3.105. (a) In this problem, a hypergeometric distribution applies because the population of individuals is finite and we sample at random without replacement. We have

$$\begin{aligned} N &= \text{total number of individuals} = 8 \\ r &= \text{number who had internships (“Class 1”)} = 5 \\ N - r &= \text{number who had student teaching (“Class 2”)} = 3. \end{aligned}$$

We sample $n = 3$ individuals from this population and let

$$Y = \text{ number of individuals who had internships (i.e., number of “Class 1 objects”).}$$

We have $Y \sim \text{hyper}(N = 8, n = 3, r = 5)$. The pmf of Y is

$$p_Y(y) = \begin{cases} \dfrac{\binom{5}{y}\binom{3}{3-y}}{\binom{8}{3}}, & y = 0, 1, 2, 3, \\ 0, & \text{otherwise.} \end{cases}$$

(b) We want $P(Y \geq 2)$. This is

$$P(Y \geq 2) = P(Y = 2) + P(Y = 3) = \frac{\binom{5}{2}\binom{3}{1}}{\binom{8}{3}} + \frac{\binom{5}{3}\binom{3}{0}}{\binom{8}{3}} = \frac{30}{56} + \frac{10}{56} \approx 0.714.$$

```
> dhyper(2,5,3,3)+dhyper(3,5,3,3)
[1] 0.7142857
```

(c) Use the formulas

$$\begin{aligned} E(Y) &= n\left(\frac{r}{N}\right) \\ V(Y) &= n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right) \end{aligned}$$

with $N = 8$, $n = 3$, and $r = 5$. The mean number of individuals who had internships in the sample is

$$E(Y) = 3\left(\frac{5}{8}\right) = \frac{15}{8} = 1.875.$$

The variance of Y is

$$3\left(\frac{5}{8}\right)\left(\frac{3}{8}\right)\left(\frac{5}{7}\right) = \frac{225}{448} \text{ (individuals)}^2.$$

The standard deviation of Y is therefore

$$\sqrt{\frac{225}{448}} \approx 0.709 \text{ individuals.}$$

3.108. In this problem, a hypergeometric distribution applies because the population of cameras is finite and we sample at random without replacement. We have

$$\begin{aligned} N &= \text{total number of cameras} = 20 \\ r &= \text{number of defective cameras (“Class 1”)} = 3 \\ N - r &= \text{number of cameras not defective (“Class 2”)} = 17. \end{aligned}$$

We sample n cameras from this population and let

$$Y = \text{ number of defective cameras (i.e., number of "Class 1 objects").}$$

We have $Y \sim \text{hyper}(N = 20, n, r = 3)$. The pmf of Y is

$$p_Y(y) = \begin{cases} \dfrac{\binom{3}{y}\binom{17}{n-y}}{\binom{20}{n}}, & y = 0, 1, 2, ..., n, \\ 0, & \text{otherwise.} \end{cases}$$

The problem asks you to find the smallest number of n such that $P(Y \geq 1) \geq 0.8$. Note that by the complement rule,

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \frac{\binom{3}{0}\binom{17}{n}}{\binom{20}{n}} = 1 - \frac{\binom{17}{n}}{\binom{20}{n}}.$$

Therefore, let's just calculate this quantity for $n = 1, 2, 3, ...,$ and find the smallest n that satisfies the ≥ 0.8 requirement. I did this in R:

```
> n = seq(1,10,1)
> prob = 1-(choose(17,n)/choose(20,n))
> cbind(n,prob)
       n      prob
 [1,]  1 0.1500000
 [2,]  2 0.2842105
 [3,]  3 0.4035088
 [4,]  4 0.5087719
 [5,]  5 0.6008772
 [6,]  6 0.6807018
 [7,]  7 0.7491228
 [8,]  8 0.8070175
 [9,]  9 0.8552632
[10,] 10 0.8947368
```

Therefore, we need to sample at least 8 cameras (from 20) to satisfy $P(Y \geq 1) \geq 0.8$.

3.113. In this problem, a hypergeometric distribution applies because the population of jurors is finite and we sample at random without replacement. We have

$$\begin{aligned} N &= \text{ total number of jurors} = 20 \\ r &= \text{ number of AA jurors ("Class 1")} = 8 \\ N - r &= \text{ number of white jurors ("Class 2")} = 12. \end{aligned}$$

We sample $n = 6$ jurors from this population and let

$$Y = \text{ number of AA jurors (i.e., number of "Class 1 objects").}$$

We have $Y \sim \text{hyper}(N = 20, n = 6, r = 8)$. The pmf of Y is

$$p_Y(y) = \begin{cases} \dfrac{\binom{8}{y}\binom{12}{6-y}}{\binom{20}{6}}, & y = 0, 1, 2, ..., 6, \\ 0, & \text{otherwise.} \end{cases}$$

The question asks you to think about how unusual it would be to get a jury of no more than 1 AA juror. We can calculate the probability of getting a jury of 1 or fewer AA jurors as follows:

$$P(Y \leq 1) = P(Y=0) + P(Y=1) = \frac{\binom{8}{0}\binom{12}{6}}{\binom{20}{6}} + \frac{\binom{8}{1}\binom{12}{5}}{\binom{20}{6}} \approx 0.024 + 0.163 = 0.187.$$

```
> dhyper(0,8,12,6)+dhyper(1,8,12,6)
[1] 0.1873065
```

Therefore, **provided that we sample at random** and without replacement, the probability of getting a jury that contains no more than 1 AA juror is nearly 1 out of 5. This is certainly not unusual.

3.120. Capture-recapture studies are commonly used to sample animal populations and to estimate the size of these populations. The basic capture-recapture design utilizes the hypergeometric distribution. Conceptualize a population of N animals (e.g., the number of female giraffes in a certain African region), where N is unknown. We can bifurcate the population into one of two classes: tagged and non-tagged. Let

$$\begin{aligned} N &= \text{number of animals in the population; i.e., the "population size"} \\ k &= \text{number of tagged animals ("Class 1")} \\ N-k &= \text{number of animals in the population not tagged ("Class 2").} \end{aligned}$$

We are told $k = 4$ animals are tagged (Class 1), so there $N - 4$ animals in the population that are not tagged (Class 2). A sample of $n = 3$ animals is later caught; let

$$Y = \text{ number of tagged animals out of 3.}$$

Then $Y \sim \text{hyper}(N, r = k = 4, n = 3)$. The pmf of Y is

$$p_Y(y) = \begin{cases} \dfrac{\binom{4}{y}\binom{N-4}{3-y}}{\binom{N}{3}}, & y = 0,1,2,3, \\ 0, & \text{otherwise.} \end{cases}$$

We want to find $P(Y = 1)$ as a function of N, the unknown population size. Using the pmf, this is

$$P(Y=1) = \frac{\binom{4}{1}\binom{N-4}{2}}{\binom{N}{3}} = \frac{4\frac{(N-4)(N-5)}{2}}{\frac{N(N-1)(N-2)}{6}} = \frac{12(N-4)(N-5)}{N(N-1)(N-2)}.$$

We now want to find the population size N that maximizes $P(Y = 1)$; i.e., if we did observe 1 tagged animal in our sample of 3, what population size N is most consistent with this outcome? Let's just calculate $P(Y = 1)$ for $N = 5, 6, 7, ...$, and find the N that makes this probability as large as possible. I did this in R:

```
> N = seq(5,15,1)
> prob = (12*(N-4)*(N-5))/(N*(N-1)*(N-2))
> cbind(N,prob)
      N      prob
 [1,]  5 0.0000000
```

```
 [2,]  6 0.2000000
 [3,]  7 0.3428571
 [4,]  8 0.4285714
 [5,]  9 0.4761905
 [6,] 10 0.5000000
 [7,] 11 0.5090909
 [8,] 12 0.5090909
 [9,] 13 0.5034965
[10,] 14 0.4945055
[11,] 15 0.4835165
```

This suggests both $N = 11$ and $N = 12$ maximize $P(Y = 1)$. Therefore, if we did observe 1 tagged animal in the sample of 3, we would estimate the population size to be 11 or 12.

3.127. Let Y denote the number of typing errors per page. Assume $Y \sim \text{Poisson}(\lambda = 4)$. The typesetter does not need to retype the page when $\{Y \leq 4\}$ occurs. Therefore,

$$\begin{aligned} P(Y \leq 4) &= P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) + P(Y = 4) \\ &= \frac{4^0 e^{-4}}{0!} + \frac{4^1 e^{-4}}{1!} + \frac{4^2 e^{-4}}{2!} + \frac{4^3 e^{-4}}{3!} + \frac{4^4 e^{-4}}{4!} \\ &\approx 0.018 + 0.073 + 0.147 + 0.195 + 0.195 = 0.628. \end{aligned}$$

In R,

```
> ppois(4,4)
[1] 0.6288369
```

3.132. Let Y denote the number of automobiles entering a mountain tunnel during a two minute period. We are basically instructed to assume $Y \sim \text{Poisson}(\lambda = 1)$. Under this assumption, we want to calculate $P(Y > 3)$. Note that

$$\begin{aligned} P(Y \leq 3) &= P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) \\ &= \frac{1^0 e^{-1}}{0!} + \frac{1^1 e^{-1}}{1!} + \frac{1^2 e^{-1}}{2!} + \frac{1^3 e^{-1}}{3!} \\ &\approx 0.368 + 0.368 + 0.184 + 0.061 = 0.981. \end{aligned}$$

By the complement rule,

$$P(Y > 3) = 1 - P(Y \leq 3) = 0.019.$$

In R,

```
> 1-ppois(3,1)
[1] 0.01898816
```

Under the $\text{Poisson}(\lambda = 1)$ assumption, it is unlikely to have more than 3 automobiles entering the tunnel during a two minute period.

Now, the question also asks us to assess (informally) whether the Poisson model seems reasonable for this problem. The way we think about this is to go back to the Poisson process assumptions; see pp 68 (notes). We would have to assume

1. the number of automobiles entering the tunnel in disjoint time periods (i.e., non-overlapping intervals of time) are independent
2. the probability of an automobile entering the tunnel in an interval of time is proportional to the length of the interval.
3. two automobiles cannot enter the tunnel at exactly the same time.

Under these three assumptions, the Poisson model applies. We might expect this is normal driving conditions; e.g., no heavy traffic, no semis holding up traffic, etc.

3.138. Here, we are being asked to find the second factorial moment; i.e., $E[Y(Y-1)]$ when $Y \sim \text{Poisson}(\lambda)$. From the definition of mathematical expectation,

$$E[Y(Y-1)] = \sum_{y=0}^{\infty} y(y-1)\ \frac{\lambda^y e^{-\lambda}}{y!}$$

The $y = 0$ and the $y = 1$ terms in this sum are zero. Therefore, we can write

$$\begin{aligned} E[Y(Y-1)] = \sum_{y=2}^{\infty} y(y-1)\ \frac{\lambda^y e^{-\lambda}}{y!} &= \sum_{y=2}^{\infty} y(y-1)\ \frac{\lambda^y e^{-\lambda}}{y(y-1)(y-2)!} \\ &= \sum_{y=2}^{\infty} \frac{\lambda^y e^{-\lambda}}{(y-2)!} \end{aligned}$$

Let $x = y - 2$. Then $y = x + 2$ and we have

$$E[Y(Y-1)] = \sum_{y=2}^{\infty} \frac{\lambda^y e^{-\lambda}}{(y-2)!} = \sum_{x=0}^{\infty} \frac{\lambda^{x+2} e^{-\lambda}}{x!} = \lambda^2 \underbrace{\sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!}}_{=\,1} = \lambda^2.$$

Therefore, $E[Y(Y-1)] = \lambda^2$. This means

$$\lambda^2 = E[Y(Y-1)] = E(Y^2 - Y) = E(Y^2) - E(Y) = E(Y^2) - \lambda \implies E(Y^2) = \lambda^2 + \lambda.$$

Finally, from the variance computing formula, we have

$$V(Y) = E(Y^2) - [E(Y)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

3.139. Let Y denote the number of defects per foot. Assume $Y \sim \text{Poisson}(\lambda = 2)$. The profit per foot is a function of Y, namely,

$$X = g(Y) = 50 - 2Y - Y^2.$$

The expected profit is

$$E(X) = E[g(Y)] = E(50 - 2Y - Y^2) = 50 - 2E(Y) - E(Y^2).$$

We have $E(Y) = 2$ and $V(Y) = 2$. We can get $E(Y^2)$ from the variance computing formula. Note that

$$V(Y) = E(Y^2) - [E(Y)]^2 \implies E(Y^2) = V(Y) + [E(Y)]^2 = 2 + 4 = 6.$$

Therefore,

$$E(X) = 50 - 2(2) - 6 = 40.$$

3.142. Suppose $Y \sim$ Poisson(λ). Suppose $y = 1, 2, 3, ...,$. In part (a), the ratio

$$\frac{p_Y(y)}{p_Y(y-1)} = \frac{\frac{\lambda^y e^{-\lambda}}{y!}}{\frac{\lambda^{y-1}e^{-\lambda}}{(y-1)!}} = \frac{\frac{\lambda^y}{y(y-1)!}}{\frac{\lambda^y \lambda^{-1}}{(y-1)!}} = \frac{\lambda}{y},$$

as claimed.
(b) Note that $p_Y(y) > p_Y(y-1)$ when the ratio

$$\frac{p_Y(y)}{p_Y(y-1)} = \frac{\lambda}{y} > 1.$$

Clearly, this ratio is larger than 1 when $y < \lambda$.
(c) From part (b), note that Poisson probabilities increase for values of $y < \lambda$. When $y > \lambda$, the probabilities start to decrease.

- If λ is not an integer, then clearly the most likely value of Y; i.e., the mode of Y, is

  $$y = \lfloor \lambda \rfloor,$$

  the floor (or greatest integer) function of λ.

- If λ is an integer, then Y has a double mode at $y = \lambda$ and $y = \lambda - 1$. This is true because

  $$\frac{p_Y(\lambda)}{p_Y(\lambda-1)} = \frac{\lambda}{\lambda} = 1.$$

  and hence $p_Y(\lambda) = p_Y(\lambda - 1)$.

3.152. Note that $m_Y(t) = e^{6(e^t-1)}$ is the mgf of $Y \sim$ Poisson$(\lambda = 6)$. For this distribution, $\mu = E(Y) = 6$ and $\sigma^2 = V(Y) = 6$. Therefore, the standard deviation is $\sigma = \sqrt{6}$. Note that

$$P(|Y - \mu| \leq 2\sigma) = P(-2\sigma \leq Y - \mu \leq 2\sigma) = P(\mu - 2\sigma \leq Y \leq \mu + 2\sigma).$$

We have

$$\begin{aligned} \mu - 2\sigma &= 6 - 2\sqrt{6} \approx 1.1 \\ \mu - 2\sigma &= 6 + 2\sqrt{6} \approx 10.9. \end{aligned}$$

Therefore,

$$P(|Y - \mu| \leq 2\sigma) = P(1.1 \leq Y \leq 10.9) = P(2 \leq Y \leq 10) = \sum_{y=2}^{10} \frac{6^y e^{-6}}{y!}$$

We can get this in R quickly using this command:

```
> sum(dpois(2:10,6))
[1] 0.9400278
```

The command `dpois(2:10,6)` calculates $P(Y = 2)$, $P(Y = 3)$, $P(Y = 4)$, ..., $P(Y = 10)$ and puts them into a vector. The `sum` function adds the elements of this vector.