Name: Drashti Mehta

Lab Assignment 6

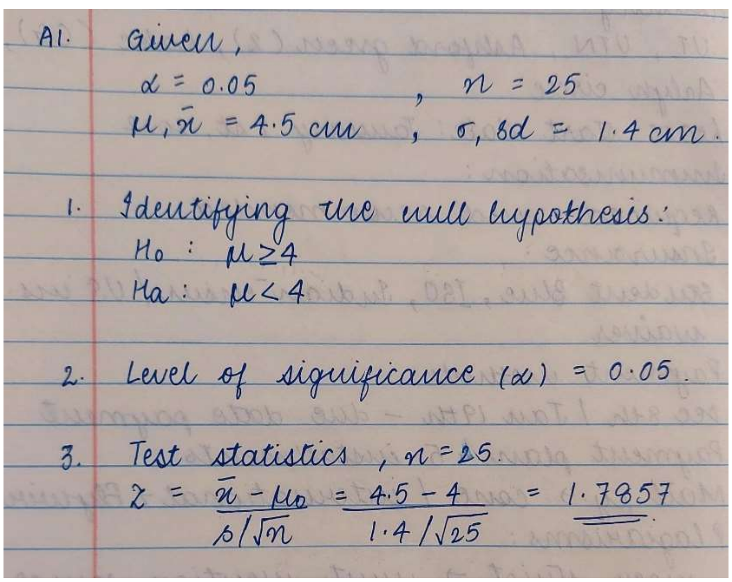
(40 points)

DIRECTIONS

- ***An uploaded copy is due by next Friday.***

A. A group of paleontologists hypothesize that insect wings evolved from modification of ancestral arthropod limbs (α=0.05). Insect wings are small, so this hypothesis requires that ancestral arthropod limbs be less than 4 cm. The scientists sampled 25 arthropods and determined that the mean limb length is 4.5cm, with a standard deviation of 1.4cm. Assume the arthropod limb length follows a normal distribution (12 pts total).

1. Using the 6 steps in hypothesis testing (don't need to show drawn figure), determine whether or not the null hypothesis of the paleontologists is true (5 pts). Be sure to find the critical value from R and compare this to the test statistic (2 pts). Make a conclusion in the context of the situation given (1 pt).



A1. Given,

$\alpha = 0.05$, $n = 25$

$\mu, \bar{x} = 4.5$ cm , σ, sd $= 1.4$ cm.

1. Identifying the null hypothesis:

$H_0 : \mu \geq 4$

$H_a : \mu < 4$

2. Level of significance $(\alpha) = 0.05$.

3. Test statistics, $n = 25$.

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{4.5 - 4}{1.4/\sqrt{25}} = 1.7857$$

R code:

```
> qt(0.05, 24)
[1] -1.710882
> pt(1.79, 24)
[1] 0.95696
```

Conclusion:

Critical value of t < 1.79 and p- value > 0.05

So, as t is not in the rejection area and p – value is greater than 0.05, we fail to reject the null hypothesis. Ancestral arthropods limbs are equal or more than 4 cm.

2. **Construct a "error table" specific for this data. If we assume that your statistical decision is correct, put a check mark where your decision falls on the truth table (1 pt). Write the specific Type I and Type II errors possible for your hypothesis test in a sentence format (1 pt).**

	True	False
Reject H0	Type I error (α)(False Positive)	Correct (1-β)(True Positive)
AcceptH0	Correct (1-α)(True Negative)	Type II error (β)(False Negative)

Type 1 error: Ancestral arthropod limbs are more than 4 cm but claimed they are less.

Type 2 error: Ancestral arthropod limbs are less than 4 cm when they are not.

3. **How can Type I error be reduced (1 pt)? How can Type II error be reduced (1 pt)?**

To avoid type i errors: We should do is to specify a tolerable probability. We can decrease the significance level.

To avoid type ii errors: We should increase the sample size for lowering type ii error.

B. Based upon beer sales in UNCC, it is expected that UNCC students consume 7.5 beers a week. In a questionnaire of 60 random UNCC students, the sample mean of beers consumed per week is 6.5, with a standard deviation of 3.4 beers consumed per week. Assume the number of beers consumed follows a normal distribution (7 pts total).

1. **State the null and alternative hypothesis. Find the test statistic and compare it with the critical value at alpha = 0.10. Determine the confidence interval at 95% (5 pts). Draw a conclusion from the confidence interval (1 pts).**

B. Given : $n = 60$, μ or $\bar{x} = 6.5$, $s = 3.4$, $\alpha = 0.1$

1. Hypothesis :

H_o : $\mu = 7.5$

H_a : $\mu \neq 7.5$.

2. $\alpha = 0.1$

3. $z = \dfrac{\bar{x} - \mu_o}{s/\sqrt{n}} = \dfrac{6.5 - 7.5}{3.4/\sqrt{60}} = \underline{-2.278}$

R code:

qnorm(0.05, lower.tail = F)
[1] 1.644854
> 2*pnorm(-2.28)
[1] 0.02260769
From the above calculation, we see that -2.28< -1.644854 and p-value < alpha, hence We can reject H0.

confidence interval at 95%:
> qnorm(0.975)
[1] 1.959964
> qnorm(0.025)
[1] -1.959964

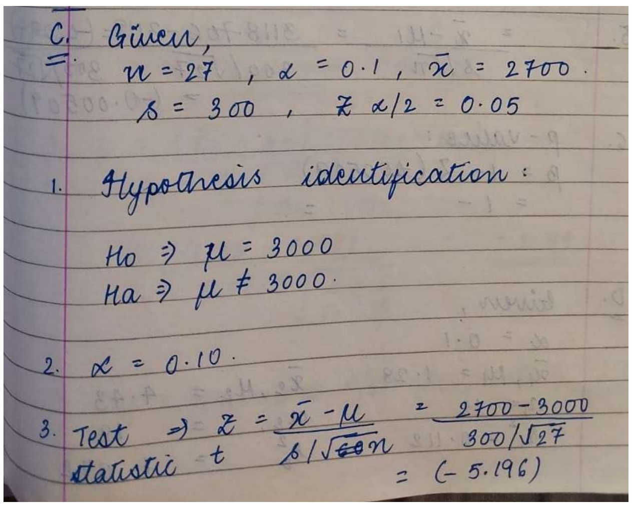
6.5 - (1.959964*3.4/sqrt(60)) < mu < 6.5+(1.959964*3.4/sqrt(60))
5.639697 < mu < 7.360303
So, UNCC students do not consume 7.5 beers a week.

2. **What is the relationship between conclusions of the hypothesis test and the confidence interval you found (1 pt)?**

Hypothesis testing uses the data from sample to determine a specified hypothesis whereas population parameter can be estimated by the confidence interval by using the data from sample. If 95% confidence interval does not contain the hypothesized parameter that means a hypothesis test at 0.05 level will reject the null hypothesis most of the time. This explains that the hypothesis test is testing the same area as the confidence interval, which is alpha. It falls outside both of them.

C. Based on recent biodiversity study, a microbiologist at UNH thinks there are 3000 bacterial species in your gut lining (α=0.10). The microbiologist sampled the gut lining of 27 humans, and determined that there was a mean number of 2700 bacterial species, with a standard deviation of 300. Assume that gut bacterial species follows a normal distribution. We want to test if the mean number of bacterial species in a human's gut lining is different from 3000 (9 pts total).

1. Using hypothesis testing (six steps - don't need to show drawn figure) and using the p-value method, determine whether or not the null hypothesis of the microbiologists is true (6 pts). Be sure to compare the p-value to alpha to make your conclusion (2 pts).
Make a conclusion in the context of the situation given (1 pt).



C. Given,
$n = 27$, $\alpha = 0.1$, $\bar{x} = 2700$.
$s = 300$, $Z\ \alpha/2 = 0.05$

1. Hypothesis identification:

$H_o \Rightarrow \mu = 3000$
$H_a \Rightarrow \mu \neq 3000$.

2. $\alpha = 0.10$.

3. Test statistic ⇒ $Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ t $= \frac{2700-3000}{300/\sqrt{27}}$ $= (-5.196)$

Alpha is 0.1 and Z alpha/2 is 1.645, which came from calculation: qnorm(+0.5)
Hence, we can see that the value for Z calculated (-5.196) < Z critical.

P-value:
> 2*pnorm(-5.19)
[1] 2.102941e-07

The p – value is < alpha. From both these results, we conclude that the null hypothesis can be rejected. The mean number of bacterial species in a human's gut lining is different from 3000.

D. Two species of Arabidopsis were measured for height using random samples. Assume both populations are normally distributed and have equal variances. We wish to know if the two population means differ. Test level α=0.10. The results were as follows (9 pts total).

Species	**a (cm)**	**b (cm)**
n =	**12**	**18**

xx =	**1.28**	**4.43**
s^2 =	**0.112**	**7.072**

1. Using hypothesis testing (six steps - don't need to show drawn figure), determine whether or not the null hypothesis of the microbiologists is true (6 pts). Find the test statistic and be sure to compare the p-value to alpha to make your conclusion (2 pts). Make a conclusion in the context of the situation given (1 pt).

D. Given,

$\alpha = 0.1$

$\bar{x}_1, \mu_1 = 1.28$ $\bar{x}_2, \mu_2 = 4.43$

$n_1 = 12$ $n_2 = 18$

$s_1^2 = 0.112$ $s_2^2 = 7.072$

1. Hypothesis testing:

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

$df = n_1 + n_2 - 2 = 12 + 18 - 2 = \underline{28}$

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

Here we are readily provided s^2 instead of s, ∴ we can directly put the values.

$$\rightarrow \therefore s_p = \sqrt{\frac{(12-1)(0.112) + (18-1)(7.072)}{28}}$$

$= \cancel{0.39359}$ $\underline{2.0825}$

t = (1.28-4.43)/(2.082718*sqrt((1/12)+(1/18))) = -4.05
tdf = t(28,0.05) = 1.701

```
> pt(4.05,28, lower.tail = FALSE)
[1] 0.0001837164
```

From above calculations, t < tdf, α/2 and p- value < alpha. So, we can reject the null hypothesis.
Conclusion: we can conclude that two population means differ.

E. The measured heights of 9 randomly sampled Arabidopsis from species "C" are 1.42, 1.22, 1.36, 1.55, 1.18, 1.47, 1.22, 1.12, 1.62. Is there evidence that the mean level of these

samples is greater than that of a species with a mean of 1.28 at α=0.01? State the hypothesis, show R commands, and use the t.test function in R to make a conclusion (3 pts).

H0 = μ <= 1.28
Ha = μ > 1.28

species_c <- c(1.42, 1.22, 1.36, 1.55, 1.18, 1.47, 1.22, 1.12, 1.62)
t.test(species_c, mu=1.28,alternative="greater")

One Sample t-test

data: species_c
t = 1.2119, df = 8, p-value = 0.1301
alternative hypothesis: true mean is greater than 1.28
95 percent confidence interval:
1.242 Inf
sample estimates:
mean of x
1.351111

p-value = 0.1742 which is > alpha = 0.01

We understand that the true mean which is 1.351 is greater than the mean 1.28 hence we **reject** the null hypothesis.