Classical statistical inference

Iterated expectations[1]

a. Given $E[X|Y] = 2Y$ and $f(Y) = .5$ with $Y \in [-3, -1]$, what is $E[X]$?

b. Given $E[Z|H] = 15H - 10$ and $H \sim$ Bernoulli(.2), what is $E[Z]$?

Properties of estimators

1. Let $X_1, \ldots, X_n \sim$ Poisson(λ) and let $\hat{\lambda} = \frac{\sum_{i=1}^{n} X_i}{n}$. Find the bias, standard error, and MSE of this estimator.[2]

[1] Grimmer 13.1
[2] Wasserman 6.6.1

2. Let $X_1, \ldots, X_n \sim \text{Uniform}(0, \theta)$ and let $\hat{\theta} = 2\bar{X}_n$. Find the bias, standard error, and MSE of this estimator.[3]

Birds of a feather get their news on Twitter

A poll conducted in 2013 found that 52% of U.S. adult Twitter users get at least some news on Twitter. The standard error for this estimate was 2.4%, and a normal distribution may be used to model the sample proportion.[4]

a. Construct a 99% confidence interval for the fraction of U.S. adult Twitter users who get some news on Twitter, and interpret the confidence interval in context.

b. Identify the follow statements as true or false. Provide an explanation to justify each of your answers.

   a. The data provide statistically significant evidence that more than half of U.S. adult Twitter users get some news through Twitter. Use a significance level of $\alpha = 0.01$.

   b. Since the standard error is 2.4%, we can conclude that 97.6% of all U.S. adult Twitter users were included in the study.

[3] Wasserman 6.6.3
[4] OI 4.8 and 4.10

a. If we want to reduce the standard error of the estimate, we should collect less data.

a. If we construct a 90% confidence interval for the percentage of U.S. adults Twitter users who get some news through Twitter, this confidence interval will be wider than a corresponding 99% confidence interval.

Choose your own death

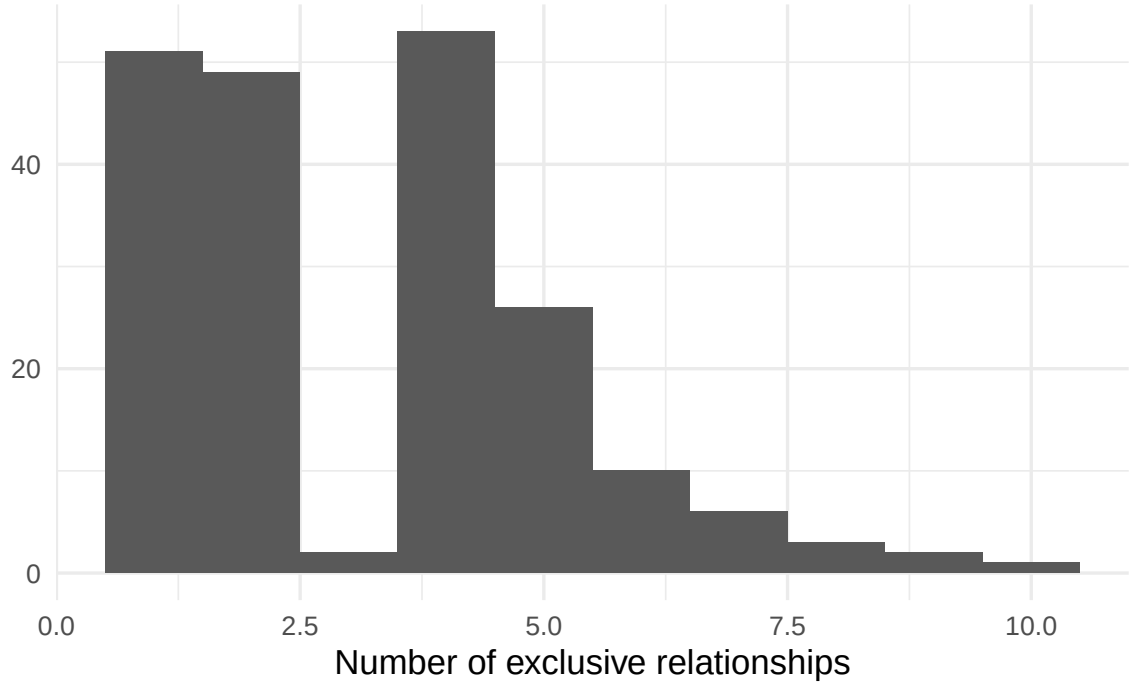
There is a theory that people can postpone their death until after an important event. To test the theory, Phillips and King (1988) collected data on deaths around the Jewish holiday Passover. Of 1919 deaths, 922 died the week before the holiday and 997 died the week after. Think of this as a binomial and test the null hypothesis that $\theta = \frac{1}{2}$. Report and interpret the p-value. Also construct a confidence interval for θ.[5]

[5]Wasserman 10.11.6

Dating on college campuses

A survey conducted on a reasonably random sample of 203 undergraduates asked, among many other questions, about the number of exclusive relationships these students have been in. The histogram below shows the distribution of the data from this sample.

The sample average is 3.2 with a standard deviation of 1.97.



Estimate the average number of exclusive relationships undergraduate students have been in using the Normal distribution and a 90% confidence interval and interpret this interval in context.[6]

Statistical significance

Determine whether the following statement is true or false, and explain your reasoning: "With large sample sizes, even small differences between the null value and the point estimate can be statistically significant."[7]

[6] OI 4.15
[7] OI 4.47

Sleep deprivation

New York is known as "the city that never sleeps". A random sample of 25 New Yorkers were asked how much sleep they get per night. Statistical summaries of these data are shown below. Do these data provide strong evidence that New Yorkers sleep less than 8 hours a night on average?[8]

n	$\bar{x}$	s	min	max
25	7.73	0.77	6.17	9.78

a. Write the hypotheses in symbols and in words.

b. Calculate the test statistic, T, and the associated degrees of freedom.

c. Find and interpret the p-value in this context.

[8] OI 5.7

d. What is the conclusion of the hypothesis test?

e. If you were to construct a 90% confidence interval that corresponded to this hypothesis test, would you expect 8 hours to be in the interval?

Interpreting public opinion polls

On June 28, 2012 the U.S. Supreme Court upheld the much debated 2010 healthcare law, declaring it constitutional. A Gallup poll released the day after this decision indicates that 46% of 1,012 Americans agree with this decision. At a 95% confidence level, this sample has a 3% margin of error. Based on this information, determine if the following statements are true or false, and explain your reasoning.[9]

a. We are 95% confident that between 43% and 49% of Americans in this sample support the decision of the U.S. Supreme Court on the 2010 healthcare law.

[9] OI 6.6

a. We are 95% confident that between 43% and 49% of Americans support the decision of the U.S. Supreme Court on the 2010 healthcare law.

b. If we considered many random samples of 1,012 Americans, and we calculated the sample proportions of those who support the decision of the U.S. Supreme Court, 95% of those sample proportions will be between 43% and 49%.

c. The margin of error at a 90% confidence level would be higher than 3%.

Survey Qs:

1. How prepared do you feel for the final exam? not prepared (1) 2 3 4 (5) very prepared
2. Which content area of the course was the easiest for you? Why? Please be specific.

3. Which content area of the course was the most challenging for you? Why? Please be specific.