Homework 4 - https://github.com/Henryblake2777/Homework4

Henry Blake - hrb987

2025-02-17

Contents

Problem 1 - Iron Bank

Null Hypothesis -

H0 -> P = .024

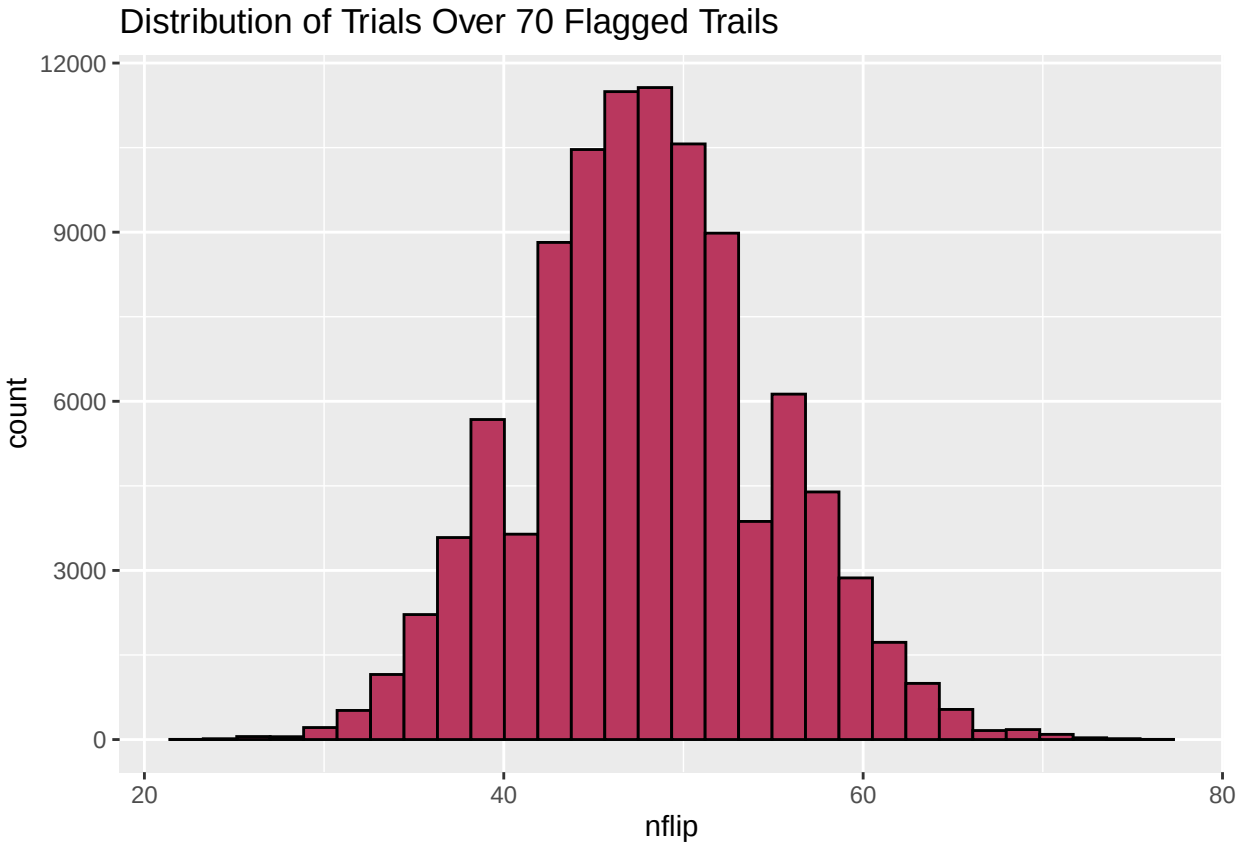
Test Statistic -

TS = 0.0346 - 0.024 = 0.0106

P-Value -

```
## [1] 0.00141
```

Probability Distribution -



Conclusion -

Due to a p-value of .0013, there is significant evidence that the 70 flagged by the detection algorithm were unusually high and not due to simple variability.

Problem 2 - Health Inspections -

Null Hypothesis -

H0 -> P = .03

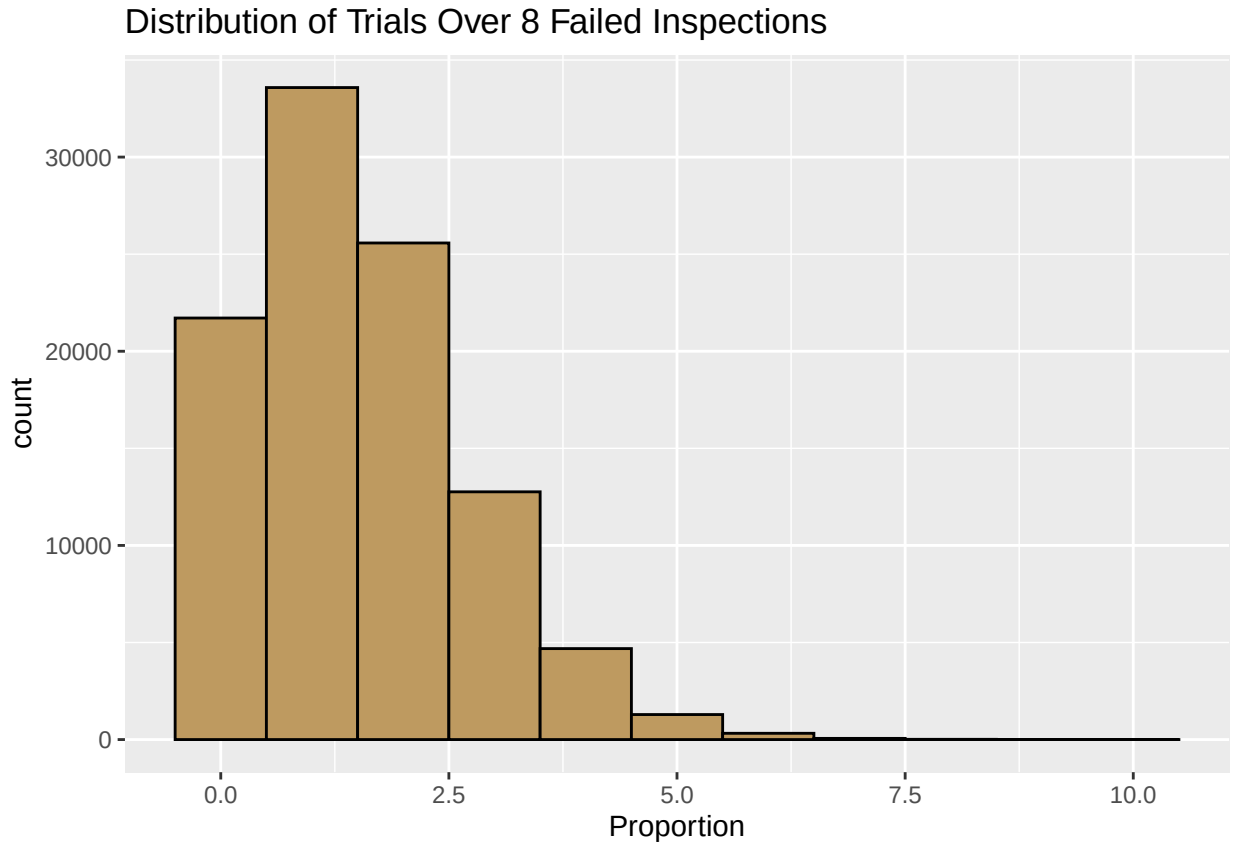
Test Statistic -

TS = .16 - .03 = .13

P-Value -

```
## [1] 1
```

Probability Distribution -



Conclusion -

Due to a p-value of practically 0, there is very significant evidence that the 8 failed health inspections by Gourmet Bites are not due to normal factors.

Problem 3 - Jury Selection

Null Hypothesis -

H0 -> Observed frequencies are consistent with the model

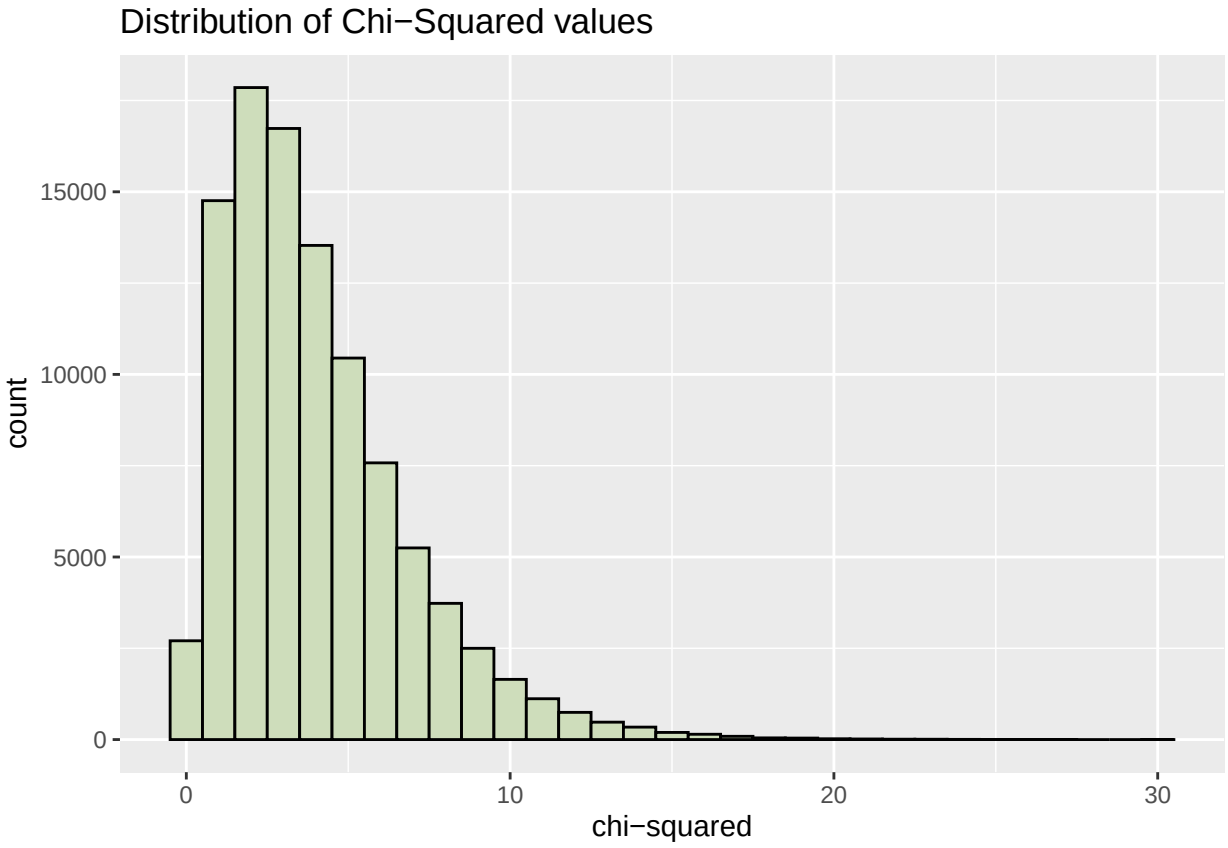
Test Statistic -

Ts = chi2 = 12.43

Chi-Squared Value -

```
## [1] 0.01457
```

Probability Distribution -

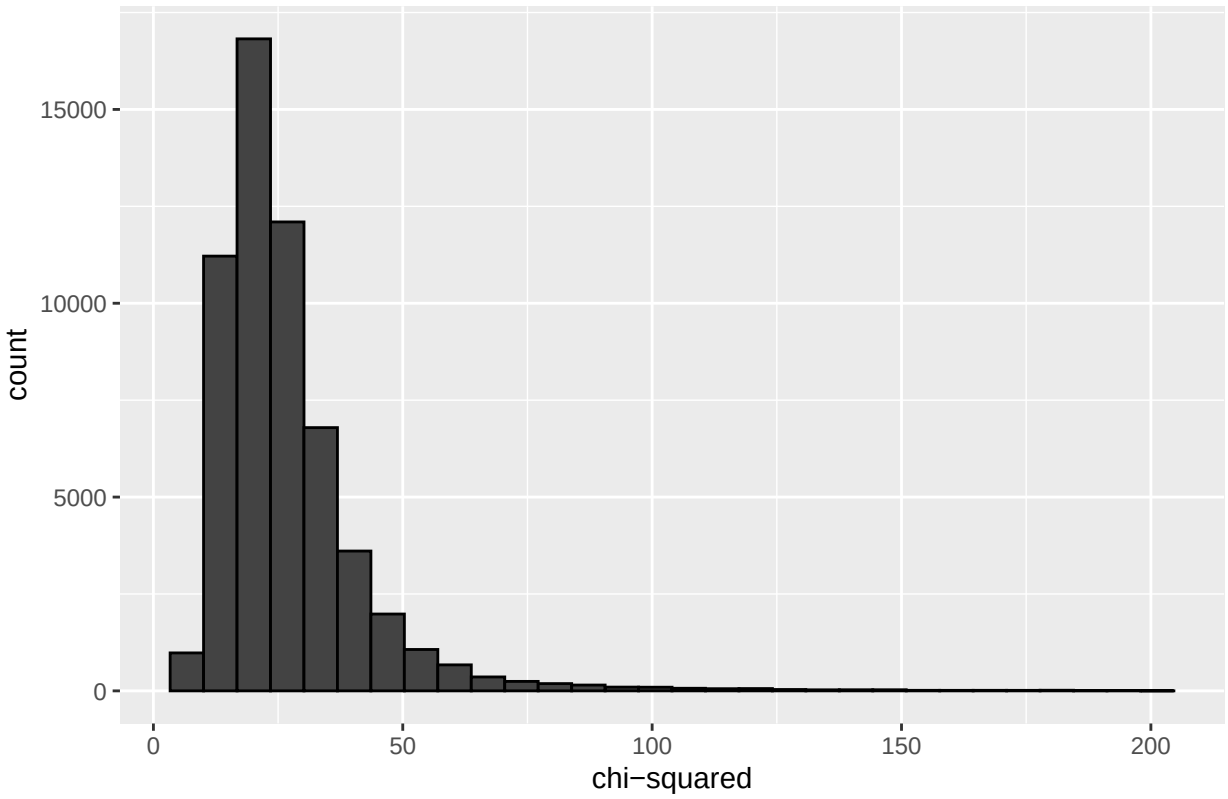


Conclusion -

Due to a p-value of .0143, there is intriguing evidence that the distribution of jurors selected by the judge does not follow the expected distribution. However, a confounding variable that may affect this could be the type of cases that the judge works.

Problem 4 - LLM Watermarking

Part A: Null Distribution -



Part B: Checking For a Watermark -

```
##  [1] 0.510 0.924 0.079 0.486 0.493 0.009 0.326 0.987 0.086 0.060
```

According to the p-values of the chi-squared statistics, sentence 6 seems to be the one that is watermarked because of the p-value of .009. It does not seem to follow the typical letter frequency distribution.