Solutions, answers, and hints for selected problems

Asterisks in "A Modern Approach to Probability Theory" by Fristedt and Gray identify the problems that are treated in this supplement. For many of those problems, complete solutions are given. For the remaining ones, we give hints, partial solutions, or numerical answers only.

For Chapter 1

1-2. Method 1: By the Binomial Theorem,

$$\sum_{k=0}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} 1^{k} = (1+1)^n = 2^n$$

and, for $n > 0$,

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} (-1)^{k} = (1-1)^n = 0\,.$$

Addition and then division by 2 gives

$$\sum_{\substack{k=0 \\ k \text{ even}}}^{n} \binom{n}{k} = 2^{n-1}\,.$$

The answer for positive n is $2^{n-1}/2^n = 1/2$. The answer for $n = 0$ is easily seen to equal 1.

Method 2: For $n \geq 1$ consider a sequence of length $(n-1)$. If it contains an even number of 'heads', adjoin a 'tails' to it to obtain a length-n sequence containing an even number of 'heads'. If it contains an odd number of 'heads', adjoin a 'heads' to it to obtain a length-n sequence containing an even number of 'heads'. Moreover, all length-n sequences containing an even number of 'heads' are obtained by one of the preceding two procedures. We have thus established, for $n \geq 1$, a one-to-one correspondence between the set of all length-$(n-1)$ sequences and the set of those length-n sequences that contain an even number of 'heads'. Therefore, there are 2^{n-1} length-n sequences that contain an even number of 'heads'. To treat the remaining case $n = 0$, we observe that the empty sequence, which is the only length-0 sequence, contains zero 'heads'. Since 0 is even, there is 1 length-0 sequence containing an even number of 'heads'.

1-4. 2^{-j}

1-10. The thirty-six points to each of which is assigned probability $\frac{1}{36}$ are the ordered pairs (r, g) for $1 \leq r \leq 6$ and $1 \leq g \leq 6$. The coordinates r and g represent the numbers showing on the red die and green die, respectively.

1-11. The set consisting of a single sample point, being the intersection of countably many events A of the form (1.2), is an event. Its probability is no larger than that of any such A. For each n and each sample point, there is such an A that has probability 2^{-n}. Thus, the probability of that sample point is no larger than 2^{-n}. Letting $n \to \infty$ we see that the probability of the sample point is 0. The process of flipping a coin until the first tails occurs terminates in a finite number of steps with probability 1.

1-12. (i) Sum the answer to Problem 4 over odd positive j to obtain $\frac{2}{3}$.

(ii) $\frac{1}{16}$

(iii) (Caution: it is common for students to use invalid reasoning in this type of problem.) We use '1' and '0' to denote heads and tails, respectively. Let S denote the set of finite sequences s of 1's and 0's terminating with 1, containing no subsequence of the form $(1, 0, 1)$ or $(1, 1, 1)$, and having the additional property that if the length of s is at least two, then the penultimate term in s is 0. For each $s \in S$, let A_s be the event consisting of those infinite sequences ω that begin with s followed by $(0, 1, 1)$, $(0, 1, 0)$, or $(1, 0, 1)$ in the next three positions, and let B_s be the event consisting of those ω that begin with s followed by $(1, 1, 1)$ or $(1, 1, 0)$ in the next three positions. Note that each A_s and B_s is a member of $\mathcal{E}$. Clearly $2P(A_s) = 3P(B_s)$.

Let $A = \bigcup_{s \in S} A_s$ and $B = \bigcup_{s \in S} B_s$. Straightforward set-theoretic arguments show that A consists of those ω in which $(1, 0, 1)$ occurs before $(1, 1, 1)$, B consists of those ω in which $(1, 1, 1)$ occurs before $(1, 0, 1)$. By writing A and B as countable unions of members of $\mathcal{E}$, we have shown that they are events. Note that in each case, these unions are taken over a family of pairwise disjoint events, from which it follows that

$$2P(A) = 2 \sum_{s \in S} P(A_s) = 3 \sum_{s \in S} P(B_s) = 3P(B).$$

Also, A and B are clearly disjoint, so

$$P(A) + P(B) = P(A \cup B) = 1 - P(A^c \cap B^c)\,.$$

We will show that $P(A^c \cap B^c) = 0$, so that the above two equalities become two equations in the two unknowns $P(A)$ and $P(B)$, the solution of which gives $P(A) = \frac{3}{5}$.

To show that $P(A^c \cap B^c) = 0$ we note that $A^c \cap B^c$ is a subset of the event D_k consisting of those ω that begin with a sequence of length $3k$ having the property that, for $1 \leq j \leq k$, the sequence $(1, 1, 1)$ does not occur in positions $3j - 2$, $3j - 1$, $3j$. The number of ways of filling the first $3k$ positions of ω with 1's and 0's is $2^{3k} = 8^k$. The number of ways of doing it so as to obtain a member of D_k is 7^k (7 choices for positions $1, 2, 3$; 7 choices for positions $4, 5, 6$ and so forth.). Thus, $P(A^c \cap B^c) \leq P(D_k) = (\frac{7}{8})^k$. Now let $k \to \infty$ to obtain the desired conclusion, $P(A^c \cap B^c) = 0$.

(iv) $\frac{5}{8}$

1-14 Let $\mathcal{B}$ denote the Borel σ-field of $\mathbb{R}$, $\mathcal{C}$ the Borel σ-field of $\mathbb{R}^+$, and

$$\mathcal{G} = \{B \in \mathcal{B} \colon B \subseteq \mathbb{R}^+\}\,.$$

The goal is to prove $\mathcal{C} = \mathcal{G}$.

We first prove that $\mathcal{G}$ is a σ-field of subsets of $\mathbb{R}^+$. Countable unions of members of $\mathcal{B}$ are members of $\mathcal{B}$ and unions of subsets of $\mathbb{R}^+$ are subsets of $\mathbb{R}^+$. Hence, $\mathcal{G}$ is closed under countable unions. The complement in $\mathbb{R}^+$ of a member G of $\mathcal{G}$ equals $\mathbb{R}^+ \cap G^c$, where G^c denotes the complement in $\mathbb{R}$. This set is clearly a subset of $\mathbb{R}^+$ and it is also a member of $\mathcal{B}$ because it is the intersection of two members of $\mathcal{B}$. Therefore, $\mathcal{G}$ is a σ-field.

The open subsets of $\mathbb{R}^+$ have the form $\mathbb{R}^+ \cap O$, where O is open in $\mathbb{R}$. Such sets, being subsets of $\mathbb{R}^+$ and intersections of two members of $\mathcal{B}$, are members of $\mathcal{G}$. Thus, the σ-field $\mathcal{G}$ contains the σ-field generated by the collection of these open subsets—namely $\mathcal{C}$.

To show that $\mathcal{G} \subseteq \mathcal{C}$ we introduce the Borel σ-field $\mathcal{D}$ of subsets of $(-\infty, 0)$ with the relative topology and set

$$\mathcal{H} = \{C \cup D \colon C \in \mathcal{C},\, D \in \mathcal{D}\}\,.$$

We can finish the proof by showing that $\mathcal{B} \subseteq \mathcal{H}$, because $\mathcal{C}$ consists of those members of $\mathcal{H}$ which are subsets of $\mathbb{R}^+$. It is clear that $\mathcal{H}$ is closed under countable unions. The formula

$$(C \cup D)^c = (\mathbb{R}^+ \backslash C) \,\cup\, ((-\infty, 0) \backslash D)$$

for $C \subseteq \mathbb{R}^+$ and $D \subseteq (-\infty, 0)$ shows that it is closed under complementation. So $\mathcal{H}$ is a σ-field. For any open set $O \in \mathbb{R}$, the representation

$$O = (\mathbb{R}^+ \cap O) \,\cup\, ((-\infty, 0) \cap O)$$

represents O as the union of open, and therefore Borel, subsets of the spaces $\mathbb{R}^+$ and $(-\infty, 0)$. Thus, the σ-field $\mathcal{H}$ contains the σ-field generated by the collection of open subsets of $\mathbb{R}$ —namely $\mathcal{B}$.

1-16 *Hint:* It suffices to show that every open set is the union of open boxes having edges of rational length and centers with rational coordinates.

For Chapter 2

2-2. Let X be a continuous function. For any open B of the target of X, $X^{-1}(B)$ is open by continuity, and thus is an event in the domain of X. Now apply Proposition 3 with $\mathcal{E}$ equal to the collection of open subsets in the target of X.

2-3. Let B be an arbitrary measurable set in the common target of X and Y. We need to show that

$$P(\{\omega \colon X(\omega) \in B\}) = P(\{\omega \colon Y(\omega) \in B\})\,.$$

Here is the relevant calculation:

$$\begin{aligned}
&P(\{\omega \colon X(\omega) \in B\}) \\
&= P(\{\omega \colon X(\omega) \in B,\, Y(\omega) = X(\omega)\}) + P(\{\omega \colon X(\omega) \in B,\, Y(\omega) \neq X(\omega)\}) \\
&= P(\{\omega \colon X(\omega) \text{ and } Y(\omega) \in B,\, Y(\omega) = X(\omega)\})\,.
\end{aligned}$$

In this calculation, we used the fact that the event in the second term of the second line is contained in a null event. To complete the proof, carry out a similar calculation with the roles of X and Y reversed.

2-9. By Problem 13 of Chapter 1 and Proposition 3 we only need show that the set $A = \{\omega : X(\omega) \leq c\}$ is a Borel set for every c (or even just for every rational c). Let a equal the least upper bound of A. We will prove that every member of the interval $(-\infty, a)$ belongs to A. Suppose $\omega_1 < a$. Since a is the least upper bound of A, there exists $\omega_2 \in A$ for which $\omega_1 < \omega_2$. Then

$$X(\omega_1) \leq X(\omega_2) \leq c\,,$$

from which it follows that $\omega_1 \in A$. Thus, A is an interval of the form $(-\infty, a)$ or $(-\infty, a]$ and is, therefore, Borel.

2-12. $\frac{5}{9}$

2-14. The distribution is uniform on the triangle $\{(v_1, v_2) : 0 < v_1 < v_2 < 1\}$. If B is a set for which area is defined, the value that the distribution assigns to B is twice its area, the factor of 2 arising because the triangle has area $\frac{1}{2}$. To prove that X is a random variable—*Hint:* Prove that X is continuous, or, alternatively, avoid the issue of continuity of a $\mathbb{R}^2$-valued function by first doing Problem 16 and then using it in conjunction with a proof that each coordinate function is continuous.

2-19. In case k is divisible by 4, the answer is

$$\binom{k/2}{k/4} 2^{-k}\,.$$

Otherwise, the answer is 0.

2-21. The Hausdorff distances are $\frac{1+\sqrt{2}}{2}$ between the first two; $\frac{1}{2}$ between the first and third; $\frac{2+\sqrt{2}}{2}$ between the second and third.

2-22. These are the probabilities: $\frac{2-\sqrt{2}}{2\pi}$, $\frac{1}{16}$, $\frac{\pi-2}{8\pi}$.

For Chapter 3

3-3. Fix ω. Since F is increasing, every member of $\{x\colon F(x) < \omega\}$ is less than every member of $\{x\colon F(x) \geq \omega\}$ and is thus a lower bound of $\{x\colon F(x) \geq \omega\}$. Hence $Y(\omega) \stackrel{\text{def}}{=} \sup\{x\colon F(x) < \omega\}$ is a lower bound of $\{x\colon F(x) \geq \omega\}$. Therefore $Y(\omega) \leq X(\omega)$.

To prove $Y(\omega) = X(\omega)$, suppose, for a proof by contradiction, that $Y(\omega) < X(\omega)$, and consider an $x \in (Y(\omega), X(\omega))$. Either $F(x) \geq \omega$ contradicting the defining property of $X(\omega)$ or $F(x) < \omega$ contradicting the defining property of $Y(\omega)$. Thus $Y = X$, and we will work with Y in the next paragraph.

Clearly, Y is increasing. Thus, to show left continuity we only need show $Y(\omega-) \geq Y(\omega)$ for every ω. Let $\delta > 0$. There exists $u > Y(\omega) - \delta$ for which $F(u) < \omega$. Hence there exists $\tau < \omega$ such that $F(u) < \tau$. Therefore

$$Y(\omega-) \geq Y(\tau) \geq u > Y(\omega) - \delta\,.$$

Now let $\delta \searrow 0$.

3-8. Whether a or b is finite or infinite,

$$Q((a,b)) = \int_a^b \frac{1}{\pi(1+x^2)}\,dx\,.$$

When a and b are finite this formula is also a formula for $Q([a,b])$, and similarly for $Q([a,b))$ and $Q((a,b])$ in case $a > -\infty$ or $b < \infty$, respectively. Note that the formula for $Q([a,b])$ is correct in the special case $a = b$.

3-12. Explanation for 'type' only. Suppose first that F_1 and F_2 are of the same type. Then there exist random variables X_1 and X_2 of the same type such that F_j is the distribution function of X_j. Then F_2 is also the distribution function of $aX_1 + b$ for some a and b with $a > 0$. Thus

$$F_2(x) = P(\{\omega : aX_1(\omega) + b \le x\}) = P(\{\omega : X_1(\omega) \le (x-b)/a\}) = F_1((x-b)/a)\,.$$

That is F_1 and F_2 must satisfy (3.2).

Conversely, suppose that F_2 and F_1 satisfy (3.2) for some a and b with $a > 0$. Let X_1 be a random variable with distribution function F_1. the above calculation then shows that $aX_1 + b$ is a random variable whose distribution function is F_2. Therefore F_2 is of the same type as F_1.

3-23. X is symmetric about b if and only if its distribution function F satisfies $F(x-b) = 1 - F((b-x)-)$ for all x.

For the standard Cauchy distribution

$$\begin{aligned}F(x) &= \frac{1}{2} + \frac{\arctan x}{\pi} = \frac{1}{2} + \frac{-\arctan(-x)}{\pi}\\ &= 1 - \Big(\frac{1}{2} + \frac{\arctan(-x)}{\pi}\Big) = 1 - \Big(\frac{1}{2} + \frac{\arctan((-x)-)}{\pi}\Big) = 1 - F((-x)-)\,.\end{aligned}$$

3-28. A random variable X having the Cauchy distribution of Problem 8 has density $x \rightsquigarrow \frac{1}{\pi(1+x^2)}$. For positive a and real b the continuous density of $aX + b$ is $x \rightsquigarrow \frac{a}{\pi(a^2+(x-b)^2)}$.

The density of the uniform distribution with support $[a,b]$ is $\frac{1}{b-a}$ on the interval $[a,b]$ and 0 elsewhere.

3-30.

$$\int_0^\infty ae^{-ax}\,dx = -e^{-ax}\Big|_0^\infty = 1$$
$$P(\{\omega\colon 2 \le X(\omega) \le 3\}) = e^{-2a} - e^{-3a}$$
$$\text{median} = a^{-1}\log 2$$

3-33. $g(x) = \frac{1}{2\sqrt{x}}[f(\sqrt{x}) + f(-\sqrt{x})]$ if $x > 0$ and $g(x) = 0$ if $x \le 0$.

3-34. (i)

$$\begin{aligned}\Gamma(\gamma+1) = \int_0^\infty u^\gamma e^{-u}\,du &= -\int_0^\infty u^\gamma\,de^{-u} \\ &= \int_0^\infty \gamma u^{\gamma-1}e^{-u}\,du = \gamma\Gamma(\gamma)\,.\end{aligned}$$

(ii) An easy calculation gives $\Gamma(1) = 0!$. For an induction proof assume that $\Gamma(\gamma) = (\gamma-1)!$ for some positive integer γ. By part (i),

$$\Gamma(\gamma+1) = \gamma\Gamma(\gamma) = \gamma[(\gamma-1)!] = \gamma!\,.$$

Note that the last step in the above calculation is valid for $\gamma = 1$. That this step be valid is one of the motivations for the definition $0! = 1$.

(iii)

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty u^{-1/2}e^{-u}\,du = \int_0^\infty \sqrt{2}\,e^{-v^2/2}\,dv\,,$$

which, by Example 1 and symmetry, equals $\sqrt{\pi}$. Now use mathematical induction.

(iv)

$$\begin{aligned}\Gamma(\alpha)\Gamma(\beta) &= \int_0^\infty\int_0^\infty u^{\alpha-1}v^{\beta-1}e^{-(u+v)}\,du\,dv \\ &= \int_0^\infty\int_v^\infty (w-v)^{\alpha-1}v^{\beta-1}e^{-w}\,dw\,dv \\ &= \int_0^\infty\int_0^w (w-v)^{\alpha-1}v^{\beta-1}e^{-w}\,dv\,dw \\ &= \int_0^\infty\int_0^1 w^{\alpha+\beta-1}(1-x)^{\alpha-1}x^{\beta-1}e^{-w}\,dx\,dw\,,\end{aligned}$$

the interchange of order of integration being valid, according to a result from advanced calculus, because the integrand is continuous and nonnegative. (The validity of the interchange in integration order is also a consequence of the Fubini Theorem, to be proved in Chapter 9.) The last expression is the product of the two desired integrals.

3-40. *Hint:* For $b = 1$ and $x \geq 0$,

$$P(\{\omega\colon -\log X(\omega) \leq x\}) = P(\{\omega\colon X(\omega) \geq e^{-x}) = 1 - e^{-x}\,,$$

the standard exponential distribution function.

3-41. Denote the three distribution functions by G_i, $i = 2, 3, 4$. For each i, $G_i(y) = 0$ when $y < 0$ and $= 1$ when $y \geq 2$. For $0 \leq y < 2$:

$$\begin{aligned}G_2(y) &= 1 - \sqrt{1 - \tfrac{y^2}{4}}\,;\\ G_3(y) &= \tfrac{2}{\pi}\arcsin\tfrac{y}{2}\,;\\ G_4(y) &= \tfrac{y^2}{4}\,.\end{aligned}$$

For Chapter 4

4-7. $n^2p^2 + npq$ (notation of Problem 39 of Chapter 3)

4-8. $7/2$

4-9. For this problem, denote the expectation operator according to Definition 1 by E_s and the expectation operator according to Definition 5 by E_p. Let X be nonnegative and simple. Thus, $E_s(X)$ and $E_p(X)$ are meaningful. Since X qualifies as an appropriate Z in the definition $E_p(X) = \sup_Z E_s(Z)$, we see that $E_p(X) \geq E_s(X)$. On the other hand, Lemma 4 implies that for all simple $Z \leq X$, $E_s(Z) \leq E_s(X)$, from which it follows immediately that $E_p(X) \leq E_s(X)$.

4-10. The random variable X defined by $X(\omega) = \frac{1}{\omega}$, defined on the probability space $((0,1], \mathcal{B}, P)$, where P denotes Lebesgue measure, has expected value ∞. This is seen by calculating $E(X_n)$ for simple random variables $X_n \leq X$ defined by $X_n(\omega) = (\lfloor X(\omega) \rfloor) \wedge n$.

4-11. We treat the case $a = b = 1$. The following calculation based on the definition of expectation for nonnegative random variables and the linearity of the expectation for simple random variables shows that $E(X) + E(Y) \leq E(X+Y)$:

$$\begin{aligned}
&E(X) + E(Y) \\
&= \sup\{E(X') \colon X' \leq X \text{ and } X' \text{ simple}\} + \sup\{E(Y') \colon Y' \leq Y \text{ and } Y' \text{ simple}\} \\
&= \sup\{E(X') + E(Y') \colon X' \leq X, Y' \leq Y \text{ and } X', Y' \text{ simple}\} \\
&= \sup\{E(X' + Y') \colon X' \leq X, Y' \leq Y \text{ and } X', Y' \text{ simple}\} \\
&\leq \sup\{E(Z) \colon Z \leq X + Y \text{ and } Z \text{ simple}\} = E(X+Y)\,.
\end{aligned}$$

To prove the opposite inequality, let Z be a simple random variable such that $Z \leq X + Y$. By the construction given in the proof of Lemma 13 of Chapter 2, we can find sequences $(X_n \colon n = 1, 2, \dots)$ and $(Y_n \colon n = 1, 2, \dots)$ of simple random variables such that for all ω and all n,

$$\begin{aligned}
X(\omega) \wedge n - \frac{1}{2^n} &\leq X_n(\omega) \leq X(\omega) \quad \text{and} \\
Y(\omega) \wedge n - \frac{1}{2^n} &\leq Y_n(\omega) \leq Y(\omega)\,.
\end{aligned}$$

It is easily checked that $X_n + Y_n \geq Z - 1/2^n$ for $n \geq \max\{Z(\omega) \colon \omega \in \Omega\}$. Thus

$$\sup_n E(X_n) + \sup_n E(Y_n) \geq E(Z)\,,$$

and the desired inequality $E(X) + E(Y) \geq E(X+Y)$ now follows from the definition of expected value.

4-14. For this problem, denote the expectation operators according to Definition 1, Definition 5, and Definition 8 by E_s, E_p, and E_g, respectively. Let X be simple (but not necessarily nonnegative). We use (4.1):

$$X = \sum_{j=1}^{n} c_j I_{C_j}\,.$$

Since $\{C_j : 1 \leq j \leq n\}$ is a partition,

$$X^+ = \sum_{j:c_j \geq 0} c_j I_{C_j}$$

and

$$X^- = -\sum_{j:c_j < 0} c_j I_{C_j} .$$

For these nonnegative simple random variables we have, using Problem 9, that

$$E_p(X^+) = E_s(X^+) = \sum_{j:c_j \geq 0} c_j P(C_j)$$

and

$$E_p(X^-) = E_s(X^-) = -\sum_{j:c_j < 0} c_j P(C_j) .$$

By these formulas and Definition 8,

$$E_g(X) = E_p(X^+) - E_p(X^-) = \sum_{j=1}^{n} c_j P(C_j) = E_s(X) .$$

4-21. The case where $E(X_1) = +\infty$ is easily treated, so we assume $E(X_1)$ is finite and, therefore, $P(\{\omega : |X_1(\omega)| = \infty\}) = 0$. Accordingly, except for ω belonging to some null set, we may define $Y_n(\omega) = X_n(\omega) - X_1(\omega)$ and $Y(\omega) = X(\omega) - X_1(\omega)$. For ω in the null set we set $Y_n(\omega) = Y(\omega) = 0$. Applying the Monotone Convergence Theorem to the sequence $(Y_1, Y_2, \dots)$, we deduce that $E(Y_n) \to E(Y)$ as $n \to \infty$. It follows, by property (iii) of Theorem 9, that

$$\lim_{n\to\infty} E(X_n - X_1) \to E(X - X_1) .$$

Since $E(X_1)$ is finite we may apply property (i) of Theorem 9 to conclude

$$\lim_{n\to\infty} [E(X_n) - E(X_1)] \to E(X) - E(X_1) .$$

Now add $E(X_1)$ to both sides.

4-22. $E(X) = \frac{p}{1-p}$, $E(X^2) = \frac{p(1+p)}{(1-p)^2}$ (notation of Problem 11 of Chapter 3)

4-23 $E(X) = \lambda$, $E(X^2) = \lambda + \lambda^2$ (notation of Problem 37 of Chapter 3)

4-26 The distributions of $X - b$ and $b - X$ are identical. By Theorem 15 they have the same mean. By properties (i) and (ii) of Theorem 9, these equal numbers are $E(X) - b$ and $b - E(X)$. It follows that $E(X) = b$.

4-29. b (notation of Example 1 of Chapter 3)

4-30. For standard beta distributions (that is, beta distributions with support $[0, 1]$), the answer is $\frac{\alpha}{\alpha+\beta}$ (notation of Example 3 of Chapter 3).

4-31. $E(X) = 1/k$, $E(\exp \circ X) = \infty$ if $k \leq 1$ and $= \frac{k}{k-1}$ if $k > 1$

4-35. $E(X_1) = E(X_3) = \frac{4}{\pi}$, $E(X_2) = \frac{\pi}{2}$, $E(X_4) = \frac{4}{3}$

For Chapter 5

5-7. $\mathrm{Var}(X_1) = \mathrm{Var}(X_3) = \frac{2(\pi^2-8)}{\pi^2}$, $\mathrm{Var}(X_2) = \frac{32-3\pi^2}{12}$, $\mathrm{Var}(X_4) = \frac{2}{9}$

5-13. $\mathrm{Cov}(aX+b, cY+d) = ac\,\mathrm{Cov}(X,Y)$

5-14. An inequality, based on the fact that φ is increasing will be useful:

$$[X - E(X)]\,(\varphi \circ X) \geq [X - E(X)]\,\varphi(E(X))\,.$$

The following calculation then completes the proof:

$$\begin{aligned}\mathrm{Cov}(X,\, \varphi \circ X) &= E\big([X - E(X)][\varphi \circ X - E(\varphi \circ X)]\big)\\ &= E\big([X - E(X)]\,(\varphi \circ X)\big)\\ &\geq E\big([X - E(X)]\,\varphi(E(X))\big) = 0\,.\end{aligned}$$

(A slightly longer, but possibly more transparent proof, consists of first reducing the problem to the case where $E(X) = 0$ and then using the above argument for that special case.)

5-17. For Example 2 of Chapter 4, the answer is 1; for Problem 18 of Chapter 4, the answer is 0 or 1 according as n is odd or even.

5-29. $s(0,0) = s(1,1) = s(2,2) = s(3,3) = 1$, $s(1,0) = s(2,0) = s(3,0) = 0$, $s(2,1) = -1$, $s(3,1) = 2$, $s(3,2) = -3$, $s(n,k) = 0$ for $k > n$; $S(0,0) = S(1,1) = S(2,2) = S(3,3) = 1$, $S(1,0) = S(2,0) = S(3,0) = 0$, $S(2,1) = S(3,1) = 1$, $S(3,2) = 3$, $S(n,k) = 0$ for $k > n$

5-32. $\rho(1-) = 1$. Thus, if ρ is the probability generating function of a distribution Q, then $Q(\{\infty\}) = 0$. To both show that ρ is a probability generating function and calculate $Q(\{k\})$ for each $k \in \mathbb{Z}^+$ we rewrite $\rho(s)$ using partial fractions:

$$\begin{aligned}\rho(s) &= \frac{-24}{2-s} + \frac{8}{(2-s)^2} + \frac{24}{3-s} + \frac{16}{(3-s)^2} + \frac{8}{(3-s)^3}\\ &= \frac{-12}{1-(s/2)} + \frac{2}{(1-(s/2))^2} + \frac{8}{1-(s/3)} + \frac{16/9}{(1-(s/3))^2} + \frac{8/27}{(1-(s/3))^3}\,.\end{aligned}$$

The first two of the last five functions are equal to their power series for $|s| < 2$ and the last three for $|s| < 3$. So we can expand in power series and collect coefficients to get a power series for $\rho(s)$ that can be differentiated term-by-term to obtain the derivatives of $\rho(s)$. Thus, we only need to show that the coefficients are nonnegative in order to conclude that $\rho(s)$ is a probability generating function, and then the coefficients are the values $Q(\{k\})$.

Formulas for the geometric series and its derivatives give

$$\begin{aligned}\rho(s) = -12\sum_{k=0}^{\infty}(\tfrac{s}{2})^k + 2\sum_{k=0}^{\infty}(k+1)(\tfrac{s}{2})^k + 8\sum_{k=0}^{\infty}(\tfrac{s}{3})^k&\\ + \tfrac{16}{9}\sum_{k=0}^{\infty}(k+1)(\tfrac{s}{3})^k + \tfrac{4}{27}\sum_{k=0}^{\infty}(k+1)(k+2)(\tfrac{s}{3})^k&\,.\end{aligned}$$

When we collect terms we get nonnegative—in fact, positive—terms, as desired:

$$Q(\{k\}) = \frac{k-5}{2^{k-1}} + \frac{4k^2+60k+272}{3^{k+3}}\,.$$

To get the mean and variance it seems best to work with $\rho(s)$ in the form originally given and use the product rule to get the first and second derivatives:

$$\rho'(s) = \frac{16}{(2-s)^3(3-s)^3} + \frac{24}{(2-s)^2(3-s)^4}$$

and

$$\rho''(s) = \frac{48}{(2-s)^4(3-s)^3} + \frac{96}{(2-s)^3(3-s)^4} + \frac{96}{(2-s)^2(3-s)^5}\,.$$

Insertion of 1 for s gives

$$\rho'(1) = \frac{7}{2} \quad \text{and} \quad \rho''(1) = 15\,.$$

Hence, the mean equals $\frac{7}{2}$ and the second moment equals $15 + \frac{7}{2} = \frac{37}{2}$. Therefore, the variance equals $\frac{74}{4} - \frac{49}{4} = \frac{25}{4}$ and the standard deviation equals $\frac{5}{2}$.

Had the problem only been to verify that ρ is a probability generating function, we could have, while calculating the first and second derivatives, seen that a straightforward induction proof would show that all derivatives are positive, and an appeal to Theorem 14 would complete the proof.

5-33. The mean is ∞ and thus the variance is undefined. The distribution Q_p corresponding to the probability generating function with parameter p satisfies $Q_p(\{\infty\}) = |1-2p|$. Also, for $0 < k = 2m < \infty$,

$$Q_p(\{2m\}) = \frac{2}{m}\binom{2m-2}{m-1}[p(1-p)]^m\,.$$

For k odd and $k = 0$, $Q_p(\{k\}) = 0$.

For Chapter 6

6-6. Method 1: Using Problem 4, we get

$$\begin{aligned}(\liminf_{n\to\infty} A_n)^c &= \left(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m\right)^c \\ &= \bigcap_{n=1}^{\infty}\left(\bigcap_{m=n}^{\infty} A_m\right)^c \\ &= \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m^c = \limsup_{n\to\infty} A_n^c\,.\end{aligned}$$

Method 2: We prove that the indicator functions of the two sets are equal:

$$\begin{aligned}I_{(\limsup A_n)^c} &= 1 - I_{\limsup A_n} = 1 - \limsup_n\{I_{A_n}\} \\ &= \liminf_n\{(1-I_{A_n})\} = \liminf_n\{I_{A_n^c}\} = I_{\liminf A_n^c}\,.\end{aligned}$$