Perspective

A newsletter for widening your point of view

2008

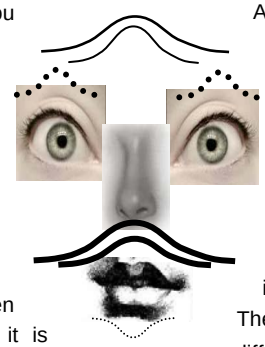
Issue 10

Richard Bach, in his book *Illusions*, states a handy aphorism: **Perspective – use it or lose it**. This periodical shares amongst recreation and tourism management professionals, and others, several tools and concepts which will help exercise your perspective. This issue considers measuring the past and predicting the future. It is also the tenth newsletter, meaning a decade of RG&A. Wahoo.

Perspective is distributed by **Rob Greenaway & Associates** as a service to the recreation and tourism industries.

Uncertain certainty

If you're not into statistics, read only this sentence: If you encounter a report which states that some result is 'statistically significant', ask the author, who cares? Or, as Nassim Taleb would say, just ignore them, or try to put a rat down their shirt.[1] Still curious? Read on.



The problem with statistics is that it's easy to be intimidated. Most of us would agree that x^2=26.8, df=1, p<.001 is pretty daunting, particularly when someone who should know better is trying to convince us that this statistical advice proves that an apparent difference between two data points is important: we should take note – it is 'statistically significant'.

My response, in the social sciences, has always been that if the difference between the two raw data points doesn't make your mouth gape, or at least raise your eyebrows, then it's up to the researcher to prove that the difference matters – not just that it might exist. I say this even if the researcher's job is merely to report 'the facts' as revealed by the data (it should never be). To this end, I have a pet paragraph that I append to peer reviews of social surveys where tests of significance have been applied:

> *Differences between data which are described as 'statistically significant' do not indicate that they are meaningful to management. Managers using the report will need to confirm the scale of effect, or difference of opinion or experience, that will trigger a management response.*

Therefore, it was with glee that I read econometricians Stephen Ziliak and Deirdre McCloskey's text, *The Cult of Statistical Significance – how the standard error costs us jobs, justice and lives.*[2] The authors revile any faith in the puffed-up spawn of the normal distribution curve: statistical significance.

Their key anxiety relates to what they call the 'sizeless stare', which is nurtured by significance tests. The fundamental issue is not that a result is likely to be the same if the test is replicated (is truly representative of the target population), but whether the research technique is sound and whether the result is important considering the issue at hand. When a researcher says an output is statistically significant (or not), Ziliak and McCloskey yell, WHO CARES? and HOW MUCH? That is, what is the size of the difference, and, is the result meaningful to the discipline? Is the result of a drug test 'clinically meaningful', rather than just 'statistically significant'? Is an economic parameter important to the economist or to the statistician?

A quoted example is Vioxx. Taking this anti-inflammatory drug had the unfortunate side-effect of increasing the chance of a heart attack, and it was withdrawn. Original research into Vioxx by the manufacturer showed that, in the test sample, only five patients taking the drug had heart attacks. In comparison, the control drug (another NSAID) was associated with only one. However, both results were statistically insignificant, no matter that Vioxx apparently caused five times the acute myocardial infarction as the control (it turned out to be eight times). The original Vioxx researchers comfortably stated that there was no difference between the two medications in relation to the risk of heart attack. The data were irrelevant – their lack of statistical significance was purported to be the only necessary measure.

Even my tear-stained copy of *Statistics Without Tears* states, "… in statistical thinking, 'significant' does not necessarily imply 'interesting' or 'important'…. For instance, suppose two related experiments each produce a difference, one that would have a 4.9% chance of arising merely because of sampling variation, and the other a 5.1% chance…. In one case it would be labelled 'significant' and in the other 'not significant'…. Perhaps, in evaluating such data, it is best to let the figures speak for themselves."[3]

However, Rowntree's only alternative proffered solution – to letting the figures speak for themselves – is to apply more stringent levels of significance to avoid Type 1 error (accepting a difference as significant when it is not. Type 2 error is the opposite, failing to recognise a real difference).

Ziliak and McCloskey go much further. They liken the familiar significance tests to bloodletting. There are more meaningful alternatives. Importantly, there is the role of the researcher in, "…thinking about your coefficients in a currency of How Much in the world as it is, or could be, and persuad[ing] a community of scientists [that the difference is meaningful]. Instead of deploying a mechanical rule about one kind of sampling error you will have to establish a minimum effect size of substantive significance in the relevant range of power, for your particular area of research, acknowledging all the sources of error."

Significance tests are misleading, nasty things. Report and rank data that are meaningful (not 'significant'), show confidence intervals and review errors. Ask, who is my audience? Why do they need these data? Am I a bland statistician or a researcher? Mostly – mostly – the *t* and *x* and *p* are a waste of ink. Worse if you square them. ❖

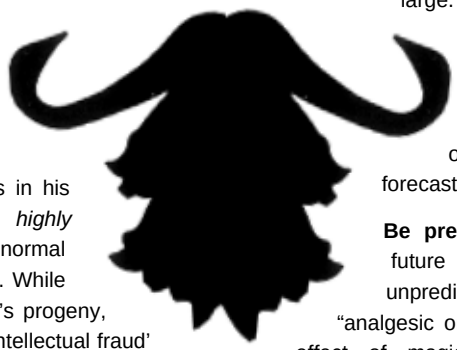
[1] See the next page.

[2] Ziliak, S.T. McCloskey, D.N. 2008. *The Cult of Statistical Significance.* University of Michigan.

[3] Rowntree, D. 1981. *Statistics Without Tears*. Pelican. p118

Certain uncertainty

Several years ago in this newsletter I wrote about 'regression to the mean' – the propensity for trends to adhere to the normal distribution curve. For large and small peas to tend to spawn average-sized peas – not ever-larger or ever-smaller peas: for biological and social phenomenon to tend towards the mean. This is apparently the case in a country that Nassim Nicholas Taleb calls Mediocristan, but not in the real world – which is increasingly from Extremistan.



Taleb describes these two metaphorical nations in his book *The Black Swan - the impact of the highly improbable.*[4] Taleb really doesn't like the normal distribution curve and the complacency it breeds. While Ziliak and McCloskey are gunning for the curve's progeny, Taleb goes for the foundation, calling it a 'great intellectual fraud' and a 'contagious and severe delusion'. The curve, amongst other things, severely under-estimates probabilities at the extremes.

His premise is that we tend to look at the future as an extension of the past, presuming such phenomena as regression to the mean to be in play. We look at the past and derive (induce) a set of rules to apply to the future, although this could lead to some odd and competing conclusions: if you survive until tomorrow, it could mean that, either, you are more likely to be immortal, or that you are closer to death.

We tend to manage our lives within the rules of mediocrity and faith in the normal distribution curve. Well, it works for dice, and, according to Taleb, it helps if you want to get promoted in academic circles (don't come up with anything too weird or challenging; have faith in quantitative modelling; apply the old rules – like statistical significance tests).

In reality, the future will be the product of wildly unpredictable and extreme events, which we will have no way of foreseeing. The world moves by large incremental, random changes. However, once these 'random' events have passed – WWI, WWII, the Internet, another share market crash, the discovery of Viagra, texting, no bird flu – we will apply a reverse logic to convince ourselves that the unexpected was not such a surprise. Had we access to all the data, and had the data behaved in a 'normal' way, we would have seen it coming. The fact that all our earlier predictions were incorrect is irrelevant – and there are more from where they came.

Interestingly, Taleb does not describe future conditions to be the result of true mathematically random events, but rather, an outcome of unknowable complexity. As with the weather, the future is written in the past, but is compounded by our free will. It is too multifarious to compute and might as well be mathematically random. For that reason, predicting the future based on historic precedents or similarities is worthless. Taleb therefore contends that our reference to specialised forecasting of social and economic trends is bunkum.

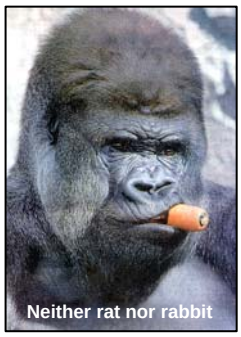
So what to do? Taleb has two simple recommendations:

Be a fool in the right places. Avoid unnecessary dependence on large-scale harmful predictions: be fooled in small matters, not in the large. Rank beliefs by the harm they may cause, not according to their plausibility. A problem here is that it takes more effort to be sceptical than it takes to believe. Humans are not natural sceptics – we have to apply ourselves. Beware of economic and social forecasters. Taleb describes them as mere entertainers.

Be prepared. Knowing that you cannot predict the future does not mean that you cannot benefit from unpredictability. Narrow-minded predictions have an, "analgesic or therapeutic effect. Be aware of the numbing effect of magic numbers. Be prepared for all relevant eventualities."

Taleb was a financial trader in a former life and worked within the world of prediction errors, applying his brand of empirical scepticism to the world's markets. His investment strategy is similar to his recommendations for life. If you accept that most risk measures are flawed, don't rely on 'medium risk' investments – how do you know they *are* medium risk? Instead put 85 to 90% of your investments in the most conservative, safe instruments. The remaining 10 to 15% goes into highly speculative areas. Your base investment is safe and your risk element is exposed to the positive effects of Black Swans – those highly improbable events. ❖

Carrot polarity


Neither rat nor rabbit

Remarkably, the Zambian common mole rat can identify the sharp end of a carrot even if you chop off both its ends (the carrot, not the rat). In a survey of 20 small mammals, the subterranean species preferred eating their carrots from the tip, without exception. Is this because they would normally encounter the sharp end when burrowing beneath the soil? This is the hypothesis of respected evolutionary palaeobioligist Simon Morris.[5] However, I have conducted exhausting interviews with a selected sub-population of rabbit owners and have a statistically significant result showing that these extra-subterranean species don't care where they start nibbling, even though they would normally first encounter the blunt end. This puts the 'first encounter' theory of carrot polarity to bed. Do you eat your toffee pops by grinding off the top layers with your incisors? (Our kids are mole rats!) I recommend Constant Vigilance when reading even the most convincing of scientific texts. ❖

For Your Interest

Beyond the realm of raising perfect children, the most exciting thing to happen this year was reading Ziliak and McCloskey. I got quite giddy and ordered a copy for the Lincoln University library. But it could have been the airport coffee. Otherwise, it's been the standard string of unusually interesting and challenging projects – numerous energy proposals for six different companies (wind and hydro), several large-scale housing and resort developments, a few irrigation proposals, some recreation user surveys (without a *p* or *t* in sight), a little territorial authority recreation planning, and 160 resource consents for helicopter landings. Eighty percent of our projects (with a CI of about 15%) are in the resource management area; preparing assessments of effect for recreation and tourism values with regard to development proposals, and normally presenting evidence in hearings and the Environment Court, as well as preparing the occasional simple resource consent application. This represents quite a shift from my originally dominant area of recreation research and planning for local government – although it remains nice to eat the carrot from both ends, so to speak.

Life in Nelson is good, despite this dismal, destructive winter (we lost a large twisted willow – ugly, but strategically located). Importantly, we have solved the local 'lack of boat' problem, and have replaced it with the 'maintaining a boat' problem.

[4] Taleb, N.N. 2007. *The Black Swan*. Random House. I wonder why he chose that publisher?

[5] Morris, S.C. 2003. *Life's Solution. Inevitable humans in a lonely universe.* Cambridge