Noise Measurement Tools

© Malcolm Moore
10-Dec-2003

Data Analysis – the Hidden Wealth

As a Senior Manager / Engineer working in Telstra, one of my roles was to work with a team of specialists that was managing customer complaints in regards to service standards (failed calls and the performance of matured calls). While in this role, the technology of the network had radically changed and common channel signalling (CCS7) was introduced, making it virtually impossible for the team to physically trace call connections as had been done for decades before.

After analysing the data from these new switches for some months we found out that this network generated data had a wealth of information in it. It took the team several months to develop in-house software that analysed this data on the fly and identified failed calls.

We could then identify and act on network faults before most customers were even aware or had complained. With this newly found technical data analysis procedure we could initiate corrections to data tables virtually anywhere in the network and this approach turned a torrent of complaints into a trickle inside a few years.

In hindsight this was world-leading technology that we developed and it increased the speed of analysis by an incredible amount. It could crunch through signalling sequences that controlled more than 30 million calls per day, every day, all hours, compared to following up on a few hundred customer reported calls per work day.

The analogy of this short story was pivotal in my thinking that computer based stock market analysis, based on historical share prices and volumes must be much faster and far more accurate than listening to a Stock Broker news report or Company report.

In practice to date this has been proven to be true and most technical analysis programs can scorch through 1000 plus stocks in less than a minute and kick out those that are not worth looking at. In practice the acceptance rate is about 1%, so that can take a lot of noise out of the analysis process!

Crunching Technical Data

I make no argument about the fact that I strongly favour the use of computers as tools for analysis. The beauty of these tools is that they are fast and unbiased, but you have to tell them exactly what you need or the result will be garbage.

There is an old computer adage that is “Garbage In – Garbage Out!” The first rule with computing is to ‘discover’ what it is that you want as the output, then search for the suitable input that has the necessary data embedded in it and then work out how to get the data extracted in the form that you want to see it.

By reducing noise, this inherently reduces uncertainty, and with the uncertainty reduced, the resolution of price, and changing prices, can be reduced. This means that better trading decisions should be made. ‘Should’ is the operative word!

The problem is if the price fluctuates then some form of approximation is necessary to group all the trade prices and trade volumes into a structure that will assist further analysis.

These fluctuations are caused by many random activities – usually humans and their emotions. When viewing the total trades in a set period, the fluctuations are simply too difficult to conceptualise and many have come up with shorthand methods to group and analyse these trades.

We can translate this engineering almost directly into Stock Exchange-ese! The term 'volatility' related to the change in prices, compared to a reference price and there are a few ways to look at this. If we assume that a price is constant, then there is no noise associated with that price. Pre-computing tools are a good start.

Noise Measurement Tools

The ways that price fluctuations can be analysed are almost limitless, but there are several similarities with electronics that cannot be ignored, and the mathematics used in electronics is very well developed and very powerful. It seems a waste to develop a new strain of mathematics for something that with a little lateral thinking, can be grafted from electronics (physics) into measuring and analysing share price fluctuations. In the people placement business this is called "transferred knowledge"!

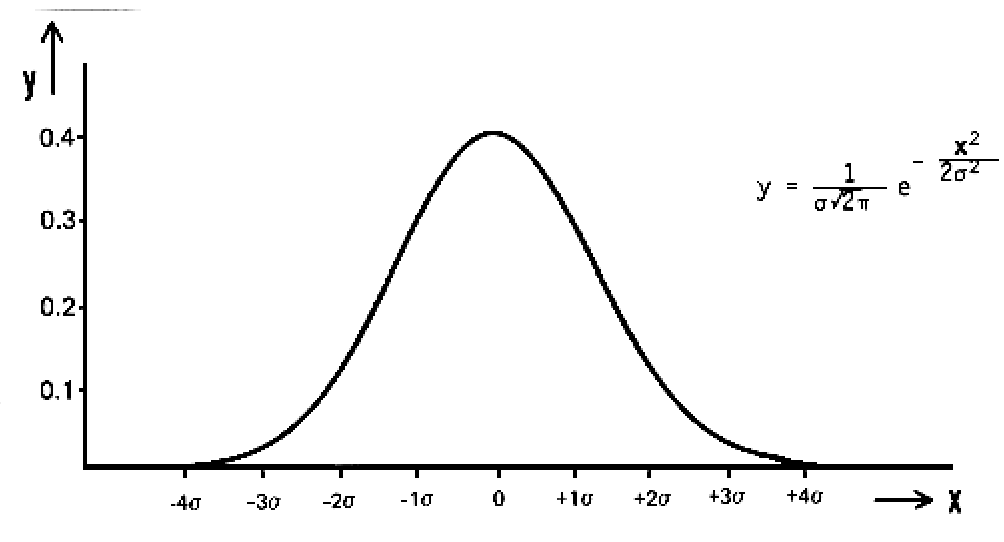
From communications theory; noise is a random quantity, and that means that the incoming values cannot be predicted, and that is the dilemma with security price and volume analysis, just as it is with communications.

If you know what signals were coming in, then you wouldn't need a receiver. Exactly the same applies in the security market (unless you have inside information), so the market is entirely reactive.

Gaussian Probability

In the world of statistics and probability, there is in practice variation from the normal (expected) and this can be recorded as a 'distribution from the normal'.

This theorem states that under very general (natural) conditions, the arithmetic mean (average) of a large number of independent random variables is approximately normally distributed.



This Standard Distribution theorem was first discovered by Abraham De Moivre in 1733, then later by Gauss in 1809 and Laplace (pronounced La Plarz) in 1812 – so it is really important! The Gaussian (probability) curve looks like the graph above:

The area under the curve and above the y = 0 line totals 1.00 or 100% (same thing) and the further that you move away from x= 0, the value of y falls off at a squared exponential rate.

In normalised terms about 68% of the area under the curve is between plus and minus one deviation (or 'sigma'). If you have moved up to four deviations/sigmas from x = 0 then the value of y is almost zero.

It is called a 'normal' curve because the equation for the curve can be normalised (or scaled) to match most naturally occurring events – including share prices!

A classic example is the thickness of a pencil line. Assume that it is accurately measured every (say) millimetre for 100 metres, the widths are then averaged to give the normal, and then the each width measurement is subtracted from that normal width. The resulting differences (or errors) can be graphed and they will form a standard or Gaussian curve.

Another classic example is firing a projectile at a target. Even if the rifle and target are both set in concrete, slight wind movement, air temperature, density, projectile deformation and charge size variations and many other physical changes all cause to change the locations of the projectiles when they hit the target.

Mathematically the distances from the centre of the grouping to each of the targets' projectile holes are related as a 'standard distribution' and mathematically the density (or count of distances from the grouping centre) will directly follow the normal or Gaussian curve!

Noise is Gaussian by nature. That is, compared to a (long-term) average reference, the count of instantaneous values of noise aligns with the Gaussian probability curve. This means that the further the value is from that average reference, the much less likely the value is to exist.

In almost no time we are talking in terms of Standard Deviations (SD) from the normal, and Oh yes, Standard Deviations are also based on exponential relationships.

In electronics and communications the universal language is based on logarithmic relationships – but we are again then talking a very similar language, as exponentials (e^x) are genetic siblings of logarithms! There is a real association here – and it relates to variation in share prices too!

Using Probability

What we do know is that the probability of the next trade price for any security will fit inside the probability curve and be displaced by some deviation from the mean value – and that is when it all becomes much too hard for the (Fundamental) analysts and they resort to waffle – like the price 'softening' and other limp statements.

The problem is the "mean value", as because the security price is varying, the mean value has to be calculated over a large range of recent trades, but if there is a trend in the recent trade prices, then the calculation process of the 'mean value' comes under scrutiny.

The mean value is actually a "moving average" of a large number of recent trades, and the value actually moves, but the bigger the number of recent trades, the less the movement of the "mean value"!

What this all boils down to is that if the trading prices are greater than the moving average, then the deviation from the normal is positive and this shows that there is a probability that the prices will continue to be higher than the moving average mean value – in other words, the trading prices should continue to rise. The converse also applies!

Also be aware that if the trading prices are some standard deviations above the moving average, then this is abnormal and the price will probably soon fall to be in a range where it normally stands.

The same cannot always be said for falling prices, as these tend to stay there and rarely recover nearly as quickly.

Noise Measurements

Noise is usually measured as a power, or as a comparison to a reference power, and initially there seems to be no correlation with security prices, or volumes except that the ranges of prices and volumes are wide, and that is more than enough to set the lateral wheels of thought spinning.

Before we measure noise, we have to understand a little about calculating power levels and the maths behind this.

It seems convoluted, but in understanding some forma of power calculations and measuring techniques, this opens the doors to some innovative ways to (visually) analyse security market history.

Power Calculations

In engineering terms, power is directly related to the heating property of energy over time, and the unit of power is the watt (W) in honour of James Watt and his exceptional work in the field of heat physics and the invention of the steam engine in 1769-1800. Electrically the watt (W) is usually evaluated in terms of its equivalent heating energy over time.

In terms of a stable current flow (direct current as per that from a battery operating a torch), the equation for power is fortunately simple and it is the product of the voltage (V) times the current (I) or $P = VI$ and in this case the impedance to the current flow is resistive.

For example, if we have a two dry cells each with a 1.5 V in series making up a potential of 3.0 V and the lamp is a 6 W variety rated at 3.0 V, then the current flow (by making I the subject of the equation so $I = P/V$) is 6/3 = 2 A (amperes). That was almost painless!

With special thanks to Nicola Tesla (*the genius mathematician and engineer who developed alternating current power generation, invented and developed the three phase power grid system, power transformers, radio transmission, fluorescent tubes, commercial filament light globes, linear motors, squirrel cage motors and thousands* *of other patents leading up to anti gravity*), our power system is based on alternating current.

That is: the current flows forwards and backwards in a cyclic fashion 50 or 60 times per second (depending on the country standard). The base reason for this is so that transformers can convert the AC power from a low voltage high current to a high voltage low current and in that form the losses in transmission over longer distances can be very greatly reduced.

This is why we have 330 kV three phase power transmission systems, and the like stretching for several 100 km! Most share prices also oscillate and so the maths from this branch of electrical engineering has application in technical analysis too.

The problem with having an alternating current (and voltage) is that if the load impedance is not resistive, then the current can lead or lag the voltage by a 'phase angle' and this leads to big power transfer inefficiencies.

To get started it is necessary to calculate the power in a resistive load (like an oven or radiator element) and work from there. As the voltage (and current) traces through a sinusoidal path every cycle, the instantaneous voltage or current can be calculated for a large number of equi-spaced time instances within a half cycle and averaged and on a sine wave that average will be about 0.637 of the peak value.

More elegantly, using integral calculus the area under the half-cycle is $2/\pi$ = 0.6366197. This is a good start, but it does not relate directly to the equivalent power!

If we put power in terms of the load resistance R and in terms of either current (I) or voltage (V), then Power (P) = $V^2/R = I^2R$ watts, then if on a sine wave we calculate the instantaneous power over a large number of equi-spaced time instances within a half cycle and averaged these then that average will be about 0.707 of the peak power value.

But look at the equations that are developed! (We will be leaning on these equations very shortly!)

Average Power for 10 points in a half cycle comes out as:

$$AveragePower = (i_1^2R + i_2^2R + i_3^2R + i_4^2R + i_5^2R + i_6^2R + i_7^2R + i_8^2R + i_9^2R + i10_{10}^2R)/10$$

In a more general form is

$$AveragePower = (i_1^2R + i_2^2R + i_3^2R + i_4^2R + i_5^2R + i_6^2R + i_7^2R.... + i_n^2R)/n$$ and

Average Power = I^2R watts so $I = \sqrt{(i_1^2 + i_2^2 + i_3^2 + ... i_n^2)/n}$

So this is the root mean square (rms) of the sinusoidal current over the half cycle, so when this is calculated out using a large number of instantaneous fractional terms as before the answer comes out as about $I_{rms} = 0.707\ I_{peak}$ for a sine wave and if this is looked at another way then $0.707 = 1/\sqrt{2}$.

Again by transforming the sinusoidal equations from the time domain into the frequency domain, then the answer falls out with brute simplicity producing $I^2R = ½ \, I_p^2R$ and cancelling out the R so the equivalent rms current is $I_{rms} = 0.707 \, I_{peak}$ as before! (The proof is not explained here.)

Comparing Power

In power generation terms power is commonly measured in kilo-watts (kW), mega-watts (MW). and giga-watts (GW). In electronic communication terms, power is usually measured in micro-watts (uW) or milli-watts (mW). Note that all these terms are in ratios of 1000:1 to another term, and comparing power on a linear scale is almost impossible!

(Thinking laterally, share prices are commonly related by percentage, and over time a share price may more than double, or halve in value making linear price scales visually misleading but we still persist with linear scales! Enter the Logarithm and semi-logarithmic scaling.

The term *Logarithm* dates back to the 1600's where exponential based equations were written in an inverse form, and the base of the logarithm could be set. A case in example is the logarithm (base 10) of 1000 is 3, or $log_{10}(1000) = 3$. The inverse of this is $10^3 = 1000$.

In 1620 Gunter came up with a ruler that had a logarithmic scale and logarithmic scales of trigonometric (triangle relationship) functions and from that the Calculator Ruler or Slide Rule was created.

This would be really boring except that you are witnessing the embryonic computer, and that logs (short for logarithms) throw open the doors to working with a wide range of values on a visually linear scale – and very few of us realise the importance of this when looking at the stock market prices beyond a few months and as the price changes.

Comparing powers is difficult at the best of times and it is much easier to compare squared voltages (that relate to power) to give a meaningful relationship.

In the case of power it is usually translated into ***rms*** (root mean square) or equivalent heating power of both signal levels. The ratio of power levels is placed as a log ratio and is called the **bel** (after Alexander Graham Bell for his acoustic and telephony work) but this logarithmic ratio is too big for normal use so we commonly use one tenth of the ratio unit and call it the decibel (dB).

decibel = 10*(log(power 1)/log(power 2)) dB

A loose but direct correlation can be obtained by applying voltages squared as follows:

decibel = 10*(log(voltage1^2)/log(voltage2^2)) dB

Then by moving the squared term to a multiply in logs:

decibel = 20*(log(voltage1)/log(voltage2)) dB

Signal to Noise Ratio

In communication terms, Signal to Noise Ratio (SNR) is the maximum signal level compared to the background signal level, both measured in the same terms, for example voltage (squared) or power.

An example of a radio system is to have an SNR of say 65 dB, meaning that the specified maximum signal level is 65 dB above the system noise level. This logarithmic ratio technique does not have to be limited just to radio communications.

When we are listening to people talk, we need to have at least a 24 dB SNR to understand what people are saying, and if there are other people in a room also talking, then their talk and other noise raises the background noise level and if you are listening, then you have to move or get the person talking to you to raise their voice level so that their voice reaches you at least 24 dB above your near local background noise.

We also need about 12 dB for dynamic range while talking so that means the peak loudness of the voice relative to the background noise will be typically 36 dB greater! This is often termed 'the hotel effect', and that is why public places are usually noisy, and I can't wait for the New York Stock Exchange (NYSE) to go electronic!

Referenced Power Levels

To alleviate some confusion, many people use the term dB as a measurement of power, when dB is a ratio of power. What really happens is that the sub term (Power 2) is set at a reference level, and then the term dB needs a suffix to show that reference, and then it is effectively a referenced power level.

In telecommunications one reference power level is 1 milliwatt (1 mW), and if a power level coming in is 1 mW then the reading is 0 dBm, for 2 mW it is +3 dBm, for 100 mW it is +23 dBm, for 17 uW it is –17.6 dBm, for 1 uW it is –30 dBm. This shows that the power levels (in watts) can vary enormously but when the power is referenced and compared logarithmically; the dBm power levels are very manageable.

In acoustics, the reference power level is a pressure of 1 dyne/cm^2 (10^{-1} newton/metre^2), which is like leaves rustling at 30 metres so it is very faint! A typical quiet room in a residential house is about 45 dBA.

Unfortunately, very few people realise that the "A" in dBA actually refers to sound pressure level at the human threshold of hearing and consequently the "A" stands for the "A" weighting, and not Acoustic.

This weighting is very predominant around the 600 Hz to 2500 Hz range and very significantly falls away outside this frequency band. The "C" weighting (dBC) is virtually flat in frequency response (20 Hz through to 20 kHz), and the "B" weighting (dBB) is effectively half the "A" weighting!

Our ears act like a Dolby noise reduction unit and 'lose' the extreme bass an treble when the sound levels are faint and as the sound levels are increased, our ears tend towards a flat frequency response, and that is why everybody tweaks the bass and treble controls – depending on the acoustic level in the room!

To cap it off, most acoustic measurements are (incorrectly) done using only the “A” weighting filter – irrespective of actual acoustic level! Speech is about 80 dBA (while the meter should be set to ‘dBB’ and a very loud band music is at about 110 dBA, (and the meter should be set to ‘dBC’) – but in general it works and gives a pessimistic result, because it is so empirical!

But you can see that using a logarithmic scale sound can cover a very wide range of acoustic power levels and these can all be on a common plot or graph.

Digital (Quantisation) Noise

In the digital world, noise caused by converting a time-varying analogue level into a stream of equivalent digitised coded data is an interesting phenomenon, as this process has a very close analogy in stock exchange data.

When an audio or video signal is ‘digitised’ the sampling process has a fast clock that holds the analogue sample, in a ‘sample and hold’ circuit then the electronics approximates a nearest equivalent digital value and the error between the sampled analogue value and the digital equivalent value is a quantisation error.

Over several thousand timed samples, these cumulative errors can be measured as noise, referenced to the maximum incoming reference signal. The noise generated in digitising is therefore called ‘quantisation noise’ and there are no prizes there!

In this type of digital conversion, each additional digital bit added to the digital converter statistically halves power of the resulting digital quantisation error, so in noise terms the quantisation noise drops by 6.02 dB for each extra digital bit.

In other words, the SNR increases by about 6 dB for each digital bit added. For an 8-bit decoder the quantisation noise would be about -42 dB, for a 12-bit decoder; -66 dB, and for a 16-bit decoder; -90 dB.

Conversely, the SNR from the maximum encoded level would be 42 dB, 66 dB and 90 dB respectively. This quantisation noise is virtually constant compared to the maximum level and has a Gaussian distribution (or spread).

These topics seem to be so far from stock market prices but look a little closer and we will see that particular stock prices fluctuate about an average price and this is in quantum steps – so it is in effect digital noise, and the maths has already been done. All we have to do is to restructure the applications and use them to our advantage!

Copyright © Malcolm Moore, 2003, 2010

Comments and Corrections are welcome