IE 33000: Probability and Statistics in Engineering II (Fall 2022)
School of Industrial Engineering, Purdue University

Homework 3

Instruction: There are 6 problems in total.

1) Problems 1 to 5 – 15 points
2) Problems 6 – 25 points

Due October 14, 2022 (11:59 pm)

For all problems, provide both hand-written solutions and R codes, wherever applicable.

Problem 1

Two machines are used for filling plastic bottles with a net volume of 16.0 ounces. The fill volume can be assumed to be normal with standard deviation $\sigma_1 = 0.0302$ and $\sigma_2 = 0.0254$ ounces. A member of the quality engineering staff suspects that both machines fill to the same mean net volume, whether or not this volume is 16.0 ounces. A random sample of 10 bottles is taken from the output of each machine.

Machine 1		Machine 2	
16.03	16.01	16.02	16.03
16.04	15.96	15.97	16.04
16.05	15.98	15.96	16.02
16.05	16.02	16.01	16.01
16.02	15.99	15.99	16.00

a. Do you think the engineer is correct? Use $\alpha = 0.05$.

b. What is the P-value for this test?

c. Calculate a 95% confidence interval on the difference in means. Provide a practical interpretation of this interval.

a) $H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 \neq 0$

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{16.015 - 16.005}{\sqrt{\frac{0.030^2}{10} + \frac{0.025^2}{10}}}$$

$Z = 0.801$

b) $p\text{val} = 2 \cdot P(Z < -z_0)$

$= 2 \cdot P(Z < -0.80)$

$= 2 \cdot 0.211$

$= 0.4229$

Since $0.423 > 0.05$ we fail to reject

c) $C.I. = \bar{x}_1 - \bar{x}_2 \pm Z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}$

$= 16.015 - 16.005 \pm 1.96 \cdot \sqrt{\frac{0.030^2 + 0.025^2}{10}}$

$C.I. = (-0.014, 0.034)$

0 is included
we have enough
evidence that
both machines
will fill to
same mean
volume

Problem 2

A polymer is manufactured in a batch chemical process. Viscosity measurements are normally made on each batch, and long experience with the process has indicated that the variability in the process is fairly stable with σ = 19.128. Fifteen batch viscosity measurements are given as follows:

724, 718, 776, 760, 745, 759, 795, 756, 742, 740, 761, 749, 739, 747, 742

A process change that involves switching the type of catalyst used in the process is made. Following the process change, eight batch viscosity measurements are taken:

735, 775, 729, 755, 783, 760, 738, 780

Assume that process variability increased due to the catalyst change to σ = 21.2833. If the difference in mean batch viscosity is 10 or less, the manufacturer would like to detect it with a high probability.

a. Formulate and test an appropriate hypothesis using α = 0.10. What are your conclusions?

b. Find the P-value.

c. Find a 90% confidence interval on the difference in mean batch viscosity resulting from the process change.

d. Compare the results of parts (a) and (c) and discuss your findings.

a) $\bar{x}_1 = \frac{1}{15} \cdot \sum x$

$= 750.2$

$\bar{x}_2 = \frac{1}{8} \sum x$

$= 756.88$

$H_0 = \mu_1 - \mu_2 \leq 10$

$H_1 = \mu_1 - \mu_2 > 10$

$$Z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{15} + \frac{\sigma_2^2}{8}}}$$

$$= \frac{(750.2 - 756.9) - 10}{\sqrt{\frac{19.128^2}{15} + \frac{21.283^2}{8}}}$$

$Z = -1.85$

b) $p\text{-val} = P(Z > z_0)$

$= P(Z > -1.85)$

$= 0.967$

Since $0.967 > 0.1$

we fail to reject
hypothesis and say
there's evidence that
mean batch viscosity is 10
or less

c) $(\bar{x}_1 - \bar{x}_2) - Z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\bar{x}_1 - \bar{x}_2) + Z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$(750.2 - 756.8) - 1.645 \cdot \sqrt{\frac{19^2}{15} + \frac{21^2}{8}} \leq \mu_1 - \mu_2 \leq (750.2 - 756.8) + 1.645 \cdot \sqrt{\frac{19^2}{15} + \frac{21^2}{8}}$

$-21.48 \leq \mu_1 - \mu_2 \leq 8.12$

d) There is no significant difference between mean batch viscosity before and after change.

Problem 3

Two catalysts may be used in a batch chemical process. Twelve batches were prepared using catalyst 1, resulting in an average yield of 86 and a sample standard deviation of 3. Fifteen batches were prepared using catalyst 2, and they resulted in an average yield of 89 with a standard deviation of 2. Assume that yield measurements are approximately normally distributed with the same standard deviation.

a. Is there evidence to support a claim that catalyst 2 produces a higher mean yield than catalyst 1? Use $\alpha = 0.01$.

b. Find a 99% confidence interval on the difference in mean yields that can be used to test the claim in part (a).

a) $\sigma_1 = \frac{3}{\sqrt{12}} = 0.866$; $\bar{x} = 89-86 \Rightarrow 3$ $\Rightarrow$ Yes there's evidence

$\sigma_2 = \frac{2}{\sqrt{15}} = 0.52$; $\sigma = \sqrt{0.866^2 + 0.52^2} \Rightarrow 1.008$

b) $(\bar{x}_1 - \bar{x}_2) \pm t_{(0.01,\, 12+13)} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

1 side upper $\Rightarrow (86-89) + 2.485 \cdot \sqrt{\frac{1}{12} + \frac{1}{15}}$

$\mu_1 - \mu_2 \leq -2.038$

Problem 4

A computer scientist is investigating the usefulness of two different design languages in improving programming tasks. Twelve expert programmers who are familiar with both languages are asked to code a standard function in both languages and the time (in minutes) is recorded. The data follow:

	Time	
Programmer	Design Language 1	Design Language 2
1	17	18
2	16	14
3	21	19
4	14	11
5	18	23
6	24	21
7	16	10
8	14	13
9	21	19
10	23	24
11	13	15
12	18	20

Find a 95% confidence interval on the difference in mean coding times. Is there any indication that one design language is preferable?

$= \mu_1 - \mu_2 \pm t_{(0.975,\, 11)} \cdot \sqrt{\frac{\sigma^2}{n}}$

$0.667 \pm 2.2 \cdot \sqrt{\frac{2.96^2}{12}}$

$= (-1.22, 2.55)$

Problem 5

Two chemical companies can supply a raw material. The concentration of a particular element in this material is important. The mean concentration for both suppliers is the same, but you suspect that the variability in concentration may differ for the two companies. The standard deviation of concentration in a random sample of $n_1 = 10$ batches produced by company 1 is $s_1 = 4.7$ grams per liter, and for company 2, a random sample of $n_2 = 16$ batches yields $s_2 = 5.8$ grams per liter. Is there sufficient evidence to conclude that the two population variances differ? Use $\alpha = 0.05$.

$F = \frac{s_1^2}{s_2^2} = \frac{4.7^2}{5.8^2}$ $\boxed{F = 0.657}$

$\alpha = 0.05$

$DF = (9, 15)$

$F_{(0.025, 9, 15)} = 3.12$

$F_{(0.975, 9, 15)} = 0.265$

Since $0.657 > 0.265$ we fail to reject null hypothesis
Thus we say there is no evidence to support claim that population variance differs

---

Problem 6

A research study quantified the absorption of electromagnetic energy and the resulting thermal effect from cellular phones. The experimental results were obtained from in vivo experiments conducted on rats. The arterial blood pressure values (mmHg) for the control group (8 rats) during the experiment are $\bar{x}_1 = 90$, $s_1 = 5$ and for the test group (9 rats) are $\bar{x}_2 = 115$, $s_2 = 10$.

a. Is there evidence to support the claim that the test group has higher mean blood pressure? Use $\alpha = 0.05$, and assume that both populations are normally distributed but the variances are not equal.

b. What is the P-value for this test?

c. Calculate a confidence interval to answer the question in part (a).

d. Do the data support the claim that the mean blood pressure from the test group is at least 15 mmHg higher than the control group? Make the same assumptions as in part (a).

e. Explain how the question in part (d) could be answered with a confidence interval.

a) μ_1 = Control mean

μ_2 = test mean

test $\mu_2 > \mu_1$

$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{90 - 115}{\sqrt{\frac{25}{8} + \frac{100}{9}}}$

$t = -6.63$

P is left tailed

DF in $(7, 8)$

$\alpha = 0.05$

b) $P_{val} = P(t_{(\alpha, df)} < -6.63) = 0.00015$

Since $P_{val} < \alpha$ we reject null hypothesis

c) $(\bar{x}_1 - \bar{x}_2) - t_{(\alpha, DF)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \mu_1 - \mu_2, (\bar{x}_1 - \bar{x}_2) + t_{(\alpha, DF)} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

$(-33.2 \le \mu_1 - \mu_2 \le -16.78)$

Since $\mu_1 - \mu_2 = 0$ and 0 isn't in interval we reject null hypothesis

D) $H_0 = \mu_2 - \mu_1 \ge 15$

$t = \frac{\bar{x}_2 - \bar{x}_1 - 15}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

$t = 2.65$

$t_{(\alpha, 8+9-2)} = 1.75$

Since $2.65 > 1.75$ we fail to reject

E) We could see if the interval of confidence included the calculated critical value.