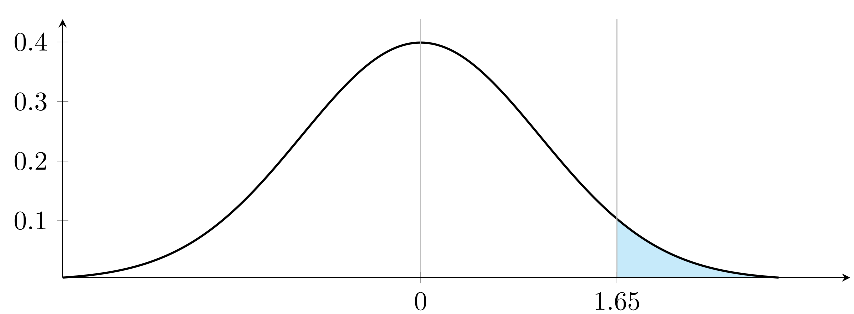
Answer: The p-value is $P(Z > 1.4) = 0.0808$

8.1.3 Let X equal the Brinell hardness measurement of ductile iron subcritically annealed. Assume that the distribution of X is $N(\mu, 100)$. We shall test the null hypothesis $H_0 : \mu = 170$ against the alternative hypothesis $H_1 : \mu > 170$, using $n = 25$ observations of X.

a) Define the test statistic and a critical region that has a significance level of $\alpha = 0.05$. Sketch a figure showing this critical region.

Answer: $Z_{0.05} = 1.645$, so the critical region is $[1.645, +\infty)$ as shaded in the figure. The test statistic is

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - 170}{10/5} = \frac{\bar{X} - 170}{2}$$



b) A random sample of $n = 25$ observations of X yielded the following measurements: 170 167 174 179 179 156 163 156 187 156 183 179 174 179 170 156 187 179 183 174 187 167 159 170 179. Calculate the value of the test statistic and state your conclusion clearly.

Answer: $\bar{X} = 172.52,\ Z = 1.26$ is not in the critical region, so we do not reject H_0

c) Give the approximate p-value of this test

Answer: The p-value is $P(Z > 1.26) = 0.1038$

8.1.6 Let X equal the forced vital capacity (FVC) in liters for a female college student. Assume the distribution of X is approximately $N(\mu, \sigma^2)$. Suppose it is known that $\mu = 3.4$ liters. A volleyball coach claims that the FVC of volleyball players is greater than 3.4. She plans to test her claim with a random sample of size $n = 9$.

a) Define the null hypothesis.

Answer: Let μ_v denote the average of FVC of volleyball players. $H_0 : \mu_v = 3.4$

b) Define the alternative hypothesis.

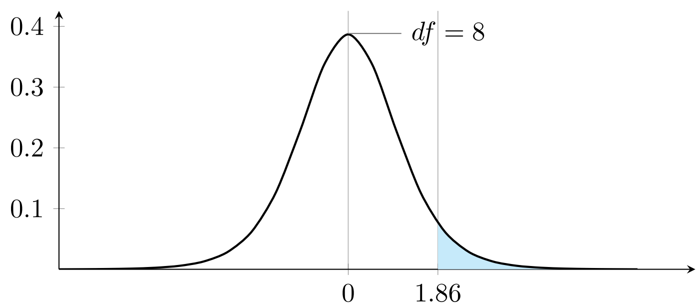
Answer: $H_1 : \mu_v > 3.4$

c) Define the test statistc.

Answer: Test statistic is $T = \dfrac{\bar{X} - \mu}{s/\sqrt{n}} = \dfrac{\bar{X} - 3.4}{s/3}$ while s is sample variation.

d) Define a critical region for which $\alpha = 0.05$. Draw a figure illustrating your critical region.

Answer: $t_{0.05}(8) = 1.860$ so the critical region is $[1.860, +\infty)$ as shaded in the figure



e) Calculate the value of the test statistic given that the random sample yielded the following FVCs: 3.4 3.6 3.8 3.3 3.4 3.5 3.7 3.6 3.7

Answer: $\bar{X} = 3.556,\ s^2 = 0.0278,\ T = \dfrac{3.556 - 3.4}{0.1667/3} = 2.8$

f) What is your conclusion?

Answer: Since $T = 2.8$ is in the critical region, we should reject H_0

g) What is the approximate p-value of this test?

Answer: Since $t_{0.025}(8) = 2.306,\ t_{0.01}(8) = 2.896$, we can estimate the p-value to be within (0.01, 0.025). The p-value calculated by software is $P(T > 2.8) = 0.0116$

8.2.3 Let X equal the weight in grams of a low-fat strawberry kudo and Y the weight of a low-fat blueberry kudo. Assume the distribution of X and Y are $N(\mu_x, {\sigma_x}^2)$ and $N(\mu_y, {\sigma_y}^2)$ respectively. Let

21.7 21.0 21.2 20.7 20.4 21.9 20.2 21.6 20.6

be $n = 9$ observations of X, and let

21.5 20.5 20.3 21.6 21.7 21.3 23.0 21.3 18.9 20.0 20.4 20.8 20.3

be $m = 13$ observations of Y. Use these observations to answer the following questions:

a) Test the null hypothesis $H_0 : \mu_x = \mu_y$ against a two-sided alternative hypothesis. You may select the significance level. Assume the variances are equal.

Answer: Set $\alpha = 0.05$, $t_{0.025}(20) = 2.086$, so the critical region is $(-\infty, -2.086] \cup [2.086, +\infty)$

$$\bar{X} = 21.033 \qquad {S_x}^2 = 0.3675 \qquad \bar{Y} = 20.892 \qquad {S_y}^2 = 1.0141 \qquad {S_p}^2 = \frac{8 \times 0.3675 + 12 \times 1.0141}{9 + 13 - 2} = 0.7555$$

$$T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{1/n + 1/m}} = \frac{21.033 - 20.892}{0.8692 \times 0.4336} = 0.3742$$

Since $T = 0.3742$ is not in the critical region, we do not reject H_0 at 5% significance level.

8.2.5 Some nurses in county public health conducted a survey of women who had received inadequate prenatal care. The mothers selected were divided into two groups: 14 mothers who said they had five or fewer prenatal visits and 14 mothers who said they had six or more prenatal visits. Let X and Y equal the respective birth weights of the babies from these two sets of mothers, and assume that the distribution of X is $N(\mu_x, \sigma^2)$ and the distribution of Y is $N(\mu_y, \sigma^2)$

a) Define the test statistic and critical region for testing $H_0 : \mu_x - \mu_y = 0$ against $H_1 : \mu_x - \mu_y < 0$. Let $\alpha = 0.05$

Answer: $n = m = 14$

$${S_p}^2 = \frac{13{S_x}^2 + 13{S_y}^2}{14 + 14 - 2} = \frac{{S_x}^2 + {S_y}^2}{2} \qquad \Rightarrow \qquad T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{1/n + 1/m}} = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{1/7}}$$

$t_{0.05}(26) = 1.706$ so the critical region is $(-\infty, -1.706]$

b) Given that the observations of X were

49 108 110 82 93 114 134 114 96 52 101 114 120 116

and the observations of Y were

133 108 93 119 119 98 106 131 87 153 116 129 97 110

calculate the value of the test statistic and state your conclusion.

Answer: Based on the data we can have

$$\bar{X} = 100.2143 \qquad {S_x}^2 = 604.489 \qquad \bar{Y} = 114.2143 \qquad {S_y}^2 = 329.258 \qquad {S_p}^2 = \frac{604.489 + 329.258}{2} = 466.874$$

$$T = \frac{100.2143 - 144.2143}{21.607\sqrt{1/7}} = \frac{-14}{8.1668} = -1.714$$

Since $T = -1.714$ is in the critical region, we should reject H_0

c) Approximate the p-value.

Answer: Since $t_{0.05}(26) = 1.706$, $t_{0.025}(26) = 2.056$, we can estimate the p-value is between (0.025, 0.05)

8.2.9 When a stream is turbid, it is not completely clear due to suspended solids in the water. The higher the turbidity, the less clear is the water. A stream was studied on 26 days, half during dry weather (X) and the other half immediately after a significant rainfall (Y). Assume the distributions of X and Y are $N(\mu_x, \sigma^2)$ and $N(\mu_y, \sigma^2)$ respectively. The following turbidities were recorded in units of NTUs:

x:	2.9	14.9	1.0	12.6	9.4	7.6	3.6	3.1	2.7	4.8	3.4	7.1	7.2
y:	7.8	4.2	2.4	12.9	17.3	10.4	5.9	4.9	5.1	8.4	10.8	23.4	9.7

a) Test the null hypothesis $H_0 : \mu_x = \mu_y$ against $H_1 : \mu_x < \mu_y$. Give bounds for the p-value and state your conclusion.

Answer: $n = m = 13$

$$\bar{X} = 6.177 \qquad {S_x}^2 = 17.250 \qquad \bar{Y} = 9.485 \qquad {S_y}^2 = 33.613$$

$$ {S_p}^2 = \frac{12{S_x}^2 + 12{S_y}^2}{13 + 13 - 2} = \frac{17.250 + 33.613}{2} = 25.432$$

$$T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{1/n + 1/m}} = \frac{-3.308}{5.043 \times 0.392} = -1.672$$

$$p - value = P(T \leq -1.672)$$

Since $t_{0.10}(24) = 1.318,\ t_{0.05}(24) = 1.711$, we can estimate the p-value is between (0.05, 0.10). Thus we should reject H_0 at 10% significance level, but do not reject at 5% significance level.