

# Chapter 2: Know your math

S520

These notes are written to accompany Trosset chapter 2.

## Sets

*Read: Trosset pp. 23–26*

Mathematical probability is based on an extension of set theory called measure theory. Measure theory is important if you want to do a Ph.D. in Statistics, but for now, set theory will do.

We use  $S$  to denote the universe of possible objects. For a set  $A$ , the **complement** of  $A$  (written  $A^c$ ) is the set of objects *not* in  $A$ .

The **union** of sets  $A$  and  $B$ , denoted  $A \cup B$ , is the set of all objects in either  $A$  or  $B$  (or both.) We can think of this as “ $A$  or  $B$  happens.” With more than two sets, the union is the set of all objects in at least one of the sets.

The **intersection** of sets  $A$  and  $B$ , denoted  $A \cap B$ , is the set of all objects in both  $A$  and  $B$ . We can think of this as “both  $A$  and  $B$  happen.” With more than two sets, the intersection is the set of all the objects that are in *every one* of the sets.

Events  $A$  and  $B$  are **disjoint** if  $P(A \cap B) = 0$ . A set of events is **pairwise disjoint** if all pairs of events in the set are disjoint.

## Counting

*Read: Trosset pp. 27–32*

### The multiplication principle

(Trosset p. 29) Suppose there are two decisions to be made and that there are  $n_1$  possible outcomes of the first decision. If, for each outcome of the first decision, there are  $n_2$  possible outcomes of the second decision, then there are  $n_1 n_2$  possible outcomes of the pair of decisions.

(If you like proving things, try proving this by induction.)

### Urns and counting

The canonical model for random sampling is the urn (Trosset p. 18.)

1. Put a bunch of balls in an urn, one for each member of the population. Mix the balls around.
2. Draw one ball from the urn. By assumption, each ball has an equal chance of being drawn.
3. If *sampling with replacement*, put the ball back in the urn. If *sampling without replacement*, the ball remains outside the urn.
4. Repeat steps 2 and 3 until a sample of the desired size has been drawn.

The key idea: If we keep track of the order of the balls, then all possible samples are equally likely. We can use the multiplication principle to count the number of possible samples; the probability of a particular sample will be one divided by this count. Let's consider sampling with and without replacement separately.

## Sampling with replacement

Suppose there are  $N$  distinct balls in the urn.

If you draw one ball, clearly there are  $N$  equally outcomes, so each ball has a probability of  $1/N$ .

Now suppose we draw **twice** from the urn, with replacement. How do we count the equally likely outcomes?

Let's start with the case where  $N = 3$ . Suppose the balls are labeled A, B, C. Then the equally likely outcomes are:

AA, AB, AC, BA, BB, BC, CA, CB, CC

Note that we count AB and BA as distinct. This means:

$$P(\text{first ball A, second ball B}) = 1/9$$

$$P(\text{first ball B, second ball A}) = 1/9$$

$$P(\text{one A and one B, in some order}) = 1/9 + 1/9 = 2/9$$

In general, when sampling with replacement, there are  $N$  equally likely ways to draw 1 ball,  $N^2$  equally likely ways to draw 2 balls, and  $N^k$  equally likely ways to draw  $k$  balls (keeping track of order.)

## Sampling without replacement

Now suppose there are three balls in the urn, and we draw all three, without replacement. How many equally likely ways are there to do this?

- There are three ways to choose the first ball.
- For each way to choose the first ball, there are two ways to choose the second ball.
- For each way to choose the first two balls, there's only one way to choose the third ball (since there's only one ball left.)

In other words, every time you take a ball out, there's one fewer ball to choose from next time. With three balls, there are  $3 \times 2 \times 1 = 6$  ways to put them in order. With eight balls, there would be  $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$  ways of putting them in order. That was a lot to write down, so instead we write  $8!$ , pronounced "eight factorial." "Factorial" means we multiply all the counting numbers from that number down to one. (NB: There's a special case: We define  $0! = 1$ . How many ways are there of putting zero objects in order? There's precisely one way: Don't do anything.)

*Example.* There are six balls in an urn, labeled A through F. How many ways are there of drawing three balls if:

1. order matters?
2. order doesn't matter?
3. There are  $6 \times 5 \times 4 = 120$  ways of drawing 3 out of 6 objects in order.
4. Out of the 120 equally likely ways of drawing 3 out of 6 objects in order, how many of them contain ABC in some order? We can list them:

ABC, ACB, BAC, BCA, CAB, CBA.

So 6 of the 120 equally likely ways are ABC. Note that we could have counted these orderings without listing them: all we're doing is putting A, B, C in an urn and sampling three letters without replacement. Hence there are  $3 \times 2 \times 1 = 6$  ways.

But the same will be true of any selection of three different letters. We're thus overcounting by a factor of 6. The number of ways if order *doesn't* matter is  $120/6 = 20$ .

We generalize these ideas using **permutations** and **combinations** (Trosset p. 29–32):

A **permutation** is a ordered choice of objects. The number of ways of choosing  $r$  out of  $n$  objects in order is:

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

A **combination** is an unordered choice of objects. There are  $r!$  permutations for each combination of length  $r$ . So the number of ways of choosing  $r$  out of  $n$  objects if order doesn't matter is:

$$C(n, r) \equiv \binom{n}{r} = \frac{P(n, r)}{r!} = \frac{n!}{r!(n-r)!}$$

## Functions

*Read: Trosset ch. 2.3*

A function is just a rule: you put an input into a function, and you get an output. For exaxple, consider the following function:

*The input may be any real number. The output is the square of the output.*

So if the input is 2, the output is 4.

We can write this function down mathematically. Call the function  $\phi$ :

$$\phi(x) = x^2$$

The **domain** of the function (the set of allowable inputs) is the set of real numbers. The **range** of the function (the set of possible outputs) is the set of all non-negative real numbers.

We can also find the **inverse** of the function: Given an output, what could the input have been? We write the inverse of the function  $\phi$  as  $\phi^{-1}$ . For example:

$$\phi^{-1}(4) = \{-2, 2\}$$

In other words, if the output of  $\phi$  was 4, the input must have been either  $-2$  or  $2$ .

Functions are crucially important in this course for a couple of reasons.

1. We'll characterize the results of experiments by using functions.

e.g. Ask a randomly selected American whether they approve of ice cream. If they say yes, call that "1". If they say no, call that "0".

2. The software R is essentially organized around functions. For example, there's a `log()` function that takes logs of things:

```
log(5)
```

```
## [1] 1.609438
```

Note that logs in this course are natural (base  $e$ ) logs unless otherwise stated. We'll learn much more powerful functions throughout this course.