10
16
9

STAT:2010/4200, Statistical Methods and Computing
Spring 2019, Instructor: Cowles
Midterm 3

Name: Solutions Section

1. A professor asked her sophomore students, "How many drinks do you typically have per session? (A drink is defined as one 12-ounce beer, one 4-ounce glass of wine, or one 1-ounce shot of hard liquor." Some of the students didn't drink. From the students who did drink, the professor obtained numeric data from 17 female students and 23 male students. She is tentatively willing to regard the students as a simple random sample of sophomore students are her college. She wishes to compare mean numbers of drinks in the populations of female students and male students at her college.

   (a) The professor's data is best described as (circle one):

   2/2

      i. single sample
      ii. paired sample
      [✓ iii. two-independent sample]
      iv. none of the above

   (b) The statistical test best suited to the professor's analysis is:

   2/2

      i. z test
      [✓ ii. t test]
      iii. chi-square test
      iv. ANOVA
      v. none of the above

   0 for any other

   Briefly justify your choice.

   3/3 We don't know population standard deviation, so z test is out. Chi square test is for proportions, not means. ANOVA is for 3 or more populations.

2. You wish to study whether the proportions of women in three occupations (accountant, actuary, graphic designer) are equal. You randomly sample local companies and obtain the following employment data:

3/3

	women	men	
accountants	17	26	43
actuaries	9	25	34
graphic designers	6	7	13
	32	58	90

   (a) In the table above, fill in the margins to complete the contingency table.

10

(b) Compute the expected count for the number of women graphic designers under the null hypothesis of equal population proportions in all three occupations. (Numeric answer; show your work.)

3 $\frac{32}{90} \times 13 = 4.62$

(c) The test statistic value is 2.159. At significance level 0.05, should we reject the null hypothesis of equal population proportions? (yes/[✓ no]) Briefly describe how you got your answer.

3 Compare to chi square distribution with 2 degrees of freedom. $p > 0.25$.

3. The Business Opportunities Handbook gives data on business startup costs in thousands of dollars for five types of businesses (pizza parlors, bakeries, shoe stores, gift shops, and pet stores).

These sample data can be used to test whether the population means of startup costs are equal in these 5 types of businesses. SAS output for this problem is attached.

(a) Why is ANOVA a better choice for this data analysis than the Chi square test?

2/2 We are concerned with population means of quantitative variable. Chi square is for proportions.

(b) List two assumptions that must be met in order for the results of ANOVA to be trustworthy. For each one, describe whether the SAS output suggests that the assumption is met.

Distributions of variable in all populations are normal. Boxplots don't give evidence of violations.

5/5

Population standard deviations are equal. From proc means, the largest sample std dev is less than twice as large as the smallest one.

(c) Write the null hypothesis that will be tested. Use conventional symbols.

3 $H_0: \mu_{pizza} = \mu_{bakeries} = \mu_{shoes} = \mu_{gifts} = \mu_{pet}$

16

(d) We wish to test the null hypothesis at significance level alpha = 0.05. Give the numeric values of the test statistic and the p-value (from the SAS output).

3

$F = 3.25$

p-value = 0.0184

(e) Should we reject H_0? ([✓ yes]/no) Briefly explain.

2

$p = 0.184 < \alpha = .05$

(f) According to the SAS output, which pair or pairs of population means are different?

2

$\mu_{pet} \neq \mu_{bakery}$

(g) In the following line from the ANOVA output, what quantity are we 95% confident falls in the interval? (Circle one)

2

```
                      Difference      Simultaneous
        type             Between     95% Confidence
     Comparison            Means         Limits

     gifts   - pet        35.38       -3.75    74.50
```

i. $\bar{x}_{gifts} - \bar{x}_{pet}$

[✓ ii. $\mu_{gifts} - \mu_{pet}$]

iii. μ_{gifts}

iv. μ_{pet}

v. None of the above

9

```
                   The UNIVARIATE Procedure
                       Variable:  cost

                       Schematic Plots

     |
 160 +            |
     |            |
     |            |           |
     |            |           |
 140 +            |           |                       |
     |            |           |                       |
     |            |           |                       |
     |            |           |                       |
 120 +         +-----+        |                       |
     |         |     |        |                       |
     |         |     |        |           |        +-----+
     |         |     |        |           |        |     |
 100 +         |     |     *-----*        |        |     |
     |         |     |     |     |        |        |     |
     |         |  +  |     |     |        |        |     |
     |         *-----*     |  +  |        |        |  +  |
  80 +         |     |     |     |        |        *-----*
     |         |     |     |     |     +-----+     |     |
     |         |     |     |     |     |     |     |     |
     |         |     |     |     |     |     |     |     |
  60 +         +-----+     |     |     |     |     +-----+
     |            |        |     |     |     |        |
     |            |        +-----+     *--+--*        |
     |            |           |        |     |        |
  40 +            |           |        |     |        |
     |                        |        |     |        |
     |                                 +-----+
     |                                    |
  20 +                                    |
      ------------+-----------+-----------+-----------+-----------
type            bakery       gifts         pet       pizza
```