School of Mathematics and Statistics

Probability Crash Course: Permutations and Combinations

Paul Hewson

Overview: This webfile is designed as a revision aid to some introductory concepts in probability. It is intended to supplement a formal encouter with a text book or a set of lectures. These notes are meant to be slightly interactive, mysterious green dots, squares and boxes appear which you can click on to answer questions and check solutions.

Last Revision Date: 7th September 2009
paul.hewson@plymouth.ac.uk
Version 0.1

1. Overview

We've spent a lot of time looking at dice and coins. If we have 3 dice, each can show one of 6 faces. The number of possible outcomes are given by $n_1 \times n_2 \times n_3 = 6^3$. We would say each outcome can occur with probability $\frac{1}{6^3}$

But more generally, what about an experiment with equally likely outcomes, the problem of finding the probability of an event reduces to counting the number of outcomes in the event and in the sample space. But with realistic numbers this can get quite tedious. Consider the manager of a small plant wishes to determine the number of ways which he can assign staff to the first shift. There are 15 staff who can serve as operators of the production equipment, 8 who can serve as maintenance personnel and 4 who can be supervisors.

If a shift requires one operator, one maintenance person and one supervisor, how many different ways can the shift the staffed? The answer is 480, which is rather tedious if one had to draw the tree diagram.

1.1. N-tuple

Definition 1 *An N-tuple is a finite ordered set with an unspecified number of members*

For example, a 5-tuple is an ordered set with 5 members.

Suppose that each outcome (i.e., element) of A takes the form of an n-dimensional vector (or n-tuple) such as $(a_1, a_2, \dots a_n)$. If there are N_1 objects that can be used for a_1, and N_2 for a_2 then:

$$n(A) = N_1 N_2 \dots N_n \tag{1}$$

assuming that sets $A_1, A_2, \dots A_k$ have respectively $n_1, n_2, \dots, n_k$ different ways in which one can first take an element of A_1, then an element of A_2 and so on.

So for our factory problem, we have $N_1 = 15$, $N_2 = 8$ and $N_3 = 4$, with $15 \times 8 \times 4 = 480$. This is rather an intuitive result, but it seems too convenient to be true that we needed exactly one maintenance person, one production person and one supervisor.

We consider two scenarios:

- Permutations (the order of the elements is important)
- Combinations (the order of the elements is not important)

1.2. Permutation

Given a population of n distinct elements, how many ordered samples of size r can be drawn:

(i) with replacement: n^r

(ii) without replacement $n(n-1)(n-2)\ldots(n-r-1)$, calculation reduces to: $\frac{n!}{(n-r)!}$

Begin Quiz For a 6 digit telephone number,

1. how many different telephone numbers can be made if each digit can be 0,1,2,3,4,5,6,7,8,9:

2. by how much is this reduced if the first digit cannot be 0:

End Quiz Correct Points: Score:

Click on the Ans box to get the correct answer; shift + click to get the solution.

1.3. Combinations

Two sets are regarded as disjoint if and only if the properties of the membership are mutually exclusive. A *partition* is defined as a set of disjoint and exhaustive subclasses of a given class that is divided in such a way that each member of the given class is a member of exactly one such subclass.

The number of ways a set of n objects can be partitioned into 2 distinct subsets of size r and $n - r$ is given by:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{2}$$

Begin Quiz

1. The UK National Lottery requires that you correctly choose 6 numbers out of possible 49. How many ways re there is selecting these numbers;

2. In a communications channel, messages are transmitted in the form of a sequence of 8 binary digits. How many distinct messages are there containing 5 zeros and 3 ones?

End Quiz Correct Points: Score:

Click on the Ans box to get the correct answer; shift + click to get the solution.

1.4. Partitioning into k-subsets

The number of ways a set of n objects can be partitioned into k distinct subsets, where set 1 has r_1 elements, set 2 has r_2 elements, ..., set k has $_k$ elements is:

$$\frac{n!}{r_1!r_2!\ldots r_k!} \tag{3}$$

Begin Quiz

1. Consider sampling from 6 Midwestern states, $\Omega =$(Iowa, Illinois, Wisconsin, Michigan, Indiana and Ohio). For the purposes of a study, we need a sample of 4 states. How many samples of size 4 are possible.

2. How many samples of size 4 contain the state of Illinois.

End Quiz Correct Points: Score:

2. Pascal's Triangle

It's not worth losing sleep over, but it's impossible at this stage not to mention Pascal's[1] triangle:

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1

Note that each number in the triangle is the sum of the two numbers either side of it in the row above. Now consider you have a pool of 7 objects, and wish to know how many different ways there are of arranging them into sets of 3. We used above $\binom{n}{r} = \binom{7}{3} = \frac{7!}{3!(7-3)!} = 35$. However, if we number the rows of Pascal's triangle from 0 (at the top) to 7, take row 7 and count along to the 4th entry, we have 35. Isn't that cute?

[1] apparently discovered by Jia Xian in 1050, published by Zhu Shije in 1303, discussed by Cardano in 1570

3. A fun exercise

This really is worth doing!

Quiz Birthdays How many students must there be in a class in order for there to be a 50% chance that two will share the same birthday

(a) 22 (b) 37 (c) 183 (d) 365

Maybe you should check this out. Can you get the dates of birth for all your fellow students (or is this a silly question)? If so, how many coincidences are there?

Calculating the paradox.

Note that there are 365^n possible birthdays in a set of n students (we ignore leap years and assume birthdays are independently distributed throughout the year).

We have $\frac{365!}{(365-n)!}$ possible permutations.

The probability of 2 coincident birthdays is therefore given by

$$\frac{365!}{(365-n)!365^n}$$

This can be rather a difficult computation (365! is rather large and we need to use logarithms).

There is an inbuilt **R** function, we can use this to plot the probabilities for a range of class sizes:

```
## create a storage vector for the results
pb <- vector("numeric",60)
## loop and calculate probability of coincidence
## for different class sizes
for (i in 1:60){
 pb[i] <- pbirthday(i)
 }
## plot
plot(c(1:60), pb, type = "l", xlab = "Class size",
  ylab = "Prob of coincidence", main = "Birthdays")
```

What do you notice about the curve? Do you get a different curve if you calculate this for yourself (in either R or Python) using a function with:

$$1 - exp\{\log(factorial(365)) - \log(factorial(365 - i)) - i \times \log(365)\}$$

Checking the **R** helpfile (?pbirthday) may be informative. What happens if you look for the probability that three or more people share a birthday `pbirthday(i,coincident=3)`

Solutions to Quizzes

Solution to Quiz: The answer is given by the first definition (we are sampling with replacement, as all the digits can be used in any position) hence we want 10^6

Click on that green button to return to the quiz $\longrightarrow$

Solution to Quiz: For five positions we have simply 10^5, but there are only 9 possibilities for the first digit hence the total number available is $10^5 \times 9 = 900,000$. This is $100,000$ less than the previous answer.

Click on that green button to return to the quiz →

Solution to Quiz: We need to calculate $\frac{49!}{6!(49-6)!}$

Click on that green button to return to the quiz $\longrightarrow$

Solution to Quiz: We need to calculate $\frac{8!}{5!(8-5)!}$

Click on that green button to return to the quiz $\longrightarrow$

Solution to Quiz: We want $\binom{n}{r} = \binom{6}{2} = \frac{6!}{(6-4)!4!} = 15$ ■

Solution to Quiz: Correct answer is 10. ■