STAT 662

Experimental Design

Qiangian Yao

1253177.

HW 1.

1. What is the experimental unit in this experiment? (NOTE: The experimental unit is ***singular***, not plural.) Be as specific as possible.
2. What is the factor in this experiment? Be as specific as possible.
3. What are the factor levels in this experiment? Be as specific as possible.
4. What are the treatments in this experiment? Be as specific as possible.
5. What is the response variable in this experiment? Be as specific as possible.

42/50

1) experimental unit: the physical entity to which a treatment is randomly assigned and applied.

⇒ Basil ~~plants~~ from each pot ok

2) Factor: An independent variable that can take any two or more values or levels.

water type ✓

3) Two levels: distilled water ✓
bottled spring water

4) Treatments: A factor level or combination of factor levels applied to experimental units.

distilled water
bottled spring water ✓

5) Response variable:

Weight of basil plants from each pot ✓
in ounces

6. Is randomization implemented in this design? Justify your answer. ← Yes

7. Is replication implemented in this design? Justify your answer. ← Yes

8. Is blocking implemented in this design? Justify your answer. ← NO.

6) Yes. 12 pots are randomly selected to be watered with distilled water while the remaining 6 pots are watered with Spring water.

example:

pots	
1	distilled water
2	
3	Spring water
4	distilled water
⋮	⋮
16	
17	
18	

randomly assigned ✓

7) Yes.
Because there are multiple pots from each treatment group. ✓

8) NO.
pots are not grouped. prior to randomization Each of all 18 pots is only assigned to only one treatment. OK

9) The two sample standard deviations are different,

$\bar{y}_1 = 2.320$ $\quad$ $\bar{y}_2 = 2.887$

$n_1 = 12$ $\quad$ $n_2 = 6$

$s_1 = 0.348$ $\quad$ $s_2 = 0.556$

→ The sample variances are different. This doesn't mean that the population variances are different.

So the assumption of equal variance in the pooled t-test may not be appropriate.

Thus use unpooled t-test assume $\sigma_1^2 \neq \sigma_2^2$.

$\sigma_1^2 = \sigma_2^2$ $\quad$ $\sigma_1^2 \neq \sigma_2^2$ $\quad$ Incorrect hypothesis test (-8)

$H_0: \mu_1 = \mu_2$, $H_a: \mu_1 \neq \mu_2$

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{2.320 - 2.887}{\sqrt{\frac{0.348^2}{12} + \frac{0.556^2}{6}}} = \frac{-0.567}{0.248} \doteq -2.286$$

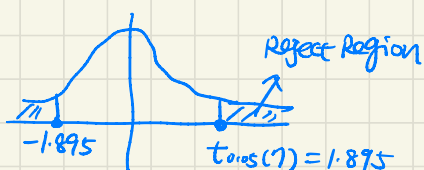
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{\left(\frac{0.348^2}{12} + \frac{0.556^2}{6}\right)^2}{\frac{(0.348^2/12)^2}{11} + \frac{(0.556^2/6)^2}{5}} = 7.027$$

$\alpha = 0.10$

$t_{\alpha/2}(df) = t_{0.05}(7) = 1.895$

Reject H_0 if $|t| > t_{0.05}(7)$.

$|t| = 2.286 > 1.895$, thus, Reject H_0.



There's enough evidence to claim $H_a: \mu_1 \neq \mu_2$ at significance level of 0.1, which is that the mean weights of basil plants watered with distilled water and spring water are different.

(10) The two sample standard deviations are different.

$\bar{y}_1 = 2.320$ $\quad$ $\bar{y}_2 = 2.887$

$n_1 = 12$ $\quad$ $n_2 = 6$

$s_1 = 0.348$ $\quad$ $s_2 = 0.556$

So the assumption of equal variance in the pooled t-test may not be appropriate.

Thus use unpooled t-test assume $\sigma_1^2 \neq \sigma_2^2$.

$H_0: \mu_1 = \mu_2$, $H_a: \mu_1 \neq \mu_2$ ✓

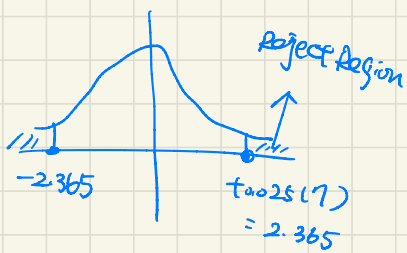
$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{2.320 - 2.887}{\sqrt{\frac{0.348^2}{12} + \frac{0.556^2}{6}}} = \frac{-0.567}{0.248} \doteq -2.286$$ ok

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{\left(\frac{0.348^2}{12} + \frac{0.556^2}{6}\right)^2}{\frac{(0.348^2/12)^2}{11} + \frac{(0.556^2/6)^2}{5}} = 7.027$$ ok

$\alpha = 0.05$

$t_{\alpha/2}(df) = t_{0.025}(7) = 2.365$ ✓

Reject H_0 if $|t| > t_{\alpha 0.025}(7)$.



$|t| = 2.286 < 2.365$,

Thus, Do Not Reject H_0.

There's no enough evidence to claim $H_a: \mu_1 \neq \mu_2$, which is that the mean weights of basil plants watered with distilled water and spring water are different, at significance level of 0.05. ok

11) C.I.

$$(\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/2}(df) \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$2.320 - 2.887$

$= -0.567$

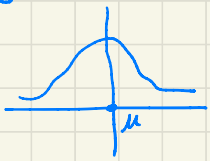
$t_{0.025(7)} \cdot \sqrt{\frac{0.348^2}{12} + \frac{0.356^2}{6}}$

$\doteq 0.587$

$(-0.567 \pm 0.587) \Rightarrow (-1.154, 0.020)$ ok

12) Normal Distribution must be met for t-test.

of what?



By Q-Q plot,
for both treatment groups,
the points fall relatively close to the corresponding straight line,
indicating that the weights from each group are at least approximately normally distributed. ✓
Thus, the normality assumption is met for this study.

13) $\bar{y}_1 = 2.320$, $n_1 = 12$, $s_1 = 0.348$; $\bar{y}_2 = 2.887$, $n_2 = 6$, $s_2 = 0.556$

pooled: $\sigma_1^2 = \sigma_2^2$

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{2.320 - 2.887}{\sqrt{0.180 \cdot \left(\frac{1}{12} + \frac{1}{6}\right)}} = -2.675$$ ✓

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

$$= \frac{11 \times 0.348^2 + 5 \times 0.556^2}{16}$$

$$\doteq 0.180$$

$df = n_1 + n_2 - 2 = 16$

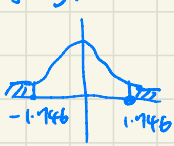
part 9. $\alpha = 0.10$

$\alpha = 0.10$, $t_{\alpha/2}(df) = t_{0.05}(16) = 1.746$

Reject H_0 if $|t| > t_{0.05}(16)$.

$|t| = 2.675 > t_{0.05}(16)$,

thus, Reject H_0

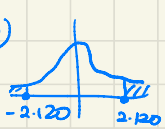


part 10. $\alpha = 0.05$

$\alpha = 0.05$, $t_{\alpha/2}(df) = t_{0.025}(16) = 2.120$ ✓

$|t| = 2.675 > t_{0.025}(16)$

thus, Reject H_0 ✓



9). (a). (i).

unpooled: $\sigma_1^2 \neq \sigma_2^2$

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \doteq -2.286$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}} \doteq 7.029$$

$\alpha = 0.10$, $t_{\alpha/2}(df) = t_{0.05}(7) = 1.895$

Reject H_0 if $|t| > t_{0.05}(7)$.

$|t| = 2.286 > 1.895$,

Thus, Reject H_0.

$\alpha = 0.05$, $t_{\alpha/2}(df) = t_{0.025}(7) = 2.365$

Reject H_0 if $|t| > t_{0.025}(7)$

$|t| = 2.286 < 2.368$

Thus, Do not Reject H_0.

In part 10, $\alpha = 0.05$, By the above calculations we can see that t-test under the two assumptions of $\sigma_1^2 \neq \sigma_2^2$ and $\sigma_1^2 = \sigma_2^2$,

The conclusions are different.

with $\sigma_1^2 \neq \sigma_2^2$ unpooled which should be used for this study,
H_0 is rejected, there's enough evidence to claim the mean weights from the two treatments differ.

with $\sigma_1^2 = \sigma_2^2$ pooled, which shouldn't be used
H_0 is not rejected, there's no enough evidence to claim the mean weights from the two treatments differ.

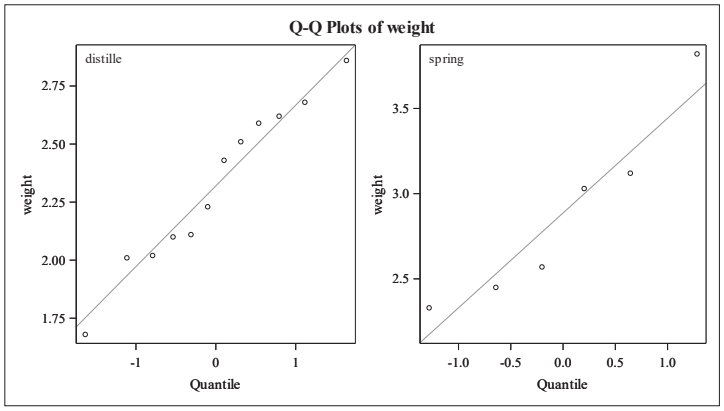
Stat 462/662 – Homework 1
Due: Tuesday, January 24 by the end of the day (11:59pm) to Blackboard

Instructions:

- Read each question and its instructions carefully!
- Perform all calculations by hand, and show all calculations used to get to your final answer.
- When necessary, round values to 3 decimals.

A pesto enthusiast would like to determine whether using distilled water or bottled spring water will result in more basil growth (on average). To investigate this question, the pesto enthusiast plants six basil seeds in each of 18 four-inch pots filled with potting soil and fertilizer. Twelve pots are randomly selected to be watered with distilled water while the remaining six pots are to be watered with spring water. All 18 plots are placed (in random order) on a large table under an LED grow light, which is turned on for 12 hours each day. After 10 weeks, the mature basil plant from each pot is harvested and weighed. The table below provides the weights (in ounces) of each plant watered with each water type as well as the sample mean weight and corresponding sample standard deviation for each water type sample. Q-Q plots of the basil plant weights for each water type are also provided below.

	Distilled Water	Spring Water
	2.51	2.45
	2.59	3.03
	2.23	3.82
	2.86	3.12
	2.10	2.33
	2.11	2.57
	2.01	
	1.68	
	2.43	
	2.62	
	2.68	
	2.02	
Mean	2.320	2.887
Standard deviation	0.348	0.556



Ultimately, the basil enthusiast would like to compare the mean weight of basil plants watered with distilled water to the mean weight of basil plants watered with spring water. Address each item below regarding the design of this experiment as well as its analysis.

1. What is the experimental unit in this experiment? (NOTE: The experimental unit is ***singular***, not plural.) Be as specific as possible.

2. What is the factor in this experiment? Be as specific as possible.

3. What are the factor levels in this experiment? Be as specific as possible.

4. What are the treatments in this experiment? Be as specific as possible.

5. What is the response variable in this experiment? Be as specific as possible.

6. Is randomization implemented in this design? Justify your answer.

7. Is replication implemented in this design? Justify your answer.

8. Is blocking implemented in this design? Justify your answer.

9. Perform the appropriate hypothesis test to determine if it is more appropriate to perform a pooled t-test or unpooled t-test to compare the mean weight of basil plants between the pots watered with distilled water and pots watered with spring water. Use a 10% significance level when performing this test. Make sure to state the hypotheses, calculate the test statistic, determine the rejection region, and state a conclusion within the context of the problem. ***Clearly indicate which t-test (pooled or unpooled) should be used for this experiment.***

10. Perform the appropriate t-test (as you determined in part 9) to determine if there is a significant difference in the mean basil weight between pots watered with distilled water and pots watered with spring water. Use a 5% significance level when performing this test. Make sure to state the hypotheses, calculate the test statistic, determine the rejection region, and state a conclusion within the context of the problem.

11. Using the appropriate method (as you determined in part 9), calculate a 95% confidence interval for the difference in the mean basil weights between pots watered with distilled water and pots watered with spring water.

12. State the main ***distributional*** assumption that must be met in order for the t-test you performed and the confidence interval you calculated to be valid. Is this assumption reasonably met? Justify your answer.

13. Suppose that the basil enthusiast used the incorrect t-test (i.e., the test you did NOT indicate in part 9) to analyze the data from their experiment. Does using this incorrect t-test result in a different conclusion from the conclusion you stated in part 10? Clearly state your final answer with justification and show any calculations you performed to help you answer this question.