Homework

• Read chapter 1, section 2.1.
• Do problems 2.1, 2.3, 2.4, 2.5, 2.7 (pp. 59–60)
• Due Wednesday, February 9

Probability so far

• Classical vs. frequentist vs. subjectivist probability

• Sum rule:
  \[ P(A \text{ or } B) = P(A) + P(B) - P(AB) \]
or, for mutually exclusive events:
  \[ P(A \text{ or } B) = P(A) + P(B) \]

• Product rule:
  \[ P(A \text{ and } B) = P(A) \times P(B|A) \]
or, for independent events:
  \[ P(A \text{ and } B) = P(A) \times P(B) \]

• Bayes Theorem:
  \[ P(A|B) = \frac{P(A) \times P(B|A)}{P(B)} \]

Expected value

• Someone offers to bet you can roll a six on one die. Is this a good bet?

• The probabilities of various outcomes don’t depend on the bet, but the payoff does.

• The payoff, or expected value of a random variable, is its long-term average value:
  \[ E[X] = \sum x_i P(x_i) \]

Expected value

• Suppose in this game you bet $1 to win $3. The payoff of the game is:
  \[ E[W] = -1 \times \frac{5}{6} + 2 \times \frac{1}{6} = -\frac{1}{2} \]
### Expected value

- We can also talk about the expected value of a function of a random variable:

$$E[f(X)] = \sum_i f(x_i)P(x_i)$$

- Say $X$ is the number of pips on the die, and $f(X)$ is the payoff for each outcome. Then to have a fair game:

$$0 = E[f(X)] = \sum_i f(x_i)P(x_i)$$

$$= -1 \times \frac{5}{6} + a \times \frac{1}{6}$$

$$\frac{5}{6} = a \times \frac{1}{6}$$

$$5 = a$$

### Expected value

- We can work with expected values like we can work with probabilities:

$$E[X + Y] = \sum_i \sum_j (x_i + y_j)P(x_i, y_j)$$

$$= \sum_i \sum_j x_iP(x_i, y_j) + \sum_j y_jP(x_i, y_j)$$

$$= \sum_i x_i \sum_j P(x_i, y_j) + \sum_j y_j \sum_i P(x_i, y_j)$$

$$= \sum_i x_iP(x_i) + \sum_j y_jP(y_j)$$

$$= E[X] + E[Y]$$

### Variance

- The expected value is the center of a probability distribution function.
- The variance indicates how much the values vary from trial to trial.
- Specifically, it’s the square of the average difference between a value and the average value:

$$\text{Var}[X] = E[(X - E[X])^2]$$

$$= E[X^2] - (E[X])^2$$

- The expected value and the variance give an indication of the ‘shape’ of a random variable’s pmf
Chuck-a-luck

- Pick a number and roll three dice. Pays 1 : 1 if your number comes up once, 2 : 1 if your number comes up twice, and 3 : 1 if your number comes up three times.

- What’s the probability of it coming up once?

\[ P(D = 1) = \binom{3}{1} \times \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} = 3 \times \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} = \frac{75}{216} = 0.3472 \]

- What’s the probability of it coming up twice?

\[ P(D = 2) = \binom{3}{2} \times \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} = \frac{15}{216} = 0.0694 \]

- What’s the probability of it coming up three times?

\[ P(D = 3) = \binom{3}{3} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{216} = 0.0046 \]

Binomial Distribution

- The same reasoning can be applied to a wide range of problems: suppose you want to perform a sequence of Bernoulli trials: independent random trials with two possible outcomes

- Call one outcome ‘success’

- The probability of getting \( r \) successes out of \( n \) trials is given by the binomial distribution:

\[ p(r; n, p) = \binom{n}{r} p^r (1 - p)^{n-r} \]

- This is a parametric distribution: \( r \) is a random variable, \( n \) and \( p \) are parameters

- \( p(r; n, p) \) is sometimes written \( p(r|n, p) \), \( p_{n,p}(r) \), or \( p(r) \)

- Suppose a couple share a recessive gene for a dread disease. For each of their children, the chance of having the disease is one in four.

- If they have three children, what is the chance that at least one of them will be affected?

\[ P(R > 0; n = 3, p = 1/4) = 1 - P(R = 0; n = 3, p = 1/4) = 1 - \binom{3}{0} (1/4)^0 (3/4)^3 = 1 - \frac{27}{64} = \frac{37}{64} \approx 0.58 \]
Another way to get the same result:

\[ P(R > 0; n = 3, p = 1/4) = \sum_i P(R = i; n = 3, p = 1/4) \]
\[ = P(R = 1; n = 3, p = 1/4) + P(R = 2; n = 3, p = 1/4) + P(R = 3; n = 3, p = 1/4) \]
\[ = \frac{27}{64} + \frac{9}{64} + \frac{1}{64} \]
\[ = \frac{37}{64} \]

We can show that for a binomial random variable \( X \) with parameters \( n \) and \( p \), the expected value is:

\[ E[X] = np \]

And, the variance is:

\[ \text{Var}[X] = np(1 - p) \]

To show this, treat the binomial random variable as the sum of \( n \) Bernoulli variables.