Risk Analysis – from the Garden of Eden to its seven most Deadly Sins

L Ferkl
Feramat Cybernetics, Czech Republic

A Dix
Scientist & Lawyer, Australia

"Probability theory is nothing but common sense reduced to calculation."
Pierre-Simon Laplace

HOWEVER,

"Common sense ain't common."
English proverb

ABSTRACT

Risk analysis is often used as the basis for making complex decisions about the optimal allocation of limited resources. Many organisations attempt to satisfy their corporate governance and other legal duties in relation to critical tunnel safety issues by relying on a risk management framework to justify their corporate decisions.

Understanding risk assessments foundations in mathematics is essential if the limitations of these techniques are to be respected and their role as a legitimate tool in assisting professional judgements respected.

Risk analysis is based on probability theory and statistics. Unfortunately, many products that claim to be "risk analysis tools" oppose several principal pillars of these branches of mathematics, and even where the tools are sound the people responsible for them often use the tools wrongly.

Furthermore practitioners often are not open and comprehensive about what has been done; especially with respect to the assumptions made, both conceptual and numerical. Like all sins such lack of detail conceals detection of the sins, and identification of the sinner.

In this paper, we will describe seven main problems of risk analysis tools of the state of the art, with mathematical background and possible impacts in practice.

1. INTRODUCTION

Risk analysis is often used to describe or justify a process of analysis of the special facts and circumstances of an aspect of the safety of complex, underground infrastructure.

Risk assessments are regularly used to determine how to prioritise limited resources (such as money and expertise) in the systematic improvement of some aspect of safety performance.

Unfortunately the techniques described as "risk analysis" are often either mathematically flawed or applied so poorly that the results are, at best, meaningless and at worst highly destructive to sound decision making. Frankly, without adherence to the fundamental mathematical principles of "risk analysis" by practitioners almost any outcome is possible – by accident or by design.

The seriousness of this situation is compounded by the fact that there is a temptation to use these techniques as a substitute for professional judgement – and not limit its use as a tool to assist in the decision making process.

Undue reliance on risk assessments to make corporate decisions is a high risk strategy and likely constitutes an abuse of the principles of sound corporate governance. Used responsibly risk assessments provide excellent support for complex decision making in both a practical and corporate governance contexts.

2. THE FIRST SIN – LACK OF DATA

Up a creek without a paddle

Risk analysis is the creek. Data is the paddle. With no data, you cannot have a precise risk analysis.

For risk analysis, a lot of input data are needed. Although lack of data might not be the fault of the analyst – the sin is not carrying out sufficient sensitivity studies. It is the failure to conduct a sensitivity analysis by varying the conceptual and numerical inputs to see the effect on the results which is at the root of this sin. A shortage of data can often tempt the analyst to devise input data which looks "reasonable", but is consciously or unconsciously aimed at recommending the option a client wants. These data can be acquired either by measurement (data collection), or by expert's opinion. The first option is always the better one, but if there is not enough empirical data, the expert's opinion is necessary.

When an expert tells us the partial probabilities, we sometimes call this *belief*, which is a more accurate term. Belief cannot be computed by classic probability theory, the more correct way would be to use a so called Bayesian approach. The Bayesian approach can be combined with the concept of relative frequency. See ref [1]. The Bayesian approach is a way of combining expert judgement with available data.

The term 'Subjective' probability generally implies a different concept and usually applies to unique events, such as the probability of a president being elected again (relative frequency has no meaning there). Although a 'guess/estimate' is 'subjective'; but the kind of initial estimate by an expert in Bayesian statistics is (or certainly may be) different in nature to that generally associated with the concepts of 'subjective probability'. In Bayesian statistics, the basic event may be regarded as related to relative frequency, at a given level of abstraction. See [3].) The Bayesian approach is an acceptable way of coping with sparse data. What do we have otherwise? Only a hunch! Of course, it has to be applied in a justifiable way, as with anything.

This means that, the fewer data there are, the less prediction we can make. Sometimes it is better to use only few data that are reliable to make a statistical prediction that is generally reliable instead of having an "exact" number that is, in fact, speculative.

An example of this in the context of tunnels is the use of data collected over the last few decades on either road or rail tunnel fires. Put simply, the data on events which occurred decades ago relates to vehicles which have little relationship to modern vehicles in terms of heat release rates and other aspects of fire performance.

If there are not enough data, it maybe better that an expert perform a qualitative analysis which gives us an intuitive expert answer ("the tunnel is very safe", "the tunnel needs some improvements"), or to look for weak points of safety.

Demanding "some number" can result in misleading answers. The main reasons are variance in the result and ill conditioning, which we will discuss as The Seventh Sin.

3. THE SECOND SIN – IMPROPER SOFTWARE

If you buy cheaply, you pay dearly

This solution is very cheap, but…

When your only tool is a hammer, every problem looks like a nail

There is no universal tool for everything.

As we have already stated, risk analysis (sometimes called fault analysis or reliability analysis) is widely used in many technical fields. There is a wide variety of specialized software to compute both quantitative and qualitative results for various applications.

The problem of common spreadsheet software (like MS Excel – widely used in tunneling) is that it cannot deal with small numbers properly.

Risk is a probability, and probability is never greater than one. Numbers in the order of 10^{-15} are not uncommon when conducting such analysis. This means that the numbers are far smaller than the spreadsheet software is designed (or capable) of competently managing.

To manipulate such small numbers properly, special techniques that specially command hardware components of the computers (particularly processor memory and RAM) is required. Common software can thereby "forget" some very small (but important) numbers, which itself can lead to misleading results of the risk analysis computations.

This following table highlights the difference between the correct result (column 4) and the software induced error (column 3).

Table 1 - MS Excel, example of computational errors

		Software Answer	Correct result
X	X + 1	$(X + 1) * 10^6 - 10^6$	$(X + 1) * 10^6 - 10^6\ 10^6$
1.0000000000E-01	1.1000000000E+00	1.0000000000E+05	1.0000000000E+05
1.0000000000E-02	1.0100000000E+00	1.0000000000E+04	1.0000000000E+04
1.0000000000E-03	1.0010000000E+00	1.0000000000E+03	1.0000000000E+03
1.0000000000E-04	1.0001000000E+00	1.0000000000E+02	1.0000000000E+02
1.0000000000E-05	1.0000100000E+00	1.0000000000E+01	1.0000000000E+01
1.0000000000E-06	1.0000010000E+00	0.9999999999E+00	1.0000000000E+00
1.0000000000E-07	1.0000001000E+00	1.0000000009E-01	1.0000000000E-01
1.0000000000E-08	1.0000000100E+00	0.9999999893E-02	1.0000000000E-02
1.0000000000E-09	1.0000000010E+00	1.0000000475E-03	1.0000000000E-03
1.0000000000E-10	1.0000000001E+00	1.0000006296E-04	1.0000000000E-04
1.0000000000E-11	1.0000000000E+00	0.9999959730E-05	1.0000000000E-05
1.0000000000E-12	1.0000000000E+00	1.0001240298E-06	1.0000000000E-06
1.0000000000E-13	1.0000000000E+00	0.9988434612E-07	1.0000000000E-07
1.0000000000E-14	1.0000000000E+00	1.0011717677E-08	1.0000000000E-08
1.0000000000E-15	1.0000000000E+00	1.1641532183E-09	1.0000000000E-09

Using the wrong software can compromise the results of a risk analysis.

4. THE THIRD SIN – PROBABILITY IS NOT NECESSITY (I.E. THE SIN IS "IMPLYING A 'RETURN TIME' IS NECESSITY")

Misery loves company

No accidents last month? We bet there will be three of them next month…

Probabilistic models are not deterministic. Probabalistic models cannot predict when events will occur – even in principle.

Probability is often mistaken for frequency theory. For example it is often thought that if you suffer a hundred-year flood, another flood will not occur for one hundred years.

The problem with the low probability events is that the numerical description of their occurrence is not a frequency. Often when risk analysis is conducted there is a discussion about "low frequency – high consequence" events. Statistically speaking these are "low probability" events. Frequency is something that repeats regularly – disasters do not regularly occur.

In fact, disasters tend to accumulate. Benoit Mandelbrot [Mandelbrot, 1963] was working on data communication errors and discovered a surprising fact – the errors tend to accumulate (according to a so called Koch set). This has lead to origin of two rapidly growing mathematical branches – fractal theory and theory of catastrophes.

Catastrophe theory has been applied to tunnels, see [2]

This sin of risk analysis is not really a sin of risk analysis itself; it is a sin of those who interpret the results of quantitative risk analysis. Quantitative risk assessment always uses probabilities – not frequencies. The results are not definite guides; the results of quantitative risk analysis give nothing more than suggestions as to the outcomes.

5. THE FOURTH SIN – IGNORING STATISTICAL DEPENDENCIES

No hoof, no horse

One may think that of a horse is a fast means of transportation. If you have a horse, you are fast. If you have a horse with hoofs, it is even faster. If you have hoofs only – well, it is not enough. The hoofs and the horse are dependant on each other for the speed of the transportation.

The only free cheese is in the mouse trap

If you want free cheese, you have to put your hand into the mouse trap. There is no other option; these two features completely depend on each other.

Statistically dependent variables are a nightmare for statisticians. Without them, all formulae are elegant and simple. Their introduction produces complicated computations. If the dependency is not very clear, it may be "transferred" into variance – that is a common approach which may be correct in some cases.

A 'serious' analyst would usually try to account for conditionalities, although they may not have the information to do so well. The estimates they put in for conditional probabilities (or any probabilities) may not be well founded. Typically, they use historical statistics and these may not apply to the current system or the future. This is problematic for tunnels because the circumstances experienced today are significantly different to those experienced even a decade ago.

Our Fourth Sin concerns the vertical relationship of probabilities – the relationship of causes. We will discuss the horizontal relationship (cause – effect) in the Fifth Sin. In general, the probability of two independent causes (A or B) is

$$P(A \cup B) = P(A) + P(B)$$

However, there are situations where the dependencies cannot be neglected (or transferred). There are even situations where the relationship of two factors is evident. In such a case, the probability of two dependent causes is

$$P(A \cup B) = P(A) + P(B) - P(A, B)$$

wherein *P (A,B)* is the probability that both A and B will happen at the same time. Fortunately, it holds that

$$P(A) + P(B) \leq P(A) + P(B) - P(A, B)$$

so ignoring statistical dependencies brings you a *pessimistic* estimate, which is expensive, but is not fatal.

For example, the number of accidents depends in part on the traffic density and speed, so introducing those two factors into any risk analysis as independent inputs is not correct.

This highlights the important role of the subject matter expert in understanding the relationship between factors when conducting a risk assessment. Failure to understand the relationships between factors can be fatal to the risk assessment outcomes.

6. THE FIFTH SIN – SUMMATION OF CAUSE – EFFECT PROBABILITIES

A chain is no stronger than its weakest link

If one chain in the safety measures of the tunnel breaks down, all the other measures are useless.

A common approach to risk estimation is to give several factors a "grade" (like in school), e.g. in the range from 0 to 10. Then all the grades are summed and an average is made – this average is considered to be a final grade of the tunnel risk.

However, one of the very basic rules (technically, one of the three axioms of probability theory, as introduced by Kolmogorov [1933] and Cox [1946]) in probability says that the (joint) probability of two events in the position of cause – effect is the product of the probability of one of them and of the second, which is conditional of the second:

$$P(A \cap B) = P(A \mid B)P(B)$$

If the two events are independent, this rule reduces to

$$P(A \cap B) = P(A)P(B)$$

From above, we can see that the "school grading" mechanism is in contradiction with one of the axioms of probability. The result of a risk analysis tool that uses this type of grading is not a probability, but something that cannot be easily (and correctly) interpreted.

A tunnel safety analysis using a risk analysis sometimes focuses upon a numerical expression of tunnel safety. A "safe tunnel" is a tunnel that has low probability of accidents. It is also clear that one safety factor that is very low (e.g., there is no emergency ventilation system) influences the safety dramatically. This fact cannot be described by the "school grading" mechanism.

This type of analysis is known as a 'points schemes' which are different very different approaches to PRA models. A 'serious' PRA practitioner would certainly distinguish clearly. EUROTAP is a points scheme. The criticism of points schemes is different to that of PRA. In points schemes the mathematical logic is not based on probability theory but the 'common sense' of the analysts. They may be of value in that they may distill the experience and judgement of experts. However, experts' estimates may differ widely, when guessing at the same thing; as we found in the Bayesian work , [1].) (See also [6] for comments on points schemes.)

Practical example: If you throw a dice, the probability of having 6 is 1:6. The probability of having two consecutive 6 is not 2:6.

7. THE SIXTH SIN – IGNORING VARIANCE (TAKING THE AVERAGE ONLY)

A miss by an inch is a miss by a mile

Sometimes, it is important to take the less probable situations into account. For example if it is important to determine that temperature tenability requirements are met (for example - not more than 80C) and the modelling says "80C" there may still be a problem in reality.

If you can't be good, be careful

Statistical computations are only estimations. They can never be 100% accurate – it is better to be cautious about the results.

In everyday life, we describe the events by one number, which is the average (or mean value). "I was in the theatre and there were 400 people" means that there were probably 400 people, maybe a few more, maybe a few less. But how large is this "few"?

In basic statistics, the data are described by features called *statistical moments*. For basic work only the first two are needed – the *average* (more technical term is *mean value*) and *variance*. Variance is the number that describes our "few". If we want to have really reliable results, we have to say how certain we are about this result. It is not wrong to have a result that is not accurate, but it is not right to treat an inaccurate result as accurate and to invite reliance of the result by decision makers.

For example in the case of tunnel fires, where we have only few data with big uncertainties, it is essential to reveal the accuracy of the computations. This also means that all the input data have to be specified and the variance and computations become much more complicated. That is not easy, but it is the right way to describe the results.

For example the Heat Release Rate in the Runehamar test of 202MW has been estimated to have been at around the 59th percentile, transformed into a Channel Tunnel equivalent; see [7]. That is, one would expect about 40 in every hundred 'replicated' fires to be greater than 202MW).

A practical example of this phenomenon is a conclusion following "risk" analysis that the consequence of an identified event occurring is Euro 10,000,000M. In fact this might mean the likely consequence is anything between Euro 1,000 and Euro 100,000,000.

Variance is, and must be disclosed to be, central to any conclusions of the risk analysis. Relying upon an average is misleading.

8. THE SEVENTH SIN – ILL NUMERICAL CONDITIONING

A pot of milk is ruined by a drop of poison

A nice algorithm can be spoiled by a single "ugly" number.

Minor change in one parameter of the risk analysis may have enough power to change the overall result dramatically. The risk of this occurring is very high especially in complex probability networks. Statisticians understand that there may be a "vertex" that has am extremely strong influence on the overall result. This raises the very real prospect that the risk analysis, may be easily (and intentionally) "tailored" or unintentionally "spoilt".

This sin is broadly related to the Sixth Sin (Ignoring Variance), because for "tailoring", you may misuse the variation as well (because from a range of probabilities the data that suit your cause is selected). The point of misused numerical conditioning is that you focus the variation into the suitable parts of your calculations.

Table 2 shows an example of a simple (non-Bayesian) network. The data are fictitious, but it shows another numerical issue. In this example, the original data (black) were altered (red), particularly the probability of a passenger car fire was doubled. The increase seems insignificant in the table – no one really cares about the difference between 0.01 and 0.02; moreover, no one really checks the passenger cars data, as everyone is concerned about the heavy goods vehicles. But the changed data result in a significant increase in the overall fire probability – from 1.9 % to 2.8 %. This may lead to installation of more expensive safety equipment which is not needed at all.

Table 2 - Numerical conditioning

				FIRES	NO FIRES
accident 1	passenger 0.85	fire 0.01/0.02		**0.0085/ 0.017**	
		no fire 0.99/0.98			**0.8415/ 0.833**
	van 0.1	fire 0.04		**0.004**	
		no fire 0.96			**0.096**
	truck 0.05	normal 0.9	fire 0.1	**0.0045**	
			no fire 0.9		**0.0405**
		dangerous 0.1	fire 0.5	**0.0025**	
			no fire 0.5		**0.0025**
			TOTAL	**1.9% / 2.8%**	**98.1% / 97.2%**

Once again this demonstrates the need for subject matter experts – to help the risk analysts. Failure to do so may produce results which in no way reflect the actual risk profile of that which is being modelled – and lead to unsound decisions' based upon very impressive looking calculations.

EXAMPLES

EuroTAP

The objective of the EuroTAP (European Tunnel Assessment Programme) was to perform tests of European tunnels in order to compare them. In the earlier analysis the tunnel operators were given data sheets, which were further processed by a specific methodology and the degree of safety of that particular tunnel was determined using this data. The methodology of the project is described e. g. in [Tetzner, 2006].

The weakness in this original approach to the risk evaluation method used in the EuroTAP project lies in the limitations of relying on the Euro tap data to compare the safety of different tunnels.

The risk evaluation is divided into two parts – Risk potential and Safety potential, which are then compared. The risk potential and safety potential are computed as a sum of points rating different categories of risks and safety measures respectively.

This represents our Second Sin – Summation of probabilities. Probabilities cannot be summed up, unless they fully exclude each other. This is not the case of tunnels. For example – if there is no traffic in the tunnel, there will be no risk of vehicle fires regardless of other factors. The K. O. criteria introduced in the method are just a minor countermeasure; they do not make the computation any clearer.

The EuroTAP method also ignores variances and statistical dependencies of the input data. It also focuses on a particular set of problems and does not address other safety measures or risk potentials than the standard ones.

This does not mean that the EuropTAP's approach to characterising the risks of different tunnels is of no value. This review illustrates that the value of the EuroTAP analysis is limited – and that the results about the comparative safety of tunnels should be used as a *guide* to comparative safety and <u>not</u> a declaration of actual tunnel safety.

9. CONCLUSIONS

Understanding the limitations of risk analysis is fundamental to using risk analysis tools and techniques to effectively understand and manage risk. There is a shortage of data for such risk analysis to be conducted in many tunnel contexts. This lack of data makes proving the results are wrong, or even discovering the nature of the errors extremely difficult – except at the conceptual level (as described herein).

A similar situation arose in risk analysis for aviation, however, the aircraft safety has been the subject of intensive research since 1920's and adequate methods have now been found to deal with the special conditions, i.e. perform a reliable risk analysis on a limited set of "positive" occurrences (see e. g. [FAA, 2000]). For qualitative risk analysis, the use of well-established general methods would be also convenient, such as [Mil-Std-756b, 1981] or [Mil-Std-1629a, 1980].

With these limiting factors in mind it is essential that risk analysis be embraced with the intellectual caution it demands for application to low frequency events which typify incidents in tunnels.

The role of experts understanding of the complex interrelation of many factors – coupled with their sense of judgement should play a critical role in any risk assessment process.

Risk assessments are a tool to making informed decisions about risk management. Risk assessments are not a substitute for sound expert decision making about low frequency complex risk issues. Risk assessments should be a tool to assist in complex decision making. Risk assessments are not a substitute for making informed decisions on complex subjects. Sound decision making is demanded of modern tunnelling professionals.

Risk assessments are a tool for decision makers – they must be used wisely as a tool for exercising judgement, not a justification for poor judgement, or a means of justifying a predetermined outcome.

REFERENCES

[1] Carvel et al, 'Variation of Heat release rate with forced longitudinal ventilation in tunnels', Fire Safety Jnl, 36, pp569-596, 2001

[2] Beard, 'A theoretical model of major fire spread in a tunnel', Fire Technology, 42, pp303-328, 2006.

[3] Beard, Chapter on 'Problems with using models for fire safety', *The Handbook of Tunnel Fire Safety*. The chapter is generic and contains many more references.

[4] Beard & Cope, Assessment of the safety of tunnels, report for the European Parliament, 2008. Available free on the web-site of the EP; follow to STOA ('science and technology options assessment', and reports.

[5] Beard, 'Requirements for acceptable model use', *Fire Safety Jnl*.,40, pp477-484, 2005

[6] Beard, Alan, Letter to the Editor, *Fire Technology*, 19, pp 69-70, 1983

[7] Beard, Alan, 'HGV Fires in tunnels: Fire Size and Spread', *Tunnel Safety Forum for Road and Rail, 2nd Int Conference*, Lyon, 20-22 April, 2009.

BIBLIOGRAPHY

[Cox, 1946] Cox, R. T., "Probability, Frequency, and Reasonable Expectation," *Am. Jour. Phys.*, vol. 14, pp. 1-13, 1946.

[Kolmogorov, 1933] Kolmogorov, A. N., *Foundations of the Theory of Probability* (in Russian). Bull. Math. Univ. Moscow, 1933.

[Mandelbrot, 1963] Berger, J. M. and Mandelbrot, B., "A New Model for Error Clustering in Telephone Circuits", *IBM Journal*, pp. 224-235, 1963.

[Tetzner, 2006] Tetzner, D., "The European Tunnel Test in the Scope of the EuroTAP Project", *In Proceedings of Tunnel Safety and Ventilation,* Graz, 2006.

[FAA, 2000] *FAA System Safety Handbook*, 2000.

[Mil-Std-756b, 1981] *Reliability Modeling and Prediction*, U. S. Military Standard, 1981.

[Mil-Std-1629a, 1980] *Procedures for Performing a Failure Mode, Effects and Criticality Analysis*, U. S. Military Standard, 1980.