Inference Concerning 2 Means or Medians Problems

Q.1. A researcher wishes to compare two species of ferns with respect to chemical uptake. She samples 10 ferns of each species, and measures the amount of the chemical uptake in each of the 20 specimens. The summary statistics and relevant quantities are given below (she finds no evidence to believe the population variances are unequal):

$$\bar{y}_1 = 4.2 \;\; \bar{y}_2 = 2.8 \;\; s_p\sqrt{\frac{1}{10}+\frac{1}{10}} = 0.5$$

a) Give a 95% confidence interval for the difference in true (population) means for the 2 species.

b) The rank sums for the 2 species are T_1=140 and T_2=70, respectively. Use the normal approximation for the Wilcoxon Rank-Sum test to test whether the population means differ at the 0.05 significance level. Give the test statistic and give the P-value for your statistic.

$$n_1 = n_2 = 10 \quad df = 10+10-2 = 18 \quad t_{.025,18} = 2.101 \quad \bar{y}_1 = 4.2 \quad \bar{y}_2 = 2.8 \quad \hat{SE}\left\{\bar{Y}_1 - \bar{Y}_2\right\} = 0.5$$

p.1.a. 95% CI for $\mu_1 - \mu_2$: $(4.2-2.8) \pm 2.101(0.5) \equiv 1.4 \pm 1.05 \equiv (0.35, 2.45)$

p.1.b. $\mu_T = \dfrac{n_1(n_1+n_2+1)}{2} = \dfrac{10(21)}{2} = 105 \quad \sigma_T = \sqrt{\dfrac{n_1 n_2 (n_1+n_2+1)}{12}} = \sqrt{\dfrac{10(10)(21)}{12}} = 13.23$

TS: $z_{obs} = \dfrac{T_1 - \mu_T}{\sigma_T} = \dfrac{140-105}{13.23} = 2.65 \qquad P = 2P(Z \geq 2.65) = 2(.0040) = .0080$

Q.2. A study was conducted to measure the effect of a new training program for new employees at a large company. A sample of 9 new employees were selected and given a test regarding ethics in the workplace before and after a 1-day training session on ethics was given. For each employee, the difference between the scores (After – Before) was obtained. The mean and standard deviation of these differences were 15.0 and 9.0, respectively. Test to determine whether the training course is effective in increasing true mean scores:

a) Null Hypothesis: $\mu_D \leq 0$ Alternative Hypothesis: $\mu_D > 0$

b) Test Statistic: $s_d / n^{1/2} = 9 / 9^{1/2} = 3$ $t_{obs} = 15 / 3 = 5$

c) Rejection Region: $t_{.05,8} = 1.860$ RR: $t_{obs} > 1.860$

Do we conclude that the training course is effective in increasing true mean scores at the α=0.05 significance level? **Yes** or No

Q.3. A researcher samples 9 computer monitors from each of two manufacturers. The lifetime of each monitor is observed and recorded (in 1000s of hours). The following table gives the summary results:

Brand	Mean	SD
A	4.20	1
B	4.60	0.8

- Test H_0: $\mu_A = \mu_B$ versus H_A: $\mu_A \neq \mu_B$ at the $\alpha = 0.05$ significance level (assuming equal variances).
  - Test Statistic:

$$s_p^2 = \frac{(9-1)(1)^2 + (9-1)(0.8)^2}{9+9-2} = \frac{13.12}{16} = 0.82$$

$$\text{TS: } t_{obs} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{4.20-4.60}{\sqrt{0.82\left(\frac{1}{9}+\frac{1}{9}\right)}} = \frac{-0.40}{0.43} = -0.937$$

  - Rejection Region: $t_{.025,16} = 2.120$ RR: $|t_{obs}| \geq 2.120$
    - Conclusion (circle best answer):
      - Conclude $\mu_A > \mu_B$
      - Conclude $\mu_A < \mu_B$
      - **Do not reject H_0: $\mu_A = \mu_B$**

Q.4. A comparison is being made to compare the amounts of food eaten per week between two bird species. Samples of 12 males from each species were observed and the amount of food consumed by each bird in a one-week period was recorded. Due to a few extreme outliers, it was decided to use the Wilcoxon Rank-Sum test. The rank sums for species A and B were T_A=170 and T_B=130. Use the normal approximation to test whether the population medians differ.

H_0: $\text{Median}_A = \text{Median}_B$ H_A: $\text{Median}_A \neq \text{Median}_B$

(Note: Under H_0 the standard deviation of T_A = 17.32)

- Expected Value of T_A under null hypothesis $\mu_T = \frac{12(12+12+1)}{2} = 150$
-
- Test Statistic: $z_{obs} = \frac{170-150}{17.32} = 1.15$
-
- P-Value: $2P(Z \geq 1.15) = 2(.1251) = .2502$

Q.5. A study was conducted to determine whether consumers' attitudes toward a product changed after seeing the product placed (subtly) in a short film. 16 Subjects were given a list of brands, and asked to give a rating to each (pre-viewing). After seeing the film, they were asked again to rate the brands. Higher ratings mean stronger preference. The mean and standard deviation of the differences (post-pre) were: mean=2.4, SD=3.0. Give a 95% confidence interval for the population mean difference.

$$n = 16 \quad \bar{d} = 2.4 \quad s_d = 3.0 \quad \hat{SE}\{\bar{D}\} = \frac{s_d}{\sqrt{n}} = 0.75 \quad df = n-1 = 15 \quad t_{.025,15} = 2.131$$

$$95\% \text{ CI for } \mu_D: \quad 2.4 \pm 2.131(0.75) \equiv 2.4 \pm 1.60 \equiv (0.8, 4.0)$$

- Does the product placement appear to significantly increase attitude? **Yes** No

Q.6. A researcher samples 7 adult males from each of two species of squirrels and measures their body mass index . The following table gives the summary results:

Species	Mean	SD
A	7.20	2.4
B	5.70	3.2

- Test H_0: $\mu_A = \mu_B$ versus H_A: $\mu_A \neq \mu_B$ at the $\alpha = 0.05$ significance level (assuming equal variances).
  - Test Statistic: **t_{obs} = 0.99**
  - Rejection Region: **$|t_{obs}| \geq 2.179$**
  - Conclusion (circle best answer):
    - Conclude $\mu_A > \mu_B$
    - Conclude $\mu_A < \mu_B$
    - **Do not reject H_0: $\mu_A = \mu_B$**

Q.7. A comparison is being made to compare the amounts of food eaten per week between two bird species. Samples of 12 males from each species were observed and the amount of food consumed by each bird in a one-week period was recorded. Due to a few extreme outliers, it was decided to use the Wilcoxon Rank-Sum test. The rank sums for species A and B were T_A=165 and T_B=135. Use the normal approximation to test whether the population medians differ.

H_0: $Median_A$ = $Median_B$ H_A: $Median_A \neq Median_B$

(Note: Under H_0 the standard deviation of T_A = 17.32)

- Expected Value of T_A under null hypothesis **μ_T = 150**
- Test Statistic: **z_{obs} = 0.87**
- P-Value: **P = .3844**

Q.8. A study was conducted to determine whether consumers' attitudes toward a product changed after seeing the product placed (subtly) in a short film. 9 Subjects were given a list of brands, and asked to give a rating to each (pre-viewing). After seeing the film, they were asked again to rate the brands. Higher ratings mean stronger preference. The mean and standard deviation of the differences (post-pre) were: mean=3.6, SD=1.2. Give a 95% confidence interval for the population mean difference.

95% CI: (2.68,4.52)

- Does the product placement appear to significantly increase attitude? **Yes** No

Q.9. A forensic researcher samples 100 adult males and 100 adult females and measures each subject's right foot length (cm). The following table gives the summary results:

Gender	Mean	Std. Dev.
Female	23.80	1.1
Male	26.30	1.6

p.9.a. Test H_0: $\mu_F = \mu_M$ ($\mu_F - \mu_M = 0$) versus H_A: $\mu_F \neq \mu_M$ ($\mu_F - \mu_M \neq 0$) at the $\alpha = 0.05$ significance level.

p.9.a.i. Test Statistic: $z_{obs} = -12.88$

p.9.a.ii. Rejection Region: $|z_{obs}| \geq 1.96$

p.9.a.iii. Conclusion (circle best answer):

- Conclude $\mu_F > \mu_M$
- **Conclude $\mu_F < \mu_M$**
- Do not reject H_0: $\mu_F = \mu_M$

p.9.b.Compute a 95% Confidence Interval for $\mu_F - \mu_M$ **(-2.88,-2.12)**

Q.10. Random samples of 5 inter-arrival times of volcano eruptions are obtained for each of 2 volcanoes in South America and are given in the following table.

p.10.a. Assign ranks to the 10 inter-arrival times (Y, in years), where 1 is the smallest and 10 is the largest.

Volcano	Y	Rank
1	2.4	3
1	3.8	5
1	12.7	8
1	14.2	9
1	17.3	10
2	1.6	1
2	2.2	2
2	3.5	4
2	4.8	6
2	7.9	7

p.10.b. Compute the rank sums for each volcano.

p.10.b.i. **$T_1 = 3 + 5 + 8 + 9 + 10 = 35$**

p.10.b.ii. **$T_2 = 1 + 2 + 4 + 6 + 7 = 20$**

p.10.c. The rejection region for a 2-sided Wilcoxon Rank-Sum test ($\alpha = 0.05$) that the population means (medians) are equal is (for $n_1 = n_2 = 5$): Reject H_0 if $\min(T_1,T_2) \leq 17$. Can we conclude the true means (medians) differ for these volcanoes? Yes / **No**

Q.11. A study compared using procyandin B-2 obtained from apples with a placebo on terminal hair growth (hairs > 60μm) in men. The B-2 sample had 19 subjects, and the placebo sample had 10 subjects. The following summary data were obtained, where the response was increase in terminal hair over 6 months (negative increases are decreases).

	B-2	Placebo
Mean	1.99	-0.82
StdDev	2.58	3.40
n	19	10
RankSum	325	110

p.11.a. Obtain a 95% Confidence Interval for μ_{B2} - μ_P assuming population variances are equal.

$$s_p^2 = \frac{(19-1)(2.58)^2 + (10-1)(3.40)^2}{19+10-2} = \frac{223.86}{27} = 8.29 \qquad \hat{SE}\left\{\overline{Y}_1 - \overline{Y}_2\right\} = \sqrt{8.29\left(\frac{1}{19}+\frac{1}{10}\right)} = 1.12 \qquad t_{.025,27} = 2.052$$

$$\text{95\% CI: } \left(1.99-(-0.82)\right) \pm 2.052(1.12) \equiv 2.81 \pm 2.31 \equiv (0.50, 5.12)$$

p.11.b. Use the large-sample Wilcoxon Rank-Sum test to test whether Median Hair growth differs for B-2 and placebo.

H_0: $M_{B2} = M_P$ H_A : $M_{B2} \neq M_P$ Note: σ_T = 18.23

$$\mu_{T_{B2}} = \frac{19(19+10+1)}{2} = 285 \qquad \text{TS: } z_{obs} = \frac{325-285}{18.23} = 2.19 \quad \text{RR: } |z_{obs}| \geq 1.96 \qquad P = 2P\left(Z \geq 2.19\right) = 2(.0143) = .0286$$

p.11.c. This was a pilot study. Suppose they want to conduct a full-scale study, and a goal to be able to detect a difference of Δ = 1.00 with power = 1-β = 0.80 and α = 0.05 (2-sided test). How many people should be assigned to each treatment assuming σ^2 = 8 for each group?

$$\text{Based on large-sample z: } n \geq \frac{2\left(z_{\alpha/2}+z_\beta\right)^2 \sigma^2}{\Delta^2} = \frac{2\left(1.96+0.84\right)^2 8}{1^2} = 125.44 \approx 126 = n_1 = n_2$$

Q.12. A study was conducted to compare stress and anger among students from 2 cultures (Indian and Iranian). Samples of size 100 were collected among males and females from each culture. One (of many) scale reported was the Anger Expression Index among females. The sum of the ranks for n1 = 100 Indian females was 8241, and the sum of ranks for the n_2 = 100 Iranian females was 11859. Use the large-sample Wilcoxon Rank-Sum Test to test whether population median anger expression scales differ among Indian and Iranian female students.

Test Statistic = z_{obs} = -4.42 Rejection Region: $|z_{obs}| \geq 1.96$ P-value > **<** 0.05

Q.13. A researcher wishes to estimate the difference $\mu_1 - \mu_2$ within E=4.0 with 95% confidence. Past experience has found that the standard deviation of each of the populations is 20. How many measurements will they need to take from each group?

$$E = 4.0 \quad z_{\alpha/2} = 1.96 \quad \sigma \approx 20 \quad n = \frac{2z_{\alpha/2}^2\sigma^2}{E^2} = \frac{2(1.96)^2(20)^2}{4^2} = 192.08 \quad \Rightarrow \quad n_1 = n_2 = 192$$

Q.14. A study measured serum androstenedione levels in a sample of n = 10 women polycystic ovarian syndrome. Measurements were made pre- and post-laparoscopic ovarian drilling (see table below).

p.14.a. Use the paired t-test to test whether the true means differ. $H_0: \mu_D = 0$ vs $H_A: \mu_D \neq 0$

Test Statistic = **t_{obs} = -4.25** Rejection Region: **$|t_{obs}| \geq t_{.025,9}$ = 2.262** P-value > **<** 0.05

Table 1
Effect of laparoscopic ovarian drilling on serum androstenedione levels in women with polycystic ovarian syndrome

Subject	Serum androstenedione (nmol/L) Pre-operative	Post-operative	Difference
1	7.9	8.2	+0.3
2	11.2	7.1	−4.1
3	13.9	8.5	−5.4
4	8.8	6.4	−2.4
5	7.7	3.4	−4.3
6	6.9	5.8	−1.1
7	9.0	7.5	−1.5
8	8.5	4.4	−4.1
9	4.9	4.2	−0.7
10	16.1	9.7	−6.4
Mean	9.49	6.52	−2.97
Standard deviation	3.35	2.06	2.21

p.14.b. Compute T^+ and T^- for the Wilcoxon Signed-Rank Test:

T^+ = 1 (smallest abs(diff))

T^- 2 + … + 10 =(10(10+1)/2)-1 = 54

Q.15. A study was conducted, comparing reaction times among 2 groups of subjects: patients with cocaine dependence (PCD) and Healthy Controls (HC). The patients were measured for the Stop Signal Reaction Time (SSRT, in milliseconds). Results are given below. Compute a 95% confidence interval for $\mu_1 - \mu_2$, assuming population variances are equal.

$$\overline{y}_{PCD} = 241 \quad s_{PCD} = 49 \quad n_{PCD} = 54 \quad \overline{y}_{HC} = 220 \quad s_{HC} = 39 \quad n_{HC} = 54$$

$$s_p = \sqrt{\frac{(54-1)49^2 + (54-1)39^2}{54+54-2}} = 44.28 \qquad \sqrt{\frac{1}{54} + \frac{1}{54}} = 0.192$$

95% Confidence Interval for $\mu_1 - \mu_2$: **(241 – 220) ± 1.983(44.28)(0.192) ≡ 21 ± 16.9 ≡ (4.1 , 37.9)**

Q.16. A researcher selects a random sample of 15 individuals and asks each to **estimate** his or her own weight. She then measures each individual's **actual** weight. She uses the data to see if there is significant evidence that individuals tend to underestimate their weights. Give her appropriate null and alternative hypotheses, where $\mu_{diff} = \mu_{est} - \mu_{act}$:

$i)\ H_0 : \mu_{\text{diff}} \geq 0\ H_A : \mu_{\text{diff}} < 0 \qquad ii)\ H_0 : \mu_{\text{diff}} \leq 0\ H_A : \mu_{\text{diff}} > 0 \qquad iii)\ H_0 : \mu_{\text{diff}} = 0\ H_A : \mu_{\text{diff}} \neq 0$

Answer = *i*)

Q.17. A study was conducted to illustrate the effect of alcohol on an individual's ability to perform a physical task. Each of twenty subjects performed a test of dexterity, the time to perform the task is (WITHOUT). Each subject was then given two drinks containing 1.5 oz. Of alcohol, and the time to perform the task is (WITH). Output is shown below, where DIFF is the value of WITH minus WITHOUT. Set-up a 95% confidence interval for the mean increase in performance time after drinking alcohol (using numbers, not symbols).

Variable	N	Mean	Std. Dev.	95% CI
WITH	20	13.20	1.856	
WITHOUT	20	8.60	1.693	
DIFF	20	4.60	1.985	

$$n = 20 \quad \bar{d} = 4.60 \quad s_d = 1.985 \quad \hat{SE}\{\bar{D}\} = \frac{1.985}{\sqrt{20}} = 0.444 \quad t_{.025,20-1} = 2.093$$

$$95\%\ \text{CI for}\ \mu_D = 4.60 \pm 2.093(0.444) \equiv 4.60 \pm 0.93 \equiv (3.67, 5.53)$$

Q.18. A researcher is comparing the breaking strengths of two types of aluminum tubes. She samples 12 tubes of each type and measures pressure necessary to bend the tubes. Labelling μ_1 as the true mean for all tubes of type 1, and μ_2 as the true mean for all tubes of type 2, conduct the test of H_0: $\mu_1 - \mu_2 = 0$ vs H_A: $\mu_1 - \mu_2 \neq 0$ at the $\alpha = 0.05$ significance level, based on the following output (assume population variances are equal):

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
y	Equal variances assumed	1.534	.229	8.574	22	.000	8.87316	1.03485	6.72702	11.01930
	Equal variances not assumed			8.574	21.174	.000	8.87316	1.03485	6.72216	11.02416

Test Statistic **$t_{obs} = 8.574$**

Reject H_0 if the test statistic falls in the range(s) **$t_{obs} \leq -2.074$, $t_{obs} \geq 2.074$**

P-value **.000**

Conclude (Circle One) Do Not Conclude $\mu_1 - \mu_2 \neq 0$ $\mu_1 - \mu_2 < 0$ **$\mu_1 - \mu_2 > 0$**

Q.19. A researcher is interested in comparing the coloring of 2 automotive painting machines. She samples *n*=12 sheets of aluminum, and she cuts each sheet into 2 parts. One part was painted with machine A, the other part was painted with machine B. A measure of paint quality was obtained from each half of each sheet. Data are given below.

Sheet	Machine A	Machine B	Diff (A-B)	rank(|diff|)	
1	61.42	62.00	-0.58	2	
2	57.64	60.63	-2.99	11	
3	60.66	59.02	1.63	6	
4	57.75	57.06	0.70	3	
5	57.85	56.05	1.80	7	
6	59.99	61.13	-1.14	5	
7	66.90	64.24	2.66	9	
8	56.45	62.39	-5.94	12	
9	57.83	56.79	1.04	4	
10	62.75	60.10	2.65	8	
11	53.48	53.33	0.14	1	
12	57.69	60.44	-2.75	10	
Mean	59.20	59.43	-0.23		
Std Dev	3.44	3.10	2.60		

p.19.a. Compute a 95% Confidence Interval for the difference between the true mean scores for the 2 machines ($\mu_1 - \mu_2 = \mu_D$) by completing the following parts:

p.19.a.i. (Point) Estimate: **-0.23** p.6.a.ii. Estimated Std. Error **$2.60 / (12)^{1/2} = 0.751$**

p.19.a.iii. Critical t-value **$t_{.025,11} = 2.201$** p.6.a.iv. 95% CI: **$-0.23 \pm 2.201(0.751) \equiv -0.23 \pm 1.65 \equiv (-1.88, 1.42)$**

p.19.b. Compute T^+ and T^- for the Wilcoxon Signed-Rank test

$T^+ = 6 + 3 + 7 + 9 + 4 + 8 + 1 = 38$ **$T^- = 2 + 11 + 5 + 12 + 10 = 40$**

p.19.c. For a 2-sided ($\alpha = 0.05$) test, we reject the null hypothesis of equal medians if $\min(T^+, T^-) \leq 13$. Do we reject the null hypothesis? Yes / **No**

Q.20. A study was conducted in n=30 subjects to determine the effect of an ultrasound treatment for carpal tunnel syndrome. Each subject received an active ultrasound on one wrist, and a sham ultrasound (placebo) on their other wrist. The following table gives the mean and standard deviation of the response: change in hand grip strength during the study, for each treatment, and for the difference within subjects.

	Mean	SD
Active	3.87	5.16
Placebo	-0.09	5.56
Difference	3.96	5.59

p.20.a. Compute a 95% Confidence Interval for $\mu_A - \mu_P$. **$3.96 \pm 2.09 \equiv (1.87, 6.05)$**

p.20.b. From this Interval, the P-value for test $H_0: \mu_A - \mu_P = 0$ vs $H_A: \mu_A - \mu_P \neq 0$ is $<$ or $>$ 0.05?

Q.21. Two researchers analyze the same set of observations from 2 samples of equal sample sizes ($n_1 = n_2 = n$).

One researcher uses the independent sample t-test, based on equal variances. The other researcher uses the independent sample t-test, based on unequal variances. Choose the correct answer:

- Their test statistics will be the same, their degrees of freedom will be the same.
- **Their test statistics will be the same, their degrees of freedom will be different.**
- Their test statistics will be different, their degrees of freedom will be the same.
- Their test statistics will be different, their degrees of freedom will be different.

Q.22. Among 2 large populations (Males and Females) who completed the Rock and Roll marathon in 2015, the **population** means and standard deviations of velocities (miles per hour) were:

$$\mu_M = 6.34 \qquad \sigma_M = 1.06 \qquad \mu_F = 5.84 \qquad \sigma_F = 0.83$$

Suppose you simultaneously took many random samples of size $n_M = n_F = 20$ from each population, and for each pair of random samples, you computed $\overline{Y}_M - \overline{Y}_F$ and saved each difference.

p.22.a. What would you expect the mean of the $\overline{Y}_M - \overline{Y}_F$ values to be. **0.50**

p.22.b. What would you expect the standard deviation of the $\overline{Y}_M - \overline{Y}_F$ values to be. **0.301**

p.22.c. Between what 2 bounds would you expect 95% of the $\overline{Y}_M - \overline{Y}_F$ values to lie between?

0.50 ± 2.024(0.301) ≡ 0.50 ± 0.609 ≡ (-0.109 , 1.109)

Q.23. Two models of video cameras are being compared for detecting animals in a wildlife setting. The cameras will film the same locations in fixed time intervals in a paired difference experiment. The parameter μ_D is the population mean difference across all possible locations in the fixed time intervals. From a pilot study, it is believed $\sigma_d = 5$. How many samples will be needed if we wish for the margin of error in estimating μ_D within E = 0.5 with 95% Confidence?

$$n = \frac{(1.96)^2 5^2}{(0.5)^2} = 384$$

Q.24. An experiment is conducted to compare breaking strengths of 2 types of fibers. The means, standard deviations, and sample sizes of random samples from each fiber type are: $\overline{y}_{1\bullet} = 50 \quad s_1 = 12 \quad n_1 = 10 \quad \overline{y}_{2\bullet} = 45 \quad s_2 = 8 \quad n_2 = 10$

Assume $\sigma_1^2 = \sigma_2^2$, Test $H_0: \mu_1 - \mu_2 = 0$ versus $H_A: \mu_1 - \mu_2 \neq 0$

$$\sqrt{\frac{(10-1)12^2 + (10-1)8^2}{10+10-2}} = 13.27$$

Test Statistic: **t_{obs} = 0.843** Reject H_0 if Test Statistic **< -2.101 or > 2.101**

Q.25. A study is conducted to compare lifetimes of 2 brands of light bulbs. Random samples of $n_1 = n_2 = 20$ are obtained from each manufacturer. Due to the highly skewed distributions of lifetimes, the large-sample Wilcoxon rank-sum test was used. The investigators tested each bulb, measuring its lifetime, and ranked all of the N = 40 bulbs. The rank sum for the two brands are $T_1 = 460$ and $T_2 = 360$. Complete the following parts to test $H_0: M_1 = M_2$ $H_A: M_1 \neq M_2$

p.25.a. Compute μ_T: **20(40+1)/2 = 410**

p.25.b. Compute σ_T: **$(20(20)(41)/12)^{1/2} = 36.97$**

p.25.c. Test Statistic: **$z_{obs} = (460 - 410) / 36.97 = 1.35$**

p.25.d. Rejection Region: **$|z_{obs}| \geq 1.96$** p.26.e. Reject H_0? Yes or **No**