Probability, Random Processes and Inference
Semester 2015-B
Dr. Ponciano Jorge Escamilla-Ambrosio
Serie 1

Naive definition of probability

1. A certain family has 6 children, consisting of 3 boys and 3 girls. Assuming that all birth orders are equally likely, what is the probability that the 3 eldest children are the 3 girls?

Solution:
Label the girls as 1, 2, 3 and the boys as 4, 5, 6: Think of the birth order is a permutation of 1, 2, 3, 4, 5, 6, e.g., we can interpret 314265 as meaning that child 3 was born first, then child 1, etc. The number of possible permutations of the birth orders is 6!. Now we need to count how many of these have all of 1, 2, 3 appear before all of 4; 5; 6. This means that the sequence must be a permutation of 1, 2, 3 followed by a permutation of 4, 5, 6. So with all birth orders equally likely, we have:

$$P(\text{the 3 girls are the 3 eldest children}) = \frac{(3!)^2}{6!} = 0.05.$$

Alternatively, we can use the fact that there are $\binom{6}{3}$ ways to choose where the girls appear in the birth order (without taking into account the ordering of the girls amongst themselves). These are all equally likely. Of these possibilities, there is only 1 where the 3 girls are the 3 eldest children. So again the probability is $1/\binom{6}{3} = 0.05$.

2. A college has 10 (non-overlapping) time slots for its courses, and blithely assigns courses to time slots randomly and independently. A student randomly chooses 3 of the courses to enrol in. What is the probability that there is a conflict in the student's schedule?

Solution:
The probability of no conflict is $(10 * 9 * 8)/10^3 = 0.72$. So the probability of there being at least one scheduling conflict is 0.28.

3. A jar contains *r* red balls and *g* green balls, where *r* and *g* are fixed positive integers. A ball is drawn from the jar randomly (with all possibilities equally likely), and then a second ball is drawn randomly.
(a) Explain intuitively why the probability of the second ball being green is the same as the probability of the first ball being green.
(b) Define notation for the sample space of the problem, and use this to compute the probabilities from (a) and show that they are the same.
(c) Suppose that there are 16 balls in total, and that the probability that the two balls are the same colour is the same as the probability that they are different colours. What are *r* and *g* (list all possibilities)?

Solution:
(a) This is true by symmetry. The first ball is equally likely to be any of the $g + r$ balls, so the probability of it being green is $g/(g+r)$. But the second ball is also equally likely to be any of the $g + r$ balls (there aren't certain

balls that enjoy being chosen second and others that have an aversion to being chosen second); once we know whether the first ball is green we have information that affects our uncertainty about the second ball, but before we have this information, the second ball is equally likely to be any of the balls.

Alternatively, intuitively it shouldn't matter if we pick one ball at a time, or take one ball with the left hand and one with the right hand at the same time. By symmetry, the probabilities for the ball drawn with the left hand should be the same as those for the ball drawn with the right hand.

(b) Label the balls as 1, 2,…, $g + r$, such that 1; 2,…, g are green and $g + 1$,…, $g + r$ are red. The sample space can be taken to be the set of all pairs (a, b) with $a, b \in \{1,\ldots, g + r\}$ and $a \neq b$ (there are other possible ways to define the sample space, but it is important to specify all possible outcomes using clear notation, and it make sense to be as richly detailed as possible in the specification of possible outcomes, to avoid losing information). Each of these pairs is equally likely, so we can use the naive definition of probability. Let G_i be the event that the ith ball drawn is green.

The denominator is $(g + r)(g + r - 1)$ by the multiplication rule. For G_1, the numerator is $g(g + r - 1)$, again by the multiplication rule. For G_2, the numerator is also $g(g + r - 1)$, since in counting favorable cases, there are g possibilities for the second ball, and for each of those there are $g + r - 1$ favorable possibilities for the first ball (note that the multiplication rule doesn't require the experiments to be listed in chronological order!); alternatively, there are $g(g - 1) + rg = g(g + r - 1)$ favorable possibilities for the second ball being green, as seen by considering 2 cases (first ball green and first ball red). Thus,

$$P(G_i) = \frac{g(g+r-1)}{(g+r)(g+r-1)} = \frac{g}{g+r},$$

for $i \in \{1, 2\}$, which concurs with (a).

(c) Let A be the event of getting one ball of each color. In set notation, we can write $A = (G1 \cap G^c_2) \cup (G^c_1 \cap G_2)$. We are given that $P(A) = P(A^c)$, so $P(A) = 1/2$. Then

$$P(A) = \frac{2gr}{(g+r)(g+r-1)} = \frac{1}{2},$$

Giving the quadratic equation:

$$g^2 + r^2 - 2gr - g - r = 0,$$

i.e.,

$$(g - r)^2 = g + r.$$

But $g + r = 16$, so $g - r$ is 4 or –4. Thus, either $g = 10$, $r = 6$, or $g = 6$; $r = 10$.

Total Probability Theorem

1. Alice and Bob have 2*n* + 1 coins, each coin with probability of heads equal to 1/2. Bob tosses *n* + 1 coins, while Alice tosses the remaining *n* coins. Assuming independent coin tosses, show that the probability that after all coins have been tossed, Bob will have gotten more heads than Alice is 1/2.

Solution:

Since Bob tosses one more coin that Alice, it is impossible that they toss both the same number of heads and the same number of tails. So Bob tosses either more heads than Alice or more tails than Alice (but not both). Since the coins are fair, these events are equally likely by symmetry, so both events have probability 1/2.

An alternative solution is to argue that if Alice and Bob are tied after 2*n* tosses, they are equally likely to win. If they are not tied, then their scores differ by at least 2, and toss 2*n*+1 will not change the final outcome. This argument may also be expressed algebraically by using the total probability theorem. Let B be the event that Bob tosses more heads. Let X be the event that after each has tossed *n* of their coins, Bob has more heads than Alice, let Y be the event that under the same conditions, Alice has more heads than Bob, and let Z be the event that they have the same number of heads. Since the coins are fair, we have P(X) = P(Y), and also P(Z) = 1 – P(X) – P(Y). Furthermore, we see that

$$\mathbf{P}(B\,|\,X) = 1, \qquad \mathbf{P}(B\,|\,Y) = 0, \qquad \mathbf{P}(B\,|\,Z) = \frac{1}{2}.$$

Now we have, using the total probability theorem,

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(X)\cdot\mathbf{P}(B\,|\,X) + \mathbf{P}(Y)\cdot\mathbf{P}(B\,|\,Y) + \mathbf{P}(Z)\cdot\mathbf{P}(B\,|\,Z)\\ &= \mathbf{P}(X) + \frac{1}{2}\cdot\mathbf{P}(Z)\\ &= \frac{1}{2}\cdot\big(\mathbf{P}(X) + \mathbf{P}(Y) + \mathbf{P}(Z)\big)\\ &= \frac{1}{2}.\end{aligned}$$

as required.

2. Two players take turns removing a ball from a jar that initially contains *m* white and *n* black balls. The first player to remove a white ball wins. Develop a recursive formula that allows the convenient computation of the probability that the starting player wins.

Solution:

Let *p*(*m*, *k*) be the probability that the starting player wins when the jar initially contains *m* white and *k* black balls. We have, using the total probability theorem,

$$p(m,k) = \frac{m}{m+k} + \frac{k}{m+k}\big(1 - p(m,k-1)\big) = 1 - \frac{k}{m+k}p(m,k-1).$$

The probabilities $p(m, 1)$, $p(m, 2)$,…, $p(m, n)$ can be calculated sequentially using this formula, starting with the initial condition $p(m, 0) = 1$.

Conditional probability and Bayes' rule

1. A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase "free money" is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention "free money". What is the probability that it is spam?

Solution: Let S be the event that an email is spam and F be the event that an email has the "free money" phrase. By Bayes' rule,

$$P(S|F) = \frac{P(F|S)P(S)}{P(F)} = \frac{0.1 \cdot 0.8}{0.1 \cdot 0.8 + 0.01 \cdot 0.2} = \frac{80/1000}{82/1000} = \frac{80}{82} \approx 0.9756.$$

2. A crime is committed by one of two suspects, A and B. Initially, there is equal evidence against both of them. In further investigation at the crime scene, it is found that the guilty party had a blood type found in 10% of the population. Suspect A does match this blood type, whereas the blood type of Suspect B is unknown.

(a) Given this new information, what is the probability that A is the guilty party?
(b) Given this new information, what is the probability that B's blood type matches that found at the crime scene?

Solution:
(a) Let *M* be the event that *A*'s blood type matches the guilty party's and for brevity, write A for "A is guilty" and B for "B is guilty". By Bayes' Rule,

$$P(A|M) = \frac{P(M|A)P(A)}{P(M|A)P(A) + P(M|B)P(B)} = \frac{1/2}{1/2 + (1/10)(1/2)} = \frac{10}{11}.$$

(We have P(MjB) = 1=10 since, given that B is guilty, the probability that A's blood type matches the guilty party's is the same probability as for the general population.)

(b) Let C be the event that B's blood type matches, and condition on whether B is guilty. This gives:

$$P(C|M) = P(C|M, A)P(A|M) + P(C|M, B)P(B|M) = \frac{1}{10} \cdot \frac{10}{11} + \frac{1}{11} = \frac{2}{11}.$$

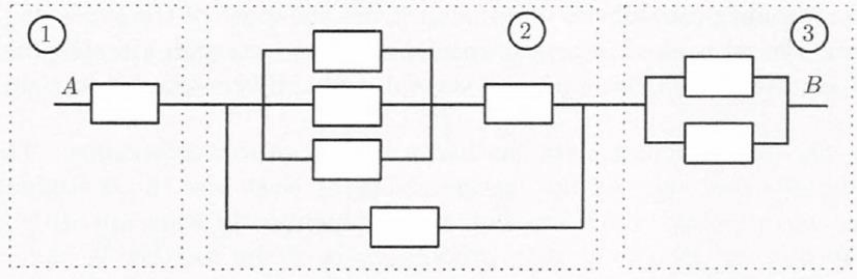
3. A batch of 100 items is inspected by testing four randomly selected items. If one of the four is defective, the batch is rejected. What is the probability that the batch is accepted if it contains five defectives?

Solution:
Let A be the event that the batch will be accepted. Then $A = A_1 \cap A_2 \cap A_3 \cap A_4$, where A_i, $i = 1,\ldots, 4$, is the event that the *i*th item is not defective. Using the multiplication rule, we have:

$$\mathbf{P}(A) = \mathbf{P}(A_1)\mathbf{P}(A_2 \mid A_1)\mathbf{P}(A_3 \mid A_1 \cap A_2)\mathbf{P}(A_4 \mid A_1 \cap A_2 \cap A_3) = \frac{95}{100} \cdot \frac{94}{99} \cdot \frac{93}{98} \cdot \frac{92}{97} = 0.812.$$

Independence

1. An electrical system consists of identical components, each of which is operational with probability p, independent of other components. The components are connected in three subsystems, as shown in the figure below. The system is operational if there is a path that starts at point A, ends at point B, and consists of operational components. What is the probability of this happening?



Solution:

The system may be viewed as a series connection of three subsystems, denoted 1, 2, and 3 in the figure. The probability that the entire system is operational is $p_1p_2p_3$, where p_i is the probability that subsystem i is operational. Using the formulas for the probability of success of a series or a parallel system, we have:

$$p_1 = p, \qquad p_3 = 1 - (1-p)^2,$$

and

$$p_2 = 1 - (1-p)\big(1 - p\big(1 - (1-p)^3\big)\big).$$

Counting

1. (a) How many ways are there to split a dozen people into 3 teams, where one team has 2 people, and the other two teams have 5 people each?

(b) How many ways are there to split a dozen people into 3 teams, where each team has 4 people?

Solution:

(a) Pick any 2 of the 12 people to make the 2 person team, and then any 5 of the remaining 10 for the first team of 5, and then the remaining 5 are on the other team of 5; this over counts by a factor of 2 though, since there is no designated “first" team of 5. So the number of possibilities is $\frac{\binom{12}{2}\binom{10}{5}}{2} = 8316$. Alternatively, politely ask the 12 people to line up, and then let the first 2 be the team of 2, the next 5 be a team of 5, and then last 5 be a team of 5. There are 12! ways for them to line up, but it does not matter which order they line up in within each group, nor does the order of the 2 teams of 5 matter, so the number of possibilities is $\frac{12!}{2!5!5! \cdot 2} = 8316$.

(b) By either of the approaches above, there are $\frac{12!}{4!4!4!}$ ways to divide the people into a Team A, a Team B, and a Team C, if we care about which team is which (this is called a multinomial coefficient). Since here it doesn't matter which team is which, this over counts by a factor of 3!, so the number of possibilities is $\frac{12!}{4!4!4!3!} = 5775$.

2. We deal from a well-shuffled 52-card deck. Calculate the probability that the 13th card is the first king to be dealt.

Solution:
The probability that the 13th card is the first king to be dealt is the probability that out of the first 13 cards to be dealt, exactly one was a king, and that the king was dealt last. Now, given that exactly one king was dealt in the first 13 cards, the probability that the king was dealt last is just 1/13, since each "position" is equally likely. Thus, it remains to calculate the probability that there was exactly one king in the first 13 cards dealt. To calculate this probability we count the "favorable" outcomes and divide by the total number of possible outcomes. We first count the favorable outcomes, namely those with exactly one king in the first 13 cards dealt. We can choose a particular king in 4 ways, and we can choose the other 12 cards in $\binom{48}{12}$ ways, therefore there are $4 \cdot \binom{48}{12}$ favorable outcomes. There are $\binom{52}{13}$ total outcomes, so the desired probability is:

$$\frac{1}{13} \cdot \frac{4 \cdot \binom{48}{12}}{\binom{52}{13}}.$$

For an alternative solution, we argue as in Example 1.10. The probability that the first card is not a king is 48/52. Given that, the probability that the second is not a king is 47/51. We continue similarly until the 12th card. The probability that the 12th card is not a king, given that none of the preceding 11 was a king, is 37/41. (There are 52 – 11 = 41 cards left, and 48 – 11 = 37 of them are not kings). Finally, the conditional probability that the 13th card is a king is 4/40. The desired probability is

$$\frac{48 \cdot 47 \cdots 37 \cdot 4}{52 \cdot 51 \cdots 41 \cdot 40}.$$

3. A well-shuffled 52-card deck is dealt to 4 players. Find the probability that each of the players gets an ace.

Solution:
The size of the sample space is the number of different ways that 52 objects can be divided in 4 groups of 13, and is given by the multinomial formula:

$$\frac{52!}{13!\,13!\,13!\,13!}.$$

There are 4! different ways of distributing the 4 aces to the 4 players, and there are:

$$\frac{48!}{12!\,12!\,12!\,12!}$$

different ways of dividing the remaining 48 cards into 4 groups of 12. Thus, the desired probability is:

$$\frac{4!\dfrac{48!}{12!\,12!\,12!\,12!}}{\dfrac{52!}{13!\,13!\,13!\,13!}}.$$

An alternative solution can be obtained by considering a different, but probabilistically equivalent method of dealing the cards. Each player has 13 slots, each one of which is to receive one card. Instead of shuffling the deck, we place the 4 aces at the top, and start dealing the cards one at a time, with each free slot being equally likely to receive the next card. For the event of interest to occur, the first ace can go anywhere; the second can go to any one of the 39 slots (out of the 51 available) that correspond to players that do not yet have an ace; the third can go to any one of the 26 slots (out of the 50 available) that correspond to the two players that do not yet have an ace; and finally, the fourth, can go to any one of the 13 slots (out of the 49 available) that correspond to the only player who does not yet have an ace. Thus, the desired probability is:

$$\frac{39 \cdot 26 \cdot 13}{51 \cdot 50 \cdot 49}.$$

By simplifying our previous answer, it can be checked that it is the same as the one obtained here, thus corroborating the intuitive fact that the two different ways of dealing the cards are probabilistically equivalent.

R

Develop the exercise 2.10 of the book Introduction to Probability by Joseph K. Blitzstein and Jessica Hwang, 2014.