LS 829 - Mathematics in Science and Civilization

Feb 6, 2010

The Surprising Consequences of Randomness

Sources and resources:

Nassim, Nicholas Taleb *Fooled by Randomness,* Second Edition, Random House, New York.
Peck, R. et al (2006) *Statistics: A Guide to the Unknown*. Fourth Edition. Thomson Nelson. Toronto.
Weldon,K.L. www.stat.sfu.ca/~weldon

Introduction: The discipline of statistics has origins in gambling, astronomy, and experimental science. Mathematics helped to clarify the properties of statistical methods and produced many useful developments: correct odds were worked out for gamblers, measurement errors were made allowance for in astronomy, and Fisher put experimental design on a firm basis. Underlying all these origins, the concept of "randomness" enabled researchers to describe "unexplained variation" unambiguously. Of course, this description was not real, only realistic. We had an abstract way of describing that part of measurement variability that defied attribution to specific causes.

Mathematics provided the tools to cope with "random" variation. Parametric models (formulas for relative frequencies calibrated by a few constants called "parameters") became the primary tool for expanding this idea of randomness to complicated applications. However, the advent of widespread computing power that has evolved over the last 50 years has changed how people study randomness: computer software in combination with these parametric models has provided methods of analysis that are accessible to a wide variety of professionals, not only mathematicians. In fact, recent trends in the development of graphical methods have moved some statisticians away from the parametric models, although the concept of randomness is still central to the proper analysis of data.

The discipline focused on coping with unexplained variation and randomness models is called "Statistics". In this talk, I will provide ten examples of how the discipline of statistics can help to reveal some surprising randomness phenomena, and for each of these examples, I will point out the general concept illustrated by the example. Many of these examples use a technique called "simulation" or sometimes "Monte Carlo simulation", and this technique requires the parametric models invented by mathematicians more than half a century ago. We will show that the mathematics is essential for research in statistics, but that it is possible to appreciate the implications of statistical thinking without mathematical detail: the ten examples provide information that is useful in everyday life.

1 – When is success "good luck"?

The book "Fooled by Randomness" explores the wide variety of situations where "success" seems to be attributable to superior intelligence or vision, but where closer analysis reveals "good luck" to be the true cause. The sports scenario used here is just an easy one to comprehend: the inclination is to attribute a winning streak to team quality, but the analysis shows that the impact of luck is under-rated.

Teams in a league are updated yearly to try to balance the quality of the teams. The reason for this is economic: the excitement that fans expect will depend on the game being fairly close in score. The biggest turnouts are for games for which the outcome is the most in doubt. In this situation, a reasonable model for the league play is to assume that each game is a 50-50 game. The surprising thing that a simulation shows is that this assumption still provides for a wide range of league points from the top to the bottom of the simulated league. Comparing the spread of league points in the 50-50 simulation of a season, with an actual league outcome, confirms that much of the apparent difference in the success of the various teams is attributable to "luck".

2 - Order from Apparent Chaos

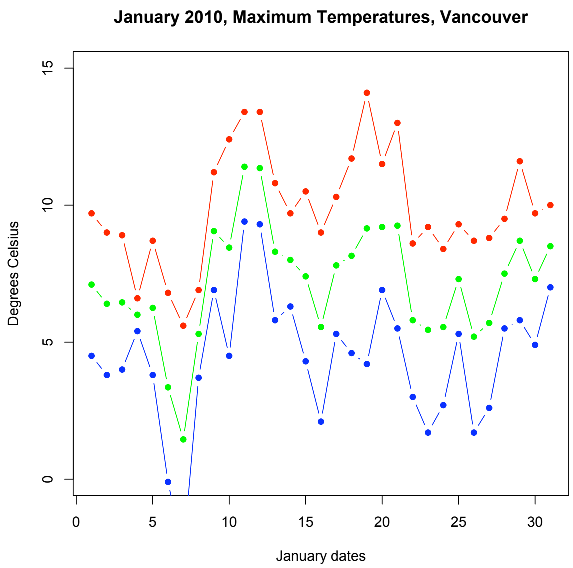
In the previous section we showed that apparent order can actually be an illusion of chaos. In this section we show that apparent chaos can actually reveal a surprising orderliness. I collected my gasoline consumption for every fill-up over a five year period. (This odd behavior may be explained by my odd profession: statistics!) The graph shows no apparent pattern. A statistics student might try to fit a line or a quadratic curve or perhaps a higher order polynomial to check for patterns in this kind of data. This approach just produces a horizontal line, and a conclusion of "no interesting pattern". However, with modern software which makes use of graphical methods to analyze data, it is not hard to find a pattern. The question then is, what causes this pattern? It is surely interesting in the context of the data to know what factors are associated with higher or lower gasoline consumption. A technical point of the example is that graphical methods are very powerful, and we should not rely on fitting simple parametric curves when the motivation for them is absent. A larger point is that statistical methods exist that are both simple to explain and have practical uses.

We also use this example to talk a bit more about "smoothing". There is always a trade-off between the amount a smoothing and the amount of detail one wants to extract. We argue that the choice is subjective, and depends on the context of the analysis, and that subjectivity in statistical analysis is a fact-of-life. Statistics can lie but so can words! Always check your sources for bias.

3 – Weather Forecasting

Vancouver weather is in the news. While the mildness of our winters is well-known, it is also a fact that the local hills are good for snow sports in the winter. But look at

January's max/min/mean temperatures. Even allowing for the cooling dues to elevation of the local hills, it will be tough to have enough snow for the Olympics.



Paradoxically, the weather is often surprising. Even apparent trends get broken abruptly and puzzle the forecasters.

In the early 1900s, meteorologists felt they could model the basic weather influences: air and water currents, effects of sun and season, and so on, but the global picture required a large number of equations to describe the interactions between such influences. Meteorologists felt that if only they could solve the equations, they could predict the weather much more accurately. However, when giant supercomputers finally were able to solve the equations, the predictions were not much better. The problem was not the equations themselves but in the sensitivity of their solutions to starting conditions.

Edward Norton Lorenz (May 23, 1917 - April 16, 2008) was the mathematician who introduced "The butterfly effect" in 1963, which explained the difficulty of forecasting the weather. Since the weather at any moment is virtually impossible to describe exactly, the wonderful meterological equations cannot be accurately initialized, and the predictions become unreliable.

Instead meteorologists today use a probability distribution of initial conditions and generate a distribution of possible outcomes. Using this aggregate of deterministic outcomes as a probability distribution is called "dynamic modeling". Not really statistical because it uses deterministic models. Purely statistical forecast (no physics) : OK for < 6 hours or >10 days (See Miller reference in the Wilks article.)

Why did people believe in the predictability of the weather for so long, in spite of the poor success of forecasters? As we will see in the next unit, purely random processes can appear to have a strong trend component when this is an illusion.

4 - Obtaining Confidential Information

If you want to know the proportion of a class that are regular marijuana users, they will likely not answer, or at least not answer truthfully, unless they are assured that their response will not be tied to their name. One way to protect respondent ID in responding to a sensitive question, is to use the randomized response technique:
1. Ask the respondent to toss a coin and to privately observe the outcome
2. If the outcome of the coin toss is a head, then the respondent should answer the sensitive question with "Yes" or "No" as appropriate. But if the outcome of the coin toss is a tail, the respondent simply answers "Yes" to the sensitive question.
When the responses are recorded, even if the name is attached, there is no way for the surveyor to know if the particular respondent is answering the sensitive question, or has simply got a tail on the coin and is not answering the sensitive question.
So suppose we have a class of 100 students, and 60 of them answer "Yes". We estimate that about 50 of them will have tossed a "Tail" and are answering Yes for this reason. But then about 10 are answering Yes to the sensitive question, out of the 50 or so who tossed "Head". Ergo, the proportion who say they are regular users is 10/50 or 20%.

Note that, of those that answer Yes, only 10 out of 60 are answering Yes to the sensitive question, so the true response of an individual answering Yes is well camouflaged.

I did this in-class survey with my 2002 STAT 100 class, which did contain about 100 students. The Yes answers were done with a show of hands! There may have been some who did not answer either question. The effect of this would be to bias my estimate downward.

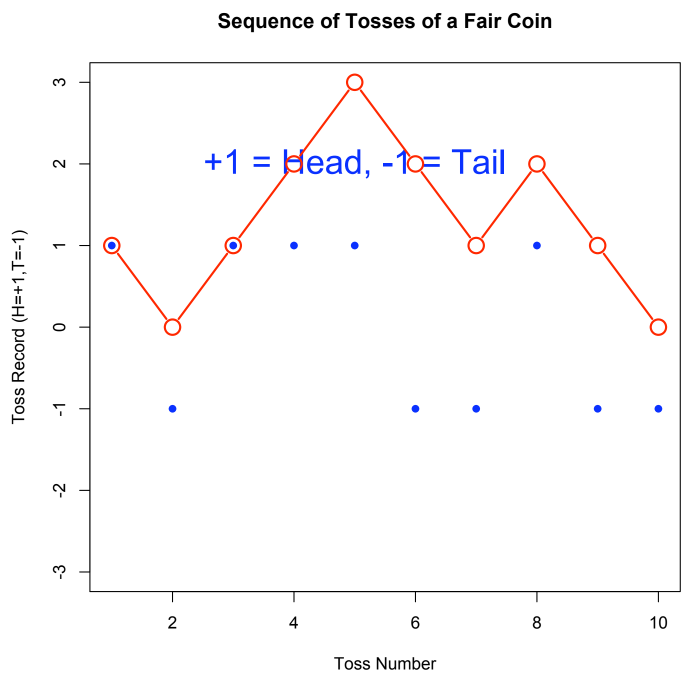
The surprising thing here is that confidential information could be obtained even though responses were tied publicly to respondents. The more general lesson is that probability can have practical uses such as preserving confidentiality while still recording useful statistical information.

5 – Randomness in the Stock Market

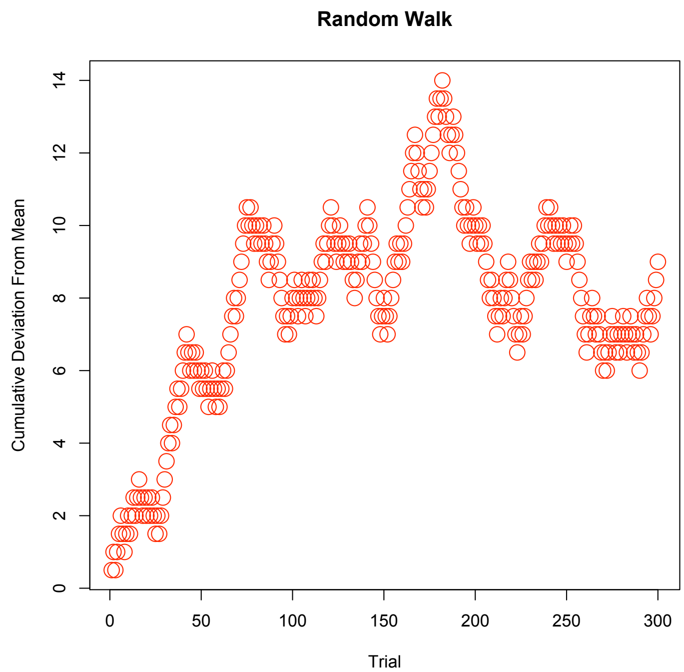
5A – Trends that deceive

Daily changes in the stock market index are explained by financial experts. E.g. “The market fell because investors worried about the possible increase in interest rates to be announced by the Bank of Canada.” Or, “The market rose in response to assurances by world leaders that climate change would not harm the economy”. Such pronouncements are reasonable, but always in retrospect! The success of the experts in predicting market moves into the future, even one day into the future, is very poor. See the evidence mentioned in the Cleary & Sharpe article.

Why do intelligent people think they can predict market moves in the short term in the face of so much evidence that they cannot? One possible answer to this is in the very subtle illusion of predictability caused by an unpredictable time series. Our demonstration of this uses an artificial series called a random walk. It is constructed as follows: At each time period, a fair coin is tossed. The outcome is recorded as a number: +1 if the coin toss results in a Head, -1 if a Tail. At each time, the sum of all previous results since time 1 is recorded. For example, if the first few tosses were H,H,T,T,T,H,T; the cumulative record would be 1,2,1,0,-1,0,-1. In the talk I will use the computer to generate a few of these random walks, but here is a typical one:



The blue dots are the +1 and -1 records. The red line is the cumulation. Now we look at the red cumulation for a much longer series of 300 tosses of a fair coin:



Note that in this case the whole series stayed on one side of 0, while we might have expected the series to jog back and forth across zero. Also note that the first 80 or so tosses suggest an upward trend. But from the way the cumulation is constructed, we know that there is no real trend. The appearance of a trend useful for prediction has been an illusion. In this cumulation series, the best predictor of the long run future is always the last value observed, because this is the average of contributions of +1 or -1.

The article by Cleary and Sharpe include the stock price of Intel as an example. Here is a recent chart of Intel stock price. Compare the nature of this chart with the nature of the random walk chart above. If you conclude that the Intel chart could be a random walk, you would find agreement from the statistical experts. Stock prices often mimic behaviour of random walks. The actual size of steps is not so important, since random walks can have many different kinds of steps – the key feature of a random walk is that changes from one time to the next average 0 and are independent of the previous change.

As Cleary and Sharpe point out, long term forecasting is much more successful than short term forecasting, in the context of the stock market. It is likely that a proper model would be an “asymmetric” random walk with a small upward trend, and this small upward trend is what makes long term investment a good investment.

1d | 2d | 5d | 1m | 3m | 6m | 1y | 3y | 5y | 10y | 20y | Max

Quotes delayed 15 minutes

1d | 2d | 5d | 1m | 3m | 6m | 1y | 3y | 5y | 10y | 20y | Max

5B-The power of diversification

"Don't put all your eggs in one basket". It does not take a statistician to point out the merit of spreading your investments around different investment types, different companies, different industries, even different countries. But quantifying the advantage to a diversified portfolio can reveal some surprising consequences of this widely known advice.

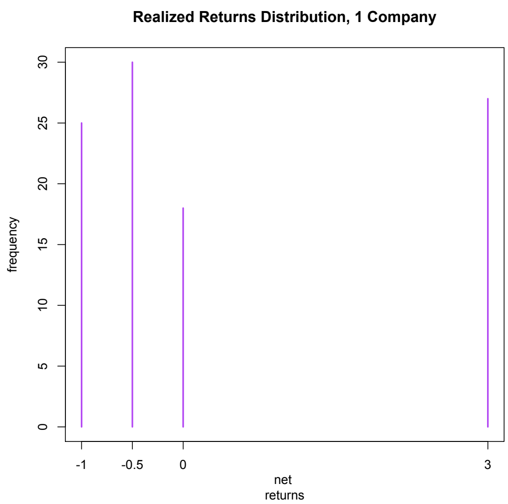
Why do some people invest in high-risk "penny" stocks that frequently become worthless? The answer is that in the unusual circumstance that the company with the penny stock survives its early years, its stock may appreciate to many times its early cost, and result in huge profits for the shareholder. So such stocks are a big gamble. Conservative investors stay away from these kinds of investments. But we will investigate a possibly profitable strategy that could be used with a portfolio of high-risk companies.

Suppose you are offered shares in a company that has the following 1-year prospects:

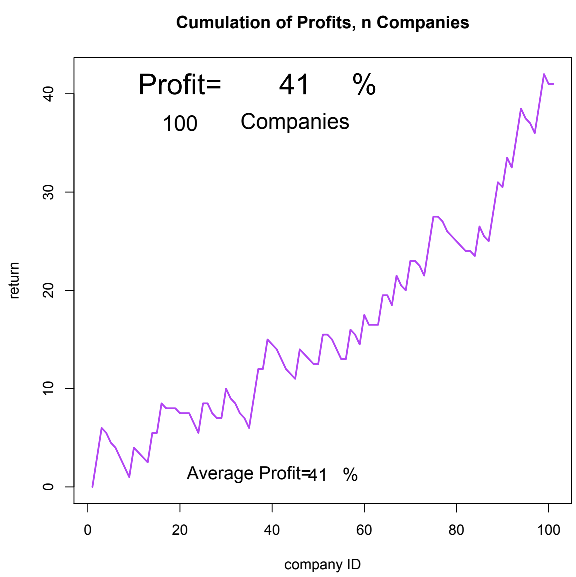
The shares are for sale for $1.00
At the end of one year, the stock will be equally likely to sell for $0, $0.50, $1.00, or $4.00. In other words, there is a 25% chance for your net income per share to be: -$1.00, -$0.50, $0.00, or $3.00. So there is a good chance (50%) that you will lose money, and a very good chance that you will not make any money (75%). There is small chance

for a fairly good payoff of $3.00 (25%). So this would normally be thought of as a risky investment.

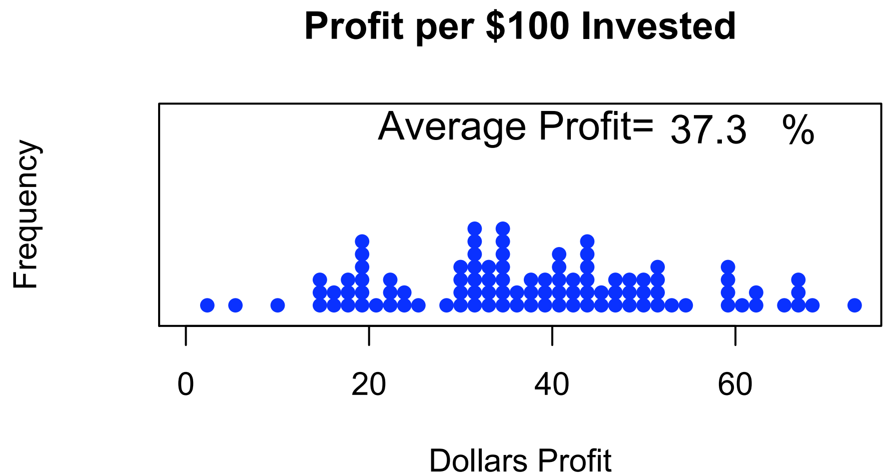
Let's look at a typical simulation of this stock: since the outcome is only a single number (-1,-1/2,0,3) it is more informative to look at 100 experiences to get an idea of what is likely.



The bars are fairly equal as one would expect. But lets look at what would happen to our net profit during a sequence of 100 such investments:



As we scan through the 100 investment outcomes, accumulating the results, we see a fairly systematic accumulation of profit: 41% more than our initial investment. But was this typical. The simulation is set up so we can do the same thing 100 times.



One does not always make 40% profit, but returns of 15%-50% are fairly typical. And all this based on an investment prospect that loses money half the time.

Is there a catch? Are all risky investments potential winners? Certainly not. We need two conditions:
1. The 100 companies (that we have each modeled in the same way) need to have *independent* outcomes. That is, whatever happens in one company is unrelated to what happens in the other companies.
2. The average return from the model company must be positive.

Re 1. The independence is hard to achieve in practice, but in forming a portfolio the best stability of returns is achieved by making the component investments as independent as possible.

Re 2. The average return in our example was 37.5 cents per dollar invested. To see this, note that $(0 + .50 + 1.00 + 4.00)/4 = 5.50/4 = 1.375$ so a $1 investment yields $1.375 on average for a positive return of 37.5%.

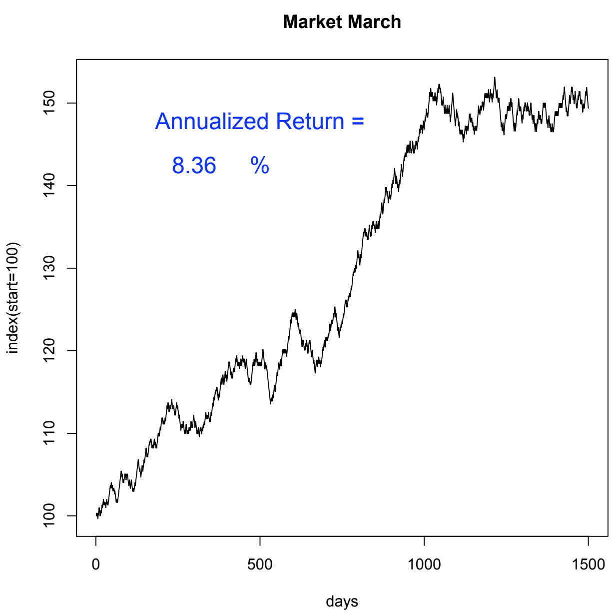
We can also compute the SD of company outcomes:
First compute the squared deviations $(0-1.375)^2+(.50-1.375)^2+(1.00-1.375)^2+(4.00-1.375)^2 = 9.7$. Then divide by n-1 = 3 and take the square root of the result.
$\sqrt{(9.7/3)} = 1.80$

The investment lesson survives these requirements: find companies with prospects that, on average, are positive, and choose companies in the portfolio for which the success is as independent as possible.

5C- Back-the-winner fallacy

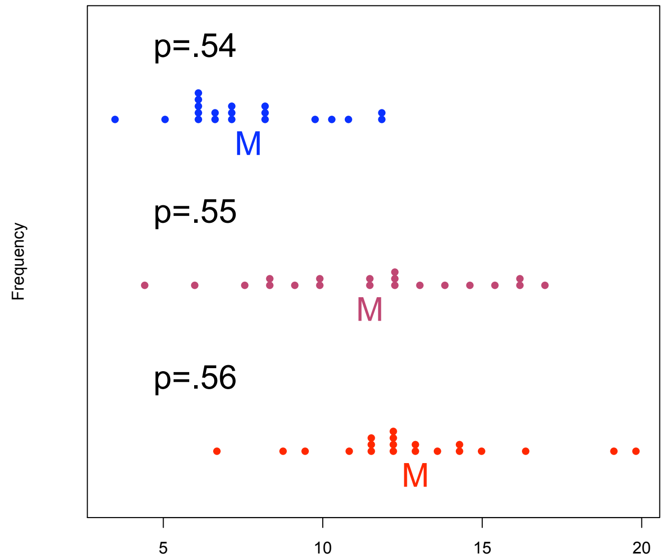
When one looks at the annualized performance of mutual fund managers over several years, the rates of return typically fall in the 5 to 15 % per annum range. We can simulate this outcome by using a daily random walk model with the probability, p, of a day-to-day increase being in the range .54 to .56. "p" is a true measure of the quality of the manager.

Here is an example of a random walk in which the advance probability is p=.55.



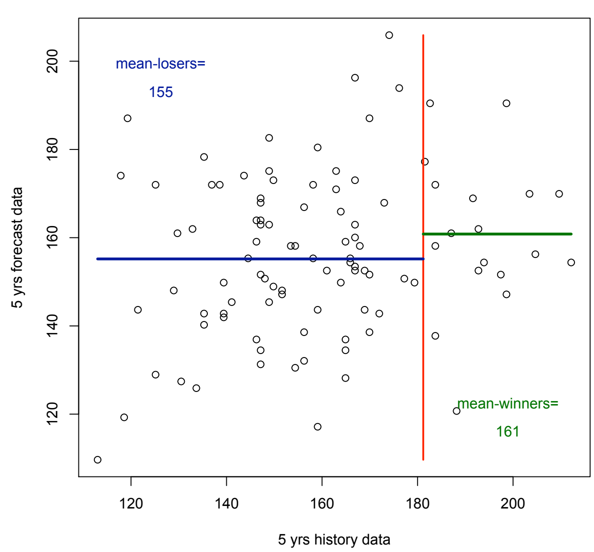
The next graph shows how the manager quality parameter p relates to the 5 year return of the fund.

Five Year Return Outcomes (Annualized)



We simulate the experience of 100 managers by assigning them a random p in this range – then we pick the "best managers" as indicated by their 5 year performance and, using the same p –value for each manager, simulate another 5 year experience. We then compare the performance of these "best managers" over the second five years with the "other managers". It turns out that the difference in performance of the "best" and "other" managers is very small. The reason is that the variability in performance over a 5 year period swamps the quality difference as measured by p. In other words, the managers chosen by performance are not really much better than the other managers. So chosing managers based on past performance is not a very useful strategy.

The manager's quality is swamped by randomness



In this simulation, the return of the winners was 9.9% per annum, while for the losers it was 9.1% per annum. This modest outcome of "playing the winner" is all the more surprising when it assumes that whatever edge a manager has during the first five years, they maintain that quality over the next five years.

As the graph displays "The managers quality is swamped by randomness".

6 – Statistics in the Courtroom

The discipline of statistics has two main claims to fame. One is that it can keep researchers from inferring apparent results that are not reproducible (i.e. are caused by random error), and two, it can often detect effects that are not otherwise apparent. The fuel consumption example we discussed earlier is an example of the latter claim. Another one in this category is the case you have read about in "Statistics in the Court Room". Note that it took 10 years of data before the evidence was strong enough for someone to notice! Without the graphical summary shown in the article, there might have been suspicions but no convincing evidence. The 74 deaths over the 1641 shifts that Kristen Gilbert worked at the hospital might not have seemed abnormal in an acute care ward.

Figure 1 in the article definitely raises suspicions! And the statistical test based on Table 1 does confirm that random variation does not explain the suspicious pattern.

It is interesting that the judge felt the statistical evidence was inappropriate for the jury, reasoning that the jury would misunderstand it. The difference between the test discounting randomness as an explanation and proving Kristen was responsible for the deaths was the issue that the jury might have missed. The other point made in the article is that the probability of the evidence given the suspect's innocence is not the same as the probability of the suspect's innocence given the evidence. This is the "Prosecutors Fallacy". The link between the two probabilities involves the prior probability that the suspect was guilty.

7 – Lotteries: Expectation and Hope

This example just shows two things:

1. Lotteries sell hope rather than realistic chances to win.
2. As an investment, public lotteries are poor vehicles, even for the lifetime participant.

Lotto 649 for example, has a jackpot awarded to one set of numbers out of approx 14,000,000. If a person bought 10 tickets every week for 60 years, their chance of winning this jackpot would be about 1/5 or 1 %, and they would have spent about $31,000 for the opportunity. If they are clever, they will chose numbers that others would not choose so that they would not have to split the prize, but actually the chance of this being a problem is pretty tiny!

When a person buys a lottery ticket, it seems not totally illogical to think about what the person would do if they won the jackpot, and this is a pleasant activity for some people. So it may be that a lottery ticket has value other than its investment value.

The surprising thing about public lotteries is how small the average return is, and also how rare the large wins are. The general lesson here is that a simple understanding of probability allows a rationale assessment of the situation.

Conclusion

I have attempted to show how a little statistical knowledge can reveal information that is not widely known and yet is useful for real life. I have also tried to provide some appreciation of the methods used in deriving this information. If anyone would like clarification at some later date, I am easily contacted by email at weldon@sfu.ca.

Recent Books of General Interest about Probability and Statistics

Mlodinow, L (2008) *The Drunkard's Walk*. Vintage Books. New York.
Levitt, S.B. and Dubner, S.J. (2005*) Freakonomics.* Harper Collins. Toronto.
Levitt, S.B. and Dubner, S.J. (2009*) Super Freakonomics.* Harper Collins. Toronto.
Rosenthal, J.S. (2005) *Struck by Lightning* Harper Perennial. Toronto.
Gardner, D. (2008) *Risk* McClelland and Stewart. Toronto.
Tanur, J.M. et al (eds) (1989) *Statistics: A Guide to the Unknown*. Third Edition. Duxbury. Belmont.
Taleb, N. N. (2008) *Fooled by Randomness: The Hidden Role of Chance in the Markets and Life,* 2nd Edition. Random House.
Gladwell, Malcolm (2008) *Outliers: The story of Success*. Little, Brown & Co.