Minitab

Science, Statistics

CA #3 –Chapter 13 - Using Statistical Software Package Use Minitab to calculate the appropriate hypothesis test for the given problems, paste the corresponding Minitab output from the session window for each problem. State the null and alternative hypotheses, state α, state the equation of the test statistic, check assumptions, state test statistic, and state p-value. Do not perform any calculations; MINITAB has done them for you! Just find the values of the test statistic and the p-value on your Minitab output. State your conclusion using words relevant to the problem.

Assignments will be graded using the following rubric:

1. _____ Word file attached in BB – 5 pts

2. _____ Minitab project file attached in BB – 5 pts

Problem 1

1. _____ Hypotheses Stated – 5 pts

2. _____ Level of Significance Stated – 1 pt

3. _____ Assumptions Verified – 5 pts

4. _____ Test Statistic Stated – 2 pts

5. _____ P-value Stated – 3 pts

6. _____ Conclusion – 4 pts

7. _____ Interpretation clearly stated in context of the problem. – 5 pts

8. _____ Minitab output in Word document (5 pts)

Problem 2

Part a)

9. _____ Part a) – 5 pts

Part b)

10. _____ Hypotheses Stated – 5 pts

11. _____ Level of Significance Stated – 1 pt

12. _____ Assumptions Verified – 5 pts

13. _____ Test Statistic Stated – 2 pts

14. _____ P-value Stated – 3 pts

15. _____ Conclusion – 4 pts

16. _____ Interpretation clearly stated in context of the problem. – 5 pts

17. _____ Minitab output in Word document – 5 pts

Part c)

18. _____ Standardized Residual Analysis – 10 pts

19. _____ Deductions

20. _____ Total Score (out of 85)

Attach your completed CA 3 under the appropriate assignment in the assignments section of BB. You will not receive full credit if that is not appropriately attached.

STA 320 Smith

CA #3 – Chapter 13 - using Minitab

Name:

1. When public opinion surveys are conducted by mail, a cover letter explaining the purpose of the survey is usually included. To determine whether the wording of the cover letter influences the response rate, three different cover letters were used in a survey of students at a Midwestern university. Suppose that each of the three cover letters accompanied questionnaires sent to an equal number of randomly selected students. Returned questionnaires were then classified according to the type of cover letter (I, II, or III). Use the accompanying data to test the hypothesis that the

true proportions of all returned questionnaires accompanied by cover letters I, II, and III are the same. Use a 0. 05 significance level.

Cover-letter Type

I

II

III

Frequency

49

41

37

Step 1The null and alternate hypotheses are

The variable has the specified distribution.

(The true proportions of all returned questionnaires accompanied by cover letters I, II, and III are the same.)

The variable does not have the specified distribution.

(The true proportions of all returned questionnaires accompanied by cover letters I, II, and III are not the same.)

The test is a right-tailed test.

Step 2The selected level of significance, α is 0. 05.

The selected test is the Chi-Square Goodness-Of-Fit Test. The equation of the test statistic is given by:

With df = c – 1 Where

The assumptions of the test are that 1) All expected frequencies are 1 or greater, 2) At most 20% of the expected frequencies are less than 5, and 3) the sample is a simple random sample.

As shown in Minitab output, all expected frequencies are greater than 1 and no expected frequency is less than 5. Furthermore, as stated in the problem, the sample is a simple random sample. Therefore, all the assumptions of the test are met.

Step 3The test statistic is

(df = 2, N = 127) = 1. 764

Step 4The P-value is

P = 0. 414

Step 5Conclusion (If P ≤ α, reject H0; otherwise, do not reject H0.)

From step 4, P = 0. 414. Because the P-value is greater than the specified significance level of 0. 05, we fail to reject H0. The test results are not statistically significant at the 5% level.

Step 6Interpretation

At the 5% significance level, the data provide sufficient evidence to conclude that the hypothesis that the true proportions of all returned questionnaires accompanied by cover letters I, II, and III are the same.

Minitab Output

2. A survey was conducted in the San Francisco Bay area in which each participating individual was classified according to the type of vehicle used most often and city of residence. A subset of the resulting data is given in the accompanying table. Do the data provide convincing evidence of an association between city of residence and vehicle type? Use a significance level of 0. 05. You may assume that it is reasonable to regard the sample as a random sample of Bay area residents.

Vehicle Type

City

Concord

Pleasant Hills

North San Francisco

Small

68

83

221

Compact

63

63

104

Midsize

88

123

142

Large

24

18

11

a. Is this a test of homogeneity or a test of independence?

It is a test of independence, as we want to test if there is an association

between city of residence and vehicle type.

b. Test the relevant hypotheses using a significance level of 0. 05.

Step 1The null and alternate hypotheses are

There is no association between city of residence and vehicle type.

There is an association between city of residence and vehicle type.

The test is a right-tailed test.

Step 2The selected level of significance, α is 0. 05.

The selected test is the Chi-Square Independence Test. The equation of the test statistic is given by:

With df = (r – 1) (c – 1) Where

The assumptions of the test are that 1) All expected frequencies are 1 or greater, 2) At most 20% of the expected frequencies are less than 5, and 3) the sample is a simple random sample.

As shown in Minitab output, all expected frequencies are greater than 1 and no expected frequency is less than 5. Furthermore, as stated in the problem, it is reasonable to regard the sample as a random sample of Bay area residents. Therefore, all the assumptions of the test are met.

Step 3The test statistic is

(df = 6, N = 1008) = 50. 18

Step 4The P-value is

P < 0. 001

Step 5Conclusion (If P ≤ α, reject H0; otherwise, do not reject H0.)

From step 4, P < 0. 001. Because the P-value is less than the specified significance level of 0. 05, we reject H0. The test results are statistically significant at the 5% level.

Step 6Interpretation

At the 5% significance level, the data provide convincing evidence of an association between city of residence and vehicle type.

Minitab Output

c. If applicable, examine the standardized residuals to examine the nature of the evidence about the association.

Standardized residuals greater than 2 in absolute value provide evidence against independence in that cell. The value of standardized residuals for Small vehicle type for all the three cities and for Large vehicle type for the city North San Francisco is negative and the absolute value is greater than 2. This suggests that the observed frequency of Small vehicle type all the three cities and for Large Vehicle type for the city North San Francisco is smaller than the expected frequency. In other words, Small vehicle type is not preferred vehicle type in all the three cities and Large vehicle type is not preferred vehicle type in the city North San Francisco. Furthermore, the value of standardized residuals for Large vehicle type for the city Concord and for Midsize vehicle type for the city Pleasant Hills is positive and the absolute value is greater than 2. This suggests that the observed frequency of Large vehicle type for the city Concord, and for Midsize vehicle type for the city Pleasant Hills is greater than the expected frequency. In other words, Large vehicle type is the preferred vehicle type for cities Concord and Midsize vehicle type is the preferred vehicle type for the city Pleasant Hills. However, there is as such no preference for Compact vehicle type across the three cities, as all the standardized residuals absolute values are less than 2 for Compact vehicle type across the three cities.