Stat 6304 Statistical Inference I
Fall 2015
Prof. Fan

Name:______________________
(print: first last)
Section #: _____ NetID #:______________

MIDTERM TWO
100 minutes

Instructions: This is a closed-book but 1 two-sized help sheet of letter size is allowed. You may use Minitab/R and a calculator (cell-phones may NOT be used as a calculator). The total score is 100 points. Except multiple-choice questions, **give concise but detailed answers for full credit.**

1. [5 points] Which statement is correct about a p-value?
   A. The smaller the p-value the stronger the evidence against the alternative hypothesis
   B. The smaller the p-value the stronger the evidence against the null hypothesis
   C. Whether a small p-value provides sufficient evidence against the alternative hypothesis depends on whether the test is one-sided or two-sided.
   D. Whether a small p-value provides sufficient evidence against the null hypothesis depends on whether the test is one-sided or two-sided.

Answer: B

2. [5 points] A 95% confidence interval for the difference between the mean handspans of men and the mean handspans of women is determined to be 2.7 centimeters to 3.3 centimeters. Which of the following statements is the best interpretation of this interval?
   A. It is likely that the difference in the population mean handspans of men and women is covered by the interval 2.7 centimeters to 3.3 centimeters.
   B. It is likely that the difference in the sample mean handspans of men and women is covered by the interval 2.7 centimeters to 3.3 centimeters.
   C. For 95% of all possible pairs of a man and a woman, the man's handspan is between 2.7 and 3.3 centimeters longer than the woman's handspan.
   D. For 95% of all possible samples (of the same size), the difference in the sample mean handspans of men and women is covered by the interval 2.7 centimeters to 3.3 centimeters.

Answer: A

3. [8 points] A randomly selected sample of n =51 men in Brazil had an average lifespan of 59 years. The standard deviation was 10 years. Calculate AND interpret a 98% confidence interval for the average lifespan for all men in Brazil.

Answer:
You can use one sample t interval: (55.63, 62.37) or one sample Z interval is (55.74, 62.26).

We are 98% confident that a Brazil man lives 55.6 to 62.4 (or 55.7 to 62.3) years on average.

4. [5 points] A cola-tasting experiment is conducted in which 100 people are each given a half-cup of Cola A and a half-cup of Cola B. All 100 participants taste both colas and then rate the taste of each cola on a 0 (horrible) to 50 (tastes great) scale. For each person, a coin is tossed to determine the order of tasting the two colas. To estimate the mean difference between ratings of the two products, what procedure should be used?
   A. A confidence interval for one proportion.
   B. A confidence interval for a difference between two proportions.
   C. A paired *t* (or 1-sample *t*) confidence interval for a mean difference.
   D. A 2-sample *t* confidence interval for a difference in two means.

Answer: C

5. [5 points] The distinction between a sampling distribution and a confidence interval is:
   A. A confidence interval gives possible values for a sample statistic when the population parameter is assumed known, while a sampling distribution gives possible values for a population parameter when only a single value of a sample statistic is known.
   B. A sampling distribution gives possible values for a sample statistic when the population parameter is assumed known, while a confidence interval gives possible values for a population parameter when only a single value of a sample statistic is known.
   C. Sampling distributions exist only for situations involving means, while confidence intervals can be computed for situations involving means and proportions.
   D. Confidence intervals exist only for situations involving means, while sampling distributions can be computed for situations involving means and proportions.

Answer: B

6. [6 points] A 95% confidence interval for the difference in the population proportions of men (group 1) and women (group 2) who exercise regularly, is –0.06 to +0.10. From this confidence interval, can we conclude: (No explanation required)

   A. men spend 10% more time exercising, on average, than do women? Yes __ No _X_
   B. the difference in proportions of men and women who exercise regularly is not significant? Yes _X_ No __
   C. the difference in proportions of men and women who exercise regularly is at least 0.05? Yes __ No _X_

7. A company studied two programs for compensating its sales staff. In program A, salespeople were paid a higher salary, plus a small commission for each item they sold. In program B, they were paid a lower salary with a larger commission. Following are the amounts sold, in thousands of dollars, for each salesperson on each program.

	Salesperson								
Program	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
A	55	22	34	22	25	61	55	36	68
B	53	24	36	28	31	61	58	38	72

(a) If the same random sample of nine salespeople were observed under program A and B, answer the following:
   i. [8 points] Perform the most powerful test to verify that the mean sales differ between the two programs. Use the level of significance of 0.05. Must state your 1) hypotheses 2) test statistic 3) p-value or rejection region 4) test result in the context of the situation.
   ii. [8 points] Construct a 95% confidence interval of the mean sales difference. According to the confidence interval, which program should the company adopt (with higher amount of sales)?
   iii. [8 points] What assumption(s) do you make in the t test? Conduct test(s) to check the assumption(s). For each test, must state your p-value and test result in the context of the situation.
   iv. [6 points] It is desired to detect, at .05 significance level, the difference between the two programs if one program can sale at least three more than the other program on average with probability of .90. What is the minimal required sample size?

(b) [30 points] If a different random sample of nine salespeople were selected for each program independently, repeat i. to iv. of part (a).

(c) [6 points] When the sample size is so small (only 9), people may not want to assume any particular shape for the population distribution. Under this situation, describe what methods (for hypothesis tests and estimation) you would use to analyze the difference in the amounts sold by the two programs.

Answer:

(a) The samples are paired; paired sample t test is one-sample t test of difference (= sales of program A - sales of program B).

Minitab: Paired T-Test and CI: A, B

```
Paired T for A - B

             N    Mean  StDev  SE Mean
A            9   42.00  17.90     5.97
B            9   44.56  16.85     5.62
```

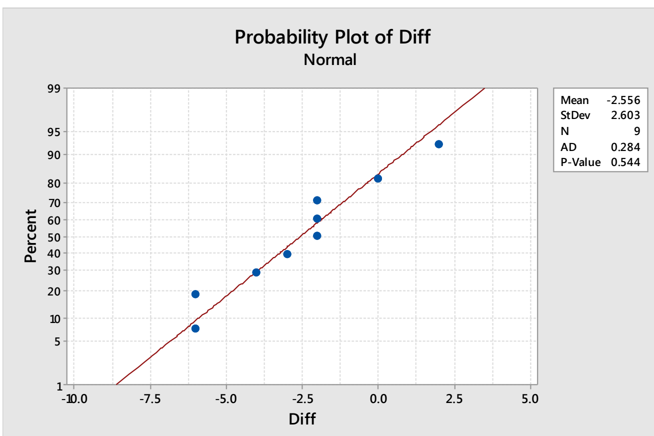
```
Difference  9  -2.556  2.603    0.868


95% CI for mean difference: (-4.557, -0.554)
T-Test of mean difference = 0 (vs ≠ 0): T-Value = -2.94  P-Value = 0.019
```

i. Let μ be the mean of differences. Conduct the one sample t test for
   - Ho: μ = 0 vs. Ha: μ ≠ 0.
   - $t = \dfrac{-2.556}{2.603/\sqrt{9}} = -2.945,$
   - Rejection region: t is larger than 2.306 the critical value of df =8 at α/2=.025, or by p-value is 0.019 < 0.05
   - so we conclude that there is a significant difference in mean sales of the two programs at 5% significant level.

ii. The 95% C. I. is -2.556 ± 2.306(2.603/√9) = (-4.557, -0.554).
This interval is strictly below zero, so the mean sales of program A is lower than that of program B. We should adopt program B

iii. The assumption of one-sample t test is the distribution of difference follows a normal curve. To verify this, we can draw a normal probability plot and conduct a normality test of difference:



P-value 0.544 is large (larger than .10), so normality of difference is a reasonable assumption.

iv.

Power and Sample Size

```
Paired t Test

Testing mean paired difference = 0 (versus ≠ 0)
Calculating power for mean paired difference = difference
α = 0.05  Assumed standard deviation of paired differences = 2.6

            Sample  Target
Difference    Size   Power  Actual Power
         3      11     0.9      0.930553
```

We need at least 11 paired samples.

(b) The samples are independent; pooled two sample t procedure is the most powerful.

Minitab: Two-Sample T-Test and CI: SALES, PROGRAM

```
Two-sample T for SALES

PROGRAM  N  Mean  StDev  SE Mean
A        9  42.0   17.9      6.0
B        9  44.6   16.9      5.6


Difference = μ (A) - μ (B)
Estimate for difference:  -2.56
95% CI for difference:  (-19.93, 14.82)
T-Test of difference = 0 (vs ≠): T-Value = -0.31  P-Value = 0.759  DF = 16
Both use Pooled StDev = 17.3857
```

i. Let μ_A be the mean sales of program A and μ_B be the mean sales of program B.

- Ho: $\mu_A = \mu_B$ vs. Ha: $\mu_A \neq \mu_B$.

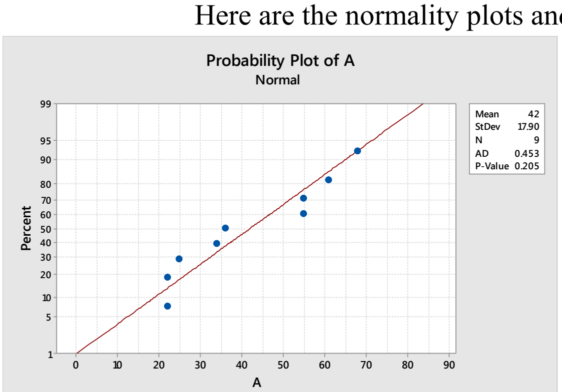
The pooled sample standard deviation is

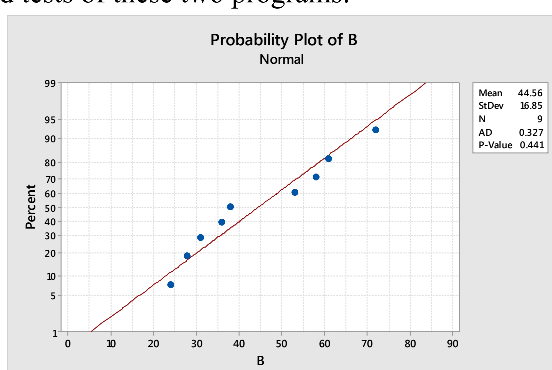
$$\sqrt{(17.903^2 \times 8 + 16.853^2 \times 8)/16} = 17.3857.$$

- $t = \dfrac{-2.556}{17.386\sqrt{2/9}} = -0.31,$
- Rejection region: t is less than 2.120 the critical value of df =9+9-2=16 at $\alpha/2$=.025 or by p-value is .759 > .05,
- so we CANNOT conclude that there is a significant difference in mean sales of the two programs at 5% significant level.

ii. The 95% C. I. is -2.556 ± 2.120(17.386√2/9) = (-19.93, 14.82).
This interval does include zero, so program A is not significantly different to program B on average. We could adopt either program.

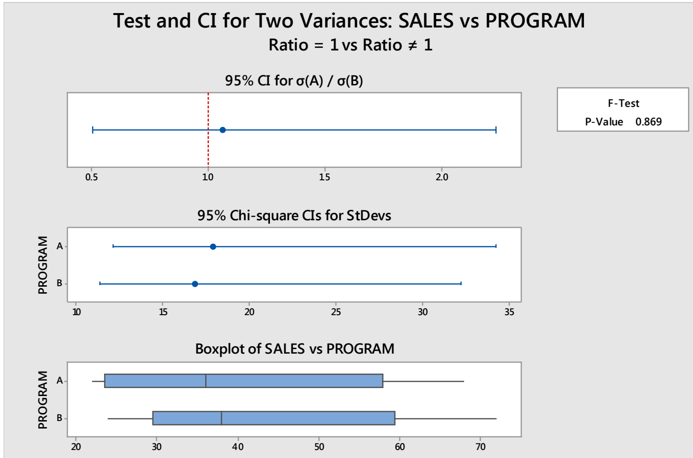
iii. The assumptions of pooled two-sample t test include
1) Both distributions of sales for the two programs are normal
Here are the normality plots and tests of these two programs:





The p-value for program A is .205 and for program B is .441; both are larger than .05, so we can conclude both distributions are not significantly different from normal distributions.

2) The two distributions have the same variances
Here is the F test of equal variance for the two programs:



p-value is .869, so we can assume equal variance of the two distributions.

iv. The required sample size is 707 (with S.D. = 17.3857) or 708 (with S.D.=17.39).

Power and Sample Size

```
2-Sample t Test

Testing mean 1 = mean 2 (versus ≠)
Calculating power for mean 1 = mean 2 + difference
α = 0.05  Assumed standard deviation = 17.39


             Sample  Target
Difference     Size   Power  Actual Power
         3      708     0.9      0.900367

The sample size is for each group.
```

(c) Without assuming normality,

1) We can still use the same formulas but the critical values of t statistics should be estimated by bootstrap methods.
2) Or we can use non-parametric procedures: Use Wilcoxon Rank Sum (or Mann-Whitney) procedure for independent samples and Wilcoxon signed rank (or 1-sample Wilcoxon) procedure for paired sample.