INTRODUCTION TO PROBABILITY

by

Dimitri P. Bertsekas and John N. Tsitsiklis

CHAPTER 2

ADDITIONAL PROBLEM SOLUTIONS

Problem 1.

Let X be the number of royal flushes that we get in n hands. We model X as a binomial random variable with parameters n and $p = 1/649740$. Let A be the event of getting at least one royal flush in n hands. Then, A^c is the event of getting no royal flush with a probability $\mathbf{P}(A^c) = \mathbf{P}(X = 0) = p_X(0) = \binom{n}{0} p^0 (1-p)^{n-0}$. Thus, $\mathbf{P}(A) = 1 - \mathbf{P}(A^c) = 1 - (1-p)^n$. Solving the inequality $1 - (1-p)^n \geq 1 - 1/e$, we get $n \geq 649744$. To understand why the threshold value of n is so close to $1/p$, note that for large n, we have

$$(1 - 1/n)^n \approx 1/e,$$

so that $1 - (1-p)^n \approx 1 - 1/e$ when $n \approx 1/p$ and p is small.

Problem 2.

A claim is first filed in year k with probability $0.05 \cdot (0.9)^{k-1}$, and the corresponding total premium is

$$1000 \cdot \big(1 + 0.9 + \cdots + (0.9)^{k-1}\big) = 1000 \cdot \frac{1 - (0.9)^k}{1 - 0.9} = 10000\big(1 - (0.9)^k\big).$$

Thus, the PMF of Y, the total premium paid up to and including the year when the first claim is filed, is

$$p_Y(y) = \begin{cases} 0.05 \cdot (0.9)^{k-1} & \text{if } y = 10000(1 - (0.9)^k),\ k = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Problem 3.

Let $Y = \max\{0, X\}$. By using the formula

$$p_Y(y) = \sum_{\{x \mid \max\{0,x\}=y\}} p_X(x),$$

we have

$$p_Y(y) = \begin{cases} 0 & \text{if } y < 0 \text{ or } b < y, \\ \frac{1-a}{1+b-a} & \text{if } y = 0, \\ \frac{1}{1+b-a} & \text{if } 0 < y \leq b. \end{cases}$$

Let $Y = \min\{0, X\}$. Similarly, we have

$$p_Y(y) = \begin{cases} 0 & \text{if } 0 < y \text{ or } y < a, \\ \frac{1+b}{1+b-a} & \text{if } y = 0, \\ \frac{1}{1+b-a} & \text{if } a \leq y < 0. \end{cases}$$

Problem 4.

(a) We must have $\sum_{x=-3}^{3} p_X(x) = 1$, so

$$K = \frac{1}{\sum_{x=-3}^{3} x^2} = \frac{1}{28}.$$

(b) Using the formula $p_Y(y) = \sum_{\{x \mid |x|=y\}} p_X(x)$, we obtain

$$p_Y(y) = \begin{cases} 2Kx^2 = \frac{x^2}{14} & \text{if } x = 1, 2, 3, \\ 0 & \text{otherwise.} \end{cases}$$

(c) If $y \geq 0$, $p_Y(y) = \sum_{\{x \mid |x|=y\}} p_X(x) = p_X(y) + p_X(-y)$. Otherwise $p_Y(y) = 0$.

Problem 5.

We have $\cos(k\pi) = 1$ for k: even and $\cos(k\pi) = -1$ for k: odd. Therefore

$$\mathbf{E}[Y] = \sum_{k=1}^{\infty} (-1)^k k \mathbf{P}(X = k) + \sum_{k=1}^{\infty} (-1)^k (-k) \mathbf{P}(X = -k) = 0.$$

where the last equality holds because, by the symmetry assumption, we have $\mathbf{P}(X = k) = \mathbf{P}(X = -k)$.

We have $\sin(k\pi) = 0$ for all integer k, so since X takes only integer values, we have that Y is equal to 0 with probability 1. Therefore, $\mathbf{E}[Y] = 0$.

Problem 6.

At any second the probability that a mosquito bit you is $0.5 \cdot 0.2 = 0.1$. So, the time T in seconds between successive bites is a geometric random variable with parameter $p = 0.1$. It follows that $\mathbf{E}[T] = 1/p = 10$.

Problem 7.

(a) Let E be the event that Fischer wins the match. We can express E as

$$E = \bigcup_{n \geq 0} E_n,$$

where E_n is the event that each of the first n games is a draw and the $(n+1)$st game is won by Fischer. Since the E_n's are disjoint, we obtain

$$\mathbf{P}(E) = \sum_{n \geq 0} \mathbf{P}(E_n) = \sum_{n \geq 0} (1 - p - q)^n p = \frac{p}{p+q}.$$

(b) Since the duration D of the match is a geometric random variable with parameter $p+q$, we obtain

$$p_D(d) = (1 - p - q)^{d-1}(p+q), \qquad d = 1, 2, \ldots,$$

$$\mathbf{E}[D] = \frac{1}{p+q},$$

and

$$\mathrm{var}(D) = \frac{1 - p - q}{(p+q)^2}.$$

Problem 8.

We know that the number of errors in n bits is a binomial random variable with parameters n and $1-p$. Its expected value is $n(1-p)$, so $\mathbf{E}[\text{number of errors}] \leq 10$ if $n(1-p) \leq 10$, or

$$p \geq 1 - \frac{10}{n}$$

Thus for $n = 10,000$, we must have $p \geq 0.999$.

Problem 9.

Let X_i be the value obtained by the ith contestant. The answers to (a) and (b) are obtained by symmetry.

(a) 1/2.

(b) 1/3.

(c) We have

$$\begin{aligned}\mathbf{P}(N = n) = \mathbf{P}(&X_1 \text{ is the smallest of the 1st } n-1, \text{ and} \\ &X_n \text{ is the smallest of the 1st } n),\end{aligned}$$

so using the extension of the symmetry argument in (a) and (b), we have

$$\mathbf{P}(N = n) = \frac{1}{n-1} \cdot \frac{1}{n}, \qquad n = 2, 3, \ldots$$

Thus,

$$\mathbf{P}(N > n) = 1 - \mathbf{P}(N \leq n) = 1 - \sum_{k=2}^{n} \frac{1}{k(k-1)}.$$

Alternatively,

$$\mathbf{P}(N > n) = \mathbf{P}(X_1 \text{ is the smallest of 1st } n) = \frac{1}{n}.$$

(d) Using the result of part (c), we have for $n > 1$,

$$\mathbf{P}(N = n) = \mathbf{P}(N > n-1) - \mathbf{P}(N > n) = \frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)}.$$

Thus,

$$\mathbf{E}[N] = \sum_{n=2}^{\infty} n\mathbf{P}(N = n) = \sum_{n=2}^{\infty} \frac{1}{n-1} = \infty.$$

Problem 10.

We prove the statement by reversing the order of summation:

$$\begin{aligned}
\sum_{i=1}^{\infty} \mathbf{P}(N \geq i) &= \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} \mathbf{P}(N = k) \\
&= \sum_{k=1}^{\infty} \sum_{i=1}^{k} \mathbf{P}(N = k) \\
&= \sum_{k=1}^{\infty} k \mathbf{P}(N = k)
\end{aligned}$$

which is the required result.

Problem 11.

Using the formula $\text{var}(X) = \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2$, we have

$$\begin{aligned}
\mathbf{E}\big[(X_1 + \cdots + X_n)^2\big] &= \text{var}(X_1 + \cdots + X_n) + \big(\mathbf{E}[X_1 + \cdots + X_n]\big)^2 \\
&= n\text{var}(X_1) + (n\mathbf{E}[X_1])^2 \\
&= n\mathbf{E}[X_1^2] - n\big(\mathbf{E}[X_1]\big)^2 + n^2\big(\mathbf{E}[X_1]\big)^2 \\
&= n\mathbf{E}[X_1^2] + n(n-1)\big(\mathbf{E}[X_1]\big)^2.
\end{aligned}$$

Thus, $c = n$ and $d = n(n-1)$.

Problem 12.

Since the outcomes of the games are independent, the joint PMF of L_1 and L_2 satisfies

$$p_{L_1,L_2}(m,n) = p_{L_1}(m) \cdot p_{L_2}(n).$$

The random variables L_1 and L_2 are identically distributed, and they have a geometric distribution shifted by 1:

$$\mathbf{P}(L_1 = m) = \mathbf{P}(L_2 = m) = (1-p)^m \cdot p.$$

Therefore

$$p_{L_1,L_2}(m,n) = p_{L_1}(m) \cdot p_{L_2}(n) = (1-p)^{n+m} \cdot p^2.$$

Problem 13.

The probability of any set of class grades where x students get an A and y students get a B is $p^x q^y (1-p-q)^{n-x-y}$. The number of possible such sets of class grades is equal to the number of partitions of the class in three groups of x, y, and $n-x-y$ students, and is given by the multinomial coefficient

$$\binom{n}{x,\, y,\, n-x-y} = \frac{n!}{x!y!(n-x-y)!}.$$

Thus,

$$p_{X,Y}(x,y) = \begin{cases} \frac{n!}{x!y!(n-x-y)!} p^x q^y (1-p-q)^{n-x-y} & \text{if } x \geq 0,\ y \geq 0,\ x+y \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Problem 14.

(a) For $i = 1, \ldots, 250$, let U_i be the random variable taking the value 1 if the ith undergraduate student get an A, and 0 otherwise. Similarly, for $i = 1, \ldots, 50$, let G_i be the random variable taking the value 1 if the ith graduate student gets an A, and 0 otherwise. Let

$$U = \sum_{i=1}^{250} U_i, \qquad G = \sum_{i=1}^{50} G_i.$$

We have $X = U + G$. The random variables U and G are binomial with PMFs

$$p_U(u) = \binom{250}{u}(1/3)^u(2/3)^{250-u}, \text{ for } u = 0, 1, \ldots, 250,$$

and

$$p_G(g) = \binom{50}{g}(1/2)^g(1/2)^{50-g}, \text{ for } g = 0, 1, \ldots, 50.$$

If follows that

$$\begin{aligned} p_X(x) &= \mathbf{P}(U + G = x) \\ &= \sum_{u=0}^{250} \mathbf{P}(U = u)\mathbf{P}(U + G = x \,|\, U = u) \\ &= \sum_{u=0}^{250} \mathbf{P}(U = u)\mathbf{P}(G = x - u). \end{aligned}$$

Therefore,

$$p_X(x) = \sum_{u=\min\{0,x-50\}}^{x} \binom{250}{u}(1/3)^u(2/3)^{250-u}\binom{50}{x-u}(1/2)^{x-u}(1/2)^{50-x+u},$$

for $x = 1, \ldots, 300$ and $p_X(x) = 0$ otherwise. If we evaluate $\sum_{x=0}^{300} x p_X(x)$ numerically, we end up with $\mathbf{E}[X] \approx 108.34$.

(b) We have

$$X = \sum_{i=1}^{250} U_i + \sum_{i=1}^{50} G_i,$$

and hence

$$\begin{aligned} \mathbf{E}[X] &= \sum_{i=1}^{250} \mathbf{E}[U_i] + \sum_{i=1}^{50} \mathbf{E}[G_i] \\ &= 250 \cdot \mathbf{P}(U_i = 1) + 50 \cdot \mathbf{P}(G_i = 1) \\ &= 250 \cdot (1/3) + 50 \cdot (1/2) \\ &\approx 108.34 \end{aligned}$$

Problem 15.

Let D and b be the numbers of tickets demanded and bought, respectively. If S is the number of tickets sold, then $S = \min\{D, b\}$. The scalper's expected profit is

$$r(b) = \mathbf{E}[150S - 75b] = 150\mathbf{E}[S] - 75b.$$

We first find $\mathbf{E}[S]$. We assume that $b \le 10$, since clearly buying more than the maximum number of demanded tickets, which is 10, cannot be optimal. We have

$$\begin{aligned}
\mathbf{E}[S] &= \mathbf{E}[S \,|\, D \le b]\mathbf{P}(D \le b) + \mathbf{E}[S \,|\, D > b]\mathbf{P}(D > b) \\
&= \sum_{i=0}^{b} i \binom{10}{i} \left(\frac{1}{2}\right)^{10} + b \sum_{i=b+1}^{10} \binom{10}{i} \left(\frac{1}{2}\right)^{10} \\
&= \left(\frac{1}{2}\right)^{10} \left(\sum_{i=0}^{b} i \binom{10}{i} + b \sum_{i=b+1}^{10} \binom{10}{i} \right).
\end{aligned}$$

Thus

$$r(b) = 150 \left(\frac{1}{2}\right)^{10} \left(\sum_{i=0}^{b} i \binom{10}{i} + b \sum_{i=b+1}^{10} \binom{10}{i} \right) - 75b.$$

A computer solution is now required to maximize the above expression over the range $0 \le b \le 10$.

Problem 16.

We first note that

$$\mathbf{P}(X = k \,|\, X + Y = n) = \frac{\mathbf{P}(X = k, X + Y = n)}{\mathbf{P}(X + Y = n)}.$$

We have

$$\begin{aligned}
\mathbf{P}(X = k, X + Y = n) &= \mathbf{P}(X = k, Y = n - k) \\
&= \mathbf{P}(X = k)\mathbf{P}(Y = n - k) \\
&= \begin{cases} p(1-p)^{k-1}p(1-p)^{n-k-1} & \text{if } k = 1, 2, \ldots, n-1, \;\; n \geq 2, \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} p^2(1-p)^{n-2} & \text{if } k = 1, 2, \ldots, n-1, \;\; n \geq 2, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

We also have

$$\begin{aligned}
\mathbf{P}(X + Y = n) &= \sum_{\{(x,y)\,|\,x+y=n\}} \mathbf{P}(X = x, Y = y) \\
&= \sum_{x=1}^{n-1} \mathbf{P}(X = x, Y = n - x) \\
&= \begin{cases} \sum_{x=1}^{n-1} p(1-p)^{x-1}p(1-p)^{n-x-1} & \text{if } n \geq 2, \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} (n-1)p^2(1-p)^{n-2} & \text{if } n \geq 2, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

The preceding equations yield

$$\begin{aligned}
\mathbf{P}(X = k \,|\, X + Y = n) &= \begin{cases} \frac{p^2(1-p)^{n-2}}{(n-1)p^2(1-p)^{n-2}} & \text{if } n \geq 2 \text{ and } k = 1, 2, \ldots, n-1, \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} \frac{1}{n-1} & \text{if } n \geq 2 \text{ and } k = 1, 2, \ldots, n-1, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

For a more intuitive line of reasoning, consider the experiment in which we toss a biased coin with probability p of getting heads until we get the second head. Let X be the number of tosses up to and including the first head, and let Y be the number of coin tosses starting with the toss after the first head and up to and including the second head. Then $X + Y$ is the number of coin tosses until we get exactly two heads, and $\mathbf{P}(X = k \,|\, X + Y = n)$ is the probability of getting a head on the kth toss given that it took exactly n tosses to get exactly two heads. This implies that the nth toss was a head and that the first through $(n-1)$st tosses contained exactly one head and the rest tails. Each of these tosses is equally likely to be the head. So the events $X = k$ given that $X + Y = n$ are equally likely as we vary k from 1 through $n - 1$. Therefore

$$\mathbf{P}(X = k \,|\, X + Y = n) = \begin{cases} \frac{1}{n-1} & \text{if } n \geq 2 \text{ and } k = 1, 2, \ldots, n-1, \\ 0 & \text{otherwise.} \end{cases}$$

Problem 17.

We first note that

$$\mathbf{P}(X = k \,|\, X+Y+Z = n) = \frac{\mathbf{P}(X = k, X+Y+Z = n)}{\mathbf{P}(X+Y+Z = n)}.$$

We have

$$\mathbf{P}(X = k, X+Y+Z = n) = \mathbf{P}(X = k, Y+Z = n-k) = \mathbf{P}(X = k)\mathbf{P}(Y+Z = n-k).$$

We consider a coin with probability of a head equal to p, and we think of X as the number of tosses up to and including the first head, Y as the number of tosses following the first head up to and including the second head, and Z is the number of tosses following the second head up to and including the third head. Then $\{Y+Z = n-k\}$ is the event whereby the second head occurs on the $(n-k)$th toss. For this event to occur, we need to get two heads and $n-k-2$ tails, and the second head must occur on toss $n-k$, but the first head could occur at any of the previous $n-k-1$ tosses. Therefore,

$$\mathbf{P}(Y+Z = n-k) = (n-k-1)p^2(1-p)^{n-k-2}$$

Combining the preceding two equations, we obtain

$$\begin{aligned}\mathbf{P}&(X = k, X+Y+Z = n)\\ &= \begin{cases} p(1-p)^{k-1}(n-k-1)p^2(1-p)^{n-k-2} & \text{if } k = 1,2,\ldots,n-2,\ n \geq 3,\\ 0 & \text{otherwise,}\end{cases}\\ &= \begin{cases} (n-k-1)p^3(1-p)^{n-3} & \text{if } k = 1,2,\ldots,n-2,\ n \geq 3,\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

Similarly, $\{X+Y+Z = n\}$ is the event whereby the third head occurs on the nth toss. For this to occur, we need to get 3 heads and $n-3$ tails. The third head must occur on the nth toss, but the first two heads could occur at any of the previous $n-1$ tosses. The number of ways two heads can occur in $n-1$ tosses is $\binom{n-1}{2}$. Therefore,

$$\mathbf{P}(X+Y+Z = n) = \binom{n-1}{2}p^3(1-p)^{n-3} \qquad n \geq 3.$$

The preceding equations yield

$$\begin{aligned}\mathbf{P}&(X = k \,|\, X+Y+Z = n)\\ &= \begin{cases} \frac{(n-k-1)p^3(1-p)^{n-3}}{\binom{n-1}{2}}p^3(1-p)^{n-3} & \text{if } n \geq 3,\ k = 1,2,\ldots,n-2,\\ 0 & \text{otherwise,}\end{cases}\\ &= \begin{cases} \frac{2(n-k-1)}{(n-1)(n-2)} & \text{if } n \geq 3,\ k = 1,2,\ldots,n-2,\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

Another way of thinking about this problem is to realize that conditioned on the event $\{X+Y+Z = n\}$, we have a discrete uniform law. Each outcome in our new universe, $\{X+Y+Z = n\}$, requires 3 heads and $n-3$ tails, so its probability is

$p^3(1-p)^{n-3}$. Since each outcome is equally likely, to compute $\mathbf{P}(X = k \mid X+Y+Z = n)$, we simply count the number of outcomes in the event $\{X = k\} \cap \{X + Y + Z = n\}$, and divide by the number of outcomes in the event $\{X + Y + Z = n\}$.

If the first head occurs on the kth toss and the third head occurs on the nth toss, the second head could occur at $n - k - 1$ different tosses. So the number of outcomes in the event $\{X = k\} \cap \{X + Y + Z = n\}$ is $n - k - 1$. The number of outcomes in the event $\{X + Y + Z = n\}$ is the number of ways to have the third head occur on the nth toss, so it is $\binom{n-1}{2}$. Therefore,

$$\mathbf{P}(X = k \mid X + Y + Z = n) = \frac{n - k - 1}{\binom{n-1}{2}} = \frac{2(n - k - 1)}{(n - 1)(n - 2)}.$$

Problem 18.

We are given that

$$p_K(k) = \begin{cases} 1/4 & \text{if } k = 1, 2, 3, 4, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$p_{N\,|\,K}(n\,|\,k) = \begin{cases} 1/k & \text{if } n = 1, \ldots, k, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Applying the equation

$$p_{N,K}(n, k) = p_{N\,|\,K}(n\,|\,k)p_K(k),$$

we obtain

$$p_{N,K}(n, k) = \begin{cases} 1/4k & \text{if } k = 1, 2, 3, 4 \text{ and } n = 1, \ldots, k, \\ 0 & \text{otherwise.} \end{cases}$$

(b) The marginal PMF $p_N(n)$ is given by

$$p_N(n) = \sum_k p_{N,K}(n, k) = \sum_{k=n}^{4} 1/4k,$$

or

$$p_N(n) = \begin{cases} 1/4 + 1/8 + 1/12 + 1/16 = 25/48 & \text{if } n = 1, \\ 1/8 + 1/12 + 1/16 = 13/48 & \text{if } n = 2, \\ 1/12 + 1/16 = 7/48 & \text{if } n = 3, \\ 1/16 = 3/48 & \text{if } n = 4, \\ 0 & \text{otherwise.} \end{cases}$$

(c) We have

$$p_{K\,|\,N}(k\,|\,2) = \frac{p_{N,K}(2, k)}{p_N(2)} = \begin{cases} 6/13 & \text{if } k = 2, \\ 4/13 & \text{if } k = 3, \\ 3/13 & \text{if } k = 4, \\ 0 & \text{otherwise.} \end{cases}$$

(d) Let A be the event $2 \leq$N≤ 3. We first find the conditional PMF of K given A. We have

$$p_{K\,|\,A}(k) = \frac{\mathbf{P}(K = k, A)}{\mathbf{P}(A)},$$

$$\mathbf{P}(A) = p_N(2) + p_N(3) = \frac{5}{12},$$

$$\mathbf{P}(K = k, A) = \begin{cases} \frac{1}{8} & \text{if } k = 2, \\ \frac{1}{12} + \frac{1}{12} & \text{if } k = 3, \\ \frac{1}{16} + \frac{1}{16} & \text{if } k = 4, \\ 0 & \text{otherwise,} \end{cases}$$

and finally

$$p_{K\,|\,A}(k) = \begin{cases} \frac{3}{10} & \text{if } k = 2, \\ \frac{2}{5} & \text{if } k = 3, \\ \frac{3}{10} & \text{if } k = 4, \\ 0 & \text{otherwise.} \end{cases}$$

The conditional PMF of K given A is symmetric around $k = 3$, so

$$\mathbf{E}[K \mid A] = 3.$$

The conditional variance of K given A is given by

$$\mathrm{var}(K \mid A) = \mathbf{E}\big[(K - \mathbf{E}[K \mid A])^2 \mid A\big] = \frac{3}{10} \cdot (2-3)^2 + \frac{2}{5} \cdot 0 + \frac{3}{10} \cdot (4-3)^2 = \frac{3}{5}.$$

(e) We are given that $\mathbf{E}[C_i] = 30$, where C_i is the cost of book i. Let T be the total cost, so that $T = C_1 + \ldots + C_N$. We find $\mathbf{E}[T]$ by using the total expectation theorem:

$$\begin{aligned}
\mathbf{E}[T] &= \mathbf{E}[T \mid N=1]p_N(1) + \mathbf{E}[T \mid N=2]p_N(2) + \mathbf{E}[T \mid N=3]p_N(3) \\
&\quad + \mathbf{E}[T \mid N=4]p_N(4) \\
&= \mathbf{E}[C_1]p_N(1) + \mathbf{E}[C_1 + C_2]p_N(2) + \mathbf{E}[C_1 + C_2 + C_3]p_N(3) \\
&\quad + \mathbf{E}[C_1 + C_2 + C_3 + C_4]p_N(4) \\
&= \mathbf{E}[C_i]p_N(1) + 2\mathbf{E}[C_i]p_N(2) + 3\mathbf{E}[C_i]p_N(3) + 4\mathbf{E}[C_i]p_N(4) \\
&= 30 \cdot \frac{25}{48} + 60 \cdot \frac{13}{48} + 90 \cdot \frac{7}{48} + 120 \cdot \frac{1}{16} \\
&= 52.5.
\end{aligned}$$

Problem 19.

(a) Let A be the event that he got a total of three pens. The event A corresponds to getting 1 pen in one of the trips and 2 in the other, or 3 pens in the first trip. So

$$\mathbf{P}(A) = \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} = \frac{5}{9}.$$

(b) Let B be the event that he visited the supply room twice on the given day. Then,

$$\mathbf{P}(B \,|\, A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} = \frac{(1/3)(1/3) + (1/3)(1/3)}{(5/9)} = \frac{2}{5}.$$

(c) We have

$$\mathbf{E}[N] = \sum_{n=2}^{5} n\mathbf{P}_N(n) = 2 \cdot \frac{1}{9} + 3 \cdot \frac{5}{9} + 4 \cdot \frac{2}{9} + 5 \cdot \frac{1}{9} = \frac{10}{3},$$

and

$$\begin{aligned}
\mathbf{E}[N \,|\, C] &= \sum_{n=4}^{5} n\mathbf{P}_{N \,|\, C}(n) \\
&= 4 \cdot \mathbf{P}(N = 4 \,|\, N > 3) + 5 \cdot \mathbf{P}(N = 5 \,|\, N > 3) \\
&= 4 \cdot \frac{2/9}{(2/9) + (1/9)} + 5 \cdot \frac{1/9}{(2/9) + (1/9)} \\
&= \frac{13}{3}.
\end{aligned}$$

(d) We have

$$\begin{aligned}
\mathbf{E}[N^2 \,|\, C] &= \sum_{n=4}^{5} n^2\mathbf{P}_{N \,|\, C}(n) \\
&= 4^2 \cdot \mathbf{P}(N = 4 \,|\, N > 3) + 5^2 \cdot \mathbf{P}(N = 5 \,|\, N > 3) \\
&= 4^2 \cdot \frac{2/9}{(2/9) + (1/9)} + 5^2 \cdot \frac{1/9}{(2/9) + (1/9)} \\
&= 19.
\end{aligned}$$

Uisng also the result from part (c), $\mathbf{E}[N \,|\, C] = 13/3$, we obtain

$$\sigma^2_{N \,|\, C} = 19 - \left(\frac{13}{3}\right)^2 = 0.2222.$$

(e) Let C_i be the event that he gets more that three pens on the ith day. Noting that the C_i's are independent, we obtain

$$\mathbf{P}\left(\bigcap_{i=1}^{16} C_i\right) = \prod_{i=1}^{16} \mathbf{P}(C_i) = \left(\frac{2}{9} + \frac{1}{9}\right)^{16} = 3^{-16}.$$

(f) Let N_i be the total number of pens he gets in the ith day, let $X = \sum_{i=1}^{16} N_i$, and let $D = \bigcap_{i=1}^{16} C_i$. Noting that conditional on D, the N_i's are still independent and that $p_{N_i \mid D} = p_{N_i \mid C_i}$, we obtain

$$\sigma^2_{X \mid D} = \sum_{i=1}^{16} \sigma^2_{N_i \mid D} = \sum_{i=1}^{16} \sigma^2_{N_i \mid C_i} = 16 \cdot 0.2222 = 3.5552.$$

Problem 20.

Let A be the event that your detection programs lead you to the correct conclusion about your computer. Let V be the event that your computer has a virus, and let V^c be the event that your computer does not have a virus. We have

$$\mathbf{P}(A) = \mathbf{P}(V)\mathbf{P}(A \,|\, V) + \mathbf{P}(V^c)\mathbf{P}(A \,|\, V^c),$$

and $\mathbf{P}(A \,|\, V)$ and $\mathbf{P}(A \,|\, V^c)$ can be found using the binomial PMF. Thus we have

$$\begin{aligned}\mathbf{P}(A \,|\, V) &= \binom{12}{9} \cdot (0.8)^9 \cdot (0.2)^3 + \binom{12}{10} \cdot (0.8)^{10} \cdot (0.2)^2 \\ &\quad + \binom{12}{11} \cdot (0.8)^{11} \cdot (0.2)^1 + \binom{12}{12} \cdot (0.8)^{12} \cdot (0.2)^0 \\ &= 0.7899.\end{aligned}$$

using a similar calculation, we find that $\mathbf{P}(A \,|\, V^c) = 0.9742$, so that

$$\mathbf{P}(A) = 0.65 \cdot 0.7899 + 0.35 \cdot 0.9742 = 0.8544.$$

Problem 21.

(a) Let L_i be the event that Joe played the lottery on week i, and let W_i be the event that he won on week i. The desired probability is

$$\mathbf{P}(L_i \,|\, W_i^c) = \frac{\mathbf{P}(W_i^c \,|\, L_i)\mathbf{P}(L_i)}{\mathbf{P}(W_i^c \,|\, L_i)\mathbf{P}(L_i) + \mathbf{P}(W_i^c \,|\, L_i^c)\mathbf{P}(L_i^c)} = \frac{(1-q)p}{(1-q)p + 1\cdot(1-p)} = \frac{p-pq}{1-pq}.$$

(b) Conditioned on X, the random variable Y is binomial

$$p_{Y\,|\,X}(y\,|\,x) = \begin{cases} \binom{x}{y} q^y (1-q)^{(x-y)} & \text{if } 0 \le y \le x, \\ 0 & \text{otherwise.} \end{cases}$$

(c) Since X has a binomial PMF, we have

$$\begin{aligned} p_{X,Y}(x,y) &= p_{Y\,|\,X}(y\,|\,x) p_X(x) \\ &= \begin{cases} \binom{x}{y} q^y (1-q)^{(x-y)} \binom{n}{x} p^x (1-p)^{(n-x)} & \text{if } 0 \le y \le x \le n, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

(d) Using the result from part (c), we could compute the marginal p_Y using the formula

$$p_Y(y) = \sum_{x=y}^{n} p_{X,Y}(x,y),$$

but the algebra is messy. An easier method is based on the fact that Y is just the sum of n independent Bernoulli random variables, each having a probability pq of being 1. Therefore Y has a binomial PMF:

$$p_Y(y) = \begin{cases} \binom{n}{y} (pq)^y (1-pq)^{(n-y)} & \text{if } 0 \le y \le n, \\ 0 & \text{otherwise.} \end{cases}$$

(e) We have

$$\begin{aligned} p_{X\,|\,Y}(x\,|\,y) &= \frac{p_{X,Y}(x,y)}{p_Y(y)} \\ &= \begin{cases} \dfrac{\binom{x}{y} q^y (1-q)^{(x-y)} \binom{n}{x} p^x (1-p)^{(n-x)}}{\binom{n}{y} (pq)^y (1-pq)^{(n-y)}} & 0 \le y \le x \le n, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

(f) From part (a), we know the probability $\mathbf{P}(L_i \,|\, W_i^c)$ that Joe played the lottery on week i given that he did not win in that week. For each of the $n-y$ weeks when Joe did not win, there are $x-y$ weeks when he played. Thus, X conditioned on $Y=y$ is binomial with parameters $n-y$ and $\mathbf{P}(L_i \,|\, W_i^c) = (p-pq)/(1-pq)$:

$$p_{X\,|\,Y}(x\,|\,y) = \begin{cases} \binom{n-y}{x-y} \left(\frac{p-pq}{1-pq}\right)^{x-y} \left(1 - \frac{p-pq}{1-pq}\right)^{(x-y)} \binom{n}{x} & 0 \le y \le x \le n, \\ 0 & \text{otherwise.} \end{cases}$$

After some algebraic manipulation, the answer to (e) can be shown equal to the above.