Math 431: Homework 7 Solutions

1. **Exercise 4.1**

   For both cases we define

   $$X = \text{Number of students born in January},$$

   so $X \sim \text{Binomial}(1200, p)$. We want to approximate $P(X > 130)$.

   (a) Under the assumption months are equally likely to contain birthdays, $p = \frac{1}{12}$. So

   $$\mu = \frac{1200}{12} = 100, \quad \sigma = \sqrt{\frac{1100}{12}} \approx 9.574.$$

   So we can approximate

   $$\begin{aligned} P(X > 13) &= 1 - P(X \leq 130) \\ &\approx 1 - P\left(Z \leq \frac{130 - 100}{9.574}\right) \\ &\approx 1 - \Phi(3.13) \\ &= 0.0009. \end{aligned}$$

   (b) Under the assumption all days are equally likely to contain birthdays, $p = \frac{31}{365}$.So

   $$\mu = \frac{1200 \cdot 31}{365} = 101.9178, \quad \sigma = \sqrt{1200 \cdot \frac{31}{365} \cdot \frac{334}{365}} \approx 9.6572.$$

   So we can approximate

   $$\begin{aligned} P(X > 13) &= 1 - P(X \leq 130) \\ &\approx 1 - P\left(Z \leq \frac{130 - 101.9175}{9.6572}\right) \\ &\approx 1 - \Phi(2.91) \\ &= 0.0018. \end{aligned}$$

2. **Exercise 4.4**

   We start defining

   $$S = \text{Number of size one steps that Liz takes after 90 rolls}.$$

   So $S \sim \text{Binomial}(90, 1/3)$, and X can be approximated using a normal distribution with

   $$\mu = 30 \quad \sigma = \sqrt{90 \cdot \frac{1}{3} \cdot \frac{2}{3}} = \sqrt{20} \approx 4.472.$$

We can express

$$X_{90} = S + 2(90 - S) = 180 - S.$$

So for our computation we have

$$\begin{aligned} P(X_{90} \geq 160) &= P(180 - S \geq 160) \\ &= P(S \leq 20) \\ &\approx \Phi\left(\frac{20-30}{4.472}\right) \\ &\approx \Phi(-2.24) \\ &\approx 0.0125. \end{aligned}$$

3. **Exercise 4.7**

Let

$$X = \text{Number of people in survey who will vote for candidate A.}$$

So $X \sim \text{Binomial}(1000, p)$. We want a 95% confidence interval, for $n = 1000$. This means we must use

$$P(|\hat{p} - p| < \varepsilon) \geq 0.95$$

to determine ε.

$$\begin{aligned} P(|\hat{p} - p| < \varepsilon) &= P\left(np - n\varepsilon < \frac{X}{n} < np + \varepsilon\right) \\ &\approx P\left(-\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}} < Z < \frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) \\ &= 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) - 1 \end{aligned}$$

So

$$2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) - 1 \geq 0.95 \quad \Rightarrow \quad \varepsilon \leq \frac{\sqrt{p(1-p)}}{\sqrt{n}} \cdot \Phi^{-1}(0.975)$$

The maximum value of $\sqrt{p(1-p)}$ is $1/2$ and $n = 1000$. So we substitue these values into our expression to find

$$\varepsilon \leq \frac{1}{2} \cdot \frac{1}{\sqrt{1000}} \cdot 1.96 = 0.031.$$

The 95% confidence interval is thus

$$(\hat{p} - \varepsilon, \hat{p} + \varepsilon) = \left(\frac{457}{1000} - 0.031, \frac{457}{1000} + 0.031\right) = (0.426, 0.488).$$

4. **Exercise 4.11**

We assume that the on each page there is a large number of opportunities for typos, but the probability of making a typo at each opportunity is small. We also must assume typos are made independently of each other. In this case we can treat the number of typos on each page as a Poisson(λ) random variable. For this problem we are given $\lambda = 6$.

We let X denote the number of typos on page 301. We thus have

$$P(X \geq 4) = 1 - P(X \leq 3) \approx 1 - \sum_{\ell=0}^{3} e^{-6} \frac{6^\ell}{\ell!} \approx 0.8488.$$

5. **Exercise 4.26**

Let X denote the number of people interview who prefer whole milk to skim milk. So $X \sim \text{Binomial}(n, p)$. Our survey is of 100 individuals, so $n = 100$. We want a confidence interval of the form

$$(\hat{p} - 0.1, \hat{p} + 0.1),$$

so our margin of error is $\varepsilon = 0.1$. This means we must use

$$P(|\hat{p} - p| < \varepsilon) \geq C$$

to determine C.

$$\begin{aligned}
P(|\hat{p} - p| < \varepsilon) &= P\left(np - n\varepsilon < \frac{X}{n} < np + \varepsilon\right) \\
&\approx P\left(-\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}} < Z < \frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) \\
&= 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) - 1 \\
&\geq 2\Phi\left(2\sqrt{n}\varepsilon\right) - 1
\end{aligned}$$

For the last inequality we used $\sqrt{p(1-p)} \leq 1/2$. So

$$2\Phi\left(2\sqrt{n}\varepsilon\right) - 1 = 2\Phi(2) - 1 \approx 0.9544.$$

To review, we have shown that

$$P(|\hat{p} - p| < 0.1) \geq 0.9544.$$

The is is 95.44% confidence interval.

6. **Exercise 4.31**

We have

$$\begin{aligned}
E\left[\frac{1}{1+X}\right] &= \sum_{k=0}^{\infty} \frac{1}{k+1} e^{-\mu} \frac{\mu^k}{k!} = \frac{1}{\mu} \sum_{k=0}^{\infty} e^{-\mu} \frac{\mu^{k+1}}{(k+1)!} \\
&= \frac{1}{\mu} \sum_{\ell=1}^{\infty} e^{-\mu} \frac{\mu^\ell}{\ell!} = \frac{1 - e^{-\mu}}{\mu}.
\end{aligned}$$

We introduced $\ell = k + 1$ and used $\sum_{\ell=1}^{\infty} e^{-\mu} \frac{\mu^\ell}{\ell!} = 1 - e^{-\mu}$.

7. **Exercise 4.35**

   We know that $X \sim \text{Binomial}(365, p)$ where p is the probability that all ten flips are heads or all ten are tails. So $p = 2/2^{10} = 1/512$.

   (a) We use the Binomial PMF.

   $$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{365}{0}\left(\frac{1}{512}\right)^0\left(\frac{511}{512}\right)^{365} + \binom{365}{1}\left(\frac{1}{512}\right)^1\left(\frac{511}{512}\right)^{365} \\ &\approx 0.1601998. \end{aligned}$$

   (b) We use the Poisson approximation with $Y \sim \text{Poisson}(365/512)$. We are motivated to use this because $p = 1/512$ is very small. We can thoroughly justify the approximation using the total error bound,

   $$np^2 = \frac{365}{2^{18}} \approx 0.001392365.$$

   So Poisson a good estimate.

   $$\begin{aligned} P(X > 1) &= 1 - P(Y \leq 1) = 1 - P(Y = 0) - P(Y = 1) \\ &= 1 - e^{-365/512}\frac{(365/512)^0}{0!} - e^{-365/512}\frac{(365/512)^1}{1!} \qquad \approx 0.160298. \end{aligned}$$

8. **Exercise 4.42**

   (a) Let X be the number of mildly defective gadgets in the box. Then $X \sim \text{Bin}(n, p)$ with $n = 100$ and $p = 0.2 = \frac{1}{5}$. We have

   $$P(A) = P(X < 15) = \sum_{k=0}^{14} \binom{100}{k} (1/5)^k (4/5)^{100-k}.$$

   (b) We have $np(1 - p) = 16 > 10$ and $np^2 = 4$. This suggests that the normal approximation is more appropriate than the Poisson approximation in this case.
   Using normal approximation we get

   $$\begin{aligned} P(X < 15) &= P\left(\frac{X - 100 \cdot \frac{1}{5}}{\sqrt{100 \cdot \frac{1}{5} \cdot \frac{4}{5}}} < \frac{15 - 100 \cdot \frac{1}{5}}{\sqrt{100 \cdot \frac{1}{5} \cdot \frac{4}{5}}}\right) \\ &= P\left(\frac{X - 100 \cdot \frac{1}{5}}{\sqrt{100 \cdot \frac{1}{5} \cdot \frac{4}{5}}} < -\frac{5}{4}\right) \\ &\approx \Phi(-1.25) = 1 - \Phi(1.25) \approx 1 - 0.8944 = 0.1056. \end{aligned}$$

   With continuity correction we would get $\Phi(-1.375) = 1 - \Phi(1.375) \approx 0.08455$ (using linear interpolation to get $\Phi(1.375)$).
   The actual value is 0.0804437 (calculated with a computer).