

Application of Sensitivity Analysis, “Worst Case”, and Maximum Possible Risk (MPR) to Adventitious Events

T. H. Taylor, Jr.¹ T. Whalen² M. Cohen³

¹Centers for Disease Control and Prevention, Atlanta, Georgia, USA (The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention)

²Georgia State University, Atlanta Georgia USA; tom@whalen3.org

³Frontline Healthcare Workers Safety Foundation, Atlanta Georgia USA

Abstract

We present here a logical progression of probability and risk analysis for adventitious events, events whose probability is not well measurably different from zero^[1] (WMDZ). We will show that such analyses culminate in maximum possible risk^[2] (MPR) and, further, that MPR is equivalent to a boundary condition for classic sensitivity analysis when applied to events which are not WMDZ. Further, we shall show that use of counter-factual probabilities provides a good estimate of these boundary conditions.

Keywords: statistics, probability, risk analysis, bioterrorism, anthrax

1. Background

Few risk-analysis techniques perform well for events or outcomes whose probability is very near zero. Such events are *adventitious*, having happened very rarely or having *never* happened. Examples include the eruption of Mt. St. Helens, and release of *Bacillus anthracis* spores from a research laboratory. Because no or insufficient data exist to well-characterize these events, the classic estimates of probability lack sufficient basis. The ob-

vious assumption of $P(\text{event})=1/N$ for some very large but arbitrary N presumably allows classic risk analysis, but such unfounded assumptions are not appropriate and ignore the legitimate advances in decision sciences.

2. What’s wrong with classical probability analysis for adventitious events?

Classic risk analysis starts with the formal mathematical definition of risk:

$$Risk(event) = P(event) \times Loss(event) \quad (1)$$

where loss is assigned in dollars, or lives, or lost-time, or any other measure of interest. Scenarios are compared by varying the probability inputs, the inputs to the Loss function, or even the events themselves. Comparison of disparate events, such as terrorist versus accidental release of pathogenic materials is possible due to the consistent definition of risk. The conventional scenarios include known confidence limits around the inputs, up to a “worst case” which might be defined statistically, e.g. low p-value. Varying the input assumptions is frequently referred to as *sensitivity analysis*.

This is the classic structure, sometimes accompanied by the assertion that the loss estimates are valid only in the context of *comparisons* between scenarios - that the *differences* between the various computations are indicative of real differences between scenarios. For scenarios which include adventitious events or events whose probabilities are not WMDZ, the classic approach is fundamentally flawed due to an inappropriate mixing of rigor with subjective “data”.

2.1. How did we come to this?

Mathematically, risk is the product of probability and loss. Loss assumes a *threat* - an event which causes harm, quantifiable in dollars, cases, lives, etc. In the non-mathematical arena risk falls into two subjective categories, segregated by presence of data:

Credible ~ data- or evidence-based belief that a particular threat *can* happen and, without countermeasures or intervention, will cause harm (loss). Example: power failure in a lab.

Plausible (or Conceivable) ~ not data- or direct-evidence-based, i.e. may have never happened, but following from a defensible, logical extension of facts. Example: earthquake in a location where none has occurred in recorded history but which is near a known fault line. For the purposes here, there is no distinction between plausible, conceivable, imaginable, or other terms which imply our belief that an event could occur but has not done so with measurable probability.

Credible and plausible threats extend naturally to credible and plausible probabilities and risks. For credible events, the probability can be legitimately estimated based on the data. Such computations are routinely performed, accurately

and precisely, in the actuarial arena for a wide range of insured threats, and in the public-utility and public-health arenas for a range of undesirable events such as power outages and disease outbreaks.

Sequential events, or sub-events, which together define an “EVENT” can be evaluated separately and combined. For dependent events:

$$P(EVENT) = P(event_1) \times \prod_{i=2}^n P(event_i | event_{i-1}) \quad (2)$$

where EVENT is defined by the simultaneous or sequential occurrence of events 1 through n (event₁ – event_n, inclusive).

For independent events, the probability of credible events may be computed as:

$$P(EVENT) = \prod_{i=1}^n P(event_i) \quad (3)$$

where events 1,2,3... are independent and together define the event of interest. For example, for power failure to cause community infection (and assuming independence for this discussion only):

$$\begin{aligned} P(\text{infection in the community}) = & \\ & P(\text{power failure at the lab}) \times \\ & P(\text{failure of fail-safe systems}) \times \\ & P(\text{an infectious agent in use}) \times \\ & P(\text{infectious exposure level in community}) \times \\ & P(\text{disease, given infection}) \times \\ & P(\text{mortality, given disease}) \end{aligned}$$

Each of the components can be estimated from historical data. So, credible events give rise to this definition of probability:

Probable ~ Measurably probable, measurably different from zero probability

This definition of probable does not mean likely, which might be inferred to be a probability greater than some threshold

such as 50% or higher. The definition of probable extends naturally to risk, assuming that a loss function can be well estimated, such as the cost of treating a case, or assigned, such as Loss(death)=1 and Loss(survival)=0. Either way, the risk is measurable.

2.2. Acceptable Standard

Once a measurable risk is computed, the final step toward decision making is comparing that risk with an acceptable standard. Establishment of a standard is a subject of substantial debate, but for the purposes here, we assume there *is* such a value, δ , which is presumed to be small. Hence the decision test is:

$$Risk(EVENT) < \delta$$

For public health issues where the outcome of an event is mortality, a loss function of 1 is commonly assigned, in which case the above test reduces to:

$$P(EVENT) < \delta$$

where δ is the probability of some event that we normally accept, such as the probability of being fatally struck by lightning. A key point is that $\delta > 0$ ^[3].

2.3. Adventitious events

For adventitious events, little or no basis for probability estimates exists; hence the classic structure is not appropriate. Instead, we must augment classic theory with our last definition:

Possible ~ credible or plausible, but with probability NOT well measurably different from zero (not WMDZ), except:

The upper bound for mathematically possible is 100% possibility, in which case, by the strong principle of probability-

possibility consistency^[4], the probability becomes measurable and classic probability analysis is appropriate. The definition of possible extends directly from probability to risk, assuming a finite loss function.

2.4. Analyzing “possible” risks

Events which are probable must be separated from events which are merely possible. Equation 3 can be re-written:

$$P(EVENT) = \prod_{i=1}^k P(event_i) \times \prod_{i=k+1}^n P(event_i) \quad (4)$$

where events 1, 2, ...k are supported by data and events k+1, k+2, ...n are not. The former group includes the k events with credible probabilities such as power outages, infection rates given exposure, mortality rates given infection. The other n-k events arise from imagination, such as: “all escaped spores from an accidental lab spill are carried by the wind in the same direction without significant dispersal”.

2.5. Sensitivity Analysis and Worst Case

Classic sensitivity analysis would lead us to vary all the terms on the right side of Equation 3. Dependent on whether an event_i is desirably high or low, lower and upper confidence limits for each term represent “worst case”. The confidence limits for events 1 through k might be based on known 95% confidence, or 99%, or 99.9999% or more, up to six-sigma for some situations. However, for terms k+1 through n, there are no data-based upper confidence bands. Hence, such an approach is not appropriate due to the mixing of rigorous statistical estimates with unsupported, subjective values. The sum A+B, where A is accurate and precise but B is assumed, does not magically achieve the precision of A; but,

rather, the sum drops to the unknown status of B.

2.6. Maximum Possible Risk

Instead of unsupported confidence limits for adventitious events, we must introduce counterfactual probabilities, which, in this case, represent the worst possible “worst case”, or the maximum possible risk (MPR). Our counterfactual assumption is that the probability of each adventitious event is 1.0; that is, we shall assume, counterfactually, that they all occur.

Equation 4 can then be re-written:

$$P(EVENT_C) = \prod_{i=1}^k P(event_i) \times \prod_{i=k+1}^n (1.0) \quad (5)$$

where EVENT_C is EVENT under the counterfactual assumptions. This can be re-written

$$P(EVENT_C) = \prod_{i=1}^k P(event_i) \quad (6)$$

where events 1, 2, ...k are supported by data. We can now legitimately compute a value for the test

$$P(EVENT_C) < \delta$$

which is equivalent to:

$$\prod_{i=1}^k P(event_i) < \delta$$

Since *all* events have probability less than or equal to one and the counterfactual events have probability equal to one, we have:

$$\prod_{i=k+1}^n P(event_i) \leq 1.0 = \prod_{i=k+1}^n P(eventC_i)$$

and

$$\prod_{i=k+1}^n P(event_i) \leq \prod_{i=k+1}^n P(eventC_i) \quad (7)$$

where eventC_i's, i=k+1 to n are the events with counterfactual probabilities. Multiplying both sides by the non-negative value $\prod_{i=1}^k [P(event_i)]$, we have

$$\prod_{i=1}^k P(event_i) \times \prod_{i=k+1}^n P(event_i) \leq \prod_{i=1}^k P(event_i) \times \prod_{i=k+1}^n P(eventC_i) \quad (8)$$

which is simply

$$P(EVENT) \leq P(EVENT_C)$$

so we have the ordered relationship:

$$P(EVENT) \leq P(EVENT_C) < \delta \quad (9)$$

So if

$$P(EVENT_C) < \delta \quad (9)$$

which can be tested with data from well-characterized history, then

$$P(EVENT) < \delta \quad (10)$$

The real event of interest, complete with adventitious elements, has a lower probability than the acceptable standard, δ , even though we can't estimate the adventitious probabilities. The standard, δ , is an upper bound for P(EVENT).

Once again assuming a Loss function assigned to 1.0, the upper bound of the probability extends to the risk:

$$Risk(EVENT) = P(EVENT) \times 1.0 = P(EVENT) < \delta \quad (11)$$

3. Decision Making

Selection of δ for the test must be made *a priori* and without consideration of the EVENT. This is a key requirement for the intellectual ethics of the analysis. The selection of δ is subjective – what can we “live with”? In addition, we must establish a Loss function. For event=mortality, many public health analyses assign the Loss function:

$$L(\text{EVENT}) = \begin{cases} 1, & \text{if the event occurs} \\ 0, & \text{otherwise.} \end{cases}$$

Using this loss function essentially equates probability and risk. This blending of the terms has obvious health-communication advantages and bypasses the sticky question of estimating the value of a human life.

The mortality rate for influenza and pneumonia (combined) for 2006 in the United States was 18.8 per 100,000^[5]. This translates to a probability of 0.000188; using a Loss function value of 1.0, the flu and pneumonia risk with which Americans live everyday is

$$\text{Risk(mortality from flu or pneumonia)} = 0.000188$$

For the period 1990-2003, the average annual U.S. human death count from lightning was 54^[6]. During the same period the U.S. population^[7] averaged approximately 270,000,000; again using Loss(death)=1.0, the risk is:

$$\text{Risk(being killed by lightning)} = 54/270,000,000 = 0.000002$$

The most dangerous cities in the U.S., as defined by murder rates, number about 50 murders per year per 100,000 population^[8]. Again, using Loss(death)=1.0, some city dwellers accept:

$$\text{Risk(being murdered)} = 50/100,000 = 0.0005$$

Selecting δ in the range of these known and accepted (begrudgingly or not) probabilities is reasonable and prudent. Similarly, selecting an arbitrarily higher δ , say corresponding to some high percentile on a Normal distribution might also be considered reasonable and prudent. For example, the traditional p-value corresponding to six-sigma for the standard Normal curve is 0.000000006. However, such a standard might impose impossible restrictions - well beyond what we accept every time we get into a car. A conscientious practitioner of risk analysis should not assign a risk limit which cannot possibly be met. *Setting unrealistically low thresholds of acceptability inevitable increases the risk of other hazards by diverting attention and resources away from them. In extreme cases it can actually increase total risk borne by a community.*

3.1. What if P(EVENT) is not < δ .

If, on the initial analysis, the risk of an EVENT is greater than the reasonable and prudent δ , then the EVENT represents unreasonable risk. In order to continue with the associated project or action, we must lower the risk. To do so is not a mathematical or analytical exercise; rather we must change something in the project.

For example, in the lab example above, we might adopt a policy that infectious agents not be used during times of likely thunderstorms and therefore likely power outages. It would then be appropriate to compute a new P(EVENT) using the lower value of P(power outage|infectious agent in use) instead of the product of the two independent terms, because the conditional event is less likely to occur.

3.2. A comment on the time variable

The values which have been cited for standards are annual probabilities – ultimately representing the risk of the specified event during one year. An additional analysis required for evaluation of many real EVENTS is consideration of multiple years. Selecting an appropriate δ for multiple years or conducting a *post facto* multi-year analysis is beyond the scope of this presentation, but is necessary and is left to the practitioner.

3.3. A note on the Loss function

Generalizing Equation 11 to loss functions other than 1.0 or going back to the definition of risk, if the Loss function has value L for a particular EVENT, then

$$Risk(EVENT) = P(EVENT) \times Loss(EVENT)$$

so the test becomes

$$P(EVENT) \times L ? < \delta$$

which for positive values of Loss becomes

$$Risk(EVENT) ? < \delta / L .$$

If the risk is not less than δ/L , which might be very much lower than δ , then different countermeasures, policies, protocols, or interventions must be determined to reduce one or more of the $P(event_i)$'s and therefore $P(EVENT)$ and $Risk(EVENT)$.

4. Summary

We have presented a logical progression from classic probability and risk analysis to MPR analysis. Though the former is appropriate for well-characterized, probable events, many real world events in-

clude adventitious components, whose probability is not well measurably different from zero. For those, classic risk analysis breaks down; an alternate approach is required and has been presented. The MPR approach more accurately provides sensitivity analyses, “worst case” analyses, and a mathematically-supported decision-making approach. When counterfactual probabilities are included, MPR is a boundary condition for classic sensitivity analysis.

5. References

- [1] Whalen et al, Possibilistic Risk and Counterfactual Probabilities, Proceedings of the 11th Joint Conference on Information Sciences, China, 2008
- [2] Schuetz et al, Maximum Possible Risk Modeling, Proceedings of the 11th Joint Conference on Information Sciences, China, 2008
- [3] Inhalation Anthrax: Dose Response and Risk Analysis, Margaret E. Coleman et al, Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science, June 1, 2008, 6(2): 147-160.
- [4] G. Klir, and B. Yuan, Fuzzy Sets and Fuzzy Logic, Prentice-Hall, NJ, 1995.
- [5] National Vital Statistics Reports, Vol. 56, No. 16, June 11, 2008 (http://www.cdc.gov/nchs/data/nvsr/nvsr56/nvsr56_16.pdf)
- [6] http://www.lightningsafety.com/nlsi_1ls/fatalities_us.html
- [7] In Census 2000, 281.4 million people were counted in the United States, a 13.2 percent increase from the 1990 census population of 248.7 million (<http://www.census.gov/prod/2001pubs/c2kbr01-2.pdf>)
- [8] http://www.nola.com/news/index.ssf/2008/01/no_murder_rate_remains_a_mong_h.html