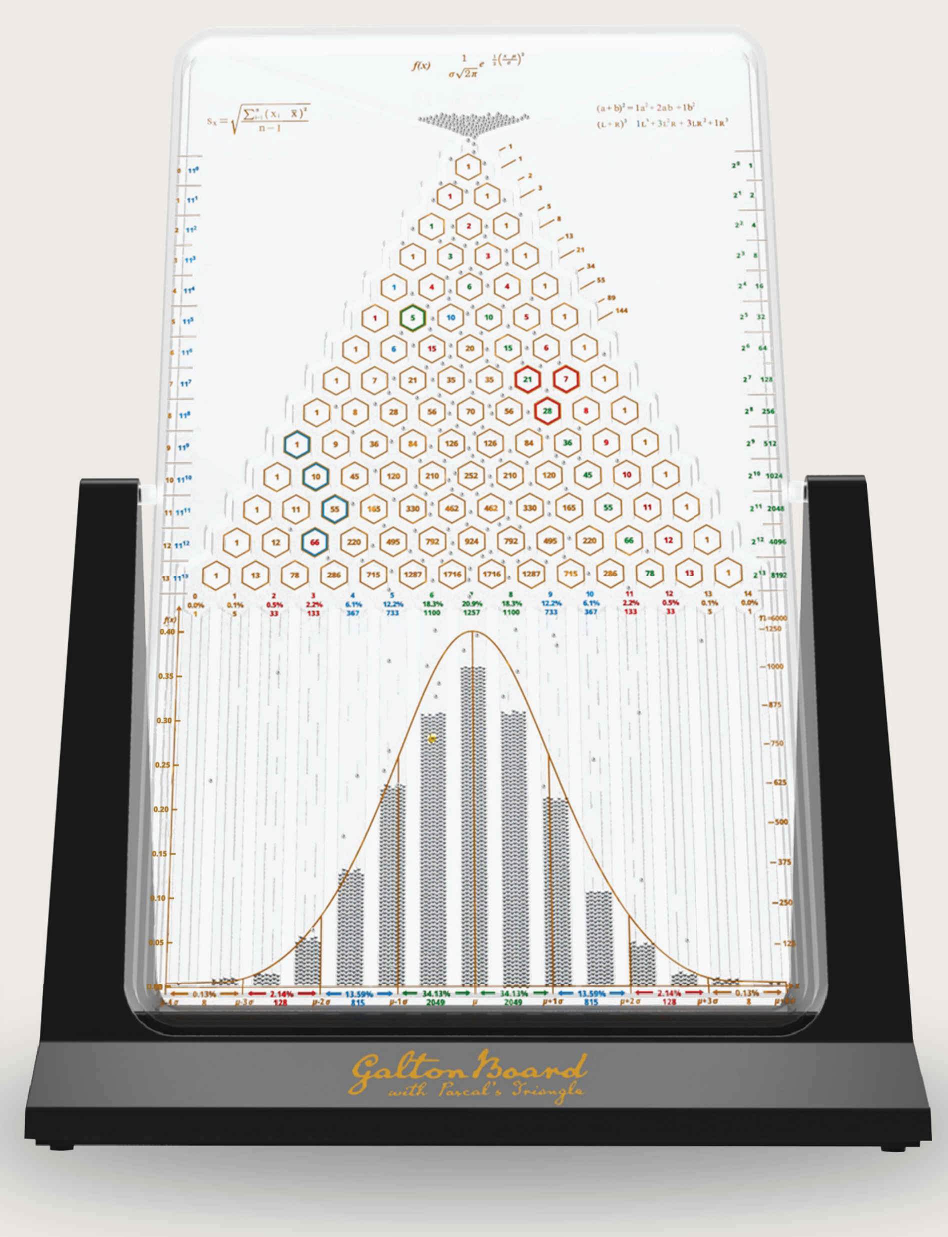
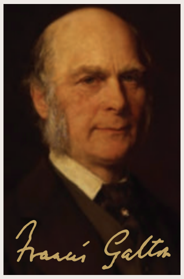
Galton Board
with Pascal's Triangle



Probability Demonstrator

Galton Board
with Pascal's Triangle



Sir Francis Galton

The Galton Board with Pascal's Triangle is a probability demonstrator providing a visualization of math in motion. It displays centuries old mathematical concepts in a innovative device. It incorporates Sir Francis Galton's (1822-1911) invention from 1873 that illustrated the binomial distribution, which for a large number of rows of hexagons and a large number of beads approximates the normal distribution, a concept known as the Central Limit Theorem. According to the Central Limit Theorem, more specifically, the de Moivre (1667-1754)– Laplace (1749-1827) theorem, the normal distribution may be used as an approximation to the binomial distribution under certain conditions. The binomial distribution is altered by the number of rows of hexagons, causing proportional changes to the standard deviation of the resulting bell-shaped curve of beads that land in the bins.

When rotated on its axis, the 6,000 steel beads cascade through rows of symmetrically placed hexagons in the Galton Board. When the device is level, each bead bounces off the hexagons with equal probability of moving to the left or right. As the beads settle into the bins at the bottom of the board, they accumulate to approximate a bell-shaped histogram. Printed on the lower part of the board is the normal distribution or bell curve, as well as the average and standard deviation lines relative to that distribution. The bell curve, also known as the Gaussian distribution (Carl Friedrich Gauss, 1777-1855), is important in statistics and probability theory. It is used in the natural and social sciences to represent random variables, like the beads in the Galton Board. You can also see the Y-axis and X-axis descriptions, and numbered bins with expected percentage and number of beads.

Printed on the top of the board are formulas for the standard deviation of a sample, normal distribution and the binomial expansion.

Overlaid on the hexagons is Pascal's triangle (Blaise Pascal, 1623-1662), which is a triangle of numbers that follows the

Blaise Pascal

rule of adding the two numbers above to get the number below. The number at each hexagon represents the number of different paths a bead could travel from the top hexagon to that hexagon. It also shows the Fibonacci numbers (Leonardo Fibonacci, 1175-1250), which are the sums of specific diagonals in Pascal's triangle. Within Pascal's triangle, mathematical properties and patterns are numerous. Those include: natural numbers, row totals, powers of 11, powers of 2, triangular numbers, Star of David theorem, and the hockey stick pattern. Other patterns in Pascal's triangle include prime numbers; square numbers; binary numbers; Catalan numbers; binomial expansion; fractals; golden ratio; and Sierpinkski's triangle.

Among the 6,000 steel beads, there is one golden bead, which identifies a single random outcome. With the percentage estimates of the probability that the golden bead will land in a specific bin, you can witness those probabilities with each flip of the Galton Board.

Embedded in this Galton Board are many statistical and mathematical concepts including probability theories, independent identically distributed (IID) random variables, the normal or bell-shaped curve, the Central Limit Theorem (the de Moivre-Laplace theorem), the binomial distribution, also known as the Bernoulli (1655-1705) distribution, regression to the mean, the law of large numbers, probabilities such as coin flipping and stock market returns, the random walk, the Gambler's Fallacy, the law of frequency of errors and what Sir Francis Galton referred to as the "law of unreason."

In his book *Natural Inheritance* (1889), Sir Francis Galton colorfully described the apparatus he created to reveal the order in apparent chaos. The following is a modified excerpt from that book. The text has been updated to correspond to the terminology used to describe our Galton Board.

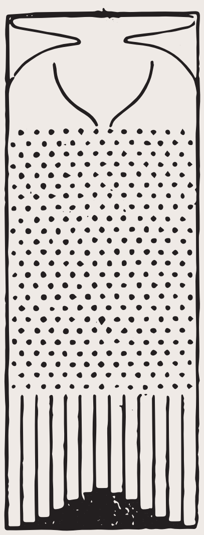
The Charms of Statistics

"It is difficult to understand why statisticians commonly limit their inquires to Averages, and do not revel in more comprehensive views. Their souls seem as dull to the charm of variety as that of the native of one of our flat English counties, whose retrospect of Switzerland was that, if its mountains could be thrown into its lakes, two nuisances would be got rid of at once. An Average is but a solitary fact, whereas if a single other fact be added to it, an entire Normal Scheme, which nearly corresponds to the observed one, starts potentially into existence.

Some people hate the very name of statistics, but I find them full of beauty and interest. Whenever they are not brutalized, but delicately handled by the higher methods, and are warily interpreted, their power of dealing with complicated phenomena is extraordinary. They are the only tools by which an opening can be cut through the formidable thicket of difficulties that bars the path of those who pursue the Science of man.**"**

Mechanical Illustrations of the Cause of the Curve of Frequency

"The Curve of Frequency, and that of Distribution, are convertible: therefore, if the genesis of either of them can be made clear, that of the other becomes also intelligible. I shall now illustrate the origin of the Curve of Frequency, by means of an apparatus (shown here) that mimics in a very pretty way the conditions on which Deviation depends.**"**



Galton's orginal drawing (1889)

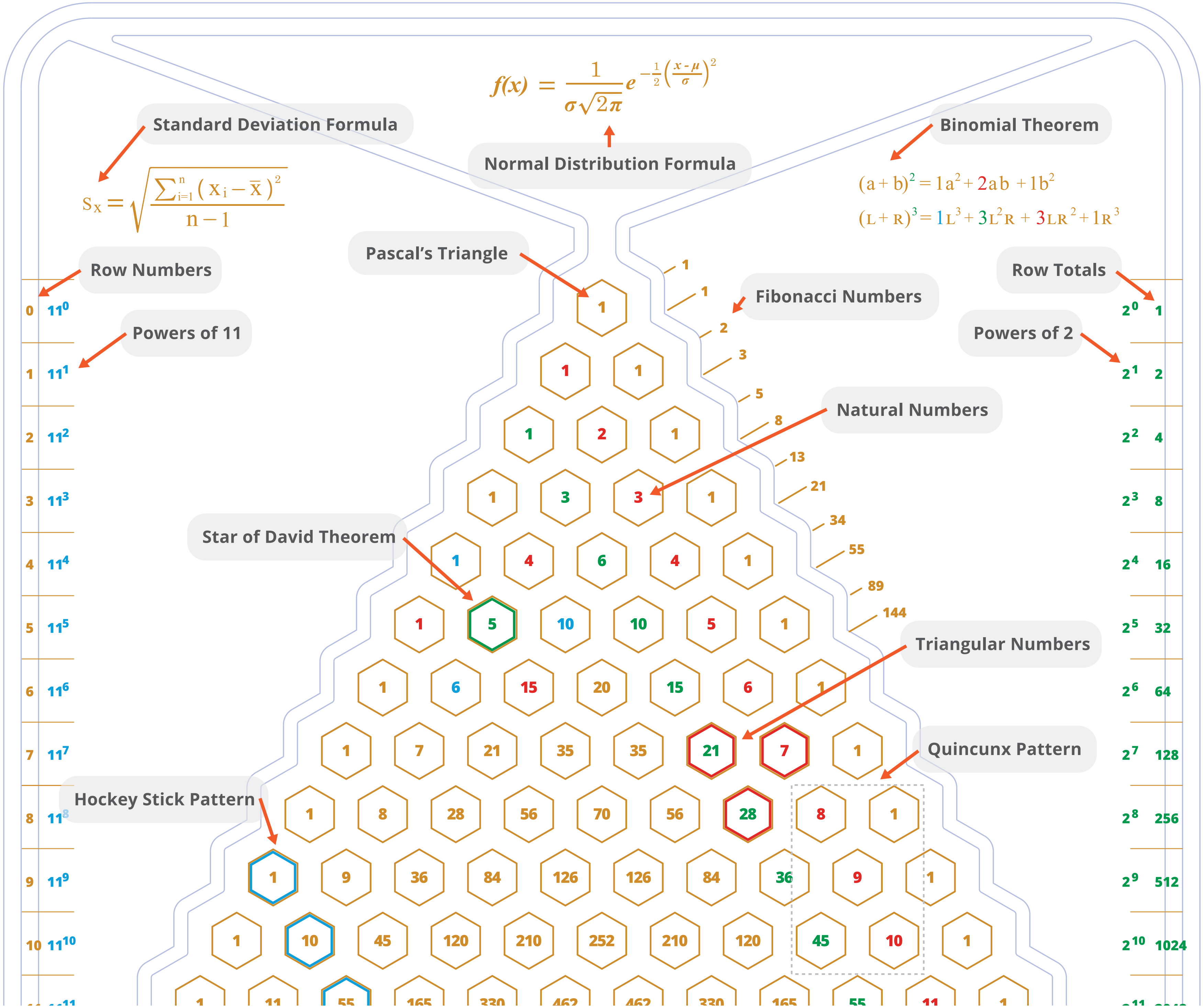
Our design of the Galton Board is constructed of a plastic frame. A reservoir is designed into the top of the board. Below the outlet of the funnel stands a succession of rows of hexagons, similar to Galton's pegs, stuck squarely into the back of the board, and below these again are a series of bins, or vertical compartments. A charge of 6,000 steel beads is enclosed in the board. When the board is flipped **"**topsy-turvy**"**, all the beads run to the upper end into the reservoir; then, when it is turned back into its working position, the desired action commences. The borders of the reservoir have the effect of directing all the beads that had collected at the upper end of the frame to run into the mouth of the funnel.

"The beads pass through the funnel and scamper deviously down through the pegs in a curious and interesting way; each of them darting a step to the right or left, as the case may be, every time it strikes a peg [hexagon]. The pegs are disposed in a quincunx fashion, so that every descending bead strikes against a peg in each successive row. The cascade issuing from the funnel broadens as it descends, and, at length every bead finds itself caught in a bin immediately after freeing itself from the last row of pegs. The outline of the distribution of beads that accumulate in the bins approximates to the Curve of Frequency, and is closely of the same shape however often the experiment is repeated.**"**

"The principle on which the action of the apparatus depends is, that a number of small and independent accidents befall each bead in its career. In rare cases, a long run of luck continues to favor the course of a particular bead towards either outside bin, but in the large majority of instances the number of accidents that cause Deviation to the right, balance in a greater or less degree those that cause Deviation to the left. Therefore most of the beads find their way into the bins that are situated near to a perpendicular line drawn from the outlet of the funnel, and the Frequency with which beads stray to different distances to the right or left of that line diminishes in a much faster ratio than those distances increase.**"**

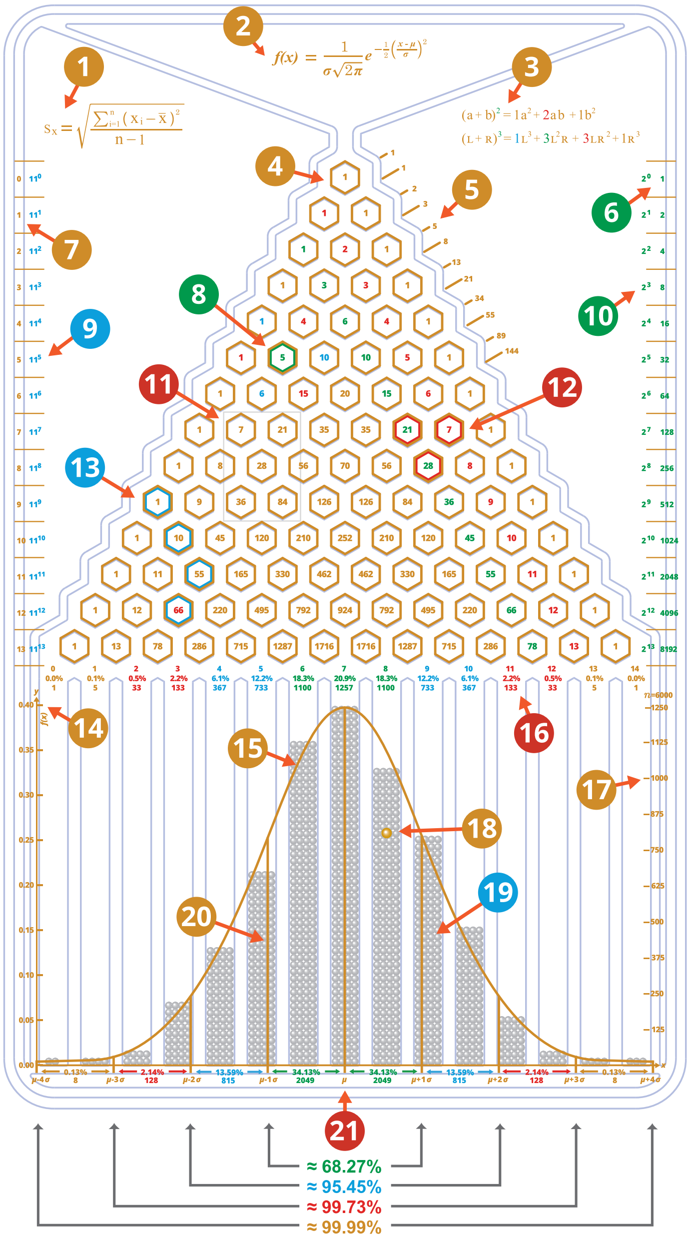
Order in Apparent Chaos

"I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "Law of Frequency of Error." The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement amidst the wildest confusion. The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshaled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to have been latent all along. The tops of the marshaled bins form a flowing curve of invariable proportions; and each element, as it is sorted into place, finds, as it were, a pre-ordained niche, accurately adapted to fit it. If the measurement at any two specified Grades in the bin are known, those that will be found at every other Grade, except towards the extreme ends, can be predicted in the way already explained, and with much precision.**"**



$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
Standard Deviation Formula
Normal Distribution Formula
Binomial Theorem
$S_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}$
$(a+b)^2 = 1a^2 + 2ab + 1b^2$
$(L+R)^3 = 1L^3 + 3L^2R + 3LR^2 + 1R^3$
Pascal's Triangle
Row Numbers
Row Totals
Fibonacci Numbers
Powers of 11
Powers of 2
Natural Numbers
Star of David Theorem
Triangular Numbers
Quincunx Pattern
Hockey Stick Pattern


Number of Beads
X-axis
6,000 Steel Beads
Bin Numbers, Expected Percentages and Beads
Golden Bead
Bell Curve
Standard Deviation Lines
Probability Density
Mean and Standard Deviations
n=6000
1250
1125
1000
875
750
625
500
375
250
125
f(x)
y
x
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0 0.0% 1
1 0.1% 5
2 0.5% 33
3 2.2% 133
4 6.1% 367
5 12.2% 733
6 18.3% 1100
7 20.9% 1257
8 18.3% 1100
9 12.2% 733
10 6.1% 367
11 2.2% 133
12 0.5% 33
13 0.1% 5
14 0.0% 1
μ-4σ
0.13% 8
μ-3σ
2.14% 128
μ-2σ
13.59% 815
μ-1σ
34.13% 2049
μ
34.13% 2049
μ+1σ
13.59% 815
μ+2σ
2.14% 128
μ+3σ
0.13% 8
μ+4σ
2^11 2048
2^12 4096
2^13 8192
11 11^11
12 11^12
13 11^13



1 Standard Deviation Formula

$$S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

How to calculate the standard deviation of a sample data set:

1. Calculate the mean of your data set ($\bar{x} = (\sum x_i)/n$), which is the estimate of μ in the normal distribution formula.
2. Subtract the mean from each of the sample data values (x_i) and list the differences. x_i's are samples of x in the normal distribution formula.
3. Square each of the differences ($\bar{x} - x_i$) from the previous step and make a list of the squares.
4. Add the squares together.
5. Subtract one from the number of data values (n) you started with.
6. Divide the sum from step four by the number from step five.
7. Take the square root of the number from the previous step. This is the standard deviation of the sample (S_x), which is the estimate of σ in the normal distribution.

2 Normal Distribution Formula

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

In probability theory, a normal distribution is a type of continuous probability distribution for a real-valued random variable. Shown here is the general form of its probability density function ($f(x)$). Normal distributions are important in statistics and are often used in the natural and social sciences to represent real-valued random variables whose distributions are not known. Included in the formula is the constant Pi ($\pi \approx 3.142$) which is the ratio of a circle's circumference to its diameter. Also included is Euler's number ($e \approx 2.718$) which is the base of the natural logarithm. The Independent Identically Distributed (IID) Central Limit Theorem states that the random variable x will be normally distributed as the sample size becomes large and sigma is finite.

3 Binomial Theorem

$$(a+b)^2 = 1a^2 + 2ab + 1b^2$$
$$(L+R)^3 = 1L^3 + 3L^2R + 3LR^2 + 1R^3$$

The Binomial Theorem describes the algebraic expansion of powers of a binomial. Pascal's triangle defines the coefficients which appear in binomial expansions. That means the n^{th} row of Pascal's triangle comprises the coefficients of the expanded expression of the polynomial $(a+b)^n$. For the Galton Board, the binomials are left and right $(L+R)^n$.

The expansion of $(a+b)^n$ is:

$$(a+b)^n = x_0a^n + x_1a^{n-1}b + x_2a^{n-2}b^2 + \ldots + x_{n-1}ab^{n-1} + x_nb^n$$

where the coefficients of the form x_k are precisely the numbers that appear in the k^{th} entry of the n^{th} row of Pascal's triangle (k and n counting starts at 0). This can be expressed as: $x_k = \binom{n}{k}$, i.e., "n choose k". The first hexagon on the Galton Board is $\binom{0}{0}$, followed below by $\binom{1}{0}$ and $\binom{1}{1}$.

Examples of binomial expressions are shown for $(a+b)^n$ for n = 2 and $(l+r)^n$ for n = 3. The numbers in each hexagon are the number of paths that a bead can take to arrive at that location.

4 Pascal's Triangle

Pascal's Triangle is a triangle of numbers that follow the rule of adding the two numbers above to get the number below. This pattern can continue endlessly. Blaise Pascal (1623-1662) used the triangle to study probability theory, as described in his mathematical treatise *Traité du triangle arithmétique* (1665). The triangle's patterns translate to mathematical properties of the binomial coefficients.

5 Fibonacci Numbers and the Golden Ratio

The sum of the numbers on the diagonal shown on Pascal's Triangle match the Fibonacci numbers. The sequence progresses in this order: **1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89** and so on. Each number in the sequence is the sum of the previous two numbers. For example: 2+3=5; 3+5=8; 5+8=13; 8+13=21… Leonardo Fibonacci (1175-1250) popularized these numbers in his book *Liber Abaci* (1202). As you progress through the Fibonacci numbers, the ratios of consecutive Fibonacci numbers approach the Golden Ratio of 1.61803398…, but never equals it. For example: 55/34=1.618; 89/55=1.618 and 144/89=1.618. This Galton Board rectangle has side lengths in the Golden Ratio of 1:1.618.

6 Row Totals

The sum of the numbers in each row is shown here and each total doubles on subsequent rows. In addition, the total of the squares of the entries of a row equals the middle entry of that row number times two. For example, if you sum the squares of the entries in row 4 ($1^2 + 4^2 + 6^2 + 4^2 + 1^2$) that equals 70, which is also the middle entry of row 8.

7 Row Numbers

The 14 rows are numbered, with the first row desginated as n=0 and first entry in each row as k=0. Fourteen rows are large enough so the resulting binomial distribution is a good discrete approximation to the continuous normal distribution.

8 Star of David Theorem

The Star of David theorem says the two sets of three numbers surrounding a number have equal products. In the example shown, the number **5** is surrounded by, in sequence, the numbers **1, 4, 10,15, 6, 1**, and taking alternating numbers, we have **1×10×6** = **4×15×1**.

9 Powers of 11

If you collapse each row into a single number by taking each element as a digit (and carrying over to the left if the element has more than one digit) you get the powers of eleven: **1, 11, 121,1331, 14641, 161051**...

10 Powers of 2

The sum of numbers in a row is equal to 2^n where n equals the row number.

11 Quincunx Pattern

The hexagons on the board are in a Quincunx pattern, which is an arrangement of five objects with four at the corners of a square or rectangle and the fifth at its center.

12 Diagonals and Triangular Numbers

The diagonals contain the figurate numbers of simplices, with the left and right edges containing only 1's. The subsequent diagonals contain natural or counting numbers, then triangular numbers (number of dots in an equilateral triangular arrangement), then tetrahedral numbers (triangular pyramidal numbers), then pentatope numbers followed by the 5, 6, and 7 simplex numbers. The square of each natural number is equal to the sum of a pair of adjacent entries on the third diagonal (Triangular Numbers). Example: $7^2 = 49 = 21 + 28$

13 Hockey Stick Pattern

The sum of the numbers in a diagonal, starting from the edge with 1, is equal to the number in the next diagonal below. Outlining these numbers reveals a hockey stick pattern, as seen here in **1** + **10** + **55** = **66**.

14 Probability Density

The probability density $f(x)$ is the relationship between observations and their probability. It defines the probability of occurrence of a discrete random variable within a particular range of continuous random variables. One very important probability density function is that of a Gaussian, or normal, random variable which looks like a bell-shaped curve. These $f(x)$ values assume a normal distribution with a sigma (σ) of 1.

15 Bell Curve

The normal distribution, often referred to as the "bell curve", is the most widely known and used of all probability distributions. Because the normal distribution approximates many natural phenomena so well, it has developed into a standard of reference for numerous probability problems. Several sets of data follow the normal distribution, such as the heights of adults, the weights of babies, classroom test scores, monthly returns of the stock market and the beads in the Galton Board.

16 Bin Numbers, Expected Percentages and Extpected Beads

Bins are numbered so the location of the golden bead can be easily identified and recorded. Expected percentage of outcomes and expected number of beads per bin base on 6,000 beads.

17 Number of Beads

This scale provides an estimate of the number of beads that are expected to land in each bin based on the binomial distribution.

18 Golden Bead

Among the 1.0 mm steel beads is a 2.2 mm golden bead. This bead demonstrates a single random outcome.

19 Steel Beads

Each steel bead represents an independent and identically distributed (IID) random variable that falls from the reservoir through a fixed pattern of hexgons. When all 6,000 steel beads are collected in the bins they form a similar distribution every time. The discrete binomial distribution of beads closely approximates the continuous normal distribution.

20 Standard Deviation Line

The standard deviation (σ) is a measure of how closely all of the data points are gathered around the mean (μ). The shape of a normal distribution is determined by the mean and the standard deviation. About 68% of the data (beads) in a bell curve fall within one standard deviation of the mean. About 95% fall within two standard deviations and about 99.7% fall within three standard deviations. The standard deviation (σ) of this 14-row binomial distribution is 1.87 bead bins. The 15 bead bins span a range of 8 standard deviations (15/1.87 = 8).

21 X-axis

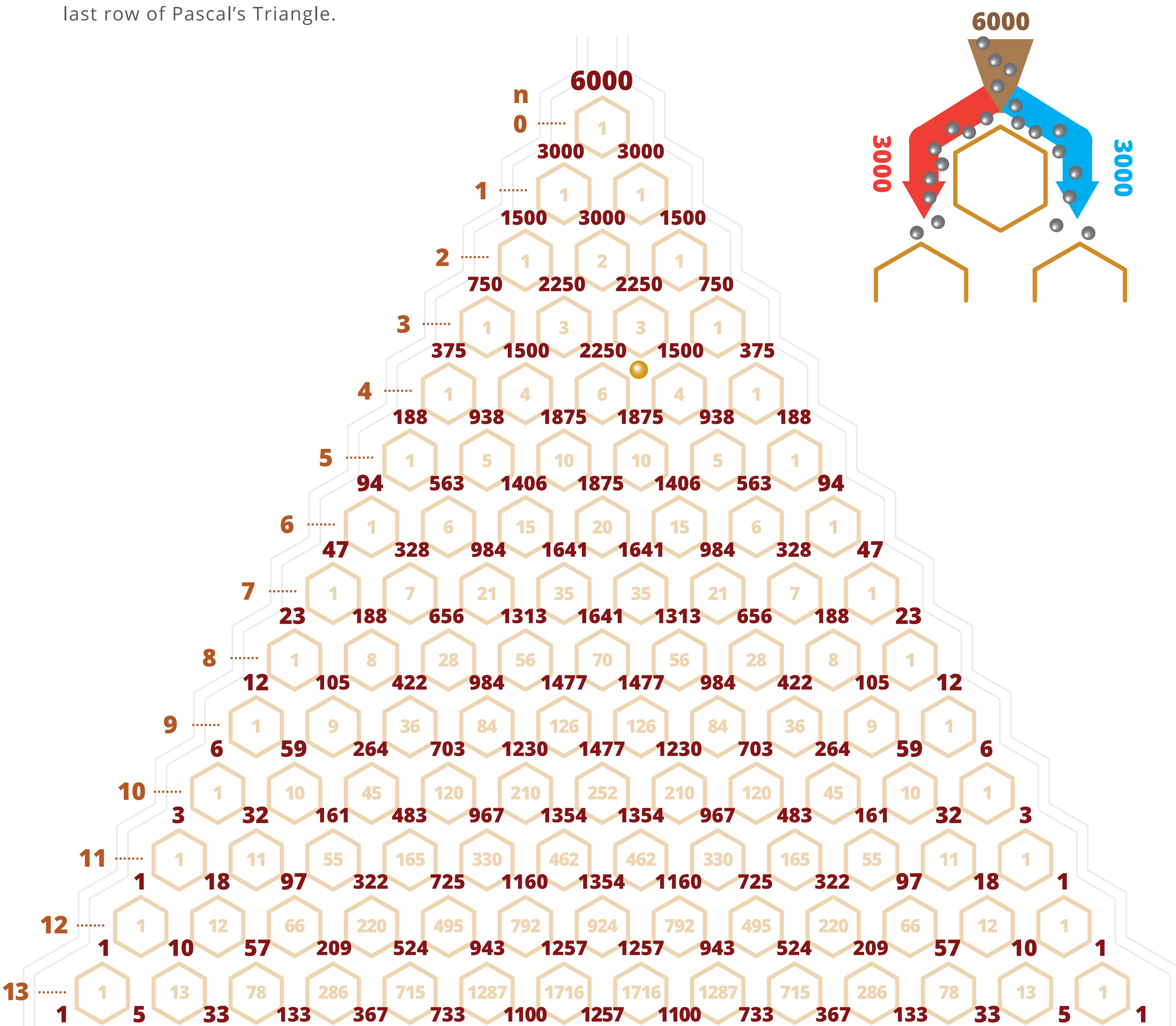
The scale shows x values from μ-4σ to μ+4σ and the percentages and numbers of beads that are expected to be between the standard deviations.

Galton Board

Binomial Distribution of Beads

Since there are approximately 6,000 beads to start with and when the Galton is level there is a 50% chance the beads will go left or right at each hexagon, this illustration shows the expected beads that will travel through each channel around the hexagons. At the first hexagon, which is considered row 0, 3,000 beads should go left and 3,000 beads should go right (assuming exactly 6,000 beads). If you follow the splitting of the beads each time you can see how many beads are expected to land in each bin on the last row of Pascal's Triangle.

For the Galton Board, the Pascal's Triangle numbers for row n can be interpreted as the number of paths to get to the k^{th} location of row n after having gone through n-1 rows. For example, for row 4, the numbers are 1, 4, 6, 4, 1, meaning that there are 1, 4, 6, 4, 1 paths to get to the five hexagonal pegs of row 4 after passing thru row 3 that has 4 hexagonal pegs. For 6,000 beads, there would be 6x6,000/16 = 2250 beads hitting the center peg in row 4.



Pascal's Triangle

Binomial Expansion

The binomial coefficient $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ appears as the k^{th} entry in the n^{th} row of Pascal's triangle (counting starts at 0). Each entry is the sum of the two above it.

The expansion is also given by the following formula:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

$$(a+b)^0 = 1$$

$$(a+b)^1 = 1a + 1b$$

$$(a+b)^2 = 1a^2 + 2ab + 1b^2$$

$$(a+b)^3 = 1a^3 + 3a^2b + 3ab^2 + 1b^3$$

$$(a+b)^4 = 1a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + 1b^4$$

$$(a+b)^5 = 1a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + 1b^5$$

$$(a+b)^6 = 1a^6 + 6a^5b + 15a^4b^2 + 20a^3b^3 + 15a^2b^4 + 6ab^5 + 1b^6$$

$$(a+b)^7 = 1a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 + 21a^2b^5 + 7ab^6 + 1b^7$$

$$(a+b)^8 = 1a^8 + 8a^7b + 28a^6b^2 + 56a^5b^3 + 70a^4b^4 + 56a^3b^5 + 28a^2b^6 + 8ab^7 + 1b^{28}$$

$$(a+b)^9 = 1a^9 + 9a^8b + 36a^7b^2 + 84a^6b^3 + 126a^5b^4 + 126a^4b^5 + 84a^3b^6 + 36a^2b^7 + 9ab^8 + 1b^9$$

$$(a+b)^{10} = 1a^{10} + 10a^9b + 45a^8b^2 + 120a^7b^3 + 210a^6b^4 + 252a^5b^5 + 210a^4b^6 + 120a^3b^7 + 45a^2b^8 + 10ab^9 + 1b^{10}$$

$$(a+b)^{11} = 1a^{11} + 11a^{10}b + 55a^9b^2 + 165a^8b^3 + 330a^7b^4 + 462a^6b^5 + 462a^5b^6 + 330a^4b^7 + 165a^3b^8 + 55a^2b^9 + 11ab^{10} + 1b^{11}$$

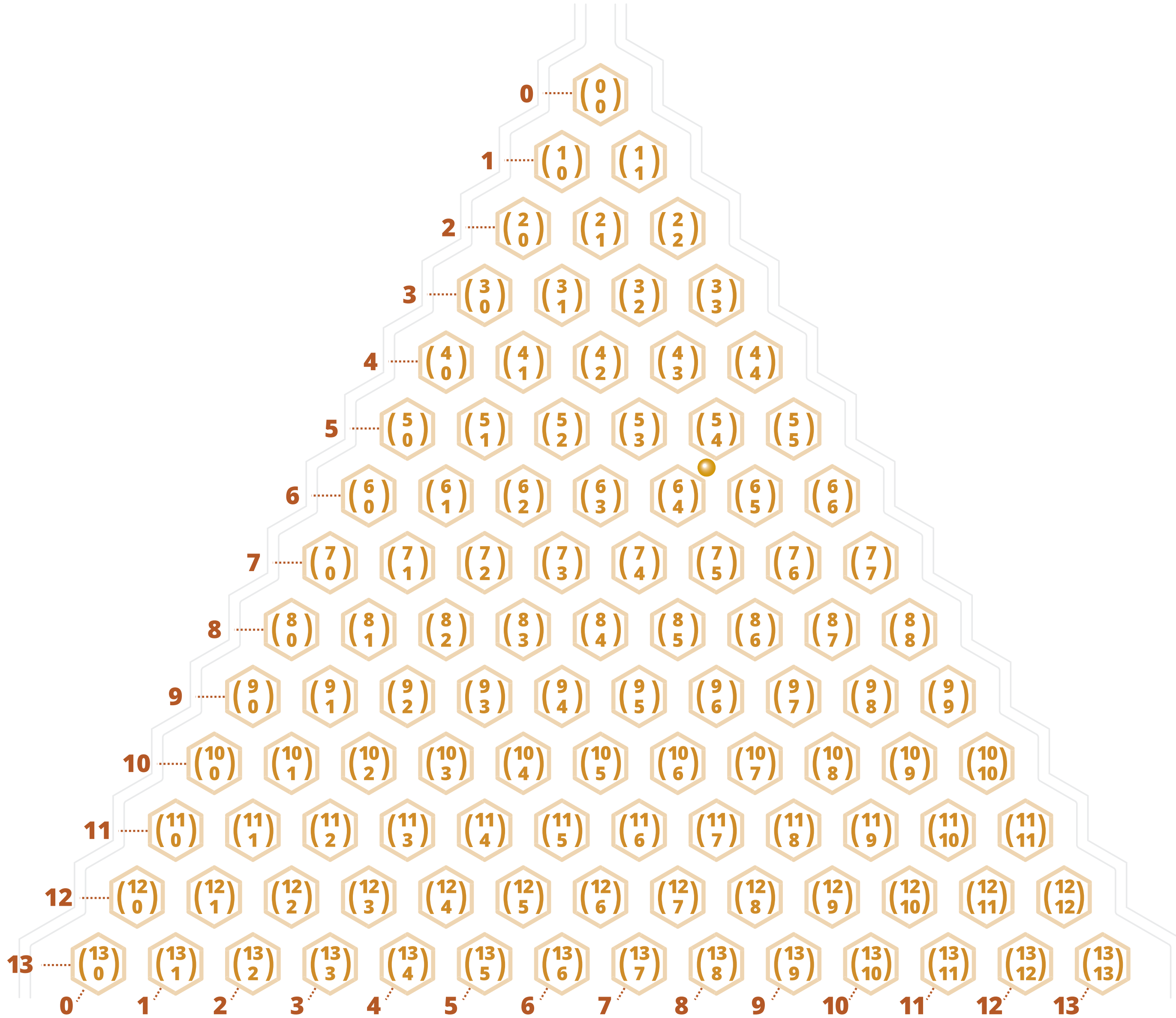
$$(a+b)^{12} = 1a^{12} + 12a^{11}b + 66a^{10}b^2 + 220a^9b^3 + 495a^8b^4 + 792a^7b^5 + 924a^6b^6 + 792a^5b^7 + 495a^4b^8 + 220a^3b^9 + 66a^2b^{10} + 12ab^{11} + 1b^{12}$$

$$(a+b)^{13} = 1a^{13} + 13a^{12}b + 78a^{11}b^2 + 286a^{10}b^3 + 715a^9b^4 + 1287a^8b^5 + 1716a^7b^6 + 1716a^6b^7 + 1287a^5b^8 + 715a^4b^9 + 286a^3b^{10} + 78a^2b^{11} + 13ab^{12} + 1b^{13}$$

Pascal's Triangle

Combinatorics

Pascal's Triangle is also an "n choose k" triangle like this one. Note how the top row is row zero and also the leftmost column is column zero. Each entry can be designated as shown in this illustration.

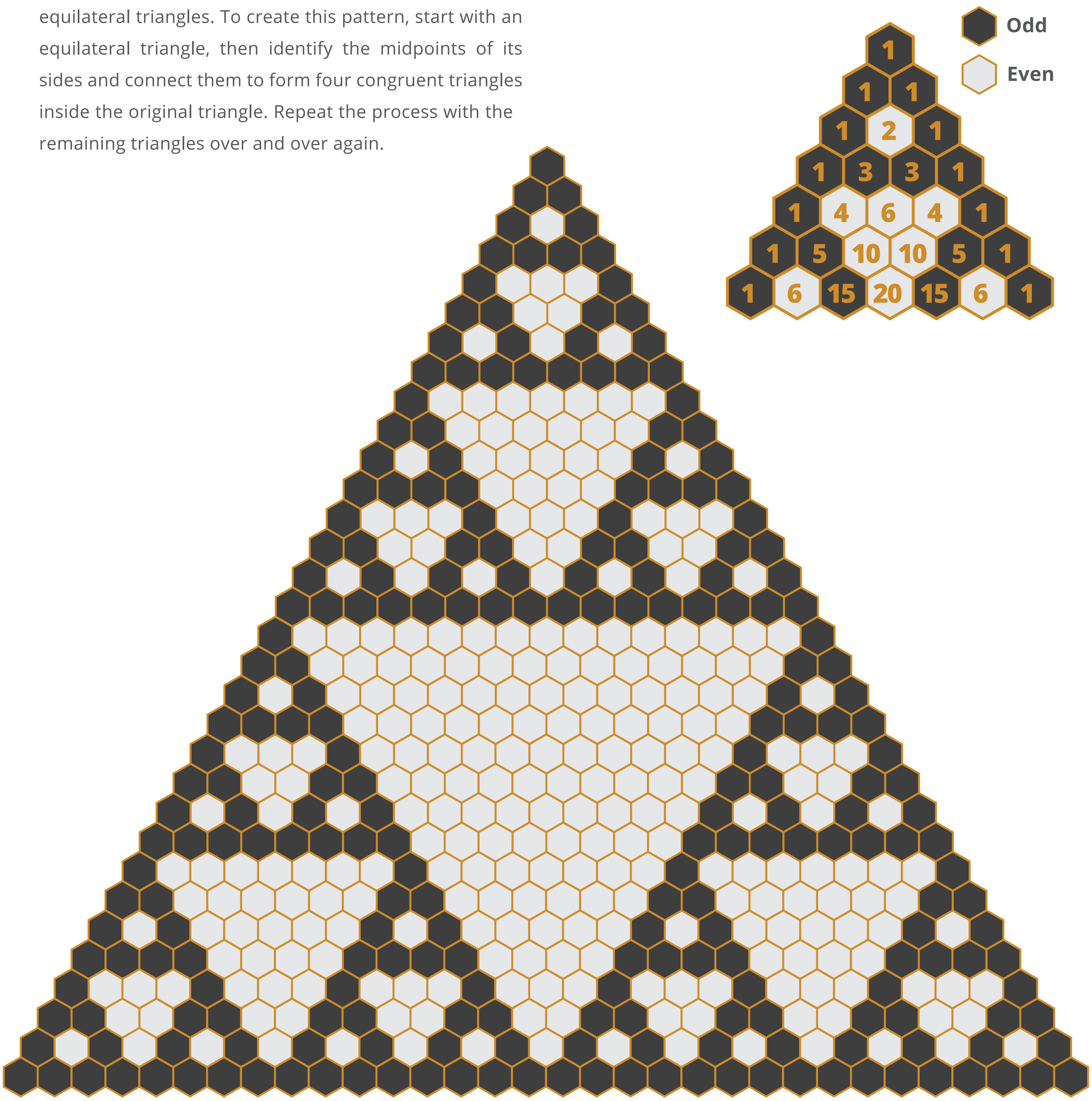


Pascal's Triangle

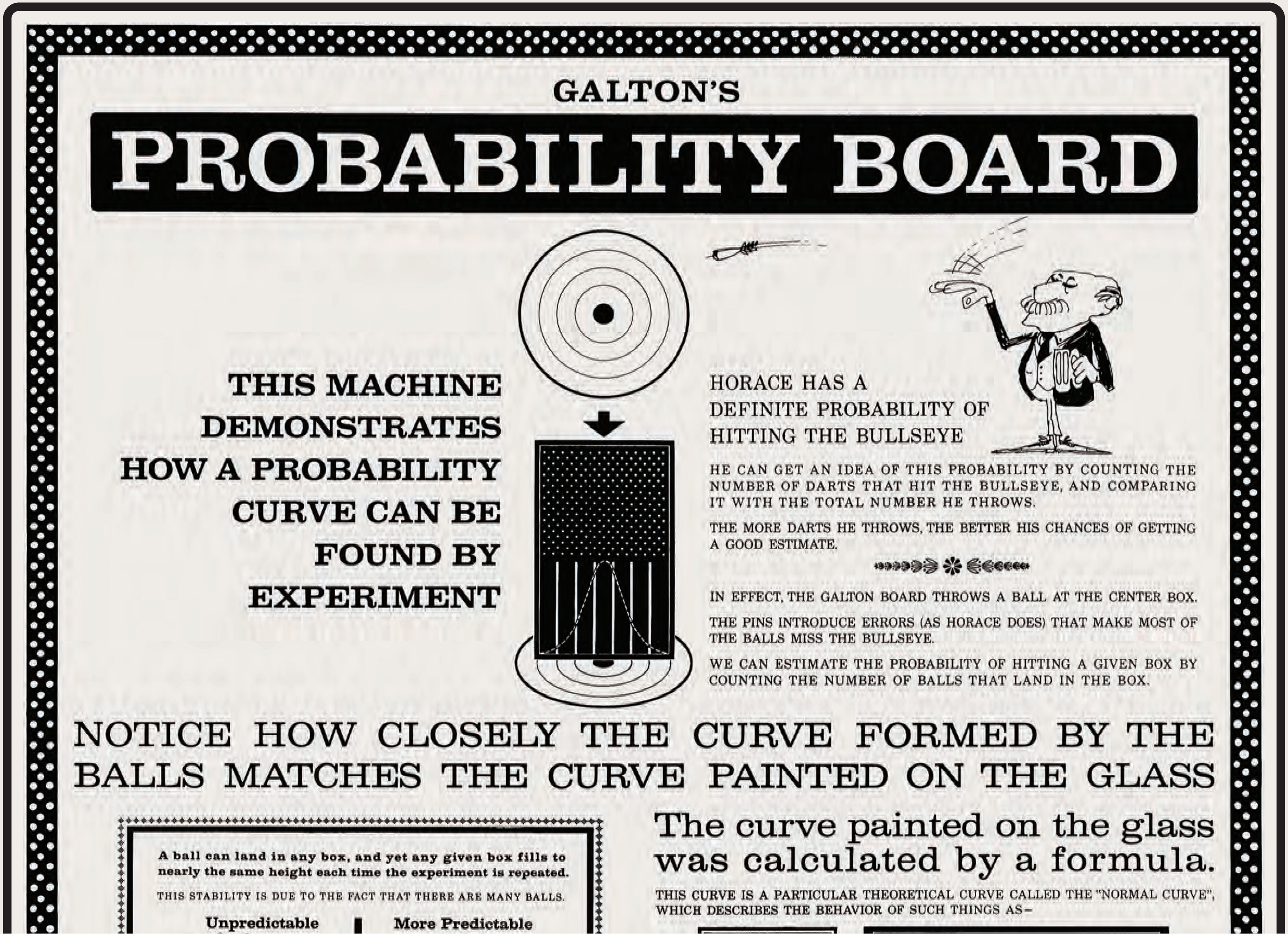
Sierpinski Triangle

The Sierpinski Triangle is a very interesting mathematical structure that is a fractal (a mathematical curve whose shape retains the same general pattern of irregularity, regardless of how high it is magnified), with the overall shape of an equilateral triangle formed by starting with an equilateral triangle and recursively (a rule that is repeated) subdividing the triangle into smaller equilateral triangles. To create this pattern, start with an equilateral triangle, then identify the midpoints of its sides and connect them to form four congruent triangles inside the original triangle. Repeat the process with the remaining triangles over and over again.

If you start with a Pascal's triangle and color the odd numbers black and leave the even numbers white, it will result in the Sierpinski Triangle, which is named after the Polish mathematician Waclaw Sierpinski (1882-1969). Sierpinski's work included three well known fractals, including the triangle, carpet and curve fractals.



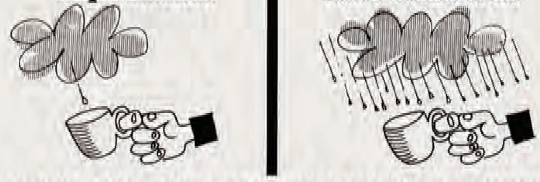
Our Galton Board is a desktop design reminescent of Charles and Ray Eames' groundbreaking 11-foot-tall "Galton's Probability Board," featured at the 1961 ***Mathematica: A World of Numbers ... and Beyond*** exhibit. Pictured below is an approximately 4-foot-tall information sign from the Mathematica Exhibit (some of the text has been enlarged so it is legible at this size). An even larger 14 ½-foot-tall Eames Probability Board was showcased at IBM's Pavilion for the 1964 World's Fair in New York.



nearly the same height each time the experiment is repeated.

THIS STABILITY IS DUE TO THE FACT THAT THERE ARE MANY BALLS.

Unpredictable | **More Predictable**



IF A RANDOM EVENT HAPPENS A GREAT MANY TIMES
THE AVERAGE RESULTS ARE LIKELY TO BE PREDICTABLE.*

*The first mathematical theorem of this kind was proved by Jacob Bernoulli.

"With the probability approaching certainty as near as we please, we may expect that the relative frequency of an event in a series of independent trials with constant probability will differ from that probability by less than any given positive number, provided the number of trials is sufficiently large."

Jacob Bernoulli (1654-1705)

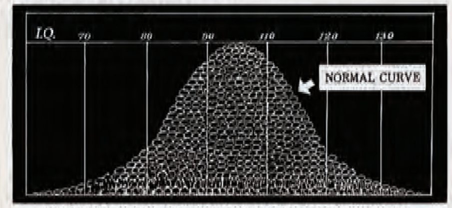
"RELATIVE FREQUENCY" is the number of times an event* occurs divided by the number of trials.**

In the Probability Board the release of a ball is a ** TRIAL.

Landing (or not landing) in a given box is an * EVENT.

THIS CURVE IS A PARTICULAR THEORETICAL CURVE CALLED THE "NORMAL CURVE", WHICH DESCRIBES THE BEHAVIOR OF SUCH THINGS AS–

I.Q. TESTS



IF PEOPLE WERE STACKED IN BOXES ACCORDING TO THEIR I.Q. SCORES, THEY WOULD FORM THE "NORMAL CURVE."

THE MEASUREMENTS OF BEAUTY CONTEST WINNERS

RUN AT ROULETTE

ERRORS IN MEASUREMENT

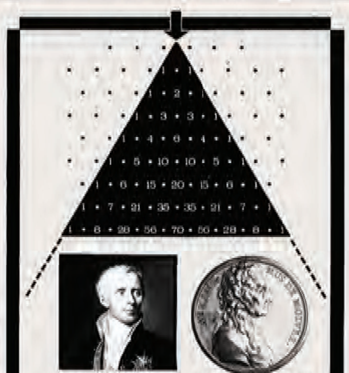
WHEN THE BALLS ARE DROPPED, THEY ARE ALL AIMED AT THE CENTER BOX. THE SUM OF ALL THE ERRORS CAUSED BY HITTING THE PINS DETERMINES THE BALLS' FINAL POSITION.

The average of many independent errors almost always leads to the Normal Curve, no matter what the underlying process may be.

THE "CENTRAL LIMIT THEOREM" IS A PRECISE STATEMENT OF CONDITIONS WHICH LEAD TO THE NORMAL CURVE.

PASCAL'S TRIANGLE

The number of possible paths to a given space in the array of pins is given by Pascal's Triangle. For the number of paths to a space is the sum of the number to the two spaces above it. The probability of a ball's dropping in any box can be found by counting the number of paths to that box, and comparing it with the total number of paths.



LAPLACE (1749-1827) DeMOIVRE (1667-1754)

As the number of trials gets large, the distribution of the balls is likely to be near normal. This idea, first stated in Abraham de Moivre's "Doctrine of Chances," was later proved by the Marquis de Laplace, and called the Laplace-de Moivre limit theorem. Hard work during the next hundred years eventually produced a much more general statement of the same sort, the "central limit theorem," universally conceded to be one of the most important results of probability theory.

PROBABILITY IS ADDITIVE

IF THE PROBABILITY OF "3" IS $\frac{1}{6}$ AND THE PROBABILITY OF "4" IS $\frac{1}{6}$

THEN THE PROBABILITY OF "3" OR "4" IS 1/6 + 1/6. $\frac{1}{6}+\frac{1}{6}$

PROBABILITY, LIKE AREA, IS ADDITIVE. HENCE, IT IS OFTEN POSSIBLE TO REPRESENT PROBABILITY AS THE AREA UNDER A CURVE.

For example, in the Galton Board, the probability of getting in –

this box / or in either of these 2 boxes / or in any of these 3 boxes

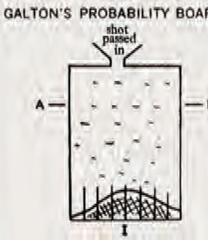
is represented by the corresponding area.

The branch of mathematics concerned with determination of lengths and areas is called "MEASURE THEORY." Probability is a branch of the Theory of Measure.

THE QUINQUNX

The pins in the Galton Board are often arranged in a figure found in nature – the four corners of a square with a pin in the center, called a Quincunx.

GALTON'S PROBABILITY BOARD ·1877



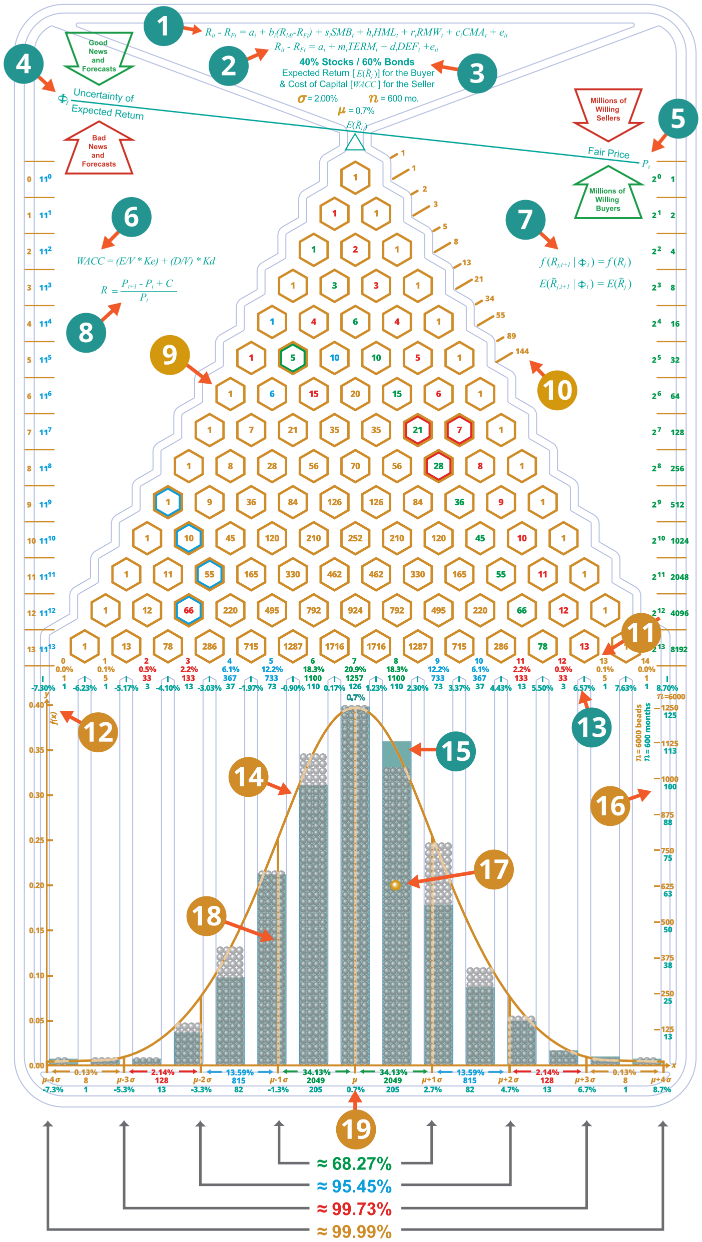
GALTONIA (Hyacinthus candicanus)

SIR FRANCIS GALTON (1822-1911)

Galton was a cousin of Charles Darwin. In addition to mathematics, he studied and wrote about Botany, Heredity, Geography, Psychology, Statistical Methods, and Mountain Climbing.

IBM

The Galton Board is also available in this Stock Market Edition



1 Fama/French Five-Factor Model For Equities

The Fama/French Five-Factor Model for Equities is an asset pricing model directed at capturing the market, size, value, profitability and investment patterns in average stock returns. It was developed in 2014 by Nobel Laureate Eugene Fama and his co-author and colleague Kenneth French. The model explains between 71% and 94% of the cross-section variance of expected returns for diversified portfolios of five factors in equities that include market, size, value, profitability and investment. It expands on the Capital Asset Pricing Model (1964) and the Fama/French Three-Factor Model (1993). The Fama/French Five-Factor Model equation is a time-series regression of a series of research indexes created by Fama and French that include long-term historical stock prices of various company characteristics. The coefficient for each factor (independent variables) indicates the exposure or tilt to that factor in the portfolio. If the exposure to the five factors, market (b_i), size (s_i), value (h_i), profitability (r_i) and investment (c_i) capture all variation in expected returns, the alpha intercept a_i in the following equation is zero for all securities and portfolios i.

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt}-R_{Ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + e_{it}$$

R_{it} is the return on the portfolio i for period t (dependent variable)

R_{Ft} is the risk-free return

R_{Mt}-R_{Ft} is the return spread between the capitalization weighted stock market and cash

SMB_t is the return on a diversified portfolio of small stocks minus the return on diversified portfolio of big stocks (i.e. the size effect)

HML_t is the difference between the return on diversified portfolios of high and low Book to Market stocks (i.e. the value effect)

RMW_t is the difference between the returns on diversified portfolios of stocks with robust and weak profitability

CMA_t is the difference between the returns on diversified portfolios of stocks of low and high investment firms, which Fama/French called conservative and aggressive

e_{it} is the error term and is a zero-mean residual

2 Fama/French Two-Factor Model For Fixed Income

$$R_{it} - R_{Ft} = a_i + m_iTERM_t + d_iDEF_t + e_{it}$$

The Fama French Two-Factor Model for fixed income aims to explain average returns on bond portfolios. The model utilizes the Term ($TERM_t$) and Default (DEF_t) risk factors. $TERM_t$ is LTG-RF, where LTG is the monthly percent long-term government bond return and RF, is the one-month Treasury bill rate, observed at the beginning of the month. DEF is CB-LTG, where CB is the return on a proxy for the market portfolio of corporate bonds. e_{it} is the error term and is a zero-mean residual.

3 40% Stocks / 60% Bonds Hypothetical Investment Portfolio

To represent market returns, we selected a hypothetical investment portfolio of **40% Stocks and 60% Bonds** because it is estimated to have a distribution close to this bell curve. The expected return of the buyer is also the cost of capital for the seller. For this portfolio, we assume a standard deviation of 2.0%, a monthly mean return of 0.7% and a sample size of 600 months.

4 Information and Uncertainty

This teeter totter illustrates Eugene Fama's Efficient Market Hypothesis which states that prices of securities fully reflect all available information. The left side of the teeter totter represents all available information (Φ_t) and the right side represents the prices (P_t) that millions of willing buyers and sellers have concluded are fair prices given that information at that time. There is a random and continuous flow of good news and forecasts and bad news and forecasts which at any point in time represents the uncertainty of the expected return of an investment that is held at a constant level of risk, such as the portfolio of **40% Stocks / 60% Bonds** shown. If uncertainty increases due to bad news, the price must make a proportional adjustment down so that the expected return remains essentially constant. The opposite is also true. This teeter totter and bell curve illustration was developed by Mark T. Hebner and is known as the Hebner Model.

5 Fair Price

The Efficient Market Hypothesis asserts that, in a well organized, reasonably transparent market, the market price (P_t) is generally equal to or close to the fair value, as investors react quickly to incorporate new information (Φ_t) about relative scarcity, utility, or potential returns in their bids.

6 Cost of Capital

$$WACC = (E/V * Ke) + (D/V) * Kd$$

In economics and accounting, the cost of capital is the cost of a company's funds (both debt and equity), or, from an investor's point of view "the required rate of return on a company's existing securities." It is used to evaluate new projects of a company. It is the minimum return that investors expect for providing capital to the company, thus setting a benchmark that a new project has to meet.

7 The Random Walk Model

$$f(R_{j,t+1} \mid \Phi_t) = f(R_j)$$
$$E(\tilde{R}_{j,t+1} \mid \Phi_t) = E(\tilde{R}_j)$$

The Efficient Markets Hypothesis states that the current price (P_t) of a security fully reflects available information (Φ_t), which implies that the successive price changes, or more usually, successive one-period returns, are independent. In addition, it assumes that successive changes, or returns, are identically distributed. Together the two assumptions constitute the Random Walk Model. Formally, the model says that: $f(R_{j,t+1} \mid \Phi_t) = f(R_j)$, which is the usual statement that the conditional and marginal probability distributions of an independent random variable are identical. In addition, the density function f must be the same for all t. If we assume that the expected return on security $E(\tilde{R}_j)$ is constant over time, we have $E(\tilde{R}_{j,t+1} \mid \Phi_t) = E(\tilde{R}_j)$.

8 Investment Return Formula

$$R = \frac{(P_{t+1} - P_t) + C}{P_t}$$

The formula for an investment return/loss (R) is the change in price ($P_{t+1} - P_t$) plus any dividends or cash paid to the investor during the period (C), divided by the original price (P_t) of the investment.

9 Pascal's Triangle

Pascal's Triangle is a triangle of numbers that follow the rule of adding the two numbers above to get the number below. This pattern can continue endlessly. Blaise Pascal (1623-1662) used the triangle to study probability theory, as described in his mathematical treatise *Traité du triangle arithmétique* (1665). The triangle's patterns translate to mathematical properties of the binomial coefficients.

10 Fibonacci Numbers

The sum of the numbers on the diagonal of Pascal's Triangle match the Fibonacci numbers (*Liber Abaci*, 1202), where each number in the sequence is the sum of the previous two numbers: **1, 1, 2, 3, 5, 8, 13, 21**, and so on. The ratios of consecutive Fibonacci numbers (55/34=1.618; 89/55=1.618 and 144/89=1.618) and the dimensions of this Galton Board approach the Golden Ratio of 1.618.

11 Bin Numbers

Bins are numbered so the location of the golden bead can be easily identified and recorded.

12 Probability Density

The probability density $f(x)$ is the relationship between observations and their probability. It defines the probability of occurrence of a discrete random variable within a particular range of continuous random variables. These $f(x)$ values assume a normal distribution with a sigma (σ) of 1.

13 Expected Percentages, Beads and Returns

Expected % of outcomes per bin for both **beads** and **returns** are shown just below the bin number, with **20.9%** expected in middle bin (**#7**). Then the expected # of beads per bin are shown, based on 6,000 beads. Below that are estimates of the # of monthly returns expected per bin based on 600 monthly returns for the **40% Stocks / 60% Bonds investment portfolio**. Finally, the scale of expected monthly returns from **-7.3%** to **8.7%** are shown on the bin dividers, which is expected to include 4 standard deviations of returns (**≈99.99%** or **$\mu \pm 4\sigma$**), with about 1 monthly return expected in each tail beyond 4 standard deviations.

14 Bell Curve

The normal distribution, often referred to as the "bell curve", is the most widely known and used of all probability distributions. Because the normal distribution approximates many natural phenomena so well, it has developed into a standard of reference for numerous probability problems. Several sets of data follow the normal distribution, such as the heights of adults, the weights of babies, classroom test scores, monthly returns of the stock market and the beads in the Galton Board.

15 Hypothetical Distribution of Monthly Returns

The blue bars printed on the back of the board behind the bell curve represent a hypothetical distribution of 600 monthly returns of the **40% Stock / 60% Bonds investment portfolio**.

16 Number of Beads and Returns

The right Y-axis provides an estimate of the number of **beads** and monthly **returns** that are expected in each bin based on the normal distribution.

17 Golden Bead

Among the 1.0 mm steel beads is a 2.2 mm golden bead. This bead demonstrates a single random outcome.

18 Standard Deviation Line

The standard deviation (σ) is a measure of how closely all of the data points are gathered around the mean (μ). The shape of a normal distribution is determined by the mean and the standard deviation. About 68% of the data (beads) in a bell curve fall within one standard deviation of the mean. About 95% fall within two standard deviations and about 99.7% fall within three standard deviations. The standard deviation (σ) of this 14-row binomial distribution is 1.87 bead bins. The 15 bead bins span a range of 8 standard deviations (15/1.87 = 8).

19 X-axis

This scale provides the mean, plus and minus 4 standard deviations from the mean (**$\mu \pm 4\sigma$**) and the percentage of **beads** and **returns** that are expected to be between the standard deviations.

"The Galton Board is a chilling reminder that out of wonderful, wild randomness, order and stability can emerge."

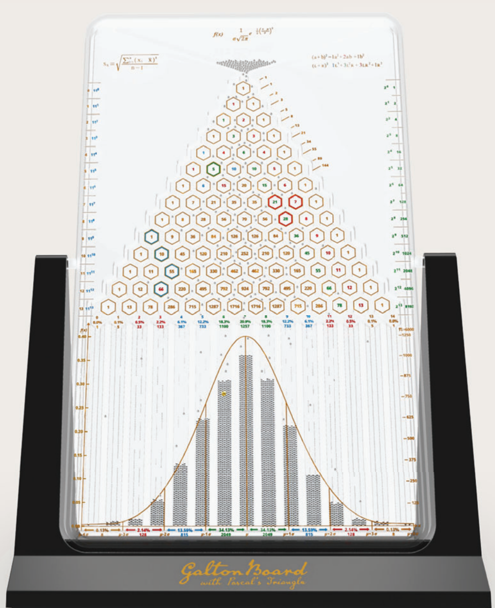
VSAUCE D!NG

Our Story

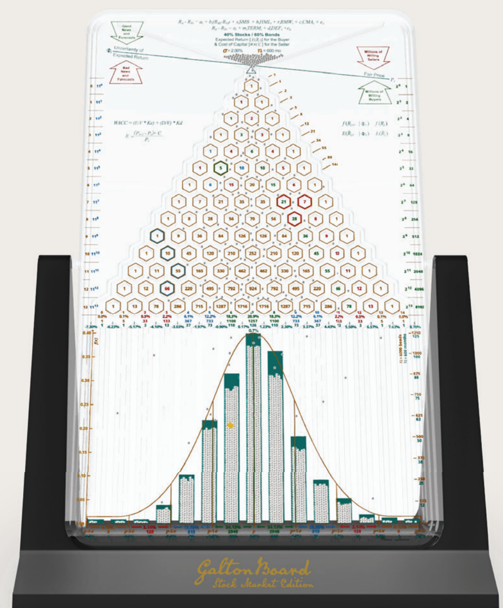
Mark T. Hebner

Mark T. Hebner is the founder and President of Four Pines Publishing, Inc. Mark's fascination with the Galton Board was ignited when he saw an Eames Office Film on the 1964 World's Fair. Charles Eames built an outdoor 14 1/2-foot-tall Galton Board for the IBM Exhibit, modeled after a previous design he had built for *Mathematica: A World of Numbers. . . . and Beyond*. Mathematica was the first fully immersive and large-scale exhibition produced by the Eames Office. It was designed for the 1961 opening of a new science wing at the California Museum of Science and Industry in Los Angeles.

Mark's first Galton Board was designed and built by the Oregon Museum of Science and Industry. It was an 8-foot-tall Galton Board for the lobby of his wealth management and taxes firm, Index Fund Advisors. Then, along with the help of Philip Poissant, Jerry Xu, Art Forster, Jackson Lin, and the Brunson family, he created his first desktop sized Galton Board that is 7 1/2 inches tall. This is the third evolution of his boards. It is 12 inches tall and more precisely captures the concepts of the binomial distribution and Pascal's Triangle, along with the many embedded mathematical concepts of Pascal's Triangle. The Stock Market Edition of this Galton Board displays and explains elements of market returns, including the Hebner Model.



Galton Board
with Pascal's Triangle



Galton Board
Stock Market Edition

FOUR PINES PUBLISHING, INC.

Educating through Innovative Design

GaltonBoard.com