1. The Law of Averages

There is something called the Law of Averages(or the Law of Large Numbers) which states that if you repeat a random experiment, such as tossing a coin or rolling a die, a **very large number of times**, (as if were a population) your outcomes, when averaged, should be equal to (or very close to) the theoretical average (a parameter).

Freedman uses the story of John Kerrich to clear up some myths about this "law" and he provides a quote from Bertrand "The roulette wheel has neither conscience nor memory". Think about his quote and then consider this situation:

If you have ever visited a casino in Las Vegas and watched people play roulette, when gamblers see a streak of "Reds" come up, some will start to bet money on "Black" because they think the law of averages means that "Black" has a better chance of coming up now because they have seen so many "Reds" show up. While it is true that in the LONG RUN the proportion of Blacks and Reds will even out, in the short run, anything is possible. So it is wrong to believe that the next few spins will "make up" for the imbalance in Blacks and Reds.

The roulette wheel has no memory (and no conscience) so it has no idea that the last say, 10 spins resulted in "Red". The chance is the same that it will land on "Red" on the 11^{th} spin. Eventually in the long run (over thousands upon thousands of spins) it will even out – but remember, we live in the short run.

Basically then, if we think of each spin (or flip or whatever you are studying) is a "trial". The larger the number of trials, the more likely it is that the overall fraction of successes will be close to the probability, p, of success in a single trial. Also with more trials: You are likely to miss the expected number of outcomes by a larger amount as measured by raw numbers, but you are likely to miss by a smaller amount in terms of percentages.

Example: A family has 7 children, all girls, and they really want a boy. Some people think they should try again because the "law of averages is going to kick in and the chance that the will have boy is really high" while other people think this particular family has the same chance (50%) of having an 8^{th} girl. Who is right? Why?

Many random processes (e.g. drawing a sample from a population, rolling a die, having kids, your commute, traffic conditions) are subject to this law.

2. Chance Error & Chance Processes

When you flip a coin as Kerrich does in the first section of chapter 16, you are, in a sense, taking a **sample** of the number of heads. You can expect half of the tosses to result in heads (with our fair coin.) So the number of heads that you actually get will follow the formula:

Observed (actual) number of heads = half the number of tosses + chance error.

If you flip a coin twice, you don't get exactly one head each time. Sometimes you get none, sometimes two. There is chance variability in the process. (recall the previous lecture)

Another point of the narrative in section 1 is that while the absolute size of the chance error increases as the number of tosses goes up, it grows more slowly than the number of tosses. If you double the number of tosses, the size of the chance error goes up only by a factor of $\sqrt{2}$ .If you increase the number of tosses by a factor of 100, the chance error only goes up by a factor of 10. The chance error obeys a *square root law.* As a result, while the size of the chance error goes up as the tosses increase, the chance error as a percentage of the number of tosses grows smaller.

3. Box Models

Now let's generalize the coin tossing process (or any other random process) to one of drawing tickets from a box (a box model).

Suppose you draw, with replacement, so that each ticket will be equally likely to be drawn.

- You can think of flipping a coin as drawing from a box with two tickets, one marked 0 (tails) and one marked 1 (heads).
- You can think of playing roulette as drawing from a box with two tickets, one marked 0 (not Red) and one marked 1 (Red).
- You can think of having children as drawing from a box with two tickets, one marked 0 (Girl) and one marked 1 (Boy).
- And you can think of rolling a die as drawing from a box with six tickets, marked 1, 2, 3, 4, 5, and 6.

When setting up these *box models*, keep four things in mind:

1. What numbers go on the faces of the tickets in the box?

- For the coin toss and roulette, we used the number 1 to represent winning and 0 to represent losing.
- For kids, we used 1=having a boy 0=having a girl if we are interested in counting the number of boys. If we were interested in the number of girls, we would reverse it.
- For the die, we used a number for each face of the die. 1, 2, 3, 4, 5, 6
- If you were playing a lottery, perhaps one ticket would have $3,000,000, all of the others would have $0.

2. How many of each kind go into the box?

You might be thinking in terms of percentages or actually counts or proportions. Example: the game of "craps" is like a box that has 36 tickets with the numbers 2 through 12 on their faces and it's like there are: 1 – "2's", 2 – "3's", 3 – "4's", 4 "5's", 5 "6's" , 6 "7's", 5 "8's", 4 "9's", 3 "10's", 2 "11's" and 1 "12's"

The boxes with a coin or a child have 2 tickets, one of each type of outcome.

3. How many draws?

If we tossed the coin five times, we would have five draws. If we spun the wheel 10 times we would have 10 draws. If we were to sample 1000 births, we would have 1000 draws.

If we spent $5 on lottery plays, we would have five draws.

Note: When we make the same bet several times in gambling, we set up our box model as follows: The tickets in the box show the amounts you can win or lose. For example, tossing a coin for a dollar, the tickets would be +$1.00 and -$1.00. The chance for winning any particular amount should equal the chance you draw a ticket with that value from the box. For our coin toss, for a fair coin, you could have just one of each ticket.

4. What do I do with the numbers on the face of each ticket?

Do I sum them up or calculate an average or calculate a percentage (or proportion)?

Summary: The Box Model (16.4) is a tool to help you calculate sums from a random process