Solutions to Exercise 1A: Probability Gymnastics

1. a) The *sample space* S is the set of all possible outcomes of an experiment. Each run of an experiment will have as its outcome one of the elements of the sample space. In that sense, the sample space defines the "universe" in which we describe the outcomes of an experiment.

   b) An *event* A is a well-defined set of possible outcomes of the experiment. It is a subset of the sample space, $A \subseteq S$. We say that an event has occurred if the actual outcome of the experiment is an element of that subset.

   c) A *random variable* is a function $X : S \mapsto \mathbb{R}$ that maps the sample space to the real numbers, so it assigns a real number to each possible outcome. It's our way to map the (generally very numerous) outcomes in the sample space to a "summary" quantity that we're interested in. For instance, a random variable could map sequences of 1,000 coin tosses to the number of heads in the sequence, or the daily transactions in a supermarket to the average amount spent on fruits and vegetables per transaction.

2. *(1.5) One ball is to be selected from a box containing red, white, blue, yellow and green balls. If the probability that the selected ball will be red is 1/5 and the probability that it will be white is 2/5, what is the probability that it will be blue, yellow, or green?*

   Since the list of 5 colours is exhaustive, the probability that the ball will have one of these colours is 1. The colours are mutually exclusive, so the events corresponding to drawing a ball of a certain colour are disjoint. Therefore, by Axiom 3 (De Groot and Schervish, p. 17):

   $$\Pr(\textsc{red} \cup \textsc{white} \cup \textsc{blue} \cup \textsc{yellow} \cup \textsc{green}) = 1 =$$
   $$\Pr(\textsc{red}) + \Pr(\textsc{white}) + \Pr(\textsc{blue} \cup \textsc{yellow} \cup \textsc{green})$$

   Solving for the required probability,

   $$\Pr(\textsc{blue} \cup \textsc{yellow} \cup \textsc{green}) = 1 - \Pr(\textsc{red}) - \Pr(\textsc{white}) = 1 - \frac{1}{5} - \frac{2}{5} = \frac{2}{5}.$$

3. *(1.6) If two balanced dice are rolled, what is the probability that the sum of the two numbers that appear will be even?*

   There are 36 possible outcomes of throwing two dice. Half of them sum to an even number (all outcomes where both die 1 and die 2 show either an odd or an even number) and half of them to an odd number (odd number on die 1 paired with even number on die 2, or vice versa). Therefore, the probability is 0.5.

4. *(1.6) If two balanced dice are rolled, what is the probability that the difference of the two numbers that appear is less than 3?*

   To achieve this result, we can pair dice as follows:
   ⚀ with ⚀, ⚁ or ⚂

⚁ with ⚀, ⚁, ⚂ or ⚃
⚂ with ⚀, ⚁, ⚂, ⚃ or ⚄
⚃ with ⚁, ⚂, ⚃, ⚄ or ⚅
⚄ with ⚂, ⚃, ⚄ or ⚄
⚅ with ⚃, ⚄ or ⚅
These are 24 outcomes out of 36 possible, so the probability is 2/3.

5. *(2.1) If* $A \subset B$ *with* $\Pr(B) > 0$*, what is the value of* $\Pr(A|B)$*?*

   Since $A \subset B$, we know that $A \cap B = A$. Starting with the definition of conditional probability (Definition 2.1.1 in De Groot and Schervish),

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)}.$$

6. *(2.1) If A and B are disjoint events and Pr(B) > 0, what is the value of Pr(A|B)?*

   Since A and B are disjoint, we know that $A \cap B = \{\}$, and (by Theorem 1.5.1) $\Pr(\{\}) = 0$. Therefore,

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(\{\})}{\Pr(B)} = \frac{0}{\Pr(B)} = 0.$$

7. *(2.1) A box contains three coins with a head on each side, four coins with a tail on each side, and two fair coins. If one of these nine coins is selected at random and tossed once, what is the probability that a head will be obtained?*

   The events of picking each type of coin are disjoint with probabilities 3/9 (two heads), 4/9 (two tails) and 2/9 (fair coin), respectively. We can therefore decompose the total probability:

$$\Pr(\text{HEAD}) = \frac{3}{9} \cdot 1 + \frac{4}{9} \cdot 0 + \frac{2}{9} \cdot \frac{1}{2} = \frac{1}{3} + 0 + \frac{1}{9} = \frac{4}{9}.$$

8. *(2.1) Suppose that a box contains one blue card and four red cards, which are labelled A, B, C, and D. Suppose also that two of these five cards are selected at random, without replacement.*

   a) *If it is known that card A has been selected, what is the probability that both cards are red?*

      After selecting the red card A, we have the blue card and cards B, C and D left in the box. Three out of these four cards are red, so the probability of selecting two red cards given that the first card was A is 3/4.

   b) *If it is known that at least one red card has been selected, what is the probability that both cards are red?*

      First of all, note that if we draw 2 cards out of 5, 4 of which are red, we are guaranteed to draw at least one red card. Therefore, the event we condition on has probability 1 and does not affect the calculcation.

The required probability then reduces to the probability of *not* drawing the blue card in any of the two draws. In the first draw, 4 out of 5 cards are red. If we first draw blue, we're out. Having drawn red in the first round, we have 4 cards left, of which 3 are red. The probability of drawing red twice is therefore $\frac{4}{5} \cdot \frac{3}{4} = \frac{3}{5}$.

9. *(2.2) If three balanced dice are rolled, what is the probability that all three numbers will be the same?*

   Three equal numbers can be achieved no matter what we get in the first roll. However, in the second and third roll, the numbers must be equal to that of the first, with an independent probability of $\frac{1}{6}$ each. The probability is therefore $\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$.

10. *(2.3) A new test has been devised for detecting a particular type of cancer. If the test is applied to a person who has this type of cancer, the probability that the person will have a positive reaction is 0.95 and the probability that the person will have a negative reaction is 0.05. If the test is applied to a person who does not have this type of cancer, the probability that the person will have a positive reaction is 0.05 and the probability that the person will have a negative reaction is 0.95. Suppose that in the general population, one person out of every 100,000 people has this type of cancer. If a person selected at random has a positive reaction to the test, what is the probability that they have this type of cancer?*

    The information in the text gives us $\Pr(\textsc{pos}|\textsc{cancer}) = 0.95$ (the probability of a positive test given this cancer is present), $\Pr(\textsc{pos}|\textsc{no cancer}) = 0.05$ (the probability of a positive test in a healthy person) and $\Pr(\textsc{cancer}) = 10^{-5}$ (the unconditional probability of a person suffering from this cancer). We can apply Bayes' theorem to find the required probability of cancer given a positive test. To determine the unconditional probability of a positive test in the denominator of the Bayes formula, we add the separate probabilities of a positive test in the affected and the healthy population.

$$\Pr(\textsc{cancer}|\textsc{pos}) = \frac{\Pr(\textsc{cancer})\Pr(\textsc{pos}|\textsc{cancer})}{\Pr(\textsc{pos})} =$$
$$\frac{\Pr(\textsc{cancer})\Pr(\textsc{pos}|\textsc{cancer})}{\Pr(\textsc{cancer})\Pr(\textsc{pos}|\textsc{cancer}) + (1-\Pr(\textsc{cancer}))\Pr(\textsc{pos}|\textsc{no cancer})} =$$
$$\frac{10^{-5} \cdot 0.95}{10^{-5} \cdot 0.95 + (1-10^{-5}) \cdot 0.05} = 0.000190$$

    The probability of a random subject with a positive test actually suffering from this cancer is only about 0.02%. This is due to the very low prevalence of the disease in the population in combination with the relatively high false positive rate of the test.

11. *(2.3) Suppose that 30 percent of the bottles produced in a certain plant are defective. If a bottle is defective, the probability is 0.9 that an inspector will notice it and remove it from the filling line. If a bottle is not defective, the probability is 0.2 that the inspector will think that it is defective and remove it from the filling line.*

a) *If a bottle is removed from the filling line, what is the probability that it is defective?*

This exercise is entirely analogous to Exercise 11.

$$\Pr(\textsc{defect}|\textsc{removed}) =$$
$$\frac{\Pr(\textsc{removed}|\textsc{defect})\Pr(\textsc{defect})}{\Pr(\textsc{removed}|\textsc{defect})\Pr(\textsc{defect}) + \Pr(\textsc{removed}|\textsc{not defect})(1-\Pr(\textsc{defect}))} =$$
$$\frac{0.9 \cdot 0.3}{0.9 \cdot 0.3 + 0.2 \cdot (1-0.3)} = 0.659$$

b) *If a customer buys a bottle that has not been removed from the filling line, what is the probability that it is defective?*

Now we are interested in the probability of a bottle being defective given that it has *not* been removed. To calculate this, we consider probabilities of the form $\Pr(\textsc{not removed}|\cdot) = 1 - \Pr(\textsc{removed}|\cdot)$. Other than that, the procedure is the same.

$$\Pr(\textsc{defect}|\textsc{not removed}) =$$
$$\frac{(1-\Pr(\textsc{removed}|\textsc{defect}))\Pr(\textsc{defect})}{(1-\Pr(\textsc{removed}|\textsc{defect}))\Pr(\textsc{defect}) + (1-\Pr(\textsc{removed}|\textsc{not defect}))(1-\Pr(\textsc{defect}))} =$$
$$\frac{(1-0.9) \cdot 0.3}{(1-0.9) \cdot 0.3 + (1-0.2) \cdot (1-0.3)} = 0.0508$$