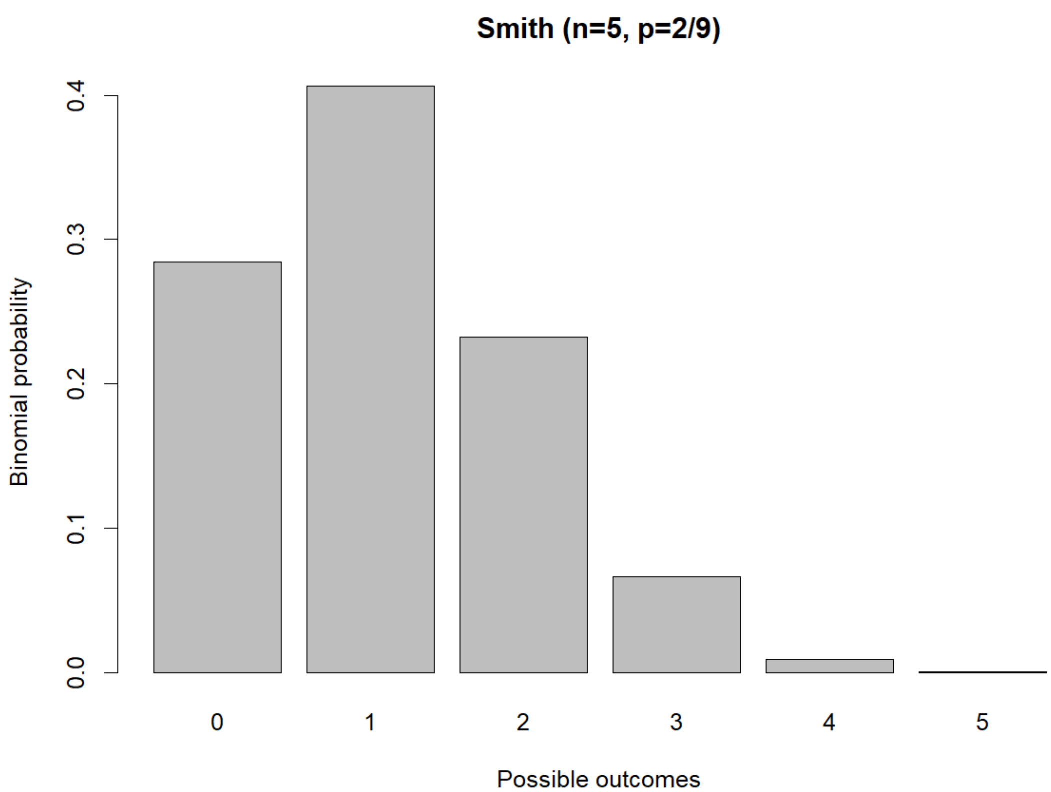
**Assignment 1)**

**Smith**

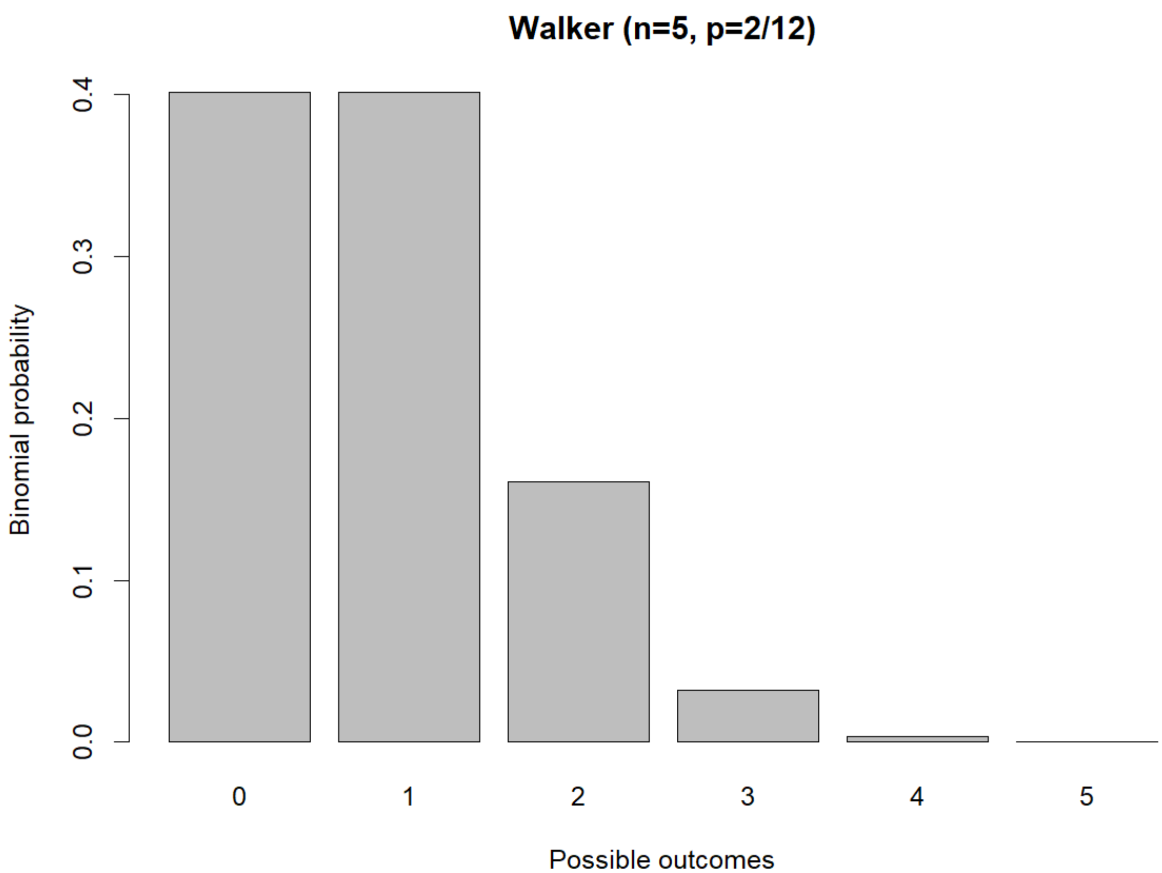
| Outcomes | BINPROB | TYPEOUT |
|---|---|---|
| 0 | 0.28463 | Suitable |
| 1 | 0.40661 | Suitable |
| 2 | 0.23235 | Unsuitable |
| 3 | 0.06639 | Unsuitable |
| 4 | 0.00948 | Unsuitable |
| 5 | 0.00054 | Unsuitable |

---



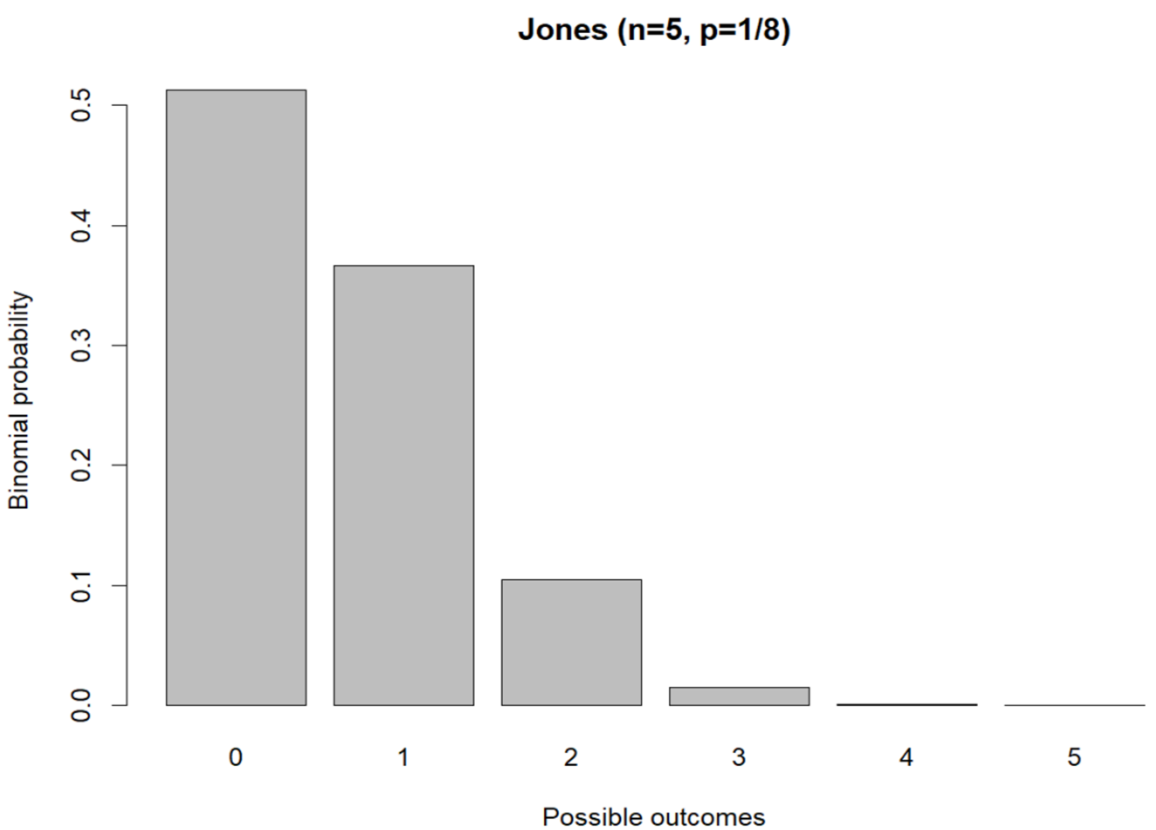
**Walker**

| OUTCOMES | BINPROB | TYPEOUT |
|---|---|---|
| 0 | 0.40188 | Suitable |
| 1 | 0.40188 | Suitable |
| 2 | 0.16075 | Unsuitable |
| 3 | 0.03215 | Unsuitable |
| 4 | 0.00322 | Unsuitable |
| 5 | 0.00322 | Unsuitable |



**Jones**

| OUTCOMES | BINPROB | TYPEOUT |
|---|---|---|
| 0 | 0.51291 | Suitable |
| 1 | 0.36636 | Suitable |
| 2 | 0.10468 | Unsuitable |
| 3 | 0.01495 | Unsuitable |
| 4 | 0.00107 | Unsuitable |
| 5 | 3.00E-05 | Unsuitable |



Total Probability of Suitable Outcomes =

- Smith: 0.69124
- Walker: 0.80376
- Jones: 0.87927

Total Probability of Unsuitable Outcomes =

- Smith: 0.30876
- Walker: 0.19624
- Jones: 0.12073

Who should proceed?

I think Walker should proceed, Walker ($\geq 0.80$). Smith ($< 0.75$) and Jones ($< 0.90$) should not, at their stated risk levels.

Formula used from lecture:

$$P(X = x) \; = \; \frac{n!}{x!(n-x)!}\, p^x (1-p)^{\,n-x}$$

n = number of trials (here, 5 years),

x = number of frost years,

p = probability of frost in a single year (from historical frequency),

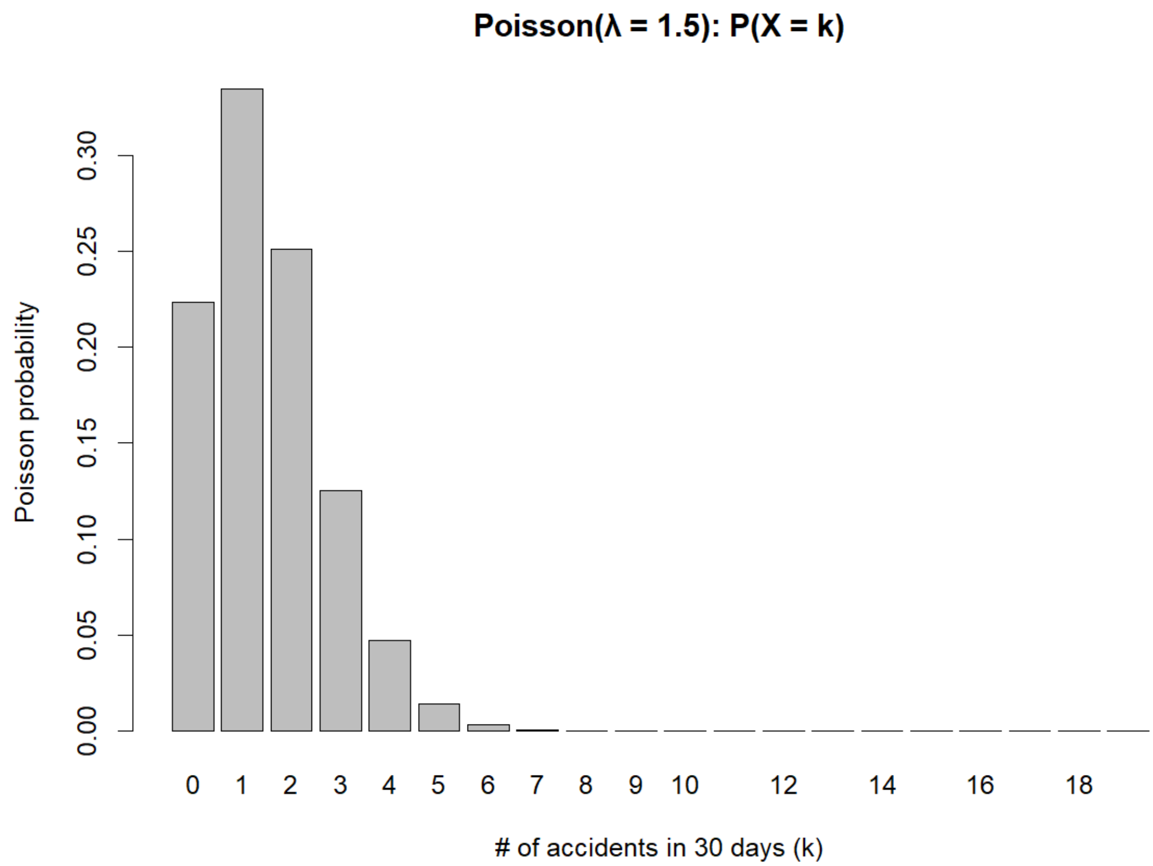
$1-p$ = probability of no frost.

Method used: Applied the binomial probability formula with dbinom in R, labeled outcomes as "Suitable" (0–1 frosts) and "Unsuitable" (more than 1). Exported the table with write.csv and made a bar chart with barplot.

**Assignment 2)**

| ACCIDENTS | POISPROB | CUMPOISPROB |
|---|---|---|
| **0** | 0.22313 | 0.22313 |
| **1** | 0.334695 | 0.557825 |
| **2** | 0.251021 | 0.808847 |
| **3** | 0.125511 | 0.934358 |
| **4** | 0.047067 | 0.981424 |
| **5** | 0.01412 | 0.995544 |
| **6** | 0.00353 | 0.999074 |
| **7** | 0.000756 | 0.99983 |
| **8** | 0.000142 | 0.999972 |
| **9** | 2.36E-05 | 0.999996 |
| **10** | 3.55E-06 | 0.999999 |
| **11** | 4.84E-07 | **1** |
| **12** | 6.04E-08 | **1** |
| **13** | 6.97E-09 | **1** |
| **14** | 7.47E-10 | **1** |
| **15** | 7.47E-11 | **1** |
| **16** | 7.00E-12 | **1** |
| **17** | 6.18E-13 | **1** |
| **18** | 5.15E-14 | **1** |
| **19** | 4.07E-15 | **1** |

| | | |
|---|---|---|




The bar chart shows the distribution of probabilities across accident counts.

The bar with the highest probability is at 1 accident, which makes sense because the mean is 1.5, so the most likely outcome is close to 1.

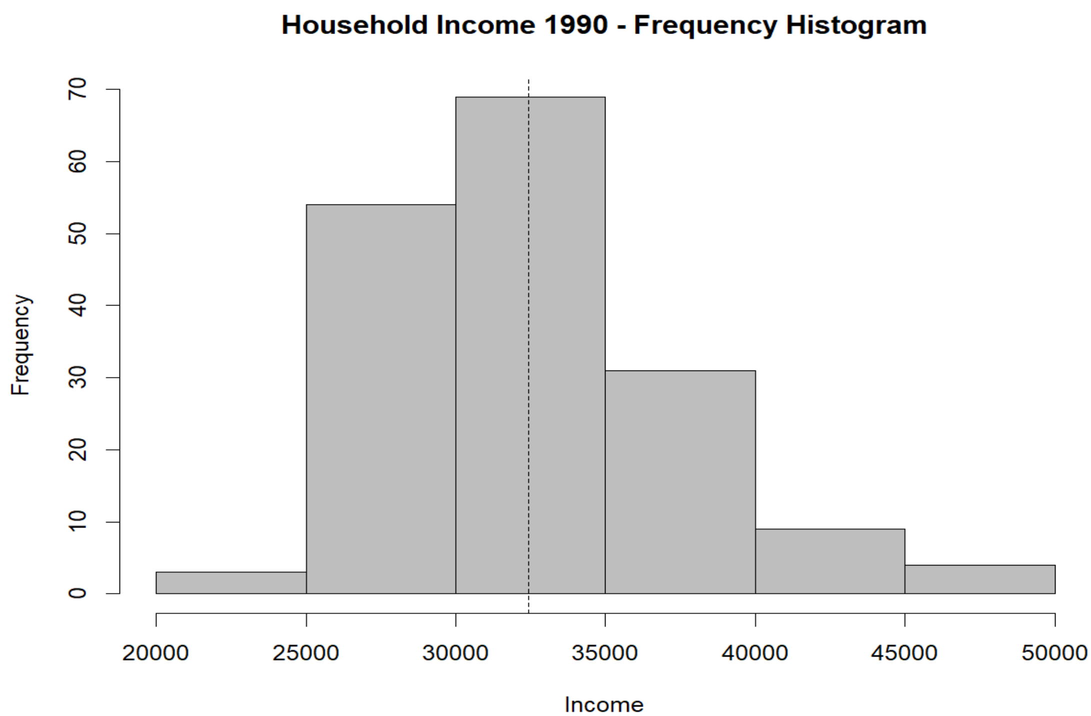
As the number of accidents increases, the probabilities quickly get smaller. The bars taper down and approach zero as the x-axis increases, which follows the Poisson pattern.

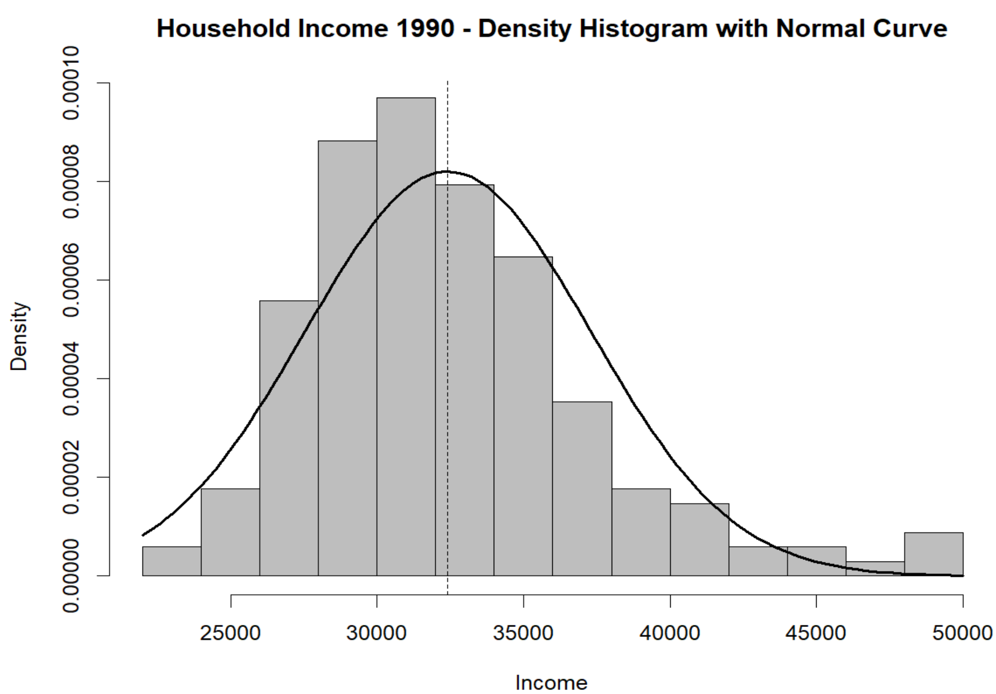
The probability of 4 or fewer accidents during the 30-day period is about 0.9814 (98.14%).

The probability of having between 5 and 10 accidents (inclusive) is about 0.0186 (1.86%).

Method used: Used dpois for Poisson probabilities and ppois for cumulative probabilities with 1.5. Saved the results as a CSV and made a bar chart with barplot. Found mode with which.max and calculated ranges with ppois

**Assignment 3)**



- Frequency histogram of household income in 1990 with the mean shown as a dashed line.
- Density histogram of 1990 household income with the fitted normal curve and mean line.

The distribution is centered around the mean income of about 32,400. The normal curve fits the density histogram well, with most incomes clustered near the center and fewer in the tails.

Household income data from 1990 were assumed to follow a normal distribution. I calculated the sample mean and standard deviation. Each value was standardized into a z-score, and probabilities were computed using the normal cumulative distribution function. Histograms were made to show the distribution, with the mean marked and the fitted normal curve overlaid.

- Mean income = 32,418.63
- Standard deviation = 4,864.05

From the z-scores and probabilities:

- The probability that 2020 income is greater than the 10th highest 1990 income = 0.04867 (4.9%).
- The probability that 2020 income is between the 41st and 100th 1990 incomes = 0.27389 (27.4%).

Method used:

Imported CSV with read.csv. Calculated mean and standard deviation using mean and sqrt(var). Standardized values into z-scores, then computed probabilities with pnorm. Plotted histograms with hist and overlaid a normal curve.