Probability and Statistics I

STAT 3600 – Fall 2021

Le Chen
lzc0090@auburn.edu

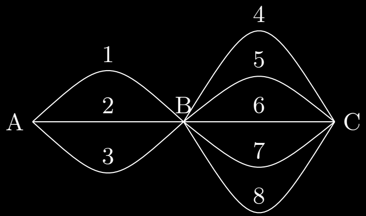
Last updated on

July 4, 2021

Auburn University
Auburn AL

Chapter 1. Probability

Example 1.2-1 (Routes between Cities) Suppose that there are three different routes from city A to city B and five different routes from city B to city C. The cities and routes are depicted in the figure, with the routes numbered from 1 to 8. We wish to count the number of different routes from A to C that pass through B. For example, one such route is 1 followed by 4, which we can denote (1, 4). Similarly, there are the routes $(1, 5), (1, 6)$, ..., $(3, 8)$. It is not difficult to see that the number of different routes $3 \times 5 = 15$.



Theorem 1.2-1 (Multiplication Principle) Suppose that an experiment (or procedure) E_1 has n_1 outcomes and, for each of these possible outcomes, an experiment (procedure) E_2 has n_2 possible outcomes. Then the has $n_1 n_2$ possible outcomes.

More generally, suppose that an experiment has k parts $(k \geq 2)$, that the i-th part of the experiment can have n_i possible outcomes $(i = 1, ..., k)$, and that all of the outcomes in each part can occur regardless of which specific outcomes have occurred in the other parts. Then the sample space S of the experiment will contain all vectors of the form $(u_1, \cdots, u_k)$, where u_i is one of the n_i possible outcomes of part $i (i = 1, ..., k)$. The total number of these vectors in S will be equal to the product $n_1 n_2 \cdots n_k$.

Example 1.2-2 A boy found a bicycle lock for which the combination was unknown. The correct combination is a four-digit number, $d_1 d_2 d_3 d_4$, where d_i, $i = 1, 2, 3, 4$, is selected from $1, 2, 3, 4, 5, 6, 7$, and 8. How many different lock combinations are possible with such a lock?

Suppose that n positions are to be filled with n different objects. There are n choices for filling the first position, $n-1$ for the second, ..., and 1 choice for the last position. So, by multiplication rule, there are

$$n(n-1)(n-2)\cdots(2)(1) = n!$$

possible arrangements. The symbol $n!$ is read "n factorial."

Definition 1.2-1 (Permutation) Each of the $n!$ arrangements (in a raw) of n different objects is called a *permutation* of the n objects.

Example 1.2-3 Order 7 books on a shelf = 7! permutations.

If only r positions are to be filled with objects selected from n different objects, $r \leq n$, then the number of possible ordered arrangements is

$$_nP_r = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}$$

Definition 1.2-2 Each of the $_nP_r$ arrangements is called a *permutation of n objects taken r at a time.*

Remark 1.2-1 Sampling without replacement, one at a time, order is important!

Example 1.2-4 (Choosing Officers) Suppose that a club consists of 25 members and that a president and a secretary are to be chosen from the membership. We shall determine the total possible number of ways in which these two positions can be filled.

Since the positions can be filled by first choosing one of the 25 members to be president and then choosing one of the remaining 24 members to be secretary, the possible number of choices is ${}_{25}P_2 = (25)(24) = 600$.

Example 1.2-4 (Choosing Officers) Suppose that a club consists of 25 members and that a president and a secretary are to be chosen from the membership. We shall determine the total possible number of ways in which these two positions can be filled.

Since the positions can be filled by first choosing one of the 25 members to be president and then choosing one of the remaining 24 members to be secretary, the possible number of choices is ${}_{25}P_2 = (25)(24) = 600$.

Sampling with Replacement

Consider a box that contains n balls numbered $1, \ldots, n$. First, one ball is selected at random from the box and its number is noted. This ball is then put back in the box and another ball is selected (it is possible that the same ball will be selected again). As many balls as desired can be selected in this way. This process is called sampling with replacement. It is assumed that each of the n balls is equally likely to be selected at each stage and that all selections are made independently of each other.

Suppose that a total of r selections are to be made, where r is a given positive integer. Then the sample space S of this experiment will contain all vectors of the form $(x_1, \ldots, x_k)$, where x_i is the outcome of the ith selection $(i = 1, \ldots, k)$. Since there are n possible outcomes for each of the r selections, the total number of vectors in S is n^r. Furthermore, from our assumptions it follows that S is a equally likely sample space. Hence, the probability assigned to each vector in S is $1/n^r$.

Example 1.2-5 (Birthday problem) In a group of k people, what is the probability that at least 2 people will have the same birthday? Assume $n = 365$ and that birthdays are equally distributed throughout the year, no twins, etc.

Solution.

1. Since there are 365 possible birthdays for each of k people, the sample space S will contain 365^k outcomes, all of which will be equally probable.
2. If $k > 365$, there are not enough birthdays for every one to be different, and hence at least two people must have the same birthday. (*Pigeonhole principle*).
3. So, we assume below that $k \leq 365$.
4. Denote

$$A_k = \{\text{at least 2 people have the same birthday in a group of } k \text{ people}\}$$

and hence,

$$A_k' = \{\text{all } k \text{ people have distinct birthdays}\}.$$

Example 1.2-5 (Birthday problem) In a group of k people, what is the probability that at least 2 people will have the same birthday? Assume $n = 365$ and that birthdays are equally distributed throughout the year, no twins, etc.

Solution.

1. Since there are 365 possible birthdays for each of k people, the sample space S will contain 365^k outcomes, all of which will be equally probable.
2. If $k > 365$, there are not enough birthdays for every one to be different, and hence at least two people must have the same birthday. (*Pigeonhole principle*)
3. So, we assume below that $k \leq 365$.
4. Denote

$$A_k = \{\text{at least 2 people have the same birthday in a group of } k \text{ people}\}$$

and hence,

$$A_k' = \{\text{all } k \text{ people have distinct birthdays}\}.$$

Example 1.2-5 (Birthday problem) In a group of k people, what is the probability that at least 2 people will have the same birthday? Assume $n = 365$ and that birthdays are equally distributed throughout the year, no twins, etc.

Solution.

1. Since there are 365 possible birthdays for each of k people, the sample space S will contain 365^k outcomes, all of which will be equally probable.
2. If $k > 365$, there are not enough birthdays for every one to be different, and hence at least two people must have the same birthday. (*Pigeonhole principle*)
3. So, we assume below that $k \leq 365$.
4. Denote

$$A_k = \{\text{at least 2 people have the same birthday in a group of } k \text{ people}\}$$

and hence,

$$A_k' = \{\text{all } k \text{ people have distinct birthdays}\}.$$

Example 1.2-5 (Birthday problem) In a group of k people, what is the probability that at least 2 people will have the same birthday? Assume $n = 365$ and that birthdays are equally distributed throughout the year, no twins, etc.

Solution.

1. Since there are 365 possible birthdays for each of k people, the sample space S will contain 365^k outcomes, all of which will be equally probable.
2. If $k > 365$, there are not enough birthdays for every one to be different, and hence at least two people must have the same birthday. (*Pigeonhole principle*)
3. So, we assume below that $k \leq 365$.
4. Denote

$$A_k = \{\text{at least 2 people have the same birthday in a group of } k \text{ people}\}$$

and hence,

$$A_k' = \{\text{all } k \text{ people have distinct birthdays}\}.$$

Example 1.2-5 (Birthday problem) In a group of k people, what is the probability that at least 2 people will have the same birthday? Assume $n = 365$ and that birthdays are equally distributed throughout the year, no twins, etc.

Solution.

1. Since there are 365 possible birthdays for each of k people, the sample space S will contain 365^k outcomes, all of which will be equally probable.
2. If $k > 365$, there are not enough birthdays for every one to be different, and hence at least two people must have the same birthday. (*Pigeonhole principle*)
3. So, we assume below that $k \leq 365$.
4. Denote

$$A_k = \{\text{at least 2 people have the same birthday in a group of } k \text{ people}\}$$

and hence,

$$A'_k = \{\text{all } k \text{ people have distinct birthdays}\}.$$

Solution(continued).

5. Counting the number of outcomes in A_k is tedious. However, the number of outcomes in S for which all k birthdays will be different, namely, the number of outcomes in A_k', is easy.
6. Indeed, $\mathcal{N}(A_k') = {}_{365}P_k$, since the first person's birthday could be any one of the 365 days, the second person's birthday could then be any of the other 364 days, and so on.
7. Hence,

$$\mathbb{P}\left(A_k'\right) = \frac{{}_{365}P_k}{365^k},$$

and therefore,

$$\mathbb{P}(A_k) = 1 - \mathbb{P}\left(A_k'\right) = 1 - \frac{{}_{365}P_k}{365^k} = 1 - \frac{365!}{(365-k)!\,365^k}.$$

$\square$

Solution(continued).

5. Counting the number of outcomes in A_k is tedious. However, the number of outcomes in S for which all k birthdays will be different, namely, the number of outcomes in A'_k, is easy.
6. Indeed, $\mathcal{N}(A'_k) = {}_{365}P_k$, since the first person's birthday could be any one of the 365 days, the second person's birthday could then be any of the other 364 days, and so on.
7. Hence,

$$\mathbb{P}\left(A'_k\right) = \frac{{}_{365}P_k}{365^k},$$

and therefore,

$$\mathbb{P}(A_k) = 1 - \mathbb{P}\left(A'_k\right) = 1 - \frac{{}_{365}P_k}{365^k} = 1 - \frac{365!}{(365-k)!\,365^k}.$$

□

Solution(continued).

5. Counting the number of outcomes in A_k is tedious. However, the number of outcomes in S for which all k birthdays will be different, namely, the number of outcomes in A_k', is easy.
6. Indeed, $\mathcal{N}(A_k') ={}_{365} P_k$, since the first person's birthday could be any one of the 365 days, the second person's birthday could then be any of the other 364 days, and so on.
7. Hence,

$$\mathbb{P}\left(A_k'\right) = \frac{{}_{365}P_k}{365^k},$$

and therefore,

$$\mathbb{P}(A_k) = 1 - \mathbb{P}\left(A_k'\right) = 1 - \frac{{}_{365}P_k}{365^k} = 1 - \frac{365!}{(365-k)!\,365^k}.$$

$\square$

Solution(continued).

5. Counting the number of outcomes in A_k is tedious. However, the number of outcomes in S for which all k birthdays will be different, namely, the number of outcomes in A_k', is easy.
6. Indeed, $\mathcal{N}(A_k') =_{365} P_k$, since the first person's birthday could be any one of the 365 days, the second person's birthday could then be any of the other 364 days, and so on.
7. Hence,

$$\mathbb{P}\left(A_k'\right) = \frac{_{365}P_k}{365^k},$$

and therefore,

$$\mathbb{P}(A_k) = 1 - \mathbb{P}\left(A_k'\right) = 1 - \frac{_{365}P_k}{365^k} = 1 - \frac{365!}{(365-k)!\,365^k}.$$

□

k	2	3	10	20	60
$\mathbb{P}(A_k)$	$\frac{1}{365}$	$\frac{1093}{133225}$	$\frac{2689423743942044098153}{22996713557917153515625}$	–	–
Approx.	0.002704	0.008204	0.1169	0.4114	0.9941

Sampling without replacement

Sample a subset of size r from n different objects, if we aren't concerned with order, the number of subsets $= {}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$.

Definition 1.2-3 (Combinations) Each of the ${}_nC_r$ unordered subsets is called a *combination of n objects taken r at a time*, where

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}. $$

Example 1.2-6 A committee of 5 persons is to selected randomly from a group of 5 men and 10 women.

(a) Find the probability that the committee consists of 2 men and 3 women.

Example 1.2-6 A committee of 5 persons is to selected randomly from a group of 5 men and 10 women.

(b) Find the probability that the committee consists of all women.

Theorem 1.2-2 (Binomial theorem) For $n \geq 0$, it holds that

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^{n-r} y^r.$$

$n=0$					$\binom{0}{0}=1$					1
$n=1$				$\binom{1}{0}=1$		$\binom{1}{1}=1$				$x+y$
$n=2$			$\binom{2}{0}=1$		$\binom{2}{1}=2$		$\binom{2}{2}=1$			$x^2+2xy+y^2$
$n=3$		$\binom{3}{0}=1$		$\binom{3}{1}=3$		$\binom{3}{2}=3$		$\binom{3}{3}=1$		$x^3+3x^2y+3xy^2+y^3$
$n=4$	$\binom{4}{0}=1$		$\binom{4}{1}=4$		$\binom{4}{2}=6$		$\binom{4}{3}=4$		$\binom{4}{4}=1$	$x^4+4x^3y+6x^2y^2+4xy^3+y^4$

Example 1.2-7 Prove $\sum_{r=0}^{n}(-1)^r \binom{n}{r} = 0$ and $\sum_{r=0}^{n} \binom{n}{r} = 2^n$.

Split objects into m groups of various sizes:

Suppose that in a set of n objects, n_1 are similar, n_2 are similar, ..., n_m are similar, where $n_1 + n_2 + \cdots + n_m = n$. Then the number of distinguishable permutations of the n objects is

$$\binom{n}{n_1, n_2, ..., n_m} = \frac{n!}{n_1! n_2! \cdots n_m!}$$

These numbers are called multinomial coefficient because of the following multinomial expansion:

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{k_1+k_2+\cdots+k_m=n} \binom{n}{k_1, k_2, \cdots, k_m} x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m}.$$

Example 1.2-8 20 members of a club need to be split into 3 committees (A, B, C) of 8, 8, and 4 people, respectively. How many ways are there to split the club into these committees?

Example 1.2-9 Suppose that three runners from team A and three runners from team B participate in a race. If all six runners have equal ability and there are no ties, what is the probability that the three runners from team A will finish first, second, and third, and the three runners from team B will finish fourth, fifth, and sixth?

Exercises from textbook: Section 1.2: 1, 3, 4, 5, 7, 8, 9, 11, 16, 17.