Solution to Exercises
Lesson 2 - Conditional probability and independence

Exercise 1. Six balls are to be randomly chosen from an urn containing 8 red, 10 green, and 12 blue balls.

(a) What is the probability at least one red ball is chosen?

(b) Given that no red balls are chosen, what is the conditional probability that there are exactly 2 green balls among the 6 chosen?

Answer: In this experiment, the sample space is

$$\Omega = \{\text{all the combinations of 6 balls selected from 30}\},$$

thus, $|\Omega| = \binom{30}{6}$.

(a) $1 - \mathbb{P}(\text{no red balls}) = 1 - \dfrac{\binom{22}{6}}{\binom{30}{6}}$.

(b) Given that no red balls are chosen, the 6 chosen are equally likely to be any of the 22 nonred balls. Thus,

$$\mathbb{P}(\text{2 green}|\text{no red}) = \frac{\binom{10}{2}\binom{12}{4}/\binom{30}{6}}{\binom{22}{6}/\binom{30}{6}} = \frac{\binom{10}{2}\binom{12}{4}}{\binom{22}{6}}.$$

Exercise 2. Three sorts of milk (A_1, A_2, A_3) are sterilized by either ash pasteurization (B_1) or cold pasteurization (B_2). A total of 2000 bottles of milk have been pasteurized by using these two procedures split according to the following table:

	A_1	A_2	A_3	total
B_1	260	350	200	810
B_2	440	450	300	1190
total	700	800	500	2000

Determine

(a) $\mathbb{P}(A_2)$ and $\mathbb{P}(B_1)$;

(b) $\mathbb{P}(A_2 \cap B_1)$ and $\mathbb{P}(A_2 \cup B_1)$;

(c) $\mathbb{P}(A_3|B_2)$, $\mathbb{P}(A_1|B_1^c)$, $\mathbb{P}(A_1|B_1 \cup B_2)$, $\mathbb{P}(A_1^c|B_2)$, $\mathbb{P}(B_1^c|A_3^c)$ and $\mathbb{P}(A_2^c|B_1 \cap B_2)$.

Answer: Note that B_1 and B_2 are mutually exclusive and their union gives the certain event (B_1 and B_2 constitute a partition of Ω). Also, A_1, A_2 and A_3 are pairwise mutually exclusive and their union is the certain event (A_1, A_2, A_3 constitute another partition of Ω). By using the data in the table, the basic theorems of probability and the definition of conditional probability we obtain:

(a) $\mathbb{P}(A_2) = \dfrac{800}{2000} = 0.4$ and $\mathbb{P}(B_1) = \dfrac{810}{2000} = 0.405$;

(b) $\mathbb{P}(A_2 \cap B_1) = \dfrac{350}{2000} = 0.175$ and
$\mathbb{P}(A_2 \cup B_1) = \mathbb{P}(A_2) + \mathbb{P}(B_1) - \mathbb{P}(A_2 \cap B_1) = 0.63$;

(c) $\mathbb{P}(A_3|B_2) = \dfrac{\mathbb{P}(A_3 \cap B_2)}{\mathbb{P}(B_2)} = \dfrac{300}{1190} = 0.252$,
$\mathbb{P}(A_1|B_1^c) = \mathbb{P}(A_1|B_2) = \dfrac{\mathbb{P}(A_1 \cap B_2)}{\mathbb{P}(B_2)} = \dfrac{440}{1190} = 0.37$,
$\mathbb{P}(A_1|B_1 \cup B_2) = \mathbb{P}(A_1) = \dfrac{700}{2000} = 0.35$, since $B_1 \cup B_2 = \Omega$;

$$\begin{aligned}\mathbb{P}(A_1^c|B_2) = \mathbb{P}(A_2 \cup A_3|B_2) &= \frac{\mathbb{P}((A_2 \cup A_3) \cap B_2)}{\mathbb{P}(B_2)} = \frac{\mathbb{P}[(A_2 \cap B_2) \cup (A_3 \cap B_2)]}{\mathbb{P}(B_2)} \\ &= \frac{\mathbb{P}(A_2 \cap B_2) + \mathbb{P}(A_3 \cap B_2) - \mathbb{P}(A_2 \cap A_3 \cap B_2)}{\mathbb{P}(B_2)} = \frac{\frac{450}{2000} + \frac{300}{2000} - 0}{\frac{1190}{2000}} = 0.63;\end{aligned}$$

$$\begin{aligned}\mathbb{P}(B_1^c|A_3^c) = \mathbb{P}(B_1^c|A_1 \cup A_2) &= \frac{\mathbb{P}(B_2 \cap (A_1 \cup A_2))}{\mathbb{P}(A_1 \cup A_2)} = \frac{\mathbb{P}((B_2 \cap A_1) \cup (B_2 \cap A_2))}{\mathbb{P}(A_1 \cup A_2)} \\ &= \frac{\mathbb{P}(B_2 \cap A_1) + \mathbb{P}(B_2 \cap A_2) - \mathbb{P}(B_2 \cap A_1 \cap A_2)}{\mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)} = \frac{\frac{440}{2000} + \frac{450}{2000} - 0}{\frac{700}{2000} + \frac{800}{2000} - 0} = 0.593;\end{aligned}$$

$\mathbb{P}(A_2^c|B_1 \cap B_2)$ cannot be defined since $B_1 \cap B_2 = \emptyset$ and thus $\mathbb{P}(B_1 \cap B_2) = 0$.

Exercise 3. Let Ω be a sample space, A, $B \subset \Omega$ be two events and $\mathbb{P}$ be a probability defined on Ω. Given $\mathbb{P}(A) = 1/4$, $\mathbb{P}(A|B) = 1/4$ and $\mathbb{P}(B|A) = 1/2$, compute $\mathbb{P}(A \cup B)$ and $\mathbb{P}(A^c|B^c)$.

Answer: We are given $\mathbb{P}(A) = 1/4$, $\mathbb{P}(A|B) = 1/4$ and $\mathbb{P}(B|A) = 1/2$. Hence we have

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1}{4}$$

and

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{1}{2}.$$

By elimination of $\mathbb{P}(A \cap B)$ and $\mathbb{P}(B \cap A)$ we obtain

$$\mathbb{P}(B) = 2\mathbb{P}(A) = \frac{1}{2}.$$

Using any of the equalities above once again yields

$$\mathbb{P}(A \cap B) = \frac{1}{8}.$$

Thus by the rule of addition,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \frac{1}{4} + \frac{1}{2} - \frac{1}{8} = \frac{5}{8}$$

and

$$\mathbb{P}(A^c|B^c) = \frac{\mathbb{P}(A^c \cap B^c)}{\mathbb{P}(B^c)} = \frac{\mathbb{P}[(A \cup B)^c]}{1 - \mathbb{P}(B)} = \frac{1 - \mathbb{P}(A \cup B)}{1 - \mathbb{P}(B)} = \frac{3/8}{1/2} = \frac{3}{4}.$$

Exercise 4. 52% of the students at a certain college are females. Five percent of the students in this college are majoring in computer science. Two percent of the students are women majoring in computer science. If a student is selected at random, find the conditional probability that

(a) this student is female, given that the student is majoring in computer science;

(b) this student is majoring in computer science, given that the student is female.

Answer: Let F denote the event that a student is a female, C that a student majoring in computer science. Then $\mathbb{P}(F) = 0.52$, $\mathbb{P}(C) = 0.05$ and $\mathbb{P}(F \cap C) = 0.02$.

(a) $\mathbb{P}(F|C) = \dfrac{\mathbb{P}(F \cap C)}{\mathbb{P}(C)} = \dfrac{0.02}{0.05} = 0.4;$

(b) $\mathbb{P}(C|F) = \dfrac{\mathbb{P}(F \cap C)}{\mathbb{P}(F)} = \dfrac{0.02}{0.52} = \dfrac{1}{26}.$

Exercise 5. A total of 500 married working couples were polled about their annual salaries, with the following information resulting.

	Husband with less than 25,000	Husband with more than 25,000
Wife with less than 25,000	212	198
Wife with more than 25,000	36	54

Thus, for instance, in 36 of the couples the wife earned more and the husband earned less than 25,000. If one of the couples is randomly chosen, what is

(a) the probability that the husband earn less than 25,000;

(b) the conditional probability that the wife earns more than 25,000 given the husband earns more than this amount;

(c) the conditional probability that the wife earns more than 25,000 given that the husband earns less than this amount?

Answer: The desired probabilities are calculated as follows:

(a)

$$\mathbb{P}(\text{the husband earns } < 25,000) = \frac{|\{\text{husband earns } < 25,000\}|}{|\{\text{all the husbands}\}|} = \frac{212+36}{500} = 0.496;$$

(b)

$$\begin{aligned}&\mathbb{P}(\text{the wife earns } > 25,000|\text{the husband earns } > 25,000)\\ =&\frac{|\{\text{the couples with both the husband and wife earning } > 25,000\}|}{|\{\text{husbands earn } > 25,000\}|}\\ =&\frac{54}{198+54} = \frac{3}{14}.\end{aligned}$$

(c)

$$\begin{aligned}&\mathbb{P}(\text{the wife earns } > 25,000|\text{the husband earns } < 25,000)\\ =&\frac{|\{\text{the couples with the husband earning } < 25,000 \text{ and wife earning } > 25,000\}|}{|\{\text{husbands earn } < 25,000\}|}\\ =&\frac{36}{212+36} = \frac{9}{62}.\end{aligned}$$

Exercise 6. Celine is undecided as to whether to take a French course or a chemistry course. She estimates that her probability of receiving an A grade would be 1/2 in a French course and 2/3 in a chemistry course. If Celine decides to base her decision on the flip of a fair coin, what is the probability that she gets an A in chemistry?

Answer: Let C denote the event that Celine takes chemistry and A denote the event that she receives an A in whatever course she takes, then the desired probability is $P(C \cap A)$, which is calculated by using the multiplication principle,

$$\mathbb{P}(C \cap A) = \mathbb{P}(C)\mathbb{P}(A|C) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}.$$

Exercise 7. Suppose that an urn contains 8 red balls and 4 white balls. We draw 2 balls from the urn without replacement.

(a) If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red?

(b) Now suppose that the balls have different weights, with each red ball having weight r and each white ball having weight w. Suppose that the probability that a given ball in the urn is the next one selected is its weight divided by the sum of the weights of all balls currently in the urn. Now what is the probability that both balls are red?

Answer:

(a) Let R_1 and R_2 denote, respectively, the events that the first and second balls drawn are red. Now, given that the first ball selected is red, there are 7 remaining red balls and 4 white balls, so $\mathbb{P}(R_2|R_1) = \frac{7}{11}$. As $\mathbb{P}(R_1)$ is clearly $\frac{8}{12}$, by the multiplication principle, the desired probability is

$$\mathbb{P}(R_1R_2) = \mathbb{P}(R_1)\mathbb{P}(R_2|R_1) = \frac{2}{3} \cdot \frac{7}{11} = \frac{14}{33}.$$

(b) Again let R_i be the event that the ith ball chosen is red and use

$$\mathbb{P}(R_1R_2) = \mathbb{P}(R_1)\mathbb{P}(R_2|R_1).$$

Now, number the red balls, and let B_i, $i = 1, \dots, 8$ be the event that the first ball drawn is red ball number i. Then

$$\mathbb{P}(R_1) = \mathbb{P}(\cup_{i=1}^{8} B_i) = \sum_{i=1}^{8} \mathbb{P}(B_i) = 8\frac{r}{8r + 4w}.$$

Moreover, given that the first ball is red, the urn then contains 7 red and 4 white balls. Thus, by an argument similar to the preceding one,

$$\mathbb{P}(R_2|R_1) = \frac{7r}{7r + 4w}.$$

Hence, the probability that both balls are red is

$$\mathbb{P}(R_1R_2) = \frac{8r}{8r + 4w} \cdot \frac{7r}{7r + 4w} = \frac{14r}{(2r + w)(7r + 4w)}.$$

Exercise 8. At a company a central computer is used in half of the working time in accountancy related to income, in 2/5 of the time in accountancy related to expenditure, and in the remaining time in accountancy related to taxes. The computer is idle in 15% of the time allocated to processing income, 9% of the time allocated to processing expenditure, and 18% of the time allocated to processing taxes.

(a) What is the probability that at a randomly chosen moment the computer is idle?

(b) What is the probability that the time is allocated to processing taxes given that the computer is idle?

(c) What is the probability that given that the computer is not idle, it is processing either income or expenditure?

Answer: Denote by B_1 the time allocated to processing income, B_2 to expenditure, B_3 to taxes, and I the event that the computer is idle. We have known from the question that $P(B_1) = 0.5$, $P(B_2) = 0.4$, $P(B_3) = 0.1$, $P(I|B_1) = 0.15$, $P(I|B_2) = 0.09$ and $P(I|B_3) = 0.18$.

(a) By the law of total probability, we have

$$\begin{aligned}\mathbb{P}(I) =& \mathbb{P}(I|B_1)\mathbb{P}(B_1) + \mathbb{P}(I|B_2)\mathbb{P}(B_2) + \mathbb{P}(I|B_3)\mathbb{P}(B_3)\\ =& 0.15 \cdot 0.5 + 0.09 \cdot 0.4 + 0.18 \cdot 0.1 = 0.129.\end{aligned}$$

(b) By using the definition of conditional probability

$$\mathbb{P}(B_3|I) = \frac{\mathbb{P}(B_3 \cap I)}{\mathbb{P}(I)} = \frac{\mathbb{P}(I|B_3)\mathbb{P}(B_3)}{\mathbb{P}(I)} = \frac{0.18 \cdot 0.1}{0.129} = 0.14.$$

(c) First note that $\mathbb{P}(B_1 \cup B_2|I^c) = \mathbb{P}(B_1|I^c) + \mathbb{P}(B_2|I^c)$ since $B_1 \cap B_2 = \emptyset$ (remember that $P(\cdot|I^c)$ is a probability for fixed condition I^c). On the other hand, by using rule of total probability(or written in short as Tot-Pro.) and the multiplication principle,

$$\begin{aligned}\mathbb{P}(B_1|I^c) =& \frac{\mathbb{P}(B_1 \cap I^c)}{\mathbb{P}(I^c)} \overset{\text{Tot-Pro}}{=} \frac{\mathbb{P}(B_1) - \mathbb{P}(I \cap B_1)}{1 - \mathbb{P}(I)} = \frac{\mathbb{P}(B_1) - \mathbb{P}(I|B_1)\mathbb{P}(B_1)}{1 - \mathbb{P}(I)}\\ =& \frac{0.5 - 0.15 \cdot 0.5}{1 - 0.129} = 0.488\end{aligned}$$

and similarly,

$$\begin{aligned}\mathbb{P}(B_2|I^c) =& \frac{\mathbb{P}(B_2 \cap I^c)}{\mathbb{P}(I^c)} \overset{\text{Tot-Pro}}{=} \frac{\mathbb{P}(B_2) - \mathbb{P}(I \cap B_2)}{1 - \mathbb{P}(I)} = \frac{\mathbb{P}(B_2) - \mathbb{P}(I|B_2)\mathbb{P}(B_2)}{1 - \mathbb{P}(I)}\\ =& \frac{0.4 - 0.09 \cdot 0.4}{1 - 0.129} = 0.418,\end{aligned}$$

These lead to $\mathbb{P}(B_1 \cup B_2|I^c) = 0.488 + 0.418 = 0.906$.

Exercise 9. There are two local factories that produce radios. Each radio produced at factory A is defective with probability 0.05, whereas each one produced at factory B is defective with probability 0.01. Suppose you purchase two radio that were produced at the same factory, which is equally likely to have been either factory A or factory B. If the first radio that you check is defective, what it the conditional probability that the other is also defective?

Answer: Let D_1 denote the event that the first radio is defective, D_2 the event that a second radio is defective, A the event that both of the radios are from factory A and B the event that both of the radios are from factory B. From the question we know that $\mathbb{P}(A) = \mathbb{P}(B) = 0.5$. By using the law of probability, we have

$$\mathbb{P}(D_1) = \mathbb{P}(D_1|A)\mathbb{P}(A) + \mathbb{P}(D_1|B)\mathbb{P}(B) = 0.05 \cdot 0.5 + 0.01 \cdot 0.5 = 0.03,$$

and

$$\mathbb{P}(D_1 \cap D_2) = \mathbb{P}(D_1 \cap D_2|A)\mathbb{P}(A) + \mathbb{P}(D_1 \cap D_2|B)\mathbb{P}(B) = 0.05^2 \cdot 0.5 + 0.01^2 \cdot 0.5 = 0.0013.$$

Thus the desired probability is

$$\mathbb{P}(D_2|D_1) = \frac{\mathbb{P}(D_1 \cap D_2)}{\mathbb{P}(D_1)} = \frac{0.0013}{0.03} = \frac{13}{300}.$$

Exercise 10. Suppose that an insurance company classifies people into one of three classes - good risks, average risks, and bad risks. Their records indicate that the probabilities that good, average, and bad risk persons will be involved in an accident over a 1-year span are, respectively, 0.05, 0.15 and 0.3. If 20 percent of the population are 'good risks', 50 percent are 'average risks', and 30 percent are 'bad risks', what proportion of people have accidents in a fixed year? If policy holder A had no accidents in 1987, what is the probability that he or she is a good (average) risk?

Answer: Denote by G, A and B the events that a person is classified as good, average or bad risks, and E denote that a person will be involved in an accident over a 1-year span. Then

$$\mathbb{P}(E|G) = 0.05,\ \mathbb{P}(E|A) = 0.15,\ \mathbb{P}(E|B) = 0.3,$$

and

$$\mathbb{P}(G) = 0.2,\ \mathbb{P}(A) = 0.5,\ \mathbb{P}(B) = 0.3.$$

Hence the proportion of people having accidents in a fixed year is

$$\mathbb{P}(E) = \mathbb{P}(E|G)\mathbb{P}(G)+\mathbb{P}(E|A)\mathbb{P}(A)+\mathbb{P}(E|B)\mathbb{P}(B) = 0.05\cdot 0.2+0.15\cdot 0.5+0.3\cdot 0.3 = 0.175.$$

The desired conditional probabilities are

$$\mathbb{P}(G|E^c) = \frac{\mathbb{P}(E^c|G)\mathbb{P}(G)}{\mathbb{P}(E^c)} = \frac{(1-0.05)\cdot 0.2}{1-0.175} = \frac{38}{165}$$

and

$$\mathbb{P}(A|E^c) = \frac{\mathbb{P}(E^c|A)\mathbb{P}(A)}{\mathbb{P}(E^c)} = \frac{(1-0.15)\cdot 0.5}{1-0.175} = \frac{17}{33}.$$

Exercise 11. You ask your neighbour to water a sickly plant while you are on vacation. Without water it will die with probability 0.8; with water it will die with probability 0.15. You are 90 percent certain that your neighbour will remember to water the plant.

(a) What is the probability that the plant will be alive when you return?

(b) If it is dead, what is the probability your neighbour forgot to water it?

Answer: There are two events in this exercise, denote the event that the plant will die by D, and the event that your neighbour remember to water the plant by W, then $\mathbb{P}(D|W) = 0.15$, $\mathbb{P}(D|W^c) = 0.8$, and $\mathbb{P}(W) = 0.9$.

(a) By using the rule of total probability and the multiplication principle, we have

$$\begin{aligned}\mathbb{P}(D^c) &= \mathbb{P}(D^c \cap W) + \mathbb{P}(D^c \cap W^c) = \mathbb{P}(D^c|W)\mathbb{P}(W) + \mathbb{P}(D^c|W^c)\mathbb{P}(W^c)\\ &= [1-\mathbb{P}(D|W)]\mathbb{P}(W) + [1-\mathbb{P}(D|W^c)](1-\mathbb{P}(W))\\ &= (1-0.15)\cdot 0.9 + (1-0.8)\cdot(1-0.9) = 0.785.\end{aligned}$$

Here we use the fact that $\mathbb{P}(\cdot|W)$ is a probability when the condition W is fixed, thus $\mathbb{P}(D^c|W) = 1-P(D|W)$. Or you can show the equation is true by directly using the definition of conditional probability and rule of total probability.

(b) The required probability can be obtained through the multiplication principle

$$\begin{aligned}\mathbb{P}(W^c|D) &= \frac{\mathbb{P}(W^c \cap D)}{\mathbb{P}(D)}\\ &= \frac{\mathbb{P}(D|W^c)\mathbb{P}(W^c)}{1-\mathbb{P}(D^c)}\\ &= \frac{0.8 \cdot (1-0.9)}{1-0.785} \approx 0.372.\end{aligned}$$

Exercise 12. Of three cards, one is painted red on both sides, one is painted black on both sides, and one is painted red on one side and black on the other. A card is randomly chosen and placed on a table. If the side facing up is red, what it the probability that the other side is also red?

Answer: The sample space for the colours of sides of cards is

$$\Omega = \{(B, B), (B, R), (R, R)\},$$

or shortly denoted by

$$\Omega = \{BB, BR, RR\},$$

and the events $\{BB\}$, $\{BR\}$, $\{RR\}$ constitute a partition of Ω.
Denote by C the condition that a randomly chosen card with the side facing up red, the event (B, B) is impossible in this case. By using Bayes' formula,

$$\mathbb{P}(RR|C) = \frac{\mathbb{P}(RR \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(RR)}{\mathbb{P}(C|BB)\mathbb{P}(BB) + \mathbb{P}(C|BR)\mathbb{P}(BR) + \mathbb{P}(C|RR)\mathbb{P}(RR)},$$

where $\mathbb{P}(BB) = \mathbb{P}(BR) = \mathbb{P}(RR) = \frac{1}{3}$, and the relevant conditional probabilities are $\mathbb{P}(C|BB) = 0$ as when both sides are coloured black there is no chance to have one side red, $\mathbb{P}(C|BR) = \frac{1}{2}$ as for a card with one side red and another side black, the chance that the side facing up red is 0,5; and $\mathbb{P}(C|RR) = 1$, as if we have both the sides of a card red, definitely we will have one side facing up red. Then substituting all these probabilities into the formula above yields

$$\mathbb{P}(RR|C) = \frac{\frac{1}{3}}{0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{2}{3}.$$

Exercise 13. A total of 600 of the 1,000 people in a retirement community classify themselves as Republicans, while the others classify themselves as Democrats. In a local election in which everyone voted, 60 Republicans voted for the Democratic candidate, and 50 Democrats voted for the Republican candidate. If a randomly chosen community member voted for the Republican, what is the probability that she or he is a Democrat?

Answer: Let r, d denote respectively the events that a person is a Republican and Democrat; R, D denote respectively the events that a person votes for the Republican and Democratic candidates. Then

$$\mathbb{P}(r) = \frac{600}{1000} = 0.6,\ \mathbb{P}(d) = \frac{1000-600}{1000} = 0.4,$$

and

$$\mathbb{P}(D|r) = \frac{\mathbb{P}(D \cap r)}{\mathbb{P}(r)} = \frac{60/1000}{600/1000} = 0.1,\ \mathbb{P}(R|d) = \frac{\mathbb{P}(R \cap d)}{\mathbb{P}(d)} = \frac{50/1000}{(1000-600)/1000} = 0.125.$$

In the experiment of choosing 1 person at random, the sample space $\Omega = \{r, d\}$ and $\{r\}$, $\{d\}$ constitute a partition of Ω. Therefore, the probability that a randomly chosen community member voting for the republican is a Democrat can be obtained by Bayes' formula

$$\mathbb{P}(d|R) = \frac{\mathbb{P}(R|d)\mathbb{P}(d)}{\mathbb{P}(R|d)\mathbb{P}(d) + \mathbb{P}(R|r)\mathbb{P}(r)} = \frac{0.125 \cdot 0.4}{0.125 \cdot 0.4 + (1-0.1) \cdot 0.6} = \frac{5}{59}.$$

Exercise 14. Prostate cancer is the most common type of cancer found in males. As an indicator of whether a male has prostate cancer, doctors often perform a test that measures the level of the PSA protein (prostate specific antigen) that is produced only by the prostate gland. Although higher PSA levels are indicative of cancer, the test is notoriously unreliable. Indeed, the probability that a noncancerous man will have an elevated PSA level is approximately 0.135, with this probability increasing to approximately 0.268 if the man does have cancer. If based on other fact, a physician is 70 percent certain that a male has prostate cancer, what is the conditional probability that he has the cancer given that

(a) the test indicates an elevated PSA level;

(b) the test does not indicate an elevated PSA level?

Repeat the preceding, this time assuming that the physician initially believes there is a 30 percent chance the man has prostate cancer.

Answer: Let H denote the event of a higher PSA level and C that there is cancer. Then we have known that

$$\mathbb{P}(H|C^c) = 0.135,\ \mathbb{P}(H|C) = 0.268,\ \mathbb{P}(C) = 0.7.$$

Consider the sample space $\Omega = \{C, C^c\}$ and then the events $\{C\}$ and $\{C^c\}$ constitute a partition of Ω.

(a) The probability that he suffers from cancer given that the test indicates an elevated PSA level is

$$\mathbb{P}(C|H) = \frac{\mathbb{P}(H|C)\mathbb{P}(C)}{\mathbb{P}(H|C)\mathbb{P}(C) + \mathbb{P}(H|C^c)\mathbb{P}(C^c)} = \frac{0.268 \cdot 0.7}{0.268 \cdot 0.7 + 0.135 \cdot 0.3} \approx 0.8224.$$

(b) By Bayes' formula the probability that he suffers from cancer given that the test does not indicate an elevated PSA level is

$$\begin{aligned}\mathbb{P}(C|H^c) =& \frac{\mathbb{P}(H^c|C)\mathbb{P}(C)}{\mathbb{P}(H^c|C)\mathbb{P}(C) + \mathbb{P}(H^c|C^c)\mathbb{P}(C^c)}\\ =& \frac{(1-0.268)\cdot 0.7}{(1-0.268)\cdot 0.7 + (1-0.135)\cdot 0.3} \approx 0.6638.\end{aligned}$$

If alternatively, we have $\mathbb{P}(C) = 0.3$, the probabilities above become

$$\mathbb{P}(C|H) = \frac{0.268\cdot 0.3}{0.268\cdot 0.3 + 0.135\cdot 0.7} \approx 0.3597$$

and

$$\mathbb{P}(C|H^c) = \frac{(1-0.268)\cdot 0.3}{(1-0.268)\cdot 0.3 + (1-0.135)\cdot 0.7} \approx 0.2661.$$

Exercise 15. Suppose that we toss 2 fair dice. Let E_1 denote the event that the sum of the dice is 6, E_2 denote the event that the sum of the dice equals 7 and F the event that the first die equals 4.

(a) Is E_1 independent of F? Justify your answer.

(b) Is E_2 independent of F? Justify your answer.

Answer:

(a)

$$\mathbb{P}(E_1 \cap F) = \mathbb{P}(\{4, 2\}) = \frac{1}{36}$$

whereas

$$\mathbb{P}(E_1)\mathbb{P}(F) = \left(\frac{5}{36}\right)\left(\frac{1}{6}\right) = \frac{5}{216}.$$

$\mathbb{P}(E_1F) = \frac{1}{36} \neq \frac{5}{216} = \mathbb{P}(E_1)\mathbb{P}(F)$, thus by the definition of two independent events, E_1 and F are not independent.

(b) The answer is yes, since

$$\mathbb{P}(E_2 \cap F) = \mathbb{P}(\{4, 3\}) = \frac{1}{36}$$

whereas

$$\mathbb{P}(E_2)\mathbb{P}(F) = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}.$$

Exercise 16. Suppose a couple is waiting for a baby to be born, both of them has the combination of ab, where a and b are certain given genes. Assume that each of the couple will transmit independently either the genes a or b to the baby, with equal probability $1/2$. The baby acquires a character $\mathcal{C}$ if and only if its parents transmit one of the combinations aa or bb to the baby. We consider the following events:

$$A = \{\text{the father transmits the gene } a\},$$
$$B = \{\text{the mother transmits the gene } b\},$$
$$C = \{\text{the baby shows the character } \mathcal{C}\}.$$

Determine the following probabilities

(a) $\mathbb{P}(A)$;

(b) $\mathbb{P}(B)$;

(c) $\mathbb{P}(C)$;

(d) $\mathbb{P}(A \cap B)$;

(e) $\mathbb{P}(A \cap C)$;

(f) $\mathbb{P}(B \cap C)$;

(g) $\mathbb{P}(A \cap B \cap C)$;

(h) Are the events A, B, C mutually independent?

Answer:

(a) $\mathbb{P}(A) = \dfrac{1}{2}$;

(b) $\mathbb{P}(B) = \dfrac{1}{2}$;

(c)

$$\begin{aligned}
\mathbb{P}(C) =& \mathbb{P}(\{\text{the baby acquires the character } \mathcal{C}\}) \\
=& \mathbb{P}(\{\text{the baby has a combination either } aa \text{ or } bb\}) \\
=& \mathbb{P}(\{\text{the baby has a combination } aa\}) + \mathbb{P}(\{\text{the baby has a combination } bb\}) \\
=& \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}, \text{ by independence of the transmission} \\
=& \frac{1}{2}.
\end{aligned}$$

(d) Since the events A and B are independent,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

(e)

$$\begin{aligned}
A \cap C =& \{\text{the father transmits the gene } a \text{ to the baby and it shows the character } \mathcal{C}\} \\
=& \{\text{ the baby acquires the combination of } aa\} \\
=& \{\text{the baby acquires a gene } a \text{ from its mother and another gene } a \text{ from its father}\}
\end{aligned}$$

Thus by independence of the transmission,

$$\begin{aligned}
& \mathbb{P}(A \cap C) \\
=& \mathbb{P}(\text{mother transmits a gene } a \text{ to the baby}) \cdot \mathbb{P}(\text{father transmits a gene } a \text{ to the baby}) \\
=& \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.
\end{aligned}$$

(f)

$$
\begin{aligned}
B \cap C = &\{\text{the mother transmits the gene } b \text{ to the baby and it shows the character } \mathcal{C}\}\\
= &\{\text{ the baby acquires the combination of } bb\}\\
= &\{\text{the baby acquires a gene } b \text{ from its mother and another gene } b \text{ from its father}\}
\end{aligned}
$$

Thus by independence of the transmission,

$$
\begin{aligned}
&\mathbb{P}(A \cap C)\\
=&\mathbb{P}(\text{mother transmits a gene } b \text{ to the baby}) \cdot \mathbb{P}(\text{father transmits a gene } b \text{ to the baby})\\
=&\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.
\end{aligned}
$$

(g) Since $A \cap B \cap C = \emptyset$,

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\emptyset) = 0.$$

(h) By Parts (a), (b), (c) and (g),

$$\mathbb{P}(A \cap B \cap C) = 0 \neq \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C),$$

so the events A, B and C are not mutually independent.

Exercise 17. A sequence of independent trials is to be performed. Each trial results in a success with probability p and a failure with probability $1 - p$. What is the probability that

(a) at least 1 success occurs in the first n trials;

(b) exactly k successes occur in the first n trials;

Answer:

(a) In order to determine the probability of at least 1 success in the first n trials, it is easiest to compute first the probability of the complementary event: that of no successes in the first n trials. If we let E_i denote the event of a failure on the ith trial, then the probability of no successes is, by independence,

$$\mathbb{P}(E_1 E_2 \cdots E_n) = \mathbb{P}(E_1)\mathbb{P}(E_2) \cdots \mathbb{P}(E_n) = (1 - p)^n.$$

Hence, the answer to Part (a) is $1 - (1 - p)^n$.

(b) Consider any particular sequence of the first n outcomes containing k successes and $n - k$ failures. Each one of these sequences will, by the assumed independence of trials, occur with probability $p^k(1 - p)^{n-k}$. Since there are $\binom{n}{k}$ such sequences (there are $n!/k!(n - k)!$ permutations of k successes and $n - k$ failures), the desired probability in Part (b) is

$$\mathbb{P}(\{\text{exactly } k \text{ successes}\} = \binom{n}{k} p^k(1 - p)^{n-k}.$$