STAT 350 Exam 2 Name: Solutions

Part 1: Multiple Choice and Short Answer (4 points each)

1) The one-factor analysis of variance (ANOVA) is used primarily to test statistical hypotheses concerning \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

1. variances
2. means
3. medians
4. standard deviations

2) In a one-way ANOVA, suppose there are four treatments with n1 = 5, n2 = 6, n3 = 5, and n4 = 4. Then, the total degrees of freedom would equal \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

1. 19 5+6+5+4 = 20 total observations so df =19
2. 20
3. 79
4. 80

3) In a one-way ANOVA, MSG is the mean square for the groups (or treatments) and MSE is the mean square for error. The null hypothesis of equal population means is rejected if:

1. MSG is much smaller than MSE
2. MSG is much larger than MSE the test statistic = MSG/MSE so large test statistic g

generally lead to rejection of the null

1. MSG is equal to MSE
2. The null hypothesis is always rejected in a one-way ANOVA

**Use the following ANONVA table to answer questions #4-#8**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | df | SS | MS | F |
| A | 3 | 2.25 | 0.75 |  |
| B | 1 | 0.95 | 0.95 |  |
| A\*B | 3 | 1.20 | 0.30 |  |
| Error | 16 | 2.1 | 0.1313 |  |
| Total | 23 | 6.5 |  |  |

4. Factor A has \_\_\_\_4\_\_\_\_\_\_\_\_\_ levels.

5. How many total observations were collected? \_\_\_\_\_\_\_\_\_\_\_\_\_\_24\_\_\_\_\_\_\_\_

6. If the researchers used a balanced design, how many replications were conducted in each treatment combination?

There are 8 possible treatment combinations and 24 total observations so if the design is balanced there would be three replications per group.

7. What are the degrees of freedom the researchers would use to determine if there was any interaction between these two factors?

a) 3,16

b) 3,23

c) 3,3

d) 16,3

8. Compute the F-ratio to test for the significance of Factor A?

.75/.1313=5.71

**Use for #9-#10.** A study was conducted to test the effectiveness of supermarket sales strategies. At one supermarket, the price level (regular, reduced price, and at cost to supermarket) and display level (normal display space, normal display space plus end of aisle, and twice the display space) were tested to determine if they had any effect on the weekly sales of a particular supermarket product. Each of the combinations of price level and display level were put in place for a randomly selected week and the weekly sales of the product were recorded. Each combination was used three times over the course of the experiment. The results of the study are shown below.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Regular Price** | **Price Reduced** | **At Cost** |
| **Normal Display** | 999  1025  1030 | 1211  1215  1182 | 1577  1559  1598 |
| **Normal Plus End** | 1191  1233  1221 | 1860  1910  1926 | 2492  2527  2511 |
| **Twice Normal** | 1226  1202  1180 | 1516  1501  1498 | 1801  1833  1852 |

ANOVA Table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | DF | SS | MS | F |
| Display | 2 | 1691393 | 845696 |  |
| Price | 2 | 3089054 | 1544527 |  |
| Display\*Price | 4 | 510705 | 127676 |  |
| Error | 18 | 8905 | 495 |  |
| Total | 26 | 53000057 |  |  |

9. Identify the treatment groups used in this experiment.

a) The three price levels used by the supermarket

b) The three display levels used by the supermarket

c) The nine combinations of price level and display level used by the supermarket

d) The weekly sales collected for each of the weeks.

10. Find the test statistic for determining whether the factor of Display is significant.

a) 0.2499

b) 495

c) 257.93

d) 1708.48 = 845696/495

Independent random samples were selected from each of two normally distributed populations. The sample sizes, means and variances are shown in the table below. Use this information to answer questions #11-#12.

|  |  |
| --- | --- |
| **Sample 1** | **Sample 2** |
| n=16 | n=25 |
| Mean = 22.5 | Mean = 28.2 |
| Variance = 2.87 | Variance = 9.85 |

11. The researchers want to test the following: Ho: versus Ha: . What is the value of the test statistic? **Show work below**.

9.85/2.87 = 3.43

12. Use the F table (at the 0.05 level of significance) to determine if they should reject or fail to reject the null hypothesis. **Show work below to include value from the F-table**.

F(24,15) @5% = 2.29

As the test statistic of 3.43 is greater than the critical value of 2.29 we reject the null hypothesis.

13) Three hundred and seven (307) diamonds were sampled and randomly sorted into three certification groups (HRD, GIA, or IGI) that certify the appraisal of diamonds. A study was conducted to determine if the average size of diamonds reported by these three certification groups differ. A completely randomized design was used and the resulting ANOVA table is shown below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | DF | SS | MS | F |
| Treatment | 2 | 8.32652 | 4.16326 | 82.93 |
| Error | 304 | 15.2604 | 0.0502 |  |
| Total | 306 | 23.5869 |  |  |

Complete the ANOVA table using the information above.

**STAT 350 Exam 2 – Part 2 Extended Response**

**Answer each of the following three questions. Each is worth 16 points each**

1. For centuries, people looked at the full moon with some trepidation. From stories of werewolves coming out, to more crime sprees, the full moon has gotten the blame. Some researchersin the early 1970s set out to actually study whether there is a “full-moon” effect on the mental health of people. The researchers collected admissions data for the emergency room at a mental health hospital for 12 months. They separated the data into **rates before the full moon** (mean number of patients seen 4–13 days before the full moon), **during the full moon** (mean number of patients seen on the full moon day), and **after the full moon** (mean number of patients seen 4–13 days after the full moon)..

i) What is the response? Admissions data from ER at mental hospital

ii) Name the factor and its number of levels.

Factor: Moon phase Levels: three levels (before, during and after)

iii) State the null and alternative hypotheses for the researchers (using words or symbols but define all symbols used)

Ho: The mean number of patients admitted to the ER of a mental hospital is the same for all three phases of the moon OR where 1 is mean number before, 2 is mean number during and 3 is mean number after full moon

Ha: The mean number of admits differs for at least two of these moon phases.

Use the ANOVA table below to answer parts (iv) and (v).

| **Source** | **DF** | **SS** | **MS** | **F-Stat** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| Moon | 2 | 41.5291 | 20.756944 | 1.1741 | 0.3217 |
| Error | 33 | 583.4025 | 17.678864 |  |  |
| Total | 35 | 624.91639 |  |  |  |

iv) Do the data provide sufficient evidence to indicate a difference among the population means? Test at Assume all assumptions for inference have been met. Justify your answer.

No since the p-value of .3217 is greater than the significance level of 0.05.

v) Write a brief summary of the finding in part iii.

There is no evidence to suggest that the mean number of patients admitted to the ER of a mental hospital is different for the three phases of the moon, at =0.05.

2. Studies conducted at the University of Melbourne indicate that there may be a difference between the mean pain thresholds of blonds and brunettes. Men and women of various ages were divided into four categories according to hair color. The purpose of the experiment was to determine whether hair color is related to the amount of pain evoked by common types of mishaps and trauma. Each person was given a pain threshold score based on his or her performance in a pain sensitivity test (where higher scores = higher pain tolerance). The data is shown in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Light Blond** | **Dark Blond** | **Light Brunette** | **Dark Brunette** |
| 62 | 63 | 42 | 32 |
| 60 | 57 | 50 | 39 |
| 71 | 52 | 41 | 51 |
| 55 | 41 | 37 | 30 |
| 48 | 43 | 43 | 35 |
| Mean = 59.2 | Mean=51.2 | Mean=42.6 | Mean = 37.4 |

i) Given that x2 = 47,700 and G2 = 906,304 compute the **SS(Total**). Show the formula and all calculations.

SS(total) = 47,700 – (906,304/20) = 2384.80

| **Source** | **DF** | **SS** | **MS** | **F-Stat** | **P-value** |
| --- | --- | --- | --- | --- | --- |
| Color | 3 | 1382.80 | 460.93333 | 7.3602129 | 0.0026 |
| Error | 16 | 1002 | 62.625 |  |  |
| Total | 19 | 2384.80 |  |  |  |

ii) Now complete the value for the SS(Color) by subtraction.

iii) As the ANOVA results indicate that there is a difference in mean pain tolerance based on hair color, use Fisher’s LSD to determine which groups vary. Show all work.

LSD = 2.120= 10.61

There is results of the six comparisons show that:

LgBl>LgBr and DkBr but not from DkBl

DkBl>DkBr but not from LgBr

LgBr is not different from DkBr

iv) Write a couple of sentences to describe the findings of this experiment.

Light blond mean pain tolerance is greater than all the brunette groups while dark blond mean tolerance varies only from the dark brunettes. There is no difference in between the brunettes or the blondes in mean pain tolerance.

3. A company has developed a new digital camera and is faced with the problem of advertising the new camera. One factor deals with the advertising approach, to focus on either price or quality, and the second factor of interest is the advertising medium to use, radio, newspaper or internet. The response variable is the number of weekly sales for each of three weeks at each treatment combination.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Radio** | **Newspaper** | **Internet** |
| **Price** | 15  20  17 | 30  32  35 | 25  28  22 |
| **Quality** | 17  20  13 | 25  27  22 | 35  37  40 |
|  | C= 102 | C=171 | C=187 |

**i) State the three null and alternative hypotheses for this test – use words not symbols.**

Ho: The factors of advertising approach and medium do not interact to affect the response mean

number of weekly sales.

Ha: The factors of advertising approach and medium do interact to affect the response mean

number of weekly sales.

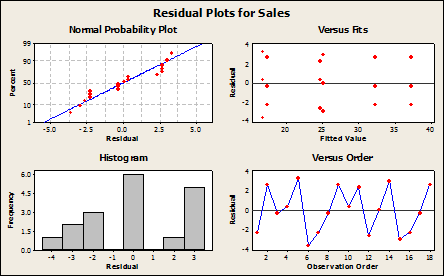
Ho: There is no difference between the two advertising approach on mean weekly sales

Ha: There is a difference between the two advertising approach on mean weekly sales

Ho: There is no difference between the three advertising mediums on mean weekly sales

Ha: There is a difference between the three advertising mediums on mean weekly sales

**ii) Check the assumptions for this test are satisfied by using the plot below.**

1) Samples are selected randomly and independently from the respective populations - assumed

2) All *k* population variances are equal - residual vs fitted order plot shows even scatter AND

3) All *k* population probability distributions are normal. – the linear pattern in the normal probability plot of residuals indicates that is condition is met.

**iii) A partial ANOVA table is shown below. Calculate the SS (medium) using the computational formula given the information in the table and that  and G = 460. You can complete the reminder of the table using subtraction.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **df** | **SS** | **MS** | **F-Stat** | **P-value** |
| Approach | 1 | 8 | 8 | 1.029 | 0.3305 |
| Medium | ??=2 | ?? = 680.11 | 340.05 | 43.72 | <0.0001 |
| Interaction | 2 | ?? = 309 | 154.5 | ??=19.86 | 0.0002 |
| Error | 12 | 93.33 | 7.78 |  |  |
| Total | ??=17 | 1090.44 |  |  |  |

Work for SS(medium) = 

iv) Use the results from the **ANOVA table above and interactions plot below** to write a brief report to the advertising manager.



Brief report:

As the null hypothesis of no interaction is rejected (p-value <0.002), there is evidence that the factors of advertising approach and medium do interact to affect the response mean number of weekly sales. The results of the means plot above indicate that the highest mean weekly sales occurs when stressing quality on the internet while stressing price is better if using the newspaper. The radio produces the lowest weekly mean sales.