

Day 2 Problem Set

Methods Camp | University of Texas at Austin

Casey Breen

Day 2 Problem Set

This problem set contains exercises from Day 2 that were originally done in small groups. To reinforce your understanding, please complete these exercises independently.

Exercise 1: Random Variables & Expectation

A survey asks respondents how many times they attended religious services last month. Suppose the distribution is:

Times attended (x)	0	1	2	4
$P(X = x)$	0.4	0.3	0.2	0.1

1.1 Is X discrete or continuous? Why?

1.2 Compute $E[X]$ by hand.

1.3 In your own words, what does $E[X]$ tell us?

1.4 If attendance is recoded as $Y = 2X + 1$, what is $E[Y]$? (Use the linearity property.)

Exercise 2: Distributions

2.1 For each scenario below, name the most appropriate distribution (Bernoulli, Binomial, Poisson, or Normal) and briefly justify your choice:

- a. Whether a single respondent owns a home (yes/no)
- b. The number of home-owners among 200 randomly sampled respondents
- c. The number of hate-crime incidents reported in a city per month
- d. Adult height in a large population

2.2 For scenario (b), if $p = 0.65$, what is $E[X]$?

Exercise 3: Regression

Suppose a researcher wants to examine occupational prestige score (Y) as a linear function of their parents' years of schooling (X) across all U.S. families.

3.1 Write the regression equation using index notation. What does the subscript i represent?

3.2 Interpret the slope coefficient.

3.3 Identify the estimand, a plausible estimator, and an example of an estimate.

3.4 Name one plausible factor that could be included in the error term, ε_i .