Math 109

Chapter Sixteen

Dr. P. K. Sahoo
Department of Mathematics
University of Louisville
The Law of Averages
John Kerrich, a South African mathematician, visited Copenhagen just before the second world war. Two days before he scheduled to fly back, the Germans invaded Denmark. Kerrich spent the rest of the war at a camp in Jutland.

To pass time he carried out a series of experiments in probability theory. One experiment involved in tossing a coin 10,000 times.
If you toss a coin, you will expect 50% times head and 50% tail since there is a 50-50 percent chance of getting either a head or a tail.

Thus, if you toss a coin 10,000 times, then you expect to get 5,000 heads (that is, the half the number of tosses).

However, you do not get exactly 5,000 heads. You only expect to get 5,000 heads. You could just as well get 5,001 or 4,998 or 5,007 heads. The amount off 5,000 is known as the chance error of the coin-tossing experiment.
Law of Averages

The law of averages was discovered by South African mathematician named John Kerrich during the second world war.

The law of averages says that the number of heads will be about half the number of tosses, give or take a chance error. As the number of tosses goes up, the chance error gets bigger. But compared to the number of tosses, it gets smaller.
The chance error of the coin-tossing experiment can be written as follows:

\[ \text{# of heads} = \text{half the number of tosses} + \text{chance error} \]

This is equivalent to saying

\[ \text{chance error} = \text{# of heads} - \text{half the number of tosses} \]
John Kerrich's coin-tossing experiment. The first column shows the number of tosses. The second shows the number of heads. The third shows the difference number of heads - half the number of tosses.

<table>
<thead>
<tr>
<th>Number of tosses</th>
<th>Number of heads</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4</td>
<td>-1</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>40</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>60</td>
<td>29</td>
<td>-1</td>
</tr>
<tr>
<td>70</td>
<td>32</td>
<td>-3</td>
</tr>
<tr>
<td>80</td>
<td>35</td>
<td>-5</td>
</tr>
<tr>
<td>90</td>
<td>40</td>
<td>-5</td>
</tr>
<tr>
<td>100</td>
<td>44</td>
<td>-6</td>
</tr>
<tr>
<td>200</td>
<td>98</td>
<td>-2</td>
</tr>
<tr>
<td>300</td>
<td>146</td>
<td>-4</td>
</tr>
<tr>
<td>400</td>
<td>199</td>
<td>-1</td>
</tr>
<tr>
<td>500</td>
<td>255</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of tosses</th>
<th>Number of heads</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>600</td>
<td>312</td>
<td>12</td>
</tr>
<tr>
<td>700</td>
<td>368</td>
<td>18</td>
</tr>
<tr>
<td>800</td>
<td>413</td>
<td>13</td>
</tr>
<tr>
<td>900</td>
<td>458</td>
<td>8</td>
</tr>
<tr>
<td>1,000</td>
<td>502</td>
<td>2</td>
</tr>
<tr>
<td>2,000</td>
<td>1,013</td>
<td>13</td>
</tr>
<tr>
<td>3,000</td>
<td>1,510</td>
<td>10</td>
</tr>
<tr>
<td>4,000</td>
<td>2,029</td>
<td>29</td>
</tr>
<tr>
<td>5,000</td>
<td>2,533</td>
<td>33</td>
</tr>
<tr>
<td>6,000</td>
<td>3,009</td>
<td>9</td>
</tr>
<tr>
<td>7,000</td>
<td>3,516</td>
<td>16</td>
</tr>
<tr>
<td>8,000</td>
<td>4,034</td>
<td>34</td>
</tr>
<tr>
<td>9,000</td>
<td>4,538</td>
<td>38</td>
</tr>
<tr>
<td>10,000</td>
<td>5,067</td>
<td>67</td>
</tr>
</tbody>
</table>
Kerrich's coin-tossing experiment. A plot of the "chance error" number of heads – half the number of tosses against the number of tosses. As the number of tosses goes up, the size of the chance error tends to go up. The horizontal axis is not drawn to scale.
One crude version of the "law of averages" is a simple principle: if you want the ordinary to happen, bet on the long run—if you want the extraordinary to happen, bet on the short run.
Example A: In tossing a fair coin, is it easier to get 64 or more heads in 100 tosses, or 640 or more heads in 1000 tosses, or are the chances the same?

Solution: Close to 50% heads is "ordinary"—64% or more is unusual. So your chances are better in the short run—it's easier to get it with 100 tosses.
Example B: A coin will be tossed either two times or 100 times. If the number of heads is exactly equal to the number of tails you will win $1. a) two tosses is better; b) 100 tosses is better; c) both are the same.

Solution: The answer is a). Although close to 50% heads is ordinary, exactly 50% heads is not—so choose the short run.

Example C: In tossing a fair coin, is it easier to get 60 or more heads in 100 tosses or 600 or more heads in 1000 tosses, or are the chances the same?
Solution: The answer is: 60 heads in 100 tosses.
Exercise Set A

1. A machine has been designed to toss a coin automatically and keep track of the number of heads. After 1,000 tosses, it has 550 heads. Express the chance error both in absolute terms and as a percentage of the number of tosses.

2. After 1,000,000 tosses, the machine in exercise 1 has 501,000 heads. Express the chance error in the same two ways.

3. A coin is tossed 100 times, landing heads 53 times. However, the last seven tosses are all heads. True or false: The chance that the next toss will be heads is somewhat less than 50%. Explain.

4. A coin will be tossed, and you win a dollar if the number of heads is exactly equal to the number of tails. Which is better: 10 tosses or 100 tosses? Or are they the same? Explain.

5. A coin will be tossed, and you win a dollar if the percentage of heads is between 40% and 60%. Which is better: 10 tosses or 100 tosses? Explain.

6. A coin is tossed, and you win a dollar if there are more than 60% heads. Which is better: 10 tosses or 100? Explain.
1. A machine has been designed to toss a coin automatically and keep track of the number of heads. After 1,000 tosses, it has 550 heads. Express the chance error both in absolute terms and as a percentage of the number of tosses.

**Answer**

\[
\text{Chance error} = \# \text{of heads} - \frac{1}{2} (\text{tosses})
\]

\[
= 550 - \frac{1}{2} (1000)
\]

\[
= 50
\]

\[
\text{Chance error in percentage} = \frac{50}{1000} \times 100
\]

\[
= 5 \%
\]
2. After 1,000,000 tosses, the machine in exercise 1 has 501,000 heads. Express the chance error in the same two ways.

**Answer**

\[
\text{Chance error} = \# \text{ of heads} - \frac{1}{2} (\text{tosses})
\]

\[
= 501,000 - \frac{1}{2} (1,000,000)
\]

\[
= 1000
\]

**Chance error in percentage**

\[
= \frac{1000}{1,000,000} \times 100 = \frac{1}{10} = 0.1\%
\]
3. A coin is tossed 100 times, landing heads 53 times. However, the last seven tosses are all heads. True or false: The chance that the next toss will be heads is somewhat less than 50%. Explain.

**Answer**  False. The last seven tosses are heads does not imply the eighth toss will be a tail since head and tail both have 50-50 chance.
4. A coin will be tossed, and you win a dollar if the number of heads is exactly equal to the number of tails. Which is better: 10 tosses or 100 tosses? Or are they the same? Explain.

Answer: The event of getting exactly number of heads equal to number of tails is an unexpected event. So go for the short run. Thus 10 tosses are better.
5. A coin will be tossed, and you win a dollar if the percentage of heads is between 40% and 60%. Which is better: 10 tosses or 100 tosses? Explain.

**Answer** 100 tosses.

As the number of tosses goes up, the percentage of heads is likely to be closer to 50%.

(This will be clear to understand after Chapter 18.)
Summary

Law of averages says that with a large number of tosses, the percentage of heads is likely to be close to 50%, although it is not likely to be exactly equal to 50%.

If you want the ordinary to happen, bet on the long run – if you want the extraordinary to happen, bet on the short run.
CHANCE PROCESSES

- Tossing a coin and observing the number of heads
- Rolling a die and observing a particular number on the die
- Spinning a roulette wheel
- Selecting a sample from a population

are examples of chance processes. A chance process is studied by using a box model.
An Example

Suppose there is a box of tickets, and each ticket has a number written on it. Then some tickets are drawn at random from the box, and the numbers on these tickets are added up. For example, take the box

\[ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \]
First draw one ticket from the box at random and note the number on the ticket. Then put the ticket back on the box and draw another ticket from the box and again note the number on it. Next you add up the two numbers.
The first draw might be 3 and the second draw might be 5, then the sum of the draws is 8. Or the first draw might be 3 and the second draw might be 3 too, then the sum of the draws is 6. There are many other possibilities. So the sum is subject to chance variability (that is, if the draws turn out one way, the sum is one thing; if they turn out differently, the sum is different too.)
Imagine taking 25 draws with replacement from the box

\[1 \ 2 \ 3 \ 4 \ 5 \ 6\]

How big is their sum going to be? We can find this out by doing a computer simulation. Computer simulation gives the following result

88 84 80 90 83 78 95 80 89

Chance variability is easy to see. The first sum is 88, the second sum is 84, the third drops to 80. The values range from 78 to 95.
In principle, the sum could have been as small as 25 or as high as 150. But in fact, the 10 simulated values are between 75 and 100.

Would this keep up with more repetitions? We will address this question in the next lecture.
Rolling a die is just like picking a number from the box.
BOX MODEL

The analogy between a chance process and drawing from a box is called a box model. There are three questions to answer when you are making a box model. These questions are:

- What numbers go into the box?
- How many of each kind?
- How many draws?
The answer to first question tells us what goes into the box. The answer to second question tells us how many of each number go into the box. The answer to third question tells us how many draws are made from the box.
BOX MODEL CONTINUES

To decide what goes into the box, you ask yourself:

- What are the possible values a trial?
- What are the chances on these values?
- How many trials are there?
Example A: A coin is tossed 50 times and the number of heads is counted. What is the box model?
Answer:

(a) What are the possible values on one trial?

On one trial, the number of heads can be 0 or 1. So the box has to contain the tickets with numbers 0 and 1.
(b) What are the chances of these values?

In one toss, getting head is 1 in 2. So also getting a tail. Hence the box is

```
  1   0   1   4
```
(c) How many trials are there?

You are tossing the coin 50 times, so the number of trials or draws has to be 50. Hence box model is the following:

```
1 0 1 1
```

50 Draws
Example B: A multiple choice test has 100 questions. Each question has five answers. Suppose you guess at random on each question and your number of correct answers is counted. What is the box model?
Answer:

(a) What are the possible values on one trial?

On one trial, the number of correct guesses can be either 0 or 1. So the box has to contain the tickets with numbers 0 and 1.
(b) What are the chances of these values?

On each question, you have 1 chance in 5 of getting the correct answer. You have 4 chances in 5 of getting the answer wrong. Hence the box is

```
4 0 1 1
```
(c) How many trials are there?

You are answering 100 questions, so the number of trials or draws has to be 100. Hence box model is the following:
Example C: A coin is tossed 500 times and the quantity, "twice the number of heads minus the number of tails" is computed. What is the box model?
Answer:

(a) What are the possible values on one trial?

On one toss, "twice the number of heads minus the number of tails" can be either \((2 \times 1) - 1\) or \((2 \times 0) - 1\). So the box has to contain the tickets with numbers 1 and +1.

\[
\begin{array}{c}
\boxed{1} \\
\boxed{+1}
\end{array}
\]
(b) What are the chances of these values?

On one trial, you have 1 chance in 2 of getting the number 1, and you have 1 chance in 2 of getting the number −1. Hence the box is

```
  1 1 1 1
```
(c) How many trials are there?

You are tossing 500 times, so the number of trials or draws has to be 500. Hence box model is the following:

```
+---+---+---+
| 1 | -1 |
+---+---+---+
```

500 Draws
Example D: A roulette table has 38 slots, of which 18 are red. If you bet a dollar on a single number, and that number comes up, you get the $1 back together with winnings of $35. If any other number comes up, you lose the dollar. What is the box model for your winnings playing roulette 100 times, betting a dollar on the number 17 each time?
(a) What are the possible values on one trial?

On one trial, you bet a dollar on number 17. If the ball drop into the pocket 17, you will win $35. If the ball drops into any other pocket, you will lose $1. So the box has to contain the tickets with numbers $35 and $-1$.

| $35$ | $-1$ |
(b) What are the chances of these values?

You have only 1 chance in 38 of winnings; so the chance of drawing the ticket with number $35 has to be 1 in 38. You have 37 chances in 38 of losing. So the chance of drawing $-1 has to be 37 in 38. The box is

\[
\begin{array}{c|c|c|c}
1 & \$35 & 37 & -1 \\
\end{array}
\]

(c) How many trials are there?

You are playing 100 times, so the number of trials or draws has to be 100.
Box Model

1 35 37 -1

100 Draws
Problems

16.6 A coin is tossed 100 times and the quantity, "the number of heads minus the number of tails" is computed. What is the appropriate box model?

16.7 A coin is tossed 100 times and the quantity, "twice the number of heads minus the number of tails" is computed. What is the appropriate box model?

16.8 A coin is tossed 100 times and the quantity, "the number of heads plus the number of tails" is computed. What is the appropriate box model?

16.9 A roulette table has 38 slots, of which 18 are red. If you bet $1 on red and red comes up, you win $1—otherwise you lose $1. What is the box model for your winnings playing roulette 100 times, betting $1 each time?
Answer

16.6 a) On one trial (i.e. one toss), this quantity is either \( +1 - 0 = +1 \) (when a head is tossed), or \( 0 - 1 = -1 \) (when a tail is tossed). Now \(+1\) and \(-1\) have equal chances, so the box model is 100 draws from the box \([+1 \ -1]\).

16.7 On one toss, "twice the number of heads minus the number of tails" can be either \((2 \times 1) - 0 = 2\) or \((2 \times 0) - 1 = -1\). So the box model is 100 draws from the box \([+2 \ -1]\).

16.8 On one trial, this quantity is \( 1 + 0 = +1 \) (when heads is tossed), or \( 0 + 1 = +1 \) (when tails is tossed), so the box model is 100 draws from the box \([+1]\).

16.9 On one play you win either \$1\) or \(-\$1\). So the box contains only the numbers "+1" and "-1". One hundred draws are made. Here "+1" has an \(18/38\) chance of occurring, so the box contains \(18/38\) "+1's" and \(20/38\) "-1's". Hence 18 "+1's" and 20 "-1's" do the trick. One hundred draws are made.
Exercise Set B

1. One hundred draws are made at random with replacement from the box \( \{1, 2\} \). Forty-seven draws turn out to be \( 1 \), and the remaining 53 are \( 2 \). How much is the sum?

2. One hundred draws are made at random with replacement from the box \( \{1, 2\} \).
   (a) How small can the sum be? How large?
   (b) About how many times do you expect the ticket \( 1 \) to turn up? the ticket \( 2 \)?
   (c) About how much do you expect the sum to be?

3. One hundred draws are made at random with replacement from the box \( \{1, 2, 9\} \).
   (a) How small can the sum be? How large?
   (b) About how much do you expect the sum to be?

4. A hundred draws will be made at random with replacement from one of the following boxes. Your job is to guess what the sum will be, and you win $1 if you are right to within 10. In each case, what would you guess? Which box is best? worst?

   (i) \( \{1, 9\} \)   (ii) \( \{4, 6\} \)   (iii) \( \{5, 5\} \)
8. A box contains 20% red marbles and 80% blue marbles. A thousand marbles are drawn at random with replacement. One of the following statements is true: which one, and why?
   (i) Exactly 200 marbles are going to be red.
   (ii) About 200 marbles are going to be red, give or take a dozen or so.

9. Repeat exercise 8, if the draws are made at random without replacement and the box contains 50,000 marbles.

10. A hundred tickets will be drawn at random from one of the two boxes shown below. On each draw, you will be paid the amount shown on the ticket, in dollars. (If a negative number is drawn, that amount will be taken away from you.) Which box is better? Or are they the same?

(i) \[\begin{bmatrix} -1 & -1 & 1 & 1 \end{bmatrix}\]

(ii) \[\begin{bmatrix} -1 & 1 \end{bmatrix}\]