

SDS 315 HW 4

<https://github.com/sophiayang5/SDS-315-HW-4>

Sophia Yang (sy23928)

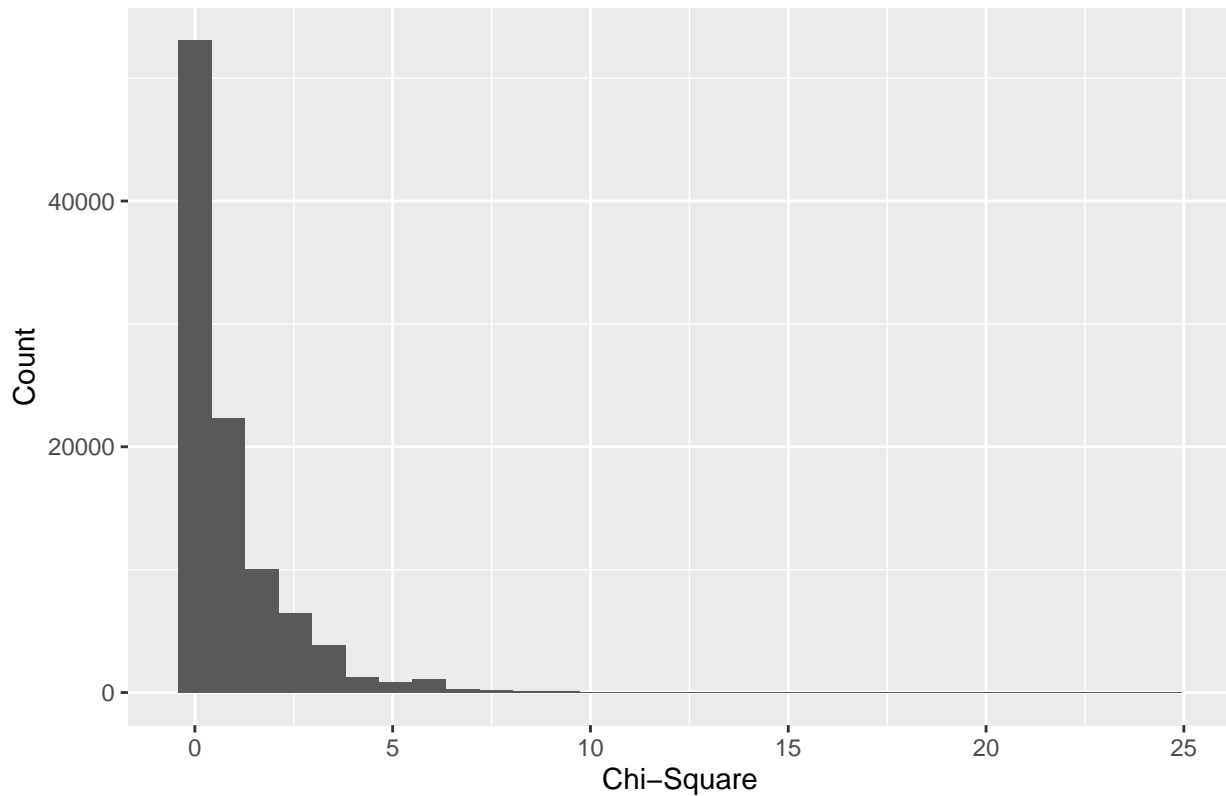
2025-02-15

Contents

Problem 1: Iron Bank	2
Problem 2: Health Inspections	3
Problem 3: Evaluating Jury Selection for Bias	4
Problem 4: LLM Watermarking	5
Part A.	5
Part B.	6

Problem 1: Iron Bank

Chi-Square Distribution of Flagged Trades



```
## count(chi2 >= trades_chi2)/n()
## 1 0.00281
```

The null hypothesis is that the rate of Iron Bank's flagged securities trades are at the same 2.4% baseline rate as other traders. I used the chi-square test statistic to measure the evidence against the null hypothesis. The probability distribution of the chi-square (assuming the null hypothesis is true) is plotted above. The p-value is 0.0024 (assuming the null hypothesis is true). Based on the p-value of the chi-square we got for the Iron Bank, the null hypothesis does not look plausible, because the p-value, and therefore the probability of having 70 flagged trades out of 2021 trades, is super low.

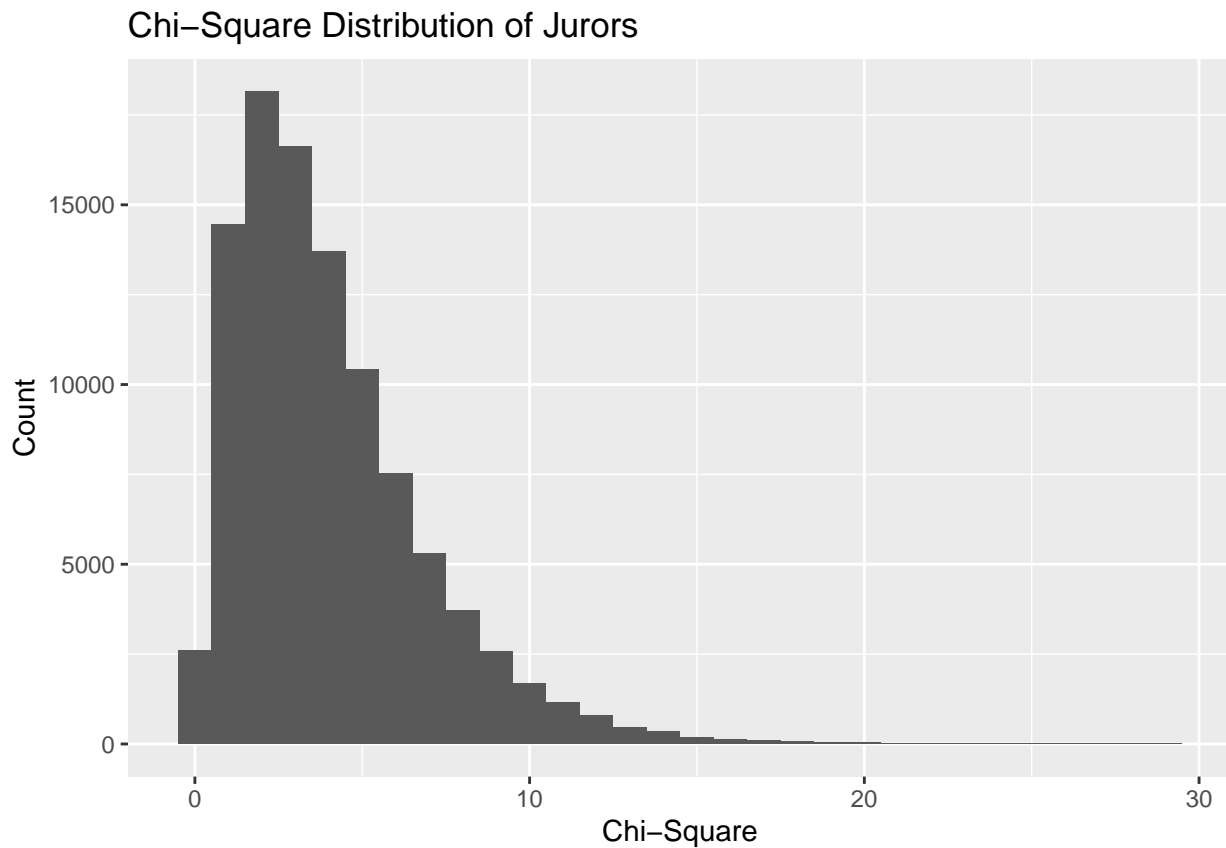
Problem 2: Health Inspections



```
## count(chi2 >= health_chi2)/n()  
## 1 0
```

The null hypothesis is that the rate of Gourmet Bites's health code violations are at the same 3% baseline rate as other restaurants. I used the chi-square test statistic to measure the evidence against the null hypothesis. The probability distribution of the chi-square (assuming the null hypothesis is true) is plotted above. The p-value is extremely close to 0 (assuming the null hypothesis is true). Based on the p-value of the chi-square we got for Gourmet Bites, the null hypothesis does not look plausible, because the p-value, and therefore the probability of having 8 health code violations out of 50 health inspections, is super low.

Problem 3: Evaluating Jury Selection for Bias



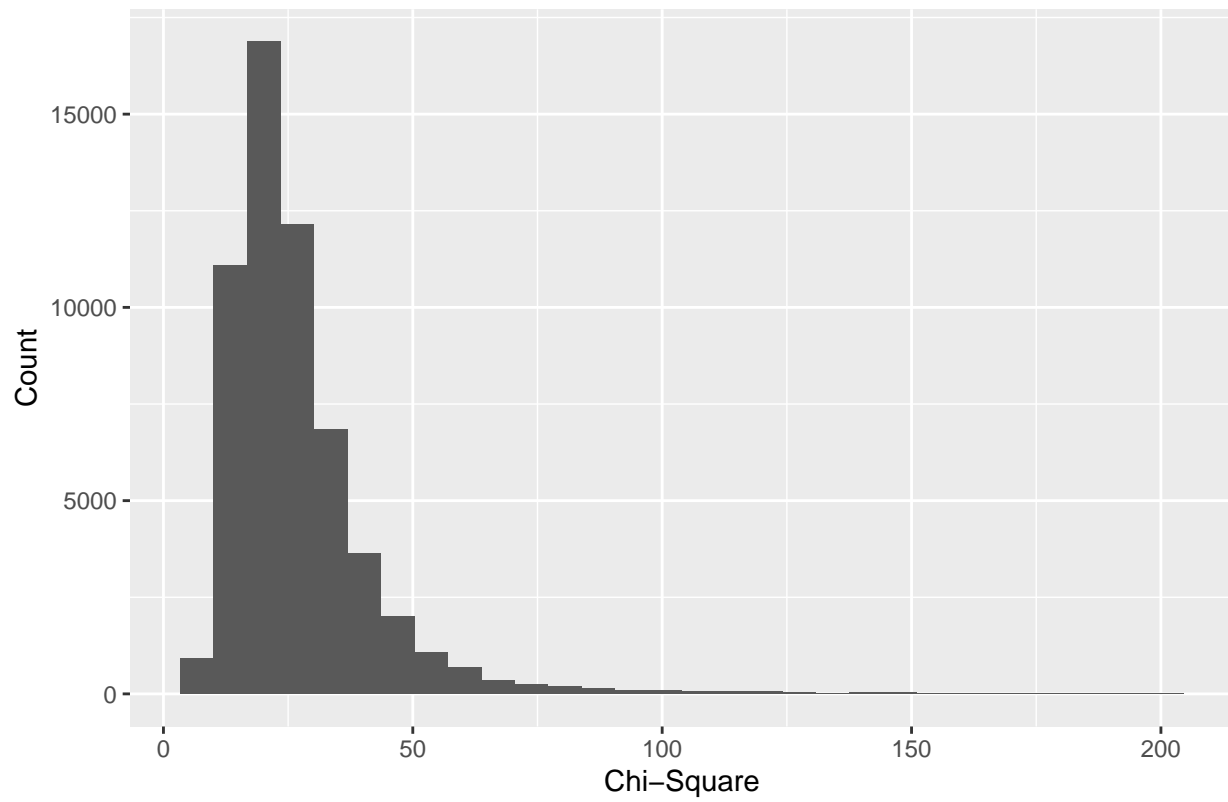
```
## count(chi2 >= jurors_chi2)/n()  
## 1 0.01458
```

The null hypothesis is that juries selected by a particular judge have proportional representation from racial/ethnic groups with the county's eligible jury population. I used the chi-square test statistic to measure the evidence against the null hypothesis. The probability distribution of the chi-square (assuming the null hypothesis is true) is plotted above. The p-value is 0.014 (assuming the null hypothesis is true). Based on the p-value of the chi-square we got for the jurors, the null hypothesis does not look plausible, because the p-value, and therefore the probability of having that racial/ethnic distribution, is low. Other explanations might be that some racial/ethnic groups are more likely to be "excused for hardship" or removed "for cause," and thus alter the proportion of that race/ethnicity represented in juries.

Problem 4: LLM Watermarking

Part A.

Null Distribution of Letter Frequencies



Part B.

Sentence	P-Value
1	0.513
2	0.926
3	0.076
4	0.489
5	0.484
6	0.009
7	0.328
8	0.988
9	0.084
10	0.059

Sentence 6 has been produced by an LLM. Due to the p-value of the sentence's chi-square statistic for its letter frequencies (0.009), I can deduce that the likelihood of getting those letter frequencies is extremely low, assuming the null hypothesis (the sentence was not made by an LLM) is true.