Chapter 7

7.1 To specify $P(X = x \mid Y = 2)$ for $x = 0, 1$, we use the results from Table 5.1. From this table, we get $P(X = 0, Y = 2) = (1-p)^2$, $P(X = 1, Y = 2) = 2p(1-p)^2$ and $P(Y = 2) = (1-p)^2 + 2p(1-p)^2$. Therefore

$$P(X = 0 \mid Y = 2) = \frac{1}{1+2p} \text{ and } P(X = 1 \mid Y = 2) = \frac{2p}{1+2p}.$$

7.2 First the marginal mass function of Y must be determined. Using the binomium of Newton, it follows that

$$\begin{aligned} P(Y = y) &= \sum_{x=0}^{y} \binom{y}{x} \left(\frac{1}{6}\right)^x \left(\frac{1}{3}\right)^{y-x} = \left(\frac{1}{3}\right)^y \sum_{x=0}^{y} \binom{y}{x} \left(\frac{1}{2}\right)^x \\ &= \left(\frac{1}{3}\right)^y \left(\frac{1}{2} + 1\right)^y = \left(\frac{1}{2}\right)^y \quad \text{for } y = 1, 2, \ldots . \end{aligned}$$

That is, the random variable X is geometrically distributed with parameter $\frac{1}{2}$. Hence the conditional mass function of X given that $Y = y$ is given by

$$\begin{aligned} P(X = x \mid Y = y) &= \frac{\binom{y}{x}\left(\frac{1}{6}\right)^x \left(\frac{1}{3}\right)^{y-x}}{\left(\frac{1}{2}\right)^y} \\ &= \binom{y}{x} \left(\frac{1}{2}\right)^x \left(\frac{2}{3}\right)^y \quad \text{for } x = 0, 1, \ldots, y. \end{aligned}$$

Note: Using the identity $\sum_{y=x}^{\infty} \binom{y}{x} a^{y-x} = (1-a)^{-x-1}$ for $|a| < 1$, it is readily verified that the marginal mass function of X is given by $P(X = 0) = \frac{1}{2}$ and $P(X = x) = \frac{3}{2}\left(\frac{1}{4}\right)^x$ for $x = 1, 2, \ldots$. The conditional mass function of Y given that $X = 0$ is

$$P(Y = y \mid X = 0) = 2\left(\frac{1}{3}\right)^y \quad \text{for } y \geq 1.$$

For $x \geq 1$ the conditional mass function of Y given that $X = x$ is

$$P(Y = y \mid X = x) = \binom{y}{x} \left(\frac{2}{3}\right)^{x+1} \left(\frac{1}{3}\right)^{y-x} \quad \text{for } y \geq x.$$

7.3 Since $P(X=1, Y=2)=\frac{1}{6}\times\frac{1}{6}$, $P(X=x, Y=2)=\frac{4}{6}\times\frac{1}{6}\times\left(\frac{5}{6}\right)^{x-3}\times\frac{1}{6}$ for $x\geq 3$ and $P(Y=2)=\frac{5}{6}\times\frac{1}{6}$, it follows that

$$P(X=x\mid Y=2)=\begin{cases}\frac{1}{5} & \text{for } x=1,\\ \frac{4}{30}\left(\frac{5}{6}\right)^{x-3} & \text{for } x\geq 3,\end{cases}$$

$$P(X=x\mid Y=20)=\begin{cases}\left(\frac{4}{6}\right)^{x-1}\frac{1}{6}\left(\frac{5}{6}\right)^{-x} & \text{for } 1\leq x\leq 19,\\ \left(\frac{4}{6}\right)^{19}\frac{1}{6}\left(\frac{5}{6}\right)^{x-21} & \text{for } x\geq 21.\end{cases}$$

7.4 In Problem 5.5 it was shown that the joint probability mass function of X and Y is

$$P(X=x, Y=y)=\frac{y-x-1}{120}\quad\text{for } 1\leq x\leq 8 \text{ and } x+2\leq y\leq 10.$$

and t the marginal distributions of X and Y are

$$P(X=x)=\frac{(10-x)(9-x)}{240}\quad\text{for } 1\leq x\leq 8.$$

$$P(Y=y)=\frac{(y-1)(y-2)}{240}\quad\text{for } 3\leq y\leq 10.$$

It now follows that, for fixed y, the conditional probability mass function of X given that $Y=y$ is

$$P(X=x\mid Y=y)=\frac{2(y-x-1)}{(y-1)(y-2)}\quad\text{for } x=1,\ldots,y-2.$$

For fixed x, the conditional probability mass function of Y given that $X=x$ is

$$P(Y=y\mid X=x)=\frac{2(y-x-1)}{(10-x)(9-x)}\quad\text{for } y=x+2,\ldots,10.$$

7.5 The joint mass function of X and Y is

$$P(X=x, Y=y)=\left(\frac{5}{6}\right)^{x-1}\frac{1}{6}\binom{x}{y}\left(\frac{1}{2}\right)^{x}\quad\text{for } 0\leq y\leq x,\ x\geq 1.$$

The marginal mass function of Y is

$$P(Y=y)=\begin{cases}\frac{1}{5}\sum_{x=y}^{\infty}\binom{x}{y}\left(\frac{5}{12}\right)^{x}=\frac{12}{35}\left(\frac{5}{7}\right)^{y} & \text{for } y\geq 1,\\ \\ \sum_{x=1}^{\infty}\left(\frac{5}{6}\right)^{x-1}\frac{1}{6}\left(\frac{1}{2}\right)^{x}=\frac{1}{7} & \text{for } y=0,\end{cases}$$

For fixed $y \geq 1$,

$$P(X = x \mid Y = y) = \frac{7}{12}\binom{x}{y}\left(\frac{5}{12}\right)^{x}\left(\frac{5}{7}\right)^{-y} \quad \text{for } x \geq y.$$

Further, $P(X = x \mid Y = 0)$ is $\frac{7}{5}\left(\frac{5}{12}\right)^{x}$ for $x \geq 0$.

7.6 The joint probability mass function of X and Y is given by

$$\begin{aligned} P(X = i, Y = i) &= \frac{1}{36} \quad \text{for } 1 \leq i \leq 6 \\ P(X = i, Y = j) &= \frac{2}{36} \quad \text{for } 1 \leq i < j \leq 6. \end{aligned}$$

The marginal mass functions of X and Y are

$$P(X = i) = \frac{1}{36} + \frac{2(6-i)}{36} \quad \text{and} \quad P(Y = j) = \frac{1}{36} + \frac{2(j-1)}{36}$$

for $1 \leq i \leq 6$ and $1 \leq j \leq 6$. This leads to

$$P(X = i \mid Y = j) = \begin{cases} \frac{1}{1+2(j-1)} & \text{for } i = j \\ \frac{2}{1+2(j-1)} & \text{for } i < j \end{cases}$$

$$P(Y = j \mid X = i) = \begin{cases} \frac{1}{1+2(6-i)} & \text{for } j = i \\ \frac{2}{1+2(6-i)} & \text{for } j > i \end{cases}$$

7.7 The joint mass function of X and Y is

$$P(X = x, Y = y) = \binom{24}{x}\left(\frac{1}{6}\right)^{x}\left(\frac{5}{6}\right)^{24-x}\binom{x}{y}\left(\frac{1}{6}\right)^{y}\left(\frac{5}{6}\right)^{x-y}.$$

for $0 \leq x \leq 24$ and $0 \leq y \leq x$. Thus, the marginal mass function of Y is

$$P(Y = y) = \left(\frac{1}{6}\right)^{y}\left(\frac{5}{6}\right)^{24-y}\sum_{x=y}^{24}\binom{24}{x}\binom{x}{y}\left(\frac{1}{6}\right)^{x}$$

for $0 \leq y \leq 24$. For fixed y, the conditional mass function of X is

$$P(X = x \mid Y = y) = \frac{\binom{24}{x}\binom{x}{y}\left(\frac{1}{6}\right)^{x}}{\sum_{k=y}^{24}\binom{24}{k}\binom{k}{y}\left(\frac{1}{6}\right)^{k}} \quad \text{for } y \leq x \leq 24.$$

7.8 The joint probability mass function of X and Y is given by

$$P(X = x,\, Y = y) = \frac{\binom{13}{x}\binom{13}{y}\binom{26}{13-x-y}}{\binom{52}{13}}$$

for $x + y \leq 13$.The marginal mass functions of X and Y are

$$P(X = x) = \frac{\binom{13}{x}\binom{39}{13-x}}{\binom{52}{13}} \text{ and } P(Y = y) = \frac{\binom{13}{y}\binom{39}{13-y}}{\binom{52}{13}}.$$

for $x = 0, 1 \ldots 13$ and $y = 0, 1, \ldots, 13$. Thus the conditional mass functions are given by

$$P(X = x \mid Y = y) = \frac{\binom{13}{x}\binom{26}{13-x-y}}{\binom{39}{13-y}},\ P(Y = y \mid X = x) = \frac{\binom{13}{y}\binom{26}{13-y-x}}{\binom{39}{13-x}}.$$

7.9 The marginal densities of X and Y are given by

$$f_X(x) = \int_0^\infty x e^{-x(y+1)}\, dy = e^{-x} \quad \text{for } x > 0$$
$$f_Y(y) = \int_0^\infty x e^{-x(y+1)}\, dx = \frac{1}{(1+y)^2} \quad \text{for } y > 0.$$

Therefore the conditional density functions of X and Y are

$$f_X(x \mid y) = (y+1)^2 x e^{-x(y+1)} \quad \text{for } x > 0,$$
$$f_Y(y \mid x) = x e^{-xy} \quad \text{for } y > 0.$$

Further, $P(Y > 1 \mid X = 1) = \int_1^\infty f_Y(y \mid 1)\, dy = \int_1^\infty e^{-y}\, dy = e^{-1}$.

7.10 The marginal densities are $f_X(x) = \int_0^1 (x - y + 1)\, dy = x + 0.5$ for $0 < x < 1$ and $f_Y(y) = \int_0^1 (x - y + 1)\, dx = 1.5 - y$ for $0 < y < 1$. Thus the conditional density functions of X and Y are

$$f_X(x \mid y) = \frac{x - y + 1}{1.5 - y} \text{ and } f_Y(y \mid x) = \frac{x - y + 1}{x + 0.5}.$$

for $0 < x < 1$ and $0 < y < 1$. This gives

$$P(X > 0.5 \mid Y = 0.25) = \int_{0.5}^1 \frac{4}{5}(x + 0.75)\, dx = 0.6$$
$$P(Y > 0.5 \mid X = 0.25) = \int_{0.5}^1 \frac{4}{3}(1.25 - y)\, dy = \frac{1}{3}.$$

7.11 The marginal probability densities of X and Y are

$$f_X(x) = \int_0^x \frac{1}{x}\,dy = 1 \quad \text{for } 0 < x < 1$$

$$f_Y(y) = \int_y^1 \frac{1}{x}\,dx = -\ln(y) \quad \text{for } 0 < y < 1.$$

Therefore, for any given y with $0 < y < 1$, the conditional density $f_X(x \mid y) = -\frac{1}{x\ln(y)}$ for $y \leq x < 1$. For any given x with $0 < x < 1$, the conditional density $f_Y(y \mid x) = \frac{1}{x}$ for $0 < y \leq x$.

7.12 The marginal density functions of X and Y are

$$f_X(x) = \int_x^\infty e^{-y}\,dy = e^{-x} \quad \text{for } x > 0$$

$$f_Y(y) = \int_0^y e^{-y}\,dx = ye^{-y} \quad \text{for } y > 0.$$

Thus the conditional density functions of X and Y are $f_X(x \mid y) = \frac{1}{y}$ for $0 < x < y$ and $f_Y(y \mid x) = e^{-(y-x)}$ for $y > x$.

7.13 We have $f_X(x) = 1$ for $0 < x < 1$ and $f_Y(y \mid x) = \frac{1}{x}$ for $1 - x < y < 1$. By $f(x,y) = f_X(x)f_Y(y \mid x)$, the joint density of X and Y is $f(x,y) = \frac{1}{x}$ for $0 < x < 1$ and $1 - x < y < 1$. Hence, using the basic formula $P((X,Y) \in C) = \iint_C f(x,y)\,dx\,dy$, we get

$$P(X + Y > 1.5) = \int_{0.5}^1 dx \int_{1.5-x}^1 \frac{1}{x}\,dy = 0.5\ln(0.5) + 0.5 = 0.1534$$

$$P(Y > 0.5) = \int_0^{0.5} dx \int_{1-x}^1 \frac{1}{x}\,dy + \int_{0.5}^1 dx \int_{0.5}^1 \frac{1}{x}\,dy = 0.5 - 0.5\ln(0.5)$$
$$= 0.8466.$$

7.14 Since $f_Y(y) = \int_0^{1-y} 3(x+y)\,dx = \frac{3}{2}(1-y)^2 + 3y(1-y)$ for $0 < y < 1$, we have for fixed y that

$$f_X(x \mid y) = \frac{x+y}{\frac{1}{2}(1-y)^2 + y(1-y)} \quad \text{for } 0 < x < 1 - y.$$

7.15 By $f(x,\,y) = f_X(x)f_Y(y \mid x)$, the joint density of X and Y is $f(x,y) = 2x \times \frac{1}{x} = 2$ for $0 < y \leq x < 1$. The marginal density of Y is $f_Y(y) = 2(1-y)$ for $0 < y < 1$. Using again $f(x,y) = f_Y(y)f_X(x \mid y)$, we get

$$f_X(x \mid y) = \frac{1}{1-y} \quad \text{for } y \leq x < 1,$$

which is the uniform density on $(y, 1)$.

7.16 The marginal density functions of X and Y are

$$f_X(x) = \int_0^x \frac{2y}{x^2}\, dy = 1 \text{ and } f_Y(y) = \int_y^1 \frac{2y}{x^2}\, dx = 2 - 2y$$

for $0 < x < 1$ and $0 < y < 1$. Therefore the conditional density functions of X and Y are

$$f_X(x \mid y) = \frac{2y/x^2}{2-2y} \text{ for } y < x < 1,\ f_Y(y \mid x) = \frac{2y/x^2}{1} \text{ for } 0 < y < x.$$

To simulate a random observation from $f(x, y)$, we use the representation $f(x, y) = f_X(x) f_Y(y \mid x)$. A random observation from $f_X(x)$ is obtained by generating a random number from $(0, 1)$. Since for fixed x the cumulative distribution function $P(Y \leq y \mid x) = \frac{y^2}{x^2}$ for $0 \leq y \leq x$ is easily inverted, a random observation from $f_Y(y \mid x) = \frac{2y}{x^2}$ can be obtained by using the inverse-transformation method. Therefore, a random observation from $f(x, y)$ can be simulated as follows: (i) generate two random numbers u_1 and u_2 from $(0, 1)$, (ii) output $x := u_1$ and $y := u_1\sqrt{u_2}$.

7.17 Put for abbreviation

$$P_{\Delta y}(x \mid y) = P\Big(X = x \mid y - \frac{1}{2}\Delta y \leq Y \leq y + \frac{1}{2}\Delta y\Big).$$

Then

$$P_{\Delta y}(x \mid y) = \frac{P\Big(y - \frac{1}{2}\Delta y \leq Y \leq y + \frac{1}{2}\Delta y \mid X = x\Big) P(X = x)}{P\Big(y - \frac{1}{2}\Delta y \leq Y \leq y + \frac{1}{2}\Delta y\Big)}.$$

Thus, for continuity points y, we have

$$P_{\Delta y}(x \mid y) \approx \frac{p_X(x) f_Y(y \mid x)\Delta y}{f_Y(y)\Delta y} = \frac{p_X(x) f_Y(y \mid x)}{f_Y(y)}.$$

Define $p_X(x \mid y)$ as $\lim_{\Delta y \to 0} P_{\Delta y}(x \mid y)$. Then, for fixed y, $p_X(x \mid y)$ as function of x is proportional to $p_X(x) f_Y(y \mid x)$. The proportionality constant is the reciprocal of $\sum_x p_X(x) f_Y(y \mid x)$. This explains the definition of the conditional mass function of X.

7.18 Assuming that the random noise N is independent of X,

$$\begin{aligned} P(Y \leq y \mid X = 1) &= P(N \leq y-1) = P\left(\frac{N-0}{\sigma} \leq \frac{y-1}{\sigma}\right) \\ &= \Phi\left(\frac{y-1}{\sigma}\right) \quad \text{for } -\infty < y < \infty. \end{aligned}$$

In the same way,

$$P(Y \leq y \mid X = -1) = \Phi\left(\frac{y+1}{\sigma}\right) \quad \text{for } -\infty < y < \infty.$$

Differentiation gives

$$f_Y(y \mid x) = \begin{cases} (1/\sigma\sqrt{2\pi})e^{-\frac{1}{2}(y-1)^2/\sigma^2} & \text{for } x = 1 \\ (1/\sigma\sqrt{2\pi})e^{-\frac{1}{2}(y+1)^2/\sigma^2} & \text{for } x = -1. \end{cases}$$

Next apply the general formula

$$p_X(x \mid y) = \frac{p_X(x) f_Y(y \mid x)}{\sum_u p_X(u) f_Y(y \mid u)}.$$

This formula was derived in Problem 7.17. Hence we find

$$P(X = 1 \mid Y = y) = \frac{pe^{-\frac{1}{2}(y-1)^2/\sigma^2}}{pe^{-\frac{1}{2}(y-1)^2/\sigma^2} + (1-p)e^{-\frac{1}{2}(y+1)^2/\sigma^2}}.$$

7.19 Let Y be the time needed to process a randomly chosen claim. For fixed $0 < x < 1$, we have $f_Y(y \mid x) = \frac{1}{x}$ for $x < y < 2x$. By the relation $f(x,y) = f_Y(y \mid x) f_X(x)$, the joint density function of X and Y satisfies $f(x,y) = 1.5(2-x)$ for $0 < x < 1$ and $x < y < 2x$, and $f(x,y) = 0$ otherwise. The sought probability is $P((X,Y) \in C)$ with $C = \{(x,y) : 0 \leq x \leq 1,\, x \leq y \leq \min(2x,1)\}$ and is evaluated as

$$\begin{aligned} &\int_0^1 dx \int_x^{\min(2x,1)} 1.5(2-x)\, dy \\ &= \int_0^{0.5} 1.5(2-x)x\, dx + \int_{0.5}^1 1.5(2-x)(1-x)\, dx = 0.5625. \end{aligned}$$

7.20 By the independence of X and Y, the joint density function $f(x,y)$ of X and Y is given by

$$f(x,y) = \lambda e^{-\lambda x} \lambda e^{-\lambda y} \quad \text{for } x, y > 0.$$

Let $V = X$ and $W = X + Y$. To obtain $f_{V,W}(v, w)$, we use the transformation rule 5.7. The functions $a(v, w)$ and $b(v, w)$ are $a(v, w) = v$ and $b(v, w) = w - v$. The Jacobian is equal to 1. Hence the joint density of V and W is

$$f_{V,W} = \lambda e^{-\lambda v} \lambda e^{-\lambda(w-v)} = \lambda^2 e^{-w} \quad \text{for } 0 < v < w < \infty.$$

The marginal density of W is given by

$$f_W(w) = \int_0^w \lambda^2 e^{-w}\, dv = w\lambda^2 e^{-w} \quad \text{for } w > 0.$$

Hence, for any fixed w, the conditional density of V given that $W = w$ is

$$f_V(v \mid w) = \frac{\lambda^2 e^{-w}}{w\lambda^2 e^{-w}} = \frac{1}{w} \quad \text{for } 0 < v < w.$$

This verifies that the conditional density of X given that $X + Y = u$ is the uniform density on $(0, u)$.

7.21 We have $P(N = k) = \int_0^1 P(N = k \mid X_1 = u)\, du$, by the law of conditional probability. Thus

$$P(N = k) = \int_0^1 u^{k-2}(1-u) du = \frac{1}{k(k-1)} \quad \text{for } k = 2, 3, \ldots.$$

The expected value of N is equal to $\sum_{k=2}^{\infty} \frac{1}{k-1} = \infty$.

7.22 The number p is a random observation from a random variable U that is uniformly distributed on $(0, 1)$. By the law of conditional probability,

$$P(X = k) = \int_0^1 P(X = k \mid U = p)\, dp = \int_0^1 \binom{n}{k} p^k (1-p)^{n-k}\, dp$$

for $k = 0, 1, \ldots, n$. Using the fact that the beta integral $\int_0^1 x^{r-1}(1 - x)^{s-1}\, dx$ is equal to $(r-1)!(s-1)!/(r+s-1)!$ for positive integers r and s, we next obtain

$$P(X = k) = \binom{n}{k} \frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1} \quad \text{for } k = 0, 1, \ldots, n.$$

7.23 Condition on the unloading time. By the law of conditional probability, the probability of no breakdown is given by

$$\int_{-\infty}^{\infty} e^{-\lambda y} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(y-\mu)^2/\sigma^2}\, dy = e^{-\mu\lambda + \frac{1}{2}\sigma^2\lambda^2}.$$

7.24 Let the random variable R denote the number of passengers who make a reservation for a given trip. Then, $P(R = r) = \frac{1}{6}$ for $r = 1, \ldots, 6$. By the law of conditional probability,

$$P(V = j) = \sum_{r=j}^{10} \binom{r}{j} 0.8^j\, 0.2^{r-j} P(R = j) \quad \text{ for } j = 0, 1, \ldots, 10$$

$$P(W = k) = \sum_{j=0}^{10} P(W = k \mid V = j) P(V = j) \quad \text{ for } k = 0, 1, 2, 3.$$

The probability $P(V = j)$ has the values 0.0001, 0.0014, 0.0121, 0.0547, 0.1397, 0.2059, 0.1978, 0.1748, 0.1286, 0.0671, and 0.0179 for $j = 0, 1, \ldots, 10$. The probability mass function of W is

$$\begin{aligned}
P(W = 0) &= 0.25[P(V = 0) + \cdots + P(V = 9)] + P(V = 10) = 0.2634,\\
P(W = 1) &= 0.45[P(V = 0) + \cdots + P(V = 8)] + 0.75P(V = 9)\\
&= 0.4621,\\
P(W = 2) &= 0.20[P(V = 0) + \cdots + P(V = 7)] + 0.30P(V = 8)\\
&= 0.1959,\\
P(W = 3) &= 0.10[P(V = 0) + \cdots + P(V = 7)] = 0.0786.
\end{aligned}$$

7.25 Let Y be the outcome of the first roll of the die. Then, by $P(X = k) = \sum_{i=1}^{6} P(X = k \mid Y = i) P(Y = i)$, we get

$$P(X = k) = \sum_{i=1}^{5} \binom{i}{k} \Big(\frac{1}{6}\Big)^k \Big(\frac{5}{6}\Big)^{i-k} \times \frac{1}{6} + \binom{6}{k-1} \Big(\frac{1}{6}\Big)^{k-1} \Big(\frac{5}{6}\Big)^{6-k+1} \times \frac{1}{6}$$

for $k = 0, 1, \ldots, 7$ with the convention $\binom{i}{k} = 0$ for $k > i$ and $\binom{6}{-1} = 0$. This probability has the numerical values 0.4984, 0.3190, 0.1293, 0.0422, 0.0096, 0.0014, 1.1×10^{-4}, and 3.6×10^{-6}.

7.26 Let $f(x)$ be the gamma density with shape parameter r and scale parameter $(1 - p)/p$. Then, $P(N = j) = \int_0^\infty P(N = j \mid X = x) f(x)\, dx$, by the law of conditional probability. Thus

$$\begin{aligned}
P(N = j) &= \int_0^\infty e^{-x} \frac{x^j}{j!} \Big(\frac{1-p}{p}\Big)^r \frac{x^{r-1}}{(r-1)!}\, e^{-x(1-p)/p}\, dx\\
&= \frac{(r+j-1)!}{j!\,(r-1)!}\, p^j (1-p)^r \int_0^\infty \Big(\frac{1}{p}\Big)^{r+j} \frac{x^{r+j-1}}{(r+j-1)!} e^{-x/p}\, dx.
\end{aligned}$$

Since the gamma density $\left(1/p\right)^{r+j}\frac{x^{r+j-1}}{(r+j-1)!}e^{-x/p}$ integrates to 1 over $(0,\infty)$, it next follows that

$$P(N=j)=\binom{r+j-1}{r-1}p^j(1-p)^r \quad \text{for } j=0,1,\ldots.$$

This can be written as $P(N=k-r)=\binom{k-1}{r-1}p^{k-r}(1-p)^r$ for $k=r,r+1,\ldots$. In other words, the random variable $N+r$ has a negative binomial distribution with parameters r and p (the random variable N gives the number of failures before the rth success occurs).

7.27 By the law of conditional probability, the probability of having k red balls among the r selected balls is

$$\sum_{n=0}^{B}\frac{\binom{n}{k}\binom{B-n}{r-k}}{\binom{B}{r}}\binom{B}{n}p^n(1-p)^{B-n}.$$

This probability can be simplified to $\binom{r}{k}p^k(1-p)^{r-k}$. This result can be directly seen by assuming that the B balls are originally non-colored and giving each of the r balls chosen the color red with probability p.

7.28 Denote by $f_1(x)$ and $f_2(x)$ the probability densities of the random variables X_1 and X_2. Let us first point out that $pf_1(x)+(1-p)f_2(x)$ is *not* the probability density of $W=pX_1+(1-p)X_2$, as many students erroneously believe. As counterexample, take $p=\frac{1}{2}$ and assume that X_1 and X_2 are independent random variables having the uniform distribution on $(0,1)$. Then $pf_1(x)+(1-p)f_2(x)$ is the uniform density on $(0,1)$, but $\frac{1}{2}X_1+\frac{1}{2}X_2$ has a triangular density rather than a uniform density.

The random variable V is distributed as X_1 with probability p and as X_2 with probability $1-p$. Then, by the law of conditional probability,

$$P(V\leq x)=pP(X_1\leq x)+(1-p)P(X_2\leq x)$$

and so $pf_1(x)+(1-p)f_2(x)$ is the probability density of V. This density is the $N(p\mu_1+(1-p)\mu_2, p^2\sigma_1^2+(1-p)^2\sigma_2^2)$ density when the $N(\mu_1,\sigma_1^2)$ distributed X_1 and the $N(\mu_2,\sigma_2^2)$ distributed X_2 are independent of each other. This result uses the fact that the sum of two independent normal random variables is normally distributed, see Rule 8.6.

7.29 By the law of conditional probability,

$$P(B^2 \geq 4AC) = \int_0^1 P\Big(AC \leq \frac{b^2}{4}\Big)\,db = \int_0^1 db\Big[\int_0^1 P\Big(C \leq \frac{b^2}{4a}\Big)\,da\Big]$$
$$= \int_0^1 db\Big[\int_0^{b^2/4} da + \int_{b^2/4}^1 \frac{b^2}{4a}\,da\Big] = \int_0^1 db\Big[\frac{b^2}{4} - \frac{b^2}{4}\ln\Big(\frac{b^2}{4}\Big)\Big]$$
$$= \frac{5}{36} + \frac{1}{6}\ln(2) = 0.2544.$$

7.30 Let the random variable X_1 be the first number picked. Also, let the random variable χ be 1 if you have to pick exactly two numbers and be 0 otherwise. By the law of conditional probability, the probability that you have to pick exactly two numbers is

$$P(\chi = 1) = \int_{0.5}^1 P(\chi = 1 \mid X_1 = x_1)\,dx_1 = \int_{0.5}^1 \frac{0.5}{x_1}\,dx_1 = -\frac{1}{2}\ln\Big(\frac{1}{2}\Big).$$

Similarly, we get that the probability that you have to pick exactly three numbers is equal to

$$\int_{0.5}^1 dx_1 \int_{0.5}^{x_1} \frac{0.5}{x_2}\,dx_2 = \frac{1}{4}\Big[\ln\Big(\frac{1}{2}\Big)\Big]^2.$$

In general, for any $0 < a < 1$, the probability that exactly n numbers must be picked in order to obtain a number less than a is $a\big[-\ln(a)\big]^{n-1}/(n-1)!$ for $n = 1, 2, \ldots$ (by writing a as $e^{-(-\ln(a))}$, this probability mass function can be seen as a shifted Poisson distribution). The expected value of the number of picks is $1 - \ln(a)$.
Note: A funny illustration of the discrete version of the problem is as follows. It is your birthday. You are asked to blow out all of the c burning candles on your birthday cake. The number of burning candles that expire when you blow while there are still d burning candles has the discrete uniform distribution on $0, 1, \ldots, d$. Then, using the law of conditional expectation, it is readily verified that the expected value of the number of attempts to blow out all c candles is $1+\sum_{k=1}^{c}\frac{1}{k}$. However, the probability mass function of the number of attempts is rather difficult to calculate.

7.31 The expected number of crossings of the zero level during the first n jumps is $\sum_{k=1}^{n-1} E(I_k)$, where $E(I_k) = P(I_k = 1)$. Denote by S_k the position of the particle just before the $(k+1)$th jump. Then S_k

is the sum of k independent standard normally distributed random variables and is thus normally distributed (see also Rule 8.6). The random variable S_k has expected value 0 and variance k. Thus S_k has the density function $\frac{1}{\sqrt{2\pi k}} e^{-\frac{1}{2}x^2/k}$. By conditioning on S_k, we get that $P(I_k = 1)$ is equal to

$$\int_0^{\infty} \frac{1}{\sqrt{2\pi k}} e^{-\frac{1}{2}x^2/k}\, dx \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy$$
$$+ \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi k}} e^{-\frac{1}{2}x^2/k}\, dx \int_{-x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy.$$

Using polar coordinates in order to evaluate these integrals, it next follows that $P(I_k = 1)$ is equal to $\frac{1}{\pi}\left[\frac{\pi}{2} - \operatorname{arctg}(\sqrt{k})\right] = \frac{1}{\pi}\operatorname{arctg}\left(\frac{1}{\sqrt{k}}\right)$. Therefore

$$\sum_{k=1}^{n-1} E(I_k) = \frac{1}{\pi} \sum_{k=1}^{n-1} \operatorname{arctg}\left(\frac{1}{\sqrt{k}}\right).$$

Note: An asymptotic expansion for this sum is $\frac{2}{\pi}\sqrt{n} + c + \frac{1}{6\pi\sqrt{n}}$, where $c = -0.68683\ldots$.

7.32 Take the minute as time unit and represent the period between 5.45 and 6 p.m. by the interval $(0, 15)$. If you arrive at time point x at the bus stop, you will take bus number 1 home if no bus number 3 will arrive in the next $15 - x$ time units. The probability of this happening is $e^{-\lambda(15-x)}$ with $\lambda = \frac{1}{15}$. This follows from the fact that the exponential distribution is memoryless. By conditioning on your arrival epoch having the uniform distribution on (0,15) with density $f(x) = \frac{1}{15}$ for $0 < x < 15$ and using the law of conditional probability, it now follows that

$$\begin{aligned} P(\text{you take bus 1 home}) &= \int_0^{15} e^{-\frac{1}{15}(15-x)} f(x)\, dx \\ &= \frac{1}{15}\int_0^{15} e^{-\frac{1}{15}(15-x)}\, dx = 1 - \frac{1}{e}. \end{aligned}$$

An intuitive explanation of why this probability is larger than $\frac{1}{2}$ is as follows. If you arrive at a random point in time at the bus stop, your average waiting time for bus number 3 is 15 minutes (by the memoryless property of the exponential distribution), while your average waiting time for bus number 1 is 7.5 minutes.

7.33 It suffices to find $P(a,b)$ for $a \geq b$. By symmetry, $P(a,b) = 1 - P(b,a)$ for $a \leq b$. For fixed a and b with $a \geq b$, let the random variables S_A and S_B be the total scores of the players A and S_B. Let $f_A(s)$ be the probability density of S_A. Then, by the law of conditional probability,

$$P(a,b) = \int_0^1 P(A \text{ beats } B \mid S_A = s) f_A(s)\, ds.$$

By conditioning on the outcome of the first draw of player A, it follows that

$$P(S_A \leq s) = \int_0^s (s-u)\, du \text{ for } 0 < s \leq a,$$

$$P(S_A > s) = \begin{cases} 1 - s + \int_0^a \big(1 - (s-u)\big)\, du & \text{for } a < s \leq 1, \\ \int_{s-1}^a \big(1 - (s-u)\big)\, du & \text{for } 1 < s < 1 + a. \end{cases}$$

Differentiation gives that the density function $f_A(s)$ of S_A is s for $0 < s \leq a$, $1 + a$ for $a < s \leq 1$ and $1 + a - s$ for $1 < s < 1 + a$. The distribution of S_B follows by replacing a by b in the distribution of S_A. Next it is a matter of tedious algebra to obtain

$$P(a,b) = \frac{1}{2} - \frac{1}{6}(a-b)(a^2b + a^2 + ab^2 + b^2 + ab + 3a - 3) \quad \text{for } a \geq b.$$

Also, by a symmetry argument,

$$P(a,b) = \frac{1}{2} + \frac{1}{6}(b-a)(b^2a + b^2 + ba^2 + a^2 + ba + 3b - 3) \quad \text{for } a \leq b,$$

using the fact that $P(a,b) = 1 - P(b,a)$ for $a \leq b$. Let a_0 be the optimal threshold value of player A. Then, $P(a_0, b) \geq 0.5$ for all b with $P(a_0, b) = 0.5$ for $b = a_0$. This leads to the equation

$$2a_0^3 + 3a_0^2 + 3a_0 - 3 = 0.$$

The solution of this equation is $a_0 = 0.5634$. If player A uses this threshold value, his win probability is at least 50%, whatever threshold player B uses.

7.34 It should be clear that each player uses a strategy of the following form: choose a new number if the original number is below a given threshold, otherwise keep the original number. Let $P(a,b)$ denote the winning probability of player A when player A uses a threshold a and player B uses a threshold b. Player A wants to use the threshold $a = a^*$, where a^* attains the maximum in $\max_a \min_b P(a,b)$. By the law of conditional probability, we have for any given a, b that

$$P(a,b) = \int_0^1 dx \int_0^1 a(x,y)dy,$$

where $a(x,y)$ is player A's winning probability if the original number of player A is x and the original number of player B is y. Consider first the case of $a \geq b$. Then, for $x \leq a$ we have that $a(x,y) = 1 - y$ for $y > b$ and $a(x,y) = \frac{1}{2}$ for $y < b$, while for $x > a$ we have that $a(x,y) = 1$ for $b < y < x$, $a(x,y) = 0$ for $y > x$ and $a(x,y) = x$ for $y < b$. This leads to

$$P(a,b) = \frac{1}{2} + \frac{1}{2}(a - b - a^2 + ab + ab^2 - a^2b) \qquad \text{for } a \geq b.$$

By a symmetry argument, we have $P(a,b) = 1 - P(b,a)$ for $a \leq b$. This gives

$$P(a,b) = \frac{1}{2} + \frac{1}{2}(a - b + b^2 - ab + ab^2 - a^2b) \qquad \text{for } a \leq b.$$

It is not necessary to invoke a numerical method for getting the number a that attains the maximum in $\max_a \min_b P(a,b)$. It is not difficult to verify by analytical means that this number is given by

$$a^* = \frac{1}{2}(\sqrt{5} - 1).$$

To prove this, we write $P(a,b)$ as $P(a,b) = \frac{1}{2} + \frac{1}{2}(a-b)(1-a-ab)$ for $a \geq b$ and $P(a,b) = \frac{1}{2} + \frac{1}{2}(a-b)(1-b-ab)$ for $a \leq b$. Using these expressions and the fact that $a^* = \frac{1}{2}(\sqrt{5}-1)$ satisfies $a^* \times a^* + a^* = 1$, it is directly verified that $P(a^*, b) > \frac{1}{2}$ both for $b > a^*$ and for $b < a^*$ (of course, $P(a^*, b) = \frac{1}{2}$ for $b = a^*$). Hence, if player A chooses his threshold as $\frac{1}{2}(\sqrt{5} - 1)$ he will win with a probability of more than 50% unless player B also uses the threshold $\frac{1}{2}(\sqrt{5} - 1)$ in which case player A wins with a probability of exactly 50%.

7.35 Denote by $S_3(a)$ [$C_3(a)$] the probability of player A being overall winner if player A gets the score a at the first draw and stops [continues] after the first draw. By conditioning on the outcome of the second draw of player A,

$$C_3(a) = \int_0^{1-a} S_3(a+v)\,dv \qquad \text{for } 0 < a < 1.$$

The function $S_3(a)$ is increasing with $S_3(0) = 0$ and $S_3(1) = 1$, whereas the function $C_3(a)$ is decreasing with $C_3(0) > 0$ and $C_3(1) = 0$. Let a_3 be defined as the solution to the equation $S_3(a) = C_3(a)$, then a_3 is the optimal stopping point for the first player A in the three-player's game. It will be shown that

$$S_3(a) = a^4 \quad \text{for } a \geq a_2 \quad \text{and} \quad C_3(a) = \frac{1}{5}(1 - a^5) \quad \text{for } a \geq a_2,$$

where $a_2(= 0.53209)$ is the optimal stopping point for the first player in the two-player's game. Taking for granted that $a_3 \geq a_2$, it follows that the optimal stopping point a_3 is the solution to the equation

$$a^4 = \frac{1}{5}(1 - a^5)$$

on the interval $(a_2, 1)$. This solution is given by

$$a_3 = 0.64865.$$

The calculation of the overall winning probability of player A is less simple and requires $S_3(a)$ for all $0 < a < 1$.

To derive $S_3(a)$ for all $0 < a < 1$, we first observe that in the three player's game the optimal strategy of the second player B is to stop after the first draw if and only if the score of this draw exceeds both the final score of player A and $a_2 = 0.53209$. Thus, given that player A's final score is a with $a > a_2$, the probability of player B getting a score below a in the first draw and next losing from player A in the second draw is equal to $\int_0^a [a - x + (1 - (1 - x))]\, dx = a^2$. Hence

$$P(A \text{ will beat } B \mid A\text{'s final score is } a) = a^2 \quad \text{for } a > a_2.$$

To obtain $P(A$ will beat $B \mid A$'s final score is $a)$ for $0 < a < a_2$, we have to add the probability that B's score is between a and a_2 in the first draw and exceeds 1 after the second draw. This probability is given by $\int_a^{a_2} x dx = \frac{1}{2}a_2^2 - \frac{1}{2}a^2$. Thus, for $0 < a < a_2$,

$$P(A \text{ will beat } B \mid A\text{'s final score is } a) = a^2 + \frac{1}{2}a_2^2 - \frac{1}{2}a^2.$$

Obviously, the conditional probability that player A will beat player C given that player's A final score is a and player A has already beaten player B is equal to a^2 for all $0 < a < 1$. This gives

$$S_3(a) = \begin{cases} a^2 \times a^2 & \text{for } a > a_2 \\ (a^2 + \frac{1}{2}a_2^2 - \frac{1}{2}a^2) \times a^2 & \text{for } 0 < a < a_2. \end{cases}$$

Next we evaluate $C_3(a)$. By

$$C_3(a) = \int_0^{1-a} S_3(a+v)\,dv \text{ for } 0 < a < 1,$$

we obtain

$$C_3(a) = \begin{cases} \frac{1}{5}(1-a^5) & \text{for } a \geq a_2 \\ \frac{1}{10}(a_2^5 - a^5) + \frac{1}{6}(a_2^5 - a_2^2 a^3) + \frac{1}{5}(1 - a_2^5)) & \text{for } 0 < a < a_2. \end{cases}$$

This result completes the derivation of the critical level a_3, but also enables us to calculate $P_3(A)$, which is defined as the probability of player A winning under optimal play of each of the players. By the law of conditional probability,

$$P_3(A) = \int_0^{a_3} C_3(a)\,da + \int_{a_3}^1 S_3(a)\,da = 0.3052.$$

Let $P_3(B)$ be the probability of player B being the overall winner when all players act optimally. To calculate $P_3(B)$, it is convenient to define $F(a)$ as the probability that the final score of player A will be no more than a for $0 \leq a \leq 1$. Then, by conditioning on the result of the first draw of player A,

$$F(a) = \int_0^{a_3} (1-(1-x))dx = \frac{1}{2}a_3^2 \quad \text{for } a = 0,$$
$$F(a) = F(0) + \int_0^a dx \int_0^{a-x} dy = F(0) + \frac{1}{2}a^2 \quad \text{for } 0 < a < a_3.$$

For $a > a_3$,

$$F(a) = F(0) + \int_{a_3}^a dx + \int_0^{a_3} dx \int_0^{a-x} dy = F(0) - \frac{1}{2}a_3^2 - a_3 + (1+a_3)a.$$

The cumulative distribution function $F(a)$ has the mass $\frac{1}{2}a_3^2$ at $a = 0$, the density a for $0 < a < a_3$ and the density $1 + a_3$ for $a_3 < a < 1$. Next we calculate $P_3(B)$ by conditioning on the final score of player A. Using the fact that $P_2(A) = 0.453802$ is the overall winning probability of the first player in the two-player game and noting that in the two-player game the first player wins with probability v^2 when the final score of the first player is v, it follows that $P_3(B)$ is given by

$$\frac{1}{2}a_3^2P_2(A) + \int_{a_3}^{1}(1+a_3)\,dx\Big[\int_x^1 v^2\,dv + \int_0^x dv\int_{x-v}^{1-v}(v+w)^2\,dw\Big]$$
$$+\int_{a_2}^{a_3} x\,dx\Big[\int_0^x dv\int_{x-v}^{1-v}(v+w)^2dw + \int_x^1 v^2dv\Big]$$
$$+\int_0^{a_2} x\,dx\Big[\int_0^x dv\int_{x-v}^{1-v}(v+w)^2\,dw$$
$$+\int_x^{a_2} dv\int_0^{1-v}(v+w)^2\,dw + \int_{a_2}^1 v^2\,dv\Big].$$

After some algebra, this leads to

$$\begin{aligned}
P_3(B) = {} & \frac{1}{2}a_3^2P_2(A) + (1+a_3)\Big[\frac{1}{3} - \frac{1}{3}a_3 - \frac{1}{12} + \frac{1}{12}a_3^4\Big] \\
& + (1+a_3)\Big[\frac{1}{6} - \frac{1}{6}a_3^2 - \frac{1}{15} + \frac{1}{15}a_3^5\Big] \\
& + \frac{1}{6}(a_3^2 - a_2^2) - \frac{1}{15}a_3^5 + \frac{1}{15}a_2^5 \\
& + \frac{1}{9}a_3^3 - \frac{1}{18}a_3^6 + \frac{1}{6}a_2^3 - \frac{1}{9}a_2^3 - \frac{1}{24}a_2^6 + \frac{1}{72}a_2^6 \\
& + \frac{1}{3}(1-a_2^3) \times \frac{1}{2}a_2^2.
\end{aligned}$$

This gives $P_3(B) = 0.3295$. Finally, the probability of player C being the overall winner is $1 - P_3(A) - P_3(B) = 0.3653$. By simulation, we found that the final score of the winning player in the three-player game has the expected value 0.836 and the standard deviation 0.149 (in the two-player game the final score of the winning player has the expected value 0.753 and the standard deviation 0.209).

Note: For the s-player game the optimal strategy for the players is easy to characterize: the first player A stops after the first draw if and only if this draw gives a score that exceeds a_s, the second player stops after the first draw if and only if this draw gives a score that exceeds both a_{s-1} and the final score of the first player; generally, the ith player stops after the first draw only if this draw gives a score that exceeds both a_{s-i+1} and the largest value of the final scores obtained so far. For any $s \geq 2$, the critical level a_s is the solution of

$$a^{2(s-1)} = \frac{1}{2s-1}(1 - a^{2s-1})$$

on the interval $(a_{s-1}, 1)$, where $a_1 = 0$. For the general s-player game, the calculation of the overall win probability of each of the players is rather cumbersome. We have used computer simulation to obtain the overall win probabilities for the cases of $s = 4$ and $s = 5$. The overall win probabilities of the players are 0.231, 0.242, 0.255, and 0.271 when $s = 4$ and are 0.186, 0.192, 0.199, 0.207, and 0.215 when $s = 5$. The optimal stopping point a_s has the values 0.71145 and 0.75225 for $s = 4$ and $s = 5$.

7.36 The conditional expected value of the number of consolation prizes given that no main prize has been won is given by

$$E(Y \mid X = 0) = \sum_{y=0}^{3} y P(Y = y \mid X = 0) = \sum_{y=0}^{3} y \frac{\binom{15}{y}\binom{30}{3-y}}{\binom{45}{3}} = 1.$$

7.37 Since

$$f_Y(y \mid s) = \frac{3(s+1)^3}{(s+y)^4} \qquad \text{for } y > 1,$$

we have

$$\begin{aligned} E(Y \mid X = s) &= \int_1^\infty y \frac{3(s+1)^3}{(s+y)^4}\, dy \\ &= 1 + (s+1)^3 \int_1^\infty (s+y)^{-3}\, dy = 1 + \frac{1}{2}(s+1). \end{aligned}$$

7.38 The joint probability mass function of X and Y is given by

$$P(X = x, Y = y) = \frac{y - x - 1}{\binom{100}{3}} \text{ for } 1 \leq x \leq 98,\ x + 2 \leq y \leq 100.$$

The marginal distributions of X and Y are given by

$$P(X = x) = \frac{(100 - x)(99 - x)}{2\binom{100}{3}},\ P(Y = y) = \frac{(y-1)(y-2)}{2\binom{100}{3}}$$

for $x = 1, 2, \ldots, 98$ and $y = 3, \ldots, 100$. Next it follows that

$$\begin{aligned} P(X = x \mid Y = y) &= \frac{2(y - x - 1)}{(y-1)(y-2)} \\ P(Y = y \mid X = x) &= \frac{2(y - x - 1)}{(100 - x)(99 - x)}. \end{aligned}$$

Hence

$$E(X \mid Y = y) = \frac{2}{(y-1)(y-2)} \sum_{x=1}^{y-2} x(y-x-1) = \frac{1}{3}y$$

$$E(Y \mid X = x) = \frac{2}{(100-x)(99-x)} \sum_{y=x+2}^{100} y(y-x-1) = \frac{1}{3}(x+202).$$

7.39 The joint density of X and Y is $f(x,y) = 6(y-x)$ for $0 < x < y < 1$, as follows from $P(x < X \leq x+\Delta x,\, y < Y \leq y+\Delta y) = 6\Delta x(y-x)\Delta y$ for $\Delta x, \Delta y$ small, see also Example 5.3 This gives $f_X(x) = 3(1-x)^2$ for $0 < x < 1$ and $f_Y(y) = 3y^2$ for $0 < y < 1$. Thus

$$f_X(x \mid y) = \frac{6(y-x)}{3y^2} \quad \text{for } 0 < x < y$$

$$f_Y(y \mid x) = \frac{6(y-x)}{3(1-x)^2} \quad \text{for } x < y < 1.$$

This gives

$$E(X \mid Y = y) = \int_0^y x\, \frac{6(y-x)}{3y^2}\, dx = \frac{1}{3}y$$

$$E(Y \mid X = x) = \int_x^1 y\, \frac{6(y-x)}{3(1-x)^2}\, dy = \frac{2+x}{3}.$$

7.40 For ease, consider the case that X and X are continuously distributed. If X and Y are independent, then their joint density function $f(x,y)$ satisfies $f(x,y) = f_X(x)f_Y(y)$. Then, by $f_X(x \mid y) = f(x,y)/f_Y(y)$, it follows that $f_X(x \mid y) = f_X(x)$ and so

$$E(X \mid Y = y) = \int_x x f_X(x \mid y)\, dx = \int_x x f_X(x)\, dx = E(X).$$

7.41 Noting that X can be written as $X = \frac{1}{2}(X+Y) + \frac{1}{2}(X-Y)$, it follows that

$$E(X \mid X+Y = v) = \frac{1}{2}v + \frac{1}{2}E(X-Y \mid X+Y = v).$$

By Problem 6.9, $X+Y$ and $X-Y$ are independent and so $E(X-Y \mid X+Y = v) = E(X-Y)$. Also, $E(X-Y) = \mu_1 - \mu_2$. Thus

$$E(X \mid X+Y = v) = \frac{1}{2}v + \frac{1}{2}(\mu_1 - \mu_2).$$

Note: The conditional distribution of X given that $X+Y=v$ is the normal distribution with mean $\frac{1}{2}(\mu_1-\mu_2+v)$ and variance $\frac{1}{2}\sigma^2(1-\rho)$. This result follows from the relation

$$P(X \leq x \mid X+Y=v) = P\Big(\frac{1}{2}(X-Y)+\frac{1}{2}v \leq x\Big)$$

and the fact that $X-Y$ is $N(\mu_1-\mu_2,\, 2\sigma^2(1-\rho))$ distributed.

7.42 The marginal density of X is $f_X(x) = \int_x^1 dy = 1-x$ for $0<x<1$ and $f_X(x) = \int_{-x}^1 dy = 1+x$ for $-1<x<0$. The marginal density of Y is $f_Y(y) = \int_{-y}^y dx = 2y$ for $0<y<1$. Therefore, for any $0<y<1$,

$$f_X(x \mid y) = \frac{1}{2y} \quad \text{for } -y<x<y$$

and $f_X(x \mid y) = 0$ otherwise. Thus

$$E(X \mid Y=y) = \int_{-y}^{y} x\,\frac{1}{2y}\,dx = 0.$$

For any $0<x<1$, we have

$$f_Y(y \mid x) = \frac{1}{1-x} \quad \text{for } x<y<1$$

and $f_Y(y \mid x) = 0$ otherwise. For any $-1<x<0$, we have

$$f_Y(y \mid x) = \frac{1}{1+x} \quad \text{for } -x<y<1$$

and $f_Y(y \mid x) = 0$ otherwise. Thus

$$E(X \mid Y=y) = \int_{-y}^{y} x\,\frac{1}{2y}\,dx = 0.$$

For $0<x<1$, we have

$$E(Y \mid X=x) = \int_x^1 y\,\frac{1}{1-x}\,dy = \frac{1}{2}(1+x).$$

For $-1<x<0$, we have

$$E(Y \mid X=x) = \int_{-x}^1 y\,\frac{1}{1+x}\,dy = \frac{1}{2}(1-x).$$

7.43 Let X be the number of trials until the first success in a sequence of Bernoulli trials and N be the number of successes in the first n trials. Then, for $1 \leq r \leq n$ and $1 \leq j \leq n-r+1$,

$$P(X = j,\, N = r) = (1-p)^{j-1} p \binom{n-j}{r-1} p^{r-1} (1-p)^{n-j-(r-1)}$$

for $1 \leq r \leq n$, $1 \leq j \leq n-r+1$. Since $P(N = r) = \binom{n}{r} p^r (1-p)^{n-r}$, we get

$$P(X = j \mid N = r) = \frac{\binom{n-j}{r-1}}{\binom{n}{r}}.$$

Thus, by $E(X \mid N = r) = \sum_{j=1}^{n-r+1} j\, P(X = j \mid N = r)$, we find

$$E(X \mid N = r) = \frac{n+1}{r+1} \quad \text{for } 1 \leq r \leq n.$$

7.44 Suppose that r dice are rolled. Define the random variable X as the total number of point gained. Let the random variable $I = 1$ if none of the r dice shows a 1 and $I = 0$ otherwise. Then $E(X) = E(X \mid I = 0)P(I = 0) + E(X \mid I = 1)P(I = 1) = E(X \mid I = 1)(\frac{5}{6})^r$. Under the condition that $I = 1$ the random variable X is distributed as the sum of r independent random variables X_k each having the discrete uniform distribution on $2, \ldots, 6$. Each of the X_k has expected value 4. Thus

$$E(X) = 4r\Big(\frac{5}{6}\Big)^r.$$

The function $4r(\frac{5}{6})^r$ is maximal for both $r = 5$ and 6. The maximal value is 8.0376. To find $\sigma(X)$, use $E(X^2) = E(X^2 \mid I = 1)P(I = 1)$ together with

$$E(X^2 \mid I = 1) = E[(X_1 + \cdots + X_r)^2] = rE(X_1^2) + r(r-1)E^2(X_1).$$

We have $E(X_1) = 4$ and $E(X_1^2) = \sum_{k=2}^{6} k^2 \frac{1}{5} = 18$. This leads to $E(X) = 8.03756$ and $\sigma(X) = 10.008$ for $r = 5$ and $E(X) = 8.0376$ and $\sigma(X) = 11.503$ for $r = 6$.

7.45 Denote by the random variables X and Y the zinc content and the iron content. The marginal density of Y is

$$f_Y(y) = \int_2^3 \frac{1}{75}(5x + y - 30)\, dx = \frac{1}{75}(y - 17.5) \quad \text{for } 20 < y < 30,$$

and so, for any $2 < x < 3$, we have $f_X(x \mid y) = \frac{5x+y-30}{y-17.5}$. Thus

$$E(X \mid Y = y) = \int_2^3 x f_X(x \mid y)\, dx = \frac{15y - 260}{6(y - 17.5)} \quad \text{for } 20 < y < 30.$$

7.46 The insurance payout is a mixed random variable: it takes on one of the discrete values 0 and 2×10^6 or a value in the continuous interval $(0, 2\times 10^6)$. To calculate its expected value we condition on the outcome of the random variable I, where $I = 0$ if no claim is made and $I = 1$ otherwise. The insurance payout is 0 if I takes on the value 0, and otherwise the insurance payout is distributed as $\min(2\times 10^6, D)$, where the random variable D has an exponential distribution with parameter $\lambda = 1/10^6$. Thus, by conditioning,

$$E(\text{insurance payout}) = 0.9 \times 0 + 0.1 \times E[\min(2 \times 10^6, D)].$$

Using the substitution rule, it follows that

$$E[\min(2 \times 10^6, D)] = \int_0^\infty \min(2 \times 10^6, x)\lambda e^{-\lambda x}\, dx$$
$$= \int_0^{2\times 10^6} x\lambda e^{-\lambda x}\, dx + \int_{2\times 10^6}^\infty (2 \times 10^6)\lambda e^{-\lambda x}\, dx.$$

This leads after some calculations to

$$E[\min(2 \times 10^6, D)] = 10^6(1 - e^{-2}) = 864{,}665 \quad \text{dollars.}$$

Hence, we can conclude that $E(\text{insurance payout}) = \$86{,}466.50$.

7.47 (**a**) We have

$$P(X \leq x \mid a < Y < b) = \frac{1}{P(a < Y < b)} \int_{-\infty}^x dv \int_a^b f(v, w)\, dw.$$

Differentiation yields that $\int_a^b f(x, w)\, dw / P(a < Y < b)$ is the conditional probability density of X given that $a < Y < b$. In the same way, we get that

$$\frac{\int_{-\infty}^x f(x, w)\, dw}{P(X > Y)}$$

is the conditional density of X given that $X > Y$.

(**b**) For (X, Y) having a standard bivariate normal distribution with correlation coefficient ρ, the formula for $E(X \mid a < Y < b)$ is obvious from (**a**). To get $E(X \mid X > Y$, note that $X = \frac{1}{2}(X+Y) + \frac{1}{2}(X-Y)$. Therefore

$$E(X \mid X > Y) = \frac{1}{2}E(X+Y \mid X-Y > 0) + \frac{1}{2}E(X-Y \mid X-Y > 0).$$

By the independence of $X+Y$ and $X-Y$ (see Problem 6.9), it follows that $E(X+Y \mid X-Y > 0) = E(X+Y) = 0$. Since $X-Y$ is $N(0,\sigma^2)$ distributed with $\sigma^2 = 2(1-\rho)$, we have

$$E(X-Y \mid X-Y > 0) = \frac{1}{P(X-Y > 0)} \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty v e^{-\frac{1}{2}v^2/\sigma^2}\, dv,$$

which yields

$$E(X-Y \mid X-Y > 0) = \sqrt{(1-\rho)\pi}.$$

7.48 For any $0 \le x \le 1$,

$$\begin{aligned} P(U_1 \le x \mid U_1 > U_2) &= \frac{P(U_1 \le x,\, U_1 > U_2)}{P(U_1 > U_2)} = \frac{\int_0^x du_1 \int_0^{u_1} du_2}{1/2} \\ &= x^2, \end{aligned}$$

$$\begin{aligned} P(U_2 \le x \mid U_1 > U_2) &= \frac{P(U_2 \le x,\, U_1 > U_2)}{P(U_1 > U_2)} = \frac{\int_0^x du_2 \int_{u_2}^{1} du_1}{1/2} \\ &= 2\Big(x - \frac{1}{2}x^2\Big). \end{aligned}$$

Thus the conditional densities of U_1 and U_2 given that $U_1 > U_2$ are $2x$ and $2(1-x)$ for $0 < x < 1$ and zero otherwise. This gives

$$\begin{aligned} E(U_1 \mid U_1 > U_2) &= \int_0^1 x\, 2x\, dx = \frac{2}{3} \\ E(U_2 \mid U_1 > U_2) &= \int_0^1 x\, 2(1-x)\, dx = \frac{1}{3}. \end{aligned}$$

7.49 By the law of conditional probability, the probability of running out of oil is given by $\frac{2}{3}P(X_1 > Q) + \frac{1}{3}P(X_2 > Q)$, where X_i is $N(\mu_i, \sigma_i^2)$ distributed. The stockout probability can be evaluated as

$$\frac{2}{3}\Big(1 - \Phi\Big(\frac{Q-\mu_1}{\sigma_1}\Big)\Big) + \frac{1}{3}\Big(1 - \Phi\Big(\frac{Q-\mu_2}{\sigma_2}\Big)\Big).$$

By the law of conditional expectation, the expected value of the shortage is

$$\frac{2}{3}E[(X_1-Q)^+]+\frac{1}{3}E[(X_2-Q)^+],$$

where $x^+ = \max(x,0)$. The expected value of the shortage can be evaluated as

$$\frac{2}{3}\sigma_1 I\Big(\frac{Q-\mu_1}{\sigma_1}\Big)+\frac{1}{3}\sigma_2 I\Big(\frac{Q-\mu_2}{\sigma_2}\Big),$$

where $I(k)$ is the so-called normal loss integral

$$I(k)=\frac{1}{\sqrt{2\pi}}\int_k^\infty (x-k)e^{-\frac{1}{2}x^2}\,dx.$$

The normal loss integral can be evaluated as

$$I(k)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}k^2}-k[1-\Phi(k)].$$

The expected value of the number of gallons left over equals the expected value of the shortage minus $\frac{2}{3}\mu_1+\frac{1}{3}\mu_2-Q$.

7.50 Denote by the random variable R the number of people who wish to make a reservation. The random variable R is Poisson distributed with expected value $\lambda = 170$. Let the random variable S be the number of people who show up for a given flight and the random variable D be the number of people who show up for a given flight but cannot be seated. By the law of conditional expectation,

$$E(S)=\sum_{r=0}^\infty E(S\mid R=r)P(R=r)=\sum_{r=0}^\infty E(S\mid R=r)e^{-\lambda}\frac{\lambda^r}{r!}$$
$$E(D)=\sum_{r=0}^\infty E(D\mid R=r)P(R=r)=\sum_{r=0}^\infty E(S\mid R=r)e^{-\lambda}\frac{\lambda^r}{r!}.$$

We have

$$E(S\mid R=r)=\sum_{k=0}^{\min(r,Q)} k\binom{\min(r,Q)}{k}(1-q)^k q^{\min(r,Q)-k}$$
$$E(D\mid R=r)=\sum_{k=N+1}^{\min(r,Q)} (k-N)\binom{\min(r,Q)}{k}(1-q)^k q^{\min(r,Q)-k}.$$

For the numerical data $Q = 165$, $N = 150$ and $q = 0.07$, we get $E(S) = 150.61$ and $E(D) = 2.71$.

7.51 Let the geometrically distributed random variable Y be the number of messages waiting in the buffer. Under the condition that $Y = y$ the random variable X is uniformly distributed on $0, 1, \ldots, y-1$. Therefore $E(X \mid Y = y) = \frac{1}{2}(y-1)$ and $E(X^2 \mid Y = y) = \frac{1}{6}(2y^2 - 3y + 1)$, see the answer to Problem 3.47. By the law of conditional expectation,

$$E(X^k) = \sum_{y=1}^{\infty} E(X^k \mid Y = y)p(1-p)^{y-1} \quad \text{for } k = 1, 2.$$

Using the relations $\sum_{k=1}^{\infty} ka^{k-1} = \frac{1}{(1-a)^2}$ and $\sum_{k=1}^{\infty} k^2 a^{k-1} = \frac{1+a}{(1-a)^3}$ for $0 < a < 1$, we find after some algebra that

$$E(X) = \frac{1}{2p}(1-p) \quad \text{and} \quad E(X^2) = \frac{p^2 - 5p + 4}{6p^2}.$$

This gives $\sigma(X) = \frac{1}{2\sqrt{3}p}\sqrt{(1-p)(p+5)}$.

7.52 Let the random variable X be the number of newly arriving messages during the transmission time T of a message. The conditional distribution of X given that $T = n$ is the binomial distribution with parameters n and p. Thus, by the law of conditional expectation,

$$E(X) = \sum_{n=1}^{\infty} E(X \mid T = n)P(T = n) = \sum_{n=1}^{\infty} npa(1-a)^{n-1} = \frac{pa}{a^2} = \frac{p}{a}$$

$$\begin{aligned} E(X^2) &= \sum_{n=1}^{\infty} E(X^2 \mid T = n)P(T = n) \\ &= \sum_{n=1}^{\infty} \big(np(1-p) + n^2p^2\big)a(1-a)^{n-1} = \frac{p(1-p)a}{a^2} + \frac{p^2a(2-a)}{a^3} \end{aligned}$$

The standard deviation of X is $\frac{1}{a\sqrt{a}}\sqrt{pa(1-p) + 2p^2(1-a)}$.

7.53 For fixed n, let $u_k(i) = E\left[X_k(i)\right]$. The goal is to find $u_n(0)$. Apply the recursion

$$u_k(i) = \frac{1}{2}u_{k-1}(i+1) + \frac{1}{2}u_{k-1}(i)$$

for i satisfying $\frac{i}{n-k} \leq \frac{1}{2}$. The boundary conditions are

$$u_0(i) = \frac{i}{n} \quad \text{and} \quad u_k(i) = \frac{i}{n-k} \quad \text{for } i > \frac{1}{2}(n-k) \text{for } 1 \leq k \leq n.$$

The sought probability $u_n(0)$ has the values 0.7083, 0.7437, 0.7675, and 0.7761 for $n = 5$, 10, 25, and 50.
Note: $u_n(0)$ tends to $\frac{\pi}{4}$ as n increases without bound, see also Example 8.4.

7.54 If you arrive at time point x at the bus stop, then your waiting time until the next bus arrival is distributed as $W(x) = \min(15 - x, T)$, where the random variable T is the time from your arrival epoch until the next bus number 3 arrives. By the memoryless property of the exponential distribution, the random variable T has the exponential density $\lambda e^{-\lambda t}$ with $\lambda = \frac{1}{15}$. By conditioning on the random variable T, the expected value of $W(x)$ is calculated as

$$E\big(W(x)\big) = \int_0^{15-x} t\lambda e^{-\lambda t}\, dt + \int_{15-x}^{\infty} (15-x)\lambda e^{-\lambda t}\, dt,$$

which leads after some algebra to $E[W(x)] = \frac{1}{\lambda}(1 - e^{-\lambda(15-x)})$. Your arrival time X at the bus stop is uniformly distributed over (0,15) and thus has density $f(x) = \frac{1}{15} = \lambda$ for $0 < x < 15$. By conditioning on your arrival time X and applying again the law of conditional expectation, we find that the expected value of your waiting time until the next bus arrival is given by

$$\int_0^{15} E\big(W(x)\big) f(x)\, dx = \int_0^{15} \Big(1 - e^{-\frac{1}{15}(15-x)}\Big)\, dx = \frac{15}{e}.$$

7.55 Let X_a be your end score when you continue for a second spin after having obtained a score of a in the first spin. Then, by the law of conditional expectation,

$$E(X_a) = \int_0^{1-a} (a+x)\, dx + \int_{1-a}^{1} 0\, dx = a(1-a) + \frac{1}{2}(1-a)^2.$$

The solution of $a(1-a) + \frac{1}{2}(1-a)^2 = a$ is $a^* = \sqrt{2} - 1$. The optimal strategy is to stop after the first spin if this spin gives a score larger than $\sqrt{2} - 1$. Your expected payoff is \$609.48.

7.56 Given that the carnival master tells you that the ball picked from the red beaker has value r, let $L(r)$ be your expected payoff when you guess a larger value and let $S(r)$ your expected payoff when you guess a smaller value. Then

$$L(r) = \frac{1}{10}\sum_{k=r+1}^{10} k + \frac{r/2}{10} = \frac{1}{20}(10-r)(r+11) + \frac{r/2}{10} = \frac{110-r^2}{20}$$

$$S(r) = \frac{1}{10}\sum_{k=1}^{r-1} k + \frac{r/2}{10} = \frac{1}{20}(r-1)r + \frac{r/2}{10} = \frac{r^2}{20}.$$

We have $L(r) > S(r)$ for $1 \le r \le 7$ and $L(r) < S(r)$ for $8 \le r \le 10$, as can be seen by noting that $110 - x^2 = x^2$ has $x^* = \sqrt{55} \approx 7.4$ as solution. Thus, given that the carnival master tells you that the ball picked from the red beaker has value r, your expected payoff is maximal by guessing a larger value if $r \le 7$ and guessing a smaller value otherwise. Applying the law of conditional expectation, it now follows that your expected payoff is

$$\sum_{k=1}^{7} \frac{110-r^2}{20} \times \frac{1}{10} + \sum_{k=8}^{10} \frac{r^2}{20} \times \frac{1}{10} = 4.375 \text{ dollars}$$

if you use the decision rule with critical level 7. The game is not fair, but the odds are only slightly in favor of the carnival master if you play optimally. Then the house edge is 2.8% (for critical levels 5 and 6 the house edge has the values 8.3% and 4.1%).

7.57 In each of the two problems, define v_i as the expected reward that can be achieved when your current total is i points. A recursion scheme for the v_i is obtained by applying the law of conditional expectation.
(**a**) For Problem 3.24, use the recursion

$$v_i = \frac{1}{6}\sum_{k=2}^{6} v_{i+k} \quad \text{for } 0 \le i \le 19,$$

where $v_j = j$ for $j \ge 20$. The maximal expected reward is $v_0 = 8.5290$.
(**b**) For Problem 3.25, use the recursion

$$v_i = \frac{1}{6}\sum_{k=1}^{6} v_{i+k} \quad \text{for } 0 \le i \le 5,$$

where $v_j = j$ for $6 \le j \le 10$ and $v_{11} = 0$. The maximal expected reward is $v_0 = 6.9988$.

7.58 Let p_r be the probability of rolling a dice total of r with two different numbers. Then, $p_r = \frac{r-1}{36}$ for $2 \leq r \leq 7$ and $p_r = p_{14-r}$ for $8 \leq r \leq 12$. To find the expected reward under the stopping rule, apply the recursion

$$v_i = \sum_{r=3}^{11} v_{i+r} p_r \qquad \text{for } i = 0, 1, \ldots, 34,$$

where $v_s = s$ for $s \geq 35$. The expected reward under the stopping rule is $v(0) = 14.215$.

Note: The stopping rule is the one-stage-look-ahead rule, see also Problem 3.28.

7.59 Define E_i as the expected value of the remaining duration of the game when the current capital of John is i dollars. Then, by conditioning, $E_i = 1 + pE_{i+1} + qE_{i-1}$ for $1 \leq i \leq a+b-1$, where $E_0 = E_{a+b} = 0$. The solution of this standard linear difference equation is

$$E_i = \frac{i}{q-p} - \frac{a+b}{q-p}\,\frac{1-(q/p)^i}{1-(q/p)^{a+b}} \text{ if } p \neq q \text{ and } E_i = a+b-i \text{ if } p = q.$$

Substituting $p = \frac{18}{37}$, $q = \frac{19}{37}$, $a = 2$ and $b = 8$ into the formula for the expected duration of the game of gambler's ruin, we get that the expected value of the number of bets is 15.083.

Note: The expected value of the number of dollars you will stake in the game is $25 \times 15.0283 = 377.07$, and the expected value of the number of dollars you will lose is $(1 - 0.1598) \times 50 - 0.1598 \times 200) = 10.2$ (the probability that you will reach a bankroll of \$250 is 0.1592). The ratio of 10.2 and 377.07 is 0.027, in agreement with the fact that in the long-run you will lose on average 2.7 dollar cents on every dollar you bet in European roulette, regardless of what roulette system you play.

7.60 Let μ_n be the expected number of clumps of cars when there are n cars on the road. Then, by conditioning on the position of the slowest car, we get the recursion

$$\mu_n = \sum_{i=1}^{n} (1 + \mu_{n-i}) \frac{1}{n} \qquad \text{for } n = 1, 2, \ldots\,,$$

where $\mu_0 = 0$. This gives that the expected number of clumps of cars is $\sum_{k=1}^{c} \frac{1}{k}$.

7.61 For fixed n, let $F(i,k)$ be the maximal expected payoff that can be achieved when still k tosses can be done and heads turned up i times so far. The recursion is

$$F(i,k) = \max\Big[\frac{1}{2}F(i{+}1,\, k{-}1){+}\frac{1}{2}F(i{-}1,k{-}1),\ \frac{i}{n-k}\Big] \text{ for } k = 1,\ldots,n$$

with $F(i,0) = \frac{i}{n}$. The maximal expected payoff $F(0,\,n)$ has the values 0.7679, 0.7780, 0.7834, and 0.7912 for $n = 25$, 50, 100, and 1,000.

7.62 Define the value function $f_k(i)$ as the expected value of the maximal score you can still reach when k rolls are still possible and the last roll of the two dice gave a score of i points. You want to find $f_6(0)$ and the optimal strategy. Let a_j denote the probability of getting a score of j in a single roll of two dice. The a_j are given by $a_j = \frac{j-1}{36}$ for $2 \le j \le 7$ and $a_j = a_{14-j}$ for $8 \le j \le 12$. The recursion is

$$f_k(i) = \max\Big[i,\ \sum_{j=2}^{12} f_{k-1}(j)a_j\Big] \quad \text{for } i = 0,1,\ldots,12.$$

applies for $k = 1,\ldots,6$ with the boundary condition $f_0(i) = i$ for all i. The recursion leads to $f_6(0) = 9.474$. The numerical calculations reveal the optimal strategy as well: if still k rolls are possible, you stop if the last roll gave s_k or more points and otherwise you continue, where $s_1 = s_2 = 8$, $s_3 = s_4 = 9$, and $s_5 = 10$.

7.63 Let state $(l,r,1)$ ($(l,r,0)$) mean that r numbers have been taken out of the hat, l is the largest number seen so far and l was obtained (not obtained) at the last pick. For $k = 0,1$, define $F_r(l,k)$ as the maximal probability of obtaining the largest number starting from state (l,r,k) when r numbers have been taken out of the hat. The maximal success probability is $\sum_{l=1}^{N} F_1(l,1)\frac{1}{N}$. The optimality equations are

$$F_r(l,0) = F_{r+1}(l,0)\frac{l-r}{N-r} + \sum_{j=l+1}^{N} F_{r+1}(j,1)\frac{1}{N-r}$$

$$F_r(l,1) = \max\Big[\frac{\binom{l-r}{n-r}}{\binom{N-r}{n-r}},\ F_{r+1}(l,0)\frac{l-r}{N-r} + \sum_{j=l+1}^{N} F_{r+1}(j,1)\frac{1}{N-r}\Big]$$

for $l = r,\ldots,N$, where $\binom{l-r}{n-r} = 0$ for $l < n$ and the boundary conditions are $F_n(l,0) = 0$ and $F_n(l,1) = 1$ for $l = n,\ldots,N$. For $n = 10$

and $N = 100$, the maximal success probability is 0.6219 and the optimal stopping rule is characterized by $l_1 = 93$, $l_2 = 92$, $l_3 = 91$, $l_4 = 89$, $l_5 = 87$, $l_6 = 84$, $l_7 = 80$, $l_8 = 72$, and $l_9 = 55$. This rule prescribes to stop in state $(l, r, 1)$ if $l \geq l_r$ and to continue otherwise.
Note: For the case of $n = 10$ and $N = 100$, we verified experimentally that l_r is the smallest value of l such that $Q_s(l, r) \geq Q_c(l, r)$, where

$$Q_s(l, r) = \frac{\binom{l-r}{10-r}}{\binom{100-r}{10-r}} \quad \text{and} \quad Q_c(l, r) = \sum_{k=1}^{10-r} \frac{1}{k} \frac{\binom{100-l}{k}\binom{l-r}{10-r-k}}{\binom{100-r}{10-r}}$$

We have that $Q_s(l, r)$ is the probability of having obtained the overall largest number when stopping in state $(l, r, 1)$, and $Q_c(l, r)$ is the probability of getting the overall largest number when continuing in state $(l, r, 1)$ and stopping as soon as you pick a number larger than l.

7.64 For $k = 0, 1$, let state (l, r, k) and the value-function $F_r(l, k)$ be defined in the same way as in Problem 7.63. Then

$$F_r(l, 0) = F_{r+1}(l, 0)\frac{l-1}{N} + \sum_{j=l}^{N} F_{r+1}(j, 1)\frac{1}{N}$$

$$F_r(l, 1) = \max\Big[\Big(\frac{l}{N}\Big)^{n-r}, \; F_{r+1}(l, 0)\frac{l-1}{N} + \sum_{j=l}^{N} F_{r+1}(j, 1)\frac{1}{N}\Big]$$

for $l = r, \ldots, N$, where the boundary conditions are $F_n(l, 0) = 0$ and $F_n(l, 1) = 1$ for $l = n, \ldots, N$. The maximal success probability is $\sum_{l=1}^{N} F_1(l, 1)\frac{1}{N}$.

7.65 Define the value function $v(i_0, i_1)$ as the maximal expected net winnings you can still achieve starting from state (i_0, i_1), where state (i_0, i_1) means that there are i_0 empty bins and i_1 bins with exactly one ball. The desired expected value $v(b, 0)$ can be obtained from the optimality equation

$$v(i_0, i_1) = \max\Big[i_1 - \frac{1}{2}(b - i_0 - i_1), \; \frac{i_0}{b}\, v(i_0 - 1, i_1 + 1) + \frac{i_1}{b}\, v(i_0, i_1 - 1) + \frac{b - i_0 - i_1}{b}\, v(i_0, i_1)\Big]$$

with the boundary condition $v(0, i_1) = i_1 - \frac{1}{2}(b - i_1)$. This equation can be solved by backwards calculations. First calculate $v(1, i_1)$ for

$i_1 = 0, \ldots, b-1$. Next calculate $v(2, i_1)$ for $i_1 = 0, \ldots, b-2$. Continuing in this way, the desired $v(b, 0)$ is obtained. Numerical investigations lead to the conjecture that the optimal stopping rule has the following simple form: you stop only in the states (i_0, i_1) with $i_1 \leq a$, where a is the smallest integer larger than or equal to $2i_0/3$. For $b = 25$, we find that the maximal expected net winnings is \$7.566. The one-stage-look-ahead rule prescribes to stop in the states (i_0, i_1) with $i_0 \leq (1+0.5)i_1$ and to continue otherwise. This stopping rule has an expected net winnings of \$7.509.
Note: The standard deviation of the net winnings is \$2.566 for the optimal stopping rule and \$2.229 for the one-stage-look-ahead rule.

7.66 Let state (i, s) correspond to the situation that your accumulated reward is i dollars and the dice total in the last role is s. Define the value-function $V(i, s)$ is the maximal achievable reward starting from state (i, s). The goal is to find $\sum_{s=2}^{12} V(s, s)\, p_s$, where p_s is the probability of getting a dice total of s in a single roll of the two dice. The p_k are given by $p_j = \frac{j-1}{36}$ for $2 \leq j \leq 7$ and $p_j = p_{14-j}$ for $8 \leq j \leq 12$. The optimality equation is

$$V(i, s) = \max\Big[i, \sum_{k=2, k\neq s}^{12} V(i+k, k)\, p_k\Big]$$

with the boundary condition $V(j, k) = j$ for $j \geq M$. Using backward calculations the values $V(s, s)$ and the optimal stopping rule are found.
Note: A heuristic rule, which is very close in performance to the optimal stopping rule, is the one-stage-look-ahead rule. The heuristic rule prescribes to stop in the states (i, s) with $i \geq N_s$ and to continue otherwise, where the threshold values N_s are given by $N_2 = 250$, $N_3 = 123$, $N_4 = 80$, $N_5 = 58$, $N_6 = 45$, $N_7 = 42$, $N_8 = 43$, $N_9 = 54$, $N_{10} = 74$, $N_{11} = 115$, and $N_{12} = 240$. The critical level N_s is the smallest integer i such that $\sum_{k=2,k\neq s}^{12} k\, p_k - i\, p_s \leq 0$ or, equivalently, $7 - (i+s)p_s \leq 0$.

7.67 Imagine that the balls are placed into the bins at times generated by a Poisson process with rate 1. Then, a Poisson process with rate $\frac{1}{b}$ generates the times at which the ith bin receives a ball. Using the independence of the Poissonian subprocesses and conditioning upon the time that the ith bin receives its first ball, it follows that

$$P(A_i) = \int_0^\infty \Big(\sum_{k=1}^{m} e^{-\frac{1}{b}t}\, \frac{(t/b)^k}{k!}\Big)^{b-1} \frac{1}{b} e^{-\frac{1}{b}t}\, dt.$$

The sought probability is $\sum_{i=1}^{b} P(A_i)$. As a sanity check, this probability is $\frac{b!}{b^b}$ for $m = 1$.

7.68 Imagine that the bottles are bought at epochs generated by a Poisson process with rate 1. Let T be the first time at which the required numbers of the letters have been collected and let N be the number of bottles needed. Then $E(N) = E(T)$. The letters A, B, R, and S are obtained at epochs generated by Poisson processes with respective rates 0.15, 0.10, 0.40, and 0.35. Moreover, these Poisson processes are independent of each other. Using the relation $E(T) = \int_0^\infty [1 - P(T \leq t)]\,dt$, we find that the expected number of bottles needed to form the payoff word is

$$\int_0^\infty \big[1 - (1 - (1 + 0.15t + 0.15^2 t^2/2!)e^{-0.15t}) \times (1 - (1 + 0.1t)e^{-0.1t}) \\ \times (1 - (1 + 0.4t)e^{-0.4t}) \times (1 - e^{-0.35t})\big]\,dt = 26.9796.$$

7.69 Imagine that rolls of the two dice occur at epochs generated by a Poisson process with rate 1. Let N be the number of rolls needed to remove all tokens and T be the first epoch at which all tokens have been removed. Then, $E(N) = E(T)$ and $E(T) = \int_0^\infty P(T > t)\,dt$. Also,

$$T = \max_{2 \leq j \leq 12} T_j,$$

where T_j is the first epoch at which all tokens in section j have been removed. The rolls resulting in a dice total of k occur according to a Poisson process with rate p_k and these Poissonian subprocesses are independent of each other. The p_k are given by $p_k = \frac{k-1}{36}$ for $2 \leq k \leq 7$ and $p_k = p_{14-k}$ for $8 \leq k \leq 12$. By the independence of the T_k,

$$P(T \leq t) = P(T_2 \leq t) \cdots P(T_{12} \leq t).$$

Also,

$$P(T_k > t) = \sum_{j=0}^{a_k - 1} e^{-p_k t} \frac{(p_k t)^j}{j!}.$$

Putting the pieces together and using numerical integration, we find $E(N) = 31.922$.

7.70 Imagine that purchases are made at epochs generated by a Poisson process with rate 1. For any $i = 1, \ldots, n$, a Poisson subprocess with rate $\frac{1}{n}$ generates the epochs at which a coupon of type i is obtained.

The Poisson subprocesses are independent of each other. Let T be the first epoch at which two complete sets of coupons are obtained and let N be the number of purchases needed to get two complete sets of coupons. Then $E(N) = E(T)$. Using the relation $E(T) = \int_0^\infty [1 - P(T \leq t)]\, dt$, we find that the expected number of purchases needed to get two complete sets of coupons is equal to

$$\int_0^\infty \left[1 - \left(1 - e^{-t/n} - \frac{t}{n} e^{-t/n}\right)^n\right] dt.$$

This integral has the value 24.134 when $n = 6$.

7.71 Imagine that the rolls of the die occur at epochs generated by a Poisson process with rate 1. Then, the times at which an odd number is rolled are generated by a Poisson process with rate $\frac{1}{2}$ and the times at which the even number k is rolled are generated by a Poisson process with rate $\frac{1}{6}$ for $k = 2$, 4, and 6. These Poisson processes are independent of each other. By conditioning on the first epoch at which an odd number is rolled, we find that the sought probability is

$$\int_0^\infty \left(1 - e^{-t/6}\right)^3 \frac{1}{2} e^{-(1/2)t}\, dt = 0.05.$$

7.72 Taking the model with replacement and mimicking the arguments used in the solution of Example 7.13, we get that the sought probability is

$$\int_0^\infty \left(1 - e^{-\frac{10}{35}t}\right)\left(1 - e^{-\frac{20}{35}t}\right) \frac{5}{35} e^{-\frac{5}{35}t}\, dt = 0.6095.$$

7.73 Imagine that the rolls of the die occur at epochs generated by a Poisson process with rate 1. Then, independent Poisson processes each having rate $\mu = \frac{1}{6}$ describe the moves of the horses. The density of the sum of $r = 6 - s_1$ independent exponentially distributed interoccurrence times each having expected value $1/\mu$ is the Erlang density $\mu^r \frac{t^{r-1}}{(r-1)!} e^{-\mu t}$, see Section 4.5.1. Thus the win probability of horse 1 with starting position $s_1 = 0$ is

$$\int_0^\infty \Big(\sum_{k=0}^{4} e^{-\frac{t}{6}} \frac{(t/6)^k}{k!}\Big)^2 \Big(\sum_{k=0}^{3} e^{-\frac{t}{6}} \frac{(t/6)^k}{k!}\Big)^2 \Big(\sum_{k=0}^{5} e^{-\frac{t}{6}} \frac{(t/6)^k}{k!}\Big) \Big(\frac{1}{6}\Big)^6 \frac{t^5}{5!} e^{-\frac{t}{6}} dt.$$

This gives the win probability 0.06280 for the horses 1 and 6. In the same way, we get the win probability 0.13991 for the horses 2 and

5, and the win probability 0.29729 for the horses 3 and 4. To find the expected duration of the game, let T_i be the time at which horse i would reaches the finish when the game would be continued until each horse has been finished. The expected duration of the game is equal to $T = \min(T_1, \ldots, T_6)$. Noting that $E(T) = \int_0^\infty P(T > t)\, dt = \int_0^\infty P(T_1 > t) \cdots P(T_6 > t)\, dt$, it follows that the expected duration of the game is

$$\int_0^\infty \Big(\sum_{k=0}^{5} e^{-t/6} \frac{(t/6)^k}{k!}\Big)^6 dt = 19.737$$

when each horse starts at panel 0.

7.74 The analysis is along the same lines as the analysis of Problem 7.73. The probability of player A winning is

$$\int_0^\infty \Big(\sum_{k=0}^{3} e^{-3t/9} \frac{(3t/9)^k}{k!}\Big) \times \Big(\sum_{l=0}^{2} e^{-2t/9} \frac{(2t/9)^l}{l!}\Big) \Big(\frac{4}{9}\Big)^5 \frac{t^4}{4!} e^{-4t/9} dt.$$

Similarly, the other win probabilities. This leads to $P(A) = 0.3631$, $P(B) = 0.3364$, and $P(C) = 0.3005$. The expected number of games is given by

$$\int_0^\infty \Big(\sum_{j=0}^{4} e^{-4t/9} \frac{(4t/9)^j}{j!}\Big) \times \Big(\sum_{k=0}^{3} e^{-3t/9} \frac{(3t/9)^k}{k!}\Big) \times \Big(\sum_{l=0}^{2} e^{-2t/9} \frac{(2t/9)^l}{l!}\Big)\, dt.$$

This integral can be evaluated as 7.3644.

7.75 Imagine that cards are picked at epochs generated by a Poisson process with rate 1. Let N be the number of picks until each card of some of the suits has been obtained and let T be the epoch at which this occurs. Then $E(T) = E(N)$. Any specific card is picked at epochs generated by a Poisson process with rate $\frac{1}{20}$. These Poisson processes are independent of each other. Let T_i be the time until all cards of the ith suit has been picked. The T_i are independent random variables and $T = \min(T_1, T_2, T_3, T_4)$. Since $P(T_i > t) = 1 - (1 - e^{-t/20})^5$ and $P(T > t) = P(T_1 > t) \cdots P(T_4 > t)$, we get

$$E(T) = \int_0^\infty \Big[1 - \Big(1 - e^{-t/20}\Big)^5\Big]^4 dt = 24.694.$$

Therefore $E(N) = 24.694$.

7.76 Imagine that a Poisson process with rate 1 generates the epochs at which a ball is placed into a randomly chosen bin. Then, for any $i = 1, \ldots, b$, the epochs at which the ith bin receives a ball are generated by a Poisson subprocess with rate $\frac{1}{b}$. These Poisson processes are independent of each other. Let T be the first time at which each bin contains m or more balls and let N be number of balls needed until each bin contains at least m balls. Then $E(N) = E(T)$. Using the relation $E(T) = \int_0^\infty [1 - P(T \le t)]\, dt$, we get

$$\begin{aligned} E(N) &= \int_0^\infty \Big[1 - \Big(\sum_{k=m}^\infty e^{-t/b} \frac{(t/b)^k}{k!}\Big)^b\Big]\, dt \\ &= b\int_0^\infty \Big[1 - \Big(1 - \sum_{k=0}^{m-1} e^{-u} \frac{u^k}{k!}\Big)^b\Big]\, du. \end{aligned}$$

7.77 Writing $x_i - \theta = x_i - \overline{x} + \overline{x} - \theta$, it follows that

$$\sum_{i=1}^n (x_i - \theta)^2 = n(\overline{x} - \theta)^2 + \sum_{i=1}^n (x_i - \overline{x})^2 - 2(\overline{x} - \theta) \sum_{i=1}^n (x_i - \overline{x}).$$

Noting that $\sum_{i=1}^n (x_i - \overline{x}) = 0$, it follows that

$$L(\mathbf{x} \mid \theta) = (\sigma\sqrt{2\pi})^{-n}\, e^{-\frac{1}{2}\sum_{i=1}^n (x_i - \theta)^2/\sigma^2}$$

is proportional to $e^{-\frac{1}{2}n(\theta - \overline{x})^2/\sigma^2}$. Thus the posterior density $f(\theta \mid \mathbf{x})$ is proportional to $e^{-\frac{1}{2}n(\theta - \overline{x})^2/\sigma^2} f_0(\theta)$, where $f_0(\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\theta - \mu_0)^2/\sigma_0^2}$. Next it is a matter of some algebra to find that the posterior density is proportional to

$$e^{-\frac{1}{2}(\theta - \mu_p)^2/\sigma_p^2},$$

where μ_p and σ_p^2 are equal to

$$\mu_p = \frac{\sigma_0^2 \overline{x} + (\sigma^2/n)\mu_0}{\sigma_0^2 + \sigma^2/n} \quad \text{and} \quad \sigma_p^2 = \frac{\sigma_0^2(\sigma^2/n)}{\sigma_0^2 + \sigma^2/n}.$$

In other words, the posterior density is the $N(\mu_p, \sigma_p^2)$ density. Inserting the data $n = 10$, $\sigma = \sqrt{2}$, $\mu_0 = 73$ and $\sigma_0 = 0.7$, it follows that the posterior density is maximal at $\theta^* = \mu_p = 73.356$. Using the 0.025 and 0.975 percentiles of the standard normal density, a 95% Bayesian credible interval for θ is

$$(\mu_p - 1.960\sigma_p,\ \mu_p + 1.960\sigma_p) = (72.617,\ 74.095).$$

7.78 The IQ of the test person is modeled by the random variable Θ. The posterior density of Θ is proportional to

$$e^{-\frac{1}{2}(123-\theta)^2/56.25} \times e^{-\frac{1}{2}(\theta-100)^2/125}.$$

Using a little algebra to rewrite this expression, we get that the posterior density is a normal density with expected value $\mu_p = 115.862$ and standard deviation $\sigma_p = 6.228$ (the normal distribution is a conjugate prior for a likelihood function of the form of a normal density, see also Problem 7.77). The posterior density is maximal at $\theta^* = 115.862$. A 95% Bayesian credible interval for θ is

$$(\mu_p - 1.960 \times \sigma_p, \mu_p + 1.960 \times \sigma_p) = (103.56, 128.01).$$

7.79 The posterior density is proportional to

$$e^{-\frac{1}{2}(t_1-\theta)^2/\sigma^2} \times e^{-\frac{1}{2}(\theta-\mu_0)^2/\sigma_0^2},$$

where $t_1 = 140$, $\sigma = 20$, $\mu_0 = 150$, and $\sigma_0 = 25$ light years. Next a little algebra shows that the posterior density is proportional to $e^{-\frac{1}{2}(\theta-\mu_p)^2/\sigma_p^2}$, where

$$\mu_p = \frac{\sigma_0^2 t_1 + \sigma^2 \mu_0}{\sigma_0^2 + \sigma^2} \quad \text{and} \quad \sigma_p^2 = \frac{\sigma_0^2 \sigma^2}{\sigma_0^2 + \sigma^2}.$$

This gives that the posterior density is a normal density with an expected value of $\mu_p = 143.902$ light years and a standard deviation of $\sigma_p = 15.617$ light years. The posterior density is maximal at $\theta = 143.902$. A 95% Bayesian credible interval for the distance is

$$(\mu_p - 1.960\sigma_p,\ \mu_p + 1.960\sigma_p) = (113.293, 174.512).$$

7.80 The prior density $f_0(\theta)$ of the proportion of Liberal voters is

$$f_0(\theta) = c\,\theta^{474-1}(1-\theta)^{501-1},$$

where c is a normalization constant. The likelihood function $L(E \mid \theta)$ is given by

$$L(E \mid \theta) = \binom{110}{527}\theta^{527}(1-\theta)^{573}.$$

The posterior density $f(\theta \mid E)$ is proportional to $L(E \mid \theta) f_0(\theta)$ and so it is proportional to

$$\theta^{1{,}000}(1-\theta)^{1{,}073}.$$

In other words, the posterior density $f(\theta \mid E)$ is the beta(1,001, 1,073) density. The posterior probability that the Liberal party will win the election is

$$\int_{0.5}^{1} f(\theta \mid E)\, d\theta = 0.5446.$$

A Bayesian 95% confidence interval for the proportion of Liberal voters can be calculated as $(0.4609, 0.5039)$.

7.81 The prior density of the parameter of the exponential lifetime of a light bulb is $f_0(\theta) = c\,\theta^{\alpha-1}e^{-\lambda\theta}$, where c is a normalization constant. Let E be the event that light bulbs have failed at times $t_1 < \cdots < t_r$ and $m-r$ light bulbs are still functioning at time T. The likelihood function $L(E \mid \theta)$ is defined as

$$\binom{m}{r} r!\,\theta^r e^{-[t_1+\cdots+t_r+(m-r)T]\theta}$$

(the rationale of this definition is the probability that one light bulb fails in each of the infinitesimal intervals $(t_i - \frac{1}{2}\Delta, t_i + \frac{1}{2}\Delta)$ and $m-r$ light bulbs are still functioning at time T). The posterior density $f(\theta \mid E)$ is proportional to

$$\theta^{\alpha+r-1}e^{-[\lambda+t_1+\cdots+t_r+(m-r)T]\theta}.$$

In other words, the posterior density $f(\theta \mid E)$ is a gamma density with shape parameter $\alpha + r$ and scale parameter $\lambda + \sum_{i=1}^{r} t_i + (m-r)T$.