Solutions Manual For
Basic Probability Theory
by Robert B. Ash

Solutions to Problems
Chapter 1

Section 1.2

2. \( D_1 = AB + AC + BC \)
\( D_2 = ABC + AB^cC + A^cBC \)
\( D_3 = A + B + C \)
\( D_4 = AB^cC + A^cBC + A^cB^cC \)
\( D_5 = (ABC)^c = A^c + B^c + C^c \)
where \( AB = A \cap B, A + B = A \cup B \)

4. (a) \( x \in A \cap (B - C) \) iff \( x \in A \) and \( x \in B - C \)
   iff \( x \in A \cap B \) and \( x \notin A \cap C \)
   iff \( x \in (A \cap B) - (A \cap C) \)

(b) \( x \in A - (B \cup C) \) iff \( x \in A \) and \( x \notin B \cup C \)
   iff \( x \in A \) and \( x \notin B \) and \( x \notin C \)
   iff \( x \in A - B \) and \( x \notin C \)
   iff \( x \in (A - B) - C \)

It is true that \( (A \cup C) - B \subseteq (A - B) \cup C \). For if \( x \in (A \cup C) - B \) and \( x \notin C \) then \( x \in A - B \). But the sets need not be equal.
For example, if \( A = B = C \) then \((A \cup C) - B = A - A = \varnothing\), and \((A - B) \cup C = \varnothing \cup A = A\).

6. \( A^c \cap B^c = (A \cup B)^c \), which will not be empty unless \( A \cup B = \Omega \).
Thus \( A^c \) and \( B^c \) will be disjoint iff \( A \cup B = \Omega \). \((A \cap C) \cap (B \cap C) \)
\( C = A \cap B = \varnothing \), hence \( A \cap C \) and \( B \cap C \) are disjoint.
\( C = (A \cup C) \cap (B \cup C) \), so \( A \cup C \) and \( B \cup C \) are not disjoint if \( C \neq \varnothing \).
7. The probability that at least one is defective is 1 - the probability that none is defective, so 1 - p = 75/100.

8. This is an application of the formula P(A U B) = P(A) + P(B) - P(A \cap B), A = \{exactly 3 kings\}, B = \{exactly 3 aces\}. Thus

\[
P(A U B) = \binom{52}{3} \binom{4}{48} + \binom{4}{4} \binom{4}{48} - \binom{4}{4} \binom{4}{4}'
\]

10. (a) A sentence of length k must start with a word of length 1 or 2; there is only one possible word of length 1, but there are two possible words of length 2. If the first word is of length j, the remainder of the sentence may be completed in N(k-j) ways; the result follows.

(b) Assume N(k) = \lambda^k; this will be a solution provided \lambda^k = \lambda^{k-1} + 2 \lambda^{k-2}, i.e. \lambda = 2 or \lambda = -1.

Thus A^2 k + B(-1)^k is a solution. Also N(0) = A+B, N(1) = 2A-B, so A and B are determined by N(0) and N(1). Since N(0) and N(1) determine N(k) for all k, any two solutions that agree when k = 0 and 1 agree everywhere, so that A^2 k + B(-1)^k is the general solution. In the present case, A = 1, B = 1/3.

11. The total number of outcomes is 365^1; the number of favorable cases is 365(364) ... (365-r+1) = 365^1. Thus p = 365^1/365^1.

13. (a) Let A be a subset of \( \Omega = \{1,2, \ldots, n\} \). Either 1 \in A or 1 \notin A; this gives us two possibilities. In general, either k \in A or k \notin A, k = 1,2, \ldots, n. This gives us 2(2) \cdots 2^n ways of choosing A. Alternately, the number of subsets with exactly k members is the number of ways of selecting k distinct integers out of n, namely \( \binom{n}{k} \). The total number of subsets is \( \sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n \).

(b) The number of ways of selecting subsets A with exactly k members is \( \binom{n}{k} \). Having chosen such an A, we have B = A + A^c. Since there are \( 2^{n-k} \) subsets of A^c, B may be chosen in \( 2^{n-k} \) ways. The number of pairs of subsets is \( \sum_{k=0}^{n} \binom{n}{k} 2^{n-k} = (1+2)^n = 3^n \).
(a) Let $\Omega = \{1, 2, \ldots, n\}$. The integer 1 belongs to a set $A_1$ of
the partition, where $A_1$ contains $j$ other elements
($j = 0, 1, \ldots, n-1$). Thus $A_1$ can be chosen in $\binom{n-1}{j}$ ways.
Having chosen $A_1$, we must partition $A_1^c$; this can be done
in $g(n-1-j)$ ways. Thus

$$g(n) = \sum_{j=0}^{n-1} \binom{n-1}{j} g(n-1-j) = \sum_{j=0}^{n-1} \binom{n-1}{j} g(n-1-j)$$

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

(b) Let $h(n) = e^{-1} \sum_{k=0}^{n-1} \frac{n}{k!}$. Then

$$h(n) = \sum_{k=0}^{n-1} \binom{n-1}{k} e^{-1} \sum_{j=0}^{k} \frac{1}{j!} \left( \sum_{k=0}^{n-1} \frac{1}{k!} \frac{n}{k} \right)$$

$$= e^{-1} \sum_{j=0}^{n-1} \frac{1}{j!} \left( \sum_{k=0}^{n-1} \frac{1}{k!} \frac{n}{k} \right) = \frac{1}{n+1}.$$ 

Thus $g$ and $h$ satisfy the difference equation of (a), and
they agree when $n = 0$. By the form of the difference
equation, they agree everywhere.

1. (continued)

follows by an induction argument. But

$$P(A_1^c \cap A_2 \cap \ldots \cap A_n) = P(A_1^c A_2 \ldots A_n) = (1-P(A_1)^c)P(A_2)\ldots P(A_n)$$

by independence

$$= P(A_1^c)P(A_2)\ldots P(A_n).$$

2. 

$$\frac{p(k+1)}{p(k)} = \frac{\binom{n}{k+1} n-k-1}{\binom{n}{k} n-k} = \frac{n-k}{k+1}$$

Thus $p(k+1)/p(k)$ is $> 1$ iff $(n-k)p > (k+1)q$, i.e. iff

$$k < (n-1)p - 1$$

$$< 1$$

$$k \geq (n-1)p - 1$$

The result follows.

3. (a) Let $A = \{\text{spade is obtained}\}$, $B = \{\text{heart is obtained}\}$;

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

(b) Let $A = \{\text{spade is obtained}\}$, $B = \{\text{ace is obtained}\}$.

$$P(A \cap B) = 1/52, P(A) = 1/4, P(B) = 1/13.$$

(c) If $A$ and $B$ are independent and mutually exclusive, then

either $A$ or $B$ must have probability zero. For

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

by disjointness, and $P(A)P(B)$ by independence. Similarly, if the events $A_i$, $i \in I$, are independent and disjoint, either all or all but one of
the events must have probability zero. (If $P(A_i) \neq 0$, apply
the above argument to each $A_j$, $j \neq i$, to conclude that

$$P(A_j) = 0 \text{ for all } j \neq i.)$$

(d) Let $A = \{\text{spade}\}$, $B = \{\text{spade or heart}\}$; $P(A \cap B) = P(A) = 1/4$, $P(A)P(B) = 1/8.$

6. There are as many terms in (1.5.2) as there are unordered
samples of size $n$ out of $k$, with replacement, i.e. \binom{k+n-1}{n}
(see 1.4.4).
7. (a) For a favorable outcome, we must select \( n_i \) of the available \( t_i \) balls for color \( C_i, i = 1, 2, \ldots, k \). The total number of outcomes is the number of ways of selecting \( n \) distinct objects from a set of \( t \); the result follows.

(b) This is a standard multinomial problem. The probability is

\[
\frac{n!}{n_1!n_2! \cdots n_k!} \cdot \frac{t_1}{p_1} \cdots \frac{t_k}{p_k}
\]

where \( p_1 = t_1/t \).

8. (a) \( P(A \cap B) = P(A)P(A) \), hence \( P(A) = (P(A))^2 \), so that \( P(A) = 0 \) or 1.

(b) If \( P(A) = 0 \), then since \( A \cap B \) is a subset of \( A \), \( P(A \cap B) = 0 \) also. Thus \( P(A \cap B) = P(A)P(B) \). If \( P(A) = 1 \), then \( P(A^c) = 0 \), hence by the above argument, \( A^c \) and \( B \) are independent. But then \( A \) and \( B \) are independent (see Remark 1 or Problem 1 of Section 1.5).

Section 1.6

1. Let \( X \) be the number of successes. Then

\[
P[\text{all successes occur consecutively} | 4 \leq X \leq 6] = \frac{6}{\sum_{k=4}^{6} P[X=k]}
\]

\[
= \frac{6}{P[X=4] + P[X=5] + P[X=6]}
\]

\[
= \frac{(7p^4q^6 + 6p^5q^5 + 5p^6q^4)}{\sum_{k=4}^{6} \binom{10}{k} p^k q^{10-k}}
\]

2. \( P[X \geq 3 | X \geq 1] = P[X \geq 3, X \geq 1] / P[X \geq 1] = P[X \geq 3] / P[X \geq 1] \)

\[
= 1 - P[X = 0] - P[X = 1] - P[X = 2]
\]

\[
= 1 - q^n - npq^{n-1} - \binom{n}{2} p^2 q^{n-2}
\]

\[
= \frac{1 - q^n}{1 - q^n}
\]

5. We may regard this problem as one of dealing two 13 card hands to players 1 and 2 from a deck with 26 cards, of which 6 are spades. In each case, we are looking for the probability that (say) player 1 received a particular number of spades. Once the number of spades for player 1 is determined, that of player 2 is determined also. Thus,

\[
(a) \frac{6}{26} \binom{20}{10} / \binom{2}{1} = 0.36
\]

\[
(b) \frac{6}{26} \binom{20}{10} + \binom{6}{2} \binom{20}{2} / \binom{2}{1} = 0.48
\]

\[
(c) \binom{6}{2} \binom{20}{10} / \binom{2}{1} = 0.15
\]

\[
(d) \binom{20}{2} / \binom{2}{1} = 0.01.
\]

6. Let \( A = \{ \text{first two balls white} \}, B = \{ \text{six white balls in the sample} \} \). If the sampling is done with replacement, then

\[
P(A|B) = P(A \cap B) / P(B) = \frac{(2/3)^2 (8/9)^6 (1/3)^4}{(10/9)^6 (2/3)^6 (1/3)^4}
\]

If the sampling is done without replacement, \( P(A|B) \) is the number of ways of selecting \( 4 \) positions out of \( 6 \) for the white balls (the first 2 positions must be occupied by white balls), divided by the number of ways of selecting \( 6 \) positions out of 10; i.e. \( \frac{4}{10} \). Note that the answer is the same with replacement as without replacement. Once it is specified that 6 white and 4 black balls are obtained, the problem is simply one of counting arrangements.

8. (a) The probability is \( P(AB + CD + AED + CEB) \) where \( A \) is the event that the switch labeled 'A' is closed, etc. and + stands for union, product for intersection. Using the expansion formula (1.4.5) for the union of \( n \) events, we obtain (writing \( ab \) for \( P(AB) \), etc.)

\[
ab + cd + ade + ceb - abcd - abed - abce - cdea - cdeb
\]

- abed + 4abcde - abce = 2p^2 + 2p^3 - 5p^4 + 2p^5.
3. (continued)

(b) \( P\{E \text{ open and signal received} \} = P\{e^c(AB + CD)\} \)

\[ = P\{ABE^c\} + P\{CDE^c\} + P\{ABEDE^c\} = 2p^2q - 4q, q = 1-p. \]

Thus

\[ P\{E \text{ open | signal received} \} = \frac{(2p^2 - 4q)a}{2p^2 + 2p^3 - 5p^4 + 2p^5} \]

Chapter 2

Section 2.2

2. \( \{w: a \leq R(w) < b\} = \{w: R(w) < b\} - \{w: R(w) < a\} \in \frac{1}{n}, \) hence

\[ \{w: a \leq R(w) \leq b\} = \bigcap_{n=1}^{\infty} \{w: a \leq R(w) < b + \frac{1}{n}\} \in \frac{1}{n} \] for all real \( a, b. \)

3. \( \{w: R_1(w) + R_2(w) < b\} = \bigcup_{r, s \text{ rational}} \{w: R_1(w) < r, R_2(w) < s\} \in \frac{1}{r+s} < b \)

hence \( R_1 + R_2 \) is a random variable.

\( \{w: aR(w) < b\} = \{w: R(w) < \frac{b}{a}\} \) if \( a > 0 \)

\( = \{w: R(w) > \frac{b}{a}\} \) if \( a < 0 \)

\( = \emptyset \) or \( \Omega \) if \( a = 0. \)

In any case, \( \{w: aR(w) < b\} \in \frac{1}{r}, \) so \( aR \) is a random variable.

\( \{w: \sqrt{R(w)} < b\} = \{w: R(w) < b^2\}, \) hence \( \sqrt{R} \) is a random variable.

Section 2.4

2. \( f_2(y) = f_1(-\lambda y) \left| \frac{d}{dy} (-\lambda y) \right| \)

\[ = \frac{1}{2y}, \quad e^{-\lambda y}, \quad e^{-\lambda y} < y < e \]

\[ = 0 \text{ elsewhere.} \]

3. \( f_2(y) = f_1 \left( \frac{1}{2} y \right) \left| \frac{d}{dy} \frac{1}{2} y \right| = 2y^{-2}, \quad 2 < y < 4 \)

\[ = f_1(\sqrt{y}) \left| \frac{d}{dy} \sqrt{y} \right| = \frac{1}{2} y^{-3/2}, \quad y > 4 \]

\[ = 0, \quad y < 2. \]

5. (a) is a special case of (b). To prove (b), let \( 0 < y < 1 \) and pick an \( x \) such that \( F_1(x) = y. \) Then \( P[R_1 \leq y] = P[R_1 \leq x] = F_1(x) = y, \) and the result follows.
1. We show that \( F(x) = \int_{-\infty}^{x} f(t) \, dt \) for all \( x \). Pick any \( x \), and let 
\[ x_1, \ldots, x_n \]
be the points of discontinuity of \( f \) (or points where \( F \) does not exist) which lie in the interval \((-\infty, x]\). Then 
\[
\int_{-\infty}^{x} f(t) \, dt = \int_{-\infty}^{x_1} f(t) \, dt + \int_{x_1}^{x_2} f(t) \, dt + \ldots + \int_{x_{n-1}}^{x_n} f(t) \, dt.
\]
Now if \( x_{i-1} < a < b < x_i \), \( f \) is continuous on \([a, b]\) and \( f = F' \) on \([a, b]\), so by the fundamental theorem of calculus, 
\[
\int_{a}^{b} f(t) \, dt = F(b) - F(a) \quad \text{for all } i.
\]
Thus 
\[
\int_{x_{i-1}}^{x_i} f(t) \, dt = F(x_i) - F(x_{i-1}).
\]
Similarly, 
\[
\int_{x_{i-1}}^{x_i} f(t) \, dt = \lim_{x \to x_{i-1}} F(x) - F(x_{i-1}).
\]
Thus 
\[
\int_{-\infty}^{x} f(t) \, dt = F(x_1) + F(x_2) - F(x_1) + \ldots + F(x_n) - F(x_{n-1}) + F(x) - F(x_n) = F(x).
\]

(a) \( R_2 = k \) \iff \( R_1 = i \) \ for some \( i = 0, 1, \ldots, 9 \) 
\iff \( 10 R_1 = i + k 10^{-1} \) \ for some \( i = 0, 1, \ldots, 9 \) 
\iff \( i + k 10^{-1} \leq 10 R_1 < i + (k+1)10^{-1} \) \ for some \( i = 0, 1, \ldots, 9 \).

(b) In this case \( f_1(y) = f(y^2) \frac{d}{dy} y^2 \) \ where \( f \) is the uniform density on \([0, 1]\); thus \( f_1(y) = 2y \). Therefore 
\[
P(R_2=k) = \frac{9}{2} \left[ (10^{-1} + 10^{-2} k + 10^{-2} k)^2 - (10^{-1} + 10^{-2} k)^2 \right]
\]
\[
= \frac{9}{2} \left[ 2(10^{-1} + 10^{-2} k)10^{-2} + \ldots + 10^{-4} \right]
\]
\[
= 10^{-4} \sum_{i=0}^{9} (20i + 2k + 1)
\]
\[
= 10^{-4} \left[ \frac{20(10)(9)}{2} + 10(2k+1) \right] = .091 + .002k.
\]

9. The equations of motion are 
\[
x = (v \cos \theta) t, \quad y = (v \sin \theta) t - \frac{1}{2} g t^2,
\]
g = acceleration of gravity. The projectile returns to earth when \( y = 0 \), i.e. at time \( t_0 = \frac{2v \sin \theta}{g} \). Thus 
\[
R = (v \cos \theta) t_0 = (v^2 \sin 2\theta)/g.
\]
Since \( 2\theta \) is uniformly distributed between 0 and \( \pi \), we obtain, as in Example 2 of Section 2.4, 
\[
f_R(y) = 2\frac{y}{v} \left[ 1 - \frac{y^2}{v^2} \right]^{-1/2}, \quad 0 < y < \frac{v^2}{g}.
\]

Section 2.5

2. \( P[a \leq R \leq b] = P[R < a] = P[R < b] = P[R < a] - P[R < b] = F(b) - F(a) \)
\[
P[a < R < b] = P[R < b] - P[R < a] = F(b) - F(a) \)
\[
P[a < R < b] = P[R < a] - P[R < b] = F(a) - F(b) \)

Section 2.6

1. \( P[a_1 < R_1 \leq b_1, a_2 < R_2 \leq b_2] = P[a_1 < R_1 < b_1, R_2 \leq b_2] \)
\[- P[a_1 < R_1 \leq b_1, R_2 < a_2] = P[R_1 < b_1, R_2 < a_2] \)
\[- P[R_1 < a_1, R_2 \leq b_2] = P[R_1 < a_1, R_2 < a_2] \)
\[- P[R_1 < a_1, R_2 < a_2] = F_{12}(b_1, b_2) - F_{12}(a_1, b_2) - F_{12}(b_1, a_2) \]
\[- F_{12}(a_1, a_2) \]

Since \( F_{12}(x,y) = \int_{0}^{x} \int_{0}^{y} f_{12}(u,v) \, dv \, du \), the result follows.
2. By an analysis similar to Problem 1, the desired probability is
\[ F(b_1, b_2, b_3) - F(a_1, b_2, b_3) - F(b_1, a_2, b_3) - F(b_1, b_2, a_3) +
F(a_1, a_2, b_3) + F(a_1, b_2, a_3) + F(b_1, a_2, a_3) - F(a_1, a_2, a_3). \]
In \( n \) dimensions,
\[ F(a_1 < x_1 \leq b_1, \ldots, a_n < x_n \leq b_n) \] is the difference operator:
\[ \Delta_{b_n}(x_1, \ldots, x_n) = F(x_1, \ldots, x_{n-1}, b_n) - F(x_1, \ldots, x_{n-1}, a_n). \]
This may be expressed as \( F_0 - F_1 + F_2 - F_3 + \ldots + (-1)^n F_n \),
where \( F_k \) is the sum of all \( \binom{n}{k} \) terms of the form \( F(c_1, \ldots, c_n) \),
such that \( c_k = a_k \) for exactly \( k \) integers \( \in [1, 2, \ldots, n] \), and
\( c_k = b_k \) for the remaining \( n-k \) integers.

3. By Problem 1, \( F(-1 < R_1 \leq 0, 0 < R_2 \leq 1) = F(0,1) - F(-1,1) =
F(0,0) + F(-1,0) = 1-1+0 = -1 < 0 \), a contradiction.

Section 2.7

4. \( F_{12\ldots n}(x_1, \ldots, x_n) = F(R_1 \leq x_1, \ldots, R_n \leq x_n) =
\prod_{i=1}^{n} F_i(x_i) \).

5. If \( R \) is degenerate at \( c \), and \( R_1 \) is an arbitrary random variable,
then \( R \) and \( R_1 \) are independent, since \( F(R \in B, R_1 \in B_1) =
P(R \in B_1) \) if \( c \in B_1 \), and 0 if \( c \notin B_1 \). In particular \( R \) and \( R_1 \)
are independent. Conversely, let \( R \) and \( R_1 \) be independent. Then
\( P[R \leq c] = P[R \leq c, R_1 \leq c] = F[R \leq c]F[R \leq c], \) i.e.
\( F_R(x) = [F_R(x)]^2 \) for all \( x \), hence \( F_R(x) = 0 \) or 1 for all \( x \).
If \( c \) is the smallest \( x \) such that \( F_R(x) = 1 \) then \( F_R(x) = 1, \)
\( x \geq c \); \( F_R(x) = 0, \) \( x < c \). Thus \( P[R=c] = 1. \)

6. Let \( g_1(x) = \sin x, g_2(x) = x \). If \( R \) and \( \sin R \) are independent,
so are \( g_1(R) \) and \( g_2(\sin R) \), i.e. \( \sin R \) and \( \sin R \) are independent,
hence by Problem 5, \( \sin R \) is degenerate. Conversely
if \( \sin R \) is degenerate, \( R \) and \( \sin R \) are independent by the
remains in Problem 5.

7. \( P[R \in B_1, \ldots, R_n \in B_n] =
\int_{B_1} \ldots \int_{B_n} f_{12\ldots n}(x_1, \ldots, x_n) \, dx_1 \ldots dx_n \)
\[ = \int_{B_1} f_1(x_1) \, dx_1 \ldots \int_{B_n} f_n(x_n) \, dx_n = P[R_1 \in B_1] \ldots P[R_n \in B_n]. \]

Section 2.8

5. The core is described by \( x^2 + y^2 \leq a^2, \) \( x^2 + y^2 + z^2 \leq 4a^2 \).
The volume of the core is, in cylindrical coordinates,
\[ \frac{2\pi}{3} a^3 (4a^2 - r^2)^{1/2} \int_0^{2\pi} \int_0^a 2 \int_0^r (r(4a^2 - r^2)^{1/2} \, dr \, dz \, d\theta \]
\[ = 4\pi \left[ \frac{1}{3} (4a^2 - r^2)^{3/2} \right]_0^a = 4\pi \left( \frac{8}{3} - \sqrt{3} \right) a^3. \]
The probability that the worm will not be eaten is
\[ \frac{4\pi \left( \frac{8}{3} - \sqrt{3} \right) a^3}{\text{volume of sphere of radius } 2a} = \frac{4\pi \left( \frac{8}{3} - \sqrt{3} \right) a^3}{4\pi \left( \frac{3}{2} a^3 \right)} = 1 - \frac{3}{8} \sqrt{3}. \]
Thus the probability that it will be eaten is \( \frac{3}{8} \sqrt{3}. \)

6. The volume of the region is
\[ \int \int_{x^2+y^2 \leq 4, x \geq 0, x \geq 0} dx \, dy \]
\[ = \int_0^{\pi/2} \int_0^2 \int_0^{2x} 3x \, dxdy = \frac{\pi}{2} \int_0^{\pi/2} \int_0^2 (3x \cos \theta) r \, dr \, d\theta = 16. \]
6. (continued)

The desired probability is \[ \frac{1}{16} \int_{x=0}^{\pi/2} \int_{r=0}^{e^{-y+z} \sin 2 \theta} 2 \theta \cdot \rho \, d\rho \, d\theta, \]

\[ \frac{1}{16} \int_{x=0}^{\pi/2} \int_{\theta=-\pi/2}^{\pi/2} (2 \cos \theta) r \, dr \, d\theta = \frac{2}{3}, \]

as would be expected intuitively since each vertical line from \( z = 0 \) to \( z = 3x \) has \( 2/3 \) of its length below the line \( z = 2x \).

7. \[
1 = \int_{x=0}^{\pi/2} \int_{\theta=-\pi/2}^{\pi/2} k^{x+y} \, dz \, dy = \int_{x=0}^{\pi/2} \int_{\theta=-\pi/2}^{\pi/2} k^{x+y} \, dz \, dy \]

\[ \frac{1}{2} \int_{x=0}^{\pi/2} \int_{\theta=-\pi/2}^{\pi/2} 9 \cos^2 \theta \, r \, dr \, d\theta = \frac{1}{27} \]

8. \[
\frac{1}{n!} \int_{x=0}^{b_1} \int_{x_1=0}^{b_2} \cdots \int_{x_{n-1}=0}^{b_n} \int_{x_n=0}^{b_n} f(x_1, x_2, \ldots, x_n) \, dx_1 \ldots dx_n \]

\[ = \frac{1}{n!L^n} [L-(n-1)d]^n \]

Thus \( P(\min |R_i - R_j| \geq d) = \frac{[L-(n-1)d]^n}{L} \) if \((n-1)d \leq L\)

\[ = 0 \) if \((n-1)d > L.\]
12. \( F \left( \mathbf{W} \in B \right) = F \left( \mathbf{Z} \in g^{-1}(B) \right) = \int_{g^{-1}(B)} f(y) \, dy \)

Let \( y = g(x), x = h(y) \) to obtain

\[ \int_{B} f(h(y)) |J_h(y)| \, dy, \]

and the result follows.

13. \( f_{12}(x, y) = \frac{1}{2\pi b^2} e^{-\left(\frac{x^2+y^2}{2b^2}\right)} \)

\[ x = r \cos \theta, \quad y = r \sin \theta, \quad \text{so} \quad J_h(r, \theta) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r. \]

Thus \( f_{12}^*(r, \theta) = \frac{1}{2\pi b^2} r e^{-r^2/2b^2}, \quad 0 < \theta < 2\pi, \quad r > 0. \)

Evaluate the individual densities of \( R_0 \) and \( \theta_0 \) by

\[ \int_0^\infty \int_0^{2\pi} \int_{-\infty}^\infty \int_{-\infty}^\infty f_{12}^*(r, \theta) \, dr \, d\theta = \]

\( f_{R_0}(r) = \frac{1}{b^2} \int_{-\infty}^\infty \int_{-\infty}^\infty f_{12}^*(r, \theta) \, dr \, d\theta = \frac{1}{b^2}, \quad 0 < r < 2b. \)

Therefore \( f_{12}^*(r, \theta) = f_{R_0}^*(r) f_{\theta_0}(\theta), \) proving independence.

14. \( f_{34}(z, w) = f_{12}(x, y) \begin{vmatrix} \frac{\partial (x, y)}{\partial (z, w)} \end{vmatrix} \)

where \( z = xy, \quad w = y, \) i.e.

\[ x = \frac{z}{w}, \quad y = w. \]

The Jacobian is

\[ \begin{vmatrix} \frac{1}{w} & -\frac{z}{w^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{w}. \]

Thus \( f_{34}(z, w) = \frac{1}{w} f_{12}(\frac{z}{w}) f_{2}(w), \quad z, w > 0. \)

Hence

\[ f_3(z) = \int_0^\infty f_{34}(z, w) \, dw = \frac{1}{w} \int_0^\infty f_{12}(\frac{z}{w}) f_{2}(w) \, dw. \]

Note: The equations \( z = xy, \quad w = y \) define a one to one mapping of \( f(x, y) \) onto \( f(z, w). \)

15. \( R = R_1 + R_2/(1+R_2). \) The density of \( R_2/(1+R_2) \) is

\[ f(y) = f_2(y/(1-y)) \frac{dy}{y(1-y)} = 1/(1-y)^2, \quad 0 \leq y \leq \frac{1}{2}, \]

hence \( R_1 \) and \( R_2/(1+R_2) \) have joint density \( f(x, y) = 1/(1-y)^2, \)

\( 0 \leq x \leq 1, \quad 0 \leq y \leq \frac{1}{2}. \) Thus

\[ P(R \leq \frac{1}{2}) = \int_0^{\frac{1}{2}} \int_0^{1-x} (1-y)^{-2} \, dy \, dx = \frac{1}{2} + \ln 2. \]

16. The speed of the particle is \( (R_1^2 + R_2^2)^{1/2} \), hence

\( T = (R_1^2 + R_2^2)^{-1/2}. \) Thus

\[ P(T \leq t) = P(R_1^2 + R_2^2 \geq 1/t^2) = \]

\[ = (2\pi)^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\left(\frac{x^2+y^2}{2}\right)} \, dx \, dy \]

\[ = (2\pi)^{-1} \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} \, dr \, d\theta = e^{-1/t^2}, \quad t > 0. \]

Thus \( f_T(t) = 2t^{-3} e^{-1/t^2}, \quad t > 0; \quad f_{T}(t) = 0, \quad t \leq 0. \)

Section 2.9

1. (a) \( \ln(1+x) = x - x^2/2 + x^3/3 - x^4/4 + \ldots \)

\[ = x + x^2[-1 + \frac{x}{3} - \frac{x^2}{4} + \ldots] \]

If \( |x| \leq \frac{1}{2}, |1 - \frac{x}{3} + \frac{x^2}{4} + \ldots| \leq \frac{1}{2} + \frac{1}{2(2)} + \frac{1}{2(2^2)} + \ldots = 1, \)

and the result follows.
1. (continued)

(b) \( N(n(1 - \frac{x_n}{n})^n = n \ln(1 - \frac{x_n}{n}) = n(-\frac{x_n}{n} + \frac{x_n^2}{n^2}) = -\lambda \).

Thus \( 1 - \frac{x_n}{n} = e^{-\lambda} \).

2. (a) \( P(R \geq 1) = 1 - P(R = 0) = 1 - e^{-\lambda} \geq .99 \), so \( e^{-\lambda} \leq .01 \),
or \( \lambda \leq .001n \geq -\ln .01 = \ln 100 = 4.6 \). Thus \( n \geq 4600 \).

(b) \( P(R < 3) = e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}) = 5e^{-2} \), hence \( P(R \geq 3) = 1 - 5e^{-2} \).

3. (a) \( P(R_1 = 1) = P(R_1 = 1, R_2 = 1) + P(R_1 = 1, R_2 = 2) \)

\[ = .4 + .3 = .7 \]

\( P(R_2 = 1) = P(R_1 = 1, R_2 = 1) + P(R_1 = 2, R_2 = 1) \)

\[ = .4 + .2 = .6 \]

\( P(R_1 = 1, R_2 = 1) = .4 \neq P(R_1 = 1)P(R_2 = 1) \), hence \( R_1 \)

and \( R_2 \) are not independent.

(b) \( P(R_1R_2 \leq 2) = 1 - p_{12}(2,2) = .9 \).

Section 3.2

1. \( E(R^n) = \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx = 0 \) if \( n \) is odd, by symmetry.

If \( n \) is even, \( E(R^n) = \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} x^n e^{-x^2/2} \, dx = (y = \frac{1}{2} x^2) \)

\[ \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} (2y)^{n/2} e^{-y(2y)^{-1/2}} \, dy = \]

\[ = \frac{2}{\sqrt{2\pi}} \frac{2^{(n-1)/2}}{\Gamma(n/2)} y(y^{-1/2})^{n/2} \frac{y^{-n/2}}{\Gamma(n/2)} \]

\[ = \frac{2}{\sqrt{2\pi}} \frac{2^{n/2}}{\Gamma(n/2)} \Gamma(n/2) \]

\[ = (n-1)(n-3)...(3)(3)(1) \).

3. (a) \( E(R_1R_2) = \int_{0}^{\infty} \int_{0}^{x} xye^{-x}e^{-y} \, dx \, dy = \int_{0}^{\infty} e^{-x} \, dx = 1 \).

(b) \( E(R_1 - R_2) = \int_{0}^{\infty} \int_{0}^{x} (x-y)e^{-x}e^{-y} \, dx \, dy = 1 - 1 = 0 \).

(c) \( E|R_1 - R_2| = \int_{0}^{\infty} \int_{0}^{x} |x-y|e^{-x}e^{-y} \, dx \, dy \)

\[ = \int_{\Lambda} (x-y)e^{-x}e^{-y} \, dx \, dy + \int_{B} (y-x)e^{-y}e^{-x} \, dx \, dy \]

where \( \Lambda: x,y \geq 0, x \geq y \), and \( B: x,y \geq 0, x < y \)

\[ = (by \ symmetry) \int_{0}^{\infty} e^{-x} \left[ \int_{0}^{x} (x-y)e^{-y} \, dy \right] \, dx \]

\[ = 2 \int_{0}^{\infty} e^{-x} \left[ x(1-e^{-x}) + xe^{-x} + e^{-x} - 1 \right] \, dx \]

\[ = 2 \int_{0}^{\infty} [xe^{-x} + e^{-2x} - e^{-x}] \, dx = 2 \left( 1 + \frac{1}{2} - 1 \right) = 1 \).
4. \[ E[\max(R_1,R_2)] = \frac{1}{2} \int_{-1}^{1} x \, dx + \frac{1}{2} \int_{1}^{\infty} x \, dx = \frac{1}{2} \int_{1}^{\infty} x \, dx = \frac{1}{3}. \]

Alternatively, \( F_1(x) = F_x(x) = \frac{1}{2} (x+1), \ -1 \leq x \leq 1. \)

Hence if \( R_3 = \max(R_1,R_2), \ F_3(x) = F_1(x)F_2(x) = \frac{1}{4} (x+1)^2, \)
\(-1 \leq x \leq 1. \) Thus \( f_3(x) = \frac{1}{2} (x+1), \ -1 \leq x \leq 1. \) Consequently
\[ E(R_3) = \int_{-1}^{1} x \, f_3(x) \, dx = \frac{1}{2} \int_{1}^{\infty} x \, dx = \frac{1}{3}. \]

5. \[ E[C(R)] = \int_{0}^{\infty} 2xe^{-x^2} \, dx + \int_{0}^{\infty} \left[ 2 + 6(x-3) \right] xe^{-x^2} \, dx \]
\[ = \int_{0}^{\infty} 2xe^{-x^2} \, dx + 6 \int_{0}^{\infty} (x-3)xe^{-x^2} \, dx \]
\[ = 2 + 6e^{-3} (2 + 3) + 2 = 2 + 30e^{-3} \approx 3.5. \]

6. (a) \[ P[\text{at least one fails}] = 1 - P[\text{neither fails}] = \]
\[ 1 - P[R_1 > T, R_2 > T] = 1 - \left( \int_{0}^{\infty} e^{-\lambda x} \, dx \right)^2 = 1 - e^{-2\lambda T}. \]

(b) If \( R \) is the "down time" then
\[ R = T - \max(R_1,R_2) \] if \( R_1 \leq T \) and \( R_2 \leq T \)
\[ = 0 \] if either \( R_1 > T \) or \( R_2 > T \)
\[ E(R) = \int_{0}^{T} \int_{0}^{T} \max(x,y) \lambda e^{-\lambda x} \lambda e^{-\lambda y} \, dx \, dy \]
\[ = (\text{by symmetry}) 2 \int_{0}^{T} \int_{0}^{T} \max(x,y) \lambda e^{-\lambda x} \lambda e^{-\lambda y} \, dx \, dy \]
\[ = 2\lambda \int_{0}^{T} (T-x) e^{-\lambda x} (1 - e^{-\lambda x}) \, dx. \]

7. \[ E(R) = np \sum_{k=1}^{n} \frac{(n-1)!}{k!(n-k)!} p^{k-1} (1-p)^{n-k} \]
\[ = np \sum_{k=1}^{n} (n-1)^n p^{k-1} (1-p)^{n-k} = np \sum_{r=0}^{n-1} (n-1)^n p^{r} (1-p)^{n-1-r} \]
\[ = np(p + 1-p)^{n-1} = np. \]

Section 3.3

2. \[ E[(2-m)^n] = \int_{-m}^{\infty} (x-m)^n \frac{1}{\sqrt{2\pi}} e^{-(x-m)^2/2} \, dx. \]

Let \( y = \frac{x-m}{\sqrt{2}} \) to obtain \( \sigma^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \, dy, \) which is \( \sigma^2 \) times the \( n \text{th} \) moment of a random variable that is normal with mean 0 and variance 1. Thus, by Problem 1, Section 3.2, \[ E[(R-m)^n] = 0, \ n \text{ odd} \]
\[ = \sigma^n (n-1)(n-3) \ldots (5)(3)(1), \ n \text{ even}. \]

3. \[ E(R_1 R_2) = E(\cos 2\sin 2) = \int_{0}^{2\pi} \cos \sin \, dx \]
\[ = \frac{1}{2\pi} \int_{0}^{2\pi} \sin 2x \, dx = 0. \]

Since \( E(R_1) = \int_{0}^{2\pi} \cos \, dx = 0, E(R_2) = \int_{0}^{2\pi} \sin \, dx = 0 \)
\[ E(R_1^2 + R_2^2) = E(R_1^2) + E(R_2^2) + 2E(R_1 R_2) = E(R_1^2) + E(R_2^2) \]
and since \( E(R_1) = E(R_2) = E(R_1 + R_2) = 0, \)
\[ \text{Var}(R_1 + R_2) = \text{Var} R_1 + \text{Var} R_2. \]
Since \( P[R_1^2 \leq 1/4, R_2^2 \leq 1/4] = 0 \) \( \neq P[R_1^2 \leq 1/4]P[R_2^2 \leq 1/4], \) \( R_1 \)
and \( R_2 \) are not independent.

4. \[ |R| \leq R \leq |R|, \text{ so by properties 2 and 3,} \]
\[ -E(|R|) \leq E(R) \leq E(|R|), \text{ i.e. } |E(R)| \leq E(|R|). \]

5. \[ R^n = (R-m + m)^n = \sum_{k=0}^{n} \binom{n}{k} (R-m)^k m^{n-k}. \]

Thus \( a_n = E(R^n) = \sum_{k=0}^{n} \binom{n}{k} m^{n-k} \beta_k, \) assuming \( \beta_1, \ldots, \beta_{n-1} \) are
finite and \( \beta_n \) exists. From this result and properties 8 and 9
we conclude that \( a_n \) is finite if \( \beta_n \) is finite.
Section 3.4

2. \( a(E(R_1) + bE(R_2) = c \), hence \( a(E(R_1) + bE(R_2)) = 0 \), and the result follows.

3. Let \( g(x) = \text{E}[(xR_1 + R_2)^2] = \text{E}(R_1)^2x^2 + 2\text{E}(R_1R_2)x + \text{E}(R_2)^2 \). Assume \( R_1 \) not essentially 0; otherwise the result is immediate. Now equality holds in the Schwarz inequality iff the discriminant of \( g \) is 0, i.e., iff the equation \( g(x) = 0 \) has a real repeated root. But this happens iff \( g(x) = 0 \) for some \( x \), i.e., iff for some \( x \) we have \( xR_1 + R_2 = 0 \) (with probability 1).

Therefore, equality holds iff \( R_1 \) and \( R_2 \) are linearly dependent.

Section 3.5

1. \( I_{A_1}, \ldots, I_{A_n} \) are independent iff

\[
P(I_{A_1} = 1, \ldots, I_{A_n} = 1) = P(I_{A_1} = 1) \cdots P(I_{A_n} = 1)
\]

for all \( i_1, \ldots, i_n = 0 \) or 1, i.e., iff

\[
P(I_{A_1} = 1, \ldots, I_{A_n} = 1) = P(B_1 \cap B_2 \cap \cdots \cap B_n) = P(B_1)P(B_2) \cdots P(B_n)
\]

where for each \( k, B_k \) is either \( A_k \) or \( A_k^c \). This is equivalent to the independence of \( A_1, \ldots, A_n \) (see Problem 1, Section 1.5).

2. (continued)

(c) \( I \cap (w) = 1 \iff \omega \in \text{exactly one} \ A_1 \) (by disjointness)

\[
U_{i=1}^{n} A_i
\]

iff \( \sum_{i=1}^{n} I_{A_i}(w) = 1 \)

(d) Let \( A_n \) expand to \( A \). If \( \omega \in A \) then eventually \( \omega \in A_n \), hence \( I_{A_n}(w) \) is eventually 1, so \( I_{A_n}(w) \to \text{P}(A_n) \). If \( \omega \not\in A \) then \( I_{A_n}(w) \equiv 0 \), hence \( I_{A_n}(w) \to \text{I}(A(w)) \). The contracting case is handled similarly.

4. Let \( A_i = \{ \text{trial } i \text{ results in success and trial } i+1 \text{ in failure} \} \).

(a) \( I_A(\omega) = 1 \) since all points \( \omega \) belong to \( A \).

\( I_A(\omega) = 0 \) since all points \( \omega \) belong to \( \emptyset \).

(b) \( I_A \cap B(\omega) = 1 \iff \omega \in A \cap B \)

\( \iff I_A(\omega) = I_B(\omega) = 1 \)

\( \iff I_A(\omega)I_B(\omega) = 1 \)

\( I_A \cup B(\omega) = 1 \iff \omega \in A \cup B \)

\( \iff I_A(\omega) = 1 \) or \( I_B(\omega) = 1 \)

\( \iff I_A(\omega) + I_B(\omega) - I_A \cap B(\omega) = 1 \)

Section 3.6

1. (a) \( P(-.5 \leq R \leq 4) = P\left[\frac{-5-1}{3} \leq R \leq \frac{4-1}{3}\right] = F^*(1) - F^*(-.5) = F^*(1) - 1 + F^*(.5) \)

(5) From the table, \( .241 = 1 + .691 = .932. \)

(b) \( P(R \geq c) = P(R \geq \frac{c-1}{3}) = 1 - F^*(\frac{c-1}{3}) = 1 - \frac{1-c}{3} = .9 \)

From the table, \( c = 1.28, \) or \( c = 2.84. \)
2. \( P[|R-m| \geq k\sigma] = P[R^* \geq k] + P[R^* \leq -k] = F^*(k) + 1 - F^*(-k) = 2(1 - F^*(k)) \), which does not depend on \( m \) or \( \sigma \). From the table, \( F^*(1.96) = .975 \), hence
\[
P[|R-m| \geq 1.96 \sigma] = 2(.025) = .05.
\]

Section 3.7

2. (a) \( P[R_n \neq 0] \to 0 \) as \( n \to 0 \)

For \( P[R_n \neq 0] = P[R_n = e^n] = \frac{1}{n} \to 0 \).

(b) \( E(R_n^k) = \sum_{x=1}^k \frac{e^n}{n^x} P[R_n = e^n] = e^n P[R_n = e^n] \)
\[
= \frac{1}{n} e^n \to 0 \quad \text{as} \quad n \to 0
\]
for any \( k > 0 \).

3. Apply the weak law of large numbers with \( e = -\frac{m}{2} > 0 \). Then
\[
P\left[\frac{R_1 + \ldots + R_n}{n} \geq \frac{m}{2}\right] = P\left[\frac{R_1 + \ldots + R_n}{n} - m \geq -\frac{m}{2}\right] \leq P\left[\left|\frac{R_1 + \ldots + R_n}{n} - m\right| \geq \epsilon\right] \text{by (1.3.9)}
\]
\[
\to 0 \quad \text{as} \quad n \to \infty.
\]

If \( K \) is any negative number, \( \frac{nm}{2} < K \) for large \( n \), hence
\[
P[R_1 + \ldots + R_n < K] \geq P[R_1 + \ldots + R_n < \frac{nm}{2}] = 1.
\]

Thus for large \( n \), the probability that your total losses after \( n \) trials will exceed \( |K| \) is overwhelming.

**Moral:** Do not gamble (at least not if your average gain on a given trial is negative). The weak law of large numbers, known colloquially as the **Law of Averages**, predicts that you are very likely to be wiped out.

Chapter 4

Section 4.2

2. Restrict \( x \) and \( y \) to be \( \geq 0 \) throughout. Then
\[
C = \{(x,y): xy \leq 2\}, \quad C_x = \{y: y \leq 2-x, \quad 0 \leq x \leq 2\};
\]
\[
C_x = \emptyset, \quad x > 2
\]
\[
P_x(C_x) = 1, \quad 0 \leq x \leq 1
\]
\[
= \frac{2-x}{x}, \quad 1 \leq x \leq 2
\]
\[
= 0, \quad x > 2.
\]

3. By (4.2.3), \( P[4 \leq R_1^* + R_2^* \leq 6] = \)
\[
\sum_{n=1}^\infty P_n \int_0^6 x e^{-ny} dy = \sum_{n=1}^\infty p_n (e^{-3} - e^{-5})
\]
\[
+ p_3 (e^{-4} - e^{-6}) + p_4 (1 - e^{-3}) + p_5 (1 - e^{-5})
\]

4. \( P[R_2 \in B|R_1 = x_1] = \frac{P[R_1 = x_1^*, R_2 \in B]}{P[R_1 = x_1]} \)
\[
= \frac{p(x_1) \int_B f_1(y) dy}{p(x_1)} \quad \text{by (4.2.2)}.
\]

Thus \( P[R_2 \in B|R_1 = x_1] = \int_B f_1(y) dy = p(x_1) (B) \).

5. If \( 0 \leq y \leq 1 \), \( P[R_2 \leq y] = \int_0^y f_1(x) P[R_2 \leq y|R_1 = x] dx \)
\[
= \int_0^y \frac{1}{2} x^2 (y-x) dx = \frac{1}{2} y.
\]
Section 4.3

1. \( f_1(x) = \int_{y=x}^{\infty} f(x,y) \, dy = \int_{y=x}^{\infty} e^{-y} \, dy = e^{-x}, \ x \geq 0. \)

Thus \( h(y|x) = \frac{f(x,y)}{f_1(x)} = e^{x-y}, \ 0 \leq x \leq y \)

= 0 elsewhere.

Therefore \( P[R_2 \leq y | R_1 = x] = \int_{-\infty}^{y} h(y|x) \, dy = \int_{x}^{y} e^{-y} \, dy = 1-e^{-y}, \ y \geq x \)

= 0, y < x.

3. (a) is a special case of (b). In (b),

\[ f_1(x) = \int_{-\infty}^{\infty} f(x,y) \, dy = \frac{1}{\text{area } C_x} \int_{C_x} \text{length } C_x \, dy. \]

Thus \( h(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{1}{\text{length } C_x} \) if \( y \in C_x \), i.e. given \( R_1 = x \), \( R_2 \) is uniformly distributed on \( C_x \).

4. If \( R_1 = x \) then \( R_2 - R_1 \leq z \) iff \( R_2 \leq x+z \). Thus

\[ P[R_3 \leq z | R_1 = x] = P[R_2 \leq x+z | R_1 = x] \]

\[ = \int_{-\infty}^{\infty} h(y|x) \, dy = \int_{-\infty}^{x+z} e^{-y} \, dy = e^{x}(e^{-x} - e^{-(x+z)}) = 1-e^{-z}. \]

The conditional density of \( R_3 \) given \( R_1 = x \) is \( \frac{d}{dz} (1-e^{-z}) = e^{-z}, \ z \geq 0, x \geq 0. \)

\[ P[1 \leq R_3 \leq 2 | R_1 = x] = \int_{1}^{2} e^{-z} \, dz = e^{-1} - e^{-2}, \ x \geq 0. \]

Note that \( R_1 \) and \( R_3 \) are independent but \( R_1 \) and \( R_2 \) are not.

Section 4.4

2. \( E(R_0^{-n} | R_1 = x_1, \ldots, R_n = x_n) = \int_{-\infty}^{\infty} \lambda^{-n} h(\lambda | x_1, \ldots, x_n) \, d\lambda \)

\[ \quad = \frac{1+nx}{n^+1} \int_{0}^{\lambda^{-1}} e^{-\lambda(1+nx)} \, d\lambda = \frac{(1+nx)^n}{n^+1}, \ x = \sum_{1=1}^{n} x_i. \]

3. \( h(y|x) = 1/2, \ 0 \leq x \leq 1, \ 0 \leq y \leq x \)

= 1, 1 \leq x \leq 2, \ 0 \leq y \leq 1

= \frac{1}{3-x}, \ 2 \leq x \leq 3, \ 0 \leq y \leq 1.

Thus \( E(R_2 | R_1 = x) = \int_{-\infty}^{x} y \cdot h(y|x) \, dy \)

= \frac{1}{x} \int_{0}^{x} y \, dy = \frac{1}{2} x, \ 0 \leq x \leq 1

= \frac{1}{2}, \ 1 \leq x \leq 2

= \frac{1}{3-x} \int_{-\infty}^{x} y \, dy = \frac{1}{2} (\frac{x-2}{3-x}), \ 2 \leq x \leq 3.

Note that all computations may be avoided by making use of Problem 3b in Section 4.3.
4. Let $R_1$ be the number of ones, $R_2$ the number of twos. Given that $R_1 = k$, $R_2$ has the binomial distribution with parameters $n-k$ and 1/5 (see Example 1, Section 2.9). Thus $E(R_2 | R_1 = k) = \frac{1}{5} (n-k)$.

5. (a) $P[R_1 \geq 1 | R_1 + R_2 \leq 3] = \frac{P[R_1 \geq 1, R_1 + R_2 \leq 3]}{P[R_1 + R_2 \leq 3]} = \frac{3}{7} = \frac{3}{7} \cdot$

(b) $E(R_1 | R_1 + R_2 \leq 3) = \frac{E(R_1 \cap \{R_1 + R_2 \leq 3\})}{P[R_1 + R_2 \leq 3]}
= \frac{\frac{2}{7} \int_{x=0}^{2} \int_{y=0}^{2-x} f(x,y) \, dx \, dy}{\frac{2}{7} \int_{y=0}^{2} \int_{x=0}^{2-y} f(x,y) \, dx \, dy}
= \frac{\frac{1}{7} \int_{x=0}^{2} \int_{y=0}^{2-x} f(x,y) \, dx \, dy}{\frac{1}{7} \int_{y=0}^{2} \int_{x=0}^{2-y} f(x,y) \, dx \, dy}
= \frac{1}{7} \left( 1 + \frac{13}{6} \right) = \frac{19}{21} \cdot$

6. $\sum_{n=1}^{\infty} P(B_n) \cdot E(R(B_n)) = \sum_{n=1}^{\infty} \frac{P(B_n) \cdot E(R(B_n))}{P(B_n)}$

$= \sum_{n=1}^{\infty} E(R(B_n)) = E(\sum_{n=1}^{\infty} R(B_n)) = E(R)$ since $\sum_{n=1}^{\infty} R(B_n) = 1$.

(It can be shown that $E(\sum_{n=1}^{\infty} R(B_n)) = \sum_{n=1}^{\infty} E(R(B_n))$ if $E(R)$ exists.)

9. $P[T - t_0 \leq x | T > t_0] = \frac{P[t_0 < T \leq t_0 + x]}{P[T > t_0]} = \frac{e^{-t_0} - (t_0 + x)}{e^{-t_0}} = 1 - e^{-x} \cdot$

Thus the (conditional) waiting time starting from $t_0$ has the same density $e^{-x}, x \geq 0$ as the original waiting time $T$, i.e. the bulb "does not remember" that it has already burned for $t_0$ units of time.

11. $E(R_1^2 + R_2^2 | R_1 = x) = x^2 + E(R_2^2 | R_1 = x)$ (cf. Problem 5, Section 4.3). Now $E(R_2^2 | R_1 = x) = \int_0^2 y^2 h(y | x) \, dy - \int_0^1 y^2 \varepsilon_2(y) \, dy = E(R_2^2)$ by independence. Thus

$E(R_1^2 + R_2^2 | R_1 = x) = x^2 + \int_0^2 \frac{1}{2} y^2 \, dy + \int_0^1 \frac{1}{2} y e^{-y} \, dy$

$= x^2 + \frac{1}{6} + 1 - x^2 + \frac{7}{6} \cdot$

12. (a) $P[R_1 = x | y < R_2 < y + dy] = \frac{P(R_1 = x, y < R_2 < y + dy)}{P(y < R_2 < y + dy)}$

$= \frac{P(R_1 = x) \cdot P(y < R_2 < y + dy | R_1 = x)}{P(y < R_2 < y + dy)}$

$= \frac{\sum_{k=1}^{y} P(R_1 = k) \cdot P(y < R_2 < y + dy | R_1 = k)}{\sum_{k=1}^{y} P(R_1 = k) \cdot h(y | x)}$

(b) $P(R_1 \in A, R_2 \in B) = \sum_{x \in A} P(R_1 = x) \int_B h(y | x) \, dy$ by (4.2.2).

But $\int_B f_2(y) P(R_1 \in A | R_2 = y) \, dy =$

$\int_B \sum_{x \in A} P(R_1 = x) \cdot h(y | x) \cdot \frac{P(R_1 = x) \cdot h(y | x)}{P(R_1 = x) \cdot h(y | x)} \, dy =$

$\sum_{x \in A} \int_B P(R_1 = x) \cdot h(y | x) \, dy.$

13. $P(R_1 \in A, R_2 \in B) = \int_A f_1(x) \cdot P(R_2 \in B | R_1 = x) \, dx$

$= \int_A f_1(x) \cdot \sum_{y \in B} P(y | x) \, dx = \sum_{y \in B} \int_A f_1(x) \cdot P(y | x) \, dx = P_2(y)$

$= \sum_{y \in B} P(R_2 = y) \cdot P(R_1 \in A | R_2 = y), which is the appropriate y \in B version of the theorem of total probability.
14. \( P[R_1=x_1, \ldots, R_n=x_n | R=\lambda] = \lambda^n e^{-(\lambda+1)x} ) \), \( x = \sum_{i=1}^n x_i \), \( x_i = 0 \) or \( 1 \).

Thus \( P[R_1=x_1, \ldots, R_n=x_n] = \frac{1}{\beta} (1+\lambda)^{n-x} \lambda^n d\lambda = \beta(1+\lambda, n-x+1). $$

By Problem 13, the conditional density of \( R \) given \( R_1=x_1, \ldots, R_n=x_n \) is
\[
\frac{\lambda^x (1-\lambda)^{n-x}}{\beta(1+\lambda, n-x+1)}, \quad 0 \leq \lambda \leq 1.
\]

Hence \( E[R | R_1=x_1, \ldots, R_n=x_n] = \int_0^1 \frac{\lambda^x (1-\lambda)^{n-x}}{\beta(1+\lambda, n-x+1)} d\lambda \)
\[
= \frac{\beta(2+x, n-x+1)}{\beta(1+\lambda, n-x+1)} = \frac{\Gamma(2+x) \Gamma(n+2)}{\Gamma(1+x) \Gamma(n+3)} = \frac{x+1}{n+2}.
\]

15. (a) The probability of error is
\[
P(\text{heads}) P[R < S | \text{heads}] + P(\text{tails}) P[R > S | \text{tails}]
\]
\[
= p \int f_0(x) dx + (1-p) \int f_1(x) dx = p \int f_0(x) dx + \int f_1(x) dx + (1-p).
\]

(b) Let \( L(x) = \frac{f_1(x)}{f_0(x)} \). If \( L(x) > \frac{p}{1-p} \), the integrand is \( < 0 \), so to minimize the probability of error, we should put \( x \in S \). If \( L(x) < \frac{p}{1-p} \), the integrand is \( > 0 \), so take \( x \notin S \).

If \( L(x) = \frac{p}{1-p} \), do anything.

For the example, \( L(x) = \frac{e^{-x^2}}{e^{-x^2}} = \frac{e^{-(x-m_1)^2/2\sigma}}{e^{-(x-m_0)^2/2\sigma}} \).

\( L(x) > \frac{p}{1-p} \) iff \( \frac{(x-m_0)^2 - (x-m_1)^2}{2\sigma^2} > \ln \frac{p}{1-p} \), i.e.,
\[
x \in S \iff x > \frac{2}{m_1 - m_0} \ln \frac{p}{1-p} + \frac{m_1 + m_0}{2}, \text{ assuming } m_0 < m_1.
\]

16. \( E[R | R \geq 2] = E[R | R \geq 2] / P[R \geq 2] = \sum_{k=2}^n \frac{k \lambda e^{-\lambda}}{1-e^{-\lambda}} \)
\[
= (np - 1) e^{-\lambda} / (1-e^{-\lambda}) - \lambda e^{-\lambda} = np - npq / (1 - npq).
\]

17. \( E[R_2 | 2 \leq R_2 \leq 4] = E[R_2 | 2 \leq R_2 \leq 4] / P[2 \leq R_2 \leq 4] = \frac{E[R_1 ^2 (\sqrt{2} \leq R_1 \leq 2)]}{P[\sqrt{2} \leq R_1 \leq 2] + P[6 \leq R_1 \leq 10]} \)
\[
eq \frac{\int_2^4 2 \sqrt{2} dx + \int_6^{10} 3 \sqrt{2} dx}{\frac{1}{10} (2 - \sqrt{2} + 4) + 6 - \sqrt{2}}
\]
\[
= \sqrt{2} \left( \frac{12 + 1}{6} \right) = \frac{12 + 1}{6} \left( 8 - 2^{3/2} \right).
\]

Alternately, \( P[R_2 = 3] = 4/10 \), and after removing this discontinuity from the distribution function of \( R_2 \), we obtain
\[
f_2(y) = \frac{dy}{dy} = 1/20y^{1/2}, \quad 0 < y \leq 36.
\]

Thus
\[
E[R_2 | 2 \leq R_2 \leq 4] = \frac{3 \int_{R_2 = 3} f_2(y) \frac{1}{20y^{1/2}} dy}{2 + \frac{1}{20y^{1/2}}}
\]
\[
= \frac{12 + 1}{10} \left( 8 - 2^{3/2} \right).
\]

19. (a) If \( R \) is absolutely continuous,
\[
E[(\theta^* - \theta)^2] = \int_{-\infty}^{\infty} E[(\theta^* - \theta)^2 | R=x] f_R(x) dx.
\]

To minimize this, it is sufficient to minimize \( E[(\theta^* - \theta)^2 | R=x] \) for each \( x \). But since \( \theta^* = d(R) \), we have
\[
E[(\theta^* - \theta)^2 | R=x] = E[(d(x) - \theta)^2 | \lambda=x] = d^2(x)
\]
\[
- 2E(\lambda | R=x) d(x) + E(\lambda^2 | R=x).
\]
19. (a) (continued)

Since $y^2 - 2Ay + B$ is a minimum when $y = A$, we have

$$d(x) = E(\theta | R=x).$$

If $R$ is discrete,

$$E[(\theta^* - \theta)^2] = \sum_x E[(\theta^* - \theta)^2 | R=x] p_R(x),$$

and the same argument applies.

(b) Clearly $d(x) = 1$ if $1 < x \leq 3$, $d(x) = -1$ if $-3 \leq x < -1$.

If $-1 \leq x \leq 1$, $P(\theta = 1 | R=x) = P(\theta = 1) f_R(x | \theta = 1) / f_R(x)$ (see Problem 12). Given $\theta = 1$, $R$ is uniformly distributed between $-1$ and $3$, so $P(\theta = 1 | R=x) = (1/2)(1/4) / (1/2) f_R(x | \theta = 1) + 1/2 f_R(x | \theta = -1) = (1/8) / (1/8 + 1/8) = 1/2$. Thus

$P(\theta = 1 | R=x) = 1/2$ also, so that $d(x) = E(\theta | R=x) = 0$.

With probability $1/2$, $(\theta^* - \theta)^2 = 0$, and with probability $1/2$, $(\theta^* - \theta)^2 = 1$, hence the minimum value of

$$E[(\theta^* - \theta)^2]$$

is $1/2$.

20. The conditional density of $\theta$ given $R=x$ is (see Problem 13)

$$h_\theta(x | \lambda) = f_\theta(\lambda) f(R=x | \theta=\lambda) / P[R=x]$$

$$= e^{\lambda x - \lambda} \lambda^{x+1} / \int_0^\infty e^{\lambda x - \lambda} \lambda^x d\lambda$$

$$= 2^{x+1} \lambda^x e^{-2\lambda} / \lambda^x.$$

Thus $E(\theta | R=x) = \int_0^\infty \lambda h_\theta(\lambda | x) d\lambda = (x+1) \int_0^\infty \lambda^{x+1} e^{-2\lambda} d\lambda$

$$= (x+1) / 2^{x+1} = 1 / 2 (x+1).$$

Chapter 5

Section 5.2

2. $N_1(s) = N_2(s) = \frac{1}{3s} (e^s - 1) + \frac{2}{3s} (1 - e^{-s})$, all $s$.

$N_0(s) = N_1(s) N_2(s) = \frac{1}{9s^2} (e^{2s} + 2e^{s+2} - 4e^{-s} + 4e^{-2s})$

$N_1(s) = \frac{1}{3} \left[ (s+2) u(x+2) + 2(x+1) u(x+1) - 3x u(x) - 4(x-1) u(x-1) + 4(x-2) u(x-2) \right].$

5. $N_0(s) = \left( \frac{\lambda}{s+\lambda} \right)^n$, so $f_0(x) = \frac{\lambda^n e^{-\lambda x}}{(n-1)!} u(x)$.

6. If $R = \tan \theta$, $f_\theta(y) = f_0(\arctan y) | \frac{d}{dy} \arctan y | = \frac{1}{\pi (1+y^2)}$

(The same result is obtained if $\theta$ is uniformly distributed between 0 and $\pi$, or 0 and $\pi/2$.)

7. For $x \geq 0$, $f(x) = u(x) - xu(x) + (x-1) u(x-1)$. Thus the Laplace transform of $f(x) u(x)$ is $N_1(s) = \frac{1}{s} - \frac{1}{s^2} + \frac{1}{s^2} e^{-s}$ (all $s$).

The Laplace transform $N_2(s)$ of $f(x) u(-x)$ is that of $f(-x) u(x)$ ($= f(x) u(x)$) with $s$ replaced by $-s$ (see Property 3, Section i.e. $N_2(s) = -\frac{1}{s} - \frac{1}{s^2} + \frac{1}{s^2} e^s$ (all $s$). Thus

$$N_R(s) = N_1(s) + N_2(s) = \frac{1}{s} \left( e^s - e^{-s} \right) - \frac{2}{s^2}.$$

Let $s = iu$ to obtain $N_\chi(u) = -\frac{1}{2u} \left( 2 \cos u \right) + \frac{2}{u^2}$

$$= \frac{2(1 - \cos u)}{u^2}. $$
8. (a) By (5.2.1), \( f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M(u) e^{ixu} \, du \), or

\[
f(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M(x) \, e^{-ixu} \, dx.
\]

(We may replace \( e^{ixu} \) by \( e^{-ixu} \) since \( M \) and \( f \) are real valued.) Multiply both sides by \( k \) to obtain the desired result.

(b) \[\int_{-\infty}^{\infty} e^{-|x|} e^{-ixu} \, dx = \frac{2}{1+u^2} \] (see the discussion after (5.2.1)).

This is nonnegative and integrable. Thus \( ke^{-|u|} \) is a characteristic function; since \( e^{0} = 1 \), the appropriate \( k \) is 1. \( M(u) = 1 - |u|, M(u) = 0, |u| > 1, \) is a characteristic function by Problem 7.

10. \( M'(u) = \int_{-\infty}^{\infty} -x \sin ux \, e^{-x^2/2} \, \frac{dx}{\sqrt{2\pi}} \)

\[
= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sin ux \, d(e^{-x^2/2}) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} \, u \cos ux \, dx = -uM(u).
\]

The differential equation \( \frac{dy}{dx} = -xy \) may be written as

\[
\frac{1}{y} \frac{dy}{dx} = -x.
\]

Integrate to obtain \( \ln y = -\frac{x^2}{2} + C \), or

\( y = Ce^{-x^2/2} \). Thus \( M(u) = e^{-u^2/2} \) (note \( M(0) \) is always 1). If \( R \) is normal \((m, \sigma^2)\) then \( R = (R-m)/\sigma \) is normal \((0, 1)\), so

\( E(e^{-iuR}) = E(e^{-iu(m+\sigma^2)}) \)

\( = e^{-iun} M_{\sigma}(u) = e^{-iun} e^{-u^2\sigma^2/2} \).

Section 5.3

1. No. If so, then \( f(x) = 2e^{-x}u(x) - [u(x) - u(x-1)] \)

\[
= 2e^{-x} - 1, \ 0 < x \leq 1
\]

\[
= 2e^{-x}, \ x > 1
\]

\( = 0, \ x < 0. \)

This is negative for \( x \) near 1 and \( < 1 \), an impossibility.

2. \( N_R(s) = \int_{a}^{b} f(x) e^{-sx} \, dx \). For any particular \( s \), \( e^{-sx} \) has some largest value for \( x \in [a,b] \), say \( K \). Then

\[
|N_R(s)| \leq \int_{a}^{b} K|f(x)| \, dx < \infty.
\]

3. No. Let \( R \) be uniformly distributed between 0 and 1. Then

\[
N_R(s) = \frac{1-e^{-s}}{s} \), so \( M_R(u) = \frac{1-e^{-iu}}{iu} \). Now
\]

\[
|N_R(u)| = \frac{1}{|u|} \left| 1 - \cos u + i \sin u \right|
\]

\[
= \sqrt{2} \left( \frac{1-\cos u}{u} \right)^{1/2} = \frac{1}{|u|} \left| \sin \frac{u}{2} \right|^{1/2} \), and thus
\]

\[
\int \left| M_R(u) \right| \, du = \infty.
\]

Section 5.4

2. Let \( \Omega = \{w_1, w_2\}, P(w_1) = P(w_2) = \frac{1}{2} \).

Let \( R(w_1) = 1, R(w_2) = 0. \)

If \( n \) is even set \( R_n(w_1) = 1, R_n(w_2) = 0. \)

If \( n \) is odd set \( R_n(w_1) = 0, R_n(w_2) = 1. \)

Then \( P(R_n = 0) = P(R_n = 1) = 1/2 \) for all \( n \), and

\( P(R = 0) = P(R = 1) = 1/2. \).
2. (continued)

Thus \( F_n(x) = F(x) \) for all \( n \) and all \( x \), so \( R_n \overset{d}{\to} R \).

But if \( 0 < \varepsilon < 1 \), \( P[|R_n - R| \geq \varepsilon] = P[R_n \neq R] = 0 \) if \( n \) is even

\[ = 1 \text{ if } n \text{ is odd.} \]

Thus \( R_n \overset{p}{\to} R \).

3. If \( \varepsilon > 0 \), then \( P[|R_n - c| \geq \varepsilon] = P[R_n \geq c + \varepsilon] + P[R_n \leq c - \varepsilon] \)

\[ = 1 - P[R_n < c + \varepsilon] + P[R_n \leq c - \varepsilon] \]

\[ \leq 1 - P[R_n \leq c + \varepsilon/2] + P[R_n \leq c - \varepsilon] = 1 - R_n(c + \varepsilon/2) \]

\[ + P_n(c - \varepsilon) + 1 + 0 = 0. \]

4. If \( \varepsilon > 0 \), \( P[|R_n| \geq \varepsilon] \) for large enough \( n \)

\[ = \frac{1}{n} \to 0. \text{ Thus } R_n \overset{p}{\to} 0. \]

But \( E(R_i - 0) = 0 \) \( P[R_i = 0] + e^{-nk} P[R_i = n] = \frac{1}{n} e^{-nk} \to 0. \)

5. (a) \( P[|Q_n - p| \leq .001] = P[|R_n - p| \leq .001] \) where \( R_n \) is the number of "A" voters

\[ = P\left[ \frac{R_n - np}{(np(1-p))^{1/2}} \leq \frac{.001 \sqrt{n/2}}{\sqrt{p(1-p)}} \right] = \frac{.001 \sqrt{n/2}}{\sqrt{p(1-p)}} \]

where \( R^* \) is normal with mean \( 0 \) and variance \( 1 \). Now

\( P[R^* \leq \varepsilon] = F^*(\varepsilon) - F^*(-\varepsilon) = 2F^*(\varepsilon) - 1. \)

Thus, with

\[ a = \frac{.001 \sqrt{n/2}}{\sqrt{p(1-p)}} \leq .99 \text{ or } F^*(a) \geq .995. \]

5. (continued)

From the table, \( a \geq 2.6 \), or \( n \geq \frac{(2600)^2}{p(1-p)} \). The largest possible value of \( p(1-p) \) occurs at \( p = 1/2, \) so

\[ n \geq \frac{(2600)^2}{1/4} = (1300)^2 = 1,690,000. \]

(b) \( 2F^*(b) = 1 \geq .95, b = \frac{.01}{\sqrt{p(1-p)}}, \) thus \( F^*(b) \geq .975, \)

or \( b \geq 1.96. \) Therefore \( n \geq (196)^2 \frac{1/4}{(1/4)} = .98^2 = 9604. \)

6. (a) \( \frac{d}{dx} \frac{1}{\sqrt{2\pi x}} e^{-x^2/2} = -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{x} \) (1 + \( \frac{1}{x^2} \)). Thus

\[ \frac{1}{\sqrt{2\pi x}} e^{-x^2/2} \int x e^{-t^2/2} (1 + \frac{1}{t^2}) dt \text{ have the same derivative, hence differ by a constant, necessarily } 0 \]

(let \( x \to \infty \)). Since

\[ \frac{1}{\sqrt{2\pi \pi}} \int x e^{-t^2/2} (1 + \frac{1}{t^2}) dt \leq \frac{1}{\sqrt{2\pi \pi}} \int x e^{-t^2/2} dt, \]

the result follows.

By (a), \( \frac{1}{\sqrt{2\pi x}} e^{-x^2/2} = 1 + \frac{1}{\sqrt{2\pi \pi}} \int x e^{-t^2/2} dt \)

\[ \frac{1}{\sqrt{2\pi \pi}} \int x e^{-t^2/2} dt \leq \frac{1}{\sqrt{2\pi \pi}} \int x e^{-t^2/2} dt \]

The ratio of integrals on the right is \( \leq \frac{1}{x} \) and therefore \( = 0 \) as \( x \to \infty \), proving the result.
8. \( P\{7940 \leq R \leq 8080 \} = P\left( \frac{7940-8000}{40} \leq R^* \leq \frac{8080-8000}{40} \right) = P\{-1.5 \leq R^* \leq 2\} = F^*(2) - F^*(-1.5) = F^*(2) + F^*(1.5) - 1 = .977 + .933 - 1 = .91.\)

---

Section 6.1

1. We specify

\[
P\{R_1, \ldots, R_n \in B_n\} = \int_{B_n} \cdots dx_1 \cdots dx_n.
\]

It follows from this that

\[
P\{R_1, \ldots, R_n \in B_k\} = \int_{B_k} \cdots dx_1 \cdots dx_k.
\]

Thus the probability measures \( P_n \) are consistent, hence we can construct a probability space on which we can define independent random variables \( R_1, R_2, \ldots \), with \( R_n \) having density \( f_n \), \( n = 1, 2, \ldots \).

2. \( P\{R_n = 1 \text{ for infinitely many } n\} = \lim_{n \to \infty} \lim_{m \to \infty} P\{U \cap \bigcup_{k=n}^m \{R_k = 1\}\} \) as in Section 6.1.

But

\[
P\{U \cup \bigcup_{k=n}^m \{R_k = 1\}\} = 1 - P\{\bigcap_{k=n}^m \{R_k = 0\}\} = 1 - q^{-n+1} = 1
\]

as \( n \to \infty \), so \( P\{R_n = 1 \text{ for infinitely many } n\} = 1. \)

\[
P\{\lim_{n \to \infty} R_n = 1\} = P\{R_n = 1 \text{ for sufficiently large } n\}
\]

\[
= 1 - P\{R_n = 0 \text{ for infinitely many } n\} = 1 - 1 = 0
\]

by the above argument.
Section 6.2

4. Let A be the average time required. If \( p \neq q \), there is a positive probability \([p-q]\) of never returning to 0 (see 6.2.7)). Thus \( A \geq (\mid p-q\mid) = A \). If \( p = q \), then regardless of the result of the first trial, the average number of trials required (after the first) to return to 0 is infinite by the remark after the statement of Problem 3. The result follows. Note: A more precise analysis may be found in Problem 5 of Section 6.3.

5. In the gambler's ruin problem starting at \( x > 0 \), the probability of eventually reaching 0 is 1 if \( q \geq p \), and \((q/p)^x\) if \( q < p \) (see (6.2.6)). By symmetry, the probability of reaching \( b \) starting from 0 is 1 if \( p \geq q \), and \((p/q)^b\) if \( p < q \).

Section 6.3

2. (a) \( h_{2n} = \frac{2n}{n} \left( \frac{2n-2}{n-1} \right)^2 \left( \frac{2n-2}{2} \right)^2 = \frac{u_{2n-2}}{2}\)

(b) \[ \frac{u_{2n}}{u_{2n-2}} = \frac{(2n)!}{n!} \left( \frac{2n-2}{n-1} \right)^2 \left( \frac{2n-2}{2} \right)^2 = \frac{(2n)(2n-1)}{n^2} \frac{1}{4} \]

Thus

\[ h_{2n} = \frac{u_{2n-2}}{2} = u_{2n-2} - \frac{u_{2n}}{u_{2n-2}} = u_{2n-2} - u_{2n}. \]

3. \( P(S_1 \neq 0, \ldots, S_{2n} \neq 0) = 1 - P[\text{at least one return in the first } 2n \text{ steps}] = 1 - h_2 - h_4 - \ldots - h_{2n} \]

\[ = 1 - (u_0 - u_2) - (u_2 - u_4) - \ldots - (u_{2n-2} - u_{2n}) \text{ by Problem 2} \]

\[ = u_{2n} \text{ (note } u_0 = 1). \]

\[ P(S_1 \neq 0, \ldots, S_{2n} \neq 0) = P(S_1 \neq 0, \ldots, S_{2n-1} \neq 0, S_{2n} = 0) + P(S_1 \neq 0, \ldots, S_{2n} \neq 0, S_{2n} \neq 0) = h_{2n} + u_{2n}. \]

Section 6.4

4. \( P(S_1 \geq 0, \ldots, S_{2n} \geq 0) = 1 - P(S_i < 0 \text{ for some } i = 1, 2, \ldots, 2n) \)

\[ = 1 - \sum_{i=1}^{2n-1} P[\text{first passage through } -1 \text{ occurs at time } i=1,3,5,\ldots] \]

But \( P[\text{first passage through } -1 \text{ at time } i] = P[\text{first passage through } +1 \text{ at time } i] \) (by (6.3.6), with \( i = 2k+1 \))

\[ \sum_{k=0}^{\infty} \frac{1}{1+2k} \left( \frac{1+2k}{2} \right)^{1+2k} h_{2k+2} = h_{2n+1} \text{ (see Problem 2a).} \]

Thus \( P(S_1 \geq 0, \ldots, S_{2n} \geq 0) = 1 - h_2 - h_4 - \ldots - h_{2n} = u_{2n} \) as in Problem 3.

8. (a) Say the insects meet after \( j \) steps. If the spider walks \( a \) steps east and \( b \) steps north, the fly must walk \( n-a \) steps west and \( n-b \) steps south. But \( a+b = j \) and \( (n-a)+(n-b) = 2n-j = j \), or \( n = j \). Thus \( a+b = n \), which means that they must meet on the diagonal \( D \).

(b) The probability that they will meet with the spider taking \( a \) steps east and \( n-a \) steps north (and the fly taking \( n-a \) steps west and a steps south) is \( \left( \frac{n}{a} \right) \left( \frac{1}{2} \right)^{n} \). Thus the probability that they will meet is

\[ \left( \frac{1}{2} \right)^{n} \sum_{a=0}^{n} \left( \frac{n}{a} \right)^2 = \left( \frac{1}{2} \right)^{2n} \text{ by Problem 7.} \]

(Note that this is really a random walk problem; tilt the picture so that the line from the spider to the fly is the axis.)
3. (continued)

Thus,

$$\lim_{z \to 1} (1-z)A(z) = \sum_{n=0}^{\infty} (a_n - a_{n+1}) = a_0 + (a_1 - a_0) + (a_2 - a_1) + \ldots$$

$$+ (a_n - a_{n-1}) + \ldots$$

$$= \lim_{n \to \infty} a_n.$$ 

4. The generating function of \( R + k \) is \( E(z^{R+k}) = z^k A(\theta) \) \( \forall \theta \).

The generating function of \( kR \) is \( E(z^{kR}) = E[z^{k}][A(\theta)] = A(z^k) \). Now

$$F(n) = P[R \leq n] = \sum_{k=0}^{n} p_k \cdot 1, \ p_k = P[R=k].$$

Thus \( \{F(n)\} \) is the convolution of \( \{p_0, p_1, \ldots\} \) and \( \{1, 1, \ldots\} \), so by Theorem 1, the generating function of \( \{F(n)\} \) is

$$A(z) \sum_{n=0}^{\infty} z^n = A(z) \left(\frac{1}{1-z}\right).$$

5. (a) \( P[R=k] = P[k-1 \text{ failures followed by a success}] = q^{k-1}p, \ k = 1, 2, \ldots \)

(b) \( N_R(s) = \sum_{k=1}^{\infty} e^{-sk} P[R=k] = \sum_{k=1}^{\infty} (qe^{-s})^k \)

$$= \frac{pe^{-s}}{1-qe^{-s}}, \ |qe^{-s}| < 1.$$ 

\( N_R \) is analytic at \( s = 0 \), hence (see Section 5.3)

$$E(R) = -N'_R(0) = -\frac{pe^{-s}}{(1-qe^{-s})^2} \bigg|_{s=0} = -\frac{p}{(1-q)^2} = \frac{1}{p}$$

$$E(R^2) = N''_R(0) = \left[ \frac{(1-qe^{-s})^2(pe^{-s})}{2} \right] \bigg|_{s=0}$$

$$= 2pe^{-s}(1-qe^{-s})(qe^{-s})/(1-qe^{-s})^4 \quad \text{at} \ s = 0, \ \text{i.e.}$$

$$= \frac{2p^2}{1-q^2}.$$ 

6. (b) (continued)

Thus,

$$\text{Var } R = E(R^2) - (E(R))^2 = \frac{1-p}{p^2}.$$ 

The generating function of \( R \) is \( N_R(z) \) with \( z = e^{-s} \), i.e.

$$A(z) = pz/1-qz. \text{ We may compute that}$$

$$A'(z) = \frac{p}{(1-qz)^2}, \ A''(z) = \frac{2pq}{(1-qz)^3}.$$ 

Thus by (6.4.1) and (6.4.2),

$$E(R) = \frac{p}{(1-q)^2} = \frac{1}{p}, \ \text{Var } R = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$ 

7. \( P[R=k] = P[R_1 = \ldots = R_k = 1, R_{k+1} = 0] = p^k q^k + q^k p^k.$$ 

Thus

$$E(R) = \sum_{k=1}^{\infty} \left( q^k p^k + q^k p^k \right).$$ 

But \( \sum_{k=1}^{\infty} \) \( kq^{k-1}p \) is the mean of a random variable with the \( k=1 \)

geometric distribution, i.e. \( 1/p \). Thus

$$E(R) = p \sum_{k=1}^{\infty} kq^{k-1}p + q \sum_{k=1}^{\infty} q^{k-1}p = \frac{2 + q}{q}.$$ 

8. \( P[N_1=j, N_2=k] = P[T_1=j, T_2=k-j] = \)

\( p^j q^k \), \( j = 1, 2, \ldots, k = 2, 3, \ldots, j < k \) (Problem 6b).

$$E(N_1 N_2) = E(T_1 T_1 + T_2) = E(T_1^2) + E(T_1)E(T_2)$$

$$= E(T_1^2) + [E(T_1)]^2$$

$$E(N_1)E(N_2) = E(N_1)E(2E(N_1)) = 2[E(T_1)]^2.$$
Section 6.5

1. If \( n > 0 \), \( \Pr[\sum_{n=1}^{\infty} T_n < N] \leq \Pr[T_1 + \ldots + T_n < \frac{n}{2\lambda}] \) if \( \frac{n}{2\lambda} \geq M \) and \( n \to \infty \) by the Weak Law of Large Numbers.

Thus

\[
\Pr[\sum_{n=1}^{\infty} T_n < \infty] = \Pr[U] \leq \Pr[\sum_{n=1}^{\infty} T_n \leq \infty] = 0.
\]

3. \( \Pr[R_1 = 1, R_{t+T} = 1] = \Pr[R_1 = 1] \Pr[R_{t+T} = 1|R_1 = 1]. \)

But \( \Pr[R_1 = 1] = \Pr[R_0 = 1] \Pr[R_1 = 1|R_0 = 1] + \Pr[R_0 = -1] \Pr[R_1 = 1|R_0 = -1] = \frac{1}{2} \Pr[\text{even number of customers in } (0,t)] + \frac{1}{2} \Pr[\text{odd number of customers in } (0,t)] = \frac{1}{2} \)

\[
\Pr[R_{t+T} = 1|R_1 = 1] = \Pr[\text{even number of customers in } (t,t+T)] = \frac{1}{2} (1 + e^{-2\lambda T}), \text{ by Problem 2.}
\]

Thus \( \Pr[R_1 = 1, R_{t+T} = 1] = \frac{1}{4} (1 + e^{-2\lambda T}). \)

Section 6.6

1. If \( \Pr(A_n^c) = 1, n = 1,2, \ldots \), then \( \Pr(\bigcap_{n=1}^{\infty} A_n) = 1 - \Pr(\bigcup_{n=1}^{\infty} A_n^c) \) and

\[
\Pr(\bigcup_{n=1}^{\infty} A_n^c) \leq \sum_{n=1}^{\infty} \Pr(A_n^c) = 0.
\]

This fails for an uncountable intersection. For example, let \( X_1, X_2, \ldots \) be uniformly distributed between 0 and 1, and take \( A_t = \{ X \neq 0 \text{ for } 0 \leq t \leq 1 \}. \) Each \( A_t \) has probability 1, but \( \bigcap_{0 \leq t \leq 1} A_t = \emptyset, \) hence has probability 0.

2. Given \( \varepsilon > 0 \), choose \( n \) so that \( \frac{1}{n} < \varepsilon. \) Then

\[
\Pr[|R_k - R| \geq \varepsilon \text{ for at least one } k \geq n] \leq \Pr[|R_k - R| \geq \frac{1}{n} \text{ for at least one } k \geq n] \to 0.
\]
3. If $0 < \epsilon \leq 1$, $P[|R_{nm}| \geq \epsilon] = P[R_{nm} \geq 1] = \frac{1}{n} \to 0$, so $R_{nn} \to 0$. But for any $w$ and any $n$, $R_{nm}(w)$ is 1 for exactly one $m = 1, 2, \ldots, n$ and 0 for the other $m$. Thus $\lim R_{nm}(w)$ never exists.

5. By Theorem 1, $\limsup A_n = [-1,1]$, $\liminf A_n = [0]$.

6. $\liminf A_n = \{(x,y): x^2+y^2 < 1\}$, $\limsup A_n = \{(x,y): x^2+y^2 \leq 1\} - \{(0,1),(0,-1)\}$.

**Proof:**

(a) If $x^2+y^2 < 1$ then eventually the distance from

$(x,y)$ to $\left(\frac{(-1)^n}{n},0\right)$ is < 1, hence $(x,y) \in A_n$; thus

$x^2+y^2 < 1$ implies $(x,y) \in \liminf A_n$.

(b) If $x^2+y^2 = 1$ but $(x,y) \not\in (0,1)$ or $(0,-1)$, say $x > 0$.

Then $(x,y) \in A_n$ for all even $n$ since the distance from

$(x,y)$ to $\left(\frac{1}{n},0\right)$ is < 1; but $(x,y) \not\in A_n$ for odd $n$ since the distance from

$(x,y)$ to $\left(\frac{-1}{n},0\right)$ is > 1. Thus

$(x,y) \in \limsup A_n$, $(x,y) \not\in \liminf A_n$ (similar reasoning for $x < 0$).

(c) If $x^2+y^2 > 1$ then eventually $(x,y) \not\in A_n$. Also, $(0,1)$ and $(0,-1)$ are in none of the $A_n$ since the distance from

$(0,1)$ and $(0,-1)$ to $\left(\frac{(-1)^n}{n},0\right)$ is > 1. Thus such points are not in $\limsup A_n$. The result follows from (a), (b) and (c).

7. Let $x = \limsup A_n$. Then $\limsup A_n = (\infty,x)$ or $(-\infty,x]$. For

if $y \in A_n$ for infinitely many $n$ then $x_n > y$ for infinitely

many $n$, hence $\limsup x_n \geq y$. Thus $\limsup A_n \subseteq (\infty,x]$. If

$y < x$ then $x_n > y$ for infinitely many $n$, so $y \in \limsup A_n$.

Thus $(-\infty,x) \subseteq \limsup A_n$, and the result follows. (The same analysis is valid for $\liminf$, with "eventually" replacing "for

infinitely many $n$".)
Chapter 7

Section 7.1

2. This follows from $\Pi^{n+1} = \Pi \Pi^n$, and an induction argument.

3. Let $S = \{-1,0,1\}$, $P_{0,-1} = P_{-1,1} = P_{1,0} = 1$, and let $\mathbb{g}(x) = x^2$.
Let the initial distribution be $\mathbb{P}_0 = P_1 = P_{-1} = \frac{1}{3}$.
$P[R_2 = 0 | R_1 = 1, R_2 = 1] = P[R_2 = 0 | R_1 = -1, R_2 = 1] = 1$.

But

$$P[R_2 = 0 | R_1 = 1, R_2 = 1] = \frac{P[R_2 = 1, R_3 = 0]}{P[R_2 = 1]} = \frac{P[R_2 = 1, R_3 = 0]}{P[R_2 = 1] + P[R_2 = -1]}$$

$$= \frac{P[R_2 = 1]}{P[R_2 = 1] + P[R_2 = -1]} < 1,$$
so $\{\mathbb{g}(R_n)\}$ does not have the Markov property.

Section 7.2

2. $P[R_n = 1 | R_{n+1} = 1, \ldots, R_{n+k} = 1] = P(A | B \cap C)$,

$$A = \{R_n = 1\}, \quad B = \{R_{n+1} = 1, \ldots, R_{n+k} = 1\}, \quad C = \{R_{n+1} = 1\},$$

But

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

Now $P(B | A \cap C) = P(B | C)$ by Problem 1, so $P(A | B \cap C) = P(A | C)$, the desired result.

Section 7.3

2. If $i$ is essential and $i$ leads to $j$, then since the equivalence class $C$ of $i$ is closed, we must have $j \in C$. But then $i$ and $j$ are equivalent, hence $j$ leads to $i$. Conversely, if the condition is satisfied and $i$ leads to $j$, then $j$ leads to $i$, so that $i$ and $j$ are equivalent. Therefore $j \in C$, so $C$ is closed.

3. Let $i \in C$, and assume $i$ leads to $j \notin C$. There is a positive probability of reaching $j$ from $i$, and once having reached $j$ we cannot return to $i$. Thus there is a positive probability of never returning to $i$, hence $\mathbb{f}_{i1} < 1$ and $i$ is transient.

5. (a) Set $p_{ij} = p_{ji} = \mathbb{P}(R_n = j)$ for all $i, j \in S$.

(b) $S$ forms a single aperiodic recurrent class. (Given that $R_0 = j$, the probability of never returning to $j$ is

$$P[R_n \neq j, n = 1,2,3, \ldots] = \lim_{k \to \infty} \mathbb{P}(R_k = j) = 0.$$}

6. $p_{ii}^{(n)} = \sum_{k=0}^{n} f^{(k)}(n-k) p_{ii}^{(n-k)}$, $n = 1,2,\ldots$ (with $f^{(0)} = 0$).

$$p_{ii}^{(0)} = 1, \quad f^{(0)}(n) = \mathbb{P}(R_1 = i) \mathbb{P}(R_2 = i) \cdots \mathbb{P}(R_n = i)$$

Thus the sequence $\{p_{ii}^{(0)} - 1, p_{ii}^{(1, n)}, p_{ii}^{(2, n)} \cdots\}$ is the convolution of $\{f^{(n)}\}$ and $\{p_{ii}\}$, so $U(z) = 1 - H(z)U(z)$.

Section 7.4

3. $\frac{1}{n} \sum_{k=1}^{n} p_{ik}^{(k)} = \frac{1}{d} \sum_{r=1}^{\lfloor \frac{n}{d} \rfloor} \frac{1}{\mu_j} \frac{\sum_{r=1}^{\lfloor \frac{n}{d} \rfloor} p_{ik}^{(k)}}{\sum_{t=0}^{\lfloor \frac{n}{d} \rfloor} p_{ij}^{(k)}}$.

By Theorem 2d and the fact that $a \to a$ implies $\frac{1}{n} \sum_{k=1}^{n} a_k = a$, this is $\frac{1}{d} \sum_{r=1}^{\lfloor \frac{n}{d} \rfloor} f_{ij}^{(r)} d/\mu_j$ where $f_{ij}^{(r)}$ is the probability of reaching $j$ from $i$ in a number of steps that is $= r$ mod $d$. Therefore

$$\frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} = \frac{1}{\mu_j} \frac{1}{d} \sum_{r=1}^{\lfloor \frac{n}{d} \rfloor} f_{ij}^{(r)} = \frac{1}{\mu_j} \frac{1}{d} \mathbb{P}(R_n = j).$$
Section 7.5

1. \( p_{i1}^{(nd)} = \sum_{k \in C_{r-1}} p_{1k}^{(nd-1)} \frac{d}{\mu_k} p_{k1} \) if \( i \in C_r \),

By Theorem 2c of Section 7.4, \( p_{1k}^{(nd-1)} \to d/\mu_k, k \in C_{r-1}. \) By Fatou's lemma,

\[
\frac{d}{\mu_k} = \liminf_{n \to \infty} p_{i1}^{(nd)} \geq \sum_{k \in C_{r-1}} \frac{d}{\mu_k} p_{k1} = \sum_{k \in C_{r-1}} \frac{d}{\mu_k} p_{k1}.
\]

But \( \sum_{i \in C_{r-1}} \frac{1}{\mu_k} = 1 \) by the discussion in the text. It follows as in Theorem 1a that

\[
\frac{1}{\mu_k} = \sum_{i \in C_{r-1}} \frac{1}{\mu_k} p_{k1}.
\]

If we assign probability 0 to states outside of \( C \), we have a stationary distribution for the chain. Now any stationary distribution for the chain must assign probability 0 to states not in \( C \). (If \( \sum_{j \in C} p_{1j}^{(n)} = v_j, j \not\in C, \) let \( n \to \infty; \) since \( j \) is transient or recurrent null, \( p_{1j}^{(n)} \to 0 \) so \( v_j = 0. \) Now a stationary distribution \( [v_j] \) for the chain also induces a stationary distribution on each cyclically moving subclass \( D \) of \( C \), relative to \( \mu_k \), namely \( [dv_j, j \in D]. \) (Note that \( \sum_{j \in D} v_j = \frac{d}{\mu_k} \) for each subclass \( D \), because of the cyclic movement.) By the argument in the text, \( dv_j = d/\mu_k \), and the result follows.

Section 8.2

2. (a) If \( \theta_1 < \theta_2 \),

\[
p_{\theta_2}(x)/p_{\theta_1}(x) = \frac{e^{-n\theta_2} x_1 \cdots x_n}{e^{-n\theta_1} x_1 \cdots x_n} = e^{n(\theta_2 - \theta_1)} x_1 \cdots x_n \]

where \( x_1, \ldots, x_n \geq 0, 1 \).

(b) If \( \theta_1 < \theta_2 \),

\[
p_{\theta_2}(x)/p_{\theta_1}(x) = \frac{\theta_2 x_1 \cdots x_n}{\theta_1 x_1 \cdots x_n} = \left( \frac{\theta_2}{\theta_1} \right)^{x_1 + \cdots + x_n}.
\]

Remark: The MLR with \( t(x) = x \) holds when \( p_\theta(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} \) the probability function of a binomial random variable with parameters \( n \) and \( \theta, 0 \leq \theta \leq 1; \) the argument is exactly as above.

For the sake of definiteness, we give the form of the UMP test at level \( \alpha \) in case (b). We have

\[
\varphi(x) = \begin{cases} 
1 & \text{if } \sum_{k=1}^n x_k > c \\
0 & \text{if } \sum_{k=1}^n x_k < c \\
a & \text{if } \sum_{k=1}^n x_k = c
\end{cases}
\]

where \( c \) is chosen \( \in \{0, 1, \ldots, n\} \) so that \( P_\theta \{x: t(x) = c\} = \alpha \), i.e.

\[
\sum \binom{n}{k} \theta^k (1 - \theta)^{n-k} = \alpha,
\]

where \( \theta \) is chosen \( \in [0, 1] \).
2. (continued)

(c) 
\[
\frac{p_{\theta+1}(x)}{p_{\theta}(x)} = \frac{\theta+1}{\theta} \binom{N-\theta-1}{n-x} \frac{\theta+1}{\theta+1-x} \frac{N-\theta-n+x}{N-\theta},
\]
which is an increasing function of \( t(x) = x \).

(d) If \( \theta_1 < \theta_2 \),
\[
\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} = \frac{(2\pi\theta_2)^{-n/2} e^{-\sum x_k^2/2\theta_2}}{(2\pi\theta_1)^{-n/2} e^{-\sum x_k^2/2\theta_1}} = \left(\frac{\theta_1}{\theta_2}\right)^{n/2} \exp\left[\frac{1}{2\theta_2} \sum x_k^2 - \frac{1}{2\theta_1} t(x)\right]
\]
where \( t(x) = \sum x_k^2/n \).

4. The test is of the form: reject \( H_0 \) if \( \sum x_k > c \), accept \( H_0 \) if \( \sum x_k < c \), where \( c = n\theta_0 + \sqrt{n} \sigma \eta_\alpha, \alpha \leq .05 \) (see Example 3 of Section 8.2). From the table of the normal distribution function, \( N_\alpha \geq 1.64 \). Also,
\[
.03 \geq \beta = F^{n-1}(c-n\theta_0)/\sqrt{n} \sigma.
\]
Let \( N_\beta \) be the number such that \( F^{n-1}(n_\beta) = \beta \).

4. (continued)

Then \( c = n\theta_0 + \sqrt{n} \sigma \eta_\beta, \eta_\beta \leq N_0 = 1.88 \). Thus
\[
\frac{c-n\theta_0}{\sqrt{n} \sigma} \geq 1.64,
\]
\[
\frac{c-n(\theta_0 + \sigma)}{\sqrt{n} \sigma} \leq -1.88.
\]
Subtract the second equation from the first to obtain \( \sqrt{n} \geq 3.52 \), or \( n \geq 12.4 \). Thus the minimum value of \( n \) is 13.

7. By Problem 6, \( \beta = 1 - Q(\beta) = 1 - (1-\alpha)(\theta_0/\theta_1)^n = (1-\alpha)2^{-n} \) so the set of admissible risk points is \( \left\{ (\alpha, (1-\alpha)2^{-n}): 0 \leq \alpha \leq 1 \right\} \). The upper boundary of the risk set is \( \left\{ (1-\alpha, 1-(1-\alpha)2^{-n}): 0 \leq \alpha \leq 1 \right\} \)
\( \left\{ (\alpha, \beta): 0 \leq \alpha \leq 1, (1-\alpha)2^{-n} \leq \beta \leq 1-\alpha 2^{-n} \right\} \).

8. By Problem 2(b), we reject if \( x_1 + x_2 + x_3 > c \), accept if \( x_1 + x_2 + x_3 < c \).

Thus we take \( c = 2 \). We reject if \( x_1 + x_2 + x_3 = 3 \), accept if \( x_1 + x_2 + x_3 = 0 \) or 1, and if \( x_1 + x_2 + x_3 = 2 \) we reject with probability \( \alpha \), where \( 1/64 + 9\alpha/64 = .1 \) or \( \alpha = .6 \). The power function is \( Q(\theta) = \theta^3 + .6(3)\theta^2(1-\theta) = (9\theta^2 - 6\theta^3)/5 \).
9. If $\varphi$ is admissible let $\varphi_\lambda$ be a LRT with the same error probabilities (Theorem 4). By the first proof of the Neyman-Pearson Lemma, $\varphi_\lambda$ is Bayes with $c_1 = c_2 = 1$, $p = \lambda / 1 + \lambda$. (When $\lambda = \infty$ we have $p = 1$ and $\alpha_\lambda = 0$, hence $B(\varphi_\lambda) = 0$, so $\varphi_\lambda$ is still Bayes in this case.) Since $\alpha = \alpha_\lambda$ and $\beta = \beta_\lambda$, $\varphi$ is also Bayes by (8.2.3). Conversely, if $\varphi$ is inadmissible and $c_1, c_2 > 0$, $0 < p < 1$, (8.2.3) shows that $\varphi$ cannot be Bayes.

Let $R$ be uniformly distributed between 0 and $\theta$, and let $H_0: \theta = 1, H_1: \theta = 2$. Let $\varphi_1 = 0$, and let $\varphi_2(x) = 0, 0 \leq x \leq 1$; $\varphi_2(x) = 1, 1 < x \leq 2$. Then $\alpha(\varphi_1) = 0$, $\beta(\varphi_1) = 1$, $\alpha(\varphi_2) = 0$, $\beta(\varphi_2) = 1 / 2$. $\varphi_1$ and $\varphi_2$ are Bayes when $p = 1$ since $\beta(\varphi_1) = \beta(\varphi_2) = 0$. But $\varphi_1$ is inadmissible since $\varphi_2$ is better than $\varphi_1$.

11. Assume first that $\beta(\varphi) > 0$, hence $\varphi$ is of size $\alpha$ by Problem 10. Since $\alpha$ is most powerful, it is admissible, hence by Problem 9, $\varphi$ is a Bayes solution for some $c_1, c_2$ and $p$. But if $\lambda = p c_1 / (1 - p) c_2$, examination of the way the Bayes solution was constructed shows that $\varphi(x)$ must be 1 for $x > \lambda$, and

$\varphi(x) = 0$ for $x < \lambda$, except for $x$ in a set of Lebesgue measure 0.

(If for example, $\varphi'(x) \leq 1 - \delta$ and $L(x) > \lambda$ on a set of positive Lebesgue measure, $B(\varphi')$ would be $> B(\varphi)$.) If $\beta(\varphi) = 0$ then $\varphi$ is a Bayes solution with $p = 0$ since in this case $B(\varphi) = 0$ by (8.2.3). Thus the above argument still applies.

12. Part (a) follows from the discussion after Theorem 3; (b) follows from (a) and Theorem 3. Part (c) holds since every LRT is Bayes (see the first proof of the Neyman-Pearson Lemma and the solution to Problem 9).

13. If $\alpha(\varphi_1) < \alpha(\varphi_2)$ but $\beta(\varphi_1) > \beta(\varphi_2)$, both statements are false. Numerical examples can be produced easily.
7. (continued)

\[ \rho_\theta(\theta) = E_\theta[(R - \theta)/2 + (1 - \theta)^2] \]

\[ = \frac{1}{4} E_\theta[(R - \theta + 1 - \theta)^2] = \frac{1}{4} [\vartheta R + (1 - \theta)^2] \]

\[ = \frac{1}{4} [\theta + (1 - \theta)^2] = \frac{1}{4} (\theta^2 - \theta + 1) \]

\[ B(\psi) = \int_0^\infty e^{-\psi} \frac{1}{4} (\theta^2 - \theta + 1) d\theta = \frac{1}{4} (2 - 1 + 1) = \frac{1}{2} \]

The maximum likelihood estimate of \( \gamma \) is found by differentiating \( n[e^{-\theta} + 1] \) with respect to \( \theta \) and setting the result equal to zero; we obtain \( \hat{\theta} = x \). The risk function using \( \hat{\theta} \) is

\[ F_0[(R - \theta)^2] = 0, \text{ hence } B(\hat{\theta}) = \int_0^\infty 0 e^{-\theta} d\theta = 1 > B(\psi). \]

Section 3.4

4. \( f_\theta(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(x - \theta)^2}{2\sigma^2} + \frac{(y - \theta)^2}{2\tau^2}\right] \)

since

\[ \frac{(x - \theta)^2}{\sigma^2} + \frac{(y - \theta)^2}{\tau^2} = \frac{x^2}{\sigma^2} + \frac{y^2}{\tau^2} - 2\theta \left( \frac{x}{\sigma^2} + \frac{y}{\tau^2} \right) + \theta^2 \left( \frac{1}{\sigma^2} + \frac{1}{\tau^2} \right) \]

the result follows.

5. (a) The results may be tabulated as follows:

<table>
<thead>
<tr>
<th>( s(x) )</th>
<th>( b(x) )</th>
<th>( c_1(x) )</th>
<th>( t_1(x) )</th>
<th>( c_2(x) )</th>
<th>( t_2(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) ( (1 - \theta)^2 )</td>
<td>( \ln \theta )</td>
<td>( -\ln(1 - \theta) )</td>
<td>( x )</td>
<td>( - )</td>
<td>( - )</td>
</tr>
<tr>
<td>(ii) ( e^{-\theta} )</td>
<td>( \frac{1}{\theta} )</td>
<td>( \ln \theta )</td>
<td>( x )</td>
<td>( - )</td>
<td>( - )</td>
</tr>
<tr>
<td>(iii) ( \frac{1}{\sqrt{2\tau}} e^{-\mu^2/2\tau^2} )</td>
<td>( \frac{1}{\theta^2} )</td>
<td>( \frac{\ln x^2}{\theta^2} )</td>
<td>( x )</td>
<td>( \frac{1}{\theta^2} )</td>
<td>( x )</td>
</tr>
<tr>
<td>(iv) ( \frac{1}{\Gamma(1, \theta_2)} )</td>
<td>( \theta_1^{-1} )</td>
<td>( \ln x )</td>
<td>( - \frac{1}{\theta_2} )</td>
<td>( x )</td>
<td></td>
</tr>
<tr>
<td>(v) ( \frac{1}{\beta(\theta_1, \theta_2)} )</td>
<td>( \theta_1^{-1} )</td>
<td>( \ln x )</td>
<td>( \theta_2^{-1} )</td>
<td>( \ln(1 - \theta) )</td>
<td>( x )</td>
</tr>
<tr>
<td>(vi) ( \frac{\theta_1^{-1}}{(1 - \theta)^{\theta_2}} )</td>
<td>( \theta_1^{-1} )</td>
<td>( \ln(1 - \theta) )</td>
<td>( x )</td>
<td>( - )</td>
<td>( - )</td>
</tr>
</tbody>
</table>

(b) It follows quickly from the Factorization Theorem that

\[ n \sum_{j=1}^n t_1(R_j), \ldots, \sum_{j=1}^n t_k(R_j) \]

is sufficient.

6. If \( \phi \) is any test let \( \phi'(x) = E[\phi(R) | T = t(x)] \). Then \( \phi' \) is a test based on \( T \), and \( E_0 \phi'(R) = E_0 \phi(R) \) for all \( \theta \), hence

\( (\alpha(\phi'), \beta(\phi')) = (\alpha(\phi), \beta(\phi)). \)

Section 8.5

1. \( \gamma(\theta) = \sum_{k=0}^\infty (-1)^k \theta^k / k! \), so the UMVE is

\[ \frac{T - \sum_{i=0}^{T-1} (-1)^i (T-1)}{(T-1) \binom{T-1}{T-1}} = (1 - \frac{1}{T}) \cdot \sum_{t=0}^T \frac{T!}{t!} \cdot \frac{(-1)^{T-t}}{(T-t) \binom{T-t}{T-1}}. \]
3. $\bar{R}$ is sufficient by Example 3, Section 8.4. Now $\bar{R}$ is normal $(\theta, \sigma^2/n)$, hence

$$F_{\theta, \sigma^2/n}(y) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{(y-t)^2}{2\sigma^2}} dt.$$ 

If $F_{\theta, \sigma^2/n}(y) = 0$ for all $\theta > 0$ then $g(y)e^{-ny^2/2\sigma^2} = 0$ for all $y$, hence $g(y) = 0$ for all $y$ (except on a set of Lebesgue measure 0). Thus as in Problem 2a, $p_{0, \sigma^2/n}(y) = 0$, hence $\bar{R}$ is complete.

Since $E(\bar{R}) = \theta$, $\bar{R}$ is a UMVUE of $\theta$; since $\frac{\theta^2}{n} = \text{Var} \bar{R} = E[(\bar{R})^2] - E(\bar{R})^2$, $(\bar{R})^2 - \frac{\theta^2}{n}$ is a UMVUE of $\theta^2$.

7. $E(R_1 \ldots R_j | \sum R_i = k) = P(R_1 = \ldots = R_j = 1 | \sum R_i = k)$

$$= F(R_1 = \ldots = R_j = 1, \sum_{i=1}^{n} R_i = k) / F(\sum R_i = k)$$

$$= \frac{\binom{n}{k-j} \binom{k-j}{(1-\theta)n-j-k} / \binom{k}{(1-\theta)n-k}}{\binom{n}{k}} = \frac{\binom{n}{k-j}}{n(n-1) \ldots (n-j)} \text{ as in Example 1.}$$

8. $E(R_1 R_2) = E(R_1)E(R_2) = \theta^2$, hence $E(R_1 R_2 | \sum R_i = k)$ is an unbiased estimate of $\theta^2$ based on a complete sufficient statistic. By Example 2, Section 8.5, $E(R_1 R_2 | \sum R_i = k) = k(k-1)/n^2$.

10. Assume $\hat{\phi}$ is a best estimate of $\theta$. Let $\hat{\phi}'(\theta) = \theta_0$. Then

$$p_{\hat{\phi}'}(\theta) \leq p_{\hat{\phi}}(\theta) = (\hat{\phi} - \theta_0)^2$$

hence $p_{\hat{\phi}'}(\theta) = E_{\theta_0}[(\hat{\phi}'(\theta) - \theta_0)^2]$.

Consequently $\hat{\phi}(\theta) = \theta_0$. But $\theta_0$ is arbitrary, so this is a contradiction.

12. (a) $E_{\theta_0} \frac{\partial}{\partial \theta} \ln f_{\theta}(x) = \int_{-\infty}^{\infty} \frac{1}{f_{\theta}(x)} \frac{\partial f_{\theta}(x)}{\partial \theta} f_{\theta}(x) dx$

$$= \int_{-\infty}^{\infty} \frac{1}{f_{\theta}(x)} \frac{\partial f_{\theta}(x)}{\partial \theta} f_{\theta}(x) dx - \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} f_{\theta}(x) dx$$

$$= \int_{-\infty}^{\infty} \frac{\partial f_{\theta}(x)}{\partial x} \frac{1}{f_{\theta}(x)} f_{\theta}(x) dx$$

$$= \int_{-\infty}^{\infty} \frac{\partial f_{\theta}(x)}{\partial x} f_{\theta}(x) dx$$

(b) $\hat{\phi}'(\theta) = \frac{\partial}{\partial \theta} E_{\theta} \hat{\phi}(R) = \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} \hat{\phi}(x) f_{\theta}(x) dx$

$$= \int_{-\infty}^{\infty} \hat{\phi}(x) \frac{\partial f_{\theta}(x)}{\partial \theta} f_{\theta}(x) dx$$

$$= \int_{-\infty}^{\infty} \hat{\phi}(x) \frac{\partial}{\partial \theta} f_{\theta}(x) dx$$

(c) By the Schwarz inequality,

$$[\text{Cov}_{\theta}(\hat{\phi}(R), \frac{\partial}{\partial \theta} \ln f_{\theta}(R))]^2 \leq \text{Var}_{\theta}(\hat{\phi}(R)) \text{Var}_{\theta} \frac{\partial}{\partial \theta} \ln f_{\theta}(R)$$

The result follows from (a) and (b).

13. The sample variance is not changed by replacing $R_i$ by $R_i/\mu$; we may assume without loss of generality that $\mu = 0$. Then

$$E[(R_1 - \bar{R})^2] = E[(R_1 - \frac{1}{n} \sum_{j=1}^{n} R_j)^2]$$

$$= E(R_1^2) - \frac{2}{n} E(R_1^2) + \frac{1}{n^2} E(R_j^2)$$

$$= \sigma^2 (1 - \frac{2}{n} + \frac{1}{n^2}) = \frac{(n-1)^2}{n^2} \sigma^2.$$
Section 8.6

3. (a) \( T = \sqrt{n} \frac{R_1}{\sqrt{R_2}} \) where \( R_1 \) is normal \((0,1)\) and \( R_2 \) is chi-square \((n)\). Thus \( T^2 = nR_1^2/R_2 = R_1^2/(R_2/n) \). \( R_1 \) is chi-square \((1)\), so that \( T^2 \) is \( F(1,n) \).

(b) \( 1/R = (R_2/n)/(R_1/m) \) where \( R_1 \) is chi-square \((n)\) and \( R_2 \) is chi-square \((n)\), and the result follows.

(c) This is immediate from the fact that a chi-square \((n)\) random variable is representable as \( W_1^2 + \ldots + W_n^2 \) where the \( W_i \) are independent and normal \((0,1)\).

4. (a) \( \sum_{i=1}^{n} \frac{R_i^2 - \mu^2}{\sigma^2} = \frac{W^2}{\sigma^2} \) is chi-square with \( n \) degrees of freedom. If \( h_n \) is the chi-square \((n)\) density and \( a \) and \( b \) are chosen so that \( \int_a^b h_n(x)dx = 1-\alpha \) then \( P[a \leq \frac{W}{\sigma} \leq b] = 1-\alpha \). Therefore \( [\frac{W}{\sigma}, \frac{W}{\sigma}] \) is a confidence interval for \( \sigma^2 \) with confidence coefficient \( 1-\alpha \).

(b) If \( V^2 \) is the sample variance, \( nV^2/\sigma^2 \) is chi-square with \( n-1 \) degrees of freedom. Thus if \( a \) and \( b \) are chosen so that \( \int_a^b h_{n-1}(x)dx = 1-\alpha \) then \( P[a \leq \frac{V^2}{\sigma^2} \leq b] = 1-\alpha \). Therefore \( [\frac{V^2}{b}, \frac{V^2}{a}] \) is a confidence interval for \( \sigma^2 \) with confidence coefficient \( 1-\alpha \).

5. \( \overline{R}_i - \mu \) is normal \((0,\sigma^2/n_i)\), \( i = 1,2 \), hence

\[
\frac{\overline{R}_1 - \overline{R}_2 - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim \frac{1}{\sqrt{2}} \]

is normal \((0,1)\). (Note that this result may be used to construct confidence intervals if \( \sigma^2 \) is known.) Since

\[
\frac{n_1V_1^2}{\sigma^2} + \frac{n_2V_2^2}{\sigma^2} \]

is chi-square \((n_1-1+n_2-1)\) by Problem 3c,

\[
\left| \frac{(n_1+n_2-2)n_1n_2}{n_1+n_2} \right|^{1/2} \frac{(\overline{R}_1 - \overline{R}_2 - (\mu_1 - \mu_2))}{(\overline{V}_1^2/\sigma^2_1 + \overline{V}_2^2/\sigma^2_2)^{1/2}}
\]

is \( t(n_1+n_2-2) \) and the result follows.

6. \( n_1V_1^2/\sigma_1^2 \) is chi-square \((n_1-1)\), \( i = 1,2 \), hence

\[
\frac{n_2V_2^2/(n_2-1)\sigma_2^2}{n_1V_1^2/(n_1-1)\sigma_1^2} \sim F(n_2-1,n_1-1).
\]

If \( S^2 \) denotes the corrected sample variance

\[
\frac{n}{n-1} V^2 = \frac{1}{n-1} \sum_{i=1}^{n} (R_i - \overline{R})^2
\]

then

\[
\frac{S^2}{\frac{V_1^2}{\sigma_1^2}} \frac{V_1^2}{\sigma_1^2} \sim F(n_2-1,n_1-1)
\]

and this allows construction of confidence intervals in the usual way.
7. (a) \( E_0 q_k(R) = P_0[k \not\in C(R)] = 1 - P_0[k \in C(R)] \).

If \( H_0 \) is true, \( k = \gamma(\theta) \) hence \( P_0[k \in C(R)] \geq 1 - \alpha \) and the result follows.

(b) \( P_0[\gamma(\theta) \in C(R)] = P_0[x: \varphi_k(x) = 0] \) where \( k = \gamma(\theta) \)

(note \( \varphi_k \) exists for each \( k \) of the form \( \gamma(\theta) \), by hypothesis)

\[ = 1 - P_0[x: \varphi_k(x) = 1] \] since the tests are nonrandomized

\[ = 1 - E_0 q_k(R). \]

But when the true parameter is \( \theta \) then the null hypothesis that \( \gamma(\theta) = k \) is true, hence \( E_0 q_k(R) \leq \alpha \) and the result follows.