

**SMMG November 4<sup>th</sup>, 2006**  
**featuring Dr. Irene Gamba**  
**“On billiards and time irreversibility...  
the legacy of Ludwig Boltzmann”**



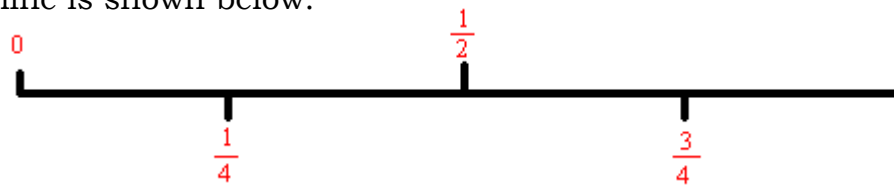
### 1. What is probability?

Have you ever wondered what the chance is that one of us might be struck by lightning or win the lottery? The likelihood or chance of something happening is called its **probability**.

Here is a way of showing this on a line:



We can say that the probability of an event occurring will be somewhere between impossible and certain. As well as words we can use fractions or decimals to show the probability of something happening. Impossible is **zero** and certain is **one**. A fraction probability line is shown below.



We can show on our line the chance that something will happen. Remember the probability of an event will **not** be more than 1. This is because 1 is certain that something will happen.

- a) Order the following events in terms of likelihood. Start with the least likely event and end with the most likely event.
- If you flip a coin it will land heads up.
  - You will not have to learn math at school.
  - You randomly select an ace from a regular deck of 52 playing cards.
  - There is a full moon at night.
  - You roll a die and 6 appears.
  - A politician fulfills all his or her campaign promises.
  - You randomly select a black card from a regular deck of 52 playing cards.
  - The sun will rise tomorrow.
  - If you have a choice of red, yellow, blue or green you will choose red.

Mathematically, we can define the probability of an event E as

$$P(E) = \frac{\text{the number of ways an event can occur}}{\text{the total number of possible outcomes}}$$

For example, if we roll a die, the probability of rolling a 2 is 1 out of 6 ( $1/6$ ), because there is only one side with 2 written on it and there are 6 total choices, the numbers 1 through 6 on each side.

Note that an event either happens or it does not happen. Thus:

$$\begin{aligned} & \text{the probability that an event E will happen} \\ &= 1 - (\text{the probability that the event E will not happen}) \end{aligned}$$

For example, if the probability of rolling a 2 is  $1/6$ , what is the probability of not rolling a 2? Now we have 5 choices that are not 2 out of 6 total choices. So the probability of not rolling a 2 is equal to  $5/6$ . This can also be written using our equation:

$$\begin{aligned} \text{Probability of not rolling a 2} &= 1 - \text{the probability of rolling a 2} \\ &= 1 - (1/6) \\ &= 5/6 \end{aligned}$$

Now what is the probability of rolling an even number, that is a 2, a 4 or a 6? This obviously should be more likely than rolling a 2 for sure. In fact,

$$\begin{aligned} & \text{the probability of one events OR another event occurring} \\ &= \text{sum of the individual probabilities} \end{aligned}$$

So the probability of rolling an even number on a fair die is  $1/6 + 1/6 + 1/6 = 1/2$ , which clearly makes sense since half the numbers on a dice are even.

We might want to know how likely it is that one thing happens and some unrelated thing also happens.

$$\begin{aligned} & \text{the probability of two events BOTH occurring} \\ &= \text{product of the individual probabilities} \end{aligned}$$

Suppose we have a die and a coin, and we are going to throw them both. Then the above becomes relevant if you want to know, for example, the probability of the outcome having a 4 on the die and heads on the coin. Using the above multiplication principle, you find that it is  $1/6 \times 1/2 = 1/12$ .

b) Find the answer to the following questions:

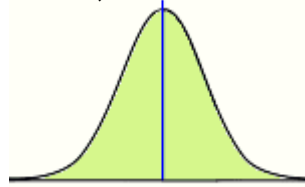
- Given a fair die, what is the probability of rolling an even number?
- If you toss a fair sided coin, what is the probability of tossing a heads?
- If you two fair sided coins, what is the probability of tossing a heads on the first coin and tails on the second?
- A fair sided cube with faces numbered 1 through 6 was tossed 300 times. Which is a reasonable prediction of the number of times the cube will land with the number 4 facing up?

- A. 25
- B. 50
- C. 100
- D. 150
- E. 200

- Jonas has a bag with exactly 18 marbles: 6 black, 8 white, 4 yellow. If he reaches into the bag and takes out a marble, what is the probability that it will be blue?
- Each of the 7 digits of the number 2,040,210 is written on a separate piece of paper and placed in a bag. If one piece of paper is selected from the bag at random, which digit would most likely be written on it?

## 2. The Normal Distribution

Other names for this distribution are the “**Gaussian distribution**,” (named after the German mathematician Karl Friedrich Gauss) and the “**bell curve**.”



Some terminology:

The value in the middle of the distribution, right where the hump is, is the **mean** (average) of the distribution. The **standard deviation** is a measure of how spread out the data is from the mean. The formula is easy: it is the square root of the variance. So now you ask, “What is variance?” The **variance** is defined as the average of the squared differences from the mean.

The normal distribution is one of the most commonly observed and is the starting point for modeling many types of real-world data, including the following:

- The **distribution of human heights**, if you consider only one gender at a time. Most people are about average height, a little bit taller than average, or a little bit shorter than average. This accounts for the “hump” in the middle of the distribution. Fewer people are very short or very tall, which accounts for the two tails on either side of the hump.
- the distribution of where the balls will fall on a Galton board;
- IQ test scores;
- some standardized test scores.



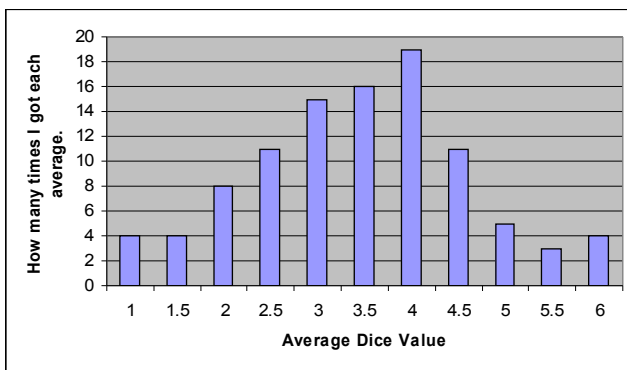
The Former German Money

10 Mark note has Karl F. Gauss printed on the back, and a small bell shaped curve and its math formula in the background

Another interesting fact about the normal distribution is that if you take a large number of samples from almost any sort of population, and find the sample means (the average value of each sample), then the means will form a normal distribution. This is something mathematicians call the “**central limit theorem**.” Here’s an example you can try at home:

1. Roll two dice and find the average value. For example, if you roll a 1 and a 4, then to find the average you would do:  $(1+4)/2 = 2.5$
2. Repeat #1 at least 10 times, but more if you can. Make sure you write down the average value each time.
3. Finally, make a histogram of your average. If you did enough dice rolls, you should start to see something that looks like a normal distribution.

I did it 100 times, and my histogram looks like this.



It’s slightly skewed, but it’s starting to look like a normal distribution already. Pretty neat!

### What is a histogram?

A histogram is a special kind of chart that shows how many of your data points fall within certain categories. The categories could be numbers, like in the example above, but they don’t have to be. For example, you could ask everyone in your math class what their favorite color is, and then create a histogram showing how many people prefer each color.

As an example, here is how I created my histogram above.

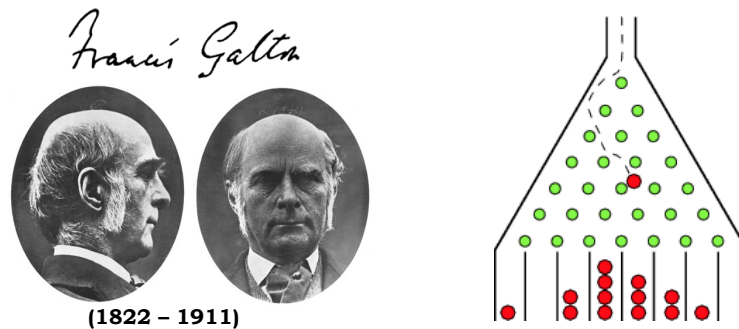
1. First, I collected my “average dice value” data
2. Then I created this table, which lists how many times I got each value in my data:

Category	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
How many times I got it.	4	4	8	11	15	16	19	11	5	3	4

3. Finally, I created my graph with the category labels on the x-axis. The height of each bar shows how many times I got each average value.
4. In this example, I used categories of width one-half, but there is no set rule for deciding how large your categories should be. Sometimes you just have to try out different sizes and see what works best in your situation.

### 3. Quincunx or Plinko at “The Price is Right!”

The Quincunx, probably better known as Galton's Board or Bean machine, is a device for statistical experiments named after the English scientist Sir Francis Galton.



It consists of a vertical (or inclined) board with a series of evenly spaced pegs arranged in a triangle in its upper half and a lower half divided into a number of evenly-spaced rectangular bins placed below the last row of pegs. Now, a quincunx board starts with 1 peg on the first row, 2 pegs on the second row, 3 pegs on the third row, and so on. Multiple balls are then dropped onto the top peg.

As each ball bounces its way down to the bottom, it strikes the pegs. If the board is built well, at each peg, half of the time the ball will fall to the right and half of the time the ball will fall to the left until it finally lands in one of the numbered (or lettered) bins.

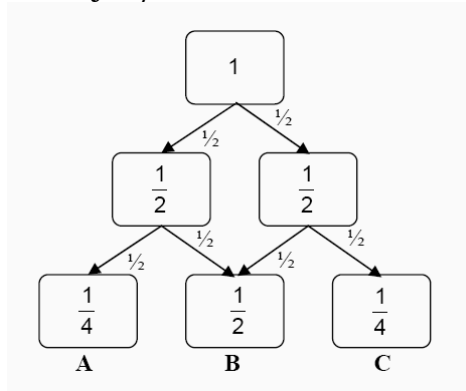
When counting the number of balls in each bin, one notices that there are more balls in the center bin than there are in the outer bins. More precisely, we have seen as part of the earlier presentation that if the number of balls is sufficiently large, the distribution of the heights of the ball heaps will approximate a [normal distribution](#)!

#### Formula

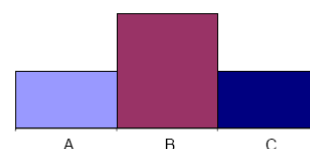
Let's investigate this a little further.

You can actually calculate probabilities here!

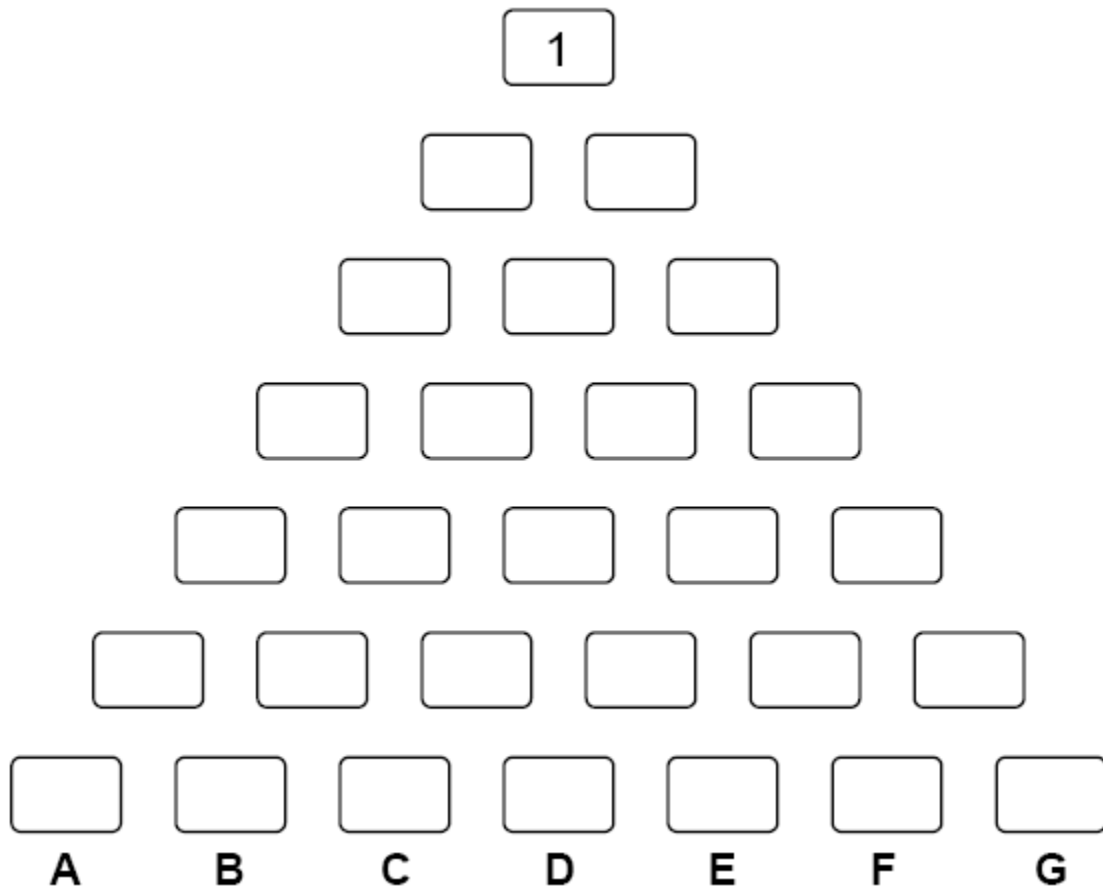
The theoretical probability of a ball following a particular path can be represented using an arrangement of boxes as shown below. At each peg, the probability is multiplied by  $1/2$ . Notice that the sum of each row of boxes equals one!



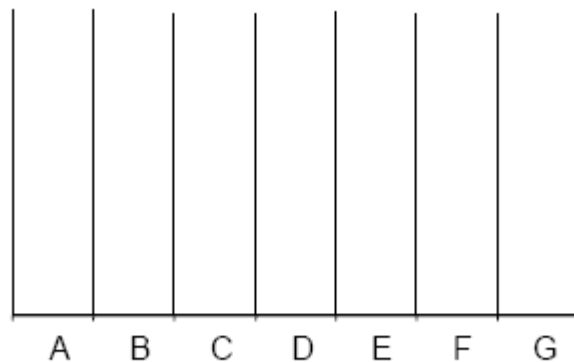
The distribution of cells A, B and C approximate a normal curve:



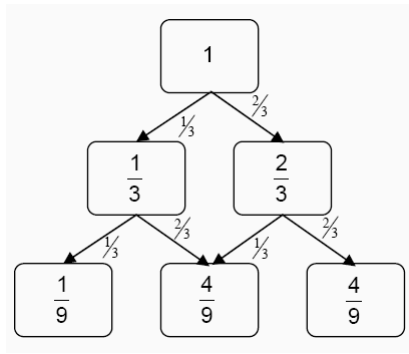
- a) The diagram below represents a Galton board with 7 rows of pins. In each box, calculate the probability that a ball will fall in that direction. Remember to use a probability of  $\frac{1}{2}$  for each direction, and remember that the sum of each row must equal 1.



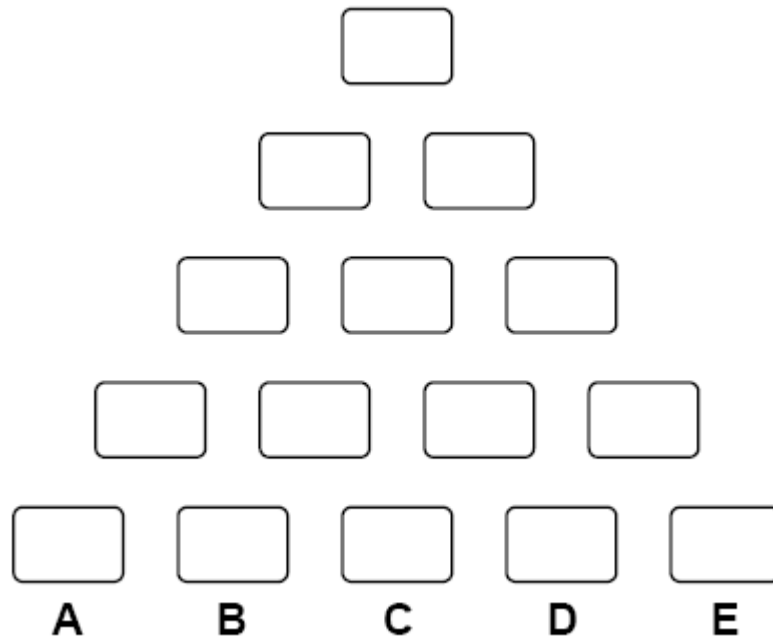
- b) Create a distribution diagram for bins A through G. What conclusions can you draw from this distribution?



- c) Now you have explored basic Galton Boards. Consider another variation, where the board would favor some outcomes more than others. For example, on the board below, it is assumed that the probability of the ball falling to the left is  $1/3$  and the probability of the ball falling to the right is  $2/3$ .



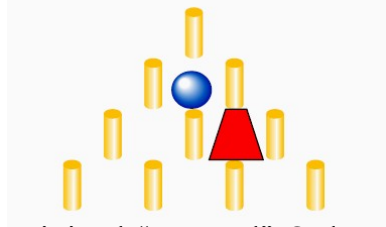
Calculate the probabilities for the biased Galton Board through five rows. Then graph the resulting distribution.



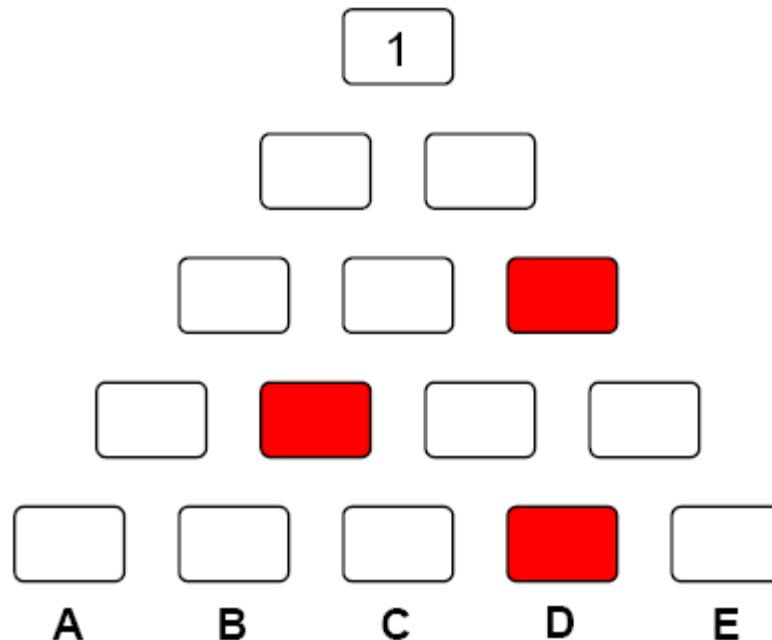
What do you notice about the symmetry of the distribution compared to an unbiased Galton Board?



- d) Consider another interesting scenario: What would happen if some of the paths were obstructed?



Going back to our original “normal” Galton Board with equal probability of bouncing left or right, look at the situation below. If a peg is blocked like in the above picture, the number of possible paths for the ball changes; the ball can now move only to the left. This means the probability that the ball will move to the left is now 1, instead of  $1/2$ .



Calculate the probabilities for this Galton Board with obstructions.  
Are there any bins in which the ball can no longer land?

#### Additional Resources

- A web site with an applet that shows a Galton board in action can be found at: <http://www.math.uah.edu/stat/applets/GaltonBoardExperiment.xhtml>
- This web site <http://galton.org/> is dedicated to the work of Sir Francis Galton, for whom the Galton board is named.
- On the game show The Price is Right, a game called “Plinko” uses a rectangular Galton Board (so you get to choose where you drop the ball) to determine what amounts of money a contestant will win. Check out a “Plinko” simulation at <http://mathdemos.gcsu.edu/mathdemos/plinko/bigboardplinko.html>

#### 4. **The Monty Hall Problem** - **“Let’s Make a Deal!”**

Suppose there are three boxes and in one of them is a grand prize while the others are empty. As a contestant, you will get to keep the prize if you pick the box that it is in. If you pick the wrong one, though, you’ll go home empty-handed.

But wait – the game is not done after your first pick! You will be given one chance to change your mind. After your first pick, we will open one of the boxes you did not choose and show you that it is empty. So you know for sure one box that does NOT contain the prize.

Your dilemma is now to decide whether you want to stick to your original choice or do you want to switch to the other remaining unopened box.

Where is the prize? How can you measure the uncertainty?

Let’s play the game a bunch of times!

Can you make a conjecture about what the best winning strategy is? Will you have a higher chance of winning if you stick to the original choice or if you switch?

Why do you think this is true? Discuss with your neighbor.

**Explore this more at home! You can use mathematics to find the best winning strategy!**

Here are some hints on how to do it:

- a) Suppose your initial guess was wrong. What would happen if you were to change your answer?
- b) Suppose instead of three boxes, there were ten. And suppose that after you guessed we were to open eight of the remaining ones that all nothing in them. Should you switch in this game? Why or why not?
- c) In order to convince yourself of what the right thing to do is, make a list of all possible ways the prize could be placed in the boxes.

Use probabilities!

Now play the game a bunch of times and keep track of the outcomes under the two possible strategies - the switch and the stick strategy. Perform the experiment about 50 times and record the results. Do the experimental data coincide with the analysis of the probabilities you did before?

## 5. The Birthday Problem

Since a year has 365 days (not counting February 29 in leap years), there would have to be 366 people gathered together in order for us to be absolutely certain that at least two people in the group have the same birthday. Why?



Here are some interesting questions, though:

- What if we were content to be just 50 percent certain of this? How many people would there have to be in a group in order for the probability to be half that at least two people in it have the same birthday?  
Make a guess!
- Among today's Saturday morning crowd, do you think there are two people that share a birthday?
- Do you think that there is someone else in the room that has the same birthday as your own? Is this more or less likely than the previous event? Why?

First try a warm-up exercise of explicitly finding the probability that two given people share the same birthday. The first person has some birthday, so the question could be rephrased: What is the probability that the second person has the same birthday as the first?

Answer: \_\_\_\_\_

Now choose two people at random. What is the probability that they were both born on December 31?

Answer: \_\_\_\_\_

Finally, back to the earlier birthday question: How many people would there have to be in a group in order for the probability to be half that at least two people in it have the same birthday?

Here comes the BIRTHDAY SURPRISE!

In a room containing only 23 people, the chance of two people sharing the same birthday is just over 50 %. That is, in a random gathering of 23 people, we will more often than not find a pair of people with the same birthday.

In a room with 183 people, the chance of finding a pair of people with the same birthday is over 99.999999%. Why?



Hint: By the multiplication principle, the number of ways in which five dates can be chosen (allowing for repetitions) is  $365 \times 365 \times 365 \times 365 \times 365$ . Of all of these  $365^5$  ways, however, only  $365 \times 364 \times 363 \times 362 \times 361$  are such that no two of the dates are the same. Thus by dividing this latter product by  $365^5$  we get the probability that five people chosen at random will have no birthday in common.

Now, if we subtract this probability from 1 (or from 100 percent) we get the complementary probability that at least two of the five people do have a birthday in common.

You need to do a similar calculation using 23 rather than 5.

## 6. Take Home Challenge: Pascal's Triangle

Think about Galton Board which is built in such a way that the probability of bouncing to the left is  $p$  (and therefore probability of bouncing right is  $1-p$ ).

A ball would end up in the bin  $k$  places from the right if it has taken  $k$  left turns. So if the Galton Board has  $n$  rows, a possible path for the ball would be  $k$  bounces to the left and  $(n-k)$  bounces to the right.

Since the probability of bouncing to the left is  $p$ , then we can calculate the probability of a certain path like this: The ball bounces  $k$  times to the left with a probability of  $p$ :  $p^k$ . And the other  $(n-k)$  bounces have the opposite probability of:  $(1-p)^{(n-k)}$ . So, the probability of following such a path is  $p^k(1-p)^{(n-k)}$ .

But there could be many such paths! For example the left turns could be the 1st and 2nd, or 1st and 3rd, or 2nd and 5th, etc.

You could list all such paths (LLRRR..., LRLRR..., LRRL...), but there is an easier way.

The number of ways of distributing  $k$  things in a sequence of  $n$  is  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

commonly called "n choose k". (The "!" means "factorial", for example  $4! = 1 \times 2 \times 3 \times 4 = 24$ .)

**So in fact, the Quincunx is just like Pascal's Triangle,  
with pegs instead of numbers. Amazing but true.**

Putting it all together, the resulting formula for the probability of the ball landing in the  $k$ th bin from the right is:

$$f(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

(Which, by the way, is the formula for the binomial distribution.)

**Example:**

For 10 rows ( $n=10$ ) and probability of bouncing left of 0.5 ( $p=0.5$ ), we can calculate the probability of being in the 3rd bin from the right ( $k=3$ ) as:

$$\text{also: } f(3; 10, 0.5) = \binom{10}{3} 0.5^3 (1 - 0.5)^{(10-3)} = \binom{10}{3} 0.5^3 0.5^7$$

$$\binom{10}{3} = \frac{10!}{3!(10-3)!} = \frac{10!}{3!7!} = 120$$

(This means there are 120 different paths that would end up with the ball in the 3rd bin from the right.)

So we get:

$$f(3; 10, 0.5) = 120 \times 0.5^3 0.5^7 = 0.1171875$$

In fact we can build a whole table for rows=10 and probability=0.5 like this:

From Right:	10	9	8	7	6	5	4	3	2	1	0
Probability:	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001
Example: 100 balls	0	1	4	12	21	24	21	12	4	1	0

Now, of course, this is a random thing so your results may vary from this ideal situation.

**Do It Yourself:**

Create the the probability table in the biased situation, were the probability of falling to the left is  $p=0.8$ .

From Right	10	9	8	7	6	5	4	3	2	1	0
Probability					0.088						
Example: 100 balls					9						

**Play with it some more!**

Use one of the many applets provided online and run 100 (or more) balls through the Quincunx and see what results you get. Have fun!