Students' Solutions Manual

for

Carlton and Devore's

Probability

with Applications in Engineering, Science, and Technology

Matthew A. Carlton
Cal Poly State University

CONTENTS

CHAPTER 1

Section 1.1

1.

a. A but not $B = A \cap B'$

b. at least one of A and $B = A \cup B$

c. exactly one hired = A and not B, or B and not $A = (A \cap B') \cup (B \cap A')$

3.

a. $\mathcal{S}$ = {1324, 1342, 1423, 1432, 2314, 2341, 2413, 2431, 3124, 3142, 4123, 4132, 3214, 3241, 4213, 4231}

b. Event A contains the outcomes where 1 is first in the list:
A = {1324, 1342, 1423, 1432}.

c. Event B contains the outcomes where 2 is first or second:
B = {2314, 2341, 2413, 2431, 3214, 3241, 4213, 4231}.

d. The event $A \cup B$ contains the outcomes in A or B or both:
$A \cup B$ = {1324, 1342, 1423, 1432, 2314, 2341, 2413, 2431, 3214, 3241, 4213, 4231}.
$A \cap B = \varnothing$, since 1 and 2 can't both get into the championship game.
$A' = S - A$ = {2314, 2341, 2413, 2431, 3124, 3142, 4123, 4132, 3214, 3241, 4213, 4231}.

5.

a. $A = \{SSF, SFS, FSS\}$.

b. $B = \{SSS, SSF, SFS, FSS\}$.

c. For event C to occur, the system must have component 1 working (S in the first position), then at least one of the other two components must work (at least one S in the second and third positions): $C = \{SSS, SSF, SFS\}$.

d. $C' = \{SFF, FSS, FSF, FFS, FFF\}$.
$A \cup C = \{SSS, SSF, SFS, FSS\}$.
$A \cap C = \{SSF, SFS\}$.
$B \cup C = \{SSS, SSF, SFS, FSS\}$. Notice that B contains C, so $B \cup C = B$.
$B \cap C = \{SSS\ SSF, SFS\}$. Since B contains C, $B \cap C = C$.

7.

a. The $3^3 = 27$ possible outcomes are numbered below for later reference.

Outcome Number	Outcome	Outcome Number	Outcome
1	111	15	223
2	112	16	231
3	113	17	232
4	121	18	233
5	122	19	311
6	123	20	312
7	131	21	313
8	132	22	321
9	133	23	322
10	211	24	323
11	212	25	331
12	213	26	332
13	221	27	333
14	222		

b. Outcome numbers 1, 14, 27 above.

c. Outcome numbers 6, 8, 12, 16, 20, 22 above.

d. Outcome numbers 1, 3, 7, 9, 19, 21, 25, 27 above.

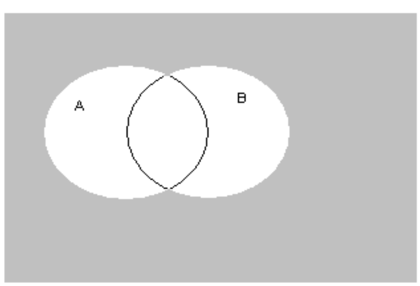
9.

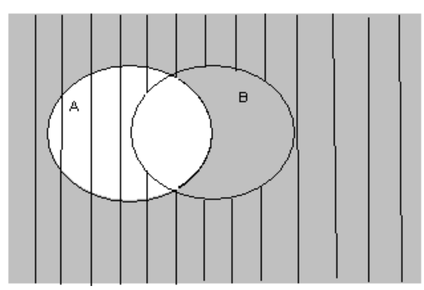
a. S = {*BBBAAAA*, *BBABAAA*, *BBAABAA*, *BBAAABA*, *BBAAAAB*, *BABBAAA*, *BABABAA*, *BABAABA*, *BABAAAB*, *BAABBAA*, *BAABABA*, *BAABAAB*, *BAAABBA*, *BAAABAB*, *BAAAABB*, *ABBBAAA*, *ABBABAA*, *ABBAABA*, *ABBAAAB*, *ABABBAA*, *ABABABA*, *ABABAAB*, *ABAABBA*, *ABAABAB*, *ABAAABB*, *AABBBAA*, *AABBABA*, *AABBAAB*, *AABABBA*, *AABABAB*, *AABAABB*, *AAABBBA*, *AAABBAB*, *AAABABB*, *AAAABBB*}.

b. *AAAABBB*, *AAABABB*, *AAABBAB*, *AABAABB*, *AABABAB*.

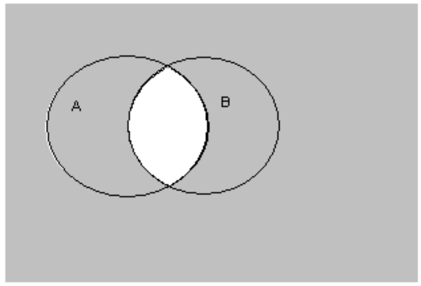
11.

a. In the diagram on the left, the shaded area is $(A \cup B)'$. On the right, the shaded area is A', the striped area is B', and the intersection $A' \cap B'$ occurs where there is both shading <u>and</u> stripes. These two diagrams display the same area.





b. In the diagram below, the shaded area represents $(A \cap B)'$. Using the right-hand diagram from (a), the union of A' and B' is represented by the areas that have either shading or stripes (or both). Both of the diagrams display the same area.



Section 1.2

13.

a. .07.

b. .15 + .10 + .05 = .30.

c. Let A = the selected individual owns shares in a stock fund. Then $P(A)$ = .18 + .25 = .43. The desired probability, that a selected customer does not shares in a stock fund, equals $P(A')$ = 1 – $P(A)$ = 1 – .43 = .57. This could also be calculated by adding the probabilities for all the funds that are not stocks.

15.

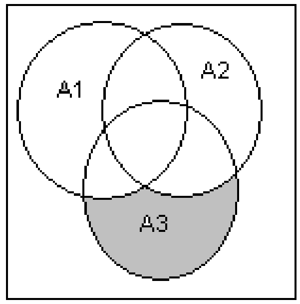
a. $A_1 \cup A_2$ = "awarded either #1 or #2 (or both)": from the addition rule,
$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$ = .22 + .25 – .11 = .36.

b. $A_1' \cap A_2'$ = "awarded neither #1 or #2": using the hint and part (a),
$P(A_1' \cap A_2') = P((A_1 \cup A_2)') = 1 - P(A_1 \cup A_2)$ = 1 – .36 = .64.

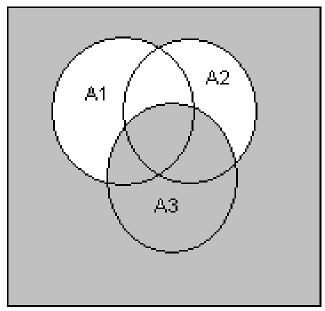
c. $A_1 \cup A_2 \cup A_3$ = "awarded at least one of these three projects": using the addition rule for 3 events,
$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3) =$
.22 +.25 + .28 – .11 – .05 – .07 + .01 = .53.

d. $A_1' \cap A_2' \cap A_3'$ = "awarded none of the three projects":
$P(A_1' \cap A_2' \cap A_3')$ = 1 – P(awarded at least one) = 1 – .53 = .47.

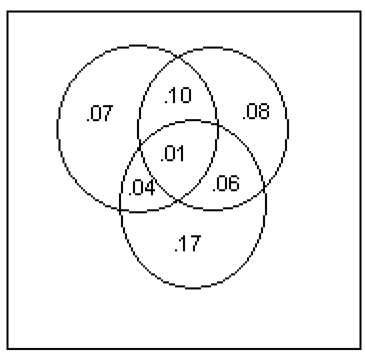
e. $A_1' \cap A_2' \cap A_3$ = "awarded #3 but neither #1 nor #2": from a Venn diagram,
$P(A_1' \cap A_2' \cap A_3) = P(A_3) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3) =$
$.28 - .05 - .07 + .01 = .17$. The last term addresses the "double counting" of the two subtractions.



f. $(A_1' \cap A_2') \cup A_3$ = "awarded neither of #1 and #2, or awarded #3": from a Venn diagram,
$P((A_1' \cap A_2') \cup A_3) = P(\text{none awarded}) + P(A_3) = .47$ (from **d**) $+ .28 = 75$.



Alternatively, answers to **a-f** can be obtained from probabilities on the accompanying Venn diagram:



17.

a. Let E be the event that at most one purchases an electric dryer. Then E' is the event that at least two purchase electric dryers, and $P(E') = 1 - P(E) = 1 - .428 = .572$.

b. Let A be the event that all five purchase gas, and let B be the event that all five purchase electric. All other possible outcomes are those in which at least one of each type of clothes dryer is purchased. Thus, the desired probability is $1 - [P(A) - P(B)] = 1 - [.116 + .005] = .879$.

19.

a. The probabilities do not add to 1 because there are other software packages besides SPSS and SAS for which requests could be made.

b. $P(A') = 1 - P(A) = 1 - .30 = .70$.

c. Since A and B are mutually exclusive events, $P(A \cup B) = P(A) + P(B) = .30 + .50 = .80$.

d. By deMorgan's law, $P(A' \cap B') = P((A \cup B)') = 1 - P(A \cup B) = 1 - .80 = .20$.
In this example, deMorgan's law says the event "neither A nor B" is the complement of the event "either A or B." (That's true regardless of whether they're mutually exclusive.)

21. Let A be that the selected joint was found defective by inspector A, so $P(A) = \frac{724}{10{,}000}$. Let B be analogous for inspector B, so $P(B) = \frac{751}{10{,}000}$. The event "at least one of the inspectors judged a joint to be defective is $A \cup B$, so $P(A \cup B) = \frac{1159}{10{,}000}$.

a. By deMorgan's law, $P(\text{neither } A \text{ nor } B) = P(A' \cap B') = 1 - \mathrm{P}(\mathrm{A} \cup \mathrm{B}) = 1 - \frac{1159}{10{,}000} = \frac{8841}{10{,}000} = .8841$.

b. The desired event is $B \cap A'$. From a Venn diagram, we see that $P(B \cap A') = P(B) - P(A \cap B)$. From the addition rule, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ gives $P(A \cap B) = .0724 + .0751 - .1159 = .0316$. Finally, $P(B \cap A') = P(B) - P(A \cap B) = .0751 - .0316 = .0435$.

23. In what follows, the first letter refers to the auto deductible and the second letter refers to the homeowner's deductible.

a. $P(MH) = .10$.

b. $P(\text{low auto deductible}) = P(\{LN, LL, LM, LH\}) = .04 + .06 + .05 + .03 = .18$. Following a similar pattern, $P(\text{low homeowner's deductible}) = .06 + .10 + .03 = .19$.

c. $P(\text{same deductible for both}) = P(\{LL, MM, HH\}) = .06 + .20 + .15 = .41$.

d. $P(\text{deductibles are different}) = 1 - P(\text{same deductible for both}) = 1 - .41 = .59$.

e. $P(\text{at least one low deductible}) = P(\{LN, LL, LM, LH, ML, HL\}) = .04 + .06 + .05 + .03 + .10 + .03 = .31$.

f. $P(\text{neither deductible is low}) = 1 - \mathrm{P}(\text{at least one low deductible}) = 1 - .31 = .69$.

25. Assume that the computers are numbered 1-6 as described and that computers 1 and 2 are the two laptops. There are 15 possible outcomes: (1,2) (1,3) (1,4) (1,5) (1,6) (2,3) (2,4) (2,5) (2,6) (3,4) (3,5) (3,6) (4,5) (4,6) and (5,6).

a. P(both are laptops) = $P(\{(1,2)\})$ = $\frac{1}{15}$ =.067.

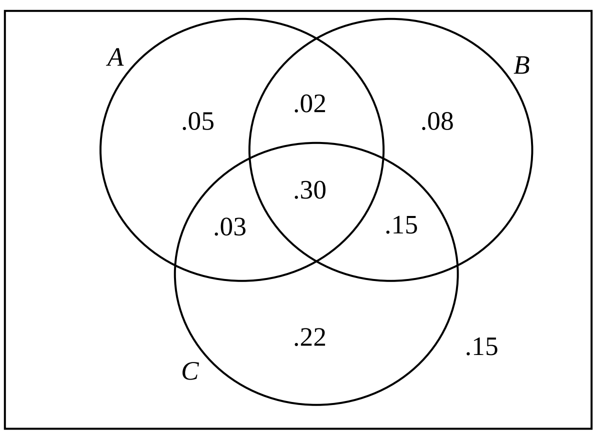
b. P(both are desktops) = $P(\{(3,4)\ (3,5)\ (3,6)\ (4,5)\ (4,6)\ (5,6)\})$ = $\frac{6}{15}$ = .40.

c. P(at least one desktop) = 1 – P(no desktops) = 1 – P(both are laptops) = 1 – .067 = .933.

d. P(at least one of each type) = 1 – P(both are the same) =
1 – [P(both are laptops) + P(both are desktops)] = 1 – [.067 + .40] = .533.

27. By rearranging the addition rule, $P(A \cap B) = P(A) + P(B) - P(A \cup B)$ = .40 + .55 – .63 = .32. By the same method, $P(A \cap C)$ = .40 + .70 – .77 = .33 and $P(B \cap C)$ = .55 + .70 – .80 = .45. Finally, rearranging the addition rule for 3 events gives
$P(A \cap B \cap C) = P(A \cup B \cup C) - P(A) - P(B) - P(C) + P(A \cap B) + P(A \cap C) + P(B \cap C)$ = .85 – .40 – .55 – .70 + .32 + .33 + .45 = .30.

These probabilities are reflected in the Venn diagram below.



a. $P(A \cup B \cup C)$ = .85, as given.

b. P(none selected) = 1 – P(at least one selected) = 1 – $P(A \cup B \cup C)$ = 1 – .85 = .15.

c. From the Venn diagram, P(only automatic transmission selected) = .22.

d. From the Venn diagram, P(exactly one of the three) = .05 + .08 + .22 = .35.

29. Recall there are 27 equally likely outcomes.

a. P(all the same station) = P((1,1,1) or (2,2,2) or (3,3,3)) = $\frac{3}{27} = \frac{1}{9}$.

b. P(at most 2 are assigned to the same station) = 1 – P(all 3 are the same) = $1 - \frac{1}{9} = \frac{8}{9}$.

c. P(all different stations) = P((1,2,3) or (1,3,2) or (2,1,3) or (2,3,1) or (3,1,2) or (3,2,1))
= $\frac{6}{27} = \frac{2}{9}$.

Section 1.3

31.

a. $(10)(10)(10)(10) = 10^4 = 10{,}000$. These are the strings 0000 through 9999.

b. Count the number of prohibited sequences. There are (i) 10 with all digits identical (0000, 1111, …, 9999); (ii) 14 with sequential digits (0123, 1234, 2345, 3456, 4567, 5678, 6789, and 7890, plus these same seven descending); (iii) 100 beginning with 19 (1900 through 1999). That's a total of 10 + 14 + 100 = 124 impermissible sequences, so there are a total of 10,000 – 124 = 9876 permissible sequences. The chance of randomly selecting one is just $\frac{9876}{10{,}000} = .9876$.

c. All PINs of the form 8xx1 are legitimate, so there are (10)(10) = 100 such PINs. With someone randomly selecting 3 such PINs, the chance of guessing the correct sequence is 3/100 = .03.

d. Of all the PINs of the form 1xx1, eleven is prohibited: 1111, and the ten of the form 19x1. That leaves 89 possibilities, so the chances of correctly guessing the PIN in 3 tries is 3/89 = .0337.

33.

a. Because order is important, we'll use ${}_8P_3 = (8)(7)(6) = 336$.

b. Order doesn't matter here, so we use $\binom{30}{6} = 593{,}775$.

c. The number of ways to choose 2 zinfandels from the 8 available is $\binom{8}{2}$. Similarly, the number of ways to choose the merlots and cabernets are $\binom{10}{2}$ and $\binom{12}{2}$, respectively. Hence, the total number of options (using the Fundamental Counting Principle) equals $\binom{8}{2}\binom{10}{2}\binom{12}{2} = (28)(45)(66) = 83{,}160$.

d. The numerator comes from part **c** and the denominator from part **b**: $\frac{83{,}160}{593{,}775} = .140$.

e. We use the same denominator as in part **d**. The number of ways to choose all zinfandel is $\binom{8}{6}$, with similar answers for all merlot and all cabernet. Since these are disjoint events, $P(\text{all same}) = P(\text{all zin}) + P(\text{all merlot}) + P(\text{all cab}) = \dfrac{\binom{8}{6}+\binom{10}{6}+\binom{12}{6}}{\binom{30}{6}} = \dfrac{1162}{593{,}775} = .002$.

35.

a. Since there are 5 receivers, 4 CD players, 3 speakers, and 4 turntables, the total number of possible selections is (5)(4)(3)(4) = 240.

b. We now only have 1 choice for the receiver and CD player: (1)(1)(3)(4) = 12.

c. Eliminating Sony leaves 4, 3, 3, and 3 choices for the four pieces of equipment, respectively: (4)(3)(3)(3) = 108.

d. From **a**, there are 240 possible configurations. From **c**, 108 of them involve zero Sony products. So, the number of configurations with at least one Sony product is 240 – 108 = 132.

e. Assuming all 240 arrangements are equally likely, $P(\text{at least one Sony}) = \frac{132}{240} = .55$.

Next, *P*(exactly one component Sony) = *P*(only the receiver is Sony) + *P*(only the CD player is Sony) + *P*(only the turntable is Sony). Counting from the available options gives
$P(\text{exactly one component Sony}) = \frac{(1)(3)(3)(3)+(4)(1)(3)(3)+(4)(3)(3)(1)}{240} = \frac{99}{240} = .413$.

37.

a. There are 12 American beers and 8 + 9 = 17 imported beers, for a total of 29 beers. The total number of five-beer samplers is $\binom{29}{5}$ = 118,755. Among them, the number with at least 4 American beers, i.e. exactly 4 or exactly 5, is $\binom{12}{4}\binom{17}{1}+\binom{12}{5}\binom{17}{0}$ = 9,207. So, the probability that you get at least 4 American beers is $\frac{9,207}{118,755}$ = .0775.

b. The number of five-beer samplers consisting of only American beers is $\binom{12}{5}\binom{8}{0}\binom{9}{0}$ = 792. Similarly, the number of all-Mexican and all-German beer samplers are $\binom{8}{5}$ = 56 and $\binom{9}{5}$ = 126, respectively. Therefore, there are 792 + 56 + 126 = 974 samplers with all beers from the same country, and the probability of randomly receiving such a sampler is $\frac{974}{118,755}$ = .0082.

39.

a. Since order doesn't matter, the number of possible rosters is $\binom{16}{6}$ = 8008.

b. The number of ways to select 2 women from among 5 is $\binom{5}{2}$ = 10, and the number of ways to select 4 men from among 11 is $\binom{11}{4}$ = 330. By the Fundamental Counting Principle, the total number of (2-woman, 4-man) teams is (10)(330) = 3300.

c. Using the same idea as in part **b**, the count is $3300+\binom{5}{3}\binom{11}{3}+\binom{5}{4}\binom{11}{2}+\binom{5}{5}\binom{11}{1}$ = 5236.

d. P(exactly 2 women) = $\frac{3300}{8008}$ = .4121; P(at least 2 women) = $\frac{5236}{8008}$ = .6538.

41. There are $\binom{5}{2}$= 10 possible ways to select the positions for *B*'s votes: *BBAAA*, *BABAA*, *BAABA*, *BAAAB*, *ABBAA*, *ABABA*, *ABAAB*, *AABBA*, *AABAB*, and *AAABB*. Only the last two have *A* ahead of *B* throughout the vote count. Since the outcomes are equally likely, the desired probability is 2/10 = .20.

43.

a. There are 6 75W bulbs and 9 other bulbs. So, P(select exactly 2 75W bulbs) = P(select exactly 2 75W bulbs and 1 other bulb) $=\dfrac{\binom{6}{2}\binom{9}{1}}{\binom{15}{3}}=\dfrac{(15)(9)}{455}=.2967$.

b. P(all three are the same rating) = P(all 3 are 40W or all 3 are 60W or all 3 are 75W) =

$$\frac{\binom{4}{3}+\binom{5}{3}+\binom{6}{3}}{\binom{15}{3}}=\frac{4+10+20}{455}=.0747\ .$$

c. P(one of each type is selected) $=\dfrac{\binom{4}{1}\binom{5}{1}\binom{6}{1}}{\binom{15}{3}}=\dfrac{120}{455}=.2637$.

d. It is necessary to examine at least six bulbs if and only if the first five light bulbs were all of the 40W or 60W variety. Since there are 9 such bulbs, the chance of this event is

$$\frac{\binom{9}{5}}{\binom{15}{5}}=\frac{126}{3003}=.042\ .$$

45.

a. If the A's were distinguishable from one another, and similarly for the B's, C's and D's, then there would be 12! possible chain molecules. Six of these are:

$A_1A_2A_3B_2C_3C_1D_3C_2D_1D_2B_3B_1$ $\quad A_1A_3A_2B_2C_3C_1D_3C_2D_1D_2B_3B_1$
$A_2A_1A_3B_2C_3C_1D_3C_2D_1D_2B_3B_1$ $\quad A_2A_3A_1B_2C_3C_1D_3C_2D_1D_2B_3B_1$
$A_3A_1A_2B_2C_3C_1D_3C_2D_1D_2B_3B_1$ $\quad A_3A_2A_1B_2C_3C_1D_3C_2D_1D_2B_3B_1$

These 6 (=3!) differ only with respect to ordering of the 3 A's. In general, groups of 6 chain molecules can be created such that within each group only the ordering of the A's is different. When the A subscripts are suppressed, each group of 6 "collapses" into a single molecule (B's, C's and D's are still distinguishable).

At this point there are (12!/3!) different molecules. Now suppressing subscripts on the B's, C's, and D's in turn gives $\frac{12!}{(3!)^4} = 369,600$ chain molecules.

b. Think of the group of 3 A's as a single entity, and similarly for the B's, C's, and D's. Then there are 4! = 24 ways to order these triplets, and thus 24 molecules in which the A's are contiguous, the B's, C's, and D's also. The desired probability is $\frac{24}{369,600} = .00006494$.

47. Label the seats <u>1</u> <u>2</u> <u>3</u> <u>4</u> <u>5</u> <u>6</u>. The probability Jim and Paula sit in the two seats to the far left is

P(J&P in 1&2) $= \frac{2\times1\times4\times3\times2\times1}{6\times5\times4\times3\times2\times1} = \frac{1}{15}$.

Similarly, P(J&P next to each other) = P(J&P in 1&2) + … + P(J&P in 5&6) = $5\times\frac{1}{15} = \frac{1}{3}$.

Third, P(at least one H next to his W) = 1 – P(no H next to his W), and we count the number of ways of no H sits next to his W as follows:

# of orderings with a H-W pair in seats #1 and 3 and no H next to his W = $6^* \times 4 \times 1^* \times 2^{\#} \times 1 \times 1 = 48$

*= pair, # =can't put the mate of seat #2 here or else a H-W pair would be in #5 and 6

# of orderings without a H-W pair in seats #1 and 3, and no H next to his W = $6 \times 4 \times 2^{\#} \times 2 \times 2 \times 1 = 192$

#= can't be mate of person in seat #1 or #2

So, the number of seating arrangements with no H next to W = 48 + 192 = 240, and

P(no H next to his W) = $= \frac{240}{6\times5\times4\times3\times2\times1} = \frac{1}{3}$. Therefore, P(at least one H next to his W) = $1 - \frac{1}{3} = \frac{2}{3}$.

49. $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!k!} = \binom{n}{n-k}$

The number of subsets of size k equals the number of subsets of size $n - k$, because to each subset of size k there corresponds exactly one subset of size $n - k$: the $n - k$ objects not in the subset of size k. The combinations formula counts the number of ways to split n objects into two subsets: one of size k, and one of size $n - k$.

Section 1.4

51. Let A be that the individual is more than 6 feet tall. Let B be that the individual is a professional basketball player. Then $P(A|B)$ = the probability of the individual being more than 6 feet tall, knowing that the individual is a professional basketball player, while $P(B|A)$ = the probability of the individual being a professional basketball player, knowing that the individual is more than 6 feet tall. $P(A|B)$ will be larger. Most professional basketball players are tall, so the probability of an individual in that reduced sample space being more than 6 feet tall is very large. On the other hand, the number of individuals that are pro basketball players is small in relation to the number of males more than 6 feet tall.

53.

a. $P(A_2 \mid A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)} = \frac{.06}{.12}$ = .50. The numerator comes from Exercise 28.

b. $P(A_1 \cap A_2 \cap A_3 \mid A_1) = \frac{P([A_1 \cap A_2 \cap A_3] \cap A_1)}{P(A_1)} = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1)} = \frac{.01}{.12}$ = .0833. The numerator simplifies because $A_1 \cap A_2 \cap A_3$ is a subset of A_1, so their intersection is just the smaller event.

c. For this example, you definitely need a Venn diagram. The seven pieces of the partition inside the three circles have probabilities .04, .05, .00, .02, .01, .01, and .01. Those add to .14 (so the chance of no defects is .86).
Let E = "exactly one defect." From the Venn diagram, $P(E)$ = .04 + .00 + .01 = .05. From the addition above, P(at least one defect) = $P(A_1 \cup A_2 \cup A_3)$ = .14. Finally, the answer to the question is
$P(E \mid A_1 \cup A_2 \cup A_3) = \frac{P(E \cap [A_1 \cup A_2 \cup A_3])}{P(A_1 \cup A_2 \cup A_3)} = \frac{P(E)}{P(A_1 \cup A_2 \cup A_3)} = \frac{.05}{.14}$ = .3571. The numerator simplifies because E is a subset of $A_1 \cup A_2 \cup A_3$.

d. $P(A_3' \mid A_1 \cap A_2) = \frac{P(A_3' \cap [A_1 \cap A_2])}{P(A_1 \cap A_2)} = \frac{.05}{.06}$ = .8333. The numerator is Exercise 28(c), while the denominator is Exercise 28(b).

55.

a. $P(\text{M} \cap \text{LS} \cap \text{PR}) = .05$, directly from the table of probabilities.

b. $P(\text{M} \cap \text{Pr}) = P(\text{M} \cap \text{LS} \cap \text{PR}) + P(\text{M} \cap \text{SS} \cap \text{PR}) = .05 + .07 = .12$.

c. $P(\text{SS})$ = sum of 9 probabilities in the SS table = .56. $P(\text{LS}) = 1 - .56 = .44$.

d. From the two tables, $P(\text{M}) = .08 + .07 + .12 + .10 + .05 + .07 = .49$. $P(\text{Pr}) = .02 + .07 + .07 + .02 + .05 + .02 = .25$.

e. $P(\text{M}|\text{SS} \cap \text{Pl}) = \dfrac{P(\text{M} \cap \text{SS} \cap \text{Pl})}{P(\text{SS} \cap \text{Pl})} = \dfrac{.08}{.04 + .08 + .03} = .533$.

f. $P(\text{SS}|\text{M} \cap \text{Pl}) = \dfrac{P(\text{SS} \cap \text{M} \cap \text{Pl})}{P(\text{M} \cap \text{Pl})} = \dfrac{.08}{.08 + .10} = .444$. $P(\text{LS}|\text{M} \cap \text{Pl}) = 1 - P(\text{SS}|\text{M} \cap \text{Pl}) = 1 - .444 = .556$.

57. We know that $P(A_1 \cup A_2) = .07$ and $P(A_1 \cap A_2) = .01$, and that $P(A_1) = P(A_2)$ because the pumps are identical. There are two solution methods. The first doesn't require explicit reference to q or r: Let A_1 be the event that #1 fails and A_2 be the event that #2 fails.
Apply the addition rule: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \Rightarrow .07 = 2P(A_1) - .01 \Rightarrow P(A_1) = .04$.

Otherwise, we assume that $P(A_1) = P(A_2) = q$ and that $P(A_1 \mid A_2) = P(A_2 \mid A_1) = r$ (the goal is to find q). Proceed as follows: $.01 = P(A_1 \cap A_2) = P(A_1)\, P(A_2 \mid A_1) = qr$ and $.07 = P(A_1 \cup A_2) =$
$P(A_1 \cap A_2) + P(A_1' \cap A_2) + P(A_1 \cap A_2') = .01 + q(1 - r) + q(1 - r) \Rightarrow q(1 - r) = .03$.
These two equations give $2q - .01 = .07$, from which $q = .04$ (and $r = .25$).

59.

a. $P(A_2 \mid A_1) = \dfrac{P(A_1 \cap A_2)}{P(A_1)} = \dfrac{.11}{.22} = .50$. If the firm is awarded project 1, there is a 50% chance they will also be awarded project 2.

b. $P(A_2 \cap A_3 \mid A_1) = \dfrac{P(A_1 \cap A_2 \cap A_3)}{P(A_1)} = \dfrac{.01}{.22} = .0455$. If the firm is awarded project 1, there is a 4.55% chance they will also be awarded projects 2 and 3.

c. $P(A_2 \cup A_3 \mid A_1) = \dfrac{P[A_1 \cap (A_2 \cup A_3)]}{P(A_1)} = \dfrac{P[(A_1 \cap A_2) \cup (A_1 \cap A_3)]}{P(A_1)}$
$= \dfrac{P(A_1 \cap A_2) + P(A_1 \cap A_3) - P(A_1 \cap A_2 \cap A_3)}{P(A_1)} = \dfrac{.15}{.22} = .682$. If the firm is awarded project 1, there is a 68.2% chance they will also be awarded at least one of the other two projects.

d. $P(A_1 \cap A_2 \cap A_3 \mid A_1 \cup A_2 \cup A_3) = \dfrac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cup A_2 \cup A_3)} = \dfrac{.01}{.53} = .0189$. If the firm is awarded at least one of the projects, there is a 1.89% chance they will be awarded all three projects.

61. $$P(A \mid B) + P(A' \mid B) = \frac{P(A \cap B)}{P(B)} + \frac{P(A' \cap B)}{P(B)} = \frac{P(A \cap B) + P(A' \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

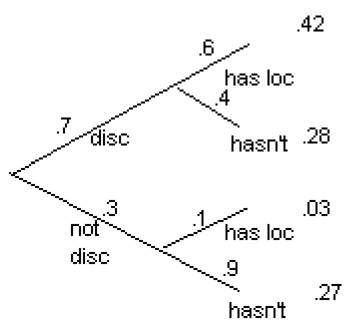
63. $$P(A \cup B \mid C) = \frac{P[(A \cup B) \cap C)}{P(C)} = \frac{P[(A \cap C) \cup (B \cap C)]}{P(C)} = \frac{P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)}{P(C)}$$
$= P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$

65. The sample space of the experiment is $\mathcal{S} = \{BB, Bb, bB, bb\}$, with all four outcomes equally likely.

a. Let A denote the event that the offspring has black fur. Then $P(A) = P(BB, Bb, bB) = \frac{3}{4} = .75$.

b. "Genotype *Bb*" really means *Bb* or *bB*. Given the hamster has black fur, the probability its genotype is *Bb* equals $P(Bb \text{ or } bB \mid A) = \frac{P(\{Bb \text{ or } bB\} \cap A)}{P(A)} = \frac{P(Bb \text{ or } bB)}{P(A)} = \frac{2/4}{3/4} = \frac{2}{3}$.

67. The tree diagram below shows the probability for the four disjoint options; e.g., *P*(the flight is discovered and has a locator) = *P*(discovered)*P*(locator | discovered) = (.7)(.6) = .42.



a. $P(\text{not discovered} \mid \text{has locator}) = \frac{P(\text{not discovered} \cap \text{has locator})}{P(\text{has locator})} = \frac{.03}{.03 + .42} = .067$.

b. $P(\text{discovered} \mid \text{no locator}) = \frac{P(\text{discovered} \cap \text{no locator})}{P(\text{no locator})} = \frac{.28}{.55} = .509$.

69. First, use the definition of conditional probability and the associative property of intersection:
$$P(A \cap B \mid C) = \frac{P((A \cap B) \cap C)}{P(C)} = \frac{P(A \cap (B \cap C))}{P(C)}$$
Second, use the Multiplication Rule to re-write the numerator:
$$\frac{P(A \cap (B \cap C))}{P(C)} = \frac{P(B \cap C)P(A \mid B \cap C)}{P(C)}$$
Finally, by definition, the ratio $\frac{P(B \cap C)}{P(C)}$ equals $P(B \mid C)$.

Substitution gives $P(A \cap B \mid C) = P(B \mid C) \cdot P(A \mid B \cap C)$, QED.

71. First, partition the sample space into statisticians with both life and major medical insurance, just life insurance, just major medical insurance, and neither. We know that P(both) = .20; subtracting them out, P(life only) = P(life) – P(both) = .75 – .20 = .55; similarly, P(medical only) = P(medical) – P(both) = .45 – .20 = .25.

a. Apply the Law of Total Probability:
P(renew) = P(life only)P(renew | life only) + P(medical only)P(renew | medical only) + P(both)P(renew | both)
= (.55)(.70) + (.25)(.80) + (.20)(.90) = .765.

b. Apply Bayes' Rule: $P(\text{both} \mid \text{renew}) = \dfrac{P(\text{both})P(\text{renew} \mid \text{both})}{P(\text{renew})} = \dfrac{(.20)(.90)}{.765} = .2353.$

73. A tree diagram can help. We know that P(day) = .2, P(1-night) = .5, P(2-night) = .3; also, P(purchase | day) = .1, P(purchase | 1-night) = .3, and P(purchase | 2-night) = .2.

Apply Bayes' rule: e.g., $P(\text{day} \mid \text{purchase}) = \dfrac{P(\text{day} \cap \text{purchase})}{P(\text{purchase})} = \dfrac{(.2)(.1)}{(.2)(.1)+(.5)(.3)+(.3)(.2)} = \dfrac{.02}{.23} = .087.$

Similarly, $P(\text{1-night} \mid \text{purchase}) = \dfrac{(.5)(.3)}{.23} = .652$ and P(2-night | purchase) = .261.

75. Let T denote the event that a randomly selected person is, in fact, a terrorist. Apply Bayes' theorem, using $P(T)$ = 1,000/300,000,000 = .0000033:

$P(T \mid +) = \dfrac{P(T)P(+ \mid T)}{P(T)P(+ \mid T)+P(T')P(+ \mid T')} = \dfrac{(.0000033)(.99)}{(.0000033)(.99)+(1-.0000033)(1-.999)} = .003289.$ That is to say, roughly 0.3% of all people "flagged" as terrorists would be actual terrorists in this scenario.

77. Let's see how we can implement the hint. If she's flying airline #1, the chance of 2 late flights is (30%)(10%) = 3%; the two flights being "unaffected" by each other means we can multiply their probabilities. Similarly, the chance of 0 late flights on airline #1 is (70%)(90%) = 63%. Since percents add to 100%, the chance of exactly 1 late flight on airline #1 is 100% – (3% + 63%) = 34%. A similar approach works for the other two airlines: the probability of exactly 1 late flight on airline #2 is 35%, and the chance of exactly 1 late flight on airline #3 is 45%.

The initial ("prior") probabilities for the three airlines are $P(A_1)$ = 50%, $P(A_2)$ = 30%, and $P(A_3)$ = 20%. Given that she had exactly 1 late flight (call that event B), the conditional ("posterior") probabilities of the three airlines can be calculated using Bayes' Rule:

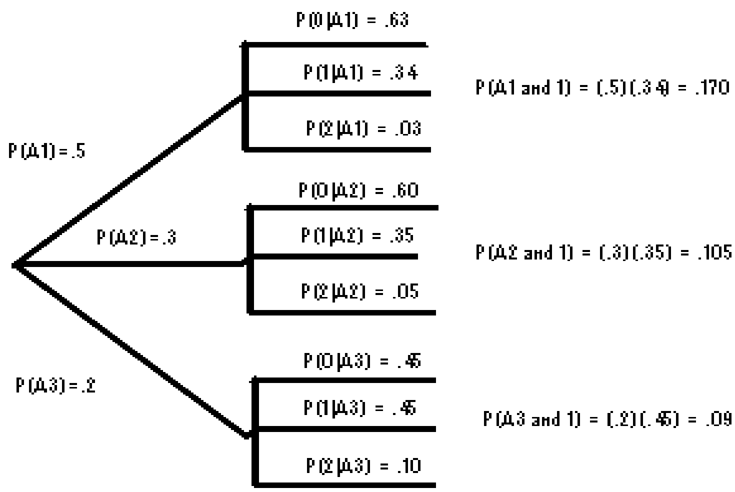
$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1)+P(A_2)P(B \mid A_2)+P(A_3)P(B \mid A_3)} = \frac{(.5)(.34)}{(.5)(.34)+(.3)(.35)+(.2)(.45)} = \frac{.170}{.365} = .4657;$$

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{P(A_1)P(B \mid A_1)+P(A_2)P(B \mid A_2)+P(A_3)P(B \mid A_3)} = \frac{(.3)(.35)}{.365} = .2877; \text{ and}$$

$$P(A_3 \mid B) = \frac{P(A_3)P(B \mid A_3)}{P(A_1)P(B \mid A_1)+P(A_2)P(B \mid A_2)+P(A_3)P(B \mid A_3)} = \frac{(.2)(.45)}{.365} = .2466.$$

Notice that, except for rounding error, these three posterior probabilities add to 1.

The tree diagram below shows these probabilities.



Section 1.5

79. Using the definition, two events A and B are independent if $P(A \mid B) = P(A)$;
$P(A \mid B) = .6125$; $P(A) = .50$; $.6125 \neq .50$, so A and B are not independent.
Using the multiplication rule, the events are independent if $P(A \cap B)=P(A)P(B)$;
$P(A \cap B) = .25$; $P(A)P(B) = (.5)(.4) = .2$. $.25 \neq .2$, so A and B are not independent.

81. $P(A_1 \cap A_2) = .11$ while $P(A_1)P(A_2) = .055$, so A_1 and A_2 are not independent.
$P(A_1 \cap A_3) = .05$ while $P(A_1)P(A_3) = .0616$, so A_1 and A_3 are not independent.
$P(A_2 \cap A_3) = .07$ and $P(A_2)P(A_3) = .07$, so A_2 and A_3 <u>are</u> independent.

83. Using subscripts to differentiate between the selected individuals,
$P(\text{O}_1 \cap \text{O}_2) = P(\text{O}_1)P(\text{O}_2) = (.45)(.45) = .2025$.
$P(\text{two individuals match}) = P(\text{A}_1 \cap \text{A}_2) + P(\text{B}_1 \cap \text{B}_2) + P(\text{AB}_1 \cap \text{AB}_2) + P(\text{O}_1 \cap \text{O}_2) =$
$.40^2 + .11^2 + .04^2 + .45^2 = .3762$.

85. Follow the same logic as in the previous exercise. With $p = 1/9{,}000{,}000{,}000$ and $n = 1{,}000{,}000{,}000$, the probability of at least one error is $1 - (1 - p)^n \approx 1 - .8948 = .1052$.

Note: For extremely small values of p, $(1 - p)^n \approx 1 - np$. So, the probability of at least one occurrence under these assumptions is roughly $1 - (1 - np) = np$. Here, that would equal $1/9 \approx .1111$.

87. $P(\text{at least one opens}) = 1 - P(\text{none open}) = 1 - (.05)^5 = .99999969$.
$P(\text{at least one fails to open}) = 1 - P(\text{all open}) = 1 - (.95)^5 = .2262$.

89. Let A_i denote the event that component #i works (i = 1, 2, 3, 4). Based on the design of the system, the event "the system works" is $(A_1 \cup A_2) \cup (A_3 \cap A_4)$. We'll eventually need $P(A_1 \cup A_2)$, so work that out first: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = (.9) + (.9) - (.9)(.9) = .99$. The third term uses independence of events. Also, $P(A_3 \cap A_4)$ = (.9)(.9) = .81, again using independence.

Now use the addition rule and independence for the system:

$$\begin{aligned} P((A_1 \cup A_2) \cup (A_3 \cap A_4)) &= P(A_1 \cup A_2) + P(A_3 \cap A_4) - P((A_1 \cup A_2) \cap (A_3 \cap A_4)) \\ &= P(A_1 \cup A_2) + P(A_3 \cap A_4) - P(A_1 \cup A_2) \times P(A_3 \cap A_4) \\ &= (.99) + (.81) - (.99)(.81) = .9981 \end{aligned}$$

(You could also use deMorgan's law in a couple of places.)

91. A = {(3,1)(3,2)(3,3)(3,4)(3,5)(3,6)} $\Rightarrow P(A) = \frac{6}{36} = \frac{1}{6}$; B = {(1,4)(2,4)(3,4)(4,4)(5,4)(6,4)} $\Rightarrow$ $P(B) = \frac{1}{6}$; and C = {(1,6)(2,5)(3,4)(4,3)(5,2)(6,1)} $\Rightarrow P(C) = \frac{1}{6}$.

a. $A \cap B$ = {(3,4)} $\Rightarrow P(A \cap B) = \frac{1}{36} = P(A)P(B)$; $A \cap C$ = {(3,4)} $\Rightarrow P(A \cap C) = \frac{1}{36} = P(A)P(C)$; and $B \cap C$ = {(3,4)} $\Rightarrow P(B \cap C) = \frac{1}{36} = P(B)P(C)$. Therefore, these three events are pairwise independent.

b. However, $A \cap B \cap C$ = {(3,4)} $\Rightarrow P(A \cap B \cap C) = \frac{1}{36}$, while $P(A)P(B)P(C) = \; = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}$, so $P(A \cap B \cap C) \neq P(A)P(B)P(C)$ and these three events are not mutually independent.

93. Let A_i denote the event that vehicle #i passes inspection (i = 1, 2, 3).

a. $P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2) \times P(A_3) = (.7)(.7)(.7) = (.7)^3$ = .343.

b. This is the complement of part **a**, so the answer is 1 – .343 = .657.

c. $P([A_1 \cap A_2' \cap A_3'] \cup [A_1' \cap A_2 \cap A_3'] \cup [A_1' \cap A_2' \cap A_3])$ = (.7)(.3)(.3) + (.3)(.7)(.3) + (.3)(.3)(.7) = $3(.3)^2(.7)$ = .189. Notice that we're using the fact that if events are independent then their complements are also independent.

d. P(at most one passes) = P(zero pass) + P(exactly one passes) = P(zero pass) + .189. For the first probability, P(zero pass) = $P(A_1' \cap A_2' \cap A_3')$ = (.3)(.3)(.3) = .027. So, the answer is .027 + .189 = .216.

e. We'll need the fact that P(at least one passes) = 1 – P(zero pass) = 1 – .027 = .973. Then,

$$P(A_1 \cap A_2 \cap A_3 \mid A_1 \cup A_2 \cup A_3) = \frac{P([A_1 \cap A_2 \cap A_3] \cap [A_1 \cup A_2 \cup A_3])}{P(A_1 \cup A_2 \cup A_3)} = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cup A_2 \cup A_3)} = \frac{.343}{.973} = .3525.$$

95.

a. $P(A) = \dfrac{2,000}{10,000}$ = .2. Using the law of total probability, $P(B) = P(A)P(B \mid A) + P(A')P(B \mid A')$ = $(.2)\dfrac{1,999}{9,999} + (.8)\dfrac{2,000}{9,999}$ = **.2** exactly. That is, $P(B)$ = $P(A)$ = .2. Finally, use the multiplication rule:

$P(A \cap B) = P(A) \times P(B \mid A) = (.2)\dfrac{1,999}{9,999}$ = .039984. Events A and B are *not* independent, since $P(B)$ = .2 while $P(B \mid A) = \dfrac{1,999}{9,999} = .19992$, and these are not equal.

b. If A and B were independent, we'd have $P(A \cap B) = P(A) \times P(B) = (.2)(.2) = .04$. This is very close to the answer .039984 from part **a**. This suggests that, for most practical purposes, we could treat events A and B in this example as if they were independent.

c. Repeating the steps in part **a**, you again get $P(A) = P(B) = .2$. However, using the multiplication rule, $P(A \cap B) = P(A) \times P(B \mid A) = \dfrac{2}{10} \times \dfrac{1}{9}$ =.0222. This is very different from the value of .04 that we'd get if A and B were independent!

The critical difference is that the population size in parts **a-b** is huge, and so the probability a second board is green *almost* equals .2 (i.e., 1,999/9,999 = .19992 ≈ .2). But in part **c**, the conditional probability of a green board shifts a lot: 2/10 = .2, but 1/9 = .1111.

97.

a. For route #1, P(late) = P(stopped at 2 or 3 or 4 crossings) = 1 – P(stopped at 0 or 1) = 1 – $[.9^4 + 4(.9)^3(.1)]$ = .0523.
For route #2, P(late) = P(stopped at 1 or 2 crossings) = 1 – P(stopped at none) = 1 – .81 = .19.
Thus route #1 should be taken.

b. Apply Bayes' Rule: P(4-crossing route | late) = $\dfrac{P(\text{4-crossing} \cap \text{late})}{P(\text{late})} = \dfrac{(.5)(.0523)}{(.5)(.0523)+(.5)(.19)} = .216$.

99.

a. Similar to several previous exercises,
P(win at least once) = 1 – P(lose all n times) = 1 – $P(\text{lose})^n$ = 1 – $(1 – 1/N)^n$.

b. The probability of rolling at least one ⚃ in n rolls is $1 – (5/6)^n$, while the "n/N" answer is $n/6$. The correct answers are .4212 (n = 3), .6651 (n = 6), and .8385 (n = 10). The wrong answer 3/6 = .5 is not very accurate for n = 3, 6/6 = 1 is absurd for n = 6, and 10/6 for n = 10 is impossible since probability cannot exceed 1! Clearly, n/N is a terrible approximation to the correct probability.

c. From Exercise 85, the probability of at least one error was .1052, compared to $n/N = 10^9/(9\times10^9) = 1/9 \approx$ .1111. Those aren't too far off, so arguably n/N might not be a bad approximation to $1 – (1 – 1/N)^n$ when N is very large (equivalently, $p = 1/N$ is small).

d. The binomial theorem states that $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = a^n + na^{n-1}b + \binom{n}{2} a^{n-2}b^2 + \cdots$. Applying the binomial theorem with a = 1 and b = –1/N gives $(1-1/N)^n = 1^n + n1^{n-1}(-1/N) + \binom{n}{2} 1^{n-2}(-1/N)^2 + \cdots =$ $1 - \dfrac{n}{N} + \dfrac{n(n-1)}{2N^2} + \cdots$. But if $n << N$, then the third term in that expression and all those thereafter will be negligible compared to the first two terms. Thus, $(1 – 1/N)^n \approx 1 – n/N$ when N is very large.
It follows that the probability of at least one success in the given setting, whose exact probability equals $1 – (1 – 1/N)^n$, can be approximated by $1 – (1 – n/N) = n/N$ when N is very large.

Section 1.6

101.

a. Let A = exactly one of B_1 or B_2 occurs = $(B_1 \cap B_2') \cup (B_2 \cap B_1')$. The Matlab and R code below has been modified from Example 1.40 to count how often, out of 10,000 independent runs event A occurs.

```
A=0;
for i=1:10000
    u1=rand; u2=rand;
    if (u1<.6 && u2>=.7)||
       (u1>=.6 && u2<.7)
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
      u1<-runif(1); u2<-runif(1)
      if((u1<.6 && u2>=.7)||
         (u1>=.6 && u2<.7)){
            A<-A+1
      }
}
```

Executing the Matlab code returned `A=4588`, so $\hat{P}(A) = \dfrac{4588}{10,000} = .4588$.

The exact probability is $P(A) = P(B_1 \cap B_2') + P(B_2 \cap B_1') = P(B_1) - P(B_1 \cap B_2) + P(B_2) - P(B_1 \cap B_2) = P(B_1) + P(B_2) - 2P(B_1 \cap B_2) = P(B_1) + P(B_2) - 2P(B_1)P(B_2) = .6 + .7 - 2(.6)(.7) = .46$.

Note: In both programs, the code `(u1<.6 && u2>=.7)||(u1>=.6 && u2<.7)` can be replaced by a single "exclusive or" command: `xor(u1<.6,u2<.7)`.

b. The estimated standard error of $\hat{P}(A)$ is $\sqrt{\dfrac{\hat{P}(A)[1-\hat{P}(A)]}{n}} = \sqrt{\dfrac{(.4588)(1-.4588)}{10,000}} \approx .00498$.

103. In the code below, seven random numbers are generated, one for each of the seven components. The sequence of and/or conjunctions matches the series and parallel ties in the system design.

```
A=0;
for i=1:10000
    u=rand(7,1);
    if (u(1)<.9 | u(2)<.9) &
       ((u(3)<.8 & u(4)<.8) |
       (u(5)<.8 & u(6)<.8)) &
       u(7)<.95
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
      u<-runif(7)
      if((u[1]<.9 | u[2]<.9) &
        ((u[3]<.8 & u[4]<.8) |
        (u[5]<.8 & u[6]<.8)) &
        u[7]<.95){
            A<-A+1
      }
}
```

Executing the Matlab code gave `A=8159`, so $\hat{P}(A) = \dfrac{8159}{10,000} = .8159$.

105. The programs below are written as functions, meaning they can receive inputs and generate outputs. Both programs take two inputs: n = the number of games to be simulated and p = the probability a contestant makes a correct guess. Each program outputs the estimated probability $\hat{P}$ of winning the game Now or Then.

```
function prob=nowthen(n,p)
win=0;
for i=1:n
    u=rand(6,1);
    x=(u<p);
    if x(1)+x(2)+x(3)==3 ||
       x(2)+x(3)+x(4)==3 ||
       x(3)+x(4)+x(5)==3 ||
       x(4)+x(5)+x(6)==3 ||
       x(5)+x(6)+x(1)==3 ||
       x(6)+x(1)+x(2)==3
        win=win+1;
    end
end
prob=win/n;
```

```
nowthen<-function(n,p){
win<-0
for(i in 1:n){
    u=runif(6);
    x=(u<p);
    if(x[1]+x[2]+x[3]==3 ||
       x[2]+x[3]+x[4]==3 ||
       x[3]+x[4]+x[5]==3 ||
       x[4]+x[5]+x[6]==3 ||
       x[5]+x[6]+x[1]==3 ||
       x[6]+x[1]+x[2]==3){
        win=win+1;
    }
}
return(win/n)
}
```

In Matlab, the above code is saved as a file **nowthen.m**; in R, the above code is executed at the command line. After this, you may call this function at the command line.

(1) Typing `nowthen(10000,.5)` at the command line gave .3993.
(2) Typing `nowthen(10000,.8)` at the command line gave .8763.

107. Modify the program from the previous exercise, as illustrated below. Of interest is whether the difference between the largest and smallest entries of the vector `dollar` is at least 5.

```
A=0;
for i=1:10000
    u=rand(25,1);
    flips=(u<.4)-(u>=.4);
    dollar=cumsum(flips);
    if max(dollar)-min(dollar)>=5
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
    u<-runif(25)
    flips<-(u<.4)-(u>=.4)
    dollar<-cumsum(flips)
    if(max(dollar)-min(dollar)>=5){
        A<-A+1
    }
}
```

Executing the Matlab code above gave `A=9189`, so $\hat{P}$ = .9189.

109. Divide the 40 questions into the four types. For the first type (two choices), the probability of correctly guessing the right answer is 1/2. Similarly, the probability of correctly guessing a three-choice question correctly is 1/3, and so on. In the programs below, four vectors contain random numbers for the four types of questions; the binary vectors `(u<1/2)`, `(v<1/3)`, and so on code right and wrong guesses with 1s and 0s, respectively. Thus, `right` represents the total number of correct guesses out of 40. A student gets at least half of the questions correct if that total is at least 20.

```
A=0;
for i=1:10000
    u=rand(10,1); v=rand(13,1);
    w=rand(13,1); x=rand(5,1);
    right=sum(u<1/2)+sum(v<1/3)
            +sum(w<1/4)+sum(x<1/5);
    if right>=20
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
    u<-runif(10); v<-runif(13)
    w<-runif(13); x<-runif(5)
    right<-sum(u<1/2)+sum(v<1/3)
             +sum(w<1/4)+sum(x<1/5)
    if(right>=20){
        A<-A+1
    }
}
```

Executing the code once gave `A=227`, so $\hat{P}$ = .0227.

111.

a. In the programs below, `test` is the vector [1 2 3 … 12]. A random permutation is generated and then compared to `test`. If <u>any</u> of the 12 numbers are in the right place, `match` will equal 1; otherwise, `match` equals 0 and we have a derangement. The scalar `D` counts the number of derangements in 10,000 simulations.

```
D=0;
test=1:12;
for i=1:10000
    permutation=randperm(12);
    match=any(permutation==test);
    if match==0
        D=D+1;
    end
end
```

```
D<-0
test<-1:12
for(i in 1:10000){
    permutation<-sample(test,12)
    match<-any(permutation==test)
    if(match==0){
        D<-D+1
    }
}
```

b. One execution of the code in **a** gave `D=3670`, so $\hat{P}(D)$ = .3670.

c. We know there are 12! possible permutations of the numbers 1 through 12. According to **b**, we estimate that 36.70% of them are derangements. This suggests that the estimated <u>number</u> of derangements of the numbers 1 through 12 is .3670(12!) = .3670(479,001,600) = 175,793,587.
(In fact, it is known that the exact number of such derangements is 176,214,841.)

113. The programs below keep a simultaneous record of whether the player wins the game and whether the game ends within 10 coin flips. These counts are stored in `win` and `ten`, respectively. The while loop insures that game play continues until the player has \$0 or \$100.

```
win=0; ten=0;
for i=1:10000
    money=20;numflips=0;
    while money>0 && money<100
        numflips=numflips+1;
        change=randsample([-10 10],1);
        money=money+change;
    end
    if money==100
        win=win+1;
    end
    if numflips<=10
        ten=ten+1;
    end
end
```

```
win<-0; ten<-0
for(i in 1:10000){
    money<-20; numflips<-0
    while(money>0 && money<100){
        numflips<-numflips+1
        change<-sample(c(-10,10),1)
        money<-money+change
    }
    if(money==100){
        win<-win+1
    }
    if(numflips<=10){
        ten<-ten+1
    }
}
```

a. One execution gave `win=2080`, so $\hat{P}$ (player wins) = .2080. (In fact, it can be shown using more sophisticated methods that the exact probability of winning in this scenario is .2, corresponding to the player starting with \$20 of a potential \$100 stake and the coin being fair.)

b. One execution gave `ten=5581`, so $\hat{P}$ (game ends within 10 coin flips) = .5581.

115.

a. Code appears below. One execution in Matlab gave `A=5224`, so $\hat{P}$ (at least one ⚅ in four rolls) = .5224. Using independence, it can be shown that the exact probability is $1 - (5/6)^4 = .5177$.

```
A=0;
for i=1:10000
    rolls=randsample(1:6,4,true);
    numsixes=sum(rolls==6);
    if numsixes>=1
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
    rolls<-sample(1:6,4,TRUE)
    numsixes<-sum(rolls==6)
    if(numsixes>=1){
        A<-A+1
    }
}
```

b. Code appears below. One execution in Matlab gave `A=4935`, so $\hat{P}$ (at least one ⚅⚅ in 24 rolls) = .4935. Using independence, it can be shown that the exact probability is $1 - (35/36)^{24} = .4914$. In particular, the probability in **a** is greater than 1/2, while the probability in **b** is less than 1/2. So, you should be willing to wager even money on seeing at least one ⚅ in 4 rolls of one die, but not on seeing at least one ⚅⚅ in 24 rolls of two dice.

```
A=0;
for i=1:10000
    die1=randsample(1:6,24,true);
    die2=randsample(1:6,24,true);
    numdblsixes=
        sum((die1==6)&(die2==6));
    if numdblsixes>=1
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
    die1<-sample(1:6,24,TRUE)
    die2<-sample(1:6,24,TRUE)
    numdblsixes<-
        sum((die1==6)&(die2==6))
    if(numdblsixes>=1){
        A<-A+1
    }
}
```

117. Let 1 represent a vote for candidate A and –1 a vote for candidate B. A randomization of the 12 A's and 8 B's can be achieved by sampling without replacement from a vector [1 … 1 –1 … –1] with 12 1's and 8 –1's. (Matlab will use the `randsample` command; R, `sample`.) To keep track of how far ahead candidate A stands as each vote is counted, employ the `cumsum` command. As long as A is ahead, the cumulative total will be positive; if A and B are ever tied, the cumulative sum will be 0; and a negative cumulative sum indicates that B has taken the lead. (Of course, the final cumulative sum will always be 4, signaling A's victory.)

```
A=0;
votes=[ones(12,1); -1*ones(8,1)];
for i=1:10000
    vote=randsample(votes,20);
    so_far=cumsum(vote);
    if all(so_far>0)
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
    die1<-sample(1:6,24,TRUE)
    die2<-sample(1:6,24,TRUE)
    numdblsixes<-
        sum((die1==6)&(die2==6))
    if(numdblsixes>=1){
        A<-A+1
    }
}
```

One execution of the code above returned `A=2013`, so $\hat{P}$ (candidate A leads throughout the count) = .2013.

119.

a. In the code below, the criterion $x^2 + y^2 \le 1$ determines whether (x, y) lies in the unit quarter-disk.

```
A=0;
for i=1:10000
    x=rand; y=rand;
    if x^2+y^2<=1
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
    x=runif(1); y=runif(1)
    if(x^2+y^2<=1){
        A<-A+1
    }
}
```

b. Since $P(A) = \pi/4$, it follows that $\pi = 4P(A) \approx 4\hat{P}(A)$. One run of the above program returned `A=7837`, which implies that $\hat{P}(A) = .7837$ and $\pi \approx 4(.7837) = 3.1348$.

(While this may seem like a silly application, since we know how to determine π to arbitrarily many decimal places, the idea behind it is critical to lots of modern applications. The technique presented here is a special case of the method called *Monte Carlo integration*.)

Supplementary Exercises

121.

a. $\binom{20}{3} = 1140$.

b. There are 19 other machinists to choose from, so the answer is $\binom{19}{3} = 969$.

c. There are 1140 total possible crews. Among them, the number that have <u>none</u> of the best 10 machinists is $\binom{10}{3} = 120$ (because you're choosing from the remaining 10). So, the number of crews having at least one of the best 10 machinists is 1140 – 120 = 1020.

d. Using parts **a** and **b**, $P(\text{best will not work}) = \frac{969}{1140} = .85$.

123.

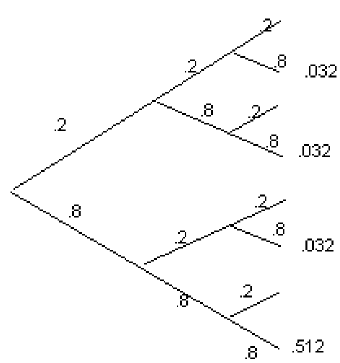
a. He will have one type of form left if either 4 withdrawals or 4 course substitutions remain. This means the first six were either 2 withdrawals and 4 subs or 6 withdrawals and 0 subs; the desired probability is
$$\frac{\binom{6}{2}\binom{4}{4}+\binom{6}{6}\binom{4}{0}}{\binom{10}{6}} = \frac{16}{210} = .0762.$$

b. He can start with the withdrawal forms or the course substitution forms, allowing two sequences: W-C-W-C or C-W-C-W. The number of ways the first sequence could arise is (6)(4)(5)(3) = 360, and the number of ways the second sequence could arise is (4)(6)(3)(5) = 360, for a total of 720 such possibilities. The <u>total</u> number of ways he could select four forms one at a time is ${}_{10}P_4 = (10)(9)(8)(7) = 5040$. So, the probability of a perfectly alternating sequence is 720/5040 = .143.

125. The probability of a bit reversal is .2, so the probability of maintaining a bit is .8.

a. Using independence, $P(\text{all three relays correctly send 1}) = (.8)(.8)(.8) = .512$.

b. In the accompanying tree diagram, each .2 indicates a bit reversal (and each .8 its opposite). There are several paths that maintain the original bit: no reversals or exactly two reversals (e.g., $1 \to 1 \to 0 \to 1$, which has reversals at relays 2 and 3). The total probability of these options is .512 + (.8)(.2)(.2) + (.2)(.8)(.2) + (.2)(.2)(.8) = .512 + 3(.032) = .608.



c. Using the answer from **b**, $P(\text{1 sent} \mid \text{1 received}) = \dfrac{P(\text{1 sent} \cap \text{1 received})}{P(\text{1 received})} =$

$$\frac{P(\text{1 sent})P(\text{1 received} \mid \text{1 sent})}{P(\text{1 sent})P(\text{1 received} \mid \text{1 sent}) + P(\text{0 sent})P(\text{1 received} \mid \text{0 sent})} = \frac{(.7)(.608)}{(.7)(.608)+(.3)(.392)} = \frac{.4256}{.5432} = .7835.$$

In the denominator, $P(\text{1 received} \mid \text{0 sent}) = 1 - P(\text{0 received} \mid \text{0 sent}) = 1 - .608$, since the answer from **b** also applies to a 0 being relayed as a 0.

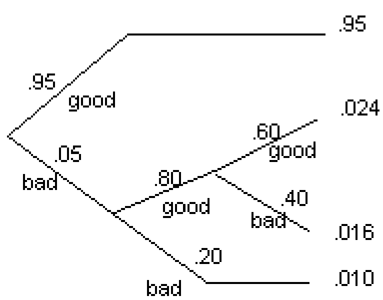
127. $P(\text{F hasn't heard after 10 times}) = P(\text{not on \#1} \cap \text{not on \#2} \cap \ldots \cap \text{ not on \#10}) = \frac{4}{5} \times \cdots \times \frac{4}{5} = \left(\frac{4}{5}\right)^{10} = .1074.$

129.

a. A sonnet has 14 lines, each of which may come from any of the 10 pages. Order matters, and we're sampling with replacement, so the number of possibilities is $10 \times 10 \times \ldots \times 10 = 10^{14}$.

b. First, consider sonnets made from only the first two pages. Each line has 2 possible choices (page 1 or page 2), so the number of sonnets you can create with the first two pages is $2 \times \ldots \times 2 = 2^{14}$. But among those, two employ just a single page: the sonnet that appears entirely on page 1, and the one entirely on page 2. So, the number of sonnets using exactly page 1 and page 2 (both pages, not no others) is $2^{14} - 2$. Next, there are ${}_{10}C_2 = 45$ pairs of pages you could choose to form the sonnet (1&2, 1&3, …, 9&10). Therefore, the total number of sonnets using exactly two pages is $45 \cdot [2^{14} - 2]$, and the probability of randomly selecting such a sonnet is $\dfrac{45 \cdot [2^{14} - 2]}{10^{14}} = \dfrac{737190}{10^{14}} = 7.3719 \times 10^{-9}$.

131. Refer to the tree diagram below.



a. $P(\text{pass inspection}) = P(\text{pass initially} \cup \text{passes after recrimping}) =$
$P(\text{pass initially}) + P(\text{fails initially} \cap \text{goes to recrimping} \cap \text{is corrected after recrimping}) =$
.95 + (.05)(.80)(.60) (following path "bad-good-good" on tree diagram) = .974.

b. $P(\text{needed no recrimping} \mid \text{passed inspection}) = \dfrac{P(\text{passed initially})}{P(\text{passed inspection})} = \dfrac{.95}{.974} = .9754$.

133. Let A = 1st functions, B = 2nd functions, so $P(B) = .9$, $P(A \cup B) = .96$, $P(A \cap B)=.75$. Use the addition rule:
$P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow .96 = P(A) + .9 - .75 \Rightarrow P(A) = .81$.

Therefore, $P(B \mid A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{.75}{.81} = .926$.

135. A tree diagram can help here.

a. The law of total probability gives $P(L) = \sum P(E_i)P(L \mid E_i) =$ (.40)(.02) + (.50)(.01) + (.10)(.05) = .018.

b. $P(E_1' \mid L') = 1 - P(E_1 \mid L') = 1 - \dfrac{P(E_1 \cap L')}{P(L')} = 1 - \dfrac{P(E_1)P(L' \mid E_1)}{1 - P(L)} = 1 - \dfrac{(.40)(.98)}{1 - .018} = .601.$

137. This question is a natural extension of the Birthday Problem. First, from the solution to the Birthday Problem, the likelihood that a birthday is shared by at least two among 10 people is

$$P(\text{birthday coincidence}) = 1 - \frac{365 \cdot 364 \cdots 356}{365 \cdot 365 \cdots 365} \text{ or } \frac{{}_{365}P_{10}}{365^{10}} = .117.$$

Second, there are 1000 possible 3-digit sequences to end a Social Security number (000 through 999). Using the same idea, the probability that at least two people in 10 share the last 3 digits of their SS number is

$$P(\text{SS coincidence}) = 1 - \frac{{}_{1000}P_{10}}{1000^{10}} = 1 - .956 = .044.$$

Assuming birthdays and SS endings are independent,

P(at least one coincidence) = P(birthday coincidence $\cup$ SS coincidence) = .117 + .044 – (.117)(.044) = .156.

139.

a. P(walks on 4th pitch) = P(first 4 pitches are balls) = $(.5)^4$ = .0625.

b. P(walks on 6th pitch) = P(2 of the first 5 are strikes $\cap$ #6 is a ball) =

P(2 of the first 5 are strikes)P(#6 is a ball) = $\binom{5}{2}(.5)^2(.5)^3(.5)$ = .15625.

c. Following the pattern from **b**, P(walks on 5th pitch) = $\binom{4}{1}(.5)^1(.5)^3(.5)$ = .125. Therefore, P(batter walks) = P(walks on 4th) + P(walks on 5th) + P(walks on 6th) = .0625 + .125 + .15625 = .34375.

d. P(first batter scores while no one is out) = P(first four batters all walk) = $(.34375)^4$ = .014.

141.

a. P(all full) = $P(A \cap B \cap C)$ = (.6)(.5)(.4) = .12.
P(at least one isn't full) = 1 – P(all full) = 1 – .12 = .88.

b. P(only NY is full) = $P(A \cap B' \cap C') = P(A)P(B')P(C')$ = (.6)(1–.5)(1–.4) = .18.
Similarly, P(only Atlanta is full) = .12 and P(only LA is full) = .08.
So, P(exactly one full) = .18 + .12 + .08 = .38.

143. $P(A_1) = P(\text{draw slip 1 or 4}) = ½$; $P(A_2) = P(\text{draw slip 2 or 4}) = ½$;
$P(A_3) = P(\text{draw slip 3 or 4}) = ½$; $P(A_1 \cap A_2) = P(\text{draw slip 4}) = ¼$;
$P(A_2 \cap A_3) = P(\text{draw slip 4}) = ¼$; $P(A_1 \cap A_3) = P(\text{draw slip 4}) = ¼$.

Hence, $P(A_1 \cap A_2) = ¼ = P(A_1)P(A_2)$, $P(A_2 \cap A_3) = ¼ = P(A_2)P(A_3)$, and $P(A_1 \cap A_3) = ¼ = P(A_1)P(A_3)$, thus there exists pairwise independence. However, $P(A_1 \cap A_2 \cap A_3) = P(\text{draw slip 4}) = ¼ \neq ⅛ = P(A_1)P(A_2)P(A_3)$, so the events are not mutually independent.

145.

a. Obviously, $a_0 = 0$ and $a_5 = 1$. Allan can't win without money, and he has won if he has all 5 dollars.

b. By the law of total probability,
$a_2 = P(\text{first flip is H})P(\text{Allan wins with \$3}) + P(\text{first flip is T})P(\text{Allan wins /w/ \$1}) = .5a_3 + .5a_1$.

c. The same logic yields four equations:
$a_1 = .5a_2 + .5a_0$, $a_2 = .5a_3 + .5a_1$, $a_3 = .5a_4 + .5a_2$, and $a_4 = .5a_5 + .5a_3$.
Rewrite the last equation as $2a_4 = a_5 + a_3$, then $a_4 - a_3 = a_5 - a_4$. Rewrite the others similarly, so we see the increment between successive a_i's is the same. Clearly then, with the initial conditions from **a**,
$a_1 = 1/5$, $a_2 = 2/5$, $a_3 = 3/5$, and $a_4 = 4/5$.
(You can also solve this system algebraically, as suggested in the problem.)

d. Following the pattern suggested in **c**, the chance Allan beats Beth with their initial fortunes being \$$a$ and \$$b$, respectively, is $\dfrac{a}{a+b}$.

147.

a. A attracts $B \Rightarrow P(B \mid A) > P(B) \Rightarrow 1 - P(B \mid A) < 1 - P(B)$, because multiplying by –1 reverses the direction of the inequality $\Rightarrow P(B' \mid A) < P(B') \Rightarrow$ by definition, A repels B'. In other words, if the occurrence of A makes B more likely, then it must make B' less likely. Notice this is really an iff statement; i.e., all of the implication arrows can be reversed.

b. This one is much trickier, since the complementation idea in **a** can't be applied here (i.e., to the conditional event A). One approach is as follows, which uses the fact that $P(B) - P(B \cap A) = P(B \cap A')$:

$$A \text{ attracts } B \Rightarrow P(B \mid A) > P(B) \Rightarrow \frac{P(A \cap B)}{P(A)} > P(B) \Rightarrow P(A \cap B) > P(A)P(B) \Rightarrow$$

$$P(B) - P(A \cap B) < P(B) - P(A)P(B) \qquad \text{because multiplying by –1 is order-reversing} \quad \Rightarrow$$

$$P(B \cap A') < P(B)[1 - P(A)] = P(B)P(A') \Rightarrow \frac{P(B \cap A')}{P(A')} < P(B) \Rightarrow P(B \mid A') < P(B) \Rightarrow$$

by definition, A' repels B. (Whew!) Notice again this is really an iff statement.

c. Apply the simplest version of Bayes' rule:

$$A \text{ attracts } B \Leftrightarrow P(B \mid A) > P(B) \Leftrightarrow \frac{P(B)P(A \mid B)}{P(A)} > P(B) \Leftrightarrow \frac{P(A \mid B)}{P(A)} > 1 \Leftrightarrow P(A \mid B) > P(A) \Leftrightarrow$$

by definition, B attracts A.

149.

a. First, the probabilities of the A_i are $P(A_1) = P(JJ) = (.6)^2 = .36$; $P(A_2) = P(MM) = (.4)^2 = .16$; and $P(A_3) = P(JM \text{ or } MJ) = (.6)(.4) + (.4)(.6) = .48$.

Second, $P(\text{Jay wins} \mid A_1) = 1$, since Jay is two points ahead and, thus has won; $P(\text{Jay wins} \mid A_2) = 0$, since Maurice is two points ahead and, thus, Jay has lost; and $P(\text{Jay wins} \mid A_3) = p$, since at that point the score has returned to deuce and the game has effectively started over. Apply the law of total probability:

$$P(\text{Jay wins}) = P(A_1)P(\text{Jay wins} \mid A_1) + P(A_2)P(\text{Jay wins} \mid A_2) + P(A_3)P(\text{Jay wins} \mid A_3)$$
$$p = (.36)(1) + (.16)(0) + (.48)(p)$$

Therefore, $p = .36 + .48p$; solving for p gives $p = \dfrac{.36}{1-.48} = .6923$.

b. Apply Bayes' rule: $P(JJ \mid \text{Jay wins}) = \dfrac{P(JJ)P(\text{Jay wins} \mid JJ)}{P(\text{Jay wins})} = \dfrac{(.36)(1)}{.6923} = .52$.

CHAPTER 2

Section 2.1

1.

$\mathscr{S}$:	*FFF*	*SFF*	*FSF*	*FFS*	*FSS*	*SFS*	*SSF*	*SSS*
X:	0	1	1	1	2	2	2	3

3. Examples include: M = the difference between the large and the smaller outcome with possible values 0, 1, 2, 3, 4, or 5; T = 1 if the sum of the two resulting numbers is even and T = 0 otherwise, a Bernoulli random variable. See the back of the book for other examples.

5. No. In the experiment in which a coin is tossed repeatedly until a H results, let Y = 1 if the experiment terminates with at most 5 tosses and Y = 0 otherwise. The sample space is infinite, yet Y has only two possible values. See the back of the book for another example.

7.

a. Possible values of X are 0, 1, 2, …, 12; discrete.

b. With n = # on the list, values of Y are 0, 1, 2, … , N; discrete.

c. Possible values of U are 1, 2, 3, 4, … ; discrete.

d. Possible values of X are $(0, \infty)$ if we assume that a rattlesnake can be arbitrarily short or long; not discrete.

e. With c = amount earned in royalties per book sold, possible values of Z are 0, c, $2c$, $3c$, … , $10{,}000c$; discrete.

f. Since 0 is the smallest possible pH and 14 is the largest possible pH, possible values of Y are [0, 14]; not discrete.

g. With m and M denoting the minimum and maximum possible tension, respectively, possible values of X are $[m, M]$; not discrete.

h. The number of possible tries is 1, 2, 3, …; each try involves 3 coins, so possible values of X are 3, 6, 9, 12, 15, …; discrete.

9.

a. Returns to 0 can occur only after an even number of tosses, so possible X values are 2, 4, 6, 8, …. Because the values of X are enumerable, X is discrete.

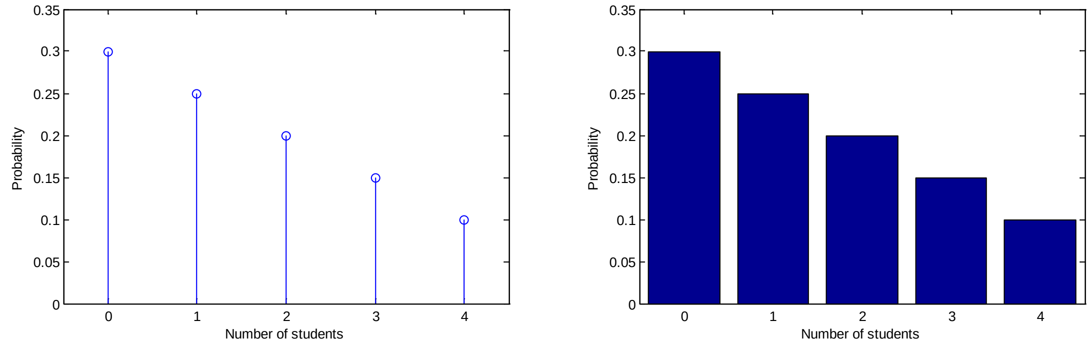
b. Now a return to 0 is possible after any number of tosses greater than 1, so possible values are 2, 3, 4, 5, …. Again, X is discrete.

Section 2.2

11.

a. The sum of the probabilities must be 1, so $p(4) = 1 - [p(0) + p(1) + p(2) + p(3)] = 1 - .90 = .10$.

b.



c. $P(X \geq 2) = p(2) + p(3) + p(4) = .20 + .15 + .10 = .45$. $P(X > 2) = p(3) + p(4) = .15 + .10 = .25$.

d. This was just a joke — professors should always show up for office hours!

13.

a. $P(X \leq 3) = p(0) + p(1) + p(2) + p(3) = .10+.15+.20+.25 = .70$.

b. $P(X < 3) = P(X \leq 2) = p(0) + p(1) + p(2) = .45$.

c. $P(X \geq 3) = p(3) + p(4) + p(5) + p(6) = .55$.

d. $P(2 \leq X \leq 5) = p(2) + p(3) + p(4) + p(5) = .71$.

e. The number of lines <u>not</u> in use is $6 - X$, and $P(2 \leq 6 - X \leq 4) = P(-4 \leq -X \leq -2) =$
$P(2 \leq X \leq 4) = p(2) + p(3) + p(4) = .65$.

f. $P(6 - X \geq 4) = P(X \leq 2) = .10 + .15 + .20 = .45$.

15.

a. (1,2) (1,3) (1,4) (1,5) (2,3) (2,4) (2,5) (3,4) (3,5) (4,5)

b. X can only take on the values 0, 1, 2. $p(0) = P(X = 0) = P(\{(3,4)\ (3,5)\ (4,5)\}) = 3/10 = .3$;
$p(2) = P(X = 2) = P(\{(1,2)\}) = 1/10 = .1$; $p(1) = P(X = 1) = 1 - [p(0) + p(2)] = .60$; and otherwise $p(x) = 0$.

c. $F(0) = P(X \leq 0) = P(X = 0) = .30$; $F(1) = P(X \leq 1) = P(X = 0 \text{ or } 1) = .30 + .60 = .90$;
$F(2) = P(X \leq 2) = 1$. Therefore, the complete cdf of X is

$$F(x) = \begin{cases} 0 & x < 0 \\ .30 & 0 \leq x < 1 \\ .90 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

17.

a. $p(2) = P(Y = 2) = P(\text{first 2 batteries are acceptable}) = P(AA) = (.9)(.9) = .81$.

b. $p(3) = P(Y = 3) = P(UAA \text{ or } AUA) = (.1)(.9)^2 + (.1)(.9)^2 = 2[(.1)(.9)^2] = .162$.

c. The fifth battery must be an A, and exactly one of the first four must also be an A.
Thus, $p(5) = P(AUUUA \text{ or } UAUUA \text{ or } UUAUA \text{ or } UUUAA) = 4[(.1)^3(.9)^2] = .00324$.

d. $p(y) = P(\text{the } y^{\text{th}} \text{ is an } A \text{ and so is exactly one of the first } y - 1) = (y - 1)(.1)^{y-2}(.9)^2$, for $y = 2, 3, 4, 5, \ldots$.

19. $p(0) = P(Y = 0) = P(\text{both arrive on Wed}) = (.3)(.3) = .09$;
$p(1) = P(Y = 1) = P((\text{W,Th}) \text{ or } (\text{Th,W}) \text{ or } (\text{Th,Th})) = (.3)(.4) + (.4)(.3) + (.4)(.4) = .40$;
$p(2) = P(Y = 2) = P((\text{W,F}) \text{ or } (\text{Th,F}) \text{ or } (\text{F,W}) \text{ or } (\text{F,Th}) \text{ or } (\text{F,F})) = .32$;
$p(3) = 1 - [.09 + .40 + .32] = .19$.

21.

a. First, $1 + 1/x > 1$ for all $x = 1, \ldots, 9$, so $\log(1 + 1/x) > 0$. Next, check that the probabilities sum to 1:
$\sum_{x=1}^{9} \log_{10}(1+1/x) = \sum_{x=1}^{9} \log_{10}\left(\frac{x+1}{x}\right) = \log_{10}\left(\frac{2}{1}\right) + \log_{10}\left(\frac{3}{2}\right) + \cdots + \log_{10}\left(\frac{10}{9}\right)$; using properties of logs, this equals $\log_{10}\left(\frac{2}{1} \times \frac{3}{2} \times \cdots \times \frac{10}{9}\right) = \log_{10}(10) = 1$.

b. Using the formula $p(x) = \log_{10}(1 + 1/x)$ gives the following values: $p(1) = .301$, $p(2) = .176$, $p(3) = .125$, $p(4) = .097$, $p(5) = .079$, $p(6) = .067$, $p(7) = .058$, $p(8) = .051$, $p(9) = .046$. The distribution specified by *B*enford's Law is <u>not</u> uniform on these nine digits; rather, lower digits (such as 1 and 2) are much more likely to be the lead digit of a number than higher digits (such as 8 and 9).

c. The jumps in $F(x)$ occur at $0, \ldots, 8$. We display the cumulative probabilities here: $F(1) = .301$, $F(2) = .477$, $F(3) = .602$, $F(4) = .699$, $F(5) = .778$, $F(6) = .845$, $F(7) = .903$, $F(8) = .954$, $F(9) = 1$. So, $F(x) = 0$ for $x < 1$; $F(x) = .301$ for $1 \le x < 2$; $F(x) = .477$ for $2 \le x < 3$; etc.

d. $P(X \le 3) = F(3) = .602$; $P(X \ge 5) = 1 - P(X < 5) = 1 - P(X \le 4) = 1 - F(4) = 1 - .699 = .301$.

23.

a. $p(2) = P(X = 2) = F(2) - F(1) = .39 - .19 = .20$.

b. $P(X > 3) = 1 - P(X \le 3) = 1 - F(3) = 1 - .67 = .33$.

c. $P(2 \le X \le 5) = F(5) - F(2\text{–}1) = F(5) - F(1) = .92 - .19 = .78$.

d. $P(2 < X < 5) = P(2 < X \le 4) = F(4) - F(2) = .92 - .39 = .53$.

25. $p(0) = P(Y = 0) = P(B \text{ first}) = p$;
$p(1) = P(Y = 1) = P(G \text{ first, then } B) = (1 - p)p$;
$p(2) = P(Y = 2) = P(GGB) = (1 - p)^2 p$;
Continuing, $p(y) = P(y\ G\text{s and then a } B) = (1 - p)^y p$ for $y = 0,1,2,3,\ldots$.

27.

a. The sample space consists of all possible permutations of the four numbers 1, 2, 3, 4:

outcome	*x* value	outcome	*x* value	outcome	*x* value
1234	4	2314	1	3412	0
1243	2	2341	0	3421	0
1324	2	2413	0	4132	1
1342	1	2431	1	4123	0
1423	1	3124	1	4213	1
1432	2	3142	0	4231	2
2134	2	3214	2	4312	0
2143	0	3241	1	4321	0

b. From the table in **a**, $p(0) = P(X = 0) = \frac{9}{24}$, $p(1) = P(X = 1) = \frac{8}{24}$, $p(2) = P(Y = 2) = \frac{6}{24}$, $p(3) = P(X = 3) = 0$, and $p(4) = P(Y = 4) = \frac{1}{24}$.

Section 2.3

29.

a. $E(X) = \sum_{\text{all } x} xp(x) = 1(.05) + 2(.10) + 4(.35) + 8(.40) + 16(.10) = 6.45$ GB. The average amount of memory across all purchased flash drives is 6.45 GB.

b. $\text{Var}(X) = \sum_{\text{all } x} (x-\mu)^2 p(x) = (1 - 6.45)^2(.05) + (2 - 6.45)^2(.10) + \ldots + (16 - 6.45)^2(.10) = 15.6475$.

c. $\text{SD}(X) = \sqrt{\text{Var}(X)} = \sqrt{15.6475} = 3.956$ GB. The amount of memory in a purchased flash drive typically differs from the average of 5.45 GB by roughly ±3.956 GB.

d. $E(X^2) = \sum_{\text{all } x} x^2 p(x) = 1^2(.05) + 2^2(.10) + 4^2(.35) + 8^2(.40) + 16^2(.10) = 57.25$. Using the shortcut formula, $\text{Var}(X) = E(X^2) - \mu^2 = 57.25 - (6.45)^2 = 15.6475$, the same answer as **b**.

31. From the table in Exercise 12, $E(Y) = 45(.05) + 46(.10) + \ldots + 55(.01) = 48.84$; similarly, $E(Y^2) = 45^2(.05) + 46^2(.10) + \ldots + 55^2(.01) = 2389.84$; thus $\text{Var}(Y) = E(Y^2) - [E(Y)]^2 = 2389.84 - (48.84)^2 = 4.4944$ and $\sigma_Y = \sqrt{4.4944} = 2.12$.
One standard deviation from the mean value of Y gives $48.84 \pm 2.12 = 46.72$ to 50.96. So, the probability Y is within one standard deviation of its mean value equals $P(46.72 < Y < 50.96) = P(Y = 47, 48, 49, 50) = .12 + .14 + .25 + .17 = .68$.

33.

a. $E(X^2) = \sum_{x=0}^{1} x^2 \cdot p(x) = 0^2(1 - p) + 1^2(p) = p$.

b. $\text{Var}(X) = E(X^2) - [E(X)]^2 = p - [p]^2 = p(1 - p)$.

c. $E(X^{79}) = 0^{79}(1 - p) + 1^{79}(p) = p$. In fact, $E(X^n) = p$ for any positive power n.

35. Let $h_3(X)$ and $h_4(X)$ equal the net revenue (sales revenue minus order cost) for 3 and 4 copies purchased, respectively. If 3 magazines are ordered (\$6 spent), net revenue is \$4 – \$6 = –\$2 if $X = 1$, 2(\$4) – \$6 = \$2 if $X = 2$, 3(\$4) – \$6 = \$6 if $X = 3$, and also \$6 if $X = 4$, 5, or 6 (since that additional demand simply isn't met). The values of $h_4(X)$ can be deduced similarly. Both distributions are summarized below.

x	1	2	3	4	5	6
$h_3(x)$	–2	2	6	6	6	6
$h_4(x)$	–4	0	4	8	8	8
$p(x)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{2}{15}$

Using the table, $E[h_3(X)] = \sum_{x=1}^{6} h_3(x)\cdot p(x) = (-2)(\frac{1}{15}) + \ldots + (6)(\frac{2}{15}) = \$4.93$.

Similarly, $E[h_4(X)] = \sum_{x=1}^{6} h_4(x)\cdot p(x) = (-4)(\frac{1}{15}) + \ldots + (8)(\frac{2}{15}) = \$5.33$.

Therefore, ordering 4 copies gives slightly higher revenue, on the average.

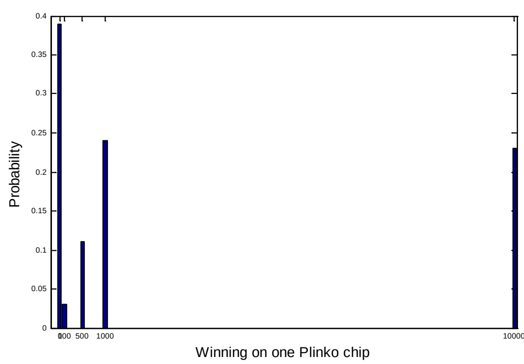
37. Using the hint, $E(X) = \sum_{x=1}^{n} x\cdot\left(\frac{1}{n}\right) = \frac{1}{n}\sum_{x=1}^{n} x = \frac{1}{n}\left[\frac{n(n+1)}{2}\right] = \frac{n+1}{2}$. Similarly,

$$E(X^2) = \sum_{x=1}^{n} x^2\cdot\left(\frac{1}{n}\right) = \frac{1}{n}\sum_{x=1}^{n} x^2 = \frac{1}{n}\left[\frac{n(n+1)(2n+1)}{6}\right] = \frac{(n+1)(2n+1)}{6}, \text{ so}$$

$$\text{Var}(X) = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}.$$

39.

a.



b. P(contestant makes money) = $P(X > 0)$ = .03 + .11 + .24 + .23 = .61. Or, $P(X > 0) = 1 - P(X \le 0)$ = 1 – .39 = .61.

c. $P(X \ge \$1000) = p(1000) + p(10{,}000)$ = .24 + .23 = .47.

d. $E(X) = \sum x \cdot p(x)$ = 0(.39) + 100(.03) + 500(.11) + 1000(.24) + 10,000(.23) = \$2598. The long-run average per Plinko chip (across all contestants across all the years of *The Price is Right*) should be around \$2598.

e. $E(X^2) = \sum x^2 \cdot p(x) = 0^2(.39) + \ldots + 10{,}000^2(.23)$ = 23,267,800, so Var(X) = 23,267,800 – $(2598)^2$ = 16,518,196 and SD(X) = $\sqrt{16{,}518{,}196}$ = \$4064.

41.

a. For a single number, $E(X) = –1(37/38) + (35)(1/38) = –2/38 = –\$1/19$, or about –5.26 cents.
For a square, $E(X) = –1(34/38) + (8)(4/38) = = –2/38 = –\$1/19$, or about –5.26 cents.

b. The expected return for a \$1 wager on roulette is the same no matter how you bet. (Specifically, in the long run players lose a little more than 5 cents for every \$1 wagered. This is how the payoffs are designed.) This seems to contradict our intuitive notion that betting on 1 of 38 numbers is riskier than betting on a color (18 of 38). The real lesson is that expected value does <u>not</u>, in general, correspond to the riskiness of a wager.

c. From Example 2.24, the standard deviation from a \$1 wager on a color is roughly \$1.
For a single number, $\text{Var}(X) = (–1 – [–1/19])^2(37/38) + (35 – [–1/19])^2(1/38) = 33.208$, so $\text{SD}(X) \approx \$5.76$.
For a square, $\text{Var}(X) = (–1 – [–1/19])^2(34/38) + (8 – [–1/19])^2(4/38) = 7.6288$, so $\text{SD}(X) \approx \$2.76$.

d. The standard deviations, in increasing order, are \$1 (color) < \$2.76 (square) < \$5.76 (single number). So, unlike expected value, standard deviation <u>does</u> correspond to our natural sense of which bets are riskier and which are safer. Specifically, low-risk/low-reward bets (such as a color) have smaller standard deviation than high-risk/high-reward bets (such as a single number).

43. Use the hint: $\text{Var}(aX + b) = E[((aX+b)-E(aX+b))^2] = E[((aX+b)-(a\mu_X+b))^2] = E[(aX-a\mu_X)^2] =$
$E[a^2(X-\mu_X)^2] = \sum a^2(x-\mu)^2 p(x) = a^2\sum(x-\mu)^2 p(x) = a^2\text{Var}(X)$.

45. With $a = 1$ and $b = –c$, $E(X – c) = E(aX + b) = a E(X) + b = E(X) – c = \mu – c$.
When $c = \mu$, $E(X – \mu) = E(X) – \mu = \mu – \mu = 0$; i.e., the expected deviation from the mean is zero.

47.

a. See the table below.

k	2	3	4	5	10
$1/k^2$	.25	.11	.06	.04	.01

b. From the table in Exercise 13, $\mu = 2.64$ and $\sigma^2 = 2.3704 \Rightarrow \sigma = 1.54$.
For $k = 2$, $P(|X – \mu| \geq 2\sigma) = P(|X – 2.64| \geq 2(1.54)) = P(X \geq 2.64 + 2(1.54)$ or $X \leq 2.64 – 2(1.54)) = P(X \geq 5.72$ or $X \leq –.44) = P(X = 6) = .04$. Chebyshev's bound of .25 is much too conservative.

For k = 3, 4, 5, or 10, $P(|X – \mu| \geq k\sigma)$ turns out to be zero, whereas Chebyshev's bound is positive. Again, this points to the conservative nature of the bound $1/k^2$.

c. X is $\pm d$ with probability 1/2 each. So, $E(X) = d(1/2) + (–d)(1/2) = 0$; $\text{Var}(X) = E(X^2) – \mu^2 = E(X^2) = d^2(1/2) + (–d)^2(1/2) = d^2$; and $\text{SD}(X) = d$.
$P(|X – \mu| < \sigma) = P(|X – 0| < d) = P(–d < X < d) = 0$, since $X = \pm d$. Here, Chebyshev's inequality gives no useful information: for $k = 1$ it states that at least $1 – 1/1^2 = 0$ percent of a distribution lies within one standard deviation of the mean.

d. $\mu = 0$ and $\sigma = 1/3$, so $P(|X – \mu| \geq 3\sigma) = P(|X| \geq 1) = P(X = –1 \text{ or } +1) = \frac{1}{18}+\frac{1}{18}=\frac{1}{9}$, identical to Chebyshev's upper bound of $1/k^2 = 1/3^2 = 1/9$.

e. There are many. For example, let $p(–1) = p(1) = \frac{1}{50}$ and $p(0) = \frac{24}{25}$.

Section 2.4

49.

a. Die rolls are independent, there is a fixed number of rolls, each roll results in a ⚃ or not in a ⚃, and the probability of rolling a ⚃ is the same on each roll. Thus, X is a binomial random variable, with parameters n = # of trials = # of rolls = 10 and $p = P$(success) $= P$(roll a ⚃) = 1/6.

b. We have a fixed number of questions/trials, the student gets each one right or wrong, "completely guessing" should imply both that his tries are independent and that the probability of getting the correct answer is the same on every question. Thus, X is a binomial random variable, with $n = 40$ and $p = 1/4$ (because there are 4 choices).

c. No: In this scenario, $p = 1/4$ for some of the questions and $p = 1/3$ for others. Since the probability is not the same for all trials, X is not binomial.

d. No: Similar to Example 2.29, the outcomes of successive trials are not independent. For example, the chance that any particular randomly selected student is a woman equals 20/35; however, the conditional probability the second selection is a woman given that the first is a woman equals $19/34 \neq 20/35$. So, X is not binomial. (The distribution of X will be studied in Section 2.6.)

e. No: Weights do not have dichotomous outcomes (success/failure). So, X is not binomial (realistically, it isn't even discrete!).

f. We have a fixed number of trials (15 apples), each either weighs more than 150 grams (success) or at most 150 grams (failure), and random sampling should imply both that the results of the 15 apple weighings are independent and that the probability any particular apple weighs more than 150 grams should be constant. Therefore, X is a binomial random variable, with $n = 15$. The numerical value of p, however, is unknown, since we don't know what proportion of all apples weigh more than 150 grams.

Two notes: (1) We're assuming these 15 applies are sampled from a very large population of apples, so that sampling with replacement is akin to Example 2.30 and not Example 2.29. (2) The fact that p is unknown does not mean X isn't binomial; rather, it just means we can't calculate probabilities pertaining to X.

51.

a. $B(4;10,.3) = .850$.

b. $b(4;10,.3) = B(4;10,.3) - B(3;10,.3) = .200$.

c. $b(6;10,.7) = B(6;10,.7) - B(5;10,.7) = .200$. (Notice that **b** and **c** are identical.)

d. $P(2 \le X \le 4) = B(4;10,.3) - B(1;10,.3) = .700$.

e. $P(2 \le X) = 1 - P(X \le 1) = 1 - B(1;10,.3) = .851$.

f. $P(X \le 1) = B(1;10,.7) = .000$.

g. $P(2 < X < 6) = P(2 < X \le 5) = B(5;10,.3) - B(2;10,.3) = .570$.

53. Let X be the number of "seconds," so $X \sim \text{Bin}(6, .1)$.

a. $P(X = 1) = \binom{6}{1}(.1)^1(.9)^5 = .3543$.

b. $P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)] = 1 - \left[\binom{6}{0}(.1)^0(.9)^6 + \binom{6}{1}(.1)^1(.9)^5\right] = 1 - [.5314 + .3543] =$ .1143.

c. Either 4 or 5 goblets must be selected.

Select 4 goblets with zero defects: $P(X = 0) = \binom{4}{0}(.1)^0(.9)^4 = .6561$.

Select 4 goblets, one of which has a defect, and the 5th is good: $\left[\binom{4}{1}(.1)^1(.9)^3\right] \times .9 = .26244$

So, the desired probability is .6561 + .26244 = .91854.

55. Let X be the number of drivers that come to a complete stop, so $X \sim \text{Bin}(20, .25)$.

a. $E(X) = np = 20(.25) = 5$.

b. $\text{Var}(X) = np(1-p) = 20(.25)(.75) = 3.75$, so $\text{SD}(X) = 1.94$.

c. $P(|X - 5| > 2(1.94)) = P(X < 1.12 \text{ or } X > 8.88) = P(X < 1.12) + P(X > 8.88) =$
$P(X < 1.12) + 1 - P(X \leq 8.88) = P(X \leq 1) + 1 - P(X \leq 8) = B(1; 20, .25) + 1 - B(8; 20, .25) = .065$.

57. Let "success" = has at least one citation and define X = number of individuals with at least one citation. Then $X \sim \text{Bin}(n = 15, p = .4)$.

a. If at least 10 have no citations (failure), then at most 5 have had at least one citation (success):
$P(X \leq 5) = B(5;15,.40) = .403$.

b. Half of 15 is 7.5, so less than half means 7 or fewer: $P(X \leq 7) = B(7;15,.40) = .787$.

c. $P(5 \leq X \leq 10) = P(X \leq 10) - P(X \leq 4) = .991 - .217 = .774$.

59. Let "success" correspond to a telephone that is submitted for service while under warranty and must be replaced. Then $p = P(\text{success}) = P(\text{replaced} \mid \text{submitted}) \cdot P(\text{submitted}) = (.40)(.20) = .08$. Thus X, the number among the company's 10 phones that must be replaced, has a binomial distribution with $n = 10$ and $p = .08$. Therefore, $P(X = 2) = \binom{10}{2}(.08)^2(.92)^8 = .1478$.

61. Let X = the number of flashlights that work, and let event B = {battery has acceptable voltage}.
Then $P(\text{flashlight works}) = P(\text{both batteries work}) = P(B)P(B) = (.9)(.9) = .81$. We have assumed here that the batteries' voltage levels are independent.
Finally, $X \sim \text{Bin}(10, .81)$, so $P(X \geq 9) = P(X = 9) + P(X = 10) = .285 + .122 = .407$.

63. Let X = the number of bits that are <u>not</u> switched during transmission ("successes"), so $X \sim$ Bin(3, .94).

a. A 3-bit sequence is decoded <u>incorrectly</u> if 2 or 3 of the bits are switched during transmission — equivalently, if the number of preserved bits is less than half (one or none). So,

P(decoded incorrectly) = $P(X = 0 \text{ or } 1) = \binom{3}{0}(.94)^0(.06)^3 + \binom{3}{1}(.94)^1(.06)^2 = .0104$.

b. While this type of repeating system triples the number of bits required to transmit a message, it reduces the likelihood of a transmitted bit being wrongly decoded by a factor of about 6 (from 6% to ~1%).

c. Similar to part **a**, let $X \sim$ Bin(5, .94) model the number of correctly decoded bits in a 5-bit message. A 5-bit message is decoded <u>incorrectly</u> if less than half the bits are preserved, and "less than half" of 5 is 0, 1, or 2:

$P(X = 0, 1, \text{ or } 2) = \binom{5}{0}(.94)^0(.06)^5 + \binom{5}{1}(.94)^1(.06)^4 + \binom{5}{2}(.94)^2(.06)^3 = .00197$.

65. In this example, $X \sim$ Bin(25, p) with p unknown.

a. P(rejecting claim when $p = .8$) = $P(X \le 15$ when $p = .8) = B(15; 25, .8) = .017$.

b. P(<u>not</u> rejecting claim when p = .7) = $P(X > 15$ when $p = .7) = 1 - P(X \le 15$ when $p = .7) =$
$= 1 - B(15; 25, .7) = 1 - .189 = .811$.
For $p = .6$, this probability is $= 1 - B(15; 25, .6) = 1 - .575 = .425$.

c. The probability of rejecting the claim when $p = .8$ becomes $B(14; 25, .8) = .006$, smaller than in **a** above. However, the probabilities of **b** above increase to .902 and .586, respectively. So, by changing 15 to 14, we're making it less likely that we will reject the claim when it's true (p really is $\ge .8$), but more likely that we'll "fail" to reject the claim when it's false (p really is $< .8$).

67. If topic A is chosen, then $n = 2$. When $n = 2$, P(at least half received) = $P(X \ge 1) = 1 - P(X = 0) =$

$1 - \binom{2}{0}(.9)^0(.1)^2 = .99$.

If topic B is chosen, then $n = 4$. When $n = 4$, P(at least half received) = $P(X \ge 2) = 1 - P(X \le 1) =$

$1 - \left[\binom{4}{0}(.9)^0(.1)^4 + \binom{4}{1}(.9)^1(.1)^3\right] = .9963$.

Thus topic B should be chosen if $p = .9$.
However, if $p = .5$, then the probabilities are .75 for A and .6875 for B (using the same method as above), so now A should be chosen.

69.

a. Although there are three payment methods, we are only concerned with S = uses a debit card and F = does not use a debit card. Thus we can use the binomial distribution. So, if X = the number of customers who use a debit card, $X \sim$ Bin($n = 100$, $p = .2$). From this,
$E(X) = np = 100(.2) = 20$, and $\text{Var}(X) = npq = 100(.2)(1–.2) = 16$.

b. With S = doesn't pay with cash, $n = 100$ and $p = .7$, so $\mu = np = 100(.7) = 70$ and $\sigma^2 = 21$.

71.

a. $np(1 – p) = 0$ if either $p = 0$ (whence every trial is a failure, so there is no variability in X) or if $p = 1$ (whence every trial is a success and again there is no variability in X).

b. $\frac{d}{dp}\left[np(1-p)\right] = n[(1)(1 – p) + p(–1)] = n[1 – 2p] = 0 \Rightarrow p = .5$, which is easily seen to correspond to a maximum value of Var(X).

73. When $n = 20$ and $p = .5$, $\mu = 10$ and $\sigma = 2.236$, so $2\sigma = 4.472$ and $3\sigma = 6.708$.
The inequality $|X – 10| \geq 4.472$ is satisfied if either $X \leq 5$ or $X \geq 15$, or
$P(|X – \mu| \geq 2\sigma) = P(X \leq 5 \text{ or } X \geq 15) = .021 + .021 = .042$. The inequality $|X – 10| \geq 6.708$ is satisfied if either $X \leq 3$ or $X \geq 17$, so $P(|X – \mu| \geq 3\sigma) = P(X \leq 3 \text{ or } X \geq 17) = .001 + .001 = .002$.

In the case $p = .75$, $\mu = 15$ and $\sigma = 1.937$, so $2\sigma = 3.874$ and $3\sigma = 5.811$.
$P(|X – 15| \geq 3.874) = P(X \leq 11 \text{ or } X \geq 19) = .041 + .024 = .065$, whereas
$P(|X – 15| \geq 5.811) = P(X \leq 9) = .004$.

All these probabilities are considerably less than the upper bounds given by Chebyshev:
for $k = 2$, Chebyshev's bound is $1/2^2 = .25$; for $k = 3$, the bound is $1/3^2 = .11$.

Section 2.5

75. All these solutions are found using the cumulative Poisson table, $P(x; \mu) = P(x; 5)$.

a. $P(X \leq 8) = P(8; 5) = .932$.

b. $P(X = 8) = P(8; 5) – P(7; 5) = .065$.

c. $P(X \geq 9) = 1 – P(X \leq 8) = .068$.

d. $P(5 \leq X \leq 8) = P(8; 5) – P(4; 5) = .492$.

e. $P(5 < X < 8) = P(7; 5) – P(5; 5) = .867 – .616=.251$.

77. Let $X \sim \text{Poisson}(\mu = 20)$.

a. $P(X \leq 10) = P(10; 20) = .011$.

b. $P(X > 20) = 1 – P(20; 20) = 1 – .559 = .441$.

c. $P(10 \leq X \leq 20) = P(20; 20) – P(9; 20) = .559 – .005 = .554$;
$P(10 < X < 20) = P(19; 20) – P(10; 20) = .470 – .011 = .459$.

d. $E(X) = \mu = 20$, so $\sigma = \sqrt{20} = 4.472$. Therefore, $P(\mu – 2\sigma < X < \mu + 2\sigma) =$
$P(20 – 8.944 < X < 20 + 8.944) = P(11.056 < X < 28.944) = P(X \leq 28) – P(X \leq 11) =$
$P(28; 20) – P(11; 20) = .966 – .021 = .945$.

79. The exact distribution of X is binomial with $n = 1000$ and $p = 1/200$; we can approximate this distribution by the Poisson distribution with $\mu = np = 5$.

a. $P(5 \le X \le 8) = P(8; 5) - P(4; 5) = .492$.

b. $P(X \ge 8) = 1 - P(X \le 7) = 1 - P(7; 5) = 1 - .867 = .133$.

81. Let X = the number of pages with typos. The exact distribution of X is Bin(400, .005), which we can approximate by a Poisson distribution with $\mu = np = 400(.005) = 2$. Based on this model,
$P(X = 1) = \dfrac{e^{-2}2^1}{1!} = .271$ and $P(X \le 3) = P(3; 2) = .857$.

83.

a. The expected number of failures on a 100-mile length of pipe in 365 days is
0.0081 failures/day × 365 days = 2.9565.
Let X = the number of failures in a year under these settings, so $X \sim$ Poisson(2.9565). Then
$P(X \ge 1) = 1 - P(X = 0) = 1 - \dfrac{e^{-2.9565}(2.9565)^0}{0!} = .948$.

b. The expected number of failures in 1500 miles (i.e. 15 100-mile segments) is 0.0864 × 15 = 1.296. Let X = the number of failures in a day under these settings, so $X \sim$ Poisson(1.296). Then
$P(X \ge 1) = 1 - P(X = 0) = 1 - \dfrac{e^{-1.296}(1.296)^0}{0!} = .726$.

85.

a. $\mu = 8$ when $t = 1$, so $P(X = 6) = \dfrac{e^{-8}8^6}{6!} = .122$; $P(X \ge 6) = 1 - P(5; 8) = .809$; and
$P(X \ge 10) = 1 - P(9; 8) = .283$.

b. t = 90 min = 1.5 hours, so $\mu = 12$; thus the expected number of arrivals is 12 and the standard deviation is $\sigma = \sqrt{12} = 3.464$.

c. t = 2.5 hours implies that $\mu = 20$. So, $P(X \ge 20) = 1 - P(19; 20) = .530$ and
$P(X \le 10) = P(10; 20) = .011$.

87.

a. For a quarter-acre (.25 acre) plot, the mean parameter is $\mu = (80)(.25) = 20$, so $P(X \le 16) = P(16; 20) = .221$.

b. The expected number of trees is $\lambda\cdot$(area) = 80 trees/acre (85,000 acres) = 6,800,000 trees.

c. The area of the circle is $\pi r^2 = \pi(.1)^2 = .01\pi = .031416$ square miles, which is equivalent to .031416(640) = 20.106 acres. Thus X has a Poisson distribution with parameter $\mu = \lambda(20.106) = 80(20.106) = 1608.5$. That is, the pmf of X is the function $p(x; 1608.5)$.

89.

a. First, consider a Poisson distribution with $\mu = \theta$. Since the sum of the pmf across all x-values (0, 1, 2, 3, …) must equal 1,

$$1 = \sum_{x=0}^{\infty} \frac{e^{-\theta}\theta^x}{x!} = \frac{e^{-\theta}\theta^0}{0!} + \sum_{x=1}^{\infty} \frac{e^{-\theta}\theta^x}{x!} = e^{-\theta} + \sum_{x=1}^{\infty} \frac{e^{-\theta}\theta^x}{x!} \Rightarrow \sum_{x=1}^{\infty} \frac{e^{-\theta}\theta^x}{x!} = 1 - e^{-\theta}$$

Also, the sum of the specified pmf across x = 1, 2, 3, … must equal 1, so

$1 = \sum_{x=1}^{\infty} k\frac{e^{-\theta}\theta^x}{x!} = k\sum_{x=1}^{\infty} \frac{e^{-\theta}\theta^x}{x!} = k[1 - e^{-\theta}]$ from above. Therefore, $k = \dfrac{1}{1-e^{-\theta}}$.

b. Again, consider a Poisson distribution with $\mu = \theta$. Since the expected value is θ,

$\theta = \sum_{x=0}^{\infty} x \cdot p(x;\theta) = 0 + \sum_{x=1}^{\infty} x \cdot p(x;\theta) = \sum_{x=1}^{\infty} x \cdot \frac{e^{-\theta}\theta^x}{x!}$. Multiply both sides by k:

$k\theta = k\sum_{x=1}^{\infty} x \cdot \frac{e^{-\theta}\theta^x}{x!} = \sum_{x=1}^{\infty} x \cdot k\frac{e^{-\theta}\theta^x}{x!}$; the right-hand side is the expected value of the specified distribution. So, the mean of a "zero-truncated" Poisson distribution is $k\theta$, i.e. $\dfrac{\theta}{1-e^{-\theta}}$.

The mean value 2.313035 corresponds to $\theta = 2$: $\dfrac{2}{1-e^{-2}}$ = 2.313035. And so, finally,

$P(X \le 5) = \sum_{x=1}^{5} k\frac{e^{-\theta}\theta^x}{x!} = \frac{e^{-2}}{1-e^{-2}}\sum_{x=1}^{5}\frac{2^x}{x!}$ = .9808.

c. Using the same trick as in part **b**, the mean-square value of our distribution is

$\sum_{x=1}^{\infty} x^2 \cdot k\frac{e^{-\theta}\theta^x}{x!} = k\sum_{x=1}^{\infty} x^2 \cdot \frac{e^{-\theta}\theta^x}{x!} = k\sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\theta}\theta^x}{x!} = k \cdot E(Y^2)$, where $Y \sim$ Poisson(θ).

For <u>any</u> rv, $\text{Var}(Y) = E(Y^2) - \mu^2 \Rightarrow E(Y^2) = \text{Var}(Y) + \mu^2$; for the Poisson($\theta$) rv, $E(Y^2) = \theta + \theta^2$. Therefore, the mean-square value of our distribution is $k \cdot (\theta + \theta^2)$, and the variance is $\text{Var}(X) = E(X^2) - [E(X)]^2 = k \cdot (\theta + \theta^2) - (k\theta)^2 = k\theta + k(1 - k)\theta^2$. Substituting $\theta = 2$ gives $\text{Var}(X) \approx 1.58897$, so $\text{SD}(X) \approx 1.2605$.

Section 2.6

91. According to the problem description, X is hypergeometric with $n = 6$, $N = 12$, and $M = 7$.

a. $P(X = 5) = \dfrac{\binom{7}{5}\binom{5}{1}}{\binom{12}{6}} = \dfrac{105}{924} = .114$.

b. $P(X \le 4) = 1 - P(X > 4) = 1 - [P(X = 5) + P(X = 6)] = 1 - \left[\dfrac{\binom{7}{5}\binom{5}{1}}{\binom{12}{6}} + \dfrac{\binom{7}{6}\binom{5}{0}}{\binom{12}{6}}\right] =$

$1 - [.114 + .007] = 1 - .121 = .879$.

c. $E(X) = n \cdot \dfrac{M}{N} = 6 \cdot \dfrac{7}{12} = 3.5$; $\text{Var}(X) = \left(\dfrac{12-6}{12-1}\right)6\left(\dfrac{7}{12}\right)\left(1-\dfrac{7}{12}\right) = 0.795$; $\sigma = 0.892$. So,
$P(X > \mu + \sigma) = P(X > 3.5 + 0.892) = P(X > 4.392) = P(X = 5 \text{ or } 6) = .121$ (from part **b**).

d. We can approximate the hypergeometric distribution with the binomial if the population size and the number of successes are large. Here, $n = 15$ and $M/N = 40/400 = .1$, so
$h(x;15, 40, 400) \approx b(x;15, .10)$. Using this approximation, $P(X \le 5) \approx B(5; 15, .10) = .998$ from the binomial tables. (This agrees with the exact answer to 3 decimal places.)

93.

a. Possible values of X are 5, 6, 7, 8, 9, 10. (In order to have less than 5 of the granite, there would have to be more than 10 of the basaltic). X is hypergeometric, with $n = 15$, $N = 20$, and $M = 10$. So, the pmf of X is

$$p(x) = h(x; 15, 10, 20) = \frac{\binom{10}{x}\binom{10}{15-x}}{\binom{20}{15}} .$$

The pmf is also provided in table form below.

x	5	6	7	8	9	10
$p(x)$	.0163	.1354	.3483	.3483	.1354	.0163

b. $P(\text{all 10 of one kind or the other}) = P(X = 5) + P(X = 10) = .0163 + .0163 = .0326$.

c. $\mu = n \cdot \dfrac{M}{N} = 15 \cdot \dfrac{10}{20} = 7.5$; $\text{Var}(X) = \left(\dfrac{20-15}{20-1}\right)15\left(\dfrac{10}{20}\right)\left(1-\dfrac{10}{20}\right) = .9868$; $\sigma = .9934$.

$\mu \pm \sigma = 7.5 \pm .9934 = (6.5066, 8.4934)$, so we want $P(6.5066 < X < 8.4934)$. That equals
$P(X = 7) + P(X = 8) = .3483 + .3483 = .6966$.

95.

a. The successes here are the top M = 10 pairs, and a sample of n = 10 pairs is drawn from among the N = 20. The probability is therefore $h(x;\, 10, 10, 20) = \dfrac{\binom{10}{x}\binom{10}{10-x}}{\binom{20}{10}}$.

b. Let X = the number among the top 5 who play east-west. (Now, M = 5.)
Then P(all of top 5 play the same direction) = $P(X = 5) + P(X = 0) = h(5;\, 10, 5, 20) + h(5;\, 10, 5, 20)$ =

$$\frac{\binom{5}{5}\binom{15}{5}}{\binom{20}{10}} + \frac{\binom{5}{0}\binom{15}{10}}{\binom{20}{10}} = .033\,.$$

c. Generalizing from earlier parts, we now have $N = 2n$; $M = n$. The probability distribution of X is hypergeometric: $p(x) = h(x;\, n, n, 2n) = \dfrac{\binom{n}{x}\binom{n}{n-x}}{\binom{2n}{n}}$ for $x = 0, 1, \ldots, n$. Also,

$$E(X) = n\cdot\frac{n}{2n} = \frac{1}{2}n \text{ and } \mathrm{Var}(X) = \left(\frac{2n-n}{2n-1}\right)\cdot n\cdot\frac{n}{2n}\cdot\left(1-\frac{n}{2n}\right) = \frac{n^2}{4(2n-1)}\,.$$

97. Let X = the number of ICs, among the 4 randomly selected ICs, that are defective. Then X has a hypergeometric distribution with n = 4 (sample size), M = 5 (# of "successes"/defectives in the population), and N = 20 (population size).

a. $P(X = 0) = h(0;\, 4, 5, 20) = \dfrac{\binom{5}{0}\binom{15}{4}}{\binom{20}{4}} = .2817$.

b. $P(X \le 1) = P(X = 0, 1) = .2817 + \dfrac{\binom{5}{1}\binom{15}{3}}{\binom{20}{4}} = .7513$.

c. Logically, if there are <u>fewer</u> defective ICs in the shipment of 20, then we ought to be <u>more</u> likely to accept the shipment. Replace M = 5 with M = 3 in the above calculations, and you will find that $P(X = 0) = .4912$ and $P(X \le 1) = .9123$.

99.

a. Let Y = the number of children the couple has until they have two girls. Then Y follows a negative binomial distribution with $r = 2$ (two girls/successes) and $p = .5$ (note: p here denotes P(girl)). The probability the family has exactly x <u>male</u> children is the probability they have $x + 2$ total children:

$P(Y = x + 2) = nb(x + 2; 2, .5) = \binom{x+2-1}{2-1}(.5)^2(1-.5)^{(x+2)-2} = (x + 1)(.5)^{x+2}$.

b. $P(Y = 4) = nb(4; 2, .5) = \binom{4-1}{2-1}(.5)^2(1-.5)^{4-2} = .1875$.

c. $P(Y \le 4) = \sum_{y=2}^{4} nb(y; 2, .5) = .25 + .25 + .1875 = .6875$.

d. The expected number of children is $E(Y) = \dfrac{r}{p} = \dfrac{2}{.5} = 4$. Since exactly two of these children are girls, the expected number of male children is $E(Y - 2) = E(Y) - 2 = 4 - 2 = 2$.

101. From Exercise 99, the expected number of male children in each family is 2. Hence, the expected total number of male children across the three families is just 2 + 2 + 2 = 6, the sum of the expected number of males born to each family.

103. Since there are 6 possible doubles, the probability of A & B together rolling a double is $p = 6/36 = 1/6$. The rv X = number of rolls to get 5 doubles has a negative binomial distribution with $r = 5$ and $p = 1/6$.

In particular, the pmf of x is $nb(x; 5, 1/6) = \binom{x-1}{5-1}(1/6)^5(5/6)^{x-5}$ for $x = 5, 6, 7, \ldots$; the expected number of rolls is $E(X) = \dfrac{r}{p} = \dfrac{5}{1/6} = 30$; and $SD(X) = \sqrt{\dfrac{r(1-p)}{p^2}} = \sqrt{\dfrac{5(5/6)}{(1/6)^2}} = \sqrt{150} = 12.2$ rolls.

105. Let X = the number of students the kinesiology professor needs to ask until she gets her 40 volunteers. With the specified assumptions, X has a negative binomial distribution with $r = 40$ and $p = .25$.

a. $E(X) = \dfrac{r}{p} = \dfrac{40}{.25} = 160$ students, and $SD(X) = \sqrt{\dfrac{r(1-p)}{p^2}} = \sqrt{\dfrac{40(.75)}{(.25)^2}} = \sqrt{480} = 21.9$ students.

b. $P(160 - 21.9 < X < 160 + 21.9) = P(138.1 < X < 181.9) = P(139 \le X \le 181) =$
$\sum_{x=139}^{181} \binom{x-1}{40-1}(.25)^{40}(.75)^{x-40} = .6756$ using software.

Section 2.7

107.

a. $M_X(t) = \sum_{x=9}^{12} e^{tx} p(x) = .01e^{9t} + .05e^{10t} + .16e^{11t} + .78e^{12t}$.

b. $M'_X(t) = .01 \cdot 9e^{9t} + .05 \cdot 10e^{10t} + .16 \cdot 11e^{11t} + .78 \cdot 12e^{12t} = .09e^{9t} + .50e^{10t} + 1.76e^{11t} + 9.36e^{12t} \Rightarrow$ by Theorem (2.21), $E(X) = M'_X(0) = .09 + .50 + 1.76 + 9.36 = 11.71$.
Next, $M''_X(t) = .09 \cdot 9e^{9t} + .50 \cdot 10e^{10t} + 1.76 \cdot 11e^{11t} + 9.36 \cdot 12e^{12t} =$
$M''_X(t) = .81e^{9t} + 5e^{10t} + 19.36e^{11t} + 112.32e^{12t} \Rightarrow E(X^2) = M''_X(0) = .81 + 5 + 19.36 + 112.32 = 137.49$.
Hence, by the variance shortcut formula, Var(X) = $E(X^2) - [E(X)]^2 = 137.49 - (11.71)^2 = 0.3659$, which gives SD($X$) = $\sqrt{0.3659} = 0.605$.

109. $M_X(t) = \sum_{x=1}^{\infty} e^{tx} p(x) = \sum_{x=1}^{\infty} e^{tx} (.5)^x = \sum_{x=1}^{\infty} (.5e^t)^x = (.5e^t)[1 + (.5e^t) + (.5e^t)^2 + \cdots] = \frac{.5e^t}{1 - .5e^t} = \frac{e^t}{2 - e^t}$, provided $|.5e^t| < 1$. (This condition is equivalent to $t < \ln(2)$ or $(-\infty, \ln(2))$ for the domain of M_X, which contains an open interval around 0.) From this, $E(X) = M'_X(0) = \left.\frac{2e^t}{(2 - e^t)^2}\right|_{t=0} = 2$.

Next, $E(X^2) = M''_X(0) = \left.\frac{2e^t(2 + e^t)}{(2 - e^t)^3}\right|_{t=0} = 6$, from which Var($X$) = $6 - 2^2 = 2$ and SD(X) = $\sqrt{2}$.

Notice that X has a geometric distribution with $p = .5$, from which we can also deduce the mean and sd.

111. It will be helpful to write the numerator of the skewness coefficient in a different form:

$$E[(X-\mu)^3] = E[X^3 - 3X^2\mu + 3X\mu^2 - \mu^3] = E[X^3] - 3\mu E[X^2] + 3\mu^2 E[X] - \mu^3$$
$$= E[X^3] - 3\mu E[X^2] + 2\mu^3 = M'''_X(0) - 3M'_X(0)M''_X(0) + 2[M'_X(0)]^3$$

Exercise	$E[X]$	$E[X^2]$	$E[X^3]$	$E[(X-\mu)^3]$	Sk. Coef.
107	11.71	137.49	1618.09	–0.49	–2.20
108	2.46	7.84	29.40	1.31	+0.54
109	2	6	26	6	+2.12
110	3.5	15.17	73.50	0	0

The grade level distribution is strongly skewed left (similar to Figure 2.9(a)), so its skewness coefficient should be negative. The defect distribution is skewed right (similar to Figure 2.9(b)), so its skewness coefficient should be positive. The geometric distribution in Exercise 109 is extremely skewed right, so its skewness coefficient should be positive. And the die roll distribution is perfectly symmetric, so its skewness coefficient is zero.

113. $E(X) = M'_X(0) = \left.\frac{2t}{(1-t^2)^2}\right|_{t=0} = 0$. Next, $E(X^2) = M''_X(0) = \left.\frac{6t^2 + 2}{(1-t^2)^3}\right|_{t=0} = 2$, from which

Var(X) = $2 - 0^2 = 2$.

115. Compare the two mgfs: they are an exact match, with $p = .75$. Therefore, by the uniqueness property of mgfs, Y must have a geometric distribution with parameter $p = .75$. In particular, the pmf of Y is given by $p(y) = (.25)^{y-1}(.75)$ for $y = 1, 2, 3, \ldots$.

117. Use the previous exercise with $a = 1/2$ and $b = -5/2$: $M_Y(t) = e^{-5/2t} M_X(t/2) = e^{-5/2t} e^{5(t/2)+2(t/2)^2} = e^{t^2/2}$.
From this, $M'_Y(t) = te^{t^2/2} \Rightarrow E(Y) = M'_Y(0) = 0$ and $M''_Y(t) = 1 \cdot e^{t^2/2} + t \cdot te^{t^2/2} \Rightarrow E(Y^2) = M''_Y(0) = 1 \Rightarrow$ $\text{Var}(Y) = 1 - 0^2 = 1$.

119. If $Y = 10 - X$, $E(Y) = E(10 - X) = 10 - E(X)$ and $\text{SD}(Y) = \text{SD}(10 - X) = |-1|\text{SD}(X) = \text{SD}(X)$. Also, $Y - \mu_Y = (10 - X) - (10 - \mu_X) = -(X - \mu_X)$. Thus, the skewness coefficient of Y is
$E[(Y - \mu_Y)^3]/\sigma_Y^3 = E[(-(X - \mu_X))^3]/\sigma_X^3 = E[(-1)^3(X - \mu_X)^3]/\sigma_X^3 = -E[(X - \mu_X)^3]/\sigma_X^3 = -c$.

121. The point of this exercise is that the natural log of the mgf is often a faster way to obtain the mean and variance of a rv, especially if the mgf has the form e^{function}.

a. $M_X(t) = e^{5t+2t^2} \Rightarrow M'_X(t) = (4t+5)e^{5t+2t^2} \Rightarrow E(X) = M'_X(0) = 5$;
$M''_X(t) = (4)e^{5t+2t^2} + (4t+5)^2 e^{5t+2t^2} \Rightarrow E(X^2) = M''_X(0) = 4 + 5^2 = 29 \Rightarrow \text{Var}(X) = 29 - 5^2 = 4$.

b. $L_X(t) = \ln[M_X(t)] = 2t^2 + 5t$. From Exercise 120,
$L'_X(t) = 4t + 5 \Rightarrow E(X) = L'_X(0) = 5$ and $L''_X(t) = 4 \Rightarrow \text{Var}(X) = L''_X(0) = 4$.

123. If $X \sim \text{Bin}(n, p)$, it's given in the section that $M_X(t) = (pe^t + 1 - p)^n$. Write Y = # of failures $= n - X = -1X+n$, and apply the rescaling proposition in the section:
$M_Y(t) = e^{nt}M_X(-t) = e^{nt}(pe^{-t} + 1 - p)^n = [e^t(pe^{-t} + 1 - p)]^n = (p + (1-p)e^t)^n$. Notice that this is the binomial mgf all over again, but with p and $1 - p$ exchanging positions.
Differentiating, $E(Y) = M'_Y(0) = n(1 - p)$; $E(Y^2) = M''_Y(0) = n(n-1)(1-p)^2 + n(1-p) \Rightarrow$
$\text{Var}(Y) = n(1 - p)p$. We observe that these formulas are the binomial mean and variance with p and $1 - p$ exchanged; equivalently, $E(Y) = n - np = n - E(X)$ and $\text{Var}(Y) = \text{Var}(X)$.

125. Apply the rescaling formula from this section with $a = 1$ and $b = -r$: $M_Y(t) = e^{-rt}M_X(t) = r$. Now, differentiate: $M'_Y(t) = -rp^r[1-(1-p)e^t]^{-r-1}[-(1-p)e^t] = rp^r(1-p)e^t[1-(1-p)e^t]^{-r-1} \Rightarrow E(Y) =$
$M'_Y(0) = rp^r(1-p)(1)[p]^{-r-1} = \dfrac{r(1-p)}{p}$. Notice that this can be re-written as $(r/p) - r$, i.e. $E(X) - r$.
The product rule then gives $E(Y^2) = M''_Y(0) = \cdots = \dfrac{r(1-p) + r^2(1-p)^2}{p^2}$, from which $\text{Var}(Y) = \dfrac{r(1-p)}{p^2}$.

127.

a. $M_X(t) = e^{\mu(e^t-1)} = e^{\mu e^t - \mu} \Rightarrow M'_X(t) = e^{\mu e^t - \mu} \cdot \mu e^t = \mu e^{\mu e^t + t - \mu} \Rightarrow E(X) = \mu e^{\mu + 0 - \mu} = \mu e^0 = \mu$. Next,
$M''_X(t) = \mu e^{\mu e^t + t - \mu} \cdot (\mu e^t + 1) \Rightarrow E(X^2) = \mu e^0 \cdot (\mu + 1) = \mu^2 + \mu$, whence $\text{Var}(X) = E(X^2) - [E(X)]^2 = \mu$.

b. $L_X(t) = \ln[M_X(t)] = \mu(e^t - 1) = \mu e^t - \mu$. $L'_X(t) = \mu e^t \Rightarrow E(X) = L'_X(0) = \mu$ and
$L''_X(t) = \mu e^t \Rightarrow \text{Var}(X) = L''_X(0) = \mu$. Just like in the previous exercises, we see here that the natural logarithm of the mgf is often much easier to use for computing mean and variance.

Section 2.8

129. Using the built-in commands of Matlab/R, our "program" can actually be extremely short. The programs below simulate 10,000 values from the pmf in Exercise 30.

In Matlab:

```
Y=randsample([0,1,2,3],10000,true,[.60,.25,.10,.05]);
mean(Y)
std(Y)
```

In R:

```
Y=sample(c(0,1,2,3),10000,TRUE,c(.60,.25,.10,.05))
mean(Y)
sd(Y)
```

One execution of the program returned $\hat{\mu} = \bar{x} = 0.5968$ and $\hat{\sigma} = s = 0.8548$. From Exercise 30, the actual mean and sd are μ = 0.60 and σ = $\sqrt{0.74} \approx 0.86$, so these estimated values are pretty close.

131. There are numerous ways to approach this problem. In our programs below, we count the number of times the event $A = \{X \le 25\}$ occurs in 10,000 simulations. Five initial tests are administered, with a positive test (denoted 1) occurring with probability .75 and a negative test (denoted 0) with probability .25. The while statement looks at the most recent five tests and determines if the sum of the indicators (1's and 0's) is 5, since that would signify five consecutive positive test results. If not, then another test is administered, and the result of that test is appended to the previous results. At the end, a vector `X` of 0's and 1's terminating in 11111 exits the while loop; if its length is ≤ 25, then the number of tests was ≤ 25, and the count for the event A is increased by 1.

```
A=0;
for i=1:10000
    k=1;
    X=randsample([0,1],5,
              true,[.25,.75]);
    while sum(X(k:k+4))<5
        k=k+1;
        test=randsample([0,1],1,
              true,[.25,.75]);
        X=[X test];
    end
    if length(X)<=25
        A=A+1;
    end
end
```

```
A<-0
for(i in 1:10000){
    k<-1
    X<-sample(c(0,1),5,
                    TRUE,c(.25,.75))
    while(sum(X[k:(k+4)])<5){
        k<-k+1
        test<-sample(c(0,1),1,
                    TRUE,c(.25,.75))
        X<-c(X,test)
    }
    if(length(X)<=25){
        A<-A+1
    }
}
```

An execution of the program in R gave A=9090, so $\hat{P}(X \le 25) = \hat{P}(A) = \dfrac{9090}{10000}$ = .9090.

133.

a. Modify the programs from Exercise 109 of Chapter 1. (See Chapter 1 solutions for an explanation of the code.)

```
X=zeros(10000,1);
for i=1:10000
    u=rand(10,1); v=rand(13,1);
    w=rand(13,1); x=rand(5,1);
    right=sum(u<1/2)+sum(v<1/3)
            +sum(w<1/4)+sum(x<1/5);
    X(i)=right;
end
```

```
X<-NULL
for(i in 1:10000){
    u<-runif(10); v<-runif(13)
    w<-runif(13); x<-runif(5)
    right<-sum(u<1/2)+sum(v<1/3)
            +sum(w<1/4)+sum(x<1/5)
    X[i]<-right
}
```

The program returns a 10,000-by-1 vector of simulated values of X. Executing this program in Matlab gave a vector with `mean(X)` = 13.5888 and `std(X)` = 2.9381.

b. The command `mean(X>=mean(X)+std(X))` computes one standard deviation above the mean, then determines which entries of `X` are at least one standard deviation above the mean (so the object in parentheses is a vector of 0's and 1's). Finally, taking the mean of that vector returns the proportion of entries that are 1's; using `sum(…)/10000` would do the same thing. Matlab returned .1562, so the estimated probability is $\hat{P}(X \geq \mu + \sigma) \approx \hat{P}(X \geq \bar{x} + s)$ = .1562.

135. The cdf of Benford's law actually has a simple form:
$F(1) = p(1) = \log_{10}(2)$; $F(2) = p(1) + p(2) = \log_{10}\left(\frac{2}{1}\right) + \log_{10}\left(\frac{3}{2}\right) = \log_{10}\left(\frac{2}{1}\cdot\frac{3}{2}\right) = \log_{10}(3)$; and using the same property of logs, $F(x) = \log_{10}(x + 1)$ for $x = 1, 2, 3, 4, 5, 6, 7, 8, 9$. Using the inverse cdf method, we should assign a value x so long as $F_{x-1} \leq u < F_x$; i.e., $\log_{10}(x) \leq u < \log_{10}(x + 1)$. But that's equivalent to the inequalities $x \leq 10^u < x + 1$, which implies that x is the greatest integer less than or equal to 10^u.
Thus, we may implement the inverse cdf method here not with a series of if-else statements, but with a simple call to the "floor" function.

```
u=rand(10000,1);
X=floor(10.^u);
```

```
u<-runif(10000)
X<-floor(10^u)
```

For one execution of this (very simple) program in R, the sample mean and standard deviation were `mean(X)` = 3.4152 and `var(X)` = 5.97.

137.

a. The functions below return a 10000-by-1 vector containing simulated values of the profit the airline receives, based on the number of tickets sold and the loss for over-booking.

```
function Profit=airline(t)
Profit=zeros(10000,1);
for i=1:10000
    X=random('bin',t,.85);
    B=max(X-120,0);
    Profit(i)=250*t-500*B;
end
```

```
airline<-function(t){
Profit<-NULL
for(i in 1:10000){
    X<-rbinom(1,t,.85)
    B<-max(X-120,0)
    Profit[i]<-250*t-500*B
}
return(Profit)
}
```

b. Execute the program above for input t = 140, 141, … 150, and determine the sample mean profit for each simulation. Matlab returned the following estimates of expected profit:

t	140	141	142	143	…	150
Avg. Profit	\$34,406	\$34,458	\$34,468	\$34,431	…	\$33,701

Average profit increased monotonically from t = 140 to t = 142 and decreased monotonically from t = 142 to t = 150 (not shown). Thus, the airline should sell 142 tickets in order to maximize their expected profit.

139. The programs below first simulate the number of chips a player earns, stored as `chips`; then, `chips` is passed as an argument in the second simulation to determine how many values from the "Plinko distribution" from Exercise 39 should be generated. The output vector `W` contains 10,000 simulations of the contestant's total winnings.

```
function W=plinkowin
W=zeros(10000,1);
c=[1,2,3,4,5];
pc=[.03,.15,.35,.34,.13];
x=[0,100,500,1000,10000];
px=[.39,.03,.11,.24,.23];
for i=1:10000
    chips=randsample(c,1,true,pc);
    wins=randsample(x,chips,true,px);
    W(i)=sum(wins);
end
```

```
plinkowin<-function(){
W<-NULL
ch<-c(1,2,3,4,5)
pc<-c(.03,.15,.35,.34,.13)
x<-c(0,100,500,1000,10000)
px<-c(.39,.03,.11,.24,.23)
for(i in 1:10000){
    chips<-sample(ch,1,TRUE,pc)
    wins<-sample(x,chips,TRUE,px)
    W[i]<-sum(wins)
}
return(W)
}
```

Using the preceding R program gave the following estimates.

a. $P(W > 11{,}000) \approx \hat{P}(W > 11{,}000)$ = `mean(W>11000)` = .2991.

b. $E(W) \approx \overline{w}$ = `mean(W)` = \$8,696.

c. $\text{SD}(X) \approx s$ = `sd(X)` = \$7,811.

d. Delete the lines creating the pmf of C, and replace the line of code creating the count chips as follows:
in Matlab: `chips=1+random('bin',4,.5);`
in R: `chips<-1+rbinom(1,4,.5)`
Running the new program gave the following estimates: $P(W > 11{,}000) \approx .2342$, $E(X) \approx \$7{,}767$, and $\text{SD}(X) \approx \$7{,}571$.

141.

a. The programs below generate a random sample of size n = 150 from the ticket pmf described in the problem. Those 150 requests are stored in `tickets`, and then the <u>total</u> number of requested tickets is stored in `T`.

```
T=zeros(10000,1);
x=[0,1,2,3,4];
px=[.05,.15,.25,.25,.30];
for i=1:10000
    tickets=randsample(x,150,true,px);
    T(i)=sum(tickets);
end
```

```
numboxes<-NULL
for(i in 1:10000){
    coupon<-rep(0,10)
    while(all(coupon>0)==0){
        new<-sample(1:10,1)
        coupon[new]<-coupon[new]+1
    }
    numboxes[i]<-sum(coupon)
}
```

b. Using the program above, $P(T \le 410) \approx$ `mean(T<=410)` = .9196 for one run. The estimated standard error of that estimate is $\sqrt{\dfrac{(.9196)(1-.9196)}{10{,}000}}$ = .00272. Hence, a 95% confidence interval for the true probability (see Chapter 5) is .9196 ± 1.96(.00272) = (.9143, .9249).

Supplementary Exercises

143.

a. We'll find $p(1)$ and $p(4)$ first, since they're easiest, then $p(2)$. We can then find $p(3)$ by subtracting the others from 1. $p(1) = P(\text{exactly one suit}) = P(\text{all } \spadesuit) + P(\text{all } \heartsuit) + P(\text{all } \diamondsuit) + P(\text{all } \clubsuit) =$

$4 \cdot P(\text{all } \spadesuit) = 4 \cdot \dfrac{\binom{13}{5}\binom{39}{0}}{\binom{52}{5}} = .00198$, since there are 13 ♠s and 39 other cards.

$p(4) = 4 \cdot P(2\ \spadesuit, 1\ \heartsuit, 1\ \diamondsuit, 1\ \clubsuit) = 4 \cdot \dfrac{\binom{13}{2}\binom{13}{1}\binom{13}{1}\binom{13}{1}}{\binom{52}{5}} = .26375$.

$p(2) = P(\text{all } \heartsuit\text{s and } \spadesuit\text{s, with} \ge \text{one of each}) + \ldots + P(\text{all } \diamondsuit\text{s and } \clubsuit\text{s with} \ge \text{one of each}) =$

$\binom{4}{2} \cdot P(\text{all } \heartsuit\text{s and } \spadesuit\text{s, with} \ge \text{one of each}) =$

$6 \cdot [P(1\ \heartsuit \text{ and } 4\ \spadesuit) + P(2\ \heartsuit \text{ and } 3\ \spadesuit) + P(3\ \heartsuit \text{ and } 2\ \spadesuit) + P(4\ \heartsuit \text{ and } 1\ \spadesuit)] =$

$$6 \cdot \left[2 \cdot \frac{\binom{13}{4}\binom{13}{1}}{\binom{52}{5}} + 2 \cdot \frac{\binom{13}{3}\binom{13}{2}}{\binom{52}{5}}\right] = 6\left[\frac{18{,}590 + 44{,}616}{2{,}598{,}960}\right] = .14592.$$

Finally, $p(3) = 1 - [p(1) + p(2) + p(4)] = .58835$.

b. $\mu = \sum_{x=1}^{4} x \cdot p(x) = 3.114$; $\sigma^2 = \left[\sum_{x=1}^{4} x^2 \cdot p(x)\right] - (3.114)^2 = .405 \Rightarrow \sigma = .636$.

145.

a. From the description, $X \sim \text{Bin}(15, .75)$. So, the pmf of X is $b(x; 15, .75)$.

b. $P(X > 10) = 1 - P(X \le 10) = 1 - B(10;15, .75) = 1 - .314 = .686$.

c. $P(6 \le X \le 10) = B(10; 15, .75) - B(5; 15, .75) = .314 - .001 = .313$.

d. $\mu = (15)(.75) = 11.75$, $\sigma^2 = (15)(.75)(.25) = 2.81 \Rightarrow \sigma = 1.68$.

e. Requests can all be met if and only if $X \le 10$ <u>and</u> $15 - X \le 8$, i.e. iff $7 \le X \le 10$. So, $P(\text{all requests met}) = P(7 \le X \le 10) = B(10; 15, .75) - B(6; 15, .75) = .310$.

147.

a. Let X = the number of bits transmitted until the third error occurs. Then $X \sim \text{NB}(r = 3, p = .05)$. Thus $P(X = 50) = nb(50; 3, .05) = \binom{50-1}{3-1}(.05)^3(.95)^{47} = .013$.

b. Using the mean of the geometric distribution, the average number of bits up to and including the first error is $1/p = 1/.05 = 20$ bits. Hence, the average number <u>before</u> the first error is $20 - 1 = 19$.

c. Now let X = the number of bit errors in a 32-bit word. Then $X \sim \text{Bin}(n = 32, p = .05)$, and $P(X = 2) = b(2; 32, .05) = \binom{32}{2}(.05)^2(.95)^{30} = .266$.

d. The rv X has a Bin(10000, .05) distribution. Since n is large and p is small, this can be approximated by a Poisson distribution, with $\mu = 10000(.05) = 500$.

149.

a. $X \sim \text{Bin}(n = 500, p = .005)$. Since n is large and p is small, X can be approximated by a Poisson distribution with $\mu = np = 2.5$. The approximate pmf of X is $p(x; 2.5) = \frac{e^{-2.5}2.5^x}{x!}$.

b. $P(X = 5) = \frac{e^{-2.5}2.5^5}{5!} = .0668$.

c. $P(X \ge 5) = 1 - P(X \le 4) = 1 - P(4; 2.5) = 1 - .8912 = .1088$.

151. Let Y denote the number of tests carried out.
For $n = 3$, possible Y values are 1 and 4. $P(Y = 1) = P(\text{no one has the disease}) = (.9)^3 = .729$ and $P(Y = 4) = 1 - .729 = .271$, so $E(Y) = (1)(.729) + (4)(.271) = 1.813$, as contrasted with the 3 tests necessary without group testing.
For $n = 5$, possible values of Y are 1 and 6. $P(Y = 1) = P(\text{no one has the disease}) = (.9)^5 = .5905$, so $P(Y = 6) = 1 - .5905 = .4095$ and $E(Y) = (1)(.5905) + (6)(.4095) = 3.0475$, less than the 5 tests necessary without group testing.

153. $p(2) = P(X = 2) = P(SS) = p^2$, and $p(3) = P(FSS) = (1 – p)p^2$.

For $x \geq 4$, consider the first $x – 3$ trials and the last 3 trials separately. To have $X = x$, it must be the case that the last three trials were FSS, and that two-successes-in-a-row was not already seen in the first $x – 3$ tries.

The probability of the first event is simply $(1 – p)p^2$.
The second event occurs if two-in-a-row hadn't occurred after 2 or 3 or … or $x – 3$ tries. The probability of this second event equals $1 – [p(2) + p(3) + … + p(x – 3)]$. (For $x = 4$, the probability in brackets is empty; for $x = 5$, it's $p(2)$; for $x = 6$, it's $p(2) + p(3)$; and so on.)

Finally, since trials are independent, $P(X = x) = (1 – [p(2) + … + p(x – 3)]) \cdot (1 – p)p^2$.

For $p = .9$, the pmf of X up to $x = 8$ is shown below.

x	2	3	4	5	6	7	8
$p(x)$	.81	.081	.081	.0154	.0088	.0023	.0010

So, $P(X \leq 8) = p(2) + … + p(8) = .9995$.

155.

a. Let event A = seed carries single spikelets, and event B = seed produces ears with single spikelets. Then $P(A \cap B) = P(A) \cdot P(B \mid A) = (.40)(.29) = .116$.
Next, let X = the number of seeds out of the 10 selected that meet the condition $A \cap B$. Then $X \sim$ Bin(10, .116). So, $P(X = 5) = \binom{10}{5}(.116)^5(.884)^5 = .002857$.

b. For any one seed, the event of interest is B = seed produces ears with single spikelets. Using the law of total probability, $P(B) = P(A \cap B) + P(A' \cap B) = (.40)(.29) + (.60)(.26) = .272$.
Next, let Y = the number out of the 10 seeds that meet condition B. Then $Y \sim$ Bin(10, .272).
$P(Y = 5) = \binom{10}{5}(.272)^5(1-.272)^5 = .0767$, while
$P(Y \leq 5) = \sum_{y=0}^{5}\binom{10}{y}(.272)^y(1-.272)^{10-y} = .041813 + … + .076719 = .97024$.

157.

a. Using the Poisson model with $\mu = 2(1) = 2$, $P(X = 0) = p(0; 2)$ or $\dfrac{e^{-2}2^0}{0!} = .135$.

b. Let S = an operator receives no requests. Then the number of operators that receive no requests follows a Bin($n = 5$, $p = .135$) distribution. So, P(exactly 4 S's in 5 trials) = $b(4; 5, .135)$ =
$\binom{5}{4}(.135)^4(.865)^1 = .00144$.

c. For any non-negative integer x, P(all operators receive exactly x requests) =

P(first operator receives x) · … · P(fifth operator receives x) = $[p(x;\,2)]^5 = \left[\dfrac{e^{-2}2^x}{x!}\right]^5 = \dfrac{e^{-10}2^{5x}}{(x!)^5}$.

Then, P(all receive the same number) = P(all receive 0 requests) + P(all receive 1 request) + P(all receive 2 requests) + … = $\displaystyle\sum_{x=0}^{\infty}\frac{e^{-10}2^{5x}}{(x!)^5}$.

159. The number of magazine copies sold is X so long as X is no more than five; otherwise, all five copies are sold. So, mathematically, the number sold is min(X, 5), and

$E[\min(x, 5)] = \displaystyle\sum_{x=0}^{\infty}\min(x,5)\,p(x;4) = 0p(0;4) + 1p(1;4) + 2p(2;4) + 3p(3;4) + 4p(4;4) + \sum_{x=5}^{\infty}5\,p(x;4) =$

$1.735 + 5\displaystyle\sum_{x=5}^{\infty}p(x;4) = 1.735 + 5\left[1-\sum_{x=0}^{4}p(x;4)\right] = 1.735 + 5[1 - P(4;4)] = 3.59.$

161.

a. No, since the probability of a "success" is not the same for all tests.

b. There are four ways exactly three could have positive results. Let D represent those with the disease and D' represent those without the disease.

Combination D	D'	Probability
0	3	$\left[\binom{5}{0}(.2)^0(.8)^5\right]\cdot\left[\binom{5}{3}(.9)^3(.1)^2\right]$ =(.32768)(.0729) = .02389
1	2	$\left[\binom{5}{1}(.2)^1(.8)^4\right]\cdot\left[\binom{5}{2}(.9)^2(.1)^3\right]$ =(.4096)(.0081) = .00332
2	1	$\left[\binom{5}{2}(.2)^2(.8)^3\right]\cdot\left[\binom{5}{1}(.9)^1(.1)^4\right]$ =(.2048)(.00045) = .00009216
3	0	$\left[\binom{5}{3}(.2)^3(.8)^2\right]\cdot\left[\binom{5}{0}(.9)^0(.1)^5\right]$ =(.0512)(.00001) = .000000512

Adding up the probabilities associated with the four combinations yields 0.0273.

163.

a. Notice that $p(x; \mu_1, \mu_2) = .5\, p(x; \mu_1) + .5\, p(x; \mu_2)$, where both terms $p(x; \mu_i)$ are Poisson pmfs. Since both pmfs are ≥ 0, so is $p(x; \mu_1, \mu_2)$. That verifies the first requirement.

Next, $\sum_{x=0}^{\infty} p(x;\mu_1,\mu_2) = .5\sum_{x=0}^{\infty} p(x;\mu_1) + .5\sum_{x=0}^{\infty} p(x;\mu_2)$ = .5 + .5 = 1, so the second requirement for a pmf is met. Therefore, $p(x; \mu_1, \mu_2)$ is a valid pmf.

b. $E(X) = \sum_{x=0}^{\infty} x \cdot p(x;\mu_1,\mu_2) = \sum_{x=0}^{\infty} x[.5\,p(x;\mu_1) + .5\,p(x;\mu_2)] = .5\sum_{x=0}^{\infty} x \cdot p(x;\mu_1) + .5\sum_{x=0}^{\infty} x \cdot p(x;\mu_2) = .5E(X_1) +$ $.5E(X_2)$, where $X_i \sim$ Poisson(μ_i). Therefore, $E(X) = .5\mu_1 + .5\mu_2$.

c. This requires using the variance shortcut. Using the same method as in **b,**

$E(X^2) = .5\sum_{x=0}^{\infty} x^2 \cdot p(x;\mu_1) + .5\sum_{x=0}^{\infty} x^2 \cdot p(x;\mu_2) = .5E(X_1^2) + .5E(X_2^2)$. For any Poisson rv,

$E(X^2) = \text{Var}(X) + [E(X)]^2 = \mu + \mu^2$, so $E(X^2) = .5(\mu_1 + \mu_1^2) + .5(\mu_2 + \mu_2^2)$.

Finally, $\text{Var}(X) = .5(\mu_1 + \mu_1^2) + .5(\mu_2 + \mu_2^2) - [.5\mu_1 + .5\mu_2]^2$, which can be simplified to equal $.5\mu_1 + .5\mu_2 + .25(\mu_1 - \mu_2)^2$.

d. Simply replace the weights .5 and .5 with .6 and .4, so $p(x; \mu_1, \mu_2) = .6\, p(x; \mu_1) + .4\, p(x; \mu_2)$.

165. Since $X \sim$ Poisson(μ) for some unknown μ, $P(X = 1) = \dfrac{e^{-\mu}\mu^1}{1!} = \mu e^{-\mu}$ and $P(X = 2) = \dfrac{e^{-\mu}\mu^2}{2!} = .5\mu^2 e^{-\mu}$.

Thus, $P(X = 1) = 4P(X = 2)$ implies $\mu e^{-\mu} = 4 \cdot .5\mu^2 e^{-\mu} \Rightarrow 1 = 2\mu \Rightarrow \mu = .5$.

167. If choices are independent and have the same probability p, then $X \sim \text{Bin}(25, p)$, where p denotes the probability that customer wants the fancy coffee maker. Using the rescaling properties of mean and standard deviation,

$E(20X + 750) = 20E(X) + 750 = 20 \cdot 25p + 750 = 500p + 750$, and

$\text{SD}(20X + 750) = 20\text{SD}(X) = 20 \cdot \sqrt{25p(1-p)} = 100\sqrt{p(1-p)}$.

Customers' choices might not be independent because, for example, the store may have a limited supply of one or both coffee makers, customers might affect each other's choices, the sales employees may influence customers, …

169.

a. Let A_1 = {voice}, A_2 = {data}, and X = duration of a call. Then $E(X) = E(X|A_1)P(A_1) + E(X|A_2)P(A_2) =$ 3(.75) + 1(.25) = 2.5 minutes.

b. Let X = the number of chips in a cookie. Then $E(X) = E(X|i = 1)P(i = 1) + E(X|\, i = 2)P(i = 2) +$ $E(X|\, i = 3)P(i = 3)$. If X is Poisson, then its mean is the specified μ — that is, $E(X|i) = i + 1$. Therefore, $E(X)$ = 2(.20) + 3(.50) + 4(.30) = 3.1 chips.

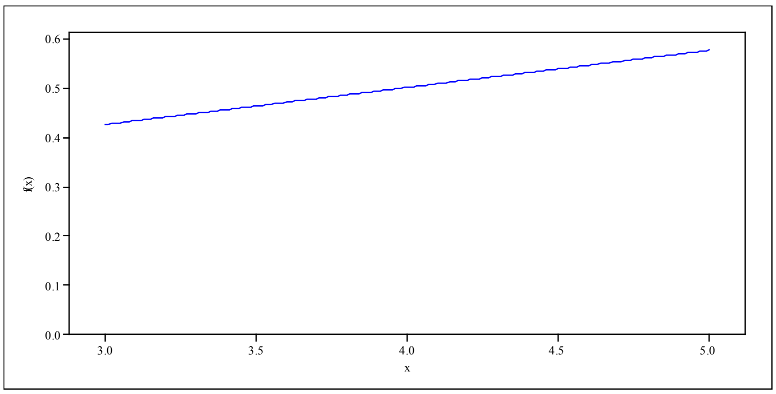
CHAPTER 3

Section 3.1

1.

a. The pdf is the straight-line function graphed below on [3, 5]. The function is clearly non-negative; to verify its integral equals 1, compute:

$$\int_3^5 (.075x + .2)\,dx = .0375x^2 + .2x \Big]_3^5 = (.0375(5)^2 + .2(5)) - (.0375(3)^2 + .2(3))$$
$$= 1.9375 - .9375 = 1$$



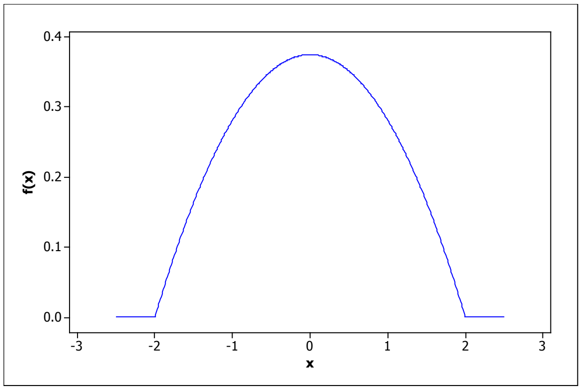
b. $P(X \le 4) = \int_3^4 (.075x + .2)\,dx = .0375x^2 + .2x \Big]_3^4 = (.0375(4)^2 + .2(4)) - (.0375(3)^2 + .2(3))$ $= 1.4 - .9375 = .4625$. Since X is a continuous rv, $P(X < 4) = P(X \le 4) = .4625$ as well.

c. $P(3.5 \le X \le 4.5) = \int_{3.5}^{4.5} (.075x + .2)\,dx = .0375x^2 + .2x \Big]_{3.5}^{4.5} = \cdots = .5$.

$P(4.5 < X) = P(4.5 \le X) = \int_{4.5}^{5} (.075x + .2)\,dx = .0375x^2 + .2x \Big]_{4.5}^{5} = \cdots = .278125$.

3.

a.



b. $P(X > 0) = \int_0^2 .09375(4-x^2)dx = .09375\left(4x - \frac{x^3}{3}\right)\Big]_0^2 = .5$.

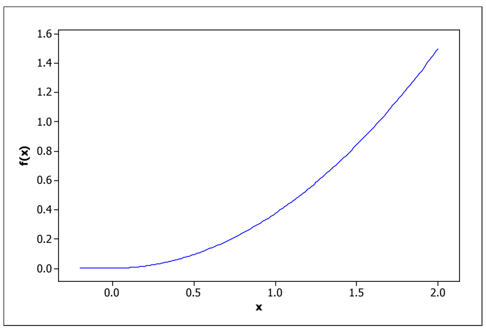
This matches the symmetry of the pdf about $x = 0$.

c. $P(-1 < X < 1) = \int_{-1}^{1} .09375(4-x^2)dx = .6875$.

d. $P(X < -.5 \text{ or } X > .5) = 1 - P(-.5 \le X \le .5) = 1 - \int_{-.5}^{.5} .09375(4-x^2)dx = 1 - .3672 = .6328.$

5.

a. $1 = \int_{-\infty}^{\infty} f(x)dx = \int_0^2 kx^2dx = \frac{kx^3}{3}\Big]_0^2 = \frac{8k}{3} \Rightarrow k = \frac{3}{8}$.



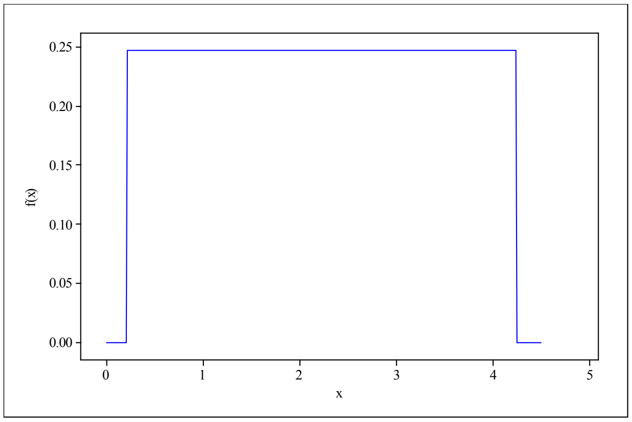
b. $P(0 \le X \le 1) = \int_0^1 \frac{3}{8}x^2dx = \frac{1}{8}x^3\Big]_0^1 = \frac{1}{8} = .125$.

c. $P(1 \le X \le 1.5) = \int_1^{1.5} \frac{3}{8}x^2dx = \frac{1}{8}x^3\Big]_1^{1.5} = \frac{1}{8}\left(\frac{3}{2}\right)^3 - \frac{1}{8}(1)^3 = \frac{19}{64} = .296875$.

d. $P(X \ge 1.5) = \int_{1.5}^{2} \frac{3}{8}x^2dx = \frac{1}{8}x^3\Big]_{1.5}^2 = \frac{1}{8}(2)^3 - \frac{1}{8}(1.5)^3 = .578125$.

7.

a. $f(x) = \dfrac{1}{B-A} = \dfrac{1}{4.25-.20} = \dfrac{1}{4.05}$ for $.20 \le x \le 4.25$ and $f(x) = 0$ otherwise.



b. $P(X > 3) = \int_3^{4.25} \dfrac{1}{4.05}dx = \dfrac{4.25-3}{4.05} \approx .3086.$

c. The median is obviously the midpoint, since the distribution is symmetric: $\eta = \dfrac{.20+4.25}{2} = 2.225$mm.

$P(|X - 2.225| < 1) = P(1.225 < X < 3.225) = \int_{1.225}^{3.225} \dfrac{1}{4.05}dx = \dfrac{3.225-1.225}{4.05} = \dfrac{2}{4.05} \approx .4938.$

d. $P(a \le x \le a + 1) = \int_a^{a+1} \dfrac{1}{4.05}dx = \dfrac{(a+1)-a}{4.05} = \dfrac{1}{4.05}$, since the interval has length 1.

9.

a. $P(X \le 6) = \int_{.5}^{6} .15e^{-.15(x-.5)}dx = .15\int_0^{5.5} e^{-.15u}du$ (after the substitution $u = x - .5$)

$= -e^{-.15u}\Big]_0^{5.5} = 1 - e^{-.825} \approx .562$

b. $P(X > 6) = 1 - P(X \le 6) = 1 - .562 = .438$. Since X is continuous, $P(X \ge 6) = P(X > 6) = .438$ as well.

c. $P(5 \le X \le 6) = \int_5^6 .15e^{-.15(x-.5)}\,dx = \int_{4.5}^{5.5} .15e^{-.15u}\,du = -e^{-.15u}\Big]_{4.5}^{5.5} = .071.$

11.

a. $P(X \le 1) = F(1) = \dfrac{1^2}{4} = .25$.

b. $P(.5 \le X \le 1) = F(1) - F(.5) = \dfrac{1^2}{4} - \dfrac{.5^2}{4} = .1875$.

c. $P(X > 1.5) = 1 - P(X \le 1.5) = 1 - F(1.5) = 1 - \dfrac{1.5^2}{4} = .4375$.

d. $.5 = F(\eta) = \dfrac{\eta^2}{4} \Rightarrow \eta^2 = 2 \Rightarrow \eta = \sqrt{2} \approx 1.414$ hours.

e. $f(x) = F'(x) = \dfrac{x}{2}$ for $0 \le x < 2$, and $= 0$ otherwise.

13.

a. $1 = \int_1^\infty \dfrac{k}{x^4}dx = k\int_1^\infty x^{-4}dx = \dfrac{k}{-3}x^{-3}\Big]_1^\infty = 0 - \left(\dfrac{k}{-3}\right)(1)^{-3} = \dfrac{k}{3} \Rightarrow k = 3$.

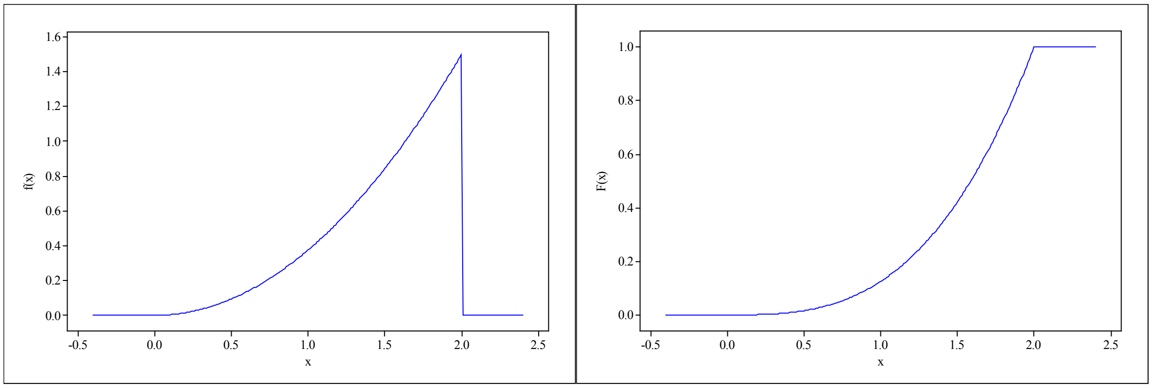
b. For $x \ge 1$, $F(x) = \int_{-\infty}^x f(y)dy = \int_1^x \dfrac{3}{y^4}dy = -y^{-3}\Big|_1^x = -x^{-3} + 1 = 1 - \dfrac{1}{x^3}$. For $x < 1$, $F(x) = 0$ since the distribution begins at 1. Put together, $F(x) = \begin{cases} 0 & x < 1 \\ 1 - \dfrac{1}{x^3} & 1 \le x \end{cases}$.

c. $P(X > 2) = 1 - F(2) = 1 - \frac{7}{8} = \frac{1}{8}$ or .125;
$P(2 < X < 3) = F(3) - F(2) = \left(1 - \frac{1}{27}\right) - \left(1 - \frac{1}{8}\right) = .963 - .875 = .088$.

15.

a. Since X is restricted to the interval [0, 2], $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 2$.

For $0 \le x \le 2$, $F(x) = \int_0^x \frac{3}{8}y^2 dy = \frac{1}{8}y^3\Big]_0^x = \frac{x^3}{8}$. Both graphs appear below.



b. $P(X \le .5) = F(.5) = \frac{(.5)^3}{8} = \frac{1}{64} = .015625$.

c. $P(.25 < X \le .5) = F(.5) - F(.25) = .015625 - .001953125 = .0137$.
Since X is continuous, $P(.25 \le X \le .5) = P(.25 < X \le .5) = .0137$.

d. The 75th percentile is the value of x for which $F(x) = .75$: $\frac{x^3}{8} = .75 \Rightarrow x = 1.817$.

17.

a. $P(X \le 1) = F(1) = .25[1 + \ln(4)] = .597$.

b. $P(1 \le X \le 3) = F(3) - F(1) = .966 - .597 = .369$.

c. For $x < 0$ or $x > 4$, the pdf is $f(x) = 0$ since X is restricted to (0, 4). For $0 < x < 4$, take the first derivative of the cdf:

$$F(x) = \frac{x}{4}\left[1 + \ln\left(\frac{4}{x}\right)\right] = \frac{1}{4}x + \frac{\ln(4)}{4}x - \frac{1}{4}x\ln(x) \Rightarrow$$

$$f(x) = F'(x) = \frac{1}{4} + \frac{\ln(4)}{4} - \frac{1}{4}\ln(x) - \frac{1}{4}x\frac{1}{x} = \frac{\ln(4)}{4} - \frac{1}{4}\ln(x) = .3466 - .25\ln(x)$$

Section 3.2

19. From part **e** of Exercise 11, $f(x) = x/2$ for $0 < x < 2$ and $f(x) = 0$ otherwise.

a. $E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_0^2 x \cdot \frac{x}{2}dx = \frac{1}{2}\int_0^2 x^2 dx = \frac{x^3}{6}\Big]_0^2 = \frac{8}{6} \approx 1.333$ hours.

b. $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^2 x^2 \frac{x}{2}dx = \frac{1}{2}\int_0^2 x^3 dx = \frac{x^4}{8}\Big]_0^2 = 2$, so $\text{Var}(X) = E(X^2) - [E(X)]^2 =$

$2 - \left(\frac{8}{6}\right)^2 = \frac{8}{36} \approx .222$, and $\sigma_X = \sqrt{.222} = .471$ hours.

c. From part **b**, $E(X^2) = 2$, so the expected charge is $2.

21.

a. $E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_0^1 x \cdot 90x^8(1-x)dx = \int_0^1 (90x^9 - 90x^{10})dx = 9x^{10} - \frac{90}{11}x^{11}\Big]_0^1 = \frac{9}{11} \approx .8182$ ft^3.

Similarly, $E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int_0^1 x^2 \cdot 90x^8(1-x)dx = \ldots = .6818$, from which
$\text{Var}(X) = .6818 - (.8182)^2 = .0124$ and $\text{SD}(X) = .11134$ ft^3.

b. $\mu \pm \sigma = (.7068, .9295)$. Thus, $P(\mu - \sigma \le X \le \mu + \sigma) = F(.9295) - F(.7068) = .8465 - .1602 = .6863$, and the probability X is more than 1 standard deviation from its mean value equals $1 - .6863 = .3137$.

23.

a. To find the $(100p)$th percentile, set $F(\eta_p) = p$ and solve for η_p: $\frac{\eta_p - A}{B - A} = p \Rightarrow \eta_p = A + (B - A)p$.

b. Set $p = .5$ to obtain $\eta = \eta_{.5} = A + (B - A)(.5) = .5B + .5A = \frac{A+B}{2}$. This is exactly the same as the mean of X, which is no surprise: since the uniform distribution is symmetric about $\frac{A+B}{2}$, $\mu = \eta = \frac{A+B}{2}$.

c. $E(X^n) = \int_A^B x^n \cdot \frac{1}{B-A}dx = \frac{1}{B-A}\frac{x^{n+1}}{n+1}\Big]_A^B = \frac{B^{n+1} - A^{n+1}}{(n+1)(B-A)}$.

25. $E(\text{area}) = E(\pi R^2) = \int_{-\infty}^{\infty} \pi r^2 f(r)dr = \int_9^{11} \pi r^2 \frac{3}{4}\left(1 - (10-r)^2\right)dr = \cdots = \frac{501}{5}\pi = 314.79$ m^2.

27. With X = temperature in °C, the temperature in °F equals $1.8X + 32$, so the mean and standard deviation in °F are $1.8\mu_X + 32 = 1.8(120) + 32 = 248$°F and $|1.8|\sigma_X = 1.8(2) = 3.6$°F. Notice that the additive constant, 32, affects the mean but does not affect the standard deviation.

29. First, $E(X)=\int_0^\infty x\cdot 4e^{-4x}\,dx$. Apply integration by parts with $u = x \to du = dx$ and $dv = 4e^{-4x}\,dx \to v = -e^{-4x}$:

$$E(X)=uv-\int v\,du = x\cdot(-e^{-4x})\Big|_0^\infty-\int_0^\infty -e^{-4x}\,dx=(0-0)+\int_0^\infty e^{-4x}\,dx=\frac{e^{-4x}}{-4}\Big|_0^\infty=\frac{1}{4}\text{ min.}$$

Similarly, $E(X^2)=\int_0^\infty x^2\cdot 4e^{-4x}\,dx$; with $u = x^2$ and $dv = 4e^{-4x}\,dx$,

$$E(X^2)=x^2\cdot(-e^{-4x})\Big|_0^\infty-\int_0^\infty(-e^{-4x})(2x)dx=(0-0)+\int_0^\infty 2xe^{-4x}\,dx=\int_0^\infty 2xe^{-4x}\,dx=\frac{1}{2}\int_0^\infty x\cdot 4e^{-4x}\,dx=$$

$\frac{1}{2}E(X)=\frac{1}{2}\cdot\frac{1}{4}=\frac{1}{8}$. Hence, Var($X$) = $\frac{1}{8}-\left(\frac{1}{4}\right)^2=\frac{1}{16}\Rightarrow$ SD(X) = $\frac{1}{4}$ min.

31. $h_1(X)=h(\mu)+h'(\mu)(X-\mu)=h'(\mu)X+h(\mu)-\mu h'(\mu)$, a linear function of X. Use the linear rescaling properties of mean and variance for parts **a** and **b**.

a. $E[h_1(X)]=E[h'(\mu)X+h(\mu)-\mu h'(\mu)]=h'(\mu)E[X]+h(\mu)-\mu h'(\mu)=\mu h'(\mu)+h(\mu)-\mu h'(\mu)=h(\mu)$

b. $\text{Var}(h_1(X))=\text{Var}(h'(\mu)X+h(\mu)-\mu h'(\mu))=\text{Var}(h'(\mu)X)=[h'(\mu)]^2\,\text{Var}(X)$

c. We have $R = h(I) = v/I$, so $h'(I) = -v/I^2$. The first-order approximation to μ_R is $h(\mu_I) = v/\mu_I = v/20$. The first-order approximation to σ_R^2 is $[h'(\mu_I)]^2\,\text{Var}(I)=\left[-\frac{v}{\mu_I^2}\right]^2\cdot\sigma_I^2=\frac{v^2}{(20^2)^2}\cdot(.5)^2=\frac{v^2}{640{,}000}$; taking the square root, the first-order approximation to σ_R is $\frac{v}{800}$.

d. From Exercise 25, the exact value of $E[\pi R^2]$ was $100.2\pi \approx 314.79$ m^2. The first-order approximation via the delta method is $h(\mu_R) = h(10) = \pi(10)^2 = 100\pi \approx 314.16$ m^2. These are quite close.

e. The derivative of $h(R) = \pi R^2$ is $h'(R) = 2\pi R$. The delta method approximation to Var[$h(R)$], therefore, is $[h'(\mu_R)]^2\cdot\text{Var}(R)=[2\pi\mu_R]^2\cdot\text{Var}(R)=[2\pi(10)]^2\cdot\frac{1}{5}=80\pi^2$. This is very close to the exact variance, given by $14008\pi^2/175 \approx 80.046\pi^2$.

33. A linear function that maps 0 to –5 and 1 to 5 is $g(x) = 10x - 5$. Let $X \sim$ Unif[0, 1], so that from Exercise 32 we know that $M_X(t) = (e^t - 1)/t$ for $t \neq 0$. Define $Y = g(X) = 10X - 5$; applying the mgf rescaling property with $a = 10$ and $b = -5$, the mgf of Y is given by $M_Y(t)=e^{-5t}M_X(10t)=e^{-5t}\cdot\frac{e^{(10t)}-1}{(10t)}=\frac{e^{5t}-e^{-5t}}{10t}$. This is an exact match to the mgf of the Unif[–5, 5] mgf based on Exercise 32. Therefore, by uniqueness of mgfs, Y must follow a Unif[–5, 5] distribution. (Equivalently, the pdf of Y is $f(y) = 1/10$ for $-5 \le y \le 5$.)

35. $f(x) = .15e^{-.15(x-.5)}$ for $x \geq .5$.

a.
$$M_X(t) = \int_{.5}^{\infty} e^{tx} \cdot .15e^{-.15(x-.5)}\,dx = .15e^{+.075}\int_{.5}^{\infty} e^{(t-.15)x}\,dx = .15e^{+.075}\left.\frac{e^{(t-.15)x}}{t-.15}\right|_{.5}^{\infty}$$
$$= .15e^{+.075}\left[0 - \frac{e^{(t-.15)(.5)}}{t-.15}\right] = \frac{.15e^{.5t}}{.15-t} \qquad \text{for } t < .15$$

The condition $t < .15$ is necessary so that $(t - .15) < 0$ and the improper integral converges.
To find the mean and variance, re-write the mgf as $M_X(t) = .15e^{.5t}(.15-t)^{-1}$ and use the product rule:

$$M_X'(t) = \frac{.075e^{.5t}}{.15-t} + \frac{.15e^{.5t}}{(.15-t)^2} \Rightarrow E(X) = M_X'(0) = \cdots = \frac{43}{6} = 7.1\overline{6}\text{ sec}$$

$$M_X''(t) = \frac{.0375e^{.5t}}{.15-t} + \frac{2(.075)e^{.5t}}{(.15-t)^2} + \frac{.3e^{.5t}}{(.15-t)^3} \Rightarrow E(X^2) = M_X''(0) = \cdots = 95.80\overline{5} \Rightarrow$$

Var(X) = (95.805) – (7.16)^2 = 44.444.

b. The mgf of the given pdf is

$$M(t) = \int_0^{\infty} e^{tx} \cdot .15e^{-.15x}\,dx = .15\int_0^{\infty} e^{(t-.15)x}\,dx = .15\left.\frac{e^{(t-.15)x}}{t-.15}\right|_0^{\infty} = \frac{.15}{.15-t}$$ for $t < .15$. Taking derivatives here

(which is much easier than in part **a**!) gives $E(X) = M'(0) = \dfrac{1}{.15} = 6.6\overline{6}$,

$E(X^2) = M''(0) = \dfrac{2}{.15^2} = 88.\overline{8}$, and Var(X) = 44.444. The time headway pdf and this pdf have exactly the same variance, while the mean of the time headway pdf is .5 more than the mean of this pdf.

c. If $Y = X - .5$, then $M_Y(t) = e^{-.5t}M_X(1t) = e^{-.5t} \cdot \dfrac{.15e^{.5t}}{.15-t} = \dfrac{.15}{.15-t}$, which is exactly the mgf in **b**. By uniqueness of mgfs, we conclude that Y follows the pdf specified in **b**: $f(y) = .15e^{-.15y}$ for $y > 0$.
In other words, the two pdfs represent "shifted" versions of two variables, X and Y. The rv X is on $[.5, \infty)$, while $Y = X - .5$ is on $[0, \infty)$. This is consistent with the moments as well: the mean of Y is .5 less than the mean of X, as suggested by $Y = X - .5$, and the two rvs have the same variance because the shift of .5 doesn't affect variance.

37.

a. For $t \leq x < \infty$, $x \cdot f(x) \geq t \cdot f(x)$. Thus,
$$\int_t^{\infty} x \cdot f(x)dx \geq \int_t^{\infty} t \cdot f(x)dx = t \cdot \int_t^{\infty} f(x)dx = t \cdot P(X \geq t) = t \cdot [1 - F(t)] .$$

b. By definition, $\mu = \int_0^{\infty} x \cdot f(x)dx = \int_0^{t} x \cdot f(x)dx + \int_t^{\infty} x \cdot f(x)dx$, from which it follows that
$$\int_t^{\infty} x \cdot f(x)dx = \mu - \int_0^{t} x \cdot f(x)dx .$$

c. Now consider the expression $t \cdot [1 - F(t)]$. Since $t > 0$ and $F(t) \le 1$, $t \cdot [1 - F(t)] \ge 0$. Combining that with part **a**, we have $0 \le t \cdot [1 - F(t)] \le \int_t^\infty x \cdot f(x)dx = \mu - \int_0^t x \cdot f(x)dx$.
As $t \to \infty$, the upper bound on $t \cdot [1 - F(t)]$ converges to 0:
$\lim_{t\to\infty}\left[\mu - \int_0^t x \cdot f(x)dx\right] = \mu - \lim_{t\to\infty}\left[\int_0^t x \cdot f(x)dx\right] = \mu - \int_0^\infty x \cdot f(x)dx = \mu - \mu = 0$.
(Those operations rely on the integral converging.)

Therefore, by the squeeze theorem, $\lim_{t\to\infty} t \cdot [1 - F(t)] = 0$ as well.

Section 3.3

39.

a. $P(0 \le Z \le 2.17) = \Phi(2.17) - \Phi(0) = .4850$.

b. $\Phi(1) - \Phi(0) = .3413$.

c. $\Phi(0) - \Phi(-2.50) = .4938$.

d. $\Phi(2.50) - \Phi(-2.50) = .9876$.

e. $\Phi(1.37) = .9147$.

f. $P(-1.75 < Z) + [1 - P(Z < -1.75)] = 1 - \Phi(-1.75) = .9599$.

g. $\Phi(2) - \Phi(-1.50) = .9104$.

h. $\Phi(2.50) - \Phi(1.37) = .0791$.

i. $1 - \Phi(1.50) = .0668$.

j. $P(|Z| \le 2.50) = P(-2.50 \le Z \le 2.50) = \Phi(2.50) - \Phi(-2.50) = .9876$.

41.

a. $\Phi(c) = .9100 \Rightarrow c \approx 1.34$, since .9099 is the entry in the 1.3 row, .04 column.

b. Since the standard normal distribution is symmetric about $z = 0$, the 9th percentile =
–[the 91st percentile] = –1.34.

c. $\Phi(c) = .7500 \Rightarrow c \approx .675$, since .7486 and .7517 are in the .67 and .68 entries, respectively.

d. Since the standard normal distribution is symmetric about $z = 0$, the 25th percentile =
–[the 75th percentile] = –.675.

e. $\Phi(c) = .06 \Rightarrow c \approx -1.555$, since .0594 and .0606 appear as the –1.56 and –1.55 entries, respectively.

43.

a. $P(X \le 50) = P\left(Z \le \frac{50-46.8}{1.75}\right) = P(Z \le 1.83) = \Phi(1.83) = .9664.$

b. $P(X \ge 48) = P\left(Z \ge \frac{48-46.8}{1.75}\right) = P(Z \ge 0.69) = 1 - \Phi(0.69) = 1 - .7549 = .2451.$

c. The mean and standard deviation aren't important here. The probability a normal random variable is within 1.5 standard deviations of its mean equals $P(-1.5 \le Z \le 1.5) =$ $\Phi(1.5) - \Phi(-1.5) = .9332 - .0668 = .8664.$

45. $X \sim N(119, 13.1)$.

a. $P(100 \le X \le 120) = \Phi\left(\frac{120-119}{13.1}\right) - \Phi\left(\frac{100-119}{13.1}\right) \approx \Phi(0.08) - \Phi(-1.45) = .5319 - .0735 = .4584.$

b. The goal is to find the speed, s, so that $P(X > s) = 10\% = .1$ (the fastest 10%). That's equivalent to $P(X \le s) = 1 - .1 = .9$ (the 90th percentile), so $.9 = \Phi\left(\frac{s-119}{13.1}\right) \Rightarrow \frac{s-119}{13.1} \approx 1.28 \Rightarrow s = 119 +$ $1.28(13.1) \approx 135.8$ kph.

c. $P(X > 100) = 1 - \Phi\left(\frac{100-119}{13.1}\right) = 1 - \Phi(-1.45) = 1 - .0735 = .9265.$

d. P(at least one is <u>not</u> exceeding 100 kph) = 1 – P(all five are exceeding 100 kph). Using independence and the answer from **c**, this equals 1 – P(first > 100 kph) × … × P(fifth > 100 kph) = $1 - (.9265)^5 =$ .3173.

e. Convert: 70 miles per hour ≈ 112.65 kilometers per hour. Thus $P(X > 70 \text{ mph}) = P(X > 112.65 \text{ kph}) =$ $1 - \Phi\left(\frac{112.65-119}{13.1}\right) = 1 - \Phi(-.48) = 1 - .3156 = .6844.$

47. $X \sim N(200, 35)$.

a. $P(X \le 250) = \Phi\left(\frac{250-200}{35}\right) = \Phi(1.43) = .9236.$

b. $P(300 \le X \le 400) = \Phi\left(\frac{400-200}{35}\right) - \Phi\left(\frac{300-200}{35}\right) = \Phi(5.71) - \Phi(2.86) = 1 - .9979 = .0021.$

c. The mean and standard deviation are actually irrelevant. For any normal distribution, $P(|X - \mu| > 1.5\sigma) = P(|Z| > 1.5) = 1 - P(-1.5 \le Z \le 1.5) = 1 - [\Phi(1.5) - \Phi(-1.5)] = 1 - [.9332 - .0668] =$ .1336.

49. Let X denote the diameter of a randomly selected cork made by the first machine, and let Y be defined analogously for the second machine.
$P(2.9 \le X \le 3.1) = P(-1.00 \le Z \le 1.00) = .6826$, while
$P(2.9 \le Y \le 3.1) = P(-7.00 \le Z \le 3.00) = .9987$.
So, the second machine wins handily.

51.

a. $P(X < 40) = P\left(Z \le \frac{40-43}{4.5}\right) = P(Z < -0.667) = .2514.$

$P(X > 60) = P\left(Z > \frac{60-43}{4.5}\right) = P(Z > 3.778) \approx 0.$

b. We desire the 25th percentile. Since the 25th percentile of a standard normal distribution is roughly $z = -0.67$, the answer is 43 + (–0.67)(4.5) = 39.985 ksi.

53. The probability X is within .1 of its mean is given by $P(\mu - .1 \le X \le \mu + .1) =$ $P\left(\frac{(\mu-.1)-\mu}{\sigma} < Z < \frac{(\mu+.1)-\mu}{\sigma}\right) = \Phi\left(\frac{.1}{\sigma}\right) - \Phi\left(-\frac{.1}{\sigma}\right) = 2\Phi\left(\frac{.1}{\sigma}\right) - 1$. If we require this to equal 95%, we find $2\Phi\left(\frac{.1}{\sigma}\right) - 1 = .95 \Rightarrow \Phi\left(\frac{.1}{\sigma}\right) = .975 \Rightarrow \frac{.1}{\sigma} = 1.96$ from the standard normal table. Thus, $\sigma = \frac{.1}{1.96} = .0510$.
Alternatively, use the empirical rule: 95% of all values lie within 2 standard deviations of the mean, so we want $2\sigma = .1$, or $\sigma = .05$. (This is not quite as precise as the first answer.)

55.

a. $P(\mu - 1.5\sigma \le X \le \mu + 1.5\sigma) = P(-1.5 \le Z \le 1.5) = \Phi(1.50) - \Phi(-1.50) = .8664.$

b. $P(X < \mu - 2.5\sigma \text{ or } X > \mu + 2.5\sigma) = 1 - P(\mu - 2.5\sigma \le X \le \mu + 2.5\sigma)$
$= 1 - P(-2.5 \le Z \le 2.5) = 1 - .9876 = .0124.$

c. $P(\mu - 2\sigma \le X \le \mu - \sigma \text{ or } \mu + \sigma \le X \le \mu + 2\sigma) = P(\text{within 2 sd's}) - P(\text{within 1 sd}) =$
$P(\mu - 2\sigma \le X \le \mu + 2\sigma) - P(\mu - \sigma \le X \le \mu + \sigma) = .9544 - .6826 = .2718.$

57.

a. $P(67 < X < 75) = P\left(\frac{67-70}{3} < \frac{X-70}{3} < \frac{75-70}{3}\right) = P(-1 < Z < 1.67) = \Phi(1.67) - \Phi(-1) = .9525 -$ $.1587 = .7938.$

b. By the Empirical Rule, c should equal 2 standard deviations. Since $\sigma = 3$, $c = 2(3) = 6$.
We can be a little more precise, as in Exercise 42, and use $c = 1.96(3) = 5.88$.

c. Let Y = the number of acceptable specimens out of 10, so $Y \sim \text{Bin}(10, p)$, where $p = .7938$ from part **a**.
Then $E(Y) = np = 10(.7938) = 7.938$ specimens.

d. Now let Y = the number of specimens out of 10 that have a hardness of less than 73.84, so $Y \sim \text{Bin}(10, p)$, where

$p = P(X < 73.84) = P\left(Z < \frac{73.84-70}{3}\right) = P(Z < 1.28) = \Phi(1.28) = .8997.$ Then

$P(Y \le 8) = \sum_{y=0}^{8} \binom{10}{y} (.8997)^y (.1003)^{10-y} = .2651.$

You can also compute $1 - P(Y = 9, 10)$ and use the binomial formula, or round slightly to $p = .9$ and use the binomial table: $P(Y \le 8) = B(8; 10, .9) = .265.$

59.

a. By symmetry, $P(-1.72 \le Z \le -.55) = P(.55 \le Z \le 1.72) = \Phi(1.72) - \Phi(.55)$.

b. $P(-1.72 \le Z \le .55) = \Phi(.55) - \Phi(-1.72) = \Phi(.55) - [1 - \Phi(1.72)]$.

No, thanks to the symmetry of the z curve about 0.

61.

a. $P(20 \le X \le 30) = P(20 - .5 \le X \le 30 + .5) = P(19.5 \le X \le 30.5) = P(-1.1 \le Z \le 1.1) = .7286$.

b. $P(X \le 30) = P(X \le 30.5) = P(Z \le 1.1) = .8643$, while
$P(X < 30) = P(X \le 29.5) = P(Z < .9) = .8159$.

63. Use the normal approximation to the binomial, with a continuity correction. With $p = .10$ and $n = 200$, $\mu = np = 20$, and $\sigma^2 = npq = 18$. So, Bin(200, .10) $\approx N(20, \sqrt{18})$.

a. $P(X \le 30) = \Phi\left(\dfrac{(30+.5)-20}{\sqrt{18}}\right) = \Phi(2.47) = .9932$.

b. $P(X < 30) = P(X \le 29) = \Phi\left(\dfrac{(29+.5)-20}{\sqrt{18}}\right) = \Phi(2.24) = .9875$.

c. $P(15 \le X \le 25) = P(X \le 25) - P(X \le 14) = \Phi\left(\dfrac{(25+.5)-20}{\sqrt{18}}\right) - \Phi\left(\dfrac{(14+.5)-20}{\sqrt{18}}\right)$

$= \Phi(1.30) - \Phi(-1.30) = .9032 - .0968 = .8064$.

65. We use a normal approximation to the binomial distribution: Let X denote the number of people in the sample of 1000 who <u>can</u> taste the difference, so $X \sim$ Bin(1000, .03). Because $\mu = np = 1000(.03) = 30$ and $\sigma = \sqrt{np(1-p)} = 5.394$, X is approximately $N(30, 5.394)$.

a. Using a continuity correction, $P(X \ge 40) = 1 - P(X \le 39) = 1 - P\left(Z \le \dfrac{39.5-30}{5.394}\right) = 1 - P(Z \le 1.76) =$

$1 - \Phi(1.76) = 1 - .9608 = .0392$.

b. 5% of 1000 is 50, and $P(X \le 50) = P\left(Z \le \dfrac{50.5-30}{5.394}\right) = \Phi(3.80) \approx 1$.

67. As in the previous exercise, $u = (x - \mu)/\sigma \rightarrow du = dx/\sigma$. Here, we'll also need $x = \mu + \sigma u$.

a. $E(X) = \int_{-\infty}^{\infty} x \cdot f(x;\mu,\sigma)dx = \int_{-\infty}^{\infty} \dfrac{x}{\sqrt{2\pi}} e^{-[(x-\mu)/\sigma]^2/2} \dfrac{dx}{\sigma} = \int_{-\infty}^{\infty} \dfrac{\mu+\sigma u}{\sqrt{2\pi}} e^{-u^2/2} du =$

$\mu \int_{-\infty}^{\infty} \dfrac{1}{\sqrt{2\pi}} e^{-u^2/2} du + \sigma \int_{-\infty}^{\infty} \dfrac{u}{\sqrt{2\pi}} e^{-u^2/2} du$. The first integral is the area under the standard normal pdf, which equals 1. The second integrand is an odd function over a symmetric interval, so that second integral equals 0. Put it all together: $E(X) = \mu(1) + \sigma(0) = \mu$.

b. $\text{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x;\mu,\sigma)dx = \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sqrt{2\pi}} e^{-[(x-\mu)/\sigma]^2/2} \frac{dx}{\sigma} = \int_{-\infty}^{\infty} \frac{(\sigma u)^2}{\sqrt{2\pi}} e^{-u^2/2} du =$

$\sigma^2 \int_{-\infty}^{\infty} u^2 \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$. We need to show the integral equals 1; toward this goal, use integration by parts with $u = u$ and $dv = u \cdot \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \rightarrow v = -\frac{1}{\sqrt{2\pi}} e^{-u^2/2}$:

$$\int_{-\infty}^{\infty} u^2 \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = -u \cdot u \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -\frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = 0 + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du .$$

Both limits required for the first evaluation are 0 because the exponential term dominates. What remains is, once again, the integral of the standard normal pdf, which is 1.
Therefore, $\text{Var}(X) = \sigma^2(1) = \sigma^2$.

69.

a. $P(Z \geq 1) \approx .5 \cdot \exp\left(\frac{83+351+562}{703+165}\right) = .1587$, which matches $1 - \Phi(1)$.

b. $P(Z < -3) = P(Z > 3) \approx .5 \cdot \exp\left(\frac{-2362}{399.3333}\right) = .0013$, which matches $\Phi(-3)$.

c. $P(Z > 4) \approx .5 \cdot \exp\left(\frac{-3294}{340.75}\right) = .0000317$, so $P(-4 < Z < 4) = 1 - 2P(Z \geq 4) \approx$ $1 - 2(.0000317) = .999937$.

d. $P(Z > 5) \approx .5 \cdot \exp\left(\frac{-4392}{305.6}\right) = .00000029$.

Section 3.4

71.

a. $E(X) = \frac{1}{\lambda} = 1$.

b. $\sigma = \frac{1}{\lambda} = 1$.

c. Using the exponential cdf, $P(X \leq 4) = 1 - e^{-(1)(4)} = 1 - e^{-4} = .982$.

d. Similarly, $P(2 \leq X \leq 5) = (1 - e^{-(1)(5)}) - (1 - e^{-(1)(2)}) = e^{-2} - e^{-5} = .129$.

73. Note that a mean value of 10 for the exponential distribution implies $\lambda = \dfrac{1}{10} = .1$. Let X denote the survival time of a mouse without treatment.

a. $P(X \geq 8) = 1 - [1 - e^{-(.1)(8)}] = e^{-(.1)(8)} = .4493$. $P(X \leq 12) = 1 - e^{-(.1)(12)} = .6988$. Combining these two answers, $P(8 \leq X \leq 12) = P(X \leq 12) - P(X < 8) = .6988 - [1 - .4493] = .1481$.

b. The standard deviation equals the mean, 10 hours. So, $P(X > \mu + 2\sigma) = P(X > 30) = 1 - [1 - e^{-(.1)(30)}] = e^{-(.1)(30)} = .0498$. Similarly, $P(X > \mu + 3\sigma) = P(X > 40) = e^{-(.1)(40)} = .0183$.

75. For **a** and **b**, we use the properties of the gamma function provided in this section.

a. $\Gamma(6) = 5! = 120$.

b. $\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{3}{2}\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right) = \left(\frac{3}{4}\right)\sqrt{\pi} \approx 1.329$.

c. $G(4; 5) = .371$ from row 4, column 5 of Table A.4.

d. $G(5; 4) = .735$.

e. $G(0; 4) = P(X \leq 0$ when $\alpha = 4) = 0$, since the gamma distribution is positive.

77. Notice that $\mu = 24$ and $\sigma^2 = 12^2 = 144 \Rightarrow \alpha\beta = 24$ and $\alpha\beta^2 = 144 \Rightarrow \beta = \dfrac{144}{24} = 6$ and $\alpha = \dfrac{24}{\beta} = 4$. Hence, the cdf of X is $G\left(\frac{x}{6}; 4\right)$.

a. $P(12 \leq X \leq 24) = G(4; 4) - G(2; 4) = .424$.

b. $P(X \leq 24) = G(4; 4) = .567$, so while the mean is 24, the median is <u>less</u> than 24, since $P(X \leq \eta) = .5$ (and .567 is more than that). This is a result of the positive skew of the gamma distribution.

c. We want a value x for which $G\left(\frac{x}{6}, 4\right) = .99$. In Table A.4, we see $G(10; 4) = .990$. So $x/6 = 10$, and the 99th percentile is $6(10) = 60$ weeks.

d. We want a value t for which $P(X > t) = .005$, i.e. $P(X \leq t) = .995$. The left-hand side is the cdf of X, so we really want $G\left(\frac{t}{6}, 4\right) = .995$. In Table A.4, $G(11; 4) = .995$, so $t/6 = 11$ and $t = 6(11) = 66$. At 66 weeks, only .5% of all transistors would still be operating.

79. To find the (100p)th percentile, set $F(x) = p$ and solve for x: $p = F(x) = 1 - e^{-\lambda x} \Rightarrow e^{-\lambda x} = 1 - p \Rightarrow -\lambda x = \ln(1 - p) \Rightarrow x = -\dfrac{\ln(1-p)}{\lambda}$.

To find the median, set $p = .5$ to get $\eta = -\dfrac{\ln(1-.5)}{\lambda} = \dfrac{.693}{\lambda}$.

81.

a. $P(X \le 5) = \int_1^5 .15e^{-.15(x-1)}\,dx = \int_0^4 .15e^{-.15u}\,du = 1 - e^{-.15(4)}$ = .4512. $P(X > 5) = 1 - .4512 = .5488$.

b. $P(2 \le X \le 5) = \int_2^5 .15e^{-.15(x-1)}\,dx = \int_1^4 .15e^{-.15u}\,du = .3119$.

c. Use the hint: we may write $Y = X - 1$ or $X = Y + 1$, where Y has an (unshifted) exponential distribution with $\lambda = .15$. Hence, $E(X) = E(Y) + 1 = \frac{1}{\lambda} + 1 = \frac{1}{.15} + 1$ = 7.667 sec.

d. Similarly, $X = Y + 1 \Rightarrow \text{SD}(X) = \text{SD}(Y) = \frac{1}{\lambda} = \frac{1}{.15}$ = 6.667 sec.

83.

a. Using (3.5), for any positive exponent k we have

$$E(X^k) = \int_0^\infty x^k \cdot \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}\,dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x^{k+\alpha-1} e^{-x/\beta}\,dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \cdot \beta^{k+\alpha}\Gamma(k+\alpha)$$
$$= \frac{\Gamma(k+\alpha)}{\Gamma(\alpha)}\beta^k$$

So, $E(X) = \frac{\Gamma(1+\alpha)}{\Gamma(\alpha)}\beta^1 = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)}\beta = \alpha\beta$; $E(X^2) = \frac{\Gamma(2+\alpha)}{\Gamma(\alpha)}\beta^2 = \frac{(\alpha+1)\Gamma(\alpha+1)}{\Gamma(\alpha)}\beta^2 = \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)}$
$= (\alpha+1)\alpha\beta^2$; and $\text{Var}(X) = (\alpha+1)\alpha\beta^2 - [\alpha\beta]^2 = \alpha\beta^2$.

b. The mgf of the gamma distribution is $M_X(t) = \frac{1}{(1-\beta t)^\alpha}$. Take derivatives:

$$M'_X(t) = \frac{-\alpha}{(1-\beta t)^{\alpha+1}}(-\beta) = \frac{\alpha\beta}{(1-\beta t)^{\alpha+1}} \Rightarrow E(X) = M'_X(0) = \alpha\beta$$
$$M''_X(t) = \frac{-\alpha\beta(\alpha+1)}{(1-\beta t)^{\alpha+2}}(-\beta) = \frac{(\alpha+1)\alpha\beta^2}{(1-\beta t)^{\alpha+2}} \Rightarrow E(X^2) = M''_X(0) = (\alpha+1)\alpha\beta^2$$

$\text{Var}(X) = (\alpha+1)\alpha\beta^2 - [\alpha\beta]^2 = \alpha\beta^2$.

Section 3.5

85.

a. $P(X \le 250) = F(250; 2.5, 200) = 1 - e^{-(250/200)^{2.5}} = 1 - e^{-1.75} = .8257$.
$P(X < 250) = P(X \le 250) = .8257$.
$P(X > 300) = 1 - F(300; 2.5, 200) = e^{-(1.5)^{2.5}} = .0636$.

b. $P(100 \le X \le 250) = F(250; 2.5, 200) - F(100; 2.5, 200) = .8257 - .162 = .6637$.

c. The question is asking for the median, η. Solve $F(\eta) = .5$: $.5 = 1 - e^{-(\eta/200)^{2.5}} \Rightarrow$
$e^{-(\eta/200)^{2.5}} = .5 \Rightarrow (\eta/200)^{2.5} = -\ln(.5) \Rightarrow \eta = 200(-\ln(.5))^{1/2.5}$ = 172.727 hours.

87. Use the substitution $y = \left(\frac{x}{\beta}\right)^{\alpha} = \frac{x^{\alpha}}{\beta^{\alpha}}$. Then $dy = \frac{\alpha x^{\alpha-1}}{\beta^{\alpha}}dx$, and $\mu = \int_0^{\infty} x \cdot \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-(x/\beta)^{\alpha}}\,dx =$ $\int_0^{\infty} (\beta^{\alpha} y)^{1/\alpha} \cdot e^{-y} dy = \beta \int_0^{\infty} y^{1/\alpha} e^{-y} dy = \beta \cdot \Gamma\left(1 + \frac{1}{\alpha}\right)$ by definition of the gamma function.

89.

a. $P(X \le 105) = F(105; 20, 100) = 1 - e^{-(105/100)^{20}} = .9295$.

b. $P(100 \le X \le 105) = F(105; 20, 100) - F(100; 20, 100) = .9295 - .6321 = .2974$.

c. Set $.5 = F(\eta)$ and solve: $.5 = 1 - e^{-(\eta/100)^{20}} \Rightarrow -(\eta/100)^{20} = \ln(.5) \Rightarrow \eta = 100[-\ln(.5)]^{1/20} = 98.184$ ksi.

91. Notice that μ_X and σ_X are the mean and standard deviation of the lognormal variable X in this example; they are not the parameters μ and σ which usually refer to the mean and standard deviation of $\ln(X)$. We're given $\mu_X = 10{,}281$ and $\sigma_X/\mu_X = .40$, from which $\sigma_X = .40\mu_X = 4112.4$.

a. To find the mean and standard deviation of $\ln(X)$, set the lognormal mean and variance equal to the appropriate quantities: $10{,}281 = E(X) = e^{\mu+\sigma^2/2}$ and $(4112.4)^2 = \text{Var}(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$. Square the first equation: $(10{,}281)^2 = e^{2\mu+\sigma^2}$. Now divide the variance by this amount:

$$\frac{(4112.4)^2}{(10{,}281)^2} = \frac{e^{2\mu+\sigma^2}(e^{\sigma^2}-1)}{e^{2\mu+\sigma^2}} \Rightarrow e^{\sigma^2} - 1 = (.40)^2 = .16 \Rightarrow \sigma = \sqrt{\ln(1.16)} = .38525$$

That's the standard deviation of $\ln(X)$. Use this in the formula for $E(X)$ to solve for μ:
$10{,}281 = e^{\mu+(.38525)^2/2} = e^{\mu+.07421} \Rightarrow \mu = 9.164$. That's $E(\ln(X))$.

b. $P(X \le 15{,}000) = P\left(Z \le \frac{\ln(15{,}000) - 9.164}{.38525}\right) = P(Z \le 1.17) = \Phi(1.17) = .8790$.

c. $P(X \ge \mu_X) = P(X \ge 10{,}281) = P\left(Z \ge \frac{\ln(10{,}281) - 9.164}{.38525}\right) = P(Z \ge .19) = 1 - \Phi(0.19) = .4247$. Even though the normal distribution is symmetric, the lognormal distribution is not a symmetric distribution. (See the lognormal graphs in the textbook.) So, the mean and the median of X aren't the same and, in particular, the probability X exceeds its own mean doesn't equal .5.

d. One way to check is to determine whether $P(X < 17{,}000) = .95$; this would mean 17,000 is indeed the 95th percentile. However, we find that $P(X < 17{,}000) = \Phi\left(\frac{\ln(17{,}000) - 9.164}{.38525}\right) = \Phi(1.50) = .9332$, so 17,000 is not the 95th percentile of this distribution (it's the 93.32%ile).

93. $.5 = F(\eta) = \Phi\left(\frac{\ln(\eta) - \mu}{\sigma}\right)$, where η refers to the median of the lognormal distribution and μ and σ to the mean and sd of the normal distribution. Since $\Phi(0) = .5$, $\frac{\ln(\eta) - \mu}{\sigma} = 0$, so $\eta = e^{\mu}$. For the load distribution, $\eta = e^{9.164} = 9547$ kg/day/km.

95.

a. Var(X) = $e^{2(2.05)+.06}(e^{.06}-1)$ = 3.96 ⇒ SD(X) = 1.99 months.

b. $P(X > 12) = 1 - P(X \le 12) = 1 - P\left(Z \le \dfrac{\ln(12)-2.05}{\sqrt{.06}}\right)$ = 1 – Φ(1.78) = .0375.

c. The mean of X is $E(X) = e^{2.05+.06/2}$ = 8.00 months, so $P(\mu_X - \sigma_X < X < \mu_X + \sigma_X) = P(6.01 < X < 9.99)$ = $\Phi\left(\dfrac{\ln(9.99)-2.05}{\sqrt{.06}}\right) - \Phi\left(\dfrac{\ln(6.01)-2.05}{\sqrt{.06}}\right)$ = Φ(1.03) – Φ(–1.05) = .8485 – .1469 = .7016.

d. In Exercise 93, it was shown that $\eta = e^{\mu}$. Here, $\eta = e^{2.05}$ = 7.77 months.

e. Φ(2.33) = .99 ⇒ $\dfrac{\ln(\eta_{.99})-2.05}{\sqrt{.06}}$ = 2.33 ⇒ $\eta_{.99} = e^{2.62}$ = 13.75 months.

f. The probability of exceeding 8 months is $P(X > 8) = 1 - \Phi\left(\dfrac{\ln(8)-2.05}{\sqrt{.06}}\right)$ = 1 – Φ(.12) = .4522, so the expected number that will exceed 8 months out of n = 10 is just 10(.4522) = 4.522.

97. Since the standard beta distribution lies on (0, 1), the point of symmetry must be ½, so we require that $f\left(\frac{1}{2}-\mu\right) = f\left(\frac{1}{2}+\mu\right)$. Cancelling out the constants, this implies $\left(\frac{1}{2}-\mu\right)^{\alpha-1}\left(\frac{1}{2}+\mu\right)^{\beta-1} = \left(\frac{1}{2}+\mu\right)^{\alpha-1}\left(\frac{1}{2}-\mu\right)^{\beta-1}$, which (by matching exponents on both sides) in turn implies that $\alpha = \beta$.

Alternatively, symmetry about ½ requires μ = ½, so $\dfrac{\alpha}{\alpha+\beta}$ = .5. Solving for α gives $\alpha = \beta$.

99.

a. Notice from the definition of the standard beta pdf that, since a pdf must integrate to 1,

$$1 = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx \Rightarrow \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Using this, $E(X) = \int_0^1 x \cdot \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha}(1-x)^{\beta-1}\,dx =$

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)} = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)} = \frac{\alpha}{\alpha+\beta}.$$

b. Similarly, $E[(1-X)^m] = \int_0^1 (1-x)^m \cdot \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx =$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha-1}(1-x)^{m+\beta-1}\,dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha)\Gamma(m+\beta)}{\Gamma(\alpha+m+\beta)} = \frac{\Gamma(\alpha+\beta)\cdot\Gamma(m+\beta)}{\Gamma(\alpha+m+\beta)\Gamma(\beta)}.$$

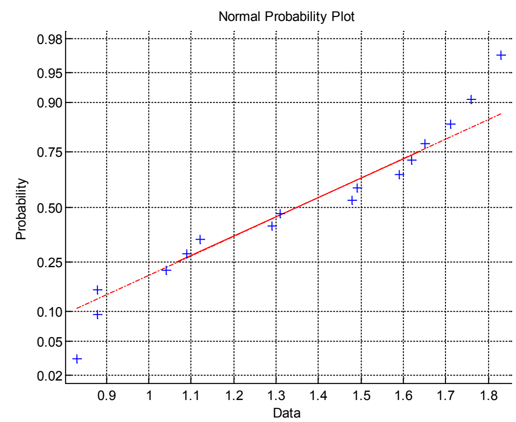
If X represents the proportion of a substance consisting of an ingredient, then 1 – X represents the proportion not consisting of this ingredient. For m = 1 above,

$$E(1-X) = \frac{\Gamma(\alpha+\beta)\cdot\Gamma(1+\beta)}{\Gamma(\alpha+1+\beta)\Gamma(\beta)} = \frac{\Gamma(\alpha+\beta)\cdot\beta\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)\Gamma(\beta)} = \frac{\beta}{\alpha+\beta}.$$

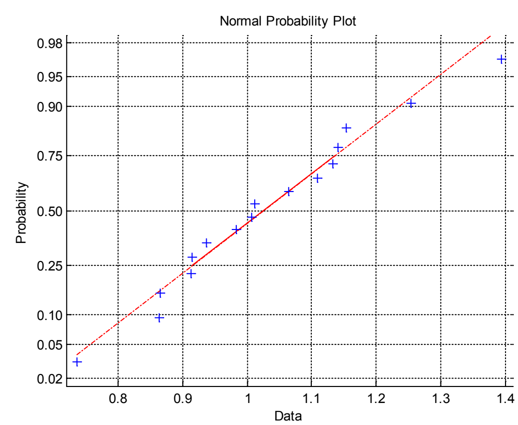
Section 3.6

101. The given probability plot is quite linear, and thus it is quite plausible that the tension distribution is normal.

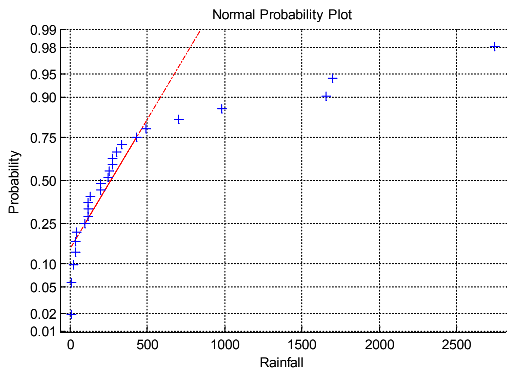
103. The *z* percentile values are as follows: –1.86, –1.32, –1.01, –0.78, –0.58, –0.40, –0.24,–0.08, 0.08, 0.24, 0.40, 0.58, 0.78, 1.01, 1.30, and 1.86. The accompanying probability plot has some curvature but (arguably) not enough to worry about. It would be reasonable to use estimating methods that assume a normal population distribution.

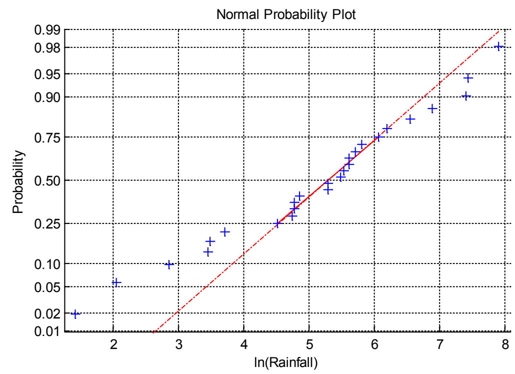


105. The *z* percentiles are the same as those in Exercise 103. The accompanying normal probability plot is straight, suggesting that an assumption of population normality is plausible.

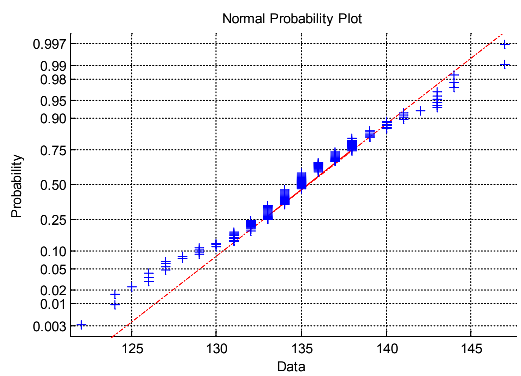


107. To check for plausibility of a lognormal population distribution for this rainfall data, take the natural logs and construct a normal probability plot. This plot and a normal probability plot for the original data appear below. Clearly the log transformation gives quite a straight plot, so lognormality is plausible. The curvature in the plot for the original data implies a positively skewed population distribution — like the lognormal distribution.



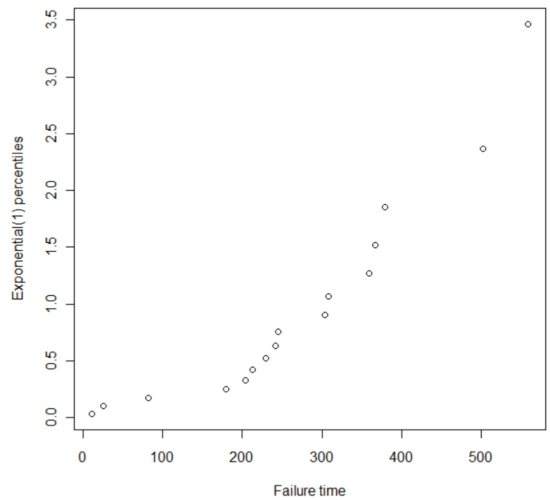


109. The pattern in the normal probability plot is reasonably linear, although with some deviation in the lower tail. By visual inspection alone, it is plausible that strength is normally distributed.



111. The $(100p)^{\text{th}}$ percentile η_p for the exponential distribution with $\lambda = 1$ is given by the formula $\eta_p = -\ln(1 - p)$. With $n = 16$, we need η_p for $p = \frac{0.5}{16}, \frac{1.5}{16}, \ldots, \frac{15.5}{16}$. These percentiles are .032, .398, .170, .247, .330, .421, .521, .633, .758, .901, 1.068, 1.269, 1.520, 1.856, 2.367, 3.466.

The accompanying plot of (failure time value, exponential(1) percentile) pairs exhibits substantial curvature, casting doubt on the assumption of an exponential population distribution.



Because λ is a scale parameter (as is σ for the normal family), λ = 1 can be used to assess the plausibility of the entire exponential family. If we used a different value of λ to find the percentiles, the slope of the graph would change, but not its linearity (or lack thereof).

Section 3.7

113. $y = 1/x \Rightarrow x = 1/y$ and $0 < x < 1 \Rightarrow 0 < 1/y < 1 \Rightarrow y > 1$. Apply the transformation theorem: $f_Y(y) = f_X(1/y)|dx/dy| = f_X(1/y)|-1/y^2| = 2(1/y)(1/y^2) = 2/y^3$ for $y > 1$. (If you're paying attention you might notice this is just the previous exercise in reverse!)

115. $y = \sqrt{x} \Rightarrow x = y^2$ and $x > 0 \Rightarrow y > 0$. Apply the transformation theorem:

$f_Y(y) = f_X(y^2)|dx/dy| = \frac{1}{2}e^{-y^2/2}|2y| = ye^{-y^2/2}$ for $y > 0$.

117. $y = \text{area} = x^2 \Rightarrow x = \sqrt{y}$ and $0 < x < 4 \Rightarrow 0 < y < 16$. Apply the transformation theorem:

$f_Y(y) = f_X(\sqrt{y})|dx/dy| = \sqrt{y}/8|1/(2\sqrt{y})| = 1/16$ for $0 < y < 16$. That is, the area Y is uniform on (0,16).

119. $y = \tan(\pi(x-.5)) \Rightarrow x = [\arctan(y)+.5]/\pi$ and $0 < x < 1 \Rightarrow -\pi/2 < \pi(x-.5) < \pi/2 \Rightarrow -\infty < y < \infty$ (since $\tan\theta \to \pm\infty$ as $\theta \to \pm\pi/2$). Also, $X \sim \text{Unif}[0, 1] \Rightarrow f_X(x) = 1$. Apply the transformation theorem:

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = 1\cdot\left|\frac{d}{dy}\left[\frac{\arctan(y)+.5}{\pi}\right]\right| = \left|\frac{1}{\pi}\frac{1}{1+y^2}\right| = \frac{1}{\pi(1+y^2)} \text{ for } -\infty < y < \infty.$$

121. Assume the target function $g(x)$ is differentiable and increasing, so that $h(y) = g^{-1}(y)$ is also differentiable and increasing. Apply the transformation theorem:

$$f_Y(y) = f_X(h(y)) \cdot |h'(y)|$$

$$1 = \frac{h(y)}{8} \cdot h'(y)$$

$$8 = h(y)h'(y)$$

Take the antiderivative of both sides to obtain $8y = (1/2)[h(y)]^2$, from which $h(y) = 4\sqrt{y}$. Now reverse the roles of x and y to find the inverse of h, aka g: $x = 4\sqrt{g(x)} \Rightarrow g(x) = x^2/16$.
As a check, apply the transformation theorem with $Y = X^2/16$ and $f_X(x) = x/8$ for $0 \le x \le 4$ and you indeed obtain $Y \sim$ Unif[0, 1].

You might notice that $x^2/16$ is the antiderivative of $x/8$; i.e., $g(x) = F_X(x)$. This is a special case of a more general result: if X is a continuous rv with cdf $F_X(x)$, then $F_X(X) \sim$ Unif[0, 1].

123. The transformation $y = x^2$ is not monotone on [–1,1], so we must proceed via the cdf method.
For $0 \le y \le 1$, $F_Y(y) = P(Y \le y) = P(X^2 \le y) = P\left(-\sqrt{y} \le X \le \sqrt{y}\right) = \dfrac{\sqrt{y}-(-\sqrt{y})}{1-(-1)} = \sqrt{y}$. [We've used the uniform cdf here.]
Thus, $f_Y(y) = \dfrac{d}{dy}\left(\sqrt{y}\right) = \dfrac{1}{2\sqrt{y}}$ for $0 < y \le 1$.

125. The transformation $y = x^2$ is not monotone and is not symmetric on [–1, 3], so we have to consider two cases. Clearly $0 \le Y \le 9$.
For $0 \le y \le 1$, $F_Y(y) = P(Y \le y) = P(X^2 \le y) = P\left(-\sqrt{y} \le X \le \sqrt{y}\right) = \dfrac{\sqrt{y}-(-\sqrt{y})}{3-(-1)} = \dfrac{\sqrt{y}}{2}$.
For $1 < y \le 9$, $F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(-1 \le X \le 1 \text{ or } 1 < X \le \sqrt{y}) = \dfrac{1}{2} + \dfrac{\sqrt{y}-1}{3-(-1)} = \dfrac{1+\sqrt{y}}{4}$.
Differentiating the two pieces, we have

$$f_Y(y) = \begin{cases} 1/(4\sqrt{y}) & 0 < y \le 1 \\ 1/(8\sqrt{y}) & 1 < y \le 9 \\ 0 & \text{otherwise} \end{cases}$$

127.

a. By assumption, the probability that you hit the disc centered at the bulls-eye with area x is proportional to x; in particular, this probability is x/[total area of target] $= x/[\pi(1)^2] = x/\pi$.
Therefore, $F_X(x) = P(X \le x) = P$(you hit disc centered at the bulls-eye with area x) $= x/\pi$. From this, $f_X(x) = d/dx[x/\pi] = 1/\pi$ for $0 < x < \pi$. That is, X is uniform on $(0, \pi)$.

b. $x = \pi y^2$ and $0 < x < \pi \Rightarrow 0 < y < 1$. Thus, $f_Y(y) = f_X(\pi y^2)|dx/dy| = 1/\pi\ |2\pi y| = 2y$ for $0 < y < 1$.

Section 3.8

129.

a. $F(x) = x^2/4$. Set $u = F(x)$ and solve for x: $u = x^2/4 \Rightarrow x = 2\sqrt{u}$.

b. The one-line "programs" below have been vectorized for speed; i.e., all 10,000 Unif[0, 1] values are generated simultaneously.

In Matlab: `x=2*sqrt(rand(10000,1))`

In R: `x<-2*sqrt(runif(10000))`

c. One execution of the R program gave `mean(x)` = 1.331268 and `sd(x)` = 0.4710592. These are very close to the exact mean and sd of X, which we can obtain through simple polynomial integrals:

$$\mu = \int_0^2 x \cdot \frac{x}{2} dx = \frac{1}{2}\int_0^2 x^2\, dx = \frac{4}{3} = 1.333;\ E(X^2) = \int_0^2 x^2 \cdot \frac{x}{2} dx = 2 \Rightarrow \sigma = \sqrt{2 - \left(\frac{4}{3}\right)^2} = \frac{\sqrt{2}}{3} = 0.4714.$$

131. $f(x) = \frac{1}{8} + \frac{3}{8}x \Rightarrow F(x) = \frac{1}{8}x + \frac{3}{16}x^2$. Set $u = F(x)$ and solve for x: $u = \frac{1}{8}x + \frac{3}{16}x^2 \Rightarrow 3x^2 + 2x - 16u = 0 \Rightarrow$

$x = \frac{-2 \pm \sqrt{4 - 4(3)(-16u)}}{2(3)} = \frac{-2 \pm 2\sqrt{1+48u}}{2(3)} = \frac{\sqrt{1+48u} - 1}{3}$. (The other root of the quadratic would place x between –2 and 0 rather than 0 and 2.) The code below implements this transformation and returns 10,000 values from the desired pdf.

In Matlab: `x=(sqrt(1+48*rand(10000,1))-1)/3`

In R: `x<-(sqrt(1+48*runif(10000))-1)/3`

133.

a. From Exercise 16**b**, the cdf of this model is $F(x) = 1 - \left(1 - \frac{x}{\tau}\right)^\theta$ for $0 < x < \tau$. Set $u = F(x)$ and solve for x: $u = 1 - \left(1 - \frac{x}{\tau}\right)^\theta \Rightarrow x = \tau \cdot \left[1 - (1-u)^{1/\theta}\right]$. This transform is implemented in the functions below.

```
function x=waittime(n,theta,tau)
u=rand(n,1);
x=tau*(1-(1-u).^(1/theta));
```

```
waittime<-function(n,theta,tau){
u<-runif(n)
x<- tau*(1-(1-u)^(1/theta))
}
```

b. Calling `x<-waittime(10000,4,80)` and `mean(x)` in R returned 15.9188, quite close to 16.

135.

a. $f(x)/g(x) \le c \Rightarrow \frac{3}{2}(1-x^2) \le c$, and this must hold for all x in [0, 1]. The left-hand side is maximized when $x = 0$, so we must have $\frac{3}{2}(1-0^2) \le c$. In other words, the smallest such c is $c = \frac{3}{2} = 1.5$.

b. The key step in the accept-reject scheme is assessing the inequality $u \cdot c \cdot g(y) \le f(y)$. Substitute $c = \frac{3}{2}$, $g(y) = 1$, and $f(y) = \frac{3}{2}(1-y^2)$; the inequality simplifies to $u \le (1 - y^2)$. This version of the key inequality is implemented in the programs below.

```
x=zeros(10000,1);
i=0;
while i<10000
    y=rand;
    u=rand;
    if u<=(1-y^2)
        i=i+1;
        x(i)=y;
    end
end
```

```
x <- NULL
i <- 0
while (i <10000){
      y <- runif(1)
      u <- runif(1)
      if (u<=(1-y^2)){
            i <- i+1
            x[i] <- y
      }
}
```

c. The expected number of candidate values required to generate a single accepted value is c, so the expected number required to generate 10,000 x-values is 10,000c = 10,000(1.5) = 15,000.

d. Executing the preceding R code returned `mean(x)` = 0.3734015. The true mean is $\mu = \int_0^1 x \cdot \frac{3}{2}(1-x^2)dx = \frac{3}{8}$ = 0.375. Our estimate of the mean is fairly close.

e. Following the instructions in the exercise, one obtains a vector `M` containing 10,000 simulated values of the "quarterly minimum" rv, M. For one simulation run in Matlab, 8760 of these 10,000 values were less than .1, so $\hat{P}(M < .1) = \frac{8760}{10{,}000}$ = .8760. It's very likely that, sometime during a quarter, weekly sales will be less than .1 ton (200 lbs).

137.

a. The cdf of the exponential pdf $g(x) = e^{-x}$ is $G(x) = 1 - e^{-x}$, from which the inverse cdf is $x = -\ln(1 - u)$. (This matches the exponential method given in the section. Of course, we could also use our software's built-in exponential simulator.)

b. The ratio $f(x)/g(x)$ is equivalent to $\sqrt{\frac{2}{\pi}}\, e^{-x^2/2} \cdot e^{+x} = \sqrt{\frac{2}{\pi}}\, e^{-x^2/2+x}$. This expression is maximized when the exponent is maximized. $\frac{d}{dx}[-x^2/2+x] = 0 \Rightarrow -x + 1 = 0 \Rightarrow x = 1$, so the absolute maximum on the interval $[0, \infty)$ occurs at $x = 0$ or $x = 1$ or as $x \to \infty$. Plugging in reveals that the maximum occurs at $x =$

1, from which we infer that $c \geq \sqrt{\frac{2}{\pi}}\, e^{-1^2/2+1} = \sqrt{\frac{2}{\pi}}\, e^{1/2} = \sqrt{\frac{2e}{\pi}} \approx 1.3155$; this is our majorization constant.

c. The expected number of candidate values required to generate a single accepted value is c, so the expected number required to generate 10,000 x-values is $10{,}000c \approx 13{,}155$.

d. The key inequality $u \cdot c \cdot g(y) \leq f(y)$ can be simplified: $u \cdot \sqrt{\frac{2e}{\pi}} \cdot e^{-y} \leq \sqrt{\frac{2}{\pi}} e^{-y^2/2} \Rightarrow$ $u \leq e^{-y^2/2+y-1/2} = e^{-(y-1)^2/2}$. This is implemented in the programs below.

```
x=zeros(10000,1);
i=0;
while i<10000
    y=-1*log(1-rand);
    u=rand;
    if u<=exp(-(y-1)^2/2)
        i=i+1;
        x(i)=y;
    end
end
```

```
x <- NULL
i <- 0
while (i <10000){
      y <- -1*log(1-runif(1))
      u <- runif(1)
      if(u<=exp(-(y-1)^2/2)){
            i <- i+1
            x[i] <- y
      }
}
```

139. By definition, the majorization constant c satisfies $f(x)/g(x) \leq c$ for all x on which f and g are positive. Suppose $c < 1$. Then $f(x) \leq c \cdot g(x) < 1 \cdot g(x)$; i.e., $f(x) < g(x)$ for all x. But both f and g are pdfs and so, in particular, both must integrate to 1. Integrating both sides of $f(x) < g(x)$ yields $1 < 1$, a contradiction.

Supplementary Exercises

141. The pdf of X is $f(x) = \frac{1}{25}$ for $0 \leq x \leq 25$ and is $= 0$ otherwise.

a. $P(10 \leq X \leq 20) = \frac{10}{25} = .4$.

b. $P(X \geq 10) = P(10 \leq X \leq 25) = \frac{15}{25} = .6$.

c. For $0 \leq x \leq 25$, $F(x) = \int_0^x \frac{1}{25} dy = \frac{x}{25}$. $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 25$.

d. $E(X) = \frac{A+B}{2} = \frac{0+25}{2} = 12.5$; $\text{Var}(X) = \frac{(B-A)^2}{12} = \frac{625}{12} = 52.083$, so $\text{SD}(X) = 7.22$.

143.

a. Clearly $f(x) \geq 0$. Now check that the function integrates to 1:

$$\int_0^\infty \frac{32}{(x+4)^3}dx = \int_0^\infty 32(x+4)^{-3}\,dx = -\frac{16}{(x+4)^2}\Bigg|_0^\infty = 0 - -\frac{16}{(0+4)^2} = 1.$$

b. For $x \leq 0$, $F(x) = 0$. For $x > 0$,

$$F(x) = \int_{-\infty}^x f(y)dy = \int_0^x \frac{32}{(y+4)^3}dy = -\frac{1}{2}\cdot\frac{32}{(y+4)^2}\Bigg]_0^x = 1 - \frac{16}{(x+4)^2}.$$

c. $P(2 \leq X \leq 5) = F(5) - F(2) = 1 - \frac{16}{81} - \left(1 - \frac{16}{36}\right) = .247$.

d. $E(X) = \int_{-\infty}^\infty x\cdot f(x)dx = \int_{-\infty}^\infty x\cdot\frac{32}{(x+4)^3}dx = \int_0^\infty (x+4-4)\cdot\frac{32}{(x+4)^3}dx$

$= \int_0^\infty \frac{32}{(x+4)^2}dx - 4\int_0^\infty \frac{32}{(x+4)^3}dx = 8 - 4 = 4$ years.

e. $E\left(\frac{100}{X+4}\right) = \int_0^\infty \frac{100}{x+4}\cdot\frac{32}{(x+4)^3}dx = 3200\int_0^\infty \frac{1}{(x+4)^4}dx = \frac{3200}{(3)(64)} = 16.67$.

145.

a. $P(39 < X < 42) = \Phi\left(\frac{42-40}{1.5}\right) - \Phi\left(\frac{39-40}{1.5}\right) = \Phi(1.33) - \Phi(-.67) = .9082 - .2514 = .6568$.

b. We desire the 85th percentile: $\Phi(z) = .85 \Rightarrow z = 1.04$ from the standard normal table, so the 85th percentile of this distribution is $40 + (1.04)(1.5) = 41.56$ V.

c. For a single diode, $P(X > 42) = 1 - P(X \leq 42) = 1 - \Phi\left(\frac{42-40}{1.5}\right) = 1 - \Phi(1.33) = .0918$.

Now let D represent the number of diodes (out of four) with voltage exceeding 42. The random variable D is binomial with $n = 4$ and $p = .0918$, so

$P(D \geq 1) = 1 - P(D = 0) = 1 - \binom{4}{0}(.0918)^0(.9082)^4 = 1 - .6803 = .3197$.

147.

a. Let X = the number of defectives in the batch of 250, so $X \sim \text{Bin}(250, .05)$. We can approximate X by a normal distribution, since $np = 12.5 \geq 10$ and $nq = 237.5 \geq 10$. The mean and sd of X are $\mu = np = 12.5$ and $\sigma = 3.446$. Using a continuity correction and realizing 10% of 250 is 25,

$P(X \geq 25) = 1 - P(X < 25) = 1 - P(X \leq 24.5) \approx 1 - \Phi\left(\frac{24.5-12.5}{3.446}\right) = 1 - \Phi(3.48) =$

$1 - .9997 = .0003$. (The exact binomial probability, from software, is .00086.)

b. Using the same normal approximation with a continuity correction, $P(X = 10) =$

$P(9.5 \leq X \leq 10.5) \approx \Phi\left(\frac{10.5-12.5}{3.446}\right) - \Phi\left(\frac{9.5-12.5}{3.446}\right) = \Phi(-.58) - \Phi(-.87) = .2810 - .1922 = .0888$.

(The exact binomial probability is $\binom{250}{10}(.05)^{10}(.95)^{240} = .0963$.)

149.

a. $E(X) = e^{3.5+1.2^2/2} = 68.03$ dB; Var(X) = $e^{2(3.5)+1.2^2}(e^{1.2^2} - 1) = 14907 \Rightarrow$ SD(X) = 122.09 dB.

b. $P(50 < X < 250) = \Phi\left(\frac{\ln(250)-3.5}{1.2}\right) - \Phi\left(\frac{\ln(50)-3.5}{1.2}\right) = \Phi(1.68) - \Phi(.34) = .9535 - .6331 = .3204$.

c. $P(X < 68.03) = \Phi\left(\frac{\ln(68.03)-3.5}{1.2}\right) = \Phi(.72) = .7642$, not .5 because the lognormal distribution is not symmetric.

151.

a. $F(x) = 0$ for $x < 1$ and $F(x) = 1$ for $x > 3$. For $1 \le x \le 3$, $F(x) = \int_1^x \frac{3}{2y^2}dy = \frac{3}{2}\left(1 - \frac{1}{x}\right)$.

b. $P(X \le 2.5) = F(2.5) = 1.5(1 - .4) = .9$; $P(1.5 \le X \le 2.5) = F(2.5) - F(1.5) = .4$.

c. $E(X) = \int_1^3 x \cdot \frac{3}{2} \cdot \frac{1}{x^2} dx = \frac{3}{2}\int_1^3 \frac{1}{x} dx = 1.5\ln(x)\Big]_1^3 = 1.648$ seconds.

d. $E(X^2) = = \int_1^3 x^2 \cdot \frac{3}{2} \cdot \frac{1}{x^2} dx = \frac{3}{2}\int_1^3 dx = 3$, so Var($X$) = $E(X^2) - [E(X)]^2 = .284$ and $\sigma = .553$ seconds.

e. From the description, $h(x) = 0$ if $1 \le x \le 1.5$; $h(x) = x - 1.5$ if $1.5 \le x \le 2.5$ (one second later), and $h(x) = 1$ if $2.5 \le x \le 3$. Using those terms,

$$E[h(X)] = \int_1^3 h(x)\,dx = \int_{1.5}^{2.5}(x-1.5)\cdot\frac{3}{2x^2}dx + \int_{2.5}^3 1\cdot\frac{3}{2x^2}dx = .267 \text{ seconds}.$$

153.

a. Since X is exponential, $E(X) = \frac{1}{\lambda} = 1.075$ and $\sigma = \frac{1}{\lambda} = 1.075$.

b. $P(X > 3.0) = 1 - P(X \le 3.0) = 1 - F(3.0) = 1 - [1 - e^{-.93(3.0)}] = .0614$.
$P(1.0 \le X \le 3.0) = F(3.0) - F(1.0) = [1 - e^{-.93(3.0)}] - [1 - e^{-.93(1.0)}] = .333$.

c. The 90th percentile is requested: $.9 = F(\eta_{.9}) = 1 - e^{-.93\eta_{.9}} \Rightarrow \eta_{.9} = \frac{\ln(.1)}{(-.93)} = 2.476$ mm.

155. We have a random variable $T \sim N(\mu, \sigma)$. Let $f(t)$ denote its pdf.

a. The "expected loss" is the expected value of a piecewise-defined function, so we should first write the function out in pieces (two integrals, as seen below). Call this expected loss $Q(a)$, to emphasize we're interested in its behavior as a function of a. We have:

$$Q(a) = E[L(a,T)] = \int_{-\infty}^{a} k(a-t)f(t)\,dt + \int_{a}^{\infty} (t-a)f(t)\,dt$$

$$= ka\int_{-\infty}^{a} f(t)\,dt - k\int_{-\infty}^{a} tf(t)\,dt + \int_{a}^{\infty} tf(t)\,dt - a\int_{a}^{\infty} f(t)\,dt = kaF(a) - k\int_{-\infty}^{a} tf(t)\,dt + \int_{a}^{\infty} tf(t)\,dt - a[1-F(a)]$$

where $F(a)$ denotes the cdf of T. To minimize this expression, take the first derivative with respect to a, using the product rule and the fundamental theorem of calculus where appropriate:

$$\begin{aligned} Q'(a) &= kaF(a) - k\int_{-\infty}^{a} tf(t)\,dt + \int_{a}^{\infty} tf(t)\,dt - a[1-F(a)] \\ &= kF(a) + kaF'(a) - kaf(a) + 0 - af(a) - 1 + F(a) + aF'(a) \\ &= kF(a) + kaf(a) - kaf(a) - af(a) - 1 + F(a) + af(a) \\ &= (k+1)F(a) - 1 \end{aligned}$$

Finally, set this equal to zero, and use the fact that, because T is a normal random variable, $F(a) = \Phi\left(\frac{a-\mu}{\sigma}\right)$:

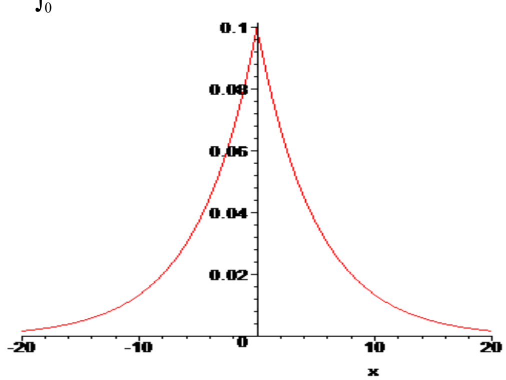
$$(k+1)F(a) - 1 = 0 \Rightarrow (k+1)\Phi\left(\frac{a-\mu}{\sigma}\right) - 1 = 0 \Rightarrow \Phi\left(\frac{a-\mu}{\sigma}\right) = \frac{1}{k+1} \Rightarrow a = \mu + \sigma \cdot \Phi^{-1}\left(\frac{1}{k+1}\right)$$

This is the critical value, a^*, as desired.

b. With the values provided, $a^* = 100{,}000 + 10{,}000\Phi^{-1}\left(\frac{1}{2+1}\right) = 100{,}000 + 10{,}000\Phi^{-1}(0.33) = 100{,}000 +$ $10{,}000(-0.44)$ from the standard normal table = 100,000 – 4,400 = \$95,600. The probability of an over-assessment equals $P(95{,}600 > T) = P(T < 96{,}500) = \Phi\left(\frac{95{,}600 - 100{,}000}{10{,}000}\right) = \Phi(-0.44) = .3300$, or 33%. Notice that, in general, the probability of an over-assessment using the optimal value of a is equal to $\frac{1}{k+1}$.

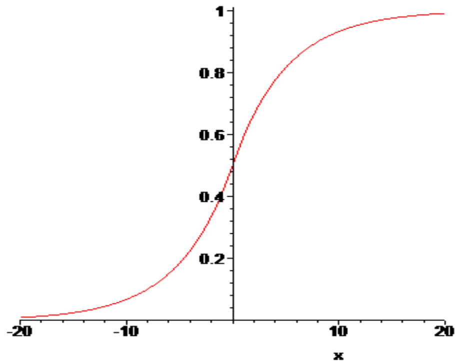
157.

a. $\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} .1e^{2x}dx + \int_{0}^{\infty} .1e^{-2x}dx = .5 + .5 = 1.$



b. For $x < 0$, $F(x) = \int_{-\infty}^{x} .1e^{.2y}dy = \frac{1}{2}e^{.2x}$.

For $x \geq 0$, $F(x) = \int_{-\infty}^{x} f(y)\,dy = \int_{-\infty}^{0} .1e^{-.2y}\,dy + \int_{0}^{x} .1e^{-.2y}dy = \frac{1}{2} + \int_{0}^{x} .1e^{-.2y}dy = 1 - \frac{1}{2}e^{-.2x}$.



c. $P(X < 0) = F(0) = .5$; $P(X < 2) = F(2) = 1 - .5e^{-.4} = .665$; $P(-1 \leq X \leq 2) = F(2) - F(-1) = .256$; and $P(|X| > 2) = 1 - (-2 \leq X \leq 2) = 1 - [F(2) - F(-2)] = .670$.

159.

a. Provided $\alpha > 1$, $1 = \int_{5}^{\infty} \frac{k}{x^{\alpha}}dx = k \cdot \frac{5^{1-\alpha}}{\alpha - 1} \Rightarrow k = (\alpha - 1)5^{\alpha-1}$.

b. For $x \geq 5$, $F(x) = \int_{5}^{x} \frac{(\alpha-1)5^{\alpha-1}}{y^{\alpha}}dy = -5^{\alpha-1}\left[x^{1-\alpha} - 5^{1-\alpha}\right] = 1 - \left(\frac{5}{x}\right)^{\alpha-1}$. For $x < 5$, $F(x) = 0$.

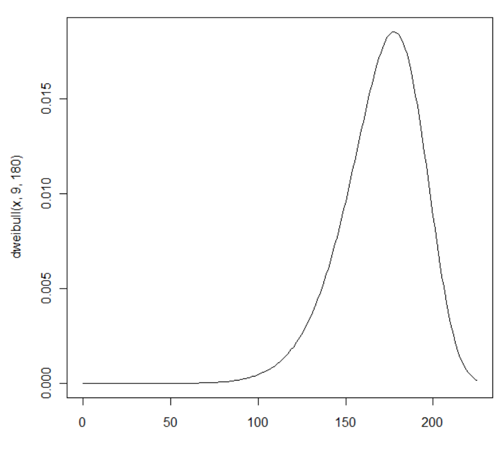
c. Provided $\alpha > 2$, $E(X) = \int_{5}^{\infty} x \cdot \frac{k}{x^{\alpha}}dx = \int_{5}^{\infty} \frac{(\alpha-1)5^{\alpha-1}}{x^{\alpha-1}}dx = 5\frac{\alpha-1}{\alpha-2}$.

d. Let $Y = \ln(X/5)$. Then $F_Y(y) = P\left(\ln\left(\frac{X}{5}\right) \leq y\right) = P\left(\frac{X}{5} \leq e^{y}\right) = P\left(X \leq 5e^{y}\right) = F(5e^{y}) = 1 - \left(\frac{5}{5e^{y}}\right)^{\alpha-1} =$

$1 - e^{-(\alpha-1)y}$, the cdf of an exponential rv with parameter $\alpha - 1$.

161.

a. The accompanying Weibull pdf plot was created in R.



b. $P(X > 175) = 1 - F(175; 9, 180) = e^{-(175/180)^9} = .4602.$
$P(150 \le X \le 175) = F(175; 9, 180) - F(150; 9, 180) = .5398 - .1762 = .3636.$

c. From **b**, the probability a specimen is <u>not</u> between 150 and 175 equals 1 – .3636 = .6364. So, P(at least one is between 150 and 175) = 1 – P(neither is between 150 and 175) = 1 – $(.6364)^2$ = .5950.

d. We want the 10th percentile: $.10 = F(x; 9, 180) = 1 - e^{-(x/180)^9}$. A small bit of algebra leads us to $x = 180(-\ln(1-.10))^{1/9} = 140.178$. Thus, 10% of all tensile strengths will be less than 140.178 MPa.

163.

a. If we let $\alpha = 2$ and $\beta = \sqrt{2}\sigma$, then we can manipulate $f(v)$ as follows:
$f(v) = \frac{v}{\sigma^2}e^{-v^2/2\sigma^2} = \frac{2}{2\sigma^2}ve^{-v^2/2\sigma^2} = \frac{2}{(\sqrt{2}\sigma)^2}v^{2-1}e^{-\left(v/\sqrt{2}\sigma\right)^2} = \frac{\alpha}{\beta^\alpha}v^{\alpha-1}e^{-(v/\beta)^\alpha}$, which is in the Weibull family of distributions.

b. Use the Weibull cdf: $P(V \le 25) = F(25; 2, \sqrt{2}\sigma) = 1 - e^{-\left(\frac{25}{\sqrt{2}\sigma}\right)^2} = 1 - e^{-\frac{625}{800}}$ = 1 – .458 = .542.

165. Let A = the cork is acceptable and B = the first machine was used. The goal is to find $P(B \mid A)$, which can be obtained from Bayes' rule:

$$P(B \mid A) = \frac{P(B)P(A \mid B)}{P(B)P(A \mid B) + P(B')P(A \mid B')} = \frac{.6P(A \mid B)}{.6P(A \mid B) + .4P(A \mid B')}.$$

From Exercise 49, $P(A \mid B) = P$(machine 1 produces an acceptable cork) = .6826 and $P(A \mid B') = P$(machine 2 produces an acceptable cork) = .9987. Therefore,

$$P(B \mid A) = \frac{.6(.6826)}{.6(.6826) + .4(.9987)} = .5062.$$

167. For $y > 0$, $F(y) = P(Y \le y) = P\left(\frac{2X^2}{\beta^2} \le y\right) = P\left(X^2 \le \frac{\beta^2 y}{2}\right) = P\left(X \le \frac{\beta\sqrt{y}}{\sqrt{2}}\right)$. Now take the cdf of X (Weibull), replace x by $\frac{\beta\sqrt{y}}{\sqrt{2}}$, and then differentiate with respect to y to obtain the desired result $f(y)$.

169. Apply the transformation theorem, or work directly with cdfs (the latter is easier because the extreme value cdf is less unpleasant than its pdf).

$$F_Y(y) = P(Y \le y) = P(\ln(X) \le y) = P(X \le e^y) = F_X(e^y) = 1 - \exp\left[-\left(\frac{e^y}{\beta}\right)^\alpha\right] = 1 - \exp\left[-e^{\alpha y}/\beta^\alpha\right].$$

Now consider the extreme value cdf from Section 3.6, with $x = y$, $\theta_1 = \ln(\beta)$ and $\theta_2 = 1/\alpha$:
$F(y; \ln(\beta), 1/\alpha) = 1 - \exp\left[-e^{(y-\ln(\beta))/(1/\alpha)}\right] = 1 - \exp\left[-e^{\alpha y - \alpha\ln(\beta)}\right]$. Using properties of e and ln,
$e^{\alpha y - \alpha\ln(\beta)} = e^{\alpha y}/e^{\alpha\ln(\beta)} = e^{\alpha y}/\left[e^{\ln(\beta)}\right]^\alpha = e^{\alpha y}/\beta^\alpha$. Hence, the extreme value distribution with the specified parameters has cdf $F(y; \ln(\beta), 1/\alpha) = 1 - \exp\left[-e^{\alpha y}/\beta^\alpha\right]$, which matches the cdf of Y, proving the desired result.

171.

a. $E(X) = 150 + (850 - 150)\dfrac{8}{8+2} = 710$ and $\text{Var}(X) = \dfrac{(850-150)^2(8)(2)}{(8+2)^2(8+2+1)} = 7127.27 \Rightarrow SD(X) \approx 84.423$.

Using software, $P(|X - 710| \le 84.423) = P(625.577 \le X \le 794.423) =$

$\displaystyle\int_{625.577}^{794.423} \frac{1}{700}\frac{\Gamma(10)}{\Gamma(8)\Gamma(2)}\left(\frac{x-150}{700}\right)^7\left(\frac{850-x}{700}\right)^1 dx = .684$.

b. $P(X > 750) = \displaystyle\int_{750}^{850} \frac{1}{700}\frac{\Gamma(10)}{\Gamma(8)\Gamma(2)}\left(\frac{x-150}{700}\right)^7\left(\frac{850-x}{700}\right)^1 dx = .376$. Again, the computation of the requested integral requires a calculator or computer.

CHAPTER 4

Section 4.1

1.

a. $P(X = 1, Y = 1) = p(1,1) = .20$.

b. $P(X \le 1 \text{ and } Y \le 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .42$.

c. At least one hose is in use at both islands. $P(X \ne 0 \text{ and } Y \ne 0) = p(1,1) + p(1,2) + p(2,1) + p(2,2) = .70$.

d. By summing row probabilities, $p_X(x) = .16, .34, .50$ for $x = 0, 1, 2$, By summing column probabilities, $p_Y(y) = .24, .38, .38$ for $y = 0, 1, 2$. $P(X \le 1) = p_X(0) + p_X(1) = .50$.

e. $p(0,0) = .10$, but $p_X(0) \cdot p_Y(0) = (.16)(.24) = .0384 \ne .10$, so X and Y are not independent.

3.

a. $p(1,1) = .15$, the entry in the 1st row and 1st column of the joint probability table.

b. $P(X_1 = X_2) = p(0,0) + p(1,1) + p(2,2) + p(3,3) = .08 + .15 + .10 + .07 = .40$.

c. $A = \{X_1 \ge 2 + X_2 \cup X_2 \ge 2 + X_1\}$, so $P(A) = p(2,0) + p(3,0) + p(4,0) + p(3,1) + p(4,1) + p(4,2) + p(0,2) + p(0,3) + p(1,3) = .22$.

d. $P(X_1 + X_2 = 4) = p(1,3) + p(2,2) + p(3,1) + p(4,0) = .17$.
$P(X_1 + X_2 \ge 4) = P(X_1 + X_2 = 4) + p(4,1) + p(4,2) + p(4,3) + p(3,2) + p(3,3) + p(2,3) = .46$.

e. $p_1(0) = P(X_1 = 0) = p(0,0) + p(0,1) + p(0,2) + p(0,3) = .19$
$p_1(1) = P(X_1 = 1) = p(1,0) + p(1,1) + p(1,2) + p(1,3) = .30$, etc.

x_1	0	1	2	3	4
$p_1(x_1)$	.19	.30	.25	.14	.12

f. $p_2(0) = P(X_2 = 0) = p(0,0) + p(1,0) + p(2,0) + p(3,0) + p(4,0) = .19$, etc.

x_2	0	1	2	3
$p_2(x_2)$	.19	.30	.28	.23

g. $p(4,0) = 0$, yet $p_1(4) = .12 > 0$ and $p_2(0) = .19 > 0$, so $p(x_1, x_2) \ne p_1(x_1) \cdot p_2(x_2)$ for every (x_1, x_2), and the two variables are <u>not</u> independent.

5. Let X_1 = the number of freshmen in the sample of 10; define X_2, X_3, X_4, analogously for sophomores, juniors, and seniors, respectively. Then the joint distribution of (X_1, X_2, X_3, X_4) is multinomial with n = 10 and (p_1, p_2, p_3, p_4) = (.20, .18, .21, .41).

a. $P((X_1, X_2, X_3, X_4) = (2, 2, 2, 4)) = \dfrac{10!}{2!2!2!4!}(.20)^2(.18)^2(.21)^2(.41)^4 = .0305$.

b. Let $Y = X_1 + X_2$ = the number of underclassmen in the sample. Then Y meets the conditions of a binomial rv, with n = 10 and p = .20 + .18 = .38. Hence, the probability the sample is evenly split among under- and upper-classmen is $P(Y = 5) = \binom{10}{5}(.38)^5(.62)^5 = .1829$.

c. The marginal distribution of X_1 is Bin(10, .20), so $P(X_1 = 0) = (.80)^{10} = .1073$.
If selections were truly random from the population of all students, there's about a 10.7% chance that no freshmen would be selected. If we consider this a low probability, then we have evidence that something is amiss; otherwise, we might ascribe this occurrence to random chance alone ("bad luck").

7.

a. $p(3, 3) = P(X = 3, Y = 3) = P$(3 customers, each with 1 package)
= P(each has 1 package | 3 customers) · P(3 customers) = $(.6)^3 \cdot (.25) = .054$.

b. $p(4, 11) = P(X = 4, Y = 11) = P$(total of 11 packages | 4 customers) · P(4 customers).
Given that there are 4 customers, there are four different ways to have a total of 11 packages: 3, 3, 3, 2 or 3, 3, 2, 3 or 3, 2, 3, 3 or 2, 3, 3, 3. Each way has probability $(.1)^3(.3)$, so $p(4, 11) = 4(.1)^3(.3)(.15) = .00018$.

9.

a. $p(1,1) = .030$.

b. $P(X \le 1 \text{ and } Y \le 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .120$.

c. $P(X = 1) = p(1,0) + p(1,1) + p(1,2) = .100$; $P(Y = 1) = p(0,1) + \ldots + p(5,1) = .300$.

d. P(overflow) = $P(X + 3Y > 5) = 1 - P(X + 3Y \le 5) = 1 - P((X,Y)$=(0,0) or …or (5,0) or (0,1) or (1,1) or (2,1)) = 1 – .620 = .380.

e. The marginal probabilities for X (row sums from the joint probability table) are $p_X(0) = .05$, $p_X(1) = .10$, $p_X(2) = .25$, $p_X(3) = .30$, $p_X(4) = .20$, $p_X(5) = .10$; those for Y (column sums) are $p_Y(0) = .5$, $p_Y(1) = .3$, $p_Y(2) = .2$. It is now easily verified that for every (x,y), $p(x,y) = p_X(x) \cdot p_Y(y)$, so X and Y are independent.

11.

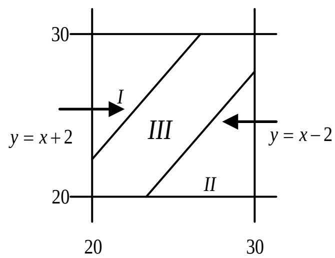
a. $1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy=\int_{20}^{30}\int_{20}^{30} k(x^2+y^2)dxdy = k\int_{20}^{30}\int_{20}^{30} x^2dydx+k\int_{20}^{30}\int_{20}^{30} y^2dxdy$

$=10k\int_{20}^{30} x^2dx+10k\int_{20}^{30} y^2dy = 20k\cdot\left(\frac{19,000}{3}\right)\Rightarrow k=\frac{3}{380,000}$.

b. $P(X < 26 \text{ and } Y < 26) = \int_{20}^{26}\int_{20}^{26} k(x^2+y^2)dxdy = k\int_{20}^{26}\left[x^2y+\frac{y^3}{3}\right]_{20}^{26} dx = k\int_{20}^{26}(6x^2+3192)\,dx =$

$k\cdot(38,304) = .3024$.

c. The region of integration is labeled *III* below.



$P(|X-Y|\le 2) = \iint_{III} f(x,y)dxdy = 1-\iint_{I} f(x,y)dxdy-\iint_{II} f(x,y)dxdy =$

$1-\int_{20}^{28}\int_{x+2}^{30} f(x,y)dydx-\int_{22}^{30}\int_{20}^{x-2} f(x,y)dydx$ = .3593 (after much algebra).

d. $f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy=\int_{20}^{30} k(x^2+y^2)dy = kx^2y+k\left.\frac{y^3}{3}\right|_{20}^{30}$ = $10kx^2 + .05$, for $20 \le x \le 30$.

e. $f_Y(y)$ can be obtained by substituting y for x in (d); clearly $f(x,y) \ne f_X(x)\cdot f_Y(y)$, so X and Y are not independent.

13.

a. Since X and Y are independent, $p(x,y) = p_X(x)\cdot p_Y(y)=\frac{e^{-\mu_1}\mu_1^x}{x!}\cdot\frac{e^{-\mu_2}\mu_2^y}{y!}=\frac{e^{-\mu_1-\mu_2}\mu_1^x\mu_2^y}{x!y!}$

for $x = 0, 1, 2, \ldots; y = 0, 1, 2, \ldots$.

b. $P(X + Y \le 1) = p(0,0) + p(0,1) + p(1,0) = \ldots = e^{-\mu_1-\mu_2}\left[1+\mu_1+\mu_2\right]$.

c. $P(X + Y= m) = \sum_{k=0}^{m} P(X=k, Y=m-k) = e^{-\mu_1-\mu_2}\sum_{k=0}^{m}\frac{\mu_1^k}{k!}\frac{\mu_2^{m-k}}{(m-k)!} = \frac{e^{-\mu_1-\mu_2}}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}\mu_1^k\mu_2^{m-k} =$

$\frac{e^{-\mu_1-\mu_2}}{m!}\sum_{k=0}^{m}\binom{m}{k}\mu_1^k\mu_2^{m-k} = \frac{e^{-\mu_1-\mu_2}}{m!}(\mu_1+\mu_2)^m$ by the binomial theorem. We recognize this as the pmf of a Poisson random variable with parameter $\mu_1+\mu_2$. Therefore, the total number of errors, $X + Y$, also has a Poisson distribution, with parameter $\mu_1+\mu_2$.

15.

a. $f(x,y) = f_X(x) \cdot f_Y(y) = \begin{cases} e^{-x-y} & x \ge 0, y \ge 0 \\ 0 & \text{otherwise} \end{cases}$

b. By independence, $P(X \le 1 \text{ and } Y \le 1) = P(X \le 1) \cdot P(Y \le 1) = (1 - e^{-1})\,(1 - e^{-1}) = .400$.

c. $P(X + Y \le 2) = \int_0^2 \int_0^{2-x} e^{-x-y}\,dy\,dx = \int_0^2 e^{-x}\left[1 - e^{-(2-x)}\right]dx = \int_0^2 (e^{-x} - e^{-2})\,dx = 1 - e^{-2} - 2e^{-2} = .594$.

d. $P(X + Y \le 1) = \int_0^1 e^{-x}\left[1 - e^{-(1-x)}\right]dx = 1 - 2e^{-1} = .264$,
so $P(1 \le X + Y \le 2) = P(X + Y \le 2) - P(X + Y \le 1) = .594 - .264 = .330$.

17.

a. Each X_i has cdf $F(x) = P(X_i \le x) = 1 - e^{-\lambda x}$. Using this, the cdf of Y is
$F(y) = P(Y \le y) = P(X_1 \le y \cup [X_2 \le y \cap X_3 \le y])$
$= P(X_1 \le y) + P(X_2 \le y \cap X_3 \le y) - P(X_1 \le y \cap [X_2 \le y \cap X_3 \le y])$
$= (1 - e^{-\lambda y}) + (1 - e^{-\lambda y})^2 - (1 - e^{-\lambda y})^3$ for $y > 0$.

The pdf of Y is $f(y) = F'(y) = \lambda e^{-\lambda y} + 2(1 - e^{-\lambda y})\left(\lambda e^{-\lambda y}\right) - 3(1 - e^{-\lambda y})^2\left(\lambda e^{-\lambda y}\right) = 4\lambda e^{-2\lambda y} - 3\lambda e^{-3\lambda y}$
for $y > 0$.

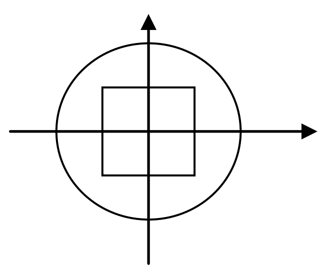
b. $E(Y) = \int_0^\infty y \cdot \left(4\lambda e^{-2\lambda y} - 3\lambda e^{-3\lambda y}\right)dy = 2\left(\frac{1}{2\lambda}\right) - \frac{1}{3\lambda} = \frac{2}{3\lambda}$.

19.

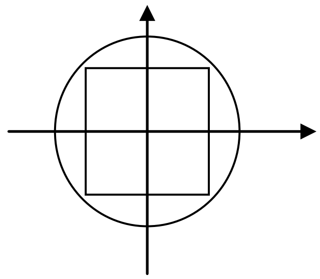
a. Let A denote the disk of radius $r/2$. Then $P((X,Y) \text{ lies in } A) = \iint_A f(x,y)\,dx\,dy$

$= \iint_A \frac{1}{\pi r^2}\,dx\,dy = \frac{1}{\pi r^2}\iint_A dx\,dy = \frac{\text{area of } A}{\pi r^2} = \frac{\pi(r/2)^2}{\pi r^2} = \frac{1}{4} = .25$. Notice that, since the joint pdf of X and Y is a constant (i.e., (X,Y) is uniform over the disk), it will be the case for any subset A that $P((X,Y)$ lies in $A) = \frac{\text{area of } A}{\pi r^2}$.

b. By the same ratio-of-areas idea, $P\left(-\frac{r}{2} \le X \le \frac{r}{2}, -\frac{r}{2} \le Y \le \frac{r}{2}\right) = \frac{r^2}{\pi r^2} = \frac{1}{\pi}$. This region is the square depicted in the graph below.



c. Similarly, $P\left(-\frac{r}{\sqrt{2}} \le X \le \frac{r}{\sqrt{2}}, -\frac{r}{\sqrt{2}} \le Y \le \frac{r}{\sqrt{2}}\right) = \frac{2r^2}{\pi r^2} = \frac{2}{\pi}$. This region is the slightly larger square depicted in the graph below, whose corners actually touch the circle.



d. $f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \frac{1}{\pi r^2} dy = \frac{2\sqrt{r^2 - x^2}}{\pi r^2}$ for $-r \le x \le r$.

Similarly, $f_Y(y) = \frac{2\sqrt{r^2 - y^2}}{\pi r^2}$ for $-r \le y \le r$. X and Y are not independent, since the joint pdf is not the product of the marginal pdfs: $\frac{1}{\pi r^2} \ne \frac{2\sqrt{r^2 - x^2}}{\pi r^2} \cdot \frac{2\sqrt{r^2 - y^2}}{\pi r^2}$.

21. Picture an inscribed equilateral triangle with one vertex at A, so the other two vertices are 120° away from A in either direction. Clearly chord AB will exceed the side length of this triangle if and only if point B is "between" the other two vertices (i.e., "opposite" A). Since that arc between the other two vertices spans 120° and the points were selected uniformly, the probability is clearly $\frac{120}{360} = \frac{1}{3}$.

Section 4.2

23.

a. $P(X > Y) = \sum\sum_{x>y} p(x, y) = p(1, 0) + p(2, 0) + p(3, 0) + p(2, 1) + p(3, 1) + p(3, 2) = .03 + .02 + .01 + .03 + .01 + .01 = .11$.

b. Adding down the columns gives the probabilities associated with the x-values:

x	0	1	2	3
$p_X(x)$	.78	.12	.07	.03

Similarly, adding across the rows gives the probabilities associated with the y-values:

y	0	1	2
$p_Y(y)$	.77	.14	.09

c. Test a coordinate, e.g. (0, 0): $p(0, 0) = .71$, while $p_X(0) \cdot p_Y(0) = (.78)(.77) = .6006 \ne .71$. Therefore, X and Y are not independent.

d. The average number of syntax errors is $E(X) = 0(.78) + 1(.12) + 2(.07) + 3(.03) = 0.35$, while the average number of logic errors is $E(Y) = 0(.77) + 1(.14) + 2(.09) = 0.32$.

e. By linearity of expectation, $E(100 - 4X - 9Y) = 100 - 4E(X) - 9E(Y) = 100 - 4(.35) - 9(.32) = 95.72$.

25. $E(X_1 - X_2) = \sum_{x_1=0}^{4}\sum_{x_2=0}^{3}(x_1 - x_2)\cdot p(x_1, x_2)$ = (0 – 0)(.08) + (0 – 1)(.07) + … + (4 – 3)(.06) = .15.

Or, by linearity of expectation, $E(X_1 - X_2) = E(X_1) - E(X_2)$, so in this case we could also work out the means of X_1 and X_2 from their marginal distributions: $E(X_1) = 1.70$ and $E(X_2) = 1.55$, so $E(X_1 - X_2) = E(X_1) - E(X_2)$ = 1.70 – 1.55 = .15.

27. The expected value of X, being uniform on $[L - A, L + A]$, is simply the midpoint of the interval, L. Since Y has the same distribution, $E(Y) = L$ as well. Finally, since X and Y are independent, $E(\text{area}) = E(XY) = E(X)\cdot E(Y) = L \cdot L = L^2$.

29. The amount of time Annie waits for Alvie, if Annie arrives first, is $Y - X$; similarly, the time Alvie waits for Annie is $X - Y$. Either way, the amount of time the first person waits for the second person is $h(X, Y) = |X - Y|$. Since X and Y are independent, their joint pdf is given by $f_X(x)\cdot f_Y(y) = (3x^2)(2y) = 6x^2y$. From these, the expected waiting time is

$$E[h(X,Y)] = \int_0^1\int_0^1 |x-y|\cdot f(x,y)\,dxdy = \int_0^1\int_0^1 |x-y|\cdot 6x^2y\,dxdy$$

$$= \int_0^1\int_0^x (x-y)\cdot 6x^2y\,dydx + \int_0^1\int_x^1 (x-y)\cdot 6x^2y\,dydx = \frac{1}{6}+\frac{1}{12}=\frac{1}{4} \text{ hour, or 15 minutes.}$$

31. $\text{Cov}(X,Y) = -\frac{2}{75}$ and $\mu_X = \mu_Y = \frac{2}{5}$.

$E(X^2) = \int_0^1 x^2\cdot f_X(x)dx = 12\int_0^1 x^3(1-x^2dx) = \frac{12}{60}=\frac{1}{5}$, so $\text{Var}(X) = \frac{1}{5}-\left(\frac{2}{5}\right)^2 = \frac{1}{25}$.

Similarly, $\text{Var}(Y) = \frac{1}{25}$, so $\rho_{X,Y} = \frac{-\frac{2}{75}}{\sqrt{\frac{1}{25}}\cdot\sqrt{\frac{1}{25}}} = -\frac{50}{75} = -\frac{2}{3}$.

33.

a. $E(X) = \int_{20}^{30} x f_X(x)dx = \int_{20}^{30} x\left[10kx^2 + .05\right]dx = \frac{1925}{76} = 25.329$; by symmetry, $E(Y) = 25.329$ also;

$E(XY) = \int_{20}^{30}\int_{20}^{30} xy\cdot k(x^2+y^2)dxdy = \frac{24375}{38} = 641.447 \Rightarrow$

$\text{Cov}(X, Y) = 641.447 - (25.329)^2 = -.1082$.

b. $E(X^2) = \int_{20}^{30} x^2\left[10kx^2+.05\right]dx = \frac{37040}{57} = 649.8246 \Rightarrow \text{Var}(X) = 649.8246 - (25.329)^2 = 8.2664$; by symmetry, $\text{Var}(Y) = 8.2664$ as well; thus, $\rho = \frac{-.1082}{\sqrt{(8.2664)(8.2664)}} = -.0131$.

35.

a. $E(XY)$ = (0)(0)(.71) + … + (3)(2)(.01) = .35 $\Rightarrow \text{Cov}(X, Y) = E(XY) - E(X)E(Y)$ = .35 – (.35)(.32) = .238.

b. Next, from the marginal distributions, $\text{Var}(X) = E(X^2) - [E(X)]^2 = .67 - (.35)^2 = .5475$ and, similarly, $\text{Var}(Y) = .3976$. Thus, $\text{Corr}(X, Y) = \frac{.238}{\sqrt{(.5475)(.3976)}} = .51$. There is a moderate, direct association between the number of syntax errors and the number of logic errors in a program. A direct association indicates that programs with a higher-than-average number of syntax errors also tend to have a higher-than-average number of logic errors, and vice versa.

37.

a. Let $H = h(X, Y)$. The variance shortcut formula states that $\text{Var}(H) = E(H^2) - [E(H)]^2$. Applying that shortcut formula here yields $\text{Var}(h(X,Y)) = E(h^2(X,Y)) - [E(h(X,Y))]^2$. More explicitly, if X and Y are discrete,

$\text{Var}(h(X,Y)) = \sum\sum [h(x,y)]^2 \cdot p(x,y) - \left[\sum\sum h(x,y)\cdot p(x,y)\right]^2$; if X and Y are continuous,

$\text{Var}(h(X,Y)) = \iint [h(x,y)]^2 \cdot f(x,y)dA - \left[\iint h(x,y)\cdot f(x,y)dA\right]^2$.

b. $E[h(X, Y)] = E[\max(X, Y)] = 9.60$, and $E[h^2(X, Y)] = E[(\max(X, Y))^2] = (0)^2(.02) + (5)^2(.06) + \ldots + (15)^2(.01) = 105.5$, so $\text{Var}(\max(X, Y)) = 105.5 - (9.60)^2 = 13.34$.

39. First, by linearity of expectation, $\mu_{aX+bY+c} = a\mu_X + b\mu_Y + c$. Hence, by definition,

$$\begin{aligned} Cov(aX+bY+c,Z) &= E[(aX+bY+c-\mu_{aX+bY+c})(Z-\mu_Z)] = E[(aX+bY+c-[a\mu_X+b\mu_Y+c])(Z-\mu_Z)] \\ &= E[(a(X-\mu_X)+b(Y-\mu_Y))(Z-\mu_Z)] \end{aligned}$$

Apply linearity of expectation a second time:

$$\begin{aligned} Cov(aX+bY+c,Z) &= E[a(X-\mu_X)(Z-\mu_Z)+b(Y-\mu_Y)(Z-\mu_Z)] \\ &= aE[(X-\mu_X)(Z-\mu_Z)]+bE[(Y-\mu_Y)(Z-\mu_Z)] \\ &= a\text{Cov}(X,Z)+b\text{Cov}(Y,Z) \end{aligned}$$

41. Use the previous exercise: $\text{Cov}(X, Y) = \text{Cov}(X, aX + b) = a\text{Cov}(X, X) = a\text{Var}(X) \Rightarrow$

so $\text{Corr}(X,Y) = \dfrac{a\text{Var}(X)}{\sigma_X \cdot \sigma_Y} = \dfrac{a\sigma_X^2}{\sigma_X \cdot |a|\sigma_X} = \dfrac{a}{|a|} = 1$ if $a > 0$, and -1 if $a < 0$.

Section 4.3

43.

a. $E(27X_1 + 125X_2 + 512X_3) = 27E(X_1) + 125E(X_2) + 512E(X_3)$
$= 27(200) + 125(250) + 512(100) = 87{,}850$.
$\text{Var}(27X_1 + 125X_2 + 512X_3) = 27^2\ \text{Var}(X_1) + 125^2\ \text{Var}(X_2) + 512^2\ \text{Var}(X_3)$
$= 27^2\ (10)^2 + 125^2\ (12)^2 + 512^2\ (8)^2 = 19{,}100{,}116$.

b. The expected value is still correct, but the variance is not because the covariances now also contribute to the variance.

c. Let V = volume. From **a**, $E(V) = 87{,}850$ and $\text{Var}(V) = 19{,}100{,}116$ assuming the X's are independent. If they are also normally distributed, then V is normal, and so

$P(V > 100{,}000) = 1 - \Phi\left(\dfrac{100{,}000 - 87{,}850}{\sqrt{19{,}100{,}116}}\right) = 1 - \Phi(2.78) = .0027$.

45. Y is normally distributed with $\mu_Y = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5) = -1$, and

$\sigma_Y^2 = \frac{1}{4}\sigma_1^2 + \frac{1}{4}\sigma_2^2 + \frac{1}{9}\sigma_3^2 + \frac{1}{9}\sigma_4^2 + \frac{1}{9}\sigma_5^2 = 3.167 \Rightarrow \sigma_Y = 1.7795$.

Thus, $P(0 \le Y) = P\left(\dfrac{0-(-1)}{1.7795} \le Z\right) = P(.56 \le Z) = .2877$ and

$P(-1 \le Y \le 1) = P\left(0 \le Z \le \dfrac{2}{1.7795}\right) = P(0 \le Z \le 1.12) = .3686$.

47. $E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 15 + 30 + 20 = 65$ min, and
$\text{Var}(X_1 + X_2 + X_3) = 1^2 + 2^2 + 1.5^2 = 7.25 \Rightarrow \text{SD}(X_1 + X_2 + X_3) = 2.6926$ min.
Thus, $P(X_1 + X_2 + X_3 \le 60) = P\left(Z \le \frac{60-65}{2.6926}\right) = P(Z \le -1.86) = .0314.$

49. Let $X_1, \ldots, X_5$ denote morning times and $X_6, \ldots, X_{10}$ denote evening times.

a. $E(X_1 + \ldots + X_{10}) = E(X_1) + \ldots + E(X_{10}) = 5E(X_1) + 5E(X_6) = 5(4) + 5(5) = 45$ min.

b. $\text{Var}(X_1 + \ldots + X_{10}) = \text{Var}(X_1) + \ldots + \text{Var}(X_{10}) = 5\text{Var}(X_1) + 5\text{Var}(X_6) = 5\left[\frac{64}{12} + \frac{100}{12}\right] = \frac{820}{12} = 68.33.$

c. $E(X_1 - X_6) = E(X_1) - E(X_6) = 4 - 5 = -1$ min, while
$\text{Var}(X_1 - X_6) = \text{Var}(X_1) + (-1)^2\text{Var}(X_6) = \frac{64}{12} + \frac{100}{12} = \frac{164}{12} = 13.67$.

d. $E[(X_1 + \ldots + X_5) - (X_6 + \ldots + X_{10})] = 5(4) - 5(5) = -5$ min, while
$\text{Var}[(X_1 + \ldots + X_5) - (X_6 + \ldots + X_{10})] = \text{Var}(X_1 + \ldots + X_5) + (-1)^2\text{Var}(X_6 + \ldots + X_{10}) = 68.33$, the same variance as for the sum in (b).

51.

a. With $M = 5X_1 + 10X_2$, $E(M) = 5(2) + 10(4) = 50$,
$\text{Var}(M) = 5^2 (.5)^2 + 10^2 (1)^2 = 106.25$ and $\sigma_M = 10.308$.

b. $P(75 < M) = P\left(\frac{75-50}{10.308} < Z\right) = P(2.43 < Z) = .0075$.

c. $M = A_1X_1 + A_2X_2$ with the A_i and X_i all independent, so
$E(M) = E(A_1X_1) + E(A_2X_2) = E(A_1)E(X_1) + E(A_2)E(X_2) = 50$.

d. $\text{Var}(M) = E(M^2) - [E(M)]^2$. Recall that for any rv Y, $E(Y^2) = \text{Var}(Y) + [E(Y)]^2$.
Thus, $E(M^2) = E\left(A_1^2X_1^2 + 2A_1X_1A_2X_2 + A_2^2X_2^2\right)$
$= E\left(A_1^2\right)E\left(X_1^2\right) + 2E\left(A_1\right)E\left(X_1\right)E\left(A_2\right)E\left(X_2\right) + E\left(A_2^2\right)E\left(X_2^2\right)$ (by independence)
$= (.25 + 25)(.25 + 4) + 2(5)(2)(10)(4) + (.25 + 100)(1 + 16) = 2611.5625$, so
$\text{Var}(M) = 2611.5625 - (50)^2 = 111.5625$.

e. $E(M) = 50$ still, but now $\text{Cov}(X_1, X_2) = (.5)(.5)(1.0) = .25$, so
$\text{Var}(M) = a_1^2\text{Var}(X_1) + 2a_1a_2\text{Cov}(X_1, X_2) + a_2^2\text{Var}(X_2) = 6.25 + 2(5)(10)(.25) + 100 = 131.25.$

53. Let X_1 and X_2 denote the (constant) speeds of the two planes.

a. After two hours, the planes have traveled $2X_1$ km and $2X_2$ km, respectively, so the second will not have caught the first if $2X_1 + 10 > 2X_2$, i.e. if $X_2 - X_1 < 5$.
$X_2 - X_1$ has a mean $500 - 520 = -20$, variance $100 + 100 = 200$, and standard deviation 14.14. Thus,
$P(X_2 - X_1 < 5) = P\left(Z < \frac{5-(-20)}{14.14}\right) = P(Z < 1.77) = .9616.$

b. After two hours, #1 will be $10 + 2X_1$ km from where #2 started, whereas #2 will be $2X_2$ from where it started. Thus, the separation distance will be at most 10 if
$|2X_2 - 10 - 2X_1| \le 10$, i.e. $-10 \le 2X_2 - 10 - 2X_1 \le 10$ or $0 \le X_2 - X_1 \le 10$. The corresponding probability is $P(0 \le X_2 - X_1 \le 10) = P(1.41 \le Z \le 2.12) = .9830 - .9207 = .0623$.

55.

a. $E(Y_i) = p = \dfrac{1}{2}$, so $E(W) = \sum_{i=1}^{n} i \cdot E(Y_i) = \dfrac{1}{2}\sum_{i=1}^{n} i = \dfrac{n(n+1)}{4}$.

b. $\text{Var}(Y_i) = p(1-p) = \dfrac{1}{4}$, so $\text{Var}(W) = \sum_{i=1}^{n} \text{Var}(i \cdot Y_i) = \sum_{i=1}^{n} i^2 \cdot \text{Var}(Y_i) = \dfrac{1}{4}\sum_{i=1}^{n} i^2 = \dfrac{n(n+1)(2n+1)}{24}$.

57. The total elapsed time between leaving and returning is $T = X_1 + X_2 + X_3 + X_4$, with $E(T) = 15 + 5 + 8 + 12 = 40$ minutes and $\text{Var}(T) = 4^2 + 1^2 + 2^2 + 3^2 = 30$. T is normally distributed, and the desired value t is the 99$^{\text{th}}$ percentile of the lapsed time distribution added to 10a.m.:

$.99 = P(T \le t) = \Phi\left(\dfrac{t-40}{\sqrt{30}}\right) \Rightarrow t = 40 + 2.33\sqrt{30} = 52.76$ minutes past 10a.m., or 10:52.76a.m.

59. Note: $\exp(u)$ will be used as alternate notation for e^u throughout this solution.

a. Using the theorem from this section,

$f_W(w) = f_X \star f_Y = \int_{-\infty}^{\infty} f_X(x) f_Y(w-x)\,dx = \int_{-\infty}^{\infty} \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{x^2}{2}\right) \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{(w-x)^2}{2}\right) dx =$

$\dfrac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\dfrac{x^2+(w-x)^2}{2}\right) dx$. Complete the square inside the exponential function to get

$\dfrac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\dfrac{x^2+(w-x)^2}{2}\right) dx = \dfrac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-(x-w/2)^2 - w^2/4\right) dx = \dfrac{e^{-w^2/4}}{2\pi}\int_{-\infty}^{\infty} e^{-(x-w/2)^2}\,dx =$

$\dfrac{e^{-w^2/4}}{2\pi}\int_{-\infty}^{\infty} e^{-u^2}\,dx$ under the substitution $u = x - w/2$.

This last integral is Euler's famous integral, which equals $\sqrt{\pi}$; equivalently, $\dfrac{1}{\sqrt{\pi}}e^{-u^2}$ is the pdf of a normal rv with mean 0 and variance 1/2, which establishes the same result (because pdfs must integrate to 1). Either way, at long last we have $f_W(w) = \dfrac{e^{-w^2/4}}{2\pi}\sqrt{\pi} = \dfrac{1}{\sqrt{4\pi}}e^{-w^2/4}$. This is the normal pdf with $\mu = 0$ and $\sigma^2 = 2$, so we have proven that $W \sim N(0, \sqrt{2})$.

b. Since X and Y are independent and normal, $W = X + Y$ is also normal. The mean of W is $E(X) + E(Y) = 0 + 0 = 0$ and the variance of W is $\text{Var}(X) + \text{Var}(Y) = 1^2 + 1^2 = 2$. This is obviously <u>much</u> easier than convolution in this case!

61.

a. Since the conditions of a binomial experiment are clearly met, $X \sim \text{Bin}(10, 18/38)$.

b. Similarly, $Y \sim \text{Bin}(15, 18/38)$. Notice that X and Y have different n's but the same p.

c. $X + Y$ is the combined number of times Matt and Liz won. They played a total of 25 games, all independent and identical, with $p = 18/38$ for every game. So, it appears that the conditions of a binomial experiment are met again and that, in particular, $X + Y \sim \text{Bin}(25, 18/38)$.

d. The mgf of a Bin(n, p) rv is $M(t) = (1-p+pe^t)^n$. Using the proposition from this section and the independence of X and Y, the mgf of $X + Y$ is

$$M_{X+Y}(t) = M_X(t)\cdot M_Y(t) = \left(\frac{20}{38}+\frac{18}{38}e^t\right)^{10}\cdot\left(\frac{20}{38}+\frac{18}{38}e^t\right)^{15} = \left(\frac{20}{38}+\frac{18}{38}e^t\right)^{25}$$. This is the mgf of a Bin(25, 18/38) rv. Hence, by uniqueness of mgfs, $X + Y \sim$ Bin(25, 18/38) as predicted.

e. Let $W = X_1 + \ldots + X_k$. Using the same technique as in **d**,

$$M_W(t) = M_{X_1}(t)\cdots M_{X_k}(t) = (1-p+pe^t)^{n_1}\cdots(1-p+pe^t)^{n_k} = (1-p+pe^t)^{n_1+\cdots+n_k}$$. This is the mgf of a binomial rv with parameters $n_1+\cdots+n_k$ and p. Hence, by uniqueness of mgfs, $W \sim$ Bin($\sum n_i$, p).

f. <u>No</u>, for two (equivalent) reasons. Algebraically, we cannot combine the terms in **d** or **e** if the p's differ. Going back to **c**, the combined experiment of 25 trials would not meet the "constant probability" condition of a binomial experiment if Matt and Liz's success probabilities were different. Hence, $X + Y$ would not be binomially distributed.

63. This is a simple extension of the previous exercise. The mgf of X_i is $\left(\dfrac{pe^t}{1-(1-p)e^t}\right)^{r_i}$. Assuming the X's are independent, the mgf of their sum is

$$M_{X_1+\cdots+X_n}(t) = M_{X_1}(t)\cdots M_{X_n}(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^{r_1}\cdots\left(\frac{pe^t}{1-(1-p)e^t}\right)^{r_n} = \left(\frac{pe^t}{1-(1-p)e^t}\right)^{r_1+\cdots+r_n}.$$

This is the mgf of a negative binomial rv with parameters $r_1+\cdots+r_n$ and p. Hence, by uniqueness of mgfs, $X_1+\cdots+X_n \sim$ NB($\sum r_i$, p).

65.

a. The pdf of X is $f_X(x) = \lambda e^{-\lambda x}$ for $x > 0$, and this is also the pdf of Y. The pdf of $W = X + Y$ is $f_W(w) = f_X \star f_Y = \int_{-\infty}^{\infty} f_X(x) f_Y(w-x)dx = \int \lambda e^{-\lambda x}\cdot\lambda e^{-\lambda(w-x)}dx = \int \lambda^2 e^{-\lambda w}dx$, where the limits of integration are determined by the pair of constraints $x > 0$ and $w - x > 0$. These are clearly equivalent to $0 < x < w$, so $f_W(w) = \int_0^w \lambda^2 e^{-\lambda w}dx = \lambda^2 w e^{-\lambda w}$ for $w > 0$ (since $x + y > 0$). Matching terms with the gamma pdf, we identify this as a gamma distribution with $\alpha = 2$ and $\beta = 1/\lambda$.

b. If $X \sim$ Exponential(λ) = Gamma(1, $1/\lambda$) and $Y \sim$ Exponential(λ) = Gamma(1, $1/\lambda$) as well, then by the previous exercise (with $\alpha_1 = \alpha_2 = 1$ and $\beta = 1/\lambda$) we have $X + Y \sim$ Gamma(1 + 1, $1/\lambda$) = Gamma(2, $1/\lambda$).

c. More generally, if $X_1, \ldots, X_n$ are independent Exponential(λ) rvs, then their sum has a Gamma(n, $1/\lambda$) distribution. This can be established through mgfs:

$$M_{X_1+\cdots+X_n}(t) = M_{X_1}(t)\cdots M_{X_n}(t) = \frac{\lambda}{\lambda-t}\cdots\frac{\lambda}{\lambda-t} = \left(\frac{\lambda}{\lambda-t}\right)^n = \frac{1}{(1-(1/\lambda)t)^n}$$, which is precisely the mgf of the Gamma(n, $1/\lambda$) distribution.

Section 4.4

67.

a. The goal is to find $P(Y > 2 \mid X = 3)$. First find the marginal pdf of X, then form the conditional pdf:

$$f_X(x) = \int_0^{10-x} c[10-(x+y)]dy = \frac{c}{2}(10-x)^2 \Rightarrow f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)} = \frac{2[10-(x+y)]}{(10-x)^2} \text{ for } 0 < y < 10-x.$$

Notice that we don't need to find the constant of integration, c, for this part!

Finally, $X = 3 \Rightarrow 0 < Y < 7$, so $P(Y > 2 \mid X = 3) = \int_2^7 f(y \mid 3)dy = \int_2^7 \frac{2(7-y)}{(10-3)^2}dy = \frac{25}{49} \approx .5102$.

b. We'll need c for this part: $1 = \int_0^{10}\int_0^{10-x} c[10-(x+y)]dydx = \frac{500c}{3} \Rightarrow c = .006$.

The probability p that one such system lasts more than three months is the probability that both of its components last more than three months:

$p = P(X > 3 \cap Y > 3) = \int_3^{10}\int_3^{10-x} .006[10-(x+y)]dydx = \cdots = .064$. Finally, the number of systems (out of 20) that meet this criterion follows a binomial distribution with $n = 20$ and $p = .064$. So, the probability that at least half of these 20 systems last more than three months equals

$$\sum_{k=10}^{20} \binom{20}{k} (.064)^k (1-.064)^{20-k} = .000000117.$$

69.

a. $Y|X=x$ is Unif$[0,x^2]$. So, $E(Y|X=x) = (0+x^2)/2 = x^2/2$ and $\text{Var}(Y|X=x) = (x^2-0)^2/12 = x^4/12$.

b. $f(x,y) = f_X(x) \cdot f_{Y|X}(y|x) = 1/(1-0) \cdot 1/(x^2-0) = 1/x^2$ for $0 < y < x^2 < 1$.

c. $f_Y(y) = \int f(x,y)dx = \int_{\sqrt{y}}^1 (1/x^2)dx = \frac{1}{\sqrt{y}} - 1$, $0 < y < 1$.

71.

a. Since all ten digits are equally likely, $p_X(x) = 1/10$ for $x = 0,1,\ldots,9$.
Next, $p_{Y|X}(y|x) = 1/9$ for $y = 0,1,\ldots,9$, $y \neq x$. (That is, any of the 9 remaining digits are equally likely.)
Combining, $p(x, y) = p_X(x) \cdot p_{Y|X}(y|x) = 1/90$ for (x,y) satisfying $x,y = 0,1,\ldots,9$, $y \neq x$.

b. $E(Y|X = x) = \sum_{y\neq x} y\, p_{Y|X}(y|x) = (1/9) \sum_{y\neq x} y = (1/9)\,[0 + 1 + \ldots + 9 - x] = (1/9)(45 - x) = 5 - x/9$.

73.

a. $f_X(x) = \int_0^x f(x,y)dy = \int_0^x 2dy = 2x$, $0 < x < 1$.

b. $f_{Y|X}(y|x) = f(x, y)/f_X(x) = 2/2x = 1/x$, $0 < y < x$. That is, $Y|X=x$ is Uniform on $(0, x)$. We will use this repeatedly in what follows.

c. From (b), $P(0 < Y < .3 \mid X = .5) = .3/.5 = .6$.

d. No, the conditional distribution $f_{Y|X}(y|x)$ actually depends on x.

e. From (b), $E(Y|X=x) = (0 + x)/2 = x/2$.

f. From (b), $\text{Var}(Y|X=x) = (x - 0)^2/12 = x^2/12$.

75.

a. For valid x and y, $p(x, y) = \dfrac{2!}{x!\,y!(2-x-y)!}(.3)^x(.2)^y(.5)^{2-x-y}$. These are displayed in the table below.

x \ y	0	1	2
0	.25	.20	.04
1	.30	.12	0
2	.09	0	0

b. The marginal distributions of X and Y are Bin(2, .3) and Bin(2, .2), respectively. From these or from summation on the table in (a), we get the following.

x	0	1	2
$p_X(x)$	.49	.42	.09

y	0	1	2
$p_Y(y)$	.64	.32	.04

c. $p_{Y|X}(y|x) = p(x, y)/p_X(x) = \dfrac{2!}{x!\,y!(2-x-y)!}(.3)^x(.2)^y(.5)^{2-x-y} \div \dfrac{2!}{x!(2-x)!}(.3)^x(.7)^{2-x} =$

$\dfrac{(2-x)!}{y!(2-x-y)!}\left(\dfrac{.2}{.7}\right)^y\left(\dfrac{.5}{.7}\right)^{2-x-y}$, which is the Bin($2 - x$, .2/.7) distribution.

d. No, their values are restricted by $x + y \le 2$.

e. From (c), $E(Y|X=x) = np = (2 - x)(.2/.7) = (4 - 2x)/7$. This can also be calculated numerically by expanding the formula in (c).

f. From (c), $\text{Var}(Y|X=x) = np(1 - p) = (2 - x)(.2/.7)(1-.2/.7) = 10(2 - x)/49$. Again, this can also be calculated numerically by expanding the formula in (c).

77.

a. $Y|X=x$ is Uniform(0, x), so $E(Y|X=x) = (0 + x)/2 = x/2$ and $\text{Var}(Y|X=x) = (x - 0)^2/12 = x^2/12$.

b. $f(x, y) = f_X(x) \cdot f_{Y|X}(y|x) = 1/(1 - 0) \cdot 1/(x - 0) = 1/x$ for $0 < y < x < 1$.

c. $f_Y(y) = \int_y^1 f(x, y)dx = \int_y^1 (1/x)dx = \ln(1) - \ln(y) = -\ln(y)$, $0 < y < 1$. [Note: since $0 < y < 1$, $\ln(y)$ is actually negative, and the pdf is indeed positive.]

79. We have $X \sim$ Poisson(100) and $Y|X=x \sim$ Bin(x, .6).

a. $E(Y|X=x) = np = .6x$, and $\text{Var}(Y|X=x) = np(1-p) = x(.6)(.4) = .24x$.

b. From **a**, $E(Y|X) = .6X$. Then, from the Law of Total Expectation, $E(Y) = E[E(Y|X)] = E(.6X) = .6E(X) = .6(100) = 60$. This is the common sense answer given the specified parameters.

c. From **a**, $E(Y|X) = .6X$ and $\text{Var}(Y|X) = .24X$. Since X is Poisson(100), $E(X) = \text{Var}(X) = 100$. By the Law of Total Variance,
$\text{Var}(Y) = E(\text{Var}(Y\mid X)) + \text{Var}(E(Y\mid X)) = E(.24X) + \text{Var}(.6X) = .24E(X) + .6^2\,\text{Var}(X)$ = .24(100) + .36(100) = 60.

81. We're give $E(Y|X) = 4X - 104$ and $\text{SD}(Y|X) = .3X - 17$. Down the road we'll need $E(X^2) = \text{Var}(X) + [E(X)]^2 = 3^2 + 70^2 = 4909$.
By the Law of Total Expectation, the unconditional mean of Y is
$E(Y) = E[E(Y|X)] = E(4X - 104) = 4E(X) - 104 = 4(70) - 104 = 176$ pounds.
By the Law of Total Variance, the unconditional variance of Y is
$\text{Var}(Y) = \text{Var}(E(Y\mid X)) + E(\text{Var}(Y\mid X)) = \text{Var}(4X - 104) + E[(.3X - 17)^2] = 4^2\text{Var}(X) + E[.09X^2 - 10.2X + 289] = 16(9) + .09(4909) - 10.2(70) + 289 = 160.81$.
Thus, $\text{SD}(Y) = 12.68$ pounds.

83.

a. $E(X) = E(1 + N) = 1 + E(N) = 1 + 4p$. $\text{Var}(X) = \text{Var}(1 + N) = \text{Var}(N) = 4p(1 - p)$.

b. Let W denote the winnings from one chip. Using the pmf, $\mu = E(W) = 0(.39) + \ldots + 10{,}000(.23) = \$2598$ and $\sigma^2 = \text{Var}(W) = 16{,}518{,}196$.

c. By the Law of Total Expectation, $E(Y) = E[E(Y|X)] = E[\mu X] = \mu\, E(X) = 2598(1 + 4p)$. By the Law of Total Variance, $\text{Var}(Y) = \text{Var}(E(Y\mid X)) + E(\text{Var}(Y\mid X)) = \text{Var}(\mu X) + E(\sigma^2 X) = \mu^2\text{Var}(X) + \sigma^2 E(X) = (2598)^2 \cdot 4p(1 - p) + 16{,}518{,}196(1 + 4p)$. Simplifying and taking the square root gives
$\text{SD}(Y) = \sqrt{16518196 + 93071200p - 26998416p^2}$.

d. When $p = 0$, $E(Y) = \$2598$ and $\text{SD}(Y) = \$4064$. If the contestant always guesses incorrectly, s/he will get exactly 1 chip and the answers from **b** apply.
When $p = .5$, $E(Y) = \$7794$ and $\text{SD}(Y) = \$7504$.
When $p = 1$, $E(Y) = \$12{,}990$ and $\text{SD}(Y) = \$9088$.
As the ability to acquire chips improves, so does the contestant's expected payout. The variability around that expectation also increases (since the set of options widens), but the standard deviation does not quite increase linearly with p.

Section 4.5

85.

a. The sampling distribution of $\bar{X}$ is centered at $E(\bar{X}) = \mu = 12$ cm, and the standard deviation of the $\bar{X}$ distribution is $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{.04}{\sqrt{16}} = .01$ cm.

b. With $n = 64$, the sampling distribution of $\bar{X}$ is still centered at $E(\bar{X}) = \mu = 12$ cm, but the standard deviation of the $\bar{X}$ distribution is $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{.04}{\sqrt{64}} = .005$ cm.

c. $\bar{X}$ is more likely to be within .01 cm of the mean (12 cm) with the second, larger, sample. This is due to the decreased variability of $\bar{X}$ that comes with a larger sample size.

87.

a. Let $\bar{X}$ denote the sample mean fracture angle of our $n = 4$ specimens. Since the individual fracture angles are normally distributed, $\bar{X}$ is also normal, with mean $E(\bar{X}) = \mu = 53$ but with standard deviation $\text{SD}(\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{4}} = .5$. Hence,

$P(\bar{X} \le 54) = \Phi\left(\frac{54-53}{.5}\right) = \Phi(2) = .9772$, and

$P(53 \le \bar{X} \le 54) = \Phi(2) - \Phi(0) = .4772$.

b. Replace 4 with n, and set the probability expression equal to .999:

$.999 = P(\bar{X} \le 54) = \Phi\left(\frac{54-53}{1/\sqrt{n}}\right) = \Phi(\sqrt{n}) \Rightarrow \sqrt{n} \approx 3.09 \Rightarrow n \approx 9.5$. Since n must be a whole number, round up: the least such n is $n = 10$.

89. We have $X \sim N(10,1)$, $n = 4$, $\mu_T = n\mu = (4)(10) = 40$ and $\sigma_T = \sigma\sqrt{n} = 2$. Hence, $T \sim N(40, 2)$. We desire the 95th percentile of T: $40 + (1.645)(2) = 43.29$ hours.

91. Let T denote the total commute time for the week. Each waiting time has mean $(0 + 10)/2 = 5$ and variance $(10 - 0)^2/12 = 100/12$. Since there are $2 \times 6 = 12$ waiting times of interest, T has mean $12(5) = 60$ and variance $12(100/12) = 100$; furthermore, T has an approximately normal distribution even for $n = 12$ in this case. Combining these facts, $P(T \le 75) \approx \Phi\left(\frac{75-60}{\sqrt{100}}\right) = \Phi(1.5) = .9332$.

93.

a. Let $\overline{X}$ denote the sample mean tip percentage for these 40 bills. By the Central Limit Theorem, $\overline{X}$ is approximately normal, with $E(\overline{X}) = \mu = 18$ and $\text{SD}(\overline{X}) = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{40}}$. Hence,

$$P(16 \le \overline{X} \le 19) \approx \Phi\left(\frac{19-18}{6/\sqrt{40}}\right) - \Phi\left(\frac{16-18}{6/\sqrt{40}}\right) = \Phi(1.05) - \Phi(-2.11) = .8357.$$

b. According to the common convention, n should be greater than 30 in order to apply the C.L.T., thus using the same procedure for $n = 15$ as was used for $n = 40$ would not be appropriate.

95.

a. Let $\overline{X}$ denote the sample average amount of gas purchased by 50 customers. By the Central Limit Theorem, $\overline{X}$ is approximately normal, with $E(\overline{X}) = \mu = 11.5$ gallons and $\text{SD}(\overline{X}) = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{50}}$ gallons. Hence $P(\overline{X} \ge 12) \approx 1 - \Phi\left(\frac{12-11.5}{4/\sqrt{50}}\right) = 1 - \Phi(0.88) = .1894$.

b. Let T denote the total amount of gas purchased by 50 customers. Notice that the customers' <u>total</u> purchase is at least 600 gallons if and only if their <u>average</u> purchase is 600/50 = 12 gallons. In other words, questions **a** and **b** are really the same! Thus, $P(T \ge 600) \approx .1894$.

c. By C.L.T., T is approximately normal with mean 50(11.5) = 575 gallons and standard deviation $\sqrt{50}(4) = 28.28$ gallons. Since the 95th percentile of the standard normal distribution is ~1.645, the 95th percentile of T is 575 + 1.645(28.28) = 621.5 gallons.

97. Let X = the number of students out of 50 whose programs contain no errors. Then $X \sim \text{Bin}(50, .4)$; the Central Limit Theorem implies that X is approximately normal with mean $\mu = np = 20$ and standard deviation $\sigma = \sqrt{npq} = 3.464$. (We saw this normal approximation in Chapter 3, but now we know it's justified by C.L.T.)

a. $P(X \ge 25) = P(X \ge 24.5) \approx 1 - \Phi\left(\frac{24.5-20}{3.464}\right) = 1 - \Phi(1.30) = .0968$.

b. Similarly, $P(15 \le X \le 25) \approx \Phi\left(\frac{25.5-20}{3.464}\right) - \Phi\left(\frac{14.5-20}{3.464}\right) = \Phi(1.59) - \Phi(-1.59) = .8882$.

99. When α is large, a Gamma(α, β) is approximately normal (because it's the sum of iid exponential rvs). Here, $E(X) = \alpha\beta = 100$, $\text{Var}(X) = \alpha\beta^2 = 200$, so $P(X \le 125) \approx \Phi\left(\frac{125-100}{\sqrt{200}}\right) = \Phi(1.77) = .9616$.

101. Assume you have a random sample $X_1, \ldots, X_n$ from an exponential distribution with parameter λ. Let $\bar{X}$ denote their sample average. Then by the law of large numbers, $\bar{X} \to E(X) = \dfrac{1}{\lambda}$ as $n \to \infty$. But our goal is a consistent estimator of λ, i.e. a quantity that converges to λ itself as $n \to \infty$.
The solution is obvious: let $g(t) = 1/t$, which is continuous for all $t > 0$. Then by the theorem cited in the exercise, $g(\bar{X}) \to g(1/\lambda)$. In other words, the consistent estimator is

$Y_n = \dfrac{1}{\bar{X}} \to \dfrac{1}{1/\lambda} = \lambda$ as $n \to \infty$.

Section 4.6

103. If X_1 and X_2 are independent, standard normal rvs, then $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) = \dfrac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \cdot \dfrac{1}{\sqrt{2\pi}} e^{-x_2^2/2} = \dfrac{1}{2\pi} e^{-(x_1^2 + x_2^2)/2}$.

a. Solve the given equations for X_1 and X_2: by adding, $Y_1 + Y_2 = 2X_1 \Rightarrow X_1 = (Y_1 + Y_2)/2$. Similarly, subtracting yields $X_2 = (Y_1 - Y_2)/2$. Hence, the Jacobian of this transformation is

$$\det(M) = \begin{vmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 \end{vmatrix} = \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} = (1/2)(-1/2) - (1/2)(1/2) = -1/2$$

Also, the sum of squares in the exponent of the joint pdf above can be re-written:

$$x_1^2 + x_2^2 = \left(\frac{y_1 + y_2}{2}\right)^2 + \left(\frac{y_1 - y_2}{2}\right)^2 = \frac{y_1^2 + 2y_1y_2 + y_2^2 + y_1^2 - 2y_1y_2 + y_2^2}{4} = \frac{y_1^2 + y_2^2}{2}$$

Finally, the joint pdf of Y_1 and Y_2 is

$$f(y_1, y_2) = \frac{1}{2\pi} e^{-((y_1^2 + y_2^2)/2)/2} \cdot |-1/2| = \frac{1}{4\pi} e^{-(y_1^2 + y_2^2)/4}$$

b. To obtain the marginal pdf of Y_1, integrate out Y_2:

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} \frac{1}{4\pi} e^{-(y_1^2 + y_2^2)/4} dy_2 = \frac{1}{\sqrt{4\pi}} e^{-y_1^2/4} \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi}} e^{-y_2^2/4} dy_2$$

The integrand is the pdf of a normal distribution with $\mu = 0$ and $\sigma^2 = 2$ (so $2\sigma^2 = 4$). Since it's a pdf, its integral is 1, and we're left with $f_{Y_1}(y_1) = \dfrac{1}{\sqrt{4\pi}} e^{-y_1^2/4}$. This is (also) a normal pdf with mean 0 and variance 2, which we know to be the distribution of the sum of two independent $N(0, 1)$ rvs.

c. Yes, Y_1 and Y_2 are independent. Repeating the math in **b** to obtain the marginal pdf of Y_2 yields $f_{Y_2}(y_2) = \dfrac{1}{\sqrt{4\pi}} e^{-y_2^2/4}$, from which we see that $f(y_1, y_2) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2)$. Thus, by definition, Y_1 and Y_2 are independent.

105. Let $Y = X_1 + X_2$ and $W = X_2 - X_1$, so $X_1 = (Y - W)/2$ and $X_2 = (Y + W)/2$. We will find their joint distribution, and then their marginal distributions to answer **a** and **b**.

The Jacobian of the transformation is $\det\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$ = 1/2.

Graph the triangle $0 \le x_1 \le x_2 \le 1$ and transform this into the (y, w) plane. The result is the triangle bounded by $y = 0$, $w = y$, and $w = 2 - y$. Therefore, on this triangle, the joint pdf of Y and W is

$$f(y,w) = 2\left(\frac{y-w}{2} + \frac{y+w}{2}\right)\cdot\left|\frac{1}{2}\right| = y$$

a. For $0 \le y \le 1$, $f_Y(y) = \int_0^y y\,dw = y^2$; for $1 \le y \le 2$, $f_Y(y) = \int_0^{2-y} y\,dw = y(2 - y)$.

b. For $0 \le w \le 1$, $f_W(w) = \int_w^{2-w} y\,dy = \ldots = 2(1 - w)$.

107. Solving for the X's gives $X_1 = Y_1$, $X_2 = Y_2/Y_1$, and $X_3 = Y_3/Y_2$. The Jacobian of the transformation is $\det\begin{bmatrix} 1 & - & - \\ 0 & 1/y_1 & - \\ 0 & 0 & 1/y_2 \end{bmatrix}$ = $1/y_1y_2$; the dashes indicate partial derivatives that don't matter. Thus, the joint pdf of the Y's is

$f(y_1, y_2, y_3) = 8y_3 \cdot |1/y_1y_2| = \dfrac{8y_3}{y_1y_2}$ for $0 < y_3 < y_2 < y_1 < 1$. The marginal pdf of Y_3 is

$$f_{Y_3}(y_3) = \int_{y_3}^1\int_{y_2}^1 \frac{8y_3}{y_1y_2}\,dy_1dy_2 = \int_{y_3}^1 -\frac{8y_3}{y_2}\ln(y_2)\,dy_2 = \int_{\ln(y_3)}^0 -8y_3u\,du = 4y_3[\ln(y_3)]^2 \text{ for } 0 < y_3 < 1.$$

109. If $U \sim$ Unif(0, 1), then $Y = -2\ln(U)$ has an exponential distribution with parameter $\lambda = ½$ (mean 2); see the section on one-variable transformations in the previous chapter. Likewise, $2\pi U$ is Uniform on $(0, 2\pi)$. Hence, Y_1, and Y_2 described here are precisely the random variables that result in the previous exercise, and the transformations $z_1 = \sqrt{y_1}\cos(y_2)$, $z_2 = \sqrt{y_1}\sin(y_2)$ restore the original independent standard normal random variables in that exercise.

Section 4.7

111.

a. Since X and W are bivariate normal, $X + W$ has a (univariate) normal distribution, with mean $E(X+W) = E(X) + E(W) = 496 + 488 = 984$ and variance given by
$\text{Var}(X + W) = \text{Var}(X) + \text{Var}(W) + 2\,\text{Cov}(X, W) = \text{Var}(X) + \text{Var}(W) + 2\,\text{SD}(X)\,\text{SD}(W)\,\text{Corr}(X, W) =$
$114^2 + 114^2 + 2(114)(114)(.5) = 38{,}988$. Equivalently, $\text{SD}(X + W) = \sqrt{38{,}988} = 197.45$.
That is, $X + W \sim N(984, 197.45)$.

b. $P(X + W > 1200) = 1 - \Phi\left(\dfrac{1200-984}{197.45}\right) = 1 - \Phi(1.09) = .1379$.

c. We're looking for the 90th percentile of the $N(984, 197.45)$ distribution:
$.9 = \Phi\left(\dfrac{x-984}{197.45}\right) \Rightarrow \dfrac{x-984}{197.45} = 1.28 \Rightarrow x = 984 + 1.28(197.45) = 1237$.

113. As stated in the section, $Y|X=x$ is normal with mean $\mu_2 + \rho\sigma_2\left(\frac{x-\mu_1}{\sigma_1}\right)$ and variance $(1-\rho^2)\sigma_2^2$.

a. Substitute into the above expressions: mean = 170 + $.9(20)\left(\frac{68-70}{3}\right)$ = 158 lbs, variance = $(1 - .9^2)(20)^2 = 76$, sd = 8.72 lbs. That is, the weight distribution of 5'8" tall American males is $N(158, 8.72)$.

b. Similarly, $x = 70$ returns mean = 170 lbs and sd = 8.72 lbs, so the weight distribution of 5'10" tall American males is $N(170, 8.72)$. These two conditional distributions are both normal and both have standard deviation equal to 7.82 lbs, but the average weight differs by height.

c. Plug in $x = 72$ as above to get $Y|X=72 \sim N(182, 8.72)$. Thus, $P(Y < 180 \mid X = 72) = \Phi\left(\frac{180-182}{8.72}\right) = \Phi(-0.23) = .4090$.

115.

a. The mean is $\mu_2 + \rho\sigma_2(x - \mu_1)/\sigma_1 = 70 + (.71)(15)(x - 73)/12 = .8875x + 5.2125$.

b. The variance is $\sigma_2^2(1 - \rho^2) = 15^2(1 - .71^2) = 111.5775$.

c. From **b**, sd =10.563.

d. From **a**, the mean when $x = 80$ is 76.2125. So, $P(Y > 90|X = 80) = 1-\Phi\left(\frac{90-76.2125}{10.563}\right) = 1 - \Phi(1.31) = .0951$.

117.

a. The mean is $\mu_2 + \rho\sigma_2(x - \mu_1)/\sigma_1 = 30 + (.8)(5)(x - 20)/2 = 2x - 10$.

b. The variance is $\sigma_2^2(1 - \rho^2) = 5^2(1 - .8^2) = 9$.

c. From **b**, sd = 3.

d. From **a**, the mean when $x = 25$ is 40. So, $P(Y > 46|X = 25) = 1-\Phi\left(\frac{46-40}{3}\right) = 1 - \Phi(2) = .0228$.

119.

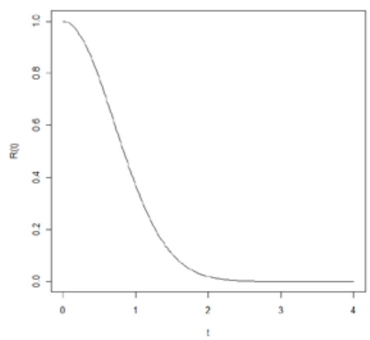
a. $P(50 < X < 100, 20 < Y < 25) = P(X < 100, Y < 25) - P(X < 50, Y < 25) - P(X < 100, Y < 20) + P(X < 50, Y < 20) = .3333 - .1274 - .1274 + .0625 = .1410$.

b. If X and Y are independent, then $P(50 < X < 100, 20 < Y < 25) = P(50 < X < 100) \cdot P(20 < Y < 25) = [\Phi(0) - \Phi(-1)]\,[\Phi(0) - \Phi(-1)] = (.3413)^2 = .1165$. This is smaller than (a). When $\rho > 0$, it's more likely that the event $X < 100$ (its mean) coincides with $Y < 25$ (its mean).

Section 4.8

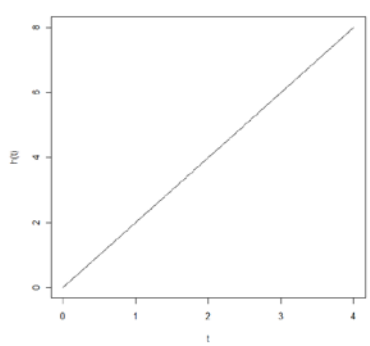
121.

a. From Chapter 3, the Weibull cdf is $F(x)=1-e^{-(x/\beta)^\alpha}$. Hence, the reliability function is $R(t) = 1 - F(t) = 1-[1-e^{-(t/\beta)^\alpha}]=e^{-(t/\beta)^\alpha} = e^{-t^2}$.



b. Let T denote the lifetime of a motor of this type, in thousands of hours. $P(T > 1.5) = R(1.5) = e^{-1.5^2}$ = .1054.

c. The Weibull pdf with $\alpha = 2$ and $\beta = 1$ is $f(t) = 2te^{-t^2}$, so the hazard function for this model is $h(t)=\dfrac{f(t)}{R(t)}=\dfrac{2te^{-t^2}}{e^{-t^2}}=2t$. This particular model has a linearly increasing hazard rate.



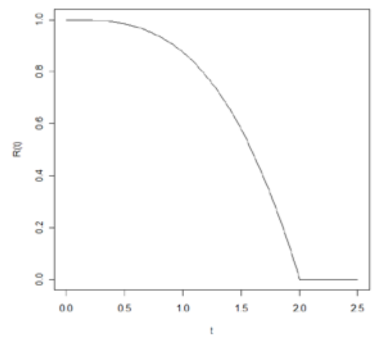
d. Using the proposition from this section, mean time to failure is given by $\mu_T = \int_0^\infty R(t)dt=\int_0^\infty e^{-t^2}\,dt=\int_0^\infty e^{-u}\frac{1}{2\sqrt{u}}du=\frac{1}{2}\int_0^\infty u^{-1/2}e^{-u}du$. Applying the gamma integral formula (with $\alpha = 1/2$ and $\beta = 1$) yields $\mu_T = \frac{1}{2}\Gamma\left(\frac{1}{2}\right)1^{1/2}=\frac{1}{2}\sqrt{\pi}=\frac{\sqrt{\pi}}{2}\approx 0.886$ thousand hours.

The mean of a Weibull rv with $\alpha = 2$ and $\beta = 1$ is $1\Gamma\left(1+\frac{1}{2}\right)=\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{\sqrt{\pi}}{2}$, confirming our answer above.

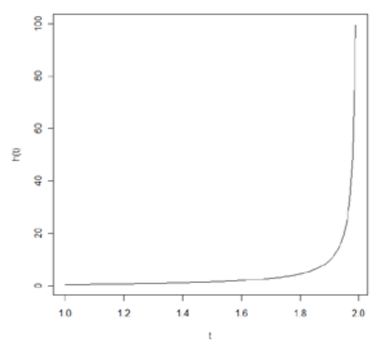
123.

a. For $0 \le t \le 2$, $R(t) = 1 - F(t) = 1 - \int_0^t .375x^2\,dx = 1 - .125t^3$. But since the pdf vanishes for $t > 2$, we must have $R(t) = P(T > t) = 0$ for $t > 2$ as well. Hence, the complete reliability function is

$$R(t) = \begin{cases} 1 - .125t^3 & 0 \le t \le 2 \\ 0 & t > 2 \end{cases}$$



b. For $0 \le t < 2$, $h(t) = \dfrac{f(t)}{R(t)} = \dfrac{.375t^2}{1 - .125t^3} = \dfrac{3t^2}{8 - t^3}$. Technically, $h(t)$ does not exist at $t = 2$.



c. Since $f(t) = 0$ and $R(t) = 0$ for $t > 2$, the hazard function is not defined for $t > 2$.

125.

a. According to the description, the bar's "system" is functional so long as at least one of the four blenders still works. That is the characteristic of a <u>parallel</u> design.

b. From Exercise 121, each of the four blenders has reliability function $R_i(t) = e^{-t^2}$. Since this is a parallel system, from the proposition in this section the system's reliability function is

$$R(t) = 1 - \prod_{i=1}^{4}[1 - R_i(t)] = 1 - (1 - e^{-t^2})^4 .$$

c. As discussed in the last proposition of this section, the hazard function may be written as

$$h(t) = -\frac{R'(t)}{R(t)} = -\frac{\frac{d}{dt}[1 - (1 - e^{-t^2})^4]}{1 - (1 - e^{-t^2})^4} = -\frac{-4(1 - e^{-t^2})^3(2te^{-t^2})}{1 - (1 - e^{-t^2})^4} = \frac{8te^{-t^2}(1 - e^{-t^2})^3}{1 - (1 - e^{-t^2})^4} .$$

d. Before attempting to integrate the reliability function, it will help to expand it out into several separate terms: $R(t) = 1-(1-e^{-t^2})^4 = 4e^{-t^2} - 6e^{-2t^2} + 4e^{-3t^2} - e^{-4t^2}$. Then, make the substitution $u = t^2$ when performing the integral, and invoke the gamma integral formula:

$$\mu_T = \int_0^\infty [4e^{-t^2} - 6e^{-2t^2} + 4e^{-3t^2} - e^{-4t^2}]dt = \int_0^\infty [4e^{-u} - 6e^{-2u} + 4e^{-3u} - e^{-4u}]\frac{1}{2\sqrt{u}}du$$

$$= 2\int_0^\infty u^{-1/2}e^{-u}\,du - 3\int_0^\infty u^{-1/2}e^{-2u}\,du + 2\int_0^\infty u^{-1/2}e^{-3u}\,du - \frac{1}{2}\int_0^\infty u^{-1/2}e^{-4u}\,du$$

$$= 2\Gamma(\tfrac{1}{2}) - 3\Gamma(\tfrac{1}{2})(1/2)^{1/2} + 2\Gamma(\tfrac{1}{2})(1/3)^{1/2} - \frac{1}{2}\Gamma(\tfrac{1}{2})(1/4)^{1/2}$$

With $\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$, this is approximately 1.39, or 1390 hours.

127.

a. Since components 1 and 2 are in parallel, the reliability function of the 1-2 subsystem is given by $1 - (1 - R_1(t))(1 - R_2(t))$. The reliability functions of the 3-4 and 5-6 subsystems are analogous. Finally, since the three subsystems are connected in series, the overall system reliability is $R(t) = [1 - (1 - R_1(t))(1 - R_2(t))][1 - (1 - R_3(t))(1 - R_4(t))][1 - (1 - R_5(t))(1 - R_6(t))]$.

b. All six components have reliability function $R_i(t) = e^{-t/100}$. Hence, $R(t) = [1 - (1 - e^{-t/100})^2]^3 = [2e^{-t/100} - e^{-2t/100}]^3 = 8e^{-3t/100} - 12e^{-4t/100} + 6e^{-5t/100} - e^{-6t/100}$. From this, the mean time to failure is $\mu_T = \int_0^\infty R(t)dt = \int_0^\infty [8e^{-3t/100} - 12e^{-4t/100} + 6e^{-5t/100} - e^{-6t/100}]dt = 8\cdot\frac{100}{3} - 12\cdot\frac{100}{4} + 6\cdot\frac{100}{5} - \frac{100}{6} =$ 70 hours.

129.

a. For $t \le \beta$, $R(t) = e^{-\int_0^t h(u)du} = e^{-\int_0^t \alpha(1-u/\beta)du} = e^{-\alpha(t - t^2/[2\beta])}$. For $t > \beta$,

$$R(t) = e^{-\int_0^t h(u)du} = e^{-\left[\int_0^\beta \alpha(1-u/\beta)du + \int_\beta^t 0\,du\right]} = e^{-\alpha\beta/2}.$$

b. Since the lifetime variable T can be infinite with positive probability, T does not have a legitimate pdf. Working purely algebraically, for $0 < t \le \beta$ we have $f(t) = h(t) \cdot R(t) = \alpha\left(1 - \frac{t}{\beta}\right)e^{-\alpha(t - t^2/[2\beta])}$. This function does not integrate to 1, but rather to $1 - e^{-\alpha\beta/2}$; as indicated by the second part of the reliability function in **a**, the remaining probability of $e^{-\alpha\beta/2}$ is assigned to the event $P(T > t$ for arbitrarily large $t) = P(T = +\infty)$.

131. Suppose $h(t) = c$ for some constant $c > 0$, for all $t > 0$. Then $R(t) = e^{-\int_0^t h(u)du} = e^{-\int_0^t c\,du} = e^{-ct}$, from which the pdf is $f(t) = -R'(t) = ce^{-ct}$ for $t > 0$. This is the exponential distribution with $\lambda = c$.

Section 4.9

133.

a. $f(x) = 1/10 \rightarrow F(x) = x/10 \rightarrow g_5(y) = 5[y/10]^4[1/10] = 5y^4/10^5$ for $0 < y < 10$. Hence, $E(Y_5) = \int_0^{10} y \cdot 5y^4/10^5\,dy = 50/6$, or 8.33 minutes.

b. By the same sort of computation as in **a**, $E(Y_1) = 10/6$, and so $E(Y_5 - Y_1) = 50/6 - 10/6 = 40/6$, or 6.67 minutes.

c. The median waiting time is Y_3; its pdf is $g_3(y) = \frac{5!}{2!1!2!}[F(y)]^2 f(y)[1 - F(y)]^2 = 30y^2(10 - y)^2/10^5$ for $0 < y < 10$. By direct integration, or by symmetry, $E(Y_3) = 5$ minutes (which is also the mean and median of the original Unif[0, 10] distribution).

d. $E(Y_5^2) = \int_0^{10} y^2 \cdot 5y^4/10^5\,dy = 500/7$, so $\text{Var}(Y_5) = 500/7 - (50/6)^2 = 125/63 = 1.984$, from which $SD(Y_5) = 1.409$ minutes.

135. $f(x) = 3/x^4$ for $x > 1 \Rightarrow F(x) = \int_1^x 3/y^4\,dy = 1 - x^{-3}$ for $x > 1$.

a. P(at least one claim > \$5000) = 1 – P(all 3 claims are ≤ \$5000) = $1 - P(X_1 \le 5 \cap X_2 \le 5 \cap X_3 \le 5) = 1 - F(5)\cdot F(5)\cdot F(5)$ by independence $= 1 - (1 - 5^{-3})^3 = .0238$.

b. The pdf of Y_3, the largest claim, is $g_3(y) = 3f(y)[F(y)]^{3-1} = 3(3y^{-4})[1 - y^{-3}]^2 = 9(y^{-4} - 2y^{-7} + y^{-10})$ for $y > 1$. Hence, $E(Y_3) = \int_1^\infty y \cdot 9(y^{-4} - 2y^{-7} + y^{-10})dy = 9\int_1^\infty (y^{-3} - 2y^{-6} + y^{-9})dy = 2.025$, or \$2,025.

137. The pdf of the underlying population distribution is $f(x) = \theta x^{\theta-1}$. The pdf of Y_i is

$g_i(y) = \frac{n!}{(i-1)!(n-i)!}[y^\theta]^{i-1}[1 - y^\theta]^{n-i}[\theta y^{\theta-1}] = \frac{n!\theta}{(i-1)!(n-i)!} y^{i\theta-1}[1 - y^\theta]^{n-i}$. Thus,

$E(Y_i) = \int_0^1 y g_i(y)dy = \frac{n!\theta}{(i-1)!(n-i)!}\int_0^1 y^{i\theta}[1 - y^\theta]^{n-i}dy$ = [via the substitution $u = y^\theta$]

$\frac{n!\theta}{(i-1)!(n-i)!}\int_0^1 u^i[1-u]^{n-i}\frac{u^{1/\theta-1}}{\theta}du = \frac{n!}{(i-1)!(n-i)!}\int_0^1 u^{i+1/\theta-1}(1-u)^{n-i}du$.

The integral is the "kernel" of the Beta($i + 1/\theta$, $n - i + 1$) distribution, and so this entire expression equals $\frac{n!}{(i-1)!(n-i)!}\frac{\Gamma(i+1/\theta)\Gamma(n-i+1)}{\Gamma(n+1/\theta+1)} = \frac{n!\Gamma(i+1/\theta)}{(i-1)!\Gamma(n+1/\theta+1)}$. Similarly, $E(Y_i^2) = \frac{n!\Gamma(i+2/\theta)}{(i-1)!\Gamma(n+2/\theta+1)}$,

from which $\text{Var}(Y_i) = \frac{n!\Gamma(i+2/\theta)}{(i-1)!\Gamma(n+2/\theta+1)} - \left[\frac{n!\Gamma(i+1/\theta)}{(i-1)!\Gamma(n+1/\theta+1)}\right]^2$.

139. The pdf of Y_{k+1} is $g_{k+1}(y) = \frac{n!}{k!k!}F(y)^k[1-F(y)]^k f(y) = \frac{n!}{k!k!}F(y)^k[F(-y)]^k f(y)$ by symmetry. Hence, $E(Y_{k+1}) = \int_{-\infty}^{\infty}\frac{n!}{k!k!}yF(y)^k[F(-y)]^k f(y)dy$. The integrand is an odd function, since $f(y)$ is even by assumption and $F(y)^k[F(-y)]^k$ is clearly also even. Hence, $E(Y_{k+1}) = 0$.

This generalizes to distributions symmetric about an arbitrary median, since we can simply translate the distribution by η, apply this result, and add back: $E(Y_{k+1}) = \eta$.

141. As suggested in the section, divide the number line into five intervals: $(-\infty, y_i]$, $(y_i, y_i + \Delta_1]$, $(y_i + \Delta_1, y_j]$, $(y_j, y_j + \Delta_2]$, and $(y_j + \Delta_2, \infty)$. For a rv X having cdf F, the probability X falls into these five intervals are $p_1 = P(X \le y_i) = F(y_i)$, $p_2 = F(y_i + \Delta_1) - F(y_i) \approx f(y_i)\Delta_1$, $p_3 = F(y_j) - F(y_i + \Delta_1)$, $p_4 = F(y_j + \Delta_2) - F(y_j) \approx f(y_j)\Delta_2$, and $p_5 = P(X > y_j + \Delta_2) = 1 - F(y_j + \Delta_2)$.

Now consider a random sample of size n from this distribution. Let Y_i and Y_j denote the ith and jth smallest values (order statistics) with $i < j$. It is unlikely that more than one X will fall in the 2nd interval or the 4th interval, since they are very small (widths Δ_1 and Δ_2). So, the event that Y_i falls in the 2nd interval and Y_j falls in the 4th interval is approximately the probability that: $i-1$ of the X's falls in the 1st interval; one X (the ith smallest) falls in the 2nd interval; $j-i-1$fall in the 3rd interval; one X (the jth smallest) falls in the 4th interval; and the largest $n-j$ X's fall in the 5th interval. Apply the multinomial formula:

$$P(y_i < Y_i \le y_i + \Delta_1, y_j < Y_j \le y_j + \Delta_2) \approx \frac{n!}{(i-1)!1!(j-i-1)!1!(n-j)!}p_1^{i-1}p_2^1 p_3^{j-i-1}p_4^1 p_5^{n-j}$$

$$\approx \frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(y_i)]^{i-1}f(y_i)\Delta_1[F(y_j)-F(y_i+\Delta_1)]^{j-i-1}f(y_j)\Delta_2[1-F(y_j+\Delta_2)]^{n-j}$$

Dividing the left-hand side by $\Delta_1\Delta_2$ and letting $\Delta_1 \to 0$, $\Delta_2 \to 0$ yields the joint pdf $g(y_i, y_j)$. Taking the same action on the right-hand side returns

$$\frac{n!}{(i-1)!(j-i-1)!(n-j)!}F(y_i)^{i-1}[F(y_j)-F(y_i)]^{j-i-1}[1-F(y_j)]^{n-j}f(y_i)f(y_j)$$, as claimed.

Section 4.10

143.

a. First, create a table of “cumulative” probabilities, starting in the upper left of the joint pmf table.

(x, y)	(0,0)	(0,1)	(0,2)	(1,0)	(1,1)	(1,2)	(2,0)	(2,1)	(2,2)
	.10	.14	.16	.24	.44	.50	.56	.70	1

Then, mimic the exhaustive search code in Figure 4.25.

```
x=zeros(10000,1); y=x;
for i=1:10000
    u=rand;
    if u<.10
        x(i)=0; y(i)=0;
    elseif u<.14
        x(i)=0; y(i)=1;
    elseif u<.16
        x(i)=0; y(i)=2;
    elseif u<.24
        x(i)=1; y(i)=0;
    elseif u<.44
        x(i)=1; y(i)=1;
    elseif u<.50
        x(i)=1; y(i)=2;
    elseif u<.56
        x(i)=2; y(i)=0;
    elseif u<.70
        x(i)=2; y(i)=1;
    else
        x(i)=2; y(i)=2;
    end
end
```

```
x <- NULL; y <- NULL
for (i in 1:10000){
      u=runif(1)
      if (u<.10){
            x[i]<-0; y[i]<-0}
      else if (u<.14){
            x[i]<-0; y[i]<-1}
      else if (u<.16){
            x[i]<-0; y[i]<-2}
      else if (u<.24){
            x[i]<-1; y[i]<-0}
      else if (u<.44){
            x[i]<-1; y[i]<-1}
      else if (u<.50){
            x[i]<-1; y[i]<-2}
      else if (u<.56){
            x[i]<-2; y[i]<-0}
      else if (u<.70){
            x[i]<-2; y[i]<-1}
      else{
            x[i]<-2; y[i]<-2}
}
```

b. For one execution of the R code, `sum(x<=1 & y<=1)` returned a count of 4154, so $\hat{P}(X \le 1, Y \le 1) = \dfrac{4154}{10,000} = .4154$. This is close to the exact answer of .42.

c. Execute the code in **a**, and then add the line `D=abs(x-y)` in Matlab or `D<-abs(x-y)` in R. Then the vector `D` contains 10,000 simulated values of $D = |X - Y|$. The sample mean and sample sd from one such run returned $\hat{\mu}_D$ = 0.4866 and $\hat{\sigma}_D$ = 0.6438.

145.

a. Since each of X and Y lies on [0, 1], the easiest choice for generating candidates is $X^*, Y^* \sim \text{Unif}[0, 1]$. (We'll have to account for the constraint $x + y \le 1$ eventually.) So, $g_1(x) = g_2(y) = 1/(1 - 0) = 1$. The majorization constant c must satisfy $\dfrac{f(x,y)}{g_1(x)g_2(y)} \le c$, i.e. $24xy \le c$, for all (x, y) in the joint domain $x \ge 0, y \ge 0, x + y \le 1$. The maximum of $24xy$ on this triangle occurs at $(x, y) = (1/2, 1/2)$, which in turn requires $c \ge 24(1/2)(1/2) = 6$. Of course, we choose to use $c = 6$.
The key requirement $u \cdot c \cdot g_1(x^*)g_2(y^*) \le f(x^*, y^*)$ is equivalent to $u(6)(1)(1) \le 24x^*y^*$, or $u \le 4x^*y^*$, provided also that $x^* + y^* \le 1$.
The accept-reject scheme then proceeds as in the accompanying code; notice that must include an indicator to insure that $f(x^*, y^*) = 0$ when the coordinates violate the bound $x^* + y^* \le 1$.

```
x=zeros(10000,1); y=zeros(10000,1);
i=0;
while i<10000
    xstar=rand;
    ystar=rand;
    u=rand;
    if u<=4*xstar*ystar*
           (xstar+ystar<=1)
        i=i+1;
        x(i)=xstar;
        y(i)=ystar;
    end
end
```

```
x <- NULL; y <- NULL
i <- 0
while (i <10000){
      xstar <- runif(1)
      ystar <- runif(1)
      u <- runif(1)
      if (u<=4*xstar*ystar*
              (xstar+ystar<=1)){
            i <- i+1
            x[i] <- xstar
            y[i] <- ystar
      }
}
```

Note: the key if statement is a single line of code: `u<=4*xstar*ystar*(xstar+ystar<=1)`.

b. The average number of iterations required to generate a single accepted value under the accept-reject scheme is $c = 6$, so the average number of iterations to generate 10,000 accepted values is 60,000. (The scheme is somewhat inefficient here because half of the (x^*, y^*) pairs are rejected immediately by violating the constraint $x^* + y^* \le 1$.)

c. Add the line of code `w=3.5+2.5*x+6.5*y`. One execution in Matlab returned `mean(w)=7.0873` and `std(w)=1.0180`. The sample mean is very close to \$7.10.

d. Using the same vector w as in **c**, $\hat{P}(W > 8)$ = `mean(w>8)` = .2080 for the aforementioned run in Matlab.

147.

a. The marginal pdf of X is $f_X(x) = \int f(x,y)dy = \int_0^{1-x} 24xy\,dy = 12x(1-x)^2$ for $0 \le x \le 1$. The conditional pdf of Y, given $X = x$, is $f(y\,|\,x) = \dfrac{f(x,y)}{f_X(x)} = \dfrac{24xy}{12x(1-x)^2} = \dfrac{2y}{(1-x)^2}$ for $0 \le y \le 1 - x$.

b. It isn't feasible to solve $F_X(x) = u$ algebraically, so we'll have to rely on accept-reject to generate x-values. Use $X^* \sim \text{Unif}[0, 1]$ as a candidate, so $g(x) = 1$; the majorization constant c must satisfy $c \ge f(x)/g(x) = 12x(1 - x)^2$ for $0 \le x \le 1$. This polynomial has its maximum at $x = 1/3$; substituting back

into the polynomial gives a majorization constant of 16/9. The key requirement $u \cdot c \cdot g(x^*) \le f(x^*)$ is equivalent to $u \le 27x^*(1 - x^*)^2/4$.
On the other hand, given $X = x$, simulation of Y is a simple matter: $F(y|x) = y^2/(1 - x)^2$; if we generate a Unif[0, 1] variate v and set $v = F(y|x)$, the solution is $y = (1-x)\sqrt{v}$. So, Y will be simulated using the inverse cdf method.

```
x=zeros(10000,1); y=x;
i=0;
while i<10000
    xs=rand;
    u=rand;
    if u<=27*xs*(1-xs)^2/4
        i=i+1;
        x(i)=xs;
        v=rand;
        y(i)=(1-xs)*sqrt(v);
    end
end
```

```
x <- NULL; y <- NULL
i <- 0
while (i <10000){
      xs <- runif(1)
      u <- runif(1)
      if (u<=27*xs*(1-xs)^2/4){
            i <- i+1
            x[i] <- xs
            v <- runif(1)
            y[i] <- (1-xs)*sqrt(v)
      }
}
```

c. The advantage of this method is that, on the average, we only require 16/9 candidates per accepted value, rather than 6 candidates per accepted value as in the earlier exercise. That more than triples the speed of the program!

149.

a. Adding across the rows of the joint pmf table, $p_X(100) = .5$ and $p_X(250) = .5$.

b. The conditional pmf of Y given $X = 100$ is obtained by dividing the joint probabilities $p(100, y)$ by the marginal probability $p_X(100) = .5$; the same goes for $Y|X = 250$. The resulting conditional pmfs are displayed below.

y	0	100	200
$p(y\|100)$	.4	.2	.4

y	0	100	200
$p(y\|250)$	.1	.3	.6

c.

```
x=zeros(10000,1); y=zeros(10000,1);
ys=[0 100 200]; py100=[.4 .2 .4];
py250=[.1 .3 .6];
for i=1:10000
    x(i)=randsample([100 250],1);
    if x(i)==100
        y(i)=randsample(ys,1,
                          true,py100);
    else
        y(i)=randsample(ys,1,
                          true,py250);
    end
end
```

```
x <- NULL; y <- NULL
ys<-c(0,100,200); py100<-
c(.4,.2,.4); py250<-c(.1,.3,.6);
for (i in 1:10000){
    x[i]<-sample(c(100,250),1)
    if(x[i]==100){
        y[i]<-
           sample(ys,1,TRUE,py100)}
    else{
        y[i]<-
           sample(ys,1,TRUE,py250)
    }
}
```

d. In Matlab, `crosstab(x,y)` will return a table (similar to the joint pmf) with the observed counts of each (x, y) pair in the simulation. For one such simulation, the result was

```
1985     1006     2022
 501     1416     3070
```

Dividing each value by n = 10,000 gives the estimated probability of that pair, e.g. $\hat{P}(X = 100, Y = 0)$ $= \frac{1985}{10{,}000} = .1985$. These six estimated probabilities are quite close to the model probabilities. The analogous command in R is `table(x,y)`.

151.

a. From Section 4.7, the marginal distribution of X is $N(\mu_1, \sigma_1)$, while the conditional distribution of Y given $X = x$ is $N(\mu_2 + \rho\sigma_2/\sigma_1(x - \mu_1),\ \sigma_2\sqrt{1-\rho^2}\,)$.

b. The functions below take 6 inputs: the five parameters and n = the desired number of simulated pairs. Note: In Matlab, it's necessary to call `[X,Y]=bivnorm(`… in order to get both vectors; just calling `bivnorm(`… will only return the first output, `X`. In R, the function will return an n-by-2 matrix; from which the individual columns `X` and `Y` could be later extracted (the Matlab program could re-written to operate the same way).

```
function [X,Y]=bivnorm(mu1,sigma1,
                   mu2,sigma2,rho,n)
X=zeros(n,1); Y=zeros(n,1);
for i=1:n
    X(i)=random('norm',mu1,sigma1);
    m=mu2+rho*sigma2/sigma1*(X(i)-mu1);
    s=sigma2*sqrt(1-rho^2);
    Y(i)=random('norm',m,s);
end
```

```
bivnorm<-function(mu1,sigma1,
              mu2,sigma2,rho,n){
X<-NULL; Y<-NULL
for (i in 1:n){
    X[i]<-rnorm(1,mu1,sigma1)
    m<-mu2+rho*sigma2/sigma1*(X[i]-mu1)
    s<-sigma2*sqrt(1-rho^2)
    Y[i]<-rnorm(1,m,s)
}
return(cbind(X,Y))
}
```

c. One run of the R program in **b** returned Y-values with sample mean 170.266 and sample standard deviation 20.29687, both reasonably close to the true mean and sd of Y.

153.

a. Let T_i denote the lifetime of the ith component (i = 1, 2, 3, 4, 5, 6). Since the 1-2 subsystem is connected in parallel, its lifetime is given by $T_{1\text{-}2} = \max(T_1, T_2)$. Likewise, $T_{3\text{-}4} = \max(T_3, T_4)$ and $T_{5\text{-}6} = \max(T_5, T_6)$. Finally, since the three subsystems are connected in series, the overall system lifetime is given by $T_{sys} = \min(T_{1\text{-}2}, T_{3\text{-}4}, T_{5\text{-}6})$. These expressions are implemented in the programs below.

```
T1=random('exp',250,[10000,1]);
T2=random('gamma',2,125,[10000,1]);
T3=random('exp',250,[10000,1]);
T4=random('gamma',2,125,[10000,1]);
T5=random('exp',300,[10000,1]);
T6=random('gamma',2,125,[10000,1]);
T12=max(T1,T2); T34=max(T3,T4);
T56=max(T5,T6);
Tsys=min([T12,T34,T56],[],2);
```

```
T1<-rexp(10000,1/250)
T2<-rgamma(10000,2,1/125)
T3<-rexp(10000,1/250)
T4<-rgamma(10000,2,1/125)
T5<-rexp(10000,1/300)
T6<-rgamma(10000,2,1/125)
T12=pmax(T1,T2); T34=pmax(T3,T4)
T56=pmax(T5,T6)
Tsys=pmin(T12,T34,T56)
```

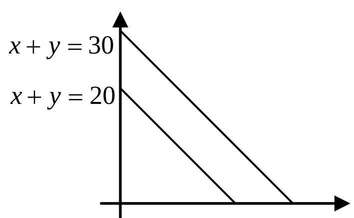
b. One execution of the Matlab code returned a vector `Tsys` with sample mean and sd 196.6193 and 104.5028, respectively. Hence, $\hat{\mu}$ = 196.6193 hours, with an estimated standard error of $\frac{s}{\sqrt{n}} = \frac{104.5028}{\sqrt{10,000}} \approx 1.045$ hours.

c. $p = P(T_{sys} < 400)$; $\hat{p} = \hat{P}(T_{sys} < 400)$ = `mean(Tsys<400)` = .9554. The estimated standard error is $\sqrt{\frac{.9554(1-.9554)}{10,000}}$ = .0021.

Supplementary Exercises

155. Let X and Y be the transmission times, so the joint pdf of X and Y is $f(x, y) = f_X(x) \cdot f_Y(y) = e^{-x}e^{-y} = e^{-(x+y)}$ for $x, y > 0$. Define $T = 2X + Y$ = the total cost to send the two messages. The cdf of T is given by $F_T(t) = P(T \le t) = P(2X + Y \le t) = P(Y \le t - 2X)$. For $X > t/2$, this probability is zero (since Y can't be negative). Otherwise, for $X \le t/2$,
$P(Y \le t - 2X) = \int_0^{t/2}\int_0^{t-2x} e^{-(x+y)}\,dydx = \cdots = 1 - 2e^{-t/2} + e^{-t}$ for $t > 0$. Thus, the pdf of T is
$f_T(t) = F_T'(t) = e^{-t/2} - e^{-t}$ for $t > 0$.

157.



a. $1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = \int_0^{20}\int_{20-x}^{30-x} kxydydx + \int_{20}^{30}\int_0^{30-x} kxydydx = \frac{81,250}{3} \cdot k \Rightarrow k = \frac{3}{81,250}$.

b. $$f_X(x) = \begin{cases} \int_{20-x}^{30-x} kxydy = k(250x - 10x^2) & 0 \le x \le 20 \\ \int_0^{30-x} kxydy = k(450x - 30x^2 + \frac{1}{2}x^3) & 20 \le x \le 30 \end{cases}$$
and, by symmetry, $f_Y(y)$ is obtained by substituting y for x in $f_X(x)$. Since $f_X(25) > 0$ and $f_Y(25) > 0$ but $f(25, 25) = 0$, $f_X(x) \cdot f_Y(y) \ne f(x, y)$ for all x, y and so X and Y are not independent.

c. $P(X + Y \le 25) = \int_0^{20}\int_{20-x}^{25-x} kxydydx + \int_{20}^{25}\int_0^{25-x} kxydydx = \frac{3}{81,250} \cdot \frac{230,625}{24} = .355$

d. $E(X + Y) = E(X) + E(Y) = 2E(X) = 2\int_0^{20} x \cdot k\left(250x - 10x^2\right)dx$

$+2\int_{20}^{30} x \cdot k\left(450x - 30x^2 + \frac{1}{2}x^3\right)dx = 2k(351,666.67) = 25.969$ lb.

e. $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot f(x,y)dxdy = \int_{0}^{20}\int_{20-x}^{30-x} kx^2y^2dydx$
$+\int_{20}^{30}\int_{0}^{30-x} kx^2y^2dydx = \frac{k}{3}\cdot\frac{33,250,000}{3} = 136.4103$, so
Cov(X, Y) = 136.4103 – $(12.9845)^2$ = –32.19. Also, $E(X^2) = E(Y^2)$ = 204.6154, so
$\sigma_x^2 = \sigma_y^2 = 204.6154 - (12.9845)^2 = 36.0182$ and $\rho = \frac{-32.19}{36.0182} = -.894$.

f. Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y) = 7.66.

159. $E(X+Y-t)^2 = \int_0^1\int_0^1 (x+y-t)^2 \cdot f(x,y)dxdy$. To find the minimizing value of t, take the derivative with respect to t and equate it to 0:
$0 = \int_0^1\int_0^1 2(x+y-t)(-1)f(x,y) = 0 \Rightarrow \int_0^1\int_0^1 tf(x,y)dxdy = t = \int_0^1\int_0^1 (x+y)\cdot f(x,y)dxdy = E(X+Y)$, so the best prediction is the individual's expected score (= 1.167).

161.

a. First, $E(W_2) = E(\rho Z_1 + \sqrt{1-\rho^2}Z_2) = \rho E(Z_1) + \sqrt{1-\rho^2}E(Z_2) = \rho(0) + \sqrt{1-\rho^2}(0) = 0.$ Second, since Z_1 and Z_2 are independent,

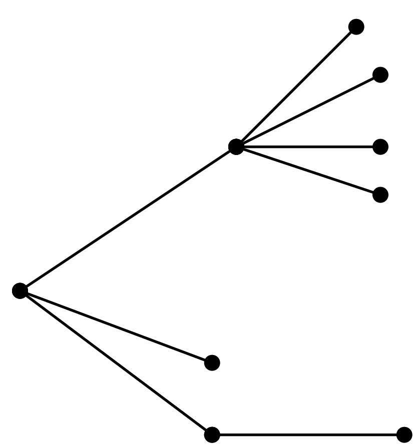
$$\text{Var}(W_2) = \text{Var}(\rho Z_1 + \sqrt{1-\rho^2}Z_2) = \rho^2\text{Var}(Z_1) + \left[\sqrt{1-\rho^2}\right]^2\text{Var}(Z_2) = \rho^2(1) + (1-\rho^2)(1) = 1.$$

b.
$$\begin{aligned}\text{Cov}(W_1,W_2) &= \text{Cov}(Z_1, \rho Z_1 + \sqrt{1-\rho^2}Z_2) = \rho\text{Cov}(Z_1,Z_1) + \sqrt{1-\rho^2}\text{Cov}(Z_1,Z_2)\\ &= \rho\text{Var}(Z_1) + \sqrt{1-\rho^2}\text{Cov}(Z_1,Z_2) = \rho(1) + \sqrt{1-\rho^2}(0) = \rho.\end{aligned}$$

c. $\text{Corr}(W_1,W_2) = \frac{\text{Cov}(W_1,W_2)}{\text{SD}(W_1)\text{SD}(W_2)} = \frac{\rho}{(1)(1)} = \rho.$

163.

a.



b. By the Law of Total Probability, $A = \bigcup_{x=0}^{\infty} A \cap \{X_1 = x\} \Rightarrow P(A) = \sum_{x=0}^{\infty} P(A \cap \{X_1 = x\})$ = $\sum_{x=0}^{\infty} P(A \mid X_1 = x) P(X_1 = x) = \sum_{x=0}^{\infty} P(A \mid X_1 = x) p(x)$. With x members in generation 1, the process becomes extinct iff these x new, independent branching processes becomes extinct. By definition, the extinction probability for each new branch is $P(A) = p^*$, and independence implies $P(A \mid X_1 = x)$ = $(p^*)^x$. Therefore, $p^* = \sum_{x=0}^{\infty} (p^*)^x p(x)$.

c. Check $p^* = 1$: $\sum_{x=0}^{\infty} (1)^x p(x) = \sum_{x=0}^{\infty} p(x) = 1 = p^*$. [We'll drop the * notation from here forward.]
In the first case, we get $p = .3 + .5p + .2p^2$. Solving for p gives $p = 3/2$ and $p = 1$; the smaller value, $p = 1$, is the extinction probability. Why will this model die off with probability 1? Because the expected number of progeny from a single individual is $0(.3)+1(.5)+2(.2) = .9 < 1$.
On the other hand, the second case gives $p = .2 + .5p + .3p^2$, whose solutions are $p = 1$ and $p = 2/3$. The extinction probability is the smaller value, $p = 2/3$. Why does this model have positive probability of eternal survival? Because the expected number of progeny from a single individual is $0(.2)+1(.5)+2(.3) = 1.1 > 1$.

165.

a. Use a sort of inclusion-exclusion principle:
$P(a \le X \le b,\ c \le Y \le d) = P(X \le b,\ Y \le d) - P(X < a,\ Y \le d) - P(X \le b,\ Y < c) + P(X < a,\ Y < c)$.
Then, since these variables are continuous, we may write
$P(a \le X \le b,\ c \le Y \le d) = F(b, d) - F(a, d) - F(b, c) + F(a, c)$.

b. In the discrete case, the strict inequalities in (a) must be re-written as follows:
$P(a \le X \le b,\ c \le Y \le d) = P(X \le b,\ Y \le d) - P(X \le a - 1,\ Y \le d) - P(X \le b,\ Y \le c - 1) +$
$P(X \le a - 1,\ Y \le c - 1) = F(b, d) - F(a - 1, d) - F(b, c - 1) + F(a - 1, c - 1)$. For the values specified, this becomes $F(10,6) - F(4,6) - F(10,1) + F(4,1)$.

c. Use the cumulative joint cdf table below. At each (x^*, y^*), $F(x^*, y^*)$ is the sum of the probabilities at points (x, y) such that $x \le x^*$ and $y \le y^*$.

	$F(x, y)$	x = 100	250
	200	.50	1
y	100	.30	.50
	0	.20	.25

d. Integrating long-hand and exhausting all possible options for (x,y) pairs, we arrive at the following:
$F(x, y) = .6x^2y + .4xy^3$, $0 \le x, y \le 1$; $F(x, y) = 0$, $x \le 0$ or $y \le 0$; $F(x, y) = .6x^2 + .4x$, $0 \le x \le 1$, $y > 1$;
$F(x, y) = .6y + .4y^3$, $x > 1$, $0 \le y \le 1$; and, obviously, $F(x, y) = 1$, $x > 1$, $y > 1$. (Whew!)
Thus, from (a), $P(.25 \le x \le .75, .25 \le y \le .75) = F(.75,.75) - F(.25,.75) - F(.75,.25) + F(.25,.25) = \ldots$ $=.23125$. [This only requires the main form of $F(x, y)$; i.e., that for $0 \le x, y \le 1$.]

e. Again, we proceed on a case-by case basis. The results are:
$F(x, y) = 6x^2y^2$, $x + y \le 1$, $0 \le x \le 1$; $0 \le y \le 1$;
$F(x, y) = 3x^4 - 8x^3 + 6x^2 + 3y^4 - 8y^3 + 6y^2 - 1$, $x + y > 1$, $x \le 1$, $y \le 1$;
$F(x, y) = 0$, $x \le 0$; $F(x, y) = 0$, $y \le 0$;
$F(x, y) = 3x^4 - 8x^3 + 6x^2$, $0 \le x \le 1$, $y > 1$;
$F(x, y) = 3y^4 - 8y^3 + 6y^2$, $0 \le y \le 1$, $x > 1$; and, obviously,
$F(x, y) = 1$, $x > 1$, $y > 1$.

167.

a. For an individual customer, the expected number of packages is 1(.4)+2(.3)+3(.2)+4(.1) = 2 with a variance of 1 (by direct computation). Given $X=x$, Y is the sum of x independent such customers, so $E(Y|X=x) = x(2) = 2x$ and $\text{Var}(Y|X=x) = x(1) = x$.

b. By the law of total expectation, $E(Y) = E[E(Y|X)] = E(2X) = 2E(X) = 2(20) = 40$.

c. By the law of total variance, $\text{Var}(Y) = \text{Var}(E(Y|X)) + E(\text{Var}(Y|X)) = \text{Var}(2X) + E(X) = 4\text{Var}(X) + E(X)$ $= 4(20) + 20 = 100$. (Recall that the mean and variance of a Poisson rv are equal.)

169. Let F_1 and F_2 denote the cdfs of X_1 and X_2, respectively. The probability of a pavement failure within t years is equal to $P(Y \le t) = P(\min(X_1, X_2) \le t) = 1 - P(\min(X_1, X_2) > t)$. The minimum of two values exceeds t iff both of the values exceed t, so we continue:
$P(Y \le t) = 1 - P(X_1 > t \cap X_2 > t) = 1 - P(X_1 > t) \cdot P(X_2 > t) = 1 - [1 - P(X_1 \le t)][1 - P(X_2 > t)] = 1 - [1 - F_1(t)][1 - F_2(t)]$. Using the cdfs provided,

$$P(Y \le t) = 1 - \left[1 - \Phi\left(\frac{-25.49 + 1.15t}{\sqrt{4.45 - 1.78t + .171t^2}}\right)\right]\left[1 - \Phi\left(\frac{-21.27 + .0325t}{\sqrt{.972 - .00028t + .00022t^2}}\right)\right]$$

Evaluate this at $t = 5$, and the argument inside the first Φ turns out to be imaginary! Interestingly, this was not noticed by the authors of the original article.

Evaluate this at $t = 10$ to get $P(Y \le 10) = 1 - [1 - \Phi(-7.22)][1 - \Phi(-21.14)] \approx 1 - [1][1] = 0$, a nonsensical answer.

171. Let $a = 1/1000$ for notational ease. W is the maximum of the two exponential rvs, so its pdf is $f_W(w) = 2F_X(w)f_X(w) = 2(1 - e^{-aw})ae^{-aw} = 2ae^{-aw}(1 - e^{-aw})$. From this, $M_W(t) = E[e^{tW}] = \int_0^\infty e^{tw}2ae^{-aw}(1 - e^{-aw})dw$ =
$2a\int_0^\infty e^{-(a-t)w}dw - 2a\int_0^\infty e^{-(2a-t)w}dw = \dfrac{2a}{a-t} - \dfrac{2a}{2a-t} = \dfrac{2a^2}{(a-t)(2a-t)} = \dfrac{2}{(1-1000t)(2-1000t)}$. From this,
$E[W] = M'_W(0) = 1500$ hours.

173. The roll-up procedure is not valid for the 75th percentile unless $\sigma_1 = 0$ and/or $\sigma_2 = 0$, as described below.
Sum of percentiles: $(\mu_1 + z\sigma_1) + (\mu_2 + z\sigma_2) = \mu_1 + \mu_2 + z(\sigma_1 + \sigma_2)$
Percentile of sums: $(\mu_1 + \mu_2) + z\sqrt{\sigma_1^2 + \sigma_2^2}$
These are equal when $z = 0$ (i.e. for the median) or in the unusual case when $\sigma_1 + \sigma_2 = \sqrt{\sigma_1^2 + \sigma_2^2}$, which happens when $\sigma_1 = 0$ and/or $\sigma_2 = 0$.

175.

a. Let $X_1, \ldots, X_{12}$ denote the weights for the business-class passengers and $Y_1, \ldots, Y_{50}$ denote the tourist-class weights. Then T = total weight = $X_1 + \ldots + X_{12} + Y_1 + \ldots + Y_{50} = X + Y$.
$E(X) = 12E(X_1) = 12(30) = 360$; $\text{Var}(X) = 12\ \text{Var}(X_1) = 12(36) = 432$.
$E(Y) = 50E(Y_1) = 50(40) = 2000$; $\text{Var}(Y) = 50\ \text{Var}(Y_1) = 50(100) = 5000$.
Thus $E(T) = E(X) + E(Y) = 360 + 2000 = 2360$,
and $\text{Var}(T) = \text{Var}(X) + \text{Var}(Y) = 432 + 5000 = 5432 \Rightarrow \text{SD}(T) = 73.7021$.

b. $P(T \le 2500) = \Phi\left(\dfrac{2500-2360}{73.7021}\right) = \Phi(1.90) = .9713$.

177. $X \sim \text{Bin}(200, .45)$ and $Y \sim \text{Bin}(300, .6)$. Because both ns are large, both X and Y are approximately normal, so $X + Y$ is approximately normal with mean $(200)(.45) + (300)(.6) = 270$, variance $200(.45)(.55) + 300(.6)(.4) = 121.40$, and standard deviation 11.02. Thus,
$P(X + Y \ge 250) = 1-\Phi\left(\dfrac{249.5-270}{11.02}\right) = 1-\Phi(-1.86) = .9686$.

179. Ann has $2(6)(16) = 192$ oz. The amount which she would consume if there were no limit is $X_1 + \ldots + X_{14}$ where each X_i is normally distributed with $\mu = 13$ and $\sigma = 2$. Thus $T = X_1 + \ldots + X_{14}$ is normal, with mean $14(13) = 182$ oz. and sd $\sqrt{14}\,(2) = 7.483$ oz., so $P(T < 192) = \Phi(1.34) = .9099$. Independence might not be a reasonable assumption here, since Ann can see how much soda she has left and might decide to drink more/less later in the two-week period based on that.

181. The student will not be late if $X_1 + X_3 \le X_2$, i.e. if $X_1 - X_2 + X_3 \le 0$. This linear combination has mean -2 and variance 4.25, so $P(X_1 - X_2 + X_3 \le 0) = \Phi\left(\dfrac{0-(-2)}{\sqrt{4.25}}\right) = \Phi(.97) = .8340$.

183.

a. $\text{Var}(X_1) = \text{Var}(W + E_1) = \sigma_W^2 + \sigma_E^2 = \text{Var}(W + E_2) = \text{Var}(X_2)$ and
$\text{Cov}(X_1, X_2) = \text{Cov}(W + E_1, W + E_2) = \text{Cov}(W, W) + \text{Cov}(W, E_2) + \text{Cov}(E_1, W) + \text{Cov}(E_1, E_2) =$
$\text{Var}(W) + 0 + 0 + 0 = \sigma_W^2$.

Thus, $\rho = \dfrac{\text{Cov}(X_1, X_2)}{\text{SD}(X_1)\text{SD}(X_2)} = \dfrac{\sigma_W^2}{\sqrt{\sigma_W^2+\sigma_E^2}\cdot\sqrt{\sigma_W^2+\sigma_E^2}} = \dfrac{\sigma_W^2}{\sigma_W^2+\sigma_E^2}$.

b. $\rho = \dfrac{1^2}{1^2 + .01^2} = .9999$.

185. $E(Y) \doteq h(\mu_1, \mu_2, \mu_3, \mu_4) = 120\left[\frac{1}{10} + \frac{1}{15} + \frac{1}{20}\right] = 26$.

The partial derivatives of $h(\mu_1, \mu_2, \mu_3, \mu_4)$ with respect to x_1, x_2, x_3, and x_4 are $-\dfrac{x_4}{x_1^2}$, $-\dfrac{x_4}{x_2^2}$, $-\dfrac{x_4}{x_3^2}$, and $\dfrac{1}{x_1} + \dfrac{1}{x_2} + \dfrac{1}{x_3}$, respectively. Substituting $x_1 = 10$, $x_2 = 15$, $x_3 = 20$, and $x_4 = 120$ gives -1.2, $-.5333$, $-.3000$, and $.2167$, respectively, so $\text{Var}(Y) = (1)(-1.2)^2 + (1)(-.5333)^2 + (1.5)(-.3000)^2 + (4.0)(.2167)^2 = 2.6783$, and the approximate sd of Y is 1.64.

187.

a. Use the joint pdf from Exercise 141 with i = 1 and j = n:

$$g(y_1, y_n) = \frac{n!}{(1-1)!(n-1-1)!(n-n)!} F(y_1)^{1-1}[F(y_n) - F(y_1)]^{n-1-1}[1-F(y_n)]^{n-n} f(y_1) f(y_n)$$
$$= n(n-1)[F(y_n) - F(y_1)]^{n-2} f(y_1) f(y_n) \quad \text{for } y_1 < y_n$$

b. This transformation can be re-written as $Y_1 = W_1$ and $Y_n = W_1 + W_2$, from which the Jacobian is $\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix}$ = 1. The constraint $y_1 < y_n$ is equivalent to $w_2 > 0$ (there's no constraint on w_1 beyond the bounds of the underlying distribution). Hence, the joint pdf of W_1 and W_2 is $f(w_1, w_2) = n(n-1)[F(w_1 + w_2) - F(w_1)]^{n-2} f(w_1) f(w_1 + w_2)$, and the marginal pdf of the sample range, W_2, is

$$f_{W_2}(w_2) = n(n-1)\int_{-\infty}^{\infty} [F(w_1 + w_2) - F(w_1)]^{n-2} f(w_1) f(w_1 + w_2)\, dw_1 \text{ for } w_2 > 0.$$

c. For the uniform distribution on [0, 1], $f(x) = 1$ and $F(x) = x$ for $0 \le x \le 1$. Hence, the integrand above simplifies to $n(n-1)[w_1 + w_2 - w_1]^{n-2}(1)(1) = n(n-1)w_2^{n-2}$. Some care must be taken with the limits of integration: the pdfs now require $0 \le w_1 \le 1$ and $0 \le w_1 + w_2 \le 1$, from which $0 \le w_1 \le 1 - w_2$. Hence, the marginal pdf of the sample range W_2 is

$f_{W_2}(w_2) = n(n-1)\int_0^{1-w_2} w_2^{n-2}\, dw_1 = n(n-1)w_2^{n-2}(1-w_2)$ for $0 \le w_2 \le 1$. Incidentally, this is the Beta distribution, with $\alpha = n - 1$ and $\beta = 2$.

189.

a. The marginal pdf of X is

$$f_X(x) = \int_0^{\infty} \frac{1}{2\pi} \frac{e^{-[(\ln x)^2 + (\ln y)^2]/2}}{xy}[1 + \sin(2\pi \ln x)\sin(2\pi \ln y)]dy$$
$$= \frac{1}{2\pi}\frac{e^{-[(\ln x)^2]/2}}{x}\int_0^{\infty} e^{-[(\ln y)^2]/2}[1 + \sin(2\pi \ln x)\sin(2\pi \ln y)]\frac{dy}{y}$$
$$= \frac{1}{2\pi}\frac{e^{-[(\ln x)^2]/2}}{x}\int_{-\infty}^{\infty} e^{-u^2/2}[1 + \sin(2\pi \ln x)\sin(2\pi u)]du$$
$$= \frac{1}{2\pi}\frac{e^{-[(\ln x)^2]/2}}{x}\left[\int_{-\infty}^{\infty} e^{-u^2/2}du + \sin(2\pi \ln x)\int_{-\infty}^{\infty} e^{-u^2/2}\sin(2\pi u)du\right]$$

The first integral is $\sqrt{2\pi}$, since the integrand is the N(0, 1) pdf without the constant. The second integral is 0, since the integral is an odd function over (–∞, ∞). Hence, the final answer is

$f_X(x) = \frac{1}{2\pi}\frac{e^{-[(\ln x)^2]/2}}{x} \cdot \frac{1}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}x}e^{-[(\ln x)^2]/2}$, the lognormal pdf with $\mu = 0$ and $\sigma = 1$.

By symmetry, this is also the marginal pdf of Y.

b. The conditional distribution of Y given $X = x$ is

$$f(y \mid x) = \frac{f(x,y)}{f_X(x)} = \frac{e^{-[(\ln y)^2]/2}}{\sqrt{2\pi}y}[1+\sin(2\pi\ln x)\sin(2\pi\ln y)]$$, from which

$$E(Y^n \mid X = x) = \int_0^\infty y^n \cdot \frac{1}{\sqrt{2\pi}} \frac{e^{-[(\ln y)^2]/2}}{y}[1+\sin(2\pi\ln x)\sin(2\pi\ln y)]dy$$

$$= \int_0^\infty y^n \cdot \frac{e^{-[(\ln y)^2]/2}}{\sqrt{2\pi}y}dy + \frac{\sin(2\pi\ln x)}{\sqrt{2\pi}}\int_0^\infty y^n \cdot e^{-[(\ln y)^2]/2}\sin(2\pi\ln y)\frac{dy}{y}$$

The first integral is $\int_0^\infty y^n \cdot \frac{e^{-[(\ln y)^2]/2}}{\sqrt{2\pi}y}dy = \int_0^\infty y^n \cdot f_Y(y)dy = E(Y^n)$. So, the goal is now to show that the second integral equals zero. For the second integral, make the suggested substitution $\ln(y) = u + n$, for which $du = dy/y$ and $y = e^{u+n}$:

$$\int_0^\infty y^n \cdot e^{-[(\ln y)^2]/2}\sin(2\pi\ln y)\frac{dy}{y} = \int_{-\infty}^\infty (e^{u+n})^n \cdot e^{-(u+n)^2/2}\sin(2\pi(u+n))du = \int_{-\infty}^\infty e^{-u^2/2+n^2/2}\sin(2\pi u+2\pi n)du$$

$= e^{n^2/2}\int_{-\infty}^\infty e^{-u^2/2}\sin(2\pi u+2\pi n)du = e^{n^2/2}\int_{-\infty}^\infty e^{-u^2/2}\sin(2\pi u)du$. The second equality comes from expanding the exponents on e; the last equality comes from the basic fact that $\sin(\theta + 2\pi n) = \sin(\theta)$ for any integer n. The integral that remains has an odd integrand (u^2 is even and sine is odd), so the integral on $(-\infty, \infty)$ equals zero. At last, we have that $E(Y^n \mid X = x) = E(Y^n)$ for any positive integer n.

c. Since the pdf is symmetric in X and Y, the same derivation will yield $E(X^n \mid Y = y) = E(X^n)$ for all positive integers n.

d. Despite the fact that the expectation of every polynomial in Y is unaffected by conditioning on X (and vice versa), the two rvs are <u>not</u> independent. From **a**, the marginal pdfs of X and Y are lognormal, from which $f_X(x)\cdot f_Y(y) = \frac{1}{\sqrt{2\pi}x}e^{-[(\ln x)^2]/2}\cdot\frac{1}{\sqrt{2\pi}y}e^{-[(\ln y)^2]/2} = \frac{1}{2\pi xy}e^{-[(\ln x)^2+(\ln y)^2]/2} \neq f(x, y)$. Therefore, by definition X and Y are not independent.

191.

a. Define success = obtaining one of the 9 "other" toys and failure = obtaining the same toy that we got in the first box. Starting with the second box, we have independent trials with success/failure defined as above, a constant probability of success (equal to 9/10), and we're searching until the first success occurs. Thus, Y_2 follows a geometric distribution with $p = 9/10$. In particular, $E(Y_2) = 1/p = 10/9$.

b. By the same reasoning, Y_3 is also a geometric rv, but now success corresponds to getting any of the <u>eight</u> toys not yet obtained. So $p = 8/10$ for Y_3, and $E(Y_3) = 1/p = 10/8$.

c. Following the pattern, each of Y_4 through Y_{10} is also geometric, but with success probability equal to (number of remaining toys)/10. The grand total number of boxes required, including our first box (which certainly yields a toy we don't have yet), is $1 + Y_2 + \ldots + Y_{10}$. Thus, the expected number of boxes required to obtain all 10 toys is

$E(1 + Y_2 + \ldots + Y_{10}) = 1 + E(Y_2) + \ldots + E(Y_{10}) = 1+\frac{10}{9}+\frac{10}{8}+\cdots+\frac{10}{1} = \frac{7381}{252} \approx 29.29$ boxes.

Notice a couple features of the solution: (1) the expected number of boxes to get the *next* new toy grows as we proceed (from 1 to 10/9 to 10/8 … to 10); (2) in general, the expected number of boxes required to get t toys is $1 + t/(t-1) + \ldots + t = t \cdot \sum_{k=1}^{t} \frac{1}{k}$.

d. Recall that the variance of a geometric rv with parameter p is $\frac{1-p}{p^2}$. Thus, the variance of the number of boxes required to obtain all 10 toys, by independence, is $\text{Var}(1 + Y_2 + \ldots + Y_{10}) = \text{Var}(Y_2) + \ldots + \text{Var}(Y_{10}) = \frac{1/10}{(9/10)^2} + \frac{2/10}{(8/10)^2} + \cdots + \frac{9/10}{(1/10)^2} \approx 125.687$, from which the sd is about 11.2 boxes.

CHAPTER 5

Section 5.1

1.

a. We use the sample mean, $\bar{x}$, to estimate the population mean μ: $\hat{\mu} = \bar{x} = \frac{\Sigma x_i}{n} = \frac{3753}{33}$ = 113.73.

b. The quantity described is the median, η. which we estimate with the sample median: $\tilde{x}$ = the middle observation when arranged in ascending order = the 17th ordered observation = 113.

c. To estimate σ, we use the sample standard deviation, $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{33}(x_i - 113.73)^2}{33-1}} =$

$= \sqrt{162.39} = 12.74$. First-graders' IQ scores typically differ from the mean IQ of 113.73 by about ±12.74 points.

d. All but three of the 33 first graders have IQs above 100. With "success" = IQ greater than 100 and x = # of successes = 33, $\hat{p} = \frac{x}{n} = \frac{30}{33}$ = .9091.

e. A sensible estimator of $100\sigma/\mu$ is $100\hat{\sigma}/\hat{\mu} = 100s/\bar{x}$ = 100(12.74)/113.73 = 11.2.

3. You can calculate for this data set that $\bar{x} = 1.3481$ and s = .3385.

a. We use the sample mean, $\bar{x} = 1.3481$.

b. Because we assume normality, the mean = median, so we also use the sample mean $\bar{x} = 1.3481$. We could also easily use the sample median.

c. For a normal distribution, the 90th percentile is equal to $\mu + 1.28\sigma$. An estimate of that population 90th percentile is $\hat{\mu} + (1.28)\hat{\sigma} = \bar{x} + 1.28s = 1.3481 + (1.28)(.3385) = 1.7814$.

d. Since we can assume normality, $P(X < 1.5) = \Phi\left(\frac{1.5-\mu}{\sigma}\right) \approx \Phi\left(\frac{1.5-\bar{x}}{s}\right) = \Phi\left(\frac{1.5-1.3481}{.3385}\right)$

$= \Phi(.45) = .6736$.

e. The estimated standard error of $\bar{x}$ is $\frac{s}{\sqrt{n}} = \frac{.3385}{\sqrt{16}} = .0846$.

5. Let θ = the total audited value. Three potential estimators of θ are $\hat{\theta}_1 = N\overline{X}$, $\hat{\theta}_2 = T - N\overline{D}$, and $\hat{\theta}_3 = T \cdot \dfrac{\overline{X}}{\overline{Y}}$.

From the data, $\bar{y}$ = 374.6, $\bar{x}$ = 340.6, and $\bar{d}$ = 34.0. Knowing N = 5,000 and T = 1,761,300, the three corresponding estimates are $\hat{\theta}_1 = (5{,}000)(340.6) = 1{,}703{,}000$, $\hat{\theta}_2 = 1{,}761{,}300 - (5{,}000)(34.0) = 1{,}591{,}300$, and $\hat{\theta}_3 = 1{,}761{,}300\left(\dfrac{340.6}{374.6}\right) = 1{,}601{,}438.281$.

7.

a. $\hat{\mu} = \bar{x} = \dfrac{\sum x_i}{n} = \dfrac{1206}{10} = 120.6.$

b. Since τ = 10,000μ, $\hat{\tau} = 10{,}000\hat{\mu}$ = 10,000(120.6) = 1,206,000.

c. 8 of 10 houses in the sample used at least 100 therms (the "successes"), so $\hat{p} = \frac{8}{10} = .80$.

d. The ordered sample values are 89, 99, 103, 109, 118, 122, 125, 138, 147, 156, from which the two middle values are 118 and 122, so $\hat{\eta} = \tilde{x}$ = (118 + 122)/2 = 120.

9.

a. $E(\overline{X}) = \mu = E(X)$, so $\overline{X}$ is an unbiased estimator for the Poisson parameter μ. Since n = 150,

$$\hat{\mu} = \bar{x} = \frac{\Sigma x_i}{n} = \frac{(0)(18) + (1)(37) + \ldots + (7)(1)}{150} = \frac{317}{150} = 2.11.$$

b. $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{\sqrt{\mu}}{\sqrt{n}}$, so the estimated standard error is $\sqrt{\dfrac{\hat{\mu}}{n}} = \dfrac{\sqrt{2.11}}{\sqrt{150}} = .119$.

11.

a. First, the mgf of each X_i is $M_{X_i}(t) = \dfrac{\lambda}{\lambda - t}$. Then, using independence, $M_{\Sigma X_i}(t) = \left(\dfrac{\lambda}{\lambda - t}\right)^n$. Finally, using $\overline{X} = \frac{1}{n}\Sigma X_i$ and the properties of mgfs, $M_{\overline{X}}(t) = M_{\Sigma X_i}(\tfrac{1}{n}t) = \left(\dfrac{\lambda}{\lambda - \frac{1}{n}t}\right)^n = \dfrac{1}{(1 - t/n\lambda)^n}$. This is precisely the mgf of the gamma distribution with $\alpha = n$ and $\beta = 1/(n\lambda)$, so by uniqueness of mgfs $\overline{X}$ has this distribution.

b. We'll use Equation (3.5): With $Y = \overline{X}$ ~ Gamma(n, $1/n\lambda$),

$$E(\hat{\lambda}) = E\left(\frac{1}{Y}\right) = \int_0^\infty \frac{1}{y} \cdot \frac{1}{\Gamma(n)(1/n\lambda)^n} y^{n-1} e^{-y/[1/n\lambda]}\, dy = \frac{1}{\Gamma(n)(1/n\lambda)^n} \int_0^\infty y^{n-2} e^{-y/[1/n\lambda]}\, dy$$

$$= \frac{1}{\Gamma(n)(1/n\lambda)^n} \Gamma(n-1)(1/n\lambda)^{n-1} = \frac{\Gamma(n-1)(n\lambda)^n}{\Gamma(n)(n\lambda)^{n-1}} = \frac{n\lambda}{n-1}$$

In particular, since $n/(n-1) > 1$, $\hat{\lambda} = 1/\overline{X}$ is a biased-high estimator of λ.

c. Similar to **b**,

$$E(\hat{\lambda}^2) = E\left(\frac{1}{Y^2}\right) = \frac{1}{\Gamma(n)(1/n\lambda)^n}\int_0^\infty y^{n-3}e^{-y/[1/n\lambda]}\,dy = \cdots = \frac{\Gamma(n-2)(n\lambda)^n}{\Gamma(n)(n\lambda)^{n-2}} = \frac{(n\lambda)^2}{(n-1)(n-2)},$$

from which $\text{Var}(\hat{\lambda}) = E(\hat{\lambda}^2) - [E(\hat{\lambda})]^2 = \dfrac{(n\lambda)^2}{(n-1)(n-2)} - \left[\dfrac{n\lambda}{n-1}\right]^2 = \dfrac{n^2\lambda^2}{(n-1)^2(n-2)}$.

13. From the description $X_1 \sim \text{Bin}(n_1, p_1)$ and $X_2 \sim \text{Bin}(n_2, p_2)$.

a. $E\left(\dfrac{X_1}{n_1} - \dfrac{X_2}{n_2}\right) = \dfrac{1}{n_1}E(X_1) - \dfrac{1}{n_2}E(X_2) = \dfrac{1}{n_1}(n_1p_1) - \dfrac{1}{n_2}(n_2p_2) = p_1 - p_2$. Hence, by definition, $\dfrac{X_1}{n_1} - \dfrac{X_2}{n_2}$ is an unbiased estimator of $p_1 - p_2$.

b. $\text{Var}\left(\dfrac{X_1}{n_1} - \dfrac{X_2}{n_2}\right) = \text{Var}\left(\dfrac{X_1}{n_1}\right) + (-1)^2\,\text{Var}\left(\dfrac{X_2}{n_2}\right) = \left(\dfrac{1}{n_1}\right)^2 \text{Var}(X_1) + \left(\dfrac{1}{n_2}\right)^2 \text{Var}(X_2) =$

$\dfrac{1}{n_1^2}(n_1p_1q_1) + \dfrac{1}{n_2^2}(n_2p_2q_2) = \dfrac{p_1q_1}{n_1} + \dfrac{p_2q_2}{n_2}$, and the standard error is the square root of this quantity.

c. With $\hat{p}_1 = \dfrac{x_1}{n_1}$, $\hat{q}_1 = 1 - \hat{p}_1$, $\hat{p}_2 = \dfrac{x_2}{n_2}$, $\hat{q}_2 = 1 - \hat{p}_2$, the estimated standard error is $\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$.

d. $(\hat{p}_1 - \hat{p}_2) = \dfrac{127}{200} - \dfrac{176}{200} = .635 - .880 = -.245$

e. $\sqrt{\dfrac{(.635)(.365)}{200} + \dfrac{(.880)(.120)}{200}} = .041$

15. $\mu = E(X) = \int_{-1}^{1} x \cdot .5(1 + \theta x)dx = \left.\dfrac{x^2}{4} + \dfrac{\theta x^3}{6}\right|_{-1}^{1} = \dfrac{1}{3}\theta \Rightarrow \theta = 3\mu$. Hence,

$\hat{\theta} = 3\bar{X} \Rightarrow E(\hat{\theta}) = E(3\bar{X}) = 3E(\bar{X}) = 3\mu = 3\left(\dfrac{1}{3}\right)\theta = \theta.$

17.

a. $E(X^2) = 2\theta$ implies that $E\left(\dfrac{X^2}{2}\right) = \theta$. Consider $\hat{\theta} = \dfrac{\sum X_i^2}{2n}$. Then

$E(\hat{\theta}) = E\left(\dfrac{\sum X_i^2}{2n}\right) = \dfrac{\sum E(X_i^2)}{2n} = \dfrac{\sum 2\theta}{2n} = \dfrac{2n\theta}{2n} = \theta$, implying that $\hat{\theta}$ is an unbiased estimator for θ.

b. $\sum x_i^2 = 1490.1058$, so $\hat{\theta} = \dfrac{1490.1058}{20} = 74.505.$

19.

a. $E(\hat{P}) = \sum_{x=r}^{\infty} \frac{r-1}{x-1} \cdot \binom{x-1}{r-1} \cdot p^r \cdot (1-p)^{x-r} = \sum_{x=r}^{\infty} \frac{(x-2)!}{(x-r)!(r-2)!} \cdot p^r \cdot (1-p)^{x-r} = \sum_{x=r}^{\infty} \binom{x-2}{r-2} \cdot p^r \cdot (1-p)^{x-r}$.

Make the suggested substitutions $y = x - 1$ and $s = r - 1$, i.e. $x = y + 1$ and $r = s + 1$:

$E(\hat{P}) = \sum_{y=s}^{\infty} \binom{y-1}{s-1} p^{s+1} (1-p)^{y-s} = p \sum_{y=s}^{\infty} \binom{y-1}{s-1} p^s (1-p)^{y-s} = p \sum_{y=s}^{\infty} nb(y; s, p) = p \cdot 1 = p$.

The last steps use the fact that the term inside the summation is the negative binomial pmf with parameters s and p, and all pmfs sum to 1.

b. For the given sequence, $x = 5$, so $\hat{p} = \frac{5-1}{5+5-1} = \frac{4}{9} = .444$.

21.

a. $\lambda = .5p + .15 \Rightarrow 2\lambda = p + .3$, so $p = 2\lambda - .3$ and $\hat{p} = 2\hat{\lambda} - .3 = 2\left(\frac{Y}{n}\right) - .3$; the estimate is $2\left(\frac{20}{80}\right) - .3 = .2$.

b. $E(\hat{p}) = E\left(2\hat{\lambda} - .3\right) = 2E\left(\hat{\lambda}\right) - .3 = 2\lambda - .3 = p$, as desired.

c. Here $\lambda = .7p + (.3)(.3)$, so $p = \frac{10}{7}\lambda - \frac{9}{70}$ and $\hat{p} = \frac{10}{7}\left(\frac{Y}{n}\right) - \frac{9}{70}$.

23.

a. Applying rules of mean and variance, $E(\hat{P}_a) = \frac{E(X) + \sqrt{n/4}}{n + \sqrt{n}} = \frac{np + \sqrt{n/4}}{n + \sqrt{n}} = \frac{\sqrt{n}\,p + 1/2}{\sqrt{n} + 1}$ and $\text{Var}(\hat{P}_a) = \frac{\text{Var}(X)}{(n + \sqrt{n})^2} = \frac{np(1-p)}{(n + \sqrt{n})^2} = \frac{p(1-p)}{(\sqrt{n} + 1)^2}$. Substitute these into the MSE formula and simplify: $\text{MSE}(\hat{P}_a) = [E(\hat{P}_a) - p]^2 + \text{Var}(\hat{P}_a) = \cdots = \frac{1}{4(\sqrt{n} + 1)^2}$. It's interesting that the MSE of this estimator does not depend on the true value of p!

b. The MSE of the usual estimator is $\frac{p(1-p)}{n}$. When p is near .5, the MSE from **a** is smaller, especially for small n. However, when p is near 0 or 1, the usual estimator has (slightly) lower MSE.

Section 5.2

25.

a. To find the mle of p, we'll take the derivative of the log-likelihood function

$\ell(p) = \ln\left[\binom{n}{x} p^x (1-p)^{n-x}\right] = \ln\binom{n}{x} + x\ln(p) + (n-x)\ln(1-p)$, set it equal to zero, and solve for p.

$\ell'(p) = \frac{d}{dp}\left[\ln\binom{n}{x} + x\ln(p) + (n-x)\ln(1-p)\right] = \frac{x}{p} - \frac{n-x}{1-p} = 0 \Rightarrow x(1-p) = p(n-x) \Rightarrow p = x/n$, so the mle of p is $\hat{p} = \frac{x}{n}$, which is simply the sample proportion of successes. For $n = 20$ and $x = 3$, $\hat{p} = \frac{3}{20} =$ .15.

b. Since X is binomial, $E(X) = np$, from which $E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}(np) = p$; thus, $\hat{p}$ is an unbiased estimator of p.

c. By the invariance principle, the mle of $(1 - p)^5$ is just $(1 - \hat{p})^5$. For $n = 20$ and $x = 3$, we have $(1 - .15)^5 = .4437$.

27. For a single sample from a Poisson distribution, $f(x_1,\ldots,x_n;\mu) = \frac{e^{-\mu}\mu^{x_1}}{x_1!}\cdots\frac{e^{-\mu}\mu^{x_n}}{x_n!} = \frac{e^{-n\mu}\mu^{\Sigma x_i}}{x_1!\cdots x_n!}$, so $\ln\left[f(x_1,\ldots,x_n;\mu)\right] = -n\mu + \sum x_i\ln(\mu) - \sum\ln(x_i!)$. Thus $\frac{d}{d\mu}\left[\ln\left[f(x_1,\ldots,x_n;\lambda)\right]\right] = -n + \frac{\sum x_i}{\mu} = 0 \Rightarrow \hat{\mu} = \frac{\sum x_i}{n} = \bar{x}$. For our problem, $f(x_1,\ldots,x_n, y_1\ldots y_n;\mu_1,\mu_2)$ is a product of the X sample likelihood and the Y sample likelihood, implying that $\hat{\mu}_1 = \bar{x}, \hat{\mu}_2 = \bar{y}$, and (by the invariance principle) $\widehat{(\mu_1 - \mu_2)} = \bar{x} - \bar{y}$.

29. The number of helmets examined, X, until r flawed helmets are found has a negative binomial distribution: $X \sim \text{NB}(r, p)$. To find the mle of p, we'll take the derivative of the log-likelihood function

$\ell(p) = \ln\left[\binom{x-1}{r-1} p^r (1-p)^{x-r}\right] = \ln\binom{x-1}{r-1} + r\ln(p) + (x-r)\ln(1-p)$, set it equal to zero, and solve for p.

$\ell'(p) = \frac{d}{dp}\left[\ln\binom{x-1}{r-1} + r\ln(p) + (x-r)\ln(1-p)\right] = \frac{r}{p} - \frac{x-r}{1-p} = 0 \Rightarrow r(1-p) = (x-r)p \Rightarrow p = r/x$, so the mle of p is $\hat{p} = \frac{r}{x}$. This is the number of successes over the total number of trials; with $r = 3$ and $x = 20$, $\hat{p} = .15$. Yes, this is the same as the mle of p based on the binomial model in Exercise 25.

In contrast, the unbiased estimator from Exercise 19 is $\hat{p} = \frac{r-1}{x-1}$, which is <u>not</u> the same as the maximum likelihood estimator. (With $r = 3$ and $x = 20$, the calculated value of the unbiased estimator is 2/19, rather than 3/20.)

31.

a. The likelihood function is $L(\theta) = f(x_1,...,x_n;\theta) = \prod_{i=1}^{n} \frac{x_i}{\theta} e^{-x_i^2/(2\theta)} = \frac{\Pi x_i}{\theta^n} e^{-\Sigma x_i^2/(2\theta)}$, so the log-likelihood function is $\ell(\theta) = \ln[L(\theta)] = \ln[\Pi x_i] - n\ln(\theta) - \frac{\Sigma x_i^2}{2\theta}$. To find the mle of θ, differentiate and set equal to zero: $0 = \ell'(\theta) = 0 - \frac{n}{\theta} + \frac{\Sigma x_i^2}{2\theta^2} \Rightarrow \theta = \frac{\Sigma x_i^2}{2n}$. Hence, the mle of θ is $\hat{\theta} = \frac{\Sigma x_i^2}{2n}$, identical to the unbiased estimator in Exercise 17. In particular, they share the same numerical value for the given data: $\hat{\theta}$ = 74.505.

b. The median of the Rayleigh distribution satisfies $.5 = \int_0^{\eta} \frac{x}{\theta} e^{-x^2/(2\theta)}\,dx = -e^{-x^2/(2\theta)}\Big|_0^{\eta} = 1 - e^{-\eta^2/(2\theta)}$; solving for η gives $\eta = \sqrt{-2\ln(.5)\theta}$. (Since ln(.5) < 0, the quantity under the square root is positive.) By the invariance principle, the mle of η is $\hat{\eta} = \sqrt{-2\ln(.5)\hat{\theta}} = \sqrt{-\ln(.5)\frac{\Sigma x_i^2}{n}}$. For the given data, the maximum likelihood estimate of η is 10.163.

33.

a. With τ known, the likelihood and log-likelihood functions of θ are

$$L(\theta) = f(x_1,...,x_n;\theta,\tau) = \prod_{i=1}^{n} \frac{\theta}{\tau}\left(1 - \frac{x_i}{\tau}\right)^{\theta-1} = \frac{\theta^n}{\tau^n}\left(\prod_{i=1}^{n}\left[1 - \frac{x_i}{\tau}\right]\right)^{\theta-1} \text{ and}$$

$$\ell(\theta) = \ln[L(\theta)] = n\ln(\theta) - n\ln(\tau) + (\theta-1)\ln\left(\prod_{i=1}^{n}\left[1 - \frac{x_i}{\tau}\right]\right).$$ Differentiate and set equal to zero:

$$0 = \ell'(\theta) = \frac{n}{\theta} - 0 + (1)\ln\left(\prod_{i=1}^{n}\left[1 - \frac{x_i}{\tau}\right]\right) \Rightarrow \hat{\theta} = -\frac{n}{\ln\left(\Pi[1 - x_i/\tau]\right)} = \frac{n}{-\Sigma\ln\left(1 - x_i/\tau\right)}.$$ (The quantity inside the natural log is < 1, so –ln( ) is actually positive.)

b. Since τ is part of the boundary of the x's, we must consider it carefully. The joint pdf (and, hence, the likelihood) is $L(\tau) = \frac{\theta^n}{\tau^n}\left(\prod_{i=1}^{n}\left[1 - \frac{x_i}{\tau}\right]\right)^{\theta-1}$ provided $0 \le x_i < \tau$ for all i. This last constraint is equivalent to $\max(x_i) < \tau$, so the likelihood function w.r.t. τ is only non-zero for $\tau > \max(x_i)$.
For $\theta = 1$, $L(\tau) = 1/\tau^n$, which is monotone decreasing in τ. So, the mle must happen at the left endpoint, i.e. $\max(x_i)$. For $\theta < 1$, the likelihood is also strictly decreasing, and so again the mle is $\max(x_i)$.
For $\theta > 1$, however, we must differentiate the log-likelihood w.r.t. τ:

$$\ell(\tau) = n\ln(\theta) - n\ln(\tau) + (\theta-1)\ln\left(\prod_{i=1}^{n}\left[1 - \frac{x_i}{\tau}\right]\right) = n\ln(\theta) - n\ln(\tau) + (\theta-1)\sum_{i=1}^{n}\ln\left[1 - \frac{x_i}{\tau}\right] \Rightarrow$$

$$\ell'(\tau) = 0 - \frac{n}{\tau} + (\theta-1)\sum_{i=1}^{n}\left[\frac{1}{1 - x_i/\tau}\cdot\frac{x_i}{\tau^2}\right] = -\frac{n}{\tau} + (\theta-1)\sum_{i=1}^{n}\frac{x_i}{\tau(\tau - x_i)} = 0 \Rightarrow (\theta-1)\sum_{i=1}^{n}\frac{x_i}{\tau - x_i} = n.$$ The mle of τ is the implicit solution to this equation, still subject to $\tau > \max(x_i)$.

35. With R_i ~ Exponential(λ) = Gamma(1, 1/λ), the gamma rescaling property implies $Y_i = t_iR_i$ ~ Gamma(1, t_i/λ) = Exponential(λ/t_i). Hence, the joint pdf of the Y's, aka the likelihood, is

$L(\lambda) = f(y_1,...,y_n;\lambda) = (\lambda/t_1)e^{-(\lambda/t_1)y_1}\cdots(\lambda/t_n)e^{-(\lambda/t_n)y_n} = \dfrac{\lambda^n}{t_1\cdots t_n}e^{-\lambda\Sigma(y_i/t_i)}$. To determine the mle, find the log-likelihood, differentiate, and set equal to zero:

$$\ell(\lambda) = \ln[L(\lambda)] = n\ln(\lambda) - \ln(t_1\cdots t_n) - -\lambda\sum_{i=1}^{n}\frac{y_i}{t_i} \Rightarrow \ell'(\lambda) = \frac{n}{\lambda} - 0 - \sum_{i=1}^{n}\frac{y_i}{t_i} = 0 \Rightarrow \lambda = \frac{n}{\sum_{i=1}^{n}(y_i/t_i)}.$$

Therefore, the mle of λ under this model is $\hat{\lambda} = \dfrac{n}{\sum_{i=1}^{n}(Y_i/t_i)}$.

Section 5.3

37. These are simply a matter of reading the t-table.

a. $t^* = 2.228$ **c.** $t^* = 2.947$ **e.** $t^* = 2.492$

b. $t^* = 2.131$ **d.** $n = 5 \Rightarrow$ df = 4; $t^* = 4.604$ **f.** $n = 38 \Rightarrow$ df = 37; $t^* \approx 2.715$

39.

a. A normal probability plot of these 20 values is quite linear.

b. For these data, $\bar{x}$ = 25.05 and s = 2.69. So, a 95% CI for μ is

$$\bar{x} \pm t^*\frac{s}{\sqrt{n}} = 25.05 \pm 2.093\frac{2.69}{\sqrt{20}} = (23.79, 26.31).$$

c. Yes: We're 95% confident the mean ACT score for calculus students is between 23.79 and 26.31, definitely higher than 21.

41.

a. A normal probability plot of these n = 26 values is reasonably linear. From software, $\bar{x}$ = 370.69 and s = 24.36. The critical value is t^* = 2.787, so a 99% CI for μ is

$$\bar{x} \pm t^*\frac{s}{\sqrt{n}} = 370.69 \pm 2.787\frac{24.36}{\sqrt{26}} = (357.38, 384.01)$$

We are 99% confident that the population mean escape time is between 357.38 sec and 384.01 sec.

b. A 90% confidence interval using the same data would necessarily be narrower.

43.

a. Based on a normal probability plot, it is reasonable to assume the sample observations came from a normal distribution.

b. With df = n – 1 = 16, the critical value for a 95% CI is t^* = 2.120, and the interval is $438.29 \pm 2.120\left(\dfrac{15.14}{\sqrt{17}}\right) = 438.29 \pm 7.785 = (430.51, 446.08)$. Since 440 is within the interval, 440 is a plausible value for the true mean. 450, however, is not, since it lies outside the interval.

45. The confidence level for each interval is 100(central area)%. The 20 df row of Table A.5 shows that 1.725 captures upper tail area .05 and 1.325 captures upper tail area .10

For the first interval, central area = 1 – sum of tail areas = 1 – (.25 + .05) = .70, and for the second and third intervals the central areas are 1 – (.20 + .10) = .70 and 1 – (.15 + .15) = 70. Thus each interval has confidence level 70%. The width of the first interval is 2.412 $s/\sqrt{n}$, whereas the widths of the second and third intervals are 2.185 and 2.128 standard errors respectively. The third interval, with symmetrically placed critical values, is the shortest, so it should be used. This will always be true for a t interval.

47. Similar to Exercise 46**b**, 99% one-sided confidence → 1% one-tail area → 2% two-tail area → 98% two-sided confidence. So, the critical values for 99% upper (and lower) confidence bounds are located in the 98% central area column of Table A.5.

For the given data set, n = 10, $\bar{x}$ = 21.90, and s = 4.13. The critical value (df = 9; 99% one-sided confidence) is t^* = 3.250, from which a 99% upper confidence bound for μ is $\bar{x} + t^* s/\sqrt{n}$ = 21.90 + 3.250(4.13)/$\sqrt{10}$ = 26.14. We are 99% confident that the population mean fat content of hot dogs is less than 26.14%.

49.

a. Since $X_1, \ldots, X_n, X_{n+1}$ are independent normal rvs, and $\bar{X} - X_{n+1}$ is a linear combination of them, $\bar{X} - X_{n+1}$ is itself normally distributed. Using the properties of mean/variance as well as the properties of the sample mean,

$E\left(\bar{X} - X_{n+1}\right) = E(\bar{X}) - E(X_{n+1}) = \mu - \mu = 0$ and

$$\text{Var}\left(\bar{X} - X_{n+1}\right) = \text{Var}(\bar{X}) + (-1)^2\,\text{Var}(X_{n+1}) = \frac{\sigma^2}{n} + \sigma^2 = \sigma^2\left(1 + \frac{1}{n}\right)$$

Thus, standardizing, $\dfrac{(\bar{X} - X_{n+1}) - 0}{\sqrt{\sigma^2\left(1 + \frac{1}{n}\right)}}$ ~ N(0, 1), and this is the same as $\dfrac{\bar{X} - X_{n+1}}{\sigma\sqrt{1 + \frac{1}{n}}}$.

b. Since $\dfrac{\bar{X} - X_{n+1}}{S\sqrt{1 + \frac{1}{n}}} \sim t_{n-1}$, $P\left(-t^* < \dfrac{\bar{X} - X_{n+1}}{S\sqrt{1 + \frac{1}{n}}} < t^*\right)$ = C, the prescribed confidence level. Solve the inequalities for X_{n+1}:

$$-t^* < \frac{\bar{X} - X_{n+1}}{S\sqrt{1 + \frac{1}{n}}} < t^* \Rightarrow -t^* S\sqrt{1 + \frac{1}{n}} < \bar{X} - X_{n+1} < t^* S\sqrt{1 + \frac{1}{n}} \Rightarrow$$

$\bar{X} - t^* S\sqrt{1 + \frac{1}{n}} < X_{n+1} < \bar{X} + t^* S\sqrt{1 + \frac{1}{n}}$. Thus, a prediction interval for the true value of the future single observation X_{n+1} has endpoints $\bar{x} \pm t^* s\sqrt{1 + \frac{1}{n}}$.

c. For the given data set, $n = 10$, $\bar{x} = 21.90$, and $s = 4.13$. At df = 10 – 1 = 9 and 95% confidence, $t^* = 2.262$. Hence, a 95% prediction interval is given by $\bar{x} \pm t^* s\sqrt{1+\frac{1}{n}} = 21.90 \pm 2.262(4.13)\sqrt{1+\frac{1}{10}} = (12.10, 31.70)$. We are 95% confident that the fat content of a future, single hot dog will be between 12.10% and 31.70%.

Section 5.4

51.

a. Yes. It is an assertion about the value of a parameter.

b. No. A sample proportion $\hat{P}$ is not a parameter.

c. No. The sample standard deviation S is not a parameter.

d. Yes. The assertion is that the standard deviation of population #2 exceeds that of population #1.

e. No. $\bar{X}$ and $\bar{Y}$ are statistics rather than parameters, so they cannot appear in a hypothesis.

f. Yes. H is an assertion about the value of a parameter.

53. In this formulation, H_0 states the welds do <u>not</u> conform to specification. This assertion will not be rejected unless there is strong evidence to the contrary. Thus the burden of proof is on those who wish to assert that the girder welds <u>are</u> strong enough; i.e., specification is satisfied. Using H_a: $\mu < 100$ results in the welds being believed in conformance unless proved otherwise, so the burden of proof is on the non-conformance claim.

55. Using $\alpha = .05$, H_0 should be rejected whenever P-value < .05.

a. P-value = .001 < .05, so reject H_0

b. .021 < .05, so reject H_0.

c. .078 is not < .05, so don't reject H_0.

d. .047 < .05, so reject H_0 (a close call).

e. .148 > .05, so H_0 can't be rejected at level .05.

57. Use appendix Table A.6.

a. $P(t > 2.0)$ at 8df = .040.

b. $P(t < –2.4)$ at 11df = .018.

c. $2P(t < –1.6)$ at 15df = 2(.065) = .130.

d. By symmetry, $P(t > –.4) = 1 – P(t > .4)$ at 19df = 1 – .347 = .653.

e. $P(t > 5.0)$ at 5df < .005.

f. $2P(t < –4.8)$ at 40df < 2(.000) = .000 to three decimal places.

59. In each case, the P-value equals $P(Z > z) = 1 - \Phi(z)$.

a. .0778 **b.** .1841 **c.** .0250 **d.** .0066 **e.** .5438

61.

a. The appropriate hypotheses are H_0: $\mu = 10$ v. H_a: $\mu < 10$.

b. P-value = $P(t > 2.3) = .017$, which is $\leq .05$, so we would reject H_0. The data indicates that the pens do not meet the design specifications.

c. P-value = $P(t > 1.8) = .045$, which is not $\leq .01$, so we would not reject H_0. There is not enough evidence to say that the pens don't satisfy the design specifications.

d. P-value = $P(t > 3.6) \approx .001$, which gives strong evidence to support the alternative hypothesis.

63.

a. A normal probability plot of the data indicates substantial positive skewness. So, it is not plausible from that the variable ALD is normal. However, since $n = 49$, normality is not required for the use of t or z inference procedures.

b. We wish to test H_0: $\mu = 1.0$ versus H_a: $\mu < 1.0$. The test statistic is $z = \dfrac{0.75 - 1.0}{.3025/\sqrt{49}} = -5.79$; at any reasonable significance level, we reject the null hypothesis. Yes, the data provides strong evidence that the true average ALD is less than 1.0.

65. The hypotheses are H_0: $\mu = 200$ versus H_a: $\mu > 200$. Since $n = 12$, we will use a one-sample t test. With the data provided, $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{249.7 - 200}{145.1/\sqrt{12}} = 1.19$. At df = 11, the P-value is roughly .128.
Since P-value = .128 > α = .05, H_0 is not rejected at the α = .05 level. We have insufficient evidence to conclude that the true average repair time exceeds 200 minutes.

67. Define μ = the true average percentage of organic matter in this type of soil. The hypotheses are H_0: $\mu = 3$ versus H_a: $\mu \neq 3$. With $n = 30$, and assuming normality, we use the t test:

$$t = \frac{\bar{x} - 3}{s/\sqrt{n}} = \frac{2.481 - 3}{.295} = \frac{-.519}{.295} = -1.759.$$ The P-value = $2[P(t > 1.759)] = 2(.041) = .082$.

At significance level .10, since .082 < .10, we would reject H_0 and conclude that the true average percentage of organic matter in this type of soil is something other than 3. At significance level .05, we would not have rejected H_0.

69. Let μ = the true average escape time, in <u>seconds</u>. Since 6 min = 360 sec, the hypotheses are H_0: $\mu = 360$ versus H_a: $\mu > 360$. The test statistic value is $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{370.69 - 360}{24.36/\sqrt{26}} = 2.24$; at df = 25, the one-tailed P-value is roughly .018. Since P-value = .018 < α = .05, H_0 should be rejected. There appears to be a contradiction of the prior belief; the evidence suggests that the true mean escape time exceeds 6 minutes.

71. The parameter of interest is μ = the true average dietary intake of zinc among males aged 65-74 years. The hypotheses are H_0: $\mu = 15$ versus H_a: $\mu < 15$. Since the sample size is large, we'll use a z-procedure here. From the summary statistics provided, $z = \frac{11.3-15}{6.43/\sqrt{115}} = -6.17$, and the P-value is $\Phi(-6.17) \approx 0$. Hence, we reject H_0 at any reasonable significance level. There is convincing evidence that average daily intake of zinc for males aged 65-74 years falls below the recommended daily allowance of 15 mg/day.

73. Let σ denote the population standard deviation. The appropriate hypotheses are H_0: $\sigma = .05$ v. H_a: $\sigma < .05$. With this formulation, the burden of proof is on the data to show that the requirement has been met (the sheaths will not be used unless H_0 can be rejected in favor of H_a. Type I error: Conclude that the standard deviation is < .05 mm when it is really equal to .05 mm. Type II error: Conclude that the standard deviation is .05 mm when it is really < .05.

75. A type I error here involves saying that the plant is not in compliance when in fact it is. A type II error occurs when we conclude that the plant is in compliance when in fact it isn't. Reasonable people may disagree as to which of the two errors is more serious. If in your judgment it is the type II error, then the reformulation H_0: $\mu = 150$ v. H_a: $\mu < 150$ makes the type I error more serious.

Section 5.5

77. The critical value for 95% confidence is $z^* = 1.96$. We have $\hat{p} = \frac{250}{1000} = .25$, from which $\tilde{p} = \frac{.25 + 1.96^2/2000}{1 + 1.96^2/1000}$ = .251. The resulting 95% CI for p, the true proportion of such consumers who never apply for a rebate, is $.251 \pm 1.96\frac{\sqrt{(.25)(.75)/1000 + 1.96^2/(4 \cdot 1000^2)}}{1+1.96^2/1000}$ = (.224, .278).

79. The critical value for 99% confidence is $z^* = 2.576$. $\hat{p} = \frac{201}{356} = .5646$; We calculate a 99% confidence interval for the proportion of all dies that pass the probe:

$$\frac{.5646 + \frac{(2.576)^2}{2(356)} \pm 2.576\sqrt{\frac{(.5646)(.4354)}{356} + \frac{(2.576)^2}{4(356)^2}}}{1 + 2.576^2/356} = (.496, .631).$$

81. With such a large sample size, we'll use the simpler CI formula. The critical value for 99% confidence is $z^* = 2.576$. We have $\hat{p}$ = 25% = .25 and n = 2003, from which the 99% CI is

$$\hat{p} \pm z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .25 \pm 2.576\sqrt{\frac{(.25)(.75)}{2003}} = (.225, .275).$$

83.

a. The margin of error for the CI (5.4) is $z^*\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$. Set this equal to the desired bound, B, for the margin of error and solve for n:

$$B = z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \Rightarrow B^2 = (z^*)^2\frac{\hat{p}(1-\hat{p})}{n} \Rightarrow nB^2 = (z^*)^2\,\hat{p}(1-\hat{p}) \Rightarrow n = \frac{z^2\,\hat{p}(1-\hat{p})}{B^2}$$

(where z and z^* denote the same thing, the critical value for the CI).

b. We have $B = .05$ (5 percentage points) and $z^* = 1.96$ for 95% confidence. Since the legislator believes $p \approx 2/3$, we can use that as an initial value for the as-yet-unknown sample proportion $\hat{p}$:

$$n = \frac{z^2\,\hat{p}(1-\hat{p})}{B^2} = \frac{(1.96)^2(2/3)(1/3)}{(.05)^2} = 341.48.$$

Since sample size must be a whole number, to be safe we should round up to $n = 342$.

c. The expression $x(1 - x)$ on $0 \le x \le 1$ has its absolute maximum at $x = .5$; you can verify this by graphing the function or using calculus. So, the most conservative approach to sample size estimation is to replace the as-yet-unknown $\hat{p}$ with the "worst-case" value of .5. Here, that yields

$n = \dfrac{z^2\,\hat{p}(1-\hat{p})}{B^2} = \dfrac{(1.96)^2(.5)(.5)}{(.05)^2} = 384.16$, which we round up to 385. (Notice that "width of at most .10" is the same as a margin of error bound equal to .05, since the width of a CI is twice the margin of error.)

85. Let p = the true proportion of all commercial chardonnay bottles considered spoiled to some extent by cork-associated characteristics. The hypotheses are H_0: $p = .15$ versus H_a: $p > .15$, the latter being the assertion that the population percentage exceeds 15%. The calculated value of the test statistic and P-value are

$$z = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{(16/91)-.15}{\sqrt{.15(.85)/91}} = 0.69$$ and P-value = $P(Z > 0.69) = 1 - \Phi(0.69) = .2451$.

Since P-value = .2451 > α = .10, we fail to reject H_0 at the .10 level.

87.

a. The parameter of interest is p = the proportion of all wine customers who would find screw tops acceptable. The hypotheses are H_0: $p = .25$ versus H_a: $p < .25$.
With $n = 106$, $np_0 = 106(.25) = 26.5 \ge 10$ and $n(1 - p_0) = 106(.75) = 79.5 \ge 10$, so the "large-sample" z procedure is applicable. From the data provided, $\hat{p} = \dfrac{22}{106} = .208$, so $z = \dfrac{.208-.25}{\sqrt{.25(.75)/106}} = -1.01$. The corresponding P-value for this one-sided test is $\Phi(-1.01) = .1562$.
Since .1562 > .10, we fail to reject H_0 at the α = .10 level. We do <u>not</u> have sufficient evidence to suggest that less than 25% of all customers find screw tops acceptable. Therefore, we recommend that the winery <u>should</u> switch to screw tops.

b. A Type I error would be to incorrectly conclude that less than 25% of all customers find screw tops acceptable, when the true percentage is 25%. Hence, we'd recommend not switching to screw tops when there use is actually justified. A Type II error would be to fail to recognize that less than 25% of all customers find screw tops acceptable when that's actually true. Hence, we'd recommend (as we did in (a)) that the winery switch to screw tops when the switch is not justified. Since we failed to reject H_0 in (a), we may have committed a Type II error.

89.

a. The parameter of interest is p = the proportion of the population of female workers that have BMIs of at least 30 (and, hence, are obese). The hypotheses are H_0: p = .20 versus H_a: p > .20.
With n = 541, $np_0 = 541(.2) = 108.2 \geq 10$ and $n(1 - p_0) = 541(.8) = 432.8 \geq 10$, so the "large-sample" z procedure is applicable.

From the data provided, $\hat{p} = \dfrac{120}{541} = .2218$, so $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \dfrac{.2218 - .20}{\sqrt{.20(.80)/541}} = 1.27$. The corresponding upper-tailed P-value is $1 - \Phi(1.27) = .1020$. Since $.1020 > \alpha = .05$, we fail to reject H_0 at the $\alpha = .05$ level. We do not have sufficient evidence to conclude that more than 20% of the population of female workers is obese.

b. A Type I error would be to incorrectly conclude that more than 20% of the population of female workers is obese, when the true percentage is 20%. A Type II error would be to fail to recognize that more than 20% of the population of female workers is obese when that's actually true.

91. The parameter of interest is p = the proportion of all college students who have maintained lifetime abstinence from alcohol. The hypotheses are H_0: p = .1, H_a: p > .1.
With n = 462, $np_0 = 462(.1) = 46.2 \geq 10$ $n(1 - p_0) = 462(.9) = 415.8 \geq 10$, so the "large-sample" z procedure is applicable.

From the data provided, $\hat{p} = \dfrac{51}{462} = .1104$, so $z = \dfrac{.1104 - .1}{\sqrt{.1(.9)/462}} = 0.74$.

The corresponding one-tailed P-value is $P(Z \geq 0.74) = 1 - \Phi(0.74) \approx .23$.
Since .23 > .05, we fail to reject H_0 at the $\alpha = .05$ level (and, in fact, at any reasonable significance level). The data does not give evidence to suggest that more than 10% of all college students have completely abstained from alcohol use.

93. Let p = true proportion of all nickel plates that blister under the given circumstances. Then the hypotheses are H_0: p = .10 versus H_a: p > .10. The observed value of the test statistic is $z = \dfrac{14/100 - .10}{\sqrt{.10(.90)/100}} = 1.33$, and the corresponding upper-tailed P-value is $1 - \Phi(1.33) = .0918$. Since $.0918 > \alpha = .05$, we fail to reject H_0. The data does not give compelling evidence for concluding that more than 10% of all plates blister under the circumstances.

The possible error we could have made is a Type II error: Failing to reject the null hypothesis when it is actually true.

95. Let p denote the true proportion of those called to appear for service who are black. We wish to test the hypotheses H_0: p = .25 vs H_a: p < .25. With a sample proportion of $\hat{p} = \dfrac{177}{1050} \approx .1686$, the calculated value of the test statistic is $z = \dfrac{.1686 - .25}{\sqrt{.25(.75)/1050}} = -6.09$. The P-value is $\Phi(-6.09) \approx 0$. If H_0 were true, there would essentially be no chance of seeing such a low proportion of blacks among jury servers. Hence, we strongly reject H_0; a conclusion that discrimination exists is very compelling.

97.

a. For testing H_0: $p = .2$ v. H_a: $p > .2$, an upper-tailed test is appropriate. The computed test statistic is $z = .97$, so the P-value $= 1 - \Phi(.97) = .166$. Because the P-value is rather large, H_0 would not be rejected at any reasonable α (it can't be rejected for any $\alpha < .166$), so no modification appears necessary.

b. Using the same methodology as the previous exercise but with an upper-tailed test, $P(\text{reject } H_0 \text{ when } p = p') = P(P\text{-value} \le \alpha \text{ when } p = p') = P(Z \ge +z_\alpha \text{ when } p = p') =$ $1 - \Phi\left(\dfrac{p_0 - p' + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$. With $p' = .5$, $p_0 = .2$, $n = 60$, and $z_\alpha = 2.33$ again, the probability is $1 - \Phi\left[\left(-.3 + 2.33(.0516)\right)/.0645\right] = 1 - \Phi(-2.79) = .9974$. Even with a moderate sample size, it's very easy to detect that $p > .2$ when in actuality $p = .5$!

Section 5.6

99.

a. For a gamma distribution with parameters α and β, the mean is $\alpha\beta$ and the standard deviation is $\sqrt{\alpha}\beta$. Here, we want the prior distribution of μ to satisfy $\alpha\beta = 15$ and $\sqrt{\alpha}\beta = 5$. Divide the two equations to get $\sqrt{\alpha} = 15/5 = 3$, so $\alpha = 9$; then, $\beta = 15/\alpha = 15/9 = 5/3$. So, the prior for μ will be Gamma(9, 5/3).

b. The prior for μ is Gamma(9, 5/3); conditional on μ, the observations $X_1, \ldots, X_n$ are assumed to be a random sample from a Poisson(μ) distribution. Hence, the <u>numerator</u> of the posterior distribution of μ is

$$\begin{aligned}\pi(\mu) f(x_1,\ldots,x_n;\mu) &= \frac{1}{\Gamma(9)(5/3)^9}\mu^{9-1}e^{-\mu/(5/3)} \cdot \frac{e^{-\mu}\mu^{x_1}}{x_1!}\cdots\frac{e^{-\mu}\mu^{x_n}}{x_n!} \\ &= C\mu^{8+\Sigma x_i}e^{-(3/5)\mu - n\mu} = C\mu^{\alpha_1 - 1}e^{-\mu/\beta_1},\end{aligned}$$

where $\alpha_1 = 9 + \Sigma x_i$ and $\beta_1 = \dfrac{1}{\frac{3}{5} + n}$. We recognize the last expression above as the "kernel" of a gamma distribution (i.e., the pdf without the constant in front). Therefore, we conclude that the posterior distribution of μ is also gamma, but with the parameters α_1 and β_1 specified above.
With the specific values provided, $n = 10$ and $\Sigma x_i = 136$, so the posterior distribution of μ given these observed data is Gamma(α_1, β_1) = Gamma(145, 5/53).

c. A 95% credibility interval is defined by the middle 95% of the posterior distribution of μ. With the aid of software, we find the .025 and .975 quantiles of the Gamma(145, 5/53) distribution are $\eta_{.025} = 11.54$ and $\eta_{.975} = 15.99$. Hence, a 95% credibility interval for μ is (11.54, 15.99).

101. The difference between this exercise and Example 5.25 is that we treat the data as n = 939 observations from a Bernoulli(θ) distribution, rather than a single observation from a binomial distribution. The Bernoulli pmf is $p(x;\theta) = \theta^x(1-\theta)^{1-x}$ for x = 0 or 1. Hence, the numerator of the posterior distribution of θ is

$$\pi(\theta)\,p(x_1,...,x_n;\theta) = 20\theta(1-\theta)^3\cdot\theta^{x_1}(1-\theta)^{1-x_1}\cdots\theta^{x_n}(1-\theta)^{1-x_n} = 20\theta^{1+\Sigma x_i}(1-\theta)^{3+\Sigma(1-x_i)}$$
$$\propto \theta^{1+\Sigma x_i}(1-\theta)^{3+n-\Sigma x_i} = \theta^{1+\Sigma x_i}(1-\theta)^{942-\Sigma x_i}$$

We recognize this as the kernel of a Beta distribution, with parameters $\alpha = 2 + \Sigma\, x_i$ and $\beta = 943 - \Sigma\, x_i$. In this contest, $\Sigma\, x_i$ = the sum of the 1's in the Bernoulli sample = the number of successes (out of 939) in the sample. If we use the same observed value of successes, 488, provided in Example 5.25, the posterior distribution of θ is a Beta distribution with α = 2 + 488 = 490 and β = 943 – 488 = 455, the same posterior distribution found in Example 5.25.

The lesson here is that we may regard the yes/no responses either as a Bernoulli random sample or as a single binomial observation; either perspective yields the exact same posterior distribution for the unknown proportion θ.

103. Using (5.5), the numerator of the posterior distribution of μ is

$$\pi(\mu)\,f(x_1,...,x_n;\mu) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\mu^{\alpha-1}e^{-\mu/\beta}\frac{e^{-\mu}\mu^{x_1}}{x_1!}\cdots\frac{e^{-\mu}\mu^{x_n}}{x_n!} = C\mu^{\alpha+\Sigma x_i-1}e^{-\mu[1/\beta+n]}$$. This is the kernel of a gamma pdf, specifically with first parameter $\alpha + \Sigma\, x_i$ and second parameter $1/[1/\beta + n]$. Therefore, the posterior distribution of μ is Gamma($\alpha + \Sigma\, x_i$, $1/(n + 1/\beta)$).

105. This is very similar to Example 5.25. Using (5.5), the numerator of the posterior distribution of p is

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}\cdot\binom{n}{x}p^x(1-p)^{n-x} \propto p^{\alpha+x-1}(1-p)^{\beta+n-x-1}$$. We recognize this as the kernel of a beta distribution with parameters $\alpha + x$ and $\beta + n - x$. Hence, the posterior distribution of p given the binomial observation $X = x$ is Beta($\alpha + x$, $\beta + n - x$).

Supplementary Exercises

107. Let x_1 = the time until the first birth, x_2 = the elapsed time between the first and second births, and so on. Then $f(x_1,...,x_n;\lambda) = \lambda e^{-\lambda x_1}\cdot(2\lambda)e^{-2\lambda x_2}...(n\lambda)e^{-n\lambda x_n} = n!\lambda^n e^{-\lambda\Sigma kx_k}$. Thus the log likelihood is $\ln(n!) + n\ln(\lambda) - \lambda\Sigma kx_k$. Taking $\frac{d}{d\lambda}$ and equating to 0 yields $\hat{\lambda} = \frac{n}{\Sigma kx_k}$.

For the given sample, n = 6, x_1 = 25.2, x_2 = 41.7 – 25.2 = 16.5, x_3 = 9.5, x_4 = 4.3, x_5 = 4.0, x_6 = 2.3; so $\sum_{k=1}^{6} kx_k = (1)(25.2)+(2)(16.5)+...+(6)(2.3) = 137.7$ and $\hat{\lambda} = \frac{6}{137.7} = .0436$.

109.

a. The likelihood is $\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x_i-\mu_i)}{2\sigma^2}}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y_i-\mu_i)}{2\sigma^2}}=\frac{1}{\left(2\pi\sigma^2\right)^n}e^{-\frac{\left(\Sigma(x_i-\mu_i)^2+\Sigma(y_i-\mu_i)^2\right)}{2\sigma^2}}$. The log likelihood is thus $-n\ln\left(2\pi\sigma^2\right)-\frac{\left(\Sigma(x_i-\mu_i)^2+\Sigma(y_i-\mu_i)^2\right)}{2\sigma^2}$. Taking $\frac{d}{d\mu_i}$ and equating to zero gives $\hat{\mu}_i=\frac{x_i+y_i}{2}$. Substituting these estimates of the $\hat{\mu}_i$s into the log likelihood gives $-n\ln\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left(\sum\left(x_i-\frac{x_i+y_i}{2}\right)^2+\sum\left(y_i-\frac{x_i+y_i}{2}\right)^2\right)=-n\ln\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left(\frac{1}{2}\Sigma\left(x_i-y_i\right)^2\right)$. Now taking $\frac{d}{d\sigma^2}$, equating to zero, and solving for σ^2 gives the desired result.

b. $E\left(\hat{\sigma}^2\right)=\frac{1}{4n}E\left(\Sigma\left(X_i-Y_i\right)^2\right)=\frac{1}{4n}\cdot\Sigma E\left(X_i-Y_i\right)^2$, but $E\left(X_i-Y_i\right)^2=\text{Var}\left(X_i-Y_i\right)+\left[E\left(X_i-Y_i\right)\right]^2=$ $2\sigma^2-[0]^2=2\sigma^2$. Thus $E\left(\hat{\sigma}^2\right)=\frac{1}{4n}\Sigma\left(2\sigma^2\right)=\frac{1}{4n}2n\sigma^2=\frac{\sigma^2}{2}$, so the mle is definitely not unbiased—the expected value of the estimator is only half the value of what is being estimated!

111.

a. $\sum_{x=-1}^{\infty}p(x;\theta)=\theta+(1-\theta)^2\sum_{x=0}^{\infty}\theta^x=\theta+(1-\theta)^2\frac{1}{1-\theta}=\theta+(1-\theta)=1$, so $p(x;\theta)$ is a legitimate pmf. Next, $E(X)=\sum_{x=-1}^{\infty}xf(x;\theta)=-\theta+(1-\theta)\sum_{x=0}^{\infty}x(1-\theta)\theta^x=-\theta+(1-\theta)\frac{\theta}{1-\theta}=0$; we've recognized the sum as the expectation of a geometric rv with $p=1-\theta$.

b. Let $Y_i=1$ if $X_i=-1$ and 0 otherwise, so $\Sigma Y_i=Y$. Then a (very devious) way to write the marginal pmf is
$p(x_i;\theta)=(1-\theta)^{2-2y_i}\theta^{x_i+2y_i}$, from which the joint pmf is $(1-\theta)^{2n-2y}\theta^{\Sigma x_i+2y}$. To find the MLE, differentiate the log-likelihood function to get $\hat{\theta}=\frac{\Sigma x_i+2y}{\Sigma x_i+2n}$.

113.

a. The pattern of points in a normal probability plot (not shown) is reasonably linear, so, yes, normality is plausible.

b. $n=18$, $\bar{x}=38.66$, $s=8.473$, and $t^*=2.567$ (the 98% confidence critical value at df = 18 – 1 = 17). The 98% confidence interval is $38.66\pm2.567\frac{8.473}{\sqrt{18}}=38.66\pm5.13=(33.53,43.79)$.

115. $\hat{p} = \frac{11}{55} = .2 \Rightarrow$ a 90% CI is $\dfrac{.2 + \frac{1.645^2}{2(55)} \pm 1.645\sqrt{\frac{(.2)(.8)}{55} + \frac{1.645^2}{4(55)^2}}}{1 + \frac{1.645^2}{55}} = \dfrac{.2246 \pm .0887}{1.0492} = (.1295, .2986)$.

117.

a. A normal probability plot (not shown) lends support to the assumption that pulmonary compliance is normally distributed.

b. For 95% confidence at df = 16 – 1 = 15, t^* = 2.131, so the CI is $209.75 \pm 2.131\frac{24.156}{\sqrt{16}} = 209.75 \pm 12.87 = (196.88, 222.62)$.

119.

a. With such a large sample size, we'll use the "simplified" CI formula (5.4). Here, $\hat{p} = \frac{1262}{2253}$ = .56, so a 95% CI for p is $\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = .56 \pm 1.96\sqrt{\frac{(.56)(.44)}{2253}}$ = .56 ± .021 = (.539, .581). We are 95% confident that between 53.9% and 58.1% of all American adults have at some point used wireless means for online access.

b. Using the "simplified" formula again, $n = \frac{z^2\hat{p}\hat{q}}{B^2} = \frac{(1.96)^2(.5)(.5)}{(.02)^2} = 2401$, where B = bound on the margin of error = half of the bound on the width = .04/2 = .02. So, roughly 2400 people should be surveyed to assure a width no more than .04 with 95% confidence.

121. Write $h(X_1, X_2) = (\theta_2 X_1 - \theta_1 X_2)/\sqrt{\theta_1^2 + \theta_2^2}$.

a. Since h is a linear function of two normal rvs, h is itself normal. $E(\theta_2 X_1 - \theta_1 X_2) = \theta_2\theta_1 - \theta_1\theta_2 = 0$, so $E(h) = 0$, and $\text{Var}(\theta_2 X_1 - \theta_1 X_2) = \theta_2^2\text{Var}(X_1) + \theta_1^2\text{Var}(X_2) = \theta_2^2 + \theta_1^2$, so Var($h$) = 1.

Divide numerator and denominator by θ_2, and h becomes $\frac{X_1 - rX_2}{\sqrt{r^2+1}}$, where $r = \theta_1/\theta_2$ is the desired ratio of means.

b. From (a), h is standard normal. So, $P(-1.96 < h < 1.96) = .95$. Solve $h = 1.96$ for r: $X_1 - rX_2 = 1.96\sqrt{r^2+1} \Rightarrow (X_1 - rX_2)^2 = (1.96)^2(r^2 + 1)$; solving this quadratic for r gives two solutions: $r = \dfrac{x_1x_2 \pm 1.96\sqrt{x_1^2 + x_2^2 - (1.96)^2}}{x_2^2 - (1.96)^2}$.

These are the confidence limits if the discriminant is non-negative; i.e., if $x_1^2 + x_2^2 \geq (1.96)^2$. Otherwise, we have to solve for the opposite ratio, $1/r$, and the resulting set of values has the form $r < a$ or $r > b$ with $a < b$.

123.

a. If A_1 and A_2 are independent, then $P(A_1 \cap A_2) = P(A_1)P(A_2) = (.95)^2 = 90.25\%$.

b. For any events A and B, $P(A \cup B) = P(A) + P(B) - P(A \cap B) \le P(A) + P(B)$. Apply that here: $P(A_1' \cup A_2') \le P(A_1') + P(A_2') = (1 - .95) + (1 - .95) = .10$, so that $P(A_1 \cap A_2) = 1 - P(A_1' \cup A_2') \ge 1 - .10 = .90$.

c. Replace .05 with α above, and you find $P(A_1 \cap A_2) \ge 100(1 - 2\alpha)\%$. In general, the simultaneous confidence level for k separate CIs is at least $100(1 - k\alpha)\%$. Importantly, this simultaneous confidence level <u>decreases</u> as k increases.

125.

a. H_0: $\mu = 2150$ v. H_a: $\mu > 2150$

b. $t = \dfrac{\bar{x} - 2150}{s/\sqrt{n}}$

c. $t = \dfrac{2160 - 2150}{30/\sqrt{16}} = \dfrac{10}{7.5} = 1.33$

d. At 15df, P-value $= P(T > 1.33) = .107$ (approximately)

e. From **d**, P-value > .05, so H_0 cannot be rejected at this significance level. The mean tensile strength for springs made using roller straightening is not significantly greater than 2150 N/mm^2.

127. Let μ = true average heat-flux of plots covered with coal dust. The hypotheses are H_0: $\mu = 29.0$ versus H_a: $\mu > 29.0$. The calculated test statistic is $t = \dfrac{30.7875 - 29.0}{6.53/\sqrt{8}} = .7742$. At df = 8 – 1 = 7, the P-value is approximately .232. Hence, we fail to reject H_0. The data does not indicate the mean heat-flux for pots covered with coal dust is greater than for plots covered with grass.

129. A normality plot reveals that these observations could have come from a normally distributed population, therefore a t-test is appropriate. The relevant hypotheses are H_0: $\mu = 9.75$ v. H_a: $\mu > 9.75$. Summary statistics are $n = 20$, $\bar{x} = 9.8525$, and $s = .0965$, which leads to a test statistic $t = \dfrac{9.8525 - 9.75}{.0965/\sqrt{20}} = 4.75$, from which the P-value ≈ 0. With such a small P-value, the data strongly supports the alternative hypothesis. The condition is not met.

131. A t test is appropriate. H_0: $\mu = 1.75$ is rejected in favor of H_a: $\mu \ne 1.75$ if the P-value < .05. The computed test statistic is $t = \dfrac{1.89 - 1.75}{.42/\sqrt{26}} = 1.70$. Since the P-value is $2P(T > 1.70) = 2(.051) = .102 > .05$, do not reject H_0; the data does not contradict prior research.
We assume that the population from which the sample was taken was approximately normally distributed.

133. Let p = the true proportion of mechanics who could identify the problem. Then the appropriate hypotheses are H_0: p = .75 vs H_a: p < .75, so a lower-tailed test should be used. With p_0= .75 and $\hat{p} = \frac{42}{72} = .583$, the test statistic is z = –3.28 and the P-value is $\Phi(-3.28)$ = .0005. Because this P-value is so small, the data argues strongly against H_0, so we reject it in favor of H_a.

135. Note: It is <u>not</u> reasonable to use a z test here, since the values of p are so small.

a. Let p = the proportion of all mis-priced purchases at all California Wal-Mart stores. We wish to test the hypotheses H_0: $p \leq .02$ v. H_a: $p > .02$.
Let X = the number of mis-priced items in 200, so $X \sim \text{Bin}(200,.02)$ under the null hypothesis. For our data, the observed value of x is .083(200) = 16.6 ≈ 17. The P-value of this upper-tailed test, with the aid of software, is $P(X \geq 17 \text{ when } X \sim \text{Bin}(200,.02)) = 7.5 \times 10^{-7}$. This P-value is absurdly small, so we clearly reject H_0 here and conclude that the NIST benchmark is <u>not</u> satisfied.

b. We'll not reject H_0 (thus siding with the NIST standard) if the P-value is suitably high; say, greater than the standard α = .05. Replacing 17 above with progressively lower numbers and calculating using the Bin(200,.02) distribution shows $P(X \geq 8) = .049$ while $P(X \geq 7) = .109$, so H_0 will not be rejected provided that the observed value of x is 7 or fewer.
If in reality p = 5%, then actually $X \sim \text{Bin}(200,.05)$ (as opposed to the null distribution). Under that model, P(observed value of X is 7 or fewer) = $B(7; 200,.05)$ = .2133. That is, there's a 21.33% chance of concluding the NIST standard is "met" (really, not realizing it hasn't been met) when p = 5%, assuming we use α = .05.

137. We wish to test H_0: $\mu = 4$ versus H_a: $\mu > 4$ using the test statistic $z = \frac{\overline{x} - 4}{\sqrt{4/n}}$. For the given sample, n = 36 and $\overline{x} = \frac{160}{36} = 4.444$, so $z = \frac{4.444 - 4}{\sqrt{4/36}} = 1.33$. The P-value is $P(Z \geq 1.33) = 1 - \Phi(1.33) = .0918$. Because the P-value is greater than α = .02, H_0 should not be rejected at this level. We do not have significant evidence at the .02 level to conclude that the true mean of this Poisson process is greater than 4.

CHAPTER 6

Section 6.1

1. The state space of the chain is {cooperative strategy, competitive strategy}. If we designate these states as 1 = cooperative and 2 = competitive, then $p_{11} = .6$, from which $p_{12} = .4$, and $p_{22} = .7$, from which $p_{21} = .3$.

$$.6 \hookrightarrow 1 \;\; \underset{.3}{\overset{.4}{\rightleftarrows}} \;\; 2 \hookleftarrow .7$$

3.

a. The state space of the chain is {full, part, broken}.

b. Identify the states as 1 = full, 2 = part, 3 = broken. Then we're told $p_{11} = .7$, $p_{12} = .2$, $p_{13} = .1$; $p_{21} = 0$, $p_{22} = .6$, $p_{23} = .4$; and $p_{31} = .8$, $p_{32} = 0$, $p_{33} = .2$.

$$
\begin{array}{ccc}
{}^{.7}\hookrightarrow 1 & \overset{.2}{\rightarrow} & 2 \hookleftarrow^{.6} \\
{}^{.1}\downarrow\uparrow^{.8} & \swarrow_{.4} & \\
{}^{.2}\hookrightarrow 3 & &
\end{array}
$$

5.

a. Either the single cell splits into 2 new cells, or it dies off. So, $X_1 = 2$ with probability p and $X_1 = 0$ with probability $1 - p$.

b. If $X_1 = 0$, then necessarily $X_2 = 0$ as well. If $X_1 = 2$, then each of these two cells may split into two more or die off. This allows for possibilities (0 or 2) + (0 or 2) = 0, 2, or 4. So, the total possible values for X_2 are 0, 2, and 4.

c. Analogous to part **b**, with x cells in the nth generation the possible values of X_{n+1} are 0, 2, …, $2x$. Let Y denote the number of cells, out of these x, that split into two rather than die. Then Y is actually a binomial rv, with parameters $n = x$ and p. Therefore, the one-step transition probabilities are

$$P(X_{n+1} = 2y \mid X_n = x) = P(\text{exactly } y \text{ of the } x \text{ cells split}) = \binom{x}{y} p^y (1-p)^{x-y} \quad \text{for } y = 0, 1, \ldots, x.$$

7.

a. If social status forms a Markov chain by generations, then a son's social status, given his father's social status, has the same probability distribution as his social status conditional on all family history. That <u>doesn't</u> mean that his grandfather's social status plays no role: it informed his father's social status, which in turn informed his own. The Markov property just means that the social prospects of the son, conditional on his father's <u>and</u> grandfather's social statuses, is the same as the son's social prospects conditional just on his father's.

b. The process is time-homogeneous if the probabilities of social status changes (e.g., poor to middle class) are the same in every generation. For example, the chance that a poor man in 1860 has a middle-class son would equal the chance a poor man in 2014 has a middle-class son. This is not realistic: social mobility is not the same now as it was in the past.

9.

a. No, X_n is not a Markov chain in this revised setting. The likelihood of X_{n+1} being cooperative v. competitive now depends directly on both X_n and X_{n-1}, rather than solely on X_n.

b. Define Y_n = (strategy at stage n, strategy at stage n+1). Then, analogous to Example 6.7, the chain of pairs Y_0, Y_1, Y_2, … forms a Markov chain. If we use 1 = cooperative and 2 = competitive again, the state space of Y_n is {(1,1), (1,2), (2,1), (2,2)}. To formulate the chain, we require at least four one-step transition probabilities for this set of pairs — eight of the 4^2 = 16 transitions are impossible, such as (2,1)→(2,2), because the inner terms must agree, and the other four can be determined as complements. See Example 6.7 for an illustration.

Section 6.2

11.

a. From the information provided, $\mathbf{P} = \begin{bmatrix} .90 & .10 \\ .11 & .89 \end{bmatrix}$.

b. $P(X_2 = \text{F} \mid X_0 = \text{F}) = P(X_2 = 1 \mid X_0 = 1) = P^{(2)}(1 \to 1)$. By the Chapman-Kolmogorov Equation, this is the (1, 1) entry of the matrix $\mathbf{P}^2$. Using matrix multiplication, $\mathbf{P}^2 = \begin{bmatrix} .8210 & .1790 \\ .1969 & .8031 \end{bmatrix}$, so $P^{(2)}(1 \to 1)$ = .8210. Similarly, $P(X_{13} = \text{F} \mid X_0 = \text{F}) = P^{(13)}(1 \to 1)$ = the (1, 1) entry of $\mathbf{P}^{13}$ = .5460 via software.

c. $P(X_2 = \text{NF} \mid X_0 = \text{NF}) = P(X_2 = 2 \mid X_0 = 2) = P^{(2)}(2 \to 2)$. By the Chapman-Kolmogorov Equation, this is the (2, 2) entry of the matrix $\mathbf{P}^2$. From above, $P^{(2)}(2 \to 2)$ = .8031. Similarly, $P(X_{13} = \text{NF} \mid X_0 = \text{NF})$ = $P^{(13)}(2 \to 2)$ = the (1, 1) entry of $\mathbf{P}^{13}$ = .5006 via software.

d. These projections are only valid if the Markov model continues to hold for the rest of the century. This assumes, for example, that long-run trends in deforestation and reforestation can be reduced to one-year transition probabilities (this might or might not be realistic). It also assumes that the likelihoods of these events do not change over time; e.g., it will always be the case that a forested area has a 90% chance of remaining forested next year. That time-homogeneity assumption does not seem realistic.

13.

a. Snow is more likely to stay on the ground in Willow City, ND than in NYC. In Willow City, the chance a snowy day is followed by another snowy day (i.e., snow stays on the ground or more snowfall occurs) is $P(S \to S)$ = .988. The same probability for New York is just .776.

b. $P(X_{n+2} = S \mid X_n = S) = P^{(2)}(S \to S)$. By the Chapman-Kolmogorov Equation, this is the (2, 2) entry of the matrix $\mathbf{P}^2$, since S is state 2. $\mathbf{P}^2 = \begin{bmatrix} .8713 & .1287 \\ .0231 & .9769 \end{bmatrix}$, so $P^{(2)}(S \to S)$ = .9769. Similarly, the chance of snow three days from now is $P^{(3)}(S \to S)$ = the (2, 2) entry of $\mathbf{P}^3$ = .9668.

c. This doesn't require matrix multiplication, but rather simply the Multiplication Rule from Chapter 1:

$$P(X_1 = S \cap \cdots \cap X_4 = S \mid X_0 = S)$$
$$= P(X_1 = S \mid X_0 = S)P(X_2 = S \mid X_1 = S, X_0 = S) \cdots P(X_4 = 1 \mid X_3 = S, ..., X_0 = S)$$
$$= P(X_1 = S \mid X_0 = S)P(X_2 = S \mid X_1 = S) \cdots P(X_4 = S \mid X_4 = S) \quad \text{Markov property}$$
$$= P(S \to S)P(S \to S) \cdots P(S \to S) = [P(S \to S)]^4$$
$$= [.988]^4 = .9529$$

15.

a. From the information provided, $\mathbf{P} = \begin{bmatrix} .6 & .4 \\ .3 & .7 \end{bmatrix}$.

b. $P(X_{n+2}$ = competitive | X_n = cooperative) = $P(X_{n+2} = 2 \mid X_n = 1) = P^{(2)}(1 \to 2)$ = the (1,2) entry of the matrix $\mathbf{P}^2$ = .52 by direct computation.

c. Once negotiations arrive in state 3, they stay in state 3. So, to find the probability that negotiations end within three time steps, we just need to see whether $X_3 = 3$ or not. That is, P(negotiations end within three time steps) = $P(X_3 = 3 \mid X_0 = 1) = P^{(3)}(1 \to 3)$ = the (1,3) entry of the new matrix $\mathbf{P}^3$ = .524.

d. Similarly, $P(X_3 = 3 \mid X_0 = 2) = P^{(3)}(2 \to 3)$ = the (2,3) entry of the new matrix $\mathbf{P}^3$ = .606. Interestingly, starting with a competitive strategy seems to end the process more quickly! (Of course, nobody said the negotiations end with an agreement.)

17. It was established in Exercise 16 that $\mathbf{P} = \begin{bmatrix} .70 & .30 & 0 & 0 & 0 \\ .14 & .62 & .24 & 0 & 0 \\ 0 & .14 & .62 & .24 & 0 \\ 0 & 0 & .14 & .62 & .24 \\ 0 & 0 & 0 & .14 & .86 \end{bmatrix}$.

a. Friday is 5 nights from now: $P(X_5 = 2 \mid X_0 = 2) = P^{(5)}(2 \to 2)$ = the (2,2) entry of $\mathbf{P}^5$ = .2740 with the aid of software. Similarly, $P(X_5 \geq 2 \mid X_0 = 2) = P(X_5 = 2,3,4 \mid X_0 = 2) = P^{(5)}(2 \to 2) + P^{(5)}(2 \to 3) + P^{(5)}(2 \to 4)$ = the sum of the (2,2) and (2,3) and (2,4) entries of $\mathbf{P}^5$ = .7747.

b. This event occurs if she has 0 umbrellas at the end of Wednesday night, three days after Sunday night: $P(X_3 = 0 \mid X_0 = 2) = P^{(3)}(2 \to 0)$ = the (2,0) entry of $\mathbf{P}^3$ = .0380.

c. Given $X_0 = 2$, the probabilities of $X_1 = 0, \ldots, X_1 = 4$ appear in the "state = 2" row of $\mathbf{P}$, which is the vector [0 .14 .62 .24 0]. Thus, $E(X_1 \mid X_0 = 2) = 0(0) + 1(.14) + 2(.62) + 3(.24) + 4(0) = 2.1$ umbrellas. Similarly, the probabilities of $X_2 = 0, \ldots, X_2 = 4$ appear in the "state = 2" row of $\mathbf{P}^2$ which is the vector [.0196 .1736 .4516 .2976 .0576]. Thus, $E(X_2 \mid X_0 = 2) = 0(.0196) + 1(.1736) + 2(.4516) + 3(.2976) + 4(.0576) = 2.2$ umbrellas.

19.

a. If $X_n = i$, then $X_{n+1} \sim$ Poisson(μ_i), so $P(i \to 0) = \dfrac{e^{-\mu_i}\mu_i^0}{0!} = e^{-\mu_i}$; $P(i \to 1) = \dfrac{e^{-\mu_i}\mu_i^1}{1!} = \mu_i e^{-\mu_i}$; $P(i \to 2) = \dfrac{e^{-\mu_i}\mu_i^2}{2!} = \dfrac{1}{2}\mu_i^2 e^{-\mu_i}$; $P(i \to 4) = \dfrac{e^{-\mu_i}\mu_i^3}{3!} = \dfrac{1}{6}\mu_i^3 e^{-\mu_i}$; and, as stated in the problem, $P(i \to 4)$ is really equal to $1 - [P(i \to 0) + \ldots + P(i \to 3)]$. These formulas define any row of the transition matrix; substituting the numerical values of μ_i returns the following matrix:

$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{bmatrix} .1439 & .2790 & .2704 & .1747 & .1320 \\ .2201 & .3332 & .2522 & .1272 & .0674 \\ .1481 & .2829 & .2701 & .1719 & .1269 \\ .0874 & .2129 & .2596 & .2109 & .2292 \\ .0319 & .1099 & .1893 & .2174 & .4516 \end{bmatrix}.$$

b. $P(X_2 = 2 \cap X_1 = 2 \mid X_0 = 2) = P(X_1 = 2 \mid X_0 = 2) \cdot P(X_2 = 2 \mid X_1 = 2, X_0 = 2)$ by the multiplication rule, $= P(X_1 = 2 \mid X_0 = 2) \cdot P(X_2 = 2 \mid X_1 = 2)$ by the Markov property $= P(2 \to 2)P(2 \to 2) = P(2 \to 2)^2 = (.2701)^2 = .07295401$.

c. $P(X_1 + X_2 = 2 \mid X_0 = 0) = P((X_1, X_2) = (0,2)$ or $(1,1)$ or $(2,0) \mid X_0 = 0) = P(0 \to 0)P(0 \to 2) + P(0 \to 1)P(1 \to 1) + P(0 \to 2)P(2 \to 0) = (.1439)(.2704) + (.2790)(.3332) + (.2704)(.1481) = .1719$.

21. It was established in Exercise 20 that
$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{bmatrix} .9477 & .0486 & .0036 & .0001 \\ .729 & .243 & .027 & .001 \\ 0 & .81 & .18 & .01 \\ 0 & 0 & .9 & .1 \end{bmatrix}.$$

a. Four operational vans corresponds to $X_0 = 0$ (none broken). The desired probabilities are $P^{(k)}(0 \to 1)$ for $k = 2, 3, 4$ (two, three, and four days after Monday). The (0,1) entry of $\mathbf{P}^2$, $\mathbf{P}^3$, and $\mathbf{P}^4$ are .0608, .0646, and .0658, respectively.

b. Looking k days ahead, $P(X_k \geq 1 \mid X_0 < 1) = P(X_k \geq 1 \mid X_0 = 0) = 1 - P(X_k = 0 \mid X_0 = 0) = 1 - P^{(k)}(0 \to 0)$. For $k = 1, 2, 3, 4$, these probabilities are $1 -$ [the (0,0) entry of $\mathbf{P}^k$] = .0523, .0664, .0709, and .0725, respectively.

c. Now the probabilities of interest are $1 - P^{(k)}(1 \to 0)$ for $k = 1, 2, 3, 4$. Using the (1,0) entries of the appropriate matrices, we determine these to be .2710, .1320, .0926, and .0798, respectively. Not surprisingly, the likelihoods of a backlog have increased by assuming a backlog already existed at the beginning of the week. Notice, though, that the backlog is likely to clear up, in the sense that the probability of a backlog continuing decreases as time increases (from .2710 Tuesday to .0798 Friday).

Section 6.3

23. The one-step transition matrix for this Markov chain is $\mathbf{P} = \begin{bmatrix} .6 & .4 \\ .3 & .7 \end{bmatrix}$. If there's a 75% chance the chain starts in cooperative strategy (state 1), the initial probability vector is $\mathbf{v}_0 = [.75\ .25]$.

a. $\mathbf{v}_1 = \mathbf{v}_0\mathbf{P} = [.75\ .25]\begin{bmatrix} .6 & .4 \\ .3 & .7 \end{bmatrix} = [.525\ .475]$. So, the (unconditional) probability that negotiator B's first strategy will be cooperative (state 1) is $v_{11} = .575$. Or, just use the law of total probability: $P(X_1 = 1) = P(X_0 = 1)P(1 \to 1) + P(X_0 = 2)P(2 \to 1) = (.75)(.6) + (.25)(.3) = .575$.

b. Negotiator B's second move occurs on time step 3 (since A and B alternate). By direct computation, $\mathbf{v}_3 = \mathbf{v}_0\mathbf{P}^3 = [.4372\ .5628]$. So, in particular, the chance negotiator B uses a cooperative strategy on his second move is .4372.

25.

a. From the information provided, $\mathbf{P} = \begin{array}{c} 0 \\ 1 \end{array}\begin{bmatrix} .96 & .04 \\ .05 & .95 \end{bmatrix}$.

b. With $\mathbf{v}_0 = [.8\ \ .2]$, $\mathbf{v}_1 = \mathbf{v}_0\mathbf{P} = [.8\ \ .2]\begin{bmatrix} .96 & .04 \\ .05 & .95 \end{bmatrix} = [.778\ \ .222]$. That is, the first relay will send 77.8% 0's and 22.2% 1's.

c. With the aid of software, $\mathbf{v}_5 = \mathbf{v}_0\mathbf{P}^5 = [.7081\ \ .2919]$. So, the bits exiting relay 5 are 70.81% 0's and 29.19% 1's.

27.

a. Since the order of the states is 1 = G (green) and 2 = S (snowy), the 20% chance of a snowy day can be expressed as $\mathbf{v}_0 = [P(G)\ \ P(S)] = [1 - .20\ \ \ .20] = [.80\ \ .20]$.

b. Using $\mathbf{v}_0$ from part **a** and $\mathbf{P}$ from Example 6.10, $\mathbf{v}_1 = \mathbf{v}_0\mathbf{P} = [.8\ .2]\begin{bmatrix} .964 & .036 \\ .224 & .776 \end{bmatrix} = [.816\ \ .184]$. That is, there's an 81.6% chance of a green day tomorrow and an 18.4% chance of a snowy day tomorrow.

c. The weather conditions 7 days from now have probabilities given by $\mathbf{v}_7 = \mathbf{v}_0\mathbf{P}^7$, which with the aid of software is found to be [.8541 .1459]. So, in particular, the chance of a green day one week from now equals .8541.

29.

a. If negotiator A always starts with a competitive strategy (state 2), then $P(X_0 = 1) = 0$ and $P(X_0 = 2) = 1$. Expressed as a probability vector, $\mathbf{v}_0 = [0\ \ 1]$.

b. Since the initial state is definitely 2 (competitive), we need only look at the distribution coming out of state 2; i.e., the second row of $\mathbf{P}$: $P(X_1 = 1) = p_{21} = .3$ and $P(X_1 = 2) = p_{22} = .7$.

c. $\mathbf{v}_2 = \mathbf{v}_1\mathbf{P} = [.3\ \ .7]\begin{bmatrix} .6 & .4 \\ .3 & .7 \end{bmatrix} = [.39\ \ .61]$. So, the probability that negotiator A's next strategy is cooperative (state 1) equals 39%, while the chance of another competitive strategy is 61%.

Section 6.4

31.

a. No, this is not a regular Markov chain. Since $P(gg \to gg) = 1$, the gg genotype always begets gg in the next generation, meaning the bottom row of every n-step transition matrix will be [0 0 1]. So, there is never a case where all 9 entries of $\mathbf{P}^n$ are positive.

b. For the matrix in Exercise 26, $\mathbf{P}^2 = \begin{bmatrix} .375 & .5 & .125 \\ .25 & .5 & .25 \\ .125 & .5 & .375 \end{bmatrix}$. Since all entries of this matrix are positive, by definition this is a regular Markov chain.

33.

a. Let $\boldsymbol{\pi} = [\pi_1 \;\; \pi_2 \;\; \pi_3]$ be the vector of steady-state probabilities. The formula $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$ gives three equations, but one is redundant with the other two. If we keep the first two equations, say, and add in the requirement $\sum \pi_j = 1$, we need to solve the system

$$.84\pi_1 + .08\pi_2 + .10\pi_3 = \pi_1,\; .06\pi_1 + .82\pi_2 + .04\pi_3 = \pi_2,\; \pi_1 + \pi_2 + \pi_3 = 1$$

Repeated substitution, or reduction of the matrix representation of this set of equations, gives the unique solution $\pi_1 = .3681$, $\pi_2 = .2153$, and $\pi_3 = .4167$.

b. China Mobile is state 3, so the long-run proportion of all Chinese cell phone users that subscribe to China Mobile is given by $\pi_3 = .4167$.

c. China Telecom is state 1. So, the average number of transitions (contract changes) between successive "visits" to state 1 is $1/\pi_1 = 1/.3681 = 2.72$.

35.

a. From the information provided, $\mathbf{P} = \begin{bmatrix} .7 & .2 & .1 \\ 0 & .6 & .4 \\ .8 & 0 & .2 \end{bmatrix}$, with 1 = full, 2 = part, and 3 = broken.

b. By direct computation, $\mathbf{P}^2 = \begin{bmatrix} .57 & .26 & .17 \\ .32 & .36 & .32 \\ .72 & .16 & .12 \end{bmatrix}$. Since all entries are positive, **P** represents a regular Markov chain.

c. Let $\boldsymbol{\pi} = [\pi_1 \;\; \pi_2 \;\; \pi_3]$ be the vector of steady-state probabilities. The formula $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$ gives three equations, but one is redundant with the other two. If we keep the first two equations and add in the requirement $\sum \pi_j = 1$, we need to solve the system

$$.7\pi_1 + .8\pi_2 = \pi_1,\; .2\pi_1 + .6\pi_2 = \pi_2,\;\; \pi_1 + \pi_2 + \pi_3 = 1$$

Repeated substitution, or reduction of the matrix representation of this set of equations, gives the unique solution $\pi_1 = 8/15$, $\pi_2 = 4/15$, and $\pi_3 = 3/15 = 1/5$.

d. The long-run proportion of days the machine is fully operational (state 1) is $\pi_1 = 8/15$, or 53.33%.

e. The average number of transitions (days) between breakdowns (visits to state 3) is $1/\pi_3 = 5$ days.

37. For ease of notation, let $\mathbf{M} = \mathbf{P}^n$, and assume that all entries of $\mathbf{M}$ are positive. If we write $\mathbf{P}^{n+1} = \mathbf{P}\mathbf{P}^n = \mathbf{P}\mathbf{M}$, then the (i, j) entry of $\mathbf{P}^{n+1}$ is given by the product $p_{i1}m_{1j} + p_{i2}m_{2j} + \cdots + p_{is}m_{sj}$, assuming $\mathbf{P}$ (and thus all its powers) is an $s \times s$ matrix. The m's are all positive; the p's are the ith row of $\mathbf{P}$, which must sum to 1, and so at least one of them must also be positive. Therefore, $p_{i1}m_{1j} + p_{i2}m_{2j} + \cdots + p_{is}m_{sj}$ is positive, and this is true for every coordinate (i, j) in $\mathbf{P}^{n+1}$. Therefore, all entries of $\mathbf{P}^{n+1}$ are positive.
Replace $\mathbf{M} = \mathbf{P}^n$ with $\mathbf{M} = \mathbf{P}^{n+1}$ in the preceding proof, and the same result holds for $\mathbf{P}^{n+2}$. Continuing in this fashion establishes the result for all higher powers.

39.

a. The one-step transition matrix for this chain is $\mathbf{P} = \begin{array}{c} 0 \\ 1 \end{array}\begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$. Assuming $0 < \alpha < 1$ and $0 < \beta < 1$, this is a regular chain. The steady-state probabilities satisfy $\boldsymbol{\pi}\,\mathbf{P} = \boldsymbol{\pi}$ and $\sum \pi_j = 1$, which are equivalent to the system $(1-\alpha)\pi_0 + \beta\pi_1 = \pi_0$, $\alpha\pi_0 + (1-\beta)\pi_1 = \pi_1$, $\pi_0 + \pi_1 = 1$. The first equation establishes that $\alpha\pi_0 = \beta\pi_1$; substitute into the second equation to get $\pi_0 + (\alpha/\beta)\pi_0 = 1 \Rightarrow \pi_0 = \dfrac{\beta}{\alpha+\beta}$. Back-substitute to get $\pi_1 = \dfrac{\alpha}{\alpha+\beta}$.

b. There are many, many cases to consider. $\alpha = \beta = 0 \Rightarrow \mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow$ the chain is constant; $\alpha = \beta = 1 \Rightarrow \mathbf{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow$ the chain alternates perfectly; $\alpha = 0, \beta = 1 \Rightarrow \mathbf{P} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \Rightarrow$ the chain arrives and stays forever in state 0; $\alpha = 1, \beta = 0 \Rightarrow \mathbf{P} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \Rightarrow$ the chain arrives and stays forever in state 1; $\alpha = 0$, $0 < \beta < 1 \Rightarrow$ the chain eventually arrives at state 0 and stays there forever; $0 < \alpha < 1$, $\beta = 0 \Rightarrow$ the chain eventually arrives at state 1 and stays there forever. Finally, $0 < \alpha < 1$ and $\beta = 1$ or $\alpha = 1$ and $0 < \beta < 1$ $\Rightarrow \mathbf{P}^2$ has all positive entries $\Rightarrow$ the chain is regular, and the steady-state probabilities in **a** still hold.

41.

a. As indicated in the hint, $P(00 \to 10) = P(0 \to 1) \cdot P(0 \to 0) = \alpha \cdot (1-\alpha)$. Using the same idea for the other fifteen possible transitions, we obtain the following transition matrix:

$$\mathbf{P} = \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array}\begin{bmatrix} (1-\alpha)^2 & \alpha(1-\alpha) & \alpha(1-\alpha) & \alpha^2 \\ \beta(1-\alpha) & (1-\alpha)(1-\beta) & \alpha\beta & \alpha(1-\beta) \\ \beta(1-\alpha) & \alpha\beta & (1-\alpha)(1-\beta) & \alpha(1-\beta) \\ \beta^2 & \beta(1-\beta) & \beta(1-\beta) & (1-\beta)^2 \end{bmatrix}$$

b. Determining the steady-state distribution looks daunting, until you realize we already know the steady-state probabilities for 0 and 1. Using independence again, we have for example $\pi_{00} = \pi_0 \cdot \pi_0 = \dfrac{\beta}{\alpha+\beta} \cdot \dfrac{\beta}{\alpha+\beta} = \dfrac{\beta^2}{(\alpha+\beta)^2}$. Similarly, $\pi_{01} = \dfrac{\beta}{\alpha+\beta} \cdot \dfrac{\alpha}{\alpha+\beta} = \dfrac{\alpha\beta}{(\alpha+\beta)^2}$. $\pi_{10} = \dfrac{\alpha\beta}{(\alpha+\beta)^2}$, and $\pi_{11} = \dfrac{\alpha^2}{(\alpha+\beta)^2}$. It's straightforward to show these satisfy $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$ and $\sum \pi = 1$.

c. A connection is feasible only if both ends are active (1 and 1), so the long-run proportion of time the system is usable is just $\pi_{11} = \alpha^2/(\alpha + \beta)^2$.

43.

a. The one- and two-step transition matrices are $\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 \end{bmatrix}$ and $\mathbf{P}^2 = \begin{bmatrix} \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{7}{9} & 0 & \frac{2}{9} \\ \frac{2}{9} & 0 & \frac{7}{9} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} \end{bmatrix}$. Just from these two matrices, we see that either $P^{(1)}(i \to j) > 0$ or $P^{(2)}(i \to j) > 0$ for all 16 pairs of states (i, j). So, the Ehrenfest chain is an irreducible chain. However, as discussed in Exercise 32, the alternating pattern of 0's and non-zero entries seen in these two matrices persists forever; that is, half of the entries of $\mathbf{P}^n$ are 0 no matter the power of n. Therefore, the Ehrenfest chain is <u>not</u> a regular Markov chain.

b. By two opposite ball exchanges (left-to-right then right-to-left, or vice versa), every state can return to itself in an even number of steps. In fact, $P^{(n)}(i \to i) > 0$ for $n = 2, 4, 6, 8, 10, \ldots$. Therefore, the period of every state is the greatest common divisor of these integers, i.e. 2.

c. By direct computation, [1/8 3/8 3/8 1/8]**P** = [3/8(1/3) (1/8)(1)+(3/8)(2/3) (3/8)(2/3)+(1/8)(1) 3/8(1/3)] = [1/8 3/8 3/8 1/8]. So, by definition, [1/8 3/8 3/8 1/8] is a stationary distribution for the Ehrenfest chain with $m = 3$.

Section 6.5

45.

a. For states $j = 1, 2, 3$, $P(j \to j + 1) = .75$ (go forward) and $P(j \to j) = .25$ (repeat the semester). On the other hand, state 4 is an absorbing state: once you graduate, you've always graduated! The corresponding one-step transition matrix is

$$\mathbf{P} = \begin{bmatrix} .25 & .75 & 0 & 0 \\ 0 & .25 & .75 & 0 \\ 0 & 0 & .25 & .75 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

b. A student graduates within 3 semesters provided that, after 3 transitions (one-semester time intervals), she has arrived in state 4: $P^{(3)}(1 \to 4)$ = the (1,4) entry of $\mathbf{P}^3 = .4219$. This is also $(.75)^3$, of course, since graduating in 3 semesters means she never repeated a semester.
Since graduation is an absorbing state, the chance of arriving there within 4 semesters is simply the probability of being graduated in 4 time steps — that is, the possibilities of graduating in exactly 3 <u>or</u> exactly 4 semesters are captured in $P^{(4)}(1 \to 4)$. The answer is the (1,4) entry of $\mathbf{P}^4 = .7383$.
Similarly, P(graduate within 5 semesters) = $P^{(5)}(1 \to 4)$ = the (1,4) entry of $\mathbf{P}^5 = .8965$.

c. The goal is to find the mean time to absorption into state 4 (graduation). The sub-matrix **Q** of non-absorbing states is $\mathbf{Q} = \begin{array}{c} 1 \\ 2 \\ 3 \end{array}\begin{bmatrix} .25 & .75 & 0 \\ 0 & .25 & .75 \\ 0 & 0 & .25 \end{bmatrix}$; by the MTTA Theorem, the mean times to absorption are given by $\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \begin{bmatrix} .75 & -.75 & 0 \\ 0 & .75 & -.75 \\ 0 & 0 & .75 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8/3 \\ 4/3 \end{bmatrix}$. So, $\mu_1 = 4$ semesters.

d. Since state 4 is the only absorbing state, the probability of eventual absorption into state 4 (graduation) must equal 1. This isn't very realistic: not all students complete the program, because obviously some drop out. But dropping out of the program is not part of this model (see the next exercise).

47.

a. Absorbing states satisfy $P(a \to a) = 1$, meaning there is a 1 in the main diagonal entry of that row. For this matrix, the absorbing states are 4 and 5. (Notice that state 1 is not absorbing — the 1 in the matrix is not along the main diagonal.)

b. $P(T_1 \le k)$ is the probability the chain gets absorbed within k steps, starting from state 1. For any fixed k, this can be calculated by $P^{(k)}(1 \to 4 \text{ or } 5) = P^{(k)}(1 \to 4) + P^{(k)}(1 \to 5)$ = the sum of the (1,4) and (1,5) entries of $\mathbf{P}^k$. These values, for $k = 1, 2, \ldots, 10$, appear in the table below.

k	1	2	3	4	5	6	7	8	9	10
$P(T_1 \le k)$	0	.46	.7108	.8302	.9089	.9474	.9713	.9837	.9910	.9949

c. Similar to Example 6.26, we can determine the (approximate) pmf of T_1 by sequentially subtracting the cumulative probabilities in part **b**; e.g., $P(T_1 = 3) = P(T_1 \le 3) - P(T_1 \le 2) = .7108 - .46 = .2508$. This approximate pmf appears below. From this table $E(T_1) \approx 1(0) + 2(.46) + \ldots + 10(.0039) = 3.1457$.

k	1	2	3	4	5	6	7	8	9	10
$P(T_1 = k)$	0	.46	.2508	.1194	.0787	.0385	.0239	.0124	.0073	.0039

d. The sub-matrix $\mathbf{Q}$ for the non-absorbing states is $\mathbf{Q} = \begin{bmatrix} 0 & 1 & 0 \\ .21 & 0 & .33 \\ .09 & .15 & 0 \end{bmatrix}$. The mean times to absorption from states 1, 2, 3, are given by the MTTA Theorem: $\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = [3.2084 \;\; 2.2084 \;\; 1.6200]^{\mathrm{T}}$. In particular, $\mu_1 = 3.2084$ transitions.

e. The sub-matrix $\mathbf{R}$ of the canonical form of $\mathbf{P}$ is $\mathbf{R} = \begin{bmatrix} 0 & 0 \\ .05 & .41 \\ .67 & .09 \end{bmatrix}$. The probabilities of eventual absorption are given by $\boldsymbol{\Pi} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R} = \begin{bmatrix} .3814 & .6186 \\ .3814 & .6186 \\ .7615 & .2385 \end{bmatrix}$. So, starting from the homepage (state 1), the probability of eventually reading Chapter 1 (state 4) is .3814, while the chance of exiting without reading Chapter 1 (state 5) is .6186.

49.

a. If the current streak is j = 0, 1, 2, or 3, there's a .5 chance of advancing the streak to j + 1 (Heads) and a .5 chance of resetting to 0 (Tails). Once four heads in a row have been achieved, however, the game is won — Michelle stops flipping the coin, so state 4 is an absorbing state. The transition matrix is

$$\mathbf{P} = \begin{bmatrix} .5 & .5 & 0 & 0 & 0 \\ .5 & 0 & .5 & 0 & 0 \\ .5 & 0 & 0 & .5 & 0 \\ .5 & 0 & 0 & 0 & .5 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

b. For any positive integer k, $P(T_0 \le k) = P$(Michelle finishes within k flips) = P(Michelle arrives in state 4 within k flips) = $P^{(k)}(0 \to 4)$ = the (0,4) entry of $\mathbf{P}^k$. Clearly, $P(T_0 \le k) = 0$ for k = 1, 2, 3, since it takes at least 4 flips to get a streak of 4 heads. Using software, the successive values of $P(T_0 \le k)$ for k = 4, 5, …, 15 are .0625, .0938, .1250, .1563, .1875, .2168, .2451, .2725, .2988, .3242, .3487, .3723. Notice that the probability of getting a 4-head streak within the first 15 flips is still extremely low.

c. From the list above, $P(T_0 \le 10) = .2451$.

d. Similar to Example 6.26, we can determine the (approximate) pmf of T_1 by sequentially subtracting the cumulative probabilities in part **b**. Clearly $P(T_0 = k)$ for k = 1, 2, 3. For k = 4, 5, …, 15, the probabilities are .0625, .03125, .03125, .03125, .03125, .0293, .0283, .0273, .0264, .0254, .0245, and .0236. Based on these, the estimated mean time to absorption is $\mu \approx 4(.0625) + 5(.03125) + \ldots + 15(.0245) =$ 3.2531 flips; this is clearly a gross under-estimate, since (1) T_0 itself is at least 4 and (2) the probabilities being used only sum to .3723. An estimate of σ^2 is given by $\sum x^2p(x) - \mu^2 = 4^2(.0625) + 5^2(.03125) + \ldots + 15^2(.0245) - (3.2531)^2$, from which $\sigma \approx 3.9897$.

e. Apply the MTTA Theorem, with $\mathbf{Q}$ equal to the upper left 4 × 4 sub-matrix of $\mathbf{P}$: $\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = [30\ \ 28\ \ 24\ \ 16]^{\mathrm{T}}$. In particular, the mean number of flips required starting from 0 is 30 flips.

51. The sub-matrix $\mathbf{Q}$ of the non-absorbing states is $\mathbf{Q} = \begin{bmatrix} .6 & .2 \\ .3 & .4 \end{bmatrix}$. Apply the MTTA Theorem: $\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = [4.44\ \ 3.89]^{\mathrm{T}}$. So, μ_{coop} = 4.44 turns, while μ_{comp} = 3.89 turns. In particular, the negotiations take longer, on average, when Negotiator A leads with a cooperative strategy. (Of course, stopping doesn't mean the negotiations were successful!)

53.

a. Applying the rules of Gambler's Ruin, $\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1-p & 0 & p & 0 & 0 \\ 0 & 1-p & 0 & p & 0 \\ 0 & 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

b. In Gambler's Ruin with \$4 total stake, both \$0 and \$4 are absorbing states. So, the sub-matrix $\mathbf{Q}$ for the non-absorbing states is the 3 × 3 central matrix of $\mathbf{P}$. Apply the MTTA Theorem:

$$\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \begin{bmatrix} 1 & -p & 0 \\ -1+p & 1 & -p \\ 0 & -1+p & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2p^2 - 2p + 1}\begin{bmatrix} 2p^2 + 1 \\ 2 \\ 2p^2 - 4p + 3 \end{bmatrix}.$$

So if Allan starts with \$1, the expected number of plays until the game ends (aka the mean time to absorption) is $\dfrac{2p^2+1}{2p^2-2p+1}$; starting with \$2, it's $\dfrac{2}{2p^2-2p+1}$; starting with \$3, it's $\dfrac{2p^2-4p+3}{2p^2-2p+1}$. (Notice these equal 1, 2, 3 when $p = 0$ and 3, 2, 1 when $p = 1$, which they should.)

c. The sub-matrix **R** of the canonical form of **P** corresponds to the \$1/\$2/\$3 rows and the \$0/\$4 columns: $\mathbf{R} = \begin{bmatrix} 1-p & 0 \\ 0 & 0 \\ 0 & p \end{bmatrix}$. The probabilities of eventual absorption are given by $\mathbf{\Pi} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}$, and the probability that Allan <u>wins</u> appears in the second column (state \$4) of this 3 × 2 matrix. The results are $\dfrac{p^3}{2p^2-2p+1}$ starting with \$1, $\dfrac{p^2}{2p^2-2p+1}$ starting with \$2, and $\dfrac{p^3-p^2+p}{2p^2-2p+1}$ starting with \$3. (Notice these all equal 0 when $p = 0$ and equal 1 when $p = 1$, as they should.)

55. To determine the average number of generations required to go from state L to state H, modify the transition matrix so that H is an absorbing state:

$$\tilde{\mathbf{P}} = \begin{array}{c} \text{L} \\ \text{M} \\ \text{H} \end{array}\begin{bmatrix} .5288 & .2096 & .2616 \\ .3688 & .2530 & .3782 \\ 0 & 0 & 1 \end{bmatrix}$$

The sub-matrix **Q** of non-absorbing states is $\mathbf{Q} = \begin{bmatrix} .5288 & .2096 \\ .3688 & .2530 \end{bmatrix}$, from which the MTTA Theorem provides $\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = [3.4825 \;\; 3.0580]^{\mathrm{T}}$. In particular, the expected number of generations to reach state H starting from L (the first state) is 3.4825 generations.

57.

a. As described in the exercise, the chain will eventually be absorbed from i into a if (1) or (2) occurs. (1): The chain transitions immediately from i to a, which occurs with probability $P(i \to a)$. (2): The chain transitions to some other non-absorbing state j, which occurs with probability $P(i \to j)$, and then from j is eventually absorbed into a. By definition, the latter probability is $\pi(j \to a)$. Outcome (2) actually covers <u>all</u> non-absorbing states j, including i itself (e.g., $i \to i \to a$), so we must sum across all such possibilities. Put together, the probability of eventual absorption from i to a, i.e. $\pi(i \to a)$, is $\pi(i \to a) = P(i \to a) + \sum_{j \in A'} P(i \to j)\pi(j \to a)$, where A' denotes the set of all non-absorbing states.

b. Use the canonical form (6.8) of **P**, so the non-absorbing states are 1, …, r. The sum above is then indexed $j = 1$ to r, and it represents the dot product of the ith row of **P** from column 1 through column r — aka the ith row of **Q** — with the column of **Π** corresponding to state a. Fix state a, and stack the preceding equation for $i = 1, 2, \ldots, r$:

$$\pi(1 \to a) = P(1 \to a) + [p_{11} \cdots p_{1r}] \cdot [\pi(1 \to a) \cdots \pi(r \to a)]^{\mathrm{T}}$$
$$\pi(2 \to a) = P(2 \to a) + [p_{21} \cdots p_{2r}] \cdot [\pi(1 \to a) \cdots \pi(r \to a)]^{\mathrm{T}}$$

$$\pi(r \to a) = P(r \to a) + [p_{r1} \cdots p_{rr}] \cdot [\pi(1 \to a) \cdots \pi(r \to a)]^{\mathrm{T}}$$

The left-most stack is column a of **Π**. The next stack is column a of **P**, but only rows 1 through r; i.e., it's column a of **R**. The dot product stack, as referenced before, is the matrix **Q** multiplied by column a of **Π**.

If we then consider these stacks for the various absorbing states a — i.e., put the aforementioned columns side-by-side — we have $\mathbf{\Pi} = \mathbf{R} + \mathbf{Q\Pi}$.

Assuming invertibility, we then have $\mathbf{R} = \mathbf{\Pi} - \mathbf{Q\Pi} = (\mathbf{I} - \mathbf{Q})\mathbf{\Pi} \Rightarrow \mathbf{\Pi} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}$, qed.

Section 6.6

59.

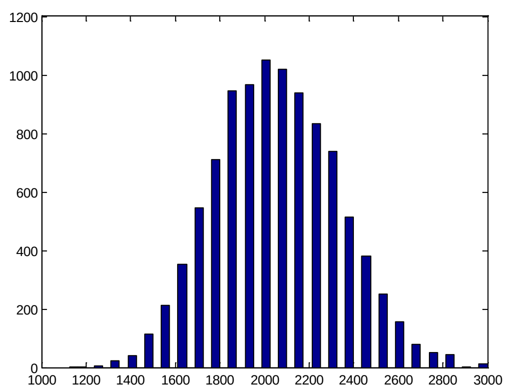
a. The programs below perform 10,000 iterations of the following process: begin with $Y = 150$ for the first day, then use the Markov chain to move into state 1, 2, or 3 for each of the next 19 days. The new state of the chain determines whether 150, 75, or 0 units are produced. The vector Y holds the values of the rv.

```
P=[.7 .2 .1;0 .6 .4; .8 0 .2];
Y=zeros(10000,1);
for iter=1:10000
Y(iter)=150; state=1;
for i=1:19
    newstate=randsample(1:3,1,

true,P(state,:));
    if newstate==1
        Y(iter)=Y(iter)+150;
    elseif newstate==2
        Y(iter)=Y(iter)+75;
    end
    state=newstate;
end
end
```

```
P<-matrix(c(.7,.2,.1,0,.6,.4,
.8,0,.2),nrow=3,ncol=3,byrow=TRUE)
Y<-NULL
for (iter in 1:10000){
Y[iter]<-150; state<-1
for (i in 1:19){
    newstate=sample(1:3,1,
                  TRUE,P[state,])
    if (newstate==1){
        Y[iter]=Y[iter]+150}
    else if (newstate==2){
        Y[iter]=Y[iter]+75
    }
    state<-newstate
}
}
```

b. Notice that the distribution of Y is somewhat bell-shaped, being a type of sum; however, the distribution has gaps because Y can only attain multiples of 75.



c. For one simulation run, $\bar{y} = 2074.4$ and $s_y = 274.82$. From Chapter 5, a 95% confidence interval for $E(Y)$ is given by $\bar{y} \pm 1.96\frac{s_y}{\sqrt{n}} = 2074.4 \pm 1.96\frac{274.82}{\sqrt{10,000}}$ = (2069.0, 2079.8).

d. For one simulation run, `sum(Y>=2000)` returned 6089, so $\hat{p} = 6089/10,000 = .6089$. From Chapter 5, a 95% confidence interval for $P(Y \geq 2000)$ is $\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .6089 \pm 1.96\sqrt{\frac{.6089(.3911)}{10,000}}$ = (.5993, .6185).

61. For simplicity, we'll use units equal to millions of dollars. So, e.g., $X_0 = 10$.

a. There are only two non-zero transition probabilities. For any possible x, $P(X_{n+1} = 10 \mid X_n = x) = .4$ and $P(X_{n+1} = 2x \mid X_n = x) = .6$, corresponding to someone winning or nobody winning this week's lottery.

b. The accompanying program simulates both the rv M from part **b** and Y from part **c**. The indicator element `winner` is 1 or 0 with probability .4 and .6, respectively. The line creating `M(i)` or `M[i]` replaces M with its current value or the current prize value, whichever is larger (that's ensures we detect the maximum).

For one simulation run, the sample mean and standard deviation of M were 2395 and 54,430, respectively. That is, on the average, the largest lottery prize available over a year will be somewhere around $2.395 billion, and that maximum value has a standard deviation of around $54.430 billion! (Answers will vary, of course; it's worth noting that because the variable M is so volatile, two runs of 10,000+ simulated values can still return wildly different results.)

```
M=zeros(10000,1); Y=zeros(10000,1);
for i=1:10000
    prize=10;
    for week=1:52
        winner=(rand<.4);
        if winner==1
            Y(i)=Y(i)+prize;
            prize=10;
        else
            prize=2*prize;
        end
        M(i)=max(M(i),prize);
    end
end
```

```
M<-rep(0,10000); Y<-rep(0,10000)
for (i in 1:10000){
    prize<-10
    for (week in 1:52){
        winner<-(runif(1)<.4)
        if (winner==TRUE){
            Y[i]<-Y[i]+prize
            prize=10}
        else{
            prize=2*prize
        }
        M[i]=max(M[i],prize)
    }
}
```

c. For one simulation run, the sample mean and standard deviation of Y were 37806 and 2,816,506, respectively. (The astronomically high standard deviation comes from the fact that, once every 10,000 years it's not impossible to see the lottery reach $670 million or more.) The resulting 95% confidence interval for $E(Y)$ is $\bar{y} \pm 1.96\frac{s_y}{\sqrt{n}} = 37806 \pm 1.96\frac{2816506}{\sqrt{10,000}}$ = (–17938, 93010). Since Y can't be negative, the realistic interval is from $0 billion to $93 billion.

d. Replace .4 with .7 in the fifth line of code. One simulation run returned sample mean and standard deviations of 623.98 and 364.82, respectively, for Y. A 95% confidence interval for $E(Y)$ under this new setting is $623.98 \pm 1.96\frac{364.82}{\sqrt{10,000}}$ = ($616.8 million, $631.1 million).

Although the CI from part **c** suggests there's a remote chance nobody will win the lottery, the potential of having to pay out money in the billions is clearly of great concern. So, comparing that CI to the one in part **d**, the lottery commission should make it <u>easier</u> for someone to win each week. (In fact, it appears they should just guarantee a winner each week: that would require a definite annual payout of $520 million, which is less than what they'd expect to pay with a .7 weekly win probability.)

63.

a. Matlab and R simulations are shown below; code to create **P** from Exercise 12 is not shown (to save space). The initial conditions `state=1` and `Y(i)=500` correspond to the first of the driver's 10 years being in state 1 (lowest-risk driver), which carries a premium of $500. Years 2 through 10 are then generated using the Markov chain.

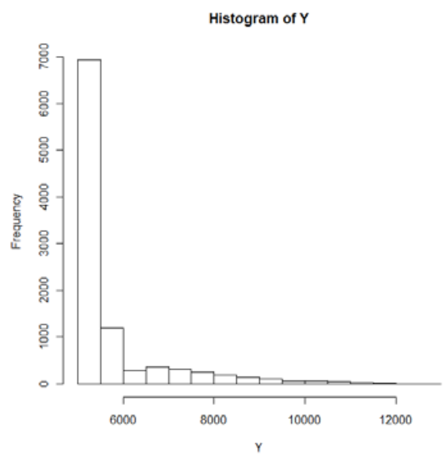
```
P=not shown;
premiums=[500 600 1000 1500];
Y=zeros(10000,1);
for i=1:10000
    state=1;
    Y(i)=500;
    for year=2:10
        newstate=randsample(1:4,1,
                        true,P(state,:));
        Y(i)=Y(i)+premiums(newstate);
        state=newstate;
    end
end
```

```
P<-not shown
premiums<-c(500,600,1000,1500)
Y<-NULL
for (i in 1:10000){
    state<-1
    Y[i]<-500
    for (year in 2:10){
        newstate=sample(4,1,
                        TRUE,P[state,])
        Y[i]<-Y[i]+premiums[newstate]
        state<-newstate
    }
}
```

For one simulation run, the sample mean and standard deviation of Y_1 were \$5637.98 and \$1197.23, respectively. From Chapter 5, a 95% confidence interval for $E(Y_1)$ is given by

$$\bar{y} \pm 1.96\frac{s_y}{\sqrt{n}} = 5637.98 \pm 1.96\frac{1197.23}{\sqrt{10,000}} =$$

(\$5614.51, \$5661.45).



b. In the code from part **a**, replace the initial state with 3 and the initial premium with 1000 in the first two lines after the "for" declaration. For one simulation run, the sample mean and standard deviation of Y_3 were \$7219.85 and \$1991.78, respectively. A 95% confidence interval for $E(Y_3)$ is given by

$$\bar{y} \pm 1.96\frac{s_y}{\sqrt{n}} = 7219.85 \pm 1.96\frac{1991.78}{\sqrt{10,000}} = (\$7180.81, \$7258.89).$$

65.

a. The programs below were written as functions, rather than simple scripts for use in parts **b** and **c**. You could also use the `cumsum` command on a vector of random ±1 values to simulate the walk.

```
function X=simpwalk(n)
X=zeros(n,1);
X(1)=randsample([-1 1],1);
for i=2:n
    s=randsample([-1 1],1);
    X(i)=X(i-1)+s;
end
```

```
simpwalk<-function(n){
X<-NULL
X[1]=sample(c(-1,1),1)
for (i in 2:n){
    s<-sample(c(-1,1),1)
    X[i]=X[i-1]+s }
return(X) }
```

b. The binary vector `Home` indicates whether any of the walk's 100 steps returned it to 0 (the origin). For one simulation run, `sum(Home)` returned 9224, so the estimated probability the walk ever returns to its origins within the first hundred steps is 9224/10,000 = .9224.

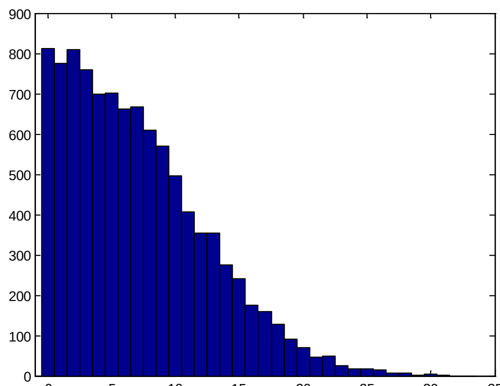
```
Home=zeros(10000,1);
for i=1:10000
    X=simpwalk(100);
    if any(X==0)
        Home(i)=1;
    end
end
```

```
Home<-rep(0,10000)
for (i in 1:10000){
    X<-simpwalk(100)
    if (any(X==0)){
        Home[i]<-1
    }
}
```

c. The code `sum((X==0))` counts the number of 0's that appear in the walk, i.e. the number of returns home during the first 100 steps. For one simulation run, the sample mean and standard deviation of the simulated values of R_0 were 7.0636 and 5.4247, respectively. A 95% confidence interval for $E(R_0)$ is therefore given by $7.0636 \pm 1.96\dfrac{5.4247}{\sqrt{10,000}}$ = (6.96, 7.17) returns.

```
R0=zeros(10000,1);
for i=1:10000
    X=simpwalk(100);
    R0(i)=sum((X==0));
end
```

```
R0<-NULL
for (i in 1:10000){
     X<-simpwalk(100)
     R0[i]<-sum((X==0))
}
```



Supplementary Exercises

67.

a. From each chamber, the hamster can only enter the two adjacent wedges, and each has probability .5. The resulting one-step transition matrix is

$$\mathbf{P} = \begin{bmatrix} 0 & .5 & 0 & 0 & 0 & .5 \\ .5 & 0 & .5 & 0 & 0 & 0 \\ 0 & .5 & 0 & .5 & 0 & 0 \\ 0 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & .5 \\ .5 & 0 & 0 & 0 & .5 & 0 \end{bmatrix}.$$

b. <u>No</u>. Starting from chamber 1, say, the hamster can only reach even-numbered chambers in an odd number of steps and odd-numbered chambers in an even number of steps. Therefore, whatever the power n may be, half the entries of $\mathbf{P}^n$ will be 0. Thus, there is no power n for which all 36 entries of $\mathbf{P}^n$ are positive, so this is not a regular Markov chain. (This is, however, an irreducible chain; rather than being regular, it has period 2.)

c. Given the design of the habitat, it seems logical that in the long run the hamster will spend an equal amount of time in each of the six chambers, suggesting $\boldsymbol{\pi} = [\tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6}]$. Verify:

$$\boldsymbol{\pi}\mathbf{P} = [\tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6}] \begin{bmatrix} 0 & .5 & 0 & 0 & 0 & .5 \\ .5 & 0 & .5 & 0 & 0 & 0 \\ 0 & .5 & 0 & .5 & 0 & 0 \\ 0 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & .5 \\ .5 & 0 & 0 & 0 & .5 & 0 \end{bmatrix} = [\tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6}] = \boldsymbol{\pi}.$$

d. Since the chain is irreducible, we may still interpret the entries of $\boldsymbol{\pi}$ as the reciprocals of the mean times to return. So, starting in chamber 3, the expected number of transitions until the hamster returns to chamber 3 is $1/\pi_3 = 1/(1/6) = 6$.

e. Since the chain is irreducible, we may apply the same technique used in the previous section: treat chamber 6 as an absorbing state. The resulting matrices are

$$\tilde{\mathbf{P}} = \begin{bmatrix} 0 & .5 & 0 & 0 & 0 & .5 \\ .5 & 0 & .5 & 0 & 0 & 0 \\ 0 & .5 & 0 & .5 & 0 & 0 \\ 0 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{Q} = \begin{bmatrix} 0 & .5 & 0 & 0 & 0 \\ .5 & 0 & .5 & 0 & 0 \\ 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & .5 & 0 \end{bmatrix}, \boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \begin{bmatrix} 5 \\ 8 \\ 9 \\ 8 \\ 5 \end{bmatrix}.$$

In particular, the mean time to "absorption" starting from state 3 is 9 transitions.

69.

a. With the specified modification, $\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$.

b. With the aid of software, determine progressively higher powers of **P**. It takes a while, but we eventually find that all 36 entries of $\mathbf{P}^6$ are positive. Thus, by definition, we have a regular chain.

c. Use software or row reduction to solve $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$ with the added constraint $\sum \pi_j = 1$. The resulting vector of probabilities is $\boldsymbol{\pi} = [1/12 \;\; 1/4 \;\; 1/4 \;\; 1/6 \;\; 1/6 \;\; 1/12]$.

d. The long-run proportion of time that Lucas spends in room 2 is $\pi_2 = 1/4$, or 25%.

e. The average number of room transitions between successive visits to room 1 is $1/\pi_1 = 12$.

71.

a. From the description provided, $P(j \to j-1) = .3$ and $P(j \to j) = .7$ for $j = 5, 4, 3, 2, 1$. Also, $P(0 \to 5) = 1$, signifying re-order 5 cars. Put it together: $\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ .3 & .7 & 0 & 0 & 0 & 0 \\ 0 & .3 & .7 & 0 & 0 & 0 \\ 0 & 0 & .3 & .7 & 0 & 0 \\ 0 & 0 & 0 & .3 & .7 & 0 \\ 0 & 0 & 0 & 0 & .3 & .7 \end{bmatrix}$.

b. Use software or row reduction to solve $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$ with the added constraint $\sum \pi_j = 1$. The resulting vector of probabilities is $\boldsymbol{\pi} = [3/53 \;\; 10/53 \;\; 10/53 \;\; 10/53 \;\; 10/53 \;\; 10/53]$. It should be no surprise that the steady-state probabilities for states 1, 2, 3, 4, 5 are equal.

c. The average time between successive orders, which correspond to visits to state 0, is $1/\pi_0 = 53/3 = 17.67$ weeks.

73.

a. From the matrix, $P^{(1)}(1 \to 3) = 0$, but $P^{(2)}(1 \to 3)$ = the (1,3) entry of $\mathbf{P}^2 = .04 > 0$. So, the minimum number time for a team to evolve from weak to strong is 2 seasons (though it's quite unlikely).

b. $P^{(4)}(3 \to 3)$ = the (3,3) entry of $\mathbf{P}^4 = .3613$.

c. Treat state 3 (strong team) as an absorbing state. Then we may apply the MTTA Theorem using the sub-matrix $\mathbf{Q} = \begin{bmatrix} .8 & .2 \\ .2 & .6 \end{bmatrix}$, from which $\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \begin{bmatrix} 15 \\ 10 \end{bmatrix}$. In particular, the expected time for a weak team (state 1) to become strong is $\mu_1 = 15$ seasons.

d. Now treat state 1 (weak) as absorbing, and apply the MTTA Theorem to $\mathbf{Q} = \begin{array}{c} 2 \\ 3 \end{array}\begin{bmatrix} .6 & .2 \\ .2 & .7 \end{bmatrix}$. Then

$\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \begin{bmatrix} 6.25 \\ 7.50 \end{bmatrix}$, from which μ_3 = 7.5 seasons.

75.

a. The condition matrix one year from now is $\mathbf{v}_1 = \mathbf{v}_0\mathbf{P}$; those for two and three years from now are given by $\mathbf{v}_2 = \mathbf{v}_0\mathbf{P}^2$ and $\mathbf{v}_1 = \mathbf{v}_0\mathbf{P}^3$, respectively. Using the transition matrix **P** and the initial condition matrix
$\mathbf{v}_0$ = [.3307 .2677 .2205 .1260 .0551] provided, we compute that
$\mathbf{v}_1$ = [.3168 .1812 .2761 .1413 .0846];
$\mathbf{v}_2$ = [.3035 .1266 .2880 .1643 .1176]; and
$\mathbf{v}_3$ = [.2908 .0918 .2770 .1843 .1561].

b. Since very bad road (state 5) is an absorbing state, we can apply the MTTA Theorem. The relevant sub-matrix **Q** consists of the first four rows and columns of **P**, from which

$\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \begin{bmatrix} 35.68 \\ 11.87 \\ 9.20 \\ 4.27 \end{bmatrix}$. That is, the expected number of years to reach very bad condition from very good condition (state 1) is 35.68 years; from good condition (2), 11.87 years; from fair condition (3), 9.20 years; and from bad (4), 4.27 years.

c. Let T = the amount of time, in years, until the randomly-selected road segment becomes very bad. Apply the law of total expectation, using the values in part **b** and the initial condition matrix $\mathbf{v}_0$:

$$\begin{aligned} E(T) &= E(T \mid X_0 = 1)P(X_0 = 1) + E(T \mid X_0 = 2)P(X_0 = 2) + \cdots + E(T \mid X_0 = 5)P(X_0 = 5) \\ &= 35.68(.3307) + 11.87(.2677) + 9.20(.2205) + 4.27(.1260) + 0(.0551) \\ &= 17.54 \text{ years} \end{aligned}$$

77.

a. Assemble **P** in its canonical form using the pieces provided and (6.8):

$$\mathbf{P} = \begin{array}{r} 0 \\ 1 \\ 2 \\ 3 \\ \text{pd} \\ \text{tbr} \end{array}\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & .959 & 0 & .041 & 0 \\ 0 & 0 & 0 & .987 & .013 & 0 \\ 0 & 0 & 0 & 0 & .804 & .196 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

b.

$$\begin{array}{ccccccccc} 0 & \rightarrow & 1 & \rightarrow & 2 & \rightarrow & 3 & \rightarrow & \text{tbr} \\ & & & \searrow & \downarrow & \swarrow & & & \\ & & & & \text{pd} & & & & \end{array}$$

c. Using the matrix **Q** provided, $\boldsymbol{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}$ = [3.8192 2.8192 1.8970 1]$^{\mathrm{T}}$. So, in particular, the mean time to absorption starting from state 0 is μ_0 = 3.8192 weeks.

d. Using the matrices provided, $\mathbf{\Pi} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R} = \begin{bmatrix} .7451 & .1686 \\ .7451 & .1686 \\ .7342 & .1758 \\ .8040 & .1960 \end{bmatrix}$. In particular, $\pi(0 \to \text{pd}) = .7451$.

e. $P(0 \to 1) = 1$, which suggests that all payments are at least 1 week late. (Perhaps the researchers only investigated payments that were at least 1 week late?) $P(3 \to \text{pd}) = .804$, meaning 80.4% of payments that are 3 weeks late do get paid, rather than falling into "to be resolved" (which means at least a month late). Especially compared to $P(1 \to \text{pd})$ and $P(2 \to \text{pd})$, this seems to suggest that clients regularly wait three weeks and then pay what's owed. Very peculiar!

79.

a. For the first relay, $P(0 \to 1) = P(1 \to 0) = .02$, from which $P(0 \to 0) = P(1 \to 1) = .98$. So, the transition matrix for the first relay is $\mathbf{P}_1 = \begin{bmatrix} .98 & .02 \\ .02 & .98 \end{bmatrix}$. Similarly, the other four matrices are $\mathbf{P}_2 = \begin{bmatrix} .97 & .03 \\ .03 & .97 \end{bmatrix}$, $\mathbf{P}_3 = \mathbf{P}_1$, and $\mathbf{P}_4 = \mathbf{P}_5 = \begin{bmatrix} .99 & .01 \\ .01 & .99 \end{bmatrix}$.

b. Analogous to the preceding exercise, the transition matrix from X_0 to X_5 is $\mathbf{P}_1\mathbf{P}_2\mathbf{P}_3\mathbf{P}_4\mathbf{P}_5$, which is directly computed to equal $\begin{bmatrix} .916 & .084 \\ .084 & .916 \end{bmatrix}$. In particular, $P(X_5 = 0 \mid X_0 = 0) = .916$.

81.

a. Let $\mathbf{v}_0$ denote the 1×10 vector of cell counts for the ten categories in 2005. Since one transition period is 8 years (e.g., 2005 to 2013), the distribution in 2013 is equal to $\mathbf{v}_0\mathbf{P}$. By direct computation, that's $\mathbf{v}_1$ = [3259 22,533 19,469 26,066 81,227 16,701 1511 211,486 171,820 56,916].

b. Similarly, the distribution in 2021, i.e. 2 transitions (16 years) after 1995, is $\mathbf{v}_0\mathbf{P}^2$, which equals $\mathbf{v}_2$ = [2683 24,119 21,980 27,015 86,100 15,117 1518 223,783 149,277 59,395]. The distribution in 2029, which is 3 transitions (24 years) after 1995, is given by $\mathbf{v}_0\mathbf{P}^3$, which equals $\mathbf{v}_3$ = [2261 25,213 24,221 27,526 89,397 13,926 1524 233,533 131,752 61,636].

For each category, the percent change from 1995 to 2029 is given by $\dfrac{\text{2029 count} - \text{1995 count}}{\text{1995 count}}$. For example, using the initial vector $\mathbf{v}_0$ and the last vector $\mathbf{v}_3$, the percent change for (1) airports is $\dfrac{2261 - 4047}{4047} = -.4413 \approx -44\%$. The complete vector of percent changes is [–44%, +24%, +46%, +12%, +20%, –26%, +1.3%, +19%, +34%, +13.4%].

c. Since all entries of **P** are positive, this is a regular Markov chain, and we may apply the Steady State Theorem. In particular, the land use distribution $8n$ years after 1995 is $\mathbf{v}_n = \mathbf{v}_0\mathbf{P}^n \to \mathbf{v}_0\mathbf{\Pi}$ as $n \to \infty$. Using software, the stationary distribution is $\boldsymbol{\pi}$ = [.0015 .0380 .0844 .0355 .1283 .0147 .0024 .4362 .1065 .1525]. This is the common row of $\mathbf{\Pi}$; multiplying $\mathbf{v}_0\mathbf{\Pi}$ returns the anticipated long-run distribution of these 610,988 cells:
$\mathbf{v}_0\mathbf{\Pi}$ = [920 23,202 51,593 21,697 78,402 8988 1445 266,505 65,073 93,160].

CHAPTER 7

Section 7.1

1.

a. Real time is continuous, so variation throughout the day is a **continuous-time** process. Temperature itself is a continuous variable, so this is a **continuous-space** process.

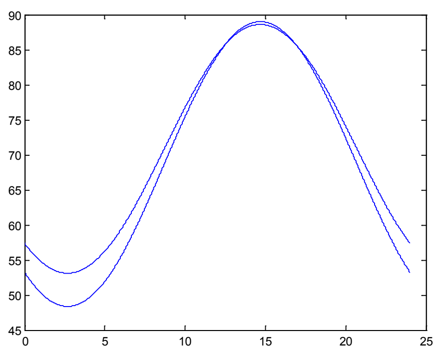
b. Again, real time is continuous, so variation throughout the day is a **continuous-time** process. But "number of customers" is a discrete variable, so this is a **discrete-space** process.

c. Now that only one measurement is taken each day, we have a **discrete-time** process (a realization of this process is a sequence of points, not a continuous curve). Temperature is still a continuous variable, however, so this is a **continuous-space** process.

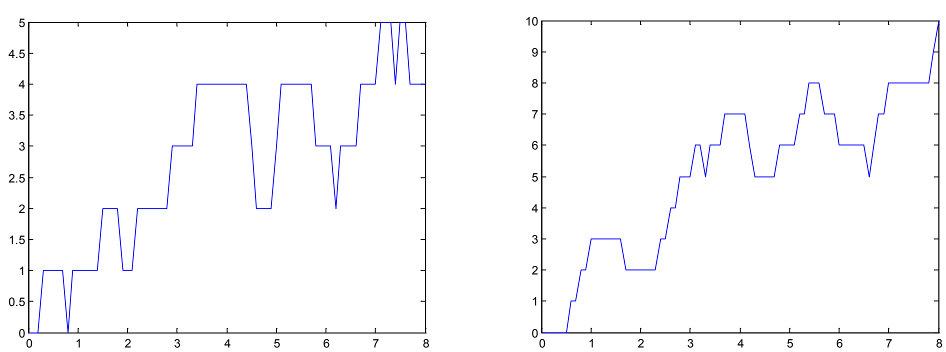
d. Again, with one measurement per day we have a **discrete-time** process. And, "total number of customers" is a discrete variable, so this is also a **discrete-space** process.

3.

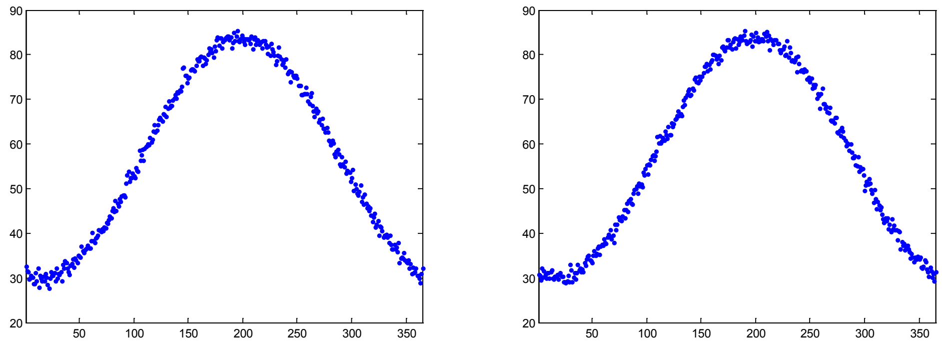
a. The sample functions below track temperatures from midnight ($t = 0$) to midnight the next day ($t = 24$).



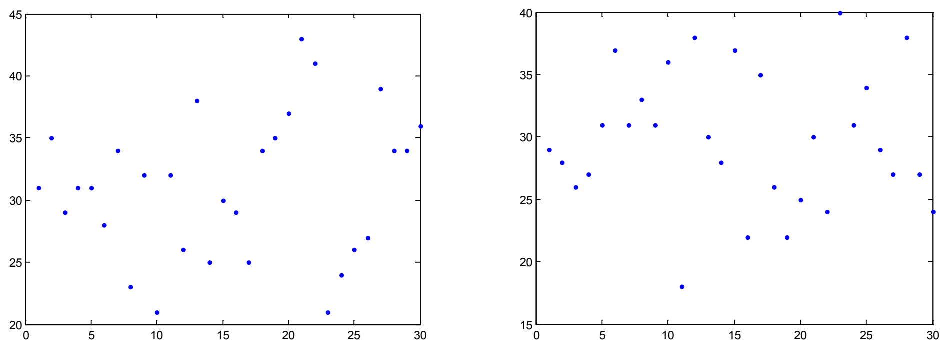
b. Two sample functions are presented separately below for ease of visualization. The diagonal connecting lines are not part of the function itself; in fact, the possible values are only non-negative integers.



c. Again, two sample functions are presented separately. Each consists of 365 daily high temperature observations in downtown Chicago, starting January 1 (n = 1) and ending December 31 (n = 365).

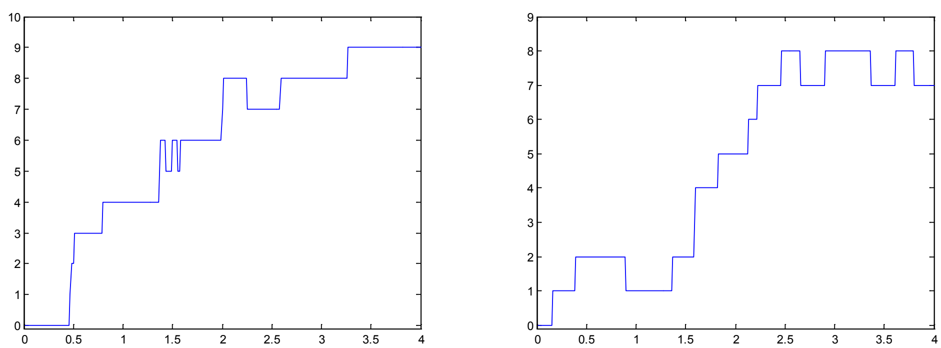


d. Both sample functions below show total number of customers for 30 consecutive days.

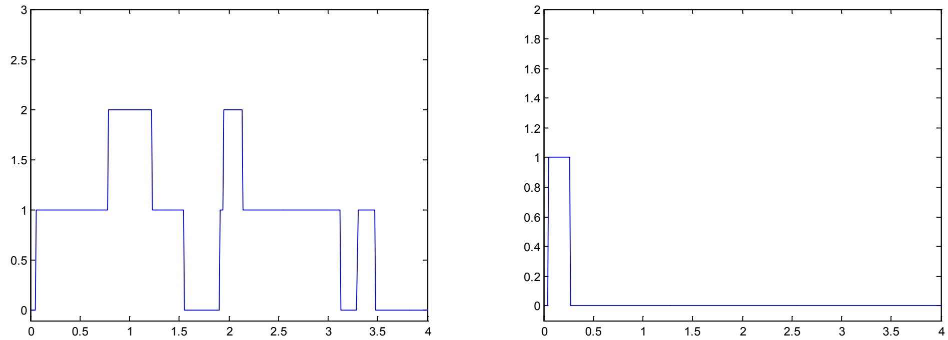


5. In the graphs below, the diagonal connecting lines are not part of the sample functions. The possible values of each sample function are non-negative integers.

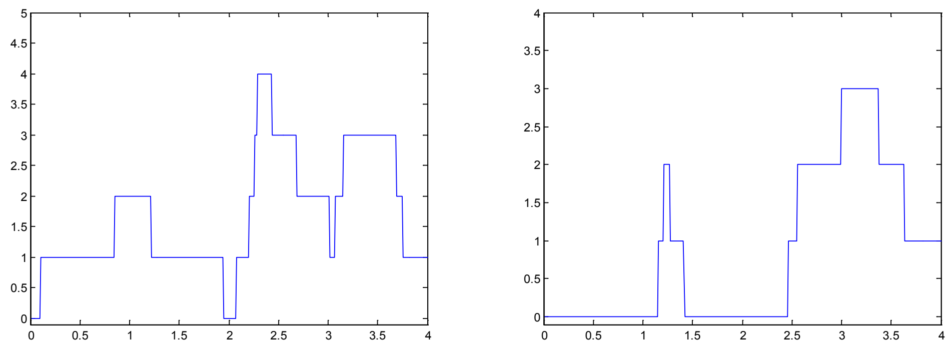
a. With login rate > logout rate, the overall trend is for the count of logged-in people to increase.



b. With logout rate > login rate, we see long stretches with nobody logged in.

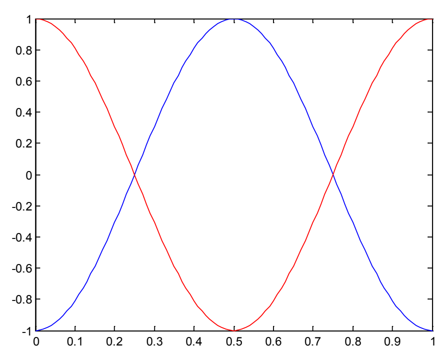


c. With logout rate = login rate, the process oscillates less systematically than in the previous two instances. In particular, there are fewer long stretches with $N(t) = 0$, and the count does not reach as high a number over the same time span as in part **a**.



7.

a. Substitute $f_0 = 1$ and $T = 1$ to obtain $x_b(t) = \cos(2\pi t + [b + 1]\pi)$ for $0 \le t \le 1$. For $b = 0$, this simplifies to $x_0(t) = \cos(2\pi t + \pi) = -\cos(2\pi t)$, while for $b = 1$ this becomes $x_1(t) = \cos(2\pi t + 2\pi) = \cos(2\pi t)$. Those two sample functions are graphed below ($x_0(t)$ in red, $x_1(t)$ in blue).



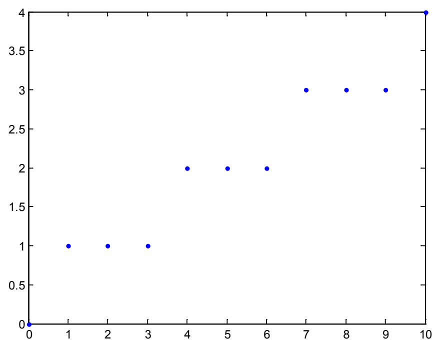
b. No: at time $t = .25$, $x_0(.25) = -\cos(\pi/2) = 0$ and $x_1(.25) = \cos(\pi/2) = 0$ (the first intersection on the graph). So, the two functions are identical, i.e. indistinguishable, there.

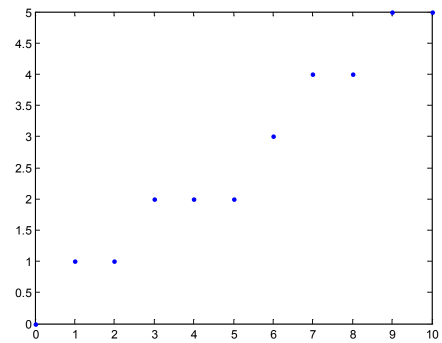
c. With probability .8, $X(t) = x_0(t)$ and with probability .2, $X(t) = x_1(t)$. Hence, at time $t = 0$, $X(0)$ equals -1 with probability .8 and +1 with probability .2. At time $t = .5$ (middle of the graph), these probabilities are reversed, and so $X(.5) = +1$ with probability .8 and -1 with probability .2.

9.

a. Since the set of all possible values of X_n across all n is {0, 1, 2, 3, …}, X_n is a discrete-space sequence.

b. Simulation in Matlab resulted in two sample functions for $(X_0, X_1, \ldots, X_{10})$: 01112223334 and 0112223445. These are graphed below.





c. For fixed n, X_n represents the number of wins in n independent and identical trials with success probability 18/38. Therefore, $X_n \sim \text{Bin}(n, 18/38)$.

Section 7.2

11.

a. It should be evident from Figure 7.2 that the sum of the four functions at any time t is zero, and hence the average is zero. Write $\beta = 2\pi f_0 t$. Since K is equally likely to 0, 1, 2, or 3,

$$\mu_X(t) = \sum_{k=0}^{3} \cos(\beta + \pi/4 + k\pi/2) \cdot \frac{1}{4} = \frac{\cos(\beta+\pi/4)+\cos(\beta+3\pi/4)+\cos(\beta+5\pi/4)+\cos(\beta+7\pi/4)}{4}$$

Use the angle addition formula to expand this into

$$\cos(\beta)[\cos(\pi/4)+\cos(3\pi/4)+\cos(5\pi/4)+\cos(7\pi/4)] - \sin(\beta)[\sin(\pi/4)+\sin(3\pi/4)+\sin(5\pi/4)+\sin(7\pi/4)]$$

The sum of the four terms inside each set of brackets is 0, whence $\mu_X(t) = \frac{0}{4} = 0$ for all t.

b. Since the mean of $X(t)$ is zero, its variance equals its mean-square value, $E[X^2(t)]$. This looks similar to the sum in part **a**:

$$\text{Var}(X(t)) = E[X^2(t)] = \sum_{k=0}^{3} [\cos(\beta + \pi/4 + k\pi/2)]^2 \cdot \frac{1}{4} = \frac{1}{4}\sum_{k=0}^{3} \cos^2(\beta + \pi/4 + k\pi/2)$$

Use the trig identity $\cos^2\theta = [1 + \cos(2\theta)]/2$:

$$\frac{1}{4}\sum_{k=0}^{3} \cos^2(\beta + \pi/4 + k\pi/2) = \frac{1}{4}\sum_{k=0}^{3} \frac{1+\cos(2\beta+\pi/2+k\pi)}{2} = \frac{1}{2} + \frac{1}{8}\sum_{k=0}^{3}\cos(2\beta+\pi/2+k\pi)$$

Similar to part **a**, angle addition shows that the summation is zero; therefore, $\text{Var}(X(t)) = 1/2$. This is surprising given Figure 7.2: it appears that there are wider and narrower distributions for $X(t)$. However, we have established that, despite varying ranges, $\text{Var}(X(t))$ is the same constant, 1/2, for all t.

13. Use the distributive properties of covariance, remembering that A is the only random variable:

$$C_{XX}(t,s) = \mathrm{Cov}(v_0 + A\cos(\omega_0 t + \theta_0), v_0 + A\cos(\omega_0 s + \theta_0)) = \mathrm{Cov}(A\cos(\omega_0 t + \theta_0), A\cos(\omega_0 s + \theta_0))$$
$$= \mathrm{Cov}(A, A)\cos(\omega_0 t + \theta_0)\cos(\omega_0 s + \theta_0)$$
$$= \mathrm{Var}(A)\cos(\omega_0 t + \theta_0)\cos(\omega_0 s + \theta_0)$$

Next, as noted in Examples 7.7-7.8, the mean function of $X(t)$ is $v_0 + E[A]\cos(\omega_0 t + \theta_0)$. Thus, the autocorrelation function is

$$R_{XX}(t,s) = C_{XX}(t,s) + \mu_X(t)\mu_X(s)$$
$$= \mathrm{Var}(A)\cos(\omega_0 t + \theta_0)\cos(\omega_0 s + \theta_0) + [v_0 + E[A]\cos(\omega_0 t + \theta_0)][v_0 + E[A]\cos(\omega_0 s + \theta_0)]$$
$$= v_0^2 + v_0 E[A][\cos(\omega_0 t + \theta_0) + \cos(\omega_0 s + \theta_0)] + \{E[A]\}^2\cos(\omega_0 t + \theta_0)\cos(\omega_0 s + \theta_0)$$
$$+ \mathrm{Var}(A)\cos(\omega_0 t + \theta_0)\cos(\omega_0 s + \theta_0)$$
$$= v_0^2 + v_0 E[A][\cos(\omega_0 t + \theta_0) + \cos(\omega_0 s + \theta_0)] + E[A^2]\cos(\omega_0 t + \theta_0)\cos(\omega_0 s + \theta_0)$$

This can also be derived directly from the definition of R_{XX}. In the last step above, we've combined the last two terms using the variance shortcut formula.

15.

a. $\sigma_N^2(t) = C_{NN}(t,t) = e^{-|t-t|} = e^0 = 1$.

b. The covariance between $N(t)$ and $N(s)$ is $e^{-|t-s|}$, a positive quantity. Hence, when $N(t)$ is above 0 (its average), we predict that $N(s)$ will exceed 0 as well.

c. $\mathrm{Cov}(N(10), N(12)) = e^{-|10-12|} = e^{-2}$. Since the variance is identically 1, meaning the standard deviation is also identically 1, $\rho = e^{-2}/(1)(1) = e^{-2} \approx .1353$.

d. $N(12) - N(10)$ is, by assumption, Gaussian. its mean is $0 - 0 = 0$, while its variance is given by $\mathrm{Var}(N(12) - N(10)) = \mathrm{Var}(N(12)) + (-1)^2\mathrm{Var}(N(10)) - 2\mathrm{Cov}(N(12), N(10)) = 1 + 1 - 2e^{-2}$. That is, $N(12) - N(10)$ has a Gaussian distribution with mean 0 and variance $2(1 - e^{-2}) \approx 1.73$.

17.

a. Since $-1 \le \cos \le 1$, $-A_0 \le X(t) \le A_0$ for any fixed t. If $X(t)$ were uniform on $[-A_0, A_0]$, then its variance would have to equal $\dfrac{(A_0 - -A_0)^2}{12} = \dfrac{A_0^2}{3}$; however, it was established in Example 7.12 that $\mathrm{Var}(X(t)) = \dfrac{A_0^2}{2}$. Therefore, $X(t)$ cannot have a uniform distribution (despite the appearance of its ensemble).

b. With the specified values, $Y = \cos(\Theta)$, where $\Theta \sim \mathrm{Unif}[-\pi, \pi]$. This is not one-to-one on $[-\pi, \pi]$; however, by symmetry we can restrict our attention to the branch of cosine on $[0, \pi]$. On that branch, write $\Theta = \arccos(Y)$. Then, by the transformation theorem,

$$f_Y(y) = f_\Theta(\arccos(y)) \cdot \left|\frac{d\theta}{dy}\right| = \frac{1}{\pi - 0} \cdot \left|\frac{d}{dy}[\arccos(y)]\right|$$
$$= \frac{1}{\pi} \cdot \left|-\frac{1}{\sqrt{1-y^2}}\right| = \frac{1}{\pi}\frac{1}{\sqrt{1-y^2}},$$

as claimed.

19.

a. $\mu_X(t) = E[X(t)] = E[S(t) + N(t)] = E[S(t)] + E[N(t)] = \mu_S(t) + \mu_N(t)$.

b. Since $S(t)$ and $N(t)$ are uncorrelated, $E[S(x)N(y)] = E[S(x)] \cdot E[N(y)]$ for any times x and y:

$$\begin{aligned} R_{XX}(t,s) &= E[X(t)X(s)] = E[(S(t)+N(t))(S(s)+N(s))] \\ &= E[S(t)S(s)] + E[S(t)N(s)] + E[N(t)S(s)] + E[N(t)N(s)] \\ &= R_{SS}(t,s) + E[S(t)]E[N(s)] + E[N(t)]E[S(s)] + R_{NN}(t,s) \\ &= R_{SS}(t,s) + \mu_S(t)\mu_N(s) + \mu_N(t)\mu_S(s) + R_{NN}(t,s) \end{aligned}$$

c. Since $S(t)$ and $N(t)$ are uncorrelated, $\text{Cov}(S(x), N(y)) = 0$ for any times x and y:

$$\begin{aligned} C_{XX}(t,s) &= \text{Cov}(X(t), X(s)) = \text{Cov}(S(t)+N(t), S(s)+N(s)) \\ &= \text{Cov}(S(t),S(s)) + \text{Cov}(S(t),N(s)) + \text{Cov}(N(t),S(s)) + \text{Cov}(N(t),N(s)) \\ &= C_{SS}(t,s) + 0 + 0 + C_{NN}(t,s) \\ &= C_{SS}(t,s) + C_{NN}(t,s) \end{aligned}$$

d. Use part **c**: $\sigma_X^2(t) = C_{XX}(t,t) = C_{SS}(t,t) + C_{NN}(t,t) = \sigma_S^2(t) + \sigma_N^2(t)$.

21.

a. $X(t)$ and $Y(t)$ uncorrelated $\Rightarrow$ by definition, $\text{Cov}(X(t), Y(s)) = 0$ for all times t and s. Apply the covariance shortcut formula: $\text{Cov}(X(t), Y(s)) = E[X(t)Y(s)] - E[X(t)]E[Y(s)] = 0 \Rightarrow$ $E[X(t)Y(s)] = E[X(t)]E[Y(s)] = \mu_X(t)\mu_Y(s)$, as claimed.

b. From **a**, $X(t)$ and $Y(t)$ uncorrelated $\Rightarrow E[X(t)Y(s)] = \mu_X(t)\mu_Y(s)$. If $X(t)$ has mean identically equal to zero, then $E[X(t)Y(s)] = 0\mu_Y(s) = 0$, which means by definition that $X(t)$ and $Y(s)$ are orthogonal.

23.

a. Use a trig identity, similar to Example 7.12:

$$\begin{aligned} R_{XY}(t,s) &= E[X(t)Y(s)] = E[\cos(\omega_0 t + \Theta)\sin(\omega_0 s + \Theta)] = \int_{-\pi}^{\pi} \cos(\omega_0 t + \theta)\sin(\omega_0 s + \theta)\cdot\frac{1}{2\pi}d\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin([\omega_0 s + \theta] + [\omega_0 t + \theta]) + \sin([\omega_0 s + \theta] - [\omega_0 t + \theta])}{2} d\theta \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi} [\sin(\omega_0(s+t) + 2\theta) + \sin(\omega_0(s-t))]d\theta \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi} \sin(\omega_0(s+t) + 2\theta)d\theta + \frac{1}{4\pi}\int_{-\pi}^{\pi} \sin(\omega_0(s-t))d\theta \\ &= \frac{1}{4\pi}(0) + \frac{1}{4\pi}\sin(\omega_0(s-t))[\pi - -\pi] = \frac{1}{2}\sin(\omega_0(s-t)) \end{aligned}$$

From Example 7.12, $X(t)$ has mean 0; a similar calculation shows that $Y(t)$ also has mean 0. Therefore,

$C_{XY}(t,s) = R_{XY}(t,s) - \mu_X(t)\mu_Y(s) = \frac{1}{2}\sin(\omega_0(s-t)) - (0)(0) = \frac{1}{2}\sin(\omega_0(s-t))$.

b. $X(t)$ and $Y(t)$ are none of these. To be orthogonal, the cross-correlation must be identically 0 (it isn't); to be uncorrelated, the cross-covariance must be identically 0 (it isn't); and non-zero correlation implies non-independence. Notice it's not sufficient for $\text{Cov}(X(t), Y(t))$ to equal zero, which happens to be true here; lack of correlation requires $\text{Cov}(X(t), Y(s)) = 0$ for all times t and s, equal or not.

Section 7.3

25. We'll rely heavily on the properties of $A_0\cos(\omega_0 t+\Theta)$ established in Example 7.14: (1) the mean of $A_0\cos(\omega_0 t+\Theta)$ is 0, and (2) the autocovariance of $A_0\cos(\omega_0 t+\Theta)$ is $(A_0^2/2)\cos(\omega_0\tau)$. From these, it follows that the autocorrelation of $A_0\cos(\omega_0 t+\Theta)$ is also $(A_0^2/2)\cos(\omega_0\tau)$.

a. $E[X(t)] = E[V + A_0\cos(\omega_0 t+\Theta)] = E[V] + E[A_0\cos(\omega_0 t+\Theta)] = \mu_V + 0 = \mu_V$.

b. Using the independence of V and Θ,

$$R_{XX}(t,s) = E[X(t)X(s)] = E[(V + A_0\cos(\omega_0 t+\Theta))\cdot(V + A_0\cos(\omega_0 s+\Theta))]$$
$$= E[V^2] + E[VA_0\cos(\omega_0 s+\Theta)] + E[VA_0\cos(\omega_0 t+\Theta)] + E[A_0^2\cos(\omega_0 t+\Theta)\cos(\omega_0 s+\Theta)]$$

The middle two terms can be separated using independence, while the third term is by definition the autocorrelation function of $A_0\cos(\omega_0 t+\Theta)$. Continuing,

$$R_{XX}(t,s) = E[V^2] + E[V]E[A_0\cos(\omega_0 s+\Theta)] + E[V]E[A_0\cos(\omega_0 t+\Theta)] + \frac{A_0^2}{2}\cos(\omega_0(t-s))$$
$$= E[V^2] + E[V](0) + E[V](0) + \frac{A_0^2}{2}\cos(\omega_0(t-s))$$

Making the substitution $s = t + \tau$ gives a final result of $E[V^2] + (A_0^2/2)\cos(\omega_0\tau)$.

c. Yes, $X(t)$ is wide-sense stationary, since its mean is a constant (see part **a**) and its autocorrelation function depends solely on the time difference τ (see part **b**).

27. For each function, we consider the properties of an autocovariance function that can be checked directly: (1) C_{XX} must be an even function of τ; (2) C_{XX} must achieve its maximum at $\tau = 0$; (3) if C_{XX} has no periodic component, it must diminish to 0 as $\tau \to \infty$.

a. $\cos(\tau)$ meets all conditions (1)-(3) above, so $\cos(\tau)$ could potentially be the autocovariance function of a WSS random process. [*Note:* The maximum does not have to be achieved *uniquely* at $\tau = 0$.]

b. $\sin(\tau)$ is not an even function, so $\sin(\tau)$ cannot be an autocovariance function.

c. $\dfrac{1}{1+\tau^2}$ meets all conditions (1)-(3) above, so $\dfrac{1}{1+\tau^2}$ could potentially be the autocovariance function of a WSS random process.

d. $e^{-|\tau|}$ meets all conditions (1)-(3) above, so $e^{-|\tau|}$ could potentially be the autocovariance function of a WSS random process.

29. No: If the mean of A and B is not zero, then $E[X(t)]$ will equal $\mu\cos(\omega_0 t) + \mu\sin(\omega_0 t)$, which is not a constant. Contrast this with Example 7.15.

31. When possible, we'll use the fact that $A(t)$ and $B(t)$ are jointly WSS to delete references to t. First, the mean function of $X(t)$ is $\mu_X(t) = E[A(t)+B(t)] = E[A(t)]+E[B(t)] = \mu_A + \mu_B$; the means of $A(t)$ and $B(t)$ are necessarily constants. Next, the autocovariance of $X(t)$ is

$$\begin{aligned}C_{XX}(t,s) &= \text{Cov}(A(t)+B(t), A(s)+B(s)) \\ &= \text{Cov}(A(t),A(s)) + \text{Cov}(A(t),B(s)) + \text{Cov}(B(t),A(s)) + \text{Cov}(B(t),B(s)) \\ &= C_{AA}(\tau) + C_{AB}(\tau) + C_{BA}(\tau) + C_{BB}(\tau)\end{aligned}$$

where, as usual, $\tau = s - t$. The property of joint wide-sense stationarity ensures that all four of these covariance functions — two autocovariance and two cross-covariance — depend solely on τ.
Since the mean of $X(t)$ is a constant and the autocovariance of $X(t)$ depends only on τ, by definition $X(t)$ is wide-sense stationary.

33.

a. Yes, $Y(t)$ must have periodic components, because its autocovariance function has periodic components (the two cosine functions).

b. $\text{Cov}(Y(0), Y(0.01)) = C_{YY}(0.01–0) = C_{YY}(0.01) = 50\cos(\pi) + 8\cos(6\pi) = –50 + 8 = –42$.

c. $R_{YY}(\tau) = C_{YY}(\tau) + \mu_Y^2 = 50\cos(100\pi\tau) + 8\cos(600\pi\tau) + 49$.

d. $E[Y^2(t)] = R_{YY}(0) = 50 + 8 + 49 = 107$.

e. $\text{Var}(Y(t)) = C_{YY}(0) = 50 + 8 = 58$.

35. It was established in Example 7.15 that $E[X(t)] = 0$. To be mean ergodic, it must also be the case that the time average of $X(t)$ is 0; i.e., $\langle X(t)\rangle = 0$. Consider just the first part of $X(t)$:

$$\langle A\cos(\omega_0 t)\rangle_T = \frac{1}{2T}\int_{-T}^{T} A\cos(\omega_0 t)dt = \frac{1}{T}\int_0^T A\cos(\omega_0 t)dt = \left.\frac{A\sin(\omega_0 t)}{\omega_0 T}\right|_0^T = \frac{A\sin(\omega_0 T)}{\omega_0 T}$$

As $T \to \infty$, the numerator diverges while the numerator is bounded, so $\langle A\cos(\omega_0 t)\rangle_T \to 0$. The time average of the second part of $X(t)$ is trivially zero since sine is odd; therefore, $\langle X(t)\rangle = \lim_{T\to\infty}\langle X(t)\rangle_T = 0$.
This random process is mean-ergodic.

37. For a WSS random process, $\text{Cov}(X(t), X(t + \tau)) = C_{XX}(\tau)$, while $\text{SD}(X(t)) = \text{SD}(X(t + \tau)) = \sigma_X$. Therefore,

$$\text{Corr}(X(t), X(t+\tau)) = \frac{C_{XX}(\tau)}{\sigma_X\sigma_X} = \frac{C_{XX}(\tau)}{\sigma_X^2} = \frac{C_{XX}(\tau)}{C_{XX}(0)}.$$

39. The key is to carefully consider the order of the arguments. For example, $R_{XY}(\tau) = E[X(t)Y(t + \tau)] = E[Y(t + \tau)X(t)] = R_{YX}(t + \tau, t)$. Since X and Y are jointly WSS, this last expression depends only on the time difference, but that's $t – (t + \tau) = –\tau$. Thus, $R_{XY}(\tau) = R_{YX}(–\tau)$. The exact same method establishes that $C_{XY}(\tau) = C_{YX}(–\tau)$.

Section 7.4

41.

a. February 28 is the 59th day of a calendar year (n = 59), so we're interested in the rv
$T_{59} = 75 + 25\sin\left(\frac{2\pi}{365}(59-150)\right) + 4\varepsilon_{59} \approx 50.000 + 4\varepsilon_{59}$. Since $\varepsilon_{59} \sim N(0, 1)$, the desired probability is
$P(T_{59} > 60) = P(50.000 + 4\varepsilon_{59} > 60) = P(\varepsilon_{59} > 2.5) = 1 - \Phi(2.5) = .0062$.

b. The only rv in the expression for T_n is ε_n, and $E[\varepsilon_n] = 0$, so
$E[T_n] = 75 + 25\sin\left(\frac{2\pi}{365}(59-150)\right) + 4E[\varepsilon_n] = 75 + 25\sin\left(\frac{2\pi}{365}(59-150)\right)$

c. Write μ_n for the mean function of T_n determined in part **b**, so that $T_n = \mu_n + 4\varepsilon_n$. Then
$C_{TT}[n,m] = \text{Cov}(T_n, T_m) = \text{Cov}(\mu_n + 4\varepsilon_n, \mu_m + 4\varepsilon_m) = \text{Cov}(4\varepsilon_n, 4\varepsilon_m) = 16\text{Cov}(\varepsilon_n, \varepsilon_m)$.
If $n = m$, then $\text{Cov}(\varepsilon_n, \varepsilon_m) = \text{Var}(\varepsilon_n) = 1$. If $n \neq m$, $\text{Cov}(\varepsilon_n, \varepsilon_m) = 0$ because the ε's are independent.
Therefore, $C_{TT}[n,m] = \begin{cases} 16 & n = m \\ 0 & n \neq m \end{cases}$. Using the Kronecker delta function, this may be written more compactly as $C_{TT}[n,m] = 16\delta[n-m]$.

d. Since the mean of T_n depends on n (see part **b**), T_n is not wide-sense stationary. And it shouldn't be: it's absurd to suggest that daytime high temperatures don't vary throughout the year!

43.

a. Let X_i = 1 if the gambler wins on the ith spin and = 0 if she loses. Then the X_i's are iid Bernoulli rvs with p = 18/38, so each one has mean p = 18/38 and variance $p(1 - p)$ = (18/38)(20/38) = 360/1444.
In this context, we recognize that the total number of wins is $S_n = X_1 + \cdots + X_n$, so that S_n is a random walk (an iid sum). Using the formulas from the proposition in this section,
$\mu_S[n] = n\mu_X = 18n/38$; $\sigma_S^2[n] = n\sigma_X^2 = 360n/1444$; $C_{SS}[n, m] = \min(n, m)\,\sigma_X^2 = 360\min(n, m)/1444$; and
$R_{SS}[n, m] = C_{SS}[n, m] + \mu_S[n]\mu_S[m] = (360\min(m, n) + 324mn)/1444$.

The mean and variance can also be determined from the realization that $S_n \sim \text{Bin}(n, 18/38)$.

b. The amount of money the gambler has won can be written in terms of the number of wins:
$Y_n = \$5 \cdot S_n + (-\$5)(n - S_n) = 10S_n - 5n$. Using the answers from part **a** and properties of mean/variance/covariance,
$\mu_Y[n] = 10\mu_S[n] - 5n = -10n/38$; $\sigma_Y^2[n] = 10^2\sigma_S^2[n] = 36{,}000n/1444$;
$C_{YY}[n,m] = \text{Cov}(Y_n, Y_m) = \text{Cov}(10S_n - 5n, 10S_m - 5m) = \text{Cov}(10S_n, 10S_m) = 100\text{Cov}(S_n, S_m)$
$= 36{,}000\min(n,m)/1444$; and
$R_{YY}[n, m] = C_{YY}[n, m] + \mu_Y[n]\mu_Y[m] = (36{,}000\min(n, m) + 100mn)/1444$.

c. Use the fact that $S_{10} \sim \text{Bin}(10, 18/38)$: $P(Y_{10} > 0) = P(10S_{10} - 5(10) > 0) = P(S_{10} > 5) =$
$\sum_{x=6}^{10}\binom{10}{x}\left(\frac{18}{38}\right)^x\left(\frac{20}{38}\right)^{10-x} = .3141$.

45. The notation can tricky here, so remember that brackets [] indicate a discrete function and parentheses () indicate a continuous function. We must determine the mean and autocovariance functions of $X[n]$.

(1) $E(X[n]) = E(X(nT_s))$. Since $X(t)$ is WSS, this mean does not depend on the argument inside parentheses; i.e., $E(X(t)) = \mu_X$ for all t. So, in particular, $E(X[n]) = E(X(nT_s)) = \mu_X$, a constant.

(2) $C_{XX}[n,n+k] = \text{Cov}(X[n], X[n+k]) = \text{Cov}(X(nT_s), X((n+k)T_s))$. Since $X(t)$ is WSS, this covariance depends only on the time difference, which is $\tau = (n+k)T_s - nT_s = kT_s$. That is, with $C_{XX}()$ denoting the autocovariance of $X(t)$, $C_{XX}[n,n+k] = C_{XX}(kT_s)$. In particular, this autocovariance function depends only on k and not on n.

Therefore, since the mean of $X[n]$ is a constant and the autocovariance of $X[n]$ depends only on the (discrete) time difference k, $X[n]$ is also WSS.

47.

a. Using linearity of expectation,

$$\mu_Y[n] = E\left(\frac{X_n + X_{n-1}}{2}\right) = \frac{E(X_n) + E(X_{n-1})}{2} = \frac{\mu_X + \mu_X}{2} = \mu_X$$

b. Using the distributive properties of covariance,

$$C_{YY}[n,m] = \text{Cov}\left(\frac{X_n + X_{n-1}}{2}, \frac{X_m + X_{m-1}}{2}\right)$$
$$= \frac{1}{4}\big(\text{Cov}(X_n, X_m) + \text{Cov}(X_{n-1}, X_m) + \text{Cov}(X_n, X_{m-1}) + \text{Cov}(X_{n-1}, X_{m-1})\big)$$

Re-write each covariance term in terms of C_{XX}, which depends only on the <u>difference</u> of indices:

$$C_{YY}[n,m] = \frac{1}{4}(C_{XX}[m-n] + C_{XX}[m-(n-1)] + C_{XX}[(m-1)-n] + C_{XX}[(m-1)-(n-1)])$$
$$= \frac{1}{4}(C_{XX}[m-n] + C_{XX}[m-n+1] + C_{XX}[m-n-1] + C_{XX}[m-n])$$
$$= \frac{1}{4}(2C_{XX}[m-n] + C_{XX}[m-n+1] + C_{XX}[m-n-1])$$

c. <u>Yes</u>, Y_n is wide-sense stationary. From part **a**, the mean of Y_n does not depend on n. From part **b**, the autocovariance of Y_n depends only on n and m through their difference, $n–m$. (Equivalently, replacing m with $n+k$ we see that the autocovariance depends only on k and not on n.) Therefore, Y_n is wide-sense stationary.

d. Use part **b**:

$$\text{Var}(Y_n) = C_{YY}[n,n] = \frac{1}{4}[2C_{XX}[0] + C_{XX}[0+1] + C_{XX}[0-1]]$$
$$= \frac{1}{4}[2C_{XX}[0] + C_{XX}[1] + C_{XX}[-1]] = \frac{1}{4}[2C_{XX}[0] + C_{XX}[1] + C_{XX}[1]]$$
$$= \frac{C_{XX}[0] + C_{XX}[1]}{2}$$

The second-to-last equality relies on the even symmetry of covariance, so $C_{XX}[–1] = C_{XX}[1]$.

49.

a. Let (*) denote the expression we're supposed to prove. Use the defining formula for Y_n iteratively:

$$\begin{aligned}
Y_n &= \alpha Y_{n-1} + X_n && \leftarrow \text{(*) for } N = 1 \\
&= \alpha[\alpha Y_{n-2} + X_{n-1}] + X_n \\
&= \alpha^2 Y_{n-2} + \alpha X_{n-1} + X_n && \leftarrow \text{(*) for } N = 2 \\
&= \alpha^2[\alpha Y_{n-3} + X_{n-2}] + \alpha X_{n-1} + X_n \\
&= \alpha^3 Y_{n-3} + \alpha^2 X_{n-2} + \alpha X_{n-1} + X_n && \leftarrow \text{(*) for } N = 3
\end{aligned}$$

etc.

You can also prove (*) more formally using mathematical induction.

b. In the first term, $|\alpha| < 1 \Rightarrow \alpha^N \to 0$ as $N \to \infty \Rightarrow \alpha^N Y_{n-N} \to 0$ as $N \to \infty$ (assuming the Y's don't diverge to ∞, which they generally will not). In the second term, the sum from 0 to $N - 1$ becomes the sum from 0 to ∞, by definition. Putting these together with part **a**,

$$\lim_{N\to\infty} Y_n = \lim_{N\to\infty}\left[\alpha^N Y_{n-N} + \sum\nolimits_{i=0}^{N-1} \alpha^i X_{n-i}\right] = 0 + \sum\nolimits_{i=0}^{\infty} \alpha^i X_{n-i} = \sum\nolimits_{i=0}^{\infty} \alpha^i X_{n-i} .$$

But the left-hand side, Y_n, does not depend on N, so the "limit" is just Y_n, and the desired result follows.

c. Use part **b**, linearity of expectation, and the fact that the X's are iid:

$$E(Y_n) = E\left(\sum_{i=0}^{\infty} \alpha^i X_{n-i}\right) = \sum_{i=0}^{\infty} \alpha^i E(X_{n-i}) = \sum_{i=0}^{\infty} \alpha^i \mu = \mu \sum_{i=0}^{\infty} \alpha^i$$

The sum is a basic geometric series equaling $1/(1 - \alpha)$, so $E(Y_n) = \dfrac{\mu}{1-\alpha}$.

d. Use part **b** and the distributive property of covariance:

$$C_{YY}[n,m] = \text{Cov}(Y_n, Y_m) = \text{Cov}\left(\sum_{i=0}^{\infty} \alpha^i X_{n-i}, \sum_{j=0}^{\infty} \alpha^j X_{m-j}\right) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \alpha^i \alpha^j \text{Cov}(X_{n-i}, X_{m-j})$$

Because the X's are independent, each covariance term is 0 unless $n - i = m - j$, i.e. $j = m - n + i$. Thus, the surviving non-zero terms are

$$\begin{aligned}
C_{YY}[n,m] &= \sum_{i=0}^{\infty} \alpha^i \alpha^{m-n+i} \text{Cov}(X_{n-i}, X_{n-i}) = \sum_{i=0}^{\infty} \alpha^{m-n+2i} \text{Var}(X_{n-i}) = \sum_{i=0}^{\infty} \alpha^{m-n+2i} \sigma^2 \\
&= \alpha^{m-n} \sigma^2 \sum_{i=0}^{\infty} \alpha^{2i} = \frac{\alpha^{m-n}\sigma^2}{1-\alpha^2}
\end{aligned}$$

where the last step uses the geometric series formula again.

e. <u>Yes</u>, Y_n is WSS. From part **c**, its mean does not depend on n. From part **d**, the substitution $m = n+k$ reveals that the autocovariance is $C_{YY}[k] = \dfrac{\alpha^k \sigma^2}{1-\alpha^2}$, which only depends on k and not on n.

f. Use the definition of the correlation coefficient:

$$\begin{aligned}
\rho(Y_n, Y_{n+k}) &= \frac{\text{Cov}(Y_n, Y_{n+k})}{\text{SD}(Y_n)\text{SD}(Y_{n+k})} = \frac{\text{Cov}(Y_n, Y_{n+k})}{\sqrt{\text{Var}(Y_n)\text{Var}(Y_{n+k})}} = \frac{C_{YY}[k]}{\sqrt{C_{YY}[0]C_{YY}[0]}} = \frac{C_{YY}[k]}{C_{YY}[0]} \\
&= \frac{\alpha^k \sigma^2 / (1-\alpha^2)}{\alpha^0 \sigma^2 / (1-\alpha^2)} = \alpha^k
\end{aligned}$$

51.

a. Using linearity of expectation,

$$E\left[\langle X[n]\rangle_N\right] = E\left[\frac{1}{2N+1}\sum_{n=-N}^{N} X[n]\right] = \frac{1}{2N+1}\sum_{n=-N}^{N} E(X[n]) = \frac{1}{2N+1}\sum_{n=-N}^{N} \mu_X$$

The indices on the sum run from $-N$ to N, a total of $2N+1$ integers ($-N, \ldots, -1, 0, 1, \ldots, N$). So, the sum is μ_X added to itself $2N+1$ times. Therefore,

$$E\left[\langle X[n]\rangle_N\right] = \frac{1}{2N+1}\cdot(2N+1)\mu_X = \mu_X$$, and this is true for any N.

b. Apply the hint, along with the distributive properties of covariance:

$$\begin{aligned}\mathrm{Var}\left(\langle X[n]\rangle_N\right) &= \mathrm{Cov}\left(\langle X[n]\rangle_N, \langle X[n]\rangle_N\right) = \mathrm{Cov}\left(\frac{1}{2N+1}\sum_{n=-N}^{N} X[n], \frac{1}{2N+1}\sum_{m=-N}^{N} X[m]\right)\\ &= \frac{1}{(2N+1)^2}\sum_{n=-N}^{N}\sum_{m=-N}^{N}\mathrm{Cov}(X[n],X[m]) = \frac{1}{(2N+1)^2}\sum_{n=-N}^{N}\sum_{m=-N}^{N} C_{XX}[m-n]\\ &= \frac{1}{(2N+1)^2}\sum_{n=-N}^{N}\sum_{k=-N-n}^{N-n} C_{XX}[k] \qquad \text{substituting } k = m-n \text{ (aka } m = k+n)\end{aligned}$$

Next, interchange the sums, and determine the new limits of the sums (drawing out a lattice will help). The now-outside variable k ranges from $-2N$ to $2N$; we'll determine the limits on the now-inside variable n shortly. Continuing from the previous step,

$$\begin{aligned}\mathrm{Var}\left(\langle X[n]\rangle_N\right) &= \frac{1}{(2N+1)^2}\sum_{n=-N}^{N}\sum_{k=-N-n}^{N-n} C_{XX}[k] = \frac{1}{(2N+1)^2}\sum_{k=-2N}^{2N}\sum_{n} C_{XX}[k]\\ &= \frac{1}{(2N+1)^2}\sum_{k=-2N}^{2N}\left\{C_{XX}[k]\cdot(\text{number of terms in the interior sum})\right\}\end{aligned}$$

since the argument of the sum, $C_{XX}[k]$, doesn't depend on n. Now we need to count indices for n. The original constraints $-N \le n \le N$ and $-N \le m \le N$ become $-N \le n \le N$ and $-N \le k + n \le N$, or $-N \le n \le N$ and $-N-k \le n \le N-k$. For $k \ge 0$, the intersection of these constraints is $-N \le n \le N-k$, a total of $(N-k) - (-N) + 1 = 2N + 1 - k$ integers. For $k \le 0$, the intersection is instead $-N-k \le n \le N$, a total of $N - (-N-k) + 1 = 2N + 1 + k$ integers.
So, in general, the interior summation above (indexed by n) has $2N + 1 - |k|$ integers; this covers both positive and negative values of k. Substitute:

$$\begin{aligned}\mathrm{Var}\left(\langle X[n]\rangle_N\right) &= \frac{1}{(2N+1)^2}\sum_{k=-2N}^{2N}\left\{C_{XX}[k]\cdot(2N+1-|k|)\right\} = \frac{1}{(2N+1)}\sum_{k=-2N}^{2N}\left\{C_{XX}[k]\cdot\frac{(2N+1-|k|)}{2N+1}\right\}\\ &= \frac{1}{(2N+1)}\sum_{k=-2N}^{2N} C_{XX}[k]\left(1-\frac{|k|}{2N+1}\right)\end{aligned}$$

completing the proof.

Section 7.5

53.

a. Let $X(2)$ denote the number of requests in a two-hour period. By assumption, $X(2) \sim$ Poisson with mean $\mu = \lambda t = 4(2) = 8$. Hence, $P(X(2)=10) = \dfrac{e^{-8}8^{10}}{10!} = .0993$.

b. Now t = 30 minutes = .5 hr, so $\mu = 4(.5) = 2$. The probability that no calls occur during that interval equals $\dfrac{e^{-2}2^{0}}{0!} = .1353$.

c. As noted in **b**, the expected number of calls in that interval is $\mu = 4(.5) = 2$.

55.

a. Let $X(t)$ denote the number of packet arrivals in a t-minute interval, so $X(t) \sim$ Poisson($10t$). Then $E(X(2)) = 10(2) = 20$, and $P(X(2)=15) = \dfrac{e^{-20}20^{15}}{15!} = .0516$.

b. $E(X(5)) = 10(5) = 50$, so $P(X(5) > 75) = 1 - P(X(5) \le 75) = 1 - \sum_{x=0}^{75} \dfrac{e^{-50}50^{x}}{x!}$. This expression can be evaluated with software to equal roughly .000372.

c. Let T = the time until the next packet arrives. By our main theorem on Poisson processes, T has an exponential distribution with parameter $\lambda = 10$. Hence, the probability T is less than 15 seconds (i.e., .25 minutes) is $P(T < .25) = \int_0^{.25} 10e^{-10t}\,dt = 1 - e^{-2.5} = .9179$.

d. Using the variable in part **c**, $E(T) = 1/\lambda = 1/10$ minutes, or 6 seconds.

e. Let Y_5 denote the arrival time of the fifth packet. By our main theorem on Poisson processes, Y_5 has a gamma distribution with $\alpha = n = 5$ and $\beta = 1/\lambda = .1$. Thus, with 45 seconds = .75 minutes,

$$P(Y_5 < .75) = \int_0^{.75} \frac{1}{\Gamma(5).1^5} y^{5-1} e^{-y/.1}\,dy = \frac{12500}{3}\int_0^{.75} y^4 e^{-10y}\,dy = .8679.$$

57. Let Y_k denote the arrival time of the kth passenger. By our main theorem on Poisson processes, Y_k has a gamma distribution with $\alpha = k$ and $\beta = 1/\lambda$. Thus, the expected duration of time until the next shuttle departs is $E(Y_k) = \alpha\beta = k/\lambda$.

59. Let $X(t)$ denote the Poisson process (t in hours), so $X(t) \sim$ Poisson($3t$).

a. Given that 4 orders are submitted during [0, 2], 10 orders are submitted in [0, 5] iff exactly 6 orders are submitted in (2, 5]. More formally, use the independent increments property of the Poisson process:

$$P(X(5)=10 \mid X(2)=4) = \frac{P(X(5)=10 \cap X(2)=4)}{P(X(2)=4)} = \frac{P(X(2)=4 \cap X(5)-X(2)=6)}{P(X(2)=4)}$$

$$= \frac{P(X(2)=4)\cdot P(X(5)-X(2)=6)}{P(X(2)=4)} \quad \text{independent increments}$$

$$= P(X(5)-X(2)=6)$$

The increment $X(5) - X(2)$ is a Poisson rv with mean $3(5 - 2) = 9$, so $P(X(5) - X(2) = 6) = \dfrac{e^{-9}9^6}{6!} = .0911$.

b. By the same principle, $P(X(t) = n \mid X(s) = m) = P(X(t) - X(s) = n - m)$. Since the increment $X(t) - X(s)$ has a Poisson distribution with mean $3(t - s)$, this probability equals $\dfrac{e^{3(t-s)}[3(t-s)]^{n-m}}{(n-m)!}$.

61. Let $X(t)$ denote the Poisson process, and let $X_1(t)$ denote the sub-process of surges that will disable the component. Assuming disabling surges are independent of each other and independent of how many surges occur, $X_1(t)$ has a Poisson distribution with mean $p\lambda t$. Thus, in particular, the probability that the component survives throughout the period $[0, t]$ is $P(X_1(t) = 0) = \dfrac{e^{-p\lambda t}(p\lambda t)^0}{0!} = e^{-p\lambda t}$.

63. Let T_1 denote the first arrival time at the north hospital, and let S_1 denote the first arrival time at the south hospital. Then Y is the longer of these two times; i.e., $Y = \max(T_1, S_1)$. Each of T_1 and S_1 are exponential rvs with parameter λ, and by assumption they are independent. So, for any time y,
$P(Y \le y) = P(\max(T_1, S_1) \le y) = P(T_1 \le y \cap S_1 \le y) = P(T_1 \le y) \cdot P(S_1 \le y) = \int_0^y \lambda e^{-\lambda t}\,dt \cdot \int_0^y \lambda e^{-\lambda s}\,ds =$ $(1-e^{-\lambda y})(1-e^{-\lambda y}) = (1-e^{-\lambda y})^2$. This is the cdf of Y: $F_Y(y) = P(Y \le y) = (1-e^{-\lambda y})^2$ for $y > 0$. To find the pdf of y, differentiate: $f_Y(y) = F_Y'(y) = 2(1-e^{-\lambda y})^1 \cdot -\lambda e^{-\lambda y} = 2\lambda e^{-\lambda y}(1-e^{-\lambda y})$ for $y > 0$.

65.

a. This is a partial converse to our main theorem on Poisson processes. The goal is to prove $P(X(t) = x)$ equals the Poisson(λt) pmf for $x = 0, 1, 2, \ldots$. The event $\{X(t) = x\}$ means exactly x arrivals occurred in $[0, t]$; framed in terms of the inter-arrival times, this means $T_1 + \cdots + T_x \le t$ while $T_1 + \cdots + T_{x+1} > t$. Write $Y_x = T_1 + \cdots + T_x$, the (absolute) arrival time of the xth event. By the assumptions of a renewal process, Y_x and T_{x+1} are independent; moreover, if the T's are exponential, then from Chapter 4 (or this section) we know that Yx has a gamma distribution. Now, use calculus to determine the desired probability: $P(T_1 + \cdots + T_x \le t \cap T_1 + \cdots + T_{x+1} > t) = P(Y_x \le t \cap Y_x + T_{x+1} > t) = P(Y_x \le t \cap T_{x+1} > t - Y_x) =$

$$\int_0^t \int_{t-y}^\infty f_{Y,T}(y,u)\,du\,dy = \int_0^t \int_{t-y}^\infty \frac{\lambda^x}{\Gamma(x)} y^{x-1} e^{-\lambda y} \cdot \lambda e^{-\lambda u}\,du\,dy = \frac{\lambda^x}{(x-1)!}\int_0^t y^{x-1}e^{-\lambda y}\left[\int_{t-y}^\infty \lambda e^{-\lambda u}\,du\right]dy$$

$$= \frac{\lambda^x}{(x-1)!}\int_0^t y^{x-1}e^{-\lambda y}\left[e^{-\lambda(t-y)}\right]dy$$

$$= \frac{\lambda^x}{(x-1)!}\int_0^t y^{x-1}e^{-\lambda t}\,dy = \frac{\lambda^x e^{-\lambda t}}{(x-1)!}\int_0^t y^{x-1}\,dy$$

$$= \frac{\lambda^x e^{-\lambda t}}{(x-1)!}\cdot\frac{t^x}{x} = \frac{e^{-\lambda t}(\lambda t)^x}{x!}$$

This is precisely the Poisson(λt) pmf.

b. For a Poisson process, $X(t) \sim$ Poisson(λt), so $E[X(t)] = \lambda t$. Also, $E[T_n] = 1/\lambda$, the mean of an exponential rv. Hence, $\lim\limits_{t\to\infty}\dfrac{E[X(t)]}{t} = \lim\limits_{t\to\infty}\dfrac{\lambda t}{t} = \lim\limits_{t\to\infty}\lambda = \lambda = \dfrac{1}{1/\lambda} = \dfrac{1}{E[T_n]}$.

67.

a. The expected number of customers in the first hour, i.e. [0, 1], is $\int_0^1 \lambda(t)dt = \int_0^1 t\,dt = .5$. Similarly, the expected number of customers in the next 6 hours, (1, 7], is $\int_1^7 1dt = 6$ and the expected number of customers in the last hour, (7, 8], is $\int_7^8 (8-t)dt = .5$. By the property of independent increments, the desired probability is

$$\begin{aligned} &P(X(1)=0 \cap X(7)-X(1)=4 \cap X(8)-X(7)=0) \\ &\quad = P(X(1)=0)\cdot P(X(7)-X(1)=4)\cdot P(X(8)-X(7)=0) \\ &\quad = \frac{e^{-.5}.5^0}{0!}\cdot\frac{e^{-6}6^4}{4!}\cdot\frac{e^{-.5}.5^0}{0!} = .0492 \end{aligned}$$

b. The probability of n customers in each of those three intervals is $\frac{e^{-.5}.5^n}{n!}\cdot\frac{e^{-6}6^n}{n!}\cdot\frac{e^{-.5}.5^n}{n!} = \frac{e^{-7}1.5^n}{(n!)^3}$. So, the probability of the same number of customers is $\sum_{n=0}^{\infty}\frac{e^{-7}1.5^n}{(n!)^3}$ = .00255.

69.

a. Given that $N(0) = +1$, $N(t)$ will equal +1 iff the telegraph switched parity an even number of times (e.g., $+1 \to -1 \to +1 \to -1 \to +1$). Otherwise, if the number of switches in $(0, t]$ were odd, $N(0)$ and $N(t)$ will have opposite parity. Thus, $P(N(t) = +1 \mid N(0) = +1) = P($<u>even</u> # of switches in $(0, t]) = p$, and $P(N(t) = +1 \mid N(0) = -1) = P($<u>odd</u> # of switches in $(0, t]) = 1 - p$.

b. Condition on the parity of $N(0)$, and use the fact that $N(0) = \pm 1$ with probability .5 each by definition: $P(N(t) = +1) = P(N(0) = +1)\, P(N(t) = +1 \mid N(0) = +1) + P(N(0) = -1)\, P(N(t) = +1 \mid N(0) = -1) = (.5)p + (.5)(1 - p) = .5$.

c. At all time points t, $N(t)$ is ± 1 with probability .5 each. Thus, $\mu_N(t) = E[N(t)] = (-1)(.5) + (+1)(.5) = 0$. Also, $\text{Var}(N(t)) = (-1 - 0)^2(.5) + (1 - 0)^2(.5) = .5 + .5 = 1$, so $\sigma_N(t) = 1$.

71. As noted in the hint, we may write $N'(t) = ½[N(t) + 1]$. From this,

- Since $N(t) = \pm 1$ with probability .5 each for all t, $N'(t) = 0$ or 1 with probability .5 each for all t.
- $E[N'(t)] = ½[E[N(t)] + 1] = ½[0 + 1] = ½$ (or, take the average of 0 and 1).
- $\text{Var}(N'(t)) = (½)^2\text{Var}(N(t)) = .25(1) = .25$.
- $\text{Cov}(N'(t), N'(t+\tau)) = \text{Cov}(½[N(t) + 1], ½[N(t+\tau) + 1]) = (½)^2\text{Cov}(N(t), N(t+\tau)) = .25e^{-2\lambda|\tau|}$.
- $E[N'(t)N'(t+\tau)] = \text{Cov}(N'(t), N'(t+\tau)) + E[N'(t)]E[N'(t+\tau)] = .25e^{-2\lambda|\tau|} + .25$.

Section 7.6

73.

a. Since $X(t)$ is WSS, $X(10)$ has mean $\mu_X = 13$ and variance $C_{XX}(0) = 9$. Thus, $P(X(10) < 5) = \Phi\left(\frac{5-13}{\sqrt{9}}\right) = \Phi(-2.67) = .0038$.

b. Since $X(t)$ is a Gaussian process, $X(10) - X(8)$ is a Gaussian rv. Its mean is $E[X(10) - X(8)] = \mu_X - \mu_X = 13 - 13 = 0$; its variance is given by

$$\begin{aligned}\text{Var}(X(10)-X(8)) &= \text{Var}(X(10)) + (-1)^2\text{Var}(X(8)) + 2(1)(-1)\text{Cov}(X(10), X(8)) \\ &= C_{XX}(0) + C_{XX}(0) - 2C_{XX}(10-8) \\ &= 9 + 9 - 2\cdot 9\cos(2/5) \approx 1.421\end{aligned}$$

Hence, $P(X(10) - X(8) < 2) = \Phi\left(\frac{2-0}{\sqrt{1.421}}\right) = \Phi(1.68) = .9535$.

75.

a. Yes, $N(t)$ is stationary. Its mean is a constant, and its autocorrelation function can be re-parameterized as follows: $R_{NN}(t, t+\tau) = 1 - |t-(t+\tau)|/10 = 1 - |\tau|/10$ for $|\tau| \le 10$ (and zero otherwise). Thus, its autocorrelation depends only on τ and not t. Therefore, $N(t)$ is wide-sense stationary. But since, $N(t)$ is also Gaussian, by the proposition in this section $N(t)$ is also (strict-sense) stationary.

b. Since $\mu_N = 0$, $C_{NN}(\tau) = R_{NN}(\tau) - 0^2 = 1 - |\tau|/10$. In particular, $\text{Var}(N(t)) = C_{NN}(0) = 1$, so $\text{SD}(N(t)) = 1$. In other words, $N(t)$ is standard normal. Therefore, $P(|N(t)| > 1) = 1 - P(-1 \le N(t) \le 1) = 1 - [\Phi(1) - \Phi(-1)] = .3174$.

c. Since $N(t)$ is a Gaussian process, the increment $N(t+5) - N(t)$ is also Gaussian. Its mean is $0 - 0 = 0$; its variance is given by

$$\begin{aligned}\text{Var}(N(t+5)-N(t)) &= \text{Var}(N(t+5)) + (-1)^2\text{Var}(N(t)) + 2(1)(-1)\text{Cov}(N(t+5), N(t)) \\ &= C_{NN}(0) + C_{NN}(0) - 2C_{YY}(5) \\ &= 1 + 1 - 2\cdot[1 - |5|/10] = 1\end{aligned}$$

So, this increment is also standard normal, and $P(|N(t+5) - N(t)| > 1) = .3174$ from part **b**.

d. Be careful here: the autocovariance function equals zero for $|\tau| > 10$. Similar to part **c**, the increment $N(t+15) - N(t)$ is Gaussian with mean 0, but its variance equals $C_{NN}(0) + C_{NN}(0) - 2C_{YY}(15) = 1 + 1 - 2(0) = 2$. Therefore,

$P(|N(t+15) - N(t)| > 1) = 1 - P(-1 \le N(t+15) - N(t) \le 1) = 1 - \left[\Phi\left(\frac{1-0}{\sqrt{2}}\right) - \Phi\left(\frac{-1-0}{\sqrt{2}}\right)\right] = 1 - [\Phi(0.71) - \Phi(-0.71)] = .4778$.

77.

a. Since $E[B(t)] = 0$ by definition, $E[X(t)] = 80 + 20\cos\left(\frac{\pi}{12}(t-15)\right)$. Next, since $80 + 20\cos\left(\frac{\pi}{12}(t-15)\right)$ is a constant w.r.t. the ensemble, $\text{Var}(X(t)) = \text{Var}(B(t)) = \alpha t = .2t$. The mean function is the expected temperature in Bakersfield t hours after midnight on August 1, which like many temperature patterns follows a sinusoid. The variance function indicates the uncertainty around that mean: at $t = 0$ we believe our temperature prediction is exactly right, but uncertainty increases as t increases.

b. The time 3pm is $t = 15$ hours after midnight, at which point $X(15) = 80 + 20\cos(0) + B(15) = 100 + B(15)$. Since Brownian motion is a Gaussian process, $X(15)$ is Gaussian with mean 100 and variance $.2(15) = 3$. Therefore, $P(X(15) > 102) = 1 - \Phi\left(\frac{102-100}{\sqrt{3}}\right) = 1 - \Phi(1.15) = .1251$.

c. The time 3pm on August 5 is $t = 4 \times 24 + 15 = 111$ hours after midnight on August 1. The mean of $X(111)$ turns out to be 100 again (the model assumes a 24-hour cycle), but the variance is $.2(111) = 22.2$. Therefore, $P(X(111) > 102) = 1 - \Phi\left(\frac{102-100}{\sqrt{22.2}}\right) = 1 - \Phi(0.42) = .3372$.

d. The goal is to determine $P(|X(111) - X(15)| \le 1)$. Write $X(15) = 100 + B(15)$ and $X(111) = 100 + B(111)$, and this simplifies to $P(|B(111) - B(15)| \le 1)$. This concerns an <u>increment</u> of Brownian motion, which is Gaussian with mean 0 and variance $\alpha\tau = .2(111 - 15) = .2(96) = 19.2$. Therefore,

$P(|B(111) - B(15)| \le 1) = P(-1 \le B(111) - B(15) \le 1) = \Phi\left(\frac{1-0}{\sqrt{19.2}}\right) - \Phi\left(\frac{-1-0}{\sqrt{19.2}}\right) =$

$\Phi(0.23) - \Phi(-0.23) = .1818$.

79. These questions use the random variables M (maximum) and T (hitting time) described for Brownian motion in this section.

a. If the meteorologist uses the mean function as her forecast, then $B(t) = (X(t) - \text{forecast})$ is her error. The goal is to determine the probability that $B(t)$ exceeds 5°F sometime between $t = 0$ and $t = 120$ (hours, aka 5 days). This is equivalent to $M > 5$, where $M = \max_{0 \le t \le 120} B(t)$. By the proposition in this section with $\alpha = .2$, $t_0 = 120$, and $x_0 = 5$,

$$P(M > 5) = 2\left[1 - \Phi\left(\frac{x_0}{\sqrt{\alpha t_0}}\right)\right] = 2\left[1 - \Phi\left(\frac{5}{\sqrt{.2(120)}}\right)\right] = 2[1 - \Phi(1.02)] = .3078.$$

b. Let T = the first time the meteorologist's forecast error exceeds 5°F. The probability she has such a forecast error <u>by</u> midnight on August 3 ($t = 2 \times 24 = 48$) is $P(T \le 48)$. From the proposition in this section,

$$P(T \le 48) = 2\left[1 - \Phi\left(\frac{x_0}{\sqrt{\alpha t}}\right)\right] = 2\left[1 - \Phi\left(\frac{5}{\sqrt{.2(48)}}\right)\right] = 2[1 - \Phi(1.61)] = .1074.$$

81. $X(t)$ and $Y(t)$ are independent Gaussian processes, each with mean 0 and variance $\alpha t = 5t$.

a. The probability the particle is more than 3 units away <u>in each dimension</u> at $t = 2$ is
$P(|X(2)| > 3 \cap |Y(2)| > 3) = P(|X(2)| > 3) \cdot P(|Y(2)| > 3)$ by independence
$= [1 - P(-3 \le X(2) \le 3)]^2$ since $X(2)$ and $Y(2)$ are identically distributed

$$= \left[1 - \left\{\Phi\left(\frac{3-0}{\sqrt{5(2)}}\right) - \Phi\left(\frac{-3-0}{\sqrt{5(2)}}\right)\right\}\right]^2 = [1 - \{\Phi(0.95) - \Phi(-0.95)\}]^2 = [.3422]^2 = .1171.$$

b. The probability the particle is more than 3 units away <u>radially</u> at $t = 2$ is

$P\left(\sqrt{[X(2)-0]^2+[Y(2)-0]^2}>3\right)=P(X^2(2)+Y^2(2)>9)$. As stated in the hint, the sum of squares of two iid $N(0, \sigma)$ rvs is exponentially distributed with $\lambda = 1/(2\sigma^2)$; here, $\sigma^2 = \alpha t = 5(2) = 10$, so $\lambda = 1/20$. Therefore, $P(X^2(2)+Y^2(2)>9)=\int_9^\infty (1/20)e^{-(1/20)x}\,dx = e^{-9/20} = .6376$.

c. At time $t = 4$, $X(4)$ and $Y(4)$ and $Z(4)$ are iid normal rvs with mean 0 and variance $.2(4) = .8$. Mimicking the solutions for parts **a** and **b** above,

i. $P(|X(4)| > 1 \cap |Y(4)| > 1 \cap |Z(4)| > 1) = \left[1-\left\{\Phi\left(\frac{1-0}{\sqrt{.8}}\right)-\Phi\left(\frac{-1-0}{\sqrt{.8}}\right)\right\}\right]^3 =$

$[1 - \{\Phi(1.12) - \Phi(-1.12)\}]^3 = [.2628]^3 = .0181$.

ii. $P(X^2(4)+Y^2(4)+Z^2(4)>1^2)=\int_1^\infty f(x)dx$, where $f(x)$ is the pdf of the gamma distribution with parameters $\alpha = 3/2$ and $\beta = 2\sigma^2 = 2(.8) = 1.6$. Continuing,

$\int_1^\infty f(x)dx = \frac{1}{\Gamma(3/2)\cdot 1.6^{3/2}}\int_1^\infty x^{3/2-1}e^{-x/1.6}dx = .7410$ with the aid of software.

It's quite surprising how different these are: while (i) implies (ii), and so the second probability must exceed the first, there's a 74% chance of the process escaping a sphere of radius 1 by $t = 4$ but only a 1.8% chance if it escaping the circumscribed cube by $t = 4$.

83.

a. Yes, Gaussian white noise is a stationary process. By definition, its mean is 0 (a constant) and its autocorrelation is $\frac{N_0}{2}\delta(\tau)$ (which depends only on τ and not t). Therefore, $N(t)$ is wide-sense stationary. But since $N(t)$ is also Gaussian, this implies that $N(t)$ is strict-sense stationary.

b. Interchange expected value and integration as necessary:

$E[X(t)] = E\left[\int_0^t N(s)ds\right] = \int_0^t E[N(s)]ds = \int_0^t 0ds = 0$; assuming $t + \tau > t > 0$,

$$R_{XX}(t,t+\tau) = E[X(t)X(t+\tau)] = E\left[\int_0^t N(s)ds\cdot\int_0^{t+\tau}N(s')ds'\right] = E\left[\int_0^t\int_0^{t+\tau}N(s)N(s')ds'ds\right]$$
$$= \int_0^t\int_0^{t+\tau}E[N(s)N(s')]ds'ds = \int_0^t\int_0^{t+\tau}R_{NN}(s'-s)ds'ds$$
$$= \int_0^t\int_0^{t+\tau}\frac{N_0}{2}\delta(s'-s)ds'ds$$

The antiderivative of $\delta(x)$ is $u(x)$, the unit step function. Continuing,

$$R_{XX}(t,t+\tau) = \frac{N_0}{2}\int_0^t[u(t+\tau-s)-u(0-s)]ds$$
$$= \frac{N_0}{2}\int_0^t u(t+\tau-s)ds \quad \text{because } u(x)=0 \text{ for } x<0$$
$$= \frac{N_0}{2}\int_0^t 1ds \quad \text{because } u(x)=1 \text{ for } x>0$$
$$= \frac{N_0}{2}t \text{ for } t+\tau>t>0$$

In particular, the autocorrelation of $X(t)$ depends on t, so $X(t)$ is not WSS. [*Note:* The autocorrelation can be re-written as $R_{XX}(t, s) = \frac{N_0}{2}t$ for $s > t > 0$, i.e. $R_{XX}(t, s) = \frac{N_0}{2}\min(t, s)$. This is precisely the autocorrelation of Brownian motion; in fact, Brownian motion can be defined as the integral of Gaussian white noise!]

85.

a. If $X(t)$ has independent increments, then the two terms in brackets on the right-hand side are independent. Hence, they have zero covariance, and
$\text{Var}(X(t+\tau) - X(0)) = \text{Var}(X(t+\tau) - X(t)) + \text{Var}(X(t) - X(0))$.
Since $X(0)$ is a constant, it may be removed:
$\text{Var}(X(t+\tau)) = \text{Var}(X(t+\tau) - X(t)) + \text{Var}(X(t))$.
Finally, if $X(t)$ has stationary increments, then $X(t+\tau) - X(t)$ has the same distribution as any other increment of time-length τ, such as $X(\tau) - X(0)$. Substitute and simplify:
$\text{Var}(X(t+\tau)) = \text{Var}(X(t+\tau) - X(t)) + \text{Var}(X(t)) = \text{Var}(X(\tau) - X(0)) + \text{Var}(X(t)) = \text{Var}(X(\tau)) + \text{Var}(X(t))$.

b. Let $g(t) = \text{Var}(X(t))$, so part **a** implies $g(t + \tau) = g(\tau) + g(t)$. As noted in the hint, the only functions that can satisfy this relationship for all t and τ are of the form $g(t) = at$, whence $\text{Var}(X(t)) = at$ for some a. [Of course, since variance is positive, $a > 0$ here.] That is, any random process with a constant initial value and stationary and independent increments must have linear variance.

Section 7.7

87.

a. This system consists of three states: 0 = empty, 1 = a person in stage 1, and 2 = a person in stage 2. (According to the problem, you can't have someone in stage 1 while another person is in stage 2.) By the main theorem of this section, the parameters of the exponential times are exactly the sojourn rates for the three states: $q_0 = \lambda$, $q_1 = \lambda_1$, and $q_2 = \lambda_2$.
Three of the instantaneous transition rates are zero: since $P(0 \to 2) = P(2 \to 1) = P(1 \to 0) = 0$, we have $q_{02} = q_{21} = q_{10} = 0$. This only leaves three other parameters. By the property $q_i = \sum_{j \neq i} q_{ij}$, we obtain $q_0 = q_{01} + q_{02} \Rightarrow q_{01} = q_0 = \lambda$ and, similarly, $q_{12} = \lambda_1$ and $q_{20} = \lambda_2$.

b. The stationary probabilities satisfy $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ and $\sum \pi_j = 1$, where from part **a** the generator matrix $\mathbf{Q}$ is
$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda & 0 \\ 0 & -\lambda_1 & \lambda_1 \\ \lambda_2 & 0 & -\lambda_2 \end{bmatrix}$. The first two columns of the matrix equation $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ yield $\pi_2 = \dfrac{\lambda}{\lambda_2}\pi_0$ and $\pi_1 = \dfrac{\lambda}{\lambda_1}\pi_0$; since the π's sum to 1, $\pi_0 + \dfrac{\lambda}{\lambda_1}\pi_0 + \dfrac{\lambda}{\lambda_2}\pi_0 = 1$, from which we can solve for π_0 and then back-substitute to find the others. The general solutions are

$$\pi_0 = \left[\frac{\lambda}{\lambda} + \frac{\lambda}{\lambda_1} + \frac{\lambda}{\lambda_2}\right]^{-1}, \quad \pi_1 = \left[\frac{\lambda_1}{\lambda} + \frac{\lambda_1}{\lambda_1} + \frac{\lambda_1}{\lambda_2}\right]^{-1}, \quad \pi_2 = \left[\frac{\lambda_2}{\lambda} + \frac{\lambda_2}{\lambda_1} + \frac{\lambda_2}{\lambda_2}\right]^{-1}$$

Plugging in $\lambda = 1$, $\lambda_1 = 3$, $\lambda_2 = 2$ returns $\pi_0 = \dfrac{6}{11}, \pi_1 = \dfrac{2}{11}, \pi_2 = \dfrac{3}{11}$.

c. Use the general solution from part **b**. Plugging in $\lambda = 1$, $\lambda_1 = 2$, $\lambda_2 = 3$ returns $\pi_0 = \dfrac{6}{11}, \pi_1 = \dfrac{3}{11}, \pi_2 = \dfrac{2}{11}$.

d. Similarly, plugging in $\lambda = 4$, $\lambda_1 = 2$, $\lambda_2 = 1$ returns $\pi_0 = \dfrac{1}{7}, \pi_1 = \dfrac{2}{7}, \pi_2 = \dfrac{4}{7}$.

89.

a. The system has the same three states: 0 = empty, 1 = a person in stage 1, and 2 = a person in stage 2. And, as in Exercise 87, the parameters of the exponential times are the sojourn rates for the three states: $q_0 = \lambda$, $q_1 = \lambda_1$, and $q_2 = \lambda_2$. Two transitions are still impossible: $P(0 \to 2) = P(1 \to 0) = 0$, so $q_{02} = q_{10} = 0$. It follows from $q_i = \sum_{j \neq i} q_{ij}$ that $q_{01} = \lambda$ and $q_{12} = \lambda_1$ as before.
The new twist is that $2 \to 0$ is now possible. The transition probabilities into states 0 and 1 upon leaving state 2 are given to be .8 and .2, respectively. By part (b) of the main theorem of this section, these conditional probabilities are q_{20}/q_2 and q_{21}/q_2, respectively. So, associated with state 2 we have $q_2 = \lambda_2$, $q_{20}/q_2 = .8$, and $q_{21}/q_2 = .2$; these give the final two instantaneous transition rates: $q_{20} = .8\lambda_2$ and $q_{21} = .2\lambda_2$.

b. The stationary probabilities satisfy $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ and $\sum \pi_j = 1$, where from part **a** the generator matrix **Q** is
$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda & 0 \\ 0 & -\lambda_1 & \lambda_1 \\ .8\lambda_2 & .2\lambda_2 & -\lambda_2 \end{bmatrix}$. The first and third columns of the matrix equation $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ yield $\pi_0 = \frac{.8\lambda_2}{\lambda}\pi_2$

and $\pi_1 = \frac{\lambda_2}{\lambda_1}\pi_2$; since the π's sum to 1, $\frac{.8\lambda_2}{\lambda}\pi_2 + \frac{\lambda_2}{\lambda_1}\pi_2 + \pi_2 = 1$, from which we can solve for π_2 and then back-substitute to find the others. The general solutions are

$$\pi_0 = .8 \times \left[\frac{.8\lambda}{\lambda} + \frac{\lambda}{\lambda_1} + \frac{\lambda}{\lambda_2}\right]^{-1}, \quad \pi_1 = \left[\frac{.8\lambda_1}{\lambda} + \frac{\lambda_1}{\lambda_1} + \frac{\lambda_1}{\lambda_2}\right]^{-1}, \quad \pi_2 = \left[\frac{.8\lambda_2}{\lambda} + \frac{\lambda_2}{\lambda_1} + \frac{\lambda_2}{\lambda_2}\right]^{-1}$$

Plugging in $\lambda = 1$, $\lambda_1 = 3$, $\lambda_2 = 2$ returns $\pi_0 = \frac{24}{49}, \pi_1 = \frac{10}{49}, \pi_2 = \frac{15}{49}$.

c. Use the general solution from part **b**. Plugging in $\lambda = 1$, $\lambda_1 = 2$, $\lambda_2 = 3$ returns
$\pi_0 = \frac{24}{49}, \pi_1 = \frac{15}{49}, \pi_2 = \frac{10}{49}$.

d. Similarly, plugging in $\lambda = 4$, $\lambda_1 = 2$, $\lambda_2 = 1$ returns $\pi_0 = \frac{2}{17}, \pi_1 = \frac{5}{17}, \pi_2 = \frac{10}{17}$.

e. For each visit to stage 1, the sojourn time is exponential with parameter $q_1 = \lambda_1$, so the average length of such a visit is $1/\lambda_1$. Similarly, the average length of time spent in one visit to stage 2 is $1/q_2 = 1/\lambda_2$. But this stage 1-stage 2 process can happen multiple times. Specifically, since the probability of a return (failure) is .2 and departure (success) is .8, the expected number of times through this two-stage process is 1/.8 = 1.25, based on the expected value of a geometric rv. Therefore, the expected total amount of time spent at the facility is $1.25(1/\lambda_1 + 1/\lambda_2)$.

91. This process is similar to the pure-birth process in the previous exercise, if we define a "birth" to be a new individual becoming infected. In the previous exercise, the following differential equation was established: $P_{ij}'(t) = -j\beta P_{ij}(t) + (j-1)\beta P_{i,j-1}(t)$. Recall that $P_{ii}(0) = 1$ and $P_{ij}(0) = 0$ for $j \neq i$; also, recall that the infinitesimal parameters can be obtained through the derivative of $P_{ij}(t)$ at $t = 0$.
Using the differential equation, the q_i's are given by
$q_i = -P_{ii}'(0) = +i\beta P_{ii}(0) - (i-1)\beta P_{i,i-1}(0) = i\beta(1) - (i-1)(0) = i\beta$. The only non-zero instantaneous transition rates are those from i to $i+1$ (one more infection); those rates are given by

$$\begin{aligned} q_{i,i+1} &= P_{i,i+1}'(0) = -(i+1)\beta P_{i,i+1}(0) + (i+1-1)\beta P_{i,i+1-1}(0) \\ &= -(i+1)\beta P_{i,i+1}(0) + i\beta P_{i,i}(0) = -(i+1)\beta(0) + i\beta(1) = i\beta \end{aligned}$$

This holds for $i = 1, \ldots, N-1$. Since the population size is zero, $q_{ij} = 0$ for any $j > N$.

93. Since this is a birth-and-death process, only three "groups" of parameters are non-zero: $q_{i,i+1}$ (birth), $q_{i,i-1}$ (death), and the q_i's.
With births resulting from a Poisson process, $q_{i,i+1} = \lambda$ for all $i \geq 0$.
As discussed in this section, exponential lifetimes correspond to a "death" rate of $\beta h + o(h)$ for each individual particle. Given that $X_t = i$, we've seen that the probability of a death in the next h time units is $P_{i,i-1}(h) = i\beta h + o(h)$, from which $q_{i,i-1} = i\beta$ for $i \geq 1$.
Finally, since $q_i = \sum_{j\neq i} q_{ij}$, $q_0 = q_{01} = \lambda$ and $q_i = q_{i,i+1} + q_{i,i-1} = \lambda + i\beta$ for $i \geq 1$.

95. Each component, A and B, mimics Example 7.35. In particular, with superscripts indicating the component, $P^A_{0,1} = \alpha_0 h + o(h), P^A_{1,0} = \alpha_1 h + o(h), P^B_{0,1} = \beta_0 h + o(h), P^B_{0,1} = \beta_1 h + o(h)$, where 0 = working and 1 = needing repair. Put the two components together, and the overall system has four states:
00 = both working; 01 = A working, B down; 10 = A down, B working; 11 = both down for repairs.
To find the transition probabilities (from which we can infer the rates), use independence: first, $P_{00,11}(h) = P^A_{0,1}(h)\cdot P^B_{0,1}(h) = (\alpha_0 h + o(h))(\beta_0 h + o(h)) = o(h)$. In fact, the probabilities of all double-switches in h time units are $o(h)$, so those instantaneous transition rates are 0. For the others,

$$\begin{aligned} P_{00,01}(h) &= P^A_{0,0}(h)\cdot P^B_{0,1}(h) = [1 - P^A_{0,1}(h)]\cdot P^B_{0,1}(h) \\ &= [1-\alpha_0 h + o(h)]\cdot[\beta_0 h + o(h)] = \beta_0 h + o(h) \Rightarrow \\ q_{00,01} &= \beta_0 \end{aligned}$$

Similarly, $q_{00,10} = \alpha_0$, $q_{01,00} = \beta_1$, etc. From these and $q_i = \sum_{j\neq i} q_{ij}$, we can construct the generator matrix:

$$\mathbf{Q} = \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array}\begin{bmatrix} -(\alpha_0+\beta_0) & \beta_0 & \alpha_0 & 0 \\ \beta_1 & -(\alpha_0+\beta_1) & 0 & \alpha_0 \\ \alpha_1 & 0 & -(\alpha_1+\beta_0) & \beta_0 \\ 0 & \alpha_1 & \beta_1 & -(\alpha_1+\beta_1) \end{bmatrix}$$

The stationary probabilities are the solution to $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ and $\sum \pi_j = 1$; with the aid of software, these are $\pi_{00} = \dfrac{\alpha_1\beta_1}{\Sigma}, \pi_{01} = \dfrac{\alpha_1\beta_0}{\Sigma}, \pi_{10} = \dfrac{\alpha_0\beta_1}{\Sigma}, \pi_{11} = \dfrac{\alpha_0\beta_0}{\Sigma}$, where $\Sigma = \alpha_1\beta_1 + \alpha_1\beta_0 + \alpha_0\beta_1 + \alpha_0\beta_0$.
Finally, since the components are connected in parallel, the system is operational unless both components are down (state 11). Therefore, the long-run probability that the system is working equals

$$1 - \pi_{11} = 1 - \frac{\alpha_0\beta_0}{\alpha_1\beta_1 + \alpha_1\beta_0 + \alpha_0\beta_1 + \alpha_0\beta_0}.$$

97.

a. By definition, $p_{ij} = P(Y_{n+1} = j \mid Y_n = i) = P(X_{(n+1)h} = j \mid X_{nh} = i) = P(X_{nh+h} = j \mid X_{nh} = i) = P_{ij}(h)$, where $P_{ij}(h)$ is the conditional transition probability for X_t. As developed in this section, $P_{ij}(h) = q_{ij}h + o(h)$ for $j \neq i$ and $P_{ii}(h) = 1 - q_i h + o(h)$, which is exactly what's claimed.

b. Let the state space of X_t and Y_n be $\{0,1,\ldots,N\}$, and let $\mathbf{P}$ be the transition matrix of Y_n. Then, from the results of part **a**,

$$\mathbf{P} = [p_{ij}] = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0N} \\ p_{10} & p_{11} & & \\ \vdots & & \ddots & \\ p_{N0} & & & p_{NN} \end{bmatrix} = \begin{bmatrix} 1-q_0 h & q_{01}h & \cdots & q_{0N}h \\ q_{10}h & 1-q_1 h & & \\ \vdots & & \ddots & \\ q_{N0}h & & & 1-q_N h \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix} + h \begin{bmatrix} -q_0 & q_{01} & \cdots & q_{0N} \\ q_{10} & -q_1 & & \\ \vdots & & \ddots & \\ q_{N0} & & & -q_N \end{bmatrix} = \mathbf{I} + h\mathbf{Q}$$

where **I** is an identity matrix. Thus, if $\boldsymbol{\pi}$ is a solution to $\boldsymbol{\pi}\mathbf{P} = \mathbf{P}$, then $\boldsymbol{\pi}\mathbf{P} - \mathbf{P} = \mathbf{0} \Rightarrow \boldsymbol{\pi}(\mathbf{P} - \mathbf{I}) = \mathbf{0} \Rightarrow \boldsymbol{\pi}(h\mathbf{Q}) = \mathbf{0} \Rightarrow h\boldsymbol{\pi}\mathbf{Q} = \mathbf{0} \Rightarrow \boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ (and vice versa), QED.

Supplementary Exercises

99.

a. For $n - 1 \le t < n$, $\mu_X(t) = E[X(t)] = E[V_n] = +1(.5) + -1(.5) = 0$. Since this holds for all $n \ge 0$, $\mu_X(t) = 0$.

b. If there exists an integer n such that t and s both lie in $[n - 1, n)$, then $X(t) = X(s) = V_n$ and $C_{XX}(t, s) = \text{Cov}(X(t), X(s)) = \text{Cov}(V_n, V_n) = \text{Var}(V_n) = 1$. If not, then $X(t) = V_n$ and $X(s) = V_m$ for some $n \ne m$, and by independence $C_{XX}(t, s) = \text{Cov}(X(t), X(s)) = \text{Cov}(V_n, V_m) = 0$.
One way to write the condition that t and s lie in the same unit interval is to require that the floor of the two numbers is the same. Thus, we may write

$$C_{XX}(t,s) = \begin{cases} 1 & \lfloor t \rfloor = \lfloor s \rfloor \\ 0 & \text{otherwise} \end{cases}.$$

101.

a. First, by linearity of expectation,

$$\mu_X(t) = E[A_k \cos(\omega_k t) + B_k \sin(\omega_k t)] = E[A_k]\cos(\omega_k t) + E[B_k]\sin(\omega_k t)$$

Since the mean of the Unif[–1, 1] distribution is 0, $\mu_X(t) = 0$. Next, since A_k and B_k are independent,

$$\begin{aligned} C_{XX}(t,s) &= \text{Cov}(A_k \cos(\omega_k t) + B_k \sin(\omega_k t), A_k \cos(\omega_k s) + B_k \sin(\omega_k s)) \\ &= \text{Var}(A_k)\cos(\omega_k t)\cos(\omega_k s) + \text{Var}(B_k)\sin(\omega_k t)\sin(\omega_k s) \\ &= \frac{1}{3}\cos(\omega_k t)\cos(\omega_k s) + \frac{1}{3}\sin(\omega_k t)\sin(\omega_k s) \end{aligned}$$

because the variance of the Unif[–1, 1] distribution is $(1 - -1)^2/12 = 1/3$. Using a trig identity, this can be simplified to $(1/3)\cos(\omega_k\tau)$. So, in particular, $X_k(t)$ is WSS.

b. First, by linearity of expectation, $\mu_Y(t) = E[X_1(t) + \cdots + X_n(t)] = E[X_1(t)] + \cdots + E[X_n(t)] = 0 + \ldots + 0 = 0$. Second, using the distributive property of covariance,

$$C_{YY}(t,s) = \text{Cov}\left(\sum_{k=1}^{n} X_k(t), \sum_{j=1}^{n} X_j(s)\right) = \sum_{k=1}^{n}\sum_{j=1}^{n} \text{Cov}(X_k(t), X_j(s))$$

Since the X_k's are independent random processes, $C_{YY}(t,s) = \sum_{k=1}^{n} \text{Cov}(X_k(t), X_k(s)) = \sum_{k=1}^{n} C_{X_k X_k}(t,s)$.

Using the answer from part **a**, $C_{YY}(t,s) = \sum_{k=1}^{n} \frac{1}{3}\cos(\omega_k \tau) = \frac{1}{3}\sum_{k=1}^{n}\cos(\omega_k \tau)$. In particular, $Y(t)$ is WSS.

103. To evaluate expectations of this mixed distribution (discrete + continuous), condition on Ω first.

a. From Example 7.12, $E[X(t) \mid \Omega = \omega] = E[\cos(\omega t + \Theta)] = 0$. So, $E[X(t)] = E[E(X(t) \mid \Omega)] = E[0] = 0$.

b. From Example 7.12, $E[X(t)X(t + \tau) \mid \Omega = \omega] = E[\cos(\omega t + \Theta)\cos(\omega(t + \tau) + \Theta)] = (1/2)\cos(\omega\tau)$. So, $R_{XX}(\tau) = E[E(X(t)X(t + \tau) \mid \Omega)] = E[(1/2)\cos(\Omega\tau)] = \sum_{k=1}^{n}(1/2)\cos(\omega_k\tau)\cdot p_k = \frac{1}{2}\sum_{k=1}^{n}\cos(\omega_k\tau)\cdot p_k$. Since the mean is zero, $C_{XX}(\tau) = R_{XX}(\tau)$.

c. Since the mean is a constant and the autocorrelation/autocovariance depends only on τ and not t, <u>yes</u>, $X(t)$ is WSS.

105.

a. S_n denotes the sum of the lifetimes of the first n rotors. Equivalently, S_n denotes the total lifetime of the machine through its use of the first n rotors.

b. For an exponential distribution with mean 125 hours, $\mu_X = \sigma_X = 125$. Using the random walk proposition from Section 7.4,
$\mu_S[n] = n\mu_X = 125n$; $\sigma_S^2[n] = n\sigma_X^2 = 15,625n$; $C_{SS}[n, m] = \min(n, m)\,\sigma_X^2 = 15,625\min(n, m)$; and $R_{SS}[n, m] = C_{SS}[n, m] + \mu_S[n]\mu_S[m] = 15,625[\min(n, m) + nm]$.

c. By the Central Limit Theorem, S_{50} is approximately normal; its mean and variance from part **b** are $125(50) = 6250$ and $15625(50) = 781,250$, respectively. Therefore,
$$P(S_{50} \geq 6240) \approx 1-\Phi\left(\frac{6240-6250}{\sqrt{781250}}\right) = 1 - \Phi(-0.01) = .5040.$$

107. The mean function of Y_n is $E[X_n - X_{n-1}] = \mu_X - \mu_X = 0$, since X_n is WSS. Next, the autocovariance of Y_n is
$$\begin{aligned} C_{YY}[n, n+k] &= \text{Cov}(Y_n, Y_{n+k}) = \text{Cov}(X_n - X_{n-1}, X_{n+k} - X_{n+k-1}) \\ &= \text{Cov}(X_n, X_{n+k}) - \text{Cov}(X_n, X_{n+k-1}) - \text{Cov}(X_{n-1}, X_{n+k}) + \text{Cov}(X_{n-1}, X_{n+k-1}) \\ &= C_{XX}[k] - C_{XX}[k-1] - C_{XX}[k+1] + C_{XX}[k] \\ &= 2C_{XX}[k] - C_{XX}[k-1] - C_{XX}[k+1] \end{aligned}$$
In the middle line, the argument of C_{XX} is the <u>difference</u> of the time indices, since X_n is WSS. Since neither the mean nor autocovariance of Y_n depends on n, Y_n <u>is</u> WSS.

109. Let $X(t)$ denote the Poisson process, i.e. the number of noise impulses in t seconds.

a. $P(X(t) = 0) = \dfrac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}$.

b. $P(X(t) \leq 1) = \dfrac{e^{-\lambda t}(\lambda t)^0}{0!} + \dfrac{e^{-\lambda t}(\lambda t)^1}{1!} = e^{-\lambda t}(1+\lambda t)$.

c. Suppose a noise impulse occurs at time t^*. Then the <u>next</u> impulse is corrected provided that it occurs more than ε seconds later; equivalently, 0 impulses occur in $(t^*, t^* + \varepsilon]$. This interval has length ε, so
$$P(X(t^* + \varepsilon) - X(t^*) = 0) = \frac{e^{-\lambda\varepsilon}(\lambda\varepsilon)^0}{0!} = e^{-\lambda\varepsilon}.$$

111. Let $X(t)$ denote the Poisson process, i.e. $X(t)$ = the number of commuters that arrive in t minutes.

a. Let $Y_1, \ldots, Y_5$ denote the arrival times of the five commuters. Conditional on $X(t_0) = 5$, the Y's are uniform rvs on $[0, t_0]$. The discounted fare for the ith customer is $2e^{-\alpha Y_i}$, so the expected value of the total discounted fare collected from all five commuters is

$E\left[2e^{-\alpha Y_1} + \cdots + 2e^{-\alpha Y_5} \mid X(t_0) = 5\right] = \sum_{i=1}^{5} 2E[e^{-\alpha Y_i} \mid X(t_0) = 5] = 5 \cdot 2E[e^{-\alpha Y_1} \mid X(t_0) = 5]$, since the Y's are identically distributed. Continuing, we have $5 \cdot 2\int_0^{t_0} e^{-\alpha y} \dfrac{1}{t_0 - 0} dy = \dfrac{10}{t_0}\int_0^{t_0} e^{-\alpha y}\, dy = \dfrac{10(1-e^{-\alpha t_0})}{\alpha t_0}$.

b. Similarly, $E[\text{total fare} \mid X(t_0) = x] = \dfrac{2x(1-e^{-\alpha t_0})}{\alpha t_0}$. Using the Poisson distribution of $X(t_0)$ and the law of total probability, $E[\text{total fare}] = \sum_{x=0}^{\infty} E[\text{total fare } \mid X(t_0) = x] \cdot P(X(t_0) = x) =$

$$\sum_{x=0}^{\infty} \frac{2x(1-e^{-\alpha t_0})}{\alpha t_0} \cdot \frac{e^{-\lambda t_0}(\lambda t_0)^x}{x!} = \frac{2(1-e^{-\alpha t_0})e^{-\lambda t_0}}{\alpha t_0} \sum_{x=0}^{\infty} \frac{x(\lambda t_0)^x}{x!} = \frac{2(1-e^{-\alpha t_0})e^{-\lambda t_0}}{\alpha t_0} \sum_{x=1}^{\infty} \frac{(\lambda t_0)^x}{(x-1)!}$$

$$= \frac{2(1-e^{-\alpha t_0})e^{-\lambda t_0}\lambda t_0}{\alpha t_0} \sum_{x=1}^{\infty} \frac{(\lambda t_0)^{x-1}}{(x-1)!} = \frac{2(1-e^{-\alpha t_0})e^{-\lambda t_0}\lambda}{\alpha} \sum_{x=1}^{\infty} \frac{(\lambda t_0)^{x-1}}{(x-1)!} = \frac{2(1-e^{-\alpha t_0})e^{-\lambda t_0}\lambda}{\alpha} \sum_{n=0}^{\infty} \frac{(\lambda t_0)^n}{n!} =$$

$$\frac{2(1-e^{-\alpha t_0})e^{-\lambda t_0}\lambda}{\alpha} \cdot e^{\lambda t_0} = \frac{2\lambda}{\alpha}(1-e^{-\alpha t_0}).$$

113. As suggested in the hint, for non-negative integers x and y write

$P(X(t) = x \text{ and } Y(t) = y) = P(X(t) = x \text{ and } M = x + y) = P(M = x + y) \cdot P(X(t) = x \mid M = x + y)$.

The first probability is, by assumption, Poisson: $P(M = x + y) = \dfrac{e^{-\mu}\mu^{x+y}}{(x+y)!}$. As for the second probability, conditional on $M = m$, each of these m loose particles has been released by time t with probability $G(t)$ and not yet released with probability $1 - G(t)$, independent of the other particles. Thus $X(t)$, the number of loose particles not yet released by time t, has a binomial distribution with parameters $n = m$ and $p = 1 - G(t)$ conditional on $M = m$.

Calculate this binomial probability and multiply by the Poisson answer above:

$$P(M = x + y) \cdot P(X(t) = x \mid M = x + y) = \frac{e^{-\mu}\mu^{x+y}}{(x+y)!} \cdot \binom{x+y}{x}(1-G(t))^x G(t)^y = \frac{e^{-\mu}\mu^{x+y}}{x!\,y!}(1-G(t))^x G(t)^y.$$

Then, to find the (marginal) distribution of $X(t)$, eliminate the variable y:

$$p_X(x) = \sum_{y=0}^{\infty} \frac{e^{-\mu}\mu^{x+y}}{x!\,y!}(1-G(t))^x G(t)^y = \frac{e^{-\mu}\mu^x}{x!}(1-G(t))^x \sum_{y=0}^{\infty} \frac{\mu^y}{y!} G(t)^y$$

$$= \frac{e^{-\mu}(\mu[1-G(t)])^x}{x!} \sum_{y=0}^{\infty} \frac{(\mu G(t))^y}{y!} = \frac{e^{-\mu}(\mu[1-G(t)])^x}{x!} \cdot e^{\mu G(t)} = \frac{e^{-\mu[1-G(t)]}(\mu[1-G(t)])^x}{x!}$$

This is precisely the Poisson pmf with parameter $\mu[1 - G(t)]$, as claimed.

CHAPTER 8

Section 8.1

1. If rect(τ) were the autocorrelation function of a WSS random process, then by the Wiener-Khinchin Theorem the power spectral density of that process would be $S_{XX}(f) = \mathscr{F}\{R_{XX}(\tau)\} = \mathscr{F}\{\text{rect}(\tau)\} = \text{sinc}(f)$. However, the sinc function takes on both positive and negative values, and a requirement of all power spectral densities is non-negativity (i.e., $S_{XX}(f) \geq 0$ for all f). Therefore, rect(τ) cannot be an autocorrelation function.

3.

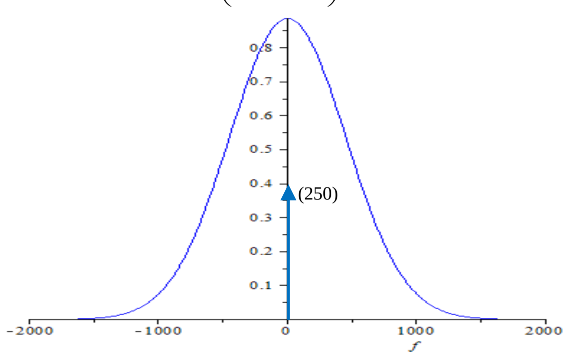
a. By the Wiener-Khinchin Theorem,

$$S_{XX}(f) = \mathscr{F}\{R_{XX}(\tau)\} = \mathscr{F}\{250 + 1000\exp(-4\times10^6\tau^2)\} = 250\mathscr{F}\{1\} + 1000\mathscr{F}\{\exp(-[2000\tau]^2)\}$$

For the second transform, use the line of the Fourier table for $g(t) = \exp(-t^2)$ along with the rescaling property $\mathscr{F}\{g(at)\} = 1/|a|G(f/a)$ with $a = 2000$:

$$250\mathscr{F}\{1\} + 1000\mathscr{F}\{\exp(-[2000\tau]^2)\} = 250\delta(f) + 1000\cdot\frac{1}{|2000|}\sqrt{\pi}\exp(-\pi^2[f/2000]^2)$$

$$= 250\delta(f) + \frac{\sqrt{\pi}}{2}\exp\left(-\frac{\pi^2 f^2}{4\times10^6}\right)$$



b. The impulse (dc power) of 250 W at $f = 0$ is not included in this interval. Using that and the symmetry of the psd,

$$P_X[500,1000] = 2\int_{500}^{1000} S_{XX}(f)\,df = 2\int_{500}^{1000}\frac{\sqrt{\pi}}{2}\exp\left(-\frac{\pi^2 f^2}{4\times10^6}\right)df = \sqrt{\pi}\int_{500}^{1000}\exp\left(-\frac{\pi^2 f^2}{4\times10^6}\right)df = 240.37\text{ W}$$

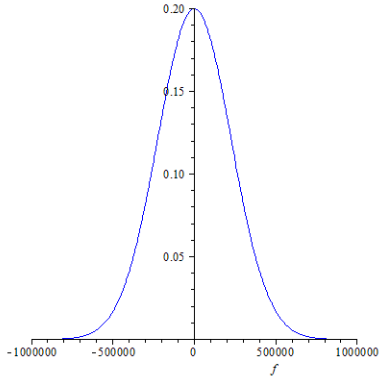
The last integral was evaluated using software.

c. The impulse at $f = 0$ <u>is</u> included here. Account for that 250 W of expected power, then use symmetry to calculate the rest:

$$P_X[0,200] = \int_{-200}^{200} S_{XX}(f)\,df = 250 + 2\int_{0}^{200}\frac{\sqrt{\pi}}{2}\exp\left(-\frac{\pi^2 f^2}{4\times10^6}\right)df = 250 + 343.17 = 593.17\text{ W.}$$

5.

a.



Using software, $P_X = 2\int_0^\infty 0.2\exp\left(-\dfrac{\pi^2 f^2}{10^{12}}\right)df$ = 112,838 W.

b. Using software, $P_X[10000,\infty) = 2\int_{10000}^\infty 0.2\exp\left(-\dfrac{\pi^2 f^2}{10^{12}}\right)df$ = 108,839 W.

c. To find the inverse Fourier transform of $S_{XX}(f)$, use the pair $e^{-t^2} \leftrightarrow \sqrt{\pi}e^{-\pi^2 f^2}$. The duality property of Fourier transforms allows us to switch sides of this pair: $e^{-f^2} \leftrightarrow \sqrt{\pi}e^{-\pi^2(-t)^2} = \sqrt{\pi}e^{-\pi^2 t^2}$. Then, use the rescaling property with $a = \pi/10^6$:

$$R_{XX}(\tau) = \mathscr{F}^{-1}\{S_{XX}(f)\} = \mathscr{F}^{-1}\left\{0.2\exp\left(-\frac{\pi^2 f^2}{10^{12}}\right)\right\} = \mathscr{F}^{-1}\left\{0.2\exp\left(-\left[\frac{\pi f}{10^6}\right]^2\right)\right\}$$

$$= 0.2\cdot\frac{1}{|\pi/10^6|}\sqrt{\pi}\exp\left(-\pi^2\left[\frac{\tau}{\pi/10^6}\right]^2\right) = \frac{200,000}{\sqrt{\pi}}\exp(-10^{12}\tau^2)$$

Therefore, $P_X = R_{XX}(0) = \dfrac{200,000}{\sqrt{\pi}} \approx$ 112,838 W, confirming the answer to **a**.

7.

a. $P_N = \int_{-\infty}^{\infty} S_{NN}(f)df = \int_{-B}^{B}\dfrac{N_0}{2}df = \dfrac{N_0}{2}(B - -B) = N_0 B$. (As a check, the units on N_0 are W/Hz and the units on B are Hz, so the units on power are (W/Hz)Hz = W, as they should be.)

b. The psd of $N(t)$ is actually a rectangle; specifically, $S_{NN}(f) = \dfrac{N_0}{2}\mathrm{rect}\left(\dfrac{f}{2B}\right)$. The denominator $2B$ insures that the width of the rectangle is $2B$. Take the inverse Fourier transform, using the rescaling property when required:

$$R_{XX}(\tau) = \mathscr{F}^{-1}\{S_{NN}(f)\} = \mathscr{F}^{-1}\left\{\frac{N_0}{2}\mathrm{rect}\left(\frac{f}{2B}\right)\right\} = \frac{N_0}{2}\cdot|2B|\,\mathrm{sinc}\left(\frac{\tau}{1/(2B)}\right)$$

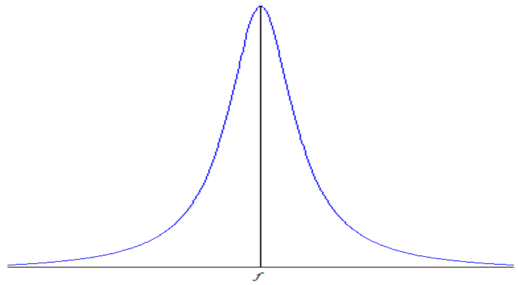
$$= N_0 B\,\mathrm{sinc}(2B\tau)$$

9.

a. From Section 7.5, the autocorrelation function of a Poisson telegraphic process $N(t)$ is $R_{NN}(\tau) = e^{-2\lambda|\tau|}$. So, the autocorrelation function of $Y(t) = A_0N(t)$ is $R_{YY}(t,t+\tau) = E[Y(t)Y(t+\tau)] = E[A_0N(t)A_0N(t+\tau)] = A_0^2E[N(t)N(t+\tau)] = A_0^2R_{NN}(\tau) = A_0^2e^{-2\lambda|\tau|}$. (In particular, $Y(t)$ is also WSS, since it's just a constant times a WSS process.)

b. By the Wiener-Khinchin Theorem,

$$S_{YY}(f) = \mathscr{F}\{R_{YY}(\tau)\} = \mathscr{F}\{A_0^2e^{-2\lambda|\tau|}\} = A_0^2 \cdot \frac{2\lambda}{\lambda^2+(2\pi f)^2} = \frac{2\lambda A_0^2}{\lambda^2+(2\pi f)^2}.$$



c. $P_Y = R_{YY}(0) = A_0^2e^{-2\lambda|0|} = A_0^2$.

d. $$P_Y[0,\lambda] = 2\int_0^{\lambda}\frac{2\lambda A_0^2}{\lambda^2+(2\pi f)^2}df = 4\lambda A_0^2\int_0^{2\pi\lambda}\frac{1}{\lambda^2+u^2}\frac{du}{2\pi} = \frac{4\lambda A_0^2}{2\pi}\cdot\frac{1}{\lambda}\arctan\left(\frac{u}{\lambda}\right)\Bigg|_0^{2\pi\lambda} = \frac{2A_0^2}{\pi}\arctan(2\pi).$$

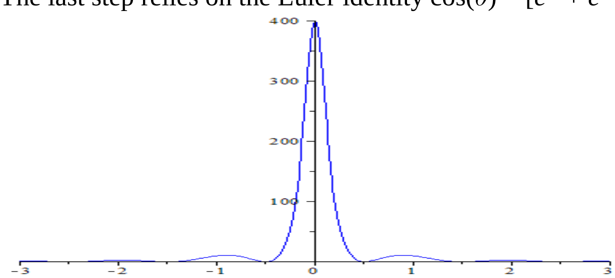
11.

a. $P_X = R_{XX}(0) = 100e^0 + 50e^{-1} + 50e^{-1} = 100(1 + e^{-1}) \approx 136.8$ W.

b. For two of the terms in the Fourier transform, use the time shift property $\mathscr{F}\{g(t - t_0)\} = G(f)e^{-j2\pi f t_0}$. By the Wiener-Khinchin Theorem,

$$\begin{aligned} S_{XX}(f) &= \mathscr{F}\{R_{XX}(\tau)\} = \mathscr{F}\{100e^{-|\tau|} + 50e^{-|\tau-1|} + 50e^{-|\tau+1|}\} \\ &= 100\cdot\frac{2}{1+(2\pi f)^2} + 50\cdot\frac{2}{1+(2\pi f)^2}\cdot e^{-j2\pi f(1)} + 50\cdot\frac{2}{1+(2\pi f)^2}\cdot e^{-j2\pi f(-1)} \\ &= 100\cdot\frac{2}{1+(2\pi f)^2}\left[1+\frac{e^{-j2\pi f}+e^{j2\pi f}}{2}\right] = \frac{200}{1+(2\pi f)^2}[1+\cos(2\pi f)] \end{aligned}$$

The last step relies on the Euler identity $\cos(\theta) = [e^{j\theta} + e^{-j\theta}]/2$.



c. $P_X[0,1] = \int_{-1}^{1}S_{XX}(f)df = 2\int_0^1\frac{200}{1+(2\pi f)^2}[1+\cos(2\pi f)]df = 126.34$ W with the aid of software.

13. If $X(t)$ is WSS, then its mean is a constant; i.e., $E[X(t)] = \mu_X$ for some constant μ_X for all t. Thus, in particular, $\mu_W(t) = E[W(t)] = E[X(t) - X(t - d)] = \mu_X - \mu_X = 0$.

Next, $R_{XX}(\tau) = E[X(t)X(t + \tau)]$ depends only on the difference τ between the two time arguments. Thus,

$$\begin{aligned}R_{WW}(t,t+\tau) &= E[W(t)W(t+\tau)] = E[(X(t)-X(t-d))(X(t+\tau)-X(t+\tau-d))]\\ &= E[X(t)X(t+\tau)] - E[X(t)X(t+\tau-d)] - E[X(t-d)X(t+\tau)] + E[X(t-d)X(t+\tau-d)]\\ &= R_{XX}(\tau) - R_{XX}(\tau-d) - R_{XX}(t+\tau-[t-d]) + R_{XX}(t+\tau-d-[t-d])\\ &= 2R_{XX}(\tau) - R_{XX}(\tau-d) - R_{XX}(\tau+d)\end{aligned}$$

In particular, neither the mean nor autocorrelation of $W(t)$ depend on t, so $W(t)$ is WSS.

Third, apply the Wiener-Khinchin Theorem and the time-shift property:

$$\begin{aligned}S_{WW}(f) &= \mathscr{F}\{R_{WW}(\tau)\} = \mathscr{F}\{2R_{XX}(\tau) - R_{XX}(\tau-d) - R_{XX}(\tau+d)\}\\ &= 2S_{XX}(f) - S_{XX}(f)e^{-j2\pi fd} - S_{XX}(f)e^{-j2\pi f(-d)}\\ &= 2S_{XX}(f)\left[1 - \frac{e^{-j2\pi fd} + e^{j2\pi fd}}{2}\right] = 2S_{XX}(f)[1-\cos(2\pi fd)]\end{aligned}$$

The last step relies on the Euler identity $\cos(\theta) = [e^{j\theta} + e^{-j\theta}]/2$.

15.

a. Yes, $X(t)$ and $Y(t)$ are jointly WSS. By definition, if $X(t)$ and $Y(t)$ are orthogonal, then $E[X(t)Y(s)] = 0$ for all times t and s. Thus, in particular, $R_{XY}(\tau) = E[X(t)Y(t + \tau)] = 0$. Since $X(t)$ and $Y(t)$ are each individually WSS and the cross-correlation of $X(t)$ and $Y(t)$ does not depend on t, by definition $X(t)$ and $Y(t)$ are jointly WSS.

b. Check the mean and autocorrelation of $Z(t)$:
$\mu_Z(t) = E[Z(t)] = E[X(t)+Y(t)] = E[X(t)] + E[Y(t)] = \mu_X + \mu_Y$, the sum of two constants. Hence, the mean of $Z(t)$ is a constant.

$$\begin{aligned}R_Z(t,t+\tau) &= E[Z(t)Z(t+\tau)] = E[(X(t)+Y(t))(X(t+\tau)+Y(t+\tau))]\\ &= E[X(t)X(t+\tau)] + E[X(t)Y(t+\tau)] + E[Y(t)X(t+\tau)] + E[Y(t)Y(t+\tau)]\\ &= R_{XX}(\tau) + 0 + 0 + R_{YY}(\tau) = R_{XX}(\tau) + R_{YY}(\tau)\end{aligned}$$

The middle two terms are 0 because $X(t)$ and $Y(t)$ are orthogonal. Since neither the mean nor autocorrelation depends on t, $Z(t)$ is indeed WSS.

c. Since $Z(t)$ is WSS, apply the Wiener-Khinchin Theorem:
$S_{ZZ}(f) = \mathscr{F}\{R_{ZZ}(\tau)\} = \mathscr{F}\{R_{XX}(\tau) + R_{YY}(\tau)\} = \mathscr{F}\{R_{XX}(\tau)\} + \mathscr{F}\{R_{YY}(\tau)\} = S_{XX}(f) + S_{YY}(f)$.

17.

a. First, $\mu_Z(t) = E[Z(t)] = E[X(t)Y(t)] = E[X(t)]E[Y(t)]$, where the last equality relies on the independence of $X(t)$ and $Y(t)$. Since each is WSS, we have $\mu_Z(t) = \mu_X \cdot \mu_Y$. Second,

$$\begin{aligned} R_{ZZ}(t,t+\tau) &= E[Z(t)Z(t+\tau)] = E[X(t)Y(t)X(t+\tau)Y(t+\tau)] = E[X(t)X(t+\tau)Y(t)Y(t+\tau)] \\ &= E[X(t)X(t+\tau)]E[Y(t)Y(t+\tau)] = R_{XX}(\tau)\cdot R_{YY}(\tau) \end{aligned}$$

where the move from the first line to the second line is again justified by independence. In particular, neither the mean nor autocorrelation of $Z(t)$ depends on t, so $Z(t)$ is WSS.

b. $S_{ZZ}(f) = \mathscr{F}\{R_{ZZ}(\tau)\} = \mathscr{F}\{R_{XX}(\tau)\cdot R_{YY}(\tau)\} = \mathscr{F}\{R_{XX}(\tau)\} \star \mathscr{F}\{R_{YY}(\tau)\} = S_{XX}(f) \star S_{YY}(f)$. That is, the psd of $Z(t)$ is the convolution of the psds of $X(t)$ and $Y(t)$.

19. <u>No</u>. Highpass white noise, like pure white noise, in not physically realizable because the model implies infinite power: $P_N = \int_{-\infty}^{\infty} S_{NN}(f)df = 2\int_B^{\infty} \frac{N_0}{2} df = +\infty$.

21.

a. From the referenced exercise, the autocorrelation function of $X(t)$ is $R_{XX}(\tau) = E[A^2]R_{YY}(\tau)$, where $E[A^2]$ is the mean-square value of A and, in particular, a constant. Thus,

$S_{XX}(f) = \mathscr{F}\{R_{XX}(\tau)\} = \mathscr{F}\{E[A^2]R_{YY}(\tau)\} = E[A^2]\mathscr{F}\{R_{YY}(\tau)\} = E[A^2]S_{YY}(f)$.

b. From the previous exercise, $S_{XX}^{\text{ac}}(f) = S_{XX}(f) - \mu_X^2\delta(f)$. We can determine the mean of $X(t)$:

$\mu_X = E[X(t)] = E[A\cdot Y(t)] = E[A]E[Y(t)] = E[A]\mu_Y$. Combine this with part **a**:

$S_{XX}^{\text{ac}}(f) = S_{XX}(f) - \mu_X^2\delta(f) = E[A^2]S_{YY}(f) - (E[A]\mu_Y)^2\delta(f)$.

c. In the answer to **b**, use the result of Exercise 20 again, but apply it this time to the psd of $Y(t)$:

$$\begin{aligned} S_{XX}^{\text{ac}}(f) &= E[A^2]S_{YY}(f) - (E[A]\mu_Y)^2\delta(f) \\ &= E[A^2](S_{YY}^{\text{ac}}(f) + \mu_Y^2\delta(f)) - (E[A])^2\mu_Y^2\delta(f) \\ &= E[A^2]S_{YY}^{\text{ac}}(f) + (E[A^2] - (E[A])^2)\mu_Y^2\delta(f) = E[A^2]S_{YY}^{\text{ac}}(f) + \sigma_A^2\mu_Y^2\delta(f) \end{aligned}$$

Notice that last term in the ac-psd of $X(t)$ is indeed an impulse at $f = 0$, which theoretically corresponds to the <u>dc</u> power offset of a WSS signal. The lesson here is that our interpretation of the elements of a power spectral density found via the Wiener-Khinchin Theorem, and specifically the idea of interpreting $a \cdot \delta(f)$ the dc power offset, is not sensible for a non-ergodic process such as $X(t)$ here. (To be clear, ergodicity is not a hypothesis of the Wiener-Khinchin Theorem — we may apply that theorem to WSS non-ergodic processes — but the "engineering interpretation" of the elements of a psd are not valid for non-ergodic processes.)

Section 8.2

23.

a. The power spectral density of $X(t)$ is actually a rectangle; specifically, $S_{XX}(f) = 0.02\text{rect}\left(\frac{f}{120,000}\right)$.

The denominator of 120,000 insures that the rectangle spans [–60000, 60000]. Thus,

$$R_{XX}(\tau) = \mathscr{F}^{-1}\{S_{XX}(f)\} = \mathscr{F}^{-1}\left\{0.02\text{rect}\left(\frac{f}{120,000}\right)\right\} = 0.02 \cdot 120,000\text{sinc}(120,000\tau)$$
$$= 2400\text{sinc}(120,000\tau).$$

b. From the autocorrelation function, $P_X = R_{XX}(0) = 2400$ W. Or, using the power spectral density, $P_X = 2\int_0^{60,000} 0.02\,df = 2400$ W.

c. Apply the transform pair for $t^k e^{-at}u(t)$ with $k = 0$ and $a = 40$:

$$H(f) = \mathscr{F}\{h(t)\} = \mathscr{F}\{40e^{-40t}u(t)\} = 40 \cdot \frac{0!}{(40 + j2\pi f)^{0+1}} = \frac{40}{40 + j2\pi f} \ .$$

d. First, the power transfer function is $|H(f)|^2 = \left|\frac{40}{40 + j2\pi f}\right|^2 = \frac{|40|^2}{|40 + j2\pi f|^2} = \frac{1600}{1600 + (2\pi f)^2}$. Then, the output power spectral density is

$$S_{YY}(f) = S_{XX}(f)|H(f)|^2 = 0.02\text{rect}\left(\frac{f}{120,000}\right)\frac{1600}{1600 + (2\pi f)^2} = \frac{32}{1600 + (2\pi f)^2}\text{rect}\left(\frac{f}{120,000}\right)$$
$$= \begin{cases} \dfrac{32}{1600 + (2\pi f)^2} & |f| \le 60 \text{ kHz} \\ 0 & \text{otherwise} \end{cases}$$

The psd (not shown) is bell-shaped and limited by ±60,000; however, the curve is so low there that the band limit makes essentially no difference.

e. $P_Y = \int_{-\infty}^{\infty} S_{YY}(f)\,df = \int_{-60000}^{60000} \frac{32}{1600 + (2\pi f)^2}\,df = 0.39997$ W. (The integral of this function from $-\infty$ to ∞ is exactly 0.4.)

25.

a. A white noise process has mean 0. Hence, $\mu_Y = \mu_X \int h(t)dt = 0$.

b. Apply the definition of the Fourier transform:

$$H(f) = \mathscr{F}\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}\,dt = \int_0^1 1e^{-j2\pi ft}\,dt = \frac{1}{-j2\pi f}e^{-j2\pi ft}\Big|_0^1 = \frac{1-e^{-j2\pi f}}{j2\pi f}.$$

Alternatively, the function $h(t) = 1$ on $[0, 1)$ is essentially the same as $h(t) = \text{rect}(t - 1/2)$, so $H(f) = \mathscr{F}\{\text{rect}(t - 1/2)\} = \text{sinc}(f) \cdot e^{-j2\pi f(1/2)} = \text{sinc}(f) \cdot e^{-j\pi f}$. Though not obvious at first glance, these are actually the same function (apply Euler's sine identity and the definition of sinc to see why).

c. Use the second expression above to find the power transfer function: $|H(f)|^2 = |\text{sinc}(f) \cdot e^{-j\pi f}|^2 = \text{sinc}^2(f)$, since the magnitude of $e^{j\theta}$ is always 1. Then the psd of $Y(t)$ is

$$S_{YY}(f) = S_{XX}(f)\,|H(f)|^2 = \frac{N_0}{2}\text{sinc}^2(f).$$

d. Using the well-known engineering fact that the area under $\text{sinc}^2(x)$ is 1,

$$P_Y = \int_{-\infty}^{\infty} S_{YY}(f)df = \int_{-\infty}^{\infty} \frac{N_0}{2}\text{sinc}^2(f)df = \frac{N_0}{2}(1) = \frac{N_0}{2}.$$

27.

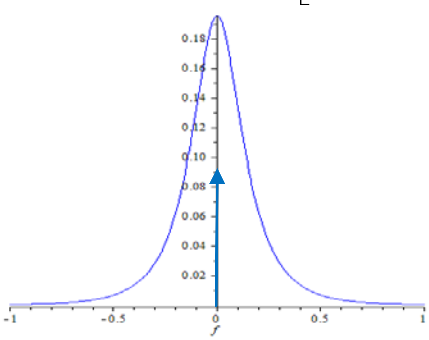
a. $S_{XX}(f) = \mathscr{F}\{R_{XX}(\tau)\} = \mathscr{F}\{100 + 25e^{-|\tau|}\} = 100\delta(f) + 25 \cdot \frac{2}{1+(2\pi f)^2} = 100\delta(f) + \frac{50}{1+(2\pi f)^2}$.

b. $P_X = R_{XX}(0) = 100 + 25 = 125$ W.

c. With $k = 1$ and $a = 4$, $H(f) = \mathscr{F}\{h(t)\} = \mathscr{F}\{te^{-4t}u(t)\} = \frac{1!}{(4 + j2\pi f)^{1+1}} = \frac{1}{(4 + j2\pi f)^2}$. The power transfer function is $|H(f)|^2 = \left|\frac{1}{(4+j2\pi f)^2}\right|^2 = \frac{|1|^2}{(|4 + j2\pi f|)^2} = \frac{1}{(16 + (2\pi f)^2)^2}$.

d. The power spectrum of $Y(t)$ is

$$S_{YY}(f) = S_{XX}(f)\,|H(f)|^2 = \left[100\delta(f) + \frac{50}{1+(2\pi f)^2}\right] \cdot \frac{1}{(16+(2\pi f)^2)^2}$$



e. The psd of $Y(t)$ consists of two parts. For the first part, use the sifting property of delta functions:
$\int_{-\infty}^{\infty} 100\delta(f)\cdot\frac{1}{(16+(2\pi f)^2)^2}df = 100\cdot\frac{1}{(16+(2\pi 0)^2)^2} = \frac{100}{256} = \frac{25}{64}$.
For the second part, use software (or some very unpleasant partial fractions!):
$\int_{-\infty}^{\infty}\frac{50}{1+(2\pi f)^2}\cdot\frac{1}{(16+(2\pi f)^2)^2}df = \int_{-\infty}^{\infty}\frac{50}{1+x^2}\cdot\frac{1}{(16+x^2)^2}\frac{dx}{2\pi} = \cdots = \frac{9}{128}$
Combined, $P_Y = \frac{25}{64}+\frac{9}{128} = \frac{59}{128} \approx 0.461$ W.

29.

a. $S_{YY}(f) = S_{XX}(f)\,|H(f)|^2 = \frac{N_0}{2}|e^{-\alpha|f|}|^2 = \frac{N_0}{2}e^{-2\alpha|f|}$.

b. Use the Fourier pair for $e^{-a|t|}$ along with the duality property:
$R_{YY}(\tau) = \mathscr{F}^{-1}\{S_{YY}(f)\} = \mathscr{F}^{-1}\left\{\frac{N_0}{2}e^{-2\alpha|f|}\right\} = \frac{N_0}{2}\cdot\frac{2(2\alpha)}{(2\alpha)^2+(2\pi[-\tau])^2} = \frac{2N_0\alpha}{4\alpha^2+4\pi^2\tau^2}$.

c. $P_Y = R_{YY}(0) = \frac{2N_0\alpha}{4\alpha^2+4\pi^2(0)^2} = \frac{N_0}{2\alpha}$.

31.

a. To implement the hint, use the convolution formula $X(t)\star h(t) = \int_{-\infty}^{\infty} X(t-s)h(s)ds$:

$$R_{XY}(\tau) = E[X(t)Y(t+\tau)] = E\left[X(t)(X(t+\tau)\star h(t+\tau))\right]$$
$$= E\left[X(t)\int_{-\infty}^{\infty} X([t+\tau]-s)h(s)ds\right] = E\left[\int_{-\infty}^{\infty} X(t)X(t+\tau-s)h(s)ds\right]$$
$$= \int_{-\infty}^{\infty} E[X(t)X(t+\tau-s)]h(s)ds \quad \text{exchanging the integral and expectation; } h(s) \text{ is not random}$$
$$= \int_{-\infty}^{\infty} R_{XX}([t+\tau-s]-t)h(s)ds = \int_{-\infty}^{\infty} R_{XX}(\tau-s)h(s)ds$$

This last integral is, by definition, the convolution of $R_{XX}(\tau)$ and $h(\tau)$, as claimed.

b. Begin with the definition of autocorrelation and the hint:

$$R_{YY}(\tau) = E[Y(t)Y(t+\tau)] = E[(X(t)\star h(t))Y(t+\tau)]$$
$$= E\left[\left(\int_{-\infty}^{\infty} X(t-u)h(u)du\right)Y(t+\tau)\right] = E\left[\int_{-\infty}^{\infty} X(t-u)Y(t+\tau)h(u)du\right]$$
$$= \int_{-\infty}^{\infty} E[X(t-u)Y(t+\tau)]h(u)du = \int_{-\infty}^{\infty} R_{XY}([t+\tau]-[t-u])h(u)du$$
$$= \int_{-\infty}^{\infty} R_{XY}(\tau+u)h(u)du$$

To make this resemble a convolution integral, make the substitution $s = -u$:
$R_{YY}(\tau) = \int_{\infty}^{-\infty} R_{XY}(\tau-s)h(-s)(-ds) = \int_{-\infty}^{\infty} R_{XY}(\tau-s)h(-s)ds = R_{XY}(\tau)\star h(-\tau)$. Substitute the result of part **a**: $R_{YY}(\tau) = (R_{XX}(\tau)\star h(\tau))\star h(-\tau) = R_{XX}(\tau)\star h(\tau)\star h(-\tau)$, which completes the proof.

33. The psd of $X(t)$ is $S_{XX}(f) = N_0/2$ for $|f| \le B$; i.e., $S_{XX}(f) = \frac{N_0}{2}\text{rect}\left(\frac{f}{2B}\right)$.

a. $S_{YY}(f) = S_{XX}(f)\,|H(f)|^2 = \frac{N_0}{2}\text{rect}\left(\frac{f}{2B}\right)|\,j2\pi f\,|^2 = 2N_0\pi^2 f^2 \text{rect}\left(\frac{f}{2B}\right)$. Equivalently,

$S_{YY}(f) = 2N_0\pi^2 f^2$ for $|f| \le B$ and equals zero otherwise.

b. There is no (tabulated) inverse Fourier transform for a band-limited polynomial, so rely on the definition of the transform:

$$R_{YY}(\tau) = \mathscr{F}^{-1}\{S_{YY}(f)\} = \int_{-\infty}^{\infty} S_{YY}(f)e^{j2\pi f\tau}\,df$$

$$= \int_{-\infty}^{\infty} S_{YY}(f)\cos(2\pi f\tau)df \quad \text{because } S_{XX}(f) \text{ is an even function}$$

$$= \int_{-B}^{B} 2N_0\pi^2 f^2 \cos(2\pi f\tau)df = 4N_0\pi^2\int_0^B f^2\cos(2\pi f\tau)df$$

This integral requires integration by parts twice. The final answer is

$$R_{YY}(\tau) = \frac{N_0}{\pi\tau^3}\left[2\pi^2B^2\tau^2\sin(2\pi B\tau) + 2\pi B\tau\cos(2\pi B\tau) - \sin(2\pi B\tau)\right]$$

c. While $R_{YY}(0)$ can be determined by taking a limit as $\tau \to 0$ using l'Hopital's rule, it's actually easier to find the area under the power spectral density of the output $Y(t)$:

$$P_Y = \int_{-\infty}^{\infty} S_{YY}(f)df = 2\int_0^B 2N_0\pi^2 f^2\,df = \frac{4N_0\pi^2}{3}f^3\Big|_0^B = \frac{4N_0\pi^2B^3}{3}$$

35.

a. First, apply the definition of autocorrelation and distribute:

$$R_{DD}(\tau) = E[D(t)D(t+\tau)] = E[(X(t)-Y(t))(X(t+\tau)-Y(t+\tau))]$$

$$= E[X(t)X(t+\tau)] - E[X(t)Y(t+\tau)] - E[X(t+\tau)Y(t)] + E[Y(t)Y(t+\tau)]$$

$$= R_{XX}(\tau) - R_{XY}(\tau) - R_{XY}(-\tau) + R_{YY}(\tau)$$

Next, apply the convolution formulas for the last three terms:

$$R_{DD}(\tau) = R_{XX}(\tau) - R_{XX}(\tau)\star h(\tau) - R_{XX}(-\tau)\star h(-\tau) + R_{XX}(\tau)\star h(\tau)\star h(-\tau)$$

$$= R_{XX}(\tau) - R_{XX}(\tau)\star h(\tau) - R_{XX}(\tau)\star h(-\tau) + R_{XX}(\tau)\star h(\tau)\star h(-\tau)$$

The second step uses the fact that $R_{XX}(-\tau) = R_{XX}(\tau)$ for any autocorrelation function.

b. Apply the Wiener-Khinchin Theorem, along with the convolution-multiplication property of Fourier transforms. Critically, we use the fact that since $h(t)$ is real-valued, $\mathscr{F}\{h(-t)\} = H(-f) = H^*(f)$:

$$S_{DD}(f) = \mathscr{F}\{R_{XX}(\tau) - R_{XX}(\tau)\star h(\tau) - R_{XX}(\tau)\star h(-\tau) + R_{XX}(\tau)\star h(\tau)\star h(-\tau)\}$$

$$= S_{XX}(f) - S_{XX}(f)\cdot H(f) - S_{XX}(f)\cdot H(-f) + S_{XX}(f)\cdot H(f)\cdot H(-f)$$

$$= S_{XX}(f) - S_{XX}(f)\cdot H(f) - S_{XX}(f)\cdot H^*(f) + S_{XX}(f)\cdot H(f)\cdot H^*(f)$$

$$= S_{XX}(f)[1 - H(f) - H^*(f) + H(f)\cdot H^*(f)]$$

$$= S_{XX}(f)[1-H(f)][1-H^*(f)] = S_{XX}(f)\left|1-H(f)\right|^2$$

The last step is justified by the fact that for any complex number, $z \cdot z^* = |z|^2$. It is clear from the last expression that $S_{DD}(f)$ is both real-valued and non-negative. To see why $S_{DD}(f)$ is symmetric in f, we again need the fact that $H(-f) = H^*(f)$ for any transfer function (because it's the Fourier transform of a real-valued function. Thus,

$$S_{DD}(-f) = S_{XX}(-f)\left|1-H(-f)\right|^2 = S_{XX}(f)\left|1-H^*(f)\right|^2 = S_{XX}(f)\left|1-H(f)\right|^2 = S_{DD}(f),$$

where the second-to-last equality comes from the fact that $|z| = |z^*|$.

37.

a. $P_X = R_{XX}(0) = 250{,}000 + 120{,}000 + 800{,}000 = 1{,}170{,}000$ W, or 1.17 MW.

b.

$$S_{XX}(f) = \mathscr{F}\{R_{XX}(\tau)\} = 250{,}000\delta(f) + 120{,}000 \cdot \frac{1}{2}[\delta(f - 35{,}000) + \delta(f + 35{,}000)] + \frac{800{,}000}{100{,}000}\text{rect}\left(\frac{f}{100{,}000}\right)$$

$$= 250{,}000\delta(f) + 60{,}000[\delta(f - 35{,}000) + \delta(f + 35{,}000)] + 8\text{rect}\left(\frac{f}{100{,}000}\right)$$

A graph of the psd (not shown) consists of a rectangle on [–50000, 50000] along with three impulses at $f = 0$ (intensity 250,000) and $f = \pm$ 35000 (intensity 60,000 each).

c. Since the power spectrum of $X(t)$ is contained in the interval $|f| \le 50{,}000$ Hz and the band limit of the ideal filter is 60,000 Hz, all of $X(t)$ passes through the filter unchanged. So, the psd of $L[X(t)]$ is the same as the answer to **b**.

d. By part **c**, $P_{L[X]} = P_X = 1.17$ MW.

e. $P_N = \int_{-\infty}^{\infty} S_{NN}(f)df = \int_{-100{,}000}^{100{,}000} 2.5\times 10^{-2}\, df = 5000$ W.

f. The power spectral density of $L[N(t)]$ will equal 2.5×10^{-2} W/Hz on the preserved band, $|f| \le 60$ kHz. So, $P_{L[N]} = \int_{-60{,}000}^{60{,}000} 2.5\times 10^{-2}\, df = 3000$ W.

g. The input and output power signal-to-noise ratios are $\text{SNR}_{\text{in}} = \frac{P_X}{P_N} = \frac{1{,}170{,}000}{5000} = 234$ and $\text{SNR}_{\text{out}} = \frac{P_{L[X]}}{P_{L[N]}} = \frac{1{,}170{,}000}{3000} = 390$. The filter achieved roughly a 67% improvement in signal-to-noise ratio, with no signal loss. However, the filter could be improved: since the upper limit of the signal frequency band is 50 kHz, an ideal lowpass filter on $|f| \le 50$ kHz would have filtered out even more noise while still preserving the signal.

Section 8.3

39. By Equation (8.6), $S_{XX}(F)=\sum_{k=-\infty}^{+\infty}R_{XX}[k]e^{-j2\pi Fk}=\sum_{k=-\infty}^{-1}R_{XX}[k]e^{-j2\pi Fk}+R_{XX}[0]e^{-j2\pi F0}+\sum_{k=1}^{+\infty}R_{XX}[k]e^{-j2\pi Fk}$. The middle term is $R_{XX}[0]$. For the left sum, make the substitution $k\to -k$:

$\sum_{k=-\infty}^{-1}R_{XX}[k]e^{-j2\pi Fk}=\sum_{k=1}^{\infty}R_{XX}[-k]e^{-j2\pi F(-k)}=\sum_{k=1}^{\infty}R_{XX}[-k]e^{+j2\pi Fk}=\sum_{k=1}^{\infty}R_{XX}[k]e^{+j2\pi Fk}$ because R_{XX} is always an even function. Now, combine this sum with the third term in $S_{XX}(F)$:

$$S_{XX}(F)=R_{XX}[0]+\sum_{k=1}^{\infty}R_{XX}[k]e^{+j2\pi Fk}+\sum_{k=1}^{\infty}R_{XX}[k]e^{-j2\pi Fk}=R_{XX}[0]+\sum_{k=1}^{\infty}R_{XX}[k](e^{+j2\pi Fk}+e^{-j2\pi Fk})$$

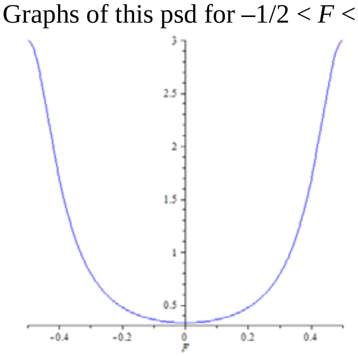
$$=R_{XX}[0]+\sum_{k=1}^{\infty}R_{XX}[k]\cdot 2\cos(2\pi Fk)=R_{XX}[0]+2\sum_{k=1}^{\infty}R_{XX}[k]\cos(2\pi kF)$$

We've used Euler's identity $\cos(\theta)=[e^{j\theta}+e^{-j\theta}]/2$.

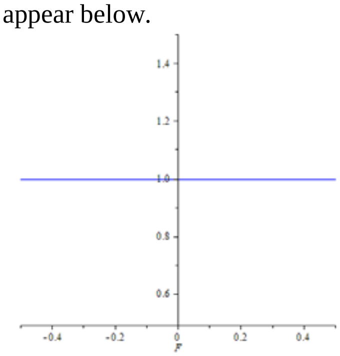
41. The discrete-time Fourier transform of $\alpha^{|n|}$ for $|\alpha|<1$ is given in Appendix B:

$$S_{XX}(F)=\sum_{k=-\infty}^{+\infty}R_{XX}[k]e^{-j2\pi Fk}=\frac{1-\alpha^2}{1+\alpha^2-2\alpha\cos(2\pi F)}$$

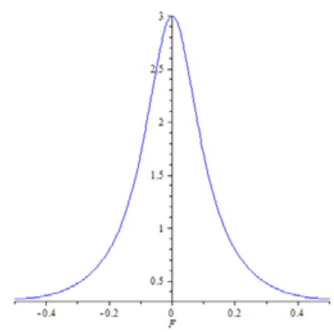
Graphs of this psd for $-1/2<F<1/2$ appear below.



$\alpha=-.5$



$\alpha=0$



$\alpha=.5$

43. With $T_s=5$, the autocorrelation of $X[k]$ and $N(t)$ are related by $R_{XX}[k]=R_{NN}(5k)$. Since the autocorrelation of Poisson telegraphic noise is $R_{NN}(\tau)=e^{-2\lambda|\tau|}$, $R_{XX}[k]=e^{-2\lambda|5k|}=e^{-10\lambda|k|}=(e^{-10\lambda})^{|k|}$. Apply the result of Exercise 41, with $\alpha=e^{-10\lambda}$:

$$S_{XX}(F)=\sum_{k=-\infty}^{+\infty}R_{XX}[k]e^{-j2\pi Fk}=\sum_{k=-\infty}^{+\infty}(e^{-10\lambda})^{|k|}e^{-j2\pi Fk}=\frac{1-(e^{-10\lambda})^2}{1+(e^{-10\lambda})^2-2(e^{-10\lambda})\cos(2\pi F)}$$

$$=\frac{1-e^{-20\lambda}}{1+e^{-20\lambda}-2e^{-10\lambda}\cos(2\pi F)}$$

45. It is extremely difficult to determine the discrete-time Fourier transform of the function $R_{YY}[k]$ specified in the problem. So, instead, relate $R_{YY}[k]$ to the autocorrelation $R_{XX}[k]$ in Example 8.10. In particular, observe that $R_{YY}[k] = \frac{\pi^2}{4} R_{XX}[k]$ for $k \neq 0$. Use that as a substitution below:

$$\begin{aligned} S_{YY}(F) &= \sum_{k=-\infty}^{+\infty} R_{YY}[k]e^{-j2\pi Fk} = R_{YY}[0] + \sum_{k\neq 0} R_{YY}[k]e^{-j2\pi Fk} = R_{YY}[0] + \sum_{k\neq 0} \frac{\pi^2}{4} R_{XX}[k]e^{-j2\pi Fk} \\ &= 1 + \frac{\pi^2}{4}\sum_{k\neq 0} R_{XX}[k]e^{-j2\pi Fk} = 1 + \frac{\pi^2}{4}\sum_{k=-\infty}^{+\infty} R_{XX}[k]e^{-j2\pi Fk} - \frac{\pi^2}{4} R_{XX}[0]e^{-j2\pi F(0)} \\ &= 1 - \frac{\pi^2}{4} R_{XX}[0] + \frac{\pi^2}{4}\sum_{k=-\infty}^{+\infty} R_{XX}[k]e^{-j2\pi Fk} = 1 - \frac{\pi^2}{4}\cdot\frac{1}{2} + \frac{\pi^2}{4} S_{XX}(F) \\ &= 1 - \frac{\pi^2}{8} + \frac{\pi^2}{4}\text{tri}(2F) \end{aligned}$$

47.

a. $P_X = \int_{-1/2}^{1/2} S_{XX}(F)dF = \int_{-1/4}^{1/4} 2P\,dF = 2P\left(\frac{1}{4} - -\frac{1}{4}\right) = P$.

b. Use the inverse Fourier transform on (–1/2, 1/2) as described in the section:

$$R_{XX}[k] = \int_{-1/2}^{1/2} S_{XX}(F)e^{j2\pi kF}\,dF = \int_{-1/2}^{1/2} S_{XX}(F)\cos(2\pi kF)dF \qquad \text{because } S_{XX}(F) \text{ is even}$$

$$= \int_{-1/4}^{1/4} 2P\cos(2\pi kF)dF = 4P\int_0^{1/4}\cos(2\pi kF)dF = \frac{4P}{2\pi k}\sin(2\pi kF)\Big|_0^{1/4}$$

$$= \frac{4P}{2\pi k}\sin(\pi k/2)$$

We can actually simplify this further: $R_{XX}[k] = \frac{4P}{2\pi k}\sin(\pi k/2) = P\frac{\sin(\pi k/2)}{\pi k/2} = P\text{sinc}(k/2)$. (This is reasonable in light of the fact that the psd of X_n is a rectangular function.)

49.

a. Use the discrete-time convolution formula:

$$Y_n = X_n \star h[n] = \sum_{k=-\infty}^{\infty} X_{n-k}h[k] = \sum_{k=0}^{M-1} X_{n-k}h[k] \qquad \text{because all other terms are zero}$$

$$= \sum_{k=0}^{M-1} X_{n-k}\cdot\frac{1}{M} = \frac{1}{M}\sum_{k=0}^{M-1} X_{n-k} = \frac{X_{n-M+1} + \cdots + X_n}{M}$$

In other words, Y_n is the simple arithmetic average of the most recent M terms of the sequence X_n.

b. First, apply the discrete-time Fourier transform:

$$H(F) = \sum_{k=-\infty}^{\infty} h[k]e^{-j2\pi kF} = \sum_{k=0}^{M-1} h[k]e^{-j2\pi kF} = \frac{1}{M}\sum_{k=0}^{M-1} e^{-j2\pi kF} = \frac{1}{M}\sum_{k=0}^{M-1}\left(e^{-j2\pi F}\right)^k$$

Next, use the formula for a finite geometric sum, $1 + r + \ldots + r^{M-1} = \frac{1-r^M}{1-r}$:

$$H(F) = \frac{1}{M}\left[1 + e^{-j2\pi F} + \cdots + (e^{-j2\pi F})^{M-1}\right] = \frac{1-e^{-j2\pi FM}}{M(1-e^{-j2\pi F})}$$

c. There are two approaches: Determine the psd of Y_n using part **b** and then find the inverse Fourier transform on (–1/2, 1/2); or, determine $R_{YY}[k]$ directly from $R_{XX}[k]$ and $h[k]$. We show the latter here.

From Exercise 44, $R_{XX}[k] = \sigma^2\delta[k]$ for some $\sigma > 0$. This immediately simplifies $R_{YY}[k]$: $R_{YY}[k] = R_{XX}[k]\star h[k]\star h[-k] = \sigma^2\delta[k]\star h[k]\star h[-k] = \sigma^2 h[k]\star h[-k]$. This convolution is given by $h[k]\star h[-k] = \sum_{\ell=-\infty}^{\infty} h[k-\ell]h[-\ell]$, and the summand is zero unless both arguments of $h[\ ]$ are in the set $0, 1, \ldots, M-1$. These joint constraints on $k-\ell$ and $-\ell$ are satisfied by exactly $M-|k|$ integers provided $|k| < M$; e.g., when $k = -1$ then ℓ must be among $\{-M+1, \ldots, -1\}$, a set of $M-1$ numbers. Otherwise, if $|k| \geq M$, then no indices result in a non-zero summand, and the entire expression is 0. For every one of those $M-|k|$ integers, the summand in the convolution is equal to $\frac{1}{M}\cdot\frac{1}{M} = \frac{1}{M^2}$, so the convolution equals $\sum_{\ell}\frac{1}{M^2} = \frac{1}{M^2}[\text{number of non-zero terms}] = \frac{M-|k|}{M^2}$ for $|k| = 0, 1, \ldots, M-1$.

Finally, $R_{YY}[k] = \sigma^2 h[k]\star h[-k] = \sigma^2\frac{M-|k|}{M^2}$ for $|k| = 0, 1, \ldots, M-1$ and zero otherwise.

http://www.springer.com/978-3-319-52400-9

Probability with Applications in Engineering, Science, and Technology
Carlton, M.A.; Devore, J.L.
2017, XXVI, 610 p. 209 illus., 178 illus. in color., Softcover
ISBN: 978-3-319-52400-9