IPDA1005 Introduction to Probability and Data Analysis

Solution to Worksheet 2

1. The combination lock on a briefcase has two dials, each marked off with sixteen notches. To open the case, a person first turns the left dial in a certain direction for two revolutions and then stops on a particular mark. The right dial is set in a similar fashion, after having been turned in a certain direction for two revolutions. How many different settings are possible? Sketch a tree diagram of the possibilities. (**Hint**: First list the operations required to operate the lock, and then their possibilities.)

   Solution. The key to doing this question is to recognize that for each lock the first step is to choose whether to rotate it clockwise or counterclockwise, which results in two choices. Then, the dial must be stopped at one of sixteen marks. Thus, the total number of possibilities is

   $$2 \times 16 \times 2 \times 16 = 1024$$

2. Alphonse Bertillon, a nineteenth-century French criminologist, developed an identification system based on eleven anatomical variables (height, head width, ear length, etc.) that presumably remain essentially unchanged during an individual's adult life. The range of each variable was divided into three subintervals: small, medium, and large.A person's Bertillon configuration is an ordered sequence of eleven letters, say,

   $$s, s, m, m, l, s, l, s, s, m, s$$

   where a letter indicates the individual's "size" relative to a particular variable.

   (a) How many possible distinct Bertillon sequences are possible?

   (b) How populated does a city have to be before it can be guaranteed that at least two citizens will have the same Bertillon configuration?

   Solution.

   (a) For each of the eleven anatomical variables there are three choices, so there are

   $$3 \times 3 \times \ldots = 3^{11}$$

   possible distinct Bertillon sequences.

   (b) To *guarantee* that at least two citizens have at least the same Bertillon sequence, we need a population of just one more person than the possible number of Bertillon sequences in part (a), i.e., $3^{11} + 1$.

3. A chemical engineer wishes to observe the effects of temperature, pressure, and catalyst concentration on the yield resulting from a certain reaction. If she intends to include two different temperatures, three pressures, and two levels of catalyst, how many different runs must she make in order to observe each temperature-pressure-catalyst combination exactly twice?

Solution. Using the multiplication rule, she needs $2 \times 3 \times 2 \times 2 = 24$ runs.

4. In a famous science fiction story by Arthur C. Clarke, *The Nine Billion Names of God*, a computer firm is hired by the lamas in a Tibetan monastery to write a program to generate all possible names of God. For reasons never divulged, the lamas believe that all such names can be written using *no more* than nine letters. If no letter combinations are ruled inadmissible, is the "nine billion" (it's actually nine *trillion*—9×10^{12}—because of the difference in English/American usage) in the story's title a large enough number to accommodate all possibilities?

   Solution. The total number of names N is the sum of all one-letter names, two-letter names, and so on. By the multiplication rule, the number of k-letter names is 26^k, so

   $$N = 26^1 + 26^2 + \ldots 26^9 = 5,646,683,826,134$$

   which is a few trillion less than nine trillion!

5. A friend is giving a dinner party. Her current wine supply includes 8 bottles of zinfandel, 10 of merlot, and 12 of cabernet (she drinks only red wine), all from different wineries.

   (a) If she wants to serve 3 bottles of zinfandel and serving order is important, how many ways are there to do this?

   (b) If 6 bottles of wine are to be randomly selected from the 30 for serving, how many ways are there to do this?

   (c) If 6 bottles are randomly selected, how many ways are there to obtain two bottles of each variety?

   (d) If 6 bottles are randomly selected, what is the probability that this results in two bottles of each variety being chosen? (**Hint**: Use the naïve definition of probability and quantities you have already calculated.)

   (e) If 6 bottles are randomly selected, what is the probability that all of them are the same variety? (**Hint**: Think carefully about how to calculate the numerator in the expression for naïve probability.)

   Solution.

   (a) $P_{3,8} = \frac{8!}{(8-3)!} = \frac{8!}{5!} = 8 \times 7 \times 6 = 336$

   (b) $\binom{30}{6} = \frac{30!}{24!6!} = 593,775$

   (c) $\binom{8}{2} \times \binom{10}{2} \times \binom{12}{2} = 83,160$

   (d) The answer is given by the ratio of (c) and (b):

   $$P_{\text{naïve}} = 0.14$$

   (e) If six bottles are chosen, the number of possible ways that they could all be of the same variety is

   $$\binom{8}{6} + \binom{10}{6} + \binom{12}{6} = 1,162$$

   and hence

   $$P_{\text{naïve}} = \frac{1162}{\binom{30}{6}} = 0.001957$$

6. The playlist on my ancient 4th generation iPod Nano (Apple stopped producing all iPods on July 27, 2017) contains 100 songs, of which 10 are by the Beatles. If I use the shuffle feature:

   (a) In how many ways can five songs be played regardless of the artist/singer?

   (b) In how many ways can the the first Beatles song be the fifth one played?

   (c) What is the probability that the first Beatles song is the fifth one played?

   (d) (**Challenging**: Use an alternative line of reasoning that involves combinations and focusing on the last 95 songs.)

   Solution.

   (a) Order is important here, so the number of ways is

   $$P_{5,100} = \frac{100!}{95!} = 100 \times 99 \times 98 \times 97 \times 96.$$

   (b) If the first four songs are not Beatles songs, but the fifth one is, the number of ways that this could occur is

   $$P_{4,90} \times 10 = \frac{90!}{86!} \times 10 = 90 \times 89 \times 88 \times 87 \times 10$$

   because there are 90 non-Beatles songs and 10 Beatles songs.

   (c) The probability that the first Beatles song is the fifth one played is simply the ratio of (c) and (b):

   $$P(1^{\text{st}}\text{Beatles song is fifth one played}) = \frac{P_{4,90} \times 10}{P_{5,100}} = 0.0679$$

   (d) Rather than focusing on selecting just the first 5 songs, think of playing all 100 songs in random order. The number of ways of choosing 10 of these songs to be Beatles songs (without regard to the order in which they are played) is $\binom{100}{10}$. Now if we choose 9 of the last 95 songs to be Bs, which can be done in $\binom{95}{9}$ ways, that leaves four non-Beatles songs and one Beatles song in the first five tunes. Finally, there is only one way for these first five songs to start with four non-Beatles tunes and then follow with a Beatles song (remember that we are considering unordered subsets). Thus

   $$P(1^{\text{st}}\text{Beatles song is fifth one played}) = \frac{\binom{95}{9}}{\binom{100}{10}}$$

   and you can verify that this expression does indeed give us the same answer as in part (c).

7. What do we perceive to be random? Most people would agree that 999999 seems less "random" than, say, 703928, but in what sense is that true? Imagine that we randomly generate a six-digit number, i.e., we make six draws with replacement from the digits 0 through 9.

   (a) What is the probability of generating 999999?

   (b) What is the probability of generating 703928?

   (c) What is the probability of generating a sequence of six identical digits?

Figure 0.1: Random numbers according to Dilbert (`http://dilbert.com/strip/2001-10-25`)

(d) What is the probability of generating a sequence with no identical digits? (Comparing the answers to (c) and (d) gives some sense of why some sequences feel intuitively more random than others.)

Solution. We'll assume here that a digit with a leading '0' is a legitimate digit.

(a) 10^{-6}

(b) 10^{-6}

(c) $10 \times 10^{-6} = 10^{-5}$

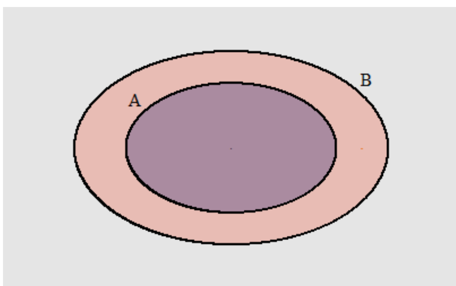
(d) The number of ways of generating a sequence with no identical digits is $P_{6,10}$, and hence the probability of generating such a sequence is $P_{6,10}/10^6 = 0.1512$, which is much larger than the probability in (c).

8. Let A and B be events defined on a sample space $\mathcal{S}$. The *difference* $B \setminus A$ is defined to be the set of all elements in B that are not in A. Show that if $A \subseteq B$, then

$$P(B \setminus A) = P(B) - P(A)$$

[**Hint**: It will help to draw a Venn diagram showing $A \subseteq B$.]

Solution.



As the figure above shows, $B - A$ and A are disjoint, so we can write, using the third axiom of probability, that

$$P(B) = P(B \setminus A) + P(A)$$

and rearranging this expression yields the desired result.

9. Show that

$$P(A \cap B) \geq 1 - P(A^c) - P(B^c)$$

for any two events A and B defined on a sample space $\mathcal{S}$.
(**Hint**: Begin by rearranging Property 3 in Lecture 2, and then use the axioms of probability thereafter.)

Solution. Using Properties 3 and 1, it's straightforward to show the following:

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 1 - P(A^c) - P(B^c) + 1 - P(A \cap B) \end{aligned}$$

But $P(A \cap B) \leq 1$, so it follows that

$$P(A \cap B) \geq 1 - P(A^c) - P(B^c)$$

(You'll find the following problems easier to do if you sketch the corresponding Venn diagram.)

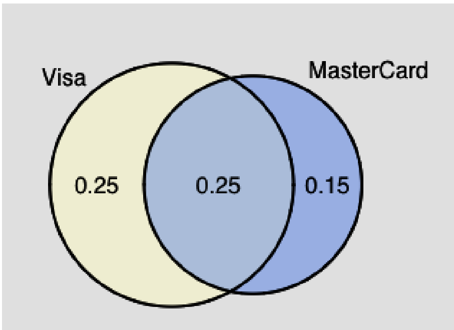
10. Consider randomly selecting a student at Curtin University, and let A denote the event that the selected individual has a Visa credit card and B be the analogous event for a MasterCard. Suppose that $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \cap B) = 0.25$.

    (a) Calculate the probability that the selected individual has at least one of the two types of cards.

    (b) What is the probability that the selected individual has neither type of card?

    (c) Describe in terms of A and B, the event that the selected student has a Visa card but not a MasterCard, and then calculate the probability of this event.

    Solution. Using the Venn diagram below, we can calculate the required probabilities, or by using a little bit of algebra.

    (a) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.5 + 0.4 - 0.25 = 0.65$

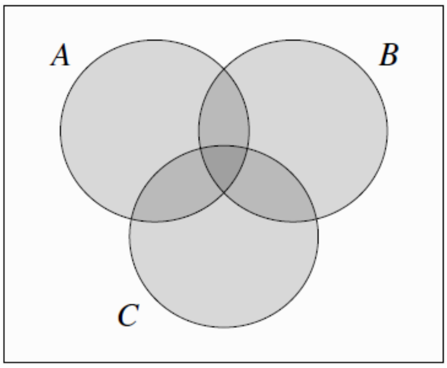
    (b) $P((A \cup B)^c) = 1 - P(A \cup B) = 0.35$

    (c) If the student has a Visa card (A) but not a MasterCard (B), this can be written as $P(A \cap B^c)$. From the Venn diagram, you can see that $P(A) = P(A \cap B) + P(A \cap B^c)$, and hence $P(A \cap B^c) = P(A) - P(A \cap B) = 0.5 - 0.25 = 0.25$.

11. Lucy is currently running two dot-com scams out of a bogus chatroom. She estimates that the chances of the first one leading to her arrest are one in ten; the "risk" associated with the second is more on the order of one in thirty. She considers the likelihood that she gets busted for both to be 0.0025. What are Lucy's chances of avoiding incarceration?

    Solution. This is similar to 1(b) above. The probability of getting caught for either or both scams is $P(A \cup B)$ and hence the probability of not getting caught at all is $1 - P(A \cup B) = 1 - (1/10 + 1/30 - 0.0025) = 0.869$.

12. Using the triple Venn diagram shown below, and your intuition, find an expression for $P(A \cup B \cup C)$. Use the expression for $P(A \cup B)$ as a guide.



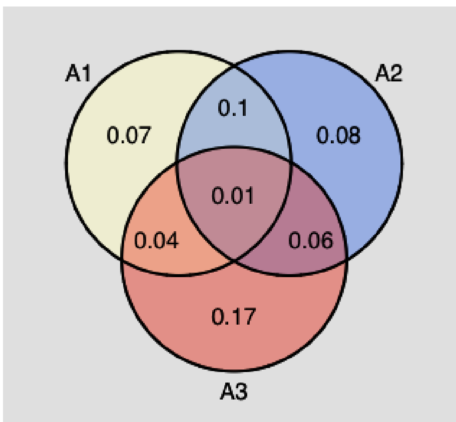
Solution. If we consider only $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, then our intuition tells us that after adding $P(A)$ and $P(B)$, we've double-counted $P(A \cap B)$, which is why we have to subtract $P(A \cap B)$. In much the same way, to get the total area of the shaded region $A \cup B \cup C$, we start by adding the areas of the three circles, $P(A) + P(B) + P(C)$. The three football-shaped regions have each been counted twice, so we then subtract $P(A \cap B) + P(A \cap C) + P(B \cap C)$. Finally, the region in the center has been added three times and subtracted three times, so in order to count it exactly once, we must add it back again. Hence,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

13. A consulting firm presently has bids out on three projects. Let $A_i = \{\text{awarded project } i\}$, for $i = 1, 2, 3$, and suppose that $P(A_1) = 0.22$, $P(A_2) = 0.25$, $P(A_3) = 0.28$, $P(A_1 \cap A_2) = 0.11)$, $P(A_1 \cap A_3) = 0.05$, $P(A_2 \cap A_3) = 0.07$, and $P(A_1 \cap A_2 \cap A_3) = 0.01$. Express in words each of the following events, and calculate their probabilities.

    (a) $A_1 \cup A_2$
    (b) $A_1^c \cap A_2^c$ (**Hint**: $(A_1 \cup A_2)^c = A_1^c \cap A_2^c$)
    (c) $A_1 \cup A_2 \cup A_3)$
    (d) $A_1^c \cap A_2^c \cap A_3^c$
    (e) $A_1^c \cap A_2^c \cap A_3$
    (f) $(A_1^c \cap A_2^c) \cup A_3$

    Solution. Using the information that has been provided, you should be able to sketch the following Venn diagram:



    (a) The firm is awarded at least one of the first two projects: $P(A_1 \cup A_2) = 0.22 + 0.25 - 0.11 = 0.36$.

(b) The firm is awarded neither of the first two projects: $P(A_1^c \cap A_2^c) = P(A_1 \cup A_2)^c = 1 - P(A_1 \cup A_2) = 0.64$.

(c) The firm is awarded at least one of the three projects: using the expression in the previous question, $P(A \cup B \cup C) = 0.22 + 0.25 + 0.28 - 0.11 - 0.05 - 0.07 + 0.01 = 0.53$.

(d) The firm is awarded none of the projects: It is straightforward to show, or illustrate using a Venn diagram, that $P(A_1^c \cap A_2^c \cap A_3^c) = P((A_1 \cup A_2 \cup A_3)^c) = 1 - P(A \cup B \cup C) = 1 - 0.53 = 0.47$.

(e) By looking at the Venn diagram above, a little bit of thinking will show that $A_1^c \cap A_2^c \cap A_3$ is that part of A_3 that is not in either A_1 or A_2 and hence represents the event that the firm is awarded only the third project. Again, using the Venn diagram, we can write that $P(A_1^c \cap A_2^c \cap A_3) = P(A_3) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cup A_2 \cup A_3) = 0.28 - 0.05 - 0.07 + 0.01 = 0.17$.

(f) $(A_1^c \cap A_2^c) \cup A_3$ can be interpreted as either the firm fails to get the first two projects or it is awarded the third, and looking at the diagram above, $P((A_1^c \cap A_2^c) \cup A_3) = 1 - (0.07 + 0.10 + 0.08) = 0.75$.

Sources: Q1, Q2, Q4, Q9 and Q11 are adapted from Larsen and Marx (2014), Q3, Q5, Q6, Q7, Q10 and Q13 are from Devore and Berk (2012) and Q8 and Q12 are from Blitzstein and Hwang (2015)

Bibliography

1. Blitzstein, J.K. and Hwang, J. (2014) *Introduction to Probability.* CRC Press/Taylor & Francis Group: Boca Raton, FL.

2. Devore, J.L. and Berk, K.N. (2012) *Modern Mathematical Statistics with Applications.* Springer: New York.

3. Larsen, R.J. and Marx, M.L. (2014) *An Introduction to Mathematical Statistics and Its Applications*, 5th ed. Prentice Hall: Boston.