

Possibilistic Risk and Counterfactual Probabilities

T. Whalen¹, T. Taylor², M. Cohen³

¹Georgia State University, Atlanta Georgia USA; tom@whalen3.org

²Taylor Engineering, Atlanta, Georgia USA

³Frontline Healthcare Workers Safety Foundation, Atlanta Georgia USA

Abstract

Possibilistic risk management focuses on reducing the relevance of historical events by countermeasures to prevent their recurrence. A new possibility measure uses the ratio of the point estimate of a probability to the width of its confidence interval. Possibility theory separates the probability and possibility of disease after a release of hazardous material from the probability and possibility of that release. Hazardous release becomes a "counter-factual assumption" treated as certainty although its probability is really not well measurably different from zero. The confidence interval and point estimate of disease probability can be calculated using appropriate toxicological and microbiological techniques.

Keywords: possibility, risk analysis, counterfactuals, laboratory safety, anthrax

1. Introduction

Comparing risks of rare, high consequence events, for example, the risk of a natural or terrorist-induced outbreak of a rare or emerging disease such as anthrax or Ebola for which a society is unprepared, versus the risk of an outbreak caused by accidental release from a research laboratory studying ways to become prepared for the former event, poses many deep methodological and epistemological problems. Nevertheless,

it is necessary to be able to assess the merits of counter-measures that are only necessary in extremely unlikely circumstances. Conventional conditional probabilities $P(A|B) = \frac{P(A \cap B)}{P(B)}$ break down when P(B) is so small relative to the uncertainty surrounding it as to make P(B) not well measurably different from zero. Possibility theory offers a solution to this dilemma.

2. Adventitious Events

Suppose event Q has happened x times over an exposure period T in the past. The point estimate of the Poisson rate of occurrence of Q, λ , is $\hat{\lambda} = \frac{x}{T}$.

The exact equal-tailed 95% confidence interval for λ , $[\lambda_L, \lambda_H]$, is based on the confidence interval for the average number of occurrences in all possible exposures of length T , $[x_L, x_H]$. [7] The values are:

$$\begin{aligned} x_L &= \frac{\chi_{.975}^2(2x)}{2}, & \lambda_L &= \frac{x_L}{T} \\ x_H &= \frac{\chi_{.025}^2(2x+2)}{2}, & \lambda_H &= \frac{x_H}{T} \end{aligned} \quad (1)$$

The point estimate for the probability that Q will never occur in time interval t for a given λ is

$$P(X=0|\lambda) = e^{-\frac{t}{T}\lambda}. \quad (2)$$

Therefore, the confidence interval for the probability that Q will not occur in a given future interval of length t is $[e^{-\frac{t}{T}\lambda_H}, e^{-\frac{t}{T}\lambda_L}]$ with point estimate $e^{-\frac{t}{T}\hat{\lambda}}$. The confidence interval for the probabil-

ity that Q will occur at least once in an interval of length t is $[1 - e^{-\frac{t}{T}\lambda_L}, 1 - e^{-\frac{t}{T}\lambda_H}]$ with point estimate $1 - e^{-\frac{t}{T}\hat{\lambda}}$.

Definition 1: the Poisson adventitiousness of event Q is

$$\delta = \frac{e^{-\frac{t}{T}\lambda_L} - e^{-\frac{t}{T}\lambda_H}}{1 - e^{-\frac{t}{T}\hat{\lambda}}} \quad (3)$$

The inverse of Poisson adventitiousness, capped at 100%, can be used as a possibility measure [2] [3]:

Definition 2: The possibility of event Q is:

$$\pi_\delta = \min(1, \frac{1}{\delta}) = \min(1, \frac{1 - e^{-\frac{t}{T}\hat{\lambda}}}{e^{-\frac{t}{T}\lambda_L} - e^{-\frac{t}{T}\lambda_H}}) \quad (4)$$

If Q has occurred fewer than 17 times historically, its possibility is less than 1; such an event has a probability not well measurably different from zero, in the sense that the uncertainty of the probability is greater than the difference between the probability and zero.

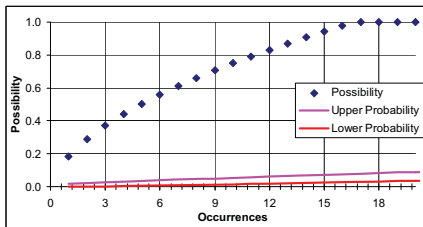


Figure 1:
Possibility and Probability of 1 Or More Occurrences in Future Interval t

The weak probability-possibility consistency principle [3, p. 206] requires that the probability of any event is less than or equal to its possibility. π_δ , the possibility measure defined above, satisfies this condition trivially when it is equal to 1; otherwise, π_δ is equal to the probability of an event divided by a

quantity that is always less than 1, so the weak probability- possibility consistency principle is always satisfied. The strong probability-possibility consistency principle requires that

$$p(A) > 0 \rightarrow \pi(A) = 1. \quad (5)$$

π_δ does not satisfy this condition in the strict sense, but we can define a "quasi-strong" consistency principle

Definition 3: A possibility measure p satisfies the quasi-strong consistence principle iff

$$p(A) >^* 0 \rightarrow \pi(A) = 1 \quad (6)$$

where $p(A) >^* 0$ means that the probability of A is well measurably different from zero in the sense defined above. Clearly, π_δ does satisfy this quasi-strong probability-possibility consistency principle.

The past history of occurrences of a rare event necessarily must extend over a long period of time. Furthermore, when a harmful event occurs, efforts are generally made to prevent its recurrence. This is especially true in engineered systems such as high containment laboratories, where changes in equipment and procedures are made, as well as changes in the architecture of future laboratories and retrofits to existing ones. Thus, often the database of past occurrences of a harmful rare event is best modeled as a fuzzy set, and the "number" of truly relevant past occurrences is a measure (such as sigma count) of this set. In such a case, the number of occurrences becomes a continuous variable and the isolated points for possibility in Figure 1 become a smooth curve.

3. Hazardous Material in a Research Facility

A proposed research laboratory has the capacity to handle a hazardous material; there is a measurable probability that it will in fact handle it. If so, it will use ultra-reliable containment measures that make the probability of a release to the environment very small, but not theoretically impossible.

There may be a history of a small number of comparable materials releases from comparable laboratories similar to the one in question. However, when a release has occurred in the past, procedures in existing and new laboratories and design features in new laboratories are changed to make releases similar to the historical ones impossible or at least much less possible. Thus, the historical data base of releases (and near misses) should be viewed as a set of fuzzily comparable cases rather than a strict count of occurrences.

4. Possibilistic Risk Analysis

Over the lifetime of the laboratory, hazardous materials such as biological "select agents" requiring BSL-4 containment, may or may not be handled. If such materials are handled, they may or may not be released to the environment. If they are released to the environment, they may or may not cause one or more cases of disease.

The probability that hazardous materials will be handled is presumably large if the facility is built with that capacity, and, by law, zero if it is not. Since the probability that hazardous material will be handled is well-defined, the possibility is 100%.

If hazardous material is indeed handled, there are generally several alternative release scenarios that can be

imagined. The chief concern is usually whether harmful or potentially harmful events will occur at least once. This implies that the probability that hazardous material will be released is the greatest probability of any of these scenarios. As discussed above, this probability will not be well measurably different from zero in a modern high containment laboratory, so the possibility will be greater than zero but less than 100%.

If it happens that hazardous material is released, there are alternate pathways by which it might or might not lead to disease. In addition to qualitative distinctions, each pathway can lead to various numbers of cases. The question for risk assessment is often "What is the probability of at least one case of disease?" Clearly this is the largest probability on the sequence 1 case, 2 cases,

Since the unconditional probability of disease caused by release of hazardous material must be \leq the probability of such a release, the unconditional possibility of disease must be less than 100%.

Relative frequency estimation of the probability release, and hence of disease, is impractical since they have been observed so rarely that reliable statistics cannot be computed without margins of error so large as to render them useless.

Subjective probability estimation is suspect because it involves asking people for opinions on probabilities far too small to be within their actual subjective experience.

4.1 Singleton Events

X_{hrs} = hazmat is not handled in the lab. $P(X_{hrs})$ is assumed to be well defined, so $r_{hrs} = r(x_{hrs}) = 1$ X_{Hrs} = hazmat is handled in the lab AND not released AND no-one gets sick. $P(X_{Hrs})$ is assumed to be well defined, so $\rho_{Hrs} = r(X_{Hrs}) = 1$
--

X_{HRS} = hazmat is handled in the lab AND released according to the most possible scenario AND no-one gets sick.
 $\rho_{HRS} = r(X_{HRS}) = \varepsilon$
 X'_{HRS} = hazmat is handled in the lab AND released according to a less possible scenario AND no-one gets sick .
 $\rho'_{HRS} = r(X'_{HRS}) \leq \varepsilon$
 X_{HRS} = hazmat is handled in the lab AND released AND one or more people get sick. The possibility of X_{HRS} is
 $\rho_{HRS} = r(X_{HRS}) = \sigma$
 X'_{HRS} = hazmat is handled in the lab AND released AND two or more people get sick . The possibility of X'_{HRS} is
 $\rho'_{HRS} = r(X'_{HRS}) \leq \sigma$
 $(X_{HRS}, X_{HRS}, \text{ and } X_{HRS} \text{ are strictly impossible}).$

4.2 Focal Events

Since release of a hazardous material is only possible if it is present in the laboratory and the material can only cause disease if it is released, the singleton events arrange themselves naturally into nested focal events.

When all the focal events in a Dempster-Shafer-Smets model [1] [5] [6] are nested, each focal event has a total possibility of 100% since each event with nonzero possibility has a non empty intersection with every other such event. More interestingly, each focal event has a "transferable belief" [Smets 1990] which is very similar to a Dempster-Shafer "basic probability." The transferable belief in a focal event A is denoted $m(A)$;

$$0 \leq m(A) \leq 1, \sum_i m(A) = 1 \tag{7}$$

The transferable belief in the largest focal event is the possibility of the most possible singleton event contained in it but not in any smaller focal event. In the example, the largest focal event is Ω , the Universal Set in which hazmat may or not be handled, released, or cause sickness. Its transferable belief is equal to σ ,

the possibility of X_{HRS} the most possible singleton event not contained in any smaller focal event.

The transferable belief in an intermediate focal event is the possibility of the most possible singleton event contained in it but not in any smaller focal event, less the possibility of the most possible singleton event not contained in it. In the example, the transferable belief in focal event A_s , (no one gets sick whether or not hazardous material is ever released) is $\varepsilon - \sigma$, the difference between the possibility of X_{HRS} and X_{HRS} .

Since the focal events are nested, the evidence model reduces to a possibility model. Each of the three focal events intersects every other focal event and shares in its transferable belief, so each of the three focal events has 100% possibility (Dempster-Shafer plausibility).

$A_r = \{X_{HRS}, X_{HRS}\}$ = focal event that hazmat is not released. $m(A_r) = 1 - \varepsilon$
 $A_s = A_r \cup \{X_{HRS}, X'_{HRS}\}$ focal event that no-one gets sick $m(A_s) = \varepsilon - \sigma$
 $\Omega = A_s \cup \{X_{HRS}, X'_{HRS}\}$ Universal Set: hazmat may or not be handled, released, or cause sickness. $\Omega = m(\Omega) = \sigma$

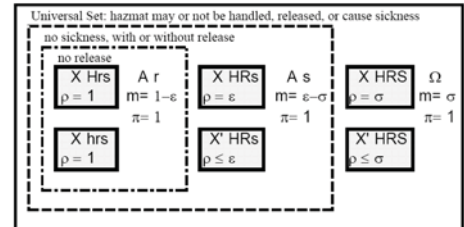


Fig. 2: Possibility Diagram for Hazardous Material

4.3 Possibilities of Important Composite Events:

$\pi(A_r) = \pi(A_s) = \pi(\Omega) = 1$ Consonant focal events are completely possible.
 $\pi(\Omega - A_r) = \pi(\{X_{HRS}, X'_{HRS}, X_{HRS}, X'_{HRS}\}) = \varepsilon$ Possibility that hazmat is released

$\pi(\Omega - A_s) = \pi(\{X_{HRS}, X'_{HRS}\}) = \sigma$ Possibility that hazmat is released and someone gets sick
 $\pi(A_s - A_r) = \pi(\{X_{HRS}, X'_{HRS}\}) = \varepsilon$ Possibility that hazmat is released and no-one gets sick

5. Numeric Example

Suppose 400 laboratories comparable to the one proposed have been doing similar research worldwide for the past 25 years. During this $T = 10,000$ laboratory-year historical period, releases of hazardous materials have occurred $x = 3$ times. This may be a simple count of a crisp set of comparable laboratories and materials, or the sigma-count of a fuzzy set.

The exact equal-tailed 95% confidence interval for the Poisson rate of occurrence of a release per 10,000 laboratory-years is $[0.6, 8.8]$ with a point estimate of 3. Assuming a 30 year useful life of the new laboratory, it is convenient to transform this to a confidence interval of $[\.0019, \.02633]$ occurrences per 30 laboratory years with a point estimate of .0090.

The corresponding probability of 1 or more releases in 30 years in $[\.0019, \.0260]$ with a point estimate of .0090, while the confidence interval for the probability of zero releases is $[\.974, \.998]$ with a point estimate of 0.9910. The possibility that hazardous material will be released during the life of the proposed laboratory is

$$\pi(\Omega - A_r) = \pi(\{X_{HRS}, X'_{HRS}, X_{HRS}, X'_{HRS}\}) = \varepsilon = \min(1, \frac{.009}{.026-.002}) = .37 \quad (8)$$

while the possibility that it will not be released is

$$\pi(\Omega) = m(\Omega) + m(A_s) + m(A_r) = \pi(\{X_{HRS}, X'_{HRS}, X_{HRS}, X'_{HRS}\}) = 1.0 \quad (9)$$

Assume that the conditional probability $p(S|R)$ that someone will get sick if

the material is released is known with negligible uncertainty. Then the probability of 1 or more cases of disease resulting from release in 30 years is between .0019 $p(S|R)$ and .0260 $p(S|R)$ with a point estimate of .0090 $p(S|R)$. The possibility measure for this is

$$\frac{.009p(S|R)}{.026p(S|R)-.002p(S|R)} = \frac{.009}{.026-.002} \frac{p(S|R)}{p(S|R)} = .37. \quad (10)$$

Similar calculations show that the following three nonfocal events all have equal possibility whenever $p(S|R)$ is known with negligible uncertainty:

- $(\Omega - A_r)$ hazmat is released,
- $(\Omega - A_s)$ hazmat is released and someone gets sick
- $(A_s - A_r)$ hazmat is released and no-one gets sick.

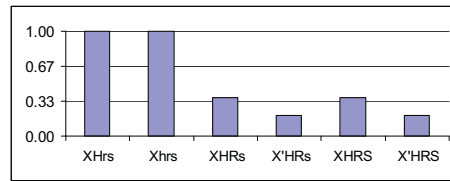


Fig. 3: Possibility Distribution of Singleton Events

5.1 Basic Distribution On Focal Events

- $m(A_r) = m(\{x_{Hrs}\}) =$ transferable belief hazmat is not released $= 1 - \varepsilon = .63$
- $m(A_s) = m(\{x_{Hrs}, x_{HRS}\}) =$ transferable belief no-one gets sick whether or not hazmat is handled $= \varepsilon - \sigma = 0$
- $m(\Omega) = m(\{X_{Hrs}, X_{hrs}, X_{HRS}, X'_{HRS}, X_{HRS}, X'_{HRS}\}) =$ unrestricted transferable belief $= \sigma = .37$

6. Counterfactual Conditional Probabilities

Possibilistic risk reduction works to reduce the relevance of historical events and near misses by countermeasures to prevent their recurrence. Guarding against release modes that are theoretically possible but have never occurred at all can be included in this framework by

suitably extending the concept of a "near miss."

A key simplifying assumption in the numerical example above is that the conditional probability of disease given release is well defined. In practice, conditional probabilities used in environmental risk assessment are often based on laboratory dose response studies. The typical study exposes animals to doses in the neighborhood "LD50," the level at which half the animals succumb. The proportion of animals affected for doses somewhat larger or smaller than LD50 is fit to a pre-specified statistical curve (e.g. lognormal) and extrapolated by several orders of magnitude to estimate the risk of exposure to the far smaller doses typical of accidental releases.

Possibility theory [9] can separate the probability and possibility of disease given that there is a release of hazardous material from the probability and possibility of the release itself. Release of hazardous material becomes a "counterfactual assumption;" [4] in effect, we treat the probability of a release as 100% even though we know it is really not well measurably different from zero. Based on this counterfactual assumption, the confidence interval and point estimate of the probability of disease can be calculated using appropriate toxicological and/or microbiological techniques.

Definition 4: A conditional probability $p(A|B)$ is counterfactual if the probability of the "given" B is not well measurable different from zero.

If the point estimate is greater than the width of the confidence interval, then the occurrence of disease assuming a release is completely possible, so the possibility of the event that hazardous material is released and causes disease is equal to the possibility that hazardous material is

released whether or not disease ensues, as in the numeric example above.

If the confidence interval width exceeds the magnitude of the point estimate, as is likely with extrapolations of several orders of magnitude, then the disease possibility assuming release can be calculated as a number strictly between zero and 100%. Then the possibility of the event that hazardous material is released and causes disease is a function of the possibility of the release and the possibility of sickness assuming release. The most straightforward function to use is the Zadeh conjunction operator [8], according to which the possibility that a release will occur and lead to disease is

$$\min \left\{ \begin{array}{l} \text{poss} \left(\begin{array}{l} \text{hazardous material} \\ \text{will be released} \end{array} \right), \\ \text{poss} \left(\begin{array}{l} \text{disease will ensue} \\ \text{assuming a release} \end{array} \right) \end{array} \right\}. \quad (11)$$

References

- [1] A.P. Dempster, "Upper and Lower Probabilities Induced by a Multivalued Mapping," *Annals of Math Statistics* 38:2, pp. 325-339, 1967.
- [2] G. Klir, and T. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall NJ, 1988.
- [3] G. Klir, and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice-Hall, NJ, 1995.
- [4] D. Lewis, *Counterfactuals. 2Rev Ed* Wiley-Blackwell, 2001.
- [5] G. Shafer, *A mathematical theory of evidence*. Princeton Univ. Press, Princeton NJ, 1976.
- [6] P. Smets, "The Combination of Evidence in the Transferable Belief Model," *IEEE Trans Pattern Analysis & Machine Intelligence* 12:5 p. 447-458, 1990.
- [7] K. Ulm, "A simple method to calculate the confidence interval of a standardized mortality ratio," *Am J Epidemiology* 131(2):373-375, 1990.
- [8] L. Zadeh, "Fuzzy sets," *Information and Control* 8, pp. 338-358, 1965.
- [9] L. Zadeh, "Fuzzy Sets as a Basis for a Theory of Possibility," *Fuzzy Sets and Systems* 1:1 pp. 3-28, 1978.