The Supernormal Properties of Normal Distributions

Brian Hildebrandt

June 3, 2009

1 Introduction

The normal distribution emerged out of probability in the first half of the eighteenth century, introduced by Abraham de Moivre, a pioneer in the field. Gauss used it to prove validate the least squares method at the beginning of the nineteenth century, and it, as if the importance of itself and Gauss was a unifying factor, soon became known as the Gaussian distribution. While acting as central to both probability and data collection in general, it also managed to become important in matters like orthogonality, density, and in several real world applications. Its applicability to real world events became irrefutable with the advent of quantum mechanics, which expresses the world probabilistically. The normal distribution has been taken into several variables and generalized to create new families of functions. Today not only mathematicians but the entire intellectual world is at least basically acquainted with normal, skew-normal, and binomial distributions.

2 Basic Properties

By far the most impressive property of the normal distribution is the way that unaffiliated items seem to converge to it. A plot of random data will usually shape into a normal distribution with enough data points. A plot detailing the probability of sums from the rolls of several dice will converge to a Gaussian shape, as will probabilities of the number of instances of a 50% chance event when repeated several times (other probabilities will give skew-normal shaped distributions first). Additionally, the sum of two normal distributions is a normal distribution, and if a normal distribution results in the sum of two independent variable each is normally distributed. Additionally, a distribution is normal if and only if the variance of any part of it is independent of the mean.[1][2] The first cumulant, μ, of a normal distribution is its mean, and the

[1] For every mean one should be able to find any variance, and all variances for a single mean, in samples from the distribution.

second cumulant, σ^2, is the standard deviation squared.[2] A normal distribution is defined in that it has all cumulants after the second zero. These parameters give a normal distribution defined by $\frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(x-\mu)^2}{2\sigma^2}$.

3 Binomial/Skew-Normal Distributions

Binomial distributions are the actual distributions for several probabilistic phenomena. They represent the ferquency of an event's occurance at some cardinal values. They are the distribution recieved by a repeated event of a certain chance. They converge to normal distributions, but those with probabilities other than 0.5 converge first to skew-normal distributions. Skew-normal distributions are of the form $\frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})(1+\text{erf}(\frac{\lambda(x-\mu)}{\sigma\sqrt{2}}))$, where λ is the skew parameter and erf is the error function, a variation on an integral of the standard normal distribution.[1]

4 Perturbation of a Random Variable[4]

$Z = X + \sqrt{a}Y$ is our perturbed system with X and Y varying independent variables with all moments finite and $a \geq 0$.[3] The cumulants are then of the form $\kappa_r(Z) = \kappa_r(X) + a^{r/2}\kappa_r(Y)$.[4] Taking the derivative as a increases from zero (one is essentially 'adding' the perturbation to an unperturbed system), $\frac{d\kappa_r}{da}|_{a=0^+} = 0$, as long as the cumulant is above 2.[5] Then, take a derivative we know must be nonzero, the derivative of $\kappa_2(Z)^{r/2}$ with respect to a at zero, and combine it with the original to obtain a factor of $\kappa_r(X)$ from the chain rule, so that the derivative of the concatenated function is not trivial with respect to r.[6] In the paper this is done with $\frac{d}{da}\frac{\kappa_r(Z)}{\kappa_2(Z)^{r/2}}|_{a=0^+} = -\frac{r\kappa_2(Y)\kappa_r(X)}{\kappa_2(X)^{r/2+1}}$. By adding a perturbation you are causing the value of the higher function cumulants to decrease proportionally to the second cumulant, by a factor of $\frac{r\kappa_2(Y)}{\kappa_2(X)}$. Since r is always positive the cumulant ratios will always decrease, unless $\kappa_r(X)$ is zero, in which case the rate of change of its ratio with $\kappa_2(X)^{r/2}$ will also be zero. Furthermore, the higher cumulant ratios will decrease faster than the lower ones, and Y distributions with high standard deviation, as well as X distributions with low standard deviations will get faster rates of change. Since all cumulants ratios above the second are decreasing, the distribution is getting closer to a corresponding normal distribution for small perturbations. Furthermore, it is easily shown that this property is unique to normal distributions, and that

[2] n^{th} cumulants are defined as the n^{th} derivative of the cumulant generating function, the natural logarithm of the expected value of the exponential of a variable times the distribution, evaluated at zero.

[3] A moment of order n for X is the expected value of X^n.

[4] κ_r denotes the r^th cumulant. The fact that $\kappa_r(cX) = c^r\kappa_r(X)$ is a well known property of cumulants.

[5] At two it goes to $\kappa_2(Y)$, and the first cumulant's derivitive is not well behaved

[6] We know that the κ_2 of Z exists and is nonzero becuase we stipulated that the variables be varying, and the second cumulant is related to the standard deviation.

perturbation does not take cumulant ratios closer to any other distribution with finite cumulants. Think of any distribution with cumulants finite. A distribution halfway this distribution and a normal distribution will have lower cumulant ratios, and these will decrease at the beginning of the perturbation. Therefore, perturbation will actually cause the character of all other distributions with finite cumulants to 'slip away'.

5 The Normal Distribution

Even though the above perturbation property works only in additive perturbation, not necesserily in any other type, it is still an extremely strong result. Essentially, we can now conclude that the normal distribution is the mathematical embodiment of randomness and of small disturbances. It is extremely special and, if you believe quantum mechanics, is not only the building block of probability, but describes everything, to a certain extent. Why it has the properties that it does, and why those properties are so important is a bit of a mystery. One of the other amazing things about the normal distribution is how many different functions it seems to be related to. It seems to have just the right balance to swing between several functions. Its generality lends itself to the creation of families as well, so it can be easily extended to several dimensions, or be deregularized slightly into a group with still powerful properties.[3]

References

[1] Cheng-Hui Chang, Jyh-Jiaun Lin, Nabendu Pal, and Miao-Chen Chiang. "A Note on Improved Approximation of the Binomial Distribution by the Skew-Normal Distribution," *The American Statistician*, **62**(2008), 167–170.

[2] E. Lukacs. "A Characterization of the Normal Distribution," *Annals of Mathematical Statistics*, **13**(1942), 91–93.

[3] Frank Sinz, Sebastian Gerwinn, Matthias Bethge. "Characterization of the p-generalized normal distribution," *Journal of Multivariate Analysis*, **100**(2009), 817–820.

[4] Robin Willink. "A Unique Property of the Normal Distribution Associated With Perturbing a General Random Variable," *The American Statistician*, **62**(2008), 144–146.