

Homework 2

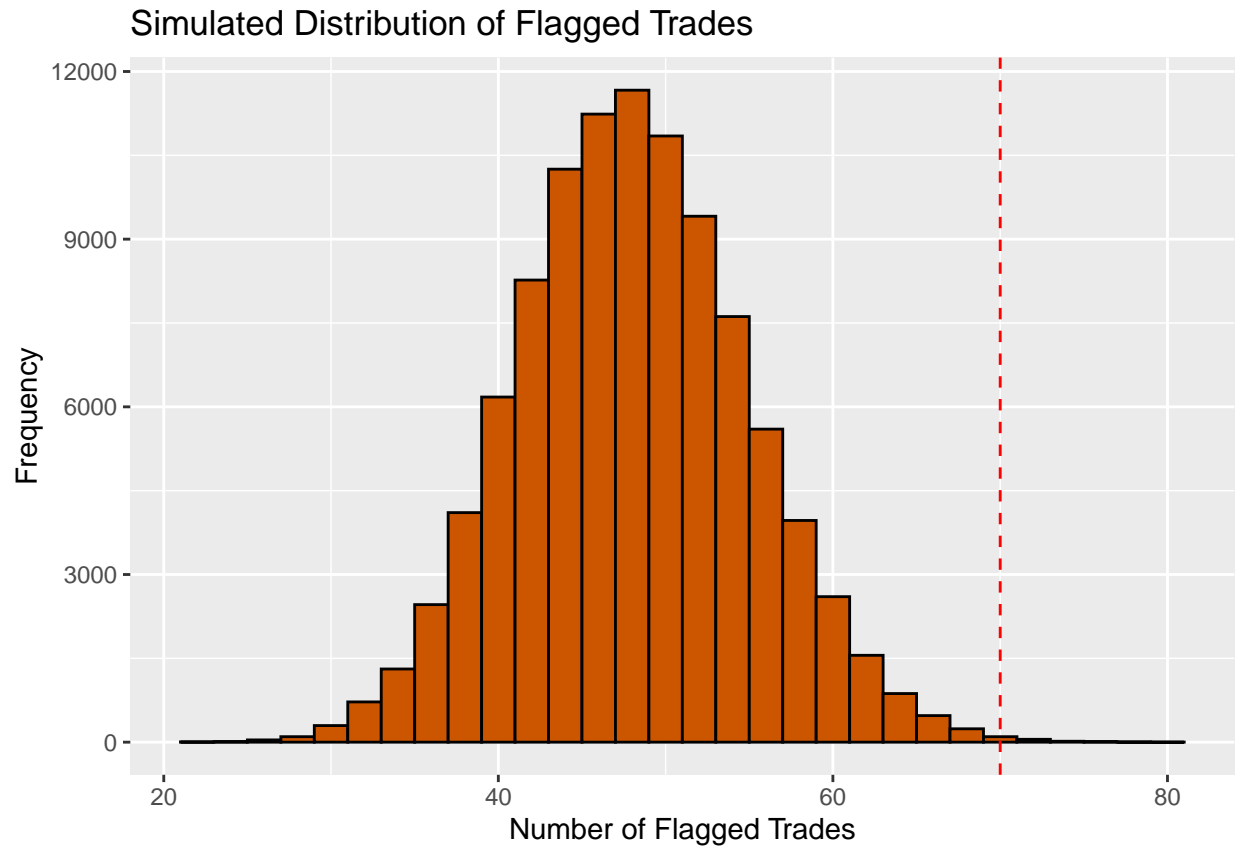
Trent Latz

2025-02-20

Name: Trent Latz **UT EID:** tjl2597 **GitHub Link:** <https://github.com/trentjlatz/SDS315-HW4>

Problem 1 - Iron Bank

The null hypothesis is that the amount of trades flagged from the Iron Bank is the same as the baseline rate of 2.4%. The test statistic used is the observed number of flagged trades: 70 out of 2021



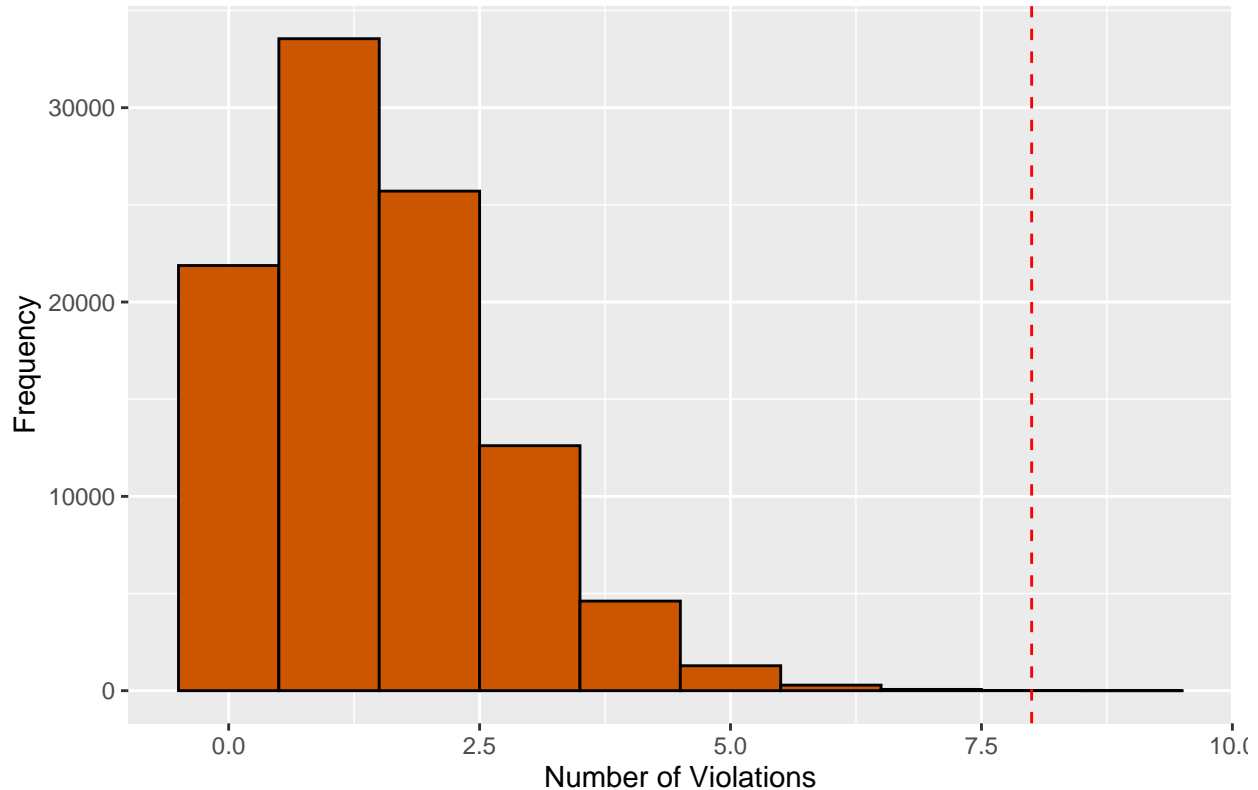
```
## [1] 0.00178
```

With a low p-value of .002, the null hypothesis that Iron Bank's rate of flagged trades is the same as the 2.4% baseline seems unlikely, suggesting that the trades are being flagged at a higher rate than expected due to random chance.

Problem 2 - Health Inspections

The null hypothesis is that the Gourmet Bites' violation rates is the same as the citywide average of 3%. The test statistic used is the observed number of health violations: 8 out of 50 inspections.

Simulated Distribution of Violations



```
## [1] 0.00016
```

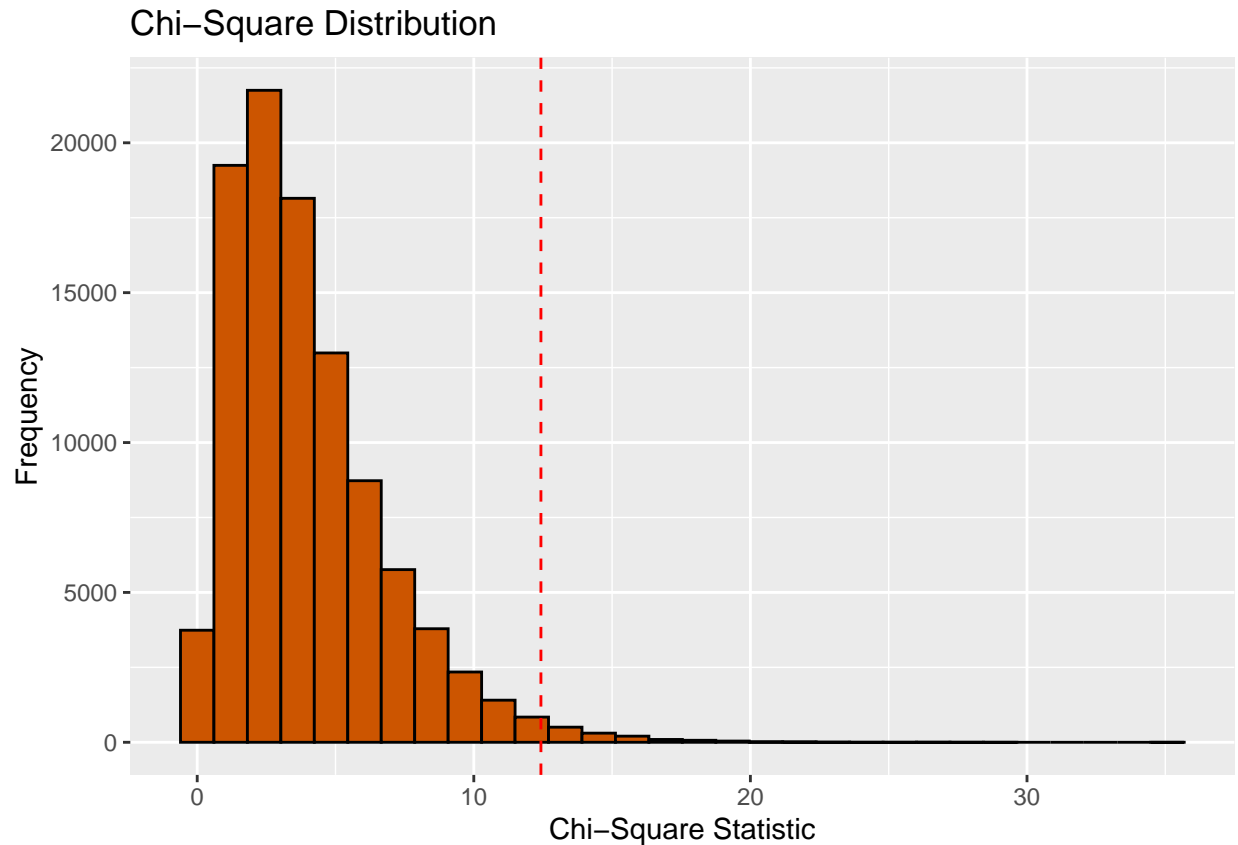
With a p-value of .0002, the null hypothesis that Gourmet Bites' violation rate is equal to the citywide average of 3% seems to be extremely unlikely, as 8 violations out of 50 inspections is very unlikely to be due to chance.

Problem 3 - Evaluating Jury Selection for Bias

H0: The null hypothesis is that the judge's jury selection process matches the county's population distribution.

T: The test statistic used is the chi-square test, measuring deviation of observed from expected results.

```
## [1] 12.42639
```

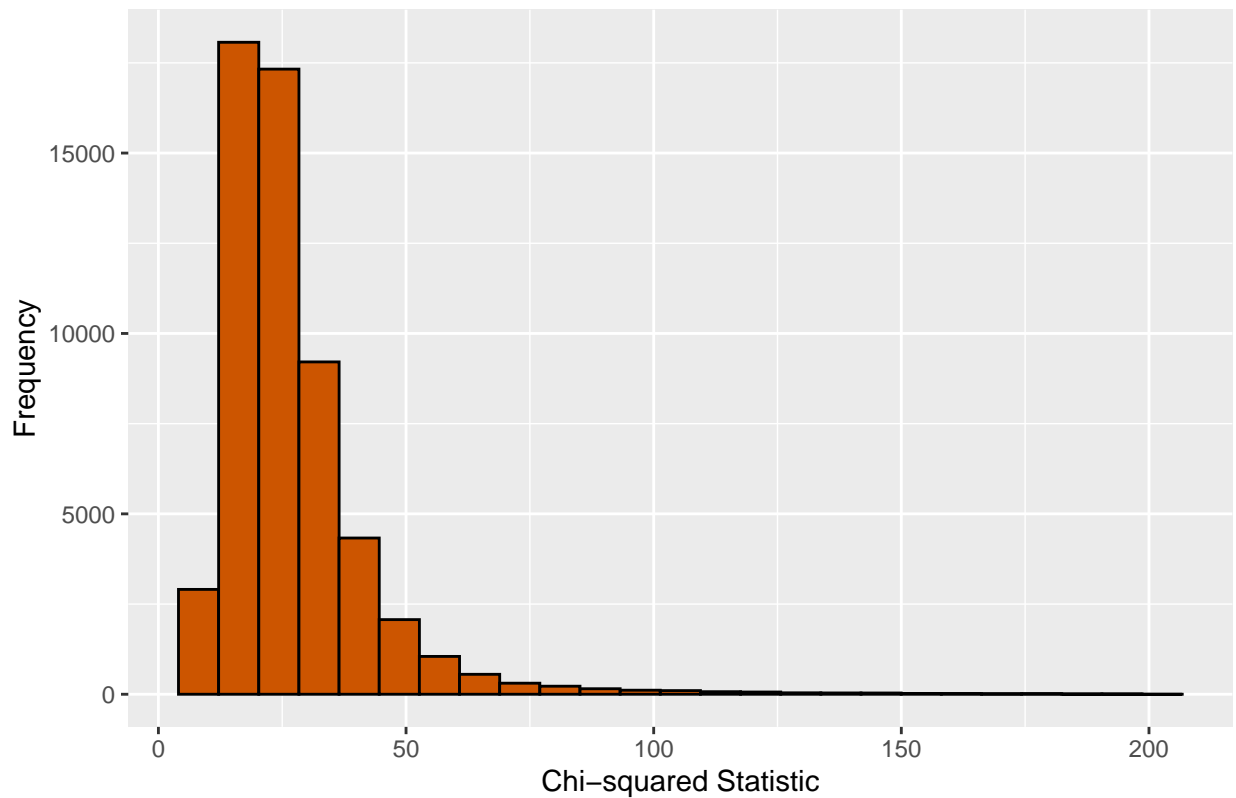


```
## [1] 0.01412
```

The graph above is our $P(T \mid H_0)$, with the red dash showing our test statistic(calculated chi-square). The p-value of .015 is less than .05, so the null hypothesis is rejected. The distribution of jurors for this judge is (statistically) significantly different from the county's population distribution, suggesting systematic bias in jury selection. While the results suggest potential bias, other factors such as the nature of the trial, the charges, or even the judge's selection methods could explain this dependency. To investigate further, we could look into those factors and compare them with other judges to see if they could explain this difference.

Problem 4: LLM watermarking

Chi-squared Distribution of Brown Sentences



##	Chi_Squared	P_Value
## 1	22.930848	0.513
## 2	13.051050	0.926
## 3	46.285861	0.076
## 4	23.546278	0.489
## 5	23.676149	0.484
## 6	96.452677	0.009
## 7	28.271419	0.328
## 8	9.635023	0.988
## 9	44.928631	0.084
## 10	49.960559	0.059

The chi-squared analysis of the letter frequency distributions reveals which sentence is LLM-generated. The sentence with the highest chi-squared value is sentence number six, “Feeling vexed after an arduous and zany day at work, she hoped for a peaceful and quiet evening at home, cozying up after a quick dinner with some TV, or maybe a book on her upcoming visit to Auckland.” This sentence is the one I originally thought was AI-generated by just lookin at it, and is also shown by being the lowest p value of .009.