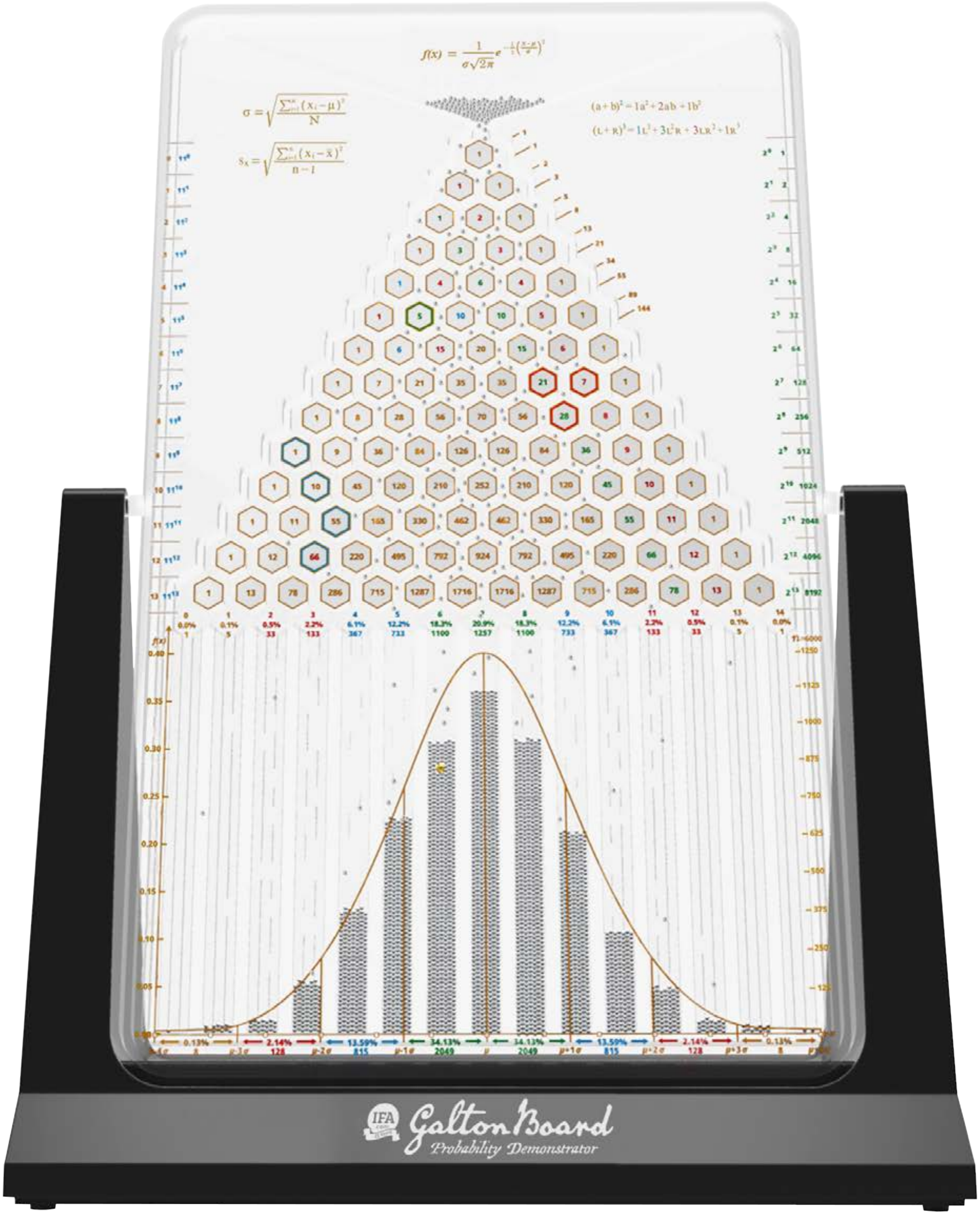
Galton Board
Probability Demonstrator

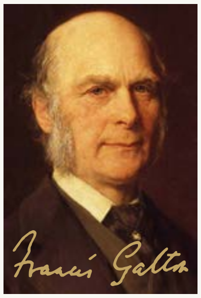


Includes Pascal's Triangle Overlay & Stock Market Clip-ons

USER GUIDE

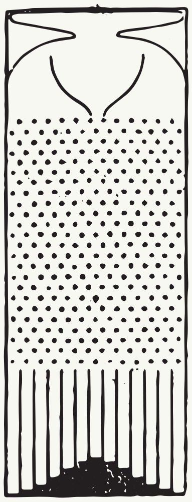
Introduction

The Galton Board with Pascal's triangle is a 12" by 9" (310mm by 218mm) probability demonstrator providing a visualization of math in motion and the powers of the probabilities and statistics. With the addition of the Stock Market Clip-ons, the board illustrates the randomness and the probabilities of various market returns.



Sir Francis Galton

The Galton Board displays centuries old mathematical concepts in an innovative, dynamic desktop 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. He was fascinated with the order of the bell curve that emerges from the apparent chaos of beads bouncing off of pegs in his board. 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.



Galton's original drawing

When rotated on its axis, the 6,000 steel beads and one large golden bead cascade through 14 rows of symmetrically placed hexagons in the Galton Board. When the device is level, beads bounce off of the 105 hexagons with equal probability of moving to the left or right. As the beads settle into one of the 15 bins at the bottom of the board, they accumulate to create a bell-shaped histogram (binomial distribution).

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 percentages and numbers of beads.

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

Blaise Pascal

Overlaid on the hexagons is Pascal's triangle (Blaise Pascal, 1623-1662), which is a triangle of numbers that follows the 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, figurate numbers, Star of David theorem, and the hockey stick pattern. Other patterns in Pascal's triangle not identified on this board include prime numbers; square numbers; binary numbers; Catalan numbers; binomial expansion; fractals; golden ratio; and the Sierpinski triangle.

Among the 6,000 steel beads, there is one golden bead, which demonstrates a single random outcome. Shown on top of each bin is the percentage estimates of the probability from Pascal's triangle that a bead will land in that bin. By following the golden bead, you can clearly observe those probabilities with each flip of the Galton Board. When a Stock Market Clip-on is in place, the golden bead can represent the likely range and probabilities of next month's stock market return. The Galton Board's probabilities as to which bin the golden bead will land in is a substitute for the prediction of stock market forecasters.

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, Bernoulli (1655-1705) trials, 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 Galton's Words

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 135 year-old book. The text has been slightly 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 Illustration 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 original board

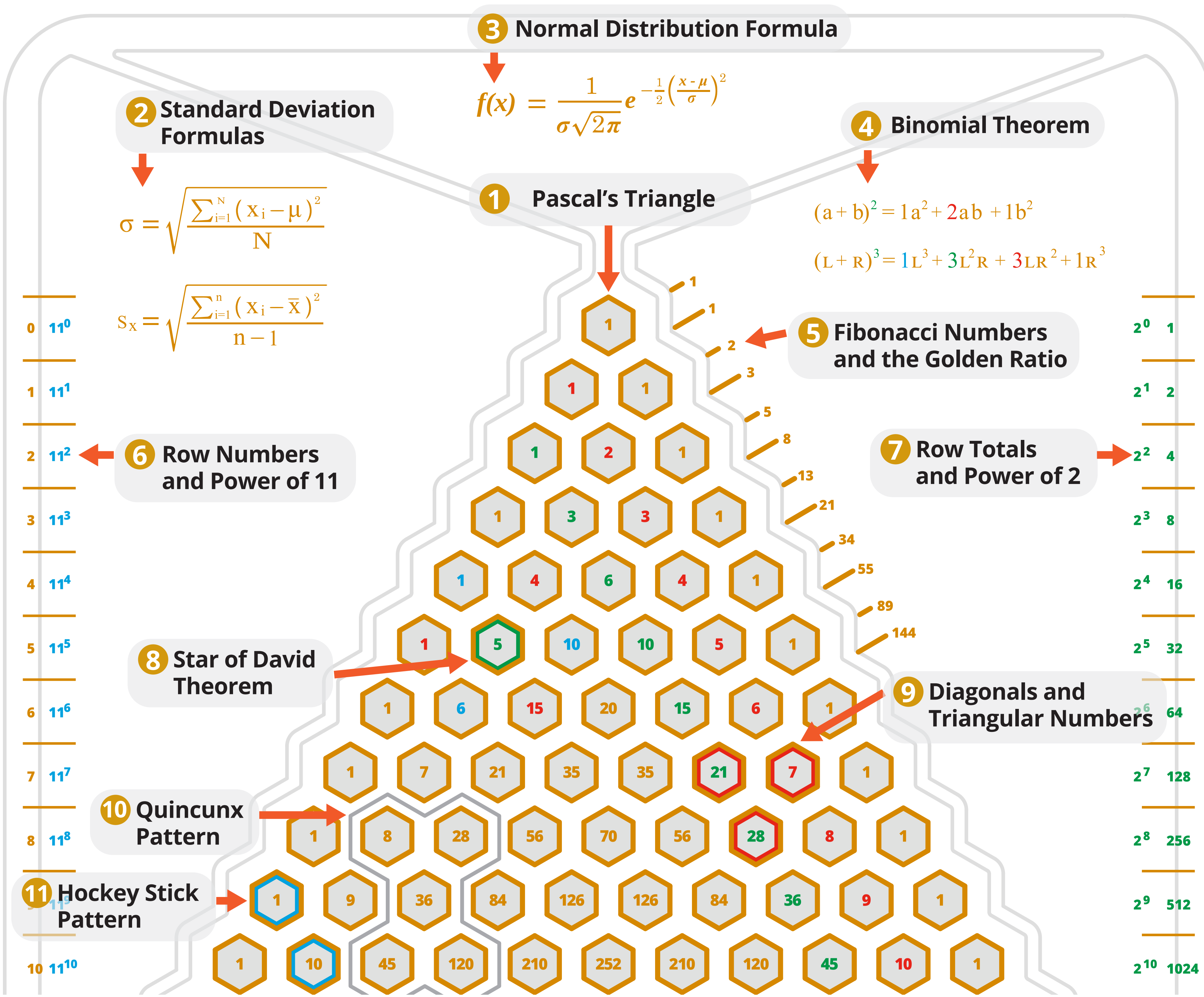
Our design of the Galton board is constructed of an anti-static plastic frame. A bead reservoir is designed into the top of the board. Below the outlet of the funnel stands a succession of 14 rows of hexagons, similar to Galton's pegs, stuck squarely into the back of the board, and below these again are a series of 15 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 [hexagons] 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. 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."

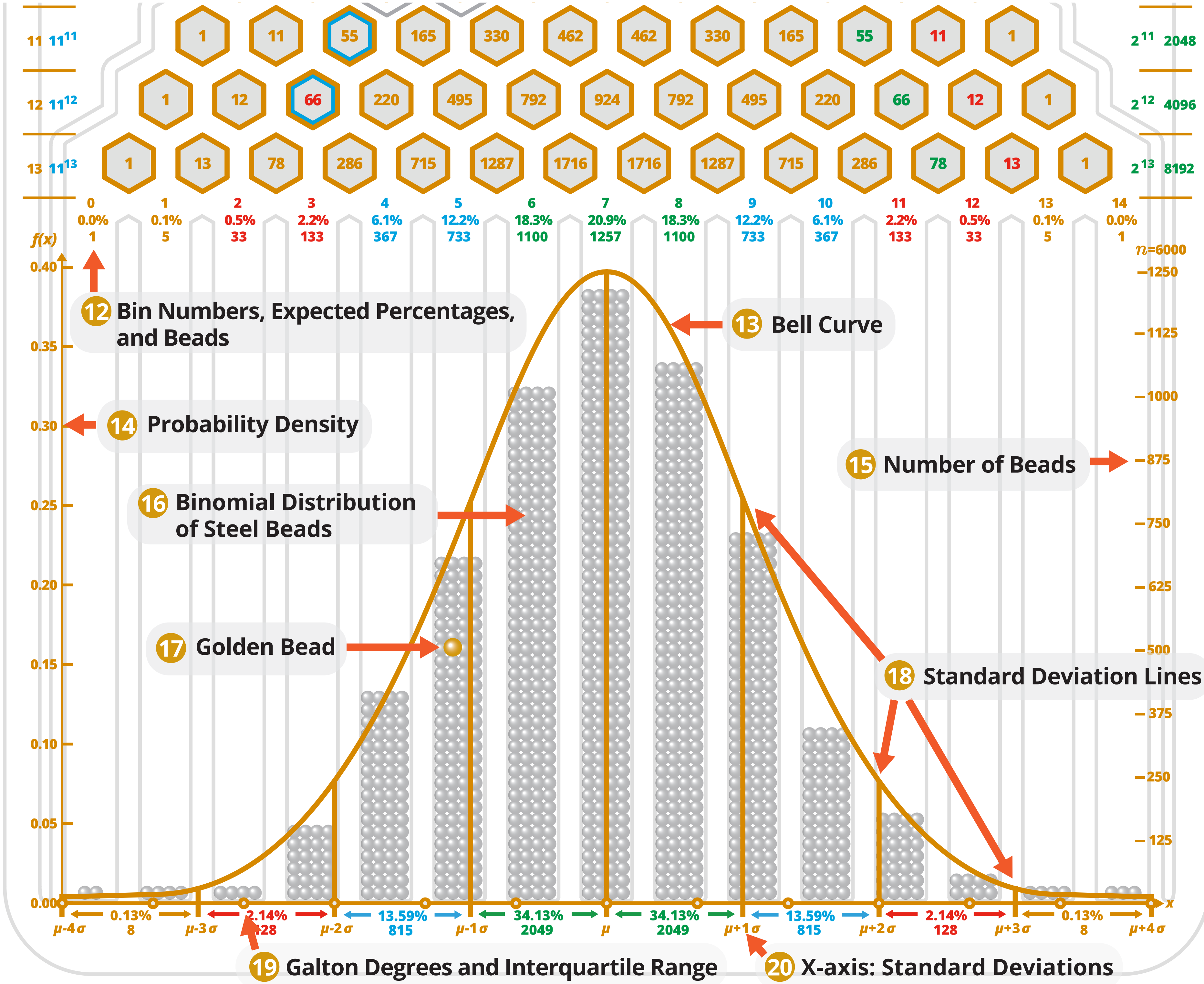
"The principle on which the action of the apparatus depends is, that a number of small and independent accidents befall each bead in its career. In rare cases, a long run of luck continues to favor the course of a particular bead towards either outside bin, but in the large majority of instances the 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."

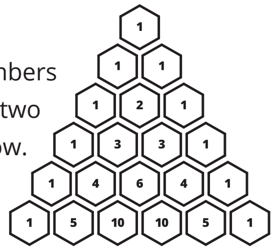


3 Normal Distribution Formula
$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
2 Standard Deviation Formulas
$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$
$S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$
4 Binomial Theorem
$(a+b)^2 = 1a^2 + 2ab + 1b^2$
$(L+R)^3 = 1L^3 + 3L^2R + 3LR^2 + 1R^3$
1 Pascal's Triangle
5 Fibonacci Numbers and the Golden Ratio
1 1 2 3 5 8 13 21 34 55 89 144
6 Row Numbers and Power of 11
7 Row Totals and Power of 2
8 Star of David Theorem
9 Diagonals and Triangular Numbers
10 Quincunx Pattern
11 Hockey Stick Pattern
0 11^0
1 11^1
2 11^2
3 11^3
4 11^4
5 11^5
6 11^6
7 11^7
8 11^8
10 11^{10}
2^0 1
2^1 2
2^2 4
2^3 8
2^4 16
2^5 32
64
2^7 128
2^8 256
2^9 512
2^{10} 1024
1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1



12 Bin Numbers, Expected Percentages, and Beads
13 Bell Curve
14 Probability Density
15 Number of Beads
16 Binomial Distribution of Steel Beads
17 Golden Bead
18 Standard Deviation Lines
19 Galton Degrees and Interquartile Range
20 X-axis: Standard Deviations
f(x)
n=6000
μ-4σ μ-3σ μ-2σ μ-1σ μ μ+1σ μ+2σ μ+3σ μ+4σ

1 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 used the triangle to study probability theory, as described in his mathematical treatise *Traité du triangle arithmétique* (1665). Other mathematicians studied it centuries before him in Persia, India, China, Germany, and Italy. The triangle's patterns translate to mathematical properties of the binomial coefficients. When placed on the Galton Board, each number on a hexagon represents the number of paths a bead can take to reach that hexagon.

2 Normal Distribution Formula

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 iid central limit theorem states that the random variable x will be normally distributed as the sample size becomes large and sigma (σ) is finite.

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

3 Standard Deviation Formulas

A standard deviation is a measure of how dispersed the data is in relation to the average (mean). To calculate the standard deviation of a sample data set of size n, follow these steps:

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 the n samples of x in the normal distribution formula.
3. Square each of the differences ($x_i - \bar{x}$) 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 for the population of size N.

Standard Deviation of a Sample	**Standard Deviation of the Population**
$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$	$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$

4 Binomial Theorem

The binomial theorem describes the algebraic expansion of powers of a binomial. Pascal's triangle defines the coefficients that 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 $(L + R)^n$.

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

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

The expansion of $(a + b)^n$ is $(a + b)^n = x_0 a^n + x_1 a^{n-1}b + x_2 a^{n-2}b^2 + \ldots + x_{n-1} ab^{n-1} + x_n b^n$ 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 on the board for $(a + b)^n$ for $n = 2$ and $(L + R)^n$ for $n = 3$.

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 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 (φ) of 1.61803398... but never equal it. For example: 55/34=1.618; 89/55=1.618; and 144/89=1.618. The rectangular dimensions of this Galton board (173mm x 280mm) approach the golden ratio of 1.618. The golden ratio was first defined by Euclid in his book *Elements*, written in 300 B.C. Leonardo Da

Vinci used the ratio to construct his masterpieces. The equation for the golden ratio is:

$$\varphi = \frac{1+\sqrt{5}}{2}$$

6 Row Numbers and Power of 11

On the left side, the fourteen rows of the Pascal's triangle are numbered, with the first row designated as n=0 and first entry in each row designated as k=0. Fourteen rows are large enough so the resulting binomial distribution is a good discrete approximation to the continuous normal distribution.

If you collapse each row into a single number by taking each element as a digit (and carry over to the left if the element has more than one digit), you get the power of eleven: 1, 11, 121, 1331, 14641... which matches to the numbers in the Pascal's triangle of that row.

7 Row Totals and Power of 2

The sum of numbers in a row is equal to 2^n where n equals the row number. For example, at row three, summing up the Pascal's numbers, 1 + 3 + 3 + 1 = 8, which also equals to 2^3.

The sum of the numbers in each row is also shown next to the power of two, 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 four ($1^2 + 4^2 + 6^2 + 4^2 + 1^2$), that equals seventy, which is also the middle entry of row eight.

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 = 60.

9 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

10 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.

11 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.

12 Bin Numbers, Expected Percentages, and Beads

The 15 bead bins are numbered from 0 to 14 so the location of the golden bead can be easily identified and recorded. Also, the probabilities from Pascal's triangle of a random outcome occurring within a certain bin can be identified by imagining a 15th row of the triangle (n=14).

Expected percentages of outcomes per bin for both beads and stock market returns are shown just below the bin number, with 20.9 percent expected in middle bin (#7). Then, the expected number of beads per bin are shown based on 6,000 beads.

13 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, large samples of monthly returns of the stock market indexes, and the beads in the Galton board.

14 Probability Density

The probability density $f(x)$ is the relationship between observations and their probability. It defines the probability of the occurrence of a random variable occurring within a particular range of continuous random variables. One 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 Number of Beads

The right Y-axis provides an estimate of the number of the beads in each bin.

16 Binomial Distribution of Steel Beads

Each steel bead represents an independent identically distributed (iid) random variable that falls from the reservoir through a fixed pattern of hexagons. A binomial distribution is created by the 6000 beads from the 14 Bernoulli trials for each bead, one trial for each hexagon hit. The discrete binomial distribution of beads closely approximates the continuous normal distribution.

17 Golden Bead

Among the 6,000 1mm steel beads is a 2.2mm golden bead. This bead demonstrates a single random outcome.

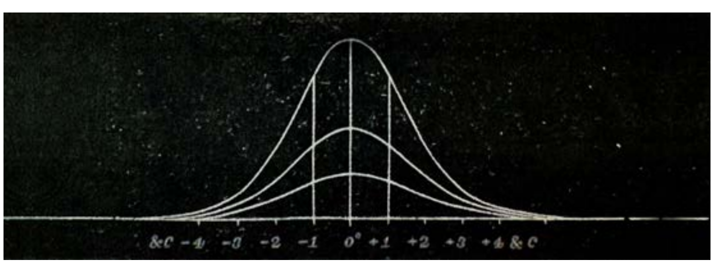
18 Standard Deviation Lines

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 percent of the data in a normal distribution falls within one standard deviation of the mean. About 95 percent falls within two standard deviations, about 99.7 percent falls within three standard deviations, and about 99.99 percent falls within four standard deviations. With 14 rows of hexagons in Pascal's triangle, there are 14 hexagons in the bottom row of the triangle. There are 15 bins, with bins on each end and between each hexagon. These 15 bins represent a total of $2 \times 15/\sqrt{14} = 8.0$ distribution standard deviations ($\mu \pm 4\sigma$). Each bin equals 0.533 standard deviations and each standard deviation equals 1.875 bins (8/15 = 0.533 or 15/8 = 1.875).

19 Galton Degrees and Interquartile Range

There are several ways to "divvy up" a distribution of data. In addition to the standard deviation (std dev), Sir Francis Galton suggested a degree scale for the x-axis, as shown by the small circles that represent and resemble degree symbols on the x-axis on our board. The Galton degrees are based on the statistical definition of the probable error, where each degree equals 0.6745 std devs for a normal distribution. The probable error defines the half-range of an interval around the central point of a distribution, such that half of the values will lie within the interval and half will lie outside. That interval of plus or minus 1 degree around 0 (the mean) in a normal distribution is 1.35 std devs wide and contains the middle 50%. Remember that 2 std devs contains the middle 68.3% of the normal distribution.

Below is Galton's original drawing of his degrees from 1877.



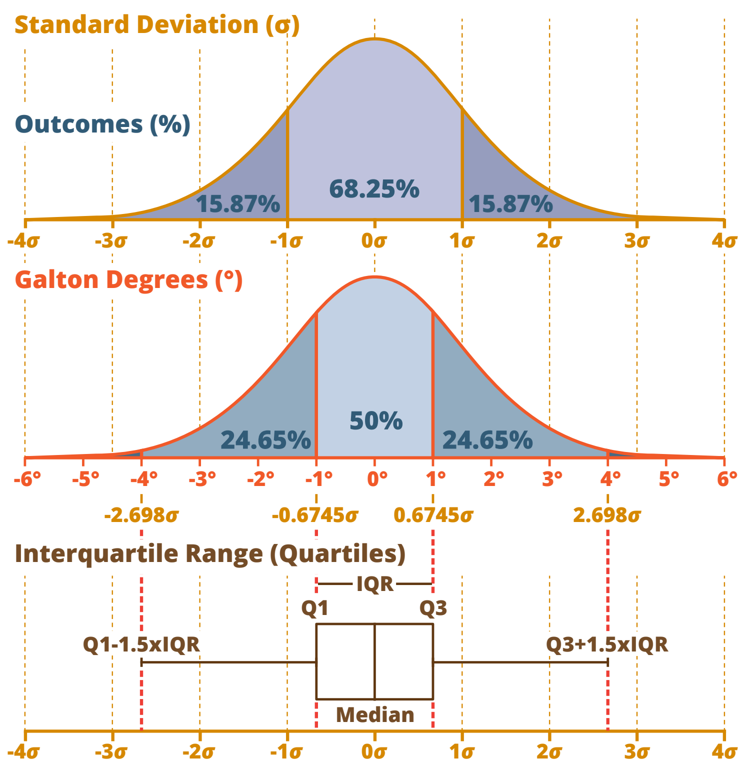
Another method to divvy up the distribution is the Interquartile Range (IQR), which measures dispersion of the data by dividing the distribution into quartiles or 25% increments using the median of the data instead of the mean. The IQR is defined as the difference between the 75th (Q3) and 25th (Q1) percentiles of the data. It is used in many statistical applications to locate where the middle 50% of the data exists without being impacted by outliers, which are data beyond IQR-related upper and lower bounds. The IQR is widely used in the financial industry to analyze the spread of data related to financial metrics, such as stock prices, exchange rates, and economic indicators.

The table below summarizes the different methods discussed above.

														TOTAL
Galton Degrees (°)	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	12
Normal Distribution (σ)	4.0	3.4	2.7	2.0	1.3	0.7	0	0.7	1.3	2.0	2.7	3.4	4.0	8.0
Outcomes b/w Ranges (%)	0.0	0.3	1.8	6.7	16.1	25.0	0	25.0	16.1	6.7	1.8	0.3	0.0	100
# of Beads Expected	2	19	108	403	968	1500	0	1500	968	403	108	19	2	6000
Interquartile Range						Q1	Q2	Q3						

Since there are 15 bead bins and 12 Galton degree bins in this Galton Board, each degree bin equals 1.25 bead bins.

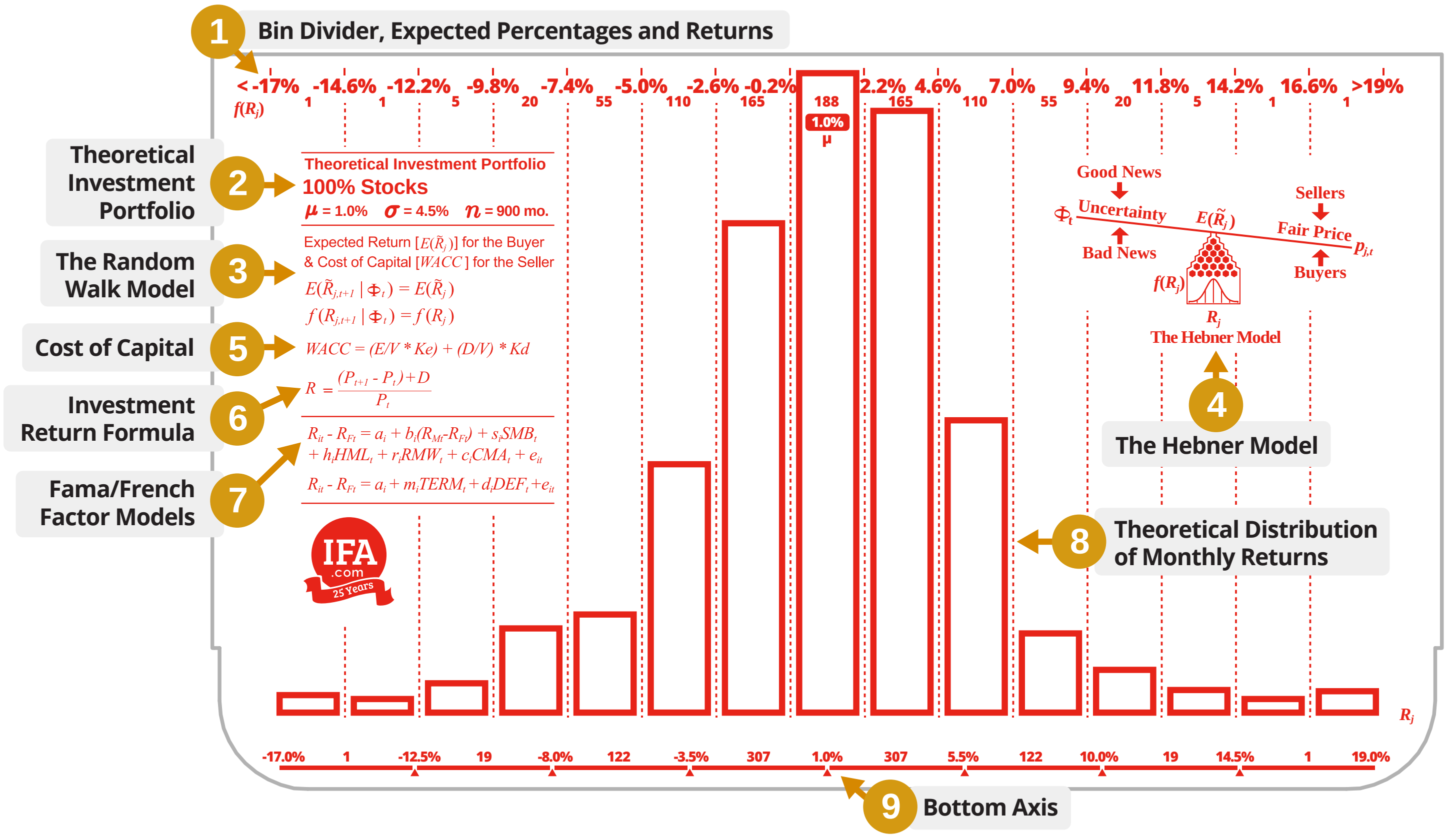
The three diagrams below show the bell curve divided up by standard deviation, Galton Degrees and Interquartile Range.

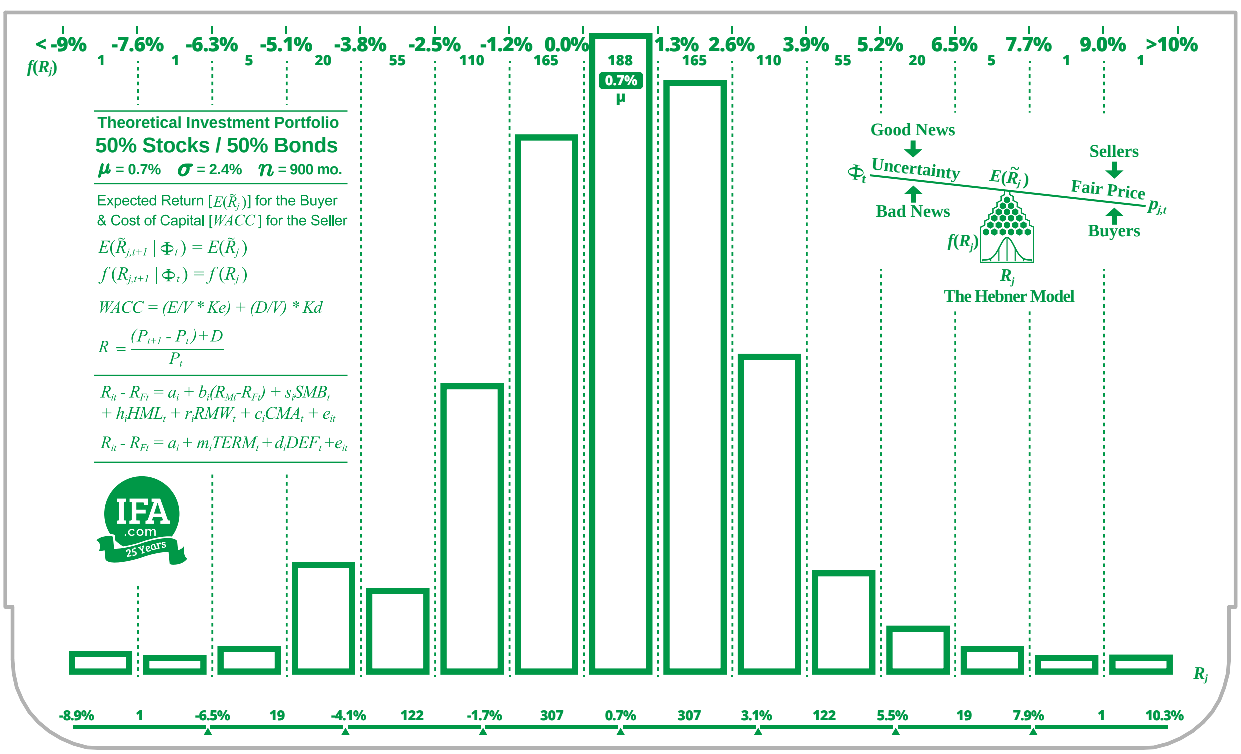


20 X-axis: Standard Deviations

The X-axis label provides information grouped by standard deviation. Labeled under each standard deviation line is its number of standard deviations from the mean up to 4 std devs ($\mu \pm 4\sigma$). Between each std dev line is the percentage of outcome that would be expected for that area of the bell curve.

Items of Interest on the Stock Market Clip-Ons





To apply a Clip-on, snap it onto the bottom portion of the main board. Use the bottom edge as a guide. To take off or switch the Clip-on, gently pry off the two sides of the Clip-on.

1 Bin Dividers with Expected Percentages and Returns

The four standard deviation range of expected monthly returns is shown on the top of bin dividers based on 900 monthly returns for the theoretical investment portfolio. The range changes according the standard deviation of the allocation of the investment portfolio. The bin dividers are scaled in such a way that the boundary of the board corresponds to the four standard deviations of returns (≈99.99 percent of outcomes or $\mu \pm 4\sigma$), with about two monthly returns expected in each tail beyond four standard deviations. The second row are estimates of the number of monthly returns expected per bin based on 900 monthly returns.

2 Investment Portfolio Illustrations

To represent market returns, we selected two theoretical investment portfolios, one aggressive and one moderate.

The red Clip-on illustrates a 100% stock portfolio (aggressive), which we assume to have a monthly average return of 1.0% and a standard deviation of 4.5%, with a sample size of 900 months. With four standard deviations, this results in a range of returns from -17% to 19%. This means with 15 bins, the return range per bin is 2.4%, with the mean of 1.0% right in the center.

The green Clip-on illustrates a theoretical 50% stock / 50% bond portfolio (moderate), which we assume to have a monthly average return of 0.7% and a standard deviation of 2.5%, with a sample size of 900 months. In keeping with the 15 bins, the range is now a tighter band of -9.3% to 10.7%, a return range of 1.33% per bin and with 0.7% mean in the center.

3 The Random Walk Model

The efficient market hypothesis states that the current price ($p_{j,t}$) of a security (j) 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 hypotheses 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 time (t)." If we assume that the expected return on a security is constant over time, we have

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

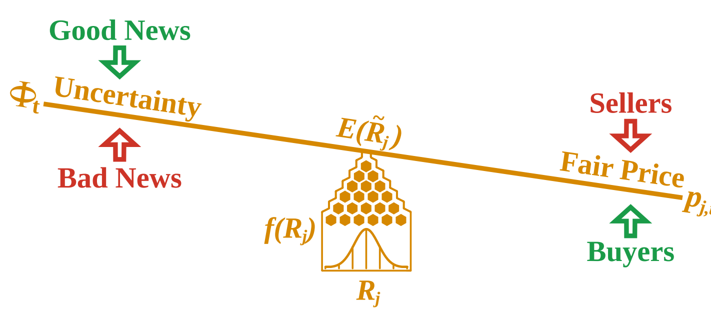
Source: *The Theory of Finance*, Eugene F. Fama & Merton H. Miller, 1972, pg. 339

4 The Hebner Model

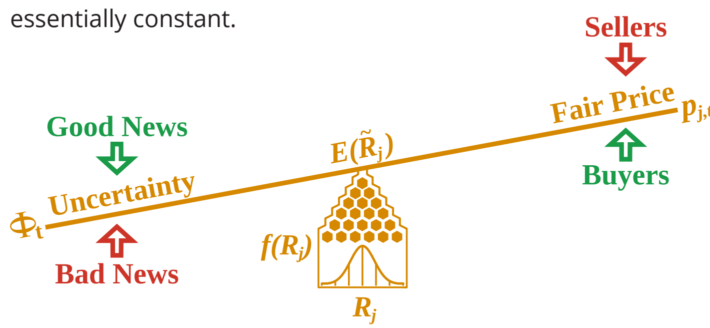
The teeter-totters below illustrate Eugene Fama's efficient market hypothesis, which states that prices of securities (j) fully reflect all available information resulting in fair prices. The left side of the teeter-totter represents whatever set of information (Φ_t) is assumed to be fully reflected in the price at that time (t) and the right side represents the prices ($p_{j,t}$) that millions of willing buyers and sellers have concluded are fair prices given the set of information at that time. 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 exchange of cash for securities.

The three components of the model occurred to Mark Hebner during the Global Financial Crisis of 2008. It starts with the teeter-totter placed at the top of the Pascal's triangle. Then, the beads bouncing around and through an array of hexagons represents the randomness of monthly stock market returns ($R_{j,t+1}$). Thirdly, the beads land in the bins representing the realized returns (R_j), which, in large samples, resemble the bell curve ($f(R_j)$).

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 ($E(\tilde{R}_j)$) that is held at a constant level of risk. If uncertainty increases due to bad news, the price must make a proportional decrease so that the expected return remains essentially constant.



If uncertainty decreases due to good news, the price must make a proportional increase so that the expected return remains essentially constant.



This model is known as the Hebner model and should be thought of as a framework for incorporating the Galton board and Pascal's triangle into how markets work.

5 Cost of Capital

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 also 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.

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

E is the market value of the firm's equity.
V is the total market value of equity and debt, or $E+D$.
Ke is the cost of equity.
D is the market value of the firm's debt.
Kd is the cost of debt.
$WACC$ is the weighted average cost of capital.
Just as a reminder, the expected return of the buyer is also the cost of capital for the seller ($E(\tilde{R}_{j,t}) = WACC$).

6 Investment Return Formula

The formula for an investment's realized 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 (D), divided by the original price (P_t) of the investment.

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

7 Fama/French Factor Models

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 percent and 94 percent of the cross-section variance of expected returns for diversified portfolios of five factors in equities. It expands on the CAPM (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).

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt}\text{-}R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it}$$

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 a 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 BtM 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.

Fama/French Two-Factor Model for Fixed Income

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 $LTG\text{-}RF$, 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 $CB\text{-}LTG$, where CB is the return on a proxy for the market portfolio of corporate bonds. Finally, e_{it} is the error term and is a zero-mean residual. Here is the equation:

$$R_{it} - R_{Ft} = a_i + m_i TERM_t + d_i DEF_t + e_{it}$$

8 Estimated Distribution of Monthly Returns

The red or green bars printed on the Clip-ons over the bell curve represent a histogram of the distribution of 900 monthly returns of a theoretical investment portfolio. The bars look similar because the returns scale has been altered to fit the portfolio and the boundary limits of the board.

... -7.4% -5.0% -2.6% -0.2% 2.2% 4.6% 7.0% 9.4% ...

... -3.8% -2.5% -1.2% 0.0% 1.3% 2.6% 3.9% 5.2% ...

9 Bottom Axis

The bottom axis shows percentage expected return at each

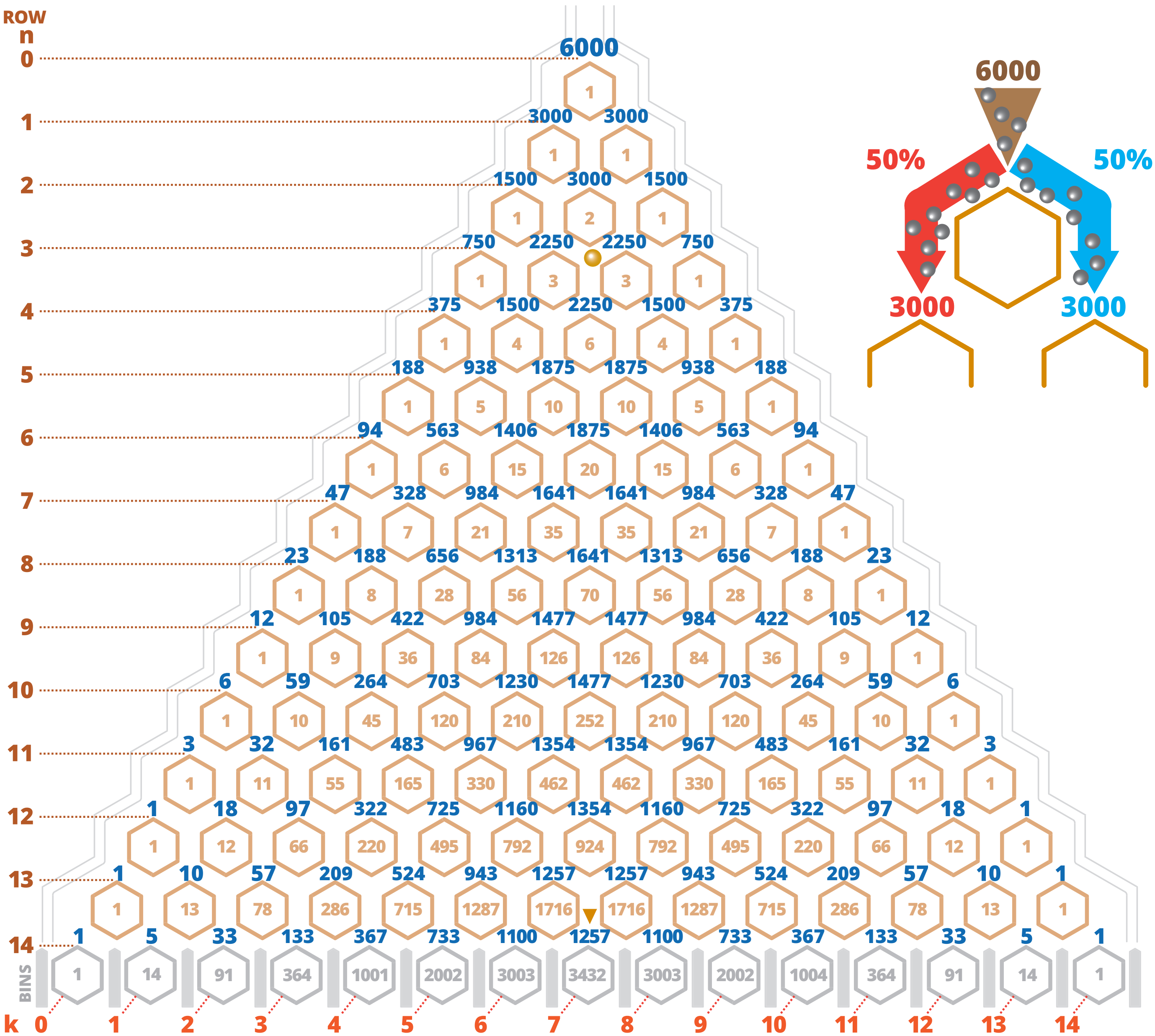
... -8.0% 122 -3.5% 307 1.0% 307 5.5% 122 10.0% ...

... -4.1% 122 -1.7% 307 0.7% 307 3.1% 122 5.5% ...

standard deviation. Shown between each standard deviation line are the expected number of monthly returns based on 900 months.

Symmetric Binomial Distribution of Beads

For a level Galton Board, there is an equal chance the beads will go either left or right at the top of each hexagon. This is an example of a Bernoulli trial. This illustration shows the expected number of beads that will travel through between each hexagon. There are approximately 6,000 beads in the bead reservoir. At the first hexagon, which is considered row 0, 3,000 beads are expected to go left and 3,000 beads expected to go right. If you follow the splitting of the beads each time you can see how many beads are expected to land in each bin after row 13. The number on all of the hexagons of Pascal's triangle can be interpreted as the number of paths to get to the **k**th location of row **n**. For example, for row 4, the numbers on the hexagons are **1, 4, 6, 4, 1** . If we add those numbers, we get a total of 16 paths to arrive at all 5 of hexagons in row 4. This also is 2 to the power of the row number ($2^4 = 16$). To determine the number of beads to arrive at the top of the middle hexagon of row 4 (**k** = 2: see golden bead), you would divide the number on that hexagon by the total of all the numbers on the hexagons in that row. So, 6/16 = 0.375 or 37.5%. Take that percentage times the total number of beads (6,000) and you get 2,250, as shown by the blue number above the middle hexagon in row 4.



Row 14 (in gray) of Pascal's triangle can be used to determine the probabilities (a symmetric binomial distribution) for a bead to fall into each of the 15 bins at the bottom of Galton board. Following the calculation for **n**=4 above, the expected percentage in the center bin (**k** = 7) of row 14 would be 3,432/16,384 = 20.9%. With 6,000 beads, that would mean 1,257 beads are expected to fall into that bin. If there were 16,384 beads, the numbers on each hexagon in row 14 would equal the beads expected to land in each bin.

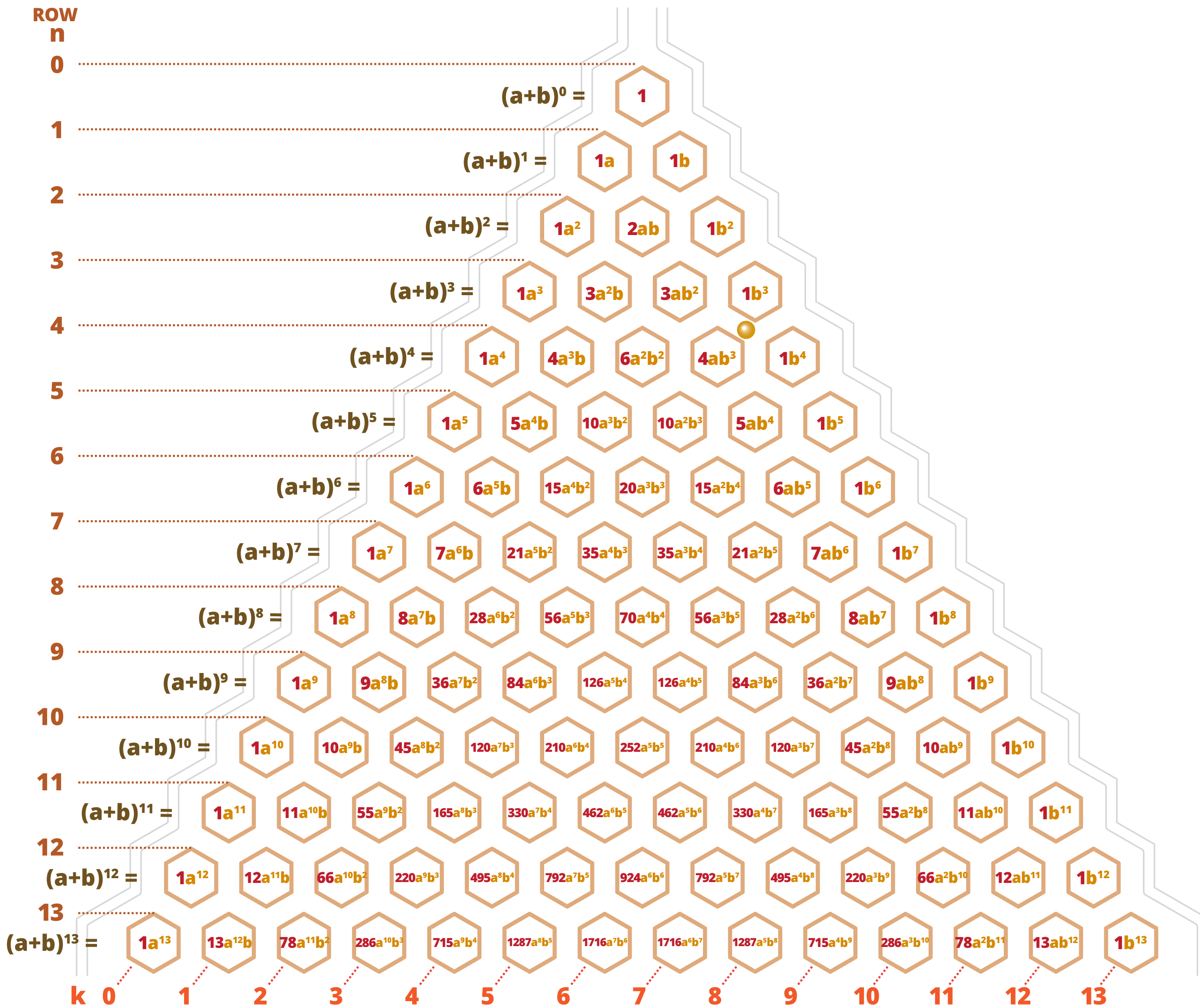
Binomial Coefficients, Expansion, and Distribution

The numbers in Pascal's triangle are the binomial coefficients of a binomial expansion, which is the algebraic expansion of a binomial raised to a power of **n**, for example **$(a+b)^n$**. For example, **$(a+b)^4$ = $1a^4$ + $4a^3b$ + $6a^2b^2$ + $4ab^3$ + $1b^4$** where **1, 4, 6, 4, 1** are the binomial coefficients shown in Pascal's triangle row 4. The general binomial formula for an arbitrary nonnegative integer **n** is given by

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

where $\binom{n}{k}$ are the binomial coefficients defined on the next page. The expected bead distribution for a level Galton Board is a symmetric binomial distribution defined by the normed binomial coefficients $\binom{n}{k}/2^n$.

Since this Galton Board has 14 rows of hexagons, the normal distribution can be used to approximate the binomial distribution (the de Moivre-Laplace theorem).



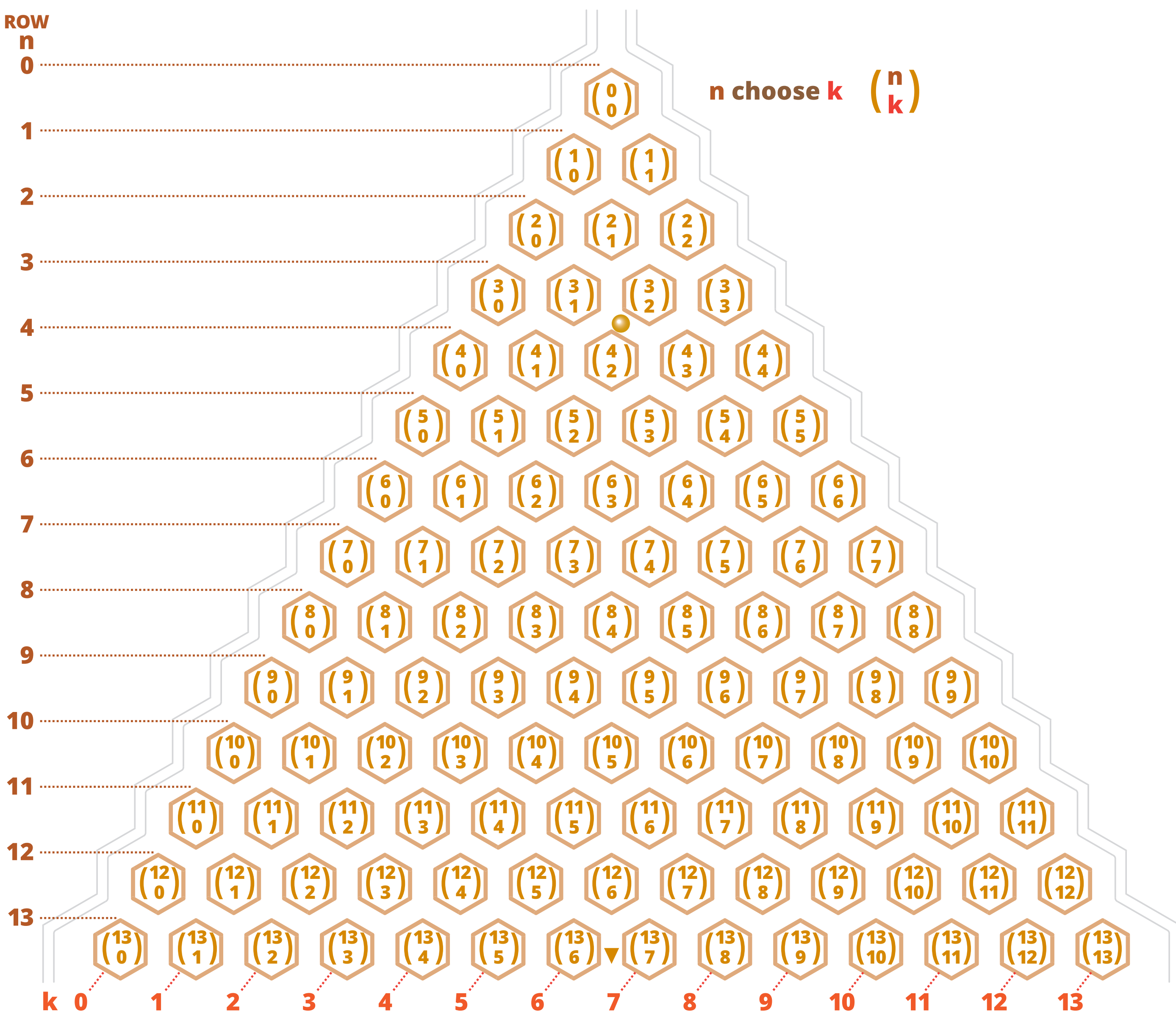
Combinatorics

The numbers in Pascal's triangle can also be used in combinatorics, which is the study of combinations and permutations. The figure shows **n** is Pascal's triangle row number and **k** (**k** = 0 to **n**) is the element location in row **n**. The number of different combinations **C** of n things taken **k** at a time (**n** choose **k**) is given by

$$C(n,k) = C^{n}_{k} = {}_{n}C_{k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

For example, if **k** is defined as the number of moves to the right that a bead makes after **n** hexagon interactions, how many ways can **k** right moves occur from **n** hexagon interactions?

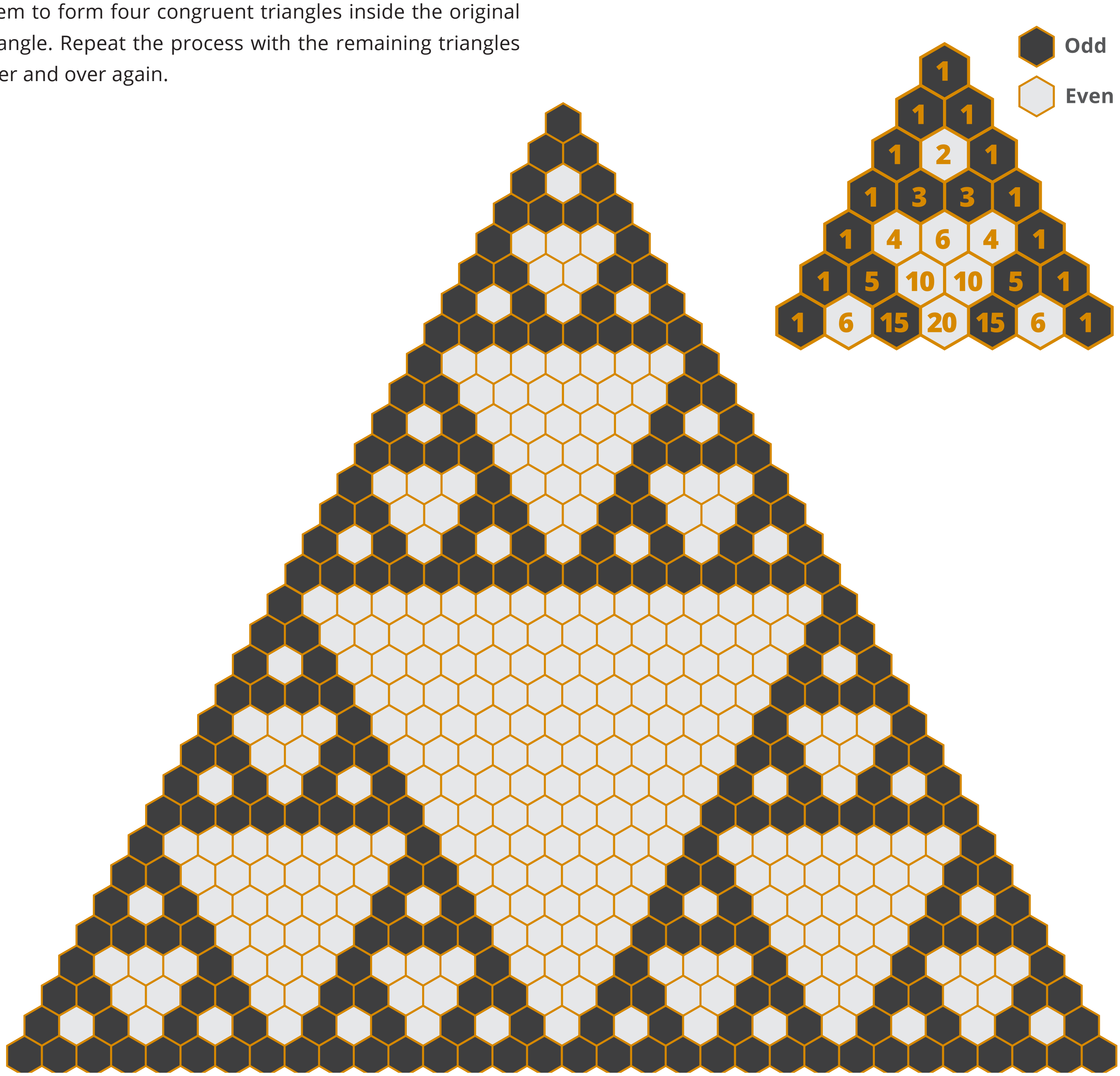
The answer for **n** choose **k** is given by $\binom{n}{k}$. If **n** = 4, then the 4 choose **k** results in Pascal's triangle row 4 are **1, 4, 6, 4, 1** for **k**'s from 0 to 4. There is only one path that a bead can take for **k** = 0 or 4 and four paths for both **k** = 1 or 3. There are six paths for **k** = 2 (see golden bead location). Pascal's triangle figure below shows the **n** choose **k** values.



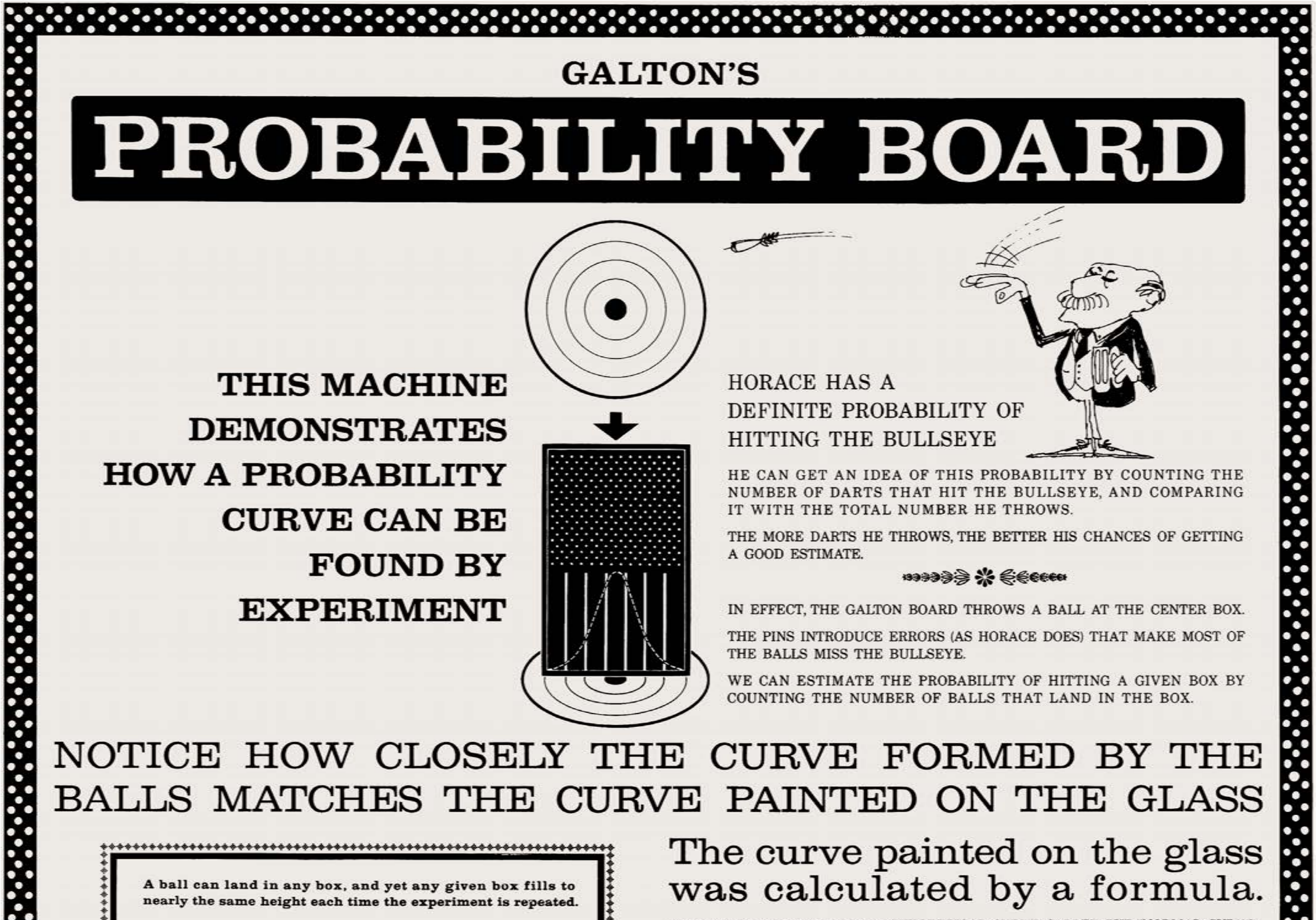
The Sierpinski Triangle

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.

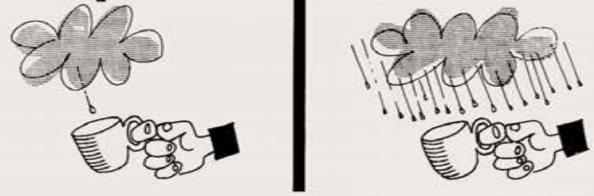
If you start with a Pascal's triangle and color the odd numbers black and leave the even numbers white, it will closely resemble the Sierpinski triangle, which is named after the Polish mathematician Waclaw Sierpinski (1882-1969). Sierpinski's work included three well known fractals, including the triangle,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



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



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."

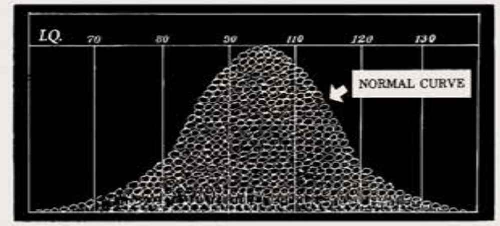
"RELATIVE FREQUENCY" is the number of times an event* occurs divided by the number of trials.**

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

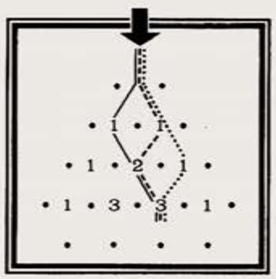
RUN AT ROULETTE

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.

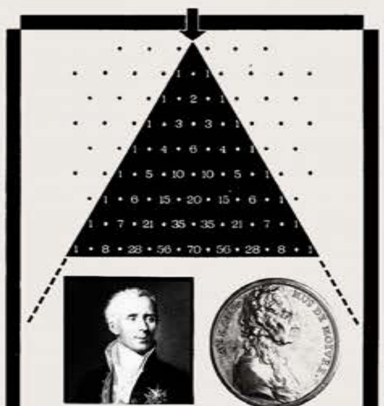
The average of many independent errors almost always leads to the Normal Curve, no matter what the underlying process may be.

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.

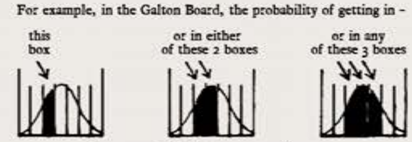
PROBABILITY IS ADDITIVE

IF THE PROBABILITY OF "5" IS $\frac{1}{6}$ AND THE PROBABILITY OF "4" IS $\frac{1}{6}$

THEN THE PROBABILITY OF "5" 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 –



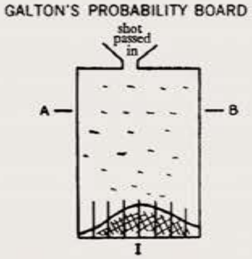
is represented by the corresponding area.

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

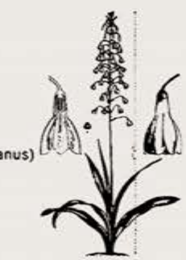
THE QUINQUNX

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

My Fascination with the Galton Board

Mark T. Hebner

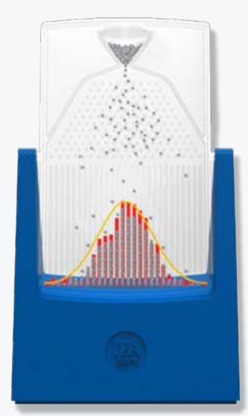
My name is Mark T. Hebner, and I am the CEO and founder of Index Fund Advisors, Inc. (IFA.com). My firm is in the wealth management and tax preparation business. I am also the creator of this Galton board.

The most common way to describe the risk and return of an investment is to estimate its average return and standard deviation of return from a large sample of historical returns, something like 900 months (75 years) of index data. If you want Excel to draw a bell curve, you only need the average and the standard deviation. They define the bell curve. As it turns out, Harry Markowitz's Nobel Prized winning scatter plot of average return versus standard deviation was just a comparison of bell curves. So imagine my excitement when I found a physical device that generates a bell curve. I realized it is a powerful demonstration of how markets work and the probability of a range of different outcomes. It also occurred to me that the Galton board simulates monthly investment returns and allows people to see the constant expected returns, the randomness of returns over a period of thirty days, and the resulting bell curve of the realized returns over very long periods. Put simply, this device helps investors understand critical investing ideas.

My fascination with the Galton board was ignited back in 2005 when I saw an Eames Office film about the 1964 World's Fair. Charles Eames built an outdoor fourteen-and-a-half-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 and sponsored by IBM. It was designed for the 1961 opening of a new science wing at the California Museum of Science and Industry in Los Angeles.

My first Galton board, shown on the right, was designed and built by the Oregon Museum of Science and Industry. The photograph depicts an eight-foot-tall by four-foot-wide museum-quality probability demonstrator that I commissioned in 2009 to educate investors about the range, probability, and shape of outcomes that result from a series of random events. This Galton board sits in the lobby of Index Fund Advisors' office and helps to portray order in the midst of chaos that is the random walk of Wall Street. The red bars behind the beads represent a large sample of monthly returns of a theoretical investment portfolio and allow the comparison of the beads to the stock market. In the stock market, random events are the news stories about a company or about capitalism in general and the prices of securities that reflect this information. The random flow of the beads, starting from a central point, simulates a series of fair prices, ultimately forming a normal distribution of monthly returns in the shape of a bell curve.

With the help of Philip Poissant, Jerry Xu, Art Forster, Jackson Lin, Mike Auchterlonie, the Brunson family, and others, I created my first seven-and-a-half-inch-tall desktop-sized Galton board, called The Random Walker®, in 2015. This compact version of the Galton Board is not only a helpful educational tool for understanding statistical concepts and stock market randomness, but also a delightful desktop device to play with. With an innovative flip-n-reset design, one can easily experience the order in chaos with just a tip of their finger. About 60,000 of these boards sit on desks all over the world.

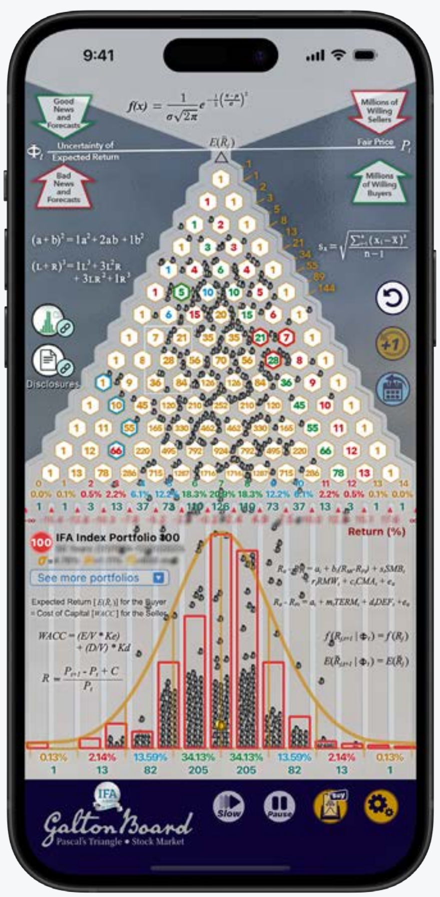


The Random Walker®

In 2022, I created this twelve-inch-tall version, that you find in this box. This much larger version is a lot easier to read, to demonstrate to groups of people, and to use in a classroom setting.

This version of the desk-top sized board incorporates many new improvements in design from the previous iteration. It more precisely captures the concepts of the binomial distribution and Pascal's triangle, along with the many embedded mathematical concepts. By adding the Clip-ons of monthly return data, one can see the incorporation of elements of the stock market, including the Hebner model, and how well they match to the bell curve of the beads.

To further promote the understanding of the principles embedded in the Galton Board and Pascal's triangle, I commissioned and created an app version of the Galton Board in 2023. This app version uses the gyrometer that allows you to turn the phone or iPad and hear and see the beads flow as if they were physical beads rolling around in your device. By tapping the settings icon, you can also overlay twenty index portfolio histograms and see the change in the returns scale of the bins as the risk changes. To get the app for iPhone and iPad, visit the Apple App Store, and search "Galton Board App" or "Index Fund Advisors". You may also visit the Mac App Store on your Mac laptop or desktop and search Galton Board app. Finally, visit the Google Play Store for android devices and search "Index Fund Advisors" and soon there will be a stand-alone "Galton Board App".



Get the App!

About Index Fund Advisors

Index Fund Advisors
WEALTH MANAGEMENT • TAXES

Index Fund Advisors, Inc. (IFA) is a fee-only advisory and wealth management firm that provides risk-appropriate, returns-optimized, globally-diversified and tax-minimized investment strategies with a fiduciary standard of care.

IFA is a registered investment adviser that provides investment advice to individuals, retirement plans, trusts, corporations, non-profits, and public and private institutions. IFA was founded in 1999, and celebrated its 25th anniversary in 2024. IFA provides investment advice to clients across United States.

The value of IFA extends beyond investment advice. As a holistic financial partner, IFA helps guide investors through all life stages. Our Wealth Advisors take a personalized approach to matching people with portfolios while offering a full-range of wealth management services and financial planning for a better overall client experience.

IFA seeks to avoid the futile, speculative, and unnecessary cost-generating activities of stock, time, manager, and style picking. Contrarily, IFA employs a disciplined, quantitative and low cost approach that emphasizes broad diversification and consistent exposure to the dimensions of returns of global securities.

IFA primarily bases its investment strategy on the highly respected research indexes designed by Eugene Fama and Kenneth French, incorporating 96 years of IFA Index Portfolio risk and return data, third generation index fund designs and 40 years of refined passive trading techniques, employed by Dimensional Fund Advisors.

To learn more, visit ifa.com

ifa.com
Mark T. Hebner
Founder and CEO
info@ifa.com
19200 Von Karman Ave. Suite 150
Irvine, CA 92612 | 888-643-3133

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

- Michael Stevens
Creator and host of YouTube channel Vsauce

D!NG VSAUCE

"This device is nothing short of wondrous - demonstrating elegantly an astonishing aspect of our universe. It is just beautiful in form and function, and the graphics along with the booklet is educational in a way that is missed in our college stat courses.
If I taught statistics I would require the purchase of this Galton Board!"

- Dr. Raymond Hall
Professor of Physics, California State University - Fresno
@physicsfun on Instagram and YouTube

IFA .com 25 Years

Galton Board
Probability Demonstrator

Visit GaltonBoard.com for videos, articles, photos and more information.

2023 WINNER EUROPEAN PRODUCT DESIGN AWARD

STEM.ORG AUTHENTICATED EDUCATIONAL PRODUCT

galtonboard.com

INCLUDES	One Galton Board Device (product and color may vary) X1; One User Guide X1; Two Stock Market Clip-ons X2
PACKAGE DIMENSIONS	14.41 x 9.13 x 4.80 inches (366 x 232 x 122 mm) 79.35 oz (2250 g)
UNIT DIMENSIONS	12.20 x 8.58 x 3.94 inches (310 x 218 x 100 mm) 36.69 oz (1040 g)

19200 Von Karman Ave Suite 150 Irvine, CA 92612 USA
888-643-3133 • #IFA-R32.1-GB-9-2024 • Made in China
 • Created by Mark T. Hebner

A special thanks to Sir Francis Galton, Blaise Pascal, Mark Hebner, Philip Poissant, Jerry Xu, Jackson Lin, Art Forster, Mike Auchterlonie, the Brunson family, Cat Phelps, Wes Long, Jesse Fulton, Robert Bray, James Duncan, Charles Eames, OMSI, Harry Markowitz, Beth Hebner, Tisley & Spiller and the many other people who contributed to the nearly 20-year long process that led to the design, manufacturing, development, review, QA, QC, IP protection and distribution of this Galton Board: Probability Demonstrator.