χ^2 Goodness-of-Fit Test

Homework 23 due Monday

Project data due Monday

Check Canvas for the reading assignment

χ^2 Goodness-of-Fit Test

- 1 categorical variable, > 2 categories
- Are the long-run proportions equal to a specific set of values?

Are pedestrians equally likely to be killed in car crashes on any day of the week?

- H_0: $\pi_{Sun} = \frac{1}{7}, \pi_{Mon} = \frac{1}{7}, \ldots \pi_{Sat} = \frac{1}{7}$
- H_a: At least one of the long-run probabilities is different from $\frac{1}{7}$.

Day	Number killed
Sun	752
Mon	734
Tues	689
Wed	674
Thurs	788
Fri	860
Sat	879
Total	**5376**

Test statistic

- $\chi^2 = \sum \frac{(Observed_i - Expected_i)^2}{Expected_i}$
- $E_i = n\pi_i$

- Example:

$$E_{Sun} = 5376 * \frac{1}{7} = 768$$

- Compute χ^2 on calculator, in Excel, or in applet

Day	Number killed
Sun	752
Mon	734
Tues	689
Wed	674
Thurs	788
Fri	860
Sat	879
Total	**5376**

How unusual is 49.05?

- χ^2 can't be negative
- Larger values -> stronger evidence against H_0

- Simulate: Spin a spinner, with probability 1/7 of landing on each day.
- Spin it 5376 times, record number of times it shows "Sun," "Mon", etc.
- Compute the simulated value of χ^2

Repeat 1000 times

- P-value = proportion of simulations with $\chi^2 \geq 49.05$

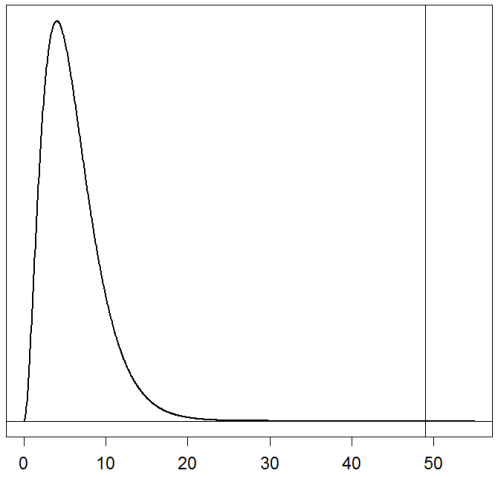
Theory-based method

- Validity conditions:
  - Same as χ^2 Test of Independence

- Degrees of freedom = $k - 1$
  - $k =$ number of categories

On TI-83

- First, compute expected counts.
- Compute χ^2 test statistic and degrees of freedom.
- 2^{nd} -> Distr -> χ^2cdf
- χ^2cdf(49.05, 10^{99}, 6)



On TI-84

- First, compute expected counts and degrees of freedom.
- Stat -> Edit -> Enter observed data in L1.
- Enter expected counts in L2.
- Stat -> Tests -> χ^2 GOF Test

Conclusion

- There is very strong evidence that at least one of the long-run probabilities of pedestrian fatalities is not equal to the specified value.

- There is very strong evidence that pedestrian fatalities are *not* equally likely on each day of the week.