SPRINGER TEXTS IN STATISTICS

*JIM PITMAN*

# PROBABILITY
## Instructor's Manual

Springer-Verlag
New York Berlin Heidelberg London Paris
Tokyo Hong Kong Barcelona Budapest

# Section 1.1

1. a) 2/3 b) 66.67% c) 0.6667 d) 4/7 e) 57.14% f) 0.5714

2. a) 7/10 = 0.7 b) 4/10 = 0.4 c) 4/10 = 0.4

3. a) If the tickets are drawn with replacement, then, as in Example 3, there are $n^2$ equally likely outcomes. There is just one pair in which the first number is 1 and the second number is 2, so $P(\text{first ticket is 1 and second ticket is 2}) = 1/n^2$.
   b) The event (the numbers on the two tickets are consecutive integers) consists of $n-1$ outcomes: $(1,2), (2,3), \ldots (n-1,n)$. So its probability is $(n-1)/n^2$.
   c) Same as Problem 3 of Example 3. Answer: $(1-1/n)/2$
   d) If the draws are made without replacement, then there are only $n^2 - n$ equally likely possible outcomes, since we have to exclude the outcomes $(1,1), (2,2), \ldots, (n,n)$. So replace the denominators in a) through c) by $n(n-1)$.

4. a) 2/38
   b) $1-(2/38) = 36/38$
   c) 1, since $P(\text{both win}) = 0$.

5. a) #(total) $= 52 \times 51 = 2652$ possibilities.
   b) #(first card ace) $= 4 \times 51 = 204$. Thus $P(\text{first card ace}) = 4/52 = 1/13$.
   c) This will be exactly the same calculation as in part b) – just substitute "second" for "first". Thus, $P(\text{second card ace}) = 1/13$. Because you have all possible ordered pairs of cards, any probability statement concerning the first card by itself must also be true for the second card by itself.
   d) $P(\text{both aces}) = \frac{\#(\text{both aces})}{\#(\text{total})} = \frac{4\times 3}{2652} = \frac{1}{221}$.
   e) $P(\text{at least one ace}) = P(\text{first card ace}) + P(\text{second card ace}) - P(\text{both cards aces}) = \frac{1}{13} + \frac{1}{13} - \frac{1}{221} = \frac{33}{221}$.

6. a) $52 \times 52 = 2704$
   b) $(52 \times 4)/(52 \times 52) = 1/13$
   c) Same as b)
   d) $(4 \times 4)/(52 \times 52) = 1/169$
   e) $P(\text{at least one ace}) = P(\text{first card ace}) + P(\text{second card ace}) - P(\text{both cards aces}) = \frac{1}{13} + \frac{1}{13} - \frac{1}{169} = \frac{25}{169}$.

7. a) $P(\text{maximum} \le 2) = P(\text{both dice} \le 2) = 4/36 = 1/9$
   b) $P(\text{maximum} \le 3) = P(\text{both dice} \le 3) = 9/36 = 1/4$
   c) $P(\text{maximum} = 3) = P(\text{maximum} \le 3) - P(\text{maximum} \le 2) = 5/36$
   d)

| Outcome | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Probability | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 |

   e) Since this covers all the possible outcomes of the experiment, and these events are mutually exclusive, you should expect $P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 1$.

8. a-d) As in Example 3, the outcome space consists of $n^2$ equally likely pairs of numbers, each number between 1 and $n$. The event (the maximum of the two numbers is less than or equal to $x$) is represented by the set of pairs having both entries less than or equal to $x$. There are $x^2$ possible pairs of this type, so for $x = 0$ to $n$: $P(\text{maximum } \le x) = x^2/n^2$ and for $x = 1$ to $n$

$$\begin{aligned} P(x) &= P(\text{maximum is exactly } x) \\ &= P(\text{maximum } \le x) - P(\text{maximum } \le x-1) \\ &= (2x-1)/n^2. \end{aligned}$$

e) As in the previous exercise, $\sum_{x=1}^{n} P(x) = 1$.

**Remark.** It follows that $\sum_{x=1}^{n}(2x-1) = n^2$. In other words, the sum of the first $n$ odd numbers is $n^2$, a fact which you can check in other ways.

9. 1/11, 1/6

10.

| Play | Chance of win | Payoff Odds = $r_{pay}$ to 1 | House Percentage |
|---|---|---|---|
| A | 18/38 | 1 to 1 | 5.26% |
| B | 12/38 | 2 to 1 | 5.26% |
| C | 12/38 | 2 to 1 | 5.26% |
| D | 6/38 | 5 to 1 | 5.26% |
| E | 5/38 | 6 to 1 | 7.89% |
| F | 4/38 | 8 to 1 | 5.26% |
| G | 3/38 | 11 to 1 | 5.26% |
| H | 2/38 | 17 to 1 | 5.26% |
| I | 1/38 | 35 to 1 | 5.26% |

Use the formula: House percentage $= (1 - P(\text{win}) \times (r_{pay} + 1)) \times 100\%$

11. Call the event $A$. By definition of fair odds, we have $r_{fair} = (1 - P(A))/P(A)$. Solve for $P(A) = 1/(r_{fair}+1)$ and substitute in the formula for the house percentage: House percentage $= (1 - P(A) \cdot (r_{pay} + 1)) \times 100\%$.

# Section 1.2

1. The relative frequency interpretation makes no sense here. Presumably, what is meant is "more likely than not in the subjective opinion of the judge".

2. Less than 1/100. Since $r_{pay} < r_{fair}$ (we assume the bookmaker wants to make a profit), it follows that
$$P(\text{win}) = 1/(r_{fair} + 1) < 1/(r_{pay} + 1).$$

3. a) Let $r_i^*$ to 1 be the fair odds against horse $i$ winning, say in the opinion of the bookmaker, and let $p_i^*$ be the probability of horse $i$ winning. Then (presuming the bookmaker wants a profit) $r_i < r_i^*$, so
$$\Sigma = \sum_{i=1}^{10} p_i = \sum_{i=1}^{10} \frac{1}{r_i + 1} > \sum_{i=1}^{10} \frac{1}{r_i^* + 1} = \sum_{i=1}^{10} p_i^* = 1.$$

   b) Yes. Bet a total of $\$B$ on the horses, with proportion $p_i/\Sigma$ of this total, i.e., $\$p_i B/\Sigma$ on horse $i$. If horse $i$ wins, you get back an amount $\$(r_i + 1)p_i B/\Sigma = \$B/\Sigma > \$B$. So, no matter which horse wins (and regardless of any probabilities), you get back more than you bet. A golden opportunity rarely found at the races!

4. a) Overall gain $= (\$8) \times 10 - (\$1) \times 90 = -\$10$.

   b) Average gain per bet $=$ (overall gain) / (# bets) $= \frac{-\$10}{100} = -10$ cents.

   c) Gambler's average gain per bet $=$ (overall gain) / (# bets)
$$= \frac{r_{pay} \times (\#\text{wins}) - 1 \times (\#\text{losses})}{\#\text{wins}+\#\text{losses}} = \frac{r_{pay} - r_{\#}}{1 + r_{\#}}.$$

   The house's average financial gain per bet is therefore $(r_{\#} - r_{pay})/(r_{\#} + 1)$ over this sequence of bets. Recall that if the fair (chance) odds against the gambler winning are $r_{fair}$ to 1, then the house percentage is $(r_{fair} - r_{pay})/(r_{fair} + 1) \times 100\%$. As the number of bets made increases, we should expect (according to the frequency interpretation of probability) $r_{\#}$ to approach $r_{fair}$, and therefore the house's average financial gain per bet should approach the house percentage. So we can interpret the house percentage as the long run average financial gain on a one dollar bet.

# Section 1.3

1. The cake is $(2 \times 2) + 2 + 1 = 7$ times as large as the piece your neighbor gets, so you get 4/7 of the cake.

2. a) $(AB^c) \cup (A^cB)$

   b) $A^cB^cC^c$

   c) Exactly one : $(AB^cC^c) \cup (A^cBC^c) \cup (A^cB^cC)$
   Exactly two : $(ABC^c) \cup (AB^cC) \cup (A^cBC)$
   Three : $ABC$

3. Take $\Omega = \{1, 2, ..., 500\}$.

   a) $\{17, 93, 202\}$

   b) $\{17, 93, 202, 4, 101, 102, 398\}^c$

   c) $\{16, 18, 92, 94, 201, 203\}$

4. a) Yes: $\{0, 1\}$

   b) Yes: $\{1\}$

   c) No. This is a subset of the event $\{1\}$ but it is not identical to $\{1\}$ because the event (first toss tails, second toss heads) also is a subset of $\{1\}$.

   d) Yes: $\{1, 2\}$

5. a) first coin lands heads

   b) second coin lands tails

   c) same as a)

   d) at least two heads

   e) exactly two tails

   f) first two coins land the same way.

6. a)

| outcome | 1 | 2 | 4 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|
| probability | 1/10 | 1/5 | 3/10 | 1/5 | 1/10 | 1/10 |

   b)

| outcome | 1 | 2 | 3 |
|---|---|---|---|
| probability | 3/5 | 1/5 | 1/5 |

7. a) As in Example 3, using the addition rule and the symmetry assumption, $P(1) = P(6) = p/2$ and $P(2) = P(3) = P(4) = P(5) = (1-p)/4$

   b) Use the additivity property: $P(3 \text{ or } 4 \text{ or } 5 \text{ or } 6) = \frac{3(1-p)}{4} + \frac{p}{2} = \frac{3-p}{4}$

8. a) $P(A \cup B) = P(A) + P(B) - P(AB) = 0.6 + 0.4 - 0.2 = 0.8$

   b) $P(A^c) = 1 - P(A) = 0.4$

   c) Similarly $P(B^c) = 0.6$

   d) $P(A^cB) = P(B) - P(AB) = 0.4 - 0.2 = 0.2$

   e) $P(A \cup B^c) = P[(A^cB)^c] = 1 - 0.2 = 0.8$

   f) $P(A^cB^c) = P[(A \cup B)^c] = 1 - 0.8 = 0.2$

9. a) 0.9 b) 1 c) 0.1

10. a) $P(\text{exactly 2 of } A, B, C) = P(ABC^c) + P(AB^cC) + P(A^cBC)$

$$= (P(AB) - P(ABC)) + (P(AC) - P(ABC)) + (P(BC) - P(ABC))$$
$$= P(AB) + P(AC) + P(BC) - 3P(ABC)$$

b) $P(\text{exactly 1 of } A,\, B,\, C) = P(AB^cC^c) + P(A^cB^cC) + P(A^cB^cC)$

$$= P(A) - P(A \cap (B \cup C)) + P(B) - P(B \cap (A \cup C)) + P(C) - P(C \cap (A \cup B))$$
$$= \dots$$
$$= P(A) + P(B) + P(C) - 2\,(P(AB) + P(AC) + P(BC)) + 3P(ABC)$$

c) $P(\text{exactly none}) = 1 - P(A \cup B \cup C)$

$$= 1 - (P(A) + P(B) + P(C)) + (P(AB) + P(AC) + P(BC)) - P(ABC)$$

11. By inclusion-exclusion for $n = 2$,

$$P(A \cup B \cup C) = P[(A \cup B) \cup C] = P(A \cup B) + P(C) - P[(A \cup B)C]$$
$$= P(A) + P(B) - P(AB) + P(C) - P[(A \cup B)C].$$

Now $(A \cup B)C = (AC) \cup (BC)$, which is clear from a Venn diagram. So by another application of inclusion-exclusion for $n = 2$,

$$P[(A \cup B)C] = P[(AC) \cup (BC)] = P(AC) + P(BC) - P[(AC)(BC)]$$
$$= P(AC) + P(BC) - P(ABC)$$

Substitute this into the previous expression.

12. Formula has been proved for $n = 1, 2, 3$. Proceed by mathematical induction. We should first prove

$$(\cup_{i=1}^n A_i) \cap A_{n+1} = \cup_{i=1}^n (A_iA_{n+1}),$$

which is easily checked, e.g. by another mathematical induction using the $n = 2$ case: $(A \cup B)C = (AC) \cup (BC)$. Next, assume that our hypothesis is true for $n$. Then

$$P(\cup_{i=1}^{n+1} A_i) = P[(\cup_{i=1}^n A_i) \cup A_{n+1}] = P(\cup_{i=1}^n A_i) + P(A_{n+1}) - P[\cup_{i=1}^n (A_iA_{n+1})]$$

(by inclusion-exclusion for $n = 2$)

$$= \sum_{i=1}^n P(A_i) - \sum_{i<j\le n} P(A_iA_j) + \sum_{i<j<k\le n} P(A_iA_jA_k) - \dots$$
$$+P(A_{n+1}) - \sum_{i\le n} P(A_iA_{n+1}) + \sum_{i<j\le n} P(A_iA_jA_{n+1}) - \dots$$

(by applying the induction hypothesis to the first and third term)

$$= \sum_{i=1}^{n+1} P(A_i) - \sum_{i<j\le n+1} P(A_iA_j) + \sum_{i<j<k\le n+1} P(A_iA_jA_k) - \dots$$

(by regrouping terms.) So the claim holds for $n + 1$.

13. For $n = 2$, $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \le P(A_1) + P(A_2)$ Now use induction by assuming true for $n$.

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\left[\bigcup_{i=1}^{n} A_i\right] \cup A_{n+1}\right) \le P\left(\bigcup_{i=1}^{n} A_i\right) + P(A_{n+1}) \le \sum_{i=1}^n P(A_i) + P(A_{n+1})$$

14. Use the identity $P(A \cup B) = P(A) + P(B) - P(AB)$ and the fact that $P(A \cup B) \le 1$.

15. Let $A_i = B_i^c$. Then

$$P(B_1B_2\cdots B_n) = P(\bigcap_{i=1}^n A_i^c) = P((\bigcup_{i=1}^n A_i)^c) = 1 - P(\bigcup_{i=1}^n A_i)$$
$$\ge 1 - \sum_{i=1}^n P(A_i) = 1 - \sum_{i=1}^n (1 - P(B_i)) = \sum_{i=1}^n P(B_i) \;-\; (n-1)$$

16. a) Use mathematical induction. The claim holds for $n = 1$ and 2. Suppose the claim holds for unions of $n$ sets. To show the claim for $n+1$ sets, use inclusion-exclusion for 2 sets to write

$$P(\cup_{i=1}^{n+1} A_i) = P\left((\cup_{i=1}^{n} A_i) \cup A_{n+1}\right) = P(\cup_{i=1}^{n} A_i) + P(A_{n+1}) - P\left((\cup_{i=1}^{n} A_i) A_{n+1}\right)$$

Use the induction hypothesis to bound the first term:

$$P(\cup_{i=1}^{n} A_i) \geq \sum_{i \leq n} P(A_i) - \sum_{i<j\leq n} P(A_i A_j)$$

and use Boole's inequality to bound the last term:

$$P\left((\cup_{i=1}^{n} A_i) A_{n+1}\right) = P(\cup_{i=1}^{n} A_i A_{n+1}) \leq \sum_{i \leq n} P(A_i A_{n+1})$$

Regroup to see the claim holds for $n+1$.

b) Define for $m \geq 1$ and sets $A_1, \ldots, A_n$

$$S(m; A_1, \ldots, A_n)$$
$$= \sum_{i \leq n} P(A_i) - \sum_{i<j\leq n} P(A_i A_j) + \cdots + (-1)^{m+1} \sum_{k_1 < \ldots < k_m \leq n} P(A_{k_1} \cdots A_{k_m})$$

The claim is that

**Claim:** For each $m \geq 1$ and each $n \geq 1$: If $A_1, \ldots, A_n$ are $n$ sets then

$$(-1)^m \left(P(\cup_1^n A_i) - S(m; A_1, \ldots, A_n)\right) \geq 0.$$

**Proof:** Use mathematical induction. The claim holds for $m = 1$ (by a)) (and also for $m = 2$, by b)). Say the claim holds for $m$. Now show that this implies the claim holds for $m+1$, i.e., that:

**Subclaim:** ($m$ is fixed) For each $n \geq 1$: If $A_1, \ldots, A_n$ are $n$ sets then

$$(-1)^{m+1} \left(P(\cup_1^n A_i) - S(m+1; A_1, \ldots, A_n)\right) \geq 0.$$

**Proof of subclaim:** Again use mathematical induction. This subclaim holds for $n = 1, \ldots, m$ because the inclusion-exclusion formula (Exercise 12) shows that the left side equals zero. Suppose the claim holds for $n$. Now show that this implies the subclaim holds for $n+1$, i.e., that

$$(-1)^{m+1} \left(P(\cup_1^{n+1} A_i) - S(m+1; A_1, \ldots, A_{n+1})\right) \geq 0. \qquad (*)$$

To do this, first observe that

$$S(m+1; A_1, \ldots, A_n) + P(A_{n+1}) - S(m; A_1 A_{n+1}, \ldots, A_n A_{n+1})$$
$$= S(m+1; A_1, \ldots, A_{n+1}). \qquad (**)$$

Therefore

$$(-1)^{m+1} \left(P(\cup_1^{n+1} A_i) - S(m+1; A_1, \ldots, A_{n+1})\right)$$
$$= (-1)^{m+1} \left(P(\cup_1^n A_i) + P(A_{n+1}) - P(\cup_1^n A_i A_{n+1}) - S(m+1; A_1, \ldots, A_{n+1})\right)$$

(by inclusion-exclusion for 2 sets)

$$= (-1)^{m+1} \{P(\cup_1^n A_i)$$
$$- P(\cup_1^n A_i A_{n+1}) + S(m; A_1 A_{n+1}, \ldots, A_n A_{n+1}) - S(m+1; A_1, \ldots, A_n)\}$$

(by identity $(**)$)

$$= (-1)^{m+1} \{P(\cup_1^n A_i) - S(m+1; A_1, \ldots, A_n)\}$$
$$+ (-1)^m \{P(\cup_1^n A_i A_{n+1}) - S(m; A_1 A_{n+1}, \ldots, A_n A_{n+1})\}$$

The first term is nonnegative by the induction hypothesis on $n$; the second term is nonnegative by the induction hypothesis on $m$. So $(*)$ holds, and we're done!

# Section 1.4

1. Let $\pi$ denote the proportion of women in the population. Then the proportion $r$ of righthanders is given by $r = .92 \times \pi + .88 \times (1 - \pi) = .88 + .04\pi$.

   a) (iii) ; since $\pi$ is unknown.

   b) (i) ; since $0 \leq \pi \leq 1$ implies $.88 \leq r \leq .92$.

   c) (i) ; $.88 + .04 \times .5 = .9$.

   d) (i) ; solve $.88 + .04 \times \pi = .9$, you'll get $\pi = .5$.

   e) (i) ; if $\pi \geq .75$, then $r \geq .91$.

2. Pick a light bulb at random. Let $D$ be the event (the light bulb is defective), and $B$ be the event (the light bulb is made in city $B$). Then $P(B) = 1/3$, $P(D|B) = .01$, and

$$P(D^cB) = P(D^c|B)P(B) = [1 - P(D|B)]P(B) = 0.99 \times 1/3 = 0.33$$

3. $P(\text{rain tomorrow} \mid \text{rain today}) = \dfrac{P(\text{rain today and tomorrow})}{P(\text{rain today})} = \dfrac{30\%}{40\%} = \frac{3}{4}$

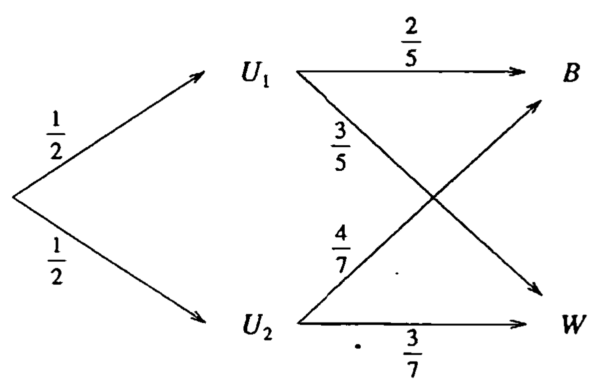
4. Call the events $A$ (having probability .1) and $B$ (having probability .3).

   a) $P(A^cB^c) = P(A^c)P(B^c) = 0.9 \times 0.7 = 0.63$

   b) $1 - P(A^cB^c) = 0.37$

   c) $P(AB^c \cup A^cB) = P(AB^c) + P(A^cB) = 0.1 \times 0.7 + 0.9 \times 0.3 = 0.34$

5. a) Let $U_1 =$ (urn 1 chosen), $U_2 =$ (urn 2 chosen), $B =$ (black ball chosen), $W =$ (white ball chosen).




   b) $P(U_1) = \frac{1}{2} = P(U_2); P(W|U_1) = \frac{3}{5}; P(B|U_1) = \frac{2}{5}; P(W|U_2) = \frac{3}{7}; P(B|U_2) = \frac{4}{7}$

   c) $P(B) = P(BU_1) + P(BU_2) = P(B|U_1)P(U_1) + P(B|U_2)P(U_2) = \frac{4}{7} \times \frac{1}{2} + \frac{2}{5} \times \frac{1}{2} = \frac{17}{35}$

6. $P(\text{second spade} \mid \text{first black})$

$$= \frac{P(\text{first black and second spade})}{P(\text{first black})}$$

$$= \frac{P(\text{first spade and second spade}) + P(\text{first club and second spade})}{P(\text{first black})}$$

$$= \frac{(13/52)(12/51) + (13/52)(13/51)}{(26/32)} = \frac{25}{102}.$$

Or you may use a symmetry argument as follows:
$P(\text{second black} \mid \text{first black}) = 25/51$,
and $P(\text{second spade} \mid \text{first black}) = P(\text{second club} \mid \text{first black})$
by symmetry. Therefore $P(\text{second spade} \mid \text{first black}) = 25/102$.
**Discussion.** The frequency interpretation is that over the long run, out of every 102 deals yielding a black card first, about 25 will yield a spade second.

7. a) $P(B) = 0.3$

b) $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$
$\Longrightarrow P(B) = \frac{P(A \cup B) - P(A)}{1 - P(A)} = \frac{0.3}{0.5} = 0.6$

8. Assume $n$ cards and all $2n$ faces are equally likely to show on top.
$P(\text{white on bottom} \mid \text{black on top})$

$$= \frac{P(\text{white on bottom and black on top})}{P(\text{black on top})} = \frac{50\% \times \frac{1}{2}}{50\% \times \frac{1}{2} + 20\%} = 5/9$$

9. By scheme A, $P(\text{student selected is from School 1}) = 100/1000 = 1/10$, while by scheme B, that chance is just $1/3$, which is the chance that school 1 is selected. So these two schemes are not probabilistically equivalent.
Consider a particular student $x$, and suppose she is from school $i$. The chance that $x$ will be selected by scheme A is $1/1000$. The chance that $x$ will be selected by scheme C is

$$P(x \text{ is selected}) = P(\text{School } i \text{ is selected, and then } x) = p_i \cdot \frac{1}{(\text{size of class } i)}.$$

For scheme C to be equivalent to scheme A, this chance should be $1/1000$. Therefore,
$p_i = (\text{size of class } i)/1000$. So $p_1 = .1$, $p_2 = .4$, $p_3 = .5$.

10. a) Let $A_i$ be the event that the $i$th source works.
$P(\text{zero work}) = P(A_1^c A_2^c) = 0.6 \times 0.5 = 0.3$
$P(\text{exactly one works}) = P(A_1 A_2^c) + P(A_1^c A_2) = 0.4 \times 0.5 + 0.6 \times 0.5 = 0.5$
$P(\text{both work}) = P(A_1 A_2) = 0.4 \times 0.5 = 0.2$

b) $P(\text{enough power}) = 0.6 \times 0.5 + 1 \times 0.2 = 0.5$

11. Let $B_1 = $ (firstborn is a boy), $G_1 = $ (firstborn is a girl), similarly for $B_2$ and $G_2$.

a) $P(B_1 B_2) = P(B_1 B_2 \mid \text{identical}) P(\text{identical}) + P(B_1 B_2 \mid \text{fraternal}) P(\text{fraternal})$

$$= \frac{1}{2} \cdot p + \frac{1}{4} \cdot (1 - p) = \frac{1+p}{4}.$$

b) $P(B_1 G_2) = 0 \cdot p + \frac{1}{4} \cdot (1 - p) = \frac{1-p}{4}$.

c) Note that the chance that the firstborn is a boy is $1/2$ whether identical or fraternal, so $P(B_1) = 1/2$. Similarly $P(B_2) = P(G_1) = P(G_2) = 1/2$. So

$$P(G_2 | B_1) = \frac{P(B_1 G_2)}{P(B_1)} = \frac{1/4 \cdot (1 - p)}{1/2} = \frac{1}{2}(1 - p).$$

d)

$$P(G_2 | G_1) = \frac{P(G_1 G_2)}{P(G_1)} = \frac{1/4 \cdot (1 + p)}{1/2} = \frac{1}{2}(1 + p).$$

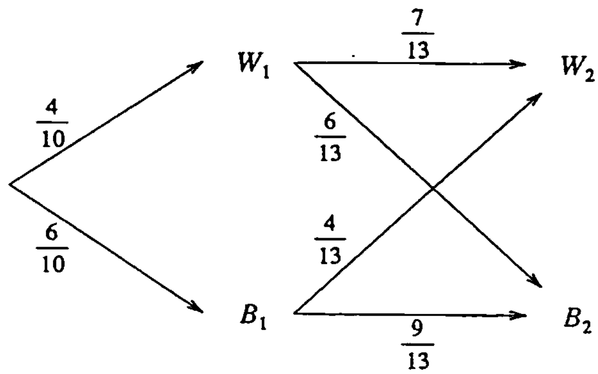
12. $\frac{P(F) - P(FG)}{1 - P(G)}$

# Section 1.5

1. a) $P(\text{black}) = P(\text{black} \mid \text{odd})P(\text{odd}) + P(\text{black} \mid \text{even})P(\text{even}) = \frac{1}{4} \times \frac{1}{2} + \frac{2}{6} \times \frac{1}{2} = \frac{7}{24}$

   b) $P(\text{even} \mid \text{white}) = \dfrac{P(\text{white} \mid \text{even})P(\text{even})}{P(\text{white})} = \frac{(2/3)\times(1/2)}{17/24} = \frac{8}{17}$

2.

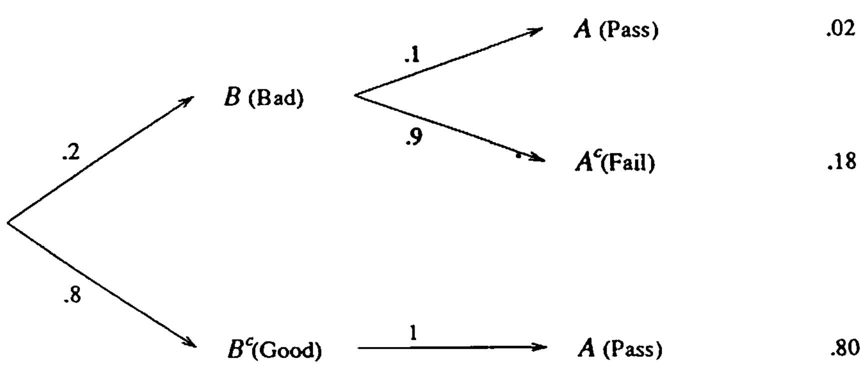



   a) $P(W_2) = P(W_1W_2) + P(B_1W_2) = \frac{4}{10} \cdot \frac{7}{13} + \frac{6}{10} \cdot \frac{4}{13} = \frac{2}{5}$

   b) $P(B_1|W_2) = \frac{P(B_1W_2)}{P(W_2)} = \frac{\frac{6}{10}\cdot\frac{4}{13}}{\frac{4}{10}} = \frac{6}{13}$.

   c) Repeat a), with symbols:
   $P(W_2) = \frac{w}{w+b} \cdot \frac{w+d}{w+b+d} + \frac{b}{w+b} \cdot \frac{w}{w+b+d} = \frac{w}{w+b}$

3. Pick at chip at random. Let $B$ = (chip is bad), let $A$ = (chip passes the cheap test). Then $P(B) = .2$, $P(A|B) = .1$, and $P(A|B^c) = 1$, which imply $P(B^c) = .8$, and $P(A^c|B) = .9$.




   a) $P(B^c|A) = \frac{P(A|B^c)P(B^c)}{P(A|B^c)P(B^c)+P(A|B)P(B)} = \frac{0.80}{0.02+0.80} = \frac{40}{41}$. Or use Bayes' rule for odds.

   b) $P(B|A) = 1 - P(B^c|A) = 1/41$.

4. a) $P(T_0|R_1) = \frac{P(R_1|T_0)P(T_0)}{P(R_1|T_0)P(T_0)+P(R_1|T_1)P(T_1)} = \frac{(.01)(.5)}{(.01)(.5)+(.98)(.5)} = \frac{1}{99}$

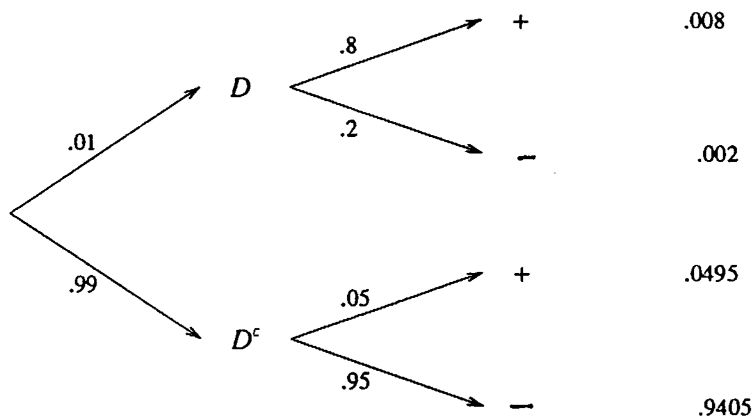
   b) $\frac{(.01)(.2)}{(.01)(.2)+(.98)(.8)} = \frac{1}{393}$

c) $P(\text{error in transmission}) = P(R_0T_1 \cup R_1T_0)$

$$= P(R_0T_1) + P(R_1T_0)$$
$$= P(R_0|T_1)P(T_1) + P(R_1|T_0)P(T_0)$$
$$= (.02)(.5) + (.01)(.5) = 3/200$$

d) $(.02)(.8) + (.01)(.2) = 9/500$

5. Let $D$ denote the event (person has the disease), $+$ denote the event (person is diagnosed as having the disease), and $-$ denote the event (person is diagnosed as healthy). Then $P(D) = .01$, $P(+|D^c) = .05$, $P(-|D) = .2$, which in turn imply that $P(D^c) = .99$, $P(-|D^c) = .95$, $P(+|D) = .8$.




a) $P(+) = P(+|D)P(D) + P(+|D^c)P(D^c) = .0575$.

b) $P(-\cap D) = P(-|D)P(D) = .002$

c) $P(-\cap D^c) = P(-|D^c)P(D^c) = .9405$

d) $P(D|+) = \frac{P(+|D)P(D)}{P(+)} = \frac{16}{115}$.
Or we may argue using odds ratios, as follows:
$\frac{P(D)}{P(D^c)} = \frac{1}{99}$, $\frac{P(+|D)}{P(+|D^c)} = \frac{0.8}{0.05} = 16 \Longrightarrow \frac{P(D|+)}{P(D^c|+)} = \frac{1}{99} \times 16 = \frac{16}{99}$
Hence $P(D|+) = \frac{16}{99+16} = \frac{16}{115}$.

e) Yes. See the explanation after Example 3.

6. a) The experimenter is assuming that before the experiment, $H_1$, $H_2$, and $H_3$ are equally likely. That is, prior probabilities are given by $P(H_1) = P(H_2) = P(H_3) = 1/3$.

b) No, because the above assumption is being made. Since the prior probabilities are subjective, so are the posterior probabilities.

c) Prior probabilities: $P(H_1) = .5$, $P(H_2) = .45$, $P(H_3) = .05$.
Likelihoods of $A$: $P(A|H_1) = .1$, $P(A|H_2) = .01$, $P(A|H_3) = .39$.
[Now all these probabilities have a long run frequency interpretation, so the posterior probabilities will as well.]
Posterior probabilities are proportional to: .05, .0045, .0195. So $H_3$ is no longer the most likely hypothesis; $H_1$ is, and $P(H_3|A) = \frac{.0195}{.05+.0045+.0195} = .263$.

7. a) As in Example 1,

| $i$ | 1 | 2 | 3 |
|---|---|---|---|
| $P(\text{Box } i \text{ and white})$ | 1/6 | 2/9 | 1/4 |
| $P(\text{Box } i \text{ and black})$ | 1/6 | 1/9 | 1/12 |
| $P(\text{Box } i \mid \text{white})$ | 6/23 | 8/23 | 9/23 |
| $P(\text{Box } i \mid \text{black})$ | 6/13 | 4/13 | 3/13 |

Since $P(\text{Box } i \mid \text{white})$ is largest when $i = 3$ and $P(\text{Box } i \mid \text{black})$ is largest when $i = 1$, the strategy is to guess Box 3 when a white ball is drawn and guess Box 1 when a black ball is drawn. The probability of guessing correctly is

$$P(\text{Box 3 and white}) + P(\text{Box 1 and black}) = 5/12.$$

b) Suppose your strategy is to guess box $i$ with probability $p_i$ ($i = 1, 2, 3$; $p_1 + p_2 + p_3 = 1$) whenever a white ball is drawn, and suppose I am picking each box with probability 1/3. Then in cases where a white ball is drawn, the probability that you guess correctly is

$$p_1 \frac{6}{23} + p_2 \frac{8}{23} + p_3 \frac{9}{23} = \frac{9}{23} - p_1 \frac{3}{23} - p_2 \frac{1}{23}.$$

Clearly the probability of your guessing correctly is greatest when $p_3 = 1$; that is, when you guess Box 3 every time that you see a white ball. A similar argument for the case where a black ball is drawn shows that the probability of your guessing correctly is greatest when you guess Box 1 every time that you see a black ball.

c) Here $P(\text{Box 1}) = 1/2$, $P(\text{Box 2}) = 1/4$, $P(\text{Box 3}) = 1/4$, so that

| $i$ | 1 | 2 | 3 |
|---|---|---|---|
| $P(\text{Box } i \text{ and white})$ | 1/4 | 1/6 | 3/16 |
| $P(\text{Box } i \text{ and black})$ | 1/4 | 1/12 | 1/16 |
| $P(\text{Box } i \mid \text{white})$ | 12/29 | 8/29 | 9/29 |
| $P(\text{Box } i \mid \text{black})$ | 12/19 | 4/19 | 3/19 |

If you continue to use the strategy found in a), i.e. guess Box 3 if a white ball is seen, and guess Box 1 if a black ball is seen, then the probability of guessing correctly is

$$P(\text{Box 3 and white}) + P(\text{Box 1 and black}) = \frac{3}{16} + \frac{1}{4} = \frac{7}{16}.$$

d) If you are convinced that I am using either the (1/3, 1/3, 1/3) strategy or the (1/2, 1/4, 1/4) strategy, then you can decide which one it is by observing the long-run proportion of trials that your guesses are correct. If I am using the (1/3, 1/3, 1/3) strategy then, by the frequency interpretation of probability, the proportion of trials that you guess correctly should be approximately 5/12, while if I am using the (1/2, 1/4, 1/4) strategy then this proportion should be closer to 7/16. If I am in fact using the (1/2, 1/4, 1/4) picking strategy, you can improve upon the guessing strategy found in a) as follows: Since $P(\text{Box } i \mid \text{white})$ is largest when $i = 1$ and $P(\text{Box } i \mid \text{black})$ is largest when $i = 1$, the strategy is to always guess Box 1. The probability of guessing correctly is then

$$P(\text{Box 1}) = \frac{1}{2}.$$

**Remark:** If you're not sure at all what strategy I am using, you can determine it approximately, as follows: Suppose I am using a $(p_1, p_2, p_3)$ strategy to pick the boxes, and suppose you decide that you will always pick (say) Box 1 when you see a white ball, Box 2 when you see a black ball. It is possible for you to determine my strategy, provided you can keep track of the proportion of trials where a white ball was chosen from Box 1 and the proportion of trials where a black ball was chosen from Box 2. (This you can do, because you can see the color of the ball drawn, and you are told whether your guess is correct.) By the frequency interpretation of probability, these two proportions should in the long run approximate the probabilities

$$P(\text{Box 1 and white}) = \frac{1}{2}p_1 \text{ and } P(\text{Box 2 and black}) = \frac{1}{3}p_2$$

respectively. Thus you should be able to determine (approximately) the values $p_1$ and $p_2$ (and hence $p_3$).

8. a) The prior probabilities (of the boxes that I pick) are as follows:

| $i$ | 1 | 2 | 3 |
|---|---|---|---|
| $\pi_i$ | 6/23 | 9/23 | 8/23 |

Use Bayes' rule $P(\text{Box } i \mid \text{black}) = \dfrac{P(\text{Box } i \text{ and black})}{P(\text{black})} = \dfrac{\pi_i \cdot \frac{1}{i+1}}{\pi_1 \cdot \frac{1}{2} + \pi_2 \cdot \frac{1}{3} + \pi_3 \cdot \frac{1}{4}}$
to determine the posterior probabilities:

| $i$ | 1 | 2 | 3 |
|---|---|---|---|
| $P(\text{Box } i \mid \text{black})$ | 3/8 | 3/8 | 1/4 |

So you should guess Box 1 or 2. For either choice, $P(\text{you guess correctly} \mid \text{black}) = 3/8$.

b) Use Bayes' rule $P(\text{Box } i \mid \text{white}) = \dfrac{P(\text{Box } i \text{ and white})}{P(\text{white})} = \dfrac{\pi_i \cdot \frac{i}{i+1}}{\pi_1 \cdot \frac{1}{2} + \pi_2 \cdot \frac{2}{3} + \pi_3 \cdot \frac{3}{4}}$ to determine the posterior probabilities:

| $i$ | 1 | 2 | 3 |
|---|---|---|---|
| $P(\text{Box } i \mid \text{white})$ | 1/5 | 2/5 | 2/5 |

[You can use $P(\text{white}) = 1 - P(\text{black}) = 15/23$.] So you should guess Box 2. For either choice, $P(\text{you guess correctly} \mid \text{white}) = 2/5$.

c) $P(\text{guess correctly})$

$$= P(\text{guess correctly} \mid \text{black})P(\text{black}) + P(\text{guess correctly} \mid \text{white})P(\text{white})$$
$$= \tfrac{3}{8} \times \tfrac{8}{23} + \tfrac{2}{5} \times \tfrac{15}{23} = \tfrac{9}{23}.$$

This is your chance of winning, no matter which of the two best choices you make in each case. For instance, you could simply always guess Box 2, in which case the event of a win for you is the same as the event that I pick Box 2.

To see why the probability of your guessing correctly is at most 9/23 whatever your strategy may be, suppose that your strategy whenever a black ball is drawn is to guess Box $i$ with probability $p_i$. Then

$P(\text{guess correctly} \mid \text{black})$

$$= p_1 P(\text{Box } 1 \mid \text{black}) + p_2 P(\text{Box } 2 \mid \text{black}) + p_3 P(\text{Box } 3 \mid \text{black})$$
$$= p_1 \tfrac{3}{8} + p_2 \tfrac{3}{8} + p_3 \tfrac{1}{4} = \tfrac{3}{8} - \tfrac{1}{8} b_3 \le \tfrac{3}{8}.$$

Similarly, whatever your strategy whenever a white ball is drawn,
$P(\text{guess correctly} \mid \text{white}) \le \frac{2}{5}$. Therefore

$P(\text{guess correctly})$

$$= P(\text{guess correctly} \mid \text{black})P(\text{black}) + P(\text{guess correctly} \mid \text{white})P(\text{white})$$
$$\le \tfrac{3}{8} \times \tfrac{8}{23} + \tfrac{2}{5} \times \tfrac{15}{23} = \tfrac{9}{23}.$$

d) $P(\text{guess correctly} \mid \text{Box } 1)$

$$= P(\text{guess } 1 \mid \text{black \& Box } 1)P(\text{black \& Box } 1 \mid \text{Box } 1)$$
$$+ P(\text{guess } 1 \mid \text{white \& Box } 1)P(\text{white \& Box } 1 \mid \text{Box } 1)$$
$$= \tfrac{18}{23} \times \tfrac{1}{2} + 0 \times \tfrac{1}{2} = \tfrac{9}{23}.$$

Similarly $P(\text{guess correctly} \mid \text{Box } 2) = \frac{5}{23} \times \frac{1}{3} + \frac{11}{23} \times \frac{2}{3} = \frac{9}{23}$ and
$P(\text{guess correctly} \mid \text{Box } 3) = 0 \times \frac{1}{4} + \frac{12}{23} \times \frac{3}{4} = \frac{9}{23}$.

# Section 1.6

1. This is just like the birthday problem:

$P(2 \text{ or more under same sign}) = 1 - P(\text{all different signs})$.

$$n = 2 : \text{chance} = 1 - \frac{11}{12} = \frac{1}{12} < \frac{1}{2}$$
$$n = 3 : \text{chance} = 1 - \frac{11}{12} \times \frac{10}{12} = \frac{17}{72} < \frac{1}{2}$$
$$n = 4 : \text{chance} = 1 - \frac{11}{12} \times \frac{10}{12} \times \frac{9}{12} = \frac{123}{288} < \frac{1}{2}$$
$$n = 5 : \text{chance} = 1 - \frac{11}{12} \times \frac{10}{12} \times \frac{9}{12} \times \frac{8}{12} = \frac{178}{288} > \frac{1}{2}$$

So, the answer is 5.

2. Assume the successive hits are independent, each occurring with chance 0.3.

a) $1 - (0.7)^2 = 0.51$

b) $1 - (0.7)^3 = 0.657$

c) $1 - (0.7)^n$

3. a) $P(\text{at least two heads} \mid \text{at least one head}) = \dfrac{P(\text{at least two H and at least one H})}{P(\text{at least one H})}$

$= \dfrac{P(\text{at least two H})}{P(\text{at least one H})} = \frac{3(2/3)^2(1/3)+(2/3)^3}{1-(1/3)^3} = \frac{.741}{.963} = .7695.$

b) $P(\text{exactly one H} \mid \text{at least one H}) + P(\text{at least two H} \mid \text{at least one H}) = 1$,
because (at least one H) = (exactly one H) $\cup$ (at least two H), and the two events are disjoint.
So $P(\text{exactly one H} \mid \text{at least one H}) = 1 - .7695 = .2305$.

4. a) $\frac{1}{20} \times \frac{9}{20} \times \frac{1}{20} = 9/8000$

b) $\left(\frac{1}{20} \times \frac{9}{20} \times \frac{19}{20}\right) + \left(\frac{1}{20} \times \frac{11}{20} \times \frac{1}{20}\right) + \left(\frac{19}{20} \times \frac{9}{20} \times \frac{1}{20}\right) = 353/8000$

c) $P(\text{jackpot}) = \frac{3}{20} \times \frac{1}{20} \times \frac{3}{20} = 9/8000$, same as before;
$P(\text{two bells}) = \left(\frac{3}{20} \times \frac{1}{20} \times \frac{17}{20}\right) + \left(\frac{3}{20} \times \frac{19}{20} \times \frac{3}{20}\right) + \left(\frac{17}{20} \times \frac{1}{20} \times \frac{3}{20}\right) = 273/8000$
The chance of the jackpot is the same on both machines, but the 1-9-1 machine encourages you to play, because you have a better chance of two bells. It will seem that you are "close" to a jackpot more frequently.

5. a) $P(\text{at least one student has the same birthday as mine})$
$= 1 - P(\text{all } n-1 \text{ other students have different birthdays from mine})$
$= 1 - (364/365)^{n-1}$
[My birthday is one particular day: in order for the event to occur, each of the other students must have been born on one of the remaining 364 days .]

b) Use the argument in Example 3. The desired event has probability at least 1/2 if and only if its complement has probability at most 1/2. This occurs if and only if $(364/365)^{n-1} \le 1/2 \iff$
$n - 1 \ge \frac{\log(1/2)}{\log(364/365)} \approx 253.6$. So $n = 254$ will do.

c) The difference between this and the standard birthday problem is that a *particular* birthday (say January 1st) must occur twice in the class. This is a much stricter requirement than in the usual birthday problem.

6. a) $p_8 = p_9 = \cdots = 0$. (Never need to roll more than seven times)
$p_1 = 0$
$p_2 = 1/6$
$p_3 = (5/6) \cdot (2/6)$ (2nd different from first, third the same as first or second)

$p_4 = (5/6) \cdot (4/6) \cdot (3/6)$
$p_5 = (5/6) \cdot (4/6) \cdot (3/6) \cdot (4/6)$
$p_6 = (5/6) \cdot (4/6) \cdot (3/6) \cdot (2/6) \cdot (5/6)$
$p_7 = (5/6) \cdot (4/6) \cdot (3/6) \cdot (2/6) \cdot (1/6) \cdot 1$

b) $p_1 + \cdots + p_{10} = 1$: you must stop before the tenth roll, and the events determining $p_1$, $p_2$, etc., are mutually exclusive.

c) Of course you can compute them and add them up. Here's another way. In general, let $A_i$ be the event that the first $i$ rolls are different, then $p_i = P(A_{i-1}) - P(A_i)$ for $i = 2, \ldots, 7$, with $P(A_1) = 1$, and $P(A_7) = 0$. Adding them up, you can easily check that the sum is 1.

7. a) $p_3(p_1 + p_2 - p_1 p_2) = p_3 p_1 q_2 + p_3 q_1 p_2 + p_3 p_1 p_2$.

b) $p_4 + P(\text{flows along top}) - p_4 \cdot P(\text{flows along top})$ where $P(\text{flows along top})$ was calculated in a).

8. a) $P(B_{12} \text{ and } B_{23}) = P(\text{all three have the same birthdates}) = \frac{365}{365} \cdot \frac{1}{365} \cdot \frac{1}{365} = \frac{1}{(365)^2}$.
$P(B_{12}) = \frac{365}{365} \cdot \frac{1}{365} = \frac{1}{365} = P(B_{23})$.
So $P(B_{12} \text{ and } B_{23}) = \frac{1}{(365)^2} = \frac{1}{365} \cdot \frac{1}{365} = P(B_{12})P(B_{23})$.
Therefore, $B_{12}$ and $B_{23}$ are independent!

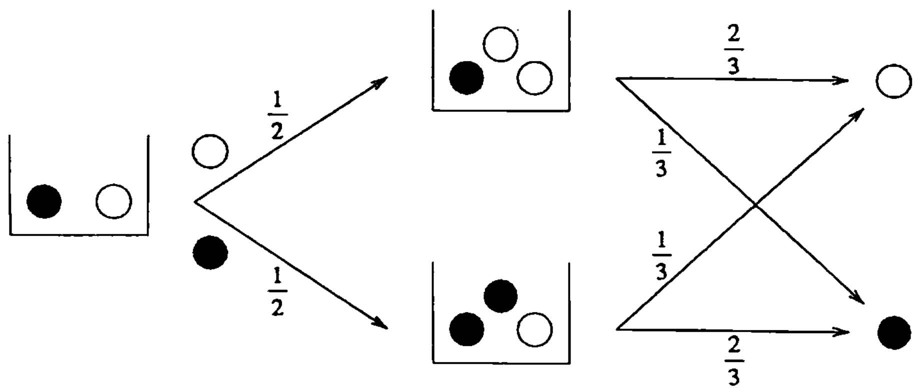
b) No. If you tell me $B_{12}$ and $B_{23}$ have occurred, then all three have the same birthday, so $B_{13}$ also has occurred. That is, $P(B_{12}B_{23}B_{31}) = \frac{1}{365^2} \neq \frac{1}{365^3} = P(B_{12})P(B_{23})P(B_{31})$.

c) Yes, each pair is independent by the same reason as a).

# Chapter 1: Review

1. $P(\text{both defective} \mid \text{item picked at random defective})$
$= \frac{P(\text{both defective})}{P(\text{item picked at random defective})} = \frac{3\%}{3\% + 5\% \times 1/2} = \frac{6}{11}$

2.




From the above tree diagram,
$P(\text{black}) = P(\text{white}) = 1/2$,
$P(\text{black, black}) = P(\text{white, white}) = 1/3$,
$P(\text{black, white}) = P(\text{white, black}) = 1/6$.
So $P(\text{white} \mid \text{at least one of two balls drawn was white}) = \frac{1/6+1/3}{1/6+1/6+1/3} = 3/4$.

3. False. Of course $P(HHH \text{ or } TTT) = 1/4$. The problem with the reasoning is that while two of the coins at least must land the same way, which two is not known in advance. Thus given say two or more $H$'s, i.e. $HHH$, $HHT$, $HTH$, or $THH$, these 4 outcomes are equally likely, so $P(HHH \text{ or } TTT \mid \text{at least two H's}) = 1/4$, not $1/2$. Similarly given at least two T's.

4. a) $P(\text{black} \mid \text{Box } 1) = 3/5 = P(\text{black} \mid \text{Box } 2)$
and $P(\text{red} \mid \text{Box } 1) = 2/5 = P(\text{red} \mid \text{Box } 2)$.
Hence $P(\text{black}) = P(\text{black} \mid \text{Box } i)$ and $P(\text{red}) = P(\text{red} \mid \text{Box } i)$ for $i = 1, 2$, and so the color of the ball is independent of which box is chosen. Or you can check that $P(\text{black, Box } 1) = P(\text{black})P(\text{Box } 1)$, etc, from the following table.

| | Box 1 | Box 2 | |
|---|---|---|---|
| black | 3/10 | 3/10 | 3/5 |
| red | 1/5 | 1/5 | 2/5 |
| | 1/2 | 1/2 | |

b)

| | Box 1 | Box 2 | |
|---|---|---|---|
| black | 3/10 | 5/18 | 26/45 |
| red | 1/5 | 2/9 | 19/45 |
| | 1/2 | 1/2 | |

Observe that $P(\text{black, Box } 1) \neq P(\text{black})P(\text{Box } 1)$, so color of ball is not independent of which box is chosen.

5. For either of the permissible orders in which to attempt the tasks, we have

$$\{\text{pass test}\} = S_1 S_2 \cup S_1^c S_2 S_3$$

where $S_i$ denotes the event {task $i$ performed successfully}; you don't have to worry abo[illegible]e third task if you already have performed the first two tasks successfully. For the first order [illegible]y, hard, easy) the probability of passing the test is therefore

$$P(\text{pass test}) = P(S_1 S_2) + P(S_1^c S_2 S_3) = zh + (1-z)hz = zh(2-z).$$

By symmetry, the probability for the second order (hard, easy, hard) is $hz(2-h)$. S[illegible] $-h > 2-z$, the second order maximizes the probability of passing the test.

6. a) Since $P(B) = P(AB) + P(A^c B)$,

$$P(A^c B) = P(B) - P(AB) = P(B) - P(A)P(B) = P(B)(1 - P(A)) = P(B)P(A^c)$$

b) Just reverse the roles of $A$ and $B$ in part a).

c) $P(A^c B^c) = P((A \cup B)^c) = 1 - P(A \cup B)$
$= 1 - P(A) - P(B) + P(AB)$
$= 1 - P(A) - P(B) + P(A)P(B)$ (by independence of $A$ and $B$)
$= (1 - P(A))(1 - P(B)) = P(A^c)P(B^c)$

7. a) The probability that there will be no one in favor of 134 is the probability that the 1st person doesn't favor 134 and the 2nd person doesn't favor 134 and the 3rd person doesn't favor 134 and the 4th person doesn't favor 134. This is just $\frac{20}{50} \times \frac{19}{49} \times \frac{18}{48} \times \frac{17}{47} = .021$

b) The probability that at least one person favors 134 is 1− the probability that no one favors 134; but the probability that no one favors 134 was done in part a). Thus the answer is $1 - .021 = .979$.

c) The probability that exactly one person favors 134 is $\binom{4}{1}$ times the probability that the 1st person favors and the 2nd person doesn't and the 3rd person doesn't and the 4th person doesn't. This is $\binom{4}{1} \times \frac{30}{50} \times \frac{19}{49} \times \frac{18}{48} \times \frac{17}{47} = 0.126$.

d) The probability that a majority favor 134 is the probability that three people favor 134 plus the probability that four people favor 134. This probability is $\binom{4}{1} \times \frac{30}{50} \times \frac{29}{49} \times \frac{28}{48} \times \frac{20}{47} + \frac{30}{50} \times \frac{29}{49} \times \frac{28}{48} \times \frac{27}{47} = 0.472$

8. a) 0.08228  b) 0.7785  c) 0.1866

9. a) By independence, $P(A \text{ and } B \text{ and } C) = P(A) \cdot P(B) \cdot P(C) = \frac{1}{60} = .01666...$

b) $P(A \text{ or } B \text{ or } C) = 1 - P(A^c B^c C^c)$
$= 1 - P(A^c) \cdot P(B^c) \cdot P(C^c) = 1 - \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} = \frac{3}{5} = .6.$
Or use inclusion-exclusion.

c) $P(\text{exactly one of the three events occurs}) = P(AB^cC^c) + P(A^cBC^c) + P(A^cB^cC)$
$= \frac{1}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} + \frac{4}{5} \cdot \frac{1}{4} \cdot \frac{2}{3} + \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{26}{60} = \frac{13}{30} = .433...$

10. a)

$$\begin{aligned} P(\text{same type}) &= P(\text{both A or both B or both AB or both O}) \\ &= P(\text{both A}) + P(\text{both B}) + P(\text{both AB}) + P(\text{both O}) \\ &= (.42)^2 + (.10)^2 + (.04)^2 + (.44)^2 = .3816; \end{aligned}$$

$$P(\text{different types}) = 1 - P(\text{same type}) = 1 - .3816 = .6184$$

b) Since $P(1) + P(2) + P(3) + P(4) = 1$, we need only evaluate three of these; the fourth can be obtained by subtraction.

$$\begin{aligned} P(1) &= P(\text{all A or all B or all AB or all O}) \\ &= P(\text{all A}) + P(\text{all B}) + P(\text{all AB}) + P(\text{all O}) \\ &= (.42)^4 + (.10)^4 + (.04)^4 + (.44)^4 = .0687 \end{aligned}$$

$$
\begin{aligned}
P(4) &= (\#\text{ arrangements})(.42)(.10)(.04)(.44) = (4!)(.0007392) = .01774 \\
P(2) &= \big[(.42)^2(.10)^2 + (.42)^2(.04)^2 + (.42)^2(.44)^2 \\
&\qquad + (.10)^2(.04)^2 + (.10)^2(.44)^2 + (.04)^2(.44)^2\big] \times 6 \\
&\quad + \big[(.42)^3(1 - .42) + (.10)^3(1 - .10) \\
&\qquad + (.04)^3(1 - .04) + (.44)^3(1 - .44)\big] \times 4 \\
&= .5973
\end{aligned}
$$

So $P(3) = 1 - .5973 - .0687 - .01774 = .3163$.

11.

$$
\begin{aligned}
P(\text{fair} \mid \text{HT}) &= \frac{P(\text{fair and HT})}{P(\text{HT})} \\
&= \frac{P(\text{HT} \mid \text{fair})P(\text{fair})}{P(\text{HT} \mid \text{fair})P(\text{fair}) + P(\text{HT} \mid \text{biased})P(\text{biased})} \\
&= \frac{(1/4) \times (f/n)}{(1/4) \times (f/n) + (2/9) \times (b/n)} = \frac{9f}{9f + 8b}
\end{aligned}
$$

12. a) If $n \geq 7$, the chance is 0: the die has only six different faces! If $n = 1$, the chance is 1.

$$
\begin{aligned}
n = 2 &: \text{chance} = \frac{5}{6} \\
3 &: \text{chance} = \frac{5}{6} \cdot \frac{4}{6} \\
4 &: \text{chance} = \frac{5}{6} \cdot \frac{4}{6} \cdot \frac{3}{6} \\
5 &: \text{chance} = \frac{5}{6} \cdot \frac{4}{6} \cdot \frac{3}{6} \cdot \frac{2}{6} \\
6 &: \text{chance} = \frac{5}{6} \cdot \frac{4}{6} \cdot \frac{3}{6} \cdot \frac{2}{6} \cdot \frac{1}{6}
\end{aligned}
$$

b) $1-$ above : the events in a) and b) are complementary.

13.

$$
\begin{aligned}
P(A|B) = P(AB|B) &= \sum_{i=1}^{n} P(AB_i|B) = \sum_{i=1}^{n} P(AB_i)/P(B) \\
&= \sum_{i=1}^{n} P(A|B_i)P(B_i)/P(B) = \sum_{i=1}^{n} P(A|B_i)P(B_i|B) \quad \text{since } B_iB = B_i
\end{aligned}
$$

14. a) Since the prior probabilities of the various boxes are the same, choose the one with the greatest likelihood: Box 100.

b) Here are the prior probabilities and likelihoods given a gold coin is drawn:

| | Box 1 | Box 2 | Box 3 | ... | Box 98 | Box 99 | Box 100 |
|---|---|---|---|---|---|---|---|
| prior: | 2/150 | 1/150 | 2/150 | ... | 1/150 | 2/150 | 1/150 |
| likelihoods: | 1/100 | 2/100 | 3/100 | ... | 98/100 | 99/100 | 100/100 |

Clearly, the product (prior odds $\times$ likelihoods) is greatest for Box 99. So Box 99 is the choice with the greatest posterior probability given a gold coin is drawn.

15. $P(\text{Box 1} \mid \text{gold}) = \frac{1}{1+0+1/2} = \frac{2}{3}$

16. a) Student 1 can choose any of the remaining $n-1$, student 2 can choose any of the eligible $n-2$, student 3 can choose any of the eligible $n-2$ except student 1, student 4 can choose any of the eligible $n-2$ other than students 1 and 2, and so on. So

$$p_r = \frac{(n-1)}{(n-1)} \times \frac{(n-2)}{(n-2)} \times \frac{(n-3)}{(n-2)} \times \frac{(n-4)}{(n-2)} \times \cdots \times \frac{(n-r)}{(n-2)}.$$

b) Use the exponential approximation:

$$\begin{aligned} \log(p_r) &= \log\left(1 - \frac{1}{n-2}\right) + \log\left(1 - \frac{2}{n-2}\right) + \cdots + \log\left(1 - \frac{r-2}{n-2}\right) \\ &\approx \left(\frac{-1}{n-2}\right) + \left(\frac{-2}{n-2}\right) + \cdots + \left(\frac{-(r-2)}{n-2}\right) \quad \text{since } \log(1+z) \approx z \text{ for small } z \\ &= -\frac{1}{n-2} \times \frac{(r-2)(r-1)}{2}. \end{aligned}$$

So for $n = 300$ and $r = 30$,

$$\log(p_{30}) \approx -\frac{28 \times 29}{298 \times 2} = -1.362416$$

so $p_{30} \approx e^{-1.362416} = .256041$.

17. False. They presumably send the same letter to 10,000 people, including the name of the winner, the name of the person to whom the letter is sent, and the name of one other person. In that case, I learn nothing from the information given, so my chances of winning the Datsun are still 1 in 10,000.

# Section 2.1

1. a) $\binom{7}{4}$; b) $\binom{7}{4}(1/6)^4(5/6)^3$

2. $P(2 \text{ boys and } 2 \text{ girls}) = \binom{4}{2}(1/2)^4 = 6/2^4 = 0.375 < 0.5$. So families with different numbers of boys and girls are more likely than those having an equal number of boys and girls, and the relative frequencies are (respectively): 0.625, 0.375.

3. a) $P(2 \text{ sixes in 5 rolls}) = \binom{5}{2}(1/6)^2(5/6)^3 = 0.160751$

   b) $P(\text{at least 2 sixes in 5 rolls}) = 1 - P(\text{at most 1 six})$
   $= 1 - [P(0 \text{ sixes}) + P(1 \text{ six})]$
   $= 1 - (5/6)^5 - \binom{5}{1}(1/6)(5/6)^4 = 0.196245$
   (This is shorter than adding the chances of 2, 3, 4, 5 sixes.)

   c) $P(\text{at most 2 sixes}) = P(0 \text{ sixes}) + P(1 \text{ six}).P(2 \text{ sixes})$
   $= (5/6)^5 + \binom{5}{1}(1/6)(5/6)^4 + \binom{5}{2}(1/6)^2(5/6)^3 = 0.964506$

   d) The probability that a single die shows 4 or greater is $3/6 = 1/2$.
   $P(\text{exactly 3 show 4 or greater}) = \binom{5}{3}(1/2)^5 = 0.3125$

   e) $P(\text{at least 3 show 4 or greater}) = \left[\binom{5}{3} + \binom{5}{4} + \binom{5}{5}\right](1/2)^5 = 0.5$
   (The binomial $(5, 1/2)$ distribution is symmetric about 2.5)

4.

$$P(2 \text{ sixes in first five rolls} \mid 3 \text{ sixes in all eight rolls})$$
$$= \frac{P(2 \text{ sixes in first 5, and 3 sixes in all eight})}{P(3 \text{ sixes in all eight})}$$
$$= \frac{P(2 \text{ sixes in first five, and 1 six in next three})}{P(3 \text{ sixes in all eight})}$$
$$= \frac{\binom{5}{2}(1/6)^2(5/6)^3 \cdot \binom{3}{1}(1/6)(5/6)^2}{\binom{8}{3}(1/6)^3(5/6)^5} = \frac{\binom{5}{2}\binom{3}{1}}{\binom{8}{3}} = \frac{10 \times 3}{56} = 0.535714$$

5. a) $\dfrac{\binom{19}{11}}{\binom{20}{12}}$ b) $\dfrac{\binom{18}{10}}{\binom{20}{12}}$ c) $1 - \left\{ \dfrac{\binom{15}{12}}{\binom{20}{12}} + 5 \times \dfrac{\binom{15}{11}}{\binom{20}{12}} \right\}$

6. a) $P(\text{exactly 4 hits}) = \binom{8}{4}(0.7)^4(0.3)^4 = 0.1361367$

   b) $P(\text{exactly 4 hits} \mid \text{at least 2 hits}) = \frac{P(\text{exactly 4 hits \& at least 2 hits})}{P(\text{at least 2 hits})}$
   $= \frac{P(\text{exactly 4 hits})}{1 - P(\text{exactly 0 hits}) - P(\text{exactly 1 hit})} = 0.1363126$

   c) $P(\text{exactly 4 hits} \mid \text{first 2 shots hit})$
   $= \frac{P(\text{exactly 4 hits \& first 2 shots hit})}{P(\text{first 2 shots hit})}$
   $= \frac{P(\text{first 2 shots hit \& exactly 2 hits in last 6 shots})}{P(\text{first 2 shots hit})}$
   $= P(\text{exactly 2 hits in last 6 shots})$ (since shots are independent)
   $= \binom{6}{2}(0.7)^2(0.3)^4 = 0.059535$

7. The chance that you win in this game is $15/36 = 5/12$ (list all 36 outcomes!), so the chance you win at least four times in five plays is

$$\binom{5}{4}(5/12)^4(7/12) + \binom{5}{5}(5/12)^5 = 0.100469$$

8. Let $n$ be a positive integer. Using the formula for the mode of the binomial $(n, p)$ distribution:

   If $np + p < 1$ then the most likely number of successes is $int\,(np + p) = 0$ .

If $np + p = 1$ then the most likely number of successes is 0 or 1 (both equally likely).

- If $np + p > 1$ then the most likely number of successes is at least 1.

Hence the most likely number of successes is zero if and only if $np+p \leq 1$, and the largest $p$ for which zero is the most likely number of successes is $p = \frac{1}{n+1}$.

9. a) Most likely number $= int\ (326 \times 1/38) = 8$.

$$P(8) = P(6) \cdot \frac{P(7)}{P(6)} \cdot \frac{P(8)}{P(7)}$$

$$= .104840 \cdot \frac{325-7+1}{7} \cdot \frac{1}{37} \cdot \frac{325-8+1}{8} \cdot \frac{1}{37} = .138724$$

b) $P(10) = P(8) \cdot \frac{P(9)}{P(8)} \cdot \frac{P(10)}{P(9)}$
$= .138724 \times \frac{325-9+1}{9} \cdot \frac{1}{37} \cdot \frac{325-10+1}{10} \cdot \frac{1}{37} = .1127847$

c) $P(10 \text{ wins in } 326 \text{ bets})$

$$= P(9 \text{ wins in } 325 \text{ bets}) \times \frac{1}{38} + P(10 \text{ wins in } 325 \text{ bets}) \times \frac{37}{38}$$

$$= .132058 \times \frac{1}{38} + .1127847 \times \frac{37}{38}$$
$$= .1132919$$

10. a) $k/(n+1)$ b) $(n-k+1)/(n+1)$.

11. a) The tallest bar in the binomial $(15, 0.7)$ histogram is at int $((n+1)p) = \text{int}(16 \times 0.7) = \text{int}(11.2) = 11$

b) $P(11 \text{ adults in sample of } 15) = \binom{15}{11}(0.7)^{11}(0.3)^4$

$$= \binom{15}{4}(0.7)^{11}(0.3)^4 = \frac{15 \times 14 \times 13 \times 12}{4 \times 3 \times 2 \times 1} = (0.7)^{11}(0.3)^4 = 0.218623$$

12. a) $P(\text{makes exactly 8 bets before stopping})$
$= P(\text{wins at 8th game, and has won 4 out of the previous seven})$
$= \binom{7}{4}(18/38)^4(20/38)^3 \cdot (18/38) = 0.1216891$

b) $P(\text{plays at least 9 times})$
$= P(\text{wins at most four bets out of the first eight})$
$= (20/38)^8 + \binom{8}{1}(18/38)(20/38)^7 + \binom{8}{2}(18/38)^2(20/38)^6 + \binom{8}{3}(18/38)^3(20/38)^5 + \binom{8}{4}(18/38)^4(20/38)^4$

$= 0.6926167$

13. a) No! Since the child must get one allele from his/her mother, which will be the dominant B, he/she must have brown eyes.

b) Taking one allele from each parent, we see that there are four (equally likely) allele pairs: Bb, Bb, bb, bb. Thus the probability of the child having brown eyes is 50% = 0.5.

c) Once again there are four possibilities: Bb, Bb, Bb, bb. Thus there ia a 75% = 0.75 chance that the child will have brown eyes.

d) Given that the mother is brown eyed and her parents were both Bb, the probability she is Bb is 2/3 and BB 1/3. Thus

$$P(\text{child brown-eyed}) = P(\text{child brown-eyed}|\text{mother Bb}) \cdot P(\text{mother Bb})$$
$$+P(\text{child brown-eyed}|\text{mother BB}) \cdot P(\text{mother BB})$$
$$= \frac{1}{2} \times \frac{2}{3} + 1 \times \frac{1}{3} = \frac{2}{3}$$

So

$$P(\text{woman Bb}|\text{child Bb}) = \frac{P(\text{child Bb}|\text{woman Bb})P(\text{woman Bb})}{P(\text{child Bb})} = \frac{(1/2)(2/3)}{2/3} = \frac{1}{2}$$

14. a) (Ts, Pw); tall and purple.

b)

| Genetic Combinations | Probability | |
|---|---|---|
| (TT, PP) | 1/16 | tall and purple |
| (TT, Pw) | 1/8 | tall and purple |
| (TT, ww) | 1/16 | tall and white |
| (Ts, PP) | 1/8 | tall and purple |
| (Ts, Pw) | 1/4 | tall and purple |
| (Ts, ww) | 1/8 | tall and white |
| (ss, PP) | 1/16 | short and purple |
| (ss, Pw) | 1/8 | short and purple |
| (ss, ww) | 1/16 | short and white |

| | tall and purple | tall and white | short and purple | short and white |
|---|---|---|---|---|
| Probability | 9/16 | 3/16 | 3/16 | 1/16 |

c) $1 - (7/16)^{10} - 10(9/16)(7/16)^{9}$.

15. a) If $0 < p < 1$, then $int\ (np + p) = np$, since $np$ is an integer.

b) Note that $np = int\ (np) + [np - int\ (np)]$.
If $[np - int\ (np)] + p \geq 1$, then $int\ (np + p) = int\ (np) + 1$, which is the integer above $np$.
If $0 < [np - int\ (np)] + p < 1$, then $int\ (np + p) = int\ (np)$, which is the integer below $np$.

c) Consider the case $n = 2, p = 1/3$, where 1 is the closest integer to $np$ and the integer above $np$, but 0 is also a mode. Also 0 is the integer below $np$, but 1 is also a mode.

# Section 2.2

1. The number of heads in 400 tosses has binomial (400, 0.5) distribution. Use the normal approximation with $\mu = 400 \times (1/2) = 200$ and $\sigma = \sqrt{400 \times \frac{1}{2} \times \frac{1}{2}} = 10$.

   a) $P(190 \le H \le 210) \approx \Phi(1.05) - \Phi(-1.05) = 0.7062$
   b) $P(210 \le H \le 220) \approx \Phi(2.05) - \Phi(0.95) = 0.1509$
   c) $P(H = 200) \approx \Phi(0.05) - \Phi(-0.05) = 0.0398$
   d) $P(H = 210) \approx \Phi(1.05) - \Phi(0.95) = 0.0242$

2. Now $\mu = 204$ and $\sigma = 9.998$.

   a) $P(190 \le H \le 210) \approx \Phi(0.65) - \Phi(-1.45) = 0.6686$
   b) $P(210 \le H \le 220) \approx \Phi(1.65) - \Phi(-0.55) = 0.2417$
   c) $P(H = 200) \approx \Phi(-0.35) - \Phi(-0.45) = 0.0368$
   d) $P(H = 210) \approx \Phi(0.65) - \Phi(0.55) = 0.0333$

3. a) Law of large numbers: the first one.
   b) Binomial (100, .5) mean 50, SD 5:

   $$1 - \Phi\left(\frac{54.5 - 50}{5}\right) = 1 - \Phi(.9) = 1 - .8159 = .1841$$

   Binomial (400, .5) mean 200, SD 10:

   $$1 - \Phi\left(\frac{219.5 - 200}{10}\right) = 1 - \Phi(1.95) = 1 - .9744 = .0256$$

4. Let $X$ be the number of patients helped by the treatment. Then $E(X) = 100$, $SD(X) = 8.16$ and $P(X > 120) = P(X \ge 120.5) \approx 1 - \Phi(2.51) = .006$.

5. Want the chance that you win at least 13 times. The number of times that you win has binomial (25, 18/38) distribution. Use the normal approximation with $\mu = 11.84, \sigma = 2.50$:
   $P(13 \text{ or more wins}) = 1 - P(12 \text{ or fewer wins}) \approx 1 - \Phi\left(\frac{12.5 - 11.84}{2.50}\right) = 1 - \Phi(0.26) = 0.3974$

6. The number of opposing voters in the sample has the binomial (200, .45) distribution. This gives $\mu = 90$ and $\sigma = \sqrt{200 \times .45 \times .55} = 7.035$. Use the normal approximation:

   a) The required chance is approximately

   $$\Phi\left(\frac{90.5 - 90}{7.035}\right) - \Phi\left(\frac{89.5 - 90}{7.035}\right) = .5283 - .4717 = .0567 \text{ (about 6\%)}$$

   b) Now the required chance is approximately

   $$1 - \Phi\left(\frac{100.5 - 90}{7.035}\right) = 1 - \Phi(1.49) = 1 - .9319 = 0.0681 \text{ (about 7\%)}$$

7. a) The city A sample has 400 people, the city B sample has 600 people, so city B accuracy is $\sqrt{6/4} = 1.22$ times greater.
   b) Both have the same accuracy, since the absolute sizes of the two samples are equal.
   c) The city A sample has 4000 people, the city B sample has 4500 people, so the city B sample is $\sqrt{4500/4000} = 1.06$ times more accurate, even though the percent of population sampled in city B is smaller than the percent sampled in city A.

8. Use the normal approximation: $\sigma = \sqrt{npq} = 9.1287$. Want relative area between 99.5 and 100.5 under the normal curve with mean $600 \times (1/6) = 100$, and $\sigma = 9.1287$. That is, want area between $-.055$ and $.055$ under the standard normal curve. Required area $= \Phi(.055) - [1 - \Phi(.055)] = .0438$.

9. a) Think of 324 independent trials, each a "success" (the person shows up) with probability 0.9. So the number of arrivals has binomial (324, 0.9) distribution. We want $P$(more than 300 successes in 324 trial: Use the normal approximation with $\mu = 324 \times 0.9 = 291.6$ and $\sigma = \sqrt{324 \times 0.9 \times 0.1} = 5.4$ to compute the required probability:
$1 - \Phi\left(\frac{300.5-291.6}{5.4}\right) = 1 - \Phi(1.65) \approx 0.0495$.

b) Increase: in the long run, each group must show up with probability .9. So effectively, traveling in groups reduces the number of trials, keeping $p$ the same. So the histogram for the proportion of successes has more mass in the tails, since $n$ is smaller.

c) Repeat a), with the 300 seats replaced by 150 pairs, and the 324 people replaced by 162 pairs. So the number of pairs that arrive has binomial (162, 0.9) distribution. Use the normal approximation with $\mu = 162 \times 0.9 = 145.8$ and $\sigma = \sqrt{162 \times 0.9 \times 0.1} = 3.82$. We want $1 - \Phi\left(\frac{150.5-145.8}{3.82}\right) = 1 - \Phi(1.23) = 0.1093$.

10. We have 30 independent repetitions of a binomial $(200, \frac{1}{2})$. For each of these repetitions, the probability of getting exactly 100 heads can be gotten by normal approximation where $\mu = (200 \times \frac{1}{2}) = 100$ and $\sigma = \sqrt{200 \times \frac{1}{2} \times \frac{1}{2}} = 7.07$.

The chance of exactly 100 heads can be approximated by

$$P(100 \text{ heads }) = \Phi\left(\frac{(100+.5)-100}{7.07}\right) - \Phi\left(\frac{(100-.5)-100}{7.07}\right) \approx .5282 - .4718 = .0564$$

Since the students are independent, the probability that all 30 students do not get exactly 100 heads is $(1 - P(100 \text{ heads }))^{30}$ and the normal approximation gives us

$$(1 - P(100 \text{ heads }))^{30} \approx (1 - .0564)^{30} = 0.175$$

11. The number of hits in the next 100 times at bat has binomial (100, 0.3) distribution. Use the normal approximation with $\mu = 30$ and $\sigma = \sqrt{100 \times 0.3 \times 0.7} = 4.58$.

a) $P(\geq 31 \text{ hits}) \approx 1 - \Phi\left(\frac{30.5-30}{4.58}\right) = 1 - \Phi(0.11) = 0.4562$

b) $P(\geq 33 \text{ hits}) \approx 1 - \Phi\left(\frac{32.5-30}{4.58}\right) = 1 - \Phi(0.545) = 0.2929$

c) $P(\leq 27 \text{ hits}) \approx \Phi\left(\frac{27.5-30}{4.58}\right) = \Phi(-0.545) = 0.2929$

d) No, independence would be lost, because if the player has been doing well the last few times at bat, he's more likely to do well the next time. Similarly, if he's been doing badly, then he's more likely to continue doing badly. This will increase all the chances above.

e) From part b), we see that the player has about a 29% chance of such a performance, if his form stayed the same (long run average .300) and if the hits were independent. So it is reasonable to conclude that this performance is "due to chance," since 29% is quite a large probability.

12. For $n = 10,000$ independent trials with success probability $p = 1/2$,

$$\mu = np = 5000 \text{ and } \sigma = \sqrt{npq} = 50$$

Hence, by the normal approximation, the probability of between $5000 - m$ and $5000 + m$ successes is approximately

$$\Phi\left(\frac{m+1/2}{50}\right) - \Phi\left(\frac{-m-1/2}{50}\right) = 2\Phi\left(\frac{m+1/2}{50}\right) - 1$$

This equals 2/3 if and only if $\Phi((m+1/2)/50) = 5/6$ which implies that $(m+1/2)/50 = 0.97$ and $m = 48$. In other words, there is about a 2/3 chance that the number of heads in 10,000 tosses of a fair coin is within about one standard deviation of the mean.

13. First find $z$ such that $\Phi(-z, z) = 95\%$. This means $\Phi(z) = 0.9750$ and by the normal table, $z = 1.96$.

$$95\% = P\left(\hat{p} \text{ in } p \pm 1.96\sqrt{\frac{pq}{n}}\right) \leq P\left(\hat{p} \text{ in } p \pm 1.96\frac{.5}{\sqrt{n}}\right)$$

We want $1.96(.5/\sqrt{n}) \leq 1\%$, so

$$n \geq \left(\frac{1.96(.5)}{.01}\right)^2 = 9604$$

14. a) # working devices in box has binomial $(400, .95)$ distribution. This is approximately normal with $\mu = 380$, $\sigma \approx 4.36$. Required percent $\approx 1 - \Phi(\frac{389.5-380}{4.36}) = 1 - \Phi(2.18) = 0.0146$ (This normal approximation is pretty rough due to the skewness of the distribution. The exact probability is 0.0094 correct to 4 decimal places. The skew-normal approximation is 0.0099 which is much better.)

b) Using the normal approximation, want largest $k$ so that $1 - \Phi(\frac{k-0.5-380}{4.36}) \geq 0.95$ so $\frac{k \quad .5}{\quad 5} =$ $-1.65$, so $k = 373$.

15. a) $\phi'(z) = \frac{1}{\sqrt{2\pi}}(-z)e^{-z^2/2} = -z\phi(z)$

b) $\phi''(z) = -\phi(z) - z(-z\phi(z)) = (z^2 - 1)\phi(z)$

c) Sketch: Outside $(-4, 4)$, they are close to zero.

d) Let $f(x) = \frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})$. Want $f''(x)$

$$f''(x) = \frac{1}{\sigma}\phi''(\frac{x-\mu}{\sigma}) = \frac{1}{\sigma}((\frac{x-\mu}{\sigma})^2 - 1)\phi(\frac{x-\mu}{\sigma})\frac{1}{\sigma^2}$$

e) For $\mu - \sigma < x < \mu + \sigma$, the second derivative in d) is negative. So the curve is concave in this region. For $x > \mu + \sigma$ or $x < \mu - \sigma$. The second derivative is positive, so curve is convex.

16. a) Using the results of Exercise 15,

$$\phi'''(z) = 2z\phi(z) - (z^2 - 1)z\phi(z) = (-z^2 + 3z)\phi(z)$$

b) Fundamental theorem of calculus; $\phi''$ vanishes at $-\infty$. The first equality of integrals is because $\phi'''$ is an odd function.

$$2\phi(\sqrt{3}) = ((\sqrt{3})^2 - 1)\phi(-\sqrt{3}) = \phi''(\sqrt{3}).$$

$$\int_0^{\sqrt{3}} \phi'''(z)dz) = \int_{-\infty}^{\sqrt{3}} \phi'''(z)dz - \int_{-\infty}^{0} \phi'''(z)dz) = 2\phi(\sqrt{3}) - (-\phi(0))$$

c) $\phi'''$ switches signs at $-\sqrt{3}$ and $\sqrt{3}$. $\phi(0) + 2\phi\sqrt{3}$ is the biggest area possible over an interval because areas are negative when the curve is negative.

17. a) $1 - \Phi(z) = 1 - \int_{-\infty}^{z} \phi(x)dx = \int_{-\infty}^{\infty} \phi(x)dx - \int_{-\infty}^{z} \phi(x)dx = \int_{z}^{\infty} \phi(x)dx.$

b) Over the range of the integral $z < x$, so $x/z$ is greater than 1. Multiplying a positive integrand by a number greater than 1 gives an integral with a larger value.

c)

$$\begin{aligned}
1 - \Phi(z) &< \int_z^{\infty} \frac{1}{\sqrt{2\pi}z} e^{-\frac{1}{2}x^2} x\,dx \\
&= \int_{\frac{1}{2}z^2}^{\infty} \frac{1}{\sqrt{2\pi}z} e^{-u} du \qquad [u = \frac{1}{2}x^2, du = x\,dx] \\
&= -\frac{1}{\sqrt{2\pi}z} e^{-u}\Big|_{\frac{1}{2}z^2}^{\infty} \\
&= \frac{1}{\sqrt{2\pi}z} e^{-\frac{1}{2}z^2} = \frac{\phi(z)}{z}
\end{aligned}$$

# Section 2.3

1. Let $p = P(0)$, then $P(k) = R(k) \cdot R(k-1) \cdots R(1)p$.
   Use $\sum_{i=1}^{n} P(k) = 1 - p$ to get the value of $p$.

2. Put $n = 10,000$, $p = 0.5$, $k = 5000$ in Formula (3). Then $\mu = 5000$, $\sigma = 50$, and the desired probability is approximately
$$\frac{1}{\sqrt{2\pi} \times 50} = 0.008 \approx 0.01$$

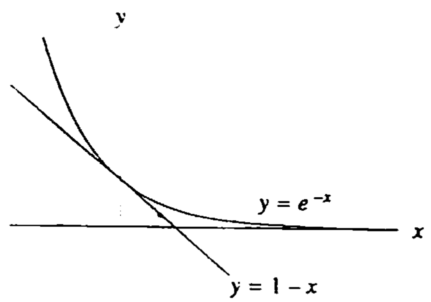
3. a) Use $P(m+1 \text{ in } 2m) = P(m \text{ in } 2m) \cdot R(m+1)$ where $R(m+1) = \frac{2m-(m+1)+1}{m+1} = \left(1 - \frac{1}{m+1}\right)$
   Similarly, $P(m-1 \text{ in } 2m) = P(m \text{ in } 2m)/R(m)$ where $R(m) = \frac{2m-m+1}{m}$. (This could also have been deduced by symmetry.)

   b)
$$\begin{aligned} &P(m+1 \text{ in } 2m+2) \\ &= P(m-1 \text{ in } 2m, \text{ then HH}) + P(m \text{ in } 2m, \text{ then HT or TH}) + P(m+1 \text{ in } 2m, \text{ then TT}) \\ &= \frac{1}{4}P(m-1 \text{ in } 2m) + \frac{1}{2}P(m \text{ in } 2m) + \frac{1}{4}P(m+1 \text{ in } 2m) \end{aligned}$$

   c) Substitute $P(m-1 \text{ in } 2m) = P(m \text{ in } 2m)\left(1 - \frac{1}{m+1}\right)$ and $P(m+1 \text{ in } 2m) = P(m \text{ in } 2m)\left(1 - \frac{1}{m+1}\right)$ into (b), and simplify. This can also be checked by cancelling factorials.

   d) Write
$$\begin{aligned} P(m \text{ in } 2m) &= \frac{P(m \text{ in } 2m)}{P(m-1 \text{ in } 2m-2)} \cdot \frac{P(m-1 \text{ in } 2m-2)}{P(m-2 \text{ in } 2m-4)} \cdots\cdots \frac{P(2 \text{ in } 4)}{P(1 \text{ in } 2)} \cdot P(1 \text{ in } 2) \\ &= \cdot\left(1 - \frac{1}{2m}\right) \cdot \left(1 - \frac{1}{2(m-1)}\right) \cdots\cdots \left(1 - \frac{1}{2 \times 2}\right) \cdot \left(1 - \frac{1}{2 \times 1}\right). \end{aligned}$$

   e) Use the fact that $1 - x \le e^{-x}$ with equality iff $x = 0$ (see diagram).




   Then the inequalities follow easily:
$$0 < P(m \text{ in } 2m) < e^{-\frac{1}{2}\left(\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{m}\right)} < e^{-\frac{1}{2}\log(m)} = \frac{1}{\sqrt{m}}$$
   (For the last inequality, draw a graph of $1/x$ and remember that the area under this graph from 1 to $m$ is $\log(m)$).

   f) By part (c),
$$\frac{\alpha_m}{\alpha_{m-1}} = 1 - \frac{1}{2m}.$$
   Square this, substitute, and simplify:
$$\frac{m+\frac{1}{2}}{m-1+\frac{1}{2}}\left(\frac{\alpha_m}{\alpha_{m-1}}\right)^2 = \frac{2m+1}{2m-1}\left(\frac{2m-1}{2m}\right)^2 = \frac{(2m-1)(2m+1)}{(2m)^2} = 1 - \frac{1}{4m^2}.$$

g) As in (d), write

$$2\left(m+\frac{1}{2}\right)\alpha_m^2 = 2\frac{\left(m+\frac{1}{2}\right)\alpha_m^2}{\left(m-1+\frac{1}{2}\right)\alpha_{m-1}^2}\cdot\frac{\left(m-1+\frac{1}{2}\right)\alpha_{m-1}^2}{\left(m-2+\frac{1}{2}\right)\alpha_{m-2}^2}\cdots\cdot\frac{\left(2+\frac{1}{2}\right)\alpha_2^2}{\left(1+\frac{1}{2}\right)\alpha_1^2}\cdot\frac{\left(1+\frac{1}{2}\right)\alpha_1^2}{\left(0+\frac{1}{2}\right)\alpha_0^2}\left(0+\frac{1}{2}\right)\alpha_0^2$$

$$= \left(1-\frac{1}{4m^2}\right)\left(1-\frac{1}{4(m-1)^2}\right)\cdots\left(1-\frac{1}{4(1)^2}\right)$$

(since $\left(0+\frac{1}{2}\right)\alpha_0^2 = 1/2$). The result follows from factoring

$$\left(1-\frac{1}{4k^2}\right) = \left(1-\frac{1}{2k}\right)\left(1+\frac{1}{2k}\right).$$

h)

$$\alpha_m = P(\text{m in 2m}) = P(\text{mode}) \sim \frac{1}{\sqrt{2\pi}\sigma}$$

with $\sigma = \frac{1}{2}\sqrt{2m}$. So $\alpha_m \sim \frac{K}{\sqrt{m}}$ where $K = \frac{1}{\sqrt{\pi}}$. But this means

$$\frac{2}{\pi} = 2K^2 = 2\lim_{m\to\infty}(m)\alpha_m^2 = 2\lim_{m\to\infty}(m+\frac{1}{2})\alpha_m^2 = \lim_{m\to\infty}\frac{2m+1}{2m}\cdot\frac{2m-1}{2m}\cdot\frac{2m-1}{2(m-1)}\cdot\frac{2m-3}{2(m-1)}\cdots\cdot\frac{3}{2}\cdot\frac{1}{2}.$$

# Section 2.4

1. a) Approximately Poisson(1).

   b) Approximately Poisson(2).

   c) Approximately Poisson(0.3284).

   d) This is the distribution of the number of successes in 1000 independent trials, where the success probability is $p = .998$. The distribution of the number of failures has binomial (1000, .002) distribution, which is approximately Poisson(2). Since #successes + #failures = 1000, it follows that the histogram for the number of successes (i.e., the desired histogram) looks like the left-to-right mirror image of a Poisson(2) histogram, with the block at 0 being sent to 1000, the block at 1 being sent to 999, etc.

2. a) The number of successes in 500 independent trials with success probability .02 has binomial (500, .02) distribution with mean $\mu = 10$. By the Poisson approximation,

$$P(1 \text{ success}) \approx e^{-\mu}\frac{\mu^1}{1!} = \mu e^{-\mu} = .000454.$$

   b) $P(2 \text{ or fewer successes}) = P(0 \text{ or } 1 \text{ or } 2 \text{ successes})$

$$\approx e^{-\mu}\frac{\mu^0}{0!} + e^{-\mu}\frac{\mu^1}{1!} + e^{-\mu}\frac{\mu^2}{2!}$$

$$= e^{-\mu}(1 + \mu + \frac{\mu^2}{2}) = .002769$$

   c) $P(4 \text{ or more successes}) = 1 - P(3 \text{ or fewer successes}) = .989664.$

3. The number of times you see 25 or more sixes has binomial distribution with $\mu = 365 \times .022 = 8.03$.

   a) $P(\text{at least once}) = 1 - P(\text{exactly } 0 \text{ times}) \approx 1 - e^{-\mu} = .999674.$

   b) $P(\text{at least twice}) = 1 - P(\text{exactly } 0 \text{ times}) - P(\text{exactly } 1 \text{ time}) \approx 1 - e^{-\mu} - \mu e^{-\mu} = .997060.$

4. Here $\mu = 365 \times 0.00068 = 0.2482$, and

   a) $1 - e^{-\mu} = .219796;$

   b) $1 - e^{-\mu} - \mu e^{-\mu} = 0.026150.$

5. The number of wins has binomial (52, 1/100) distribution with mean $\mu = 52/100$, and by the Poisson approximation $P(k \text{ wins}) \approx e^{-\mu}\mu^k/k!$.
   $k = 0 : P(0 \text{ wins}) \approx e^{-\mu} = .594521;$
   $k = 1 : P(1 \text{ win}) \approx \mu e^{-\mu} = .309151;$
   $k = 2 : P(2 \text{ wins}) \approx \frac{\mu^2}{2}e^{-\mu} = .080379.$

6. a) The number of black balls seen in a series of 100 draws with replacement has binomial (1000, 2/1000) distribution with mean $\mu = 2$. By the Poisson approximation,

$$P(\text{fewer than 2 black balls}) \approx e^{-\mu}\frac{\mu^0}{0!} + e^{-\mu}\frac{\mu^1}{1!} = e^{-\mu}(1 + \mu) = .406006.$$

$$P(\text{exactly 2 black balls}) = e^{-\mu}\frac{\mu^2}{2!} = .270671.$$

   Calculate the probability of more than 2 black balls by subtraction, conclude that getting fewer than 2 black balls is most likely.

b) $P$(both series see same number of black balls) $= \sum_{k=0}^{\infty} P$(both series see k black balls)

$$\approx \sum_{k=0}^{\infty} e^{-\mu}\frac{\mu^k}{k!}e^{-\mu}\frac{\mu^k}{k!} \text{ by independence}$$

$$= e^{-4}\sum_{k=0}^{\infty}\frac{4^k}{(k!)^2} = 0.207002$$

**7.** a) 2 b) .2659 c) .2475 d) .2565 e) $m = 250$. Normal approximation 0.0266 f) $m = 2$. Poisson approximation: .2565

**8.**

$$P(k) = e^{-\mu}\frac{\mu^k}{k!}$$

$$P(k-1) = e^{-\mu}\frac{\mu^{k-1}}{(k-1)!}$$

so $R(k) = P(k)/P(k-1)$ is given by

$$R(k) = \frac{\mu}{k}$$

which is decreasing as $k$ increases. The maximimum probability will occur at the largest value of $k$ for which $R(k) \geq 1$; after this $P(k)$ will decrease. Thus the maximum occurs at $int(\mu)$. There is a double maximum if and only if $R(k) = 1$ for some $k$. This can only occur if $\mu$ is an integer. The two values of $k$ that maximize are then $\mu$ and $\mu - 1$. There can never be a triple maximum since this would imply that $R(k) = 1$ and $R(k-1) = 1$ for some $k$.

**9.** Assume that each box of cereal has a prize with chance .95, independently of all others. The number of prizes collected by the family in 52 weeks has the binomial distribution with parameters 52 and .95. This gives $\mu = 52 \times .95 = 49.4$, and $\sigma = \sqrt{52 \times .95 \times .05} = 1.57$. Since $\sigma$ is very small (less than 3), the normal approximation is not good. Use the Poisson instead. The number of "dud" boxes has the Poisson distribution with paramter $52 \times .05 = 2.6$. We want the chance of 46 or more prizes, that is, 6 or less duds. This is

$$e^{-2.6}\{1 + 2.6 + 2.6^2/2 + 2.6^3/3! + 2.6^4/4! + 2.6^5/5! + 2.6^6/6!\} = .982830.$$

[For comparison, the exact binomial probability is .985515. The normal approximation gives .993459.]

**10.** Distribution of the number of successes is

$$\text{binomial } (n, 1/N) \approx \text{ Poisson } (n/N) \approx \text{ Poisson } (5/3).$$

$$P(\text{at least two}) = 1 - P(0) - P(1) \approx 1 - e^{-5/3}(1 + 5/3) = 1 - e^{-5/3}\cdot\frac{8}{3} \approx 0.49633 \approx 0.5$$

# Section 2.5

1. a) $\frac{\binom{20}{4}\binom{30}{6}}{\binom{50}{10}}$ b) $\binom{10}{4}(2/5)^4(3/5)^6$

2. a) $\frac{26}{52}\cdot\frac{26}{51}\cdot\frac{25}{50}$ b) $3\times$ a) c) $1-\frac{26}{52}\cdot\frac{25}{51}\cdot\frac{24}{50}$

3. a) $\frac{\binom{4}{4}\binom{48}{9}}{\binom{52}{13}}$ b) $\frac{\binom{3}{3}\binom{48}{9}}{\binom{51}{12}}$ c) $\frac{\binom{4}{4}\binom{48}{9}}{\binom{52}{13}-\binom{48}{13}}$ d) 0

4. The exact chance is $\frac{\sum_{k=45}^{100}\binom{40,000}{k}\binom{60,000}{100-k}}{\binom{100,000}{100}}$

   For the approximation, use the normal curve with $\mu = 40$, $\sigma = 4.9$. Chance is approximately $1-\Phi(\frac{44.5-40}{4.9}) = 0.1788$

5. Solve $\frac{-.05}{\sqrt{.55\times.45/n}} \le -2.326$, then $n \ge 537$.

6. a) $\frac{\binom{36}{13}}{\binom{52}{13}}$ b) $\frac{\binom{40}{13}}{\binom{52}{13}}-\frac{\binom{36}{13}}{\binom{52}{13}}$ c) (a) + 4(b).

7. Denote by $B_i$ the event ($i$th ball is black), similarly for $R_i$.

   a) $P(B_1B_2B_3B_4) = P(B_1)P(B_2|B_1)P(B_3|B_1B_2)P(B_4|B_1B_2B_3) = \frac{50}{80}\frac{49}{79}\frac{48}{78}\frac{47}{77} = .1456.$

   b) This is four times $P(B_1B_2B_3R_4)$, so by d) equals .3716.

   c) $P(B_1B_2B_3R_4) = P(B_1)P(B_2|B_1)P(B_3|B_1B_2)P(R_4|B_1B_2B_3) = \frac{50}{80}\frac{49}{79}\frac{48}{78}\frac{30}{77} = .0929.$

8. Let the outcome space be the set of all 3-element subsets from the set $\{1,...,100\}$, so that, e.g., the 3-set $\{23, 21, 1\}$ means that the winning tickets were tickets #1, #21, and #23. Each of the $\binom{100}{3}$ such 3-sets is equally likely.

   a) The event (one person gets all three winning tickets) is the disjoint union of the 10 equally likely events (person $i$ gets all three tickets). The event (person $i$ gets all three tickets) corresponds to all 3-sets consisting entirely of tickets bought by person $i$. There are $\binom{10}{3}$ such 3-sets, so the desired probability is

$$10\times\frac{\binom{10}{3}}{\binom{100}{3}} = .007421.$$

   b) The event (there are three different winners) is the disjoint union of the $\binom{10}{3}$ equally likely events (the winning tickets were bought by persons $i$, $j$, and $k$) ($i, j, k$ all different). The event (the winning tickets were bought by persons $i$, $j$, and $k$) consists of $\binom{10}{1}\times\binom{10}{1}\times\binom{10}{1}$ 3-sets, so the desired probability is

$$\binom{10}{3}\times\frac{\binom{10}{1}\binom{10}{1}\binom{10}{1}}{\binom{100}{3}} = .742115.$$

   c) By subtraction, .250464. Just to be sure, use the above technique to obtain the desired probability:

$$10\times9\times\frac{\binom{10}{2}\binom{10}{1}}{\binom{100}{3}} = .250464.$$

9. a) Let $A_i$ be the event that the $1^{st}$ sample contains exactly $i$ bad items, $i = 0, 1, 2, 3, 4, 5$. And let $B_j$ be the event that the $2^{nd}$ sample contains exactly $j$ bad items, $j = 0, 1, ..., 10$. Then $P$(2nd sample drawn and contains more than one bad item)

$$= P\left[A_1 \cap \{\cup_{j=2}^{10} B_j\}\right]$$

$$= P\left[\cup_{j=2}^{10} B_j \mid A_1\right] P(A_1)$$

$$= P(A_1) \cdot \{1 - P[B_0 \cup B_1 \mid A_1]\}$$
$$= P(A_1) \cdot \{1 - P(B_0|A_1) - P(B_1|A_1)\}$$
$$= \frac{\binom{10}{1}\binom{40}{4}}{\binom{50}{5}} \cdot \left\{1 - \frac{\binom{9}{0}\binom{36}{10}}{\binom{45}{10}} - \frac{\binom{9}{1}\binom{36}{9}}{\binom{45}{10}}\right\}$$
$$= 0.431337 \cdot (1 - 0.079678 - 0.265592) = 0.282409$$

b) $P(\text{lot accepted})$

$$= P[A_0 \cup \{A_1 \cap (B_0 \cup B_1)\}]$$
$$= P(A_0) + P(A_1) \cdot P[B_0 \cup B_1 \mid A_1]$$
$$= \frac{\binom{10}{0}\binom{40}{5}}{\binom{50}{5}} + \frac{\binom{10}{1}\binom{40}{4}}{\binom{50}{5}} \cdot \left\{\frac{\binom{9}{0}\binom{36}{10}}{\binom{45}{10}} + \frac{\binom{9}{1}\binom{36}{9}}{\binom{45}{10}}\right\}$$
$$= 0.310563 + 0.431337 \cdot (0.079678 + 0.265592) = 0.459491.$$

10. Every sequence of $k_1$ good, $k_2$ bad and $k_3$ indifferent elements has the same probability:

$$(G/N)^{k_1}(B/N)^{k_2}(I/N)^{k_3}$$

This is clear if you think of the sample as $n$ independent trials. The probability of each pattern of $k_1$ good, $k_2$ bad and $k_3$ indifferent elements is a product of $k_1$ factors of $G/N$, $k_2$ factors of $B/N$, and $k_3$ factors of $I/N$. The number of different patterns of $k_1$ good, $k_2$ bad and $k_3$ indifferent elements is $n!/(k_1!k_2!k_3!)$, which proves the formula.

11. The set is $\{g : \max\{0, n - N + G\} \le g \le \min\{n, G\}\}$. This is because the maximum possible number of good elements in sample is $\min\{n, G\}$. And the minimum possible number of good elements in the sample is $n$ minus the max number of bad elements in sample, i.e.

$= n - \min\{n, N - G\} = \max\{0, n - N + G\}$

Formula is correct for all $g$ because $\binom{a}{b} = 0$ for $b < 0$ or $b > a$.

12. There are $\binom{52}{5} = 2,598,960$ possible *distinct* poker hands.

a) There are 52 cards in a pack: $A, 2, 3, ..., 10, J, Q, K$ in each of the four suits: spades, clubs, diamonds and hearts. A straight is five cards in sequence. An $A$ can be at the begining or the end of a sequence, but never in the middle. i.e. $A, 2, 3, 4, 5$ and $10, J, Q, K, A$ are legitimate straights, but $K, A, 2, 3, 4$ (a "round-the-corner" straight) is not. So there are are 10 possible starting points for the straight flush, ($A$ through 10) and 4 suits for it to be in, giving a total of 40 hands.

Thus $P(\text{straight flush}) = \frac{40}{\binom{52}{5}} = 0.0000154.$

b) $P(\text{four of a kind}) = \frac{13 \times 48}{\binom{52}{5}} = \frac{624}{2598960} = 0.000240$

c) $P(\text{full house}) = \frac{13 \times 12 \times \binom{4}{3} \times \binom{4}{2}}{\binom{52}{5}} = \frac{3744}{2598960} = 0.00144$

d) $P(\text{flush}) = P(\text{all same suit}) - P(\text{straight flush})$
$= \frac{4 \times \binom{13}{5} - 4 \times 10}{\binom{52}{5}} = \frac{5108}{2598960} = 0.00197$

e) $P(\text{straight}) = P(5 \text{ consec. ranks}) - P(\text{straight flush}) = \frac{10 \times \binom{4}{1}^5 - 10 \times 4}{\binom{52}{5}} = \frac{10200}{2598960} = 0.00392$

f) $P(\text{three of a kind}) = \frac{13 \times \binom{4}{3} \times \binom{12}{2} \times \binom{4}{1} \times \binom{4}{1}}{\binom{52}{5}} = \frac{54912}{2598960} = 0.0211$

g) $P(\text{two pairs}) = \frac{\binom{13}{2} \times \binom{4}{2} \times \binom{4}{2} \times 11 \times \binom{4}{1}}{\binom{52}{5}} = \frac{123552}{2598960} = 0.0475$

h) $P(\text{one pair}) = \frac{13 \times \binom{4}{2} \times \binom{12}{3} \times \binom{4}{1} \times \binom{4}{1} \times \binom{4}{1}}{\binom{52}{5}} = \frac{1098240}{2598960} = 0.423$

i) Since the events are mutually exclusive, the probability of none of the above is
$1 -$ sum of (a) through (h) $= 0.501.$

13.
$$P(\text{pass}) = 1 - P(\text{fail}) = 1 - [P(>\ 10 \text{ defectives})]^2$$
$$P(>\ 10 \text{ defectives}) \approx 1 - \Phi\left(\frac{10.5 - 500(.05)}{\sqrt{500(.05)(.95)}}\right) \approx 1 - \Phi(-2.98) = .9986$$
$$P(\text{pass}) = 1 - (.9986)^2 = .0028$$

# Chapter 2: Review

1. a) $\binom{10}{4}(1/6)^4(5/6)^6$
   b) $\binom{10}{4}(1/5)^4(4/5)^6$
   c) $\frac{10!}{4!3!2!}/6^{10}$
   d) $\frac{\binom{6}{3}}{\binom{10}{3}} = \frac{\binom{7}{4}}{\binom{10}{4}}$

2. 0.007

3. a) $P(3H) = P(3H|3 \text{ spots})P(3 \text{ spots}) + ... + P(3H|6 \text{ spots})P(6 \text{ spots})$
   $= \left\{\binom{3}{3}(1/2)^3 + \binom{4}{3}(1/2)^4 + \binom{5}{5}(1/2)^5 + \binom{6}{3}(1/2)^6\right\} \times \frac{1}{6}$
   $= \left(\frac{1}{2^3} + \frac{4}{2^4} + \frac{10}{2^5} + \frac{20}{2^6}\right) \times \frac{1}{6} = \frac{1}{6}.$
   b)
   $$P(4 \text{ spots}|3H) = \frac{P(3H|4 \text{ spots})P(4 \text{ spots})}{P(3H)} = \frac{\binom{4}{3}(1/2)^4(1/6)}{(1/6)} = \frac{1}{4}$$

4. $P(\text{exactly 9 tails} \mid \text{at least 9 tails})$
   $= P(\text{exactly 9 tails and at least 9 tails} \mid \text{at least 9 tails})$
   $= P(\text{exactly 9 tails} \mid \text{at least 9 tails})$
   $= \frac{10(1/2)^{10}}{10(1/2)^{10}+(1/2)^{10}} = 10/11$

5. $\frac{\binom{97}{57}}{\binom{100}{60}}$

6. a) $(7/10)^4 - (6/10)^4$
   b) The six must be one of the four numbers drawn. The remaining three numbers must be selected from $\{0, ..., 5\}$. The desired probability is therefore $\binom{6}{3}/\binom{10}{4}$.

7. $k \approx 1025$

8. $\sum_{k=0}^{10} \left[\binom{10}{k}(1/6)^k(5/6)^{10-k}\right]^2$

9. a) 80%
   b) $P(\# \text{ of kids} = x| \text{ at least 2 girls})$
   $$= \frac{0.4 \times \frac{1}{4}}{P(\geq 2 \text{ girls})} \quad \text{for } x = 2$$
   $$= \frac{0.3 \times \frac{4}{8}}{P(\geq 2 \text{ girls})} \quad \text{for } x = 3$$
   $$= \frac{0.1 \times \frac{11}{16}}{P(\geq 2 \text{ girls})} \quad \text{for } x = 4$$
   So $x = 3$ is most likely.
   c) $P(\text{1G,3B \& pick G}) + P(\text{2G,2B \& pick G}) + P(\text{3G,1B \& pick G})$
   $$= \frac{4}{16} \cdot \frac{1}{4} + \frac{6}{16} \cdot \frac{2}{4} + \frac{4}{16} \cdot \frac{3}{4} = 0.4375$$

10. a) We can use the binomial approximation to the hypergeometric distribution with $n = 10$, $p = .15$. So $1-(1-.15)^{10} = .8031$. And the histogram will be that of the binomial $(10, 0.15)$ distribution.
    b) No. Presumably some machines are more reliable than others. Then results of successive tests on a machine picked at random are not independent. So the independence assumption required for the binomial distribution is not satisfied.

11.

$$P(\text{bad}) = \frac{2}{3} \times \frac{1}{100} + \frac{1}{3} \times \frac{2}{100} = 0.0133$$

$$P(2 \text{ bad out of } 12) = \binom{12}{2}(.0133)^2(1 - .0133)^{10} = 0.0102$$

Assume items are bad independent of each other. Reasonable since both A and B produce a large number of items each day.

12. a) 0.423 (See solution to Exercise 2.5.12).

b) Want chance of at most 149 "one pair"s in the first 399 deals

$$= \sum_{k=0}^{149} \binom{399}{k} (0.423)^k (0.577)^{399-k}$$

c) Use normal approximation: $\mu = 168.78$, $\sigma = 9.87$.
Want $\Phi(\frac{149.5 - 168.78}{9.87}) = \Phi(-1.95) = .0256$

13. The number of "dud" seeds in each packet has the binomial (50, .01) distribution, which is very well approximated by the Poisson (.5) distribution. The chance that a single packet has to be replaced is therefore

$$1 - e^{-0.5}\{1 + .5 + .5^2/2\} = .0144.$$

Assuming that packets are independent of each other, the number of replaceable packets out of the next 4000 has the binomial distribution with parameters 4000 and .0144. This gives $\mu = 57.6$ and $\sigma = 7.535$, which is much bigger than 3. So the normal approximation will work well. The chance that more than 40 out of the next 4000 packets have to be replaced is very close to

$$1 - \Phi\left(\frac{40.5 - 57.6}{7.535}\right) = 1 - \Phi(-2.33) = \Phi(2.33) = .9901.$$

14. a) 1/5  b) 2/9
c) line: $\frac{n-k+1}{\binom{n}{k}} = \frac{(n-k+1)!k!}{n!}$; circle: $\frac{n}{\binom{n}{k}} = \frac{(n-k)!k!}{(n-1)!}$.

15. a) $\binom{20}{5}(0.4)^5(0.6)^{15}$

b) $\frac{20!}{2!4!6!8!}(0.1)^2(0.2)^4(0.3)^6(0.4)^8$

c) $P(25th$ ball is red, and there are 2 red balls in first 24 draws)

$$= 0.1 \times \binom{24}{2}(0.1)^2(0.9)^{22}$$

16. a) $\frac{\binom{48}{4}}{\binom{52}{8}}$  b) $\frac{1}{\binom{52}{8}}$  c) $13 \times \frac{\binom{48}{4}}{\binom{52}{8}} - \binom{13}{2} \times \frac{1}{\binom{52}{8}}$

17. a) $\frac{\binom{6}{1}}{6^4}$  b) $\binom{6}{1} \times \binom{5}{1} \times \frac{\binom{4}{1}}{6^4}$  c) $\binom{6}{2} \times \frac{\binom{4}{2}}{6^4}$

18. a) $\binom{7}{3}(1/6)^3(5/6)^4$  b) $6 \times 5 \times \frac{\binom{7}{4}}{6^7}$  c) $\frac{\binom{7}{2}\binom{5}{2}}{6^7}$  d) $6 \times \frac{\binom{7}{2} \times 5!}{6^7}$
e) $P(\text{sum} \geq 9) = 1 - P(\text{sum} < 9)$
$= 1 - P(\text{seven 1's}) - P(\text{six 1's and a 2})$
$= 1 - (1/6)^7 - 7(1/6)^7$.

19. a) $(2/3)^4$;  b) $\binom{4}{1}(2/3)^4(1/3) + (2/3)^4$

20. a) $\sum_{x=0}^{k} \binom{n}{x} q^x p^{n-x}$

b) $(0.99)^8 + 8 \cdot 0.01(0.99)^7 + \binom{8}{2}(0.01)^2(0.99)^6$.

21. For $n$ odd, $\sum_{x=0}^{(n-1)/2} \binom{n}{x} q^x p^{n-x}$.

22. a) Assume each person who buys a ticket shows up independently of all others with probability 0.97. If $n$ tickets are sold, then the number of people who show up has binomial $(n, 0.97)$ distribution. Find $n$ so that

$$.95 \le P(0 \text{ to } 400) \approx \Phi\left(\frac{400.5 - .97n}{\sqrt{(.97)(.03)n}}\right).$$

Therefore $\frac{400.5-.97n}{\sqrt{(.97)(.03)n}} \ge 1.645$, which implies $n \le 407$.

23.

$$P(\text{at least one girl} \mid \text{at least one boy}) = \frac{P(\text{at least one girl and at least one boy})}{P(\text{at least one boy})}$$

$$= \frac{P(\text{not all girls or all boys})}{(.2 \times 1/2) + (.4 \times 3/4) + (.2 \times 7/8) + (.1 \times 15/16)}$$

$$= \frac{(.4 \times 1/2) + (.2 \times 6/8) + (.1 \times 14/16)}{(.2 \times 1/2) + (.4 \times 3/4) + (.2 \times 7/8) + (.1 \times 15/16)}.$$

24. Denote by $\pi(n)$, $n = 0$ to 5, the proportion of families having $n$ children.

a) We assume that in every family, every child is equally likely to be a boy or a girl. In this case, if a family is chosen at random then

$$P(\text{family has } n \text{ children and exactly 2 girls})$$

$$= P(n \text{ children})P(\text{exactly 2 girls} \mid n \text{ children})$$

$$= \pi(n) \times \binom{n}{2}\frac{1}{2^n}$$

(interpret $\binom{n}{2}$ as 0 if $n < k$). Hence

$$P(\text{family has exactly 2 girls}) = \sum_{n=0}^{5} \pi(n)\binom{n}{2}\frac{1}{2^n} = .203125.$$

b) Since each child is equally likely to be chosen, we have

$$P(\text{child comes from a family having exactly 2 girls}) = \frac{\#\text{such children}}{\#\text{children in population}}.$$

Suppose the population has $N$ families, $N$ assumed large. Denominator: There are $N\pi(n)$ *families* having exactly $n$ children, so there are $nN\pi(n)$ *children* belonging to $n$-child families; hence there are $\sum_{n=0}^{5} nN\pi(n)$ children in the population.

Numerator: In part (a) we showed that the proportion of families having $n$ children and exactly 2 girls is $\pi(n)\binom{n}{2}\frac{1}{2^n}$. Thus there are $nN\pi(n)\binom{n}{2}\frac{1}{2^n}$ children who come from families having $n$ children and exactly 2 girls, and $\sum_{n=0}^{5} nN\pi(n)\binom{n}{2}\frac{1}{2^n}$ children who come from families having exactly 2 girls.

Therefore the chance that a child chosen at random from the children in this population comes from a family having exactly 2 girls is

$$\frac{\sum_{n=0}^{5} n\pi(n)\binom{n}{2}\frac{1}{2^n}}{\sum_{n=0}^{5} n\pi(n)} = .294207.$$

25. a) $P(\text{win in 3 sets}) = P(\text{win all three}) = p^3$

$P(\text{win in exactly 4 sets}) = P(\text{win 2 out of first three, then win fourth}) = \binom{3}{2}p^2q \cdot p = 3p^3q$

$P(\text{win in exactly 5 sets}) = P(\text{win 2 out of first four, then win fifth}) = \binom{4}{2}p^2q^2 \cdot p = 6p^3q^2$

b) $P(\text{player A wins the match})$

$= P(\text{wins in exactly 3 sets}) + P(\text{wins in exactly 4 sets}) + P(\text{wins in exactly 5})$

$= p^3 + 3p^3q + 6p^3q^2$

c)

$$P(\text{match lasts only 3 sets}|\text{A won}) = \frac{P(\text{A won in only 3 sets})}{P(\text{A won})}$$

$$= \frac{p^3}{p^3 + 3p^3q + 6p^3q^2} = \frac{1}{1 + 3q + 6q^2}$$

d) If $p = 2/3$, the answer in c) is 0.375.

e) No: if player A has lost the first two sets, he may be nervous about saving the match, and that could affect his performance. In general, the assumption of independence of a player's performance in successive games is rather suspect.

26. a) $8/\binom{10}{3}$

b) Using the inclusion-exclusion formula, the probability in question is

$$P(A\&B \cup B\&C \cup \cdots \cup J\&A)$$

$$= P(A\&B) + P(B\&C) + \cdots + P(J\&A) - \sum P(\cap\text{'s of two pairs}) + \sum P(\cap\text{'s of three pairs}) - \ldots$$

$$= 10 \times 8/\binom{10}{3} - 10/\binom{10}{3} = 7/12$$

because the probability of each intersection of 3 or more pairs is 0.

27. Assume that every quadruplet of 4 exam groups appears equally often in the population of students, so that the proportion of students having a specific set of 4 groups is $1/\binom{18}{4}$.

By equally likely outcomes: There are $\binom{6}{4}3^4$ quadruplets which correspond to different exam days (count the ways to pick the 4 days, then the ways to pick an exam time from each of those 4 days). Hence the desired proportion is $\binom{6}{4}3^4/\binom{18}{4} = .3971$.

By conditioning: Pick a student at random. Let $D_i$ denote the day of the $i$th exam. Then

$$P(D_1, D_2, D_3, D_4 \text{ different}) = P(D_1, D_2 \text{ different}) \times$$

$$P(D_1, D_2, D_3 \text{ different}|D_1, D_2 \text{ different}) \times$$

$$P(D_1, D_2, D_3, D_4 \text{ different}|D_1, D_2, D_3 \text{ different})$$

$$= \frac{15}{17} \times \frac{12}{16} \times \frac{9}{15} = .3971.$$

28. a) $1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + \frac{(-1)^{n-1}}{n!}$

b) $1 - e^{-1}$

29. What kinds of throws are wimpouts? Firstly, they cannot have any numbers in them. Of the the throws with only letters the only ones that contain a $W$ either have four different other letters or 2 or more letters of the same type. Both of these are scoring throws. Thus the only throws that do not score consist entirely of the letters $A, B, C, D$ with no more than two of any letter. Condition on the dice with the $W$ because it has only $A, B, C$. Suppose it is $A$. Then the outcomes of the other dice that lead to a wimpout are: ABCD, ABCC, ABBC, ABDD, ABBD, ACDD, ACCD, BBCC, BBCD, BBDD, BCDD, BCCD, CCDD. Taking into account the different possible orderings and multiplying by 3 we get a total of 450 outcomes. Since there are $6^5 = 7776$ equally likely outcomes possible when you roll 5 dice,

$$P(\text{wimpout}) = \frac{450}{7776} = \frac{25}{432} = 0.05787$$

30. Draw the curve $y = \log x$ from 1 to $n$. Between $x$ and $x + 1$, draw a box with height $\log x$. (The first "box" has 0 height.) Connect the upper left corners of the boxes by straight lines. The area under the curve is the sum of three parts; the area of the boxes, the area of the triangles above the boxes, and the area of the slivers above the triangles and below the curve. The area under the curve is

$$\int_1^n \log x\,dx = n \log n - n + 1 = \log\left[\left(\frac{n}{e}\right)^n\right] + 1$$

The area of the boxes is

$$\sum_{x=1}^{n-1} \log x = \log (n-1)!$$

The area of the triangles is

$$\sum_{x=1}^{n-1} \frac{1}{2}(\log(x+1) - \log x) = \frac{1}{2} \log n$$

By moving the slivers so that they all have their right ends on the point $(2, \log 2)$, you can see that none of the slivers overlap and all of the slivers fit in a box between 1 and 2 with height $\log 2$. So the area of the slivers is some number $c_n < \log 2$ for all $n$. We see that the area under the curve is also equal to

$$\log (n-1)! + \frac{1}{2} \log n + c_n = \log n! - \frac{1}{2} \log n + c_n$$

By setting the two expressions for the area under the curve to be equal, and solving for $\log n!$, we get

$$\log n! = \log \left[ \left( \frac{n}{e} \right)^n \right] + 1 + \frac{1}{2} \log n - c_n = \log \left[ \left( \frac{n}{e} \right)^n \sqrt{n} e^{1-c_n} \right]$$

where $c_n$ increases to a limit $c$ with $c < \log 2$. So

$$n! = \left( \frac{n}{e} \right)^n \sqrt{n} e^{1-c_n} \sim C \left( \frac{n}{e} \right)^n \sqrt{n}$$

where $C = e^{1-c}$.

The exact probability of getting $m$ heads in $2m$ coin tosses is

$$\begin{aligned} \binom{2m}{m} \left( \frac{1}{2} \right)^{2m} &= \frac{(2m)!}{(m!)^2} \left( \frac{1}{2} \right)^{2m} \\ &\sim \frac{C(2m/e)^{2m} \sqrt{2m}}{C^2 (m/e)^{2m} m} \left( \frac{1}{2} \right)^{2m} \\ &= \frac{\sqrt{2}}{C\sqrt{m}} \end{aligned}$$

By the normal approximation, this is approximately

$$\frac{1}{\sqrt{2m(1/2)^2}\sqrt{2\pi}} = \frac{1}{\sqrt{m\pi}}$$

Setting these expressions equal gives $C = \sqrt{2\pi}$.

31. No Solution

32. Those who have computed the probabilities of all the 5 card poker hands and know that a straight is about twice as likely as a flush may be surprised to learn that the answer to this question depends on the size of the hand, $h$. There are $\binom{52}{h}$ distinct $h$ card hands possible. Of these, $4\binom{13}{h}$ are *flushes*: 4 suits (hearts, diamonds, clubs, spades) and we want to choose $h$ cards from any one of these suits. There are $(13-h+1)$ cards that a *straight* of length $h$ could start on, and for each card in the sequence $d, d+1, \ldots, d+h$ there are four possible suits to choose from, giving a total of $(13-h+1)4^h$ possible *straights*. For $h = 1, 2, 3, 4$ , $4\binom{13}{h} > (13-h+1)4^h$ and for $5 \le h \le 13$, $4\binom{13}{h} < (13-h+1)4^h$. The case $h = 1$ is trivial because any single card is both a *straight* and a *flush*. For $h > 13$ neither a *straight* nor a *flush* is possible, so they both have zero probability.

33. a) Note that $P(HH|HH \text{ or } HT \text{ or } TH) = 1/3$. So toss the fair coin twice. Report 1 if the outcome is $HH$, and 0 if the outcome is $HT$ or $TH$, keep trying otherwise.

    b) Note that $P(HT|HT \text{ or } TH) = P(TH|HT \text{ or } TH) = 1/2$ independent of $p$. So toss the biased coin twice. Report 1 if the outcome is $HT$, and 0 if the outcome is $TH$, keep trying otherwise.

34. a) Note that the number of heads and the number of tails have the same distribution, therefore
$P(\#$ of heads from my toss $= \#$ of heads from your toss$)$
$= P(\#$ of heads from my toss $- \#$ of heads from your toss $= 0)$
$= P(\#$ of heads from my toss $+ \#$ of tails from your toss $- m = 0)$
$= P(\#$ of heads from my toss $+ \#$ of tails from your toss $= m)$
And the distribution of # of heads from my toss + # of tails from your toss is binomial $(2m, 1/2)$ by symmetry and the independence of tosses.

b) Let $M_n$ be the number of heads I get on $n$ tosses, and $Y_n$ be the number of heads you get on $n$ tosses. Then
$P(Y_{m+1} > M_m) = P(Y_{m+1} > M_m \mid \text{your } m+1\text{st toss is head}) \times 1/2$
$+ P(Y_{m+1} > M_m \mid \text{your } m+1\text{st toss is tail}) \times 1/2$
$= P(Y_m \geq M_m) \times 1/2 + P(Y_m > M_m) \times 1/2$
$= P(Y_m \geq M_m) \times 1/2 + P(Y_m < M_m) \times 1/2$ by symmetry
$= 1/2.$

35. a) $\sum_{k=20}^{35} \binom{1000}{k} (\frac{1}{38})^k (\frac{37}{38})^{1000-k}$

b) The SD is $\sigma = 5.06$. Use normal approximation.

$$\Phi(\frac{35.5 - 26.316}{5.06}) - \Phi(\frac{19.5 - 26.316}{5.06})$$
$$= \Phi(1.815) - \Phi(-1.347)$$
$$\approx .965 - (1 - .9115) = 0.8765$$

36. c) As $n \to \infty$, $H(k)$ is asymptotic to $e^{-\frac{1}{2}(k-np)^2/npq}$. Hence if $H(k) < \epsilon$, then $k$ satisfies approximately

$$e^{-\frac{1}{2}(k-np)^2/npq} < \epsilon$$
$$-\frac{1}{2}(k-np)^2/npq < \log \epsilon$$
$$(k-np)^2 > 2npq \log \frac{1}{\epsilon}$$
$$k < np - \sqrt{2npq \log \frac{1}{\epsilon}} \text{ or } k > np + \sqrt{2npq \log \frac{1}{\epsilon}}.$$

So the $a, b$ such that $a < m < b$ and both $H(a)$ and $H(b)$ are less than $\epsilon$ are approximately

$$a \approx np - \sqrt{2npq \log \frac{1}{\epsilon}}$$
$$b \approx np + \sqrt{2npq \log \frac{1}{\epsilon}},$$

and

$$b - a \approx 2\sqrt{2npq \log \frac{1}{\epsilon}}.$$

The run time to perform the calculation is then approximately $2K \times \sqrt{2npq \log \frac{1}{\epsilon}}$.

d) Given $p$ and $\epsilon$, the run time to compute every probability in the binomial $(n, p)$ distribution to within $\epsilon$ is proportional to $\sqrt{n}$. Thus it should take $\sqrt{10}$ times as long to compute when $n$ is multiplied by 10. Hence if it takes 2 seconds to compute the binomial (100, 18/38) distribution correct to 3 decimal places, it should take $2\sqrt{10} \approx 6.3$ seconds to compute the binomial (1000, 18/38) distribution correct to 3 decimal places.

37. No Solution

# Section 3.1

1. $X$ has binomial $(3, 1/2)$ distribution. So $P(X = k) = \binom{3}{k}(1/2)^3$ for $k = 0, 1, 2, 3$.

a)

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | 1/8 | 3/8 | 3/8 | 1/8 |

b) Write $Y = \lvert X - 1\rvert$.

| $y$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(Y = y)$ | 3/8 | 1/2 | 1/8 |

2. a) *Joint distribution table for (X, Y) (with replacement)*

| | | *possible values x for X* | | | | *distn* |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | *of Y* |
| *possible* | 1 | 1/16 | 1/16 | 1/16 | 1/16 | 1/4 |
| *values* | 2 | 1/16 | 1/16 | 1/16 | 1/16 | 1/4 |
| *y* | 3 | 1/16 | 1/16 | 1/16 | 1/16 | 1/4 |
| *for Y* | 4 | 1/16 | 1/16 | 1/16 | 1/16 | 1/4 |
| *distn of X* | | 1/4 | 1/4 | 1/4 | 1/4 | 1 |

$P(X \le Y) = 10 \times \frac{1}{16} = \frac{5}{8}$

b) *Joint distribution table for (X, Y) (without replacement)*

| | | *possible values x for X* | | | | distn |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | *of Y* |
| *possible* | 1 | 0 | 1/12 | 1/12 | 1/12 | 1/4 |
| *values* | 2 | 1/12 | 0 | 1/12 | 1/12 | 1/4 |
| *y* | 3 | 1/12 | 1/12 | 0 | 1/12 | 1/4 |
| *for Y* | 4 | 1/12 | 1/12 | 1/12 | 0 | 1/4 |
| *distn of X* | | 1/4 | 1/4 | 1/4 | 1/4 | 1 |

$P(X \le Y) = 6 \times \frac{1}{12} = \frac{1}{2}$

3. a) { 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 }

b) Distribution Table for $S$:

| $s$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(S = s)$ | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

4. a) *Joint distribution table for* $(X_1, X_2)$

| | | possible values $x_1$ for $X_1$ | | | | | | distn |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | of $X_2$ |
| possible | 1 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| values | 2 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| $x_2$ | 3 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| for $X_2$ | 4 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| | 5 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| | 6 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| distn of $X_1$ | | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1 |

b) *Joint distribution table for* $(Y_1, Y_2)$

| | | possible values $y_1$ for $Y_1$ | | | | | | distn |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | of $Y_2$ |
| possible | 1 | 1/36 | 2/36 | 2/36 | 2/36 | 2/36 | 2/36 | 11/36 |
| values | 2 | 0 | 1/36 | 2/36 | 2/36 | 2/36 | 2/36 | 9/36 |
| $y_2$ | 3 | 0 | 0 | 1/36 | 2/36 | 2/36 | 2/36 | 7/36 |
| for $Y_2$ | 4 | 0 | 0 | 0 | 1/36 | 2/36 | 2/36 | 5/36 |
| | 5 | 0 | 0 | 0 | 0 | 1/36 | 2/36 | 3/36 |
| | 6 | 0 | 0 | 0 | 0 | 0 | 1/36 | 1/36 |
| distn of $Y_1$ | | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 | 1 |

5. Distribution of $X_1X_2$:

| $z$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $P(X_1X_2 = z)$ | 1/36 | 2/36 | 2/36 | 3/36 | 2/36 | 4/36 | 2/36 | 1/36 | 2/36 |

| $z$ | 12 | 15 | 16 | 18 | 20 | 24 | 25 | 30 | 36 |
|---|---|---|---|---|---|---|---|---|---|
| $P(X_1X_2 = z)$ | 4/36 | 2/36 | 1/36 | 2/36 | 2/36 | 2/36 | 1/36 | 2/36 | 1/36 |

6. There are 8 equally likely outcomes for three fair coin tosses:

| outcome | probability | $X$ | $Y$ | $X+Y$ |
|---|---|---|---|---|
| HHH | 1/8 | 2 | 2 | 4 |
| HHT | 1/8 | 2 | 1 | 3 |
| HTH | 1/8 | 1 | 1 | 2 |
| HTT | 1/8 | 1 | 0 | 1 |
| THH | 1/8 | 1 | 2 | 3 |
| THT | 1/8 | 1 | 1 | 2 |
| TTH | 1/8 | 0 | 1 | 1 |
| TTT | 1/8 | 0 | 0 | 0 |

a) Joint distribution table for $(X, Y)$

| $Y$ \ $X$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 1/8 | 1/8 | 0 |
| 1 | 1/8 | 2/8 | 1/8 |
| 2 | 0 | 1/8 | 1/8 |

b) $X$ and $Y$ are not independent, since, for instance, $P(X = 2, Y = 0) = 0$, while $P(X = 2)P(Y = 0) = (1/4)(1/4)$.

c)

| $z$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(X+Y = z)$ | 1/8 | 2/8 | 2/8 | 2/8 | 1/8 |

7. a) $(N = 2) = (ABC^c) \cup (AB^cC) \cup (A^cBC)$

b) $P(N = 2) = ab(1-c) + a(1-b)c + (1-a)bc$.

8. a)

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P(X = x)$ | 0.4 | 0.3 | 0.2 | 0.1 |

b)

| $y$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| $P(Y = y)$ | 0.1 | 0.2 | 0.3 | 0.4 |

c) Let $\hat{X}$ be the number of cards until the first ace when dealing from the bottom of the deck. Then $\hat{X}$ has the same distribution as $X$, and

$$Y = 5 - (\hat{X} - 1) = 6 - \hat{X}$$

So

$$P(Y = y) = P(6 - \hat{X} = y) = P(\hat{X} = 6 - y) = P(X = 6 - y)$$

9.

| $x$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| $P(X = x)$ | 5/35 | 10/35 | 12/35 | 8/35 |

10. a) binomial $(n, p)$ b) binomial $(m, p)$ c) binomial $(n + m, p)$

d) yes; functions of disjoint blocks of independent variables are independent.

11. a) By the change of variable principle, $U_n + V_m$ has the same distribution as $S_n + T_m$ in Exercise 10, and this is binomial $(n + m, p)$

b) By (a), $P(U_n + V_m = k) = \binom{n+m}{k} p^k (1 - p)^{n+m-k}$. But also,

$$P(U_n + V_m = k) = \sum_{j=0}^{k} P(U_n = j, V_m = k - j) = \sum_{j=0}^{k} P(U_n = j)P(V_m = k - j)$$

$$= \sum_{j=0}^{k} \binom{n}{j} p^j (1 - p)^{n-j} \binom{m}{k-j} p^{k-j} (1 - p)^{m-(k-j)}$$

$$= [\sum_{j=0}^{k} \binom{n}{j} \binom{m}{k-j}] p^k (1 - p)^{n+m-k}.$$

Now equate the two expressions for $P(U_n + V_m = k)$. [To get the sum from 0 to $n$, note $\binom{x}{y} = 0$ if $y < 0$ or $y > x$]

c) Suppose you have $n + m$ objects, $n$ of which are red, and $m$ blue. You want to choose $k$ out of the $n + m$ objects. Among the $k$ selected objects, some $j$ will be red, where $j$ could be $0, 1, 2, \ldots, n$. So you could just as well pick $j$ red objects first, then $(k - j)$ blue objects.

d) This is similar to b).

e) $\binom{2n}{n}$

12. a) If we think of $N_i$ as just counting the number of times we get category $i$ in $n$ trials, we don't care what happens when we don't get this category. We get category $i$ with probability $p_i$, and don't get it with probability $1 - p_i$, so the distribution of $N_i$ is just binomial $(n, p_i)$.

b) Similarly, $N_i + N_j$ counts the number of times we get either category $i$ or $j$, which happens with probability $p_i + p_j$, so the distribution of $N_i + N_j$ is binomial $(n, p_i + p_j)$.

c) Now we just consider the three categories $i$,$j$, and everything else, which gives a joint distribution which is multinomial $(n, p_i, p_j, 1 - p_i - p_j)$.

13. a)

$$P(X > k) = P(\text{first } k \text{ balls are different color}) = 1 \cdot \frac{2n-2}{2n} \cdot \frac{2n-4}{2n} \cdot \cdots \cdot \frac{2n-2(k-1)}{2n}$$

b)

$$\log P(X > k) = \sum_{j=0}^{k-1} \log(\frac{n-j}{n}) = \sum_{j=1}^{k-1} \log(1 - j/n) \approx \sum_{j=1}^{k-1} -j/n = -\frac{k(k-1)}{2n}$$

So $k(k - 1) = [-\log(\frac{1}{2})]2n$. Roughly $k^2 = n \log 4$, so $k = \sqrt{n \log 4}$. And $k = 1177$ for $n = 10^6$.

14. a) $\binom{g-1}{3} p^3 q^{g-1-3} \cdot p = \binom{g-1}{3} p^4 q^{g-4}$ for $g = 4, 5, 6, 7$.

b) $\sum_{g=4}^{7} \binom{g-1}{3} p^4 q^{g-4}$

c) $P(\text{A wins}) = 1808 / 2187$

d) The outcome of the World series would be the same if the teams played all 7 games. Let $X$ be the number of times that A wins if all 7 games are played. If $X \geq 4$, then A has won the World series, since B can have won at most 3 games. And if A won the World Series, then A won four games before B did, so $X \geq 4$. So $P(\text{A wins}) = P(X \geq 4)$. Of course $X$ has binomial $(7, p)$ distribution!

e) $G$ has range $\{4, 5, 6, 7\}$.
$P(G = g) = P(\text{A wins in } g \text{ games}) + P(\text{B wins in } g \text{ games})$
$= \binom{g-1}{3} p^4 q^{g-4} + \binom{g-1}{3} q^4 p^{g-4}$.
$p = 1/2$ makes $G$ and the winner independent.

15. a) $P(X = Y) = P(X = 1, Y = 1) + P(X = 2, Y = 2) + \cdots + P(X = n, Y = n)$
$= P(X = 1)P(Y = 1) + \ldots + P(X = n)P(Y = n)$ ( by independence)
$= n \cdot \frac{1}{n^2} = \frac{1}{n}$.

b&c) Notice that by symmetry, $P(X > Y) = P(X < Y)$. Moreover,
$1 = P(X > Y) + P(X < Y) + P(X = Y) = 2P(X > Y) + 1/n$.
So $P(X > Y) = \frac{n-1}{2n} = P(X < Y)$.

d) $P(\max(X, Y) = k)$
$= P(X = k, Y < k) + P(X < k, Y = k) + P(X = k, Y = k)$
$= \frac{1}{n} \cdot \frac{k-1}{n} + \frac{k-1}{n} \cdot \frac{1}{n} + \frac{1}{n^2}$
$= \frac{2k-1}{n^2}$

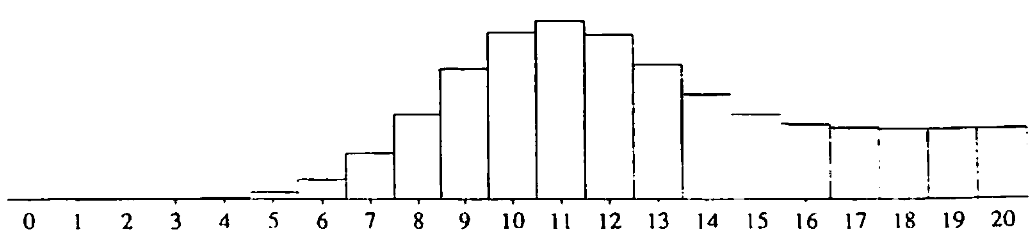
e) $P(\min(X, Y) = k) = P(\max(n + 1 - X, n + 1 - Y) = n + 1 - k)$
$= P(\max(X^*, Y^*) = n + 1 - k)$
$= \frac{2(n+1-k)-1}{n^2}$.
where $X^*$ and $Y^*$ are copies of $X$ and $Y$ respectively.

f) Note that the distribution of $X + Y$ is symmetric about $n + 1$ (recall the sum of two dice). For $2 \leq k \leq n + 1$,
$P(X + Y = k) = \sum_{j=1}^{k-1} P(X = j)P(Y = k - j) = \frac{k-1}{n^2}$.
On the other hand, for $n + 1 < k \leq 2n$
$P(X + Y = k) = P[2(n + 1) - (X + Y) = 2(n + 1) - k] = \frac{[2(n+1)-k]-1}{n^2} = \frac{2n-k+1}{n^2}$.

16. a) $P(X + Y = n) = \sum_{k=0}^{n} P(X = k, X + Y = n)$
$= \sum_{k=0}^{n} P(X = k, Y = n - k)$
$= \sum_{k=0}^{n} P(X = k)P(Y = n - k)$

b) $P(X + Y = 8) = \sum_{k=0}^{8} P(X = k)P(Y = 8 - k) = \sum_{k=0}^{6} P(X = k)P(Y = 8 - k)$
$= \frac{1}{36} \times \frac{5}{36} + \frac{2}{36} \times \frac{4}{36} + \frac{3}{36} \times \frac{3}{36} + \frac{4}{36} \times \frac{2}{36} + \frac{5}{36} \times \frac{1}{36} = \frac{35}{1296} = 0.027$

17. a) $P(Z = k) = P(Y < k, X = k) + P(Y = k, X < k) + P(Y = k, X = k)$
$= (k/21)\binom{20}{k}(1/2)^{20} + (1/21)\sum_{i=0}^{k}\binom{20}{i}(1/2)^{20}$




b) Left tail comes from binomial: very thin. Right tail comes from uniform: thick and flat.

18. a) The number of spots has a distribution symmetric about $E(\text{number of spots}) = 3 \times 3.5 = 10.5$, so $P(11 \text{ or more spots}) = P(10 \text{ or fewer spots}) = 1/2$.

b) The number of spots has a distribution symmetric about $5 \times 3.5 = 17.5$, so $P(18 \text{ or more spots}) = P(17 \text{ or fewer spots}) = 1/2$.

**19.** a) $P(S = k) = \sum_{i+j=k} p_i r_j$ for all $k = 2$ to 12, so in particular
$P(S = 2) = p_1 r_1$;
$P(S = 7) = p_1 r_6 + p_2 r_5 + p_3 r_4 + p_4 r_3 + p_5 r_2 + p_6 r_1$;
$P(S = 12) = p_6 r_6$.

b) $P(S = 7) > p_1 r_6 + p_6 r_1 = P(S = 2)\frac{r_6}{r_1} + P(S = 12)\frac{r_1}{r_6}$.

c) Suppose the values are equally likely. Then from (b) and calculus, $1 > \frac{r_6}{r_1} + \frac{r_1}{r_6} \geq 2$: contradiction!

d) Yes; for example, let $X$ take values 1, 2 with probability 1/2 each, and $Y$ take values 1, 3 with probability 1/2 each. Then $X + Y$ has uniform distribution over 2, 3, 4, 5.

**20.** No. Counterexample: Let $X_1, X_2, X_3$ be the indicators of the events $H_1, H_2, S$ in Example 1.6.8.

**21.** Yes. Proof by mathematical induction.

**22.** a) $P(X = x) = \sum_y P(X = x, Y = y) = \sum_y f(x)g(y) = f(x)\sum_y g(y)$
Similarly $P(Y = y) = g(y)\sum_x f(x)$.

b) If $P(X = x, Y = y) = f(x)g(y)$, then
$P(X = x)P(Y = y) = f(x)\sum_j g(j) \cdot g(y)\sum_i f(i) = f(x)g(y)\left(\sum_j g(j)\right)\left(\sum_i f(i)\right)$.
So we just need to show $\left(\sum_j g(j)\right)\left(\sum_i f(i)\right) = 1$:

$$1 = \sum_i \sum_j P(X = i, Y = j) = \sum_i \sum_j f(i)g(j) = \left(\sum_i f(i)\right)\left(\sum_j g(j)\right)$$

Note: In this calculation $x$ and $y$ are fixed. The sums are over dummy variables which are not called $x$ and $y$, to avoid confusion with the fixed values.

**23.** If $X \leq T$, then $Y \leq T$, since $Y \leq X$. Hence

$$(X \leq T) \subset (Y \leq T) \text{ and } P(X \leq T) \leq P(Y \leq T).$$

This argument still works if $T$ is a random variable.

**24.** a) Let $p_i = P(X = i \bmod 2)$, $i = 0, 1$. Then $p_0 + p_1 = 1$,

$$P(X + Y \text{ is even}) = p_0^2 + p_1^2 = 1 - 2p_0(1 - p_0) \geq 1 - 2 \cdot \frac{1}{4}.$$

b) Let $p_i = P(X = i \bmod 3)$, $i = 0, 1, 2$. Then $p_0 + p_1 + p_2 = 1$,

$$P(\text{X + Y + Z is a multiple of 3}) = p_0^3 + p_1^3 + p_2^3 + 6p_0p_1p_2.$$

Now write simply $p, q, r$ for $p_0$, $p_1$, $p_2$. So

$$p + q + r = 1, 0 \leq p, q, r \leq 1.$$

To show: $p^3 + q^3 + r^3 + 6pqr \geq \frac{1}{4}$. Consider

$$1 = (p + q + r)^3 = p^3 + q^3 + r^3 + 6pqr + 3[p^2q + p^2r + q^2p + q^2r + r^2p + r^2q].$$

Notice that $p(q + r) = pq + pr \leq 1/4$, $q(p + r) = qp + qr \leq 1/4$, and $r(p + q) = rp + rq \leq 1/4$. The probability in question is thus

$$1 - 3[p[pq + pr] + q[qp + qr] + r[rp + rq]]$$
$$\geq 1 - 3[(p + q + r) \cdot 1/4] \geq 1 - 3/4 = 1/4.$$

# Section 3.2

1. $15 \times 1 + 25 \times .2 + 50 \times .7 = 41.5$

2. Average of list $1 = 1 \times .2 + 2 \times .8 = 1.8$; Average of list $2 = 3 \times .5 + 5 \times .5 =$

   a) $1.8 + 4 = 5.8$ (distributive property of addition)

   b) $1.8 - 4 = 2.2$

   c&d) Can't do it: need to know the order of the numbers in the two lists.

3. $E(\#\text{ sixes in 3 rolls}) = 3 \times \frac{1}{6} = \frac{1}{2}$.

   $E(\#\text{ odd numbers in 3 rolls}) = 3 \times \frac{1}{2} = \frac{3}{2}$.

4. If 25 of the numbers are 8, then all the others must be 0, since the average must be 2. If 26 or more of the numbers are 8 or more, there is no way the average can be 2, since some of the other numbers would have to be negative.

5. Suppose he bets on 6, and let $N$ be his net gain.

| $k$ | $P(\text{get } k \text{ 6's})$ | $N$ | $n \times P(N = n)$ |
|---|---|---|---|
| 3 | $(1/6)^3$ | 3 | $3 \times (1$ [illegible] $)^3$ |
| 2 | $3(5/6)(1/6)^2$ | 2 | $2 \cdot 3(5/$ [illegible] $/6)^2$ |
| 1 | $3(5/6)^3(1/6)$ | 1 | $3(5/6)$ [illegible] $(1/6)$ |
| 0 | $(5/6)^3$ | $-1$ | $-(5/6)^3$ |

   Therefore $E(N) = -.078705$, and the gambler expects to lose about 8 cents per game in the long run.

6. $X = I_1 + I_2 + \cdots + I_7$ where $I_i$ is an indicator random variable indicating whether the $i$th card is a spade. Then $E(X) = 7P(\text{first card is a spade}) = \frac{7}{4}$

7. $E(X) = \sum_1^n p_i$, by linearity of $E$. No more assumptions required.

8. $E\left[(X+Y)^2\right] = E(X^2) + 2E(XY) + E(Y^2) = 17$.

9. $E(X-Y)^2 = E(X^2) - 2E(XY) + E(Y^2) = p - 2pr + r$

10. a) Write $Y = (I_A + I_B)^2$.
    If $I_A = 0$ and $I_B = 0$ (this occurs with prob $(1 - P(A)) \cdot (1 - P(B))$ ), then $Y = (0+0)^2$.
    If $I_A = 1$ and $I_B = 0$ (this occurs with prob $P(A) \cdot (1 - P(B))$ ), then $Y = (1+0)^2$.
    If $I_A = 0$ and $I_B = 1$ (this occurs with prob $(1 - P(A)) \cdot P(B)$ ), then $Y = (0+1)^2$.
    If $I_A = 1$ and $I_B = 1$ (this occurs with prob $P(A) \cdot P(B)$ ), then $Y = (1+1)^2$.
    So

| $y$ | $P(Y = y)$ |
|---|---|
| 0 | $(1 - P(A)) \cdot (1 - P(B))$ |
| 1 | $P(A) \cdot (1 - P(B)) + (1 - P(A)) \cdot (P(B))$ |
| 4 | $P(A) \cdot (P(B))$ |

    b) Use properties of expectation described in this section: Note that $Y = (I_A + I_B)^2 = I_A + 2I_{AB} + I_B$ (using $x^2 = x$ for $x = 0$ and $x = 1$, so $I_A^2 = I_A$ and $I_B^2 = I_B$) so

$$E(Y) = P(A) + 2P(AB) + P(B) = P(A) + 2P(A)P(B) + P(B).$$

    Alternatively (and this is more cumbersome), use part (a) and the basic definition of expectation.

11. Use Markov's inequality:

$$P(\text{at least one win}) \le \text{expected \# of wins} = \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = .3$$

Or you can use Boole's inequality.
Actually, $P(\text{at least one win}) = 1 - P(\text{no wins}) = 1 - \frac{900}{1000} \cdot \frac{899}{999} \cdot \frac{898}{998} = .271$. The bound is close because the bound pretends the events ($i$th ticket wins) are mutually exclusive. Well, they almost are, because $P(\text{more than 1 win})$ is tiny.

12.

$$E(X) = \sum_x xP(x) \le \sum_x bP(x) \quad (\text{since } P(X \le b) = 1)$$
$$= b\sum_x P(x) = b.$$

Similarly you can show the other inequality.

13. a) By linearity of expectation, $10 \times 3.5 = 35$.

b) Let $X_1, X_2, X_3$ be the three numbers, and let $Min$ denote the minimum of the first three numbers.
We want $E(X_1 + X_2 + X_3 - Min) = E(X_1 + X_2 + X_3) - E(Min) = 3 \times 3.5 - E(Min)$. Exactly as in Example 9,

$$\begin{aligned} E(Min) &= q_1 + q_2 + q_3 + q_4 + q_5 + q_6 \\ &= 1 + (\frac{5}{6})^3 + (\frac{4}{6})^3 + (\frac{3}{6})^3 + (\frac{2}{6})^3 + (\frac{1}{6})^3 \\ &= \frac{441}{216} = 2.042 \end{aligned}$$

where $q_m = P(Min \ge m)$. So the required expectation is $3 \times 3.5 - 2.042 = 8.458$.

c) Let $Max$ be the maximum of the numbers on the first five rolls. For each $m$ between 1 and 6,

$$P(Max = m) = P(Max \le m) - P(Max \le m - 1) = (\frac{m}{6})^5 - \left(\frac{m-1}{6}\right)^5.$$

So

$$E(Max) = 1 \cdot (\frac{1}{6})^5 + 2\left((\frac{2}{6})^5 - (\frac{1}{6})^5\right) + \cdots + 6\left((\frac{6}{6})^5 - (\frac{5}{6})^5\right) = 5.43.$$

Or: by symmetry, $E(Max) = 7 - E(min(X_1, \ldots, X_5))$.

d) The number of multiples of 3 in the first ten rolls has binomial (10, 1/3) distribution, so its expectation is $10/3 = 3.3333$.

e) The number of faces which fail to appear in the first ten rolls is is $I_1 + I_2 + \cdots + I_6$, where $I_i$ is the indicator of the event (face $i$ fails to appear in the first 10 rolls). Now for each $i$,

$$E(I_i) = P(\text{face } i \text{ fails to appear in the first ten rolls}) = (5/6)^{10}.$$

So the required expectation is $6 \times E(I_1) = 6 \times (5/6)^{10} = 0.969024$.

f) The number of different faces in the first ten rolls equals 6 minus the number of faces which fail to appear. So the required expectation is $6 - 6(5/6)^{10} = 5.030976$.

14. We want $E(N)$, where $N$ is the number of floors at which the elevator makes a stop to let out one or more of the people. $N$ is a counting variable. It's the sum of the ten indicators

$$I(\text{at least one person chooses floor } i), i = 1, \ldots, 10.$$

So by linearity,

$$E(N) = \sum_{i=1}^{10} P(\text{at least one person chooses floor } i).$$

Now for each $i$

$$P(\text{at least one person chooses floor } i) = 1 - P(\text{nobody chooses floor } i) = 1 - (9/10)^{12}$$

by the independence of the people's choices. Hence

$$E(N) = 10 \times \left[1 - (9/10)^{12}\right] \approx 7.18.$$

15. a) If $y \le b$, then $\pi y$ will be the profit, and $\lambda(b-y)$ will be the loss incurred from the items stocked but unsold. So $-\pi y + \lambda(b-y)$ will be the loss in this case. And, if $y > b$, then $\pi b$ will be the profit, since there are only $b$ items to be sold.

b) Write $p(y) = P(Y = y)$. Then

$$\begin{aligned} r(b) = E[L(Y,b)] &= \sum_{y \le b} [-\pi y + \lambda \cdot (b-y)]\, p(y) - \sum_{y>b} \pi b p(y) \\ &= \sum_{y\le b} [-\pi y + \lambda \cdot (b-y)]\, p(y) + \sum_{y \le b} \pi b p(y) - \pi b \\ &= (\lambda + \pi) \sum_{y \le b} (b-y) p(y) - \pi b. \end{aligned}$$

How to minimize over all integers $b$?

**Method 1:** Argue that if we buy $b$ items, then the expected loss will be

$$r(b) = (\lambda + \pi) \sum_{y \le b-1} (b-y)p(y) - \pi b.$$

while if we buy $b-1$ items, then the expected loss will be

$$r(b-1) = (\lambda+\pi) \sum_{y \le b-1} (b-1-y)p(y) - \pi \cdot (b-1)$$

Therefore

$$r(b) - r(b-1) = (\lambda+\pi) \sum_{y\le b-1} p(y) - \pi = (\lambda+\pi) P(Y \le b-1) - \pi.$$

So buying $b$ items will be better than buying $b-1$ items whenever

$$P(Y \le b-1) < \frac{\pi}{\lambda + \pi}. \qquad (*)$$

Let $b*$ be the largest value of $b$ for which $(*)$ holds. (Such a one exists, since the left-hand side tends to 1 as $b$ increases). Note that $b*$ is also the smallest integer for which $P(Y \le b) \ge \pi/(\lambda+\pi)$. Argue that

$$r(0) > \cdots > r(b*-1) > r(b*) \le r(b*+1) \le \cdots$$

so $r(y)$ is minimized at $y = b*$ (and possibly elsewhere).

**Method 2:** View $r(b)$ as a function of a real variable $b$. Argue that $r(b)$ is continuous, and that if $b$ is not an integer, then $r$ is differentiable at $b$ with derivative

$$r'(b) = (\lambda+\pi) \sum_{y \le b} p(y) - \pi = (\lambda+\pi) P(Y \le b) - \pi.$$

The right-hand side is a nondecreasing function of $b$. Let $b*$ denote the smallest integer for which the right-hand side is nonnegative. Argue that if $b < b*$, then $r(b) > r(b*)$, and if $b \ge b*$, then $r(b) \ge r(b*)$. Hence $r$ attains its minimum over real $b$ at $b*$ (and possibly elsewhere).

16. a) $P(X_i = k) = P(X_1 = k) = P(\text{first } k \text{ cards are non-aces, next is ace}) = \frac{(48)_k \cdot 4}{(52)_{k+1}}$.

b) $X_1 + X_2 + X_3 + X_4 + X_5 + 4 = 52 \Longrightarrow 5E(X_1) = 48 \Longrightarrow E(X_i) = E(X_1) = 9.6$.

c) No. For instance, $P(X_1 = 30, X_2 = 30) = 0$, but $P(X_1 = 30) = P(X_2 = 30) > 0$.

**17.** a) $P(D \le 9) = P(3 \text{ red balls are among the first 9 draws}) = \frac{\binom{10}{9-3}}{\binom{13}{9}}$.

b) $P(D = 9) = P(D \le 9) - P(D \le 8) = \frac{\binom{10}{9-3}}{\binom{13}{9}} - \frac{\binom{10}{8-3}}{\binom{13}{8}}$.

c) Label the blue balls and the green balls, say, $b_1, \ldots, b_{10}$. Then
$D = 3 + \sum_{i=1}^{10} I(b_i \text{ is drawn before the third red ball})$, so
$E(D) = 3 + \sum_{i=1}^{10} P(b_i \text{ is drawn before the third red ball}) = 3 + 10 \times \frac{3}{4} = 10.5$.

**18.** Solve the system of equations

$$p(a) + p(b) = 1$$
$$ap(a) + bp(b) = \mu$$

for $p(a) = \frac{\mu - b}{a - b}$ and $p(b) = \frac{a - \mu}{a - b}$.

**19.** Let $x = \#$ blues. Then $\#$ reds $= 2x$, $\#w = x$, $\#g = 3x$. So

$$p_b = p_w = \frac{1}{7}, \qquad p_n = \frac{2}{7}, \qquad p_g = \frac{3}{7}$$

a) $P(X \ge 4) = P(X = 4) = \left(\frac{5!}{1!1!1!2!}(\frac{1}{7})^3(\frac{2}{7})(\frac{3}{7})\right) \times 2 + \frac{5!}{1!1!1!2!}(\frac{1}{7})^2(\frac{2}{7})^2(\frac{3}{7}) + \frac{5!}{1!1!1!2!}(\frac{1}{7})^2(\frac{2}{7})(\frac{3}{7})^2$.

b) $E(X) = E(I_b) + E(I_w) + E(I_r) + E(I_g)$ where e.g. $I_g =$ indicator "green appears"

$$= (1 - (\frac{6}{7})^5) + (1 - (\frac{6}{7})^5) + (1 - (\frac{5}{7})^5) + (1 - (\frac{4}{7})^5)$$

**20.** Write $p(x) = P(X = x)$. Solve the system of equations

$$p(0) + p(1) + p(2) = 1$$
$$p(1) + 2p(2) = \mu_1$$
$$p(1) + 4p(2) = \mu_2$$

for $p(2) = \frac{\mu_2 - \mu_1}{2}$, $p(1) = 2\mu_1 - \mu_2$, $p(0) = 1 - \left(\frac{\mu_2 - \mu_1}{2}\right) - (2\mu_1 - \mu_2)$.

**21.** a) The indicator of $A^c$ must be 1 when $A^c$ occurs and 0 otherwise; clearly this is true for $1 - I_A$: $A^c = 1 \Leftrightarrow I_A = 0 \Leftrightarrow 1 - I_A = 1$.

b) The indicator of $AB$ must be 1 when $A$ and $B$ both occur and 0 otherwise. Since $I_A I_B = 1 \Leftrightarrow (I_A = 1 \text{ and } I_B = 1) \Leftrightarrow I_A I_B = 1$ we have $I_{AB} = I_A I_B$.

c) We wish to show that the indicator of the union can be found by the given formula, which can be shown by the following application of the rules in a) and b).

$$\begin{aligned} I_{A_1 \cup A_2 \cup \cdots \cup A_n} &= I_{(A_1^c A_2^c \cdots A_n^c)^c} \\ &= 1 - I_{(A_1^c A_2^c \cdots A_n^c)} \\ &= 1 - (I_{A_1^c} I_{A_2^c} \cdots I_{A_n^c}) \\ &= 1 - (1 - I_{A_1})(1 - I_{A_2}) \cdots (1 - I_{A_n}) \end{aligned}$$

d)

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) &= E(I_{A_1 \cup A_2 \cup \cdots \cup A_n}) \\ &= E(1 - (1 - I_{A_1})(1 - I_{A_2}) \cdots (1 - I_{A_n})) \\ &= E(\sum_i I_{A_i} - (\sum_{i<j} I_{A_i} I_{A_j}) + \cdots + ((-1)^{n+1} I_{A_1} I_{A_2} \cdots I_{I_n})) \\ &= \sum_i E(I_{A_i}) - (\sum_{i<j} E(I_{A_i} I_{A_j})) + \cdots + ((-1)^{n+1} E(I_{A_1} I_{A_2} \cdots I_{I_n}))) \\ &= \sum_i P(A_i) - \sum_{i<j} P(A_i A_j) + \cdots + (-1)^{n+1} P(A_1 \cdots A_n) \end{aligned}$$

22. a) Let $I_{3,i}$ be the indicator of a run of length 3 starting at the $i$th trial. Let $P(3,i) = E(I_{3,i})$ be the probability of this event. Then for $n > 3$

$$\begin{aligned}
E(R_{3,n}) &= E\left(\sum_{i=1}^{n-2} I_{3,i}\right) = \sum_{i=1}^{n-2} E(I_{3,i}) = \sum_{i=1}^{n-2} P(3,i) \\
&= P(3,1) + \sum_{i=2}^{n-3} P(3,i) + P(3,n-2) \\
&= p^3(1-p) + (n-4)(1-p)(p^3)(1-p) + (1-p)p^3 \\
&= 2p^3(1-p) + (n-4)(p^3)(1-p)^2
\end{aligned}$$

b) Similarly, let $I_{m,i}$ be the indicator of a run of length 3 starting at the $i$th trial. The for $m < n$

$$\begin{aligned}
E(R_{m,n}) &= \sum_{i=1}^{n-m+1} P(m,i) \\
&= 2p^m(1-p) + (n-m-1)(p^m)(1-p)^2
\end{aligned}$$

For $m = n$, $E(R_{n,n}) = p^n$.

c)

$$R_n = \sum_{m=1}^{n} R_{m,n}$$

so

$$\begin{aligned}
E(R_n) &= \sum_{m=1}^{n} E(R_{m,n}) \\
&= p^n + \sum_{m=1}^{n-1} 2p^m(1-p) + (n-m-1)(p^m)(1-p)^2 \\
&= p^n + 2(p-p^n) + (n-1)(1-p)^2 \sum_{m=1}^{n-1}(p^m) - (1-p)^2 \sum_{m=1}^{n-1} mp^m \\
&= 2p - p^n + \sum_{m=1}^{n-1}(n-m-1)(p^m)(1-p)^2
\end{aligned}$$

Let $\Sigma_1 = \sum_{m=1}^{n-1}(n-m-1)(p^m)(1-p)^2$. Then

$$\Sigma_1 - p\Sigma_1 = (n-1)p(1-p)^2 - \sum_{m=1}^{n-1} p^m(1-p)^2 = (n-1)p(1-p)^2 - (p-p^n)(1-p)$$

so

$$\Sigma_1 = (n-1)p(1-p) - (p-p^n)$$

and finally

$$E(R_n) = 2p - p^n + \Sigma_1 = p + (n-1)p(1-p)$$

d) Let $I_j$ be the indicator that a run of some length starts on the $j$th trial. Then $R_n = \sum_{j=1}^{n} I_j$ and if we let $P(j)$ be the probability that a run of some length starts on the $j$th trial, we have

$$E(R_n) = E\left(\sum_{j=1}^{n} I_j\right) = \sum_{j=1}^{n} E(I_j) = \sum_{j=1}^{n} P(j)$$

where $P(1) = p$ and $P(j) = (1-p)p$ for $j > 1$. Thus

$$E(R_n) = p + (n-1)p(1-p)$$

# Section 3.3

1. a)

| $x$ | 28 | 30 | 31 |
|---|---|---|---|
| $P(X=x)$ | 1/12 | 1/3 | 7/12 |

$E(X) = 30.42$, $SD(X) = 0.86$.

b)

| $x$ | 28 | 30 | 31 |
|---|---|---|---|
| $P(X=x)$ | 28/365 | $4 \times 30/365$ | $7 \times 31/365$ |

$E(X) = 30.44$, $SD(X) = 0.88$

2. $E(Y^2) = 3$, $Var(Y^2) = 15/2$

3. a) $E(2X+3Y) = 2E(X) + 3E(Y) = 5$.
   b) $Var(2X+3Y) = 4Var(X) + 9Var(Y) = 26$.
   c) $E(XYZ) = [E(X)][E(Y)][E(Z)] = [E(X)]^3 = 1$.
   d)

$$\begin{aligned} Var(XYZ) &= E[(XYZ)^2] - [E(XYZ)]^2 = E[X^2Y^2Z^2] - [E(XYZ)]^2 \\ &= E[X^2]E[Y^2]E[Z^2] - [E(X)E(Y)E(Z)]^2 \\ &= [E(X^2)]^3 - [E(X)]^6 = [Var(X) + [E(X)]^2]^3 - [E(X)]^6 = 26. \end{aligned}$$

4. $Var(X_1X_2) = E[(X_1X_2)^2] - [E(X_1X_2)]^2$
   $= E(X_1^2) \cdot E(X_2^2) - [E(X_1X_2)]^2$
   $= (\mu_1^2 + \sigma_1^2) \cdot (\mu_2^2 + \sigma_2^2) - (\mu_1\mu_2)^2$
   $= \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2$

5. By the computational formula for variance,

$$Var(X-a) = E\left[(X-a)^2\right] - [E(X-a)]^2 = E\left[(X-a)^2\right] - (\mu - a)^2.$$

But $Var(X-a) = Var(X) = \sigma^2$, so

$$E\left[(X-a)^2\right] = \sigma^2 + (\mu - a)^2.$$

Thus the mean square distance between $X$ and the constant $a$ is $\sigma^2$ + something non-negative. The mean square distance is minimized if this 'something non-negative' is actually 0, i.e., $a = \mu$. And in this case, $E\left[(X-\mu)^2\right] = \sigma^2 = Var(X)$.

6. Intuitively, 1 and 6 are the two most extreme values, so increasing the probability of these values should increase the variance.

$$\begin{aligned} Var(X_p) &= \sum_{x=1}^{6} x^2 P(X_p = x) - [E(X_p)]^2 \\ &= 1^2\frac{p}{2} + 2^2\frac{1-p}{4} + 3^2\frac{1-p}{4} + 4^2\frac{1-p}{4} + 5^2\frac{1-p}{4} + 6^2\frac{p}{2} - (3.5)^2 \\ &= \frac{2p - 4p - 9p - 16p - 25p + 72p}{4} + \frac{4+9+16+25}{4} - (3.5)^2 \\ &= 5p + \frac{5}{4} \end{aligned}$$

7. $X = X_1 + X_2 + X_3$, where $X_i$ has binomial $(n_i, p_i)$ distribution.
   (a) No! It is binomial only when the $p_i$'s are all the same.
   (b) $E(X) = \sum_{i=1}^{3} n_i p_i$, $Var(X) = \sum_{i=1}^{3} n_i p_i q_i$.

8. a) $N = X_1 + X_2 + X_3$ b) $E(N) = \frac{1}{5} + \frac{1}{4} + \frac{1}{3} = \frac{47}{60}$

c) Note that $N$ is the indicator of the event $A_1 \cup A_2 \cup A_3$, since $A_i$'s are disjoint. So $Var(N) = (\frac{1}{5} + \frac{1}{4} + \frac{1}{3})(1 - \frac{1}{5} - \frac{1}{4} - \frac{1}{3}) = \frac{611}{3600}$

d) $Var(N) = Var(X_1) + Var(X_2) + Var(X_3) = \frac{1}{5} \cdot \frac{4}{5} + \frac{1}{4} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{2}{3} = \frac{2051}{3600}$

e) $Var(N) = \left[\frac{1}{5} \cdot 3^2 + (\frac{1}{4} - \frac{1}{5}) \cdot 2^2 + (\frac{1}{3} - \frac{1}{4}) \cdot 1^2\right] - (\frac{1}{5} + \frac{1}{4} + \frac{1}{3})^2 = \frac{5291}{3600}$

9. The number $N_1$ of individuals who vote Republican in both elections has binomial $(r, 1 - p_1)$ distribution, and the number $N_2$ of individuals who vote Democratic in the first election and Republican in the second has binomial $(n - r, p_2)$ distribution. The number of Republican votes in the second election is then $N_1 + N_2$. So

$$E(N_1 + N_2) = E(N_1) + E(N_2) = r(1 - p_1) + (n - r)p_2$$

Since $N_1$ and $N_2$ are independent, we have

$$Var(N_1 + N_2) = Var(N_1) + Var(N_2) = r(1 - p_1)p_1 + (n - r)p_2(1 - p_2)$$

10. a) $E(X^k) = \sum_{x=1}^n x^k P(X = x) = \sum_{x=1}^n x^k \cdot \frac{1}{n} = \frac{1}{n}(1^k + 2^k + \cdots + n^k) = s(k, n)/n$.

$E\left[(X + 1)^k\right] = \sum_{x=1}^n (x + 1)^k P(X = x) = \frac{1}{n}\left[2^k + 3^k + \cdots + (n + 1)^k\right] = \frac{1}{n}\left[s(k, n + 1) - 1\right]$.

b) By the binomial expansion, $kX^{k-1} + \binom{k}{2}X^{k-2} + \cdots + 1 = (X + 1)^k - X^k$, so

$$E\left[kX^{k-1} + \binom{k}{2}X^{k-2} + \cdots + 1\right] = E\left[(X + 1)^k\right] - E(X^k)$$

$$= \frac{2^k + 3^k + \cdots + (n + 1)^k}{n} - \frac{1 + 2^k + \cdots + n^k}{n} = \frac{(n + 1)^k - 1}{n}.$$

c) Put $k = 2$ in b):

$E(2X + 1) = \frac{(n+1)^2 - 1}{n} = n + 2 \Longrightarrow E(X) = \frac{n+1}{2}$

By a), $s(1, n)/n = E(X)$. So $s(1, n) = \frac{n(n+1)}{2}$.

d) Put $k = 3$ in b):

$E(3X^2 + 3X + 1) = \frac{(n+1)^3 - 1}{n} = n^2 + 3n + 3$.

$\Longrightarrow E(X^2) = \frac{1}{3}\left[n^2 + 3n + 3 - 3E(X) - 1\right] = \frac{1}{3}\left[n^2 + 3n + 3 - \frac{3(n+1)}{2} - 1\right] = \frac{1}{6}(n + 1)(2n + 1)$.

By a), $\frac{s(2,n)}{n} = E(X^2)$. So $s(2, n) = \frac{1}{6}n(n + 1)(2n + 1)$.

e) $Var(X) = E(X^2) - [E(X)]^2 = \frac{1}{6}(n + 1)(2n + 1) - [(n + 1)/2]^2 = \frac{n^2 - 1}{12}$.

f) For $n = 6$, $E(X) = 7/2$ agrees, and $Var(X) = 35/12$ agrees.

g) Put $k = 4$ in b):

$$E(4X^3 + 6X^2 + 4X + 1) = \frac{(n + 1)^4 - 1}{n} = n^3 + 4n^2 + 6n + 4$$

Use part a) to conclude

$$s(3, n) = nE(X^3) = \frac{n}{4}\left(n^3 + 4n^2 + 6n + 4 - 6E(X^2) - 4E(X) - 1\right)$$

$$\frac{n}{4}\left[n^3 + 4n^2 + 6n + 4 - (n + 1)(2n + 1) - 2(n + 1) - 1\right] = \frac{n^2(n + 1)^2}{4} = [s(1, n)]^2.$$

11. $Y = (a - b) + bX$. So

$$E(Y) = a - b + bE(X) = a - b + b \cdot \frac{n + 1}{2} = a + b\left(\frac{n - 1}{2}\right)$$

$$Var(Y) = b^2 Var(X) = b^2 \cdot \left(\frac{n^2 - 1}{12}\right)$$

**12.** a) According to the Chebychev inequality,

$$P(X \geq 20) \leq P(|X-10| \geq 10) \leq \frac{5^2}{10^2} = 1/4.$$

b) If X has binomial $(n,p)$ distribution, then $E(X) = np$ and $SD(X) = \sqrt{np(1-p)}$. Try $np = 10$, and $np(1-p) = 25$, which implies that $1-p = 25/10 = 2.5$. This is impossible!

**13.** Let $n$ denote the number of scores (among the million individuals) exceeding 130, and let $X$ denote the score of one of the million individuals picked at random. Then

$$P(X > 130) = \frac{n}{10^5} \iff n = 10^6 P(X > 130).$$

a) We have, by Chebychev's inequality,
$P(X > 130) = P(X - 100 > 30) \leq P(|X-100| > 30) \leq \frac{Var(X)}{30^2} = \frac{1}{9}$.
So $n \leq 10^6 \times \frac{1}{9} < 111112$.

b) Since the distribution of scores is symmetric about 100,
$P(X - 100 > 30) = P(100 - X > 30)$. Therefore
$P(X > 130) = P(X - 100 > 30) = \frac{1}{2}P(|X-100| > 30) \leq \frac{1}{18}$
and $n \leq 10^6 \times \frac{1}{18} \leq 55556$.

c) $P(X > 130) = P(\frac{X-100}{10} > 3) \approx 1 - \Phi(3) = 0.0013$, hence $n = 10^6 \times 0.0013 = 1300$.

**Note:** We can get a sharper bound in a) using Cantelli's inequality, which says: if $X$ is a random variable with mean $m$ and variance $\sigma^2$, then for $a > 0$

$$P(X - m > a) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

**Proof of Cantelli's inequality:** Without loss of generality, we may assume $m = 0$, $\sigma = 1$. Note that if $X \geq a$ then $(1 + aX)^2 \geq (1 + a^2)^2$. So

$1_{(X>a)} \leq \frac{(1+aX)^2}{(1+a^2)^2}$,

where $1_{(X>a)}$ is the indicator of the event $(X > a)$. Take expectations, and get the result.

Application to part a): $P(X > 130) = P(X - 100 > 30) \leq \frac{10^2}{10^2+30^2} = \frac{1}{10}$.

**14.** a) Let $X$ be the income in thousands of dollars of a family chosen at random. Then $E(X) = 10$ and, by Markov's inequality,

$$P(X \geq 50) \leq \frac{E(X)}{50} = \frac{1}{5}.$$

b) If $SD(X) = 8$ then by Chebychev's inequality

$$P(X \geq 50) \leq P(|X-10| > 5 \cdot 8) \leq \frac{1}{5^2} = \frac{1}{25}.$$

**15.** b) $10\sqrt{8}$

**16.** a)

| $x$ | -2 | -1 | 0 | 3 |
|---|---|---|---|---|
| $P(X=x)$ | .25 | .25 | .25 | .25 |

Thus $E(X) = \frac{-2-1+0+3}{4} = 0$ and $Var(X) = \frac{(-2)^2+(-1)^2+0^2+3^2}{4} - 0^2 = \frac{7}{2} = 3.5$

b) Let $S_n = \sum_{i=1}^n X_i$ where each $X_i$ is one play of this game. Then we wish to know $P(S_{100} \geq 25)$. $E(S_{100}) = 100E(X) = 0$ and $Var(S_{100}) = 100Var(X) = 350$. Furthermore, the sum of 100 draws should look pretty close to normal, so

$$P(S_{100} \geq 25) \approx 1 - \Phi\left(\frac{25 - .5 - 0}{\sqrt{350}}\right) = 1 - \Phi\left(\frac{24.5}{18.71}\right) = .0952$$

17. $E(X) = 1/4$, $SD(X) = \frac{1}{4}\sqrt{11}$; $E(S) = 6.25$, $SD(S) = \frac{5}{4}\sqrt{11}$.

a) $P(S < 0) \approx \Phi\left(\frac{-0.5-6.25}{(5/4)\sqrt{11}}\right) \approx .05$.

b) $P(S = 0) \approx \Phi\left(\frac{0.5-6.25}{(5/4)\sqrt{11}}\right) - \Phi\left(\frac{-0.5-6.25}{(5/4)\sqrt{11}}\right) \approx .03$.

c) $P(S > 0) \approx 1 - \Phi\left(\frac{0.5-6.25}{(5/4)\sqrt{11}}\right) \approx .92$.

18. Let $X_i$ be the amount won on the $i$th bet. Thus $X_i = 6$ if you win on the $i$th bet, and $X_i = -1$ if you lose. Let $S$ be the sum of 300 dollar bets, and we wish to find $P(S > 0)$.

$$\begin{aligned} E(X_i) &= 6 \times \frac{5}{38} + -1 \times \frac{33}{38} = -0.07895 \\ SD(X_i) &= \sqrt{36 \times \frac{5}{38} + 1 \times \frac{33}{38} - (-0.07895)^2} = 2.366 \\ E(S) &= 300 \times -0.07895 = -23.68 \\ SD(S) &= \sqrt{300} \times 2.366 = 40.98 \end{aligned}$$

By the normal approximation,

$$P(S > 0) \approx 1 - \Phi\left(\frac{0 - (-23.68)}{40.98}\right) = 0.2817$$

19. Let $X_i$ be the weight of the $i$th guest in the sample, and let $S = X_1 + X_2 + \cdots + X_{30}$. Want $P(S > 5000)$. Use the normal approximation, with $E(S) = 4500$, and

$$SD(S) = \sqrt{30} \times 55 = 301.25$$

Need the area to the right of $(5000 - 4500)/301.25 = 1.66$ under the standard normal curve. This is $1 - \Phi(1.66) = 1 - 0.9515 = 0.0485$.

20. a) The probability that your profit is \$8,000 or more is the probability that your stock gives you a profit of either \$200 per \$1000 invested or \$100 per \$1000 invested; this probability is .5.

b) Let $S_{100}$ be the sum of the profits for the 100 stocks invested in, and let $X$ be the random variable representing one of the profits.
$E(X) = \frac{200+100+0-100}{4} = 50$ and $Var(X) = \frac{200^2+100^2+0^2+(-100)^2}{4} - 50^2 = 12,500$.
$E(S_{100}) = 100E(X) = 5000$, and since the investments are independent, $Var(S_{100}) = 100Var(X) = 1,250,000$. Using normal approximation,

$$P(S_{100} \geq 8000) = 1 - \Phi\left(\frac{8000 - 50 - 5000}{\sqrt{1,250,000}}\right) = 1 - \Phi\left(\frac{2950}{1118}\right) = .0042$$

In the numerator of the normal approximation term, note that the $8000 - 50$ is just the familiar continuity correction since $S_{100}$ can only take on values which are multiples of \$100.

21. a) Let $X_i$ be the error caused by the $i$th transaction; thus $X_i$ is uniformly distributed between $-49$ and 50. Let $S_{100}$ be the total accumulated error for 100 transactions, and then we wish to know $P(S_{100} > 500$ or $S_{100} < -500)$. We have

$$E(X_i) = \sum_{j=-49}^{50} \frac{j}{100} = .5$$

$$Var(X_i) = \sum_{j=-49}^{50} \frac{j^2}{100} - .5^2 = 833.25$$

So $E(S_{100}) = 50$ and $Var(S_{100}) = 83,325$. Using normal approximation,

$$P(|S_{100}| > 500) \approx 1 - \left(\Phi\left(\frac{500 + .5 - 50}{\sqrt{83,325}}\right) - \Phi\left(\frac{-500 - .5 - 50}{\sqrt{83,325}}\right)\right) = 0.0876$$

b) Using the same notation as in a), we now have that

$$E(X_i) = \sum_{j=-49}^{50} \frac{j \times .75}{99} = .3788$$

except that the coefficient when $j = 0$ is wrong, but this does not affect the answer. Similarly,

$$Var(X_i) = \sum_{j=-49}^{50} \frac{j^2 \times .75}{99} - .3788^2 = 631.30$$

Thus $E(S_{100}) = 37.88$ and $Var(S_{100}) = 63,130$. Using normal approximation,

$$P(|S_{100}| > 500) \approx 1 - \left(\Phi\left(\frac{500 + .5 - 37.88}{\sqrt{63,130}}\right) - \Phi\left(\frac{-500 - .5 - 37.88}{\sqrt{63,130}}\right)\right) = 0.0489$$

22. a) Let $X_i$ be the face showing on the $i$th roll; then $E(X_i) = 7/2$, and $SD(X_i) = \sqrt{35/12}$. Let $\bar{X}_n$ be the average of the first $n$ rolls. Then
$E(\bar{X}_n) = E(X_1) = 7/2$, and
$SD(\bar{X}_n) = SD(X_1)/\sqrt{n} = \sqrt{35/12n}$, and

$$P\left(3\frac{5}{12} \le \bar{X}_n \le 3\frac{7}{12}\right) = P\left(|Z_n| \le \sqrt{\frac{n}{12 \times 35}}\right)$$

where

$$Z_n = \frac{\bar{X}_n - E(\bar{X}_n)}{SD(\bar{X}_n)} = \frac{\bar{X}_n - (7/2)}{\sqrt{35/12n}}$$

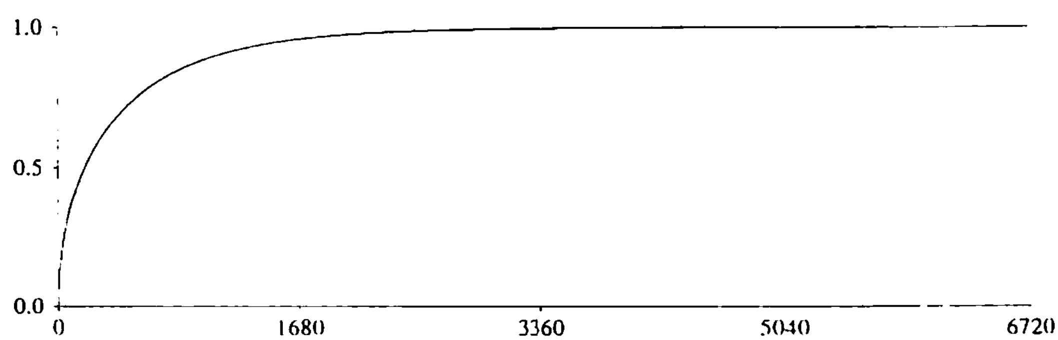
is $\bar{X}_n$ standardized, which will be approximately standard normal for large $n$. When $n = 105$,

$$P\left(|Z_n| \le \sqrt{\frac{n}{12 \times 35}}\right) \approx P(|Z| \le 0.5) = 0.383$$

where $Z$ is standard normal.
The same argument gives the numbers in the following table:

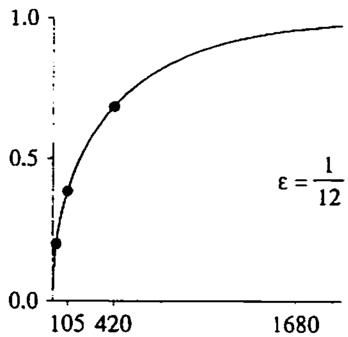
| $n$ | $P\left(3\frac{5}{12} \le \bar{X}_n \le 3\frac{7}{12}\right)$ |
|---|---|
| 105 | $P(\|Z\| \le .5) = .383$ |
| 420 | $P(\|Z\| \le 1) = .6826$ |
| 1680 | $P(\|Z\| \le 2) = .9544$ |
| 6720 | $P(\|Z\| \le 4) \approx 1$ |

b)

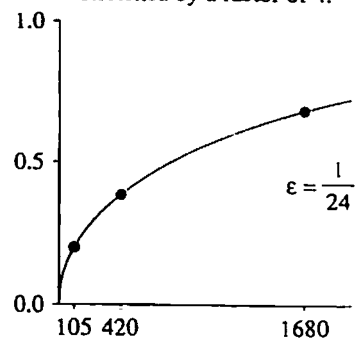


c) $P\left(|\bar{X}_n - 7/2| \leq 1/24\right) \approx P\left(|Z_n| \leq \sqrt{\frac{n}{4\times12\times35}}\right)$.

See that if $\epsilon$ is replaced by $\epsilon/2$, then probabilities that were previously obtained at $n$ are now obtained at $4n$. So the scale on the $n$-axis is stretched by a factor of 4. In general, dividing $\epsilon$ by a factor of $f$ results in a graph with the scale on the $n$-axis stretched by a factor of $f^2$.




Same curve with horizontal axis stretched by a factor of 4.




**23.** Let $X_i$ be the lifetime in weeks of the $i$th battery. Assume $S = X_1 + \cdots + X_{27}$ is approximately normally distributed. Using $E(S) = 27 \times 4 = 108$ and $SD(S) = \sqrt{27} \times 1 = \sqrt{27}$, we have

$$
\begin{aligned}
&P(\text{more than 26 replacements in a 2-year period})\\
&= P(X_1 + \cdots + X_{27} < 2 \times 52)\\
&= P(S < 104)\\
&= P\left(\frac{S-108}{\sqrt{27}} < \frac{-4}{\sqrt{27}}\right)\\
&\approx \Phi(-.77)\\
&\approx .22
\end{aligned}
$$

**24.** a)

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(S_2 = x)$ | .0625 | .25 | .375 | .25 | .0625 |

b) $E(S_{50}) = 50$ and $Var(S_{50}) = 50(.5) = 25$. Thus by normal approximation,

$$P(S_{50} = 50) \approx \Phi\left(\frac{50.5 - 50}{5}\right) - \Phi\left(\frac{49.5 - 50}{5}\right) = 0.0797$$

c) $S_n = X_1 + \cdots + X_n$ where the $X_i$ are independent, each with binomial $(2, 1/2)$ distribution. It follows (Exercise 3.1.11) that $S_n$ has binomial $(2n, 1/2)$ distribution:

$$P(S_n = k) = \binom{2n}{k}\left(\frac{1}{2}\right)^{2n}$$

**25.** a) $k = 1$:

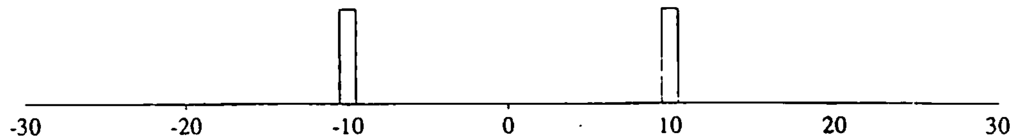
| $x$ | $-10$ | $10$ |
|---|---|---|
| $P(X = x)$ | 1/2 | 1/2 |

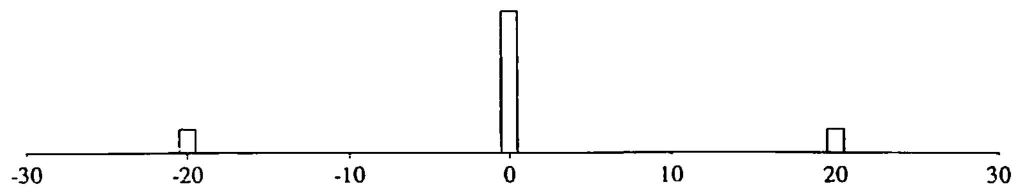
$k = 2$:

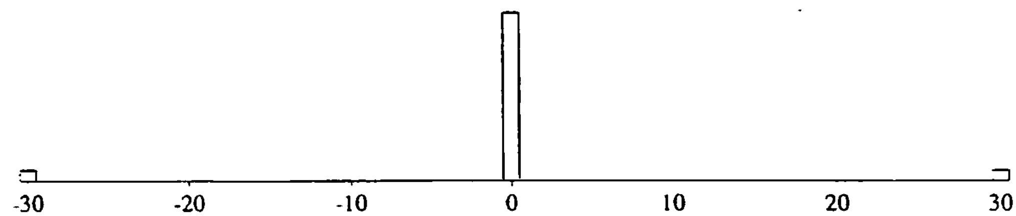
| $x$ | $-20$ | $0$ | $20$ |
|---|---|---|---|
| $P(X = x)$ | 1/8 | 3/4 | 1/8 |

$k = 3$:

| $x$ | $-30$ | $0$ | $30$ |
|---|---|---|---|
| $P(X = x)$ | $1/18$ | $8/9$ | $1/18$ |




b) $E(X) = \mu$ by symmetry of the distribution about $\mu$.
$Var(X) = E[(X-\mu)^2] = 2 \cdot (k\sigma)^2 \cdot \frac{1}{2k^2} + 0^2(1 - \frac{1}{k^2}) = \sigma^2$
$P(|X - \mu| \geq k\sigma) = P(X = \mu + k\sigma) + P(X = \mu - k\sigma) = 2 \cdot \frac{1}{2k^2} = \frac{1}{k^2}$,
which is the Chebychev bound on the above probability.

c) $P(|Y - \mu| < \sigma) = 0 \Longrightarrow P(|Y - \mu| \geq \sigma) = 1$. So absolute deviations from the mean are $\geq \sigma$ with probability 1. On the other hand, the average of squared deviations (variance) $= \sigma^2$. So the deviations must actually equal $\pm\sigma$. Suppose $P(Y - \mu = \sigma) = p$, so $P(Y - \mu = -\sigma) = 1 - p$. Then
$0 = E(Y - \mu) = p\sigma + (1-p)(-\sigma)$,
so $p = 1/2$. That is, $P(Y = \mu + \sigma) = P(Y = \mu - \sigma) = \frac{1}{2}$, which is the above distribution for $k = 1$.

26. a) 1.5 b) Note that
$0 \leq Var(|X - \mu|) = E(|X - \mu|^2) - [E(|X - \mu|)]^2$. From this, the result follows immediately. That the equality holds if and only if $|X - \mu|$ is a constant follows from the first inequality.

27. a) Use the fact that expectation preserves inequalities:
(i) $0 \leq X \leq 1 \Longrightarrow 0 \leq E(X) \leq 1$.
(ii) Note that $X^2 \leq X$, so $E(X^2) \leq E(X)$ and
$0 \leq Var(X) = E(X^2) - [E(X)]^2 \leq E(X) - [E(X)]^2 = \mu - \mu^2$ .
The inequality $\mu - \mu^2 \leq 1/4$ follows from calculus, or from completing the square:
$\mu - \mu^2 = \frac{1}{4} - \left(\mu - \frac{1}{2}\right)^2 \leq \frac{1}{4}$.

b) The results are trivial if $a = b$. So assume $a < b$. Let $Y = \frac{X-a}{b-a}$. Then $0 \le Y \le 1$.

(i) Use part i) of a) to conclude $0 \le E(Y) \le 1$. Substitute $E(Y) = \frac{\mu - a}{b-a}$ and rearrange.

(ii) Use part ii) of a) to conclude

$$0 \le Var(Y) \le E(Y)\,(1 - E(Y)) \le 1/4.$$

Substitute $E(Y) = \frac{\mu-a}{b-a}$ and $Var(Y) = \frac{Var(X)}{(b-a)^2}$ and rearrange.

(iii) Immediate from ii).

c) Consider a random variable $X$ taking values $0, \ldots, 9$ with chances equal to the proportions of each digit among the list. We have $0 \le X \le 9$ and $Var(X) = \frac{1}{4}(9)^2$. By part b), we must have

$$E(X^2) - (E(X))^2 = E(X)\,(9 - E(X)) = \frac{1}{4}(9)^2.$$

The second equality yields $E(X) = 9/2$ and the first yields $9E(X) = E(X^2)$. But $9X - X^2 \ge 0$, so $9X - X^2 = 0$ with probability 1 (see below). Hence $X$ takes values only 0 or 9. Conclude from $E(X) = 9/2$ that $P(X = 0) = P(X = 9) = 1/2$, that is, that half the list are 0's and the rest are 9's.

**Claim:** If $Y \ge 0$ and $E(Y) = 0$ then $Y = 0$ with probability 1.

Reason: Suppose not. Then $P(Y > a) > 0$ for some $a > 0$ . But by Markov's inequality, we have $P(Y > a) \le E(Y)/a = 0$, contradiction!

28. $Var(S) = \sum_i p_i(1 - p_i) = np(1-p) - \sum_i (p_i - p)^2$.

29. $\bar{D}_n = (D_1 + \cdots + D_n)/n$, where each $D_i$ is uniformly distributed on $\{0, \ldots, 9\}$: $P(D_i = k) = 1/10$ for each $k$.
So $E(D_i) = \frac{1}{10}(1 + 2 + \cdots + 9) = 9/2$.
To get the variance of $D_i$, note that $D_i + 1$ is uniform on $\{1, 2, \ldots, 10\}$. So by a previous problem (Moments of the Uniform Distribution), $Var(D_i) = Var(D_i + 1) = \frac{10^2 - 1^2}{12} = 33/4$ and $SD(D_i) = \sqrt{33}/2$.

a) So guess $int\ (9/2) = 4$.

b) If $n = 1$, then the chance of correct guessing is $P(D_1 = 4) = 1/10$. The chance of being correct would be the same no matter what we guessed.
If $n = 2$, the chance of correct guessing is

$$P\,(4 \le (D_1 + D_2)/2 < 5) = P(D_1 + D_2 = 8 \text{ or } 9) = P(D_1 + D_2 = 8) + P(D_1 + D_2 = 9) = \frac{9}{100} + \frac{10}{100} = \frac{1}{1}$$

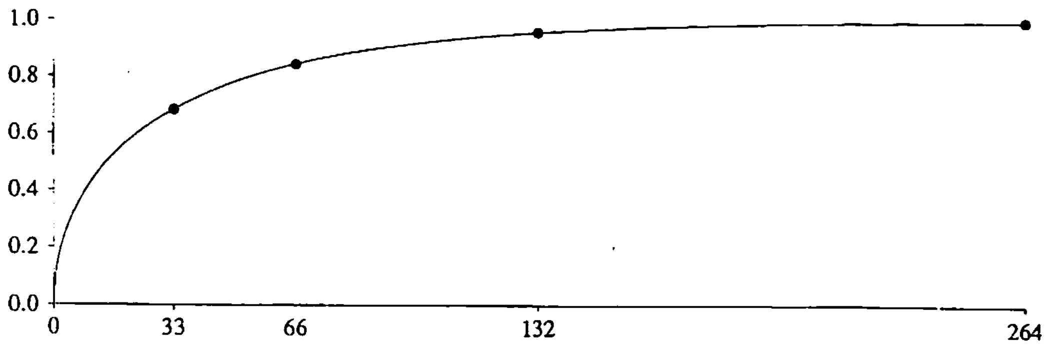
Note that the distribution of $D_1 + D_2$ is triangular with peak at $(D_1 + D_2 = 9)$, so guessing $\bar{D}_2 = 4$ indeed maximizes the chance of guessing correctly.

For large $n$, the standardized variable $Z_n = \frac{2\sqrt{n}(\bar{D}_n - 9/2)}{\sqrt{33}}$ has approximately standard normal distribution. So the chance of correct guessing is

$$P\left(4 \le \bar{D}_n < 5\right) \approx P\left(\left|\frac{2\sqrt{n}(\bar{D}_n - 9/2)}{\sqrt{33}}\right| \le \sqrt{\frac{n}{33}}\right) = P\left(|Z_n| \le \sqrt{\frac{n}{33}}\right).$$

The normal approximation now gives:
$n = 33$: $P(4 \le \bar{D}_n < 5) \approx P(|Z| \le 1) = 0.6826$
$n = 66$: $P(4 \le \bar{D}_n < 5) \approx P(|Z| \le \sqrt{2}) = 0.8414$

$n = 132$: $P(4 \le \bar{D}_n < 5) \approx P(|Z| \le 2) = 0.9544$




c) $P(4 \le \bar{D}_n \le 5) \ge 0.99 \iff P(|Z_n| \le \sqrt{n/33}) \ge 0.99 \iff \sqrt{n/33} \ge 2.58 \iff n \ge 220.$

**30.** $D_i^2$ takes the values 0, 1, 4, 9, 16, 25, 36, 49, 64, 81 with equal probability, so the $X_i$ are independent with common distribution

| $x$ | 0 | 1 | 4 | 5 | 6 | 9 |
|---|---|---|---|---|---|---|
| $P(X_i = x)$ | .1 | .2 | .2 | .1 | .2 | .2 |

So $E(X_i) = 9/2$ and $Var(X_i) = 9.05 \Longrightarrow SD(X_i) = 3.008$.

a) By the law of averages, you expect $\bar{X}_n$ to be close to $E(\bar{X}_n) = 4.5$ for large $n$. So predict 4.5.

b) $P(|\bar{X}_n - 4.5| > \epsilon)$

$= P\left(\frac{\sqrt{n}|\bar{X}_n - 4.5|}{3.008} > \frac{\sqrt{n}\epsilon}{3.008}\right) \approx P\left(|Z| > \frac{\sqrt{n}\epsilon}{3.008}\right) = 2P\left(Z > \frac{\sqrt{n}\epsilon}{3.008}\right),$

where $Z$ has standard normal distribution. For $n = 10000$, we need $\epsilon$ such that

$P\left[Z > \frac{\sqrt{n}\epsilon}{3.008}\right] = \frac{1}{400} = 0.0025 \Longrightarrow \frac{\sqrt{n}\epsilon}{3.008} = 2.81,$

therefore $\epsilon = 2.81 \times 3.008/100 = 0.085$.

c) Need $n$ such that $P(|\bar{X}_n - 4.5| \le 0.01) \ge 0.99$ i.e.,

$P\left[|Z| \le \frac{\sqrt{n} \times 0.01}{3.008}\right] \ge 0.99 \Longrightarrow \frac{\sqrt{n}}{300.8} \ge 2.58$

therefore $n \ge 602276$.

d) We have calculated $E(X_i) = 9/2$ and $Var(X_i) = 9.05$. From the previous problem, we have $E(D_i) = 9/2$ and $Var(D_i) = 33/4 = 8.25$. Since $D_i$ has smaller variance than does $X_i$, the value of $\bar{D}_n$ can be predicted more accurately.

e) Since $E(\bar{X}_{100}) = 4.5$, you should predict the first digit of $\bar{X}_{100}$ to be 4. The chance of being correct is

$P(4 \le \bar{X}_{100} < 5) \approx P\left(\left|\frac{\sqrt{100}(\bar{X}_{100} - 4.5)}{3.008}\right| \le \frac{\sqrt{100}}{2 \times 3.008}\right) \approx P(|Z| \le 1.66) = 0.903.$

**31.** a) $9/2, \sqrt{33}/2$

d)

$$P(|S_n - 4\tfrac{1}{2}n| \le b\sqrt{n}) = P\left(\frac{|S_n - 4\frac{1}{2}n|}{\frac{\sqrt{33}}{2}\sqrt{n}} \le \frac{2b}{\sqrt{33}}\right)$$
$$\approx 2\Phi(2b/\sqrt{33}) - 1$$

**32.** No Solution

**33.** No Solution

# Section 3.4

1. a) Binomial probability: $\binom{9}{5}p^5(1-p)^4$

   b) $P(\text{first 6 tosses are tails, seventh is head}) = (1-p)^6 \cdot p$

   c) $P(\text{exactly 4 heads among first 11 tosses and 12th toss is heads}) = \binom{11}{4}p^4(1-p)^7 \cdot p$

   d) $\sum_{k=0}^{5} P(k \text{ heads among first 8 tosses and } k \text{ among next 5})$
   $= \sum_{k=0}^{5} \binom{8}{k}p^k(1-p)^{8-k} \cdot \binom{5}{k}p^k(1-p)^{5-k}$

2. a) $D$ is distributed as 1+ a geometric random variable with parameter $\frac{1}{2}$. Whatever the first ball is, we then wait until we draw a ball of the other color.

   b) Since the expected value of a geometric random variable is $\frac{1}{p}$, $E(D) = 1 + \frac{1}{p} = 3$.

   c) The SD of a geometric random variable is $\frac{\sqrt{q}}{p}$ so in this case we have $SD(D) = \sqrt{2}$.

3. Assuming $X$ has geometric distribution on $\{1, 2, \ldots\}$ with $p = 1/12$, which must be approximately correct, $E(X) = 12$.

4. a) The probability of some person being the "odd one out" is 1− the probability of having the three coins be HHH or TTT. Thus the probability is $1 - \left((\frac{1}{2})^3 + (\frac{1}{2})^3\right) = \frac{3}{4}$.

   b) Let the length of play be the random variable $X$, then for $r = 1, 2, \ldots$

   $$P(X = r) = \left(\frac{1}{4}\right)^{r-1}\frac{3}{4}$$

   c) Since $X$ has geometric(3/4) distribution, and the geometric ($p$) distribution has mean $1/p$, $E(X) = 4/3$.

5. a) Let $q_i = 1 - p_i$. The probability that Mary takes more than $n$ tosses to get a head is the same as the probability that Mary gets tails the first $n$ times, or $q_2{}^n$.

   b) The probability that the first person to get a head tosses more than $n$ times is the same as the probability that everyone gets tails for the first $n$ times, or $(q_1q_2q_3)^n$

   c) In order for the first person to get a head to toss exactly $n$ times, there must be no heads in the first $n-1$ tosses and then at least one head on the $n$th toss. Thus the probability is $(q_1q_2q_3)^{n-1}(1 - q_1q_2q_3) = (q_1q_2q_3)^{n-1} - (q_1q_2q_3)^n$.

   d) Condition on the first head occuring at time $n$. Now in order for neither Bill nor Tom to get a head before Mary, Mary must get a head at time $n$; it does not matter what Bill and Tom get. Thus we want the probability that Mary gets a head given at least one head. Let $M$ be the event that Mary gets a head and let $H$ be the event that at least one head occurs. Then

   $$P(M|H) = \frac{P(M \text{ and } H)}{P(H)} = \frac{p_2}{(1 - q_1q_2q_3)}$$

6. a) $P(W = k) = P(T = k+1) = q^{(k+1)-1}p = q^k p$.

   b) $P(W > k) = P(T > k+1) = q^{k+1}$.

   c) $E(W) = E(T-1) = E(T) - 1 = \frac{1}{p} - 1 = q/p$.

   d) $Var(W) = Var(T-1) = Var(T) = q/p^2$.

7. a) Use the craps principle. Imagine the following game: A tosses the biased coin once, then B tosses it once. If A's toss lands heads, then say that A wins the game; if A's toss does not land heads but B's does, then say that B wins the game; otherwise the game ends in a draw. The chance that A wins this game is $p$; the chance that B wins this game is $qp$; and the chance of a draw is $q^2$.

You can see that A and B are really repeating this game independently, over and over, until either A wins or B wins, and that the desired probability is $P(\text{A wins before B does})$. By the craps principle, this is $\frac{p}{p+qp} = \frac{1}{1+q}$.

Another way: In terms of $T$, the number of rolls required to produce the first head, the desired probability is

$P(T \text{ is odd}) = P(T=1) + P(T=3) + P(T=5) + \ldots = p + q^2p + q^4p + \ldots = \frac{p}{1-q^2}$.

b) Apply the craps principle to the situation in a): the desired probability is $P(\text{B wins before A does}) = \frac{qp}{p+qp} = \frac{q}{1+q}$. Or compute $P(T \text{ is even}) = \frac{qp}{1-q^2}$. Or subtract the answer in a) from 1, because one of the players must see a head sometime.

c) Apply the craps principle, this time with the game consisting of A tossing the coin once, then B tossing it twice. The chance that A wins this game is $p$, while the chance that B wins is $q - q^3$. So the chance that A gets the first head is $p/(1-q^3)$, and the chance that B gets the first head is $(q-q^3)/(1-q^3)$. Note that $1 - q^3 = (1-q)(1+q+q^2)$, so $P(\text{A gets the first head})$ simplifies to $1/(1+q+q^2)$. Alternatively, in terms of the random variable $T$, the chance that A gets the first head is the probability that $T$ is of the form $1+3k$ for some $k = 0, 1, 2, \ldots$, similarly for B.

d) Solve $\frac{1}{1+q+q^2} = .5$ to get $q = \frac{\sqrt{5}-1}{2}$ and $p = \frac{3-\sqrt{5}}{2} = .381966$.

e) If $p$ is very small, then the fact that A gets to toss first doesn't confer much of an advantage to A. However, since B tosses twice as often as A does, you would expect that the chance that B gets the first head is close to 2/3. Indeed, as $q$ tends to 1:

$P(\text{A gets the first head}) = \frac{1}{1+q+q^2} \to \frac{1}{3}$,

$P(\text{B gets the first head}) \to \frac{2}{3}$.

8. a,b) Suppose that the player's point $X_0$ is $x = 4, 5, 6, 8, 9$, or $10$. Then $P(\text{win}|X_0 = x)$ is the probability that in repeated throws of a pair of dice, the sum $x$ appears before the sum 7 does. In the notation of the Example 2, imagine that A and B are repeating, independently, a competition consisting of the throw of a pair of dice. On each throw, if $x$ appears, then A wins; if 7 appears, then B wins; otherwise a draw results. The desired probability is the chance that A wins before B does, which, by the craps principle, is $\frac{P(x)}{P(x)+P(7)}$.

| $x$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(x)$ | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |
| $P(\text{win} \mid X_0 = x)$ | 0 | 0 | 3/9 | 4/10 | 5/11 | 1 | 5/11 | 4/10 | 3/9 | 1 | 0 |

c) $P(\text{win}) = \sum_{x=2}^{12} P(\text{win}|X_0 = x)P(X_0 = x)$
$= 0 \cdot \frac{4}{36} + \frac{3}{9} \cdot \frac{3}{36} \cdot 2 + \frac{4}{10} \cdot \frac{4}{36} \cdot 2 + \frac{5}{11} \cdot \frac{5}{36} \cdot 2 + 1 \cdot \frac{6}{36} + 1 \cdot \frac{2}{36}$
$= \frac{0+220+352+500+660+220}{36 \times 10 \times 11} = \frac{1952}{36 \times 10 \times 11}$

9. Let $X$ be the payoff. Then

$P(X = n^2) = (\frac{1}{2})^n$ for $n = 1, 2, \ldots$,.

That is, $X = W^2$, where $W$ has geometric $(p)$ distribution on $\{1, 2, \ldots\}$ for $p = 1/2$. Since $W$ has mean $1/p = 2$ and variance $q/p^2 = 2$,

$E(X) = E(W^2) = Var(W) + [E(W)]^2 = 6$.

Your net gain is given by $X - 10$, so $E(\text{net gain}) = -4$. That is, you are going to lose \$4 per game in the long run.

10. a) Condition on the first outcome and use the rule of average conditional probabilities: Let $S$ = (first trial results in success), $F$ = (first trial results in failure). Then

$P(X = n) = P(X = n|S)P(S) + P(X = n|F)P(F)$

Note that

$P(X = n|S) = P(W_F = n-1) = p^{n-2}q,\ n \geq 2$, where $W_F$ denotes the number of tosses until the first failure, which has a geometric distribution on $\{1, 2, ...\}$ with parameter $q$. Similarly,

$P(X = n|F) = P(W_S = n-1) = q^{n-2}p,\ n \geq 2,$

Therefore

$P(X = n) = p^{n-1}q + q^{n-1}p$ for $n = 2, 3, ....$

b) Duplicate the method for finding the moments of the geometric distribution: $E(X) = \sum_{n=2}^{\infty} np^{n-1}q + \sum_{n=2}^{\infty} nq^{n-1}p$

$= q\left(\sum(1,p) - 1\right) + p\left(\sum(1,q) - 1\right)$

$= q\left(\frac{1}{(1-p)^2} - 1\right) + p\left(\frac{1}{(1-q)^2} - 1\right)$

$= \frac{1}{q} + \frac{1}{p} - 1$ , where $\sum(1,p) = \sum_{n=1}^{\infty} np^{n-1}$ .

c) Similarly $E(X^2) = \sum_{n=2}^{\infty} n^2p^{n-1}q + \sum_{n=2}^{\infty} n^2q^{n-1}p$

$= q\left(\sum(2,p) - 1\right) + p\left(\sum(2,q) - 1\right)$

$= q\left(\frac{1+p}{(1-p)^3} - 1\right) + p\left(\frac{1+q}{(1-q)^3} - 1\right)$

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

Finally, use $Var(X) = E(X^2) - [E(X)]^2$ .

11. Write $q_A = 1 - p_A,\ q_B = 1 - p_B$.

a) $P(\text{A wins})$

$= P(\text{A tosses H, B tosses T}) + P(\text{A tosses TH, B tosses TT}) + P(\text{A tosses TTH, B tosses TTT}) + \dots$

$= p_Aq_B + (q_Aq_B)p_Aq_B + (q_Aq_B)^2p_Aq_B + \dots$

$= \frac{p_Aq_B}{1-q_Aq_B}$.

b) $P(\text{B wins}) = \frac{q_Ap_B}{1-q_Aq_B}$. (Interchange A and B.)

c) $P(\text{draw}) = 1 - P(\text{A wins}) - P(\text{B wins}) = \frac{p_Ap_B}{1-q_Aq_B}$ .

d) Let $N$ be the number of trials (in which A and B both toss) required until at least one H is seen (by either A or B). $N$ has range $\{1, 2, 3, \dots\}$. For $k = 1, 2, 3, ...$

$$P(N = k) = P(\text{first } k-1 \text{ trials see TT, } k\text{th trial does not see TT})$$
$$= (q_Aq_B)^{k-1}(1 - q_Aq_B).$$

Or: Each trial is a Bernoulli trial with success corresponding to the event (at least one H is seen among the tosses of A and B). The probability of success on each trial is then $1 - P(\text{no H is seen}) = 1 - q_Aq_B$. Therefore $N$, the waiting time until the first success, is geometric with parameter $(1 - q_Aq_B)$.

12. Write $q_1 = 1 - p_1, q_2 = 1 - p_2$.

a) $P(W_1 = W_2) = \sum_{k=1}^{\infty} P(W_1 = k, W_2 = k)$

$$= \sum_{k=1}^{\infty} P(W_1 = k)P(W_2 = k) = \sum_{k=1}^{\infty} q_1^{k-1}p_1q_2^{k-1}p_2 = \frac{p_1p_2}{1 - q_1q_2} = \frac{p_1p_2}{p_1 + p_2 - p_1p_2}$$

b) $P(W_1 < W_2) = \sum_{k=1}^{\infty} P(W_1 = k, W_1 < W_2)$

$= \sum_{k=1}^{\infty} P(W_1 = k, W_2 > k) = \sum_{k=1}^{\infty} q_1^{k-1}p_1q_2^k = \frac{p_1q_2}{1-q_1q_2}$

c) By symmetry, it's $\frac{p_2q_1}{1-q_2q_1}$. Check: (a) + (b) + (c) = 1.

d) Put $X = \min(W_1, W_2)$. For $k = 0, 1, 2, ...$ we have

$$P(X > k) = P(W_1 > k \text{ and } W_2 > k) = P(W_1 > k)P(W_2 > k) = q_1^kq_2^k = (q_1q_2)^k.$$

So $X$ is geometric with parameter $1 - q_1q_2 = p_1 + p_2 - p_1p_2$.

c) Put $Y = \max(W_1, W_2)$. $Y$ has range $\{1, 2, 3, ...\}$. For $n = 0, 1, 2, ...$ we have

$$P(Y \le n) = P(W_1 \le n \text{ and } W_2 \le n) = P(W_1 \le n)P(W_2 \le n)$$
$$= [1 - P(W_1 > n)][1 - P(W_2 > n)] = (1 - q_1^n)(1 - q_2^n).$$

For $n = 1, 2, 3, ...$ we then have

$$P(Y = n) = P(Y \le n) - P(Y \le n - 1) = (1 - q_1^n)(1 - q_2^n) - (1 - q_1^{n-1})(1 - q_2^{n-1}).$$

**13.** a) Let $q = 1 - p$. Condition on the first two trials, and use recursion, as follows: Let $B$ = (first draw is black), and $WB$ = (first draw is white, second is black). Then

$P(\text{Black wins}) = P(\text{Black wins}|B) \times P(B) + P(\text{Black wins}|WB) \times P(WB)$

$= 1 \times p + P(\text{Black wins}) \times qp$

Therefore

$P(\text{Black wins}) = \frac{p}{1-qp}$

$P(\text{White wins}) = 1 - P(\text{Black wins}) = \frac{q^2}{1-qp}$.

Or use the craps principle: You can imagine that Black and White are repeating over and over, independently, a competition which consists of each player drawing *once* at random with replacement from the box. This competition results in a "win" for Black with probability $p$, and a "win" for White with probability $q^2$. The competition is repeated until one player "wins"; the chance that Black "wins" before White does is $p/(p + q^2)$, which is equal to the previously calculated chance.

b) Set $P(\text{Black wins}) = 1/2$. Solving the quadratic equation gives $p = (3 - \sqrt{5})/2 = 0.381966$ . (The other root is greater than 1.)

c) No, $\frac{3-\sqrt{5}}{2}$ is irrational.

d) No player has more than a 51% chance of winning if and only if

$$.49 \le P(\text{Black wins}) \le .51$$

This happens if and only if

$$.375139206... \le p \le .388805725...$$

Reason: Observe that $f(p) = p/[1 - (1 - p)p]$ is an increasing function of $p \in [0, 1]$. If $0 < c < 1$, then the equation $f(p^*) = c$ has one solution, namely $p^* = \left(c + 1 - \sqrt{(c + 1)^2 - 4c^2}\right)/2c$ .

Calculate all possible values of $p$ for each value of $b + w$ . The smallest value of $b + w$ for which $p$ lies in the above range is $b + w = 13$ (with $b = 5$, $p = 5/13 = .384615$).

**14.** Let $n \ge 1$. $V_n$ is a random variable having range $\{n, \dots, 2n - 1\}$. For $k = n, \dots, 2n - 1$ we have

$(V_n = k) = (V_n = k, k\text{th trial is success}) \cup (V_n = k , k\text{th trial is failure})$

The two events on the right are mutually exclusive. The first event is

$(V_n = k, k\text{th trial is success}) = (\text{exactly } n - 1 \text{ successes in first } k - 1 \text{ trials}, k\text{th trial is success})$

and has probability

$P(\text{exactly } n - 1 \text{ successes in first } k - 1 \text{ rials}, k\text{th trial is success})$

$= P(\text{exactly } n - 1 \text{ successes in first } k - 1 \text{ trials})P(k\text{th trial is success})$

$= \binom{k-1}{n-1}p^{n-1}q^{k-n} \cdot p = \binom{k-1}{n-1}p^n q^{k-n}$ . Similarly the second event has probability $\binom{k-1}{n-1}q^n p^{k-n}$ . Hence

$P(V_n = k) = \binom{k-1}{n-1}(p^n q^{k-n} + q^n p^{k-n})$ , $k = n, \dots, 2n - 1$.

**15.** a) Let $k \ge 0$ and $m \ge 0$. Use problem 1:

$P(F - k = m|F \ge k) = \frac{P(F=m+k)}{P(F\ge k)} = \frac{q^{m+k}p}{q^k} = q^m p = P(F = m)$

b) Let $F$ assume the values $0, 1, 2, \ldots$ with probabilities $p_0, p_1, p_2, \ldots$. We claim that $p_k = (1 - p_0)^k p_0$.

To prove this claim, introduce the tail probabilities

$t_k = P(F \geq k) = p_k + p_{k+1} + p_{k+2} + \ldots$ .

Put $m = 0$ in the Property to see

$P(F = k | F \geq k) = P(F = 0)$ for all $k \geq 0$,

which is equivalent to

$\frac{p_k}{t_k} = p_0$ for all $k \geq 0$.

Use $p_k = t_k - t_{k+1}$ and $1 - p_0 = t_1$ to get

$\frac{t_{k+1}}{t_k} = t_1$ for all $k \geq 0$.

This leads to the solution

$t_k = (t_1)^k$ for all $k \geq 0$

and

$p_k = (t_1)^k p_0 = (1 - p_0)^k p_0$, for all $k \geq 0$,

as claimed.

16. a) If $k \geq 1$ then

$\frac{P(k)}{P(k-1)} = \frac{\binom{k+r-1}{r-1} p^r (1-p)^k}{\binom{k+r-2}{r-1} p^r (1-p)^{k-1}} = \frac{k+r-1}{kq}$.

b,c) If $m$ is a mode, then $m \geq 0$ and

$\frac{P(m+1)}{P(m)} \leq 1$ and $\frac{P(m)}{P(m-1)} \geq 1$.

It follows that

$(r-1)\frac{q}{p} - 1 \leq m \leq (r-1)\frac{q}{p}$.

So if $(r-1)\frac{q}{p}$ is not an integer, then there is a unique mode, namely $int\left((r-1)\frac{q}{p}\right)$. If $(r-1)\frac{q}{p}$ is zero, then there is a unique mode, namely zero. If $(r-1)\frac{q}{p}$ is an integer greater than 0, then there are two modes, namely $(r-1)\frac{q}{p}$ and $(r-1)\frac{q}{p} - 1$.

17. Let $X$ be the number of boys in a family and $Y$ the number of children in a family. Observe that given $Y = n$, $X$ has binomial$(n, p = 1/2)$ distribution; therefore

$$P(X = k) = \sum_{n=k}^{\infty} P(X = k | Y = n) P(Y = n) = \sum_{n=k}^{\infty} \binom{n}{k} (1/2)^n p^n (1-p).$$

Set $j = n - k$ :

$$P(X = k) = \frac{(1-p)p^k}{2^k} \sum_{j=0}^{\infty} \binom{k+j}{k} (p/2)^j = \frac{(1-p)p^k}{2^k} \sum_{j=0}^{\infty} \binom{k+j}{j} (p/2)^j$$

Set $s = k + 1$ and multiply and divide by $\left(1 - \frac{p}{2}\right)^s$:

$$P(X = k) = \frac{(1-p)p^k}{2^k} \left(1 - \frac{p}{2}\right)^{-k-1} \sum_{j=0}^{\infty} \binom{s-1+j}{j} (p/2)^j \left(1 - \frac{p}{2}\right)^s = \frac{2(1-p)p^k}{(2-p)^{k+1}} \sum_{j=0}^{\infty} P(T_s = j)$$

where $T_s$ denotes the number of failures before the $s$th success in Bernoulli$(p/2)$ trials. Thus the last sum equals 1, and

$$P(X = k) = \frac{2(1-p)p^k}{(2-p)^{k+1}}, \quad k \geq 0.$$

18. a) $P(G = n) = \begin{cases} (p^2 + q^2)(2pq)^{\frac{n}{2}-1} & n = 2, 4, \ldots \\ 0 & \text{otherwise} \end{cases}$

b) $G = 2X$ where $X$ has geometric $(p^2+q^2)$ distribution on $\{1, 2, ...\}$. So $E(G) = 2E(X) = \frac{2}{p^2+q^2}$.

c) $Var(G) = 4Var(X) = \frac{8pq}{(p^2+q^2)^2}$.

**19.** a) $r$ plus expectation of negative binomial.

b) Let $S_n$ = number of heads in $n$ tosses.

$$P(T_r < 2r) = P(T_r \le 2r-1)$$
$$= P(S_{2r-1} \ge r) = 1/2$$

by symmetry about $r - 1/2$ of the binomial $(2r-1, 1/2)$ distribution.

c) From b)

$$\frac{1}{2} = P(T_r < 2r) = P(T_r - r < r) = \sum_{i=0}^{r-1} \binom{i+r-1}{r-1} (1/2)^r (1/2)^i,$$

which gives c) with $n = r - 1$.

**20.** a)

$$\sum_{n=1}^{\infty} P(X \ge n) = \sum_{n=1}^{\infty}\sum_{x=n}^{\infty} P(X=x) = \sum_{x=1}^{\infty}\sum_{n=1}^{x} P(X=x) = \sum_{x=1}^{\infty} xP(X=x) = E(X)$$

b)

$$\sum_{n=1}^{\infty} nP(X \ge n) = \sum_{n=1}^{\infty} n \sum_{x=n}^{\infty} P(X=x) = \sum_{x=1}^{\infty} P(X=x) \sum_{n=1}^{x} n = \sum_{x=1}^{\infty} P(X=x)\frac{x(x+1)}{2} = E\left[\frac{X(X+1)}{2}\right]$$

c) $Var(X) = E(X^2) - [E(X)]^2 = 2\Sigma_2 - \Sigma_1 - \Sigma_1^2$

**21.** Let $F_r$ be negative binomial $(r, p)$ and $N$ be Poisson$(\mu)$.

a) $E(F_r) = \frac{rq}{p} \to rq = \mu$ in the limit.

b) $Var(F_r) = \frac{rq}{p^2} \to rq = \mu$ in the limit.

c)

$$\begin{aligned}
P(F_r = k) &= \binom{k+r-1}{r-1} p^r q^k \\
&= \frac{(k+r-1) \times \cdots \times r}{k!} p^r q^k \\
&\approx \frac{p^r}{k!}(rq)^k \\
&= (1-q)^{\frac{\mu}{q}} \frac{\mu^k}{k!} \\
&= \left((1-q)^{\frac{1}{q}}\right)^{\mu} \frac{\mu^k}{k!} \\
&\to e^{-\mu}\frac{\mu^k}{k!} \\
&= P(N = k)
\end{aligned}$$

**22.** No Solution

**23.** No Solution

24. a) Let $N_1 = 1$, and for $i \geq 1$ let $N_{i+1}$ be the additional number of trials (boxes) required t obtain an animal different from all previous. Then $T_n = N_1 + \ldots + N_n$, and $N_i$ is a geometri ındom variable on $\{1, 2, \ldots\}$ with parameter $p_i = (n-i+1)/n$ $(i = 1, \ldots, n)$. The variables $N$ .., $N_n$ are independent, so

$$Var(T_n) = Var(\sum_{i=1}^{n} N_i) = \sum_{i=1}^{n} Var(N_i) = \sum_{i=1}^{n} \frac{q_i}{p_i^2}$$

$$= \sum_{i=1}^{n} \left(\frac{1}{p_i^2} - \frac{1}{p_i}\right) = \sum_{i=1}^{n} \left(\frac{n^2}{(n-i+1)^2} - \frac{n}{n-i+1}\right)$$

$$= \sum_{j=1}^{n} \frac{n^2}{j^2} - \sum_{j=1}^{n} \frac{n}{j} \qquad (j = n-i+1)$$

Hence $\sigma_n = SD(T_n) = \left(n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j}\right)^{1/2}$.

b) Note that $\sigma_n^2 \leq n^2 \sum_{j=1}^{n} \frac{1}{j^2}$ for all $n$. The series $\sum_{j=1}^{\infty} \frac{1}{j^2}$ converges, so there exists a constant $c$ $(0 < c < \infty)$ such that $\sum_{i=1}^{n} \frac{1}{j^2} < c^2$ for all $n$. Hence $\sigma_n \leq nc$ for all $n$.

c) Chebychev's inequality states that $T_n$ is unlikely to be more than a few standard deviations away from its expected value. Here $E(T_n) \approx n \log n$, and $SD(T_n) \leq cn$.

d) As $n \to \infty$,

$$P\left(\frac{T_n - n\log n}{n} \leq x\right) \to e^{-e^{-x}}.$$

See solution of 3.Rev.41 c) for the proof.

# Section 3.5

1. The distribution is approximately binomial (200, .01), which is approximately Poisson (2). So the probability is approximately

$1 - e^{-2}\left[1 + 2 + \frac{2^2}{2} + \frac{2^3}{6}\right] = .1428.$

2. Suppose cookies contain $\lambda$ raisins on average. $P(\text{cookie contains at least one raisin}) = 1 - e^{-\lambda}$, assuming a Poisson distribution for the number of raisins. We want $\lambda$ so that $1 - e^{-\lambda} \geq .99$, or $\lambda \geq -\log(.01) \approx 4.6$. So an average of 5 raisins per cookie will do.

3. a) Use Poisson model. Let $X$ be the number of raisins in a cookie. Then $X$ has Poisson (4) distribution. So $P(X = 0) = e^{-4}$. Let $N$ be the number of raisin-less cookies per bag. Then $N$ has binomial $(12, e^{-4})$ distribution, which is approximately Poisson $(12 \times e^{-4})$. The long run proportion of complaint bags is

   $P(N > 0) \approx 1 - e^{-12e^{-4}} \approx 12e^{-4} \approx 0.222.$

   b) Proceed as in a). Let $a$ be the average required. Want

   $12e^{-\frac{a \times 2}{16}} = 0.05.$

   Take logs to get $a \approx 44$.

4. Assume the number of misprints on each page has the Poisson (1) distribution. Then

   $P(\text{more than 5 misprints per page}) = 1 - \sum_{x=0}^{4} \frac{e^{-1}}{x!} = .0037.$

   Assume the number of misprints on any one page is independent of the number on the other pages. Then in a book of 300 pages, the number of pages having more than 5 misprints has the binomial (300, .0037) distribution, which can be approximated by the Poisson ($300 \times .0037$) distribution. Therefore

   $P(\text{at least one page contains more than 5 misprints}) = 1 - e^{-.0037 \times 300} = 1 - .33 = .67.$

5. Assume that $N$, the number of microbes in the viewing field, is Poisson with parameter $\lambda$. Since the average density in an area of $10^{-4}$ square inches is $5000 \times 10^{-4} = 0.5$ microbes, we have $\lambda = 0.5$ and

   $P(N \geq 1) = 1 - P(N = 0) = 1 - e^{-\lambda} = 1 - e^{-0.5} = 0.39.$

6. Assume the number of drops falling in a given square inch in 1 minute is a Poisson random variable with mean 30. Then the number of drops falling in a given square inch in 10 seconds will be a Poisson random variable with mean 5. So

$$P(0) = e^{-5} = 0.00674$$

7. a) The number of fresh raisins per muffin has Poisson (3) distribution. The number of rotten raisins per muffin has Poisson (2) distribution. The total number of raisins per muffin has Poisson (5) distribution, assuming the number of fresh raisins per muffin is independent of the number of rotten ones.

   b) The number of raisins in 20% of a muffin has Poisson (1) distribution, so

$$P(\text{no raisins}) = e^{-1}\frac{1^0}{0!} = e^{-1} \approx .3679.$$

8. Once again assume that the number of pulses received by the Geiger counter in a given one minute period is a Poisson random variable with mean 10. Then the number received in a given half minute period will be a Poisson random variable with mean 5.

$$P(3) = \frac{e^{-5}5^3}{3!} = 0.1404$$

9. a) $P(X = 1, Y = 2) = P(X = 1)P(Y = 2) = e^{-1}\frac{e^{-2}2^2}{2!} = .09959.$

b) Note that $X+Y$ has Poisson (3) distribution. Therefore
$P((X+Y)/2 \geq 1) = P(X+Y > 2) = 1 - P(X+Y \leq 1) = 1-(e^{-3}+e^{-3}3) = .8008$

c) $P(X=1|X+Y=4) = \binom{4}{1}\frac{1}{3}(\frac{2}{3})^3 = .3951.$

10. a) $E(3X+5) = 3 \times \lambda + 5.$

b) $Var(3X+5) = 9 \times \lambda.$

c) $E\left[\frac{1}{1+X}\right] = \sum_{x=0}^{\infty} \frac{1}{1+x}\frac{e^{-\lambda}\lambda^x}{x!}$
$= \frac{1}{\lambda}\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!} = \frac{1}{\lambda}\left[-e^{-\lambda}+\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}\right] = \frac{1}{\lambda}(1-e^{-\lambda}).$

11. a) $X+Y$ is Poisson with mean 2, so $P(X+Y=4) = e^{-2}2^4/4!$

b) Again, $X+Y$ is Poisson with mean 2, so

$$E\left[(X+Y)^2\right] = Var(X+Y) + (E(X+Y))^2 = \mu + \mu^2 = 6$$

c) $X+Y+Z$ is Poisson with mean 3, so $P(X+Y+Z=4) = e^{-3}3^4/4!$

12. The total number of particles reaching the counter is the sum of two independent Poisson random variables, one with parameter 3.87, the other with parameter 5.41 (these numbers come from a famous experiment by Rutherford, Chadwick, and Ellis, in the 1920's). So the total number of particles reaching the counter follows the Poisson (9.28) distribution. So the required probability is

$e^{-9.28}\left[1+9.28+\frac{(9.28)^2}{2}+\frac{(9.28)^3}{6}+\frac{(9.28)^4}{24}\right] = 0.04622.$

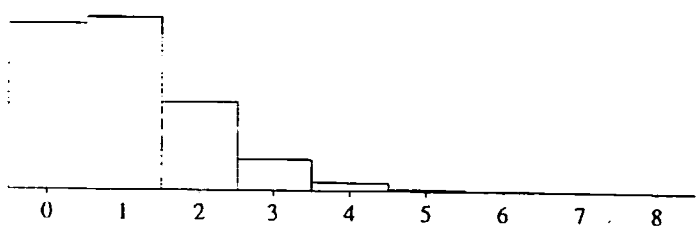
13. a) $\mu(x) = \frac{6.023\times10^{23}}{22.4\times10^3} \times x^3 \approx 2.69\times10^{19}x^3$

$\sigma(x) = \sqrt{\mu(x)} \approx 5.19\times10^9\cdot x^{3/2}$

b) $\sigma(x) = \frac{\mu(x)}{100} \Longrightarrow x = 7.19\times10^{-6}$ cm.

14. a) Using the random scatter theorem, the number of tumors a single person gets in a week has approximately Poisson ($\lambda = 10^{-5}$) distribution. Since different people may be regarded as independent, and different weeks on the same person are also independent, the total number of tumors observed in the population over a year is a sum of $52\times2000$ independent Poisson($10^{-5}$) random variables, which is a Poisson(1.04) random variable.

b) The number of tumors observed on a given person in a year will be a Poisson random variable with rate parameter $\lambda = 52\times10^{-5}$. Thus
$P(\text{at least one tumor}) = 1 - P(\text{none}) = 1 - e^{-0.00052} = 0.00051986$
Thus the number of people getting 1 or more tumors has distribution
binomial $(2000, 0.00051986) \approx$ Poisson $(2000\times0.00051986) \approx$ Poisson $(1.04)$ .




The histograms are essentially the same: Poisson with parameter 1.04. The means and standard deviations are 1.04 and $\sqrt{1.04} \approx 1.02$ respectively.

Why are the answers to a) and b) nearly the same when they are counting different things? The reason is that the chance of a person getting 2 or more tumors is insignificant, even when compared with the chance of getting 1 tumor.

15. (a) The chance that a given pages has no mistakes is $\frac{e^{-.01}(.01)^0}{0!} = e^{-.01} = .99005$ and thus the expected number of pages with no mistakes is $200 \times .99005 = 198.01$. The number of pages with mistakes is distributed as a binomial(200, .99005), so the variance is $200 \times .99005 \times .00995 = 1.97$.

(b) The mistakes found on a given page is distributed as a Poisson(.009), so the chance that at least one mistake will be found on a given page is $1 - \frac{e^{-.009}(.009)^0}{0!} = e^{-.009} = .00896$. The expected number of pages on which at least one mistake is found is then $200 \times .00896 = 1.79$.

(c) The number of pages with mistakes can be well approximated as a Poisson(1.99) since 1.99 is the expected number of pages with mistakes. Let $X$ = number of pages with mistakes and use the Poisson approximation to get

$$P(X \geq 2) = 1 - (P(X = 0) + P(X = 1)) = 1 - (e^{-1.99} + 1.99e^{-1.99}) = 0.59$$

16. a) Assume that chocolate chips are distributed in cookies according to a Poisson scatter. Let $X$ be the number of chocolate chips in a three cubic inch cookie. Then $X$ has Poisson (6) distribution, so

$P(X \leq 4) = \sum_{i=0}^{4} \frac{e^{-6}6^i}{i!} = .2851$.

b) Let $Z_1$, $Z_2$, and $Z_3$ denote the total number of goodies (either chocolate chips or marshmallows) in cookies 1, 2, and 3 respectively. Assume that marshmallows are distributed in cookies according to a Poisson scatter. By the independence assumption between marshmallows and chocolate chips, $Z_1$ has Poisson (6) distribution, $Z_2$ and $Z_3$ each have Poisson (9) distribution, and the $Z_i$'s are independent. We have

$P(Z_1 = 0) = e^{-6}$, $P(Z_2 = 0) = P(Z_3 = 0) = e^{-9}$.

The complement of the desired event has probability

$P(Z_1 = Z_2 = Z_3 = 0) + P(Z_1 > 0, Z_2 = Z_3 = 0) + P(Z_2 > 0, Z_1 = Z_3 = 0) + P(Z_3 > 0, Z_1 = Z_2 = 0)$

$= e^{-6}(e^{-9})^2 + (1 - e^{-6})(e^{-9})^2 + 2(1 - e^{-9})e^{-6}e^{-9} = 6.27 \times 10^{-7}$

and the desired event has probability virtually 1.

17. Assuming the distribution of raindrops over a particular square inch during a given ten second period is a Poisson random scatter, then the number of drops hitting this square inch during the ten second period has Poisson($\lambda = 5$) distribution.

a) So $P(\text{no hit}) = e^{-5} = 0.006738$ .

b) Argue that if $N_1$ denotes the number of big drops and $N_2$ the number of small, then $N_1$ and $N_2$ are independent and $N_1$ has Poisson($\frac{2}{3} \times 5$) distribution, $N_2$ has Poisson($\frac{1}{3} \times 5$) distribution. Hence

$$P(N_1 = 4, N_2 = 5) = P(N_1 = 4)P(N_2 = 5) = e^{-10/3}\frac{(10/3)^4}{4!}e^{-5/3}\frac{(5/3)^5}{5!} = 0.003714.$$

18. a) Let $S_n$ denote the number of survivors between time $n$ and $n+1$, and let $I_n$ denote the number of immigrants between time $n$ and $n+1$. Then $X_{n+1} = S_n + I_n$ where $I_n$ has Poisson ($\mu$) distribution, independent of $S_n$, and

$$P(S_n = k | X_n = x) = \binom{x}{k}(1-p)^k p^{x-k}, 0 \leq k \leq x.$$

**Claim:** $X_n$ has Poisson $\left(\mu \sum_{k=0}^{n}(1-p)^k\right)$ distribution, $n = 0, 1, 2 \ldots$.

**Proof.** Induction. True for $n = 0$. If the claim holds for $n = m$, where $m \geq 0$, then, putting $\lambda = \mu \sum_{k=0}^{m}(1-p)^k$, we have for all $k \geq 0$

$$P(S_m = k) = \sum_{x=0}^{\infty} P(S_m = k | X_m = x)P(X_m = x)$$

$$= \sum_{x=k}^{\infty} \binom{x}{k}(1-p)^k p^{x-k} e^{-\lambda}\frac{\lambda^x}{x!}$$

$$= \frac{e^{-\lambda}[\lambda(1-p)]^k}{k!} \sum_{x=k}^{\infty} \frac{(\lambda p)^{x-k}}{(x-k)!} = e^{-\lambda(1-p)} \frac{[\lambda(1-p)]^k}{k!}$$

so $S_m$ has Poisson distribution with parameter $\lambda(1-p) = \mu \sum_0^m (1-p)^{k+1}$. Since $I$ ; inde- pendent of $S_m$ and has Poisson ($\mu$) distribution, it follows that $X_{m+1} = S_m + I_m$ h 'oisson distribution with parameter $\mu + \mu \sum_{k=0}^{m} (1-p)^{k+1} = \mu \sum_{k=0}^{m+1} (1-p)^k$. So clai: ,lds for $n = m+1$.

b) As $n \to \infty$, $\mu \sum_0^n (1-p)^k \to \frac{\mu}{p}$, so the distribution tends to the Poisson ($\frac{\mu}{p}$) dist: .ion.

19. a) $G(z) = \sum_{i=0}^{\infty} p_i z^i = \sum_{i=0}^{\infty} e^{-\mu} \frac{\mu^i}{i!} z^i = e^{-\mu} e^{\mu z} = e^{-\mu(1-z)}$

b) $G'(z) = \mu e^{-\mu+\mu z}$, $G''(z) = \mu^2 e^{-\mu+\mu z}$, $G'''(z) = \mu^3 e^{-\mu+\mu z}$.
So $E(X) = \mu$, $E(X(X-1)) = \mu^2$, $E(X(X-1)(X-2)) = \mu^3$

c) $E(X^2) = \mu^2 + \mu$, $E(X^3) = \mu^3 + 3(\mu^2 + \mu) - 2\mu = \mu^3 + 3\mu^2 + \mu$

d) $E[(X-\mu)^3] = E(X^3) - 3E(X^2)\mu + 3E(X)\mu^2 - \mu^3 = \mu$.
Skewness $(X) = \frac{\mu}{\mu^{3/2}} = \frac{1}{\sqrt{\mu}}$

20. No Solution

21. c) 0.58304 d) 0.5628 e) 0.58306

# Section 3.6

1. a) $1/13$
   b) $4/50$
   c) $4 \times \frac{\binom{13}{5}}{\binom{52}{5}}$
   d) $1 - \frac{\binom{4}{0}\binom{48}{5}}{\binom{52}{5}} - \frac{\binom{4}{1}\binom{48}{4}}{\binom{52}{5}}$

2. a) $1/13$ b) $1/4$ c) $(13 \times 12 \times 11 \times 10 \times 9)/(52 \times 51 \times 50 \times 49 \times 48)$ d) $(48 \times 47 \times 46 \times 45 \times 4)/(52 \times 51 \times 50 \times 49 \times 48)$

3. a) $8/47$ b) $(12 \times 11 \times 10 \times 9 \times 8)/(51 \times 50 \times 49 \times 48 \times 47)$ c) $1/4$ d) $1/13$ e) $1/13$ f) $1/4$

4. a)

| $n$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P(T_1 = n)$ | 0.4 | 0.3 | 0.2 | 0.1 |

b)

| $n$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P(6 - T_2 = n)$ | 0.4 | 0.3 | 0.2 | 0.1 |

c)

| $n$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| $P(T_2 = n)$ | 0.1 | 0.2 | 0.3 | 0.4 |

d)

| | | $T_1$ | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| $T_2$ | 2 | 0.1 | 0 | 0 | 0 |
| | 3 | 0.1 | 0.1 | 0 | 0 |
| | 4 | 0.1 | 0.1 | 0.1 | 0 |
| | 5 | 0.1 | 0.1 | 0.1 | 0.1 |

e) $T_1 = 1+$ number of non-defectives before first defective, $T_2 - T_1 = 1+$ number of non-defectives between first and second defective, and $6 - T_2 = 1+$ number of non-defectives after second defective. $P(T_1 = n_1, T_2 - T_1 = n_2, 6 - T_2 = n_3) = .1$ where it is not zero, which is a symmetric function, so the random variables are exchangeable.

f) By e), $T_2 - T_1$ has the same distribution as $T_1$, so see a).

5. $P(\text{one fixed box empty}) = \left(\frac{b-1}{b}\right)^n$ so $EX = b\left(\frac{b-1}{b}\right)^n$

$$EX^2 = b\left(\frac{b-1}{b}\right)^n + b(b-1)\left(\frac{b-2}{b}\right)^n$$

So

$$\begin{aligned} Var(X) &= EX^2 - (EX)^2 \\ &= b\left(\frac{b-1}{b}\right)^n + b(b-1)\left(\frac{b-2}{b}\right)^n - b^2\left(\frac{b-1}{b}\right)^{2n} \end{aligned}$$

6. a) Consider $M = \sum_{i=1}^n I_i$ where $I_i$ is the indicator of the $i$th ball being in the $i$th box. Thus $E(M) = nE(I_i) = 1$.

b)

$$\begin{aligned} E(M^2) &= E\left(\left(\sum_{i=1}^n I_i\right)^2\right) \\ &= E\left(\sum_{i=1}^n I_i^2\right) + 2E\left(\sum_{i<j} I_i I_j\right) \end{aligned}$$

$$= 1 + 2\binom{n}{2}\frac{1}{n}\frac{1}{n-1}$$
$$= 2$$

Thus $SD(M) = \sqrt{2-1} = 1$

c) For large $n$, the distribution of $M$ is approximately Poisson(1). Intuitively, the distribution is very much like a binomial($n, \frac{1}{n}$) except for the dependence between the draws, but as the number of draws gets large the dependence between draws becomes small, and the Poisson(1) becomes a good approximation.

7. a) $n \cdot \frac{26}{52}$

b) $\left(\frac{52-n}{52-1}\right) \cdot n \cdot \frac{26}{52} \cdot \frac{26}{52}$

8. a) $P(X = k) = \frac{\binom{26}{k}\binom{26}{26-k}}{\binom{52}{26}}, \qquad k = 0, 1, 2, \ldots, 26.$

b) $E(X) = 26 \cdot \frac{1}{2} = 13.$

c) $SD(X) = \sqrt{\frac{52-26}{52-1}}\sqrt{26 \times \frac{1}{2} \times \frac{1}{2}} = 1.82.$

d) $P(X \geq 15) = P(X \geq 14.5) \approx 1 - \Phi(.824) = .2061.$

The normal approximation gives a fairly good answer. You can check that

$$P(X = 13) = .21812, P(X = 14) = .18807;$$

therefore the exact answer is, by symmetry,

$$\frac{1 - (.21812 + 2 \times .18807)}{2} = .2029.$$

9. Let $b_1, b_2, \ldots, b_B$ denote the $B$ bad elements in the population, and define

$$I_i = \begin{cases} 1 & \text{if bad element } b_i \text{ appears before the first good element} \\ 0 & \text{otherwise.} \end{cases}$$

Then $X - 1 = \#$ of bad elements before the first good element $= I_1 + I_2 + \cdots + I_B$.

a) $E(X-1) = E(I_1) + E(I_2) + \cdots + E(I_B)$. By symmetry, all the expectations on the right hand side are equal, and equal to $P(b_1$ appears before first good element). This probability equals $1/(G+1)$:

Let $g_1, g_2, \ldots, g_G$ denote the $G$ good elements. Consider the $G+1$ elements $g_1$, $g_2$, ..., $g_G$, $b_1$. We are interested in the position of $b_1$ relative to the $g$'s. We can choose this position in $G+1$ ways, all equally likely. Exactly one of these choices puts $b_1$ before all the $g$'s. So $P(b_1$ appears before first good element$) = 1/(G+1)$.

Conclude:
$$E(X) = E(X-1) + 1 = \frac{B}{G+1} + 1 = \frac{B+G+1}{G+1} = \frac{N+1}{G+1}.$$

b) We have $Var(X) = Var(X-1) = E\left[(X-1)^2\right] - [E(X-1)]^2$ . Now

$$\begin{aligned} E\left[(X-1)^2\right] &= E\left[(I_1 + I_2 + \cdots + I_B)^2\right] \\ &= \sum_{i=1}^{B} E(I_i^2) + \sum\sum\nolimits_{i \neq j} E(I_i I_j) \\ &= B \cdot \frac{1}{G+1} + B(B-1)E(I_1 I_2) \end{aligned}$$

by symmetry and the fact that $I_i{}^2 = I_i$; and

$$E(I_1 I_2) = P(b_1 \text{ and } b_2 \text{ both appear before the first red card}) = \frac{1}{\binom{G+2}{2}}.$$

(Reasoning is similar to above: Consider the positions of $b_1$ and $b_2$ relative to the $g$'s. There are $G+2$ positions to fill, and we can choose the pair of positions for $b_1$ and $b_2$ in $\binom{G+2}{2}$ ways. Only the pair $(1,2)$ gives the event we want.)

So

$$\begin{aligned} Var(X) &= \frac{B}{G+1} + B(B-1)\frac{1}{\binom{G+2}{2}} - \left(\frac{B}{G+1}\right)^2 \\ &= \frac{B}{G+1}\left[1 + \frac{2(B-1)}{G+2} - \frac{B}{G+1}\right] \\ &= \frac{B}{G+1}\left[\frac{G^2+3G+2+2BG+2B-2G-2-BG-2B}{(G+2)(G+1)}\right] \\ &= \frac{B}{G+1}\left(\frac{G^2+G+BG}{(G+2)(G+1)}\right) = \frac{BG(N+1)}{(G+1)^2(G+2)} \end{aligned}$$

and

$$SD(X) = \sqrt{\frac{BG(N+1)}{(G+1)^2(G+2)}}.$$

10. No Solution

11. a) $P(x_1,\ldots,x_n) = 1/\binom{n}{g}$ if $x_1+\cdots+x_n = g$ and 0 otherwise
b) no c) yes

12. a) There are $\binom{N}{n}$ possible ways to draw $n$ numbers from the set $\{1\cdots N\}$. Thus the process of taking a simple unordered random sample from this set gives the chance of any given subset of $n$ numbers as $\frac{n!(N-n)!}{N!}$. If we consider the process of the exhaustive sample, where the $X_i$ 's are the draws in order, the chance of getting any particular ordering is, for example,

$$P(X_1 = B, X_2 = G, X_3 = B, \cdots, X_N = B) = \frac{N-n}{N}\frac{n}{N-1}\frac{N-n-1}{N-2}\cdots\frac{1}{1}$$

where the ordering of the numerators may be different but the product must be equal to $\frac{(N-n)!n!}{N!}$ which is the same as it was for the sample of size $n$ from $\{1\cdots N\}$.

b) This is essentially done in the previous problem;

$$\begin{aligned} P(T_1 = t_1,\ldots,T_n = t_n) &= \frac{(N-n)(N-n-1)\cdots(N-n-t_1+2)(n)\cdots}{N!} \\ &= \frac{n!(N-n)!}{N!} \end{aligned}$$

c) The object here is to count the number of samples have $i-1$ elements less than $t$, $t$, and $n-i$ elements bigger than $t$. This number is $\binom{t-1}{i-1}\binom{N-t}{n-i}$ and the total number of samples is $\binom{N}{n}$, so the chance is $\frac{\binom{t-1}{i-1}\binom{N-t}{n-i}}{\binom{N}{n}}$.

d) Given D, writing out the probability of any given sequence of $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ gives the same formula as that in part b). $P(D) = \frac{N-1}{N}\frac{N-2}{N}\cdots\frac{N-n+1}{N}$ so $P(D) > \left(\frac{N-n}{N}\right)^n \to 1$ as $N \to \infty$ for $n$ fixed.

13. a) uniform over ordered $(n+1)$−tuples of non-negative integers that sum to $N-n$.

c) $\frac{\binom{N-n}{w}}{\binom{N}{w}} \cdot \frac{n}{N-w}$

d) $E(W_i) = (N-n)\frac{1}{n+1}$, $E(T_i) = i((N-n)\frac{1}{n+1} + 1)$. For $N = 52$ and $n = 4$, $E(W_i) = 9.6$, $E(T_1) = 10.6$, $E(T_2) = 21.2$, $E(T_3) = 31.8$, $E(T_4) = 42.4$

e) $P(W_1 + W_2 = t) = P(T_2 = t+2) = \frac{\binom{n}{1}\binom{N-n}{t}}{\binom{N}{t+1}} \cdot \frac{n-1}{N-t-1}$, $0 \le t < N$ (for $t = N$, prob is 0 except in the trivial case $n = 0$ when prob. is 1).

f) Because the $W_i$ are exchangable,
$D_n$ has the same distribution as $N - (W_1 + W_2) - 2$.
So $P(D_n = d) = P(W_1 + W_{n+1} = N - 2 - d)$. Now use e).
$E(D_n) = E(T_n) - E(T_1) - 1 = (n-1)(\frac{N-n}{n+1} + 1) - 1$

14. No Solution

15. a) $W_2$ through $W_n$ are spacing between two aces ($W_1$ and $W_{n+1}$ are not). To get two consecutive aces, at least one of $W_1, W_2, \ldots, W_n$ must be 0. The expression in the problem $= 1 - P($not all $w_i = 0,\ i = 2, \ldots, n)$.

b) Put down $N$ places in a row, and color $t$ of them blue, the rest red. $\binom{N-t}{n}$ is the number of ways to place the good elements avoiding all the blue places. There is a 1-1 correspondence between such choices and ways to make ($W_i \geq t_i$ for each $i$).

c) $t_1 = 0 = t_{n+1}$, $t_2 = t_3 = \ldots = t_n = 1$. So $t = n - 1$.

16. No Solution

# Chapter 3: Review

1. a) $1-(5/6)^{10}$ b) $10/6$ c) 35 d) $\frac{\binom{4}{2}\binom{6}{3}}{\binom{10}{5}} = \frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}}$

   e) $\frac{1}{2}(1 - P(\text{same number of sixes in first five rolls as in second five rolls}))$

   $= \frac{1}{2}\left\{1 - \sum_{k=0}^{5}\left[\binom{5}{k}(1/6)^k(5/6)^{5-k}\right]^2\right\}$

2. a)

$$P(\text{first 6 before tenth roll}) = P(\text{at least one 6 in first 9 rolls}) = 1-(5/6)^9 = .806194$$

   b)

$$P(\text{third 6 on tenth roll}) = P(\text{two sixes on first nine rolls, 6 on tenth roll})$$
$$= \binom{9}{2}(1/6)^2(5/6)^7(1/6) = .046507$$

   c)

$$P(\text{three 6's in first ten rolls} \mid \text{six 6's in first twenty rolls})$$
$$= P(\text{three 6's in first ten rolls, three sixes in last ten rolls})\big/P(\text{six 6's in first twenty rolls})$$
$$= \frac{\binom{10}{3}(1/6)^3(5/6)^7\binom{10}{3}(1/6)^3(5/6)^7}{\binom{20}{6}(1/6)^6(5/6)^{14}} = \frac{\binom{10}{3}^2}{\binom{10}{6}} = .371517.$$

   d) Want the expectation of the sum of six geometric $(1/6)$ random variables, each of which has expectation 6. So answer: 36.

   e) Coupon collector's problem: the required number of rolls is 1 plus a geometric $(5/6)$ plus a geometric $(4/6)$ plus etc. up to a geometric $(1/6)$, so the expectation is

$$1+(6/5)+(6/4)+(6/3)+(6/2)+(6/1) = 14.7.$$

3. $X : \max(D_1, D_2)$ $Y = \min(D_1, D_2)$

$$P(X=x) = P(X \le x) - P(X \le x-1) = \left(\frac{x}{6}\right)^2 - \left(\frac{x-1}{6}\right)^2 = \frac{2x-1}{36} \quad (x = 1, \ldots, 6)$$

$$P(Y=y, X=3) = \begin{cases} \frac{2}{36} & \text{for } y = 1,2 \\ \frac{1}{36} & \text{for } y = 3 \\ 0 & \text{else} \end{cases}$$

$$\text{So } P(Y=y \mid X=3) = \begin{cases} \frac{2}{5} & \text{for } y = 1,2 \\ \frac{1}{5} & \text{for } y = 3 \\ 0 & \text{else} \end{cases}$$

$$P(X=x \mid Y=y) = \begin{cases} \frac{2}{36} & \text{for } 1 \le y < x \le 6 \\ \frac{1}{36} & \text{for } 1 \le y \le x \le 6 \\ 0 & \text{else} \end{cases}$$

   $E(X+Y) = E(D_1 + D_2) = 7$

4. Use the fact that $S$ has the same distribution as $200 - S$.

a) Use the convolution formula:

$$P(S=n)=P(X+Y=n)=\sum_{all\,x}P(X=x)P(Y=n-x).$$

If $0\le n\le 100$, then

$$P(S=n)=\sum_{x=0}^{n}P(X=x)P(Y=n-x)=\sum_{x=0}^{n}\frac{1}{101}\times\frac{1}{101}=\frac{n+1}{101^2}.$$

If $100<n\le 200$, then $0\le 200-n\le 100$ so by the previous case

$$P(S=n)=P(200-S=200-n)=\frac{201-n}{101^2}.$$

b) $P(S\le n)=\sum_{k=0}^{n}P(S=k)$ . If $0\le n\le 100$ then

$$P(S\le n)=\sum_{k=0}^{n}\frac{k+1}{101^2}=\sum_{j=1}^{n+1}\frac{j}{101^2}=\frac{(n+1)(n+2)}{2\cdot 101^2}.$$

The above formula works for $-2\le n\le 100$. If $100\le n\le 200$ then $-1\le 199-n\le 100$ so by the previous case

$$\begin{aligned}P(S\le n)&=1-P(S\ge n+1)=1-P(200-S\le 199-n)\\&=1-P(S\le 199-n)\\&=1-\frac{(199-n+1)(199-n+2)}{2\cdot 101^2}\\&=1-\frac{(200-n)(201-n)}{2\cdot 101^2}.\end{aligned}$$

5. a) Let $G$ = gain on one spin

$$EG=\sum_{i=1}^{6}i\frac{18}{38}\cdot\frac{1}{6}-\sum_{i=1}^{6}i\frac{20}{38}\cdot\frac{1}{6}$$

$$=-\frac{1}{38\cdot 6}\sum_{i=1}^{6}2i=-\frac{6\times 7}{38\times 6}=-0.1842$$

b) $\frac{38}{18}=2.111.$

c) $\frac{1}{\frac{1}{6}\cdot\frac{18}{38}}=12.666$

6. a) Let $Y$ be the number of times the gambler wins in 50 plays. Then $Y$ has binomial (50, 9/19) distribution, and

$P(\text{ahead after 50 plays})=P(Y>25)=\sum_{i=26}^{50}\binom{50}{i}(9/19)^i(10/19)^{50-i}$.

Note that the gambler's capital, in dollars, after 50 plays is $100+10Y+(-10)(50-Y)=20Y-400$. So

$P(\text{not in debt after 50 plays})=P(20Y-400>0)=P(Y>20)=\sum_{i=21}^{50}\binom{50}{i}(9/19)^i(10/19)^{50-i}$.

b) $E(\text{capital})=E(20Y-400)=20E(Y)-400=20(50)(9/19)-400=1400/19$ ;

$Var(\text{capital})=Var(20Y-400)=400Var(Y)=400(50)(9/19)(10/19)=1800000/19$

c) Use the normal approximation to the binomial: $E(Y)=23.7$, and $SD(Y)=3.53$, so

$P(Y>25)\approx 1-\Phi\left(\frac{25.5-23.7}{3.53}\right)=1-\Phi(.51)=.305$ .

$P(Y>20)\approx 1-\Phi\left(\frac{20.5-23.7}{3.53}\right)=1-\Phi(-.91)=.8186$ .

7. a) $P(\ge 4\text{ heads in 5 tosses})=\frac{6}{32}=0.1875$

b) $P(0, 1, \text{or } 2 \text{ heads in 5 tosses}) = \frac{16}{32} = 0.5$

c)

| poss. vals | 0 | 1 | 2 | |
|---|---|---|---|---|
| probs | $\frac{26}{32}$ | $\frac{5}{32}$ | $\frac{1}{32}$ | $EX = 0.21875$ |

8. a) Let $X = D_1 + \cdots + D_b$, where $D_1$ is the number of balls until the first black ball, $D_2$ is the number of balls drawn after the first black until the second black ball, and so on. When drawing with replacement, the $D_i$ are independent geometric($\frac{b}{w+b}$) random variables, so $X$ has the negative binomial($b$, $\frac{b}{w+b}$) distribution on $b, b+1, \ldots$.

$$P(X = k) = \binom{k-1}{b-1}\left(\frac{b}{b+w}\right)^b\left(\frac{w}{b+w}\right)^{k-b} \qquad (k \geq b)$$

b) When drawing without replacement, it is possible to draw the $b$th black ball on any draw from $b$ to $b+w$.

$$\begin{aligned} P(X = k) &= P(\text{b-1 blacks in first k-1 draws}) \\ &\quad \times P(k\text{th draw is black} \mid \text{b-1 blacks in first k-1 draws}) \\ &= \frac{\binom{b}{b-1}\binom{w}{k-b}}{\binom{b+w}{k-1}} \times \frac{1}{b+w-k+1} \\ &= \frac{bw!(k-1)!}{(k-b)!(b+w)!} \qquad (b \leq k \leq b+w) \end{aligned}$$

9. Let $X, Y$ be the numbers rolled from the two doubling cubes, and let $U, V$ be the numbers rolled from two ordinary dice. Then $(\log_2 X, \log_2 Y)$ has the same distribution as $(U, V)$ .

a) $P(XY < 100) = P(\log_2 XY < \log_2 100) = P(U + V < 6.64) = \frac{1}{2}[1 - P(U + V = 7)] = 5/12$.

b) $P(XY < 200) = P(U + V < 7.64) = 5/12 + 1/6 = 7/12$.

c) $E(X) = 21$, so by independence $E(XY) = E(X)E(Y) = 441$.

d) $E(X^2) = 910$, so $Var(XY) = E[(XY)^2] - [E(XY)]^2 = 633619$ and $SD(XY) \approx 796$.

10. a) $\frac{1}{n} \times \left(1 - \frac{1}{n}\right)$ if $i \neq j$.

b) number of matches $= \sum_{j=1}^{n} I(\text{match occurs at place } j)$ $E(\text{number of matches}) = n \times \frac{1}{n} = 1$.

11. Assume $X$ has Poisson(2) distribution.

$P(X < 2) = .406, P(X = 2) = .271, P(X > 2) = .323$.

12. Following the hint,

$$P_1 = p^{s-1} + (1 + p + \ldots + p^{s-2})qP_0 = p^{s-1} + (1 - p^{s-1})P_0$$
$$P_0 = (1 + q + \ldots + q^{f-2})pP_1 = (1 - q^{f-1})P_1,$$

hence

$$P_1 = p^{s-1} + (1 - p^{s-1})(1 - q^{f-1})P_1 = \frac{p^{s-1}}{1 - (1 - p^{s-1})(1 - q^{f-1})} = \frac{p^{s-1}}{p^{s-1} + q^{f-1} - p^{s-1}q^{f-1}}$$

and

$$P_0 = \frac{p^{s-1}(1 - q^{f-1})}{p^{s-1} + q^{f-1} - p^{s-1}q^{f-1}}$$

and finally

$$P(A) = pP_1 + qP_0 = \frac{p^{s-1}(1 - q^f)}{p^{s-1} + q^{f-1} - p^{s-1}q^{f-1}}$$

13. Note that $E(X^2) = VarX + [E(X)]^2 = \sigma^2 + \mu^2$, same for $E(Y)^2$. By independence we have

$$E[(XY)^2] = E(X^2)E(Y^2) = (\sigma^2 + \mu^2)^2$$
$$E(XY) = [E(X)][E(Y)] = \mu^2$$

and so

$$Var(XY) = E[(XY)^2] - [E(XY)]^2 = (\sigma^2 + \mu^2)^2 - (\mu^2)^2 = \sigma^2(\sigma^2 + 2\mu^2).$$

14. a) This is 1 minus the chance that current flows along none of the lines. The chance that current does not flow on any particular line is the chance that at least one of the switches on that line doesn't work. So answer:

$$1-(1-p_1)(1-p_2^2)(1-p_3^3)(1-p_4^4).$$

b) Let $X$ be the number of working switches. Then $X$ is the sum of 10 indicators, corresponding to whether each switch works or doesn't work. All the indicators are independent.

$$E(X)=p_1+2p_2+3p_3+4p_4,$$

$$Var(X)=p_1q_1+2p_2q_2+3p_3q_3+4p_4q_4,$$

and $SD(X)$ is the square root of this.

Alternatively, the number of switches working in line $i$ has the binomial distribution with parameters $i$ and $p_i$, independently of all other lines. $X$ is the sum of these ten binomials. Formulae for $E(X)$ and $Var(X)$ can be read off the binomial formulae.

15. a) Binomial (100, 1/38). b) Poisson (100/38)

c) Negative binomial (3, 1/38) shifted to $\{3,4,...\}$. d) $3\times 38$

16.

| $d$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $P(D=d)$ | .27 | .04 | .12 | .04 | .12 | .09 | .12 | .04 | .12 | .04 |

The distribution of $D$ on $\{1,2,\ldots,9\}$ is symmetric about 5, so $E(D|D>0)=5$. Therefore

$$E(D)=E(D|D=0)P(D=0)+E(D|D>0)P(D>0)=0+5\times(73/100)=3.65.$$

17. $P(\text{one six}\mid\text{all different})=\frac{P(\text{one six and all different})}{P(\text{all different})}$.

a) Numerator $=\binom{5}{N-1}\frac{N!}{6^N}$; denominator $=\binom{6}{N}\frac{N!}{6^N}$, so the required probability is $N/6$.

b) The numerator is

$$\sum_{n=1}^{6}P(\text{one six and all different}\mid N=n)P(N=n)=\sum_{n=1}^{6}\binom{5}{n-1}\frac{n!}{6^n}\frac{1}{6}=1/6;$$

the denominator is

$$\sum_{n=1}^{6}P(\text{all different}\mid N=n)P(N=n)=\sum_{n=1}^{6}\binom{6}{n}\frac{n!}{6^n}\frac{1}{6}=.46245,$$

so the required probability is .3604.

18. a) The return (in cents) from the game is

$$X=I_1+I_2+\cdots+I_{100},$$

where

$$I_i=\begin{cases}1 & \text{if the number on deal } i \text{ is greater than those of all previous cards dealt}\\ 0 & \text{otherwise}\end{cases}$$

That is, $I_i$ is the indicator of the event that a *record value* occurs at deal $i$ (a record is considered to have occured at the first deal), and $X$ simply counts the number of records seen. A counting argument shows that $P(I_i=1)=\frac{1}{i}$, so

$$E(X)=\sum_{i=1}^{100}\frac{1}{i}\approx\log 100+\gamma+\frac{1}{2\times 100}=5.19,$$

where $\gamma=.57721$ (Euler's constant). [This approximation was used in the Collector's Problem of Example 3.4.5]. So 5.19 cents is the fair price to pay in advance.

b) The gain (in cents) from 25 plays is

$$S = X_1 + X_2 + \cdots + X_{25},$$

where the $X_i$'s are independent copies of $X$. The $I_i$'s are independent so that

$$Var(X) = \sum_{i=1}^{100} Var(I_i) = \sum_{i=1}^{100} \frac{1}{i}\left(1 - \frac{1}{i}\right) \approx 5.19 - \frac{\pi^2}{6} = 3.54 \quad \Longrightarrow \quad SD(X) = 1.88.$$

Using the normal approximation, obtain $P(S > 25 \times 10) \approx 1 - \Phi(12.8) \approx 0$.

19.

$Y$ : number of failures before first success, $P(Y \geq y) = q^y$
$X$ : Poisson $(\mu)$ independent of $Y$

$$P(Y \geq X) = \sum_{k=0}^{\infty} q^k e^{-\mu} \frac{\mu^k}{k!} = e^{-\mu} \sum_{k=0}^{\infty} \frac{(q\mu)^k}{k!}$$

$$e^{-\mu} e^{q\mu} = e^{-\mu(1-q)} = 0.6065 \text{ if } p = 1/2,\ \mu = 1$$

20. a) Let $p = 1/2 + x$ (where $x$ could be negative). So $1 - p = 1/2 - x$, and

$$p(1-p) = (1/2 + x)(1/2 - x) = 1/4 - x^2 \leq 1/4$$

since $x^2 \geq 0$.

b) The margin of error in the estimate is $\sqrt{p(1-p)/n}$, where $n$ is the sample size, and $p$ is the proportion of part time employed students. This assumes sampling without replacement. For sampling with replacement the margin of error would be smaller, due to the correction factor. No matter what $p$ is,

$$\sqrt{\frac{p(1-p)}{n}} \leq \sqrt{\frac{1}{4n}}$$

by part a), so we should take the smallest $n$ so that $\sqrt{1/4n} \leq .05$, i.e., $n = 100$.

21. $X$ has negative binomial $(1, p)$ distribution, $Y$ has negative binomial $(2, p)$ distribution, and $X$ and $Y$ are independent. Now suppose a coin having probability $p$ of landing heads is tossed repeatedly and independently. Let $X'$ denote the number of tails observed until the first head is observed; let $Y'$ denote the additional number of failures until the third head. Then $(X, Y)$ has the same distribution as does $(X', Y')$. So $Z$ has the same distribution as $X' + Y'$ = the number of tails seen until the third head: negative binomial distribution on $\{0, 1, \ldots\}$ with parameters $r = 3$ and $p$.

22. Suppose the daily demand $X$ has Poisson distribution with parameter $\lambda > 0$. (Here $\lambda = 100$.) Suppose the newsboy buys $n$ (constant) papers per day, $n \geq 1$. Then $min(X, n)$ papers are sold in a day; the daily profit $\pi_n$ in dollars is

$$\pi_n = \frac{1}{4} \min(X, n) - \frac{1}{10} n;$$

and the long run average profit per day is

$$E[\pi_n] = \frac{1}{4} E[\min(X, n)] - \frac{1}{10} n.$$

**Claim:** For each $n = 1, 2, 3, \ldots$:

$$(1) E[\min(X, n)] = \lambda P(X \leq n - 1) + n P(X \geq n + 1)$$

$$(2) E[\min(X, n)] = \sum_{k=1}^{n} P(X \geq k).$$

Part (2) holds no matter what distribution $X$ has.

Proof: (1)

$$E[\min(X,n)] = \sum_{k=0}^{\infty} \min(k,n)P(X=k)$$
$$= \sum_{k=0}^{n} kP(X=k) + \sum_{k=n+1}^{\infty} nP(X=k)$$
$$= \sum_{k=1}^{n} ke^{-\lambda}\frac{\lambda^k}{k!} + nP(X \ge n+1)$$
$$= \lambda \sum_{k=1}^{n} e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!} + nP(X \ge n+1)$$
$$= \lambda P(X \le n-1) + nP(X \ge n+1).$$

(2) Use the tail sum formula for expectation:

$$E[\min(X,n)] = \sum_{k=1}^{n} P(\min(X,n) \ge k) = \sum_{k=1}^{n} P(X \ge k).$$

a) Say $\lambda = 100$ and the newsboy buys $n = 100$ papers per day. Then his long run average daily profit (in dollars) is

$$E[\pi_{100}] = \frac{1}{4}E[\min(X,100)] - \frac{1}{10}\cdot 100$$
$$= \frac{1}{4}[100P(X \le 99) + 100P(X \ge 101)] - 10 \text{ by Claim (1)}$$
$$= 25[1 - P(X = 100)] - 10$$
$$= 15 - 25P(X = 100) \approx 14$$

since

$$P(X = 100) = e^{-100}\frac{(100)^{100}}{100!} \approx \frac{1}{\sqrt{2\pi \cdot 100}} \approx 0.04 \text{ by Stirling's approximation.}$$

b) If $\lambda = 100$ and the newsboy buys $n$ papers per day, then his long run average profit per day is

$$E[\pi_n] = \frac{1}{4}E[\min(X,n)] - \frac{1}{10}n$$
$$= \frac{1}{4}\sum_{k=1}^{n} P(X \ge k) - \frac{1}{10}n \text{ by Claim (2)}$$
$$= \frac{1}{4}\sum_{k=1}^{n}\left(P(X \ge k) - \frac{4}{10}\right).$$

Since the function $k \to P(X \ge k) - \frac{4}{10}$ is decreasing, is positive at $k = 1$, and is negative as $k \to \infty$, it follows that $E[\pi_n]$ is maximized at $n^* =$ the largest $k$ such that $P(X \ge k) - \frac{4}{10} \ge 0$, i.e., $n^* = 102$. Thus the newsboy should buy 102 papers per day in order to maximize his daily profit - assuming that demand has Poisson(100) distribution.

23. a) The drawing process can be described in another way as follows: think of the box as initially containing $2n$ half toothpicks in $n$ pairs. Then half toothpicks are simply being drawn at random without replacement. The problem is to find the distribution of $H$, the number of halves remaining after the last pair is broken. By the symmetry of sampling without replacement, $H$ has the same distribution as $H'$, where $H'$ is the number of draws preceding (i.e., not including) the first time that the remaining half of some toothpick is drawn. This may be clearer by an analogy. Think of $2n$ cards in a deck, 2 of each of $n$ colors, and imagine dealing cards one by one off the top of the deck. Then $H$ corresponds to the number of cards remaining in the deck after each color has been seen at least once. If the cards had been dealt from the bottom of the deck, this number would have been the number (corresponding to $H'$) of cards dealt preceding

the first card having a color already seen. Since the order in which we deal the cards makes no difference, it follows that $H$ and $H'$ have the same distribution. Thus for $1 \le k \le n$

$$P(H = k) = P(H' = k)$$

$$= P(\text{first } k \text{ colors are different and } (k+1)\text{st color has been seen before})$$

$$= \frac{2n}{2n} \cdot \frac{2n-2}{2n-1} \cdots \frac{2n-2(k-1)}{2n-(k-1)} \cdot \frac{k}{2n-k}$$

$$= \frac{2^k (n)_k k}{(2n)_{k+1}}.$$

Another way to obtain $P(H = k)$ is through $P(H \ge k)$: For each $k = 1, 2, ..., n+1$ we have

$$P(H \ge k) = P(H' \ge k) = P(\text{first } k \text{ colors are all different})$$

$$= \frac{2n}{2n} \cdot \frac{2n-2}{2n-1} \cdots \frac{2n-2(k-1)}{2n-(k-1)} = \frac{2^k (n)_k}{(2n)_k}.$$

b) As $n \to \infty$,

$$P(H \ge k) = 1 \cdot \frac{1-\frac{1}{n}}{1-\frac{1}{2n}} \cdot \frac{1-\frac{2}{n}}{1-\frac{2}{2n}} \cdot \ldots \cdot \frac{1-\frac{k-1}{n}}{1-\frac{k-1}{2n}}$$

$$\approx \exp\left[\left(-\frac{1}{n} + \frac{1}{2n}\right) + \left(-\frac{2}{n} + \frac{2}{2n}\right) + \ldots + \left(-\frac{(k-1)}{n} + \frac{(k-1)}{2n}\right)\right],$$

$$= \exp\left(-\frac{1+2+\ldots+(k-1)}{2n}\right) \approx \exp\left(-\frac{k^2}{4n}\right).$$

Thus, putting $k = r\sqrt{2n}$,

$$P(\frac{H}{\sqrt{2n}} \ge r) = P(H \ge r\sqrt{2n}) \approx \exp(-r^2/2).$$

This limiting distribution of $H/\sqrt{2n}$ is called the Rayleigh distribution (see Section 5.3).

c) This approximation suggests

$$E(H) = \sum_{k=1}^{\infty} P(H \ge k) \sim \sum_{k=1}^{\infty} \exp\left(-\frac{k^2}{4n}\right) = \sqrt{2n} \sum_{k=1}^{\infty} \frac{1}{\sqrt{2n}} e^{-\frac{1}{2}(k/\sqrt{2n})^2}$$

$$\sim \sqrt{2n} \int_0^\infty e^{-\frac{1}{2}x^2} dx = \sqrt{2n}\sqrt{\pi/2}$$

That is to say

$$E(H) \sim \sqrt{\pi n} \text{ as } n \to \infty.$$

d) Expect about $\sqrt{100 \cdot \pi} \approx \sqrt{314} \approx 17$ or so.

24. a) If $P(X = 3, Y = 2, Z = 1) = 1/3$,
$P(X = 2, Y = 1, Z = 3) = 1/3$,
$P(X = 1, Y = 3, Z = 2) = 1/3$,
then $P(X > Y) = P(Y > Z) = P(Z > X) = 2/3$.

b) Let $p = \min\{P(X > Y), P(Y > Z), P(Z > X)\}$. Then
$p \le \frac{1}{3}[P(X > Y) + P(Y > Z) + P(Z > X)]$
$= \frac{1}{3}E[I(X > Y) + I(Y > Z) + I(Z > X)]$
$\le 2/3$, since $I(X > Y) + I(Y > Z) + I(Z > X) \le 2$.

c) Say each voter assigns each candidate a numerical score 1, 2, or 3, 3 for the most preferred candidate, 1 for the least preferred. Pick a voter at random, and let $X$ be the score assigned to $A$, $Y$ the score of $B$, $Z$ the score of $C$ for this voter. Then $P(X > Y) =$"Proportion of voters who prefer A to B" , etc. So a population of 3 voters with preferences as in a) above would make each of these proportions 2/3.

d) Let
$P(X_1 = n, X_2 = n-1, X_3 = n-2, ..., X_n = 1) = 1/n,$
$P(X_1 = n-1, X$ [illegible] $= n-2, X_3 = n-3, ..., X_n = n) = 1/n,$
$\dots$
$P(X_1 = 1, X_2$ [illegible] $X_3 = n-1, ..., X_n = 2) = 1/n.$
Then all the p[illegible]bilities are $(n-1)/n$, and the minimum probability $p$ is bounded above by $(n-1)/n$.

e) We have $P(X > Y) = 1 - (1-p_1)p_2$
$P(Y > Z) = p_2$
$P(Z > X) = 1 - p_1.$
These will all equal $q$ if $1 - p_1 = q$, $p_2 = q$, and $1 - (1-p_1)p_2 = q$, that is, if

$$1 - q^2 = q,$$

the equation of the golden mean. (L. Dubins)

25. a)

| $z$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(Y_1 + Y_2 = z)$ | 9/36 | 12/36 | 10/36 | 4/36 | 1/36 |

b) 10/3

c) One possibility: $f(x) = \begin{cases} 0 & \text{if } x = 1,2,3 \\ 1 & \text{if } x = 4,5 \\ 2 & \text{if } x = 6 \end{cases}$

26. a) We wish to find $n$ such that $(.99)^n = .5$, so $n = \frac{\log .5}{\log .99} \approx 69$.

b) This is essentially a negative binomial with the roles of success and failure reversed. Let $F_4$ be the number of failures (successful honks) before the $4th$ success (failure to honk). Then $E(F_4) = \frac{4(1-.01)}{.01} = 396$, but for our problem we must also add 3 non-honks to get an answer of 399.

27.

$$E(X) = \sum_{x=1}^{5} 100xq^{x-1}p + \sum_{x=1}^{\infty}(500 + 40x)q^{5+x-1}p$$

$$= 100(p + 2qp + 3q^2p + 4q^3p + 5q^4p) + 500\sum_{x=1}^{\infty} q^{5+x-1}p + 40\sum_{x=1}^{\infty} xq^{5+x-1}p$$

$$= 100(p + 2qp + 3q^2p + 4q^3p + 5q^4p) + (500q^5) + (40q^5\frac{1}{p})$$

$$= 186.43 + 156.62$$
$$= 343.047$$

28. a) If $w \geq 1$ and $y \in A$ then

$$\begin{aligned} P(W_1 = w, Y_1 = y) &= P(X_1 \notin A, \dots, X_{w-1} \notin A, X_w = y) \\ &= P(X_1 \notin A) \cdots P(X_{w-1} \notin A)P(X_w = y) \\ &= (1 - P_1(A))^{w-1} P_1(y) \end{aligned}$$

Sum over $w$ to get $P(Y_1 = y) = P_1(y)/P_1(A)$ and sum over $y$ to get $P(W_1 = w) = (1 - P_1(A))^{w-1} P_1(A)$.
Continue in this way to show that for all $k \geq 1$

$P(W_1 = w_1, Y_1 = y_1, \dots, W_k = w_k, Y_k = y_k) = P(W_1 = w_1)P(Y_1 = y_1) \cdots P(W_k = w_k)P(Y_k = y_k).$

(Idea: Express the event on the left in terms of the independent variables $X_1, X_2, \dots$ )

b) By the weak law of large numbers: As $n \to \infty$,

$$\frac{\sum_{i=1}^{n} I(X_i \in AB)}{\sum_{i=1}^{n} I(X_i \in A)} = \frac{\sum_{i=1}^{n} I(X_i \in AB)/n}{\sum_{i=1}^{n} I(X_i \in A)/n} \to \frac{P_1(AB)}{P_1(A)} = P_1(B|A).$$

**29.** c) uniform on $\{0, 1, \ldots, n\}$ d) no, yes e) $\frac{b}{b+w}$ f) $\frac{b+d}{b+w+d}$

**30.** No Solution

**31.** No Solution

**32.** No Solution

**33.** a) If $n \geq k \geq 1$:

$$P(\text{all } n \text{ red in bag}|\text{pick all } k \text{ red}) = \frac{P(\text{pick all } k \text{ red}|\text{all } n \text{ red in bag})P(\text{all } n \text{ red in bag})}{P(\text{pick all } k \text{ red})}$$

$$= \frac{1 \cdot (1/2)^n}{\sum_{j=0}^{n} P(\text{pick all } k \text{ red}|j \text{ red in bag})P(j \text{ red in bag})}$$

$$= \frac{(1/2)^n}{\sum_{j=0}^{n} \frac{j^k}{n^k}\binom{n}{j}(1/2)^n}$$

$$= \frac{n^k}{2^n \sum_{j=0}^{n} j^k \binom{n}{j}(1/2)^n}$$

$$= \frac{n^k}{2^n E(X^k)},$$

where $X$ has binomial $(n, 1/2)$ distribution.

b) If $k = 1$: then $E(X) = n/2$, so

$$P(\text{all } n \text{ red in bag}|\text{pick red}) = \frac{1}{2^{n-1}}.$$

If $k = 2$: then $E(X^2) = Var(X) + [E(X)]^2 = \frac{n}{4} + \frac{n^2}{4}$, so

$$P(\text{all } n \text{ red in bag}|\text{pick both red}) = \frac{1}{2^{n-2}\left(1 + \frac{1}{n}\right)}.$$

c) If the $k$ balls are drawn without replacement from the bag, follow the argument in a), but replace $\frac{j^k}{n^k}$ with $\frac{(j)_k}{(n)_k}$:

$$P(\text{all } n \text{ red in bag}|\text{pick all } k \text{ red}) = \frac{(1/2)^n}{\sum_{j=0}^{n} \frac{(j)_k}{(n)_k}\binom{n}{j}(1/2)^n} = \frac{(n)_k}{2^n E(X)_k}$$

where again $X$ has binomial $(n, 1/2)$ distribution. On the other hand, if the $k$ balls are drawn without replacement from the bag, then the number of red balls seen among the $k$ has binomial $(k, 1/2)$ distribution: it's as if we drew the $k$ balls directly from the very large collection. So

$$P(\text{all } n \text{ red in bag}|\text{pick all } k \text{ red}) = \frac{P(\text{pick all } k \text{ red}|\text{all } n \text{ red in bag})P(\text{all } n \text{ red in bag})}{P(\text{pick all } k \text{ red})} = \frac{1 \cdot (1/2)^n}{(1/2)^k}.$$

d) If $k = 3$: Since

$$(X)_3 = X(X-1)(X-2) = X^3 - 3X^2 + 2X,$$

compute

$$E(X^3) = E(X)_3 + 3E(X^2) - 2E(X) = \frac{(n)_3}{8} + 3\left(\frac{n}{4} + \frac{n^2}{4}\right) - 2\left(\frac{n}{2}\right) = \frac{n^2(n+3)}{8}$$

and the result in a) simplifies to

$$P(\text{all } n \text{ red in bag}|\text{pick all 3 red}) = \frac{1}{2^{n-3}\left(1 + \frac{3}{n}\right)}.$$

e) Let $p$ be the proportion of red balls in the original collection of balls. Repeat the argument in c) to get

$$P(\text{all } n \text{ red in bag}|\text{pick all } k \text{ red}) = \frac{p^n}{\sum_{j=0}^{n} \frac{(j)_k}{(n)_k}\binom{n}{j}p^j(1-p)^{n-j}} = \frac{(n)_k p^n}{E(X)_k}$$

(where $X$ has binomial $(n,p)$ distribution) and

$$P(\text{all } n \text{ red in bag}|\text{pick all } k \text{ red}) = \frac{p^n}{p^k}.$$

Therefore $E(X)_k = (n)_k p^k$.
Check: This formula gives $E(X) = E(X)_1 = (n)_1 p = np$ and $E[X(X-1)] = E(X)_2 = (n)_2 p^2 = n(n-1)p^2$ from which it follows $Var(X) = E[X(X-1)] + E(X) - [E(X)]^2 = np(1-p)$.

34. No Solution

35. No Solution

36. a) The $k$th binomial moment $b_k$ is just given by

$$\sum_{i_1<\cdots<i_k} p^k = \binom{n}{k}p^k$$

b)

$$\mu = b_1 = \binom{n}{1}p^1 = np$$

$$\sigma^2 = 2b_2 + b_1 - b_1^2 = 2\binom{n}{2}p^2 + np - (np)^2 = n(n-1)p^2 + np - (np)^2 = np - np^2 = np(1-p)$$

37. a) $\binom{n}{k}\frac{(G)_k}{(N)_k}$

38. a) Using the $k$th binomial moment results, we observe that the $k$th factorial moment is just $k!$ times the $k$th binomial moment. Let $A_i$ be the event that the $i$th letter gets the correct address. Then $M_n = \sum_{i=1}^{n} A_i$ and the $k$th factorial moment is just

$$\begin{aligned}
f_k(M_n) &= k! \sum_{i_1<i_2<\cdots<i_k} P(A_{i_1}A_{i_2}\cdots A_{i_k}) \\
&= k! \sum_{i_1<i_2<\cdots\leq i_k} \frac{1}{n} \times \frac{1}{n-1} \times \cdots \times \frac{1}{n-k+1} \\
&= k!\binom{n}{k}\frac{1}{n} \times \frac{1}{n-1} \times \cdots \times \frac{1}{n-k+1} \\
&= 1
\end{aligned}$$

b) Let $X$ be Poisson(1), then the $k$th binomial moment is

$$\begin{aligned}
f_k(X) &= E((X)(X-1)\cdots(X-k+1)) \\
&= \sum_{i=0}^{\infty}(i) \times (i-1) \times \cdots \times (i-k+1)P(X=i) \\
&= \sum_{i=k}^{\infty}(i) \times (i-1) \times \cdots \times (i-k+1)\frac{e^{-1}}{i!} \\
&= e^{-1}\sum_{i=k}^{\infty}\frac{1}{(i-k)!} \\
&= 1
\end{aligned}$$

c) Use the expression of Exercise 3.4.22 for ordinary moments in terms of factorial moments.
As $n \to \infty$, the $k$th moment of $M_n$ are equals the $k$th moment of X for all $k \leq n$, so the convergence is immediate.

d) Using the Sieve formula,

$$\begin{aligned}
P(M_n = k) &= \sum_{j=k}^{n} \binom{j}{k} (-1)^{k-j} b_j(M_n) \\
&= \sum_{j=k}^{n} \binom{j}{k} (-1)^{k-j} \left(\frac{1}{j!}\right) \\
&= \sum_{j=k}^{n} (-1)^{k-j} \frac{1}{(k!)(j-k)!} \\
&= \frac{1}{k!} \sum_{j=0}^{n-k} (-1)^j \frac{1}{j!}
\end{aligned}$$

and as $n \to \infty$, this goes to $\frac{e^{-1}}{k!}$ as desired.

39. No Solution

40. The following identities hold:

$$\begin{aligned}
p_1 + p_2 + \cdots + p_n &= 1 \\
x_1 p_1 + x_2 p_2 + \cdots + x_n p_n &= \mu_1 \\
x_1^2 p_1 + x_2^2 p_2 + \cdots + x_n^2 p_n &= \mu_2 \\
&\cdots \\
&\cdots \\
x_1^{n-1} p_1 + x_2^{n-1} p_2 + \cdots + x_n^{n-1} p_n &= \mu_{n-1}
\end{aligned}$$

This is the same as $\mu = pM$, where $M$ is the $n \times n$ matrix

$$M = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}.$$

To see that $M$ must have an inverse, it suffices to show that $M$ has rank $n$; equivalently, that the columns of $M$ are linearly independent.

**Proof:** Let $c_0, c_1, \ldots, c_{n-1}$ be real constants such that

$$\begin{aligned}
0 &= c_0 + c_1 x_1 + c_2 x_1^2 + \cdots + c_{n-1} x_1^{n-1} \\
0 &= c_0 + c_1 x_2 + c_2 x_2^2 + \cdots + c_{n-1} x_2^{n-1} \\
&\cdots \\
&\cdots \\
&\cdots \\
0 &= c_0 + c_1 x_n + c_2 x_n^2 + \cdots + c_{n-1} x_n^{n-1}
\end{aligned}$$

Then the polynomial $f$ defined by $f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1}$ has $n$ distinct roots, namely $x_1, x_2, \ldots, x_n$. But $f$ has degree at most $n - 1$. Therefore $f$ must be identically zero. Conclude: all the $c$'s must be zero.

**41.** Rephrase this problem in terms of the collector's problem: Suppose there are $M$ objects in a complete collection, and suppose a trial consists of selecting an object at random with replacement from the set of $M$ objects. So on each trial, the object selected is equally likely to be any one of the $M$ possible, independently of what the other trials yielded. Now find $n$ so that $P(T \le n) = 0.9$, where $T$ is the number of trials required to see

a) an object specified in advance
b) at least one of the possible objects twice
c) every possible object at least once
d) at least once every object in a specified set comprising half of all possible objects
e) half of all objects.

In the present context, $M = 1024$, since each repetition of the ten toss experiment results in one of $2^{10}$ possible sequences.

a) $T$ has geometric distribution on $\{1, 2, \ldots\}$ with parameter $1/M$, so

$$P(T > n) = \left(1 - \frac{1}{M}\right)^n \quad n = 1, 2, 3, \ldots.$$

Therefore

$$P(T \le n) = 0.9 \iff \left(1 - \frac{1}{M}\right)^n = 0.1 \iff n = \frac{\log 0.1}{\log\left(1 - \frac{1}{M}\right)}.$$

For example, if $M = 1024$, then require $n \approx 2350$.

b) Let $n$ be a positive integer. Observe that

$$(T > n) = (\text{the first } n \text{ trials yielded all different objects}).$$

By comparison with the birthday problem (see index) we obtain

$$P(T > n) = \left(1 - \frac{1}{M}\right)\left(1 - \frac{2}{M}\right)\cdots\left(1 - \frac{n-1}{M}\right).$$

If $M$ is large, the right-hand side is approximately $e^{-\frac{n(n-1)}{2M}}$, and this approximation is good over all values of $n$. Thus if $M$ is large,

$$P(T > n) \approx e^{-\frac{n(n-1)}{2M}} \quad n = 1, 2, 3, \ldots.$$

and

$$\begin{aligned} P(T \le n) = 0.9 &\iff \exp -\frac{n(n-1)}{2M} \approx 0.1 \\ &\iff n^2 - n \approx 2M \log 10 \\ &\iff n \approx \frac{1 + \sqrt{1 + 8M \log 10}}{2}. \end{aligned}$$

For example, if $M = 1024$, then require $n \approx 70$.

c) We claim that for each real $x$,

$$\lim_{M \to \infty} P(T \le M(\log M + x)) = e^{-e^{-x}}.$$

First, a heuristic justification: We have $T = max(T_1, T_2, \ldots, T_M)$ where $T_i$ is the number of trials (from the beginning) required to see object $i$. Then each $T_i$ has geometric distribution on $\{1, 2, \ldots\}$ with parameter $1/M$, so

$$\begin{aligned} P(T_i \le M(\log M + x)) &= 1 - P(T_i > M(\log M + x)) \\ &\approx 1 - \left(1 - \frac{1}{M}\right)^{M(\log M + x)} \\ &\approx 1 - (e^{-1})^{\log M + x} \quad \text{if } M \text{ large} \end{aligned}$$

$$= 1 - \frac{e^{-x}}{M}.$$

Now if $M$ is large, the random variables $T_1, T_2, ..., T_M$ are almost independent; hence

$$P(T \leq M(\log M + x)) \approx \Pi_{i=1}^{M} P(T_i \leq M(\log M + x))$$

$$= \left(1 - \frac{e^{-x}}{M}\right)^M \approx e^{-e^{-x}}.$$

Now, a more rigorous justification: Let $k$ be a positive integer. We have

$$(T > k) = (\text{one of the } M \text{ objects has not been seen in the first} k \text{trials})$$

$$= \cup_{i=1}^{M}(\text{object } i \text{ not seen in } k \text{ trials}).$$

By inclusion-exclusion,

$$P(T > k) = P\left\{\cup_{i=1}^{M}(\text{object } i \text{ not seen in } k \text{ trials})\right\}$$

$$= \sum_{i=1}^{M} P(\text{object } i \text{ not seen in } k \text{ trials}) - \sum_{i<j} P(\text{objects } i,\ j \text{ not seen in } k \text{ trials})$$

$$+ ... + (-1)^{M-1} P(\text{objects } 1,\ \ldots,\ M \text{ not seen in } k \text{ trials})$$

$$= \binom{M}{1}\left(\frac{M-1}{M}\right)^k - \binom{M}{2}\left(\frac{M-2}{M}\right)^k + ... + (-1)^{M-1}\binom{M}{M}\left(\frac{M-M}{M}\right)^k$$

$$= \sum_{j=1}^{M}(-1)^{j-1}\binom{M}{j}\left(1 - \frac{j}{M}\right)^k.$$

Hence

$$P(T \leq k) = \sum_{j=0}^{M}(-1)^j\binom{M}{j}\left(1 - \frac{j}{M}\right)^k.$$

Let $x$ (real) be fixed. Replace $k$ by $M(\log M + x)$ and investigate the behavior of the probability as $M \to \infty$:

$$P(T \leq M(\log M + x)) = \sum_{j=0}^{M}(-1)^j\binom{M}{j}\left(1 - \frac{j}{M}\right)^{\lfloor M(\log M + x)\rfloor} = \sum_{j=0}^{\infty} a_{M,j},$$

where

$$a_{M,j} = \begin{cases} (-1)^j\binom{M}{j}\left(1 - \frac{j}{M}\right)^{\lfloor M(\log M + x)\rfloor} & 0 \leq j \leq M \\ 0 & j > M \end{cases}.$$

Note that for each $j = 0, 1, 2, ...$

$$\lim_{M\to\infty} a_{M,j} = \frac{(-1)^j e^{-jx}}{j!} = \frac{(-e^{-x})^j}{j!}.$$

We evaluate the limit as $M \to \infty$ of $P(T \leq M(\log M + x))$ as follows:

$$\lim_{M\to\infty} P(T \leq M(\log M + x)) = \lim_{M\to\infty}\sum_{j=0}^{\infty} a_{M,j} = \sum_{j=0}^{\infty}\lim_{M\to\infty} a_{M,j} = \sum_{j=0}^{\infty}\frac{(-e^{-x})^j}{j!} = e^{-e^{-x}};$$

the movement of the limit inside the summation is justified e.g. by the Weierstrass M-test (no relation to the present $M$).

Set $e^{-e^{-x}} = 0.9 \iff x \approx 2.25$. Then for $M$ large, we have approximately

$$P(T \leq n) = 0.9 \iff n \approx M(\log M + 2.25).$$

For example, if $M = 1024$, then require $n \approx 9400$.

d) Without loss of generality, the desired subcollection consists of objects $1, 2, 3, ..., M/2$. By a similar analysis to (c), we obtain

$$P(T > k) = P\left\{\cup_{i=1}^{M/2}(\text{object } i \text{ has not been seen in } k \text{ trials})\right\}$$

$$= \sum_{j=1}^{M/2}(-1)^j \binom{M/2}{j}\left(1 - \frac{j}{M}\right)^k;$$

$$P(T \le k) = \sum_{j=0}^{M/2}(-1)^j \binom{M/2}{j}\left(1 - \frac{j}{M}\right)^k;$$

$$P(T \le M(\log\frac{M}{2} + x)) = \sum_{j=0}^{M}(-1)^j \binom{M/2}{j}\left(1 - \frac{j}{M}\right)^{\lfloor M(\log\frac{M}{2}+x)\rfloor} \to e^{-e^{-x}} \text{ as } M \to \infty.$$

(Heuristically, we may view the collection of $M$ tickets as consisting of half desirable and half undesirable. $T$ is the time required to obtain all the desirable tickets. Since roughly half of our trials result in undesirable tickets, it would take twice as long to obtain all the desirable tickets as it normally would had there been no undesirable tickets at all. In other words, if $T'$ denotes the number of trials needed to obtain a complete collection of $M/2$ when drawing from the set of $M/2$ objects, then $T$ is distributed approximately like $2T'$; hence if $M$ is large, then

$$P(T \le M(\log\frac{M}{2} + x)) \approx P(2T' \le M(\log\frac{M}{2} + x))$$

$$= P(T' \le \frac{M}{2}(\log\frac{M}{2} + x))$$

$$\approx e^{-e^{-x}}.)$$

Thus for $M$ large,

$$P(T \le n) = 0.9 \iff n \approx M(\log\frac{M}{2} + 2.25).$$

For example, if $M = 1024$, then require $n \approx 8700$.

e) $T = X_1 + X_2 + ... + X_{M/2}$ where $X_1 = 1$, and for $i = 2, 3, ..., M$, $X_i$ is the additional number of trials [after obtaining the $(i-1)$st different object] required to obtain a new object. Then $X_i$ has geometric distribution on $\{1, 2, ...\}$ with parameter $p_i = (M - i + 1)/M$, $i = 1, ..., M$, and the $\{X_i\}$ are mutually independent. Hence

$$E(T) = \sum_{i=1}^{M/2} E(X_i) = \sum_{i=1}^{M/2}\frac{1}{p_i} = \sum_{i=1}^{M/2}\frac{M}{M-i+1}$$

$$= M\left(\frac{1}{M} + \frac{1}{M-1} + ... + \frac{1}{\frac{M}{2}+1}\right) \sim M\log 2 \text{ as } M \to \infty,$$

and

$$Var(T) = \sum_{i=1}^{M/2} Var(X_i) = \sum_{i=1}^{M/2}\left(\frac{i}{p_i^2} - \frac{1}{p_i}\right)$$

$$= \sum_{i=1}^{M/2}\left(\frac{M}{M-i+1}\right)^2 - \sum_{i=1}^{M/2}\left(\frac{M}{M-i+1}\right)$$

$$= M^2\left(\frac{1}{M^2} + \frac{1}{(M-1)^2} + ... + \frac{1}{(\frac{M}{2}1)^2}\right) - M\left(\frac{1}{M}\frac{1}{M-1} + ... + \frac{1}{\frac{M}{2}1}\right)$$

$$\sim M^2\left(\frac{1}{M}) - M(\log 2)\right) = M(1 - \log 2) \text{ as } M \to \infty.$$

Observe that the variables $X_1, X_2, ..., X_{M/2}$ are roughly identically distributed, or at least roughly of the same size (they have geometric distribution, with parameters between $1/2$ and $1$;

contrast this with the full set of variables $X_1, X_2, ..., X_M$ , where the latter variables have very high expectation). Hence

$$\frac{T - M\log 2}{\sqrt{M(1-\log 2)}} \approx \frac{T - E(T)}{SD(T)}$$

has approximately normal (0,1) distribution. In fact, it can be shown that for all $x$:

$$\lim_{M\to\infty} P\left(T \le M\log 2 + x\sqrt{M(1-\log 2)}\right) = \lim_{M\to\infty} P\left(\frac{T - M\log 2}{\sqrt{M(1-\log 2)}} \le x\right) = \Phi(x).$$

Set $\Phi(x) = 0.9 \iff x \approx 1.282$. Then for $M$ large, we have

$$P(T \le n) = 0.9 \iff n \approx M\log 2 + 1.282\sqrt{M(1-\log 2)}.$$

For example, if $M = 1024$, then require $n \approx 730$.

# Section 4.1

1. a) The desired probability is the area under the standard normal density $y = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$ between $z = 0$ and $z = 0.001$. The density is very nearly constant over this interval, so the desired probability is approximately

$$\text{(width of rectangle)(approx height of rectangle)} = 0.001 \times \frac{1}{\sqrt{2\pi}} = \frac{1}{1000\sqrt{2\pi}} = .000399.$$

b) Similarly the desired probability is approximately

$$0.001 \times \frac{1}{\sqrt{2\pi}} \times e^{-1/2} = \frac{1}{1000\sqrt{2\pi e}} = .000242.$$

2. a)

$$\int_1^\infty \frac{c}{x^4}dx = \frac{-c}{3x^3}\Big|_1^\infty = \frac{c}{3}$$

and since $f(x)$ is a density function, it must integrate to 1, so $c = 3$.

b)

$$E(X) = \int_1^\infty x\frac{3}{x^4}dx = \frac{-3}{2x^2}\Big|_1^\infty = \frac{3}{2}$$

c)

$$E(X^2) = \int_1^\infty x^2\frac{3}{x^4}dx = \frac{-3}{x}\Big|_1^\infty = 3$$

Thus $Var(X) = E(X^2) - (E(X))^2 = 3 - \frac{9}{4} = \frac{3}{4}$.

3. a) Since $f$ is a probability density, $\int_{-\infty}^{\infty} f(x)dx = 1$, so
$1 = c\int_0^1 x(1-x)dx = c\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big|_0^1 = \frac{c}{6} \Longrightarrow c = 6.$

b) $P(X \le \frac{1}{2}) = \int_0^{\frac{1}{2}} 6x(1-x)dx = 3x^2 - 2x^3\Big|_0^{\frac{1}{2}} = \frac{1}{2}$.

**Remark.** Once you note that $f(x)$ is symmetric about $\frac{1}{2}$ (draw a picture), the answer is clear without calculation.

c) $P(X \le \frac{1}{3}) = \int_0^{\frac{1}{3}} 6x(1-x)dx = 3x^2 - 2x^3\Big|_0^{\frac{1}{3}} = \frac{7}{27}$.

d) By the difference rule of probabilities,
$P(\frac{1}{3} < X \le \frac{1}{2}) = P(X \le \frac{1}{2}) - P(X \le \frac{1}{3}) = \frac{1}{2} - \frac{7}{27} = \frac{13}{54}$.
**Remark.** Of course you can obtain the same result by integration.

e) $E(X) = \int f(x)dx = \int_0^1 6x^2(1-x)dx = 2x^3 - \frac{3}{2}x^4\Big|_0^1 = \frac{1}{2}$.

Again, this is obvious by symmetry. But you must compute an integral for the variance:

$E(X^2) = \int_0^1 x^2 f(x)dx = \int_0^1 6x^3(1-x)dx = \frac{3}{2}x^4 - \frac{6}{5}x^5\Big|_0^1 = \frac{3}{10}$

$Var(X) = E(X^2) - [E(X)]^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}$

4. a) We know that $\int_0^1 cx^2(1-x)^2dx = 1$, and

$$\int_0^1 cx^2(1-x)^2dx = c\int_0^1 (x^2 - 2x^3 + x^4)dx = c\left(\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5}\right)\Big|_0^1 = \frac{c}{30}$$

so $c = 30$.

b)

$$E(X) = \int_0^1 30x^3(1-x)^2 dx = 30\int_0^1 (x^3 - 2x^4 + x^5)dx = 30\left(\frac{1}{4} - \frac{2}{5} + \frac{1}{6}\right) = \frac{1}{2}$$

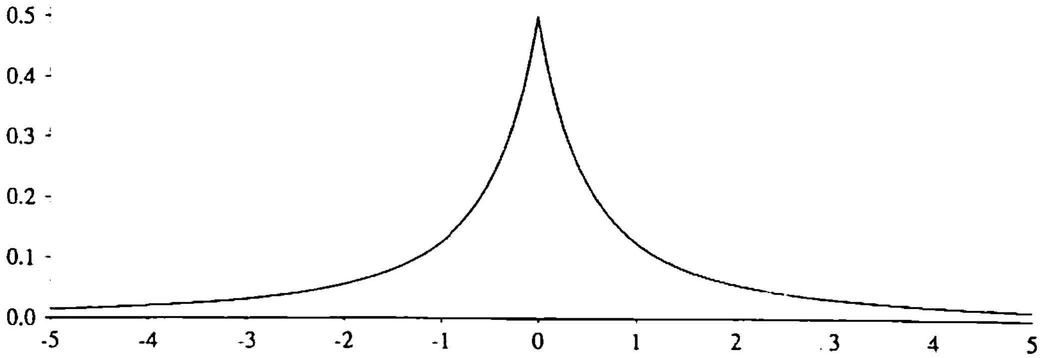
c)

$$E(X^2) = \int_0^1 30x^4(1-x)^2 dx = 30\int_0^1 (x^4 - 2x^5 + x^6)dx = 30\left(\frac{1}{5} - \frac{1}{3} + \frac{1}{7}\right) = \frac{2}{7}$$

and so

$$Var(X) = \frac{2}{7} - \frac{1}{4} = \frac{1}{28}$$

5. a) Graph:




b)

$$P(-1 < X < 2) = \frac{1}{2}\left(\int_{-1}^0 \frac{1}{(1-x)^2}dx + \int_0^2 \frac{1}{(1+x)^2}dx\right) = \frac{1}{2}\left(\left[\frac{1}{1-x}\right]\Big|_{-1}^{0} + \left[-\frac{1}{1+x}\right]\Big|_{0}^{2}\right) = \frac{7}{12}$$

c) $P(|X| > 1) = 2 \times \frac{1}{2}\int_1^\infty \frac{1}{(1+x)^2}dx = -\frac{1}{1+x}\Big|_1^\infty = \frac{1}{2}$

d) No, because $E(|X|) = \int_{-\infty}^\infty \frac{|x|}{2(1+|x|)^2}dx = \int_0^\infty \frac{x}{(1+x)^2}dx = \infty$ (see Example 3).

6. a) $\frac{1}{3} = P(X \le 0) = P\left(\frac{X-\mu}{\sigma} < -\frac{\mu}{\sigma}\right) = \Phi\left(-\frac{\mu}{\sigma}\right) \iff -\frac{\mu}{\sigma} = -.4303$;
$\frac{2}{3} = P(X \le 1) = P\left(\frac{X-\mu}{\sigma} < \frac{1-\mu}{\sigma}\right) = \Phi\left(\frac{1-\mu}{\sigma}\right) \iff \frac{1-\mu}{\sigma} = .4303$.
Subtract: $\frac{1}{\sigma} = .8606 \iff \sigma = 1.162$. Hence $\mu = .4303\sigma = 0.5$. Or you may easily see that, by symmetry, $\mu$ must be located halfway between 0 and 1.

b) If $P(X \le 1) = \frac{3}{4}$ then $\Phi\left(\frac{1-\mu}{\sigma}\right) = \frac{3}{4} \iff \frac{1-\mu}{\sigma} = .6742$. This implies that $\frac{1}{\sigma} = .6742 + .4303 = 1.1045$, and $\sigma = .9054$, and $\mu = .3896$.

7. Let $X$ be the height of an individual picked at random from this population. We know that the distribution of $X$ is approximately normal $(\mu, \sigma^2)$, with $\mu = 70$ (inches), and $P(X > 72) = .1$ . That is,

$.9 = P(X \le 72) = P\left(\frac{X-70}{\sigma} \le \frac{72-70}{\sigma}\right) \approx P\left(Z \le \frac{2}{\sigma}\right) = \Phi\left(\frac{2}{\sigma}\right)$

(where $Z$ has standard normal distribution). Hence $2/\sigma = 1.28$ from the normal table, and the chance that the height of an individual picked at random exceeds 74 inches is

$P(X > 74) = P\left(\frac{X-70}{\sigma} > \frac{74-70}{\sigma}\right) \approx P\left(Z > \frac{4}{\sigma}\right) = P(Z > 2.56) = 1 - \Phi(2.56) = .0052$.

In a group of 100, the number of individuals who are over 74 inches tall therefore has binomial(100, .0052) distribution. By the Poisson approximation, the chance that there are 2 or more such individuals is approximately (with $\mu = 100 \times .0052 = .52$)

$1 - (e^{-\mu} + e^{-\mu}\mu) = 1 - e^{-.52}(1 + .52) = .096$.

8. a)

$$\Phi\left(\frac{12.2-12}{1.1}\right)-\Phi\left(\frac{11.8-12}{1.1}\right)=.1443$$

b)

$$\Phi\left(\frac{12.2-12}{.11}\right)-\Phi\left(\frac{11.8-12}{.11}\right)=.9307$$

Since the Central Limit Theorem says that the average of a large number of measurements will be normal, it is not necessary for the measurements themselves to be normal, although if the measurements are extremely skewed then 100 may not be a large enough number.

9. The distribution of $S_4$ is approximately normal with mean 2 and variance $4 \times \frac{1}{12} = \frac{1}{3}$.

$P(S_4 \geq 3) \approx 1 - \Phi\left(\frac{3-2}{1/\sqrt{3}}\right) = 1 - \Phi(1.73) = 1 - .9582 = .0418$

10. a) $\Phi(\frac{9.800-9.7800}{0.0031}) - \Phi(\frac{9.7840-9.7800}{0.0031}) = \Phi(6.45) - \Phi(1.29) = 0.0985$

b) $\Phi(\frac{9.7794-9.7800}{0.0031}) = \Phi(-0.19) = 1 - \Phi(0.19) = 0.4246$

c) $\Phi(1.28) \approx 0.90$, so the weight $= 9.7800 + (1.28 \times 0.0031) = 9.7840$ gm.

11. a) $\Phi(0.43) \approx \frac{2}{3}$, so $\frac{1.1-1}{\sigma} \approx 0.43$ and $\sigma \approx 0.2325$.

b) $\Phi\left(\frac{0.20}{0.2325}\right) - \Phi\left(-\frac{0.20}{0.2325}\right) = 2\Phi(0.86) - 1 = 0.6102$

c) $\Phi(0.675) \approx 0.75$, so the diameter is 0.675 multiples of $\sigma$ below the mean. $1 - (0.675 \times 0.2325) \approx 0.84$

12. a) Range of $X$: $[-2, 2]$

If $-2 \leq x \leq 2$, then

$f(x)dx = P(X \in dx) = \frac{2\times(2-|x|)dx}{4\times(\frac{1}{2}\times 2\times 2)} = \frac{1}{4}(2-|x|)dx$,

so $f(x) = (2-|x|)/4$. Elsewhere $f(x) = 0$ .

b) Range of $X$: $[-2, 1]$.

If $-2 \leq x < 0$, then

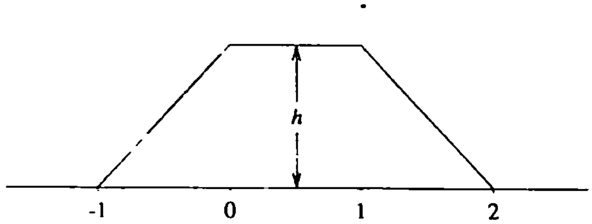
$f(x)dx = P(X \in dx) = \frac{(2+x)dx}{\frac{1}{2}\times 3\times 2} = \frac{1}{3}(2+x)dx$ .

If $0 \leq x \leq 1$, then $f(x)dx = \frac{2(1-x)dx}{3}$.

Elsewhere $f(x) = 0$.

c) Range of $X$ : $[-1, 2]$.

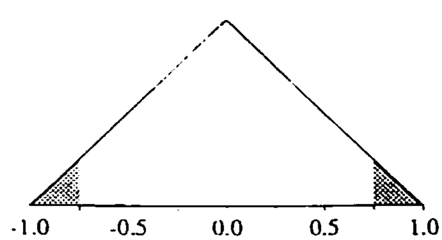
The density will be linear on $[-1, 0]$, constant on $[0, 1]$, linear on $[1, 2]$:




Area = $2h$

To make area $= 1$, $h$ must satisfy $2h = 1$, or $h = \frac{1}{2}$.

13. Let $X$ denote the length of a rod produced. Note that the probabilities of interest remain the same under a linear change of scale. So, without loss of generality, assume $X$ has a triangular density from $-1$ to $1$, as in the figure.




a) $P(|X| > .75) = 1/16$ from the figure.

b) $P(|X| \le .5 \,\|X| \le .75) = \frac{1-(1/2)^2}{1-(1/4)^2} = .8$ .

If the customer buys $n$ rods, the number $N$ of rods which meet his specifications (a··uming independence of rod lengths) has binomial ($n$, .8) distribution. Need $n$ such that

$P(N \ge 100) \ge .95 \iff P(N < 100) \le .05.$

Now by the normal approximation to the binomial,

$P(N < 100) \approx \Phi\left(\frac{99.5-(n)(.8)}{\sqrt{(n)(.8)(.2)}}\right)$ .

So solve

$\Phi\left(\frac{99.5-(n)(.8)}{\sqrt{(n)(.8)(.2)}}\right) \le .05 \iff \frac{99.5-(.8)n}{(.4)\sqrt{n}} \le -1.645 \iff n \ge 134.$

14. (continued from Exercise 13)
Again, the probabilities remain the same under a linear change of scale. Let $Y$ be the length of a rod produced by the current manufacturing process. To get the mean and standard deviation of $Y$, note that the density of $X$ is given by

$$f_X(x) = \begin{cases} 1-|x| & |x| \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

Therefore $E(Y) = E(X) = 0$ by symmetry; $Var(Y) = Var(X) = E(X^2) = \int_{-1}^{1} x^2(1-|x|)dx = 2\int_0^1 (x^2 - x^3)dx = \frac{1}{6}$.

So $Y$ has normal (0, 1/6) distribution.

a) $P(|Y| > .75) = 2 \times \left[1 - \Phi(\sqrt{6} \times .75)\right] = .066193$

b) $P(|Y| \le .5) = 2 \times \Phi(\sqrt{6} \times .5) - 1 = .779328;$
therefore $P(|Y| \le .5 \,\|Y| \le .75) = \frac{.779328}{1-.066193} = .834571.$
So you should choose the manufacturer of this exercise, since each rod that you buy would have a greater chance of meeting your specifications (although not by much).

15. a) $(0, 1/2)$ b) $\text{erf}(x) = 2\Phi(\sqrt{2}x) - 1$ c) $\Phi(z) = (\text{erf}(z/\sqrt{2}) + 1)/2$

# Section 4.2

1. Let $X$ denote the lifetime of an atom. Then $X$ has exponential distribution with rate $\lambda = \log 2$.

   a) $P(X > 5) = e^{-5\lambda} = (1/2)^5 = 1/32$ .

   b) Find $t$ (years) such that

   $$P(X > t) = .1 \iff e^{-t\lambda} = .1 \iff t = \frac{\log 10}{\lambda} = 3.32$$

   c) Assuming that the lifetimes of atoms are independent, the number $N_t$ of atoms remaining after $t$ years has binomial $(1024,\ e^{-\lambda t})$ distribution. So find $t$ such that

   $$E(N_t) = 1 \iff 1024e^{-\lambda t} = 1 \iff t = \frac{\log 1024}{\lambda} = 10.$$

   d) $N_{10}$ has binomial $(1024,\ 1/1024)$ distribution, which is approximately Poisson (1). So by the Poisson approximation,

   $$P(N_{10} = 0) \approx e^{-1} = .3679.$$

2. a) Let $X$ denote the lifetime of an atom, then $X$ is exponentially distributed with rate $\lambda = \frac{\log 2}{.5} = \log 4$ per century. Now we have $10^{20}$ atoms, and we wish to find the time such that we expect 1 out of $10^{18}$ atoms to survive, which we can do by solving $P(X > t) = \frac{1}{10^{18}}$ for $t$. We know that $P(X > t) = e^{-\lambda t}$ and so $-(\log 4)(t) = -18 \log 10$ and finally $t = 18\frac{\log 10}{\log 4} = 29.9 \approx 30$ centuries.

   b) This is equivalent to saying that there is about a 50% chance that no atoms are left. Let $N_t$ be the number of atoms left at time $t$, and we wish to find $t$ such that $P(N_t = 0) = .5$. Note that $N_t$ will be a binomial $(n, p)$ where $n = 10^{20}$. Note further that this will be approximately a Poisson with $\mu = np$. Thus we observe that

   $$P(N_t = 0) = \frac{e^{-np}(np)^0}{0!} = .5$$

   and so $np = \log 2$ and finally $p = \frac{\log 2}{n}$. We also know that $p = .5^t$ since $p$ is the probability that a given atom will survive for $t$ centuries, and so

   $$.5^t = \frac{\log 2}{n}$$

   $$t = \frac{\log\left(\frac{\log 2}{n}\right)}{\log .5} = 66.97 \approx 67$$

   So after 67 centuries there is about a 50% chance that no atoms are left.

3. Let $T$ be the time until the next earthquake, then we have in general that $P(T < t) = 1 - e^{-\lambda t}$.

   a) The probability of an earthquake in the next year is $1 - e^{-1} = 0.6321$.

   b) Similarly, $P(T < .5) = 1 - e^{-.5} = 0.3935$.

   c) $P(T < 2) = 1 - e^{-2} = 0.8647$

   d) $P(T < 10) = 1 - e^{-10} = 0.99995$

4. Let $W$ be the lifetime of a component. Then $W$ has exponential distribution with rate $\lambda = 1/10$.

   a) $P(W > 20) = e^{-20\lambda} = e^{-2} \approx 0.135$ .

   b) The median lifetime $m$ satisfies

   $$1/2 = P(W > m) = e^{-m\lambda} \iff m = \frac{(\log 2)}{\lambda} = 6.93.$$

c) $SD(W) = 1/\lambda = 10$ .

d) If $X$ denotes the average lifetime of 100 independent components, then $E(X) = 10$ and $SD(X) = 10/\sqrt{100} = 1$ so by the normal approximation

$$P(X > 11) = 1 - P(X \le 11) \approx 1 - \Phi(1) = .1586$$

e) Let $W_1$ and $W_2$ denote the lifetimes of the first and second components respectively. Then

$$P(W_1 + W_2 > 22) = P(N < 2) = e^{-2.2} + 2.2e^{-2.2} = .35457$$

where $N$ has Poisson $(22\lambda = 2.2)$ distribution.

5. a) $P(W_4 \le 2) = 1 - P(W_4 > 2) = 1 - e^{-2} \approx .86$.

b) $P(T_4 \le 5) = P(N(0,5] \ge 4) = 1 - P(N(0,5] \le 3) = 1 - e^{-5}(1 + 5 + 25/2 + 125/6) \approx .73$

c) $E(T_4) = E(W_1 + W_2 + W_3 + W_4) = E(W_1) + E(W_2) + E(W_3) + E(W_4) = 1 + 1 + 1 + 1 = 4$

Note. $T_4$ has gamma (4, 1) distribution.

6. Let $N_2$ be the number of hits during the first 2 minutes, and $N_4$ be the number of hits during the first 4 minutes. Then $N_2$ has Poisson (2) distribution and $N_4$ has Poisson (4) distribution, and

$$P(2 \le T_3 \le 4) = P(T_3 > 2) - P(T_3 > 4)$$

$$= P(N_2 \le 2) - P(N_4 \le 2)$$

$$= \left(e^{-2} + 2e^{-2} + \frac{e^{-2}2^2}{2!}\right) - \left(e^{-4} + 4e^{-4} + \frac{e^{-4}4^2}{2!}\right) = 5e^{-2} - 13e^{-4} = 0.43857.$$

7. $1 - e^{-t_p\lambda} = p \iff t_p = -\frac{1}{\lambda}\log(1-p)$

8. To compute the density $f$ of $X$, argue infinitesimally:

$f(t)dt = P(X \in (t, t+dt))$

$= P(X \in (t, t+dt) | \lambda = \frac{1}{100})P(\lambda = \frac{1}{100}) + P(X \in (t, t+dt) | \lambda = \frac{1}{200})P(\lambda = \frac{1}{200})$

$= f_{Y_1}(t)dt \cdot \frac{1}{3} + f_{Y_2}(t)dt \cdot \frac{2}{3}$ ,

where $f_{Y_1}$ is the density of an exponential random variable $Y_1$ having rate $1/100$, and $f_{Y_2}$ is the density of an exponential random variable $Y_2$ having rate $1/200$. Hence

$f(t) = \frac{1}{3}f_{Y_1}(t) + \frac{2}{3}f_{Y_2}(t)(t > 0)$.

a) $P(X \ge 200) = \frac{1}{3}P(Y_1 \ge 200) + \frac{2}{3}P(Y_2 \ge 200) = \frac{1}{3}e^{-200/100} + \frac{2}{3}e^{-200/200} \approx .29$

b) $E(X) = \frac{1}{3}E(Y_1) + \frac{2}{3}E(Y_2) = \frac{1}{3} \times 100 + \frac{2}{3} \times 200 = \frac{500}{3}$

c) $E(X^2) = \frac{1}{3}E(Y_1^2) + \frac{2}{3}E(Y_2^2) = \frac{1}{3} \times 2 \times 100^2 + \frac{2}{3} \times 2 \times 200^2 = 6 \times 100^2$. Therefore $Var(X) = E(X^2) - [E(X)]^2 = \frac{290{,}000}{9}$.

9. a) $\Gamma(r+1) = \int_0^\infty x^r e^{-x}dx = -x^r e^{-x}\Big|_0^\infty + r\int_0^\infty x^{r-1}e^{-x}dx = r\Gamma(r)$

b) Note that $\Gamma(1) = \int_0^\infty e^{-x}dx = 1$.

So $\Gamma(r+1) = r\Gamma(r) = r(r-1)\Gamma(r-1) = \cdots = r!\Gamma(1) = r!$.

c) $E(T^n) = \int_0^\infty t^n e^{-t}dt = \Gamma(n+1) = n!$.

$Var(T) = E(T^2) - [E(T)]^2 = 2 - 1 = 1 \Longrightarrow SD(T) = 1$.

d) $P(\lambda T > u) = P(T > u/\lambda) = e^{-\lambda u/\lambda} = e^{-u}$, so $\lambda T$ has exponential (1) distribution, and $E[(\lambda T)^n] = n! \Longrightarrow E(T^n) = n!/\lambda^n$.

$SD(\lambda T) = 1 \Longrightarrow SD(T) = 1/\lambda$.

10. a) Let $X = int\ (T)$. $X$ takes values in $\{0, 1, 2, \ldots\}$ and

$P(X = k) = P(k \le T < k + 1)$
$= P(T < k + 1) - P(T < k)$
$= (1 - e^{-\lambda(k+1)}) - (1 - e^{-\lambda k})$
$= e^{-\lambda k}(1 - e^{-\lambda})$
$= q^k p$, where $p = 1 - e^{-\lambda}, q = 1 - p$.

b) If $T$ has exponential distribution on $(0, \infty)$ with rate $\lambda$, then $mT$ has exponential distribution with rate $\lambda/m$ (why?) so by a) $mT_m = int\ (mT)$ has geometric distribution on $\{0, 1, \ldots\}$ with success parameter $p_m = 1 - e^{-\lambda/m}$.

Conversely, if for each $m = 1, 2, \ldots$ we have that $int\ (mT)$ has geometric distribution on $\{0, 1, \ldots\}$ with success parameter $p_m$, then: argue that there exists $\lambda > 0$ such that $p_m = 1 - e^{-\lambda/m}$ (set $q_1 = e^{-\lambda}$ and consider $P(T \ge 1)$); next argue that for each rational $t > 0$ we have $P(T \ge t) = e^{-\lambda t}$; finally argue that this last must hold for all real $t$ by using the fact that $P(T > t)$ is a nonincreasing function of $t$.

c) $E(T_m) = \frac{1}{m}E(int\ (mT)) = \frac{1}{m} \cdot \frac{q_m}{p_m} = \frac{1}{m} \cdot \frac{e^{-\lambda/m}}{(1-e^{-\lambda/m})} \to \frac{1}{\lambda}$ as $m \to \infty$
(Use l'Hôpital's rule, or series expansions).

Since $T_m \le T \le T_m + \frac{1}{m}$, we have $E(T_m) \le E(T) \le E(T_m) + \frac{1}{m}$. Let $m \to \infty$ to see $E(T) = \frac{1}{\lambda}$.
Similarly argue that $E(T_m^2) \to \frac{2}{\lambda^2}$ as $m \to \infty$, and therefore that $E(T^2) = \frac{2}{\lambda^2}$.

11. a) Want to show $P(T \ge t) \approx e^{-\lambda t}$. Write $\Delta = 10^{-6}$ seconds, and consider $t = n\Delta$ for $n = 0, 1, 2, \ldots$. By assumption,

$$P(T \le (n+1)\Delta \mid T > n\Delta) = \lambda\Delta \quad \text{for all } n = 0, 1, 2, \ldots; \quad \text{equivalently}$$

$$P(T > (n+1)\Delta | T > n\Delta) = 1 - \lambda\Delta \quad \text{for all } n = 0, 1, 2, \ldots.$$

Therefore

$$P(T \ge 0) = 1,$$

$$P(T > \Delta) = P(T > 0)P(T > \Delta | T > 0) = 1 \times (1 - \lambda\Delta) = 1 - \lambda\Delta,$$

$$P(T > 2\Delta) = P(T > \Delta)P(T > 2\Delta | T > \Delta) = (1 - \lambda\Delta)(1 - \lambda\Delta) = (1 - \lambda\Delta)^2.$$

In general, $P(T > n\Delta) = (1 - \lambda\Delta)^n$. Use the approximation $1 - \lambda\Delta \approx e^{-\lambda\Delta}$ to conclude

$$P(T > n\Delta) \approx (e^{-\lambda\Delta})^n = e^{-n\lambda\Delta}.$$

Put $t = n\Delta : P(T > t) \approx e^{-\lambda t}$.

b) $P(1 < T \le 2) \approx \int_1^2 \lambda e^{-\lambda t} dt = e^{-\lambda} - e^{-2\lambda}$.

12. a) Differentiate the gamma $(r, \lambda)$ density:

$$\frac{d}{dt}f_{r,\lambda}(t) = \frac{\lambda^r}{\Gamma(r)}\left[(r-1)t^{r-2}e^{-\lambda t} - \lambda e^{-\lambda t}t^{r-1}\right] = \frac{\lambda^r}{\Gamma(r)}e^{-\lambda t}t^{r-2}(r - 1 - \lambda t).$$

If $r \le 1$, then the derivative is negative for all $t > 0$, so $f_{r,\lambda}$ is maximized at 0.
If $r > 1$, then the derivative is zero at $t^* = (r-1)/\lambda$, is positive to the left of $t^*$, and is negative to its right. So $t^*$ yields a local maximum for the density. But the density is zero at $t = 0$ and tends to zero as $t \to \infty$; hence the density achieves its overall maximum at $t^*$.
If $r < 1$, then the density blows up to $\infty$ as $t$ approaches zero.

b) For $k = 0, 1, 2, \ldots$ we have :

$$E(T^k) = \int_0^\infty t^k \frac{\lambda^r}{\Gamma(r)}t^{r-1}e^{-\lambda t}dt = \frac{\lambda^r}{\Gamma(r)}\int_0^\infty t^{(r+k)-1}e^{-\lambda t}dt = \frac{\lambda^r}{\Gamma(r)}\frac{\Gamma(r+k)}{\lambda^{r+k}} = \frac{1}{\lambda^k}\frac{\Gamma(r+k)}{\Gamma(r)}.$$

Hence

$E(T) = \frac{1}{\lambda}\frac{\Gamma(r+1)}{\Gamma(r)} = \frac{r}{\lambda}$

$E(T^2) = \frac{1}{\lambda^2}\frac{\Gamma(r+2)}{\Gamma(r)} = \frac{(r+1)r}{\lambda^2}$
$Var(T) = \frac{r^2+r}{\lambda^2} - (\frac{r}{\lambda})^2 = \frac{r}{\lambda^2}$
$SD(T) = \frac{\sqrt{r}}{\lambda}$.

**13.** a) Estimate $\lambda = 1/20 = 5\%$ per day.

b) $N_d$ has binomial $(10,000, e^{-d/20})$ distribution. Therefore

$$E(N_d) = 10,000e^{-d/20}, SD(N_d) = 100\sqrt{e^{-d/20}(1 - e^{-d/20})}.$$

From this calculate
$E(N_{10}) = 6065$; $SD(N_{10}) = 49$;
$E(N_{20}) = 3679$; $SD(N_{20}) = 48$;
$E(N_{30}) = 2231$; $SD(N_{30}) = 42$.

**14.** Option b) is correct: The probability that a component fails in its first day of use is $1 - e^{-1/20}$, which is approximately $1/20 = 5\%$, because $1 - e^{-x} \approx x$ as $x \to 0$; and is less than 5% because $e^{-x} \geq 1 - x$ for all $x$ (See Appendix III). Exact value is 4.877...%.

**15.** a) 80 days

b) 40 days

c) $P(T_{total} \geq 60) = P(\text{at most 3 failures in 60 days}) = e^{-3}(1 + 3 + \frac{3^2}{2} + \frac{3^3}{6}) = 13e^{-3} \approx .64723$ since the number of failures in 60 days has Poisson $(.05 \times 60)$ distribution.

**16.** Say a total of $k$ components will do. Since $P(T_{total} \geq 60) = P(N_{60} \leq k - 1)$, we require

$$P(N_{60} \leq k - 1) \geq 0.9$$

By trial and error, we find $P(N_{60} \leq 5) \approx 0.91608$, so a total of six components (five spares) will do.

**17.** Redoing the satellite problem:

a) 80 days

b) $20\sqrt{2}$ days

c) Guess the answer to c) should be larger, because now 60 days is more standard deviations below the mean of 80 days. In fact, $T_{total}$ has the same distribution as the sum of 8 independent exponential (.1) random variables, so $P(T_{total} \geq 60) = P(N_{60} < 8) \approx .744$ where now $N_{60}$ has Poisson (6) distribution.

Redoing the preceding problem: Since $P(N_{60} \not< 10) \approx .9161$, it follows that four spare components will do.

# Section 4.3

1. a) $P(T \le b) = 1 - P(T > b) = 1 - G(b)$.
   b) $P(a \le T \le b) = P(T \ge a) - P(T > b) = G(a) - G(b)$. (Since $T$ is continuous, $P(T \ge a)$ equals $P(T > a)$ equals $G(a)$.)

2. Suppose $T$ has constant hazard rate: Say $\lambda(t) = c$ for all $t > 0$. Use (7) to get

$$G(t) = e^{-\lambda t}, t > 0.$$

Then the density of $T$ is, by (5),

$$f(t) = -\frac{dG(t)}{dt} = ce^{-ct}, t > 0,$$

so $T$ has exponential distribution with rate $c$.

Conversely, if $T$ has exponential distribution with rate $\lambda$, then for each $t > 0$:

$$f(t) = \lambda e^{-\lambda t}; G(t) = P(T > t) = e^{-\lambda t}; \lambda(t) = \frac{f(t)}{G(t)} = \lambda.$$

3. a) $\left(\frac{b}{b+t}\right)^a, t > 0$.
   b) $\left(\frac{a}{b+t}\right)\left(\frac{b}{b+t}\right)^a = \frac{ab^a}{(b+t)^{a+1}}, t > 0$.

4. (i) $\Longrightarrow$ (ii):

$$\exp\left(-\int_0^t \lambda(u)du\right) = \exp\left(-\int_0^t \lambda\alpha u^{\alpha-1}du\right) = \exp(-\lambda t^\alpha) = G(t).$$

(ii) $\Longrightarrow$ (iii): differentiate $G$ with respect to $t$:

$$-\frac{d}{dt}G(t) = -\frac{d}{dt}e^{-\lambda t^\alpha} = \lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha} = f(t).$$

(iii) $\Longrightarrow$ (ii):

$$\int_t^\infty f(u)du = \int_t^\infty \lambda\alpha u^{\alpha-1}e^{-\lambda u^\alpha}du = -e^{-\lambda u^\alpha}\Big|_{u=t}^\infty = e^{-\lambda t^\alpha} = G(t).$$

(iii) & (ii) $\Longrightarrow$ (i):

$$\frac{f(t)}{G(t)} = \frac{\lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha}}{e^{-\lambda t^\alpha}} = \lambda\alpha t^{\alpha-1} = \lambda(t).$$

5. Let $k \ge 0$.
   a) $E(T^k) = \int_0^\infty t^k f(t)dt = \int_0^\infty t^k\lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha}dt = \int_0^\infty \left(\frac{x}{\lambda}\right)^{k/\alpha}e^{-x}dx = \lambda^{-k/\alpha}\int_0^\infty x^{k/\alpha}e^{-x}dx = \lambda^{-k/\alpha}\Gamma\left(\frac{k}{\alpha}+1\right)$.
   (Substitute $x = \lambda t^\alpha$.)
   b) By (a),

$$E(T) = \lambda^{-1/\alpha}\Gamma\left(\frac{1}{\alpha}+1\right);$$

$$E(T^2) = \lambda^{-2/\alpha}\Gamma\left(\frac{2}{\alpha}+1\right);$$

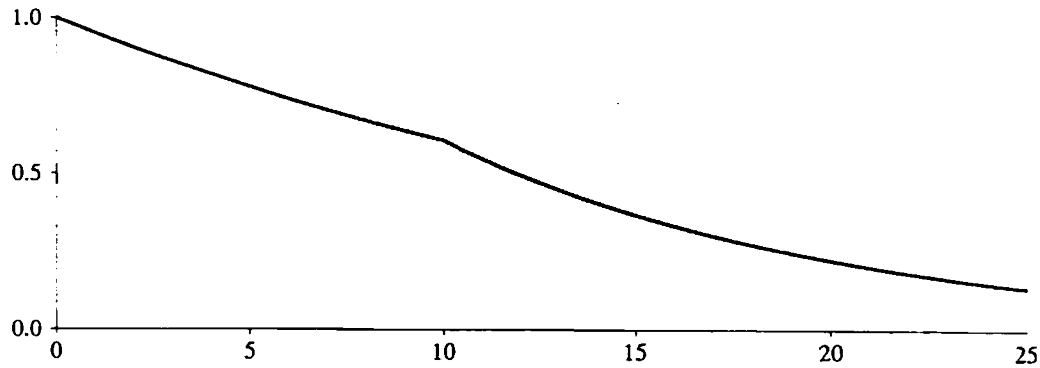
hence $$Var(T) = \lambda^{-2/\alpha}\left\{\Gamma\left(\frac{2}{\alpha}+1\right) - \left[\Gamma\left(\frac{1}{\alpha}+1\right)\right]^2\right\}.$$

6. We have $\lambda(t) = 1/20$ if $0 \le t \le 10$, and $\lambda(t) = 1/10$ if $t > 10$.

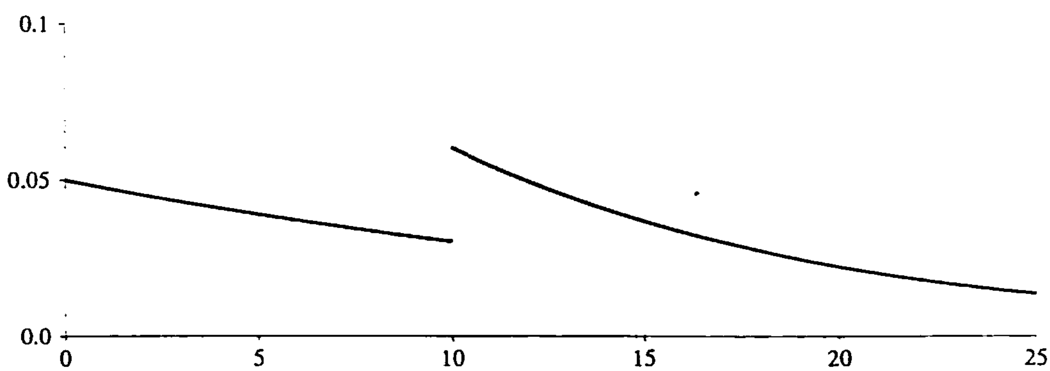
a) $P(T > 15) = G(15) = \exp\left(-\int_0^{15}\lambda(u)du\right) = \exp\left\{-[\int_0^{10}(1/20)du + \int_{10}^{15}(1/10)du\right. = e^{-1} \approx$ .3679.

b) If $0 \le t \le 10$ then $G(t) = \exp\left(-\int_0^t (1/20)du\right) = e^{-t/20}$;

if $t > 10$ then $G(t) = \exp\left\{-[\int_0^{10}(1/20)du + \int_{10}^{t}(1/10)du]\right\} = e^{-(\frac{1}{2}+\frac{t-10}{10})}$.




c) $f(t) = -\frac{dG(t)}{dt} = \begin{cases} \frac{1}{20}e^{-t/20} & 0 < t < 10 \\ \frac{1}{10}e^{-(t/10-1/2)} & t > 10 \end{cases}$




d) $E(T) = \int_0^\infty G(t)dt = \int_0^{10} e^{-t/20}dt + \int_{10}^\infty e^{-(\frac{1}{2}+\frac{t-10}{10})}dt = 20\int_0^{1/2} e^{-u}du + 10e^{-1/2}\int_0^\infty e^{-v}dv =$ 13.93.

(Substitute $u = t/20$, $v = (t-10)/10$ .)

7. a) Integrate by parts the relation $E(T^2) = \int_0^\infty t^2 f(t)dt$.

b) $E(T^2)$ is 400. So $SD(T)$ is $\sqrt{400 - 100\pi} \approx 9.265$.

c) If $\bar{T}$ denotes the average lifetime of 100 components, then $E(\bar{T}) = E(T_1) = 17.7245$ and $SD(\bar{T}) = SD(T_1)/\sqrt{100} = 0.9265$ so by the normal approximation

$P(\bar{T} > 20) \approx 1 - \Phi(2.456) = 0.007$ .

8. a) Only for $a \ge 0$, $b \ge 0$, and either $a > 0$ or $b > 0$.

b) $G(t) = \exp -\left(\frac{at^2}{2} + bt\right)$.

c) $f(t) = (at+b)\exp-\left(\frac{at^2}{2}+bt\right)$

d) $E(T) = \int_0^\infty G(t)dt$

$= \int_0^\infty \exp-\left(\frac{at^2}{2}+bt\right)dt$

$= e^{b^2/2a}\int_0^\infty \exp-\frac{a}{2}\left(t+\frac{b}{a}\right)^2 dt$

$= e^{b^2/2a}\int_{b/\sqrt{a}}^\infty e^{-z^2/2}\frac{dz}{\sqrt{a}}$ ( $z=\sqrt{a}\left(t+\frac{b}{a}\right)$ )

$= \sqrt{\frac{2\pi}{a}}e^{b^2/2a}\int_{b/\sqrt{a}}^\infty \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$

$= \sqrt{\frac{2\pi}{a}}e^{b^2/2a}\left(1-\Phi(b/\sqrt{a})\right)$

e) Compute $E(T^2)$ using

$E\left(\frac{a}{2}T^2+bT\right) = \int_0^\infty\left(\frac{at^2}{2}+bt\right)(at+b)\exp-\left(\frac{at^2}{2}+bt\right)dt = \int_0^\infty ue^{-u}du = 1;$

then use $Var(T) = E(T^2)-[E(T)]^2$.

**9.** a) Put (5) in (6) to get

$$\lambda(t) = \frac{f(t)}{G(t)} = -\frac{1}{G(t)}\frac{dG(t)}{dt} = -\frac{d}{dt}\log G(t).$$

b) $\int_0^t \lambda(u)du = \int_0^t -\frac{d}{du}\log G(u)du = -\log G(t)$

$$\iff \log G(t) = -\int_0^t \lambda(u)du$$

$$\iff G(t) = \exp\left\{-\int_0^t \lambda(u)du\right\}$$

**10.** a,b) Smaller, since for each $s,t$ positive:

$$\begin{aligned}
P(T>s+t|T>s) &= \frac{P(T>s+t, T>s)}{P(T>s)}\\
&= \frac{P(T>s+t)}{P(T>s)}\\
&= \exp\left(-\int_0^{s+t}\lambda(u)du\right)\Big/\exp\left(-\int_0^s\lambda(u)du\right)\\
&= \exp\left(-\int_s^{s+t}\lambda(u)du\right)\\
&= \exp\left(-\int_0^t\lambda(x+s)dx\right)\\
&\le \exp\left(-\int_0^t\lambda(x)dx\right) \qquad \text{since } \lambda \text{ is increasing}\\
&= P(T>t).
\end{aligned}$$

c) If $\lambda$ is decreasing, then the inequality reverses.

# Section 4.4

1. Exponential $(\lambda/c)$.

2. Use the linear change of variable formula. If $T$ has gamma $(r, \lambda)$ distribution, with density
$$f_{r,\lambda}(t) = \frac{\lambda^r}{\Gamma(r)} t^{r-1} e^{-\lambda t}, t > 0,$$
then the density of $cT$ $(c > 0)$ is
$$f_{cT}(u) = \frac{1}{c} f_{r,\lambda}(u/c) = \frac{1}{c} \frac{\lambda^r}{\Gamma(r)} \frac{u^{r-1}}{c^{r-1}} e^{-\lambda u/c} = \frac{(\lambda/c)^r}{\Gamma(r)} u^{r-1} e^{-(\lambda/c)u}, u > 0.$$
So $cT$ has gamma$(r, \lambda/c)$ distribution. Apply this twice to see that $T$ has gamma $(r, \lambda)$ distribution iff $\lambda T$ has gamma $(r, 1)$ distribution.

3. The range of $Y = U^2$ is $(0, 1)$. The function $y = u^2$ is strictly increasing and has derivative $\frac{dy}{du} = 2u$ for $u \in (0, 1)$. So by the one to one change of variable formula for densities, the density of $Y$ is
$$f_Y(y) = f_U(u) \Big/ \left|\frac{dy}{du}\right| = \frac{1}{2u} = \frac{1}{2\sqrt{y}}, y \in (0, 1).$$

4. The density of $X$ is $f_X(x) = 1/2$ for $x \in (-1, 1)$, so by Example 5 the density of $Y$ is
$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} = \frac{\frac{1}{2} + \frac{1}{2}}{2\sqrt{y}} = \frac{1}{2\sqrt{y}}, y \in (0, 1).$$
Note: This distribution is called the beta (1/2, 1) distribution.

5. The range of $Y = X^2$ is $(0, 4)$. The density of $X$ is $f_X(x) = 1/3$ for $x \in (-1, 2)$, so by Example 5 the density of $Y$ is
$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}}.$$
If $y \in (0, 1)$ then this simplifies to $\frac{\frac{1}{3}+\frac{1}{3}}{2\sqrt{y}} = \frac{1}{3\sqrt{y}}$.

   If $y \in (1, 4)$ then this simplifies to $\frac{\frac{1}{3}+0}{2\sqrt{y}} = \frac{1}{6\sqrt{y}}$.

   For other $y$ this simplifies to 0.

6. Notice that $Y = \tan\Phi$. Its range is $(-\infty, \infty)$. The function $y = \tan\phi$ is strictly increasing with derivative $\sec^2\phi$ for $\phi \in (-\pi/2, \pi/2)$. So by the one to one change of variable formula for densities, the density of $Y$ is
$$\begin{aligned} f_Y(y) &= f_\Phi(\phi) \Big/ \left|\frac{dy}{d\phi}\right| \\ &= \frac{1}{\pi} / \sec^2\phi \\ &= \frac{1}{\pi(1+y^2)}, y \in (-\infty, \infty). \end{aligned}$$
The Cauchy distribution is symmetric since $f_Y(y) = f_Y(-y)$.

   The expectation of a Cauchy random variable is undefined since
$$\begin{aligned} \int_{-\infty}^{\infty} \Big| y f_Y(y) \Big| dy &= \int_{-\infty} \frac{\text{[illegible]}}{\pi(1+y^2)} dy \\ &\geq \int_1^{\infty} \frac{1}{2\pi|y|} dy \\ &= \infty \end{aligned}$$
so the required integral does not converge absolutely.

7. If $U$ has uniform $(0,1)$ distribution, then $\pi U$ has uniform $(0,\pi)$ distribution, so $\pi U - \pi/2$ has uniform $(-\frac{\pi}{2},\frac{\pi}{2})$ distribution. (See Example 1.) Now apply Exercise 6.

8. a) Apply the many-to-one change of variable formula for densities: If $z = g(y) = \frac{1}{1+y^2}$, then the density $f_Z(z)$ of $Z = g(Y)$ is given by

$$\begin{aligned}
f_Z(z) &= \sum_{\{y:g(y)=z\}} f_Y(y) \Big/ \left|\frac{dz}{dy}\right| \\
&= \sum_{\{y=\pm\sqrt{\frac{1}{z}-1}\}} f_Y(y) \Big/ \left| -2y/(1+y^2)^2 \right| \\
&= \frac{f_Y\left(-\sqrt{\frac{1}{z}-1}\right) + f_Y\left(\sqrt{\frac{1}{z}-1}\right)}{2z\sqrt{z-z^2}} \\
&= \frac{1}{\pi\sqrt{z(1-z)}}.
\end{aligned}$$

So $c = 1/\pi$ , $\alpha = -1/2$ , $\beta = -1/2$.

b)

$$P(Z \le x) = \int_0^x \frac{dz}{\pi\sqrt{z(1-z)}}$$

Now if we let $z = \sin^2(u)$ then $dz = 2\sin(u)\cos(u)du$ and we have

$$\begin{aligned}
P(Z \le x) &= \frac{1}{\pi}\int_0^{\arcsin\sqrt{x}} \frac{2\sin(u)\cos(u)du}{\sqrt{\sin^2(u)(1-\sin^2(u))}} \\
&= \frac{1}{\pi}\int_0^{\arcsin\sqrt{x}} \frac{2\sin(u)\cos(u)du}{\sqrt{\sin^2(u)\cos^2(u)}} \\
&= \frac{2}{\pi}\int_0^{\arcsin\sqrt{x}} du \\
&= \frac{2}{\pi}\arcsin\sqrt{x}
\end{aligned}$$

NOTE: using a different trigonometric substitution will give answer that may look very different. For example, using $z = \frac{1}{2} + \frac{1}{2}\sin u$ gives $P(Z \le x) = \frac{1}{\pi}\arcsin(2x-1) + \frac{1}{2}$. It is true, though not obvious, that this is equal to $\frac{2}{\pi}\arcsin\sqrt{x}$.

c)

$$\begin{aligned}
E(Z) &= E\left(\frac{1}{1+Y^2}\right) \\
&= \frac{2}{\pi}\int_0^\infty \frac{dy}{(1+y^2)^2} \\
&= \frac{2}{\pi}\int_0^{\pi/2} \cos^2\theta\, d\theta \\
&= 1/2
\end{aligned}$$

(substitute $y = \tan\theta$); this is obvious by the symmetry of the density of $Z$.

d)

$$\begin{aligned}
E(Z^2) &= E\left(\frac{1}{(1+Y^2)^2}\right) \\
&= \frac{2}{\pi}\int_0^\infty \frac{dy}{(1+y^2)^3}
\end{aligned}$$

$$= \frac{2}{\pi}\int_0^{\pi/2} \cos^4\theta d\theta$$
$$= 3/8$$

(substitute $y = \tan\theta$); so $Var(Z) = 1/8$.

9. Let $\alpha > 0$.

a) The range of $Y = T^\alpha$ is $(0,\infty)$. The function $y = t^\alpha$ is strictly increasing and has derivative $\frac{dy}{dt} = \alpha t^{\alpha-1}$ for $t > 0$. So by the one to one change of variable formula for densities, the density of $Y$ is

$$f_Y(y) = f_T(t)\Big/\left|\frac{dy}{dt}\right|$$
$$= \lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha}\Big/\alpha t^{\alpha-1}$$
$$= \lambda e^{-\lambda t^\alpha}$$
$$= \lambda e^{-\lambda y}, y > 0.$$

b) By Example 4, $X = -\lambda^{-1}\log U$ has exponential $(\lambda)$ distribution. So it suffices to show that $T = X^{1/\alpha}$ has Weibull $(\lambda,\alpha)$ distribution. $T$ has range $(0,\infty)$. The function $t = x^{1/\alpha}$ is strictly increasing, so by the change of variable formula, the density of $T$ is

$$f_T(t) = f_Y(y)\Big/\left|\frac{dt}{dx}\right|$$
$$= \lambda e^{-\lambda x}\Big/\frac{1}{\alpha}x^{(1/\alpha)-1}$$
$$= \lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha}, \quad t > 0.$$

10. a) The range of $Y = |Z|$ is $(0,\infty)$. Use the change of variable formula for many to one functions to get the density of $Y$: for $y > 0$

$$f_Y(y) = \sum_{z:|z|=y} f_Z(z) = f_Z(y) + f_Z(-y).$$

Here $f_Z(z) = \phi(z)$ is symmetric, so

$$f_Y(y) = 2\phi(y) = \sqrt{\frac{2}{\pi}}e^{-y^2/2}, y > 0.$$

b) The range of $Y = Z^2$ is $(0,\infty)$. By Example 5, the density of $Y$ is, for $y > 0$,

$$f_Y(y) = \frac{f_Z(\sqrt{y}) + f_Z(-\sqrt{y})}{2\sqrt{y}}$$
$$= \frac{\phi(\sqrt{y})}{\sqrt{y}}$$
$$= \frac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2}.$$

c) The range of $Y = 1/Z$ is $(-\infty,0)\cup(0,\infty)$. If $z \neq 0$ then the function $y = 1/z$ is strictly decreasing with derivative $\frac{dy}{dz} = -\frac{1}{z^2}$. So by the one to one change of variable formula, the density of $Y$ is

$$f_Y(y) = f_Z(z)\Big/\left|\frac{dy}{dz}\right|$$
$$= \phi(1/y)/y^2$$
$$= \frac{1}{\sqrt{2\pi}}y^{-2}e^{-\frac{1}{2y^2}}, y \neq 0.$$

d) The range of $Y = 1/Z^2$ is $(0, \infty)$. By Example 5, the density of $Y = (1/Z)^2$ is, for $y > 0$,

$$\begin{aligned} f_Y(y) &= \frac{f_{1/Z}(\sqrt{y}) + f_{1/Z}(-\sqrt{y})}{2\sqrt{y}} \\ &= \frac{f_{1/Z}(\sqrt{y})}{\sqrt{y}} \\ &= \frac{1}{\sqrt{2\pi}} y^{-3/2} e^{-\frac{1}{2y}}. \end{aligned}$$

11. Duplicate the solution to Problem 1 of Example 7 for a sphere of radius $r$:

$$P(\Theta \in d\theta) = \frac{2\pi r \cos\theta r d\theta}{A}$$

where $A$ = total area. Integrate from $\theta = -\pi/2$ to $\pi/2$ to get

$$\begin{aligned} A &= 2\pi r^2 \int_{-\pi/2}^{\pi/2} \cos\theta d\theta \\ &= 4\pi r^2. \end{aligned}$$

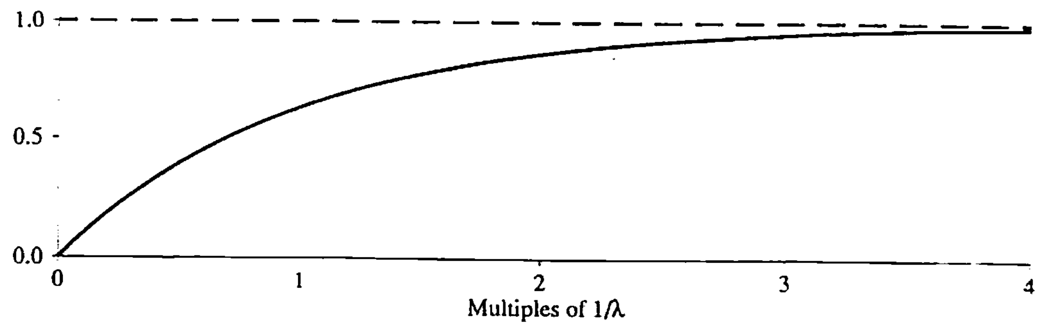
Or: If we accept that the area of a sphere between two parallel planes is proportional to the distance between the planes, then for a sphere of radius $r$, the surface area between two parallel planes a distance $\Delta$ apart must be $2\pi r\Delta$ (take the planes very close together and near the equator). So the total surface area of a sphere of radius $r$ is $2\pi r \cdot 2r = 4\pi r^2$.

# Section 4.5

**1.** a) If $x \geq 0$ then $F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$.

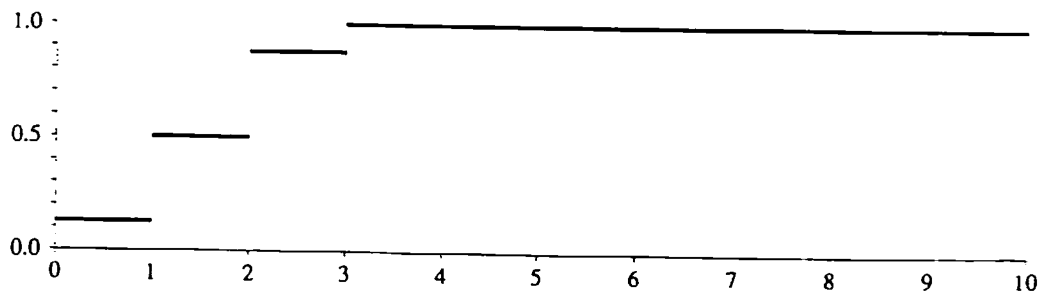
If $x < 0$ then $F(x) = 0$.

b)




**2.** a)

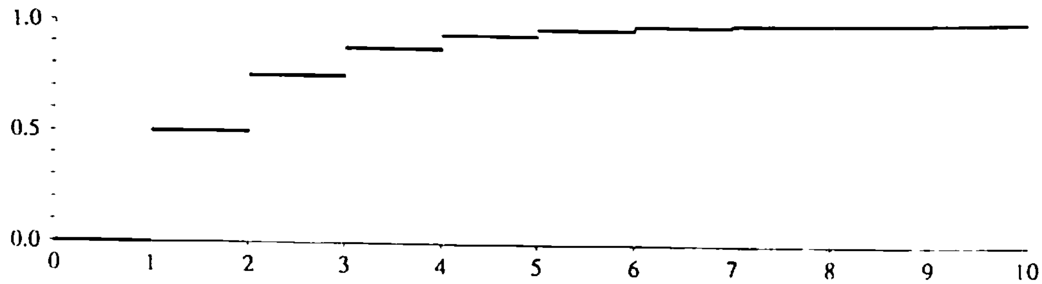
| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $F_X(x)$ | 1/8 | 1/2 | 7/8 | 1 |




b) Since $P(k) = (1/2)^k$ for $k = 1, 2, 3, \ldots$, we have:

If $x \geq 1$ then $F(x) = \sum_{k \leq x} P(k) = \sum_{k=1}^{int\,(x)} (1/2)^k = 1 - (1/2)^{int\,(x)}$ ;

If $x < 1$ then $F(x) = 0$.



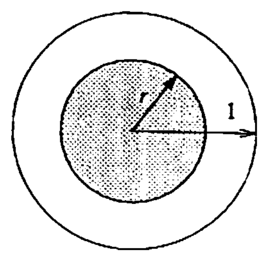
**3.** a) $Y$ has the same distribution as $X$, so

$$f_Y(y) = \begin{cases} \frac{2}{\pi}\sqrt{1-y^2} & |y| \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & y < -1 \\ \frac{1}{2} + \frac{1}{\pi}\left[y\sqrt{1-y^2} + arcsin y\right] & |y| \le 1 \\ 1 & y > 1 \end{cases}$$

b) If $0 < r < 1$ then

$$\begin{aligned} F_R(r) &= P(R \le r) \\ &= \frac{\text{Area inside radius } r}{\text{Area of circle}} \\ &= \frac{\pi r^2}{\pi} \\ &= = r^2 \end{aligned}$$




Differentiate to get $f_R(r) = \frac{d}{dr}F_R(r) = 2r$.

**4.** If $a > 0$, then

$$\begin{aligned} F_{aX+b}(y) &= P(aX + b \le y) \\ &= P\left(X \le \frac{y-b}{a}\right) \\ &= F_X\left(\frac{y-b}{a}\right) \end{aligned}$$

If $a < 0$, then

$$\begin{aligned} F_{aX+b}(y) &= P(aX + b \le y) \\ &= P\left(X \ge \frac{y-b}{a}\right) \\ &= 1 - F_X\left(\frac{y-b}{a}\right) \end{aligned}$$

assuming $F_X(x)$ is a continuous function of $x$.

**5.** If $x \le 0$ then

$F_X(x) = \int_{-\infty}^{x} \frac{1}{2}e^{-|y|}dy = \int_{-\infty}^{x} \frac{1}{2}e^{y}dy = \frac{1}{2}e^{x}$.

If $x \ge 0$ then

$F_X(x) = F_X(0) + P(0 < X \le x) = \frac{1}{2} + \int_0^x \frac{1}{2}e^{-|y|}dy = \frac{1}{2} + \frac{1}{2}(1 - e^{-x}) = 1 - \frac{1}{2}e^{-x}$.

**6.** a) $P(X \ge 1/2) = 1 - F(1/2) = 7/8$ .

b) $f(x) = \frac{d}{dx}F(x) = \begin{cases} 0 & x \le 0 \\ 3x^2 & 0 \le x < 1 \\ 0 & x \ge 1 \end{cases}$ .

c) $E(X) = \int x f(x)dx = \int_0^1 x3x^2dx = \int_0^1 3x^3dx = 3/4$ .

d) Let $Y_1$, $Y_2$, $Y_3$ be independent uniform $(0, 1)$ random variables. Then for $i = 1, 2, 3$

$$P(Y_i \le x) = \begin{cases} 0 & x \le 0 \\ x & 0 \le x \le 1 \\ 1 & x \ge 1 \end{cases}$$

so if $X = max(Y_1, Y_2, Y_3)$ , then

$$\begin{aligned} P(X \le x) &= P(Y_1 \le x, Y_2 \le x, Y_3 \le x) \\ &= [P(Y_1 \le x)]^3 \\ &= \left\{ \begin{matrix} 0 & x \le 0 \\ x^3 & 0 \le xl \\ 1 & x \ge 1 \end{matrix} \right\} \\ &= F(x). \end{aligned}$$

7. $f_Y(y)dy = f_T(t)dt$ where $t = y^2$

a) $f_Y(y) = \lambda e^{-\lambda y^2} 2y,\ y \ge 0$

b) $EY = \int_0^\infty t^{1/2} \lambda e^{-\lambda t} dt = \lambda \int_0^\infty t^{3/2-1} e^{-\lambda t} dt = \frac{\lambda \Gamma(3/2)}{\lambda^{3/2}} = \frac{\frac{1}{2}\sqrt{\pi}}{\sqrt{\lambda}} = 0.51$ when $\lambda = 3$.

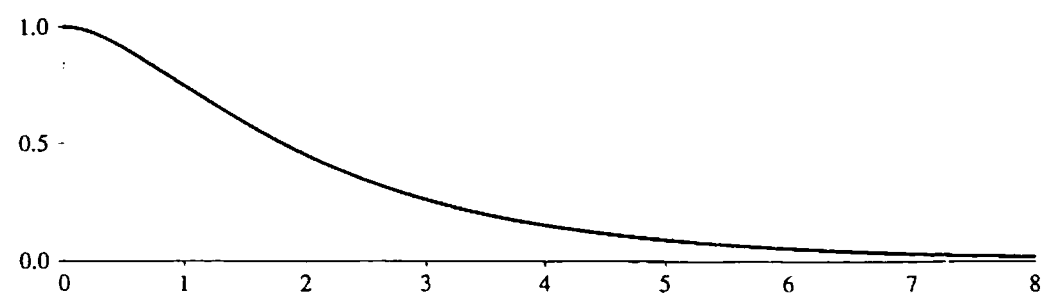
c) Find the inverse c.d.f. $P(Y \le y) = P(T \le y^2) = 1 - e^{-\lambda y^2} = u$, say. So $1 - u = e^{-\lambda y^2}$ so $y = \sqrt{\frac{\log(1-u)}{-\lambda}}$

$$Y = \sqrt{\frac{\log(1-u)}{-\lambda}}$$

8. Let $L_i$ denote the lifetime component $i$, and let $L$ denote the lifetime of the entire system.

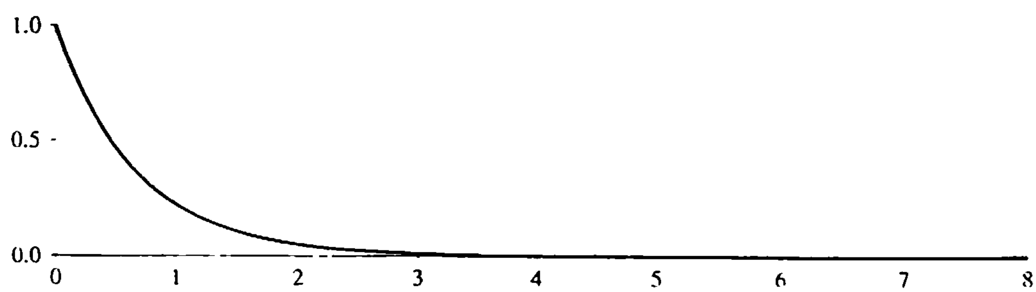
a) The system fails when and only when both components fail, so $L = \max(L_1, L_2)$. If $t > 0$ then

$$\begin{aligned} P(L > t) &= 1 - P(\max(L_1, L_2) \le t) \\ &= 1 - P(L_1 \le t, L_2 \le t) \\ &= 1 - (1 - e^{-t/\mu_1})(1 - e^{-t/\mu_2}). \end{aligned}$$



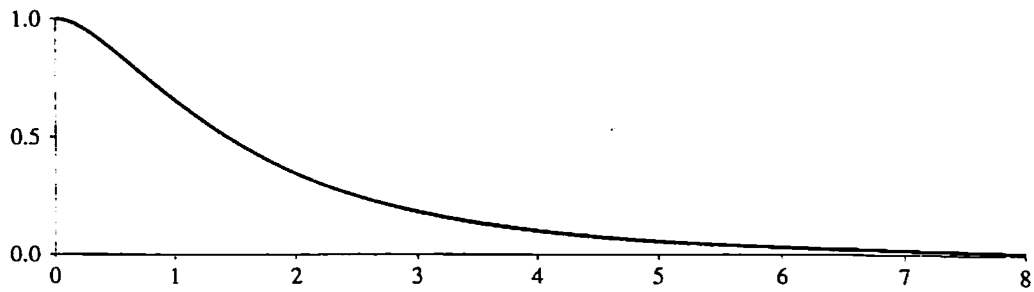

b) Here $L = \min(L_1, L_2)$. By Example 3, $L$ has exponential distribution with rate $1/\mu_1 + 1/\mu_2$. So if $t > 0$ then

$$P(L > t) = e^{-(1/\mu_1 + 1/\mu_2)t}.$$



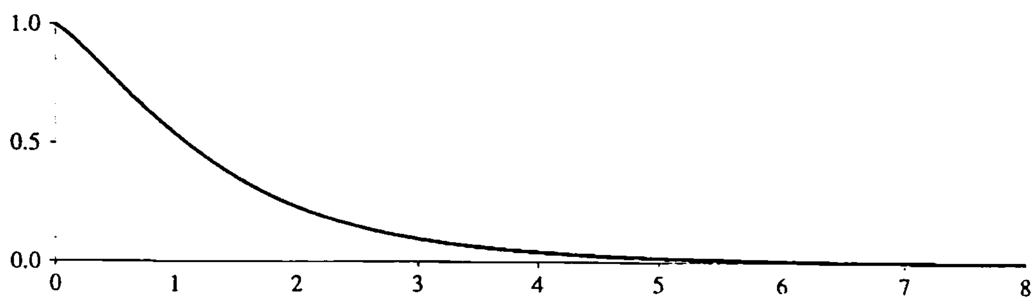
c) Here $L = \max(L_{top}, L_{bottom})$, where, by Example 3, $L_{top}$ has exponential distribution with rate $(1/\mu_1 + 1/\mu_2)$ and $L_{bottom}$ has exponential distribution with rate $(1/\mu_3 + 1/\mu_4)$. Put $a = 1/\mu_1 + 1/\mu_2$ and $b = 1/\mu_3 + 1/\mu_4$. By part a), $L$ has survival function

$$P(L > t) = 1 - (1 - e^{-at})(1 - e^{-bt}), t > 0.$$




d) Here $L = \min(\max(L_1, L_2), L_3)$. If $t > 0$ then

$$\begin{aligned} P(L > t) &= P(\max(L_1, L_2) > t, L_3 > t) \\ &= P(\max(L_1, L_2) > t)\, P(L_3 > t) \\ &= \left(1 - (1 - e^{-t/\mu_1})(1 - e^{-t/\mu_2})\right) e^{-t/\mu_3}. \end{aligned}$$




9. a) $X$ has the same distribution as $F^{-1}(U)$, so

$$\begin{aligned} E(X) &= E(F^{-1}(U)) \\ &= \int F^{-1}(y) f_U(y)\,dy \\ &= \int_0^1 F^{-1}(y)\,dy \\ &= \text{shaded area} \end{aligned}$$

(Integrate horizontal strips.) But the shaded area can also be represented, by integrating vertical strips, as

$$\int_0^\infty [1 - F(x)]\,dx = \int_0^\infty P(X > x)\,dx.$$

b) If $X$ is a discrete random variable with values $0, 1, 2, \ldots$, then

$$E(X) = \int_0^\infty P(X > x)\,dx$$

$$= \sum_{n=1}^{\infty} \int_{n-1}^{n} P(X > x)dx$$
$$= \sum_{n=1}^{\infty} P(X \geq n)$$

Or: the region above the distribution function of $X$ consists of rectangles of width 1 and height

$$\begin{aligned} 1 - F(0) &= P(X \geq 1), \\ 1 - F(1) &= P(X \geq 2), \\ 1 - F(2) &= P(X \geq 3), \\ &\text{etc.} \end{aligned}$$

c) If $X$ has exponential distribution with rate $\lambda$, then $X \geq 0$ so

$$E(X) = \int_0^{\infty} P(X > x)dx = \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}.$$

If $X$ has geometric $(p)$ distribution on $\{1, 2, 3, \ldots\}$ then by b)

$$E(X) = \sum_{n=1}^{\infty} P(X \geq n) = \sum_{n=1}^{\infty} q^{n-1} = \frac{1}{1-q} = \frac{1}{p}.$$

d) $X = X[I(X > 0) + I(X < 0)] = XI(X > 0) - (-X)I(X < 0) = X_+ - X_-$
where $X_+ = XI(X > 0)$ and $X_- = (-X)I(X < 0)$. So $E(X) = E(X_+) - E(X_-)$.
Since $X_+ \geq 0$,

$$\begin{aligned} E(X_+) &= \int_0^{\infty} P(X_+ > x)dx \\ &= \int_0^{\infty} P(X > x)dx \\ &= (+)\text{area in diagram} \end{aligned}$$

Since $X_- \geq 0$,

$$\begin{aligned} E(X_-) &= \int_0^{\infty} P(X_- > x)dx \\ &= \int_0^{\infty} P(X < -x)dx \\ &= \int_{-\infty}^{0} P(X < t)dt \\ &= (-)\text{area in diagram.} \end{aligned}$$

# Section 4.6

1. Let $X_1, X_2, X_3, X_4$ be the arrival times, in minutes, of persons 1, 2, 3, and 4, translated so that 12 noon $= 0$. Then $X_1, X_2, X_3, X_4$ are independent with common normal $(0, 5^2)$ distribution.

   a) $P(X_{(1)} < -10) = 1 - P(X_{(1)} \geq -10)$
   $= 1 - P(X_1 \geq -10, X_2 \geq -10, X_3 \geq -10, X_4 \geq -10)$
   $= 1 - [P(X_1 \geq -10)]^4$
   $= 1 - (.9772)^4 = .0881$
   since $P(X_1 \geq -10) = P(X_1/5 \geq -2) = \Phi(2) = .9772$.

   b) $P(X_{(4)} > 15) = 1 - P(X_{(4)} \leq 15)$
   $= 1 - P(X_1 \leq 15, X_2 \leq 15, X_3 \leq 15, X_4 \leq 15)$
   $= 1 - [P(X_1 \leq 15)]^4$
   $= 1 - (.9986)^4 = .0056$
   since $P(X_1 \leq 15) = P(X_1/5 \leq 3) = \Phi(3) = .9986$.

   c) Let $f$ and $F$ denote the common density and distribution function of the $X_i$'s. The desired probability is
   $$P(-1/6 \leq X_{(2)} \leq 1/6) = \int_{-1/6}^{1/6} f_{X_{(2)}}(x)dx,$$
   where $f_{X_{(2)}}$ is the density of the second order statistic:
   $$\begin{aligned} f_{X_{(2)}}(x) &= 4f(x)\binom{4-1}{2-1}[F(x)]^{2-1}[1-F(x)]^{4-2} \\ &= 12f(x)F(x)[1-F(x)]^2. \end{aligned}$$
   Over the interval $[-1/6, 1/6]$, this density is roughly constant, so the desired probability is approximately
   $$\begin{aligned} (\text{length of interval}) \times (\text{value of } f_{X_{(2)}} \text{ at } 0) &= \frac{1}{3} f_{X_{(2)}}(0) \\ &= 4f(0)F(0)[1-F(0)]^2 \\ &= 4 \cdot \frac{1}{5\sqrt{2\pi}} \cdot (1/2) \cdot (1/2)^2 \\ &= \frac{1}{10\sqrt{2\pi}} \approx .0399. \end{aligned}$$

   **Remark.** To compute the probability exactly, you may integrate by parts. Or you may argue as follows: Let $N_1$ denote the number of $X$'s which are $\leq 1/6$, and let $N_2$ denote the number of $X$'s which are $< -1/6$. Then $N_1$ has binomial $(4, F(1/6))$ distribution, and $N_2$ has binomial $(4, F(-1/6))$ distribution. Now argue that
   $$P(-1/6 \leq X_{(2)} \leq 1/6) = P(X_{(2)} \leq 1/6) - P(X_{(2)} < -1/6) = P(N_1 \geq 2) - P(N_2 \geq 2).$$

2. b) Let $k \geq 1$ integer. If $X$ has beta distribution with integer parameters $r, s$, then
   $E(X^k) = \int_0^1 x^k \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}dx$
   $= \frac{1}{B(r,s)} \int_0^1 x^{r+k-1}(1-x)^{s-1}dx$
   $= \frac{1}{B(r,s)} B(r+k, s)$
   $= \frac{(r+k-1)!(s-1)!}{(r+k+s-1)!} / \frac{(r-1)!(s-1)!}{(r+s-1)!}$
   $= \frac{(r+k-1)!/(r-1)!}{(r+s+k-1)!/(r+s-1)!}$
   $= \frac{r(r+1)\cdots(r+k-1)}{(r+s)(r+s+1)\cdots(r+s+k-1)}$.

a) Hence $Var(X) = E(X^2) - [E(X)]^2$
$= \frac{r(r+1)}{(r+s)(r+s+1)} - \left(\frac{r}{r+s}\right)^2$
$= \frac{r}{r+s} \cdot \frac{(r+1)(r+s)-r(r+s+1)}{(r+s+1)(r+s)}$
$= \frac{rs}{(r+s)^2(r+s+1)}$.

3. a) $(y-x)^n$
b) $(1-x)^n - (y-x)^n$
c) $y^n - (y-x)^n$
d) $1-(1-x)^n - y^n + (y-x)^n$
e) $\binom{n}{k}x^k(1-y)^{n-k}$
f) $\binom{n}{k+1}x^{k+1}(1-y)^{n-k-1} + \binom{n}{k}x^k(1-y)^{n-k} + \frac{n!}{(k)!1!(n-k)!}x^k(y-x)(1-y)^{n-k-1}$

4. a) $P(Z=1) = P(Z=0) = 1/2$; b) yes, yes, yes; c) the $n!$ possible orders of the $n$ variables are equally likely, independent of the $n$ order statistics.

5. Let $F$ be the common c.d.f. of the $X$'s.

a) $X_{(k)}$ is less than or equal to $x$ if and only if at least $k$ of the $X$'s are less than or equal to $x$ (and the remainder are greater than $x$). Since each $X$ has chance $F(x)$ of being less than or equal to $x$, and the $X$'s are independent, it follows that the number of $X$'s which are less than or equal to $x$ is a binomial $(n, F(x))$ random variable, so

$$P(X_{(k)} \le x) = \sum_{i=k}^{n} \binom{n}{i}[F(x)]^i[1-F(x)]^{n-i}.$$

b) If $r, s$ are positive integers: The $r$th order statistic of $r+s-1$ independent uniform $(0,1)$ random variables has beta $(r,s)$ distribution. But by part (a), its c.d.f. must be

$$\sum_{i=r}^{r+s-1} \binom{r+s-1}{i}[F(x)]^i[1-F(x)]^{r+s-1-i}.$$

Finally note that the c.d.f. of a uniform (0, 1) random variable satisfies $F(x) = x$, $0 \le x \le 1$.

c) If $0 \le u \le 1$ then

$$\begin{aligned} f(u) &= \frac{1}{B(r,s)}u^{r-1}(-u+1)^{s-1} \\ &= \frac{1}{B(r,s)}u^{r-1}\sum_{i=0}^{s-1}\binom{s-1}{i}(-u)^i \\ &= \frac{1}{B(r,s)}\sum_{i=0}^{s-1}\binom{s-1}{i}(-1)^i u^{r-1}u^i. \end{aligned}$$

# Chapter 4: Review

1. $X_t$ has binomial $(n, e^{-\lambda t})$ distribution, so

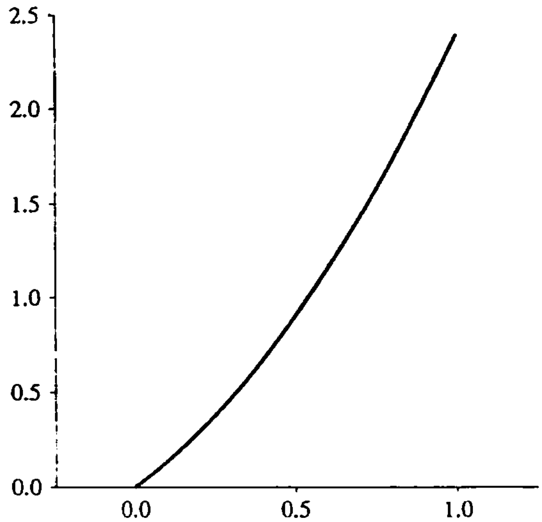
   a) $E(X_t) = ne^{-\lambda t}$

   b) $Var(X_t) = ne^{-\lambda t}(1 - e^{-\lambda t})$

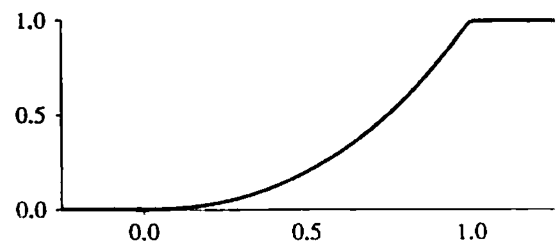
2. $\int_0^1 (x + x^2)dx = \frac{5}{6} \Longrightarrow c = \frac{6}{5}$.

   $$F(x) = c\int_{-\infty}^{x}(u + u^2)du = \begin{cases} 0 & x < 0 \\ 6/5(x^2/2 + x^3/3) & 0 \le x \le 1 \\ 1 & x > 1. \end{cases}$$

   $E(X) = 7/10$, $SD(X) = \sqrt{5}/10$.




Density




Distribution function

3. $X$ is the maximum of 3 independent uniform $(0, 1)$ variables:

   $$F_X(x) = P(X \le x) = P(Y_1 \le x, Y_2 \le x, Y_3 \le x) = \begin{cases} 0 & \text{if } x < 0 \\ x^3 & \text{if } 0 < x < 1 \\ 1 & \text{if } x > 1 \end{cases}$$

   $$f_X(x) = \frac{d}{dx}F_X(x) = \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

   $$E(X) = \int x f_X(x)dx = \int_0^1 x 3x^2 dx = \frac{3}{4}.$$

4. a) $P(X < 1) = \int_{-\infty}^{1} f(x)dx = \int_{-\infty}^{0}(1/2)e^x dx + \int_0^1 (1/2)e^{-x}dx = 1 - (1/2)e^{-1}$ .

   b) $E(X) = 0$ by symmetry;
   $Var(X) = E(X^2) = \int x^2 f(x)dx = 2\int_0^\infty x^2 \cdot (1/2)e^{-x}dx = \int_0^\infty x^2 e^{-x}dx = 2$ .

   c) $Y = X^2$ has range $[0, \infty)$. For each $y \ge 0$

   $$F_Y(y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) = 2P(0 \le X \le \sqrt{y})$$
   $$= 2\int_0^{\sqrt{y}} f(x)dx = 2\int_0^{\sqrt{y}} (1/2)e^{-x}dx = 1 - e^{-\sqrt{y}}.$$

5. a) $P(T > 30) = \frac{40}{100} = .4$.

b) If $0 < t \leq 30$, then $P(T > t) = \frac{100-2t}{100}$ . If $30 < t \leq 70$, then $P(T > t) = \frac{70-t}{100}$ .

c) If $0 < t \leq 30$, then $f_T(t) = 2/100$ . If $30 < t \leq 70$, then $f_T(t) = 1/100$ .

d) $E(T) = \int_0^{30} \frac{2t}{100} dt + \int_{30}^{70} \frac{t}{100} dt = 29$.

$E(T^2) = \int_0^{30} \frac{2t^2}{100} dt + \int_{30}^{70} \frac{t^2}{100} dt = \frac{3700}{3}$,

$Var(T) = E(T^2) - [E(T)]^2 = 392.33 \Longrightarrow SD(T) \approx 19.81$.

e) Let $l$ be the distance from one end of the road to the ambulance station, and $T_l$ be the response time. We may assume, without loss of generality, that $l \leq 50$. Then

$$P(T_l > t) = \begin{cases} 1 & t \leq 0 \\ \frac{100-2t}{100} & 0 < t \leq l \\ \frac{(100-l)-t}{100} & l < t \leq 100 - l \\ 0 & t > 100 - l \end{cases}$$

After a calculation similar to that of part d),

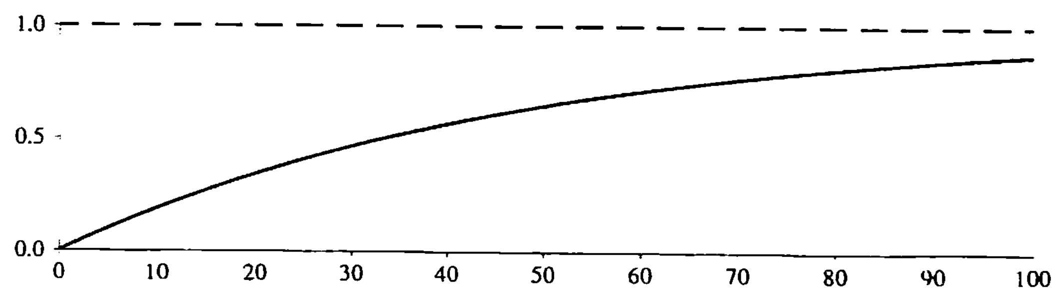
$E(T_l) = \frac{l^2}{100} - l + 50$,

which is minimized when $l = 50$. So the expected response time is minimized when the station is located at the midpoint of the road.

6. $T$ has exponential distribution with rate $\lambda = 1/48$.

a) The c.d.f. of $T$ is, for $t > 0$

$$F_T(t) = P(T \leq t) = 1 - P(T > t) = 1 - e^{-\lambda t} = 1 - e^{-t/48}.$$



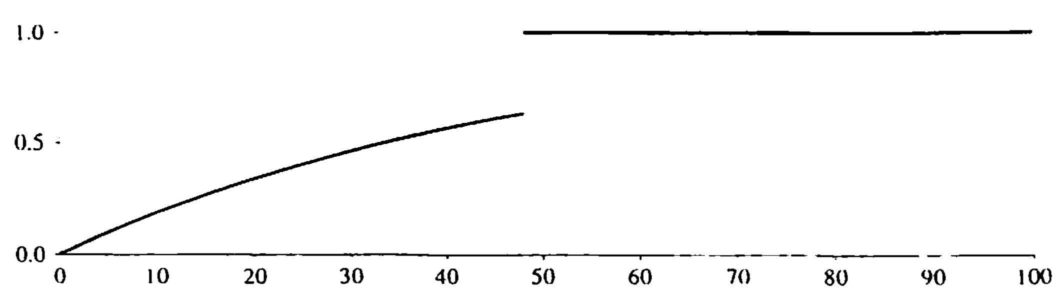

b) $U$ is distributed as $\min(T, 48)$. If $t \geq 48$ then

$$F_U(t) = P(\min(T, 48) \leq t) = 1$$

and if $0 \leq t < 48$ then

$$F_U(t) = P(\min(T, 48) \leq t) = P(T \leq t) = 1 - e^{-t/48}.$$

If $t < 0$ then $F_U(t) = 0$.
$U$ is neither discrete nor continuous.



c) $E(T) = 48$ by assumption, and

$$\begin{aligned} E(U) &= E\min(T,48) \\ &= \int_{t=0}^{\infty} \min(t,48) f_T(t)dt \\ &= \int_0^{48} t f_T(t)dt + \int_{48}^{\infty} 48 f_T(t)dt \\ &= \int_0^{48} t(\frac{1}{48}e^{-t/48})dt + 48P(T>48) \\ &= 48(1-2e^{-1}) + 48e^{-1} = 48(1-e^{-1}) = 30.34. \end{aligned}$$

d) No. The policy of replacement at 48 hours is unnecessary, since, by the memoryless property of the exponential distribution, the remaining time till failure of components which last beyond 48 hours has the same distribution as the lifetime of brand new components.

7. a) $1 = \int f(x)dx = \int_{-\infty}^{\infty} \alpha e^{-\beta|x|}dx = 2\alpha \int_0^{\infty} e^{-\beta x}dx = \frac{2\alpha}{\beta}$ . So $\alpha = \frac{\beta}{2}$.

   b) $E(X) = \int x f(x)dx = 0$ since $f$ is an even function;
   $Var(X) = E(X^2) = \frac{\beta}{2}\int_{-\infty}^{\infty} x^2 e^{-\beta|x|}dx = \frac{2}{\beta^2}$ .

   c) If $y > 0$ then $P(|X| > y) = 2P(X > y) = \int_y^{\infty} \beta e^{-\beta x}dx = e^{-\beta y}$.

   d) $P(X \le x) = \begin{cases} 1-(1/2)e^{-\beta x} & x \ge 0 \\ (1/2)e^{\beta x} & x < 0 \end{cases}$

8. a) $1/\int h(x)dx$ c) $1/\sqrt{2\pi}$, 0, 1 d) 2, 2/3, 1/18 e) 1/10, 5, 100/12 f) 5, 1/5, 1/25

9. a) normal, mean $a$, variance $\frac{1}{2}$, factor $= \frac{1}{\sqrt{\pi}}$

   b) normal, mean $a$, variance $b^2/2$, factor $= \frac{1}{\sqrt{\pi}b}$

   c) gamma $(6, a)$

$$\text{mean} = \frac{6}{a}, \quad \text{variance} = \frac{6}{a^2}, \quad \text{factor} = \frac{a^6}{5!}$$

   d) bilateral exponential, parameter $a$

$$\text{mean} = 0, \quad \text{variance} = \frac{2}{a^4}, \quad \text{factor} = \frac{1}{2a}$$

   e) beta $(8, 10)$

$$\text{mean} = \frac{8}{18}, \quad \text{variance} = \frac{80}{18^2 \times 19}, \quad \text{factor} = \frac{17!}{7!9!}$$

   f)

$$\begin{aligned} \int_0^b x^7(b-x)^9dx &= \int_0^1 b^7u^7(b-bu)^9 b\,du, \quad u = \tfrac{x}{b}, \quad x = bu, \quad dx = b\,du \\ &= \int_0^1 b^7u^7b^9(1-u)^9 b\,du = b^{17}\int_0^1 u^7(1-u)^9du \end{aligned}$$

$$\text{mean} = \frac{8b}{18}, \quad \text{variance} = \frac{80b^2}{18^2 \times 19}, \quad \text{factor} = (\frac{17!}{7!9!})\frac{1}{b^{17}}$$

10. Note that $(1/\sqrt{\pi})e^{-x^2}$ is the density function of a normal $(0, 1/2)$ random variable, say $X$.

   a) $\int_0^{\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}}e^{-x^2}dx = \frac{\sqrt{\pi}}{2}\cdot 1 = .886.$

   b) $\int_0^1 e^{-x^2}dx = \sqrt{\pi}P(0 < X < 1) = \sqrt{\pi}\left[\Phi(\sqrt{2}) - \Phi(0)\right] = .746.$

   c) $\int_0^{\infty} xe^{-x^2}dx = -\frac{1}{2}e^{-x^2}\Big|_0^{\infty} = \frac{1}{2}.$

   d) $\int_0^{\infty} x^2e^{-x^2}dx = \frac{\sqrt{\pi}}{2}\int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{\pi}}e^{-x^2}dx = \frac{\sqrt{\pi}}{2}Var(X) = \frac{\sqrt{\pi}}{4} = .433.$

11. a) 1/2

b) Compare with the gamma $(r, \lambda)$ density: if $r > 0, \lambda > 0$

$$1 = \int_0^\infty \frac{\lambda^r}{\Gamma(r)} t^{r-1} e^{-\lambda t} dt \iff \int_0^\infty t^{r-1} e^{-\lambda t} dt = \frac{\Gamma(r)}{\lambda^r}.$$

Put $r = 8, \lambda = 2$, obtain

$$\int_0^\infty t^7 e^{-2t} dt = \frac{\Gamma(8)}{2^8} = \frac{7!}{2^8}.$$

c) Compare with the beta function for positive integers: if $r$ and $s$ are positive integers, then

$$B(r, s) = \int_0^1 x^{r-1}(1-x)^{s-1} dx = \frac{(r-1)!(s-1)!}{(r+s-1)!}.$$

So

$$\int_0^{100} x^2(100-x)^2 dx = 100^5 \int_0^1 u^2(1-u)^2 du = 100^5 B(3,3) = 100^5 \frac{2!2!}{5!} = \frac{100^5}{30}.$$

12. $P(1 < T < 2) = P(T > 1) - P(T > 3) = P(N_1 \leq 2) - P(N_3 \leq 2)$
where $N_t$ denotes the number of hits in the time interval $[0, t]$. Since $N_t$ has Poisson $(2t)$ distribution, we have

$$P(N_1 \leq 2) = \sum_{k=0}^{2} e^{-2} \frac{2^k}{k!} = e^{-2}(1 + 2 + \frac{2^2}{2}) = 5e^{-2}$$

$$P(N_3 \leq 2) = \sum_{k=0}^{2} e^{-6} \frac{6^k}{k!} = e^{-6}(1 + 6 + \frac{6^2}{2}) = 25e^{-6}$$

and the desired probability is $5e^{-2} - 25e^{-4} = 0.6147$

13. a) $\frac{e^{-\lambda_{loc}}(\lambda_{loc})^5}{5!} \frac{e^{-\lambda_{dis}}(\lambda_{dis})^3}{3!}$ b) $e^{-3(\lambda_{loc}+\lambda_{dis})} \frac{[3(\lambda_{loc}+\lambda_{dis})]^{50}}{50!}$ c) $\left(\frac{\lambda_{loc}}{\lambda_{dis}+\lambda_{loc}}\right)^{10}$

14. a) This is equilavent to finding $P(T_4 > 2)$ where $T_r$ is gamma(4,3), so

$$P(T_4 > 2) = \sum_{k=0}^{3} e^{-6} \frac{6^k}{k!} = e^{-6}(1 + 6 + 18 + 36) = 0.1512$$

b) Because of the memoryless property of the Poisson process, $T_1$, $T_2 - T_1$, and $T_3 - T_2$ are all independent exponentials, so this is just $P(T_1 < 1)^3 = (1 - e^{-3})^3 = 0.858$

c) Given that there were 10 arrivals in the first 4 minutes, the times of the actual arrivals are uniform on (0,4), so the distribution of the number of arrivals in the first 2 minutes is binomial (10,.5).

15. $\frac{8!}{4!4!} \frac{3}{7}^4 \frac{4}{7}^4$

16. a) As soon as the first car comes, we are waiting for two more cars to come which gives us a gamma(2,3), call it $T_2$.

$$P(T_2 < 3) = 1 - \sum_{k=0}^{1} e^{-9} \frac{9^k}{k!} = 1 - e^{-9}(1 + 9) = 0.9988$$

b) In a ten-minute interval, the probability of 15 cars is $e^{-30} \frac{30^{15}}{15!} = 0.001027$. Given that there were 15 cars, the probability of 10 Japanese cars is $\binom{15}{10} .6^{10} .4^5 = 0.0007378$, assuming that the chance of being Japanese is independent from car to car. Further assuming that the number of cars which appear in a given time interval is independent of whether the cars are Japanese, the chance of both events occuring is $0.001027 \times 0.0007378 = 7.576 \times 10^{-7}$.

**17.** If $T$ has exponential distribution with rate $\lambda$, then for all $0 \le t < \infty$ we have

$$P(T \le t) = \int_{-\infty}^{t} f(u)du = \int_{0}^{t} \lambda e^{-\lambda u} du = 1 - e^{-\lambda t}.$$

Conversely, if $T$ is such that $P(T \le t) = 1 - e^{-\lambda t}$ for all $0 \le t < \infty$, then for $0 \le a < b < \infty$ we have

$$P(a < T \le b) = P(T \le b) - P(T \le a) = 1 - e^{-\lambda b} - (1 - e^{-\lambda a}) = e^{-\lambda a} - e^{-\lambda b}$$

**18.** a) The lines are independent, so we just multiply the probabilities of each of the events to get $e^{-(\lambda_A+\lambda_B+\lambda_C)}\frac{\lambda_A\lambda_B^2\lambda_C}{2}$

b) The overall rate for all buses is $\lambda_A + \lambda_B + \lambda_C$, so we have $e^{-2(\lambda_A+\lambda_B+\lambda_C)}\frac{(2(\lambda_A+\lambda_B+\lambda_C))^7}{7!}$

c) $P(A > t) = e^{-\lambda_A t}$

**19.** a) $(20-2)\log_2 10$
b) $20\log_2 10 - \log_e 2$

**20.** The assumption indicates $T = T_1 + T_2$, where $T_1$ and $T_2$ are independent exponential ($\lambda$) random variables. Hence $T$ has gamma $(2, \lambda)$ distribution.

a) $f(t) = \lambda^2 t e^{-\lambda t}, t \ge 0.$

b) $G(t) = (1 + \lambda t)e^{-\lambda t}, t \ge 0.$

c) $\text{hazard}(t) = -\frac{f(t)}{G(t)} = \frac{\lambda^2 t}{(1+\lambda t)}, t \ge 0.$

d) For $\lambda = 1$, $t = 2$ the hazard rate is $\frac{1^2 \times 2}{1+1\times 2} = \frac{2}{3}$ per hour.
Given survival to $t = 2$ hours, the probability of failure in the next minute (1/60 hour) is approximately
(hazard rate) $\times$ (length of interval) $= \frac{2}{3} \times \frac{1}{60} = \frac{1}{90}$.

**21.** $P(R_1 \le r) = 1 - e^{-\frac{1}{2}r^2}$

a) $f_Y(y) = 2ye^{-y^2}, y \ge 0$

b) exponential (1)

c) 1

**22.** a) The range of $Y$ is $[0, a/2]$.
If $0 < y < a/2$ then $P(Y \le y) = P(X \le y) = y/a$;
if $y \ge a/2$ then $P(Y \le y) = 1$.

b) Not continuous, since the distribution function is not a continuous function of $y$.

c) $E(Y) = E(\min(X, a/2)) = \int \min(x, a/2) f_X(x)dx = \int_0^{a/2} x \cdot \frac{1}{a}dx + \int_{a/2}^{a} \frac{a}{2} \cdot \frac{1}{a}dx = \frac{3}{8}a$ .

Or use $E(Y) = \int_0^\infty P(Y > y)dy = \int_0^{a/2}(1 - \frac{y}{a})dy = \frac{3}{8}a$.

**23.** a) $EM = E(M-3) + 3 = 5, \quad Var(M) = Var(M-3) = 4.$

b) $X = e^M = e^3 e^{M-3} = e^3 e^W$, where $W = M - 3$. So $X \ge e^3$ and

$$f_X(x) = \frac{1}{|\frac{dx}{dw}|}\frac{1}{2}e^{-\frac{1}{2}w} = \frac{1}{x}\frac{1}{2}e^{-\frac{1}{2}\log(xe^{-3})} = \frac{1}{x}\frac{1}{2}\left(xe^{-3}\right)^{-\frac{1}{2}} = \frac{1}{2}\left(\frac{e}{x}\right)^{\frac{3}{2}} \quad (x \ge e^3)$$

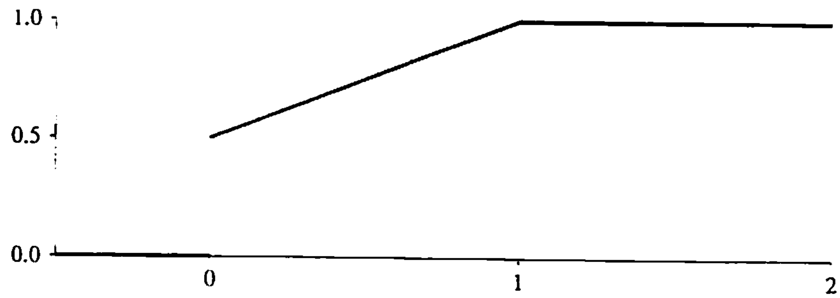
c) $P(M > 4) = P(M - 3 > 1) = e^{-\frac{1}{2}}$ so the probability is $\left(e^{-\frac{1}{2}}\right)^2 = e^{-1} = 0.3679$.

**24.** Let $Y$ denote the time elapsed in minutes between the instant the lights last turned red and the instant the car arrives at the lights. By assumption, $Y$ has uniform $(0, 2)$ distribution. And

$$X = \begin{cases} 1 - Y & \text{if } 0 < Y < 1 \\ 0 & \text{if } 1 < Y < 2. \end{cases}$$

a) The range of $X$ is $[0,1]$. If $0 < x < 1$ then

$$\begin{aligned} P(X \le x) &= P(X \le x, 0 < Y < 1) + P(X \le x, 1 < Y < 2) \\ &= P(1 - Y \le x, 0 < Y < 1) + P(0 \le x, 1 < Y < 2) \\ &= P(1 - x \le Y < 1) + P(1 < Y < 2) \\ &= \frac{1 - (1 - x)}{2} + \frac{2 - 1}{2} \\ &= \frac{1 + x}{2}. \end{aligned}$$




b) neither.

c) $X = g(Y)$ where $g(y) = 1 - y$ if $0 < y < 1$, $= 0$ if $1 < y < 2$. So

$$E(X) = E(g(Y)) = \int g(y) f_Y(y) dy = \int_0^1 (1 - y) \cdot \frac{1}{2} dy + 0 = \frac{1}{4};$$

Similarly $E(X^2) = \int_0^1 (1 - y)^2 \cdot \frac{1}{2} dy = \frac{1}{6}$, so $Var(X) = E(X^2) - [E(X)]^2 = \frac{5}{48}$ .

d) The total delay is $D = X_1 + X_2 + \cdots + X_{10}$, where $X_i$ is the delay at light $i$. Assume the $X$'s are independent. Then $E(D) = 10E(X_1) = 2.5$, and $Var(D) = 10Var(X_1) = 50/48$. So by the normal approximation,

$$P(D > 4) = P\left(\frac{D - 2.5}{\sqrt{50/48}} > \frac{4 - 2.5}{\sqrt{50/48}}\right) \approx 1 - \Phi(1.47) \approx 7\%.$$

**25.** a) $Y$ is uniform on $(0, 1/2]$.

b) uniform on $(0, 1]$

c) $EY = \frac{1}{4}$ $Var(Y) = \frac{1}{48}$.

**26.** a) $E(Xe^{tY}) = E(X)E(e^{tY}) = 2 \times \left[2 \times \int_1^{1.5} e^{tu} du\right] = \frac{4}{t}(e^{1.5t} - e^t)$ .

b) $E(W_t^2) = E(X^2)E(e^{2tY}) = \frac{5}{t}(e^{3t} - e^{2t})$.

$SD(W_t) = \sqrt{\frac{5}{t}(e^{3t} - e^{2t}) - \frac{16}{t^2}(e^{3t} + e^{2t} - 2e^{2.5t})}$.

**27.** a) For each $n \ge 2$ we have

$$\begin{aligned} P(U_1 \le u \text{ and } N = n) &= P(U_1 \le u \text{ and } N > n - 1) - P(U_1 \le u \text{ and } N > n) \\ &= P(U_{n-1} \le \cdots \le U_1 \le u) - P(U_n \le U_{n-1} \le \cdots \le U_1 \le u). \end{aligned}$$

Now for each $n \ge 1$ we have

$$P(U_n \le \cdots \le U_1 \le u) = \frac{1}{n!} P(U_1 \le u, U_2 \le u, \ldots, U_n \le u) = \frac{u^n}{n!}$$

since each of the $n!$ orderings of $U_1, \ldots, U_n$ is equally likely. Therefore

$$P(U_1 \le u \text{ and } N = n) = \frac{u^{n-1}}{(n-1)!} - \frac{u^n}{n!}.$$

b)

$$\begin{aligned}
P(U_1 \le u \text{ and } N \text{ is even}) &= \sum_{k=1}^{\infty} P(U_1 \le u \text{ and } N = 2k) \\
&= \sum_{k=1}^{\infty} \left( \frac{u^{2k-1}}{(2k-1)!} - \frac{u^{2k}}{(2k)!} \right) \\
&= \frac{u}{1!} - \frac{u^2}{2!} + \frac{u^3}{3!} - \frac{u^4}{4!} + \cdots \\
&= 1 - e^{-u}.
\end{aligned}$$

c) From a) with $u = 1$, $P(N \ge n) = 1/(n-1)!$ for $n \ge 2$, and this is true also for $n = 1$. So by the tail sum formula of Exercise 3.4.20

$$E(N) = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} = \sum_{k=0}^{\infty} \frac{1}{k!} = e$$

28. Let $\Theta \in [0, \pi]$ be the angle subtended by the (shorter) arc between the random point and the arbitrary fixed point. Then by assumption, $\Theta$ has uniform $(0, \pi)$ distribution. Observe that $\sin(\Theta/2) = X$ (or use the Law of Cosines); therefore $\Theta = 2\arcsin(X)$.

a) If $0 < x < 1$ then $P(X \le x) = P(\Theta \le 2\arcsin(x)) = \frac{2}{\pi}\arcsin(x)$.

b) $E(X) = E(\sin(\Theta/2)) = \frac{1}{\pi}\int_0^{\pi} \sin(\theta/2)d\theta = \frac{2}{\pi}$.

c) $E(X^2) = E(1 - \cos\Theta) = \frac{1}{\pi}\int_0^{\pi}(1 - \cos\theta)d\theta = 1$; $Var(X) = 1 - (\frac{2}{\pi})^2$.

29. a). Suppose $c > 0$ has been set, and you choose $b > 0$. Let $G$ be your net gain. Then

$$G = \begin{cases} 1 - cb & \text{if } X > 0 \\ -1 - cb & \text{if } X < 0 \end{cases}$$

so

$$\begin{aligned}
E(G) &= (1 - cb)P(X > 0) + (-1 - cb)P(X < 0) \\
&= P(X > 0) - P(X < 0) - cb \\
&= 2P(X > 0) - 1 - cb \\
&= 2\Phi(b) - 1 - cb
\end{aligned}$$

since $P(X > 0) = P\left(\frac{X-b}{1} > \frac{0-b}{1}\right) = \Phi(b)$.
So the option is advantageous when and only when there exists $b > 0$ such that $E(G) > 0$, i.e. when and only when there exists $b > 0$ such that

$$2\Phi(b) - 1 - cb > 0 \iff \Phi(b) > \frac{1 + cb}{2}.$$

Such a $b$ exists when and only when $\frac{c}{2} < \Phi'(0)$. [One way to see this: Plot $\Phi(b)$ and $\frac{1+cb}{2}$ as functions fo $b$.] Since $\Phi'(0) = \phi(0) = 1/\sqrt{2\pi}$, conclude that the option is advantageous to you when $c < \sqrt{\frac{2}{\pi}}$.
b). For such $c$, the expected net gain is maximized (why not minimized?) at $b$ satisfying

$$\frac{\partial}{\partial b}E(G) = 0 \iff \phi(b) = \frac{c}{2} \iff e^{-b^2/2} = \sqrt{\frac{\pi}{2}}c.$$

Note that the right side of the last equation is less than 1 (why?), so the equation has a solution $b = \sqrt{-2\log(\sqrt{\frac{\pi}{2}}c)}$.

30. a) Let $X$ be the diameter of a generic ball bearing. Then $X$ has normal distribution with mean $\mu = .250$ and standard deviation $\sigma = .001$, so the chance that the bearing meets specification (i) is

$$P(|X - \mu| < \sigma) = 0.6828.$$

We want $E(T_{16})$, where $T_{16}$ is the number of trials until the 16th success in Bernoulli($p = .6828$) trials; so $E(T_{16}) = 16/0.6828 \approx 23.4$.

b) Let $Y$ be the diameter of a ball bearing that meets specification (i). Then $Y$ has expectation

$$E(Y) = E(X||X-\mu| < \sigma) = E(\sigma Z + \mu||Z| < 1) = \mu$$

(where $Z = \frac{X-\mu}{\sigma}$ has standard normal distribution) and second moment

$$E(Y^2) = E\left(X^2||X-\mu| < \sigma\right) = E\left((\sigma Z+\mu)^2||Z| < 1\right) = \sigma^2 E(Z^2||Z| < 1) + \mu^2.$$

Therefore

$$Var(Y) = \sigma^2 E(Z^2||Z| < 1) \Longrightarrow SD(Y) = \sigma\sqrt{E(Z^2||Z| < 1)}.$$

But

$$\begin{aligned} E(Z^2||Z| < z) &= \frac{E(Z^2 I(|Z| < z))}{P(|Z| < z)} \\ &= \frac{\int_{-z}^{z} x^2\phi(x)dx}{P(|Z| < z)} \\ &= \frac{-2z\phi(z) + \int_{-z}^{z}\phi(x)dx}{P(|Z| < z)} \\ &= 1 - \frac{2z\phi(z)}{P(|Z| < z)}; \end{aligned}$$

hence $SD(Y) = \sigma\sqrt{1 - \frac{2\phi(1)}{P(|Z|<1)}} = 5.397 \times 10^{-4}$.
The required probability is

$$P(3.995 < Y_1 + \cdots + Y_{16} < 4.005)$$

where $Y_1, \ldots, Y_{16}$ are independent copies of $Y$. By the normal approximation, this is

$$\approx P\left(|Z| < \frac{4.005 - 16\mu}{4\sigma}\right) = P(|Z| < 2.316) \approx .98.$$

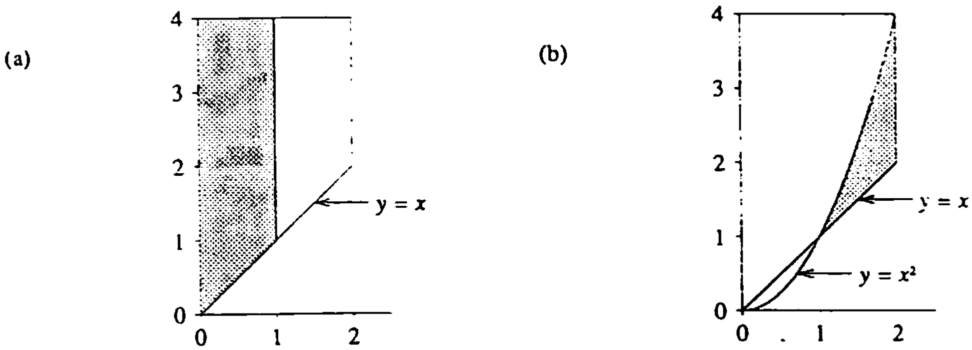
31. No Solution

# Section 5.1

1. The area of the set of interest is 6.

   a) $P(X < 1) = \frac{\text{indicated area}}{\text{area of set}} = \frac{3\frac{1}{2}}{6} = \frac{7}{12}$ . (Or divide the set into 12 triangles.)

   b) $P(Y < X^2) = \frac{\text{indicated area}}{\text{area of set}} = \frac{1}{6}\int_1^2 (x^2 - x)dx = \frac{5}{36}$ .

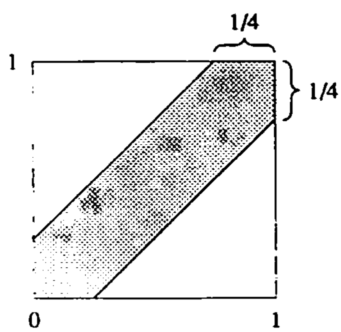



2. a) $P(\text{within 0.01 inches of } l) = \dfrac{0.02}{0.2} = 0.1$

   b) $P(\text{two measurements within 0.01 inches}) = 1 - \dfrac{2(1/2)(0.19)^2}{0.2^2} = 0.0975$

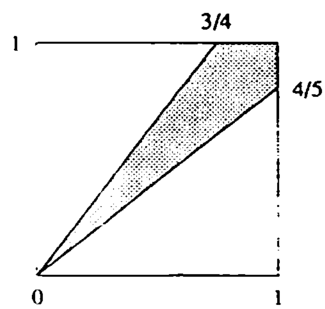
3.
$$\frac{3\frac{1}{2}}{6} = \frac{7}{12}$$

4. a)




$P(|X - Y| \leq .25)$
$= \text{indicated area}$
$= 1 - \left(\frac{3}{4}\right)^2$
$= \frac{7}{16}$

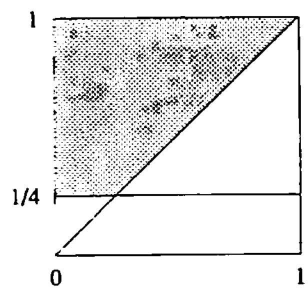
b)




$P\left(|\frac{X}{Y} - 1| \leq .25\right)$
$= P\left(\frac{4}{5}X \leq Y \leq \frac{4}{3}X\right)$
$= \text{indicated area}$
$= 1 - \frac{1}{2}\left(\frac{3}{4} + \frac{4}{5}\right)$
$= \frac{9}{40}.$

shaded region = $\{Y \geq X, Y \geq .25\}$

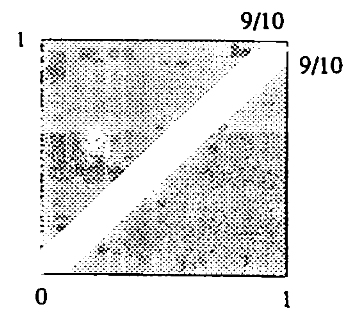
c)




$P(Y \geq \ldots | Y \geq .25)$
$= \text{indicated area}/\frac{3}{4}$
$= \left(\frac{1}{2} - \frac{1}{32}\right)/\frac{3}{4}$
$= \frac{5}{8}.$

5. a) The percentile rank $X$ of a student picked at random has uniform distribution on $(0,1)$, so $P(X > 0.9) = 0.1$ .

b) The percentile ranks $X, Y$ of students picked independently at random are independent uniform $(0,1)$ random variables, so
$P(|X - Y| > 0.1) = (9/10)^2 = \frac{81}{100}$ .

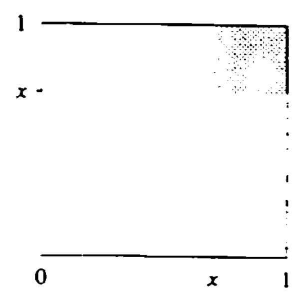



6. a) $P(\text{Jack arrives at least two minutes before Jill}) = \frac{(1/2)13^2}{15^2} \approx 0.376$

b) Let $F = \{\text{first person arrives before 12:05}\}$ and $L = \{\text{last person arrives after 12:10}\}$ Then,

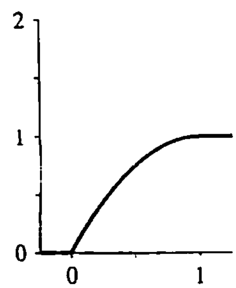
$$\begin{aligned} P(FL) &= 1 - P((FL)^c) = 1 - P(F^c \cup L^c) \\ &= 1 - P(F^c) - P(L^c) + P(F^c \cap L^c) \\ &= 1 - \left(\frac{10}{15}\right)^{10} - \left(\frac{10}{15}\right)^{10} + \left(\frac{5}{15}\right)^{10} \\ &= 0.965 \end{aligned}$$

7. a) $P(M \geq x) = \text{shaded area} = (1-x)^2$.

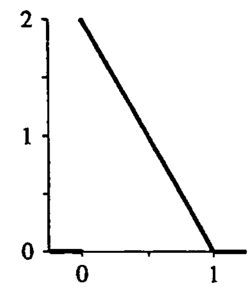


b) If $0 < x < 1$ then $P(M \le x) = 1 - P(M > x) = 1 - (1-x)^2$;
if $x \ge 1$ then $P(M \le x) = 1$;
if $x \le 0$ then $P(M \le x) = 0$.

So the density of $M$ is given by $f_M(x) = \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & \text{elsewhere.} \end{cases}$




Distribution function




Density

8. a) Let $U_1, \ldots, U_n$ be the unordered independent uniforms.

$$\begin{aligned} P(U_{(1)} > x \text{ and } U_{(n)} < y) &= P(x < U_1, U_2, \ldots U_n < y) \\ &= P(x < U_1 < y) \cdots P(x < U_n < y) \\ &= (y-x)^n \end{aligned}$$

b)

$$\begin{aligned} P(U_{(1)} \le x \text{ and } U_{(n)} < y) &= P(U_{(n)} < y) - P(U_{(1)} > x \text{ and } U_{(n)} < y) \\ &= y^n - (y-x)^n \end{aligned}$$

9. Standardise the length of the stick to be 1 (the solution clearly will not depend on the length of the stick!). Look at one end of the stick, and let $X$ and $Y$ denote the distances from that end of the stick to the break points. Then $X$ and $Y$ are independent uniform $(0,1)$ random variables. Let $L$ denote the minimum of $X$ and $Y$, and $R$ the maximum. ($L$ to suggest left, $R$ to suggest right.) Then the lengths of the broken pieces can be expessed as $L, R-L, 1-R$. To form a triangle, the maximum of these three should be less than the sum of the rest. That is,

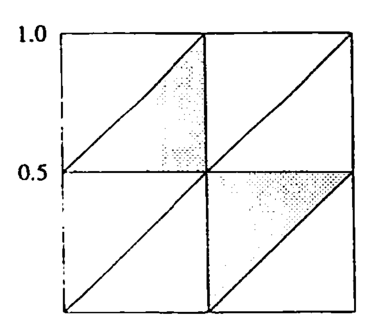
triangle $\iff \max\{L, R-L, 1-R\} < 1 - \max\{L, R-L, 1-R\}$

$\iff \max\{L, R-L, 1-R\} < 1/2$

$\iff L < 1/2$ and $R - L < 1/2$ and $1/2 < R$

$\iff$ ($X < 1/2$ and $Y - X < 1/2$ and $1/2 < Y$) or ($Y < 1/2$ and $X - Y < 1/2$ and $1/2 < X$).

Conclude: $P(\text{triangle}) = \dfrac{\text{shaded area}}{\text{total area}} = \frac{1}{4}$.



# Section 5.2

1. a) $f_{X,Y}(x,y) = \begin{cases} 1 & \text{if} 0 < |y| < x < 1 \\ 0 & \text{elsewhere.} \end{cases}$

   b) If $0 < x < 1$ then $f_X(x) = \int_{-x}^{x} 1dy = 2x$; otherwise $f_X(x) = 0$.

   If $0 < y < 1$ then $f_Y(y) = \int_{|y|}^{1} 1dx = 1 - |y|$; otherwise $f_Y(y) = 0$.

   c) No, because $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$.

   d) $E(X) = \int_0^1 2x^2dx = 2/3$, $E(Y) = 0$ by symmetry.

2. a) $f_{X,Y}(x,y) = \begin{cases} 1/2 & \text{if } 0 < |x| + |y| < 1 \\ 0 & \text{elsewhere} \end{cases}$

   b) If $|x| < 1$ then $f_X(x) = \int_{-(1-|x|)}^{1-|x|}(1/2)dy = 1 - |x|$. Otherwise $f_X(x) = 0$.

   By symmetry, $Y$ has the same marginal density as $X$:

   $f_Y(y) = \begin{cases} 1 - |y| & \text{if } |y| < 1 \\ 0 & \text{elsewhere.} \end{cases}$

   c) No, because $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$.

   d) $E(X) = E(Y) = 0$.

3. a) $1 = \int\int f(x,y)dxdy = \int_{y=0}^{1} dy \int_{x=0}^{1} c(x^2 + 4xy)dx = c\int_{y=0}^{1}(\frac{1}{3} + 2y)dy = \frac{4}{3}c$. So $c = 3/4$.

   b) If $0 < a < 1$, then

   $$P(X \leq a) = \int_{y=0}^{1} dy \int_0^a c(x^2+4xy)dx = c\int_{y=0}^{1}\left(\frac{a^3}{3} + 2a^2y\right)dy = c\left(\frac{a^3}{3} + a^2\right) = \frac{3}{4}\left(\frac{a^3}{3} + a^2\right).$$

   c) If $0 < b < 1$, then

   $$P(Y \leq b) = \int_{x=0}^{1} dx \int_0^b c(x^2 + 4xy)dy = c\int_{x=0}^{1}(x^2b + 2xb^2)dx = c\left(\frac{b}{3} + b^2\right) = \frac{3}{4}\left(\frac{b}{3} + b^2\right).$$

4. a)

   $$\begin{aligned} P(X \leq x, Y \leq y) &= \int_0^x \int_0^y 6e^{-2x-3y}dydx \\ &= \int_0^x 6e^{-2x}(-\frac{1}{3}e^{-3y})\Big|_0^y dx \\ &= (1 - e^{-3y})\int_0^x 2e^{-2x}dx \\ &= (1 - e^{-3y})(1 - e^{-2x}) \end{aligned}$$

   b)

   $$f_X(x) = \int_0^\infty 6e^{-2x-3y}dy = 2e^{-2x}$$

   c)

   $$f_Y(y) = \int_0^\infty 6e^{-2x-3y}dx = 3e^{-3y}$$

   d) Yes, they are independent since $f(x,y) = f_X(x)f_Y(y)$.

5. $\frac{\mu}{3\lambda+\mu}$

6. a)

$$\begin{aligned}
P(Y > 2X) &= \int_0^{.5} \int_{2x}^{1} 90(y-x)^8 dy dx \\
&= \int_0^{.5} 10(y-x)^9 \Big|_{2x}^{1} dx \\
&= \int_0^{.5} 10[(1-x)^9 - x^9] dx \\
&= -(1-x)^{10} - x^{10} \Big|_0^{.5} \\
&= 1 - (.5)^9 \\
&= \frac{511}{512}
\end{aligned}$$

b)

$$f_X(x) = \int_x^1 90(y-x)^8 dy = 10(y-x)^9 \Big|_x^1 = 10(1-x)^9$$

c) minimum, maximum

7. Without loss of generality, the circle has radius 1.

$$P(R_2 \le (1/2)R_1) = \int_0^1 \int_0^{r_1/2} (2r_1)(2r_2) dr_1\, dr_2 = \int_0^1 (r_1/2)^2 2r_1 dr_1 = \frac{1}{8}.$$

8. a)

$$\begin{aligned}
f_Y(y) &= \int_{-y}^{y} c(y^2 - x^2)e^{-y} dx = ce^{-y}\left[xy^2 - \frac{x^3}{3}\right]_{x=-y}^{y} \\
&= \frac{4c}{3} y^3 e^{-y} \quad (y > 0)
\end{aligned}$$

Therefore, $Y \sim$ gamma$(4,1)$ and

$$\frac{4c}{3} = \frac{1}{\Gamma(4)} = \frac{1}{6} \quad \text{so} \quad c = \frac{1}{8}$$

b) Let $Z = 4Y^3$.

$$\begin{aligned}
f_Z(z) &= \frac{1}{\left|\frac{dz}{dy}\right|} f_Y\left(\left(\frac{z}{4}\right)^{1/3}\right) \\
&= \frac{1}{12(\frac{z}{4})^{2/3}} \frac{1}{6} \frac{z}{4} e^{-(\frac{z}{4})^{1/3}} \\
&= \frac{z^{1/3} e^{-(z/4)^{1/3}}}{72 \cdot 4^{1/3}} \quad (z > 0)
\end{aligned}$$

c) Since given $Y = y$, $|X| \le y$,

$$E(|X|) = E\left[E\left(|X| \,\Big|\, Y\right)\right] \le E[Y] = 4$$

9. a) $2\lambda^2 e^{-\lambda(x+y)}$ $(0 < x < y)$, no; b) $2\lambda^2 e^{-2\lambda x - \lambda z}$ $(x > 0, z > 0)$, yes; c) $X$ is exponential $(2\lambda)$ and $Z$ is exponential $(\lambda)$.

10. For $0 < v < w$,

$$\begin{aligned} P(V \in dv, W \in dw) &= P(\text{no } X_i \text{ in } (0,v),\ 1 \text{ in } dv,\ n-2 \text{ in } (v,w),\ 1 \text{ in } dw) \\ &= n(n-1)P(X_1 \in dv, v < X_2, \ldots, X_{n-1} < w, X_n \in dw) \\ &= n(n-1)\left(\lambda e^{-\lambda v}dv\right)\left(e^{-\lambda v} - e^{-\lambda w}\right)^{n-2}\left(\lambda e^{-\lambda w}dw\right) \\ f(v,w) &= n(n-1)\lambda^2 e^{-\lambda(v+w)}\left(e^{-\lambda v} - e^{-\lambda w}\right)^{n-2} \quad (0 < v < w) \end{aligned}$$

11. a) $E(X+Y) = E(X) + E(Y) = (1/2) + 1 = 3/2$
b) $E(XY) = E(X)E(Y) = (1/2) \times 1 = 1/2$
c) $E[(X-Y)^2] = E(X^2) - 2E(X)E(Y) + E(Y^2)$ by linearity and independence
$= \frac{1}{3} - 2 \times \frac{1}{2} \times 1 + 2$ using $Var(X) = 1/12$, and $Var(Y) = 1$
$= 4/3$.
d) $E(e^{2Y}) = \int_0^\infty e^{2y} \cdot e^{-y}dy = \infty$. So $E(X^2e^{2Y}) = E(X^2)E(e^{2Y}) = \infty$.

12. For $0 < s < t$,

$$\begin{aligned} P(T_1 \in ds, T_5 \in dt) &= P(\text{no arrivals in } (0,s),\ 1 \text{ in } ds,\ 3 \text{ in } (s,t),\ 1 \text{ in } dt) \\ &= \left(e^{-\lambda s}\right)(\lambda ds)\left(\frac{e^{-\lambda(t-s)}[\lambda(t-s)]^3}{3!}\right)(\lambda dt) \\ f(s,t) &= \frac{1}{6}\lambda^5(t-s)^3e^{-\lambda t} \quad (0 < s < t) \end{aligned}$$

13. The distributions are all the same, with density $2(1-x)$ for $0 < x < 1$.

14. a) Let $V = \min(U_1, \ldots, U_5)$ and $W = \max(U_1, \ldots, U_5)$. Then $R = W - V$.

$$\begin{aligned} P(W \geq w) &= 1 - P(W < w) = 1 - w^5 \\ E(W) &= \int_0^1 P(W \geq w)dw = \int_0^1 1 - w^5 dw = \frac{5}{6} \\ P(V \geq v) &= (1-v)^5 \\ E(V) &= \int_0^1 P(V \geq v)dv = \int_0^1 (1-v)^5 dv = \frac{1}{6} \\ E(R) &= E(W) - E(V) = \frac{2}{3} \end{aligned}$$

b) For $0 < v < w < 1$,

$$\begin{aligned} P(V \in dv, W \in dw) &= P(\text{no } U_i \text{ in } (0,v),\ 1 \text{ in } dv,\ 3 \text{ in } (v,w),\ 1 \text{ in } dw) \\ &= 20P(U_1 \in dv, v < U_2, U_3, U_4 < w, U_5 \in dw) \\ &= 20dv(w-v)^3dw \\ f(v,w) &= 20(w-v)^3 \quad (0 < v < w < 1) \end{aligned}$$

c)

$$\begin{aligned} P(R > 0.5) &= \int_{0.5}^1 \int_0^{w-0.5} 20(w-v)^3 dv dw \\ &= 20\int_{0.5}^1 \left[-\frac{(w-v)^4}{4}\right|_{v=0}^{w-0.5} dw \\ &= 5\int_{0.5}^1 w^4 - (0.5)^4 dw \\ &= 5\left[\frac{w^5}{5} - w(0.5)^4\right|_{w=0.5}^1 \\ &= \frac{31}{32} \end{aligned}$$

15. a) $F(b,d) - F(a,d) - F(b,c) + F(a,c)$

b) $F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)dudv$.

c) $f(x,y) = \frac{d}{dx}\frac{d}{dy}F(x,y)$.

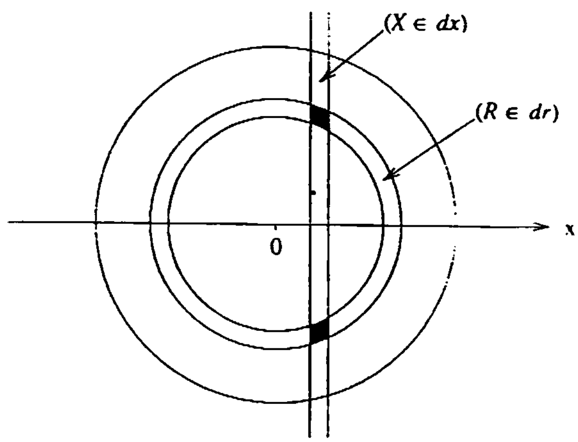
d) $F(x,y) = F_X(x)F_Y(y)$.

e) $F(x,y) = P(\min \le x \text{ and } \max \le y) = P(\max \le y) - P(\min > x \text{ and } \max \le y)$
$= y^n - (y-x)^n, \qquad 0 < x < y < 1.$
$f(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}F(x,y) = n(n-1)(y-x)^{n-2}, \qquad 0 < x < y < 1.$

16. Considering the three-dimensional space which corresponds to triplets $(X_1, X_2, X_3)$ we wish to find the volume which gives us $X_1 < X_2 < X_3$. This volume is given by:

$$
\begin{aligned}
P(X_1 < X_2 < X_3) &= \int_0^\infty \int_{x_1}^\infty \int_{x_2}^\infty f(x_3)f(x_2)f(x_1)dx_3dx_2dx_1 \\
&= \int_0^\infty \int_{x_1}^\infty \int_{x_2}^\infty \lambda_3 e^{-\lambda_3 x_3}\lambda_2 e^{-\lambda_2 x_2}\lambda_1 e^{-\lambda_1 x_1} dx_3dx_2dx_1 \\
&= \int_0^\infty \int_{x_1}^\infty e^{-\lambda_3 x_2}\lambda_2 e^{-\lambda_2 x_2}\lambda_1 e^{-\lambda_1 x_1} dx_2dx_1 \\
&= \int_0^\infty \int_{x_1}^\infty \lambda_2 e^{-(\lambda_3+\lambda_2)x_2}\lambda_1 e^{-\lambda_1 x_1} dx_2dx_1 \\
&= \int_0^\infty \frac{\lambda_1\lambda_2}{\lambda_2+\lambda_3} e^{-(\lambda_3+\lambda_2)x_1} e^{-\lambda_1 x_1} dx_1 \\
&= \frac{\lambda_1\lambda_2}{(\lambda_2+\lambda_3)(\lambda_1+\lambda_2+\lambda_3)}
\end{aligned}
$$

17. a) Note that $-R \le X \le R$. Let $0 \le r \le 1$, and $-r \le x \le r$. If $X \in dx$ and $R \in dr$, then the point $(X, Y)$ lies in one of two (almost) parallelograms:




So

$$
\begin{aligned}
P(X \in dx, R \in dr) &= \frac{2 \times \text{area of parallelogram}}{\text{area of circle}} \\
&= 2 \times dx \times \frac{dr}{\sqrt{1-(x/r)^2}} \Big/ \pi = \frac{2}{\pi}\frac{r}{\sqrt{r^2-x^2}}dxdr
\end{aligned}
$$

b) Note again that $-R \le X \le R$. Let $0 \le r \le 1$, and $-r \le x \le r$. If $X \in dx$ and $R \in dr$, then the point $(X, Y, Z)$ lies in an "inner tube" formed by rotating the parallelogram in (a) about the $x$-axis. Hence

$$P(X \in dx, R \in \quad = 2 \times \pi \times (\text{distance to } x\text{-axis}) \times (\text{area of parallelogram}) \Big/ (\text{volum} \quad \text{here})$$

$$= 2 \times \pi \times \sqrt{r^2 - x^2} \times \frac{dx\,dr}{\sqrt{1-(x/r)^2}} \Big/ \frac{4}{3}\pi = \frac{3}{2} r\,dx\,dr$$

So $f(x, r) = \frac{3}{2} r, \quad 0 \le r \le 1, \quad -r \le x \le r.$

18. a) By symmetry, you can argue that the probability must be $\frac{1}{2}$, since $P(X_1 = X_2) = 0$ if they are independent continuous random variables.

$$P(X_1 < X_2) = P(X_2 < X_1) \quad \text{and} \quad P(X_1 < X_2) + P(X_2 < X_1) = 1$$

Here is an alternative proof using an explicit calculation. Let $X_1$ and $X_2$ have range $(a, b)$. (Possibly $a = -\infty$ or $b = +\infty$.) Then,

$$\begin{aligned} P(X_1 < X_2) &= \int_a^b \int_a^{x_2} f(x_1) f(x_2) dx_1\, dx_2 \\ &= \int_a^b f(x_2)[F(x_2) - F(a)] dx_2 \\ &= \int_a^b f(x_2) F(x_2) dx_2 - 0 \\ &= E[F(X_2)] = \frac{1}{2} \end{aligned}$$

since $F(X_2)$ has the uniform(0, 1) distribution.

b) Again, using symmetry, you can argue that each of the 6 possible orders is equally likely, so the probability must be $\frac{1}{6}$.

19. a) Lon clearly has uniform distribution on $(-180, 180)$: the event $(\text{Lon} \in dx)$ has the same probability for all $x$. Hence

$$f_{Lon}(x) = \begin{cases} 1/360 & -180 < x < 180 \\ 0 & \text{otherwise.} \end{cases}$$

b) Let $\Theta$ be as in Example 4.4.7; then Lat $= \frac{180}{\pi}\Theta$. Since the density of $\Theta$ is

$$f_\Theta(\theta) = \begin{cases} (1/2)\cos\theta & -\pi/2 < \theta < \pi/2 \\ 0 & \text{otherwise,} \end{cases}$$

it follows by the linear change of variable formula that Lat has density

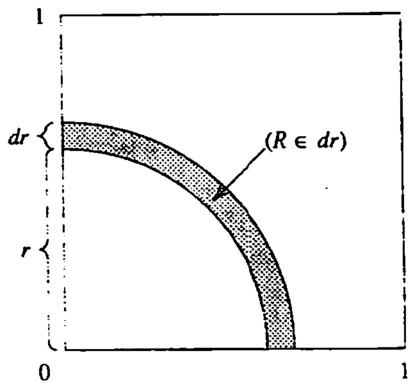
$$\begin{aligned} f_{Lat}(y) &= \frac{\pi}{180} f_\Theta\left(\frac{\pi}{180} y\right) \\ &= \begin{cases} \frac{\pi}{360}\cos\left(\frac{\pi}{180} y\right) & \text{if } -\frac{\pi}{2} < \frac{\pi}{180} y < \frac{\pi}{2} \iff -90 < y < 90 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

c) $P(\text{Lon} \in dx, \text{Lat} \in dy) = P(\text{Lon} \in dx \mid \text{Lat} \in dy) P(\text{Lat} \in dy) = P(\text{Lon} \in dx) P(\text{Lat} \in$ since Lon still has uniform distribution on $(-180, 180)$ given the value of Lat. Henc e joint density of (Lon, Lat) is

$$\begin{aligned} f(x, y) &= f_{Lon}(x) f_{Lat}(y) \\ &= \begin{cases} \frac{1}{360} \cdot \frac{\pi}{360}\cos\left(\frac{\pi}{180} y\right) & \text{if } -180 < x < 180 \quad \text{and} \quad -90 < y < 90 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

d) Yes, since the joint density equals the product of the marginal densities.

20. Since $(X, Y)$ has uniform distribution on the unit square, it follows that the probability that $(X, Y)$ lies in a given subset of the unit square is the area of that subset.




a) If $0 < r < 1$ then

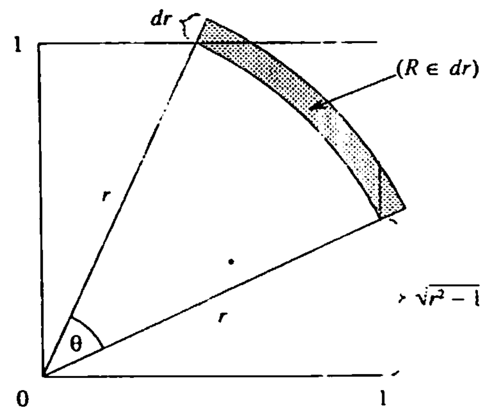
$$f_R(r)dr = P(R \in dr) = r \times \frac{\pi}{2} \times dr$$

(Recall that arc length = radius × angle subtended) while if $1 < r < \sqrt{2}$ then

$$f_R(r) = P(R \in dr) = r \times \theta \times dr,$$

where $\frac{\pi}{2} = \theta + 2\arccos\frac{1}{r}$ (arccos has range $[0, \pi]$ ). So

$$f_R(r) = \begin{cases} (\pi/2)r & 0 < r < 1 \\ r\,(\pi/2 - 2\arccos(1/r)) & 1 < r < \sqrt{2} \end{cases}$$




b) Integrate $f_R$, or observe that if $0 < r < 1$ then

$$F_R(r) = P(R \le r) = P((X, Y) \text{ is within } r \text{ of } (0,0)) = \frac{1}{4}\pi r^2$$

and if $1 < r < \sqrt{2}$ then

$$\begin{aligned} F_R(r) &= \text{area of sector of circle} + \text{area of 2 triangles} \\ &= \frac{\theta}{2\pi} \times \pi r^2 + 2 \times \frac{1}{2}\sqrt{r^2 - 1} \\ &= \left(\frac{\pi}{4} - \arccos\frac{1}{r}\right) r^2 + \sqrt{r^2 - 1}. \end{aligned}$$

c) The following inequality holds for each $a, b \geq 0$:

$$\sqrt{a^2+b^2} \geq \frac{a+b}{\sqrt{2}}$$

(Proof: Rearrange $(a-b)^2 \geq 0$ to get $2(a^2+b^2) \geq (a+b)^2$, then take square roots. Or: Let $W$ be a random variable taking values $a, b$ with probability $1/2$. Then $\frac{a^2+b^2}{2} = E(W^2) \geq (EW)^2 = (\frac{a+b}{2})^2$.) Apply the above inequality to $X$ and $Y$, take expectations to get

$$E(R) = E\sqrt{X^2+Y^2} \geq E\left(\frac{X+Y}{\sqrt{2}}\right) = \frac{E(X)+E(Y)}{\sqrt{2}} = \frac{1}{\sqrt{2}}.$$

In fact, the inequality is strict, since $P(\sqrt{X^2+Y^2} = \frac{X+Y}{\sqrt{2}}) = P(X = Y) = 0$, hence $R = \sqrt{X^2+Y^2} > \frac{X+Y}{\sqrt{2}}$ with probability 1, and so $E(R) > E(\frac{X+Y}{\sqrt{2}})$.
For the second inequality, note that $R > 0$ implies

$$(ER)^2 < E(R^2) = E(X^2+Y^2) = 2E(X^2) = \frac{2}{3}.$$

d)

$$\begin{aligned}
E(R) &= \int_0^1 \frac{\pi}{2} r^2 dr + \int_1^{\sqrt{2}} 2r^2 \left(\frac{\pi}{4} - \arccos\frac{1}{r}\right) dr \\
&= \frac{\pi}{2}\int_0^1 r^2 dr + \frac{\pi}{2}\int_1^{\sqrt{2}} r^2 dr - 2\int_1^{\sqrt{2}} r^2 \arccos\frac{1}{r} dr \\
&= \frac{\pi}{2}\int_0^{\sqrt{2}} r^2 dr - 2\int_{\theta=0}^{\pi/4} \theta \sec^3\theta \tan\theta d\theta (\theta = \arccos\frac{1}{r})
\end{aligned}$$

To evaluate the second integral, use integration by parts:

$$\begin{aligned}
\int \theta \sec^3 \tan\theta d\theta &= \int \frac{1}{3}\theta \cdot d(\sec^3\theta) \\
&= \frac{1}{3}\theta\sec^3\theta - \frac{1}{3}\int \sec^3\theta d\theta \\
&= \frac{1}{3}\theta\sec^3\theta - \frac{1}{3}\cdot\frac{1}{2}[\sec\theta\tan\theta + \log|\sec\theta+\tan\theta|]
\end{aligned}$$

Hence

$$\begin{aligned}
E(R) &= \frac{\pi}{2}\left\{\frac{(\sqrt{2})^3}{3}\right\} - 2\left\{\frac{1}{3}\cdot\frac{\pi}{4}(\sqrt{2})^3 - \frac{1}{6}\left[\sqrt{2}+\log(\sqrt{2}+1)\right]\right\} \\
&= \frac{1}{3}\left[\sqrt{2}+\log(\sqrt{2}+1)\right] = 0.7651957.
\end{aligned}$$

You can check that $\sqrt{\frac{1}{2}} < .765 < \sqrt{\frac{2}{3}}$.

21. a) $E(D_{corner}) = .765$, $E(D_{center}) = E(D_{corner})/2 = .3825$.

b) Let the two points be $(X_1, Y_1)$, and $(X_2, Y_2)$.

$$\begin{aligned}
E(D^2) &= E[(X_1-X_2)^2 + (Y_1-Y_2)^2] \\
&= 4E[(X_1)^2] - 4E(X_1X_2) \qquad \text{(by linearity and symmetry)} \\
&= 4\int_0^1 x^2 dx - 4E(X_1)E(X_2) \qquad \text{(by independence)} \\
&= 4\cdot 1/3 - 4\cdot 1/2\cdot 1/2 = 1/3
\end{aligned}$$

c) Since $Var(D) = E(D^2) - [E(D)]^2 \geq 0$, we have $E(D) \leq \sqrt{E(D^2)} = 1/\sqrt{3} = 0.577$.

d) If $\sigma^2 = Var(D)$ were known, then a 95% confidence interval for $E(D)$, based on $n$ independent observations of $D$, would be

$$\bar{D} \pm (1.96) \times \frac{\sigma}{\sqrt{n}}$$

where $\bar{D}$ = average of $n$ observations. Here $n = 10,000$, $\bar{D} = 0.5197$, and we estimate $\sigma$ by

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n} D_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} D_i\right)^2} = \sqrt{0.3310 - (0.5197)^2} = 0.2468.$$

So an approximate 95% confidence interval for $E(D)$ is

$$0.5197 \pm \frac{(1.96) \times (0.2468)}{\sqrt{10,000}} \iff 0.5197 \pm 0.0048.$$

# Section 5.3

1. Using the same notation as in Example 1,

   a) $P(R \le 1/2) = F_R(1/2) = 1 - e^{-\frac{1}{2}(\frac{1}{2})^2} = 1 - e^{-\frac{1}{8}} = .1175.$

   b) $\frac{1}{4}P(1 \le R \le 2) = \frac{1}{4}[F_R(2) - F_R(1)] = \frac{1}{4}\left[(1 - e^{-\frac{1}{2}2^2}) - (1 - e^{-\frac{1}{2}})\right] = .1178.$

   c) $E(|Y|) = \int_{-\infty}^{\infty} |y| \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = 2\int_0^{\infty} y \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = -\sqrt{\frac{2}{\pi}} e^{-y^2/2}\Big|_0^{\infty} = \sqrt{\frac{2}{\pi}}.$

   d) $P(|X| \le r) = 2P(0 \le X \le \sqrt{2\log 2}) = 2 \times (.881 - .5) = .762$

   e) By independence of $X$ and $Y$, $P(|X| \le r, |Y| \le r) = P(|X| \le r)P(|Y| \le r) = .762^2 = .58.$

   f) By the rotational symmetry of the joint distribution of $(X, Y)$,
   $P(|X| \le \frac{\sqrt{2}}{2}r, |Y| \le \frac{\sqrt{2}}{2}r) = \left[P(|X| \le \sqrt{\log 2})\right]^2 = (2 \times .2967)^2 = .3521.$

   g) $P(|X| \le r, 0 \le Y \le r) = P(|X| \le r)P(0 \le Y \le r) = .762 \times .381 = .29.$

2. a) $E(X^2) = Var(X) + E(X)^2 = 4$ and $E(Y^2) = Var(Y) + E(Y)^2 = 8$, so
   $E(10X^2 + 8Y^2 - XY + 8X + 5Y - 1) = 40 + 64 - 0 + 8 + 10 - 1 = 121.$

   b) $2X - (3Y - 5)$ is normal with mean 1 and variance 48, so

$$P(2X > 3Y - 5) = P(2X - 3Y + 5 > 0) = 1 - \Phi\left(\frac{0-1}{\sqrt{48}}\right) = 0.5574$$

3. a) $1 - \Phi(0.5)$ b) $1/2$ c) 5 d) $\sqrt{14}$

4. a) By symmetry, this must be 0.5.

   b) $P(Y - .2 < X < Y) = P(-.2 < X - Y < 0)$ and $X - Y$ is normal with mean 0 and variance .4, so

$$P(-.2 < X - Y < 0) = \Phi\left(\frac{0}{\sqrt{.4}}\right) - \Phi\left(\frac{-.2}{\sqrt{.4}}\right) = .1241.$$

5. Say $Y$ has standard deviation $\sigma$. Then $X$ has normal (0,1) distribution, $Y$ has normal $(1, \sigma^2)$ distribution, and $X - Y$ has normal $(-1, 1 + \sigma^2)$ distribution; hence

$$\begin{aligned} \frac{2}{3} &= P(X \le Y) \\ &= P(X - Y \le 0) \\ &= P\left(\frac{X - Y + 1}{\sqrt{1+\sigma^2}} \le \frac{1}{\sqrt{1+\sigma^2}}\right) \\ &= \Phi\left(\frac{1}{\sqrt{1+\sigma^2}}\right) \end{aligned}$$

   So $\frac{1}{\sqrt{1+\sigma^2}} = .43$ and $\sigma = \left((1/.43)^2 - 1\right)^{1/2} \approx 2.1$ .

6. a) $3X + 2Y$ has the normal$(0, 13)$ distribution.

$$\begin{aligned} P(3X + 2Y > 5) &= P\left(\frac{3X + 2Y}{\sqrt{13}} > \frac{5}{\sqrt{13}}\right) \\ &= 1 - \Phi(1.39) = 0.0823 \end{aligned}$$

   b) By using inclusion-exclusion,

$$\begin{aligned} P(\min(X, Y) < 1) &= P(X < 1) + P(Y < 1) - P(X < 1 \text{ and } Y < 1) \\ &= 2\Phi(1) - [\Phi(1)]^2 = 0.9748 \end{aligned}$$

c) By using the symmetry of the standard bivariate normal distribution,

$$\begin{aligned}P(|\min(X,Y)|<1) &= P(-1<X<1 \text{ and } Y>-1)+P(-1<Y<1 \text{ and } X>1)\\ &= P(-1<X<1 \text{ and } Y>-1)+P(-1<X<1 \text{ and } Y<-1)\\ &= P(-1<X<1)\\ &= \Phi(1)-\Phi(-1)=0.6826\end{aligned}$$

d) By using the rotational symmetry of the standard bivariate normal distribution,

$$\begin{aligned}P(\min(X,Y)>\max(X,Y)-1) = P(\max(X,Y)-\min(X,Y)<1) &= P(|X-Y|<1)\\ &= P\left(-\frac{1}{\sqrt{2}}<X<\frac{1}{\sqrt{2}}\right)\\ &= \Phi(0.71)-\Phi(-0.71)=0.5222\end{aligned}$$

7. Let $X$ be the (actual) arrival time of the bus, measured in minutes relative to 8 AM (so $X=1$ means that the bus arrived at 8:01 AM). Let $Y$ be my arrival time, also measured in minutes relative to 8 AM. Then $X$ has normal distribution with mean 10 and standard deviation 2/3, and $Y$ has normal distribution with mean 9 and standard deviation 1/2.

a) $P(Y<10)=P\left(\frac{Y-9}{1/2}<\frac{10-9}{1/2}\right)=\Phi(2)=0.9772.$

b) Assume $X$ and $Y$ are independent. Then $Y-X$ has normal distribution with mean $-1$ and variance $(2/3)^2+(1/2)^2=25/36$, and

$$P(Y<X)=P(Y-X<0)=P\left(\frac{Y-X+1}{\sqrt{25/36}}<\frac{1}{\sqrt{25/36}}\right)=\Phi(\sqrt{36/25})=\Phi(1.2)=0.8849.$$

c) $P(X<9 \mid X\notin[9,12])=\frac{P(X<9)}{P(X<9 \text{ or } X>12)}=\frac{P(X<9)}{P(X<9)+P(X>12)}=.9795$
since $P(X<9)=P\left(\frac{X-10}{2/3}<\frac{9-10}{2/3}\right)=\Phi(-\frac{3}{2})=.0668$
and $P(X>12)=P\left(\frac{X-10}{2/3}>\frac{12-10}{2/3}\right)=1-\Phi(3)=.0014.$

8. Let $X$ be Peter's arrival time and $Y$ be Paul's arrival time, both measured in minutes after noon.

a) Since $X-Y$ has the normal$(-2,34)$ distribution,

$$P(X<Y)=P(\frac{X-Y+2}{\sqrt{34}}<\frac{2}{\sqrt{34}})=\Phi(0.34)=0.6331$$

b) By independence,

$$\begin{aligned}P(-3<X<3 \text{ and } -3<Y<3) &= P(-3<X<3)P(-3<Y<3)\\ &= \left[\Phi\left(\frac{3}{5}\right)-\Phi\left(\frac{-3}{5}\right)\right]\left[\Phi\left(\frac{5}{3}\right)-\Phi\left(\frac{-1}{3}\right)\right]\\ &= 0.2626\end{aligned}$$

c)

$$\begin{aligned}P(|X-Y|<3)=P(-3<X-Y<3) &= \Phi\left(\frac{3+2}{\sqrt{34}}\right)-\Phi\left(\frac{-3+2}{\sqrt{34}}\right)\\ &= \Phi(0.86)-\Phi(-0.17)=0.3726\end{aligned}$$

9. Let $X_i$ denote the height in inches of the $i$th person. Then each $X_i$ has normal distribution with mean 70 and variance 4.

a) $P(\max(X_1,X_2,\dots,X_{100})>76)=1-P(\max(X_1,\dots,X_{100})\le 76)$
$=1-[P(X_1\le 76)]^{100}$
$=1-[\Phi(3)]^{100}=1-(.9986)^{100}=.1307.$

b) Write $Y = \frac{X_1+\cdots+X_{100}}{100}$. We want $P(Y > 70.5)$. Now $E(Y) = E(X_1) = 70$, $SD(Y) = \frac{SD(X_1)}{10} = \frac{2}{10} = 0.2$, and $Y$ has normal distribution since it is a linear combination of independent normal random variables. Hence the desired probability is

$$P(Y > 70.5) = P\left(\frac{Y-70}{0.2} > \frac{70.5-70}{0.2}\right) = 1 - \Phi(2.5) = 1 - 0.9938 = .0062.$$

c) The answer to (b) will be approximately the same as before, since $Y$ still has expected value 70 and SD 0.2; we assume the normal approximation is good. On the other hand, the answer to (a) will be exactly $1 - [P(X_1 \le 76)]^{100}$, but now there is no guarantee that $P(X_1 \le 76)$ is still .9986; hence raising the actual probability to the 100th power could make the answer very different from before.

10. a) Let $X_1$ and $X_2$ be the two incomes in dollars.

$$P\left(\frac{X_1+X_2}{2} > 65,000\right) = P\left(\frac{\bar{X} - 60,000}{10,000/\sqrt{2}} > \frac{65,000-60,000}{10,000/\sqrt{2}}\right)$$
$$= 1 - \Phi(0.71) = 0.2389$$

b) Let $X$ and $Y$ be the younger and older persons' income respectively.
$X - Y$ has the normal$(-\$20,000, \$14,142^2)$ distribution.

$$P(X > Y) = P\left(\frac{X - Y + 20,000}{\sqrt{14,142}} > \frac{20,000}{\sqrt{14,142}}\right)$$
$$= 1 - \Phi(1.41) = 0.0793$$

c) By independence,

$$\begin{aligned} P(\min(X,Y) > 50,000) &= P(X > 50,000)P(Y > 50,000) \\ &= (1 - \Phi(1))(1 - \Phi(-1)) \\ &= 0.1335 \end{aligned}$$

11. a) Say that $X_t$ has normal distribution with mean $m(t)$ and variance $v(t)$. Clearly $X_0 = 0$, so $m(0) = v(0) = 0$. Next argue that $m(s) + m(t) = m(s+t)$ and $v(s) + v(t) = v(s+t)$ for all $s, t > 0$; conclude that $m(t) = 0$ for all $t > 0$ and that $v(t) = t\sigma^2$ for all $t > 0$. Therefore $X_t$ has normal distribution with mean 0 and variance $t\sigma^2$.

b) $X_t$ and $Y_t$ are independent normal $(0, t\sigma^2)$ random variables, so $R = \frac{R_t}{\sigma\sqrt{t}}$ has Rayleigh distribution. Therefore

$$E(R_t) = \sigma\sqrt{t}E(R) = \sigma\sqrt{\frac{t\pi}{2}}$$

and

$$Var(R_t) = t\sigma^2 Var(R) = t\sigma^2\left(\frac{4-\pi}{2}\right) \quad SD(R_t) = \sigma\sqrt{t\left(\frac{4-\pi}{2}\right)}.$$

c) $P(R_1 > 2) = P(R > 2) = e^{-\frac{1}{2}2^2} = 0.1353$ .

12. Let the coordinates of the point where the first shot strikes be $(X_1, Y_1)$. Let the coordinates of the point where the second shot strikes be $(X_2, Y_2)$. Assume the two shots strike independently. Then $X_1, X_2, Y_1, Y_2$ are mutually independent normal $(\mu, 1)$ random variables (for some $\mu$), and the distance between the points where the two shots strike, say $D$, is given by

$$D = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} = \sqrt{2}\sqrt{\left(\frac{X_1 - X_2}{\sqrt{2}}\right)^2 + \left(\frac{Y_1 - Y_2}{\sqrt{2}}\right)^2} = \sqrt{2}\sqrt{U^2 + V^2},$$

where $U = (X_1 - X_2)/\sqrt{2}$, $V = (Y_1 - Y_2)/\sqrt{2}$. But $U$ and $V$ are independent normal $(0,1)$ random variables, so $\sqrt{U^2 + V^2}$ has Rayleigh distribution. Therefore

a) $E(D) = \sqrt{2}E(\sqrt{U^2 + V^2}) = \sqrt{2}\sqrt{\frac{\pi}{2}} = \sqrt{\pi} \approx 1.7725$.

b) $E(D^2) = 2E(U^2 + V^2) = 4 \Longrightarrow Var(D) = E(D^2) - [E(D)]^2 = 4 - \pi$.

**13.** We may assume $\sigma = 1$ by considering $X^* = X/\sigma$ and $Y^* = Y/\sigma$.

a) Refine the argument used to derive the density of $R$: The event $(R \in dr, \Theta \in d\theta)$ corresponds to $(X, Y)$ falling in an infinitesimal region having area $\frac{d\theta}{2\pi} \times 2\pi r dr = r dr d\theta$. (Argue that $(X, Y)$ must fall in an annulus around the origin of infinitesimal width $dr$ and radius $r$, but the angle subtended by $(X, Y)$ is restricted to be between $\theta$ and $\theta + d\theta$.) The joint density of $(X, Y)$ has nearly constant value $c^2 e^{-\frac{1}{2}r^2}$ over this region, so

$$P(R \in dr, \Theta \in d\theta) = c^2 e^{-\frac{1}{2}r^2} \times r dr d\theta = \frac{1}{2\pi} r e^{-\frac{1}{2}r^2} dr d\theta.$$

Conclude: $(R, \Theta)$ has joint density

$$f_{R,\Theta}(r, \theta) = \frac{1}{2\pi} r e^{-\frac{1}{2}r^2} r > 0, 0 < \theta < 2\pi.$$

Integrate out to get the marginal densities of $R$ and $\Theta$, see that $R$ has Rayleigh distribution, $\Theta$ has uniform$(0, 2\pi)$ distribution, and $R$ and $\Theta$ are independent [ $f_{R,\Theta}(r, \theta) = f_R(r) f_\Theta(\theta)$ ].

b) Let $X$ and $Y$ be independent normal $(0, \sigma^2)$ random variables. Define $R$ and $\Theta$ by $R = \sqrt{X^2 + Y^2}$ and $\Theta = \arctan(Y/X)$. By part (a), we know $R$ and $\Theta$ are independent, and $R/\sigma$ has Rayleigh distribution, $\Theta$ has uniform$(0, 2\pi)$ distribution. Since $X$ and $Y$ are independent normal $(0, \sigma^2)$ variables, and since the joint distribution of $X$ and $Y$ is determined by the joint distribution of $R$ and $\Theta$ (namely, via the transformation $X = R\cos\Theta$, $Y = R\sin\Theta$), the same conclusion (that $X$ and $Y$ are indep normal $(0, \sigma^2)$ variables) holds whenever $R$ and $\Theta$ have this joint distribution.

c) Using the results of (b), it suffices to find $h, k$ such that $h(U)$ has Rayleigh distribution and $k(V)$ has uniform$(0, 2\pi)$ distribution. Clearly we may take $k(v) = 2\pi v$. As for $h$, note that the distribution function of a random variable having Rayleigh distribution is

$$F(r) = 1 - e^{-r^2/2} \quad \text{for } r > 0.$$

The inverse of $F$ is $F^{-1}(u) = \sqrt{-2\log(1-u)}$ . By the result given in Section 4.5 on inverse distribution functions and simulation of random variables, $F^{-1}(U)$ is a random variable having distribution function $F$. In other words, $F^{-1}(U)$ has Rayleigh distribution. So it suffices to take $h = F^{-1}$, that is, $h(u) = \sqrt{-2\log(1-u)}$ .

**14.** a) Since the standard bivariate normal density is rotationally symmetric, $\Theta$ is distributed uniform$(0, 2\pi)$. This implies that $2\Theta$ *mod* $2\pi$ is also distributed uniform$(0, 2\pi)$.

b) Since $\cos(a \text{ mod } 2\pi) = \cos(a)$ and $\sin(a \text{ mod } 2\pi) = \sin(a)$, Exercise 13 part b) implies that, $R\cos 2\Theta$ and $R\sin 2\Theta$ are independent standard normals.

c) Since $R = \sqrt{X^2 + Y^2}$,

$$\frac{2XY}{\sqrt{X^2 + Y^2}} = \frac{2(R\cos\Theta)(R\sin\Theta)}{R} = R\sin 2\Theta$$

$$\frac{X^2 - Y^2}{\sqrt{X^2 + Y^2}} = \frac{R^2\cos^2\Theta - R^2\sin^2\Theta}{R} = R\cos 2\Theta$$

They are independent standard normals.

**15.** a) Let $Y = Z^2$. Clearly $f_Y(y) = 0$ for $y \le 0$. For $y > 0$, use Example 5 of Section 4.4:
$f_Y(y) = \frac{f_Z(\sqrt{y}) + f_Z(-\sqrt{y})}{2\sqrt{y}} = \frac{2\phi(\sqrt{y})}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} = \frac{1}{\sqrt{\pi}} (1/2)^{1/2} y^{1/2-1} e^{-y/2}$ ;
that is, $Y$ has gamma$(1/2, 1/2)$ distribution. The fact that $\Gamma(1/2) = \sqrt{\pi}$ follows from $\int_0^\infty f_Y(y)dy = 1$.

b) Write $n = 2k-1$, and use induction on $k = 1, 2, \ldots$. It's true for $k = 1$ by part (a). If true for $k$, then

$$\begin{aligned}
\Gamma\left(\frac{2k+1}{2}\right) &= \Gamma\left(\frac{2k-1}{2}+1\right) \\
&= \frac{2k-1}{2}\Gamma\left(\frac{2k-1}{2}\right) \\
&= \frac{2k-1}{2}\cdot\frac{\sqrt{\pi}(2k-2)!}{2^{2k-2}(k-1)!} \\
&= \frac{2k(2k-1)\cdot\sqrt{\pi}(2k-2)!}{2\cdot 2^{2k-2}\cdot 2k(k-1)!} \\
&= \frac{\sqrt{\pi}(2k)!}{2^{2k}k!},
\end{aligned}$$

therefore true for $k+1$.

c) $(X/\sigma)^2$ has gamma $(1/2, 1/2)$ distribution by (a). That is, $Y = X^2/\sigma^2$ has density $f_Y(y) = \frac{1}{\sqrt{\pi}}(\frac{1}{2})^{1/2}y^{1/2-1}e^{-y/2}, y > 0$.

So by a linear change of variable, $X^2 = \sigma^2 Y$ has density

$f_{X^2}(w) = \frac{1}{\sigma^2}f_Y(\frac{w}{\sigma^2}) = \frac{1}{\sqrt{\pi}}(\frac{1}{2\sigma^2})^{1/2}w^{1/2-1}e^{-w/2\sigma^2}$;

that is, $X^2$ has gamma $(1/2, \frac{1}{2\sigma^2})$ distribution.

d) $Z_1^2, \ldots, Z_n^2$ are independent gamma $(1/2, 1/2)$ variables, by (a). Since $Z_1^2 + \cdots + Z_n^2$ has gamma $(n/2, 1/2)$ distribution.

e) For each $j$, $Y_j$ has gamma $(k_j/2, 1/2)$ distribution (by definition of $\chi^2$). The $Y_j$'s are independent, $Y_1 + \cdots + Y_n$ has gamma $((k_1 + \cdots + k_n)/2, 1/2)$ distribution, i.e., has $\chi^2$ distribution with $k_1 + \cdots + k_n$ degrees of freedom.

16. Since $R_{2m}^2$ has the gamma$(m, 1/2)$ distribution, it has the same distribution as the time of the $m$th arrival of a Poisson process with rate $1/2$.

Let $T_m$ be the time of the $m$th arrival of a Poisson process with rate $1/2$. For $x \geq 0$, $R_{2m}^2$ has c.d.f.

$$\begin{aligned}
F(x) &= P(R_{2m}^2 \leq x) = P(T_m \leq x) \\
&= 1 - P(T_m > x) = 1 - P(\text{at most } m-1 \text{ arrivals in } (0,x)) \\
&= 1 - \sum_{k=0}^{m-1}\frac{e^{-\frac{x}{2}}(\frac{x}{2})^k}{k!}
\end{aligned}$$

b) Since $P(R_{2m} \leq x) = P(R_{2m}^2 \leq x^2)$, $R_{2m}$ has c.d.f.

$$P(R_{2m} \leq x) = 1 - \sum_{k=0}^{m-1}\frac{e^{-\frac{x^2}{2}}(\frac{x^2}{2})^k}{k!}$$

c) For $m = 1$,

$$P(R_2^2 \leq x) = 1 - e^{-\frac{x}{2}}$$
$$P(R_2 \leq x) = 1 - e^{-\frac{x^2}{2}}$$

| $r$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(R_4 \leq r)$ | 0.0902 | 0.5940 | 0.9389 | 0.9970 | 0.9999 |

17. a) Skew-normal approximations: 0.1377, 0.5940, 0.9196, 0.9998, 1.0000
Compare to the exact values: 0.0902, 0.5940, 0.9389, 0.9970, 1.0000
b) 0.441, 0.499. Skew-normal is better.

18. The $X_i$'s are independent and identically distributed uniform $[-1/2, 1/2]$ random variables, so $E(X_i) = 0$ and $Var(X_i) = 1/12$. Hence $E(X) = 0$ and $Var(X) = 1/12n$ . Since $X$ is the average of iid random variables having finite variance, it follows that for large $n$, $X$ has approximately normal distribution with mean 0 and variance $1/12n$, and $\sqrt{12n}X$ has approximately standard normal distribution. This argument applies to $Y$ and $Z$, too. Since $X$, $Y$, and $Z$ are independent, we then have that

$$(\sqrt{12n}X)^2 + (\sqrt{12n}Y)^2 + (\sqrt{12n}Z)^2 = 12n(X^2 + Y^2 + Z^2) = 12nR^2$$

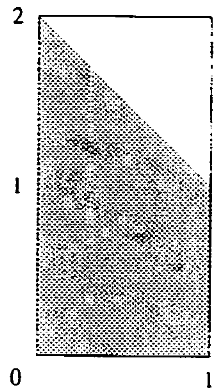
has approximately chi-squared distribution with 3 degrees of freedom. Hence

$$0.95 = P(12nR^2 \leq 7.82) = P\left(R \leq \sqrt{\frac{7.82}{12n}}\right).$$

So $R$ is 95% sure to be smaller than $\sqrt{\frac{7.82}{12n}} = \frac{0.807}{\sqrt{n}}$ .

# Section 5.4

1. a) From the picture: $P(X_1 + X_2 \le 2) = \frac{\text{shaded area}}{\text{area of rectangle}} = \frac{3}{4}$.




b) See the pictures below: If $0 \le z \le 1$,

$P(X_1 + X_2 \in dz) = \frac{(1/2)\cdot(z+dz)^2 - (1/2)\cdot z^2}{2} \approx \frac{z \cdot dz}{2}$ (ignoring $(dz)^2$ as negligible in comparison to $dz$);

if $1 \le z \le 2$,

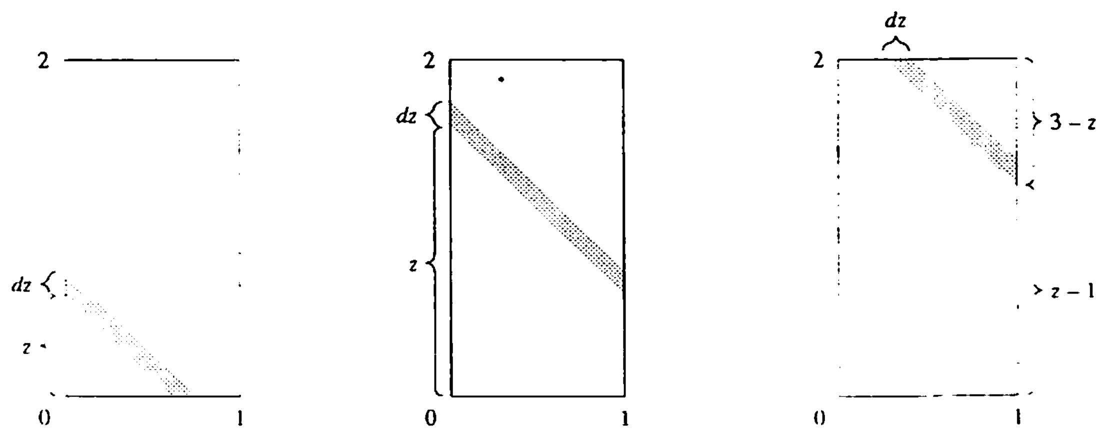
$P(X_1 + X_2 \in dz) = \frac{dz}{2}$ (area of a parallelogram);

if $2 \le z \le 3$,

$P(X_1 + X_2 \in dz) = \frac{(1/2)\cdot(3-z)^2 - (1/2)\cdot(3-z-dz)^2}{2} \approx \frac{(3-z)dz}{2}$ (again ignoring $(dz)^2$ ).

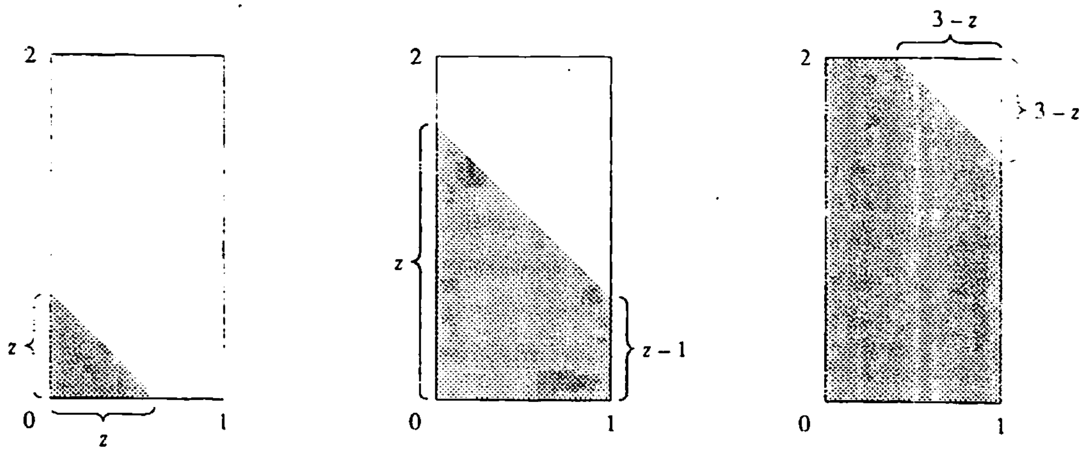
Therefore

$$f_{X_1+X_2}(z) = \begin{cases} z/2 & 0 \le z \le 1 \\ 1/2 & 1 \le z \le 2 \\ (3-z)/2 & 2 \le z \le 3 \\ 0 & \text{elsewhere} \end{cases}$$




c) You could either integrate b) or argue from the pictures below:

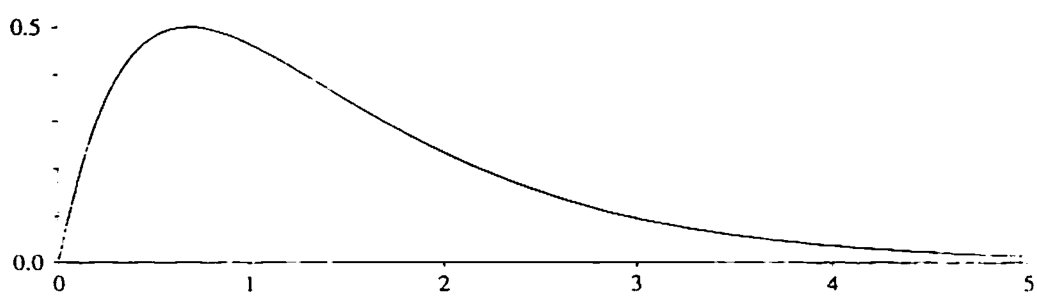
$$F_{X_1+X_2}(z) = \begin{cases} 0 & z \le 0 \\ z^2/4 & 0 \le z \le 1 \\ (2z-1)/4 & 1 \le z \le 2 \\ 1-(3-z)^2/4 & 2 \le z \le 3 \\ 1 & z \ge 3 \end{cases}$$




2. a) $\int_0^1 t\,dt + \int_1^{1.5}(2-t)dt = 7/8$.
   Or: $P(S_2 \le 1.5) = 1 - P(S_2 \le 0.5) = 1 - \int_0^{1/2} t\,dt = 7/8$.
   b) $1/2$, by symmetry of the density of $S_3$.
   c) $\int_0^{1/2} t^2 dt + \int_1^{1.1} \left(1 - \frac{1}{2}(2-t)^2 - \frac{1}{2}(t-1)^2\right) dt = 0.2213$.
   d) Approximately $f_{S_3}(1) \times 0.001 = (1/2) \times 0.001 = 0.0005$.

3. Let $X$ be the waiting time in the first queue, $Y$ the waiting time in the second. Then $X$ and $Y$ are independent exponential random variables with rates $\alpha$ and $\beta$ respectively.

   a) From the convolution formula,
   $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_0^z \alpha e^{-\alpha x} \beta e^{-\beta(z-x)}dx$, $z > 0$.
   If $\alpha = \beta$, the integral becomes
   $\alpha^2 e^{-\alpha z} \int_0^z dx = \alpha^2 z e^{-\alpha z}$ (density of gamma $(2, \alpha)$).
   If $\alpha \ne \beta$, the integral becomes
   $\int_0^z \alpha e^{-\alpha x} \beta e^{-\beta(z-x)}dx = \alpha\beta e^{-\beta z} \int_0^z e^{-(\alpha-\beta)x}dx = \frac{\alpha\beta}{\alpha-\beta}(e^{-\beta z} - e^{-\alpha z})$. Sketch of the density in case $\alpha = 1$ and $\beta = 2$:




   b) $E(X+Y) = E(X) + E(Y) = \frac{1}{\alpha} + \frac{1}{\beta}$.

   c) $Var(X+Y) = Var(X) + Var(Y) = \frac{1}{\alpha^2} + \frac{1}{\beta^2} \Longrightarrow SD(X+Y) = \sqrt{\frac{\alpha^2+\beta^2}{\alpha\beta}}$.

4. a) Let $T_1$ be the time to the first failure, $T_2$ the time till both fail. Then $T_1$ has constant hazard rate $2\lambda$, so has exponential distribution with rate $2\lambda$. Also, $T_2 - T_1$ is independent of $T_1$ with the same distribution. So $T_2 = T_1 + (T_2 - T_1)$ has gamma $(2, 2\lambda)$ distribution.

   b) Use the fact that $T_2$ is the sum of two independent exponential $(2\lambda)$ random variables.
   $E(T_2) = 2 \times \frac{1}{2\lambda} = \frac{1}{\lambda}$,
   $Var(T_2) = 2 \times \frac{1}{(2\lambda)^2} = \frac{1}{2\lambda^2}$.

   c) For each $t > 0$:
   $P(T \le t) = \int_0^t (2\lambda)^2 te^{-2\lambda t} dt = \int_0^{2\lambda t} xe^{-x} dx = 1 - (1 + 2\lambda t)e^{-2\lambda t}$.
   We want $t$ such that
   $1 - (1 + 2\lambda t)e^{-2\lambda t} = .9 \iff e^{2\lambda t} = 10(1 + 2\lambda t)$.
   The equation $e^y = 10(1 + y)$ has one positive solution, namely $y^* \approx 3.88972$, so the desired $t$ is $t = \frac{y^*}{2\lambda} \approx \frac{1.94486}{\lambda}$.

5. a) The range of $X + Y$ is the interval $[1, 7]$. If $1 < t < 7$ then

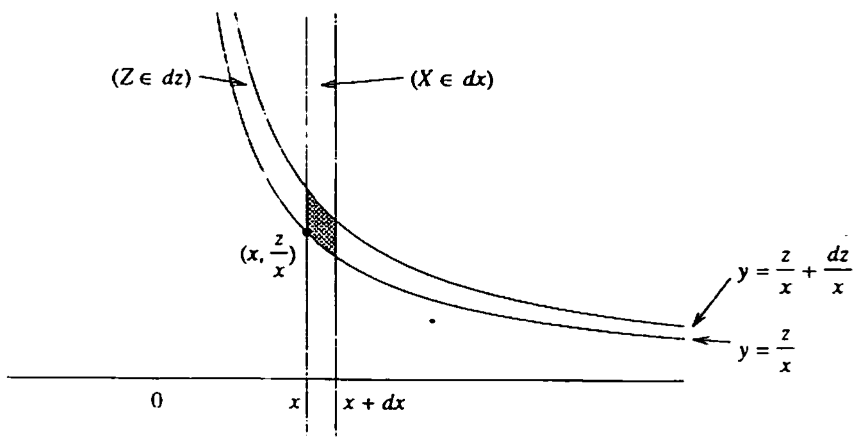
$$P(X + Y \in dt) = P(X = \text{int}(t), t - \text{int}(t) < Y < t - \text{int}(t) + dt) = \frac{1}{6}dt.$$

   So $X + Y$ has uniform distribution over $[1, 7]$ and $10(X + Y)$ has uniform distribution over $[10, 70]$.

   b) $\frac{58-29}{70-10} = 0.483$

6. gamma $(r_1 + r_2 + \cdots + r_n, \lambda)$. Use the result for two variables given in Example 2, and induction on $n$.

7. a) Write $Z = XY$. Let $z > 0$:



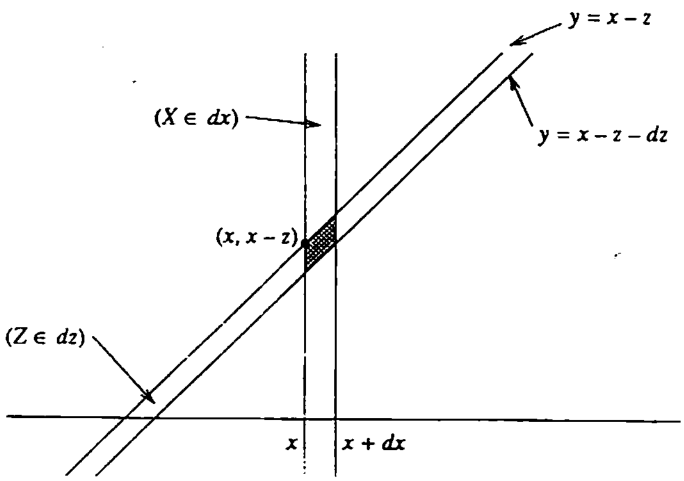

The event $(X \in dx, Z \in dz)$ is represented by the shaded region. This region is approximately a parallelogram, with left side having length $\frac{dz}{|x|}$, and distance between the vertical sides $dx$, so the area of the parallelogram is approximately $\frac{1}{|x|}dxdz$. The joint density of $X$ and $Y$ on this small parallelogram has nearly constant value $f_{X,Y}(x, \frac{z}{x})$. Conclude:

$$P(X \in dx, XY \in dz) = \frac{1}{|x|} f_{X,Y}(x, \frac{z}{x}) dxdz.$$

This argument works for $z < 0$ too. Integrate out $x$ to get

$$f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_{X,Y}(x, \frac{z}{x}) dx.$$

b) Write $Z = X - Y$. Let $z$ be arbitrary.



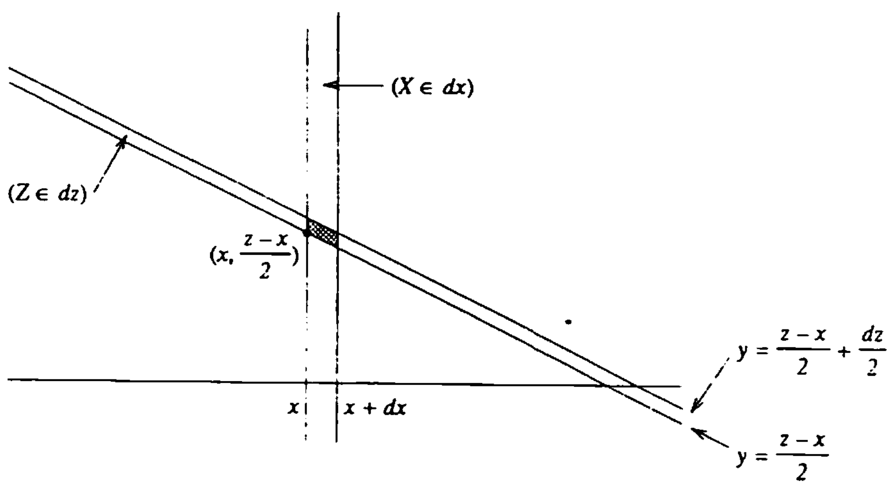

From diagram, we get

$$P(X \in dx, X - Y \in dz) = f_{X,Y}(x, x - z)dxdz,$$

whence

$$f_{X-Y}(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, x - z)dx.$$

c) Let $Z = X + 2Y$. For arbitrary $z$ we have:




From diagram, obtain

$$P(X \in dx, X + 2Y \in dz) = f_{X,Y}(x, \frac{z - x}{2})\frac{dxdz}{2}$$

and

$$f_{X+2Y}(z) = \int_{-\infty}^{\infty} \frac{1}{2} f_{X,Y}(x, \frac{z - x}{2})dx.$$

8. If $z > 0$, then

$$F_{X/Y}(z) = P\left(\frac{X}{Y} \le z\right) = P(Y \ge \frac{1}{z}X) \quad \text{(since } Y > 0\text{)}$$

$$= \int_{x=0}^{\infty}\int_{y=x/z}^{\infty} f_X(x)f_Y(y)dydx = \int_{x=0}^{\infty} \alpha e^{-\alpha x}\left(\int_{y=x/z}^{\infty}\beta e^{-\beta y}dy\right)dx$$

$$= \int_{x=0}^{\infty} \alpha e^{-\alpha x}e^{-\frac{\beta}{z}x}dx = \frac{\alpha}{\alpha+\frac{\beta}{z}} = \frac{\alpha z}{\alpha z+\beta}.$$

If $z \le 0$ then $F_{X/Y}(z) = 0$ .

9. $f_X(x) = -\log(x) \qquad (0 < x < 1)$

10.

$$f_Y(y) = \begin{cases} 1/2 & \text{if } 0 < y < 1 \\ 1/(2y^2) & \text{if } y \ge 1 \\ 0 & \text{otherwise} \end{cases}$$

11. uniform $(0,1)$

12. $X = -\log\{U(1-V)\} = -\log U - \log(1-V)$. Since $1-V$ is also uniform on $[0,1]$, we just need to find the distribution of $-\log U$. For $x \ge 0$,

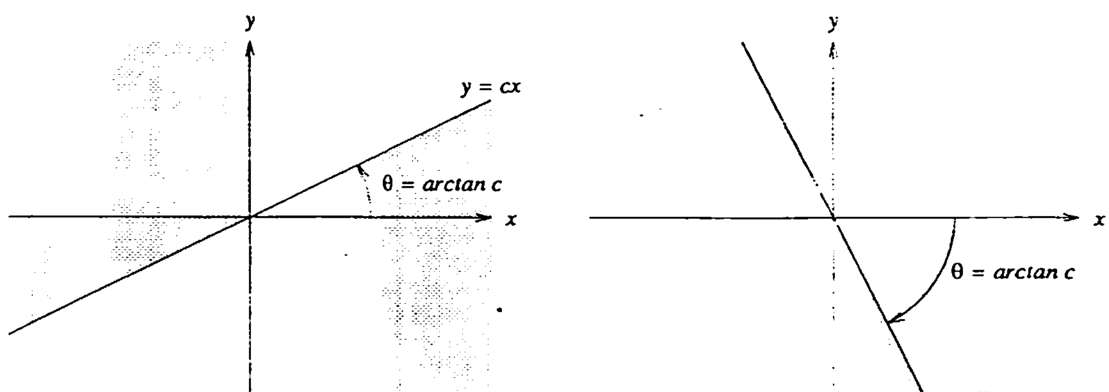
$$P(-\log U \le x) = P(U \ge e^{-x}) = 1 - e^{-x}$$

so $-\log U$ is an exponential random variable with rate 1. Since $X$ is the sum of two independent exponentials with rate 1, it has the gamma$(2,1)$ distribution. Its density is $f(x) = xe^{-x}$ $(x > 0)$. $E(X) = ?$ and $Var(X) = 2$.

13. $f_Z(z) \qquad \lambda e^{-\lambda|z|}$

14. If the joint distribution of $(X,Y)$ is symmetric under rotations, then the probability that $(X,Y)$ falls in a sector centered at $(0,0)$ subtending an angle $\theta$ is $\theta/2\pi$. So let $c \in R$. We have

$$P\left(\frac{Y}{X} \le c\right) = P(Y \le cX, X > 0) + P(Y \ge cX, X < 0).$$




Case c > 0 Case c < 0

From diagram, the corresponding region in the plane is a pair of sectors centered at $(0,0)$, each sector subtending an angle of $\frac{\pi}{2} + \arctan c$, where $\arctan c$ is that unique angle in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ having tangent $c$. Hence

$$P\left(\frac{Y}{X} \le c\right) = 2P\left(\ (X,Y) \text{ falls in a sector subtending an angle of } \tfrac{\pi}{2} + \arctan c\ \right)$$

$$= \frac{2(\frac{\pi}{2} + \arctan c)}{2\pi} = \frac{1}{2} + \frac{1}{\pi}\arctan c.$$

So $Y/X$ is a continuous random variable having density

$$f_{Y/X}(z) = \frac{d}{dz}P(Y/X \le z) = \frac{1}{\pi}\frac{1}{1+z^2}$$

15. a) Let $X_1 = \lambda X$ and $Y_1 = \lambda Y$. Then $X_1$ and $Y_1$ are independent exponential (1) variables, and $Z = \min(X_1, Y_1)/\max(X_1, Y_1)$.

b) $P(Z \le z) = 2z/(1+z)$

c) $f_Z(z) = 2/(1+z)^2$ for $0 < z < 1$.

16. a) Fix $r$ and $t$.

$F(r, \lambda, t) = P(T \le t) = P(\lambda T \le \lambda t) = P(T_1 \le \lambda t)$

where $T_1$ has gamma $(r, 1)$ distribution. This last quantity is an increasing function of $\lambda$, since any c.d.f. is nondecreasing.

b) Fix $\lambda$ and $t$. Let $0 < r' < r$. Let $T$ have gamma $(r, \lambda)$ distribution, and $T'$, $T''$ be independent with gamma $(r', \lambda)$ and gamma $(r - r', \lambda)$ distribution respectively. Since $T$ has the same distribution as $T' + T''$, and since $T' + T'' \ge T'$, we have

$$F(r, \lambda, t) = P(T \le t) = P(T' + T'' \le t) \le P(T' \le t) = F(r', \lambda, t)$$

17. a) Let $0 < t < 1$. The plane that cuts at $(t/3, t/3, t/3)$ has equation $x + y + z = t$. The plane that cuts at $((t+h)/3, (t+h)/3, (t+h)/3)$, $h$ small positive, has equation $x + y + z = t + h$. These two planes determine a "slab" of uniform thickness $h/\sqrt{3}$. The desired cross-sectional area is

$$\begin{aligned}\lim_{h\to 0} \frac{\text{volume of slab}}{\text{thickness of slab}} &= \lim_{h \to 0} \frac{\text{volume}\{(x, y, z) : t \le x + y + z \le t + h\}}{h/\sqrt{3}} \\ &= \lim_{h\to 0} \frac{P(t \le X + Y + Z \le t + h)}{h/\sqrt{3}} \\ &= \sqrt{3} f_{X+Y+Z}(t)\end{aligned}$$

where $X, Y, Z$ are independent uniform $[0, 1]$ random variables. Hence the area of the cross-section at $(t/3, t/3, t/3)$ is

$$\begin{cases} \frac{\sqrt{3}}{2} t^2 & 0 < t < 1 \\ \sqrt{3}(-t^2 + 3t - \frac{3}{2}) & 1 < t < 2 \\ \frac{\sqrt{3}}{2}(3 - t)^2 & 2 < t < 3 \end{cases}$$

b) If $t \le 1$, the plane $x + y + z = t$ intersects the edges of the cube at the three points

$$(t, 0, 0)(0, t, 0)(0, 0, t)$$

(find these coordinates e.g. using the equations of the edges of the cube); hence the cross section is an equilateral triangle having side length $t\sqrt{2}$ and area $\frac{\sqrt{3}}{4} \cdot 2t^2 = \frac{\sqrt{3}}{2} t^2$.
If $t = 3/2$, the plane $x + y + z = 3/2$ intersects the edges of the cube at six points forming the vertices of a (plane) regular hexagon having side length $1/\sqrt{2}$; hence the cross-sectional area is $6 \cdot \frac{\sqrt{3}}{4} \frac{1}{2} = \frac{3}{4\sqrt{3}}$.

18. a) Duplicate the derivation of the density of the sum of three independent uniform random variables: If $n \ge 2$ then

$$\begin{aligned} f_n(x) &= f_{S_{n-1}+X_n}(x) \\ &= \int_{-\infty}^{\infty} f_{n-1}(u) f_{X_n}(x - u) du \\ &= \int_{x-1}^{x} f_{n-1}(u) du \\ &= F_{n-1}(x) - F_{n-1}(x - 1). \end{aligned}$$

b) Claim: For each $n \ge 1$: For all $i = 1, \dots, n$:

(I): On the interval $(i - 1, i)$, $f_n(x)$ is a polynomial having leading term $\frac{(-1)^{i-1}}{(n-1)!}\binom{n-1}{i-1} x^{n-1}$;

(II): On the interval $(i - 1, i)$, $F_n(x)$ is a polynomial having leading term $\frac{(-1)^{i-1}}{n!}\binom{n-1}{i-1} x^n$.

Proof: Induction. True for $n = 1$. If holds for $n = k$, where $k \geq 1$, then

(I): For each $i = 1, \ldots, k+1$: On the interval $(i-1, i)$ we have

$$\begin{aligned}
f_{k+1}(x) &= F_k(x) - F_k(x-1) \\
&= \frac{(-1)^{i-1}}{k!}\binom{k-1}{i-1}x^k - \frac{(-1)^{i-2}}{k!}\binom{k-1}{i-2}(x-1)^k + \text{ terms of degree at most } k-1 \\
&= \frac{(-1)^{i-1}}{k!}\left\{\binom{k-1}{i-1}\binom{k-1}{i-2}\right\}x^k + \text{ terms of degree at most } k-1 \\
&= \frac{(-1)^{i-1}}{k!}\binom{k}{i-1}x^k + \text{ terms of degree at most } k-1 \\
&= \text{ a polynomial having leading term } \frac{(-1)^{i-1}}{(k+1-1)!}\binom{k+1-1}{i-1}x^{k+1-1}.
\end{aligned}$$

Note that this argument works for the "boundary" cases $i = 1$ and $i = k+1$.

(II): For each $i = 1, \ldots, k+1$: On the interval $(i-1, i)$ we have

$$\begin{aligned}
F_{k+1}(x) &= \int_0^{i-1} f_{k+1}(t)dt + \int_{i-1}^{x} f_{k+1}(t)dt \\
&= \text{const} + \int_{i-1}^{x}\left\{\frac{(-1)^{i-1}}{k!}\binom{k}{i-1} \text{ terms of degree at most } k-1\right\}dt \\
&= \frac{(-1)^{i-1}}{(k+1)!}\binom{k}{i-1}x^{k+1} + \text{ terms of degree at most } k \\
&= \text{a polynomial having leading term } \frac{(-1)^{i-1}}{(k+1)!}\binom{k+1-1}{i-1}x^{k+1}.
\end{aligned}$$

So claim holds for $n = k+1$.

c) Claim: For each $n = 1, 2, \ldots$ we have: If $0 < x < 1$ then

$$(i) \quad f_n(x) = \frac{x^{n-1}}{(n-1)!}; \qquad (ii) \quad F_n(x) = \frac{x^n}{n!}.$$

Proof: Holds for $n = 1$. If holds for $n = k$, where $k \geq 1$, then for $x \in (0, 1)$

$$f_{k+1}(x) = F_k(x) - F_k(x-1) = \frac{x^k}{k!} - 0 = \frac{x^{k+1-1}}{(k+1-1)!}$$

and

$$F_{k+1}(x) = \int_0^x f_{k+1}(t)dt = \int_0^x \frac{t^k}{k!}dt = \frac{x^k}{(k+1)!}$$

so claim holds for $n = k+1$.

d) Claim: For each $n = 1, 2, \ldots$ we have: If $n-1 < x < n$ then

$$(i) \quad f_n(x) = \frac{(n-x)^{n-1}}{(n-1)!}; \qquad (ii) \quad F_n(x) = 1 - \frac{(n-x)^n}{n!}.en(x) = 1 - \frac{(n-x)^n}{n!}.$$

Proof: Holds for $n = 1$. If holds for $n = k$, where $k \geq 1$, then for $x \in (k, k+1)$

$$f_{k+1}(x) = F_k(x) - F_k(x-1) = 1 - F_k(x-1) = 1 - \left(1 - \frac{(k-(x-1))^k}{k!}\right) = \frac{(k+ \quad -x)^k}{k!}$$

and

$$F_{k+1}(x) = P(S_{k+1} \leq x) = 1 - P(S_{k+1} > x) = 1 - \int_x^{k+1} f_{k+1}(t)dt = 1 - \frac{(k+1-x)^{k+1}}{(k+1)!}$$

so claim holds for $n = k+1$.

e) $P(0 \le S_4 \le 1) = F_4(1) = \frac{1}{4!}$ by (c).

f) $P(1 \le S_4 \le 2) = P(S_4 \le 2) - P(S_4 \le 1) = \frac{1}{2} - \frac{1}{4!}$ (by symmetry and part (e))

g) $P(1.5 \le S_4 \le 2) = \int_{1.5}^{2} f_4(t)dt$ . Now if $1 < t < 2$ then

$$\begin{aligned}
f_3(t) &= F_2(t) - F_2(t-1) = 1 - \frac{(2-t)^2}{2!} - \frac{(t-1)^2}{2!};\\
F_3(t) &= F_3(1) + \int_1^t f_3(u)du\\
&= (t-1) + \frac{(2-t)^3}{3!} - \frac{(t-1)^3}{3!};\\
f_4(t) &= F_3(t) - F_3(t-1) = (t-1) + \frac{(2-t)^3}{3!} - \frac{2(t-1)^3}{3!}.
\end{aligned}$$

Hence

$$P(1.5 \le S_4 \le 2) = \int_{1.5}^{2} f_4(t)dt = \int_{1.5}^{2}\left[(t-1) + \frac{(2-t)^3}{3!} - \frac{2(t-1)^3}{3!}\right]dt = 0.29948.$$

Aside: Integrate $f_4$ to get, for $1 < t < 2$,

$$F_4(t) - F_4(1) = \int_1^t f_4(u)du = \frac{(t-1)^2}{2!} - \frac{(2-t)^4}{4!} - \frac{2(t-1)^4}{4!} + \frac{1}{4!}.$$

**19.** If $X, Y$ have joint density $f_{X,Y}(x,y)$ then argue as in Exercise 7 that $X$ and $X+Y$ have joint density

$$f_{X,X+Y}(x,z) = f_{X,Y}(x, z-x)$$

and $X/Y$ , $Y$ have joint density

$$f_{X/Y,Y}(u,y) = |y| f_{X,Y}(uy, y).$$

Therefore the joint density of $X/(X+Y)$ and $X+Y$ is

$$f_{X/(X+Y),X+Y}(w,z) = |z| f_{X,X+Y}(wz, z) = |z| f_{X,Y}(wz, z-wz).$$

Here

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, x > 0, \quad and \quad f_Y(y) = \frac{\lambda^s}{\Gamma(s)} y^{s-1} e^{-\lambda y}, y > 0.$$

So the joint density of $X/(X+Y)$ and $X+Y$ is

$$\begin{aligned}
f_{X/(X+Y),X+Y}(w,z) &= |z| f_{X,Y}(wz, (1-w)z)\\
&= z f_X(wz) f_Y((1-w)z)\\
&= z \frac{\lambda^r}{\Gamma(r)} (wz)^{r-1} e^{-\lambda wz} \frac{\lambda^s}{\Gamma(s)} [(1-w)z]^{s-1} e^{-\lambda(1-w)z}\\
&= \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} w^{r-1}(1-w)^{s-1} \frac{\lambda^{r+s}}{\Gamma(r+s)} z^{r+s-1} e^{-\lambda z}
\end{aligned}$$

if $0 < w < 1$ and $z > 0$ ; the density is zero otherwise. Integrate out to see that

$$f_{X+Y}(z) = \frac{\lambda^{r+s}}{\Gamma(r+s)} z^{r+s-1} e^{-\lambda z}, z > 0$$

(we knew this already); and that

$$f_{X/(X+Y)}(w) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} w^{r-1}(1-w)^{s-1}, 0 < w < 1;$$

and that $f_{X/(X+Y),X+Y}(w,z) = f_{X/(X+Y)}(w) \cdot f_{X+Y}(z)$. Thus $X/(X+Y)$ has beta distribution with parameters $r, s$, and is independent of $X+Y$.

# Chapter 5: Review

1. Draw a picture of the unit square; then compute

$$P(Y \geq 1/2 | Y \geq X^2) = \frac{P(Y \geq 1/2, Y \geq X^2)}{P(Y \geq X^2)} = 1 - \frac{\sqrt{2}}{4},$$

since numerator $= \int_{1/2}^{1} \sqrt{y}dy = \frac{2}{3}(1 - \frac{\sqrt{2}}{4})$

and denominator $= \int_0^1 \sqrt{y}dy = \frac{2}{3}$ .

2. a) 3/4.

b) Density of $Z = |X + Y|$ is $f_Z(z) = 1 - z/2$, $0 < z < 2$. So

$$E|X + Y| = E(Z) = \int_0^2 z f_Z(z)dz = \int_0^2 (z - \frac{z^2}{2})dz = \frac{2}{3}.$$

3. a) Assume center of coin lands uniformly on disc of radius 2.5"

$$P(\text{lands in disc of radius } \frac{1}{2}'') = \frac{\pi(\frac{1}{2})^2}{\pi(2\frac{1}{2})^2} = 0.04$$

b) Center lands uniformly in square of side 4.5

$$\frac{\pi(0.25)}{(4.5)^2} = 0.03878\ldots$$

c) $1 - \frac{\text{area of square with side } (4.5 - \frac{1}{2}\sqrt{2})}{(4.5)^2} = 0.2896$

4. a) $P(X^2 + Y^2 \leq r^2)$ is one fourth the area of a circle of radius $r$ inside a square with the same center and side length 2. For $r > \sqrt{2}$, this is 1. For $r \leq 1$, this is $\frac{\pi r^2}{4}$. For $1 < r \leq \sqrt{2}$, the circle extends beyond the square on each side. On each side, the square intersects the circle in a chord with length $2\sqrt{r^2 - 1}$. For one side, the area of excess is the area of a wedge of the circle minus the area of a triangle. That is

$$\text{Area of excess} = \frac{2\arccos(1/r)}{2\pi}(\pi r^2) - \frac{1}{2}(2\sqrt{r^2 - 1})(1) = r^2 \arccos(\frac{1}{r}) - \sqrt{r^2 - 1}$$

Using this we see that for $1 \leq r \leq \sqrt{2}$,

$$\begin{aligned} P(X^2 + Y^2 \leq r^2) &= \frac{1}{4}\left(\pi r^2 - 4(r^2 \arccos(\frac{1}{r}) - \sqrt{r^2 - 1})\right) \\ &= \frac{\pi r^2}{4} - r^2 \arccos(\frac{1}{r}) + \sqrt{r^2 - 1} \end{aligned}$$

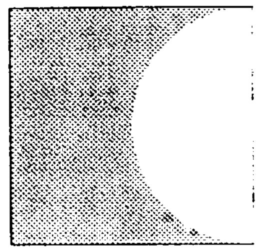
b)

$$F(x) = P(R^2 \leq x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{\pi x}{4} & \text{for } 0 \leq x \leq \\ \frac{\pi x}{4} - x\arccos(\frac{1}{\sqrt{x}}) + \sqrt{x - 1} & \text{for } 1 < x \leq . \\ 1 & \text{for } x > 2 \end{cases}$$

c) Differentiate the c.d.f. to find the density.

$$f(x) = F'(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{\pi}{4} & \text{for } 0 \leq x \leq 1 \\ \frac{\pi}{4} - \arccos(\frac{1}{\sqrt{x}}) & \text{for } 1 < x \leq 2 \\ 0 & \text{for } x > 2 \end{cases}$$

5. a) $P(D \geq 1/2) =$ shaded area $= 1 - \frac{1}{2} \cdot \pi \cdot (\frac{1}{2})^2 = 1 - \pi/8$ .



b) Let the coordinates of the point be $(X, Y)$, where without loss of generality $X$ has uniform distribution on $(0,1)$, and $Y$ has uniform distribution on $(-1/2, 1/2)$. Take the midpoint of the given side to be $(0,0)$. Then $D^2 = X^2 + Y^2$ by Pythagoras' theorem, and

$$E(D^2) = E(X^2) + E(Y^2) = \int_0^1 x^2 dx + \int_{-1/2}^{1/2} y^2 dy = 5/12.$$

6. The charge $C$ per call in dollars is

$$C = \begin{cases} 1 & \text{if } T \leq 1 \\ 1 + 0.01 \times 60(T-1) = .4 + .6T & \text{if } T \geq 1 \end{cases}$$

where $T$ is the call duration in minutes. So

$$E(C) = \int_{t=0}^{1} 1 f(t)dt + \int_{t=1}^{\infty} (.4 + .6t)dt = 1 + .6 \int_{t=1}^{\infty} (t-1) f(t) dt$$

where $f$ is the density of $T$.

a) Here $f(t) = e^{-t}$, so $\int_{t=1}^{\infty} (t-1)f(t)dt = e^{-1}$. So $E(C) = 1 + .6e^{-1} = 1.22$.

b) Here $f(t) = (1/2)e^{-t/2}$, so $\int_{t=1}^{\infty} (t-1)f(t)dt = 2e^{-1/2}$. So $E(C) = 1 + 1.2e^{-1/2} = 1.73$.

c) Here $f(t) = 4te^{-2t}$, so $\int_{t=1}^{\infty} (t-1)f(t)dt = 2e^{-2}$. So $E(C) = 1 + 1.2e^{-2} = 1.16$.

7. $\bar{X}$ has normal distribution with mean $\mu$ and SD 0.1, so

$$P(|\bar{X} - \mu| \geq .25) = 2(1 - \Phi(.25/.1)) = 2(1 - \Phi(2.5)) = 2(1 - .9938) = .0124.$$

8. By the central limit theorem, the distribution of $\bar{X}$ is *approximately* normal, with the same mean and variance. So the required probability is now approximately 0.0124 . [$n$ is 100, which is pretty large. So the normal approximation should be good.]

9. a)

$$f_{X+U}(x) = \begin{cases} 1/4 & 0 < x < 1 \\ 1/2 & 1 < x < 2 \\ 1/4 & 2 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

b) Try a uniform $(-1/2, 1/2)$ distribution for $U$.

10. a) For $z > 0$ ,

$$P(Z > z) = P(X > z, Y > z) = P(X > z)P(Y > z) = e^{-\lambda z}e^{-\mu z} = e^{-(\lambda+\mu)z};$$

hence $P(Z \leq z) = 1 - e^{-(\lambda+\mu)z}$ and

$$f_Z(z) = (\lambda + \mu)e^{-(\lambda+\mu)z}, z > 0.$$

b) $P(X \geq Y) = \int_{y=0}^{\infty} \int_{x=y}^{\infty} \lambda e^{-\lambda x} \mu e^{-\mu y} dx dy = \int_{y=0}^{\infty} \mu e^{-\mu y} e^{-\lambda y} dy = \frac{\mu}{\lambda + \mu}$.

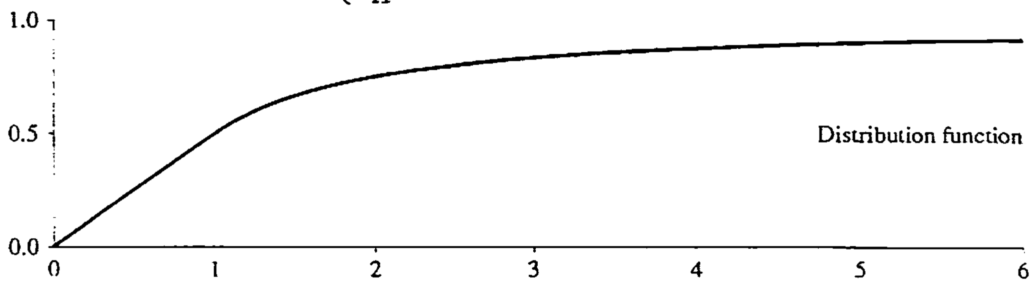
c) $P\left(\frac{1}{2} < \frac{X}{Y} < 2\right) = 1 - \left[P(\frac{1}{2} \ge \frac{X}{Y}) + P(\frac{X}{Y} \ge 2)\right] = 1 - 2P(X \ge 2Y)$ by symmetry.
But $X$ has exponential distribution with rate $\lambda$, and $2Y$ has exponential distribution with rate $\lambda/2$, so by part (b) the desired probability is

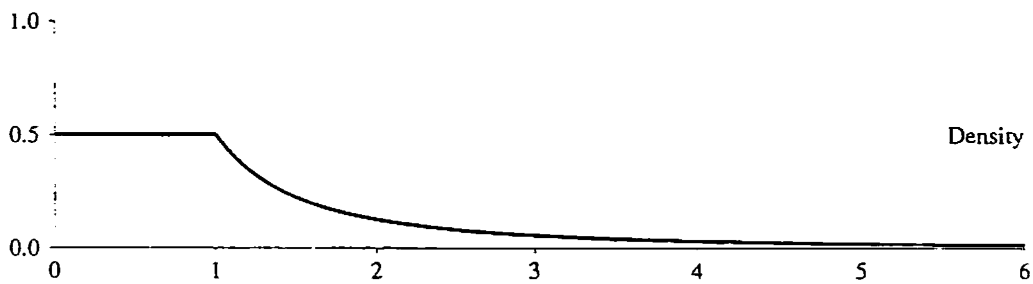
$$1 - 2 \cdot \frac{(1/2)\lambda}{\lambda + (1/2)\lambda} = \frac{1}{3}.$$

11. a) If $0 < x < 1$ then $P(X > x) = P(U/V > x) = P(U > xV) = 1 - (1/2)x$. [Draw a picture.]

b) $F_X(x) = \begin{cases} \frac{1}{2}x & 0 < x < 1 \\ 1 - \frac{1}{2x} & x > 1 \end{cases}$

c) $f_X(x) = \frac{d}{dx}F_X(x) = \begin{cases} \frac{1}{2} & 0 < x < 1 \\ \frac{1}{2x^2} & x > 1 \end{cases}$




12. The distance from the first marksman's shot to the center is

$$\sqrt{X^2 + Y^2} = a\sqrt{\left(\frac{X}{a}\right)^2 + \left(\frac{Y}{a}\right)^2} = aR$$

and the distance from the second marksman's shot to the center is

$$\sqrt{(X')^2 + (Y')^2} = b\sqrt{\left(\frac{X'}{b}\right)^2 + \left(\frac{Y'}{b}\right)^2} = bR'$$

where $R$, $R'$ are independent Rayleigh random variables. The desired probability is therefore

$$\begin{aligned} P(bR' < aR) &= P(R > \frac{b}{a}R') \\ &= \int_{y=0}^{\infty} \int_{x=\frac{b}{a}y}^{\infty} f_R(x) f_{R'}(y)\,dx\,dy \\ &= \int_{y=0}^{\infty} f_{R'}(y) P\left(R > \frac{b}{a}y\right) dy \\ &= \int_{y=0}^{\infty} y e^{-\frac{1}{2}y^2} e^{-\frac{1}{2}(\frac{b}{a}y)^2} dy \\ &= \int_{y=0}^{\infty} y e^{-\frac{1}{2}[1+(\frac{b}{a})^2]y^2} dy = \frac{a^2}{a^2+b^2}. \end{aligned}$$

Intuitively, if $a << b$ then the desired probability should be small.

13. Let $\Theta \in [0, \pi]$ be the angle subtended by the (shorter) arc between $(X, Y)$ and the point $(0, 1)$. Then $\Theta$ has uniform distribution on $[0, \pi]$, and $X = \cos\Theta$.

a) If $|x| \le 1$ then

$$F_X(x) = P(X \le x) = P(\cos\Theta \le x) = P(\Theta \ge \arccos(x)) = 1 - \frac{1}{\pi}\arccos(x),$$

where arccos has range $[0, \pi]$.

b) $Y$ has the same distribution function as $X$.

c) $P(X+Y \le z) = P(X \le z/\sqrt{2})$ by rotational symmetry of the uniform distribution on the circle. So

$$F_{X+Y}(z) = F_X(z/\sqrt{2}) = \begin{cases} 0 & \text{if } z < -\sqrt{2} \\ 1 - \frac{1}{\pi}\arccos\frac{z}{\sqrt{2}} & \text{if } |z| \le \sqrt{2} \\ 1 & \text{if } z > \sqrt{2} \end{cases}$$

14. a) $1/6$, by symmetry: all $3!$ orderings of $U_1, U_2, U_3$ are equally likely.

b) $E(U_1U_2U_3) = E(U_1)\cdot E(U_2)\cdot E(U_3) = [E(U_1)]^3 = \frac{1}{8}$ .

c) $Var(U_1U_2U_3) = E[(U_1U_2U_3)^2] - [E(U_1U_2U_3)]^2 = [E(U_1^2)]^3 - (\frac{1}{8})^2 = \frac{1}{27} - \frac{1}{64}$ .

d) If $0 < z < 1$ then $P(U_1U_2 > z) = \int_{x=z}^{1}(1 - \frac{z}{x})dx = 1 - z + z\log z$, so

$$\begin{aligned} P(U_1U_2 > U_3) &= \iint_{z>u} f_{U_1U_2}(z) f_{U_3}(u)\,dz\,du \\ &= \int_{u=-\infty}^{\infty}\int_{z=u}^{\infty} f_{U_1U_2}(z) f_{U_3}(u)\,dz\,du \\ &= \int_{u=-\infty}^{\infty} P(U_1U_2 > u) f_{U_3}(u)\,du = \int_0^1 (1 - u + u\log u)\,du = \frac{1}{4}. \end{aligned}$$

e) $P(\max(U_1, U_2) > U_3) = 1 - P(\max(U_1, U_2) < U_3) = 1 - P(U_3 \text{ is largest}) = 1 - \frac{1}{3} = \frac{2}{3}$

15. a) $P(Z_1 < Z_2 < Z_3) = 1/6$ by the same argument as in (a).

b) $E(Z_1Z_2Z_3) = E[(Z_1)^3] = 0$.

c) $Var(Z_1Z_2Z_3) = [E(Z_1^2)]^3 - [E(Z_1)]^6 = [Var(Z_1)]^3 - 0 = 1$.

d) $P(Z_1Z_2 > Z_3) = 1/2$, since $Z_1Z_2 - Z_3$ is a symmetric (about 0) random variable .

e) $P(\max(Z_1, Z_2) > Z_3) = 2/3$ by the same argument as in (e) .

f) $P(Z_1^2 + Z_2^2 > 1) = e^{-\frac{1}{2}}$ by Section 5.3.

g) $Z_1 + Z_2 + Z_3$ is normally distributed with mean 0 and variance 3, so the desired probability is

$$P\left(\frac{Z_1 + Z_2 + Z_3}{\sqrt{3}} < \frac{2}{\sqrt{3}}\right) = \Phi(2/\sqrt{3}) = 0.8759.$$

h) $P(Z_1/Z_2 < 1) = P(Z_1 < Z_2, Z_1 > 0) + P(Z_1 < Z_2, Z_1 < 0) = \frac{3}{4}$ , by the rotational symmetry of the joint density (draw a picture).

i) $3Z_1 - 2Z_2 - 4Z_3$ is normally distributed with mean 0 and variance 29, so the desired probability is

$$P\left(\frac{3Z_1 - 2Z_2 - 4Z_3}{\sqrt{29}} < \frac{1}{\sqrt{29}}\right) = \Phi(1/\sqrt{29}) = 0.5737.$$

16. a) $R^2$ has the chi-square distribution with 3 degrees of freedom, so it also has the gamma$(3/2, 1/2)$ distribution. Hence it has density

$$f_{R^2}(t) = \frac{1}{\sqrt{2\pi}}\sqrt{t}e^{-\frac{t}{2}} \qquad (t > 0)$$

b) By the change of variable formula,

$$f_R(r) = \frac{1}{(1/2\sqrt{r^2})} f_{R^2}(r^2) = \frac{2r}{\sqrt{2\pi}}\sqrt{r^2}e^{-\frac{r^2}{2}} = \sqrt{\frac{2}{\pi}}r^2 e^{-\frac{r^2}{2}} \qquad (r > 0)$$

c) Letting $u = r^2/2$,

$$E(R) = \int_0^\infty r\sqrt{\frac{2}{\pi}}r^2 e^{-\frac{r^2}{2}}\,dr = \sqrt{\frac{2}{\pi}}\int_0^\infty 2u^{2-1}e^{-u}\,du = 2\sqrt{\frac{2}{\pi}}\Gamma(2) = 2\sqrt{\frac{2}{\pi}}$$

d)

$$\begin{aligned} E(R^2) &= E(X^2+Y^2+Z^2) = 3 \\ Var(R^2) &= 3 - (2\sqrt{2/\pi})^2 \\ &= 3 - \frac{8}{\pi} \end{aligned}$$

**17.** $X_1^2, X_2^2, \ldots, X_{100}^2$ are independent identically distributed random variables having mean 1 and variance 2. [In fact, $X^2$ has gamma (1/2, 1/2) distribution: see the section on gamma densities.] Hence the sum has mean 100 and variance 200; and by the normal approximation, the sum is approximately normally distributed.

a) So if $Z$ denotes a standard normal variable, then

$$\begin{aligned} P(X_1^2+\cdots+X_{100}^2 \ge 80) &= P\left(\frac{\text{sum}-100}{\sqrt{200}} \ge \frac{80-100}{\sqrt{200}}\right) \\ &\approx P(Z \ge -\sqrt{2}) = \Phi(\sqrt{2}) \approx .92. \end{aligned}$$

b) Find $c$ so that

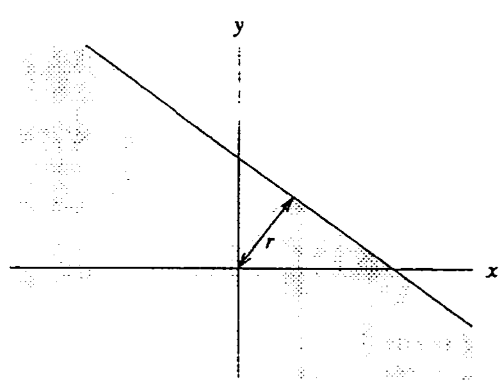
$$\begin{aligned} .95 &= P(100 - c \le X_1^2+\cdots+X_{100}^2 \le 100+c) \\ &= P\left(-\frac{c}{\sqrt{200}} \le \frac{\text{sum}-100}{\sqrt{200}} \le \frac{c}{\sqrt{200}}\right) \\ &\approx P(|Z| \le c/\sqrt{200}) = 2\Phi(c/\sqrt{200}) - 1 \end{aligned}$$

$$\iff \Phi(c/\sqrt{200}) = .975 \iff \frac{c}{\sqrt{200}} = 1.96 \iff c = 27.7.$$

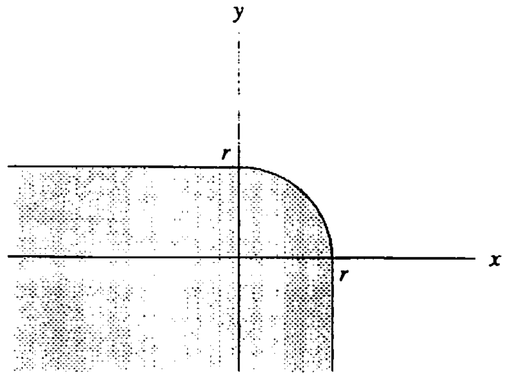
**18.** Let $r > 0$. Note that if $a, b \ge 0$ then the event

$$aX + bY \le r\sqrt{a^2+b^2}$$

occurs if and only if $(X, Y)$ lies in the shaded region:




where the line has slope $-a/b$ and the distance from the origin to the line is $r$. Hence the desired event occurs if and only if $(X, Y)$ lies in the region formed by intersecting all such regions:




This region is the union of 3 (infinite) "rectangles" and one quarter circle. The desired probability is then

$$P(X<0, Y<0)+P(0<X<r, Y<0)+P(X<0, 0<Y<r)+\frac{1}{4} P(\sqrt{X^2+Y^2} \leq r),$$

the last term by rotational symmetry of $(X, Y)$. Use independence of $X$, $Y$, and the fact that $\sqrt{X^2+Y^2}$ has Rayleigh distribution to evaluate:

$$\begin{aligned}
& P(X<0) P(Y<0)+2 P(0<X<r) P(Y<0)+\frac{1}{4} P(R \leq r) \\
& =\Phi(0) \Phi(0)+2[\Phi(r)-\Phi(0)] \Phi(0)+\frac{1}{4}\left(1-e^{-\frac{1}{2} r^2}\right) \\
& =\frac{1}{4}+2\left(\Phi(r)-\frac{1}{2}\right) \frac{1}{2}+\frac{1}{4}\left(1-e^{-\frac{1}{2} r^2}\right)=\Phi(r)-\frac{1}{4} e^{-\frac{1}{2} r^2} .
\end{aligned}$$

**19.** a) $(K_1=k)=(W_k<\min_{i \neq k} W_i)$; b) $p_k=\lambda_k /(\lambda_1+\cdots \lambda_d)$; c) use the memoryless property of the exponential waiting times; d) the answer to g) must be $p_k$ by the law of large numbers; e) $\lambda_k T$; f) $(\lambda_1+\cdots \lambda_d) T$; g) $\lambda_k /(\lambda_1+\cdots \lambda_d)$.

**20.** a) If $t>0$ then

$P(T_{min}>t)=P(T_1>t, T_2>t)=P(T_1>t) P(T_2>t)=e^{-\lambda_1 t} e^{-\lambda_2 t}=e^{-(\lambda_1+\lambda_2) t}$

so $T_{min}$ has exponential distribution with rate $\lambda_1+\lambda_2$.

b) $P(X_{min}=1)=P(T_1<T_2)=\int_{y=0}^{\infty} \int_{x=y}^{\infty} \lambda_2 e^{-\lambda_2 x} \lambda_1 e^{-\lambda_1 y} d x d y=\int_{y=0}^{\infty} \lambda_1 e^{-\lambda_1 y} e^{-\lambda_2 y} d y= \frac{\lambda_1}{\lambda_1+\lambda_2}$.

Get $P(X_{min}=2)$ by interchanging 1 and 2, or by subtraction from 1.

c)

$$\begin{aligned}
P(X_{min}=1, T_{min}>t) & =P(T_1<T_2, T_1>t) \\
& =\int_{y=t}^{\infty} \int_{x=y}^{\infty} \lambda_2 e^{-\lambda_2 x} \lambda_1 e^{-\lambda_1 y} d x d y \\
& =\int_t^{\infty} \lambda_1 e^{-\lambda_1 y} e^{-\lambda_2 y} d y \\
& =\frac{\lambda_1}{\lambda_1+\lambda_2} e^{-(\lambda_1+\lambda_2) t} \\
& =P(X_{min}=1) P(T_{min}>t) .
\end{aligned}$$

Similarly $P(X_{min}=2, T_{min}>t)=P(X_{min}=2) P(T_{min}>t)$.

d) **Claim:** For each $n = 2, 3, 4, \ldots$:
If $T_1, T_2, \ldots, T_n$ are independent exponential random variables with rates $\lambda_1, \lambda_2, \ldots \lambda_n$ respectively, and if $T_{min}$ denotes the minimum of $T_1, \ldots$ and if $X_{min}$ denotes the such that $T_i = T_{min}$, then:

(i) The distribution of $T_{min}$ is exponential with rate $\lambda_1 + \cdots + \lambda_n$ ;
(ii) $P(X_{min} = i) = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n}$ ;
(iii) $T_{min}$ and $X_{min}$ are independent.

**Proof:** Induction. True for $n = 2$, by (a), (b), (c).
If holds for $n = k$, where $k \geq 2$, then: for each $i = 1, \ldots, k+1$, define
$U_i = \min\{T_j : 1 \leq j \leq k+1, j \neq i\}$.
That is, $U_i$ is the minimum of $k$ independent exponentials; so by the induc. hypothesis, $U_i$ has exponential distribution with parameter $(\sum_1^{k+1} \lambda_j - \lambda_i)$. Moreover, for each $i = 1, \ldots, k+1$, $T_{min}$ is the same as $\min(T_i, U_i)$, and $T_i$ and $U_i$ are independent.

(i) By (a), $T_{min} = \min(T_{k+1}, U_{k+1})$ has exponential distribution with rate $\sum_1^{k+1} \lambda_j$.
(ii) For each $i = 1, \ldots, k+1$:

$$P(X_{min} = i) = P(T_i < U_i) \overset{\text{by (b)}}{=} \frac{\lambda_i}{\lambda_i + (\sum_1^{k+1} \lambda_j - \lambda_i)}.$$

(iii) For each $i = 1, \ldots, k+1$: For each $t > 0$ :

$$P(X_{min} = i, T_{min} > t) = P(T_i < U_i, T_{min} > t)$$

$$\overset{\text{by (c)}}{=} P(T_i < U_i)P(T_{min} > t) = P(X_{min} = i)P(T_{min} > t).$$

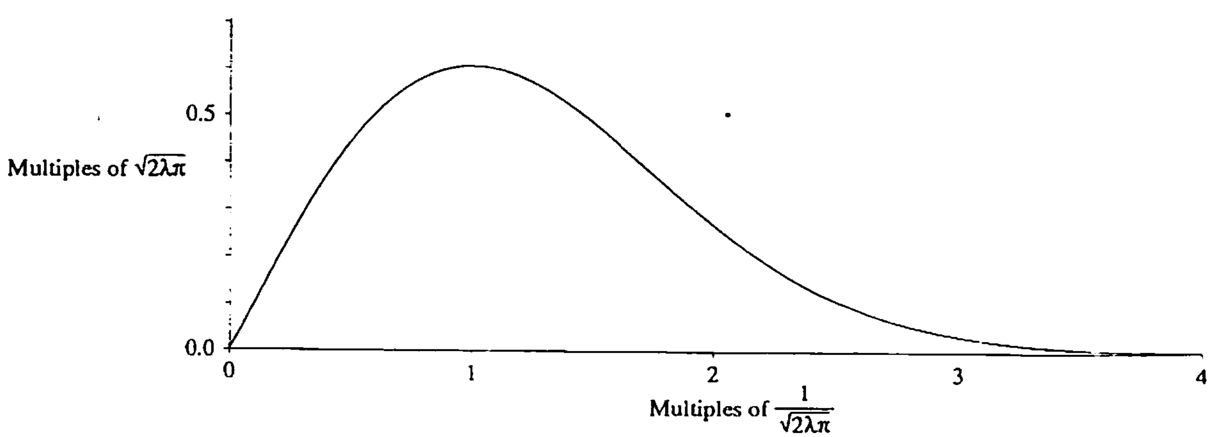
So claim holds for $n = k+1$ .

**21.** a) The distribution function is, for $r > 0$,

$$P(R \leq r) = 1 - P(R > r) = 1 - P(\text{no point inside circle of radius } r) = 1 - e^{-\lambda \pi r^2}.$$

The density is, for $r > 0$,

$$\frac{d}{dr}(R \leq r) = 2\lambda\pi r e^{-\lambda\pi r^2}$$




b) Write $R_1 = \sqrt{2\lambda\pi}R$. Then

$$P(R_1 \leq r_1) = P(R \leq r_1/\sqrt{2\lambda\pi}) = 1 - e^{-\frac{1}{2}(r_1)^2}$$

This is the distribution function of the Rayleigh distribution.

c) From Section 5.3 $E(R_1^2) = 2$, and $E(R_1) = \sqrt{\frac{\pi}{2}}$. (See Example 1). So

$$E(R) = E(R_1)/\sqrt{2\lambda\pi} = \frac{1}{2\sqrt{\lambda}}.$$

$$SD(R) = SD(R_1)/\sqrt{2\lambda\pi} = \frac{1}{2\sqrt{\lambda}}\sqrt{\frac{4-\pi}{\pi}}$$

d) The density is maximized at the mode. Suppose first $\sqrt{2\lambda\pi} = 1$, so $R = R_1$. The density is $re^{-\frac{1}{2}r^2}$, and

$$\frac{d}{dr}re^{-\frac{1}{2}r^2} = e^{-\frac{1}{2}r^2} - r^2e^{-\frac{1}{2}r^2}$$

which is zero for $r = 1$, clearly giving a maximum. Thus $r = 1$ is the mode of $R_1$, so $r = 1/\sqrt{2\lambda\pi}$ is the mode of $R = R_1/\sqrt{2\lambda\pi}$.

To find the median value $r$, solve

$$P(R \le r) = 1/2 \iff 1 - e^{-\lambda\pi r^2} = 1/2 \iff r = \sqrt{\frac{\log 2}{\lambda\pi}}$$

22. a) Note that $\left(\frac{2}{m\sigma^2}K\right)^{1/2}$ is distributed like $W = \sqrt{X^2+Y^2+Z^2}$ , where $X, Y, Z$ are independent normal (0, 1) random variables. To find the density of $W$, duplicate the argument in the text for the density of $\sqrt{X^2+Y^2}$ :

The joint density of $X, Y, Z$ is

$$\phi(x)\phi(y)\phi(z) = \frac{1}{(2\pi)^{3/2}}e^{-\frac{1}{2}(x^2+y^2+z^2)}.$$

The event $(W \in dw)$ corresponds to $(X, Y, Z)$ falling in a spherical shell about the origin of infinitesimal width $dw$, radius $w$, surface area $4\pi w^2$ and volume $4\pi w^2 dw$. And $P(W \in dw)$ is the "volume" over this infinitesimal shell beneath the joint density. Over this shell, the joint density has nearly constant value $\frac{1}{(2\pi)^{3/2}}e^{-\frac{1}{2}w^2}$; conclude that

$$P(W \in dw) = 4\pi w^2\frac{1}{(2\pi)^{3/2}}e^{-\frac{1}{2}w^2}dw$$

and the density of $W$ is

$$f_W(w) = \sqrt{\frac{2}{\pi}}w^2e^{-\frac{1}{2}w^2}, \quad w > 0.$$

The density of $K = \frac{m\sigma^2}{2}W^2$ is, by a change of variable,

$$\begin{aligned} f_K(y) &= \sqrt{\frac{2}{m\sigma^2}}\cdot\frac{1}{2\sqrt{y}}\cdot f_W\left(\sqrt{\frac{2y}{m\sigma^2}}\right) \\ &= \sqrt{\frac{2}{m\sigma^2}}\cdot\frac{1}{2\sqrt{y}}\sqrt{\frac{2}{\pi}}\frac{2y}{m\sigma^2}e^{-\frac{1}{2}\frac{2y}{m\sigma^2}} \\ &= \frac{2}{\sqrt{\pi}}\frac{1}{(m\sigma^2)^{3/2}}y^{1/2}e^{-y/m\sigma^2}, y > 0. \end{aligned}$$

b) $E(K) = \frac{m}{2}E(V_x^2) + \frac{m}{2}E(V_y^2) + \frac{m}{2}E(V_z^2) = \frac{3m}{2}E(V_x^2) = \frac{3m}{2}Var(V_x) = \frac{3m}{2}\sigma^2$. For the mode of the distribution, note that

$$f_K(y) = \text{const}\cdot y^{1/2}e^{-y/m\sigma^2};$$

$$f_K{}'(y) = \text{const}\cdot\left[\tfrac{1}{2}y^{-1/2}e^{-y/m\sigma^2} - y^{1/2}\frac{1}{m\sigma^2}e^{-y/m\sigma^2}\right] = \text{const}\cdot y^{-1/2}e^{-y/m\sigma^2}\left(\tfrac{1}{2} - \frac{y}{m\sigma^2}\right).$$

The derivative of the density equals zero only at $y = \frac{1}{2}m\sigma^2$ . This value of $y$ must correspond to a global maximum (why?), so the mode of the energy distribution is $\frac{1}{2}m\sigma^2$.

23. a) $\frac{P(Y>1/3, Z<2/3)}{P(Y\geq 1/3)} = \frac{(1/3)^3}{(2/3)^3} = 1/8$.

b) $\frac{P(Z<2/3)-P(Y>1/3,Z<2/3)}{1-P(Y\geq 1/3)} = 7/19$.

24. a) Assume the center of the coin is uniformly distributed relative to the grid. The probability is then the ratio of relevant areas: $\frac{(s-d)^2}{s^2}$

b) Assuming also that the coin is fair, and lands heads or tails independent of where its center is located: $\frac{1}{2} \cdot \frac{(s-d)^2}{s^2}$

c) Assume the tosses are independent of each other, with assumptions as above for each toss. Then $X$ and $Y$ are independent, $X$ has binomial $(n = 4, p = 1/4)$ distribution, and $Y$ has binomial $(n = 4, p = 1/2)$ distribution. Compute probabilities by considering all possible pairs of values of $(X, Y)$ and adding over the appropriate pairs: $P(X = Y) = 886/4096$

d) $P(X < Y) = 2685/4096$

e) $P(X > Y) = 525/4096$ (subtract the others from 1).

25. a) As in Example 5.3.3, for $0 < x < y < 1$,

$$\begin{aligned}
&P(V_k \in dx, V_m \in dy)\\
&= P(k-1 \text{ in } (0,x), 1 \text{ in } dx, m-k-1 \text{ in } (x,y), 1 \text{ in } dy, n-m-1 \text{ in } (y,1))\\
&= \frac{n!x^{k-1}dx(y-x)^{m-k-1}dy(1-y)^{n-m-1}}{(k-1)!(m-k-1)!(n-m-1)!}
\end{aligned}$$

Dropping the $dx$ and $dy$ leaves the joint density for $0 < x < y < 1$. The joint density is zero elsewhere.

b) beta$(m-k, n-m_k+1)$

c) beta$(k, m-k+1)$

26. See solution to Exercise 6.Rev.30.

27. No Solution

28. No Solution

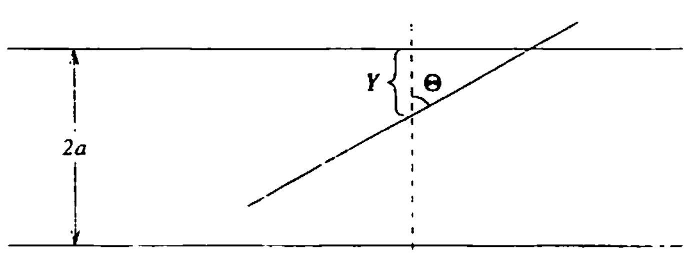
29. a) Say that the spacing between the parallel lines is $2a$. Let $Y$ be the distance from the center of the needle to the closest line. Let $\Theta$ be the (acute) angle between the needle and a line running perpendicular to the parallel lines. Then $(Y, \Theta)$ is uniformly distributed on the rectangle

$$R = \{(y, \theta) : 0 \leq y \leq a, 0 \leq \theta \leq \frac{\pi}{2}\}.$$

For all positive $x$ we have

$$\{X \leq x\} = \{\cos\Theta \geq \frac{Y}{x}\} = \{Y \leq x\cos\Theta\}.$$

(Fix the value of $Y$ first; then decide what $\Theta$ can be.)




Hence if $0 < x \leq a$ then

$$P(X \leq x) = \frac{\int_0^{\pi/2} x\cos\theta d\theta}{\text{area of } R}$$

$$= \frac{x}{\frac{\pi}{2}a} = \frac{2}{\pi a}x$$

and if $x \geq a$ then

$$P(X \leq x) = \frac{\int_0^{\arccos\frac{a}{x}} a d\theta + \int_{\arccos\frac{a}{x}}^{\pi/2} x\cos\theta d\theta}{\text{area of } R}$$
$$= \frac{2}{\pi a}\left[a \arccos\frac{a}{x} + x - \sqrt{x^2-a^2}\right],$$

where arccos has domain $[0,1]$ and range $[0, \frac{\pi}{2}]$ .

b)

$$fx(x) = \begin{cases} \frac{2}{\pi a} & \text{if } 0 < x \leq a \\ \frac{2}{\pi a}\left[1 - \sqrt{\frac{x^2-a^2}{x}}\right] & \text{if } x \geq a \end{cases}$$

30. a) $X_n$ is the sum of $n$ independent uniform$(-1,1)$ variables, with mean 0 and variance 1/3. So $X_n$ has mean 0 and SD $\sqrt{n/3} = 10$ for $n = 300$. Using the normal appoximation, $P(|X_n| > 10) \approx$ 32% for $n = 300$.

b) Same as a): 32%

c) This event occurs if and only if both the events in a) and b) occur. Because the components of a point picked at random from a square are independent, the $n$ uniform variables whose sum is $X_n$ are independent of the $n$ uniform variables whose sum is $Y_n$. There for $X_n$ and $Y_n$ are independent, and the required probability is

$$P(|X_n| > 10 \text{ and } |Y_n| > 10) = P(|X_n| > 10)P(|Y_n| > 10) \approx (0.32)^2$$

d) From the previous part, for $n = 300$ the joint distribution of $X_n/10$ and $Y_n/10$ is approximately that of $(X, Y)$ where $X$ and $Y$ are independent standard normal. Let $R = \sqrt{X^2+Y^2}$. Then the required probability is

$$P(X_n^2 + Y_n^2 > 100) = P((X_n/10)^2 + (Y_n/10)^2 > 1) \approx P(X^2+Y^2 > 1) = P(R^2 > 1) = e^{-1/2}$$

31. a) For large $n$, $X_n$ in this problem as well as $X_n$ in the previous problem will be approximately normal. It is also clear that they will have the same mean, so it remains only to find $r$ so that they have the same variance, or, equivalently, that $E[X_1^2]$ is the same in both problems. For the previous problem,

$$E[X_1^2] = \int_{-1}^{1} x^2 \frac{1}{2} dx = \frac{1}{6}x^3\Big|_{-1}^{1} = \frac{1}{3}$$

And for the current problem, if we consider $R^2 = X_1^2 + Y_1^2$, and we observe that $E[X_1^2] = E[Y_1^2]$, then we see that $E[X_1^2] = \frac{1}{2}E[R^2]$.

$$E[R^2] = \frac{1}{\pi r^2}\int_0^r x^2 2\pi x dx = 2\frac{1}{4r^2}x^4\Big|_0^r = \frac{r^2}{2}$$

And so $E[X_1^2] = \frac{r^2}{4}$, but we want this to be equal to $\frac{1}{3}$, so $r = \sqrt{(\frac{4}{3})}$.

b) No, since $0 = P(Y_n = \frac{nr}{2} | X_n = nr) \neq P(Y_n = \frac{nr}{2} | X_n = 0) > 0$.

c) For both the circle and the square, we let $R = X_n \cos\theta + Y_n \sin\theta$, and we have

$$E(R) = \cos\theta E(X_n) + \sin\theta E(Y_n) = 0$$

and

$$E(R^2) = \cos^2\theta E(X_n^2) + \sin\theta\cos\theta E(X_n Y_n) + \sin^2 E(Y_n^2) = E(X_n^2)$$

since $E(X_n Y_n) = 0$ by symmetry and $E(X_n^2) = E(Y_n^2)$. Now from a) we know that for large $n$ the distribution of $X_n$ from the squares is nearly the same as $X_n$ from the circles, so $R = X_n\cos\theta + Y_n\sin\theta$ must have nearly the same distribution for both cases.

d) Since the distribution of $X_n \cos\theta + Y_n \sin\theta$ is approximately normal, we can use the result to see that the answer must be the same as the answer to part d) of the previous problem, using squares, which is $e^{-1/2}$.

**32.** a) $r = \sqrt{2/3}$ b) no, $X_1$ and $Y_1$ are not independent d) $e^{-1/2}$

**33.** a)$r = \sqrt{1/2}$ b) no, $X_1$ and $Y_1$ are not independent d) $e^{-1/2}$ g) $2^{-n}$

**34.** No Solution

# Section 6.1

1. a) The unconditional distribution of $X$ is binomial (3, 1/2):

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | 1/8 | 3/8 | 3/8 | 1/8 |

b) Given $X = x$, $Y$ is distributed like $x$ plus a binomial $(3 - x, 1/2)$ variable. That is, $Y$ has the binomial $(3 - x, 1/2)$ distribution shifted to $\{x, x + 1, \ldots 3\}$.

| $y$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(Y = y \mid X = 0)$ | 1/8 | 3/8 | 3/8 | 1/8 |

| $y$ | 1 | 2 | 3 |
|---|---|---|---|
| $P(Y = y \mid X = 1)$ | 1/4 | 1/2 | 1/4 |

| $y$ | 2 | 3 |
|---|---|---|
| $P(Y = y \mid X = 2)$ | 1/2 | 1/2 |

| $y$ | 3 |
|---|---|
| $P(Y = y \mid X = 3)$ | 1 |

c) Use $P(X = x, Y = y) = P(X = x)P(Y = y \mid X = x)$:

*Joint distribution table for* $(X, Y)$

| | | *values x for X* | | | | *Marginal distn of Y* |
|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | |
| *values y for Y* | 0 | $\frac{1}{8} \cdot \frac{1}{8}$ | 0 | 0 | 0 | $\frac{1}{64}$ |
| | 1 | $\frac{1}{8} \cdot \frac{3}{8}$ | $\frac{3}{8} \cdot \frac{1}{4}$ | 0 | 0 | $\frac{9}{64}$ |
| | 2 | $\frac{1}{8} \cdot \frac{3}{8}$ | $\frac{3}{8} \cdot \frac{1}{2}$ | $\frac{3}{8} \cdot \frac{1}{2}$ | 0 | $\frac{27}{64}$ |
| | 3 | $\frac{1}{8} \cdot \frac{1}{8}$ | $\frac{3}{8} \cdot \frac{1}{4}$ | $\frac{3}{8} \cdot \frac{1}{2}$ | $\frac{1}{8} \cdot 1$ | $\frac{27}{64}$ |
| *Marginal distn of X* | | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ | 1 |

d) Use $P(Y = y) = \sum_x P(X = x, Y = y)$ and the result of c):

| $y$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(Y = y)$ | 1/64 | 9/64 | 27/64 | 27/64 |

e) Use $P(X = x \mid Y = y) = P(X = x, Y = y)/P(Y = y)$:

| $x$ | 0 |
|---|---|
| $P(X = x \mid Y = 0)$ | 1 |

| $x$ | 0 | 1 |
|---|---|---|
| $P(X = x \mid Y = 1)$ | 1/3 | 2/3 |

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X = x \mid Y = 2)$ | 1/9 | 4/9 | 4/9 |

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x \mid Y = 3)$ | 1/27 | 2/9 | 4/9 | 8/27 |

f) For each $y = 0, 1, 2, 3$: Guess the value $x$ that gives the highest probability $P(X = x|Y = y)$.

| $y$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Guess | 0 | 1 | 1 or 2 | 2 |

g)

$$
\begin{aligned}
P(\text{guess correctly}) &= \sum_y P(\text{guess correctly, } Y = y) \\
&= \sum_y P(\text{guess correctly} \mid Y = y) P(Y = y) \\
&= \frac{1}{64} + \frac{6}{64} + \frac{12}{64} + \frac{12}{64} = \frac{31}{64}.
\end{aligned}
$$

2. Condition on the value of $T$:

$$P(G = g) = \sum_{t=0}^{4} P(G = g|T = t)P(T = t)$$

Now given $T = t$, $G$ has binomial $(t, 1/2)$ distribution, so

$$P(G = g|T = t) = \binom{t}{g}(1/2)^t, \quad g = 0, \ldots, t.$$

Conclude:

$$
\begin{aligned}
P(G = 0) &= 1 \cdot 0.1 + 1/2 \cdot 0.2 + (1/2)^2 \cdot 0.4 + (1/2)^3 \cdot 0.2 + (1/2)^4 \cdot 0.1 = 0.33125 \\
P(G = 1) &= 0 + 1/2 \cdot 0.2 + 2 \cdot (1/2)^2 \cdot 0.4 + 3 \cdot (1/2)^3 \cdot 0.2 + 4 \cdot (1/2)^4 \cdot 0.1 = 0.4 \\
P(G = 2) &= 0 + 0 + (1/2)^2 \cdot 0.4 + 3 \cdot (1/2)^3 \cdot 0.2 + 6 \cdot (1/2)^4 \cdot 0.1 = 0.2125 \\
P(G = 3) &= 0 + 0 + 0 + (1/2)^3 \cdot 0.2 + 4 \cdot (1/2)^4 \cdot 0.1 = 0.05 \\
P(G = 4) &= 0 + 0 + 0 + 0 + (1/2)^4 \cdot 0.1 = 0.00625
\end{aligned}
$$

| $G$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(G = g)$ | 0.33125 | 0.4 | 0.2125 | 0.05 | 0.00625 |

3. a) Suppose there are $n$ families in all. Then there are a total of

$$0 \times .1 \times n + 1 \times .2 \times n + 2 \times .4 \times n + 3 \times .2 \times n + 4 \times .1 \times n = 2n$$

children in the town. There are $0 \times .1 \times n$ children who come from 0-child families, and $1 \times .2 \times n$ children who come from 1-child families, and $2 \times .4 \times n$ children who come from 2-child families, and so on. So the distribution of $U$ is given by:

| $U$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(U = u)$ | $0/2n = 0$ | $.2n/2n = .1$ | $.8n/2n = .4$ | $.6n/2n = .3$ | $.4n/2n = .2$ |

The distribution of $T$ is different from the distribution of $U$ because we are choosing children at random rather than families at random; so families with many children are more likely than before.

b) The number of families having two girls and a boy is $.2 \times n \times (3/8)$ (This is $n$ times the chance that a family picked at random has two girls and a boy.) So there are three times that many children who come from a family consisting of two girls and a boy. The desired probability is therefore $\frac{3 \times .2 \times n \times (3/8)}{2n} = .1125$.

Or: Let $A$ = (child's family has 3 kids), $B$ = (child's family has two girls and a boy). We want

$$P(AB) = P(A)P(B|A) = .3 \times \frac{3}{8} = .1125.$$

c) The probability that a family picked at random has two girls and a boy is $(.2)(3/8) = .075$. This is different from the probability in b) because the two experiments are different.

4. $\frac{\binom{10}{5}\binom{10}{7}}{\binom{20}{12}}$

5. a) Given $X_1 + X_2 = n$, the possible values of $X_1$ are $0, 1, \ldots, n$. For $0 \le k \le n$, justify the following:

$$\begin{aligned} P(X_1 = k | X_1 + X_2 = n) &= \frac{P(X_1 = k)P(X_2 = n-k)}{P(X_1 + X_2 = n)} \\ &= \frac{e^{-\lambda_1}\frac{(\lambda_1)^k}{k!} e^{-\lambda_2}\frac{(\lambda_2)^{n-k}}{(n-k)!}}{e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}} \\ &= \frac{n!}{k!(n-k)!}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k} \end{aligned}$$

So given $X_1 + X_2 = n$, $X_1$ has binomial $\left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$ distribution.

b) Assume that the number of eggs laid by one insect is independent of the number laid by the other; and that they both lay the same expected number of eggs. Letting $X_i$ represent the number of eggs laid by insect $i$, we have $X_1$, $X_2$ independent, each with Poisson $(\lambda)$ distribution for some $\lambda$. By a),

$$P(X_1 \ge 90 | X_1 + X_2 = 150) = \sum_{k=90}^{150} \binom{150}{k} (1/2)^{150}$$

To evaluate approximately, use the normal approximation to the binomial. Here $\mu = np = 75$, $\sigma = \sqrt{npq} = 6.124$ and the desired probability is approximately

$$1 - \Phi\left(\frac{89.5 - 75}{6.124}\right) = .0089.$$

6. a) The probability is only positive when $i_1 + \cdots + i_m = n$.
Let $N = N_1 + \cdots + N_m$ and $\lambda = \lambda_1 + \cdots + \lambda_m$.

$$\begin{aligned} P(N_1 = i_1, \ldots, N_m = i_m \mid N = n) &= \frac{P(N_1 = i_1, \ldots, N_m = i_m, N = n)}{P(N = n)} \\ &= \prod_{k=1}^{m} P(N_k = i_k) \Big/ \frac{e^{-\lambda}(\lambda)^n}{n!} \\ &= \prod_{k=1}^{m}\left(\frac{e^{-\lambda_k}\lambda_k^{i_k}}{i_k!}\right) \times \frac{n!}{e^{-\lambda}(\lambda)^n} \\ &= \left(\frac{n!}{i_1!\cdots i_m!}\right)\left(\frac{\lambda_1}{\lambda}\right)^{i_1}\cdots\left(\frac{\lambda_m}{\lambda}\right)^{i_m} \end{aligned}$$

So the conditional distribution is multinomial $\left(n, \frac{\lambda_1}{\sum \lambda_i}, \ldots, \frac{\lambda_m}{\sum \lambda_i}\right)$

b) Assuming that $\lambda = \lambda_1 + \cdots + \lambda_m$,

$$\begin{aligned} P(N_1 = i_1, \ldots, N_m = i_m) &= \sum_{n=0}^{\infty} P(N_1 = i_1, \ldots, N_m = i_m \mid N = n)P(N = n) \\ &= \left(\frac{(i_1 + \cdots + i_m)!}{i_1!\cdots i_m!}\right)\left(\frac{\lambda_1}{\lambda}\right)^{i_1}\cdots\left(\frac{\lambda_m}{\lambda}\right)^{i_m}\frac{e^{-\lambda}\lambda^{\sum i_k}}{(\sum i_k)!} \\ &= \frac{\lambda_1^{i_1}\cdots\lambda_m^{i_m}}{i_1!\cdots i_m!}e^{-(\lambda_1+\cdots+\lambda_m)} = \prod_{k=1}^{m}\frac{e^{-\lambda_k}\lambda_k^{i_k}}{i_k!} \end{aligned}$$

The unconditional distribution is independent Poisson$(\lambda_k)$.
In general, even if $\lambda \ne \lambda_1 + \cdots + \lambda_m$, a similar calculation shows that the $N_k$ are unconditionally independent Poisson$(\lambda\lambda_k/(\lambda_1 + \cdots + \lambda_m))$.

7. a) For each $k \geq 0$:

$$
\begin{aligned}
P(X = k) &= \sum_{n=k}^{\infty} P(X = k, N = n) \\
&= \sum_{n=k}^{\infty} P(N = n)P(X = k|N = n) \\
&= \sum_{n=k}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\
&= e^{-\lambda}\frac{p^k}{k!}\cdot\lambda^k\sum_{n=k}^{\infty}\frac{\lambda^{n-k}(1-p)^{n-k}}{(n-k)!} \\
&= e^{-\lambda}\frac{(p\lambda)^k}{k!}e^{\lambda(1-p)} = e^{-\lambda p}\frac{(\lambda p)^k}{k!}
\end{aligned}
$$

So $X$ has Poisson $(\lambda p)$ distribution.

b) Let $N$ denote the number of breakages, and $X$ the number of healed breakages. So $N$ has Poisson $(.4)$ distribution. And given $N = n$, $X$ has binomial $(n, .2)$ distribution. By a) the unconditional distribution of $X$ is Poisson $(.08)$. So the required probability is $e^{-.08}\frac{(.08)^4}{24} = .0000016$.

8. The event $(X = x, Y = y)$ is contained in the event $(N = x + y)$. Hence

$$(X = x, Y = y) = (N = x + y, X = x, Y = y)$$

Now for $x \geq 0, y \geq 0$,

$$
\begin{aligned}
P(X = x, Y = y) &= P(N = x + y, X = x, Y = y) \\
&= P(N = x + y)P(X = x, Y = y|N = x + y) \\
&= e^{-\lambda}\frac{\lambda^{x+y}}{(x+y)!}\frac{(x+y)!}{x!y!}(1/6)^x(5/6)^y \\
&= \left[e^{-\lambda/6}\frac{(\lambda/6)^x}{x!}\right]\left[e^{-5\lambda/6}\frac{(5\lambda/6)^y}{y!}\right] \\
&= P(X = x)P(Y = y)
\end{aligned}
$$

because by the previous exercise, $X$ has Poisson $(\lambda/6)$ distribution, and $Y$ has Poisson $(5\lambda/6)$ distribution. So $X$ and $Y$ are independent.

9. Write (for example) $P(y|x, z)$ for $P(Y = y|X = x, Z = z)$.

$$
\begin{aligned}
& P(x, y|z) = P(x|z)P(y|z) \\
\Longleftrightarrow\quad & \frac{P(x,y,z)}{P(z)} = \frac{P(x,z)}{P(z)}\frac{P(y,z)}{P(z)} \quad \text{(definition of conditional probability)} \\
\Longleftrightarrow\quad & \frac{P(x,y,z)}{P(x,z)} = \frac{P(y,z)}{P(z)} \quad \text{(cross multiply)} \\
\Longleftrightarrow\quad & P(y|x,z) = P(y|z) \quad \text{(definition of conditional probability)}
\end{aligned}
$$

The symmetry of the first equation indicates that we may obtain a further equivalent condition by interchanging $x$ and $y$ in the last equation:

$$P(x|y, z) = P(x|z).$$

10. a) By conditional independence,

$$
\begin{aligned}
P(R = r, X_1 = x_1, X_2 = x_2) &= P(R = r)P(X_1 = x_1 \mid R = r)P(X_2 = x_2 \mid R = r) \\
&= \pi_r\binom{n}{x_1}(r/10)^{x_1}(1 - r/10)^{n-x_1}\binom{n}{x_2}(r/10)^{x_2}(1 - r/10)^{n-x_2}
\end{aligned}
$$

b)

$$P(R=r|X_1=x_1) = P(R=r, X_1=x_1)\Big/P(X_1=x_1)$$

$$= \pi_r\binom{n}{x_1}(r/10)^{x_1}(1-r/10)^{n-x_1}\Big/\sum_{s=0}^{10}\pi_s\binom{n}{x_1}(s/10)^{x_1}(1-s/10)^{n-x_1}$$

c)

$$P(X_2=x_2|R=r, X_1=x_1) = P(X_2=x_2|R=r) \qquad \text{(by conditional independence)}$$

$$= \binom{n}{x_2}(r/10)^{x_2}(1-r/10)^{n-x_2}$$

d) $P(X_2=x_2|X_1=x_1) = \sum_r P(R=r|X_1=x_1)P(X_2=x_2|R=r, X_1=x_1)$

e) If $\pi_r = 1/11, n=1$: Put $x_2 = x_1 = 1$ in b) and c):

$$P(R=r|X_1=1) = (1/11)(r/10)\Big/\sum_{s=0}^{10}(1/11)(s/10) = r/55;$$

$$P(X_2=1|R=r, X_1=1) = r/10.$$

So

$$P(X_2=1|X_1=1) = \sum_{r=0}^{10}\frac{r}{55}\cdot\frac{r}{10} = 385/550 = .7.$$

Similarly

$$P(X_2=1|X_1=0) = \sum_{r=0}^{10}\frac{10-r}{55}\cdot\frac{r}{10} = .3.$$

By taking complements, see $P(X_2=0|X_1=1) = .3$; and $P(X_2=0|X_1=0) = .7$. By symmetry, $P(X_2=1) = 1/2$. So $X_1$ and $X_2$ are not independent: $P(X_2=1|X_1=1) \neq P(X_2=1)$.

# Section 6.2

1. a) For each $x = 1$ to 6 we have

$$P(X = x) = P(X \geq x) - P(X \geq x+1) = (\frac{6-x+1}{6})^2 - (\frac{6-x}{6})^2 = \frac{13-2x}{36}$$

$$P(Y = y|X = x) = \begin{cases} \frac{P(X=x,Y=y)}{P(X=x)} = \frac{1/36}{(13-2x)/36} = \frac{1}{13-2x} & \text{if } y = x \\ \frac{P(X=x,Y=y)}{P(X=x)} = \frac{2/36}{(13-2x)/36} = \frac{2}{13-2x} & \text{if } y > x. \end{cases}$$

$$E(Y|X = x) = \frac{x}{13-2x} + \sum_{y=x+1}^{6} y \cdot \frac{2}{13-2x}$$

$$= \frac{1}{13-2x}\left[x + 2 \cdot \frac{(x+7)(6-x)}{2}\right] = \frac{42-x^2}{13-2x}.$$

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $E(Y\|X = x)$ | 41/11 | 38/9 | 33/7 | 26/5 | 17/3 | 6/1 |

b) For each $y = 1$ to 6 we have

$$P(Y = y) = P(Y \leq y) - P(Y \leq y-1) = (\frac{y}{6})^2 - (\frac{y-1}{6})^2 = \frac{2y-1}{36}$$

$$P(X = x|Y = y) = \begin{cases} \frac{P(X=y,Y=y)}{P(Y=y)} = \frac{1/36}{(2y-1)/36} = \frac{1}{2y-1} & \text{if } x = y \\ \frac{P(X=x,Y=y)}{P(Y=y)} = \frac{2/36}{(2y-1)/36} = \frac{2}{2y-1} & \text{if } x < y. \end{cases}$$

$$E(X|Y = y) = \sum_{x=1}^{y-1} x \cdot \frac{2}{2y-1} + \frac{y}{2y-1} = \frac{2}{2y-1} \cdot \frac{(y-1)y}{2} + \frac{y}{2y-1} = \frac{y^2}{2y-1}.$$

| $y$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $E(X\|Y = y)$ | 1 | 4/3 | 9/5 | 16/7 | 25/9 | 36/11 |

2. a) For each $x = 1$ to $n$ we have

$$P(Y = y|X = x) = \begin{cases} 0 & y < x \\ \frac{1}{2(n-x)+1} & y = x \\ \frac{2}{2(n-x)+1} & y > x \end{cases}$$

and

$$E(Y|X = x) = \sum_{y=1}^{n} yP(Y = y|X = x) = \frac{x}{2(n-x)+1} + \sum_{y=x+1}^{n} \frac{2y}{2(n-x)+1}$$

$$= \frac{x + 2\left(\frac{n(n+1)}{2} - \frac{x(x+1)}{2}\right)}{2(n-x)+1} = \frac{n(n+1) - x^2}{2(n-x)+1}.$$

b) For each $y = 1$ to $n$ we have

$$P(X = x|Y = y) = \begin{cases} \frac{2}{2y-1} & x < y \\ \frac{1}{2y-1} & x = y \\ 0 & x > y \end{cases}$$

and

$$E(X|Y = y) = \sum_{x=1}^{y-1} \frac{2x}{2y-1} + \frac{y}{2y-1} = \frac{(y-1)y}{2y-1} + \frac{y}{2y-1} = \frac{y^2}{2y-1}.$$

Notice that this conditional expectation is the same for all $n$.

3. Assume $n$ is at least 2. There are $\binom{n}{2}$ possible values of $(X, Y)$, all equally likely: they consist of all pairs of numbers of the form $(x, y)$, where $1 \le x < y \le n$ . So

$$P(X = x, Y = y) = \begin{cases} 1/\binom{n}{2} & \text{if } 1 \le x < y \le n \\ 0 & \text{otherwise} \end{cases}$$

$$P(X = x) = \sum_{y=x+1}^{n} P(X = x, Y = y) = \begin{cases} (n-x)/\binom{n}{2} & \text{if } 1 \le x \le n-1 \\ 0 & \text{otherwise} \end{cases}$$

$$P(Y = y) = \sum_{x=1}^{y-1} P(X = x, Y = y) = \begin{cases} (y-1)/\binom{n}{2} & \text{if } 2 \le y \le n \\ 0 & \text{otherwise.} \end{cases}$$

a) If $1 \le x \le n-1$ then $P(Y = y|X = x) = \frac{1}{n-x}$, $\quad x+1 \le y \le n$ , so

$$E(Y|X = x) = \sum_{\text{all } y} yP(Y = y|X = x) = \sum_{y=x+1}^{n} y\frac{1}{n-x} = \frac{1}{n-x}\left[\sum_{1}^{n} y - \sum_{1}^{x} y\right]$$
$$= \frac{1}{n-x}\left[\frac{n(n+1)}{2} - \frac{x(x+1)}{2}\right] = \frac{n+x+1}{2}.$$

Intuitively, given $X = x$, $Y$ has uniform distribution on $\{x+1, \ldots, n\}$ , so the expected value of $Y$ is the average of $x+1$ and $n$.

b) If $2 \le y \le n$ then $P(X = x|Y = y) = \frac{1}{y-1}$, $1 \le x \le y-1$, so

$$E(X|Y = y) = \sum_{\text{all } x} xP(X = x|Y = y) = \sum_{x=1}^{y-1} x\frac{1}{y-1} = \frac{(y-1)(y)}{2(y-1)} = \frac{y}{2}.$$

Intuitively, given $Y = y$, $X$ has uniform distribution on $\{1, \ldots, y-1\}$ , so the expected value of $X$ is the average of 1 and $y-1$.

4. a) $E(Y) = E[E(Y \mid X)] = E\left(\frac{1+X}{2}\right) = \frac{1}{2} + \frac{1}{2}\left(\frac{1+n}{2}\right) = \frac{n+3}{4}$

b)

$$E(Y^2) = E[E(Y^2 \mid X)] = E\left(\sum_{y=1}^{X} y^2 \frac{1}{X}\right)$$
$$= E\left(\frac{X(2X+1)(X+1)}{6X}\right) = E\left(\frac{(2X+1)(X+1)}{6}\right) = \frac{1}{3}E(X^2) + \frac{1}{2}E(X) + \frac{1}{6}$$
$$= \frac{1}{3}\left(\frac{n(n+1)(2n+1)}{6n}\right) + \frac{1}{2}\left(\frac{1+n}{2}\right) + \frac{1}{6} = \frac{4n^2+15n+17}{36}$$

c) $Var(Y) = E(Y^2) - [E(Y)]^2 = \frac{7n^2+6n-13}{144}$, so $SD(Y) = \frac{\sqrt{7n^2+6n-13}}{12}$

d)

$$P(X+Y=2) = \sum_{k=1}^{n} P(X+Y=2 \mid X=x)P(X=x) = P(X+Y=2 \mid X=1)P(X=1) = \frac{1}{n}$$

5. a) Observe that given $A$, we have $X = X_1$, while given $A^c$, we have $X = X_2$.

$$\begin{aligned} F(x) = P(X \le x) &= P(X \le x|A)P(A) + P(X \le x|A^c)P(A^c) \\ &= P(X_1 \le x)P(A) + P(X_2 \le x)P(A^c) \\ &= F_1(x)p + F_2(x)(1-p). \end{aligned}$$

b) $E(X) = E(X|A)P(A) + E(X|A^c)P(A^c) = E(X_1)p + E(X_2)(1-p)$.

c) $E(X^2) = E(X^2|A)P(A) + E(X^2|A^c)P(A^c)$
$= [(E(X_1))^2 + Var(X_1)]p + [(E(X_1))^2 + Var(X_1)](1-p)$. Hence

$$Var(X) = E(X^2) - [E(\cdot \quad = Var(X_1)p + Var(X_2)(1-p) + p(1-p)(E(X_1) - E\cdot\cdot_2))^2$$

6. a) $P(\text{all } X\text{'s} < t \mid N = n) = t^{\cdot} \quad ) < t < 1)$

b)

$$P(\text{all } X\text{'s} < t) = \sum_{n=0}^{\infty} P(\text{all } X\text{'s} < t \mid N = n)P(N = n)$$
$$= \sum_{n=0}^{\infty} t^n \frac{e^{-\mu}\mu^n}{n!} = e^{-\mu}\sum_{n=0}^{\infty}\frac{(t\mu)^n}{n!}$$
$$= e^{-\mu}e^{t\mu} = e^{-(1-t)\mu}$$

c) $P(S_N = 0) = P(N = 0) = e^{-\mu}$, since in the event $N > 0$, the probability that all $X$'s will be 0 is 0.

.i) $E(S_N) = E[E(S_N \mid N)] = E\left(\frac{N}{2}\right) = \frac{\mu}{2}$

.. Condition on $N$ : If $1 \le k \le n$ then

$$E(S_N|N = k) = E(X_1 + \cdots + X_k|N = k) = E(X_1|N = k) + \cdots + E(X_k|N = k) = k\mu.$$

If $k = 0$ then $E(S_N|N = k) = 0$ . Hence $E(S_N|N = k) = k\mu, k = 0, \ldots, n$ , and

$$E(S_N) = \sum_{k=0}^{n} E(S_N|N = k)P(N = k) = \sum_{k=0}^{n} k\mu P(N = k) = \mu E(N).$$

8. Define $D_0 = 1$ and for $n = 1, 2, 3, \ldots$ let $D_n$ be the number of individuals in the $n$th generation. Then for each $n \ge 0$,

$$D_{n+1} = \begin{cases} X_1^{(n)} + \cdots + X_k^{(n)} & \text{if } D_n = k,\ k = 1, 2, \ldots \\ 0 & \text{if } D_n = 0 \end{cases}$$

where $X_1^{(n)}, X_2^{(n)}, \ldots$ are independent and identically distributed random variables having mean $\mu$. [ $X_i^{(n)}$ is the number of children produced by the $i$th individual of generation $n$; generation 0 corresponds to the original individual. We assume that for each $n$, each $X_i^{(n)}$ is independent of $D_n$. ]

**Claim:** $E(D_n) = \mu^n, n = 1, 2, 3, \ldots$.

**Proof:** Induction on $n$. True for $n = 1$, since $D_1 = X_1^{(1)}$. If true for $n = k$, where $k \ge 1$, then

$$E(D_{k+1}) = E(X_1^{(k)} + \cdots + X_{D_k}^{(k)})$$
$$= \mu \cdot E(D_k) \quad \text{by Exercise 7}$$
$$= \mu \cdot \mu^k = \mu^{k+1},$$

so claim holds for $n = k + 1$ .

9. a) Given $T_2 = j$, $T_1$ is distributed like the place at which the first good element appears in a random ordering of $j - 1$ elements, of which 1 is good; hence

$$E(T_1|T_2 = j) = 1 \times \left(\frac{j-1+1}{1+1}\right) = \frac{j}{2}.$$

b) Given $T_1 = j$, $T_2$ is distributed like $j$ plus the place at which the first good element appears in a random ordering of $N - j$ elements, of which $k - 1$ are good; hence

$$E(T_2|T_1 = j) = j + 1 \times \left(\frac{N-j+1}{k-1+1}\right) = j + \frac{N-j+1}{k}.$$

c) If $h < i$ : Given $T_i = j$, $T_h$ is distributed like the place at which the $h$th good element appears in a random ordering of $j-1$ elements, of which $i-1$ are good; hence

$$E(T_h|T_i = j) = h \times \left(\frac{j-1+1}{i-1+1}\right) = h\frac{j}{i}.$$

As a check, note that (a) is a special case.

If $h > i$ : Given $T_i = j$, $T_h$ is distributed like $j$ plus the place at which the $(h-i)$th good element appears in a random ordering of $N-j$ elements, of which $k-i$ are good; hence

$$E(T_h|T_i = j) = j + (h-i)\left(\frac{N-j+1}{k-i+1}\right).$$

As a check, note that (b) is a special case.

10. Since the number of black balls among the $n$ balls drawn from the box is the sum of $n$ indicators, we have

$$E(\text{number of black balls among the } n \text{ balls drawn}) = n \times P(\text{first ball drawn is black}).$$

Condition on $X$, the number of black balls among the $d$ balls drawn from the other box (which contains $b_0$ black balls and $w_0$ white balls):

$$\begin{aligned} P(\text{first ball drawn is black}) &= \sum_{k=0}^{d} P(\text{first ball drawn is black}|X=k)P(X=k) \\ &= \sum_{k=0}^{d} \frac{k+b}{b+w+d} P(X=k) \end{aligned}$$

since the box of $b+w+d$ has $k+b$ black balls when $X=k$. The last sum is $E\left(\frac{X+b}{b+w+d}\right)$. Since $X$ has hypergeometric distribution ($X$ is the number of good elements [black balls] in a sample of size $d$ from a population of size $b_0 + w_0$ containing $b_0$ goods), we have $E(X) = d \times \frac{b_0}{b_0+w_0} = \frac{db_0}{b_0+w_0}$. Hence

$$E(\text{number of black balls among the } n \text{ balls drawn}) = n \times \left(\frac{b'+b}{b+w+d}\right)$$

where $b' = \frac{db_0}{b_0+w_0}$.

11. The number of aces among the 5 cards drawn can be written as

$$X_1 + X_2 + X_3 + X_4 + X_5,$$

where

$$X_i = \begin{cases} 1 & \text{if } i\text{th card drawn is an ace} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore the expected number of aces among the 5 cards drawn is $5E(X_1)$. The number of aces among the pack of 39 cards is $H+1$, where $H$ is the number of aces contributed by the 13 cards coming from the top half of 26 cards. Then $H$ has hypergeometric distribution with $G=3$, $B=23$, and $n=13$, so $E(H) = 13 \times \frac{3}{26} = 3/2$. Use this fact to compute

$$\begin{aligned} E(X_1) = P(X_1 = 1) &= \sum_{k=0}^{3} P(X_1 = 1|H=k)P(H=k) \\ &= \sum_{k=0}^{3} \frac{k+1}{39} \cdot P(H=k) \\ &= E\left(\frac{H+1}{39}\right) \\ &= \frac{1}{39} \times \frac{3}{2} + \frac{1}{39} \\ &= \frac{5}{78}. \end{aligned}$$

So the expected number of aces among the 5 cards drawn is $5 \times E(X_1) = 25/78$ .

12. a) $E(B_{n+1}|B_n) = B_n\left(\frac{n+2-B_n}{n+2}\right) + (B_n+1)\left(\frac{B_n}{n}\right) = \frac{n+3}{n+2}B_n$

b) Use induction on $n$. Assume $E(B_n) = \frac{n+2}{3}$. $E(B_1) = 1 = \frac{1+2}{3}$
$E(B_{n+1}) = E[E(B_{n+1}|B_n)] = E\left(\frac{n+3}{n+2}B_n\right) = \frac{(n+1)+2}{3}$

c) $E\left(\frac{B_n}{n+2}\right) = \left(\frac{1}{n+2}\right)\left(\frac{n+2}{3}\right) = \frac{1}{3}$

13. c) $\frac{(n-m)}{(n-1)}m\frac{k}{n}\frac{(n-k)}{n}$

14. $P(X_1 = i_1, \ldots, X_n = i_n \mid S_n = k)$ will be 0 unless $\sum_{j=1}^n i_j = k$.

$$\begin{aligned} P(X_1 = i_1, \ldots, X_n = i_n \mid S_n = k) &= \frac{P(X_1 = i_1, \ldots, X_n = i_n, S_n = k)}{P(S_n = k)} \\ &= \frac{p^{\Sigma i_j}(1-p)^{n-\Sigma i_j}}{\binom{n}{k}p^k(1-p)^{n-k}} \\ &= \frac{1}{\binom{n}{k}} \end{aligned}$$

15. a) $E(S) = nE(\Pi)$

b) $Var(S) = nE(\Pi)(1-E(\Pi)) + n(n-1)Var(\Pi)$

c) $\Pi$ constant makes $Var(S)$ smallest, $\Pi$ with values only 0 and 1 makes $Var(S)$ largest. This follows from b) and

$$0 \le Var(\Pi) = E(\Pi^2) - E(\Pi)^2 \le E(\Pi) - E(\Pi)^2,$$

(since $0 \le \Pi \le 1$, so $\Pi^2 \le \Pi$).

16. Use the formula for $E(h(X)Y)$ by conditioning on $X$:

$$E(h(X)Y) = \sum_{all\ x} E(h(X)Y|X = x)P(X = x).$$

But given $X = x$, $h(X)$ equals $h(x)$, which can be treated as a constant:

$$E(h(X)Y) = \sum_{all\ x} h(x)E(Y|X = x)P(X = x) = E\left[h(X)E(Y|X)\right]$$

17. We know that

$$E[(Y - g(X))^2] = \sum_x E[(Y - g(X))^2|X = x]P(X = x).$$

All the terms in this sum are non-negative. So $E[(Y - g(X))^2]$ will be minimized if each of the terms $E[(Y - g(X))^2|X = x]$ is minimized. Given $X = x$, the random variable $g(X)$ can be treated as a constant $g(x)$, so

$$\begin{aligned} E[(Y - g(X))^2|X = x] &= E[(Y - g(x))^2|X = x] \\ &= E(Y^2|X = x) - 2g(x)E(Y|X = x) + [g(x)]^2. \end{aligned}$$

Taking $g(x) = E(Y|X = x)$ minimizes this last quantity. Therefore

$$E[(Y - g(X))^2] \ge \sum_x E[(Y - E(Y|X = x))^2|X = x]P(X = x) = E[(Y - E(Y|X))^2].$$

18. Write

$$Y - E(Y) = (Y - E(Y|X)) + (E(Y|X) - E(Y)).$$

Next, square and take expectations in both sides of the equality:

$$Var(Y) = E[(Y - E(Y))^2] = E[(Y - E(Y|X))^2] + E[(E(Y|X) - E(Y))^2].$$

Observe that the cross term vanishes, since applying Exercise 16 with $h(x) = E(Y|X) - E(Y)$ gives

$$E(Y - E(Y|X))(E(Y|X) - E(Y)) = E(Yh(X)) - E[E(Y|X)h(X)] = 0.$$

Finally, note that $E(Z) = E(E(Z|X))$ for any random variable $Z$; therefore

$$\begin{aligned} Var(Y) &= E[E((Y - E(Y|X))^2|X)] + E[(E(Y|X) - E(E(Y|X)))^2] \\ &= E(Var(Y|X)) + Var(E(Y|X)). \end{aligned}$$

# Section 6.3

1. By the integral conditioning formula,

$$P(A) = \int P(A|X = x) f_X(x)dx = \int_0^1 x^2 \cdot 1dx = 1/3.$$

2. a) $f_X(x) = \int_0^1 2x + 2y - 4xy\,dy = 2xy + y^2 - 2xy^2 \Big|_{y=0}^{1} = 1$

   By symmetry, $Y$ is also uniform(0,1).

   b) $f_Y(y \mid X = \frac{1}{4}) = \frac{f(\frac{1}{4},y)}{f_X(\frac{1}{4})} = \frac{2}{4} + 2y - \frac{4}{4}y = \frac{1}{2} + y \quad (0 < y < 1)$

   c) $E(Y \mid X = \frac{1}{4}) = \int_0^1 y(\frac{1}{2} + y)dy = \frac{y^2}{4} + \frac{y^3}{3}\Big|_{y=0}^{1} = \frac{7}{12}$

3. Let $0 < x < 2$ . Given $X = x$, it is intuitively clear that $Y$ is uniformly distributed on the interval $(0, 2 - x)$, so

$$P(Y \le y|X = x) = \begin{cases} 0 & \text{if } y < 0 \\ y/(2-x) & \text{if } 0 < y < 2 - x \\ 1 & \text{if } y > 2 - x \end{cases}$$

   You can be more rigorous: follow the approach of Example 1.

4. a) $f_Y(y) = \int_0^y \lambda^3 x e^{-\lambda y}dx = \frac{\lambda^3}{2} y^2 e^{-\lambda y}$ So, $Y$has the Gamma$(3, \lambda)$ and $E(Y) = \frac{3}{\lambda}$.

   b) $f(x \mid Y = 1) = \frac{f(x,1)}{f_Y(1)} = \frac{\lambda^3 x e^{-\lambda}}{(\lambda^3/2)e^{-\lambda}} = 2x \quad (0 < x < 1)$.

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

5. a) $P(Y \ge 1/2|X = x) = \begin{cases} \frac{x+1-1/2}{x+1} & \text{if } -1/2 \le x \le 0 \\ \frac{-x+1-1/2}{-x+1} & \text{if } 0 \le x \le 1/2 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \frac{1/2-|x|}{1-|x|} & \text{if } |x| \le 1/2 \\ 0 & \text{otherwise} \end{cases}$

   b) $P(Y < 1/2|X = x) = 1 - P(Y \ge 1/2|X = x)$.

   c,d) Let $-1 \le x \le 1$. Given $X = x$, $Y$ is uniformly distributed on $(0, 1 - |x|)$, hence has mean $(1 - |x|)/2$ and variance $(1 - |x|)^2/12$.

6. a) By making the substitution $x = ys$,

$$f_Y(y) = \int_0^y \frac{e^{-y/2}}{2\pi\sqrt{x(y-x)}}dx = \frac{e^{-y/2}}{2\pi}\int_0^1 \frac{1}{\sqrt{s(1-s)}}ds = \frac{e^{-y/2}}{2\pi}\int_0^1 x^{\frac{1}{2}-1}(1-x)^{\frac{1}{2}-1}dx$$

   Evaluating this beta integral gives

$$f_Y(y) = \frac{e^{-y/2}}{2\pi}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)} = \frac{1}{2}e^{-y/2}$$

   So $Y$ is exponential$(\frac{1}{2})$.

   b) First find the conditional density of $X$ given $Y = 1$.

$$f_X(x \mid Y = 1) = \frac{f(x,1)}{f_Y(1)} = \frac{e^{-\frac{1}{2}}}{2\pi\sqrt{x(1-x)}} \Big/ \frac{1}{2}e^{-\frac{1}{2}} = \frac{1}{\pi\sqrt{x(1-x)}} \quad (0 < x < 1)$$

   Since this density is symmetric around $\frac{1}{2}$, $E(X \mid Y = 1) = \frac{1}{2}$.

7. We require

$$1 = \iint f(y,z)dz\,dy = \int_{z=0}^{1}\int_{y=0}^{z} k(z-y)dydz = k\int_{z=0}^{1}\left[\frac{-(z-y)^2}{2}\right]\Bigg|_{y=0}^{z} dz = k\int_0^1 \frac{z^2}{2}dz = \frac{k}{6},$$

so $k = 6$.

a) $f_Y(y) = \int f(y,z)dz = \int_{z=y}^{1} 6(z-y)dz = \begin{cases} 3(1-y)^2 & 0 < y < 1 \\ 0 & \text{elsewhere.} \end{cases}$

b) The conditional density of $Z$ given $Y = y$ $(0 < y < 1)$ is

$$f_Z(z|Y=y) = \frac{f_{Y,Z}(y,z)}{f_Y(y)} = \frac{6(z-y)}{3(1-y)^2}, \qquad y \le z \le 1.$$

Hence the conditional density of $Z$ given $Y = 1/2$ is

$$f_Z(z|Y=1/2) = 8(z-1/2), \qquad 1/2 \le z \le 1,$$

and

$$P(Z < 2/3|Y = 1/2) = \int_{1/2}^{2/3} 8(z-1/2)dz = 1/9$$

8. a) $E(Y) = E[E(Y \mid X)] = E(5X) = \frac{5}{2}$
$E(Y^2) = E[E(Y^2 \mid X)] = E[(5X)^2 + 5X(1-X)] = 20E(X^2) + 5E(X) = \frac{55}{6}$

b) $P(Y = y, x < X < x+dx) = \binom{5}{y}x^y(1-x)^{5-y}dx$

c) $P(x < X < x+dx \mid Y = y) = P(Y = y, x < X < x+dx)/P(Y = y)$
The conditional density of $X$ given $Y = y$ is proportional to $P(Y = y, x < X < x+dx)$, so it must be Beta$(y+1, 6-y)$.

9. a) $P(A|B) = \frac{P(AB)}{P(B)} = \frac{\int P(AB|Y=p)P(Y\in dp)}{\int P(B|Y=p)P(Y\in dp)} = \frac{\int_{p=0}^{1} p^2\cdot 1dp}{\int_{p=0}^{1} p\cdot 1dp} = \frac{1/3}{1/2} = \frac{2}{3}$.

b) If $0 < p < 1$ then

$$\begin{aligned} P(Y \in dp|AB^c) &= \frac{P(Y\in dp, AB^c)}{P(AB^c)} \\ &= \frac{P(AB^c|Y\in dp)P(Y\in dp)}{\int P(AB^c|Y\in dp)P(Y\in dp)} = \frac{p(1-p)dp}{\int_0^1 p(1-p)dp} = 6p(1-p)dp. \end{aligned}$$

10. Let $N$ be the number of arrivals in $(0,1)$.

a) Since disjoint intervals of a Poisson process are independent,

$$\begin{aligned} P(N = 10, t < T_1 < t+dt) &= P(0 \text{ arrivals in } (0,t),\ 1 \text{ in } dt,\ 9 \text{ in } (t,1))) \\ &= e^{-\lambda t}\lambda dt \frac{e^{-\lambda(1-t)}(\lambda(1-t))^9}{9!} \\ &= \frac{e^{-\lambda}\lambda^{10}}{9!}t^{1-1}(1-t)^{10-1}dt \quad (0 < t < 1) \end{aligned}$$

Since the conditional density is proportional to this probability, $T_1$ given $N = 10$ has the Beta$(1,10)$ density.

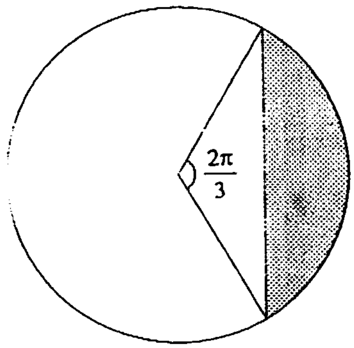
b)

$$\begin{aligned} P(N = 10, t < T_5 < t+dt) &= P(4 \text{ arrivals in } (0,t),\ 1 \text{ in } dt,\ 5 \text{ in } (t,1))) \\ &= \frac{e^{-\lambda t}(\lambda t)^4}{4!}\lambda dt \frac{e^{-\lambda(1-t)}(\lambda(1-t))^5}{5!} \\ &= \frac{e^{-\lambda}\lambda^{10}}{4!5!}t^{5-1}(1-t)^{6-1}dt \quad (0 < t < 1) \end{aligned}$$

The conditional density of $T_5$ given $N = 10$ is Beta$(5,6)$.

11. No: if $(X, Y)$ truly is uniformly distributed over the unit disk, then $X$ is *not* uniformly distributed on $(-1, 1)$. In the present situation (for instance) the chance that $X$ is near $\pm 1$ is higher than it should be when $(X, Y)$ is uniform over the unit disk. Here's an explicit calculation: In the present situation, the chance that $X$ is greater than $1/2$ is $1/4$, while if $(X, Y)$ were uniform over the disk, this chance would be

$$\frac{\frac{1}{2}\left(\frac{2\pi}{3} - \sin\frac{2\pi}{3}\right)}{\pi} \approx .1955$$




Another way: Compute the joint density of $X$ and $Y$:

$$f(x, y) = \begin{cases} \frac{1}{4\sqrt{1-x^2}} & \text{if } x^2 + y^2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Compare this with the density of random variables which are truly uniformly distributed over the unit disk:

$$f(x, y) = \begin{cases} \frac{1}{\pi} & \text{if } x^2 + y^2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

12. a) $T_1$ is the minimum of ten exponential (1) random variables, hence

$$f_{T_1}(x) = 10e^{-10x}, x > 0.$$

b) Suppose the first emission occurred at time $x$. Then each of the remaining particles is known to have lived at least $x$ units of time. Given that a particle has lived $x$ units of time, its remaining lifetime still has exponential distribution with rate 1 (by the memoryless property). Thus the remaining time (beyond $x$) until a new emission occurs is the minimum of 9 independent exponentials. That is,

$$P(T_2 - T_1 > y | T_1 = x) = e^{-9y}, \quad y > 0.$$

So given $T_1 = x$ $(x > 0)$, $T_2$ has conditional survival function

$$P(T_2 > z | T_1 = x) = P(T_2 - T_1 > z - x | T_1 = x) = e^{-9(z-x)}, \quad z > x;$$

and conditional density

$$f_{T_2}(z | T_1 = x) = \begin{cases} 9e^{-9(z-x)} & \text{if } z > x \\ 0 & \text{otherwise.} \end{cases}$$

c) By the average conditional probability formula, we have, for $z > 0$ :

$$\begin{aligned} f_{T_2}(z) &= \int f_{T_2}(z | T_1 = x) f_{T_1}(x) dx \\ &= \int_0^z 9e^{-9(z-x)} 10e^{-10x} dx \\ &= 90e^{-9z} \int_0^z e^{-x} dx \\ &= 90e^{-9z}(1 - e^{-z}). \end{aligned}$$

Or you may use the formula for the density of the $k$th order statistic.

**13.** a) Use the formula for average conditional probabilities:

$$\begin{aligned}
P(X \geq Y) &= \sum_{y=0}^{\infty} P(X \geq y | Y = y) P(Y = y) \\
&= P(X \geq 0)P(Y = 0) + P(X \geq 1)P(Y = 1) + P(X \geq 2)P(Y = 2) \\
&= 1 \times e^{-\lambda} + \frac{2}{3} \times \lambda e^{-\lambda} + \frac{1}{3} \times \frac{\lambda^2 e^{-\lambda}}{2}.
\end{aligned}$$

Therefore

$$P(X < Y) = 1 - \frac{1}{3} e^{-\lambda}(3 + 2\lambda + \lambda^2/2).$$

b) $P(X \in dx | X < Y) = \frac{P(X \in dx, x < Y)}{P(X < Y)} = \frac{P(X \in dx) P(Y > x)}{P(X < Y)}$ .

If $0 < x < 3$ then $P(X \in dx) = \frac{dx}{3}$ ; otherwise $P(X \in dx) = 0$ .

If $0 \leq x < 1$ then $P(Y > x) = P(Y \geq 1) = 1 - e^{-\lambda}$;

If $1 \leq x < 2$ then $P(Y > x) = P(Y \geq 2) = 1 - e^{-\lambda}(1 + \lambda)$;

If $2 \leq x < 3$ then $P(Y > x) = P(Y \geq 3) = 1 - e^{-\lambda}(1 + \lambda + \lambda^2/2)$.

Therefore

$$P(X \in dx | X < Y) = \begin{cases} \frac{1 - e^{-\lambda}}{3 - e^{-\lambda}(3 + 2\lambda + \frac{\lambda^2}{2})} dx & \text{if } 0 \leq x < 1 \\ \frac{1 - e^{-\lambda}(1+\lambda)}{3 - e^{-\lambda}(3 + 2\lambda + \frac{\lambda^2}{2})} dx & \text{if } 1 \leq x < 2 \\ \frac{1 - e^{-\lambda}(1+\lambda+\frac{\lambda^2}{2})}{3 - e^{-\lambda}(3 + 2\lambda + \frac{\lambda^2}{2})} dx & \text{if } 2 \leq x < 3 \end{cases}$$

Conditional Density of $X$ given $X < Y$

| | $0 \leq x < 1$ | $1 \leq x < 2$ | $2 \leq x < 3$ |
|---|---|---|---|
| $\lambda = 1$ | .65 | .27 | .08 |
| $\lambda = 2$ | .49 | .33 | .18 |
| $\lambda = 3$ | .41 | .34 | .25 |

**Remark:** Having computed $P(Y > x)$ , we can redo (a):

$$\begin{aligned}
P(X < Y) &= \int P(Y > X | X = x) P(X \in dx) \\
&= \int P(Y > x) P(X \in dx) \qquad \text{(independence)} \\
&= \frac{1}{3} \int_0^3 P(Y > x) dx \\
&= \frac{1}{3} [P(Y \geq 1) + P(Y \geq 2) + P(Y \geq 3)] \\
&= \frac{1}{3} \left[ 3 - e^{-\lambda}(3 + 2\lambda + \frac{\lambda^2}{2}) \right].
\end{aligned}$$

c) Apply (b):

$$\begin{aligned}
E(X | X < Y) &= \int x P(X \in dx | X < Y) \\
&= \frac{\int_0^3 x P(X \in dx) P(Y > x)}{P(X < Y)} \qquad \text{from (b)} \\
&= \frac{P(Y \geq 1) \int_{x=0}^{1} x dx + P(Y \geq 2) \int_{x=1}^{2} x dx + P(Y \geq 3) \int_{x=2}^{3} x dx}{3 P(X < Y)} \\
&= \frac{P(Y \geq 1) + 3 P(Y \geq 2) + 5 P(Y \geq 3)}{6 P(X < Y)} \\
&= \frac{9 - e^{-\lambda}(9 + 8\lambda + \frac{5}{2}\lambda^2)}{6 - 2e^{-\lambda}(3 + 2\lambda + \frac{1}{2}\lambda^2)}.
\end{aligned}$$

14. Let $x_1 + \cdots + x_n = k$.

$$\begin{aligned} P(\Pi \in dp \mid X_1 = x_1, \ldots, X_n = x_n) &= \frac{P(\Pi \in dp, X_1 = x_1, \ldots, X_n = x_n)}{P(X_1 = x_1, \ldots, X_n = x_n)} \\ &= \frac{p^k(1-p)^{n-k} f(p)dp}{P(X_1 = x_1, \ldots, X_n = x_n)} \end{aligned}$$

The posterior density of $\Pi$ given $X_1 = x_1, \ldots, X_n = x_n$ is proportional to $p^k(1-p)^{n-k}f(p)$, which depends only on $k$.

15. a)

$$\begin{aligned} P(\pi \in dp, S_n = k) &= P(\pi \in dp)P(S_n = k \mid \pi = p) \\ &= \tfrac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} p^{r-1}(1-p)^{s-1} dp \tbinom{n}{k} p^k (1-p)^{n-k} \\ &= \tfrac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} \tbinom{n}{k} p^{r+k-1}(1-p)^{s+n-k-1} dp \end{aligned}$$

Hence

$$\begin{aligned} f_\pi(p \mid S_n = k) = P(\pi \in dp \mid S_n = k)/dp &= \tfrac{P(\pi \in dp, S_n = k)}{dp P(S_n = k)} \\ &= c(r, s, n, k) p^{r+k-1}(1-p)^{s+n-k-1} \end{aligned}$$

for some constant $c(r, s, n, k)$. Now use the hint.

b)

$$\begin{aligned} P(S_n = k) &= \textstyle\int_0^1 P(\pi \in dp, S_n = k) \\ &= \tfrac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} \tbinom{n}{k} \tfrac{\Gamma(r+k)\Gamma(s+n-k)}{\Gamma(r+s+n)} \end{aligned}$$

c) Take $r = s = 1$. Use $\Gamma(m) = (m-1)!$ for integer $m$:

$$P(S_n = k) = \frac{1}{1 \cdot 1} \frac{n!}{k!(n-k)!} \frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1}$$

d) If $X$ has beta $(a, b)$ distribution,

$$E(X) = \frac{a}{a+b} \quad \text{and} \quad Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$$

Applied $a = r + k$ and $b = s + n - k$ this gives

$$\begin{aligned} E(\pi \mid S_n = k) &= \tfrac{r+k}{r+s+n} \\ Var(\pi \mid S_n = k) &= \tfrac{(r+k)(s+n-k)}{(r+s+n)^2(r+s+n+1)} \end{aligned}$$

16. a) By using the substitution $u = \lambda(t + \alpha)$,

$$\begin{aligned} P(N = k) &= \int_0^\infty P(N = k \mid \Lambda = \lambda) f_\Lambda(\lambda) d\lambda \\ &= \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^k}{k!} \frac{\alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma(r)} d\lambda \\ &= \frac{\alpha^r t^k}{\Gamma(r) k! (t+\alpha)^{r+k}} \int_0^\infty e^{-u} u^{r+k-1} du \\ &= \frac{\Gamma(r+k)}{\Gamma(r) k!} p^r q^k \end{aligned}$$

where $p = \frac{\alpha}{t+\alpha}$ and $q = \frac{t}{t+\alpha}$.

b) Using $\Gamma(x+1) = x\Gamma(x)$,

$$\frac{\Gamma(r+k)}{\Gamma(r)} = (r+k-1)\frac{\Gamma(r+k-1)}{\Gamma(r)} = \cdots = \prod_{i=0}^{k-1}(r+i)$$

When $r$ is a postive integer,

$$\frac{\Gamma(r+k)}{\Gamma(r)k!} = \frac{(r+k-1)!}{(r-1)!k!} = \binom{r+k-1}{r-1}$$

c)

$$\begin{aligned} E(z^N) &= E[E(z^N \mid \Lambda)] = E\left[\sum_{k=0}^{\infty} z^k \frac{e^{-\Lambda t}(\Lambda t)^k}{k!}\right] \\ &= E[e^{-\Lambda t}e^{z\Lambda t}] = E[e^{-\Lambda t(1-z)}] \\ &= \int_0^\infty e^{-\lambda t(1-z)}\frac{\alpha^r\lambda^{r-1}e^{-\alpha\lambda}}{\Gamma(r)}d\lambda \end{aligned}$$

Substituting $u = \lambda(t(1-z)+\alpha)$,

$$\begin{aligned} &= \frac{\alpha^r}{\Gamma(r)(t(1-z)+\alpha)^r}\int_0^\infty e^{-u}u^{r-1}du \\ &= \left(\frac{\alpha}{t+\alpha-zt}\right)^r \\ &= \left(\frac{\alpha}{t+\alpha}\right)^r\left(\frac{1}{1-zt/(t+\alpha)}\right)^r \\ &= p^r(1-zq)^{-r} \end{aligned}$$

d) $E(N) = E[E(N \mid \Lambda)] = E(\Lambda t) = \frac{rt}{\alpha} = \frac{r(1-p)}{p}$
$E(N^2) = E[E(N^2 \mid \Lambda)] = E[\Lambda t + (\Lambda t)^2] = \frac{rt}{\alpha} + (\frac{r}{\alpha^2} + (\frac{r}{\alpha})^2) = \frac{r(1-p)}{p} + r(r+1)(\frac{1-p}{p})^2$
$Var(N) = E(N^2) - [E(N)]^2 = \frac{r(1-p)}{p^2}$

e) Let $g(z) = E(z^N)$. Then,

$$E(N) = g'(1) = p^r q r(1-qz)^{-(r+1)}\Big|_{z=1} = \frac{r(1-p)}{p}$$

$$E[N(N-1)] = g''(1) = p^r q^2 r(r+1)(1-qz)^{-(r+2)}\Big|_{z=1} = \frac{r(r+1)(1-p)^2}{p^2}$$

$$Var(N) = g''(1) + g'(1) - [g'(1)]^2 = \frac{r(1-p)}{p^2}$$

f)

$$\begin{aligned} P(\Lambda \in d\lambda \mid N=k) &= \frac{P(\Lambda \in d\lambda, N=k)}{P(N=k)} = \left(\frac{\alpha^r\lambda^{r-1}e^{-\alpha\lambda}d\lambda}{\Gamma(r)}\right)\left(\frac{e^{-\lambda t}(\lambda t)^k}{k!}\right)\Big/P(N=k) \\ &= (\text{constant})\lambda^{r+k-1}e^{-\lambda(\alpha+t)}d\lambda \end{aligned}$$

So given $N=k$, $\Lambda$ has the gamma$(r+k, \alpha+t)$ distribution.

17. a)Independent negative binomial $(r_i, p)$ $(i = 1,2)$ b) negative binomial $(r_1+r_2, p)$ c) negative binomial $(\sum_i r_i, p)$

# Section 6.4

1. a) 0.5 b) positively dependent c) 0.2 d) 0.356

2. a) If $P(A|B) = P(A|B^c)$ then

$$\begin{aligned} P(A) &= P(A|B)P(B) + P(A|B^c)P(B^c) \\ &= P(A|B)[P(B) + P(B^c)] = P(A|B), \end{aligned}$$

which implies that $A$ and $B$ are independent.

b) If $P(A|B) < P(A|B^c)$ then

$$\begin{aligned} P(A) &= P(A|B)P(B) + P(A|B^c)P(B^c) \\ &< P(A|B)[P(B) + P(B^c)] = P(A|B) \end{aligned}$$

c) By interchanging the roles of $B$ and $B^c$ in part b), we get the result immediately.

d) Independence implies $P(A|B) = P(A)$. Plugging this into the given formula, we get

$$\begin{aligned} & P(A|B) = P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) \\ \iff & P(A|B)[1 - P(B)] = P(A|B^c)P(B^c) \\ \iff & P(A|B) = P(A|B^c) \qquad \text{so long as } P(B) < 1. \end{aligned}$$

e) Positive dependence implies

$$\begin{aligned} & P(A|B) > P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) \\ \iff & P(A|B)[1 - P(B)] > P(A|B^c)P(B^c) \\ \iff & P(A|B) > P(A|B^c) \qquad \text{so long as } P(B) < 1. \end{aligned}$$

f) This is just the same as the previous part with the inequality reversed.

3. Let $F_1$ denote the event that the first component fails, and $W_1$ denote the event that the first component works, etc. Then $F_1^c = W_1$, and $P(F_2|F_1) > P(F_2)$ implies that

$$P(W_2|F_1) = 1 - P(F_2|F_1) < 1 - P(F_2) = P(W_2)$$

Also,

$$\begin{aligned} P(W_2|W_1) &= \frac{P(W_1 W_2)}{P(W_1)} \\ &= \frac{1 - P(F_1) - P(F_2) + P(F_1 F_2)}{1 - P(F_1)} \\ &> \frac{1 - P(F_1) - P(F_2) + P(F_1)P(F_2)}{1 - P(F_1)} \\ &= 1 - P(F_2) \\ &= P(W_2) \end{aligned}$$

since $P(F_1 F_2)/P(F_2) = P(F_1|F_2) > P(F_1)$.

4. *Joint distribution table for* $(X, Y)$

| | | values $x$ for $X$ | | | Marginal distn of $Y$ |
|---|---|---|---|---|---|
| | | -1 | 0 | 1 | |
| values | -1 | 0 | 1/4 | 0 | 1/4 |
| $y$ | 0 | 1/4 | 0 | 1/4 | 1/2 |
| for $Y$ | 1 | 0 | 1/4 | 0 | 1/4 |
| Marginal distn of $X$ | 1/4 | 1/2 | 1/4 | 1 | |

Now you can easily check that $X$ and $Y$ are uncorrelated but not independent:

$$E(XY) = E(X) = E(Y) = 0 \rightarrow Cov(X,Y) = 0, \quad \text{but}$$

$$P(X=0, Y=0) \neq P(X=0)P(Y=0).$$

5. *Joint distribution table for* $(X, Y)$

| | | values $x$ for $X$ | | | Marginal distn of $Y$ |
|---|---|---|---|---|---|
| | | -1 | 0 | 1 | |
| values $y$ | 0 | 0 | 1/3 | 0 | 1/3 |
| for $Y$ | 1 | 1/3 | 0 | 1/3 | 2/3 |
| Marginal distn of $X$ | 1/3 | 1/3 | 1/3 | 1 | |

So compute $E(X) = E(XY) = 0 \Longrightarrow Cov(X,Y) = 0$.

But $P(X=-1, Y=0) = 0 \neq P(X=-1)P(Y=0)$, etc., so $X$ and $Y$ are not independent. In fact, $Y = X^2$ implies that $X$ and $Y$ are not independent.

6.

$$\begin{aligned} Cov(X,Y) &= E[(X_1 - X_2)(X_1 + X_2)] - E(X_1 - X_2)E(X_1 + X_2) \\ &= E(X_1^2 - X_2^2) \quad \text{since} E(X_1) = E(X_2) \\ &= 0 \quad \text{since} E(X_1^2) = E(X_2^2). \end{aligned}$$

So $X, Y$ are uncorrelated. But $X, Y$ are not independent: For example, you can easily check that $P(X_1 - X_2 = 0, X_1 + X_2 = 3) = 0$, while $P(X_1 - X_2 = 0) = 1/6$, $P(X_1 + X_2 = 3) = 1/18$.

Intuitively, if we know that $X_1 - X_2 = 0$, then $X_1 + X_2$ must be even, since the two dice landed the same way. So given that the difference is 0, we have some information about the sum.

7. a) *Joint distribution table for* $(X_2 + X_3, X_2 - X_3)$

| | | values for $X_2 + X_3$ | | | Marginal distn of $X_2 - X_3$ |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | |
| values | -1 | 0 | 1/6 | 0 | 1/6 |
| for | 0 | 1/3 | 0 | 1/6 | 1/2 |
| $X_2 - X_3$ | 1 | 0 | 1/3 | 0 | 1/3 |
| Marginal distn of $X_2 + X_3$ | | 1/3 | 1/2 | 1/6 | 1 |

b) $E(X_2 - X_3)^3 = (-1)^3 \cdot (1/6) + (0)^3 \cdot (1/2) + (1)^3 \cdot (1/3) = 1/6$

c) $E(X_2X_3) = P(X_2 = 1, X_3 = 1) = P(X_2 = 1)P(X_3 = 1)$ by independence $= E(X_2)E(X_3)$, so $X_2$ and $X_3$ are uncorrelated.

8. Begin by noting that $X_1$ and $X_N$ each are distributed binomial($k$, $1/N$).

$$\begin{aligned}
E(X_1X_N) &= E[E(X_1X_N \mid X_1)] = E[X_1E(X_N \mid X_1)] \\
&= E[X_1\frac{k-X_1}{N-1}] = \frac{kE(X_1)}{N-1} - \frac{E(X_1^2)}{N-1} \\
&= \frac{k^2}{N(N-1)} - \left(\frac{k}{N^2} + \frac{k^2}{N^2(N-1)}\right) = \frac{k(k-1)}{N^2} \\
corr(X_1, X_N) &= \frac{E(X_1X_N) - E(X_1)E(X_n)}{SD(X_1)SD(X_N)} = \frac{E(X_1X_N) - (E(X_1))^2}{Var(X_1)} \\
&= \frac{(k(k-1)/N^2) - (k/N)^2}{k(1/N)((N-1)/N)} = -\frac{1}{N-1}
\end{aligned}$$

9. Use the results of Example 7 to obtain

   a) $k(n+1)/2$

   b) $\frac{k(n^2-1)(n-k)}{12(n-1)}$

   Note: Brief solution in text is incorrect.

10. Condition on $H_{100}$.

$$\begin{aligned}
E(H_{100}H_{300}) &= E[E(H_{100}H_{300} \mid H_{100})] = E[H_{100}E(H_{300} \mid H_{100})] \\
&= E[H_{100}(H_{100} + 200 \cdot \frac{1}{2})] = E(H_{100}^2) + 100E(H_{100}) \\
&= \left(100 \cdot \frac{1}{2} \cdot \frac{1}{2} + 50^2\right) + 100(50) = 7525 \\
corr(H_{100}, H_{300}) &= \frac{E(H_{100}H_{300}) - E(H_{100})E(H_{300})}{SD(H_{100})SD(H_{300})} \\
&= \frac{7525 - 50(150)}{(5)(5\sqrt{3})} = \frac{1}{\sqrt{3}}
\end{aligned}$$

11. $\sqrt{1/3}$

12. $X$, $Y$, and $Z$ are indicator random variables.

   $E(X) = P(\text{select a Democrat}) = \alpha, \quad E(Y) = P(\text{select a Republican}) = \beta$

   b) $Var(X) = \alpha(1-\alpha), \quad Var(Y) = \beta(1-\beta)$

   c) $Cov(X,Y) = E(XY) - E(X)E(Y) = 0 - \alpha\beta$

   d) $E(D_n - R_n) = nE(X) - nE(Y) = n(\alpha - \beta)$

   e) $D_n - R_n = \sum_{i=1}^n (X_i - Y_i)$.

$$\begin{aligned}
Var(D_n - R_n) &= nVar(X - Y) = nVar(X) + nVar(Y) - 2nCov(X,Y) \\
&= n\alpha(1-\alpha) + n\beta(1-\beta) + 2n\alpha\beta = n(\alpha + \beta - (\alpha - \beta)^2)
\end{aligned}$$

13. For each $i = 1, \ldots, n$ let $A_i$ denote the event that result $A$ occurs on trial $i$, similarly for $B_i$. Then

$$N_A = \sum_{i=1}^n 1_{A_i};$$

$$N_B = \sum_{j=1}^n 1_{B_j};$$

$$N_AN_B = \sum_{i=1}^n \sum_{j=1}^n$$

$$1_{A_i}1_{B_j} = \sum_{i=1}^n \sum_{j=1}^n 1_{A_iB_i}.$$

Hence

$$
\begin{aligned}
E(N_A) &= \sum_{i=1}^{n} P(A_i) = nP(A_1) \\
E(N_B) &= nP(B_1) \\
E(N_A N_B) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \\
P(A_i B_i) &= \sum_{i=1}^{n} P(A_i B_j) + 2\sum_{i} < jP(A_i)P(B_j) \\
&= nP(A_1 B_1) + n(n-1)P(A_1)P(B_1).
\end{aligned}
$$

Therefore $N_A$ and $N_B$ are uncorrelated if and only if $E(N_A N_B) = E(N_A)E(N_B)$ if and only if $P(A_1 B_1) = P(A_1)P(B_1)$, i.e. if and only if $1_{A_i}$ and $1_{B_i}$ are independent for all $i$.

It is intuitively clear that $N_A$ and $N_B$ must then be independent. In more detail, if $1_{A_i}$ and $1_{B_i}$ are independent for each $i$, and the $n$ pairs $(1_{A_1}, 1_{B_1}), \ldots, (1_{A_n}, 1_{B_n})$ are independent, then the $2n$ variables $1_{A_1}, 1_{B_1}, \ldots, 1_{A_n}, 1_{B_n}$ are independent. So $N_A$ and $N_B$ must be independent, since they are functions of disjoint blocks of independent variables.

14. From the inequality $-1 \le Corr(X,Y) \le 1$ we get the following equivalent statements:

$$
\begin{gathered}
-SD(X)SD(Y) \le Cov(X,Y) \le SD(X)SD(Y) \\
-2SD(X)SD(Y) \le 2Cov(X,Y) \le 2SD(X)SD(Y) \\
-2SD(X)SD(Y) \le Var(X+Y) - Var(X) - Var(Y) \le 2SD(X)SD(Y) \\
Var(X) - 2SD(X)SD(Y) + Var(Y) \le Var(X+Y) \le Var(X) + 2SD(X)SD(Y) + Var(Y) \\
(SD(X) - SD(Y))^2 \le Var(X+Y) \le (SD(X) + SD(Y))^2 \\
|SD(X) - SD(Y)| \le SD(X+Y) \le |SD(X) + SD(Y)|
\end{gathered}
$$

15. a), b) are special cases of c), so it suffices to prove c) only.

c)

$$
\begin{aligned}
Cov\Big(\sum_i X_i, \sum_j Y_j\Big) &= E\left[\sum_i X_i - E\Big(\sum_i X_i\Big)\right]\left[\sum_j Y_j - E\Big(\sum_j Y_j\Big)\right] \\
&= E\left[\sum_i X_i - \sum_i E(X_i)\right]\left[\sum_j Y_j - \sum_j E(Y_j)\right] \\
&= E\left\{\sum_i [X_i - E(X_i)]\right\}\left\{\sum_j [Y_j - E(Y_j)]\right\} \\
&= E\left\{\sum_{i,j} [X_i - E(X_i)][Y_j - E(Y_j)]\right\} \\
&= \sum_{i,j} E[X_i - E(X_i)][Y_j - E(Y_j)] \\
&= \sum_{i,j} Cov(X_i, Y_j).
\end{aligned}
$$

d) We have $N_R = \sum_{i=1}^{n} X_i$ and $N_B = \sum_{j=1}^{n} Y_j$, where $X_i = 1$ if the $i$th spin is red, $= 0$ otherwise; and $Y_j = 1$ if the $j$th spin is black, $= 0$ otherwise. Note that $E(X_i) = r$, $E(Y_j) = b$, and that

$$
E(X_i Y_j) = P(i\text{th spin is red and } j\text{th spin is black}) = \begin{cases} rb & \text{if } i \ne j \\ 0 & \text{if } i = j \end{cases}
$$

Hence

$$Cov(X_i, Y_j) = \begin{cases} 0 & \text{if } i \neq j \\ -rb & \text{if } i = j \end{cases}$$

Now use c) to obtain

$$Cov(N_R, N_B) = \sum_{i=1}^{n} Cov(X_i, Y_i) + \sum_{i \neq j} Cov(X_i, Y_j) = -nrb.$$

16. Note that

$$Cov(aX + b, cY + d) = acCov(X, Y), \quad SD(aX + b) = |a|SD(X), \quad SD(cY + d) = |c|SD(Y).$$

Therefore

$$Corr(aX + b, cY + d) = \frac{ac}{|ac|}Corr(X, Y) = \begin{cases} Corr(X, Y) & \text{if } ac > 0 \\ -Corr(X, Y) & \text{if } ac < 0 \end{cases}$$

17. Note that $1_{A^c} = 1 - 1_A$, and $1_{B^c} = 1 - 1_B$, and apply the result of Exercise 16.

18. If $X_1, \dots, X_n$ are exchangeable, then $(X_1, \dots, X_n)$ has the same distribution as $(X_{\pi(1)}, \dots, X_{\pi(n)})$ for any permutation $\pi$ of $(1, 2, \dots, n)$. In particular, by considering the first coordinate of the above relation, each $X_i$ has the same distribution as $X_1$; and by considering the first two coordinates of the above relation, each pair $(X_i, X_j)$ $(i \neq j)$ has the same distribution as $(X_1, X_2)$. Therefore

$$\begin{aligned} Var(\sum X_i) &= \sum_{i=1}^{n} Var(X_i) + \sum_{i \neq j} Cov(X_i, X_j) \\ &= \sum_{i=1}^{n} Var(X_1) + \sum_{i \neq j} Cov(X_1, X_2) \\ &= nVar(X_1) + n(n-1)Cov(X_1, X_2). \end{aligned}$$

19. a,b) Let $Y$ denote the value (in cents) obtained on any particular draw. Then $E(Y) = 18.75$, $SD(Y) = 8.1968$, so (measuring $X$ in cents)

$$E(X) = 20; \qquad E(Y) = 375$$

$$SD(X) = \sqrt{\frac{40 - 20}{40 - 1}} \cdot \sqrt{20} \cdot SD(Y) = 26.25.$$

c) $P(X \leq 300) \approx \Phi(-2.86) \approx .0021$. Exact value: .002152141.

d) The quarters are bigger than the other coins, so maybe they'll be drawn more readily. This suggests that $E(X)$ will be higher than calculated, and $P(X \leq 300)$ will be lower than calculated.

20. a) $x_1 p_1 + x_2 p_2$

b) $\mu_1 p_1 + \mu_2 p_2$

c) $|x_1 - x_2|\sqrt{p_1 p_2}$

d) $\sqrt{\sigma_1^2 p_1 + \sigma_2^2 p_2 + (\mu_1 - \mu_2)^2 p_1 p_2} = \sqrt{E[Var(Y|X)] + Var[E(Y|X)]}$

e) $(x_1 - x_2)(\mu_1 - \mu_2)p_1 p_2$

f) $\dfrac{(x_1 - x_2)(\mu_1 - \mu_2)p_1 p_2}{|x_1 - x_2|\sqrt{p_1 p_2}\sqrt{\sigma_1^2 p_1 + \sigma_2^2 p_2 + (\mu_1 - \mu_2)^2 p_1 p_2}}$

In general:

$$E(X) = \sum_{i=1}^{n} x_i p_i$$

$$E(Y) = \sum_{i=1}^{n} \mu_i p_i$$

$$Var(X) = \sum_{i<j} (x_i - x_j)^2 p_i p_j$$

$$Var(Y) = \sum_{i=1}^{n} \sigma_i^2 p_i + \sum_{i<j} (\mu_i - \mu_j)^2 p_i p_j$$

$$Cov(X,Y) = \sum_{i<j} (x_i - x_j)(\mu_i - \mu_j) p_i p_j$$

**21.** a) 15/13 b) 24/13 c) 100/169

d) Use the fact that $X + Y = 3$ to conclude

$$0 = Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y).$$

But $X + Y = 3$ implies that $Var(X) = Var(Y)$, so $Cov(X,Y) = -Var(X) = -100/169$.

**22.** $X = X_1 + X_2 + \cdots + X_m$, where $X_i = I(i\text{th couple survived})$, and

$$P(X_i = 1) = \frac{\binom{2m-2}{d}}{\binom{2m}{d}}, \qquad P(X_i = 1, X_j = 1) = \frac{\binom{2m-4}{d}}{\binom{2m}{d}}.$$

a) $E(X) = m \times \frac{\binom{2m-2}{d}}{\binom{2m}{d}} = \frac{(2m-d)(2m-d-1)}{2(2m-1)}$.

b)

$$\begin{aligned} Var(X) &= m \times \frac{\binom{2m-2}{d}}{\binom{2m}{d}} \left[1 - \frac{\binom{2m-2}{d}}{\binom{2m}{d}}\right] + m(m-1) \times \left\{ \frac{\binom{2m-4}{d}}{\binom{2m}{d}} - \left[\frac{\binom{2m-4}{d}}{\binom{2m}{d}}\right]^2 \right\} \\ &= \frac{d(d-1)(2m-d)(2m-d-1)}{2(2m-1)^2(2m-3)}. \end{aligned}$$

**23.** a) $\text{MSE} = E[Y - (\beta X + \gamma)]^2 = E(Y^2) - 2E[Y(\beta X + \gamma)] + E[(\beta X + \gamma)^2]$
$= \beta^2 E(X^2) - 2\beta E(XY) + 2\beta\gamma E(X) - 2\gamma E(Y) + \gamma^2 + E(Y^2)$

b) The derivative of MSE with respect to $\gamma$ is $2\beta E(X) - 2E(Y) + 2\gamma$. Set to zero, obtain $\hat{\gamma} = E(Y) - \beta E(X)$. The second derivative with respect to $\gamma$ is positive, so MSE is minimised at $\hat{\gamma}$. If $\beta = 0$, the MSE is minimised at $\hat{\gamma} = E(Y)$ and the resulting minimal MSE is

$$E[(Y - (\beta X + \hat{\gamma}))]^2 = E[Y - E(Y)]^2 = Var(Y).$$

c) Substitute $\gamma = \hat{\gamma}(\beta) = E(Y) - \beta E(X)$ in part a) to obtain

$$\begin{aligned} \text{MSE} &= \beta^2 E(X^2) - 2\beta E(XY) + 2\beta E(X)[E(Y) - \beta E(X)] \\ &- \quad 2E(Y)[E(Y) - \beta E(X)] + [E(Y) - \beta E(X)]^2 + E(Y^2). \end{aligned}$$

The derivative with respect to $\beta$ is

$$2\beta E(X^2) - 2[E(XY) - E(X)E(Y)] - 4\beta[E(X)]^2 + 2E(X)E(Y) - 2[E(Y) - \beta E(X)]E(X)$$

$$\begin{aligned} &= -2Cov(X,Y) + 2\beta\left[E(X^2) - [E(X)]^2\right] \\ &= -2Cov(X,Y) + 2\beta Var(X). \end{aligned}$$

The second derivative with respect to $\beta$ is $2Var(X) > 0$.

d) Let $\beta$ and $\gamma$ be arbitrary. From parts b) and c) we have

$$\text{MSE}(\hat{\beta}, \hat{\gamma}(\hat{\beta})) \le \text{MSE}(\beta, \hat{\gamma}(\beta)) \le \text{MSE}(\beta, \gamma),$$

so the value of MSE at $(\hat{\beta}, \hat{\gamma}(\hat{\beta}))$ is at least as small as the value at any other pair $(\beta, \gamma)$. But from part a) we know that MSE is a quadratic function of $\beta$ and $\gamma$, so it has a unique minimum.

e) Since $\hat{\beta}$ and $\hat{\gamma}$ can be treated as constants,

$$E(\hat{Y}) = E(\hat{\beta}X + \hat{\gamma}) = \hat{\beta}E(X) + [E(Y) - \hat{\beta}E(X)] = E(Y)$$

and

$$Var(\hat{Y}) = Var(\hat{\beta}X + \hat{\gamma}) = \hat{\beta}^2 Var(X).$$

For the last equality note that

$$\begin{aligned} E(\hat{Y}Y) &= E[(\hat{\beta}X + \hat{\gamma})Y] = \hat{\beta}E(XY) + \hat{\gamma}E(Y) \\ &= \hat{\beta}[E(XY) - E(X)E(Y)] + [E(Y)]^2 = \hat{\beta}^2 Var(X) + [E(Y)]^2 \end{aligned}$$

and

$$E(\hat{Y}^2) = Var(\hat{Y}) + [E(\hat{Y})]^2 = \hat{\beta}^2 Var(X) + [E(Y)]^2.$$

f) We have

$$\begin{aligned} Var(Y) &= E(Y^2) - [E(Y)]^2 = E[Y - \hat{Y} + \hat{Y}]^2 - [E(Y)]^2 \\ &= E(Y - \hat{Y})^2 + 2E[\hat{Y}(Y - \hat{Y})] + E(\hat{Y})^2 - [E(Y)]^2. \end{aligned}$$

By part c), $E[\hat{Y}(Y-\hat{Y})] = 0$ and $E(Y) = E(\hat{Y})$. Conclude that $Var(Y) = E[(Y-\hat{Y})^2] + Var(\hat{Y})$. Finally observe that by part c),

$$\begin{aligned} Var(\hat{Y}) &= \hat{\beta}^2 Var(X) = \frac{[Cov(X,Y)]^2}{[Var(X)]^2} Var(X) \\ &= \frac{[Cov(X,Y)]^2}{Var(X)Var(Y)} Var(Y) = \rho^2 Var(Y). \end{aligned}$$

g) By part c),

$$\hat{\beta} = \frac{Cov(X,Y)}{Var(X)} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \cdot \frac{\sqrt{Var(Y)}}{\sqrt{Var(X)}} = \rho \frac{SD(Y)}{SD(X)}$$

If $\hat{\gamma}$ is such that the line $y = \hat{\beta}x + \hat{\gamma}$ passes through the point $(E(X), E(Y))$, then this requirement yields the equation

$$E(Y) = \hat{\beta}E(X) + \hat{\gamma} \quad \Longleftrightarrow \quad \hat{\gamma} = E(Y) - \hat{\beta}E(X).$$

h) By the previous part, the best linear predictor of $Y^*$ based on $X^*$ has slope

$$Corr(X^*, Y^*) \cdot \frac{SD(Y^*)}{SD(X^*)} = Corr(X,Y) \cdot \frac{1}{1} = \rho$$

(since the correlation coefficient is invariant under linear transformations), and passes through the point $(E(X^*), E(Y^*)) = (0, 0)$.

# Section 6.5

1. Let $X$ and $Y$ denote the PSAT and SAT scores in standard units.

   a) $P(Y > 0 \mid X = -2) = P\left[\frac{Y-(0.6)\cdot(-2)}{\sqrt{1-0.6^2}} > \frac{0-(0.6)\cdot(-2)}{\sqrt{1-0.6^2}} \mid X = -2\right] = P(Z > 1.5) = 0.0668.$

   b) $P(Y > 0|X < 0) = \frac{P(X<0,Y>0)}{1/2} = 2P(X < 0, Y > 0).$
   Transform to independent random variables, as in Example 2.

   $$P(X < 0, Y > 0) = P(X < 0, 0.6X + 0.8Z > 0) = \frac{\arctan 4/3}{2\pi} = 0.1476.$$

   (since the corresponding region in the $(X, Z)$ plane is a sector centered at $(0,0)$ subtending an angle of $\arctan 4/3$) . So the desired probability is 0.2952.

   c) We need

   $$\begin{aligned} P(\text{SAT score} > \text{PSAT score} + 50) &= P(90Y + 1300 > 100X + 1200 + 50) \\ &= P\left[90(0.6X + 0.8Z) - 100X > -50\right] \\ &= P(-46X + 72Z > -50). \end{aligned}$$

   Now $-46X+72Z$ has normal distribution with mean zero and SD 85.44 (it's a linear combination of independent normals); hence the desired probability is $\Phi(50/85.44) = 0.72$.

2. Let $X$ and $Y$ be the heights in standard units of a randomly chosen daughter-mother pair from this population. Then $Y = \frac{1}{2}X + \frac{\sqrt{3}}{2}Z$, where $X$ and $Z$ are independent standard normals.

   $$\begin{aligned} P(X < Y \mid X > 0) &= \frac{P(0 < X, X < Y)}{P(X > 0)} \\ &= P(0 < X, X < \frac{1}{2}X + \frac{\sqrt{3}}{2}Z) \Big/ P(X > 0) \\ &= P(0 < X, \frac{1}{\sqrt{3}}X < Z) \Big/ P(X > 0) \end{aligned}$$

   By using the rotational symmetry of the bivariate normal distribution, the probability in the numerator is $\frac{(\pi/3)}{2\pi} = \frac{1}{6}$. Hence, $\frac{1}{3}$ of the daughters with above average height are shorter than their mothers.

3. Let $X$ and $Y$ denote the weight and height in standard units of a person chosen at random. The event (person is above 90th percentile in height) is the same as $(Y > 1.282)$, and the event (person is in 90th percentile in weight) is the same as $(X = 1.282)$. Hence the desired probability is $P(Y > 1.282|X = 1.282)$. Now given $X = x$, $Y$ has normal $(\rho x, 1 - \rho^2)$ distribution, so

   $$\begin{aligned} P(Y > x|X = x) &= P\left(\frac{Y - \rho x}{\sqrt{1-\rho^2}} > \frac{x - \rho x}{\sqrt{1-\rho^2}}|X = x\right) \\ &= P\left(Z > \sqrt{\frac{1-\rho}{1+\rho}}x\right) \end{aligned}$$

   If $x = 1.282$ and $\rho = 0.75$ then the desired probability is $P(Z > 0.484) = .314$ .

4. $Var(X + 2Y) = Var(X) + 4Var(Y) + 2Cov(X, Y)$. If $X$ and $Y$ are independent, then the variance of this sum is 5. If the correlation of $X$ and $Y$ is $\frac{1}{2}$, then the variance of the sum is 6. Therefore, $P(X + 2Y \leq 3)$ is either $\Phi(\frac{3}{\sqrt{5}})$ or $\Phi(\frac{3}{\sqrt{6}})$.

5. a) As in Example 2, transform to independent variables $X$ and $Z$:

$$q = P(X \geq 0, Y \geq 0) = P\left[X \geq 0, Z \geq \frac{-\rho}{\sqrt{1-\rho^2}}X\right]$$

This last quantity is the probability that $(X, Z)$ lies in the sector centered at $(0,0)$ subtending an angle of

$$\frac{\pi}{2} + \arctan\left[\frac{\rho}{\sqrt{1-\rho^2}}\right] = \frac{\pi}{2} + \arcsin(\rho),$$

(where arcsin and arctan have range $[-\pi/2, \pi/2]$ ). Hence, by the rotational symmetry of $(X, Z)$, this probability is

$$q = \frac{\text{angle subtended}}{2\pi} = \frac{1}{4} + \frac{1}{2\pi}\arctan\left[\frac{\rho}{\sqrt{1-\rho^2}}\right] = \frac{1}{4} + \frac{1}{2\pi}\arcsin(\rho)$$

Note that $0 \leq q \leq 1/2$.

b) Solve for $\rho$ in the above formula for $q$. The arcsin version is most convenient for this:

$$\rho = \sin\left[2\pi(q - \frac{1}{4})\right], 0 \leq q \leq \frac{1}{2}.$$

This can be obtained from the arctan version, too, but you have to argue about the sign of $\rho$.

6. a) $X - kY$ is normal with mean 0 and variance $1 + k^2$.

$$P(X < kY) = P(X - kY < 0) = \frac{1}{2}$$

b) $U = \sqrt{3}X + Y, V = X - \sqrt{3}Y$.

$$P(U > kV) = P(\sqrt{3}X + Y > k(X - \sqrt{3}Y)) = P((\sqrt{3} - k)X + (1 + k\sqrt{3})Y > 0) = \frac{1}{2}$$

c)

$$\begin{aligned}
P(U^2 + V^2 < 1) &= P\left((\sqrt{3}X + Y)^2 + (X - \sqrt{3}Y)^2 < 1\right) \\
&= P\left(X^2 + Y^2 < \frac{1}{4}\right) = P\left(\sqrt{X^2 + Y^2} < \frac{1}{2}\right) \\
&= \int_0^{\frac{1}{2}} re^{-\frac{1}{2}r^2}dr = \int_0^{\frac{1}{8}} e^{-u}du = -e^{-u}\Big|_{u=0}^{\frac{1}{8}} \\
&= 1 - e^{-\frac{1}{8}}
\end{aligned}$$

since $\sqrt{X^2 + Y^2}$ has the Rayleigh distribution (or using polar co-ordinates).

d) $Var(V) = Var(X - \sqrt{3}Y) = 4$, so $V/2$ has the standard normal distribution

$$\begin{aligned}
Cov(X, V) &= E[X(X - \sqrt{3}Y)] = E(X^2) - \sqrt{3}E(XY) = 1 \\
corr(X, V/2) &= corr(X, V) = \frac{Cov(X, V)}{SD(X)SD(V)} = \frac{1}{2} \\
f_X(x \mid V = v) &= f_X(x \mid V/2 = v/2)
\end{aligned}$$

Because $X$ and $V/2$ have the standard bivariate normal distribution with $\rho = \frac{1}{2}$ $(X \mid V = v) = \frac{1}{2}\frac{v}{2}$ and $Var(X \mid V = v) = 1 - (\frac{1}{2})^2$. Therefore given $V = v$, the conditional distribution of $X$ is normal$(\frac{v}{4}, \frac{3}{4})$.

7. a) As always (Section 6.4) independent implies uncorrelated. Conversely, let $X^* = (X - \mu_X)/\sigma_X$ and $Y^* = (Y - \mu_Y)/\sigma_Y$ be the standardised versions of $X$ and $Y$. Then $X^*$ and $Y^*$ have standard bivariate normal distribution with $\rho = 0$ . (See box on bivariate normal.) But in the case $\rho = 0$ , inspection shows that the joint density of $X^*$ and $Y^*$ equals the product of the marginal densities, so $X^*$ and $Y^*$ are independent. (Another way: appeal to the definition of the standard bivariate normal distribution in terms of independent standard normal variables; if $\rho = 0$, then we return to the original pair of independent variables.) Being functions of $X^*$ and $Y^*$ respectively, it follows that $X$ and $Y$ are also independent.

b)

$$\begin{aligned} E(Y|X = x) &= E\left(\mu_Y + \sigma_Y Y^* \mid X^* = \frac{x - \mu_X}{\sigma_X}\right) \\ &= \mu_Y + \sigma_Y E\left(Y^* \mid X^* = \frac{x - \mu_X}{\sigma_X}\right) \\ &= \mu_y + \sigma_Y \cdot \rho \cdot \left(\frac{x - \mu_X}{\sigma_X}\right), \end{aligned}$$

since given $X^* = t, Y^*$ has normal $(\rho t, 1 - \rho^2)$ distribution.

c)

$$\begin{aligned} Var(Y|X = x) &= Var\left(\mu_Y + \sigma_Y Y^* \mid X^* = \frac{x - \mu_X}{\sigma_X}\right) \\ &= \sigma_Y^2 Var\left(Y^* \mid X^* = \frac{x - \mu_X}{\sigma_X}\right) \\ &= \sigma_Y^2(1 - \rho^2). \end{aligned}$$

d) In terms of the standardized versions, we have

$$\begin{aligned} aX + bY + c &= a(\mu_X + \sigma_X X^*) + b(\mu_Y + \sigma_Y Y^*) + c \\ &= a(\mu_X + \sigma_X X^*) + b\left[\mu_Y + \sigma_Y(\rho X^* + \sqrt{1 - \rho^2}Z)\right] + c \\ &= (a\sigma_X + b\rho\sigma_Y)X^* + b\sigma_Y\sqrt{1 - \rho^2}Z + a\mu_X + b\mu_Y + c, \end{aligned}$$

where $X^*$ and $Z$ are independent standard normal variables. Hence $aX + bY + c$ is a linear combination of independent normals, so is itself normally distributed. Its parameters are:

$$E(aX + bY + c) = aE(X) + bE(Y) + c = a\mu_X + b\mu_Y + c,$$

$$Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y.$$

e) Let $V = X\cos\theta + Y\sin\theta$ , $W = -X\sin\theta + Y\cos\theta$ . Then $V$ and $W$ are each normally distributed by part (d). To show independence, it suffices to show, by part (a), that $E(VW) = 0$. Note that

$$\begin{aligned} VW &= -X^2\sin\theta\cos\theta + XY\cos^2\theta - XY\sin^2\theta + Y^2\sin\theta\cos\theta \\ &= XY\cos 2\theta - \frac{1}{2}(X^2 - Y^2)\sin 2\theta. \end{aligned}$$

Therefore $E(VW) = \cos 2\theta E(XY) - \sin 2\theta E(X^2 - Y^2)/2 = 0$ if and only if

$$\begin{aligned} \cot 2\theta &= \frac{E(X^2 - Y^2)}{2E(XY)} = \frac{\sigma_X^2 - \sigma_Y^2}{2\rho\sigma_X\sigma_Y} \\ \iff \theta &= \frac{1}{2}\text{arccot}\left[\frac{\sigma_X^2 - \sigma_Y^2}{2\rho\sigma_X\sigma_Y}\right]. \end{aligned}$$

Geometric significance of $\theta$? If a point $(x, y)$ in the plane is rotated clockwise by an angle $\phi$ about the origin, its coordinates become $(x\cos\phi + y\sin\phi, y\cos\phi - x\sin\phi)$. Thus if a bivariate normal distribution in the plane, centered at the origin, is rotated clockwise by the above angle $\theta$ , one obtains a bivariate normal distribution whose axes of symmetry coincide with the coordinate axes. (Reason: By the above calculation, the new variables are independent, so the ellipses of constant density now have the form $ax^2 + by^2 =$ const .) So $\theta = (1/2)\text{arccot}[(\sigma_X^2 - \sigma_Y^2)/2\rho\sigma_X\sigma_Y]$ is the inclination of one of the axes of the ellipses of constant density for $(X, Y)$ .

8. a) Since $E(Y_1) = E(Y_2) = 0$,

$$\begin{aligned} Cov(Y_1, Y_2) &= E[(X_1 + X_2)(\alpha X_1 + 2X_2)] \\ &= \alpha E(X_1^2) + 2E(X_2^2) + (\alpha + 2)E(X_1 X_2) \\ &= \alpha + 2 = 0 \end{aligned}$$

so, $\alpha = -2$. Therefore $Y_2$ has the normal$(0, 16)$ density.

b) $Cov(X_2, Y_2) = E[X_2(2X_2 - 2X_1)] = 2E(X_2^2) - 2E(X_1 X_2) = 2$

9. a) Means: $\mu$, $a\mu + b$. Variances: $\sigma^2$, $a^2\sigma^2 + \tau^2$. Correlation: $a\sigma/\sqrt{a^2\sigma^2 + \tau^2}$ b) normal $(a\mu + b, a^2\sigma^2 + \tau^2)$ c) normal $(\mu + a\sigma^2(z - a\mu - b)/(a^2\sigma^2 + \tau^2), \sigma^2\tau^2/(a^2\sigma^2 + \tau^2))$

10. a) Let $corr(V, W) = \rho$.

$$\frac{W - \mu_W}{\sigma_W} = \rho\left(\frac{V - \mu_V}{\sigma_V}\right) + \sqrt{1 - \rho^2}Z$$

so,

$$W = \mu_W + \sigma_W\rho\left(\frac{V - \mu_V}{\sigma_V}\right) + \sigma_W\sqrt{1 - \rho^2}Z$$

Now show that $aV + bW$ is a sum of independent normal random variables (plus a constant), and hence is normal.

$$\begin{aligned} aV + bW &= aV + b\mu_W + b\sigma_W\rho\left(\frac{V - \mu_V}{\sigma_V}\right) + b\sigma_W\sqrt{1 - \rho^2}Z \\ &= \left(a + \frac{b\sigma_W\rho}{\sigma_V}\right)V + b\sigma_W\sqrt{1 - \rho^2}Z + \left(b\mu_W - \frac{b\sigma_W\rho\mu_V}{\sigma_V}\right) \end{aligned}$$

b) We need to show that $cV + dW = \alpha(aV + bW) + \beta(Z')$ where $aV + bW$ and $Z'$ are independent.

11. See solution of Exercise 6.Rev.18.

12. a) $Cov(S, V) = E[(S - E(S))(V - E(V))] = E[(bV + W)(V)] = E(bV^2 + VW) = b\sigma_V{}^2 + 0$; hence $Corr(S, V) = \frac{Cov(S,V)}{SD(S)SD(V)} = \frac{b\sigma_V}{\sigma_S} = \frac{b\sigma_V}{\sqrt{b^2\sigma_V{}^2 + \sigma_W{}^2}}$.

b) Since $S$ and $V$ are two different linear combinations of independent normal variables, they have a bivariate normal distribution with

$$E(S) = \mu_S = a \qquad SD(S) = \sigma_S = \sqrt{b^2\sigma_V{}^2 + \sigma_W{}^2}$$

$$E(V) = \mu_V = 0 \qquad SD(V) = \sigma_V$$

$$Corr(S, V) = \rho = \frac{b\sigma_V}{\sqrt{b^2\sigma_V{}^2 + \sigma_W{}^2}}$$

Let $X$ and $Y$ be standardized versions of $S$ and $V$, i.e.,

$$X = \frac{S - \mu_S}{\sigma_S}, \qquad Y = \frac{V - \mu_V}{\sigma_V}.$$

Then $X$ and $Y$ have standard bivariate normal distribution with correlation $\rho$, so given $X = x$, $Y$ has normal $(\rho x, 1 - \rho^2)$ distribution, and $V = \sigma_V Y$ has normal $(\rho\sigma_V x, \sigma_V{}^2(1 - \rho^2))$ distribution. The event $(S = s)$ is the same as the event $\left(X = \frac{s - \mu_S}{\sigma_S}\right)$; thus, given that $S = s$, the distribution of $V$ is normal with mean

$$\rho\sigma_V\left(\frac{s - \mu_S}{\sigma_S}\right) = \frac{b\sigma_V{}^2(s - b)}{b^2\sigma_V{}^2 + \sigma_W{}^2}$$

and variance

$$\sigma_V{}^2(1 - \rho^2) = \frac{\sigma_V{}^2\sigma_W{}^2}{b^2\sigma_V{}^2 + \sigma_W{}^2}.$$

c) The best estimate of $V$ given $S = s$ is $E(V|S = s) = \rho\sigma_V\left(\frac{s - \mu_S}{\sigma_S}\right) = \frac{b\sigma_V{}^2(s - b)}{b^2\sigma_V{}^2 + \sigma_W{}^2}$.

d) $E\left|V - \rho\sigma_V\left(\frac{S - \mu_S}{\sigma_S}\right)\right| = \sigma_V E|Y - \rho X| = \sigma_V\sqrt{1 - \rho^2}E|Z| = \frac{\sigma_V\sigma_W}{\sqrt{b^2\sigma_V{}^2 + \sigma_W{}^2}}\sqrt{\frac{2}{\pi}}$.

13. The standard bivariate normal distribution with correlation $\rho$ can be obtained by taking two independent standard normal variables $X$ and $Z$ and transforming $(X, Z)$ into $(X, Y)$ via the transformation

$$\begin{aligned} x &= x \\ y &= x\cos\theta + z\sin\theta \end{aligned}$$

where $\theta$ is that unique angle in $[0, \pi]$ whose cosine is $\rho$; the joint distribution of $(X, Y)$ is then the required distribution. Under this transformation, the image of the point $(x, z) = (\cos(\theta/2), \sin(\theta/2))$ is

$$(x, y) = (\cos(\theta/2), \cos(\theta/2)\cos\theta + \sin(\theta/2)\sin\theta) = (\cos(\theta/2), \cos(\theta/2)),$$

a point on the 45° line in the $(X, Y)$ plane. Similarly, under this transformation, the image of the point $(x, z) = (\cos(\theta/2 + \pi/2), \sin(\theta/2 + \pi/2)) = (-\sin(\theta/2), \cos(\theta/2))$ is

$$(x, y) = (-\sin(\theta/2), -\sin(\theta/2)\cos\theta + \cos(\theta/2)\sin\theta) = (-\sin(\theta/2), \sin(\theta/2)),$$

a point on the -45° line in the $(X, Y)$ plane. These two images determine the axes of an ellipse of constant density for the distribution of $(X, Y)$, so the ratio of the lengths of these axes is

$$\frac{\sqrt{2\cos^2(\theta/2)}}{\sqrt{2\sin^2(\theta/2)}} = \frac{\sqrt{1+\cos\theta}}{\sqrt{1-\cos\theta}} = \sqrt{\frac{1+\rho}{1-\rho}}.$$

# Chapter 6: Review

1. a) $\{0, 1, 2, \dots, 100\}$.

   b) For k between 0 and 100,

$$\begin{aligned}
P(X = k \mid X + Y = 100) &= \frac{P(X = k, X + Y = 100)}{P(X + Y = 100)} \\
&= \frac{P(X = k, Y = 100 - k)}{P(X + Y = 100)} \\
&= \frac{e^{-\lambda_1}\frac{\lambda_1^k}{k!} \times e^{-\lambda_2}\frac{\lambda_2^{100-k}}{(100-k)!}}{e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^{100}}{100!}} \\
&= \frac{100!}{k!(100-k)!}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{100-k}
\end{aligned}$$

   This is the binomial distribution with parameters 100 and $\lambda_1/(\lambda_1 + \lambda_2)$.

   c) Given $X + Y = 100$, $X$ has the binomial distribution with parameters 100 and $1/100$, by part b). This is approximately the Poisson (1) distribution. The chance is about

$$e^{-1}\{1/4! + 1/5! + 1/6!\} = e^{-1}\{1/24 + 1/120 + 1/720\} = 0.018472$$

2. For each integer $x \geq 0$, $y \geq 0$ with $x + y \geq 1$ we have

$$\begin{aligned}
P(X = x, Y = y) &= P(X = x, N = x + y) \\
&= P(X = x|N = x + y)P(N = x + y) \\
&= \binom{x+y}{y}(1/2)^{x+y}(1/3)(2/3)^{x+y-1} \\
&= \binom{x+y}{y}(1/3)^{x+y}(1/2).
\end{aligned}$$

   Hence

$$\begin{aligned}
P(X = x, Y = 0) &= (1/3)^x(1/2), \quad x = 1, 2, 3, \dots \\
P(Y = 0) &= \sum_{x=1}^{\infty}(1/3)^x(1/2) = 1/4 \\
P(X = x|Y = 0) &= \frac{P(X = x, Y = 0)}{P(Y = 0)} = 2(1/3)^x = (2/3)(1/3)^{x-1}, x = 1, 2, 3, \dots
\end{aligned}$$

   So $X$ has geometric $(p = 2/3)$ distribution on $\{1, 2, 3, \dots\}$ given $Y = 0$. The most likely value of $X$ given $Y = 0$ is then 1, and $E(X|Y = 0) = 3/2$.

3. First note that $A_n + B_n = 2\mu$ , hence $A_n - B_n = 2A_n - 2\mu$ ; then apply the results of Example 6.4.7 to get $E(A_n)$ and $SD(A_n)$. Hence

$$\begin{aligned}
E(A_n - B_n) &= 2E(A_n) - 2\mu = 2\mu - 2\mu = 0; \\
SD(A_n - B_n) &= 2SD(A_n) = 2\frac{\sigma}{\sqrt{n}}\sqrt{\frac{2n - n}{2n - 1}} = \frac{2\sigma}{\sqrt{2n - 1}}.
\end{aligned}$$

4. No Solution

5. The joint density looks like that of a bivariate normal distribution. Indeed, try to find positive constants $a, b$ so that the distribution of $(X/a, Y/b)$ is standard bivariate normal with correlation $\rho$. The joint density of $(X/a, Y/b)$ is

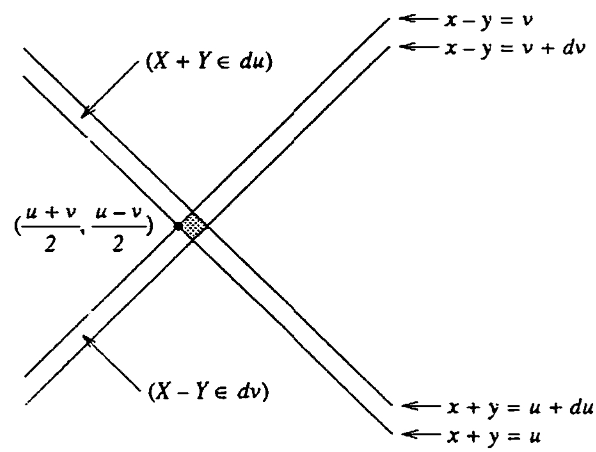
$$abf(ax, by) = abc\exp\left\{-(a^2x^2 + abxy + b^2y^2)\right\}.$$

Comparing this with the standard bivariate normal joint density yields the equations

$$a^2 = \frac{1}{2(1-\rho^2)}$$
$$ab = \frac{-2\rho}{2(1-\rho^2)}$$
$$b^2 = \frac{1}{2(1-\rho^2)}$$
$$abc = \frac{1}{2\pi\sqrt{1-\rho^2}}$$

which have solution $a = b = \sqrt{\frac{2}{3}}$ , $\rho = -\frac{1}{2}$ , $c = \sqrt{\frac{3}{2\pi}}$. The correlation between $X$ and $Y$ is the same as (since correlation is invariant under linear transformations) the correlation between $X/a$ and $Y/b$ , which is $\rho = -1/2$.

6. a) Use the technique of Section 5.4 (Operations):




$$P(X + Y \in du, X - Y \in dv) = f_{X,Y}\left(\frac{u+v}{2}, \frac{u-v}{2}\right)\frac{du}{\sqrt{2}}\frac{dv}{\sqrt{2}}.$$

Since the joint density of $(X, Y)$ is $f_{X,Y}(x, y) = e^{-x}e^{-y}(x \geq 0, y \geq 0)$ , the joint density of $(X + Y, X - Y)$ is

$$f(u,v) = \frac{1}{2}f_{X,Y}\left(\frac{u+v}{2}, \frac{u-v}{2}\right)$$
$$= \begin{cases} (1/2)e^{-(\frac{u+v}{2})}e^{-(\frac{u-v}{2})} & \text{if } \frac{u+v}{2} \geq 0 \text{ and } \frac{u-v}{2} \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

This simplifies to

$$f(u,v) = \begin{cases} (1/2)e^{-u} & \text{if } u \geq 0 \text{ and } -u \leq v \leq u \\ 0 & \text{otherwise.} \end{cases}$$

Check: Compute the marginals! Do you get what you expect?

b) If $X$ and $Y$ are independent with the same distribution, then

$$Cov(X + Y, X - Y) = E((X + Y)(X - Y)) - E(X + Y)E(X - Y)$$
$$= E(X^2 - Y^2) - E(X + Y)E(X - Y) = 0,$$

so the correlation between $X$ and $Y$ is zero.

7. The joint density of $(X, Y)$ is

$$f(x,y) = \begin{cases} 2 & \text{if } 0 < x < 1 \text{ and } 0 < y < 1 \text{ and } x + y < 1 \\ 0 & \text{otherwise} \end{cases}$$

a) The marginal density of $X$ is

$$f_X(x) = \begin{cases} 2(1-x) & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

so $E(X) = 1/3$, $Var(X) = 1/18$.

b) Given $Y = 1/3$, $X$ has uniform distribution on $(0, 2/3)$, so
$E(X|Y = 1/3) = 1/3$, $\quad Var(X|Y = 1/3) = 1/27$.

c) The density of $\max(X, Y)$ is

$$f(z) = \begin{cases} 4z & \text{if } 0 < z < 1/2 \\ 4(1-z) & \text{if } 1/2 < z < 1 \\ 0 & \text{otherwise} \end{cases}$$

(i.e., it is triangular on $(0, 1)$), so $max(X, Y)$ has mean $1/2$ and variance $1/24$.

d) The density of $\min(X, Y)$ is

$$f(w) = \begin{cases} 4(1-2w) & \text{if } 0 < w < 1/2 \\ 0 & \text{otherwise} \end{cases}$$

so $\min(X, Y)$ has mean $1/6$ and variance $1/72$.

8. a) Given:

$$\begin{aligned} f_Y(y) &= 2e^{-2y}, & y > 0 \\ f_X(x|Y = y) &= \tfrac{1}{y}e^{-x/y}, & x > 0 \end{aligned}$$

Hence $f_{X,Y}(x,y) = f_X(x|Y = y)f_Y(y) = \frac{2}{y}e^{-x/y}e^{-2y}$, $x > 0, y > 0$. Recall that if $T$ has exponential distribution, then $Var(T) = [E(T)]^2$ and $E(T^2) = 2(ET)^2$. So

$$E(Y) = 1/2, \quad E(Y^2) = 2(1/2)^2 = 1/2;$$

and

$$E(X|Y = y) = y, \quad E(X^2|Y = y) = 2y^2.$$

b) $E(X|Y) = Y$, so $\quad E(X) = EE(X|Y) = EY = 1/2$

c) $E(XY|Y = y) = E(Xy|Y = y) = yE(X|Y = y) = y^2$, so $\quad E(XY|Y) = Y^2$. Hence

$$E(XY) = EE(XY|Y) = EY^2 = 1/2.$$

Next, $E(X^2|Y) = 2Y^2$, so

$$E(X^2) = EE(X^2|Y) = E(2Y^2) = 2EY^2 = 1$$

and

$$Var(X) = E(X^2) - (EX)^2 = 1 - (1/2)^2 = 3/4.$$

Therefore

$$Corr(X,Y) = \frac{Cov(X,Y)}{SD(X)SD(Y)} = \frac{E(XY) - (EX)(EY)}{\sqrt{Var(X)Var(Y)}} = \frac{(1/2) - (1/2)(1/2)}{\sqrt{(3/4)(1/4)}} = \frac{1}{\sqrt{3}}.$$

9. Condition on $Z$.

$$P[(X/Y) > (Y/Z)] = P\left(X > \frac{Y^2}{Z}\right)$$

$$= \int_0^1 P\left(X > \frac{Y^2}{Z}\Big| Z = z\right) dz = \int_0^1 P\left(X > \frac{1}{z}Y^2\right) dz$$

$$P\left(X > \frac{1}{z}Y^2\right) = \int_0^{\sqrt{z}} 1 - \frac{y^2}{z} dy = y - \frac{y^3}{3z}\Big|_{y=0}^{\sqrt{z}} = \frac{2}{3}z^{\frac{1}{2}}$$

Substituting into the first integral,

$$P[(X/Y) > (Y/Z)] = \int_0^1 \frac{2}{3}z^{\frac{1}{2}} dz = \frac{2}{3}\frac{z^{\frac{3}{2}}}{(3/2)}\Big|_{z=0}^{1} = \frac{4}{9}$$

10. a) exponential $(\alpha + \beta + \gamma)$

b) $\gamma/(\alpha + \beta + \gamma)$

c) By the memoryless property of the exponential distributions of $T_A$ and $T_B$, it's the same as the distribution of $min(T_A, T_B)$, i.e., exponential $(\alpha + \beta)$.

d) Call this time $T_2$. To get the c.d.f. of $T_2$, use $P(T_2 \le t) = 1 - P(T_2 > t)$. and condition on which component fails first:

$$P(T_2 > t) = \frac{\alpha}{\alpha + \beta + \gamma}e^{-(\beta+\gamma)t} + \frac{\beta}{\alpha + \beta + \gamma}e^{-(\gamma+\alpha)t} + \frac{\gamma}{\alpha + \beta + \gamma}e^{-(\alpha+\beta)t}$$

11. a) Let $N_t$ = number of claims in time $t$. Given $N_t = n$,

$$X_t = Y_1 + \cdots + Y_n$$

where the $Y_i$ are independent exponential $(\mu)$, with mean $1/\mu$. So

$$E[X_t | N_t = n] = \frac{n}{\mu}$$

$$E[X_t] = E[E[X_t|N_t]] = E(\frac{N_t}{\mu}) = \frac{E(N_t)}{\mu} = \frac{\lambda t}{\mu}$$

b) Similarly

$$\begin{aligned} E[X_t^2 | N_t = n] &= E(Y_1 + \cdots + Y_n)^2 \\ &= (\tfrac{n}{\mu})^2 + Var(Y_1 + \cdots + Y_n) \\ &= (\tfrac{n}{\mu})^2 + \tfrac{n}{\mu^2} = \tfrac{n(n+1)}{\mu^2} \\ E(X_t^2) &= E[\tfrac{N_t(N_t+1)}{\mu^2}] = \tfrac{1}{\mu^2}[Var(N_t) + [E(N_t)]^2 + E(N_t)] \\ &= (\lambda^2 t^2 + 2\lambda t)/\mu^2 \end{aligned}$$

c) $SD(X_t) = \sqrt{2\lambda t/\mu^2} = \frac{\sqrt{2\lambda t}}{\mu}$

d)

$$\begin{aligned} Cov(X_s, X_t) &= Cov(X_s, X_s + X_t - X_s) \\ &= Cov(X_s, X_s) + Cov(X_s, X_t - X_s) \\ &= Var(X_s) \end{aligned}$$

because $X_s$ and $X_t - X_s$ are independent. So

$$Corr(X_s, X_t) = \frac{[SD(X_s)]^2}{SD(X_s)SD(X_t)} = \frac{SD(X_s)}{SD(X_t)} = \sqrt{\frac{s}{t}}$$

12. a) Let $X_i$ be the weight of the $i$th person.

$$\begin{aligned} P(\text{total weight} > 4000) &= P\left(\sum_{i=1}^{26} X_i > 4000\right) \\ &= 1 - \Phi\left(\frac{4000 - 26(150)}{\sqrt{26} \times 30}\right) = 1 - \Phi(0.65) = 1 - 0.7422 = 0.2578 \end{aligned}$$

b) Let $W_i$ be the weight of the object the $i$th person is carrying.

$$\begin{aligned} E(X_i + W_i) &= E(X_i) + E[E(W_i \mid X_i)] = 150 + E(0.05X_i) = 157.5 \\ E[(X_i + W_i)^2] &= E(X_i^2) + 2E(X_iW_i) + E(W_i^2) \\ &= E(X_i^2) + 2E[E(X_iW_i \mid X_i)] + E[E(W_i^2 \mid X_i)] \\ &= E(X_i^2) + 2E(0.05X_i^2) + E[(0.05X_i)^2 + 4] \\ &= (1 + 2(0.05) + (0.05)^2)(150^2 + 900) + 4 = 25802.5 \\ Var(X_i + W_i) &= 25802.5 - (157.5)^2 = 996.25 \\ P(\text{total weight} > 4000) &= P\left(\sum_{i=1}^{26}(X_i + W_i) > 4000\right) \\ &= 1 - \Phi\left(\frac{4000 - 26(157.5)}{\sqrt{26 \times 996.25}}\right) \\ &= 1 - \Phi(-0.59) = \Phi(0.59) = 0.7224 \end{aligned}$$

c) Ignore any packing constraints and assume that the people will fit as long as the total area they need is at most $54 \times 92 = 4968$ square inches.

$$\Phi\left(\frac{4968 - 20\mu}{\sqrt{20}(0.1\mu)}\right) = 0.99 = \Phi(2.33)$$

Thus,

$$\mu = \frac{4968}{2.33(0.1)\sqrt{20} + 20} = 236 \text{ square inches}$$

13.

$$\begin{aligned} Cov(X - Y, X + Y) &= Var(X) - Var(Y) - Var(X) - Var(Y) \\ &= 0 \text{ iff } Var(X) = Var(Y) \end{aligned}$$

a) By a), $X - Y$ and $X + Y$ are uncorrelated because $Var(X) = 1 = Var(Y)$. So $X - Y$, $X + Y$ are uncorrelated (hence independent) normals,

$$\begin{aligned} P(X - Y < 1, X + Y > 2) &= P(X - Y) < 1)P(X + Y > 2) \\ &= \Phi(\tfrac{1}{\sqrt{0.8}})[1 - \Phi(\tfrac{2}{\sqrt{3.2}})] \\ &= (0.8687)(1 - 0.8687) \\ &= 0.114 \end{aligned}$$

14. Pick a person at random from the population. Let $H$ be that person's height, and let $G$ $(= m, f)$ be the gender of that person.

a) $E(H) = E(H|G = m)P(G = m) + E(H|G = f)P(G = f) = 67 \cdot (1/2) + 63 \cdot (1/2) = 65$.

b)

$$\begin{aligned} E(H^2) &= E(H^2|G = m)P(G = m) + E(H^2|G = f)P(G = f) \\ &= (3^2 + 67^2) \cdot (1/2) + (3^2 + 63^2) \cdot (1/2) \\ &= 4238 \end{aligned}$$

so $Var(H) = 4238 - (65)^2 = 13$. Or use

$$Var(H) = E[Var(H|G)] + Var[E(H|G)] = (3^2 + 3^2) \cdot (1/2) + (\frac{67 - 63}{2})^2 = 9 + 4 = 13.$$

c)

$$\begin{aligned} P(63 < H < 67) &= P(63 < H < 67|G = m)P(G = m) + P(63 < H < 67|G = f)P(G = f) \\ &= (1/2)P\left(\frac{63-67}{3} < Z < \frac{67-67}{3}\right) + (1/2)P\left(\frac{63-67}{3} < Z < \frac{67-63}{3}\right) \\ &= \Phi(4/3) - \Phi(0) \\ &= .4088 \end{aligned}$$

d) $P(63 < H < 67) = \Phi\left(\frac{67-65}{\sqrt{13}}\right) - \Phi\left(\frac{63-65}{\sqrt{13}}\right) = .4209.$

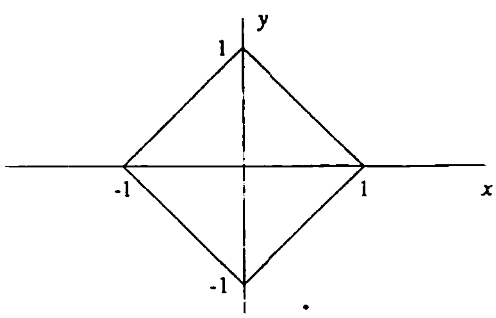
The answers are different because in c), the distribution of $H$ is *not* normal: it's a mixture of normals, which results in a normal distribution only if the two distributions being mixed are identical. The answers to c) and d) differ only slightly because the distributions in c) and d) are both unimodal and symmetric about 65, with the same standard deviation, and we are seeking the probability of an event which is symmetric about 65.

**Remark:** You can show by calculus that a half-and-half mixture of a normal $(a, \sigma^2)$ distribution with a normal $(b, \sigma^2)$ distribution will have a bimodal density if and only if $|a - b| > 2\sigma$.

e) Let $X$ be the man's height, and $Y$ the woman's height. Argue that $X - Y = X + (-Y)$ has normal distribution with mean $67 - 63 = 4$, and variance $3^2 + 3^2 = 18$, and therefore

$$P(X > Y) = P(X - Y > 0) = P\left(\frac{X - Y - 4}{\sqrt{18}} > \frac{-4}{\sqrt{18}}\right) = \Phi(4/\sqrt{18}) \approx .83.$$

15. a) By rotational symmetry, the desired probability is one-quarter the chance that $(X, Y)$ falls in the square:




The square has side length $\sqrt{2}$, so after a 45° rotation its probability is seen to be

$$\left[\Phi(\frac{\sqrt{2}}{2}) - \Phi(-\frac{\sqrt{2}}{2})\right]^2 = \left[2\Phi(1/\sqrt{2}) - 1\right]^2.$$

Hence the desired probability is

$$P(X \geq 0, Y \geq 0, X + Y \leq 1) = \frac{1}{4}\left[2\Phi(1/\sqrt{2}) - 1\right]^2 = \left[\Phi(1/\sqrt{2}) - 1/2\right]^2.$$

b) By the same argument, we have for $t > 0$

$$P(X + Y \leq t, X \geq 0, Y \geq 0) = \left[\Phi(t/\sqrt{2}) - 1/2\right]^2$$

and

$$P(X + Y \leq t | X \geq 0, Y \geq 0) = \left[2\Phi(t/\sqrt{2}) - 1\right]^2.$$

Differentiate to get

$$P(X + Y \in dt | X \geq 0, Y \geq 0) = 2\sqrt{2}\phi(t/\sqrt{2})\left[2\Phi(t/\sqrt{2}) - 1\right] dt.$$

c) For the median, find $t^*$ so that

$$\left[2\Phi(t^*/\sqrt{2}) - 1\right]^2 = \frac{1}{2}$$
$$\iff \quad \Phi(t^*/\sqrt{2}) = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) = .8536$$
$$\iff \quad \frac{t^*}{\sqrt{2}} = 1.05 \quad \iff \quad t^* = 1.49.$$

d) For the mode, ignore the $2\sqrt{2}$, let $x = t/\sqrt{2}$ in the density, differentiate with respect to $x$ and set the result equal to 0: We have

$$\frac{d}{dx}\phi(x)\,[2\Phi(x) - 1] = \phi'(x)\,[2\Phi(x) - 1] + 2\phi(x)^2,$$

where $\phi'(x) = -x\phi(x)$. Thus $x$ solves

$$\Phi(x) = \frac{\phi(x)}{x} + \frac{1}{2}.$$

Trial and error gives $x = .876901$, hence the mode is $t = \sqrt{2}x = 1.240125$ .

16. a) Let $X$ denote the rainfall in any particular year. $X$ has gamma($r = 3, \lambda = 3/20$) distribution. So

$$P(X > 35) = P(N_{35} \le 2) = e^{-5.25}(1 + 5.25 + \frac{1}{2}(5.25)^2) \approx .1051,$$

where $N_{35}$ has Poisson($\frac{3}{20} \times 35$) distribution.

b) $(1 - .1051)^{10} \approx .3294$.

c) $1 - [1 - P(X > R_{20})]^{10}$; calculate $P(X > R_{20})$ as in part a).

d) Let $M_1$ be the maximum rainfall over the period from 20 years ago to 10 years ago, let $M_2$ be the maximum rainfall over the last 10 years, and let $M_3$ be the maximum rainfall over the next 10 years. The $M$'s are independent and identically distributed, by assumption. Want $P(M_3 > \max[M_1, M_2])$, which is 1/3 by symmetry.

17. **Fact:** $(X, Y)$ has a joint distribution which is symmetric under rotations if and only if for all $\theta$, $(X, Y)$ has the same distribution as $(X\cos\theta - Y\sin\theta, Y\cos\theta + X\sin\theta)$ .
**Proof:** The point $(x, y)$, when rotated by $\theta$ about the origin, is mapped to the point $(x\cos\theta - y\sin\theta, y\cos\theta + x\sin\theta)$ .

a) $X$ and $Y$ are not necessarily independent. For example, (anticipating part (b)) let $(X, Y)$ be a point (in the plane) picked uniformly at random from the unit circle $\{(x, y) : x^2 + y^2 = 1\}$ . Then, given $X = x$ , the value of $Y$ is constrained to at most two points, whereas unconditionally, $Y$ has a continuous distribution on the interval $(-1, 1)$ .

In general, $X$ and $Y$ are uncorrelated: Put $\theta = \pi$ in the above Fact to see that $(X, Y)$ has the same distribution as $(-X, -Y)$ ; hence $E(X) = E(Y) = 0$ (this is intuitively clear). Put $\theta = \pi/2$ to see that $(X, Y)$ has the same distribution as $(-Y, X)$ ; hence $X$ has the same distribution as $Y$. (Again, this is intuitively clear.) Put $\theta = \pi/4$ to see that $(X, Y)$ has the same distribution as $\left(\frac{X-Y}{\sqrt{2}}, \frac{X+Y}{\sqrt{2}}\right)$ . Hence

$$E(XY) = E\left(\frac{X - Y}{\sqrt{2}}\right)\left(\frac{X + Y}{\sqrt{2}}\right) = \frac{E(X^2 - Y^2)}{2} = 0 = E(X)E(Y).$$

(The third equality follows from the fact that $X$ has the same distribution as $Y$.)

b) The distribution of $(X, Y)$ is clearly symmetric under rotations. Since $X$ has the same distribution as $Y$ (see part (a)), we have $E(X^2) = E(Y^2)$. But $X^2 + Y^2 = 1$. Hence $E(X^2) + E(X^2) = 1$ , and $E(X^2) = E(Y^2) = 1/2$. Since $X$ and $Y$ are uncorrelated, we have $E(XY) = 0$ (see part (a)).

c) $(X, Y)$ is uniformly distributed on the unit circle, as in (b).
[Reason: Let $\Theta$ ( $-\pi < \Theta < \pi$ ) be the angle subtended by the point $(X, Y)$ . Show that $\Theta$ is uniform on $(-\pi, \pi)$: If $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ then

$$\begin{aligned} P(\theta < \Theta < \theta + d\theta) &= P(\theta < \tan^{-1}(Y/X) < \theta + d\theta) \\ &= P(\tan\theta < Y/X < \tan(\theta + d\theta)) \\ &= P(\tan\theta < \tan 2\pi U < \tan(\theta + d\theta)) \\ &= P(\theta < 2\pi U < \theta + d\theta) \\ &= \frac{1}{2\pi} d\theta; \end{aligned}$$

and similarly for the remaining $\theta$ . ]
$(X, Y)$ are uncorrelated, not independent: see part (a).

18. a) $P(\max \in dx) = 2\Phi(x)\phi(x)dx$, hence

$$\begin{aligned} E(\max) &= \int_{-\infty}^{\infty} 2x\Phi(x)\phi(x)dx \\ &= \int_{-\infty}^{\infty} 2[\frac{d}{dx} - \phi(x)]\Phi(x)dx \\ &= \int_{-\infty}^{\infty} 2\phi(x)\phi(x)dx \\ &= 1/\sqrt{\pi}. \end{aligned}$$

By symmetry, $E(\min) = -1/\sqrt{\pi}$.

b) By scaling from a),

$$E(\max) = \mu + \frac{\sigma}{\sqrt{\pi}}, \qquad E(\min) = \mu - \frac{\sigma}{\sqrt{\pi}}.$$

c)

$$P(\max \in dx) = 2P(X \in dx)P(Y \leq x | X = x) = 2\phi(x)\Phi\left(\frac{x - \rho x}{\sqrt{1 - \rho^2}}\right) dx$$

Integrating by parts as in a) gives

$$E(\max) = \sqrt{\frac{1 - \rho}{\pi}}.$$

19. a) The fraction $F$ of elements seen in the sample is

$$F = \frac{1}{n} \sum_{i=1}^{n} I(\text{element } i \text{ is seen at least once}).$$

The indicators are identically distributed, so $F$ has expectation

$$\begin{aligned} E(F) &= \frac{1}{n} \times n \times P(\text{element 1 is seen at least once}) \\ &= 1 - P(\text{element 1 is not seen}) = 1 - \left(1 - \frac{1}{n}\right)^n \end{aligned}$$

b) About $1 - e^{-1}$.

c) The number of elements *not* seen is $N = \sum_{i=1}^{n} J_i$ where $J_i = I(\text{element } i \text{ not seen})$. The $J_i$'s are identically distributed (not independent). Argue that

$$Var(J_i) = E(J_i) - [E(J_i)]^2 \leq 1/4$$

and (if $i \neq k$)

$$Cov(J_i, J_k) = E(J_i J_k) - E(J_i)E(J_k) = \left(1 - \frac{2}{n}\right)^n - \left(1 - \frac{1}{n}\right)^{2n}.$$

Note that $Cov(J_i, J_k) \leq 0$ for all $n$ (if $n \geq 2$, use the fact that $0 \leq 1 - 2/n \leq (1 - 1/n)^2$ ); so

$$Var(N) = Var(\sum_{i=1}^{n} J_i) = nVar(J_1) + n(n-1)Cov(J_1, J_2) \leq n/4,$$

and

$$Var(F) = Var\left(\frac{n-N}{n}\right) = \frac{Var(N)}{n^2} \leq \frac{1}{4n}.$$

d) Chebychev gives

$$P(|F - E(F)| > .01) \leq \frac{Var(F)}{(.01)^2} \leq \frac{2500}{n}.$$

If $n = \binom{52}{5}$, this last is .00096.

20. Let $X_0$ be the result of the first roll. Then

$E(Y) = \sum_{x=2}^{12} E(Y|X_0 = x)P(X_0 = x)$.

If $x = 2, 3, 7, 11$, or $12$, then $E(Y|X_0 = x)$ is 1. Otherwise it equals $1 + E(G_x)$, where $G_x$ denotes the number of throws of a pair of dice required to see either an $x$ or a 7. $G_x$ has the geometric distribution on { 1, 2, ... } with parameter $p = P(x) + P(7) = (x+5)/36$; therefore $E(G_x) = 1/p = 36/(x+5)$, and

$E(Y) = 1 \cdot \frac{12}{36} + \left(1 + \frac{36}{9}\right) \cdot \frac{3}{36} \cdot 2 + \left(1 + \frac{36}{10}\right) \cdot \frac{4}{36} \cdot 2 + \left(1 + \frac{36}{11}\right) \cdot \frac{5}{36} \cdot 2 = 3.3757576.$

21. a) $E(W_H) = \frac{1}{p}$

b) Let $E(W_{HH}) = x$. Then

$$\begin{aligned} x &= (1-p)(1+x) + p(2p + (2+x)(1-p)) \\ &= q + qx + 2p^2 + 2pq + xpq \\ x &= \frac{q + 2pq + 2p^2}{1 - q - pq} \end{aligned}$$

c) $W_{HHH} = \begin{cases} k + W & \text{if } W_T = k,\ k = 1, 2, 3 \\ 3 & \text{else} \end{cases}$ where given that $W_T = k$, the r.v. $W$ has the same distribution as $W_{HHH}$.

$$\begin{aligned} \text{So } x &= (1+x)q + (2+x)pq + (3+x)p^2q + 3p^3 \\ &= q + 2pq + 3p^2q + 3p^3 + (q + pq + p^2q)x \\ x &= \frac{q + 2pq + 3p^2q + 3p^3}{1 - q - pq - p^2q} \end{aligned}$$

d)

$$x = \sum_{k=1}^{m} (k + x)p^{k-1}q + mp^m$$

$$\text{So } x = \frac{\sum_{k=1}^{m} kp^{k-1}q + mp^m}{1 - \sum_{k=1}^{m} p^{k-1}q}$$

22. (i) Let $W_n$ denote the number of trials until $n$ heads in a row will be observed. Then

$$W_n = \begin{cases} W_{n-1} + 1 & \text{with probability } 1/2 \\ W_{n-1} + 1 + W_n^* & \text{with probability } 1/2 \end{cases}$$

where $W_n^*$ is an independent copy of $W_n$. Therefore

$E(W_n) = \frac{1}{2}E(W_{n-1} + 1) + \frac{1}{2}E(W_{n-1} + 1 + W_n^*)$,

which implies

$\frac{1}{2}E(W_n) = E(W_{n-1}) + 1$ .

Solve the above difference equation with $E(W_1) = 2$: get $E(W_n) = 2^{n+1} - 2$.

Put $n = 100$: then $E(W_{100}) = 2^{101} - 2$ seconds $\approx 8.04 \times 10^{22}$ years.

(ii) Use Markov's inequality to get a rough answer: since $P(W_{100} > n) \le \frac{E(W_{100})}{n}$, it will do to take $n$ such that

$$\frac{E(W_{100})}{n} \le .01 \longleftrightarrow n \ge 10^2 \times (2^{101} - 2).$$

So, roughly $8.04 \times 10^{24}$ years.

**23.** Let $X$ denote the total number of eggs laid, $Y$ the number that hatch and $Z$ the number that don't hatch. Then $X = Y + Z$.

$$\begin{aligned} P(Y = y) &= \sum_{x=y}^{\infty} \binom{x}{y} p^y (1-p)^{x-y} \frac{e^{-\lambda}\lambda^x}{x!} \\ &= \frac{e^{-\lambda} p^y \lambda^y}{y!} \sum_{x=y}^{\infty} \frac{(1-p)^{x-y}\lambda^{x-y}}{(x-y)!} \\ &= \frac{e^{-p\lambda}(p\lambda)^y}{y!} \end{aligned}$$

So $Y$ has a Poisson($p\lambda$) distribution. Similarly $Z$ has a Poisson $((1-p)\lambda)$ distribution.

Independence:

$$\begin{aligned} P(Y = y, Z = z) &= P(Y = y, Z = z, X = y + z) \qquad \text{since } (Y = y, Z = z) \subset (X = y + z) \\ &= P(Y = y, Z = z | X = y + z) P(X = y + z) \\ &= \binom{y+z}{y} p^y (1-p)^z \frac{e^{-\lambda}\lambda^{y+z}}{(y+z)!} \\ &= \frac{e^{-\lambda}\lambda^{y+z} p^y (1-p)^z}{y!z!} \\ &= P(Y = y) P(Z = z) \end{aligned}$$

**24.** $T = \sum_{i=1}^{6} iN_i$, $N_i$ = number of dice that land $i$
$ET = 42$, $Var(T) = 364$, $SD(T) = 19.08$

**25.** a) $N_1$ and $N_2$ are independent Poisson variables having parameters $\lambda p_1$ and $\lambda p_2$ respectively: Let $n_1$ and $n_2$ be nonnegative integers. Then for each $n \ge n_1 + n_2$ we have

$$P(N_1 = n_1, N_2 = n_2 | N = n)$$

$= P(n_1$ of the accidents result in 1 injury and $n_2$ of the accidents result in 2 injuries $| N = n)$

$$= \frac{n!}{n_1! n_2! (n - n_1 - n_2)!} p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n - n_1 - n_2}$$

and

$$\begin{aligned} &P(N_1 = n_1, N_2 = n_2) \\ &= \sum_{n=n_1+n_2}^{\infty} P(N = n) P(N_1 = n_1, N_2 = n_2 | N = n) \\ &= \sum_{n=n_1+n_2}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} \frac{n!}{n_1! n_2! (n - n_1 - n_2)!} p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n - n_1 - n_2} \\ &= e^{\frac{-\lambda}{n_1! n_2!}} p_1^{n_1} p_2^{n_2} \lambda^{n_1+n_2} \sum_{n=n_1+n_2} [\lambda(1 - p_1 - p_2)]^{\frac{n - n_1 - n_2}{(n - n_1 - n_2)!}} \\ &= \frac{e^{-\lambda}(\lambda p_1)^{n_1} (\lambda p_2)^{n_2}}{n_1! n_2!} e^{\lambda(1 - p_1 - p_2)} \\ &= \frac{e^{-\lambda p_1}(\lambda p_1)^{n_1}}{n_1!} \frac{e^{-\lambda p_2}(\lambda p_2)^{n_2}}{n_2!}. \end{aligned}$$

b) We have $M = \sum_{k=0}^{\infty} kN_k$ . Generalize part (a) to see that the variables $N_0, N_1, N_2, \ldots$ are independent, and $N_k$ has Poisson $(\lambda p_k)$ distribution. Therefore

$$E(M) = \sum_{k=0}^{\infty} kE(N_k) = \sum_{k=0}^{\infty} k(\lambda p_k) = \lambda \sum_{k=0}^{\infty} kp_k;$$

$$Var(M) = \sum_{k=0}^{\infty} k^2 Var(N_k) = \sum_{k=0}^{\infty} k^2(\lambda p_k) = \lambda \sum_{k=0}^{\infty} k^2 p_k;$$

$$SD(M) = \left(\lambda \sum_{k=0}^{\infty} k^2 p_k\right)^{\frac{1}{2}}.$$

26. Let $N$ denote the number of points falling in the unit square. The key idea is that given $N$ with $N \geq 1$, the $N$ points of the scatter are distributed like $N$ points picked uniformly and independently at random from the square. (See following exercise for a calculation which verifies this). In particular, given $N = 2$, the two points are independently uniformly distributed on the unit square. Draw a circle of radius $\delta$ about one of them. The chance that the other is outside is $1 - \pi\delta^2$, provided the first point is not within $\delta$ of a side of the square; if the first point *is* within $\delta$ of a side, then the chance is bigger than $1 - \pi\delta^2$. So overall, $P(D|N = 2) \geq 1 - \pi\delta^2$.

Given $N = 3$, the three points are independent and uniformly distributed on the unit square. So

$$\begin{aligned}
&P(D|N = 3)\\
&= P(\text{two points not within } \delta) \times\\
&\qquad P(\text{third point more than } \delta \text{ away from both} \mid \text{first two not within } \delta)\\
&\geq (1 - \pi\delta^2)[1 - P(\text{third point within } \delta \text{ of one of the first two} \mid \text{first two not within } \delta)]\\
&\geq (1 - \pi\delta^2)(1 - 2\pi\delta^2)\\
&\geq 1 - (\pi\delta^2 + 2\pi\delta^2) = 1 - 3\pi\delta^2.
\end{aligned}$$

In general,

$$\begin{aligned}
P(D|N = n) &\geq \left(1 - (\pi\delta^2 + 2\pi\delta^2 + \cdots + (n-2)\pi\delta^2)\right)(1 - (n-1)\pi\delta^2)\\
&\geq 1 - \pi\delta^2(1 + 2 + \cdots + n - 1) = 1 - \frac{n(n-1)}{2}\pi\delta^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
P(D) &= \sum_{n=0}^{\infty} P(D|N = n)P(N = n)\\
&\geq \sum_{n=0}^{\infty} [1 - \frac{n(n-1)}{2}\pi\delta^2]e^{-\lambda}\frac{\lambda^n}{n!}\\
&= 1 - \frac{\pi\delta^2}{2}\sum_{n=0}^{\infty} e^{-\lambda} n(n-1)\frac{\lambda^n}{n!}\\
&= 1 - \frac{\pi\delta^2}{2}\lambda^2 \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!}\\
&= 1 - \frac{\pi}{2}\delta^2\lambda^2.
\end{aligned}$$

27. By considering $B_{j+1} = S - \cup_{k=1}^{j} B_k$, we may consider, without loss of generality, a partition $B_1, \ldots, B_j$ of $S$. Since $X_i$'s are conditionally independent and identically distributed, given $N = n$, $(N(B_1), \ldots, N(B_j))$ has a multinomial distribution with parameters $P(X_i \in B_1) = Q(B_1), \ldots, P(X_i \in B_j) = Q(B_j)$. Therefore

$$P\left[N(B_1) = n_1, \ldots, N(B_j) = n_j\right]$$

$$= P[N(B_1) = n_1, \ldots, N(B_j) = n_j, N = n_1 + \cdots + n_j], \quad \text{(since } B_k\text{'s partition } S)$$
$$= P[N(B_1) = n_1, \ldots, N(B_j) = n_j \mid N = n_1 + \cdots + n_j] \times P(N = n_1 + \cdots + n_j)$$
$$= \frac{n!}{n_1! \cdots n_j!}[Q(B_1)]^{n_1} \cdots [Q(B_j)]^{n_j} \times \frac{e^{-\lambda}\lambda^{n_1+\cdots+n_j}}{(n_1 + \cdots + n_j)!}$$
$$= \frac{e^{-\lambda Q(B_1)}[\lambda Q(B_1)]^{n_1}}{n_1!} \cdots \frac{e^{-\lambda Q(B_j)}[\lambda Q(B_j)]^{n_j}}{n_j!}.$$

28. a) $P(X = Y) = P(X - Y = 0)$. Now $X$ and $Y$ are independent, each approximately normally distributed with mean $N/2$ and variance $N/4$, so $X - Y$ is approximately normally distributed with mean 0 and variance $N/2$. Therefore

$$P(X - Y = 0) \approx \frac{1}{\sqrt{2\pi(N/2)}} = \frac{1}{\sqrt{\pi N}}.$$

To have $P(X = Y) \approx 1/10$ we need $N \approx 32$.

b) $P(X = k|X = Y) = \frac{P(X=k,X=Y)}{P(X=Y)} = \frac{P(X=k,Y=k)}{P(X=Y)} = \frac{P(X=k)^2}{P(X=Y)}$.
By the normal approximation, this is roughly

$$\left(\frac{1}{\sqrt{2\pi(N/4)}}e^{-(k-(N/2))^2/2(N/4)}\right)^2 \Big/ \frac{1}{\sqrt{\pi N}} = \frac{1}{\sqrt{2\pi(N/8)}}e^{-(k-(N/2))^2/2(N/8)}.$$

That is, given $X = Y$, $X$ has approximately normal distribution with mean $N/2$ and variance $N/8$ — same as unconditionally, but with variance half as big. So the conditional probability that $|X - N/2| \le \sqrt{N}/2$ is somewhat larger than 68%.

**Another way:** The probability in question is approximately $P(|X\star| \le 1|X\star = Y\star)$ where $X\star$ and $Y\star$ are independent standard normal variables. Slice the standard bivariate normal density along the line $\{(x, y) : x = y\}$. The resulting curve is again a standard normal density, by rotational symmetry. The probability in question is approximately the area under this curve above the line segment $\{(x, y) : x = y, |x| \le 1\}$, divided by the total area under the curve (which is 1). So the probability in question is approximately the area under the standard normal density from $-\sqrt{2}$ to $\sqrt{2}$, which is greater than the area from $-1$ to 1.

**Yet another way:** Let $X\star$, $Y\star$ be independent standard normal variables.
Let $U = (X\star + Y\star)/\sqrt{2}$, $V = (X\star - Y\star)/\sqrt{2}$. Then $U$ and $V$ are again independent standard normal variables, so

$$P(X\star \in dx|X\star = Y\star) = P\left(\frac{U+V}{\sqrt{2}} \in dx|V = 0\right) = P\left(\frac{U}{\sqrt{2}} \in dx\right).$$

That is, conditional on $X\star = Y\star$, $X\star$ has normal distribution with mean 0 and variance 1/2.

29. a) Fix $k \ge 1$, and suppose $N \ge k + 1$. Let $\text{Good}_1$ denote the event that there is a good element in the first place, $\text{Bad}_1$ the event that there is a bad element in the first place. Given $\text{Good}_1$, $T_1 - 1$ equals zero; and given $\text{Bad}_1$, $T_1 - 1$ has the same distribution as $T_1^*$, where $T_1^*$ is the place of the first good element in a random arrangement of $k$ good elements among $N - 1$ elements. So

$$E[(T_1 - 1)(T_2 - 2)|\text{Bad}_1] = E[T_1^*(T_1^* - 1)] = \alpha(k, N - 1),$$

and

$$\begin{aligned}
\alpha(k, N) &= E[T_1(T_1 - 1)] \\
&= E[T_1(T_1 - 1)|\text{Bad}_1] \cdot P(\text{Bad}_1) + 0 \\
&= E[(T_1 - 1)(T_1 - 2) + 2(T_1 - 1)|\text{Bad}_1] \cdot P(\text{Bad}_1) \\
&= \{E[(T_1 - 1)(T_1 - 2)|\text{Bad}_1] + 2E[T_1 - 1|\text{Bad}_1]\} \cdot P(\text{Bad}_1) \\
&= [\alpha(k, N - 1) + 2\mu(k, N - 1)]\frac{N - k}{N} \\
&= \left[\alpha(k, N - 1) + 2\frac{N}{k + 1}\right]\frac{N - k}{N}.
\end{aligned}$$

b) **Claim:** For each $k \geq 1$, the formula holds for each $N = k, k+1, k+2, \dots$.
**Proof:** Induction on $N$. If $N = k$, then $T_1 = 1$ so $T_1(T_1 - 1) = 0$ , so

$$\alpha(k, N) = ET_1(T_1 - 1) = 0 = \frac{2(N+1)(N-k)}{(k+1)(k+2)}.$$

If claim holds for $N = m$, where $m \geq k$, then

$$\begin{aligned}\alpha(k, m+1) &= (m+1-k)\left[\frac{2}{k+1} + \frac{\alpha(k,m)}{m+1}\right] \\ &= (m+1-k)\left[\frac{2}{k+1} + \frac{2(m-k)}{(k+1)(k+2)}\right] \\ &= \frac{2(m+2)(m+1-k)}{(k+1)(k+2)}\end{aligned}$$

so claim holds for $N = m + 1$ .

c)

$$\begin{aligned}Var(T_1) &= E(T_1^2) - (ET_1)^2 \\ &= ET_1(T_1 - 1) - (ET_1)^2 + (ET_1) \\ &= \frac{2(N+1)(N-k)}{(k+1)(k+2)} - \left(\frac{N+1}{k+1}\right)^2 + \frac{N+1}{k+1} \\ &= \frac{(N+1)}{(k+1)^2(k+2)}[2(N-k)(k+1) - (N+1)(k+2) + (k+1)(k+2)] \\ &= \frac{(N+1)(N-k)k}{(k+1)^2(k+2)}.\end{aligned}$$

d) If $k = 1$ then $T_1$ has uniform distribution on $1, \dots, N$, so $Var(T_1) = \frac{N^2-1}{12}$, which agrees with the formula in (c).

e) Let $k \geq 1$ . If $1 \leq i \leq k+1$ then

$$\begin{aligned}Var(T_i) &= Var\left(\sum_{k=1}^{i} W_k\right) \\ &= \sum_{k=1}^{i} Var(W_k) + 2\sum_{1 \leq k < l \leq i} Cov(W_k, W_l) \\ &= iVar(W_1) + 2\frac{i(i-1)}{2}Cov(W_1, W_2) \\ &= iVar(T_1) + i(i-1)Cov(W_1, W_2)\end{aligned}$$

by exchangeability. Put $i = k + 1$, get

$$0 = Var(N+1) = Var(T_{k+1}) = (k+1)Var(T_1) + (k+1)kCov(W_1, W_2)$$

$$\iff \qquad Cov(W_1, W_2) = -[Var(T_1)]/k.$$

So

$$\begin{aligned}Var(T_i) &= iVar(T_1) + i(i-1)Cov(W_1, W_2) \\ &= iVar(T_1)\left[1 - \frac{i-1}{k}\right] \\ &= \frac{i(k+1-i)}{k}\frac{(N+1)(N-k)k}{(k+1)^2(k+2)}.\end{aligned}$$

f) The distribution of $T_i$ is the same as the distribution of $N + 1 - T_{k+1-i}$ since $T_1, \dots, T_k$ would have the same distribution if we had "read" the $N$ elements in the reverse order.

g) Here $k = 4$, $N = 52$, so

$$E(T_i) = i\frac{N+1}{k+1} = \frac{53}{5}i \quad \text{and} \quad Var(T_i) = i(k+1-i)\frac{(N+1)(N-k)}{(k+1)^2(k+2)} = \frac{424}{25}i(5-i).$$

| $i$ | $E(T_i)$ | $Var(T_i)$ | $SD(T_i)$ |
|---|---|---|---|
| 1 | 10.6 | 67.84 | 8.24 |
| 2 | 21.2 | 101.76 | 10.09 |
| 3 | 31.8 | 101.76 | 10.09 |
| 4 | 42.4 | 67.84 | 8.24 |

30. a) For large $n$, $A_{all}$ has approximately the normal$(\frac{1}{2}, \frac{1}{12n})$ distribution.

$$P(|A_{all} - \frac{1}{2}| < \epsilon) = P\left(\left|\frac{A_{all} - (1/2)}{\sqrt{1/12n}}\right| < \frac{\epsilon}{\sqrt{1/12n}}\right)$$
$$\approx \Phi(\sqrt{12n}\epsilon) - \Phi(-\sqrt{12n}\epsilon) \to 1 \quad \text{as } n \to \infty$$

The distribution of order statistics of independent uniforms is described in Section 4.6: $V_k$ has beta$(k, n-k+1)$ distribution.

$$\begin{aligned}
E(A_{ext}) &= \frac{1}{2}[E(V_1 + V_n)] = \frac{1}{2}\left(\frac{1}{n+1} + \frac{n}{n+1}\right) = \frac{1}{2} \\
E(A_{ext}^2) &= \frac{1}{4}E(V_1^2) + \frac{1}{2}E(V_1V_n) + \frac{1}{4}E(V_n^2) \\
&= \frac{1}{4}\frac{2}{(n+1)(n+2)} + \frac{1}{2}E(V_1V_n) + \frac{1}{4}\frac{n}{n+2} \\
E(V_1V_n) &= E[E(V_1V_n \mid V_1)] = E[V_1 E(V_n \mid V_1)] \\
&= E[V_1(V_1 + \frac{n-1}{n}(1-V_1))] = \frac{1}{n}E(V_1^2) + \frac{n-1}{n}E(V_1) \\
&= \frac{1}{n}\frac{2}{(n+1)(n+2)} + \frac{n-1}{n}\frac{1}{n+1} = \frac{1}{n+2} \\
E(A_{ext}^2) &= \frac{1}{2(n+1)(n+2)} + \frac{1}{2(n+2)} + \frac{n}{4(n+2)} = \frac{n^2+3n+4}{4(n+1)(n+2)} \\
Var(A_{ext}) &= \frac{n^2+3n+4}{4(n+1)(n+2)} - \left(\frac{1}{2}\right)^2 = \frac{1}{2(n+1)(n+2)}
\end{aligned}$$

Now use Chebychev's inequality to establish,

$$P(|A_{ext} - \frac{1}{2}| < \epsilon) \geq 1 - \frac{Var(A_{ext})}{\epsilon^2} = 1 - \frac{1}{2(n+1)(n+2)\epsilon^2} \to 1 \quad \text{as } n \to \infty$$

$$\begin{aligned}
E(A_{mid}) &= \frac{(n+1)/2}{n+1} = \frac{1}{2} \\
E(A_{mid}^2) &= \frac{((n+1)/2)((n+1)/2+1)}{(n+1)(n+2)} = \frac{n+3}{4(n+2)} \\
Var(A_{mid}) &= \frac{n+3}{4(n+2)} - \left(\frac{1}{2}\right)^2 = \frac{1}{4(n+2)}
\end{aligned}$$

Again, using Chebychev's inequality,

$$P(|A_{mid} - \frac{1}{2}| < \epsilon) \geq 1 - \frac{Var(A_{mid})}{\epsilon^2} = 1 - \frac{1}{4(n+2)\epsilon^2} \to 1 \quad \text{as } n \to \infty$$

b) $A_{ext}$ is expected to be much closer to $\frac{1}{2}$, since it's variance is much smaller.

c) The Central Limit Theorem allows you to use the normal approximation for $A_{all}$. The beta$(r, s)$ distribution is approximately normal for $r$ and $s$ large so $A_{mid}$ can also be approximated by the normal distribution..

$$\begin{aligned}
P(0.49 < A_{all} < 0.51) &\approx \Phi\left(\frac{0.01}{\sqrt{1/(12\cdot 101)}}\right) - \Phi\left(\frac{-0.01}{\sqrt{1/(12\cdot 101)}}\right) \\
&= \Phi(0.35) - \Phi(-0.35) = 0.2736 \\
P(0.49 < A_{mid} < 0.51) &\approx \Phi\left(\frac{0.01}{\sqrt{1/(4\cdot 103)}}\right) - \Phi\left(\frac{-0.01}{\sqrt{1/(4\cdot 103)}}\right) \\
&= \Phi(0.21) - \Phi(0.21) = 0.1664
\end{aligned}$$

Another method is necessary for $A_{ext}$ because its distribution is not close to normal. The pair $(V_1, V_n)$ has joint density

$$f(x,y) = n(n-1)(y-x)^{n-2} \quad (0 < x < y < 1)$$

Let $Z = V_1 + V_n$. Then for $0 \le z \le 1$,

$$\begin{aligned}
f_Z(z) &= \int_0^{\frac{z}{2}} n(n-1)(z-2x)^{n-2}\,dx \\
&= -\frac{n}{2}(z-2x)^{n-1}\Big|_{x=0}^{z/2} = \frac{n}{2}z^{n-1}
\end{aligned}$$

For $1 \le z \le 2$,

$$\begin{aligned}
f_Z(z) &= \int_{z-1}^{\frac{z}{2}} n(n-1)(z-2x)^{n-2}\,dx \\
&= -\frac{n}{2}(z-2x)^{n-1}\Big|_{x=z-1}^{z/2} = \frac{n}{2}(2-z)^{n-1}
\end{aligned}$$

Since $A_{ext} = Z/2$, it has density

$$f_{A_{ext}}(x) = \begin{cases} n2^{n-1}x^{n-1} & \text{for } 0 \le x \le \frac{1}{2} \\ n2^{n-1}(1-x)^{n-1} & \text{for } \frac{1}{2} \le x \le 1 \end{cases}$$

Now the probability can be computed directly.

$$\begin{aligned}
P(0.49 < A_{ext} < 0.51) &= 2\int_{0.49}^{0.50} (101)2^{100}x^{100}\,dx = 2^{101}x^{101}\Big|_{x=0.49}^{0.50} \\
&= 1 - (0.98)^{101} = 0.8700
\end{aligned}$$

31. No Solution

32. No Solution

33. No Solution

34. No Solution

35. No Solution