Solutions to HW2

Note: These solutions were generated by R. D. Yates and D. J. Goodman, the authors of our textbook. I have added comments in italics where I thought more detail was appropriate.

Problem 1.4.3 ■

The basic rules of genetics were discovered in mid-1800s by Mendel, who found that each characteristic of a pea plant, such as whether the seeds were green or yellow, is determined by two genes, one from each parent. Each gene is either dominant d or recessive r. Mendel's experiment is to select a plant and observe whether the genes are both dominant d, both recessive r, or one of each (hybrid) h. In his pea plants, Mendel found that yellow seeds were a dominant trait over green seeds. A yy pea with two yellow genes has yellow seeds; a gg pea with two recessive genes has green seeds; a hybrid gy or yg pea has yellow seeds. In one of Mendel's experiments, he started with a parental generation in which half the pea plants were yy and half the plants were gg. The two groups were crossbred so that each pea plant in the first generation was gy. In the second generation, each pea plant was equally likely to inherit a y or a g gene from each first generation parent. What is the probability $P[Y]$ that a randomly chosen pea plant in the second generation has yellow seeds?

Problem 1.4.3 Solution

The first generation consists of two plants each with genotype yg or gy. They are crossed to produce the following second generation genotypes, $S = \{yy, yg, gy, gg\}$. Each genotype is just as likely as any other so the probability of each genotype is consequently 1/4. A pea plant has yellow seeds if it possesses at least one dominant y gene. The set of pea plants with yellow seeds is

$$Y = \{yy, yg, gy\}. \tag{1}$$

So the probability of a pea plant with yellow seeds is

$$P[Y] = P[yy] + P[yg] + P[gy] = 3/4. \tag{2}$$

Problem 1.5.5 ■

You have a shuffled deck of three cards: 2, 3, and 4 and you deal out the three cards. Let E_i denote the event that ith card dealt is even numbered.

(a) What is $P[E_2|E_1]$, the probability the second card is even given that the first card is even?

(b) What is the conditional probability that the first two cards are even given that the third card is even?

(c) Let O_i represent the event that the ith card dealt is odd numbered. What is $P[E_2|O_1]$, the conditional probability that the second card is even given that the first card is odd?

(d) What is the conditional probability that the second card is odd given that the first card is odd?

Problem 1.5.5 Solution

The sample outcomes can be written ijk where the first card drawn is i, the second is j and the third is k. The sample space is

$$S = \{234, 243, 324, 342, 423, 432\}. \tag{1}$$

and each of the six outcomes has probability 1/6. The events $E_1, E_2, E_3, O_1, O_2, O_3$ are

$$E_1 = \{234, 243, 423, 432\}, \qquad O_1 = \{324, 342\}, \tag{2}$$
$$E_2 = \{243, 324, 342, 423\}, \qquad O_2 = \{234, 432\}, \tag{3}$$
$$E_3 = \{234, 324, 342, 432\}, \qquad O_3 = \{243, 423\}. \tag{4}$$

(a) The conditional probability the second card is even given that the first card is even is

$$P[E_2|E_1] = \frac{P[E_2E_1]}{P[E_1]} = \frac{P[243, 423]}{P[234, 243, 423, 432]} = \frac{2/6}{4/6} = 1/2. \tag{5}$$

(b) The probability the first two cards are even given the third card is even is

$$P[E_1E_2|E_3] = \frac{P[E_1E_2E_3]}{P[E_3]} = 0. \tag{6}$$

(c) The conditional probabilities the second card is even given that the first card is odd is

$$P[E_2|O_1] = \frac{P[O_1E_2]}{P[O_1]} = \frac{P[O_1]}{P[O_1]} = 1. \tag{7}$$

(d) The conditional probability the second card is odd given that the first card is odd is

$$P[O_2|O_1] = \frac{P[O_1O_2]}{P[O_1]} = 0. \tag{8}$$

Problem 1.5.6 ⧫

Deer ticks can carry both Lyme disease and human granulocytic ehrlichiosis (HGE). In a study of ticks in the Midwest, it was found that 16% carried Lyme disease, 10% had HGE, and that 10% of the ticks that had either Lyme disease or HGE carried both diseases.

(a) What is the probability $P[LH]$ that a tick carries both Lyme disease (L) and HGE (H)?

(b) What is the conditional probability that a tick has HGE given that it has Lyme disease?

Problem 1.5.6 Solution

The problem statement yields the obvious facts that $P[L] = 0.16$ and $P[H] = 0.10$. The words "10% of the ticks that had either Lyme disease or HGE carried both diseases" can be written as

$$P[LH|L \cup H] = 0.10. \tag{1}$$

(a) Since $LH \subset L \cup H$,

$$P[LH|L \cup H] = \frac{P[LH \cap (L \cup H)]}{P[L \cup H]} = \frac{P[LH]}{P[L \cup H]} = 0.10. \tag{2}$$

Thus,

$$P[LH] = 0.10 P[L \cup H] = 0.10 \left(P[L] + P[H] - P[LH]\right). \tag{3}$$

Since $P[L] = 0.16$ and $P[H] = 0.10$,

$$P[LH] = \frac{0.10\,(0.16 + 0.10)}{1.1} = 0.0236. \tag{4}$$

(b) The conditional probability that a tick has HGE given that it has Lyme disease is

$$P[H|L] = \frac{P[LH]}{P[L]} = \frac{0.0236}{0.16} = 0.1475. \tag{5}$$

Problem 1.6.3 ■

In an experiment, A, B, C, and D are events with probabilities $P[A] = 1/4$, $P[B] = 1/8$, $P[C] = 5/8$, and $P[D] = 3/8$. Furthermore, A and B are disjoint, while C and D are independent.

(a) Find $P[A \cap B]$, $P[A \cup B]$, $P[A \cap B^c]$, and $P[A \cup B^c]$.

(b) Are A and B independent?

(c) Find $P[C \cap D]$, $P[C \cap D^c]$, and $P[C^c \cap D^c]$.

(d) Are C^c and D^c independent?

Problem 1.6.3 Solution

(a) Since A and B are disjoint, $P[A \cap B] = 0$. Since $P[A \cap B] = 0$,

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] = 3/8. \tag{1}$$

A Venn diagram should convince you that $A \subset B^c$ so that $A \cap B^c = A$. This implies

$$P[A \cap B^c] = P[A] = 1/4. \tag{2}$$

It also follows that $P[A \cup B^c] = P[B^c] = 1 - 1/8 = 7/8$.

(b) Events A and B are dependent since $P[AB] \neq P[A]P[B]$.

(c) Since C and D are independent,

$$P[C \cap D] = P[C]\,P[D] = 15/64. \tag{3}$$

The next few items are a little trickier. From Venn diagrams, we see

$$P[C \cap D^c] = P[C] - P[C \cap D] = 5/8 - 15/64 = 25/64. \tag{4}$$

It follows that

$$\begin{aligned} P[C \cup D^c] &= P[C] + P[D^c] - P[C \cap D^c] & (5)\\ &= 5/8 + (1 - 3/8) - 25/64 = 55/64. & (6)\end{aligned}$$

Using DeMorgan's law, we have

$$P[C^c \cap D^c] = \left(\frac{3}{8}\right)\left(\frac{5}{8}\right) = 15/64 = 1 - P[C \cup D] = P[(C \cup D)^c]. \tag{7}$$

(d) Since $P[C^cD^c] = P[C^c]P[D^c]$, C^c and D^c are independent.

Problem 1.6.4 ■

In an experiment, A, B, C, and D are events with probabilities $P[A \cup B] = 5/8$, $P[A] = 3/8$, $P[C \cap D] = 1/3$, and $P[C] = 1/2$. Furthermore, A and B are disjoint, while C and D are independent.

(a) Find $P[A \cap B]$, $P[B]$, $P[A \cap B^c]$, and $P[A \cup B^c]$.

(b) Are A and B independent?

(c) Find $P[D]$, $P[C \cap D^c]$, $P[C^c \cap D^c]$, and $P[C|D]$.

(d) Find $P[C \cup D]$ and $P[C \cup D^c]$.

(e) Are C and D^c independent?

Problem 1.6.4 Solution

(a) Since $A \cap B = \emptyset$, $P[A \cap B] = 0$. To find $P[B]$, we can write

$$\begin{aligned} P[A \cup B] &= P[A] + P[B] - P[A \cap B] & (1)\\ 5/8 &= 3/8 + P[B] - 0. & (2)\end{aligned}$$

Thus, $P[B] = 1/4$. Since A is a subset of B^c, $P[A \cap B^c] = P[A] = 3/8$. Furthermore, since A is a subset of B^c, $P[A \cup B^c] = P[B^c] = 3/4$.

(b) The events A and B are dependent because

$$P[AB] = 0 \neq 3/32 = P[A]P[B]. \quad (3)$$

(c) Since C and D are independent $P[CD] = P[C]P[D]$. So

$$P[D] = \frac{P[CD]}{P[C]} = \frac{1/3}{1/2} = 2/3. \quad (4)$$

In addition, $P[C \cap D^c] = P[C] - P[C \cap D] = 1/2 - 1/3 = 1/6$. To find $P[C^c \cap D^c]$, we first observe that

$$P[C \cup D] = P[C] + P[D] - P[C \cap D] = 1/2 + 2/3 - 1/3 = 5/6. \quad (5)$$

By De Morgan's Law, $C^c \cap D^c = (C \cup D)^c$. This implies

$$P[C^c \cap D^c] = P[(C \cup D)^c] = 1 - P[C \cup D] = 1/6. \quad (6)$$

Note that a second way to find $P[C^c \cap D^c]$ is to use the fact that if C and D are independent, then C^c and D^c are independent. Thus

$$P[C^c \cap D^c] = P[C^c]P[D^c] = (1 - P[C])(1 - P[D]) = 1/6. \quad (7)$$

Finally, since C and D are independent events, $P[C|D] = P[C] = 1/2$.

(d) Note that we found $P[C \cup D] = 5/6$. We can also use the earlier results to show

$$P[C \cup D^c] = P[C] + P[D] - P[C \cap D^c] = 1/2 + (1 - 2/3) - 1/6 = 2/3. \quad (8)$$

(e) By Definition 1.7, events C and D^c are independent because

$$P[C \cap D^c] = 1/6 = (1/2)(1/3) = P[C]P[D^c]. \quad (9)$$

Problem 1.6.5 ■

In an experiment with equiprobable outcomes, the event space *(I would have said "the sample space")* is $S = \{1, 2, 3, 4\}$ and $P[s] = 1/4$ for all $s \in S$. Find three events in S that are pairwise independent but are not independent. (Note: Pairwise independent events meet the first three conditions of Definition 1.8).

Problem 1.6.5 Solution

For a sample space $S = \{1, 2, 3, 4\}$ with equiprobable outcomes, consider the events

$$A_1 = \{1, 2\} \qquad A_2 = \{2, 3\} \qquad A_3 = \{3, 1\}. \quad (1)$$

Each event A_i has probability 1/2. Moreover, each pair of events is independent since

$$P[A_1A_2] = P[A_2A_3] = P[A_3A_1] = 1/4. \quad (2)$$

However, the three events A_1, A_2, A_3 are not independent since

$$P[A_1A_2A_3] = 0 \neq P[A_1]P[A_2]P[A_3]. \quad (3)$$

Problem 1.6.6 ■

(Continuation of Problem 1.4.3) One of Mendel's most significant results was the conclusion that genes determining different characteristics are transmitted independently. In pea plants, Mendel found that round peas are a dominant trait over wrinkled peas. Mendel crossbred a group of (rr, yy) peas with a group of (ww, gg) peas. In this notation, rr denotes a pea with two "round" genes and ww denotes a pea with two "wrinkled" genes. The first generation were either (rw, yg), (rw, gy), (wr, yg), or (wr, gy) plants with both hybrid shape and hybrid color. Breeding among the first generation yielded second-generation plants in which genes for each characteristic were equally likely to be either dominant or recessive. What is the probability $P[Y]$ that a second-generation pea plant has yellow seeds? What is the probability $P[R]$ that a second-generation plant has round peas? Are R and Y independent events? How many visibly different kinds of pea plants would Mendel observe in the second generation? What are the probabilities of each of these kinds?

Problem 1.6.6 Solution

There are 16 distinct equally likely outcomes for the second generation of pea plants based on a first generation of $\{rwyg, rwgy, wryg, wrgy\}$. They are listed below

$$\begin{array}{cccc} rryy & rryg & rrgy & rrgg \\ rwyy & rwyg & rwgy & rwgg \\ wryy & wryg & wrgy & wrgg \\ wwyy & wwyg & wwgy & wwgg \end{array} \tag{1}$$

A plant has yellow seeds, that is event Y occurs, if a plant has at least one dominant y gene. Except for the four outcomes with a pair of recessive g genes, the remaining 12 outcomes have yellow seeds. From the above, we see that

$$P[Y] = 12/16 = 3/4 \tag{2}$$

and

$$P[R] = 12/16 = 3/4. \tag{3}$$

To find the conditional probabilities $P[R|Y]$ and $P[Y|R]$, we first must find $P[RY]$. Note that RY, the event that a plant has rounded yellow seeds, is the set of outcomes

$$RY = \{rryy, rryg, rrgy, rwyy, rwyg, rwgy, wryy, wryg, wrgy\}. \tag{4}$$

Since $P[RY] = 9/16$,

$$P[Y|R] = \frac{P[RY]}{P[R]} = \frac{9/16}{3/4} = 3/4 \tag{5}$$

and

$$P[R|Y] = \frac{P[RY]}{P[Y]} = \frac{9/16}{3/4} = 3/4. \tag{6}$$

Thus $P[R|Y] = P[R]$ and $P[Y|R] = P[Y]$ and R and Y are independent events. There are four visibly different pea plants, corresponding to whether the peas are round (R) or not (R^c), or yellow (Y) or not (Y^c). These four visible events have probabilities

$$P[RY] = 9/16 \qquad P[RY^c] = 3/16, \tag{7}$$

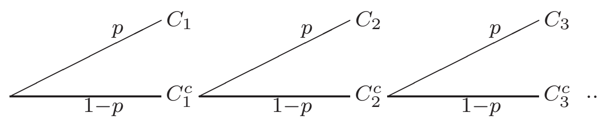
$$P[R^cY] = 3/16 \qquad P[R^cY^c] = 1/16. \tag{8}$$

Problem 1.7.10 ■

Each time a fisherman casts his line, a fish is caught with probability p, independent of whether a fish is caught on any other cast of the line. The fisherman will fish all day until a fish is caught and then he will quit and go home. Let C_i denote the event that on cast i the fisherman catches a fish. Draw the tree for this experiment and find $P[C_1]$, $P[C_2]$, and $P[C_n]$.

Problem 1.7.10 Solution

The experiment ends as soon as a fish is caught. The tree resembles



From the tree, $P[C_1] = p$ and $P[C_2] = (1-p)p$. Finally, a fish is caught on the nth cast if no fish were caught on the previous $n-1$ casts. Thus,

$$P[C_n] = (1-p)^{n-1}p. \tag{1}$$

Problem 1.9.1 •

Consider a binary code with 5 bits (0 or 1) in each code word. An example of a code word is 01010. In each code word, a bit is a zero with probability 0.8, independent of any other bit.

(a) What is the probability of the code word 00111?

(b) What is the probability that a code word contains exactly three ones?

Problem 1.9.1 Solution

(a) Since the probability of a zero is 0.8, we can express the probability of the code word 00111 as 2 occurrences of a 0 and three occurrences of a 1. Therefore

$$P[00111] = (0.8)^2(0.2)^3 = 0.00512. \tag{1}$$

(b) The probability that a code word has exactly three 1's is

$$P[\text{three 1's}] = \binom{5}{3}(0.8)^2(0.2)^3 = 0.0512. \tag{2}$$

Problem 1.9.3 •

Suppose each day that you drive to work a traffic light that you encounter is either green with probability 7/16, red with probability 7/16, or yellow with probability 1/8, independent of the status of the light on any other day. If over the course of five days, G, Y, and R denote the number of times the light is found to be green, yellow, or red, respectively, what is the probability that $P[G = 2, Y = 1, R = 2]$? Also, what is the probability $P[G = R]$?

Problem 1.9.3 Solution

We know that the probability of a green and red light is 7/16, and that of a yellow light is 1/8. Since there are always 5 lights, G, Y, and R obey the multinomial probability law:

$$P[G=2, Y=1, R=2] = \frac{5!}{2!1!2!}\left(\frac{7}{16}\right)^2\left(\frac{1}{8}\right)\left(\frac{7}{16}\right)^2. \tag{1}$$

The probability that the number of green lights equals the number of red lights

$$P[G=R] = P[G=1, R=1, Y=3] + P[G=2, R=2, Y=1] + P[G=0, R=0, Y=5] \tag{2}$$

$$= \frac{5!}{1!1!3!}\left(\frac{7}{16}\right)\left(\frac{7}{16}\right)\left(\frac{1}{8}\right)^3 + \frac{5!}{2!1!2!}\left(\frac{7}{16}\right)^2\left(\frac{7}{16}\right)^2\left(\frac{1}{8}\right) + \frac{5!}{0!0!5!}\left(\frac{1}{8}\right)^5 \tag{3}$$

$$\approx 0.1449. \tag{4}$$

Problem 1.10.2 ■

We wish to modify the cellular telephone coding system in Example 1.41 in order to reduce the number of errors. In particular, if there are two or three zeroes in the received sequence of 5 bits, we will say that a deletion (event D) occurs. Otherwise, if at least 4 zeroes are received, then the receiver decides a zero was sent. Similarly, if at least 4 ones are received, then the receiver decides a one was sent. We say that an error occurs if either a one was sent and the receiver decides zero was sent or if a zero was sent and the receiver decides a one was sent. For this modified protocol, what is the probability $P[E]$ of an error? What is the probability $P[D]$ of a deletion?

Problem 1.10.2 Solution

Suppose that the transmitted bit was a 1. We can view each repeated transmission as an independent trial. We call each repeated bit the receiver decodes as 1 a success. Using $S_{k,5}$ to denote the event of k successes in the five trials, then the probability k 1's are decoded at the receiver is

$$P[S_{k,5}] = \binom{5}{k} p^k (1-p)^{5-k}, \quad k = 0, 1, \ldots, 5. \tag{1}$$

The probability that *[the transmitted]* bit is decoded correctly is

$$P[C] = P[S_{5,5}] + P[S_{4,5}] = p^5 + 5p^4(1-p) = 0.91854. \tag{2}$$

The probability a deletion occurs is

$$P[D] = P[S_{3,5}] + P[S_{2,5}] = 10p^3(1-p)^2 + 10p^2(1-p)^3 = 0.081. \tag{3}$$

The probability of an error is

$$P[E] = P[S_{1,5}] + P[S_{0,5}] = 5p(1-p)^4 + (1-p)^5 = 0.00046. \tag{4}$$

Note that if a 0 is transmitted, then 0 is sent five times and we call decoding a 0 a success. You should convince yourself that this a symmetric situation with the same deletion and error probabilities. Introducing deletions reduces the probability of an error by roughly a factor of 20. However, the probability of successful decoding is also reduced.

Problem 1.10.3 ■

Suppose a 10-digit phone number is transmitted by a cellular phone using four binary symbols for each digit, using the model of binary symbol errors and deletions given in Problem 1.10.2. If C denotes the number of bits sent correctly, D the number of deletions, and E the number of errors, what is $P[C = c, D = d, E = e]$? Your answer should be correct for any choice of c, d, and e.

Problem 1.10.3 Solution

Note that each digit 0 through 9 is mapped to the 4 bit binary representation of the digit. That is, 0 corresponds to 0000, 1 to 0001, up to 9 which corresponds to 1001. Of course, the 4 bit binary numbers corresponding to numbers 10 through 15 go unused, however this is unimportant to our problem. the 10 digit number results in the transmission of 40 bits. For each bit, an independent trial determines whether the bit was correct, a deletion, or an error. In Problem 1.10.2, we found the probabilities of these events to be

$$P[C] = \gamma = 0.91854, \qquad P[D] = \delta = 0.081, \qquad P[E] = \epsilon = 0.00046. \tag{1}$$

Since each of the 40 bit transmissions is an independent trial, the joint probability of c correct bits, d deletions, and e erasures has the multinomial probability

$$P[C = c, D = d, E = d] = \begin{cases} \frac{40!}{c!d!e!}\gamma^c\delta^d\epsilon^e & c+d+e = 40; c,d,e \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Problem 1.11.2 ■

Build a Matlab simulation of 50 trials of the experiment of Example 1.27. Your output should be a pair of 50×1 vectors $\mathbf{C}$ and $\mathbf{H}$. For the ith trial, H_i will record whether it was heads ($H_i = 1$) or tails ($H_i = 0$), and $C_i \in \{1, 2\}$ will record which coin was picked.

Problem 1.11.2 Solution

Rather than just solve the problem for 50 trials, we can write a function that generates vectors $\mathbf{C}$ and $\mathbf{H}$ for an arbitrary number of trials n. The code for this task is

```
function [C,H]=twocoin(n);
C=ceil(2*rand(n,1));
P=1-(C/4);
H=(rand(n,1)< P);
```

The first line produces the $n \times 1$ vector `C` such that `C(i)` indicates whether coin 1 or coin 2 is chosen for trial i. Next, we generate the vector `P` such that `P(i)=0.75` if `C(i)=1`; otherwise, if `C(i)=2`, then `P(i)=0.5`. As a result, `H(i)` is the simulated result of a coin flip with heads, corresponding to `H(i)=1`, occurring with probability `P(i)`.

Problem 1.11.3 ■

Following Quiz 1.9, suppose the communication link has different error probabilities for transmitting 0 and 1. When a 1 is sent, it is received as a 0 with probability 0.01. When a 0 is sent, it is received as a 1 with probability 0.03. Each bit in a packet is still equally likely to be a 0 or 1. Packets have been coded such that if five or fewer bits are received in error, then the packet can be decoded. Simulate the transmission of 100 packets, each containing 100 bits. Count the number of packets decoded correctly.

Problem 1.11.3 Solution

Rather than just solve the problem for 100 trials, we can write a function that generates n packets for an arbitrary number of trials n. The code for this task is

```
function C=bit100(n);
% n is the number of 100 bit packets sent
B=floor(2*rand(n,100));
P=0.03-0.02*B;
E=(rand(n,100)< P);
C=sum((sum(E,2)<=5));
```

First, `B` is an $n \times 100$ matrix such that `B(i,j)` indicates whether bit `i` of packet `j` is zero or one. Next, we generate the $n \times 100$ matrix `P` such that `P(i,j)=0.03` if `B(i,j)=0`; otherwise, if `B(i,j)=1`, then `P(i,j)=0.01`. As a result, `E(i,j)` is the simulated error indicator for bit `i` of packet `j`. That is, `E(i,j)=1` if bit `i` of packet `j` is in error; otherwise `E(i,j)=0`. Next we sum across the rows of `E` to obtain the number of errors in each packet. Finally, we count the number of packets with 5 or more errors.

For $n = 100$ packets, the packet success probability is inconclusive. Experimentation will show that `C=97`, `C=98`, `C=99` and `C=100` correct packets are typical values that might be observed. By increasing n, more consistent results are obtained. For example, repeated trials with $n = 100,000$ packets typically produces around $C = 98,400$ correct packets. Thus 0.984 is a reasonable estimate for the probability of a packet being transmitted correctly.