

Assignment 5 Example

LAST NAME: STRUTHERS

FIRST NAME: CYNTHA

USERID: castruth

UWaterloo ID: 20456458

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 | 8 | 23 | 43 | 63 | 13 |
| 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

= 69.33145

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

The p-value = 3.141931e-14

Insert your conclusion regarding the hypothesis here.

The observed value of the Pearson Goodness of Fit statistic for testing this hypothesis = 69.33333

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

The p-value = 3.141931e-14

Insert your conclusion regarding the hypothesis here.

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

Original table of observed and expected frequencies:

Table of Observed and Expected Frequencies

| | [, 1] | [, 2] | [, 3] | [, 4] | [, 5] | [, 6] | [, 7] | [, 8] | [, 9] | [, 10] |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| y | 0.000 | 1.00 | 2.00 | 3.0 | 4.0 | 5.00 | 6.00 | 7.000 | 8.000 | 9.000 |
| observed | 3.000 | 8.00 | 20.00 | 35.0 | 28.0 | 28.00 | 15.00 | 6.000 | 4.000 | 3.000 |
| expected | 2.822 | 11.21 | 22.27 | 29.5 | 29.3 | 23.29 | 15.42 | 8.753 | 4.347 | 3.087 |

Table of observed and expected frequencies after collapsing:

Table of Observed and Expected Frequencies

| | [, 1] | [, 2] | [, 3] | [, 4] | [, 5] | [, 6] | [, 7] | [, 8] |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|
| y | 1.00 | 2.00 | 3.0 | 4.0 | 5.00 | 6.00 | 7.000 | 8.000 |
| observed | 11.00 | 20.00 | 35.0 | 28.0 | 28.00 | 15.00 | 6.000 | 7.000 |
| expected | 14.03 | 22.27 | 29.5 | 29.3 | 23.29 | 15.42 | 8.753 | 7.434 |

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

The observed value of the likelihood ratio statistic for testing this hypothesis
= 3.882476

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

The p-value = 0.6925769

Insert your conclusion regarding the hypothesis here.

The observed value of the Pearson Goodness of Fit statistic for testing this hypothesis = 3.82811

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

The p-value = 0.6999245

Insert your conclusion regarding the hypothesis here.

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 = 109

Insert the table of observed frequencies here.

Table of Observed Frequencies:

| Smoker | Indicator | Height Indicator | | |
|--------|------------|------------------|-------|------|
| | | Average | Short | Tall |
| | Non-smoker | 23 | 6 | 26 |
| | Smoker | 13 | 31 | 10 |

Insert the table of expected frequencies here.

Table of Expected Frequencies:

| Smoker | Indicator | Height Indicator | | |
|--------|------------|------------------|-------|-------|
| | | Average | Short | Tall |
| | Non-smoker | 18.17 | 18.67 | 18.17 |
| | Smoker | 17.83 | 18.33 | 17.83 |

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

= 28.66472

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

The p-value = 5.963973e-07

Insert your conclusion regarding the hypothesis here.

The observed value of the Pearson Goodness of Fit statistic for testing this hypothesis = 26.77386

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

The p-value = 1.535077e-06

Insert your conclusion regarding the hypothesis here.

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?