ASSIGNMENT 1 — Hypothesis Tests, Effect Size, Power

For all questions below, provide all programming code and plots in the report. Unless stated otherwise, assume $\alpha = 0.05$

Hypothesis Testing

1. A study by Thienprasiddhi et al. examined a sample of 16 subjects with open-angle glaucoma and unilateral hemifield defects. The ages (years) of the subjects were: [62, 62, 68, 48, 51, 60, 51, 57, 57, 41, 62, 50, 53, 34, 62, 61]. Can we conclude that the mean age of the population from which the sample may be presumed to have been drawn is less than 60 years? Let $\alpha = 0.05$. (7 marks).

   a. State the hypotheses (1 mark)

   b. Explain why you selected either a 1-sided or 2-sided test (1 mark)

   c. Compute the t-statistic (2 mark)

   d. Find the critical value (2 mark)

   e. Calculate the p-value (2 mark)

   f. State the statistical decision ("Reject H0" or "Do Not Reject H0") and the conclusion (1 mark)

   g. Perform the same test using a built-in function in R (or some other language) (1 mark)

2. We wish to know if we can conclude that the mean daily caloric intake in the adult rural population of a developing country is different than 2000. Draw from a Normal distribution with a mean of 1985 and a standard deviation of 210. Do this analysis separately or an n = 10, n = 500, and n = 10,000. (9 marks)

   a. State the hypotheses (1 mark)

   b. Explain why you selected either a 1-sided or 2-sided test (1 mark)

   c. Plot a histogram of the simulated data (1 mark)

   d. Compute the t-statistic (1 mark)

   e. Find the critical value (1 mark)

   f. Calculate the p-value (1 mark)

g. State the statistical decision (“Reject H0” or “Do Not Reject H0”) and the conclusion (1 mark)

h. Perform the same test using a built-in function in R (or some other language) (1 mark)

i. Interpret what happens as n increases and explain why? (1 mark)

3. A test designed to measure mothers’ attitudes toward their labor and delivery experiences was given to two groups of new mothers. Sample 1 (attenders: $\mu = 4.75, SD = 1.0, n = 15$) had attended prenatal classes held at the local health department. Sample 2 (nonattenders: $\mu = 3.0, SD = 1.5, n = 22$) did not attend the classes. Do these data provide sufficient evidence to indicate that attenders, on the average, score higher than nonattenders? Use Welch’s t-test. $\alpha = 0.01$ (6 marks).

a. State the hypotheses (1 mark)

b. Explain why you selected either a 1-sided or 2-sided test (1 mark)

c. Compute the t-statistic (1 mark)

d. Find the critical value (1 mark)

e. Calculate the p-value (1 mark)

f. State the statistical decision (“Reject H0” or “Do Not Reject H0”) and the conclusion (1 mark)

4. Kindergarten students were the participants in a study conducted by Susan Bazyk et al. The researchers studied the fine motor skills of 37 children receiving occupational therapy. They used an index of fine motor skills that measured hand use, eye–hand coordination, and manual dexterity before and after 7 months of occupational therapy. Higher values indicate stronger fine motor skills. Can one conclude on the basis of these data that after 7 months, the fine motor skills in a population of similar subjects would be stronger? Data is shown in the table below (6 marks).

Subject	Pre	Post	Subject	Pre	Post
1	91	94	20	76	112
2	61	94	21	79	91
3	85	103	22	97	100
4	88	112	23	109	112
5	94	91	24	70	70
6	112	112	25	58	76
7	109	112	26	97	97
8	79	97	27	112	112
9	109	100	28	97	112
10	115	106	29	112	106
11	46	46	30	85	112
12	45	41	31	112	112
13	106	112	32	103	106
14	112	112	33	100	100
15	91	94	34	88	88
16	115	112	35	109	112
17	59	94	36	85	112
18	85	109	37	88	97
19	112	112			

Source: Data provided courtesy of Susan Bazyk, M.H.S.

a. State the hypotheses (1 mark)

b. Explain why you selected either a 1-sided or 2-sided test (1 mark)

c. Compute the t-statistic (1 mark)

d. Find the critical value (1 mark)

e. Calculate the p-value (1 mark)

f. State the statistical decision (“Reject H0” or “Do Not Reject H0”) and the conclusion (1 mark)

5. Let’s assume we know two groups of data are non-normally distributed but we want to see if they are significantly different from one another (6 marks).

a. Draw from a Beta probability distribution for both Group 1 ($\alpha = 1, \beta = 9, n = 20$) and Group 2 ($\alpha = 2, \beta = 9, n = 20$) (1 mark).

b. Plot a histogram for each group (1 mark)

c. Show a QQ plot for each group (1 mark)

d. Perform a Shapiro-Wilk test for each group and interpret, while highlighting whether a Welch’s t-test or Mann-Whitney U test should be used (1 mark)

e. Perform a Mann-Whitney U test using a built in function in R (or other language) (1 mark)

f. State the statistical decision (“Reject H0” or “Do Not Reject H0”) and the conclusion (1 mark)

6. An assumption of performing a paired t-test is that the distribution of paired differences (i.e., $\bar{X}_D$) is normally distributed. Use the data from **Question 4** to (5 marks):

a. Plot a histogram of the paired differences (1 mark)

b. Show a QQ plot (1 mark)

c. Perform a Shapiro-Wilk test and interpret, while highlighting whether a paired t-test or Wilcoxon test should be used (1 mark)

d. Perform a Wilcoxon test using a built in function in R (or other language) (1 mark)

e. State the statistical decision (“Reject H0” or “Do Not Reject H0”) and the conclusion (1 mark)

7. You have three different groups (G1, G2, G3) and an ANOVA reports there is a significant main effect (don’t worry, we’ll get into ANOVA soon). You then follow-up with multiple mean comparisons using a Welch’s t-test and find the following p-values: [G1 vs G2 = 0.06, G1 vs G3 = 0.049, G2 vs G3 = 0.01]. But you are concerned about Type I error.

a. Perform and Bonferroni correction (1 mark)

b. Perform and Bonferroni-Holm correction (1 mark)

c. Interpret the findings, while considering Type 1 and Type 2 error (1 mark)

8. Here you are going to make p-curves, which gives insight on what frequency you can expect p-values due to chance. You can do this by simulating n experiments (e.g., n = 100,000) and looking at the histogram of p-values. To do this, initialize an array, make a for-loop, calculate a p-value based on some statistic, store the p-value in an array, and do this repeatedly n number of times (See TIPS directly below). Lets simulate 100,000 experiments by comparing 2 groups.

a. For each of the 100,000 iterations, compare Group 1 ($\mu = 10.0, SD = 2.0, n = 10$) and Group 2 ($\mu = 10.0, SD = 2.0, n = 10$) by drawing each group's values from a normal distribution and then performing a Welch's t-test (built in function) (1 mark)

b. Plot the histogram of the 100,000 p-values (1 mark)

c. What proportion of the p-values are under 0.05? (1 mark)

d. For each of the 100,000 iterations, compare Group 1 ($\mu = 10.0, SD = 2.0, n = 10$) and Group 2 ($\mu = 12.0, SD = 2.0, n = 10$) by drawing each group's values from a normal distribution and then performing a Welch's t-test (built in function) (1 mark)

e. Plot the histogram of the 100,000 p-values (1 mark)

f. What proportion of the p-values are under 0.05? (1 mark)

g. Interpret the results from the simulations above(1 mark)

```
TIPS
PVALUES1 = array(NA,100000) # initialize nan array
for (i in 1:100000) { # this makes a for loop in R
# insert group1 here and draw values from normal distribution
# insert group2 here and draw values from normal distribution
pval = t.test(group1, group2, alternative = "two.sided")$p.value
# this saves the p-value and stores in the initialized array
PVALUES1[i] = pval
}
```

Effect Sizes

Power Analyses