**Assignment 5 Template**

**LAST NAME: Yuan**

**FIRST NAME: Feng**

**USERID: y87feng**

**UWaterloo ID:20600787**

**Problem 1: Fill in the information below based on your data which were generated using your ID number as the seed for the random number generator.**

**Insert the table of observed and expected frequencies here.**

**Category.1 Category.2 Category.3 Category.4 Category.5**

**Observed 25 9 43 20 53**

**Expected 30 30 30 30 30**

**The hypothesis of interest is that the data arise from a Multinomial model with equal probabilities.**

**The observed value of the likelihood ratio statistic for testing this hypothesis**

**= 44.27806**

**The degrees of freedom for the Chi-squared distribution =4**

**The p-value =5.616762e-09**

**Insert your conclusion regarding the hypothesis here.**

**The p-value is less than 0.001 and it’s a strong evidence based on data against the hypothesis**

**The observed value of the Pearson Goodness of Fit statistic for testing this hypothesis = 42.13333**

**The degrees of freedom for the Chi-squared distribution =4**

**The p-value =1.565307e-08**

**Insert your conclusion regarding the hypothesis here.**

**The p-value is less than 0.001 and it’s a strong evidence based on data against the hypothesis**

**Problem 2: Fill in the information below based on your data which were generated using your ID number as the seed for the random number generator.**

**Model = 1**

**Insert the original table of observed and expected frequencies here.**

**[,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11]**

**y 0.000 1.000 2.00 3.00 4.00 5.00 6.00 7.00 8.000 9.000 10.000**

**observed 3.000 6.000 19.00 30.00 34.00 19.00 19.00 6.00 5.000 8.000 1.000**

**expected 2.063 8.842 18.95 27.08 29.02 24.88 17.77 10.88 5.832 2.778 1.899**

**Insert the table of observed and expected frequencies after collapsing here.**

**[,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]**

**y 1.0 2.00 3.00 4.00 5.00 6.00 7.00 8.00**

**observed 9.0 19.00 30.00 34.00 19.00 19.00 6.00 14.00**

**expected 10.9 18.95 27.08 29.02 24.88 17.77 10.88 10.51**

**The hypothesis of interest is the data arise from a Poisson model.**

**The observed value of the likelihood ratio statistic for testing this hypothesis**

**= 6.736063**

**The degrees of freedom for the Chi-squared distribution =6**

**The p-value =0.3459471**

**Insert your conclusion regarding the hypothesis here.**

**The p-value is greater than 0.1 and has no evidence based on data against the hypothesis**

**The observed value of the Pearson Goodness of Fit statistic for testing this hypothesis = 6.328437**

**The degrees of freedom for the Chi-squared distribution =6**

**The p-value =0.3874214**

**Insert your conclusion regarding the hypothesis here.**

**The p-value is greater than 0.1 and has no evidence based on data against the hypothesis .**

**Problem 3: Fill in the information below based on your data which were generated using your ID number as the seed for the random number generator.**

**Number of observations = 159**

**Insert the table of observed frequencies here.**

**Smoker Indicator Average Short Tall**

**Non-smoker 29 14 37**

**Smoker 24 39 16**

**Insert the table of expected frequencies here.**

**Height Indicator**

**Smoker Indicator Average Short Tall**

**Non-smoker 26.67 26.67 26.67**

**Smoker 26.33 26.33 26.33**

**The hypothesis of interest is that the variate smoking and the variate height are independent variates.**

**The observed value of the likelihood ratio statistic for testing this hypothesis**

**= 21.29361**

**The degrees of freedom for the Chi-squared distribution =2**

**The p-value =2.377674e-05**

**Insert your conclusion regarding the hypothesis here.**

**The p-value is less than 0.001 and has a strong evidence based on data against the hypothesis**

**The observed value of the Pearson Goodness of Fit statistic for testing this hypothesis = 20.57943**

**The degrees of freedom for the Chi-squared distribution =2**

**The p-value =3.398079e-05**

**Insert your conclusion regarding the hypothesis here.**

**The p-value is less than 0.001 and has a strong evidence based on data against the hypothesis**

**Suppose for your data you found evidence of a relationship between smoking and height. Can you conclude that a person’s height affects whether they smoke or not? Why or why not?**