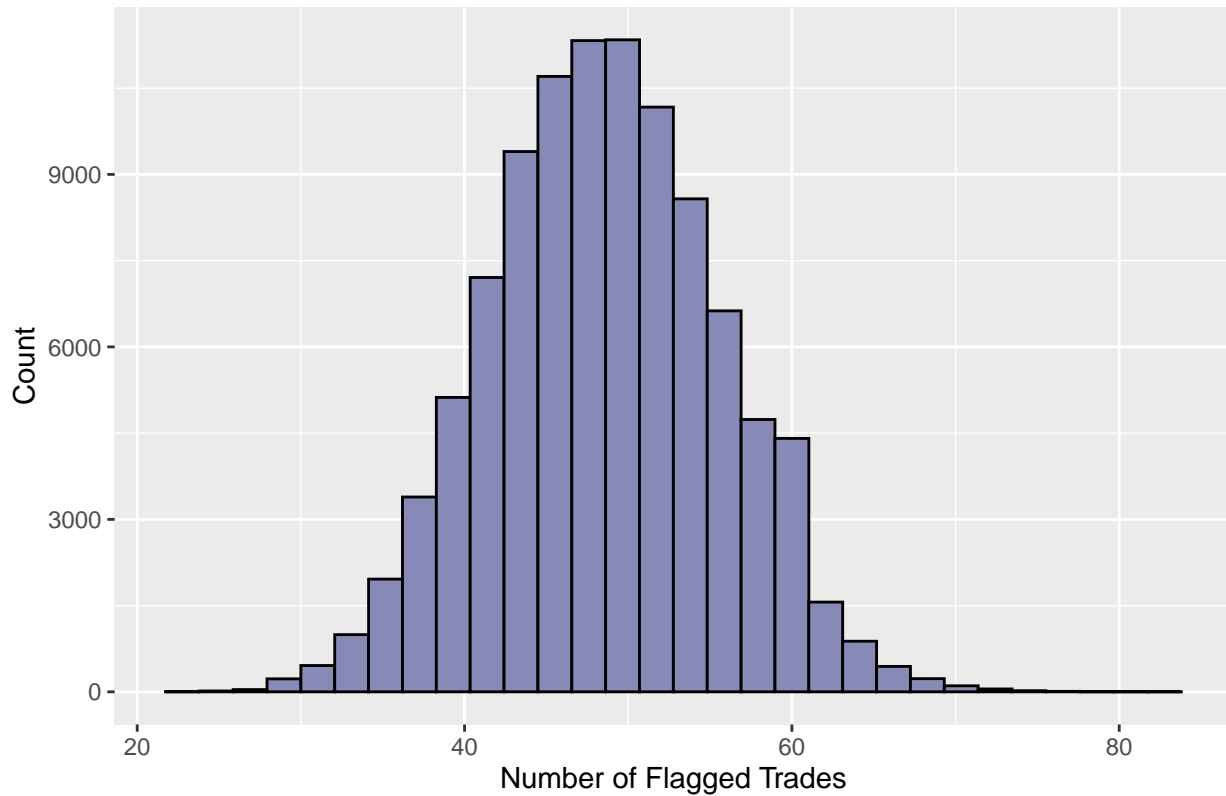# HW4

Brian Liang

2025-02-16

UT EID: bl29375

GitHub Link: https://github.com/brianliangg/SDS315_HW4

Problem 1

Monte Carlo Simulation of Flagged Trades Under the Null Hypothesis

The null hypothesis we are using is that over the long run, securities trades from the Iron Bank are flagged at the same 2.4% baseline rate as that of other traders. The test statistic is the number of flagged trades which in this case is the 70 out of 2021 trades. Above is the probability distribution of the test statistic given that the null hypothesis is true by using 100,000 Monte Carlo Simulations. There was found to be a p-value of 0.00187 after the Monte Carlo simulation under the null hypothesis. This means there was around a 0.187% chance of getting 70 flagged trades or more if the null hypothesis is true. And because this probability is unlikely, the null hypothesis that the securities trades from Iron Bank are flagged at the same 2.4% baseline rate as that of other traders is not consistent with the flagged securities trade from Iron Bank.

Problem 2

Monte Carlo Simulation of Health Code Violations Under the Null Hypoth...
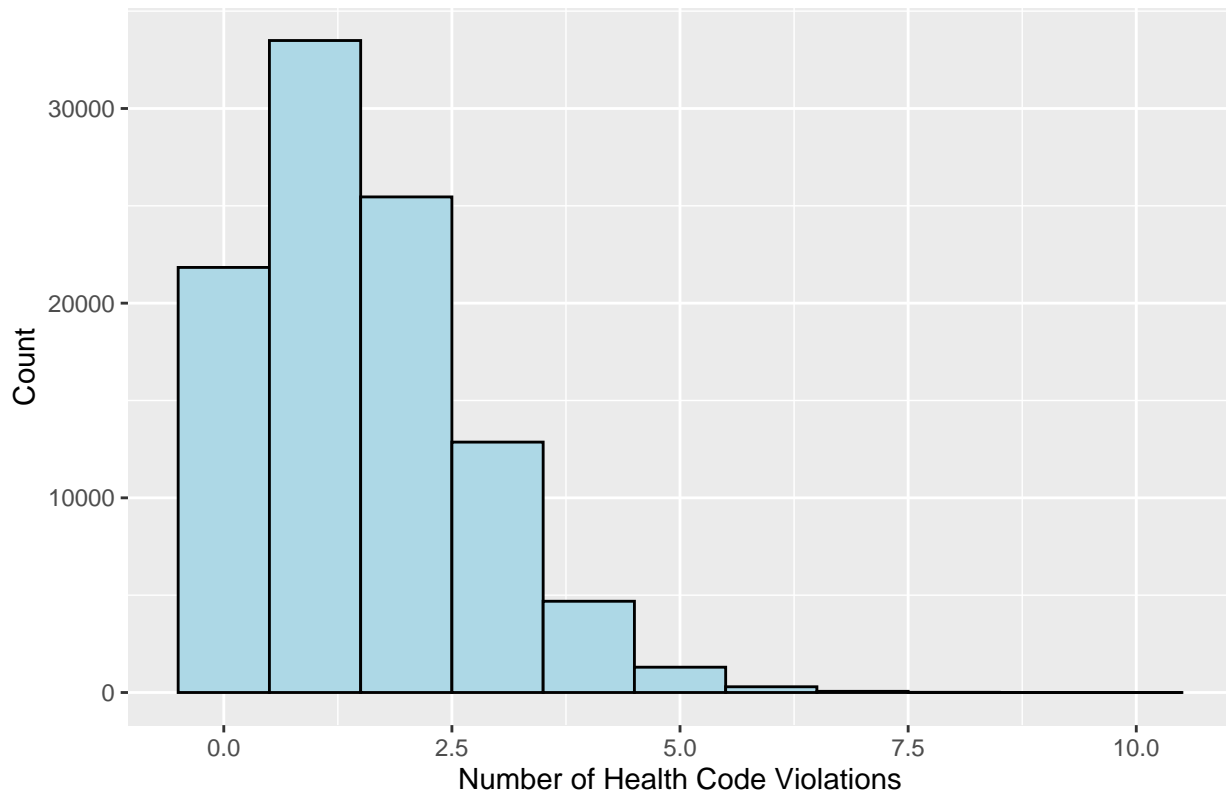
The null hypothesis we are using is that n average, restaurants in the city are cited for health code violations at the same 3% baseline rate. The test statistic is the number of health code violations reported which in this case is the 8 out of 50 from Gourmet Bites. Above is the probability distribution of the test statistic given that the null hypothesis is true by using 100,000 Monte Carlo Simulations. There was found to be a p-value of 1.3×10^{-4} after the Monte Carlo simulation under the null hypothesis. This means there was around a 0.013% chance of getting 8 health code violations or more if the null hypothesis is true. And because this probability is so unlikely, the null hypothesis that restaurants in the city are cited for health code violations at the same 3% baseline rate is not consistent with the health code violations from Gourmet Bites.

Problem 3

Monte Carlo Simulation of Chi–Square Distribution Under the Null Hypoth
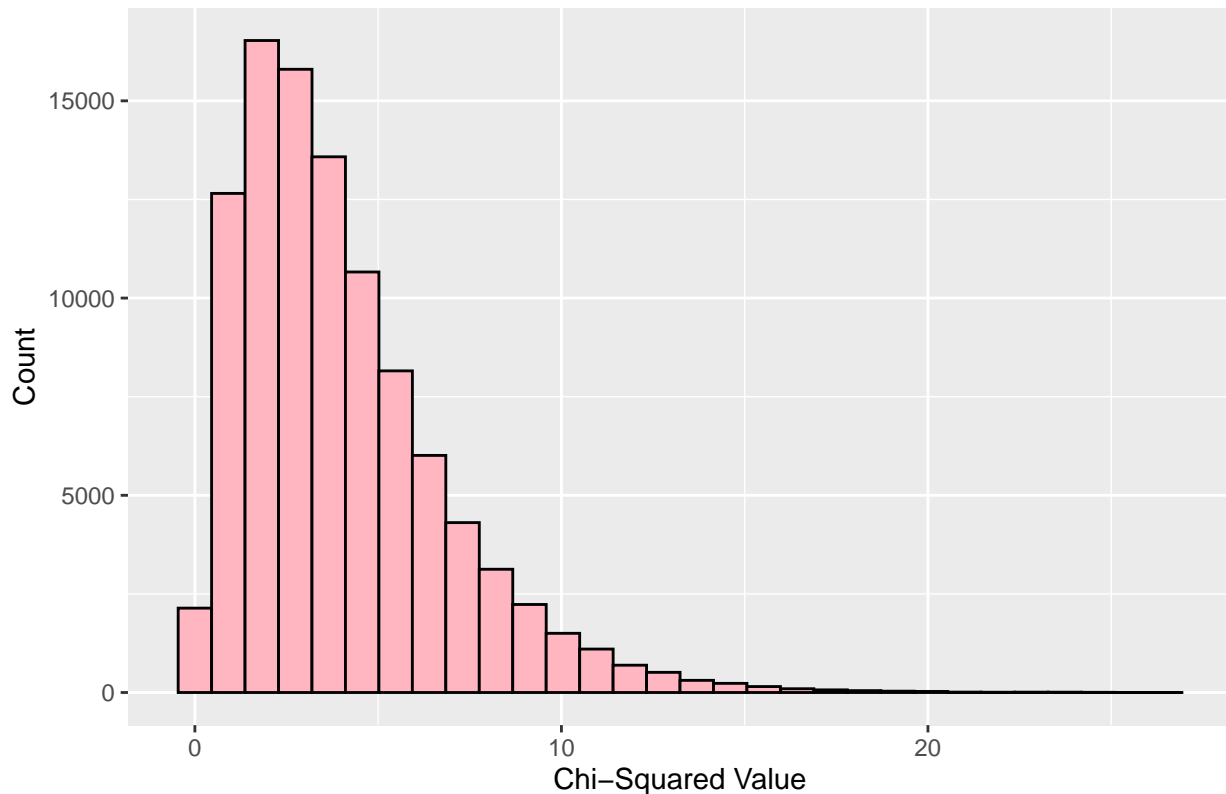
The hypothesis test that was used establishes the null hypothesis is that the distribution of jurors empaneled by judges across the country is 30% for Group 1, 25% for Group 2, 20% for Group 3, 15% for Group 4, and 10% for Group 5. The test statistic we are comparing to the null hypothesis is the chi-square statistic of corresponding group counts for the empaneled jurors in 20 trials overseen by the judge in question, which is 85 for Group 1, 56 for Group 2, 59 for Group 3, 27 for Group 4, and 13 for Group 5. Above is the probability distribution or chi-square distribution of the test statistic given that the null hypothesis is true by using 100,000 Monte Carlo Simulations. There was found to be an observed chi-squared statistic of 12.4263889 for the judge in question, leading to a a p-value of 0.01445. This means there was around a 1.445% chance of observing these counts or counts more extreme if the null hypothesis is true. And because this probability is unlikely, it does suggest that may have been systematic bias in the jury selection. However, this could also be affected by other confounding variables such as through the peremptory challenges that the attorneys have. We could investigate this further observing the groups that were excluded my the peremptory challenges and see if that is consistent with a majority of other attorneys.

Problem 4

Part A

Chi–Square Distribution Under the Null Hypothesis
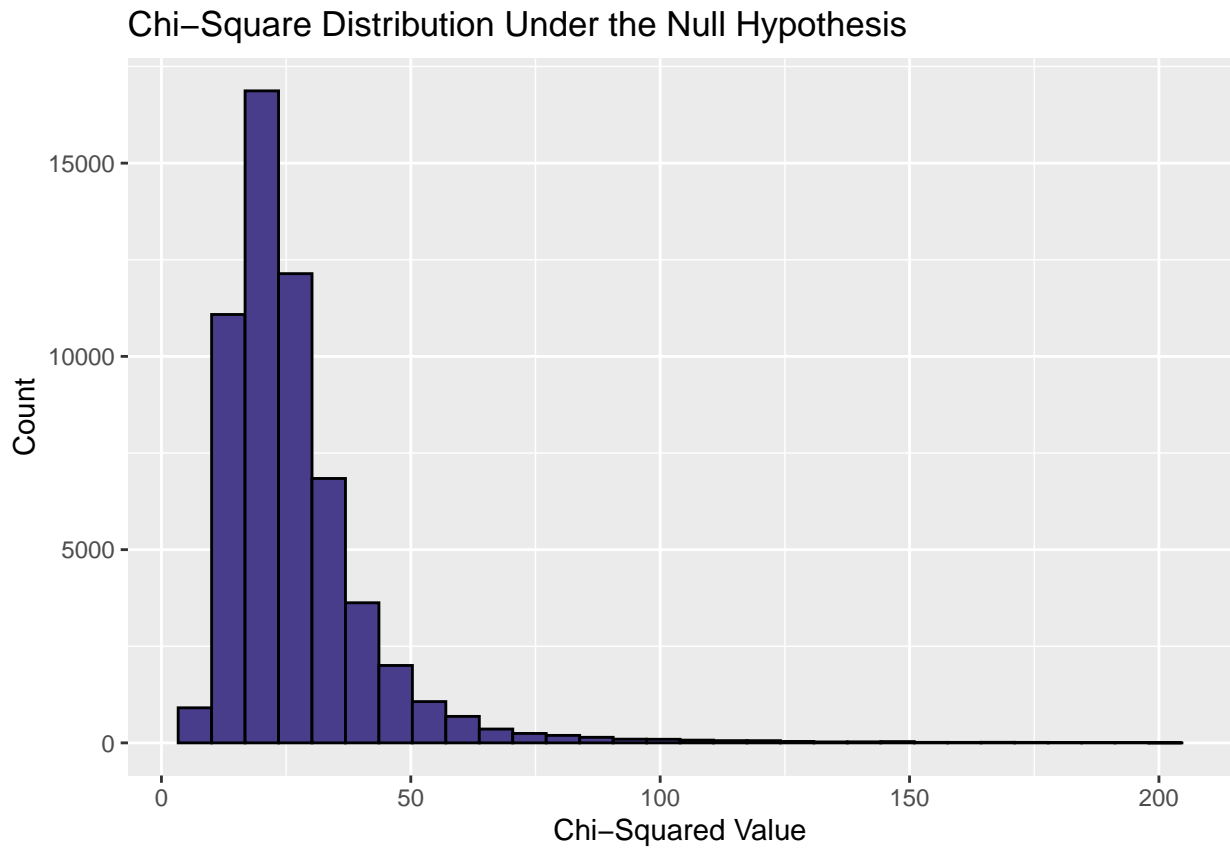
Displayed is the what the chi-squared statistic looks like across lots of normal English sentences not generated by an LLM. This will act as the null distribution for Part B.

Part B

Table 1: Table of P-Values For Each Sentence

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

Above is the table of p-values for each sentence based on their letter frequencies and if they line up with our null distribution. Based on these p-values, the sentence that has been watermarked by an LLM is sentence 6 because it has the lowest p-value of 0.009. This would mean there is around a 0.9% of observing a sentence

4

with that letter frequency or more extreme which is unlikely.