STA 3032 - Practice set 2

True or False

1. In conducting a hypothesis test regarding a mean, the larger the p-value, the stronger the evidence against the null hypothesis.
   False

2. The (large-sample) sampling distribution for the sample mean is normally distributed only if the underlying population of measurements is normal.
   False

3. Nonparametric/distribution free procedures should ONLY be used for non-normal data.
   False

4. All else being equal, as we increase the sample size, we increase the width of a Confidence Interval for μ.
   False

Open Ended Problems

1. A method of spraying for rust mites has been conducted on a random sample of 100 10-acre plots in citrus groves in Florida. From the sample, the mean and standard deviation for the yield of fruit is 750 and 200 tons, respectively. Construct a 95% confidence interval for μ, the population mean yield per 10-acre plot.
   Answer: 95% CI for μ is

$$750 \mp \underbrace{t_{1-0.025,99}}_{1.984217} \frac{200}{\sqrt{100}} \to (710.3157, 789.6843)$$

2. A scientist wishes to estimate the mean breaking strength of a certain type of steel rod within 3.0 psi with 95% confidence. Based on experience with a similar type of steel, she believes the standard deviation is approximately 15.0 psi in individual breaking strengths. How many steel rods should she test?
   Answer: width is 6,

$$n \geq \left(2(1.96)\frac{15}{6}\right)^2 = \left((1.96)\frac{15}{3}\right)^2 = 96.04$$

so minimum $n = 97$.

3. A textile engineer is interested whether the variation in breaking strength of Yarn Type A is larger than the variation in Type B. Random samples of $n_A = n_B = 7$ measurements were obtained, and the sample variances were $s_A^2 = 5.0$ and $s_B^2 = 1.25$. Assuming both types are normally distributed, do the data provide sufficient evidence to conclude $\sigma_A^2 > \sigma_B^2$? Infer by using $\alpha = 0.05$.
   Answer:

```
(1) p-value for one sided test
> 1-pf(5/1.25,6,6)
[1] 0.05792

(2) Equivalent one sided CI
> 5/1.25*(1/qf(0.95,6,6))
[1] 0.9337361 to infty
```

   Conclusion: (1) p-value > 0.05, (2) 1 is in the CI, so fail to reject null of $\sigma_A^2 \leq (=)\sigma_B^2$.

4. A forensic researcher samples 100 adult males and 100 adult females and measures each subjects right foot length (cm). The following table gives the summary results:

Gender	$\bar{x}$	s
Female	23.80	1.1
Male	26.30	1.6

(a) Test $\text{H}_0 : \mu_F = \mu_M$ versus $\text{H}_A : \mu_F \neq \mu_M$ at the $\alpha = 0.05$ significance level.
Answer:

$$T.S. = \frac{23.80 - 26.30 - 0}{\sqrt{1.1^2/100 + 1.6^2/100}} = -12.87566$$

Since T.S. is about 12.9 standard deviations below the null hypothesis of 0 for the difference ($\mu_F - \mu_M = 0$, as the inference methodology is set up for differences), we know p-value is tiny and $\mu_F \ll \mu_M$. To find p-value we must first find the degrees of freedom, ν which is 175.4964. Then,

```
> 2*pt(-12.87566,175.4964)
[1] 3.873543e-27
```

(b) Compute a 95% CI for $\mu_F - \mu_M$

```
> (23.80-26.30)+c(1,-1)*qt(0.025,175.4964)*
sqrt(1.1^2/100+1.6^2/100)
[1] -2.883199 -2.116801
```

(c) Assuming equal variances, recompute the 95% CI. How does it differ from the previous interval?

```
> (23.80-26.30)+c(1,-1)*qt(0.025,198)*
sqrt((99*1.1^2+99*1.6^2)/198)*sqrt(1/100+1/100)
[1] -2.882897 -2.117103
```

5. Suppose a sample of 20 students were given a diagnostic test before studying a particular module and then again after completing the module. We want to find out if, in general, our teaching leads to improvements in students knowledge/skills (i.e. test scores). We can use the results from our sample of students to draw conclusions about the impact of this module in general.

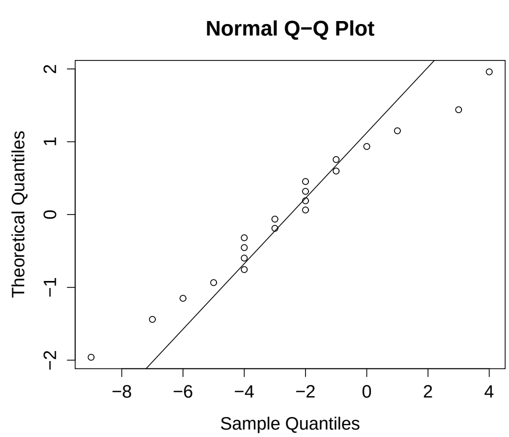
Student	Pre-module score	Post-module score
1	18	22
2	21	25
3	16	17
4	22	24
5	19	16
6	24	29
7	17	20
8	21	23
9	23	19
10	18	20
11	14	15
12	16	15
13	16	18
14	19	26
15	18	18
16	20	24
17	12	18
18	22	25
19	15	19
20	17	26

(a) Is the data paired? **Yes, 2 observations on the same subject.**

(b) Check the normality assumption.
Answer: Since dealing with paired data, we need to work with differences:

$$-4, -4, -1, -2, 3, -5, -3, -2, 4, -2, -1, 1, -2, -7, 0, -4, -6, -3, -4, -9$$

A qqplot of this data shows heavier tailed than normal.



(c) Test the null hypothesis (using both parametric and nonparametric/distribution free approaches) that the true mean difference is zero. What is your final conclusion?

Answer: Parametric should be done purely for practise, but here will only do Wilcoxon signed rank test. There is a zero that we will deleted yielding $T_{-} = 166$ and $T_{+} = 24$ for $n = 19$. The test statistic will be the minimum of the 2 (since 2 sided) which is 24.

```
> wilcox.test(x[,2],x[,3],paired=TRUE,exact=TRUE,correct=FALSE)

Wilcoxon signed rank test

data:  x[, 2] and x[, 3]
V = 24, p-value = 0.004128
alternative hypothesis: true location shift is not equal to 0
```

Note: In the context of the problem maybe a one sided test would be more appropriate since of interest is whether Post scores are greater than Pre scores, i.e. the alternative being that the differences are symmetrically distributed around $D_0 < 0$. In that case, try,

```
> wilcox.test(x[,2],x[,3],alternative="less",mu=0)
```

6. Is cocaine deadlier than heroin? A study was conducted where rats were randomly assigned to receive unlimited supplies of cocaine or heroin. Of the 200 rats assigned to cocaine, 80% died in 30 days. Of the 200 rats assigned to heroin, 50% died in 30 days.

   (a) Compute a 95% Confidence Interval for $p_C - p_H$. Difference in population mean 30-day death rates for rats assigned to cocaine versus heroin.
   Answer: Since the sample sizes are large, Agresti-Coull or classical approach will give very similar answers.

$$\begin{aligned}\tilde{p}_C = \frac{161}{202},\ \tilde{p}_H = \frac{101}{202} &\Rightarrow \tilde{p}_C - \tilde{p}_H \pm 1.96\sqrt{\frac{\tilde{p}_C(1-\tilde{p}_C)}{202} + \frac{\tilde{p}_H(1-\tilde{p}_H)}{202}} \\ &\rightarrow (0.2085366, 0.3855228)\end{aligned}$$

   (b) Do you feel that from this C.I. that there is evidence to conclude that cocaine is deadlier than heroin in this setting?
   Answer: Since the interval is strictly greater than zero, $p_C - p_H > 0 \Rightarrow p_C > p_H$. However, maybe more relevant would be to perform a one-sided C.I. or one-sided hypothesis test since this a one side question.

7. Recent studies of the private practices of physicians who saw no Medicaid patients suggested that the 40th percentile length of each patient visit was 22 minutes. It is believed that the 40th percentile visit length in practices with a large Medicaid load is shorter than 22 minutes. A random sample of 20 visits in practices with a large Medicaid load yielded, in order, the following visit lengths:

$$9.4, 13.4, 15.6, 16.2, 16.4, 16.8, 18.1, 18.7, 18.9, 19.1,$$

$$19.3, 20.1, 20.4, 21.6, 21.9, 23.4, 23.5, 24.8, 24.9, 26.8$$

Based on these data, is there sufficient evidence to conclude that the 40th percentile visit length in practices with a large Medicaid load is shorter than 22 minutes?
Answer: The test statistic is $B = 5$. The p-value is $P(B \leq 5 | B \sim \text{Bin}(20, 0.6))$

```
> pbinom(5,20,0.6)
[1] 0.001611525
```

Remark: Equivalently, someone could also find the same p-value by using $B = 15$ and calculating $P(B \geq 15 | B \sim \text{Bin}(20, 0.4))$

8. A built in dataset in `R` named InsectSprays contains the counts of insects in agricultural experimental units treated with (6) different insecticides. Test whether the variance in counts of the insecticides is equal. (If using `R` just type `InsectSprays` and you will be able to see the data.

```
library(car)

leveneTest(count ~ spray, data=InsectSprays)
#> Levene's Test for Homogeneity of Variance (center = median)
#>       Df F value   Pr(>F)
#> group  5  3.8214 0.004223 **
#>       66
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

9. Students were asked to classify themselves according to smoking (Heavy, Regularly, Occasionally, Never) and according to exercise (Frequently, Some, None)

Smoking/Exercise	Frequently	Some	None
Heavy	7	3	1
Regularly	9	7	1
Occasionally	12	4	3
Never	87	84	18

Test whether the variables Smoking and Exercise are dependent. If so, interpret using standardized residuals.

Answer: Pearson's chi-square test statistic and p-value are calculated. Since the p-value is large the null (independence between smoking and exercise) is not rejected. However, 4 out of the expected 12 counts have expected values less than 5 so Pearson's chi-square is not a good approximation. Instead Fisher's exact test should be used, which also yields a large p-value.

```
> smoking=matrix(c(7,3,1,9,7,1,12,4,3,87,84,18),4,3,byrow=TRUE)
> dimnames(smoking)=list(c("Heavy","Regularly","Occasionally","Never"),
+ c("Frequently","Some","None"))
> c2=chisq.test(smoking,correct=FALSE);c2
Warning message:
In chisq.test(smoking, correct = FALSE) :
  Chi-squared approximation may be incorrect

Pearson's Chi-squared test

data:  smoking
X-squared = 5.4885, df = 6, p-value = 0.4828

> c2$expected
             Frequently      Some      None
Heavy          5.360169  4.567797  1.072034
Regularly      8.283898  7.059322  1.656780
Occasionally   9.258475  7.889831  1.851695
Never         92.097458 78.483051 18.419492
> c2$stdres
             Frequently        Some       None
Heavy          1.013066 -0.98246603 -0.0750004
Regularly      0.360707 -0.03030993 -0.5575547
Occasionally   1.312236 -1.88859760  0.9263282
Never         -1.662265  1.82488110 -0.2305460
> fisher.test(smoking)
Fisher's Exact Test for Count Data
data:  smoking
p-value = 0.4138
alternative hypothesis: two.sided
```