Chapter 1

Probability Theory

"If any little problem comes your way, I shall be happy, if I can, to give you a hint or two as to its solution."

Sherlock Holmes
The Adventure of the Three Students

1.1 a. Each sample point describes the result of the toss (H or T) for each of the four tosses. So, for example THTT denotes T on 1st, H on 2nd, T on 3rd and T on 4th. There are $2^4 = 16$ such sample points.

b. The number of damaged leaves is a nonnegative integer. So we might use $S = \{0, 1, 2, \ldots\}$.

c. We might observe fractions of an hour. So we might use $S = \{t : t \geq 0\}$, that is, the half infinite interval $[0, \infty)$.

d. Suppose we weigh the rats in ounces. The weight must be greater than zero so we might use $S = (0, \infty)$. If we know no 10-day-old rat weighs more than 100 oz., we could use $S = (0, 100]$.

e. If n is the number of items in the shipment, then $S = \{0/n, 1/n, \ldots, 1\}$.

1.2 For each of these equalities, you must show containment in both directions.

a. $x \in A\backslash B \Leftrightarrow x \in A$ and $x \notin B \Leftrightarrow x \in A$ and $x \notin A \cap B \Leftrightarrow x \in A\backslash(A \cap B)$. Also, $x \in A$ and $x \notin B \Leftrightarrow x \in A$ and $x \in B^c \Leftrightarrow x \in A \cap B^c$.

b. Suppose $x \in B$. Then either $x \in A$ or $x \in A^c$. If $x \in A$, then $x \in B \cap A$, and, hence $x \in (B \cap A) \cup (B \cap A^c)$. Thus $B \subset (B \cap A) \cup (B \cap A^c)$. Now suppose $x \in (B \cap A) \cup (B \cap A^c)$. Then either $x \in (B \cap A)$ or $x \in (B \cap A^c)$. If $x \in (B \cap A)$, then $x \in B$. If $x \in (B \cap A^c)$, then $x \in B$. Thus $(B \cap A) \cup (B \cap A^c) \subset B$. Since the containment goes both ways, we have $B = (B \cap A) \cup (B \cap A^c)$. (Note, a more straightforward argument for this part simply uses the Distributive Law to state that $(B \cap A) \cup (B \cap A^c) = B \cap (A \cup A^c) = B \cap S = B$.)

c. Similar to part a).

d. From part b).
$A \cup B = A \cup [(B \cap A) \cup (B \cap A^c)] = A \cup (B \cap A) \cup A \cup (B \cap A^c) = A \cup [A \cup (B \cap A^c)] = A \cup (B \cap A^c)$.

1.3 a. $x \in A \cup B \;\Leftrightarrow\; x \in A$ or $x \in B \;\Leftrightarrow\; x \in B \cup A$
$x \in A \cap B \;\Leftrightarrow\; x \in A$ and $x \in B \;\Leftrightarrow\; x \in B \cap A$.

b. $x \in A \cup (B \cup C) \Leftrightarrow\; x \in A$ or $x \in B \cup C \Leftrightarrow\; x \in A \cup B$ or $x \in C \Leftrightarrow\; x \in (A \cup B) \cup C$.
(It can similarly be shown that $A \cup (B \cup C) = (A \cup C) \cup B$.)
$x \in A \cap (B \cap C) \Leftrightarrow\; x \in A$ and $x \in B$ and $x \in C \Leftrightarrow\; x \in (A \cap B) \cap C$.

c. $x \in (A \cup B)^c \Leftrightarrow\; x \notin A$ or $x \notin B \Leftrightarrow\; x \in A^c$ and $x \in B^c \Leftrightarrow\; x \in A^c \cap B^c$
$x \in (A \cap B)^c \Leftrightarrow\; x \notin A \cap B \Leftrightarrow\; x \notin A$ and $x \notin B \Leftrightarrow\; x \in A^c$ or $x \in B^c \Leftrightarrow\; x \in A^c \cup B^c$.

1.4 a. "A or B or both" is $A \cup B$. From Theorem 1.2.9b we have $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

b. "A or B but not both" is $(A\cap B^c)\cup(B\cap A^c)$. Thus we have

$$\begin{aligned} P((A\cap B^c)\cup(B\cap A^c)) &= P(A\cap B^c)+P(B\cap A^c) && \text{(disjoint union)}\\ &= [P(A)-P(A\cap B)]+[P(B)-P(A\cap B)] && \text{(Theorem1.2.9a)}\\ &= P(A)+P(B)-2P(A\cap B). \end{aligned}$$

c. "At least one of A or B" is $A\cup B$. So we get the same answer as in a).

d. "At most one of A or B" is $(A\cap B)^c$, and $P((A\cap B)^c)=1-P(A\cap B)$.

1.5 a. $A\cap B\cap C=\{$a U.S. birth results in identical twins that are female$\}$

b. $P(A\cap B\cap C)=\frac{1}{90}\times\frac{1}{3}\times\frac{1}{2}$

1.6

$$p_0=(1-u)(1-w),\quad p_1=u(1-w)+w(1-u),\quad p_2=uw,$$

$$\begin{aligned} p_0=p_2 &\Rightarrow u+w=1\\ p_1=p_2 &\Rightarrow uw=1/3. \end{aligned}$$

These two equations imply $u(1-u)=1/3$, which has no solution in the real numbers. Thus, the probability assignment is not legitimate.

1.7 a.

$$P(\text{scoring } i \text{ points})=\begin{cases} 1-\frac{\pi r^2}{A} & \text{if } i=0\\ \frac{\pi r^2}{A}\left[\frac{(6-i)^2-(5-i)^2}{5^2}\right] & \text{if } i=1,\ldots,5. \end{cases}$$

b.

$$\begin{aligned} P(\text{scoring } i \text{ points}|\text{board is hit}) &= \frac{P(\text{scoring } i \text{ points}\cap\text{board is hit})}{P(\text{board is hit})}\\ P(\text{board is hit}) &= \frac{\pi r^2}{A}\\ P(\text{scoring } i \text{ points}\cap\text{board is hit}) &= \frac{\pi r^2}{A}\left[\frac{(6-i)^2-(5-i)^2}{5^2}\right] \qquad i=1,\ldots,5. \end{aligned}$$

Therefore,

$$P(\text{scoring } i \text{ points}|\text{board is hit})=\frac{(6-i)^2-(5-i)^2}{5^2}\qquad i=1,\ldots,5$$

which is exactly the probability distribution of Example 1.2.7.

1.8 a. $P(\text{scoring exactly } i \text{ points})=P(\text{inside circle } i)-P(\text{inside circle } i+1)$. Circle i has radius $(6-i)r/5$, so

$$P(s\text{scoring exactly } i \text{ points})=\frac{\pi(6-i)^2r^2}{5^2\pi r^2}-\frac{\pi\left((6-(i+1))\right)^2r^2}{5^2\pi r^2}=\frac{(6-i)^2-(5-i)^2}{5^2}.$$

b. Expanding the squares in part a) we find $P(\text{scoring exactly } i \text{ points})=\frac{11-2i}{25}$, which is decreasing in i.

c. Let $P(i)=\frac{11-2i}{25}$. Since $i\le 5$, $P(i)\ge 0$ for all i. $P(S)=P(\text{hitting the dartboard})=1$ by definition. Lastly, $P(i\cup j)=\text{area of } i \text{ ring}+\text{area of } j \text{ ring}=P(i)+P(j)$.

1.9 a. Suppose $x\in(\cup_\alpha A_\alpha)^c$, by the definition of complement $x\notin\cup_\alpha A_\alpha$, that is $x\notin A_\alpha$ for all $\alpha\in\Gamma$. Therefore $x\in A_\alpha^c$ for all $\alpha\in\Gamma$. Thus $x\in\cap_\alpha A_\alpha^c$ and, by the definition of intersection $x\in A_\alpha^c$ for all $\alpha\in\Gamma$. By the definition of complement $x\notin A_\alpha$ for all $\alpha\in\Gamma$. Therefore $x\notin\cup_\alpha A_\alpha$. Thus $x\in(\cup_\alpha A_\alpha)^c$.

b. Suppose $x \in (\cap_\alpha A_\alpha)^c$, by the definition of complement $x \notin (\cap_\alpha A_\alpha)$. Therefore $x \notin A_\alpha$ for some $\alpha \in \Gamma$. Therefore $x \in A_\alpha^c$ for some $\alpha \in \Gamma$. Thus $x \in \cup_\alpha A_\alpha^c$ and, by the definition of union, $x \in A_\alpha^c$ for some $\alpha \in \Gamma$. Therefore $x \notin A_\alpha$ for some $\alpha \in \Gamma$. Therefore $x \notin \cap_\alpha A_\alpha$. Thus $x \in (\cap_\alpha A_\alpha)^c$.

1.10 For $A_1, \ldots, A_n$

$$(i) \quad \left(\bigcup_{i=1}^n A_i\right)^c = \bigcap_{i=1}^n A_i^c \qquad (ii) \quad \left(\bigcap_{i=1}^n A_i\right)^c = \bigcup_{i=1}^n A_i^c$$

Proof of (i): If $x \in (\cup A_i)^c$, then $x \notin \cup A_i$. That implies $x \notin A_i$ for any i, so $x \in A_i^c$ for every i and $x \in \cap A_i$.

Proof of (ii): If $x \in (\cap A_i)^c$, then $x \notin \cap A_i$. That implies $x \in A_i^c$ for some i, so $x \in \cup A_i^c$.

1.11 We must verify each of the three properties in Definition 1.2.1.

a. (1) The empty set $\emptyset \in \{\emptyset, S\}$. Thus $\emptyset \in \mathcal{B}$. (2) $\emptyset^c = S \in \mathcal{B}$ and $S^c = \emptyset \in \mathcal{B}$. (3) $\emptyset \cup S = S \in \mathcal{B}$.

b. (1) The empty set $\emptyset$ is a subset of any set, in particular, $\emptyset \subset S$. Thus $\emptyset \in \mathcal{B}$. (2) If $A \in \mathcal{B}$, then $A \subset S$. By the definition of complementation, A^c is also a subset of S, and, hence, $A^c \in \mathcal{B}$. (3) If $A_1, A_2, \ldots \in \mathcal{B}$, then, for each i, $A_i \subset S$. By the definition of union, $\cup A_i \subset S$. Hence, $\cup A_i \in \mathcal{B}$.

c. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be the two sigma algebras. (1) $\emptyset \in \mathcal{B}_1$ and $\emptyset \in \mathcal{B}_2$ since $\mathcal{B}_1$ and $\mathcal{B}_2$ are sigma algebras. Thus $\emptyset \in \mathcal{B}_1 \cap \mathcal{B}_2$. (2) If $A \in \mathcal{B}_1 \cap \mathcal{B}_2$, then $A \in \mathcal{B}_1$ and $A \in \mathcal{B}_2$. Since $\mathcal{B}_1$ and $\mathcal{B}_2$ are both sigma algebra $A^c \in \mathcal{B}_1$ and $A^c \in \mathcal{B}_2$. Therefore $A^c \in \mathcal{B}_1 \cap \mathcal{B}_2$. (3) If $A_1, A_2, \ldots \in \mathcal{B}_1 \cap \mathcal{B}_2$, then $A_1, A_2, \ldots \in \mathcal{B}_1$ and $A_1, A_2, \ldots \in \mathcal{B}_2$. Therefore, since $\mathcal{B}_1$ and $\mathcal{B}_2$ are both sigma algebra, $\cup_{i=1}^\infty A_i \in \mathcal{B}_1$ and $\cup_{i=1}^\infty A_i \in \mathcal{B}_2$. Thus $\cup_{i=1}^\infty A_i \in \mathcal{B}_1 \cap \mathcal{B}_2$.

1.12 First write

$$\begin{aligned} P\left(\bigcup_{i=1}^\infty A_i\right) &= P\left(\bigcup_{i=1}^n A_i \cup \bigcup_{i=n+1}^\infty A_i\right) \\ &= P\left(\bigcup_{i=1}^n A_i\right) + P\left(\bigcup_{i=n+1}^\infty A_i\right) \qquad (A_i\text{s are disjoint}) \\ &= \sum_{i=1}^n P(A_i) + P\left(\bigcup_{i=n+1}^\infty A_i\right) \qquad (\text{finite additivity}) \end{aligned}$$

Now define $B_k = \bigcup_{i=k}^\infty A_i$. Note that $B_{k+1} \subset B_k$ and $B_k \to \phi$ as $k \to \infty$. (Otherwise the sum of the probabilities would be infinite.) Thus

$$P\left(\bigcup_{i=1}^\infty A_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^\infty A_i\right) = \lim_{n\to\infty}\left[\sum_{i=1}^n P(A_i) + P(B_{n+1})\right] = \sum_{i=1}^\infty P(A_i).$$

1.13 If A and B are disjoint, $P(A \cup B) = P(A) + P(B) = \frac{1}{3} + \frac{3}{4} = \frac{13}{12}$, which is impossible. More generally, if A and B are disjoint, then $A \subset B^c$ and $P(A) \le P(B^c)$. But here $P(A) > P(B^c)$, so A and B cannot be disjoint.

1.14 If $S = \{s_1, \ldots, s_n\}$, then any subset of S can be constructed by either including or excluding s_i, for each i. Thus there are 2^n possible choices.

1.15 Proof by induction. The proof for $k = 2$ is given after Theorem 1.2.14. Assume true for k, that is, the entire job can be done in $n_1 \times n_2 \times \cdots \times n_k$ ways. For $k+1$, the $k+1$th task can be done in n_{k+1} ways, and for each one of these ways we can complete the job by performing the remaining k tasks. Thus for each of the n_{k+1} we have $n_1 \times n_2 \times \cdots \times n_k$ ways of completing the job by the induction hypothesis. Thus, the number of ways we can do the job is $\underbrace{(1 \times (n_1 \times n_2 \times \cdots \times n_k)) + \cdots + (1 \times (n_1 \times n_2 \times \cdots \times n_k))}_{n_{k+1}\text{terms}} = n_1 \times n_2 \times \cdots \times n_k \times n_{k+1}$.

1.16 a) 26^3. b) $26^3 + 26^2$. c) $26^4 + 26^3 + 26^2$.

1.17 There are $\binom{n}{2} = n(n-1)/2$ pieces on which the two numbers do not match. (Choose 2 out of n numbers without replacement.) There are n pieces on which the two numbers match. So the total number of different pieces is $n + n(n-1)/2 = n(n+1)/2$.

1.18 The probability is $\frac{\binom{n}{2}n!}{n^n} = \frac{(n-1)(n-1)!}{2n^{n-2}}$. There are many ways to obtain this. Here is one. The denominator is n^n because this is the number of ways to place n balls in n cells. The numerator is the number of ways of placing the balls such that exactly one cell is empty. There are n ways to specify the empty cell. There are $n-1$ ways of choosing the cell with two balls. There are $\binom{n}{2}$ ways of picking the 2 balls to go into this cell. And there are $(n-2)!$ ways of placing the remaining $n-2$ balls into the $n-2$ cells, one ball in each cell. The product of these is the numerator $n(n-1)\binom{n}{2}(n-2)! = \binom{n}{2}n!$.

1.19 a. $\binom{6}{4} = 15$.

b. Think of the n variables as n bins. Differentiating with respect to one of the variables is equivalent to putting a ball in the bin. Thus there are r unlabeled balls to be placed in n unlabeled bins, and there are $\binom{n+r-1}{r}$ ways to do this.

1.20 A sample point specifies on which day (1 through 7) each of the 12 calls happens. Thus there are 7^{12} equally likely sample points. There are several different ways that the calls might be assigned so that there is at least one call each day. There might be 6 calls one day and 1 call each of the other days. Denote this by 6111111. The number of sample points with this pattern is $7\binom{12}{6}6!$. There are 7 ways to specify the day with 6 calls. There are $\binom{12}{6}$ to specify which of the 12 calls are on this day. And there are 6! ways of assigning the remaining 6 calls to the remaining 6 days. We will now count another pattern. There might be 4 calls on one day, 2 calls on each of two days, and 1 call on each of the remaining four days. Denote this by 4221111. The number of sample points with this pattern is $7\binom{12}{4}\binom{6}{2}\binom{8}{2}\binom{6}{2}4!$. (7 ways to pick day with 4 calls, $\binom{12}{4}$ to pick the calls for that day, $\binom{6}{2}$ to pick two days with two calls, $\binom{8}{2}$ ways to pick two calls for lowered numbered day, $\binom{6}{2}$ ways to pick the two calls for higher numbered day, 4! ways to order remaining 4 calls.) Here is a list of all the possibilities and the counts of the sample points for each one.

pattern	number of sample points	
6111111	$7\binom{12}{6}6! =$	4,656,960
5211111	$7\binom{12}{5}6\binom{7}{2}5! =$	83,825,280
4221111	$7\binom{12}{4}\binom{6}{2}\binom{8}{2}\binom{6}{2}4! =$	523,908,000
4311111	$7\binom{12}{4}6\binom{8}{3}5! =$	139,708,800
3321111	$\binom{7}{2}\binom{12}{3}\binom{9}{3}5\binom{6}{2}4! =$	698,544,000
3222111	$7\binom{12}{3}\binom{6}{3}\binom{9}{3}\binom{7}{2}\binom{5}{2}3! =$	1,397,088,000
2222211	$\binom{7}{5}\binom{12}{2}\binom{10}{2}\binom{8}{2}\binom{6}{2}\binom{4}{2}2! =$	314,344,800
		3,162,075,840

The probability is the total number of sample points divided by 7^{12}, which is $\frac{3{,}162{,}075{,}840}{7^{12}} \approx .2285$.

1.21 The probability is $\frac{\binom{n}{2r}2^{2r}}{\binom{2n}{2r}}$. There are $\binom{2n}{2r}$ ways of choosing $2r$ shoes from a total of $2n$ shoes. Thus there are $\binom{2n}{2r}$ equally likely sample points. The numerator is the number of sample points for which there will be no matching pair. There are $\binom{n}{2r}$ ways of choosing $2r$ different shoes styles. There are two ways of choosing within a given shoe style (left shoe or right shoe), which gives 2^{2r} ways of arranging each one of the $\binom{n}{2r}$ arrays. The product of this is the numerator $\binom{n}{2r}2^{2r}$.

1.22 a) $\frac{\binom{31}{15}\binom{29}{15}\binom{31}{15}\binom{30}{15}\cdots\binom{31}{15}}{\binom{366}{180}}$ b) $\frac{\frac{336}{366}\frac{335}{365}\cdots\frac{316}{336}}{\binom{366}{30}}$.

1.23

$$\begin{aligned}
P(\text{ same number of heads }) &= \sum_{x=0}^{n} P(1^{st} \text{ tosses } x, 2^{nd} \text{ tosses } x) \\
&= \sum_{x=0}^{n}\left[\binom{n}{x}\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{n-x}\right]^2 = \left(\frac{1}{4}\right)^{n}\sum_{x=0}^{n}\binom{n}{x}^2.
\end{aligned}$$

1.24 a.

$$\begin{aligned}
P(A \text{ wins}) &= \sum_{i=1}^{\infty} P(A \text{ wins on } i^{th} \text{ toss}) \\
&= \frac{1}{2}+\left(\frac{1}{2}\right)^2\frac{1}{2}+\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)+\cdots = \sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^{2i+1} = 2/3.
\end{aligned}$$

b. $P(A \text{ wins}) = p+(1-p)^2p+(1-p)^4p+\cdots = \sum_{i=0}^{\infty} p(1-p)^{2i} = \frac{p}{1-(1-p)^2}$.

c. $\frac{d}{dp}\left(\frac{p}{1-(1-p)^2}\right) = \frac{p^2}{[1-(1-p)^2]^2} > 0$. Thus the probability is increasing in p, and the minimum is at zero. Using L'Hôpital's rule we find $\lim_{p\to 0}\frac{p}{1-(1-p)^2} = 1/2$.

1.25 Enumerating the sample space gives $S' = \{(B,B),(B,G),(G,B),(G,G)\}$,with each outcome equally likely. Thus $P(\text{at least one boy}) = 3/4$ and $P(\text{both are boys}) = 1/4$, therefore

$$P(\text{ both are boys} \mid \text{at least one boy }) = 1/3.$$

An ambiguity may arise if order is not acknowledged, the space is $S' = \{(B,B),(B,G),(G,G)\}$, with each outcome equally likely.

1.27 a. For n odd the proof is straightforward. There are an even number of terms in the sum $(0,1,\cdots,n)$, and $\binom{n}{k}$ and $\binom{n}{n-k}$, which are equal, have opposite signs. Thus, all pairs cancel and the sum is zero. If n is even, use the following identity, which is the basis of Pascal's triangle: For $k>0$, $\binom{n}{k} = \binom{n-1}{k}+\binom{n-1}{k-1}$. Then, for n even

$$\begin{aligned}
\sum_{k=0}^{n}(-1)^k\binom{n}{k} &= \binom{n}{0}+\sum_{k=1}^{n-1}(-1)^k\binom{n}{k}+\binom{n}{n} \\
&= \binom{n}{0}+\binom{n}{n}+\sum_{k=1}^{n-1}(-1)^k\left[\binom{n-1}{k}+\binom{n-1}{k-1}\right] \\
&= \binom{n}{0}+\binom{n}{n}-\binom{n-1}{0}-\binom{n-1}{n-1} = 0.
\end{aligned}$$

b. Use the fact that for $k>0$, $k\binom{n}{k} = n\binom{n-1}{k-1}$ to write

$$\sum_{k=1}^{n} k\binom{n}{k} = n\sum_{k=1}^{n}\binom{n-1}{k-1} = n\sum_{j=0}^{n-1}\binom{n-1}{j} = n2^{n-1}.$$

c. $\sum_{k=1}^{n}(-1)^{k+1}k\binom{n}{k} = \sum_{k=1}^{n}(-1)^{k+1}\binom{n-1}{k-1} = n\sum_{j=0}^{n-1}(-1)^{j}\binom{n-1}{j} = 0$ from part a).

1.28 The average of the two integrals is

$$\begin{aligned}[(n\log n - n) + ((n+1)\log(n+1) - n)]/2 &= [n\log n + (n+1)\log(n+1)]/2 - n \\ &\approx (n+1/2)\log n - n.\end{aligned}$$

Let $d_n = \log n! - [(n+1/2)\log n - n]$, and we want to show that $\lim_{n\to\infty} md_n = c$, a constant. This would complete the problem, since the desired limit is the exponential of this one. This is accomplished in an indirect way, by working with differences, which avoids dealing with the factorial. Note that

$$d_n - d_{n+1} = \left(n+\frac{1}{2}\right)\log\left(1+\frac{1}{n}\right) - 1.$$

Differentiation will show that $((n+\frac{1}{2}))\log((1+\frac{1}{n}))$ is increasing in n, and has minimum value $(3/2)\log 2 = 1.04$ at $n = 1$. Thus $d_n - d_{n+1} > 0$. Next recall the Taylor expansion of $\log(1+x) = x - x^2/2 + x^3/3 - x^4/4 + \cdots$. The first three terms provide an upper bound on $\log(1+x)$, as the remaining adjacent pairs are negative. Hence

$$0 < d_n d_{n+1} < \left(n+\frac{1}{2}\right)\left(\frac{1}{n}\frac{1}{2n^2}+\frac{1}{3n^3}\right) - 1 = \frac{1}{12n^2}+\frac{1}{6n^3}.$$

It therefore follows, by the comparison test, that the series $\sum_1^\infty d_n - d_{n+1}$ converges. Moreover, the partial sums must approach a limit. Hence, since the sum telescopes,

$$\lim_{N\to\infty}\sum_{1}^{N} d_n - d_{n+1} = \lim_{N\to\infty} d_1 - d_{N+1} = c.$$

Thus $\lim_{n\to\infty} d_n = d_1 - c$, a constant.

1.29 a.

Unordered	Ordered
{4,4,12,12}	(4,4,12,12), (4,12,12,4), (4,12,4,12)
	(12,4,12,4), (12,4,4,12), (12,12,4,4)

Unordered	Ordered
	(2,9,9,12), (2,9,12,9), (2,12,9,9), (9,2,9,12)
{2,9,9,12}	(9,2,12,9), (9,9,2,12), (9,9,12,2), (9,12,2,9)
	(9,12,9,2), (12,2,9,9), (12,9,2,9), (12,9,9,2)

b. Same as (a).

c. There are 6^6 ordered samples with replacement from $\{1,2,7,8,14,20\}$. The number of ordered samples that would result in $\{2,7,7,8,14,14\}$ is $\frac{6!}{2!2!1!1!} = 180$ (See Example 1.2.20). Thus the probability is $\frac{180}{6^6}$.

d. If the k objects were distinguishable then there would be $k!$ possible ordered arrangements. Since we have $k_1,\ldots,k_m$ different groups of indistinguishable objects, once the positions of the objects are fixed in the ordered arrangement permutations within objects of the same group won't change the ordered arrangement. There are $k_1!k_2!\cdots k_m!$ of such permutations for each ordered component. Thus there would be $\frac{k!}{k_1!k_2!\cdots k_m!}$ different ordered components.

e. Think of the m distinct numbers as m bins. Selecting a sample of size k, with replacement, is the same as putting k balls in the m bins. This is $\binom{k+m-1}{k}$, which is the number of distinct bootstrap samples. Note that, to create all of the bootstrap samples, we do not need to know what the original sample was. We only need to know the sample size and the distinct values.

1.31 a. The number of ordered samples drawn with replacement from the set $\{x_1,\ldots,x_n\}$ is n^n. The number of ordered samples that make up the unordered sample $\{x_1,\ldots,x_n\}$ is $n!$. Therefore the outcome with average $\frac{x_1+x_2+\cdots+x_n}{n}$ that is obtained by the unordered sample $\{x_1,\ldots,x_n\}$

has probability $\frac{n!}{n^n}$. Any other unordered outcome from $\{x_1, \dots, x_n\}$, distinct from the unordered sample $\{x_1, \dots, x_n\}$, will contain m different numbers repeated $k_1, \dots, k_m$ times where $k_1 + k_2 + \cdots + k_m = n$ with at least one of the k_i's satisfying $2 \le k_i \le n$. The probability of obtaining the corresponding average of such outcome is

$$\frac{n!}{k_1!k_2!\cdots k_m!n^n} < \frac{n!}{n^n}, \text{ since } k_1!k_2!\cdots k_m! > 1.$$

Therefore the outcome with average $\frac{x_1+x_2+\cdots+x_n}{n}$ is the most likely.

b. Stirling's approximation is that, as $n \to \infty$, $n! \approx \sqrt{2\pi}n^{n+(1/2)}e^{-n}$, and thus

$$\left(\frac{n!}{n^n}\right) \Big/ \left(\frac{\sqrt{2n\pi}}{e^n}\right) = \frac{n!e^n}{n^n\sqrt{2n\pi}} = \frac{\sqrt{2\pi}n^{n+(1/2)}e^{-n}e^n}{n^n\sqrt{2n\pi}} = 1.$$

c. Since we are drawing with replacement from the set $\{x_1, \dots, x_n\}$, the probability of choosing any x_i is $\frac{1}{n}$. Therefore the probability of obtaining an ordered sample of size n without x_i is $(1 - \frac{1}{n})^n$. To prove that $\lim_{n\to\infty}(1 - \frac{1}{n})^n = e^{-1}$, calculate the limit of the log. That is

$$\lim_{n\to\infty} n \log\left(1 - \frac{1}{n}\right) = \lim_{n\to\infty} \frac{\log\left(1 - \frac{1}{n}\right)}{1/n}.$$

L'Hôpital's rule shows that the limit is -1, establishing the result. See also Lemma 2.3.14.

1.32 This is most easily seen by doing each possibility. Let $P(i)$ = probability that the candidate hired on the ith trial is best. Then

$$P(1) = \frac{1}{N}, \quad P(2) = \frac{1}{N-1}, \quad \dots \quad , P(i) = \frac{1}{N-i+1}, \quad \dots \quad , P(N) = 1.$$

1.33 Using Bayes rule

$$P(M|CB) = \frac{P(CB|M)P(M)}{P(CB|M)P(M) + P(CB|F)P(F)} = \frac{.05 \times \frac{1}{2}}{.05 \times \frac{1}{2} + .0025 \times \frac{1}{2}} = .9524.$$

1.34 a.

$$\begin{aligned} &P(\text{Brown Hair}) \\ &= P(\text{Brown Hair}|\text{Litter 1})P(\text{Litter 1}) + P(\text{Brown Hair}|\text{Litter 2})P(\text{Litter 2}) \\ &= \left(\frac{2}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{3}{5}\right)\left(\frac{1}{2}\right) = \frac{19}{30}. \end{aligned}$$

b. Use Bayes Theorem

$$P(\text{Litter 1}|\text{Brown Hair}) = \frac{P(BH|L1)P(L1)}{P(BH|L1)P(L1) + P(BH|L2)P(L2} = \frac{\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)}{\frac{19}{30}} = \frac{10}{19}.$$

1.35 Clearly $P(\cdot|B) \ge 0$, and $P(S|B) = 1$. If $A_1, A_2, \dots$ are disjoint, then

$$\begin{aligned} P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) &= \frac{P\left(\bigcup_{i=1}^{\infty} A_i \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap B)\right)}{P(B)} \\ &= \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B). \end{aligned}$$

1.37 a. Using the same events A, B, C and $\mathcal{W}$ as in Example 1.3.4, we have

$$\begin{aligned} P(\mathcal{W}) &= P(\mathcal{W}|A)P(A) + P(\mathcal{W}|B)P(B) + P(\mathcal{W}|C)P(C) \\ &= \gamma\left(\frac{1}{3}\right) + 0\left(\frac{1}{3}\right) + 1\left(\frac{1}{3}\right) \quad = \quad \frac{\gamma+1}{3}. \end{aligned}$$

Thus, $P(A|\mathcal{W}) = \frac{P(A\cap\mathcal{W})}{P(\mathcal{W})} = \frac{\gamma/3}{(\gamma+1)/3} = \frac{\gamma}{\gamma+1}$ where,

$$\begin{cases} \frac{\gamma}{\gamma+1} = \frac{1}{3} & \text{if } \gamma = \frac{1}{2} \\ \frac{\gamma}{\gamma+1} < \frac{1}{3} & \text{if } \gamma < \frac{1}{2} \\ \frac{\gamma}{\gamma+1} > \frac{1}{3} & \text{if } \gamma > \frac{1}{2}. \end{cases}$$

b. By Exercise 1.35, $P(\cdot|\mathcal{W})$ is a probability function. A, B and C are a partition. So

$$P(A|\mathcal{W}) + P(B|\mathcal{W}) + P(C|\mathcal{W}) = 1.$$

But, $P(B|\mathcal{W}) = 0$. Thus, $P(A|\mathcal{W}) + P(C|\mathcal{W}) = 1$. Since $P(A|\mathcal{W}) = 1/3$, $P(C|\mathcal{W}) = 2/3$. (This could be calculated directly, as in Example 1.3.4.) So if A can swap fates with C, his chance of survival becomes $2/3$.

1.38 a. $P(A) = P(A\cap B) + P(A\cap B^c)$ from Theorem 1.2.11a. But $(A\cap B^c) \subset B^c$ and $P(B^c) = 1 - P(B) = 0$. So $P(A\cap B^c) = 0$, and $P(A) = P(A\cap B)$. Thus,

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(A)}{1} = P(A)$$

.

b. $A \subset B$ implies $A\cap B = A$. Thus,

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1.$$

And also,

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(A)}{P(B)}.$$

c. If A and B are mutually exclusive, then $P(A\cup B) = P(A) + P(B)$ and $A\cap(A\cup B) = A$. Thus,

$$P(A|A\cup B) = \frac{P(A\cap(A\cup B))}{P(A\cup B)} = \frac{P(A)}{P(A)+P(B)}.$$

d. $P(A\cap B\cap C) = P(A\cap(B\cap C)) = P(A|B\cap C)P(B\cap C) = P(A|B\cap C)P(B|C)P(C)$.

1.39 a. Suppose A and B are mutually exclusive. Then $A\cap B = \emptyset$ and $P(A\cap B) = 0$. If A and B are independent, then $0 = P(A\cap B) = P(A)P(B)$. But this cannot be since $P(A) > 0$ and $P(B) > 0$. Thus A and B cannot be independent.

b. If A and B are independent and both have positive probability, then

$$0 < P(A)P(B) = P(A\cap B).$$

This implies $A\cap B \neq \emptyset$, that is, A and B are not mutually exclusive.

1.40 a. $P(A^c\cap B) = P(A^c|B)P(B) = [1 - P(A|B)]P(B) = [1 - P(A)]P(B) = P(A^c)P(B)$, where the third equality follows from the independence of A and B.

b. $P(A^c\cap B^c) = P(A^c) - P(A^c\cap B) = P(A^c) - P(A^c)P(B) = P(A^c)P(B^c)$.

1.41 a.

$$\begin{aligned}
&P(\text{ dash sent} \mid \text{dash rec}) \\
&= \frac{P(\text{ dash rec} \mid \text{dash sent})P(\text{ dash sent})}{P(\text{ dash rec} \mid \text{dash sent})P(\text{ dash sent}) + P(\text{ dash rec} \mid \text{dot sent})P(\text{ dot sent})} \\
&= \frac{(2/3)(4/7)}{(2/3)(4/7)+(1/4)(3/7)} \quad = \quad 32/41.
\end{aligned}$$

b. By a similar calculation as the one in (a) $P(\text{dot sent}|\text{dot rec}) = 27/434$. Then we have $P(\text{ dash sent}|\text{dot rec}) = \frac{16}{43}$. Given that dot-dot was received, the distribution of the four possibilities of what was sent are

Event	Probability
dash-dash	$(16/43)^2$
dash-dot	$(16/43)(27/43)$
dot-dash	$(27/43)(16/43)$
dot-dot	$(27/43)^2$

1.43 a. For Boole's Inequality,

$$P(\cup_{i=1}^n) \le \sum_{i=1}^n P(A_i) - P_2 + P_3 + \cdots \pm P_n \le \sum_{i=1}^n P(A_i)$$

since $P_i \ge P_j$ if $i \le j$ and therefore the terms $-P_{2k} + P_{2k+1} \le 0$ for $k = 1, \ldots, \frac{n-1}{2}$ when n is odd. When n is even the last term to consider is $-P_n \le 0$. For Bonferroni's Inequality apply the inclusion-exclusion identity to the A_i^c, and use the argument leading to (1.2.10).

b. We illustrate the proof that the P_i are increasing by showing that $P_2 \ge P_3$. The other arguments are similar. Write

$$\begin{aligned}
P_2 = \sum_{1\le i<j\le n} P(A_i \cap A_j) &= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} P(A_i \cap A_j) \\
&= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[\sum_{k=1}^{n} P(A_i \cap A_j \cap A_k) + P(A_i \cap A_j \cap (\cup_k A_k)^c)\right]
\end{aligned}$$

Now to get to P_3 we drop terms from this last expression. That is

$$\begin{aligned}
&\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[\sum_{k=1}^{n} P(A_i \cap A_j \cap A_k) + P(A_i \cap A_j \cap (\cup_k A_k)^c)\right] \\
&\ge \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[\sum_{k=1}^{n} P(A_i \cap A_j \cap A_k)\right] \\
&\ge \sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} P(A_i \cap A_j \cap A_k) \quad = \sum_{1\le i<j<k\le n} P(A_i \cap A_j \cap A_k) \quad = \quad P_3.
\end{aligned}$$

The sequence of bounds is improving because the bounds $P_1, P_1 - P_2 + P_3, P_1 - P_2 + P_3 - P_4 + P_5, \ldots$, are getting smaller since $P_i \ge P_j$ if $i \le j$ and therefore the terms $-P_{2k} + P_{2k+1} \le 0$. The lower bounds $P_1 - P_2, P_1 - P_2 + P_3 - P_4, P_1 - P_2 + P_3 - P_4 + P_5 - P_6, \ldots$, are getting bigger since $P_i \ge P_j$ if $i \le j$ and therefore the terms $P_{2k+1} - P_{2k} \ge 0$.

c. If all of the A_i are equal, all of the probabilities in the inclusion-exclusion identity are the same. Thus

$$P_1 = nP(A), \quad P_2 = \binom{n}{2}P(A), \quad \ldots \quad , P_j = \binom{n}{j}P(A),$$

and the sequence of upper bounds on $P(\cup_i A_i) = P(A)$ becomes

$$P_1 = nP(A), \quad P_1 - P_2 + P_3 = \left[n - \binom{n}{2} + \binom{n}{3}\right] P(A), \ldots$$

which eventually sum to one, so the last bound is exact. For the lower bounds we get

$$P_1 - P_2 = \left[n - \binom{n}{2}\right] P(A), \quad P_1 - P_2 + P_3 - P_4 = \left[n - \binom{n}{2} + \binom{n}{3} - \binom{n}{4}\right] P(A), \ldots$$

which start out negative, then become positive, with the last one equaling $P(A)$ (see Schwager 1984 for details).

1.44 $P(\text{at least 10 correct}|\text{guessing}) = \sum_{k=10}^{20} \binom{20}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} = .01386$.

1.45 $\mathcal{X}$ is finite. Therefore $\mathcal{B}$ is the set of all subsets of $\mathcal{X}$. We must verify each of the three properties in Definition 1.2.4. (1) If $A \in \mathcal{B}$ then $P_X(A) = P(\cup_{x_i \in A}\{s_j \in S : X(s_j) = x_i\}) \geq 0$ since P is a probability function. (2) $P_X(\mathcal{X}) = P(\cup_{i=1}^m \{s_j \in S : X(s_j) = x_i\}) = P(S) = 1$. (3) If $A_1, A_2, \ldots \in \mathcal{B}$ and pairwise disjoint then

$$\begin{aligned} P_X(\cup_{k=1}^\infty A_k) &= P(\bigcup_{k=1}^\infty \{\cup_{x_i \in A_k}\{s_j \in S : X(s_j) = x_i\}\}) \\ &= \sum_{k=1}^\infty P(\cup_{x_i \in A_k}\{s_j \in S : X(s_j) = x_i\}) \quad = \quad \sum_{k=1}^\infty P_X(A_k), \end{aligned}$$

where the second inequality follows from the fact the P is a probability function.

1.46 This is similar to Exercise 1.20. There are 7^7 equally likely sample points. The possible values of X_3 are 0, 1 and 2. Only the pattern 331 (3 balls in one cell, 3 balls in another cell and 1 ball in a third cell) yields $X_3 = 2$. The number of sample points with this pattern is $\binom{7}{2}\binom{7}{3}\binom{4}{3}5 = 14{,}700$. So $P(X_3 = 2) = 14{,}700/7^7 \approx .0178$. There are 4 patterns that yield $X_3 = 1$. The number of sample points that give each of these patterns is given below.

pattern	number of sample points	
34	$7\binom{7}{3}6$	$= 1{,}470$
322	$7\binom{7}{3}\binom{6}{2}\binom{4}{2}\binom{2}{2}$	$= 22{,}050$
3211	$7\binom{7}{3}6\binom{4}{2}\binom{5}{2}2!$	$= 176{,}400$
31111	$7\binom{7}{3}\binom{6}{4}4!$	$= 88{,}200$
		288,120

So $P(X_3 = 1) = 288{,}120/7^7 \approx .3498$. The number of sample points that yield $X_3 = 0$ is $7^7 - 288{,}120 - 14{,}700 = 520{,}723$, and $P(X_3 = 0) = 520{,}723/7^7 \approx .6322$.

1.47 All of the functions are continuous, hence right-continuous. Thus we only need to check the limit, and that they are nondecreasing

a. $\lim_{x\to-\infty} \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x) = \frac{1}{2} + \frac{1}{\pi}\left(\frac{-\pi}{2}\right) = 0$, $\lim_{x\to\infty} \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x) = \frac{1}{2} + \frac{1}{\pi}\left(\frac{\pi}{2}\right) = 1$, and $\frac{d}{dx}\left(\frac{1}{2}+\frac{1}{\pi}\tan^{-1}(x)\right) = \frac{1}{1+x^2} > 0$, so $F(x)$ is increasing.

b. See Example 1.5.5.

c. $\lim_{x\to-\infty} e^{-e^{-x}} = 0$, $\lim_{x\to\infty} e^{-e^{-x}} = 1$, $\frac{d}{dx}e^{-e^{-x}} = e^{-x}e^{-e^{-x}} > 0$.

d. $\lim_{x\to-\infty}(1 - e^{-x}) = 0$, $\lim_{x\to\infty}(1 - e^{-x}) = 1$, $\frac{d}{dx}(1 - e^{-x}) = e^{-x} > 0$.

e. $\lim_{y\to-\infty}\frac{1-\epsilon}{1+e^{-y}}=0$, $\lim_{y\to\infty}\epsilon+\frac{1-\epsilon}{1+e^{-y}}=1$, $\frac{d}{dx}(\frac{1-\epsilon}{1+e^{-y}})=\frac{(1-\epsilon)e^{-y}}{(1+e^{-y})^2}>0$ and $\frac{d}{dx}(\epsilon+\frac{1-\epsilon}{1+e^{-y}})>0$, $F_Y(y)$ is continuous except on $y=0$ where $\lim_{y\downarrow 0}(\epsilon+\frac{1-\epsilon}{1+e^{-y}})=F(0)$. Thus is $F_Y(y)$ right continuous.

1.48 If $F(\cdot)$ is a cdf, $F(x)=P(X\le x)$. Hence $\lim_{x\to\infty}P(X\le x)=0$ and $\lim_{x\to-\infty}P(X\le x)=1$. $F(x)$ is nondecreasing since the set $\{x: X\le x\}$ is nondecreasing in x. Lastly, as $x\downarrow x_0$, $P(X\le x)\to P(X\le x_0)$, so $F(\cdot)$ is right-continuous. (This is merely a consequence of defining $F(x)$ with "$\le$".)

1.49 For every t, $F_X(t)\le F_Y(t)$. Thus we have

$$P(X>t)=1-P(X\le t)=1-F_X(t)\ge 1-F_Y(t)=1-P(Y\le t)=P(Y>t).$$

And for some t^*, $F_X(t^*)<F_Y(t^*)$. Then we have that

$$P(X>t^*)=1-P(X\le t^*)=1-F_X(t^*)>1-F_Y(t^*)=1-P(Y\le t^*)=P(Y>t^*).$$

1.50 Proof by induction. For $n=2$

$$\sum_{k=1}^{2}t^{k-1}=1+t=\frac{1-t^2}{1-t}.$$

Assume true for n, this is $\sum_{k=1}^{n}t^{k-1}=\frac{1-t^n}{1-t}$. Then for $n+1$

$$\sum_{k=1}^{n+1}t^{k-1}=\sum_{k=1}^{n}t^{k-1}+t^n=\frac{1-t^n}{1-t}+t^n=\frac{1-t^n+t^n(1-t)}{1-t}=\frac{1-t^{n+1}}{1-t},$$

where the second inequality follows from the induction hypothesis.

1.51 This kind of random variable is called hypergeometric in Chapter 3. The probabilities are obtained by counting arguments, as follows.

x	$f_X(x)=P(X=x)$	
0	$\binom{5}{0}\binom{25}{4}\Big/\binom{30}{4}$	$\approx .4616$
1	$\binom{5}{1}\binom{25}{3}\Big/\binom{30}{4}$	$\approx .4196$
2	$\binom{5}{2}\binom{25}{2}\Big/\binom{30}{4}$	$\approx .1095$
3	$\binom{5}{3}\binom{25}{1}\Big/\binom{30}{4}$	$\approx .0091$
4	$\binom{5}{4}\binom{25}{0}\Big/\binom{30}{4}$	$\approx .0002$

The cdf is a step function with jumps at $x=0,1,2,3$ and 4.

1.52 The function $g(\cdot)$ is clearly positive. Also,

$$\int_{x_0}^{\infty}g(x)dx=\int_{x_0}^{\infty}\frac{f(x)}{1-F(x_0)}dx=\frac{1-F(x_0)}{1-F(x_0)}=1.$$

1.53 a. $\lim_{y\to-\infty}F_Y(y)=\lim_{y\to-\infty}0=0$ and $\lim_{y\to\infty}F_Y(y)=\lim_{y\to\infty}1-\frac{1}{y^2}=1$. For $y\le 1$, $F_Y(y)=0$ is constant. For $y>1$, $\frac{d}{dy}F_Y(y)=2/y^3>0$, so F_Y is increasing. Thus for all y, F_Y is nondecreasing. Therefore F_Y is a cdf.

b. The pdf is $f_Y(y)=\frac{d}{dy}F_Y(y)=\begin{cases}2/y^3 & \text{if } y>1\\ 0 & \text{if } y\le 1.\end{cases}$

c. $F_Z(z)=P(Z\le z)=P(10(Y-1)\le z)=P(Y\le(z/10)+1)=F_Y((z/10)+1)$. Thus,

$$F_Z(z)=\begin{cases}0 & \text{if } z\le 0\\ 1-\left(\frac{1}{[(z/10)+1]^2}\right) & \text{if } z>0.\end{cases}$$

1.54 a. $\int_0^{\pi/2} \sin x dx = 1$. Thus, $c = 1/1 = 1$.

b. $\int_{-\infty}^{\infty} e^{-|x|} dx = \int_{-\infty}^{0} e^{x} dx + \int_0^{\infty} e^{-x} dx = 1 + 1 = 2$. Thus, $c = 1/2$.

1.55

$$P(V \le 5) = P(T < 3) = \int_0^3 \frac{1}{1.5} e^{-t/1.5}\, dt = 1 - e^{-2}.$$

For $v \ge 6$,

$$P(V \le v) = P(2T \le v) = P\left(T \le \frac{v}{2}\right) = \int_0^{\frac{v}{2}} \frac{1}{1.5} e^{-t/1.5}\, dt = 1 - e^{-v/3}.$$

Therefore,

$$P(V \le v) = \begin{cases} 0 & -\infty < v < 0, \\ 1 - e^{-2} & 0 \le v < 6\ , \\ 1 - e^{-v/3} & 6 \le v \end{cases} \quad .$$