

# CMSC 634

## Assignment 3

Due: start of class on October 7th.

Please answer the following 4 questions.

Note: All data reported is fictional.

1. 9 participants are recruited to try a new image processing system. They are asked to complete a series of manipulations to a sample image.

Completion times for the 9 participants are 4 min, 3 min, 12 min, 5 min, 7 min, 8 min, 10 min, 5 min, and 9 min respectively.

- a. Compute the mean of the sample.
  - b. Compute the variance of the sample.
  - c. What is the 80% confidence interval for the population mean?
  - d. What does "80% confidence interval" mean?
2. A grad student wants to compare pictorial passwords (pic) to traditional textual passwords (txt). She recruits two groups of participants. Group 1 ( $n=7$ ) chooses 15 pictorial passwords, and Group 2 ( $n=5$ ) chooses 15 textual passwords. A week later, all participants are asked to recall their passwords. The sample with pictorial passwords remembered {12, 10, 6, 9, 15, 12, 14} passwords. The sample with textual passwords remembered {8, 5, 9, 4, 6} passwords.
    - a. Can you infer a statistically significant difference between pic users' recall and txt users' recall? Use a two-tailed t-test with  $\alpha = 0.05$ .
    - b. What is the null hypothesis of your t-test?
    - c. What is the smallest absolute value of the t-statistic needed to disprove the null hypothesis?
    - d. What is the value of the t-statistic in this case?
  3. 7 participants complete a set of problems using Mathematica and a similar set of problems using Maple. The Mathematica and Maple conditions were counterbalanced. The number of mistakes are shown in the table below.

Participant	Number of Mistakes	
	Mathematica	Maple
A	0	4
B	0	0
C	3	14
D	3	23
E	2	9
F	0	8
G	0	6

- a. Does the software package have a significant effect on number of mistakes? Test with  $\alpha = 0.05$ , two tails.
- b. What is the null hypothesis?
- c. What is the smallest absolute value of the t-statistic needed to disprove the null hypothesis?
- d. What is the value of the t-statistic in this case?

- e. To compute the t-statistic by hand, you need to approximate  $\sigma_{differences}$  (std. devn. of the distribution of differences for the whole population). You can approximate  $\sigma_{differences}$  by using  $s_{differences}$  (an estimate derived from the standard deviation of one sample's differences). What is the value of  $s_{differences}$ ?
4. We are interested in the effect of certain loop optimizations on compilation time. Our test subjects are four different programs. We compile each of them three times: once without optimization, once with loop unrolling, and once with loop splitting. The compilation times are:

Program	No Optimization	Loop Unrolling	Loop Splitting
A	74	88	75
B	76	90	77
C	78	89	76
D	76	85	76

Does the optimization condition significantly affect compilation times? Use a repeated measures ANOVA with  $\alpha = 0.05$ . Draw an ANOVA table like the one drawn in lecture.