THE GAUSSIAN

If one constructs an array of numbers starting with 1 in the first row, ½ and ½ in the second row, and 1/4,1/2, and 1/4 in the third row, and then carries on the procedure indefinitely, a Pascal like triangle will result looking as follows-

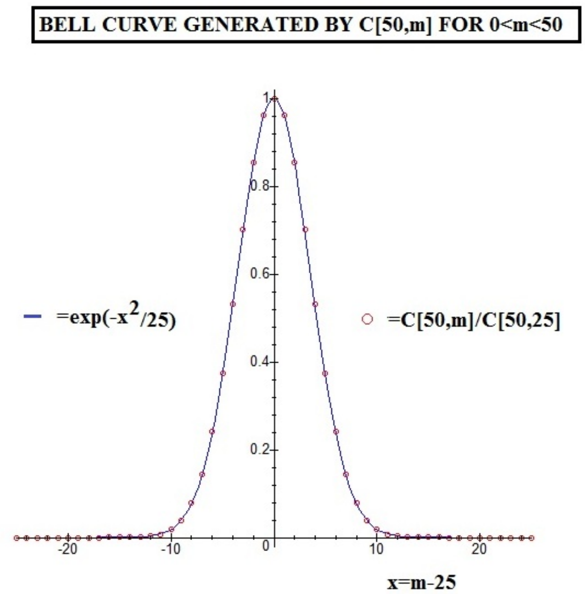
$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & \frac{1}{2} & & \frac{1}{2} & & & \\
 & & \frac{1}{4} & & \frac{1}{2} & & \frac{1}{4} & & \\
 & \frac{1}{8} & & \frac{3}{8} & & \frac{3}{8} & & \frac{1}{8} & \\
\frac{1}{16} & & \frac{1}{4} & & \frac{3}{8} & & \frac{1}{4} & & \frac{1}{16}
\end{array}
$$

Each row has its elements add up to exactly one and the element C[n,m] in the nth row and mth column is given by

$$C[n,m]=\frac{n!}{m!(n-m)!2^n}$$

We recognize this to be the standard binomial coefficient divided by the nth power of two. One has C[4,3]=4!/{3!(1!)16}=1/4 , C[6,0]=1/64, and C[50,25]= 0.1122751727… .

If we move down to row 50 and then make a point plot of the elements C[50,m] /C[50,25]over the range -50<m<50 we get the following peculiar looking bell shaped curve-



The red circles represent the normalized C[50,m]points. Superimposed on these we have placed the exponential curve y=exp(-x^2/25). The two curves almost coincide, with the discrepancy becoming less and less as n gets still larger. For this case the area under the curve for the red circles equals 1/C[50,25]=8.906.. and that under the blue exponential curve is 8.862. In the limit of n=infinity, the areas under the red-circle curve and the exponential curve should be equal. So if we divide things by C[50,25], the area under both curves will be one in that limit. This means-

$$\frac{\lim}{n\to\infty}\{1/C[n,n/2]\}=1=\int_{-\infty}^{\infty}\exp(-ax^2)dx=\sqrt{\frac{\pi}{a}}$$

Upon setting a=π, we get a new integral which, upon the substitution of z=xsqrt(a) , yields-

$$\frac{1}{\sqrt{\pi}}\int_{z=-\infty}^{\infty}\exp(-z^2)dz=1$$

From it one finds a new continuous function-

$$F(z)=\frac{1}{\sqrt{\pi}}\exp(-z^2)$$

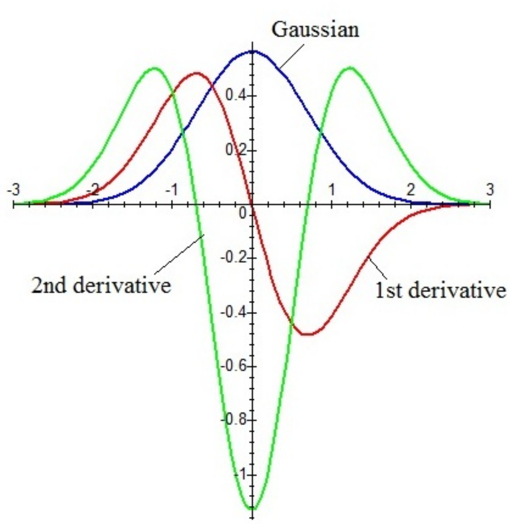
This is known as the **Gaussian normal distribution**, In view of the above derivation, one sees that this function is already contained in both the binomial theorem and the Pascal triangle. Both of these mathematical constructs were known to both Newton and Leibnitz some hundred years earlier. However, it was not until Gauss, in the early 18 hundreds , that the function was formalized. It plays a major role in statistics and probability theory including studies on life longevity and IQ distributions in populations. Also we have encountered the function in our partial differential equations class when studying the spreading of heat from a hot spot in an insulated bar of infinite length (see http://www2.mae.ufl.edu/~uhk/MATHFUNC.htm). I'll come back later in this article to discuss the related Probability Density Function which is just this Gaussian after a simple transformation..

When integrated over the infinite range -∞<z<+∞, the Gaussian yields an area of one and its first and second derivatives are-

$$\frac{dF(z)}{dz} = -\frac{2z}{\sqrt{\pi}}\exp(-z^2) \quad and \quad \frac{d^2F}{dz^2} = \frac{1}{\sqrt{\pi}}\{-2+4z^2\}\exp(-x^2)$$

A zero slope point occurs at z=0 and F(0)=1/sqrt(π) there. Inflection points occur at z= ±1/sqrt(2) . A graph of the Gaussian and its derivatives follows-

THE GAUSSIAN AND ITS DERIVATIVES



There are numerous integrals involving the Gaussian with many of these relating to error functions and Laplace transforms. If one sets $t=z^2$, then one can easily derive the values of the following integrals-

$$I = \int_0^\infty z^{2m}\exp(-z^2)dz = \frac{\lim}{s\to 1} laplace(t^m) = \Gamma(m+1)$$

$$J = \int_0^\infty \sin(at)\exp(-bt)dt = \frac{a}{a^2+b^2}$$

and

$$K = \int_0^M \exp(-x^2)dx = \frac{\sqrt{\pi}}{2}erf(M)$$

Let us next recall the form of the Gaussian function given above and introduce the new independent variable z=(w- μ)/{σsqrt(2)}. This allows us to write-

$$1 = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-z^2)dz = \frac{1}{\sigma\sqrt{2\pi}}\int_{w=-\infty}^{\infty}\exp(-\frac{(w-\mu)^2}{2\sigma^2})dw$$

The integrand in this last integral, namely,-

$$P(w) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(w-\mu)^2}{2\sigma^2}$$

represents the continuous **Probability Density Distribution** encountered in probability and statistics. Here σ^2 is the variance and μ the mean value of a set of continuous data input. The standard deviation σ is the distance from w=0 to the inflection point at z=1/sqrt(2).This means the area underneath the P(w) curve to the first standard deviation is-

$$\text{Area to } \sigma = \frac{1}{\sqrt{\pi}}\int_{0}^{1/\sqrt{2}}\exp(-z^2)dz = \frac{1}{2}erf(\frac{1}{\sqrt{2}}) = 0.3413447...$$

We can continue on and ask for the areas under the curve between σ and 2σ , and the area under the curve between 2σ and 3σ. This produces-

$$\text{Area from } \sigma \text{ to } 2\sigma = \frac{1}{2}\left\{erf(\sqrt{2}) - erf(\frac{1}{\sqrt{2}})\right\} = 0.1359051...$$

and

$$\text{Area from } 2\sigma \text{ to } 3\sigma \quad \frac{1}{2}\left\{erf(3/\sqrt{2}) - erf(\sqrt{2}) = 0.021400234...\right\}$$

Thus a departure from the mean at μ=0 by two standard deviations has a 1-(0.4772498/0.5)=4.55% chance. The problem with the Gaussian is that the values drop off too quickly and thus do not allow for black swan occurrences where unpredictable things do occur once in a while and are much more likely to do so than predicted by a Gaussian. An excellent book on this topic is “The Black Swan” by Nassim Taleb(2010).