CHAPTER 6

Some discrete distributions

6.1. Examples: Bernoulli, binomial, Poisson, geometric distributions

Bernoulli distribution

A random variable X such that $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$ is said to be a *Bernoulli random variable* with parameter p. Note $\mathbb{E}X = p$ and $\mathbb{E}X^2 = p$, so $\operatorname{Var} X = p - p^2 = p(1 - p)$.

We denote such a random variable by $X \sim \operatorname{Bern}(p)$.

Binomial distribution

A random variable X has a *binomial distribution* with parameters n and p if $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

We denote such a random variable by $X \sim \operatorname{Binom}(n, p)$.

The number of successes in n Bernoulli trials is a binomial random variable. After some cumbersome calculations one can derive $\mathbb{E}X = np$. An easier way is to realize that if X is binomial, then $X = Y_1 + \cdots + Y_n$, where the Y_i are independent Bernoulli variables, so $\mathbb{E}X = \mathbb{E}Y_1 + \cdots + \mathbb{E}Y_n = np$.

We have not defined yet what it means for random variables to be independent, but here we mean that the events such as $(Y_i = 1)$ are independent.

Proposition 6.1

Suppose $X := Y_1 + \cdots + Y_n$, where $\{Y_i\}_{i=1}^n$ are independent Bernoulli random variables with parameter p, then

$$\mathbb{E}X = np, \operatorname{Var} X = np(1 - p).$$

Proof. First we use the definition of expectation to see that

$$\mathbb{E}X = \sum_{i=0}^{n} i \binom{n}{i} p^i (1 - p)^{n-i} = \sum_{i=1}^{n} i \binom{n}{i} p^i (1 - p)^{n-i}.$$

Then

$$
\begin{aligned}
\mathbb{E}X &= \sum_{i=1}^{n} i \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} \frac{(n-1)!}{(i-1)!((n-1)-(i-1))!} p^{i-1}(1-p)^{(n-1)-(i-1)} \\
&= np \sum_{i=0}^{n-1} \frac{(n-1)!}{i!((n-1)-i)!} p^i (1-p)^{(n-1)-i} \\
&= np \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{(n-1)-i} = np,
\end{aligned}
$$

where we used the Binomial Theorem (Theorem 1.1).

To get the variance of X, we first observe that

$$\mathbb{E}X^2 = \sum_{i=1}^{n} \mathbb{E}Y_i^2 + \sum_{i \neq j} \mathbb{E}Y_i Y_j.$$

Now

$$
\begin{aligned}
\mathbb{E}Y_i Y_j &= 1 \cdot \mathbb{P}(Y_i Y_j = 1) + 0 \cdot \mathbb{P}(Y_i Y_j = 0) \\
&= \mathbb{P}(Y_i = 1, Y_j = 1) = \mathbb{P}(Y_i = 1)\mathbb{P}(Y_j = 1) = p^2
\end{aligned}
$$

using independence of random variables $\{Y_i\}_{i=1}^n$. Expanding $(Y_1 + \cdots + Y_n)^2$ yields n^2 terms, of which n are of the form Y_k^2. So we have $n^2 - n$ terms of the form $Y_i Y_j$ with $i \neq j$. Hence

$$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = np + (n^2 - n)p^2 - (np)^2 = np(1-p).$$

□

Later we will see that the variance of the sum of independent random variables is the sum of the variances, so we could quickly get $\operatorname{Var} X = np(1-p)$. Alternatively, one can compute $\mathbb{E}(X^2) - \mathbb{E}X = \mathbb{E}(X(X-1))$ using binomial coefficients and derive the variance of X from that.

Poisson distribution

A random variable X has the *Poisson distribution* with parameter λ if

$$\mathbb{P}(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}.$$

We denote such a random variable by $X \sim \operatorname{Pois}(\lambda)$. Note that

$$\sum_{i=0}^{\infty} \lambda^i / i! = e^{\lambda},$$

so the probabilities add up to one.

Proposition 6.2

Suppose X is a Poisson random variable with parameter λ, then

$$\mathbb{E}X = \lambda,$$
$$\operatorname{Var} X = \lambda.$$

PROOF. We start with the expectation

$$\mathbb{E}X = \sum_{i=0}^{\infty} i e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \lambda \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = \lambda.$$

Similarly one can show that

$$\begin{aligned}\mathbb{E}(X^2) - \mathbb{E}X = \mathbb{E}X(X-1) &= \sum_{i=0}^{\infty} i(i-1)e^{-\lambda}\frac{\lambda^i}{i!} \\ &= \lambda^2 e^{-\lambda} \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} \\ &= \lambda^2,\end{aligned}$$

so $\mathbb{E}X^2 = \mathbb{E}(X^2 - X) + EX = \lambda^2 + \lambda$, and hence $\operatorname{Var} X = \lambda$. □

Example 6.1. Suppose on average there are 5 homicides per month in a given city. What is the probability there will be at most 1 in a certain month?

Solution: If X is the number of homicides, we are given that $\mathbb{E}X = 5$. Since the expectation for a Poisson is λ, then $\lambda = 5$. Therefore $\mathbb{P}(X = 0) + \mathbb{P}(X = 1) = e^{-5} + 5e^{-5}$.

Example 6.2. Suppose on average there is one large earthquake per year in California. What's the probability that next year there will be exactly 2 large earthquakes?

Solution: $\lambda = \mathbb{E}X = 1$, so $\mathbb{P}(X = 2) = e^{-1}(\frac{1}{2})$.

We have the following proposition connecting binomial and Poisson distributions.

Proposition 6.3 (Binomial approximation of Poisson distribution)

If X_n is a binomial random variable with parameters n and p_n and $np_n \to \lambda$, then $\mathbb{P}(X_n = i) \to \mathbb{P}(Y = i)$, where Y is Poisson with parameter λ.

6.1 (Approximation of Poisson by binomials)

Note that by setting
$$p_n := \lambda/n \quad \text{for} \quad n > \lambda$$
we can approximate the Poisson distribution with parameter λ by binomial distributions with parameters n and p_n.

This proposition shows that the Poisson distribution models binomials when the probability of a success is small. The number of misprints on a page, the number of automobile accidents, the number of people entering a store, etc. can all be modeled by a Poisson distribution.

Proof. For simplicity, let us suppose that $\lambda = np_n$ for $n > \lambda$. In the general case we can use $\lambda_n = np_n \xrightarrow[n\to\infty]{} \lambda$. We write

$$\begin{aligned}
\mathbb{P}(X_n = i) &= \frac{n!}{i!(n-i)!} p_n^i (1-p_n)^{n-i} \\
&= \frac{n(n-1)\cdots(n-i+1)}{i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\
&= \frac{n(n-1)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}.
\end{aligned}$$

Observe that the following three limits exist

$$\begin{aligned}
&\frac{n(n-1)\cdots(n-i+1)}{n^i} \xrightarrow[n\to\infty]{} 1, \\
&(1-\lambda/n)^i \xrightarrow[n\to\infty]{} 1, \\
&(1-\lambda/n)^n \xrightarrow[n\to\infty]{} e^{-\lambda},
\end{aligned}$$

which completes the proof. □

In Section 2.2.3 we considered **discrete uniform distributions** with $\mathbb{P}(X = k) = \frac{1}{n}$ for $k = 1, 2, \ldots, n$. This is the distribution of the number showing on a die (with $n = 6$), for example.

Geometric distribution

A random variable X has the geometric distribution with parameter p, $0 < p < 1$, if
$$\mathbb{P}(X = i) = (1-p)^{i-1} p \quad \text{for} \quad i = 1, 2, \ldots.$$

Using a geometric series sum formula we see that

$$\sum_{i=1}^{\infty} \mathbb{P}(X = i) = \sum_{i=1}^{\infty} (1-p)^{i-1} p = \frac{1}{1-(1-p)} p = 1.$$

In Bernoulli trials, if we let X be the first time we have a success, then X will be a geometric random variable. For example, if we toss a coin over and over and X is the first time we get a heads, then X will have a geometric distribution. To see this, to have the first success occur on the k^{th} trial, we have to have $k-1$ failures in the first $k-1$ trials and then a success. The probability of that is $(1-p)^{k-1}p$.

Proposition 6.4

If X is a geometric random variable with parameter p, $0<p<1$, then

$$\begin{aligned}\mathbb{E}X &= \frac{1}{p},\\ \operatorname{Var}X &= \frac{1-p}{p^2},\\ F_X(k) &= \mathbb{P}(X \leqslant k) = 1-(1-p)^k.\end{aligned}$$

Proof. We will use

$$\frac{1}{(1-r)^2} = \sum_{n=0}^{\infty} nr^{n-1}$$

which we can show by differentiating the formula for geometric series $1/(1-r) = \sum_{n=0}^{\infty} r^n$. Then

$$\mathbb{E}X = \sum_{i=1}^{\infty} i \cdot \mathbb{P}(X=i) = \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1}p = \frac{1}{(1-(1-p))^2} \cdot p = \frac{1}{p}.$$

Then the variance

$$\operatorname{Var}X = \mathbb{E}(X-\mathbb{E}X)^2 = \mathbb{E}\left(X-\frac{1}{p}\right)^2 = \sum_{i=1}^{\infty}\left(i-\frac{1}{p}\right)^2 \cdot \mathbb{P}(X=i)$$

To find the variance we will use another sum. First

$$\frac{r}{(1-r)^2} = \sum_{n=0}^{\infty} nr^n,$$

which we can differentiate to see that

$$\frac{1+r}{(1-r)^3} = \sum_{n=1}^{\infty} n^2 r^{n-1}.$$

Then

$$\mathbb{E}X^2 = \sum_{i=1}^{\infty} i^2 \cdot \mathbb{P}(X=i) = \sum_{i=1}^{\infty} i^2 \cdot (1-p)^{i-1}p = \frac{(1+(1-p))}{(1-(1-p))^3} \cdot p = \frac{2-p}{p^2}.$$

Thus

$$\operatorname{Var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2}.$$

The cumulative distribution function (CDF) can be found by using the geometric series sum formula

$$1-F_X(k)=\mathbb{P}(X>k)=\sum_{i=k+1}^{\infty}\mathbb{P}(X=i)=\sum_{i=k+1}^{\infty}(1-p)^{i-1}p=\frac{(1-p)^k}{1-(1-p)}p=(1-p)^k.$$

$\square$

Negative binomial distribution

A random variable X has *negative binomial distribution* with parameters r and p if

$$\mathbb{P}(X=n)=\binom{n-1}{r-1}p^r(1-p)^{n-r},\quad n=r,r+1,\ldots.$$

A negative binomial represents the number of trials until r successes. To get the above formula, to have the r^{th} success in the n^{th} trial, we must exactly have $r-1$ successes in the first $n-1$ trials and then a success in the n^{th} trial.

Hypergeometric distribution

A random variable X has hypergeometric distribution with parameters m, n and N if

$$\mathbb{P}(X=i)=\frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}.$$

This comes up in sampling without replacement: if there are N balls, of which m are one color and the other $N-m$ are another, and we choose n balls at random without replacement, then X represents the probability of having i balls of the first color.

Another model where the hypergeometric distribution comes up is the probability of a success changes on each draw, since each draw decreases the population, in other words, when we consider sampling without replacement from a finite population). Then N is the population size, m is the number of success states in the population, n is the number of draws, that is, quantity drawn in each trial, i is the number of observed successes.

6.2. Further examples and applications

6.2.1. Bernoulli and binomial random variables.

Example 6.3. A company prices its hurricane insurance using the following assumptions:

(i) In any calendar year, there can be at most one hurricane.
(ii) In any calendar year, the probability of a hurricane is 0.05.
(iii) The numbers of hurricanes in different calendar years are mutually independent.

Using the company's assumptions, find the probability that there are fewer than 3 hurricanes in a 20-year period.

Solution: denote by X the number of hurricanes in a 20-year period. From the assumptions we see that $X \sim \operatorname{Binom}(20, 0.05)$, therefore

$$\begin{aligned}\mathbb{P}(X<3) &= \mathbb{P}(X \leqslant 2) \\ &= \binom{20}{0}(0.05)^0(0.95)^{20} + \binom{20}{1}(0.05)^1(0.95)^{19} + \binom{20}{2}(0.05)^2(0.95)^{18} \\ &= 0.9245.\end{aligned}$$

Example 6.4. Phan has a 0.6 probability of making a free throw. Suppose each free throw is independent of the other. If he attempts 10 free throws, what is the probability that he makes at least 2 of them?

Solution: If $X \sim \operatorname{Binom}(10, 0.6)$, then

$$\begin{aligned}\mathbb{P}(X \geqslant 2) &= 1 - \mathbb{P}(X=0) - \mathbb{P}(X=1) \\ &= 1 - \binom{10}{0}(0.6)^0(0.4)^{10} - \binom{10}{1}(0.6)^1(0.4)^9 \\ &= 0.998.\end{aligned}$$

6.2.2. The Poisson distribution. Recall that a Poisson distribution models well events that have a low probability and the number of trials is high. For example, the probability of a misprint is small and the number of words in a page is usually a relatively large number compared to the number of misprints.

(1) The number of misprints on a random page of a book.
(2) The number of people in community that survive to age 100.
(3) The number of telephone numbers that are dialed in an average day.
(4) The number of customers entering post office on an average day.

Example 6.5. Levi receives an average of two texts every 3 minutes. If we assume that the number of texts is Poisson distributed, what is the probability that he receives five or more texts in a 9-minute period?

Solution: Let X be the number of texts in a 9−minute period. Then $\lambda = 3 \cdot 2 = 6$ and

$$\begin{aligned}\mathbb{P}(X \geqslant 5) &= 1 - \mathbb{P}(X \leqslant 4) \\ &= 1 - \sum_{n=0}^{4} \frac{6^n e^{-6}}{n!} \\ &= 1 - 0.285 = 0.715.\end{aligned}$$

Example 6.6. Let X_1, ..., X_k be independent Poisson random variables, each with expectation λ. What is the distribution of the random variable $Y := X_1 + ... + X_k$?

Solution: The distribution of Y is Poisson with the expectation $\lambda = k\lambda$. To show this, we use Proposition 6.3 and (6.1) to choose $n = mk$ Bernoulli random variables with parameter $p_n = k\lambda_1/n = \lambda_1/m = \lambda/n$ to approximation the Poisson random variables. If we sum them all together, the limit as $n \to \infty$ gives us a Poisson distribution with expectation $\lim\limits_{n\to\infty} np_n = \lambda$. However, we can re-arrange the same $n = mk$ Bernoulli random variables in k groups, each group having m Bernoulli random variables. Then the limit gives us the distribution of $X_1 + ... + X_k$. This argument can be made rigorous, but this is beyond the scope of this course. Note that we do not show that the we have convergence in distribution.

Example 6.7. Let $X_1, \ldots, X_k$ be independent Poisson random variables, each with expectation $\lambda_1, \ldots, \lambda_k$, respectively. What is the distribution of the random variable $Y = X_1 + ... + X_k$?

Solution: The distribution of Y is Poisson with expectation $\lambda = \lambda_1 + ... + \lambda_k$. To show this, we again use Proposition 6.3 and (6.1) with parameter $p_n = \lambda/n$. If n is large, we can separate these n Bernoulli random variables in k groups, each having $n_i \approx \lambda_i n/\lambda$ Bernoulli random variables. The result follows if $\lim\limits_{n\to\infty} n_i/n = \lambda_i$ for each $i = 1, ..., k$.

This entire set-up, which is quite common, involves what is called *independent identically distributed Bernoulli random variables* (i.i.d. Bernoulli r.v.).

Example 6.8. Can we use binomial approximation to find the mean and the variance of a Poisson random variable?

Solution: Yes, and this is really simple. Recall again from Proposition 6.3 and (6.1) that we can approximate Poisson Y with parameter λ by a binomial random variable $\text{Binom}(n, p_n)$, where $p_n = \lambda/n$. Each such a binomial random variable is a sum on n independent Bernoulli random variables with parameter p_n. Therefore

$$\begin{aligned}\mathbb{E}Y &= \lim_{n\to\infty} np_n = \lim_{n\to\infty} n\frac{\lambda}{n} = \lambda, \\ \operatorname{Var}(Y) &= \lim_{n\to\infty} np_n(1-p_n) = \lim_{n\to\infty} n\frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right) = \lambda.\end{aligned}$$

6.2.3. Table of distributions. The following table summarizes the discrete distributions we have seen in this chapter. Here $\mathbb{N}$ stands for the set of positive integers, and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ is the set of nonnegative integers.

Name	Notation	Parameters	PMF ($k \in \mathbb{N}_0$)	$\mathbb{E}[X]$	$\mathrm{Var}(X)$
Bernoulli	$\mathrm{Bern}(p)$	$p \in [0,1]$	$\binom{1}{k} p^k (1-p)^{1-k}$	p	$p(1-p)$
Binomial	$\mathrm{Binom}(n,p)$	$n \in \mathbb{N}$ $p \in [0,1]$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	$\mathrm{Pois}(\lambda)$	$\lambda > 0$	$e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ
Geometric	$\mathrm{Geo}(p)$	$p \in (0,1)$	$\begin{cases} (1-p)^{k-1} p, & \text{for } k \geqslant 1, \\ 0, & \text{else.} \end{cases}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Negative binomial	$\mathrm{NBin}(r,p)$	$r \in \mathbb{N}$ $p \in (0,1)$	$\begin{cases} \binom{k-1}{r-1} p^r (1-p)^{k-r}, & \text{if } k \geq r, \\ 0, & \text{else.} \end{cases}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
Hyper-geometric	$\mathrm{Hyp}(N,m,n)$	$N \in \mathbb{N}_0$ $n, m \in \mathbb{N}_0$	$\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$	$\frac{nm}{N}$	$\frac{nm(N-n)}{N(N-1)}(1-\frac{m}{N})$

6.3. Exercises

Exercise 6.1. A UConn student claims that she can distinguish Dairy Bar ice cream from Friendly's ice cream. As a test, she is given ten samples of ice cream (each sample is either from the Dairy Bar or Friendly's) and asked to identify each one. She is right eight times. What is the probability that she would be right exactly eight times if she guessed randomly for each sample?

Exercise 6.2. A Pharmaceutical company conducted a study on a new drug that is supposed to treat patients suffering from a certain disease. The study concluded that the drug did not help 25% of those who participated in the study. What is the probability that of 6 randomly selected patients, 4 will recover?

Exercise 6.3. 20% of all students are left-handed. A class of size 20 meets in a room with 18 right-handed desks and 5 left-handed desks. What is the probability that every student will have a suitable desk?

Exercise 6.4. A ball is drawn from an urn containing 4 blue and 5 red balls. After the ball is drawn, it is replaced and another ball is drawn. Suppose this process is done 7 times.

(a) What is the probability that exactly 2 red balls were drawn in the 7 draws?
(b) What is the probability that at least 3 blue balls were drawn in the 7 draws?

Exercise 6.5. The expected number of typos on a page of the new Harry Potter book is 0.2. What is the probability that the next page you read contains

(a) 0 typos?
(b) 2 or more typos?
(c) Explain what assumptions you used.

Exercise 6.6. The monthly average number of car crashes in Storrs, CT is 3.5. What is the probability that there will be

(a) at least 2 accidents in the next month?
(b) at most 1 accident in the next month?
(c) Explain what assumptions you used.

Exercise 6.7. Suppose that, some time in a distant future, the average number of burglaries in New York City in a week is 2.2. Approximate the probability that there will be

(a) no burglaries in the next week;
(b) at least 2 burglaries in the next week.

Exercise 6.8. The number of accidents per working week in a particular shipyard is Poisson distributed with mean 0.5. Find the probability that:

(a) In a particular week there will be at least 2 accidents.

(b) In a particular two week period there will be exactly 5 accidents.
(c) In a particular month (i.e. 4 week period) there will be exactly 2 accidents.

Exercise 6.9. Jennifer is baking cookies. She mixes 400 raisins and 600 chocolate chips into her cookie dough and ends up with 500 cookies.

(a) Find the probability that a randomly picked cookie will have three raisins in it.
(b) Find the probability that a randomly picked cookie will have at least one chocolate chip in it.
(c) Find the probability that a randomly picked cookie will have no more than two bits in it (a bit is either a raisin or a chocolate chip).

Exercise 6.10. A roulette wheel has 38 numbers on it: the numbers 0 through 36 and a 00. Suppose that Lauren always bets that the outcome will be a number between 1 and 18 (including 1 and 18).

(a) What is the probability that Lauren will lose her first 6 bets.
(b) What is the probability that Lauren will first win on her sixth bet?

Exercise 6.11. In the US, albinism occurs in about one in 17,000 births. Estimate the probabilities no albino person, of at least one, or more than one albino at a football game with 5,000 attendants. Use the Poisson approximation to the binomial to estimate the probability.

Exercise 6.12. An egg carton contains 20 eggs, of which 3 have a double yolk. To make a pancake, 5 eggs from the carton are picked at random. What is the probability that at least 2 of them have a double yolk?

Exercise 6.13. Around 30,000 couples married this year in CT. Approximate the probability that at least in one of these couples

(a) both partners have birthday on January 1st.
(b) both partners celebrate birthday in the same month.

Exercise 6.14. A telecommunications company has discovered that users are three times as likely to make two-minute calls as to make four-minute calls. The length of a typical call (in minutes) has a Poisson distribution. Find the expected length (in minutes) of a typical call.

6.4. Selected solutions

Solution to Exercise 6.1: This should be modeled using a binomial random variable X, since there is a sequence of trials with the same probability of success in each one. If she guesses randomly for each sample, the probability that she will be right each time is $\frac{1}{2}$. Therefore,

$$\mathbb{P}(X=8) = \binom{10}{8}\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^2 = \frac{45}{2^{10}}.$$

Solution to Exercise 6.2: $\binom{6}{4}(0.75)^4(0.25)^2$

Solution to Exercise 6.3: For each student to have the kind of desk he or she prefers, there must be no more than 18 right-handed students and no more than 5 left-handed students, so the number of left-handed students must be between 2 and 5 (inclusive). This means that we want the probability that there will be 2, 3, 4, or 5 left-handed students. We use the binomial distribution and get

$$\sum_{i=2}^{5}\binom{20}{i}\left(\frac{1}{5}\right)^i\left(\frac{4}{5}\right)^{20-i}.$$

Solution to Exercise 6.4(A):

$$\binom{7}{2}\left(\frac{5}{9}\right)^2\left(\frac{4}{9}\right)^5$$

Solution to Exercise 6.4(B):

$$\begin{aligned}\mathbb{P}(X\geqslant 3) &= 1-\mathbb{P}(X\leqslant 2)\\ &= 1-\binom{7}{0}\left(\frac{4}{9}\right)^0\left(\frac{5}{9}\right)^7-\binom{7}{1}\left(\frac{4}{9}\right)^1\left(\frac{5}{9}\right)^6-\binom{7}{2}\left(\frac{4}{9}\right)^2\left(\frac{5}{9}\right)^5\end{aligned}$$

Solution to Exercise 6.5(A): $e^{-0.2}$

Solution to Exercise 6.5(B): $1-e^{-0.2}-0.2e^{-0.2}=1-1.2e^{-0.2}$.

Solution to Exercise 6.5(C): Since each word has a small probability of being a typo, the number of typos should be approximately Poisson distributed.

Solution to Exercise 6.6(A): $1-e^{-3.5}-3.5e^{-3.5}=1-4.5e^{-3.5}$

Solution to Exercise 6.6(B): $4.5e^{-3.5}$

Solution to Exercise 6.6(C): Since each accident has a small probability it seems reasonable to suppose that the number of car accidents is approximately Poisson distributed.

Solution to Exercise 6.7(A): $e^{-2.2}$

Solution to Exercise 6.7(B): $1-e^{-2.2}-2.2e^{-2.2}=1-3.2e^{-2.2}$.

Solution to Exercise 6.8(A): We have

$$\mathbb{P}(X \geqslant 2) = 1 - \mathbb{P}(X \leqslant 1) = 1 - e^{-0.5}\frac{(0.5)^0}{0!} - e^{-0.5}\frac{(0.5)^1}{1!}.$$

Solution to Exercise 6.8(B): In two weeks the average number of accidents will be $\lambda = 0.5 + 0.5 = 1$. Then $\mathbb{P}(X = 5) = e^{-1}\frac{1^5}{5!}$.

Solution to Exercise 6.8(C): In a 4 week period the average number of accidents will be $\lambda = 4 \cdot (0.5) = 2$. Then $\mathbb{P}(X = 2) = e^{-2}\frac{2^2}{2!}$.

Solution to Exercise 6.9(A): This calls for a Poisson random variable R. The average number of raisins per cookie is 0.8, so we take this as our λ . We are asking for $\mathbb{P}(R = 3)$, which is $e^{-0.8}\frac{(0.8)^3}{3!} \approx 0.0383$.

Solution to Exercise 6.9(B): This calls for a Poisson random variable C. The average number of chocolate chips per cookie is 1.2, so we take this as our λ. We are asking for $\mathbb{P}(C \geqslant 1)$, which is $1 - \mathbb{P}(C = 0) = 1 - e^{-1.2}\frac{(1.2)^0}{0!} \approx 0.6988$.

Solution to Exercise 6.9(C): This calls for a Poisson random variable B. The average number of bits per cookie is $0.8 + 1.2 = 2$, so we take this as our λ. We are asking for $\mathbb{P}(B \leqslant 2)$, which is $\mathbb{P}(B = 0) + \mathbb{P}(B = 1) + \mathbb{P}(B = 2) = e^{-2}\frac{2^0}{0!} + e^{-2}\frac{2^1}{1!} + e^{-2}\frac{2^2}{2!} \approx .6767$.

Solution to Exercise 6.10(A): $\left(1 - \frac{18}{38}\right)^6$

Solution to Exercise 6.10(B): $\left(1 - \frac{18}{38}\right)^5 \frac{18}{38}$

Solution to Exercise 6.11 Let X denote the number of albinos at the game. We have that $X \sim \text{Binom}(5000, p)$ with $p = 1/17000 \approx 0.00029$. The binomial distribution gives us

$\mathbb{P}(X = 0) = \left(\frac{16999}{17000}\right)^{5000} \approx 0.745$ $\qquad$ $\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - \left(\frac{16999}{17000}\right)^{5000} \approx 0.255$

$\mathbb{P}(X > 1) = \mathbb{P}(X \geqslant 1) - \mathbb{P}(X = 1) =$

$$= 1 - \left(\frac{16999}{17000}\right)^{5000} - \begin{pmatrix} 5000 \\ 1 \end{pmatrix} \left(\frac{16999}{17000}\right)^{4999} \left(\frac{1}{17000}\right)^1 \approx 0.035633$$

Approximating the distribution of X by a Poisson with parameter $\lambda = \frac{5000}{17000} = \frac{5}{17}$ gives

$\mathbb{P}(Y = 0) = \exp\left(-\frac{5}{17}\right) \approx 0.745$ $\qquad$ $\mathbb{P}(Y \geqslant 1) = 1 - \mathbb{P}(Y = 0) = 1 - \exp\left(-\frac{5}{17}\right) \approx 0.255$

$\mathbb{P}(Y > 1) = \mathbb{P}(Y \geqslant 1) - \mathbb{P}(Y = 1) = 1 - \exp\left(-\frac{5}{17}\right) - \exp\left(-\frac{5}{17}\right)\frac{5}{17} \approx 0.035638$

Solution to Exercise 6.12: Let X be the random variable that denotes the number of eggs with double yolk in the set of chosen 5. Then $X \sim \text{Hyp}(20, 3, 5)$ and we have that

$$\mathbb{P}(X \geqslant 2) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \frac{\binom{3}{2} \cdot \binom{17}{3}}{\binom{20}{5}} + \frac{\binom{3}{3} \cdot \binom{17}{2}}{\binom{20}{5}}.$$

Solution to Exercise 6.13: We will use Poisson approximation.

(a) The probability that both partners have birthday on January 1st is $p = \frac{1}{365^2}$. If X denotes the number of married couples where this is the case, we can approximate the distribution of X by a Poisson with parameter $\lambda = 30,000 \cdot 365^{-2} \approx 0.2251$. Hence, $\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - e^{-0.2251}$.

(b) In this case, the probability of both partners celebrating birthday in the same month is $1/12$ and therefore we approximate the distribution by a Poisson with parameter $\lambda = 30,000/12 = 2500$. Thus, $\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - e^{-2500}$.

Solution to Exercise 6.14: Let X denote the duration (in minutes) of a call. By assumption, $X \sim \mathrm{Pois}(\lambda)$ for some parameter $\lambda > 0$, so that the expected duration of a call is $\mathbb{E}[X] = \lambda$. In addition, we know that $\mathbb{P}(X = 2) = 3\mathbb{P}(X = 4)$, which means

$$e^{-\lambda}\frac{\lambda^2}{2!} = 3e^{-\lambda}\frac{\lambda^4}{4!}.$$

From here we deduce that $\lambda^2 = 4$ and hence $\mathbb{E}[X] = \lambda = 2$.