

# **Extensions to Conventional Surety Analysis NST 460**

**John Darby  
Sandia National Laboratories  
October 6, 2014**

**Update of SAND2009-6649P**

**Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration  
under contract DE-AC04-94AL85000.**



# Objectives

---

- **Present advanced surety/reliability analysis techniques as an extension of traditional probabilistically based surety/reliability analysis techniques**
- **Provide understanding of need for advanced techniques**
- **Summarize advanced techniques by simple examples**
- **Discuss software tools available for implementing advanced techniques for not so simple real world problems**



# To the Student

---

- Pay attention to the **concepts**
  - e.g., epistemic uncertainty vs. aleatory uncertainty
- Understand that the techniques have an **axiomatic basis** developed by mathematicians
  - Don't try to “roll your own”
- Understand that there are people, references, and software **available to help you**
- We will cover a **lot of material quickly**
  - Not a college course
  - Will not cover all the material included here
    - Backup material included with more details
- Relax and enjoy the course😊



# Course Topics

---

- **Part I: Traditional Techniques**

- Risk measure for safety
- Probability measure of uncertainty
- Probabilistic Risk Analysis (PRA)

## How to Treat Uncertainty

- Bayesian concepts
- Epistemic uncertainty
- Belief/Plausibility measure of uncertainty
- Fuzzy sets: vagueness
- Approximate reasoning
- Linguistic evaluations



## Information Available

### Quantitative

### Quantitative, some Qualitative

### Qualitative

**Lightning strikes at Pantex**

**New strong link**

**Abnormal environments  
Terrorist attacks**

**Objective  
Aleatory  
Uncertainty  
Classical  
Probability  
Classical  
Statistics**

**Bayesian  
Concepts**

**Subjective  
Epistemic  
Uncertainty  
Belief/Plausibility  
Fuzzy Sets**

**Progression of Course Topics**

**Traditional Techniques**

**Advanced Techniques**



# Part I: Traditional Techniques

---



# Risk Measure for Safety

---

- Risk is a combination of the **Likelihood** of an undesirable event and the **Consequence** of that event
- **Product** Definition of Risk
  - Risk = Likelihood \* Consequence
  - High likelihood low consequence event can have same risk as low likelihood high consequence event

Accident Sequence	Likelihood (per year)	Consequence (equivalent \$ Loss)	Risk: Likelihood * Consequence (\$ per year)
A	3	100	300
B	1	400	400
C	$10^{-4}$	$10^7$	1000
D	$10^{-2}$	$10^4$	100
E	0.1	700	70
			Total: \$1870 per year



# Risk Measure for Safety

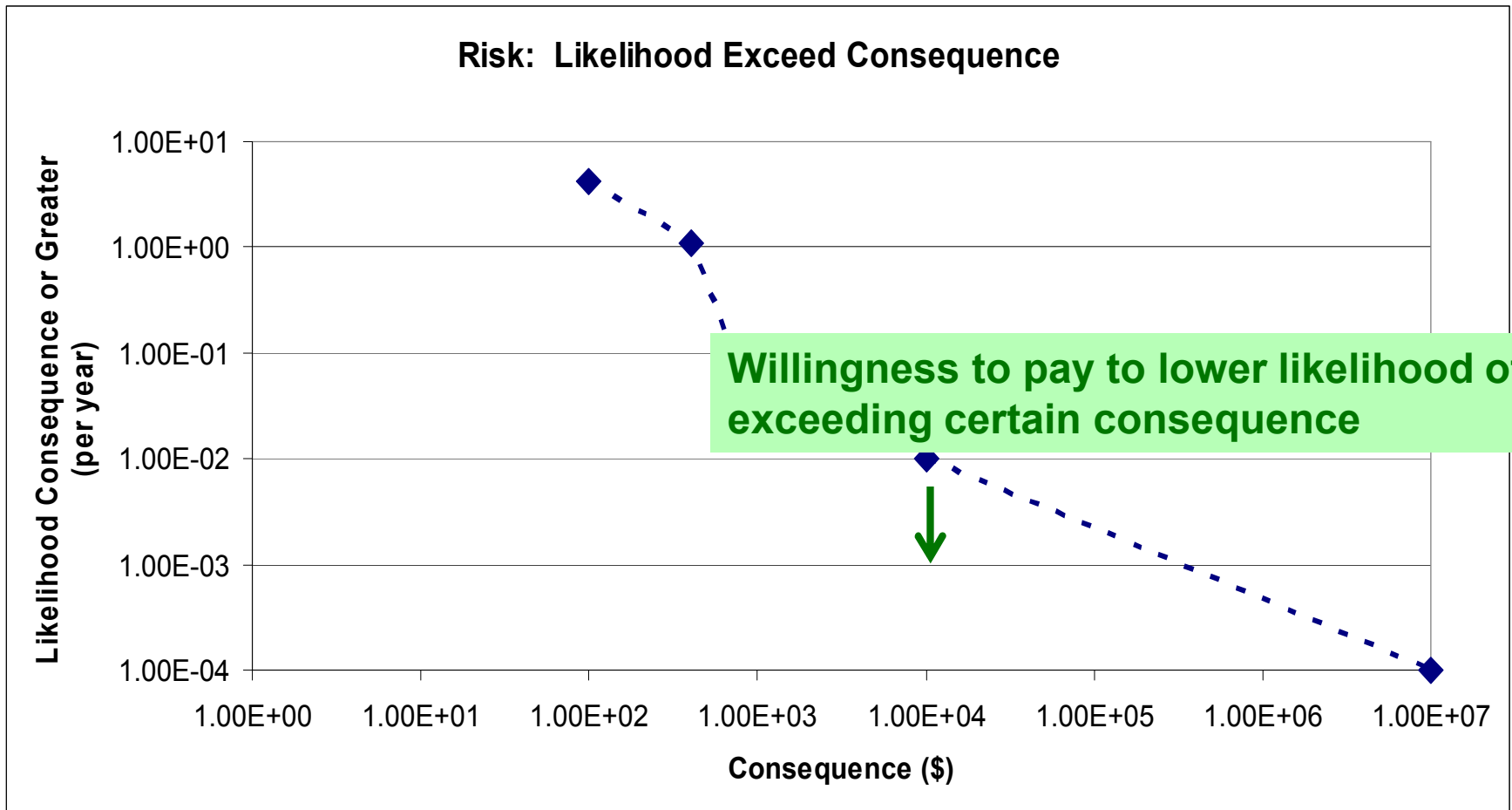
---

- Product definition of risk loses distinction between likelihood and consequence
  - Typically more concerned with higher consequence sequences
- Risk as **likelihood of exceeding consequence**
  - Kaplan and Garrick “On the Quantitative Definition of Risk”, Risk Analysis Vol. 1 No. 1, 1981

Accident Sequence ordered by increasing consequence	Likelihood (per year)	Consequence (equivalent \$ Loss)	Risk: Likelihood of Consequence or Greater (per year)
A	3	100	$3 + 1 + 0.1 + 10^{-2} + 10^{-4} = 4.1101$
B	1	400	$1 + 0.1 + 10^{-2} + 10^{-4} = 1.1101$
E	0.1	700	$0.1 + 10^{-2} + 10^{-4} = 0.1101$
D	$10^{-2}$	$10^4$	$10^{-2} + 10^{-4} = 0.0101$
C	$10^{-4}$	$10^7$	0.0001



# Risk Measure for Safety





# Likelihood

---

- **Frequency or Probability?**
- Attacks against US soldiers deployed in Iraq after war ended
  - Frequency: 5000 per year
  - Probability of attack in 1 year about 1.0
  - Consequence: 3 deaths per 10 attacks: 0.3 deaths per attack
- Product measure for Risk with **likelihood as frequency**
  - Risk =  $5000 * 0.3 = 1500$  deaths per year
- Product measure for Risk with **likelihood as probability**
  - Risk =  $1.0 * 0.3 = 0.3$  deaths in one year
  - **WRONG!**
- Probability is “*one **or more** attacks per year*”
  - 1 attack has consequence 0.3 deaths
  - 2 attacks have consequence  $2 * 0.3 = 0.6$  deaths
  - 3 attacks have consequence  $3 * 0.3 = 0.9$  deaths
  - ...

# Likelihood

- Probability within time T for event with frequency f
  - Exponential distribution
  - $P(T) = 1 - e^{-fT}$
  - If  $fT$  small  $P(T) \approx 1 - (1 - fT) = fT$ 
    - For f in per year and T of 1 year  $P(1) = f$
    - f and P numerically equal, but not the same concept
- Earlier problem using Probability instead of Frequency
  - Erlangian distribution gives P of **exactly** n = 1, 2, 3... occurrences (consequences)
    - $P(n, T) = e^{-fT} (fT)^n / n!$
    - Risk =  $\sum_{n=0}^{\infty} e^{-fT} (fT)^n / n! * (nC)$ 
      - F = 5000 per year
      - T = 1 year
      - C = 0.3 deaths
    - Solution using Mathematica software
 

```
In[15]:= Risk = Sum[E^(-f T) (f T)^n / n! n C, {n, 0, Infinity}]
```

```
Out[15]= C f T
```

```
In[16]:= Risk /. {C -> 0.3, f -> 5000, T -> 1}
```

```
Out[16]= 1500.
```
  - **1500 deaths in 1 year**



# Probability Measure of Uncertainty

---

- Frequency,  $F$ , and Consequence,  $C$ , for Risk,  $R$ , have uncertainty
  - $F$  and  $C$  are random variables
- Risk has Uncertainty
  - Risk is a function of the two random variables  $F$  and  $C$
- Digress to discuss Probability, then apply to provide Uncertainty for Risk



# Classical Probability is Objective

---

- **Classical probability is a specific value (one value)**
  - Event E, N identical trials
  - Probability of Event:  $P(E)$ 
    - $P(E) = \lim_{N \rightarrow \infty} (\text{number of time E occurs} / N)$
    - $P(E)$  is **fixed but perhaps unknown** with certainty
    - To know  $P(E)$  precisely requires infinite number of identical trials
  - **Classical probability is an Objective concept**
    - Probability is a Frequency
      - Not a physical rate but a dimensionless ratio
- For now assume we know the probability
  - As introductory courses on probability assume
- (If do not know probability can infer it from a sample using statistics; discussed later)



# Probability Concepts

---

- Sample space is **set** of **all unique** outcomes
  - Set: No repeats, order does not matter
  - Toss a Die
  - Valid Sample Space
    - {Even, Odd}
      - All outcomes and each outcome is unique
  - Invalid Sample Space
    - {4 or Greater}
      - Does not have all outcomes: 1, 2, and 3 not included
    - {4 or Greater, 5 or less}
      - Outcomes not unique: 4 and 5 are in both outcomes
- Failure to understand that **outcomes must be unique** has led to many incorrect analyses



# Probability Concepts

---

- Basic courses in probability focus on Random

**Many standard probability distributions**

**For a random variable:**

- **Binomial**
- **Normal**
- **Exponential**
- **Beta**
- **...**

...

– Outcomes are **mutually exclusive**



# Probability Concepts

---

- **Event is a subset of the sample space**
  - Random variable for die value  $\{1, 2, 3, 4, 5, 6\}$
  - Define Event A as “greater than 3” =  $\{4, 5, 6\}$
  - Define Event B as “less than 5”  $\{1, 2, 3, 4\}$
  - Events are **NOT** mutually exclusive
    - A and B share outcome 4
  - Outcomes **ARE** mutually exclusive
    - 1, 2, 3, 4, 5, 6
    - Outcomes sometimes called “elementary events”





# Probability Concepts

---

- **More complicated Sample Space**
  - Toss a coin twice, each toss is Heads, H, or Tails, T
  - Sample Space =  $\{<H, T>, <T, H>, <H, H>, <T, T>\}$
  - Sample space is a set  $\{\}$  of **tuples**  $<>$ 
    - Set: no repeats, order does not matter
    - Tuple: repeats allowed, order does matter



---

***If you cannot describe the sample space,  
you do not understand the problem.***



# Probability Concepts

---

**Probability Measure,  $P$ , for Sample Space,  $S$ :**  
**Kolmogorov axioms**  
**(the mathematics for probability)**

1. For any event  $E$ ,  $0 \leq P(E) \leq 1$
2.  $P(S) = 1$
3. For any set of **mutually exclusive** events  $\{E_1, E_2, E_3, \dots, E_n\}$  the Probability of the union (or) of all the events is the **sum** of the probabilities of each event

$$P(E_1 \text{ or } E_2 \text{ or } E_3 \text{ or } \dots \text{ or } E_n) = P(E_1) + P(E_2) + P(E_3) + \dots + P(E_n)$$



# Probability Concepts

---

- Outcomes are mutually exclusive so probabilities of outcomes add (third Kolmogorov axiom)
  - $P(S) = \text{Sum of } P(\text{all outcomes for } S) = 1.0$
- For any two events A and B
  - $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
  - If A and B are **mutually exclusive**  $P(A \text{ and } B) = 0$
  - If A and B are **independent**  $P(A \text{ and } B) = P(A) * P(B)$
- Mutually exclusive events are **NOT** independent events
- For any real-world problem
  - What is the sample space?
  - What events are of concern?
    - **Do not implicitly assume mutually exclusive or independent**



---

***Failure to consider dependence in  
real-world problems will  
under-estimate risk.***



# Uncertainty for Risk

---

- Risk a function of Likelihood and Consequence
- Likelihood as Frequency,  $F$ , a random variable
- Consequence,  $C$ , a random variable
- Risk as a function of the random vector  $F \times C$ 
  - Backup material discusses random vector



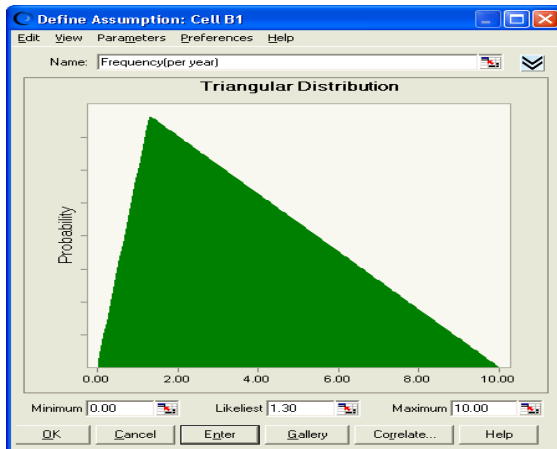
# Probability Concepts

---

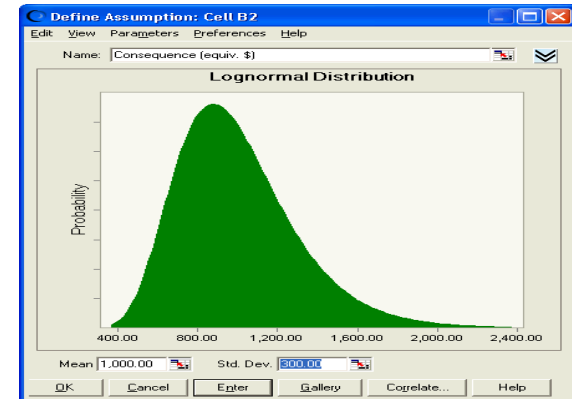
- **Probability Distributions**
  - **R** a continuous or discrete random **variable**
  - **r** a **specific value** of R
- **Three Probability Distributions for a random variable R**
  - **Cumulative Distribution Function (CDF)**
    - $\text{CDF}(r) = \text{Probability}(R \leq r)$
  - **Complementary Cumulative Distribution Function (CCDF)**
    - $\text{CCDF}(r) = \text{Probability}(R > r) = 1 - \text{CDF}(R)$
  - **Probability Density Function (PDF)**
    - $\text{Probability}(r \text{ in } [a,b]) = \int_{a \text{ to } b} \text{PDF}(r) \, dr$   
for continuous R
    - $\text{Probability}(r) = \text{PDF}(r)$  for discrete R

# Risk as a Product: $R = F * C$ with Uncertainty

- Using **Crystal Ball software (Oracle)**
  - An overlay on Excel that treats Excel variables as random variables with probability distributions
  - Convolution by sampling
    - Monte Carlo or Latin Hypercube
- Assume F (per year) is triangular over [0, 10]

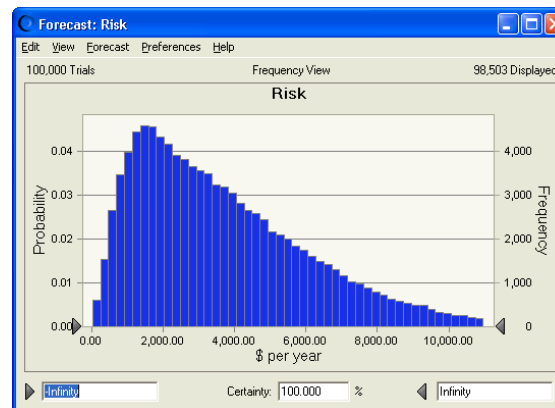


- Assume C (equiv. \$) is lognormal

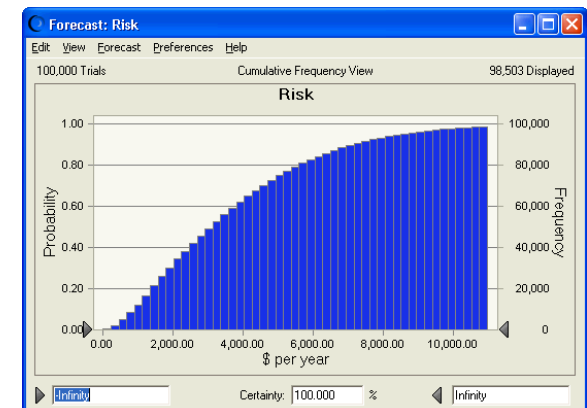


- Risk (\$ per year) = F \* C per convolution

PDF for Risk

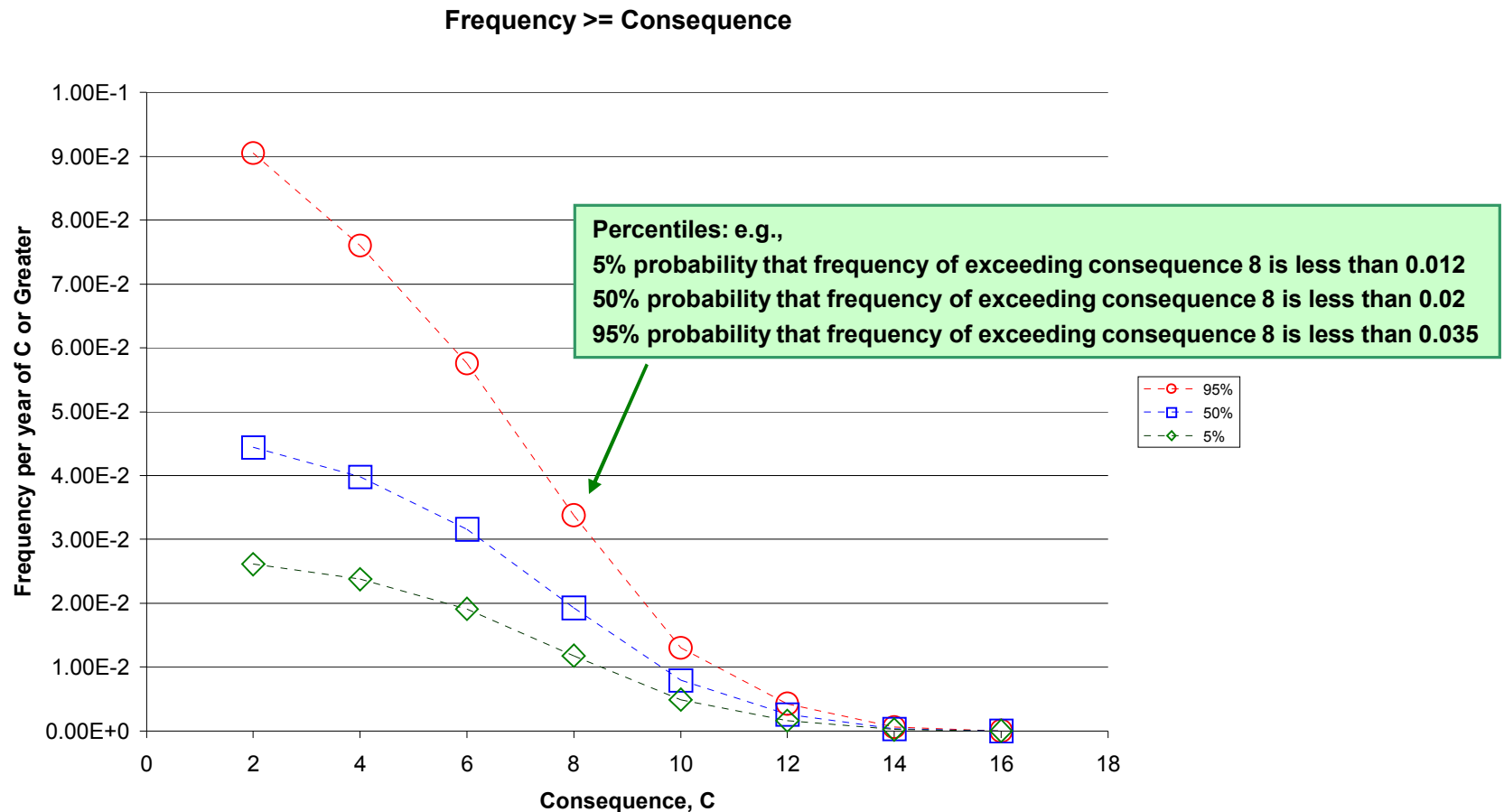


CDF for Risk





# Risk as Likelihood of Exceedance with Uncertainty





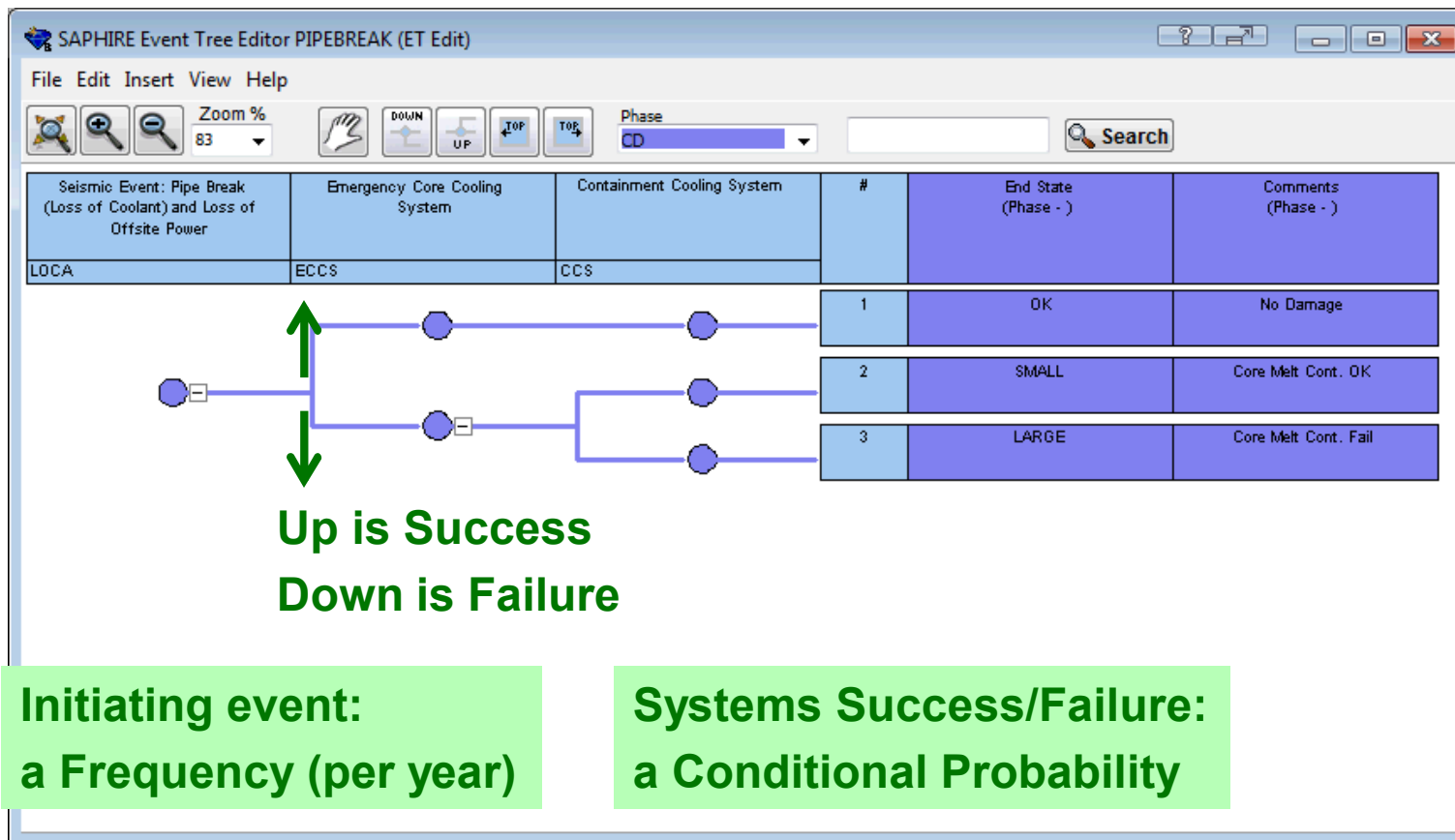
# Probabilistic Risk Analysis (PRA)

---

- Evaluate risk of complex system of systems using probability measure
  - Event trees for systems interaction logic: sequences
  - Fault trees for system failure logic
  - Link fault trees for accident sequences of concern
    - Shared components among systems
  - Consider uncertainty using probability
  - Main application to date: commercial nuclear power plants
- Risk of one sequence more complicated than combining F and C
  - F (initiating event) \*  $\prod P_{\text{conditional failure mitigating systems}}$   
**Combined with many different Consequences**
- Many Sequences, Many Systems, Many Components per System
- Software required for real world complex applications
  - **SAPHIRE software (written by INEL for NRC)**

# PRA Event Tree Example

- Seismic Event at Commercial Nuclear Power Plant
  - Pipe break: cool core and cool containment
  - Loss off offsite power: need onsite emergency electrical power
    - Core and containment cooling systems share emergency electrical power

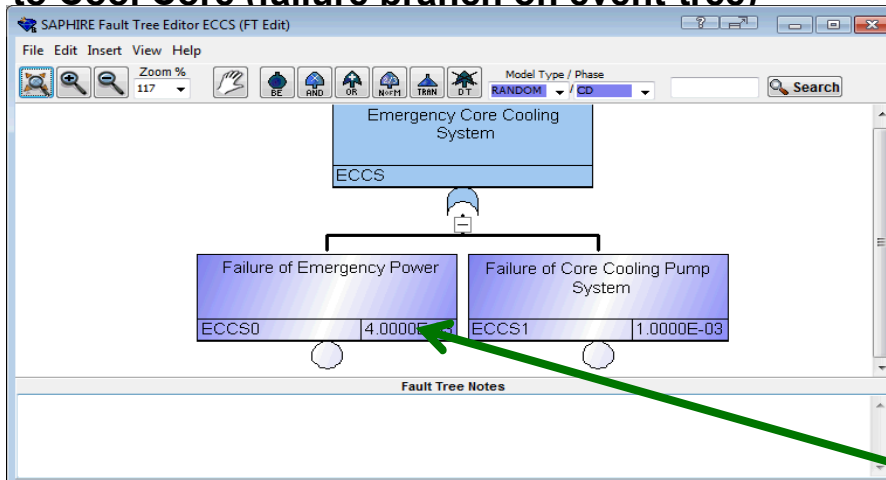


Two Sequences  
of Concern:

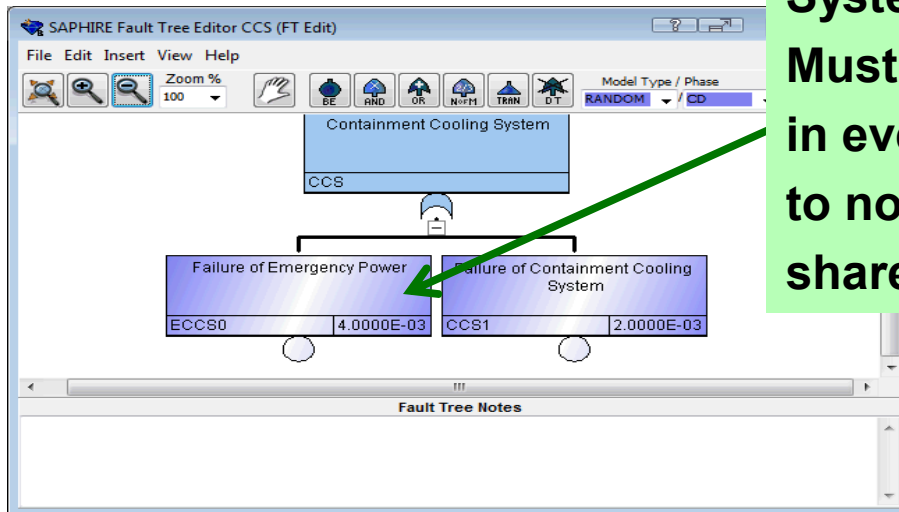
- (1) Fail Cool Core:  
Small Release
- (2) Fail Cool Core  
and Fail Cool  
Containment:  
Large Release

# PRA Fault Tree Example

- Fail to Cool Core (failure branch on event tree)



- Fail to Cool Containment (failure branch on event tree)



**Systems share power.  
Must link fault trees  
in event tree sequences  
to not double-count failures in  
shared components!**

# Cut Sets for Event Tree

HTML Viewer

**Detailed Cut Sets (by Sequence)**  
9/23/2014 6:56:03 PM

Note: Cut sets that contribute  $\geq 0.01\%$  are reported

SEQUENCE/CS#	PROB./FREQ.	TOTAL %	BASIC EVENT	DESCRIPTION	PROBABILITY
PIPEBREAK - 2					
Displaying 1 of 1 cut sets					
1	1.000E-7	100%	LOCA	Seismic Event: Pipe Break (Loss of Coolant) and Loss of Offsite Power	1.000E-4
			ECCS1	Failure of Core Cooling Pump System	1.000E-3
PIPEBREAK - 3					
Displaying 2 of 2 cut sets					
1	4.002E-7	99.95%	LOCA	Seismic Event: Pipe Break (Loss of Coolant) and Loss of Offsite Power	1.000E-4
			ECCS0	Failure of Emergency Power	4.000E-3
2	2.000E-10	0.05%	LOCA	Seismic Event: Pipe Break (Loss of Coolant) and Loss of Offsite Power	1.000E-4
			CCS1	Failure of Containment Cooling System	2.000E-3
				Core Cooling Pump System	1.000E-3

Mode Software Version: Sapphire 8.1.0

Print... Save As... Close

**Failure of emergency power fails both core and containment cooling: large release**



# Classical Statistical Inference

---

- **Probability: predict result of a sample based on knowledge about the population**
  - Probability that pump fails to start on demand is “known” to be a binomial probability distribution with probability of failure of 0.01
    - Probability exactly  $x$  failures in  $n$  trials is:  $n!/(x! * (n - x)!) p^x (1 - p)^{n-x}$
  - Probability a specific pump fails to start on demand is 0.01
    - $n$  and  $x$  are 1, and  $p$  is 0.01
- **Statistics: characterize the population based on taking a sample**
  - Probability that pump fails to start on demand is assumed to be a binomial probability distribution but probability of failure is not known
  - Take a sample of pumps, estimate probability of failure for the population of all pumps
- **Population has parameters**
- **Sample has statistics used to estimate parameters**
- **Inference mean infer parameters of population from statistics of sample**
- **Triola, Elementary Statistics is an excellent introductory text with a lot of interesting applications**
  - Prussian soldiers killed from being kicked by horses
  - Voltaire and friends became rich buying all lottery tickets: cost of tickets less than value of prize!



# Classical Statistical Inference

---

- Typically we do not know the parameters of a probability distribution
- Infer parameters from statistics
  - Take a **sample**
- Example: Probability of failure of a component is assumed to be described by the binomial distribution
  - $p$ , probability of failure, is a parameter of the binomial distribution
  - What is  $p$ ?
    - $p$  is fixed but unknown
    - Infer  $p$  from sample



# Classical Statistical Inference

---

- From population of  $N$ , sample  $n$  and observe  $x$  failures
  - **Sample with replacement** (binomial distribution)
  - **Sample without replacement** (hypergeometric distribution)
    - If  $N$  is large (greater than about 10 times  $n$ ) hypergeometric is well approximated by binomial
    - Sample without replacement more efficient for small population
  - From the sample we can establish a **confidence interval for the parameter  $p$  of the binomial distribution**  
**Confidence interval represents the Uncertainty in  $p$  due to finite size of sample**



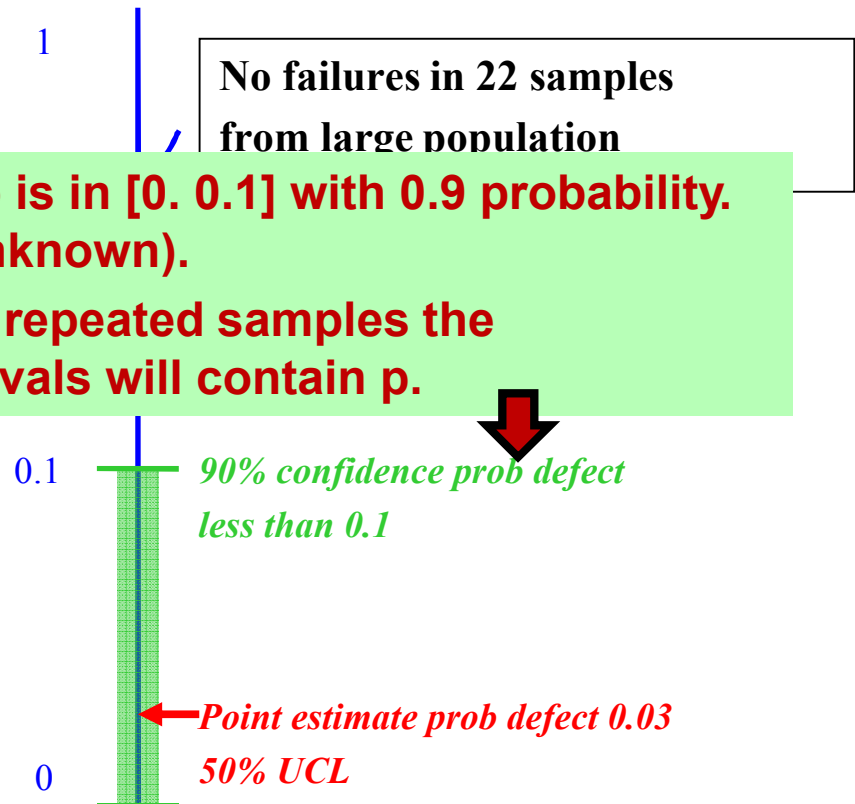
# Classical Statistical Inference

- **Weapon Component Failures:  
SNL Point Estimates**

- Sample  $n$  of  $N$  and observe  $x$  failures

- **This does NOT mean that  $p$  is in  $[0, 0.1]$  with 0.9 probability.  $p$  is a specific value (but unknown).**
- **This means that for 90% of repeated samples the calculated confidence intervals will contain  $p$ .**

- $x = 0$  from 2 samples has same estimate as  $x = 0$  from  $10^6$  samples
- If  $x$  is 0 use 50% UCL for point estimate to consider that larger  $n$  provides less uncertainty





# Classical Statistical Inference

---

- Surveillance of Stockpile for Estimate of Warhead Reliability
- Sample without replacement:  $n$  of  $N$ 
  - Hypergeometric distribution

**Number of Samples  
for 90% confidence  
for 90% reliability  
No failures in Sample**

Population	Sample Size
10	9
20	14
30	16
40	17
50	18
70	19
100	20
200	21
300+	22

**For  $n$  small and  $N$  large,  
Hypergeometric dist.  
has same result as  
binomial dist.**

**Number of Samples  
for 90% confidence  
for 95% reliability  
No failures in Sample**

Population	Sample Size
10	9
20	18
30	21
50	27
80	35
100	37
200	41
500	43
1500+	45



## Part II: Advanced Techniques

---



# Bayesian Concepts

---

- **Conditional probability**
  - Probability of event A *given* event B:  $P(A|B)$
  - $P(A|B) = P(A \text{ and } B) / P(B)$
  - $P(B|A) = P(B \text{ and } A) / P(A)$
- **Since  $P(A \text{ and } B) = P(B \text{ and } A)$** 
  - $P(A|B) = P(B|A) * P(A) / P(B)$



# Bayesian Concepts

---

- **Bayes theorem: discrete case**
  - **S a sample space**
  - **$\{A_1, A_2, A_3, \dots, A_n\}$  a partition over S**
    - **The A's are mutually exclusive and their union is S**
  - **B any event in S**
  - **Law of total probability**
    - **$P(B) = \sum_{k=1 \text{ to } n} P(B|A_k) * P(A_k)$**
- **$P(A_i|B) = P(B|A_i) * P(A_i) / \sum_{k=1 \text{ to } n} P(B|A_k) * P(A_k)$**



# Bayesian Concepts

---

- Example
  - Test for disease is 99% accurate given you have the disease
  - Test has  $10^{-4}$  false positive (falsely says you have disease)
  - 1 in  $10^6$  people have the disease
  - You test positive
    - Probability you have the disease is 0.99?
    - NO
  - $P(T|D) = 0.99$  is probability Test T is correct given you have the Disease D
  - $P(T|ND) = 10^{-4}$  is probability Test T is false positive given you do not have the disease ND
  - $P(D) = 10^{-6}$  is probability an individual selected at random has the disease
  - $P(D|T)$  is the probability you have the disease given you test positive
  - $P(D|T) = P(T|D) * P(D) / \{P(T|D)*P(D) + P(T|ND)*P(ND)\} =$   
 $0.99 * 10^{-6} / \{0.99 * 10^{-6} + 10^{-4} * (1 - 10^{-6})\} \approx 10^{-6} / 10^{-4} = 0.01$
  - Probability you have the disease given you test positive is 0.01, not 0.99



# Bayesian Concepts

---

**So far we have just used the properties of conditional probability.**

**The Bayesian approach is revolutionary in its *interpretation* of conditional probability.**



# Bayesian Concepts

---

- $P(A_i|B) = P(B|A_i) * P(A_i) / \sum_{k=1 \text{ to } n} P(B|A_k) * P(A_k)$
- Let  $P(A_i)$  be our initial probability distribution for event  $A_i$ 
  - $P(A_i)$  is our **prior** probability distribution for  $A_i$  before updating with information
- Let  $B$  be new information
- $P(A_i|B)$  is our updated probability distribution for  $A_i$  given the new information  $B$ 
  - $P(A_i|B)$  is our **posterior** probability distribution for  $A_i$  after updating with information  $B$
- Technique to update given new information
- We discussed discrete case, can also address continuous case

**Probability is SUBJECTIVE based on your state of knowledge. Totally different from classical, objective concept of probability.**





# Bayesian Concepts

---

- Bayesian Inference
  - Treat probability as a **random variable**
    - Different from classical statistical inference
      - Probability treated as **fixed (but unknown)**
- Reference: **Martz and Waller Bayesian Reliability Analysis**
- Example: Binomial distribution with unknown failure probability  $p$
- Assume  $p$  described by a beta distribution

**Objective probability is a fixed (typically unknown) value.**

**Subjective Probability is a Random Variable. NOT fixed.**

**distribution**

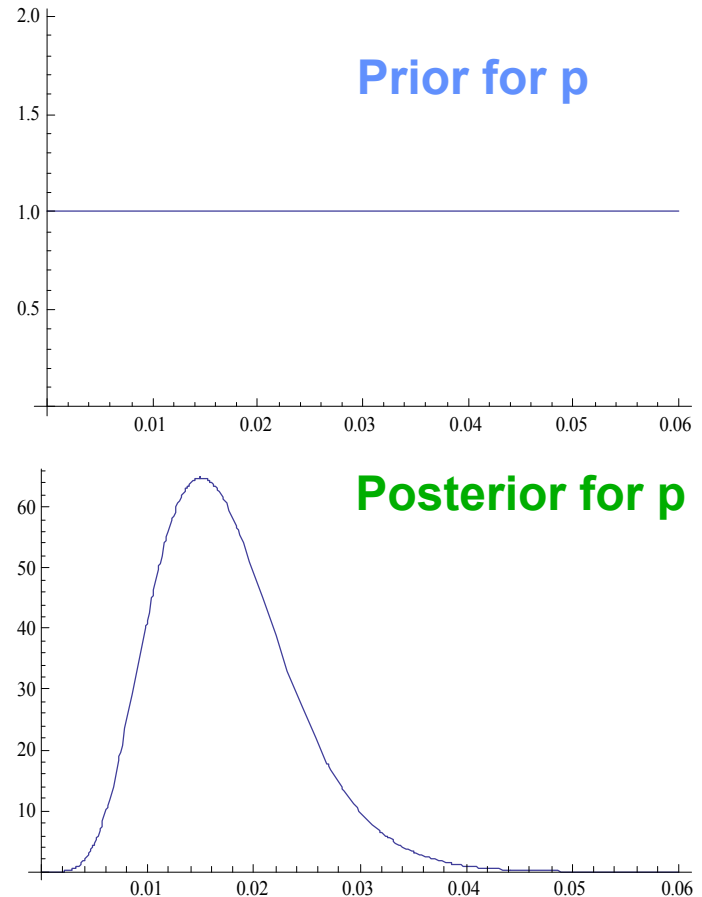
- This means the updated (posterior) distribution for  $p$  will also be a beta distribution
- Caution: two different sets of parameters used in literature for beta distribution, may need to transform variables
  - Beta[ $x, n$ ] and Beta[ $\alpha, \beta$ ]
    - $x$  failures in  $n$  trials (discrete)
    - Shape parameters  $\alpha$  and  $\beta$  (continuous)
    - Transformation:  $\alpha = x$  and  $\beta = n - x$

# Bayesian Concepts

- Using Beta[ $\alpha$ ,  $\beta$ ] convention
  - Beta[ $x_0$ ,  $n_0 - x_0$ ] is the prior given  $x_0$  failures in  $n_0$  trials
  - Beta[ $x + x_0$ ,  $n - x + n_0 - x_0$ ] is the posterior given new information  $x$  failures in  $n$  trials
- Assume prior distribution for  $p$  is uniform
  - No information as to what  $p$  really is
  - $p$  equally likely to be any value in  $[0, 1]$
  - Beta[1, 1] is the uniform prior for  $p$ 
    - $x_0$  is 1 and  $n_0$  is 2
- New information
  - $x$  failures in  $n$  trials
- Beta[ $x + 1$ ,  $n - x + 2 - 1$ ] is the posterior (updated) distribution for  $p$
- Using Mathematica

```
PosteriorDistribution[x0_, n0_, x_, n_] =  
BetaDistribution[x + x0, n - x + n0 - x0];
```

- with  $x = 6$  and  $n = 400$ , the posterior probability distribution (PDF) for  $p$  is

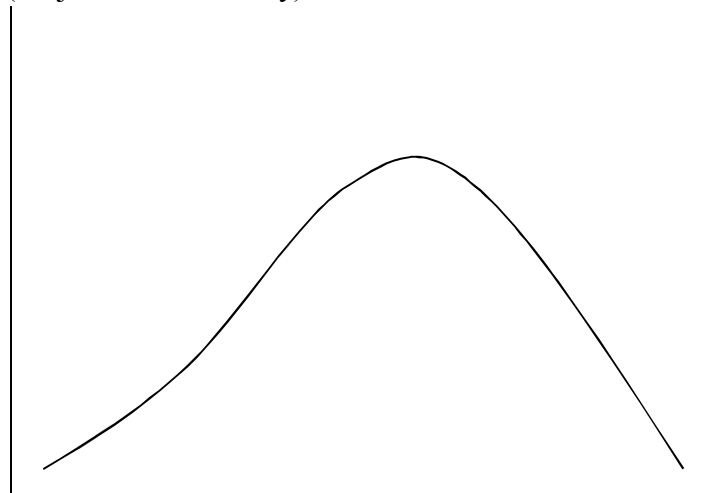


# Bayesian Approach is Subjective

- Probability is a state of knowledge and can be estimated even without sufficient data to evaluate the classical frequency
  - **Subjective** concept of probability
- Treat probability itself as a random variable instead of a fixed, but perhaps unknown, frequency
- Probability of a Probability means
  - Subjective probability of the objective probability (the frequency)
  - Confusion is that Probability used to mean **two different concepts**
    - Both concepts obey Kolmogorov axioms
- See earlier reference: Kaplan and Garrick 1981 paper in Risk Analysis
- Update  $P(E)$  with information:  $P(E | \text{Information})$  as discussed earlier

Probability (subjective Probability)

Probability of Probability means  
Subjective Probability (state of knowledge)  
of Objective Probability  
(classical frequency)



Frequency (objective Probability)

jldarby@sandia.gov



# Epistemic Uncertainty

---

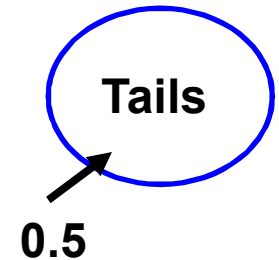
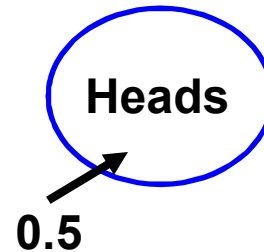
- Probability mostly deals with **aleatory (stochastic or random)** uncertainty
- Probability has difficulty dealing with **epistemic (state of knowledge)** uncertainty
- **Example of difference between aleatory and epistemic uncertainty.** Consider a fair coin, heads on one side, tails on the other, with each side equally likely. The uncertainty as to the outcome of a toss—heads or tails—is aleatory. The probability of heads is one half and the probability of tails is one half. The uncertainty is due to the randomness of the toss. Suppose, however, that we do not know the coin is fair; the coin could be biased to come up heads, or the coin could even be two-tailed. Now we have epistemic uncertainty; our state of knowledge is insufficient to assign a probability to heads or tails: all we can say is the likelihood of heads (or tails) is somewhere between 0 and 1. This is an example of “total ignorance”.



# Epistemic Uncertainty: Total Ignorance

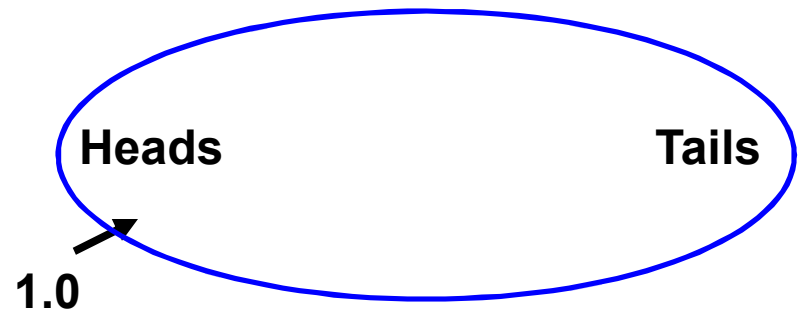
---

Probability Approach: **Assume**  
a probability distribution;  
typically assume uniform



With total ignorance we have *assumed*  
the same probability as if we knew the coin to be fair!  
We have **thrown away all the epistemic uncertainty!**

Subset (interval) Approach:  
The probability  
is somewhere in {Heads, Tails}





# **Belief/Plausibility Measure of Uncertainty**

---

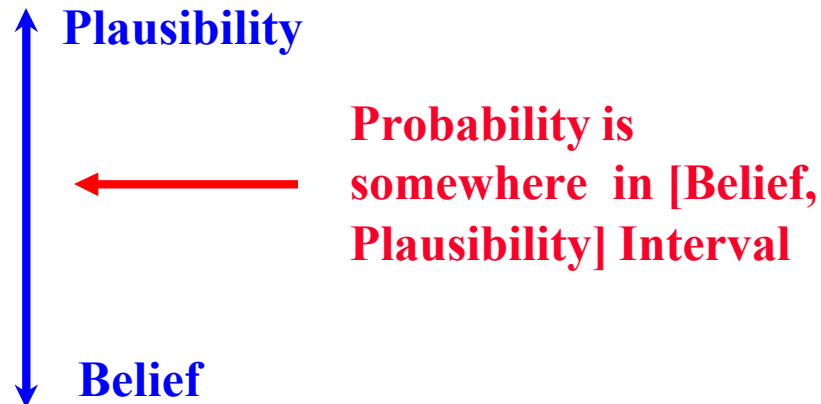
- **“Probability assigned to intervals”**
- **Belief is a lower bound for probability**
- **Plausibility is an upper bound for probability**
- **Also called Dempster/Shافر approach**



# Belief/Plausibility

---

- Belief / Plausibility form a Lower / Upper Bound for Probability



- Similar to a Confidence Interval for a Parameter of a probability distribution; a confidence measure that parameter is in interval, but exactly where in interval is not known
- Belief/Plausibility both reduce to Probability if Evidence is Specific
  - Subsets (intervals) with evidence are singletons (points)
- For coin that cannot be observed, Belief / Plausibility for both Heads and Tails is 0 / 1



# Belief / Plausibility

---

- **More general than Bayesian probability**
  - **Bayesian probability**
    - Assume prior probability distribution
    - Update with data to form posterior probability distribution
  - **Belief and Plausibility**
    - Do not know prior probability distribution
    - Little data for performing update
    - Total Ignorance easily addressed
- **Belief and plausibility are both probability if no epistemic uncertainty (evidence is specific)**
- **Useful for Formalizing Expert Judgment**





## Belief / Plausibility

---

- **m is Evidence assigned to a subset (or interval)**
  - Any subset with evidence is called a Focal Element
  - All evidence sums to 1.0
- **For B a Focal Element with evidence m(B)**
- **For A any Subset**

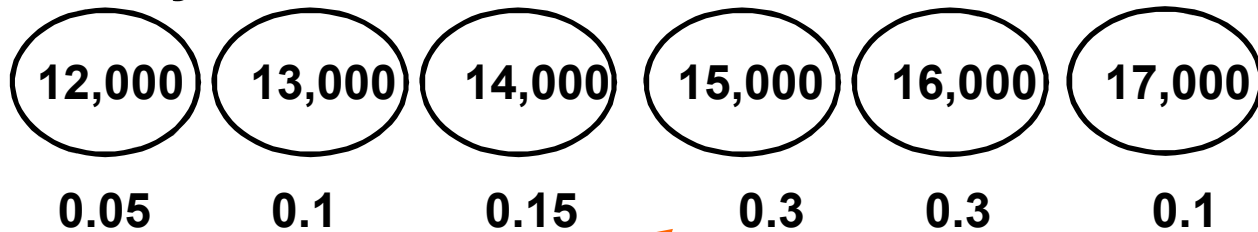
$$Bel(A) = \sum_{B|B \subseteq A} m(B)$$

$$Pl(A) = \sum_{B|A \cap B \neq \emptyset} m(B)$$

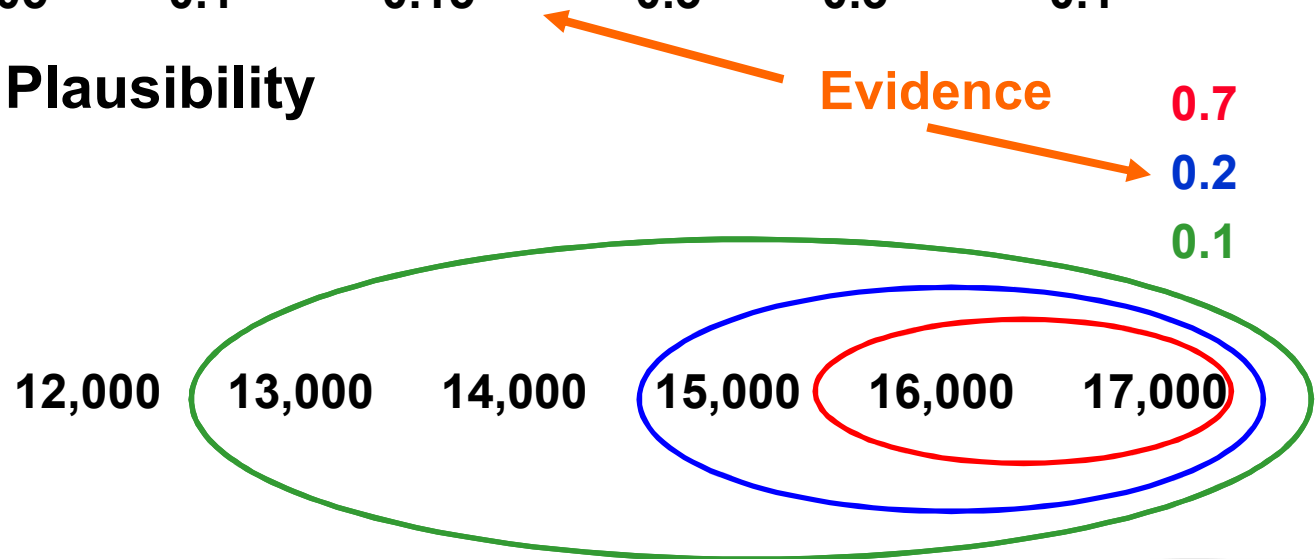
# Belief / Plausibility

- Example of **Evidence**: In January, 2014 Predict Stock Market Close Dec. 31, 2014

– Probability

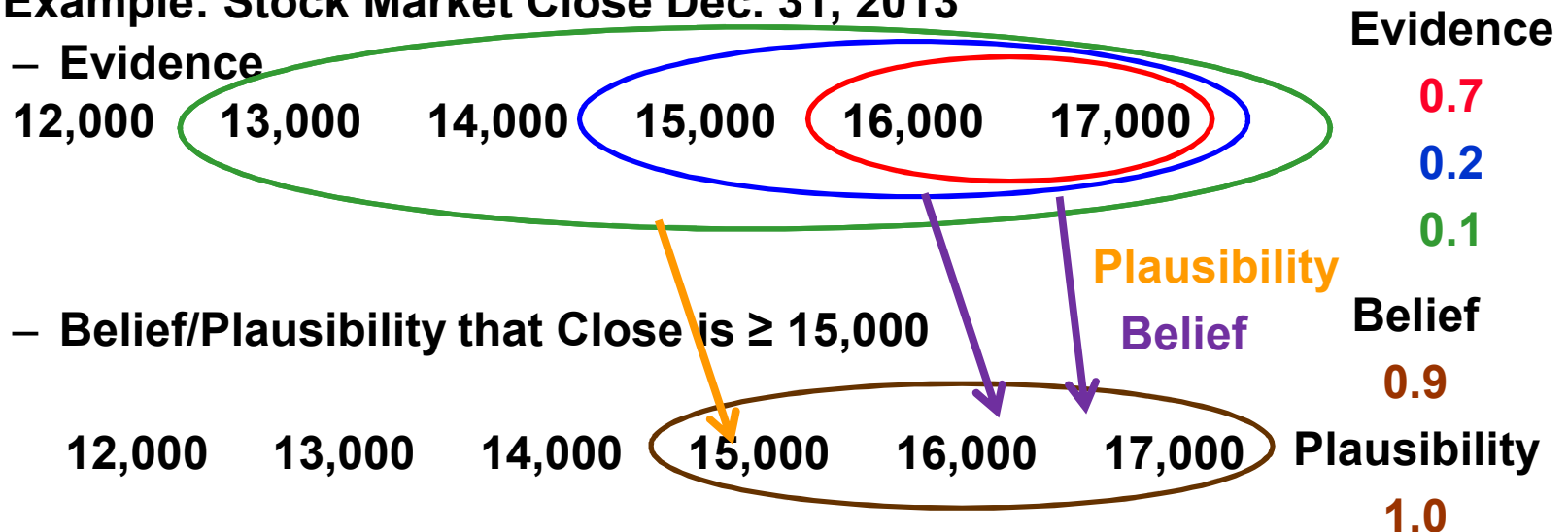


– Belief / Plausibility



# Graphical Interpretation

- Evidence is weighted information that outcome is equally likely to be **any outcome somewhere** in set of selected outcomes
- Belief is sum of all evidence contained **within** set of selected outcomes
- Plausibility is sum of all evidence that **overlaps** set of selected outcomes
- Example: Stock Market Close Dec. 31, 2013





# Belief / Plausibility

---

- **Sample space can be discrete or continuous**
  - **Earlier examples were discrete**
    - Evidence over subsets
  - **Can apply to intervals of reals**
    - Evidence over intervals



## Example of Evidence

---

- On Dec. 31, 2005, Grandma and Grandpa are trying to figure out the age of a distant relative, Jack
- Grandpa says “I think Jack was not born before 1980.”
- Grandma says “I think Jack is a teenager.”
- Jack has definite age, but there is uncertainty as to his age.
- Jack’s age is somewhere in  $[0, 150]$  years
- We have two pieces of **Evidence**: what Grandpa says and what Grandma says.



## Example of Evidence

---

- **Evidence #1: What Grandpa says. This is evidence for Jack's age being somewhere *exactly* in  $(0, 26)^*$**



- \* **Nomenclature.  $[a, b]$  contains all values between  $a$  and  $b$ , including  $a$  and  $b$ ;  $[a, b)$  contains all values between  $a$  and  $b$ , including  $a$  but not including  $b$ .**



## Example of Evidence

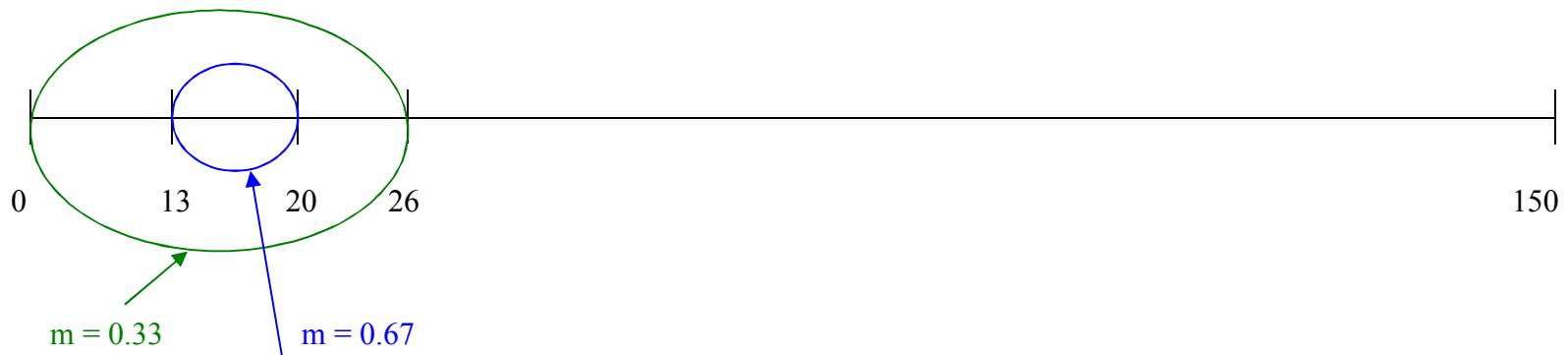
---

- **Evidence #2: What Grandma says. This is evidence for Jack's age being somewhere *exactly* in  $[13, 20)$**



# Example of Evidence

- We know that Grandma has a better memory than Grandpa about relatives, so we decide to weight Grandma's evidence twice as much as Grandpa's evidence:  $2/3$  for Grandma and  $1/3$  for Grandpa
- Our focal elements are as follows.  $m$  is a degree of evidence





# Example of Evidence

- For the interval **[13, 20)**, the belief that Jack's age is in  $[13, 20)$  is 0.67 and the plausibility that Jack's age is in  $[13, 20)$  is 1.0.
- For the Interval **(0, 26)**, the belief that Jack's age is in  $[0, 26)$  is 1.0 and the plausibility that Jack's age is in  $[0, 26)$  is 1.0
- For the interval  $[26, 150]$ , the belief is 0 and the plausibility is 0.
- Based on the evidence, Jack is not 26 years old or older; we are certain Jack's age is in  $(0, 26)$ . The probability that Jack is a teenager, age in  $[13, 20)$ , is somewhere in the belief/plausibility interval 0.67 to 1.0.





## Example of Evidence

---

- **A given expert can provide more than one piece of evidence**
  - **Grandma could have provided both pieces of evidence**
    - **My best recollection is that Jack is a teenager**
    - **I think I remember Aunt Maude telling me that Jack is not yet 26**
    - **Grandma assigns evidence 2/3 to teenager**
    - **Grandma assigns evidence 1/3 to not yet 26**



# Assigning Evidence is an Art

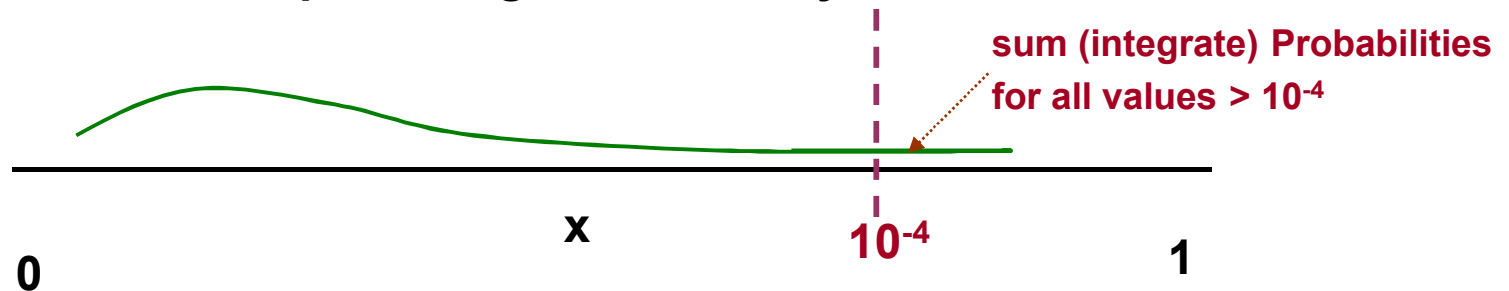
---

- **What is the Evidence?**
- **What Weight is given Each Piece of Evidence?**

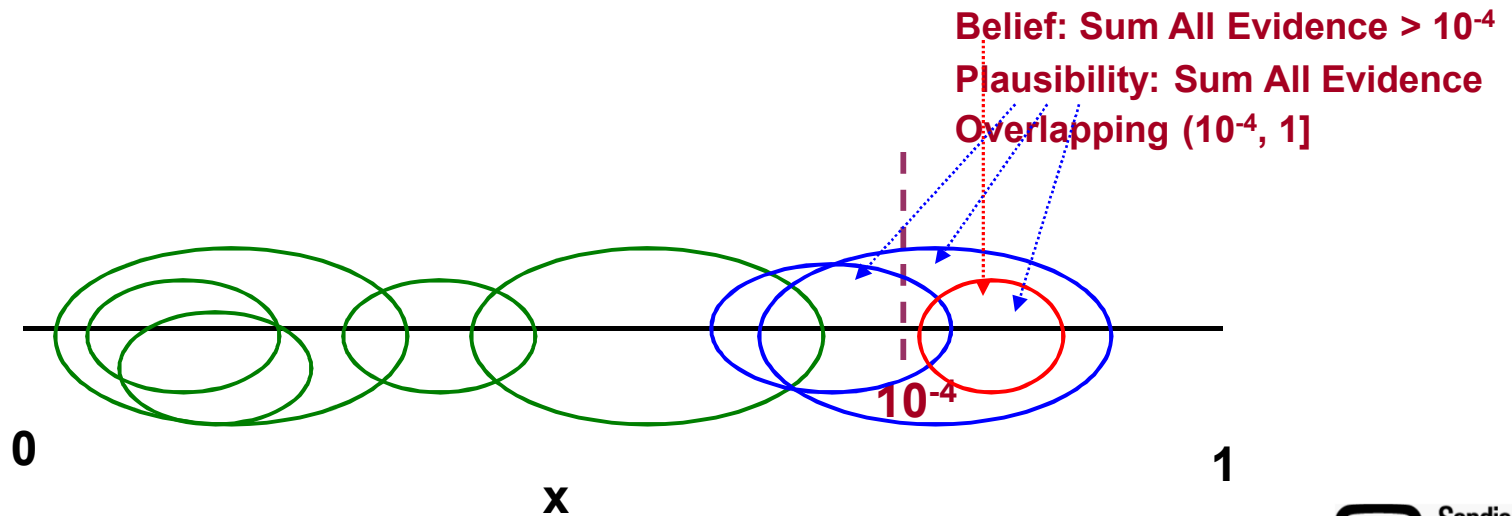
# Uncertainty for a Random Variable

**X a Random Variable in  $[0, 1]$ : Concern is X exceeds  $10^{-4}$**

- Subject Matter Expert assigns Probability Distribution To X



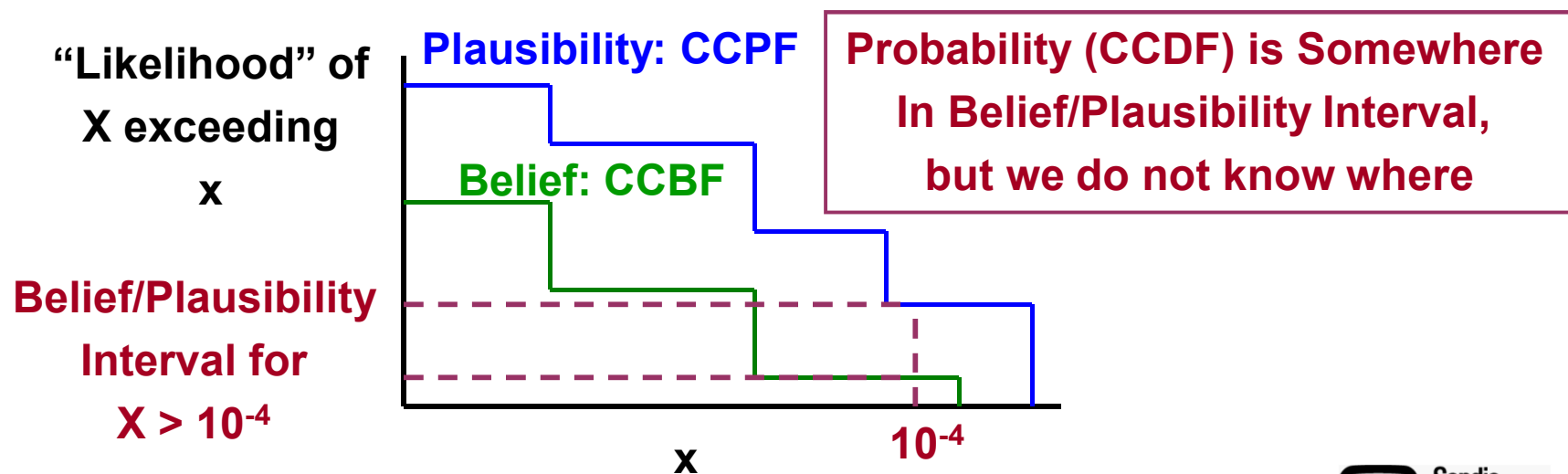
- Subject Matter Expert assigns Evidence to Intervals of X



# Belief/Plausibility Viewed as Lower/Upper Bound on Probability

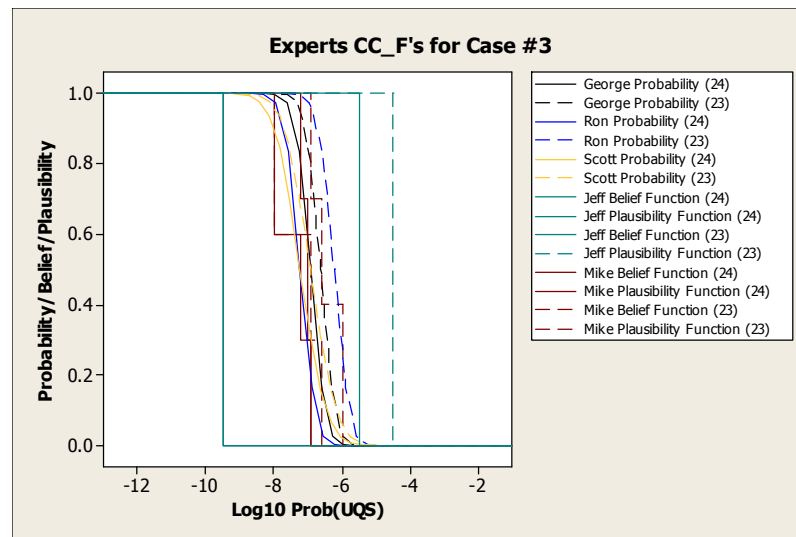
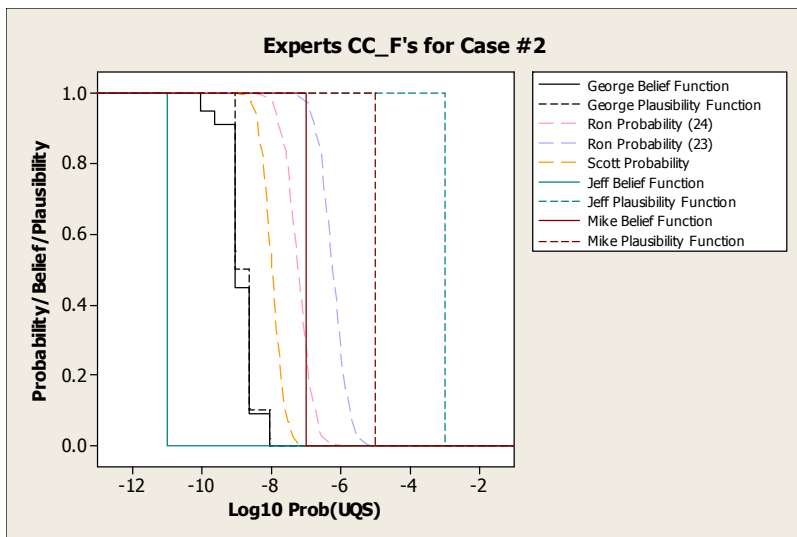
- Belief/plausibility distribution for  $X$  obtained from Expert assigning evidence to intervals in range for  $X$ . All evidence must sum to 1.0. Result presented as belief/plausibility of exceedance: Complementary Cumulative Belief Function (CCBF) and Complementary Cumulative Plausibility Function (CCPF).

x a specific value of Random Variable  $X$  with range  $[0, 1]$ .



# Belief / Plausibility

- Evaluation of nuclear weapon strong link UQS issue





# Belief / Plausibility References

---

- **Introductory References**
  - Klir and Yuan, Fuzzy Sets and Fuzzy Logic
  - Appendices in Darby “Evaluation of Risk from Acts of Terrorism: The Adversary/Defender Model Using Belief and Fuzzy Sets”, SAND2006-5777
- **Advanced References**
  - Shafer, A Mathematical Theory of Evidence, 1976, Princeton University Press
  - Helton, Jon et al “An exploration of alternative approaches to the representation of uncertainty in model predictions”, Reliability Engineering and System Safety, Vol. 85 Nos. 1 – 3, July –Sept, 2004
  - Helton, Jon “Conceptual and Computational basis for the Quantification of Margins and Uncertainty”, SAND2009-3055, June, 2009



# Belief / Plausibility Software

---

- **BeliefConvolution custom SNL Java code, J. Darby, SNL**
  - Also evaluates fuzzy numbers
- **RAMAS RiskCalc software, S. Ferson, Applied Biomathematics**
- **Go see Dr. Jon Helton (on-site consultant at SNL)**
  - Sampling techniques
  - Non-algebraic functions
    - e.g.,  $(a + b)^a$





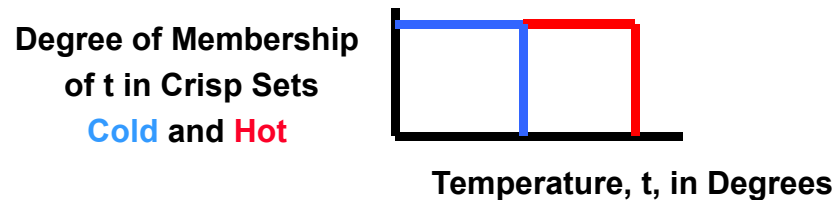
# Fuzzy Sets: Vagueness

---

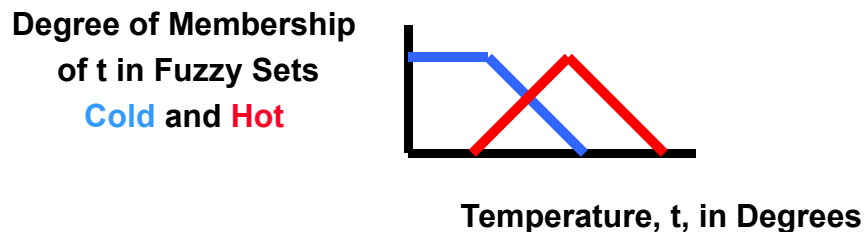
- So far our Uncertainty has been **Ambiguity**
  - Uncertainty as to what will occur **in the future**
    - Dow Jones Industrial Average Close on Dec. 31, 2009
      - Will be one value
      - Ambiguity as to what that value will be
- **Vagueness** is another type of Uncertainty
  - Uncertainty as how to categorize a **known** outcome
    - Dow Jones close is 9,876 on Dec. 31, 2009
      - Is this “High” ?
      - What do you mean by “High”?
  - Vagueness can be expressed with words: fuzzy sets

# What is a Fuzzy Set

- **Classical or crisp set**
  - Element is either totally within or not within a crisp set
    - Membership value either 0 or 1



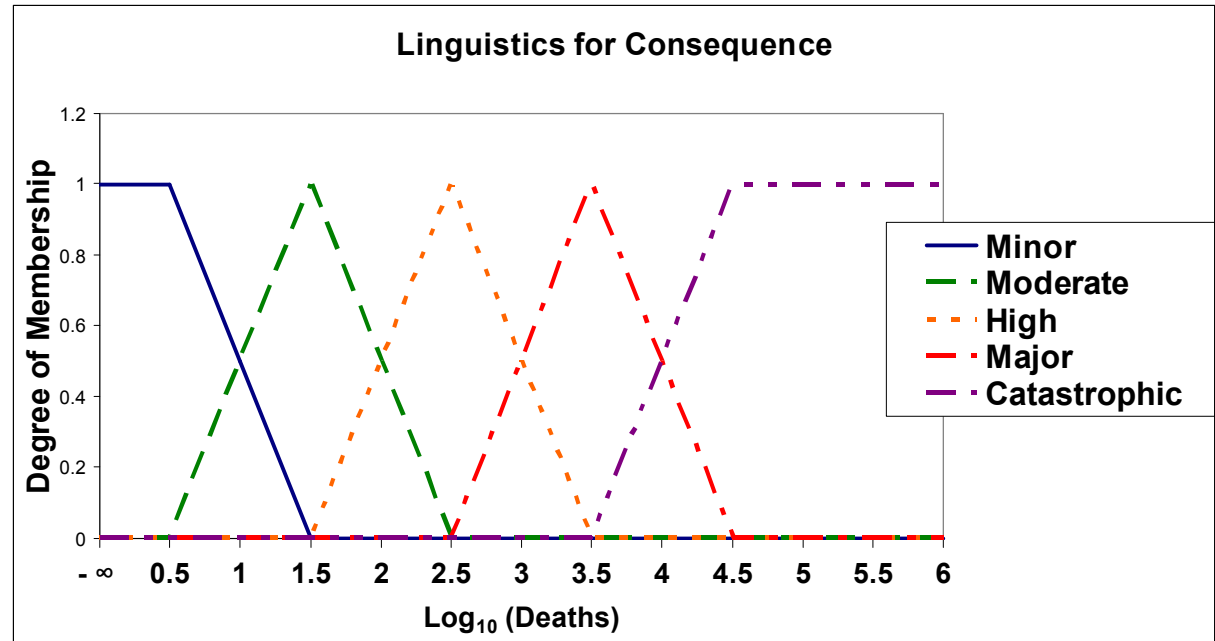
- **Fuzzy Set**
  - Element can be partially within more than one set
    - Membership value can be any value in  $[0, 1]$



# Vagueness

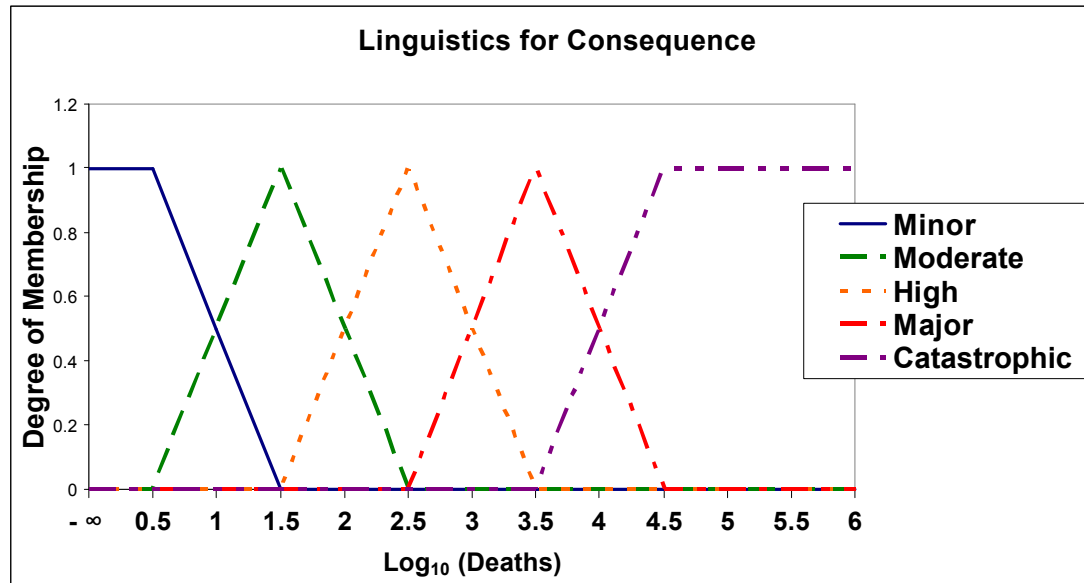
## Fuzzy Sets for Numeric Variable

- Represent Variable with **Sets** to **reason at Fidelity Desired**. Above 30,000 deaths is “Catastrophic”.
- Use **Fuzzy Sets** to **Avoid Sharp Distinction**. “Major” Deaths is Between *About* 1000 and *About* 10,000. 999 and 1001 deaths are each part “High” and part “Major”.

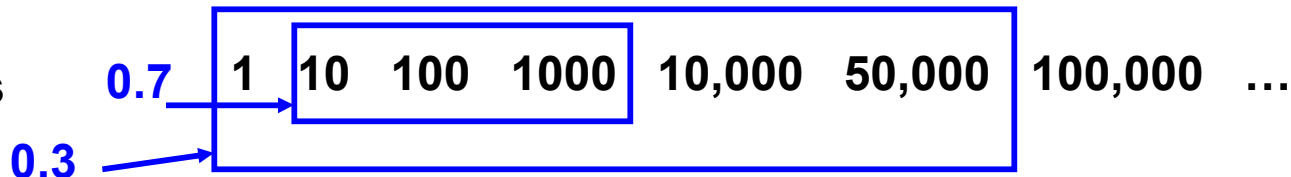


# Uncertainty for Fuzzy Sets: Numeric Variable

## Fuzzy Sets for Deaths



## Evidence For Deaths



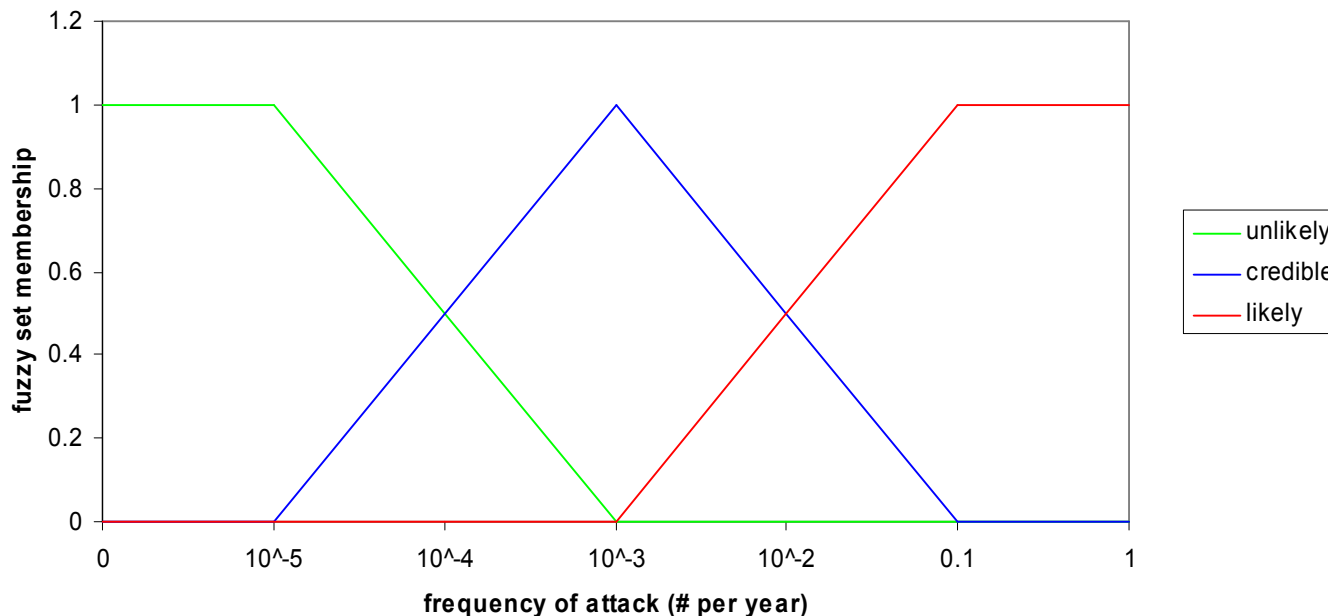
## Uncertainty Distribution

for Deaths:	Minor	Moderate	High	Major	Catastrophic
Belief / Plausibility	0 / 0.65	0 / 1	0 / 1	0 / 0.65	0 / 0.3

# Why is this Important?

- What is Likelihood of Bio-Terror attack against a Major US City
  - Evidence is about 1 major Attack every 5 years (0.2/year)
    - Assume Expert Opinion is: 10% Chance Attack is bio (0.02/year)
- Assume Following Fuzzy Sets for Evaluating Frequency of Attack

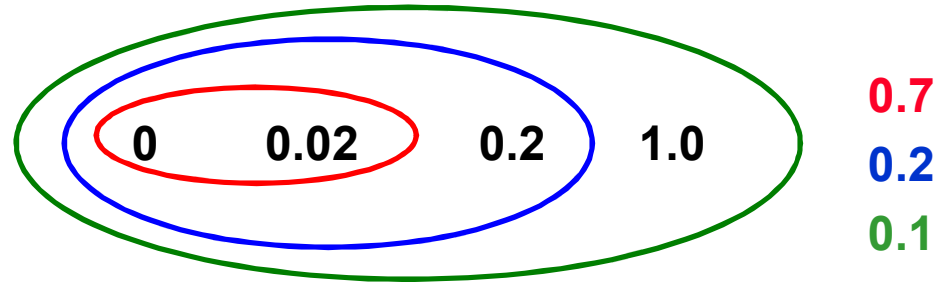
Defender Fuzzy Sets for Threat



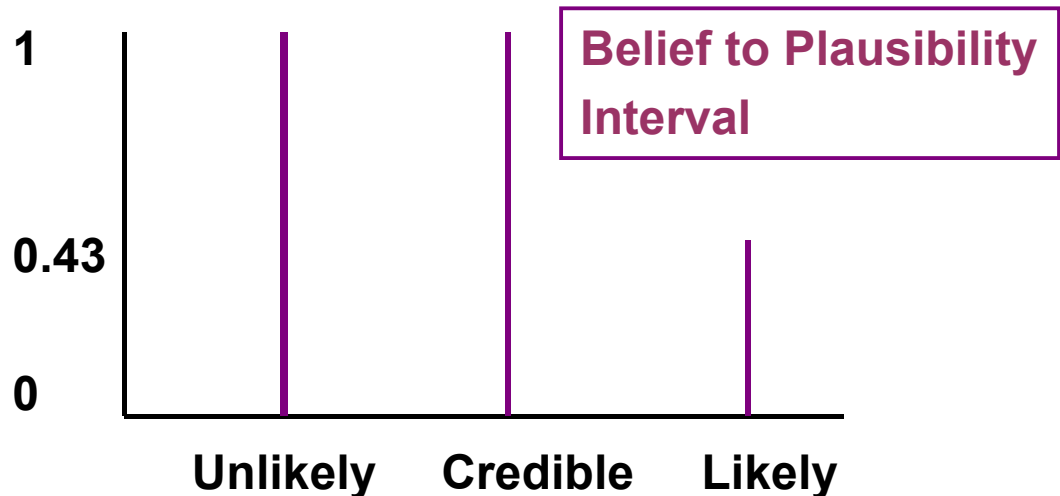


# Why is this Important?

Evidence



Likelihood  
For Frequency



*(Calculated with BeliefConvolution  
code using Yager Method)*



# Qualitative Variables

---

- Variable Segregated into **Purely Linguistic Fuzzy Sets**
  - Variable: “Health”
  - Fuzzy Sets: “Bad”, “Moderate”, “Excellent”
- Why Pure Linguistics?
  - **Numeric Scale is Unknown**
    - Is “Health”  $[0, 10]$ ,  $[0, 10^6]$ ,  $[-700, \text{square root of } 42]$ ?
  - **Scaling is Un-Manageable when Combine Variables**
    - Combine “Health” with “Wealth” to Evaluate “Quality of Life”
      - “Wealth” can be Numeric:  $[\$0, \$50\text{B}]$
      - What is Numeric Scale for “Health”?
      - What is Numeric Scale for “Quality of Life”?



# Fuzzy Sets for Non-Numeric Variable

---

**Adversary Level of Technical Training:**

**High School**

**Bachelors**

**Advanced**

Do **NOT** Force Numeric Measure: Requires Arbitrary Scale

**Adversary Level of Technical Training:**

**High School = 1?**

**Bachelors = 2?**

**Advanced = 3?**

**Adversary Level of Technical Training:**

**High School = 10?**

**Bachelors = 100?**

**Advanced = 1000?**





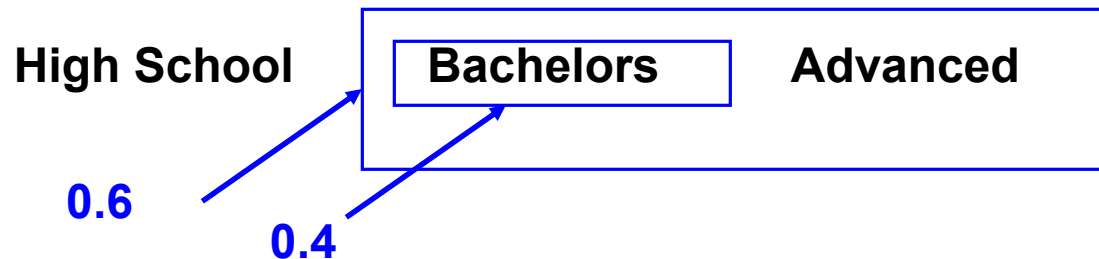
# Uncertainty for Fuzzy Sets: Non-Numeric Variable

---

- **Fuzzy Sets for Adversary Level of Technical Training**

– High School                      Bachelors                      Advanced

- **Evidence**



- **Uncertainty Distribution: Belief / Plausibility**

High School	Bachelors	Advanced
0 / 0	0.4 / 1	0 / 0.6



# Combining Variables: Convolution of Uncertainty Distributions

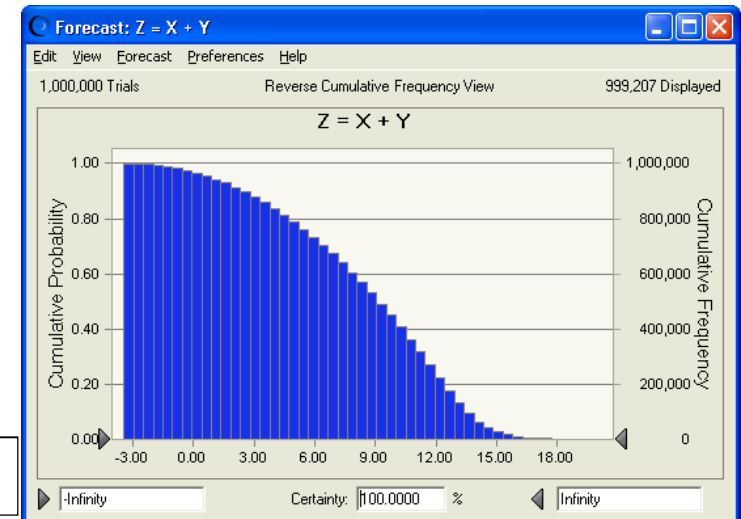
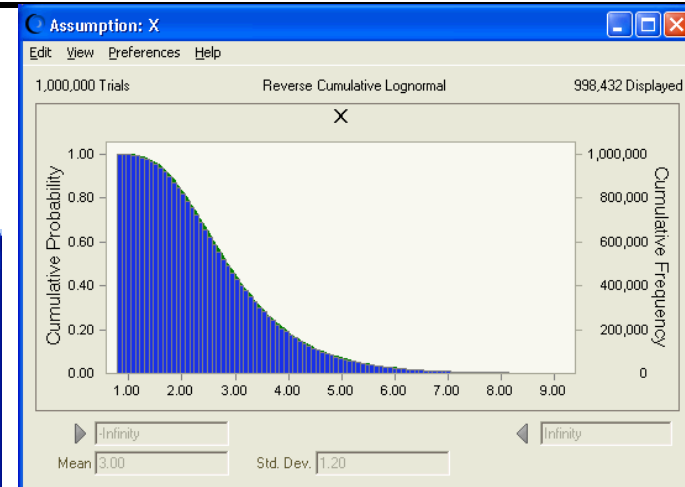
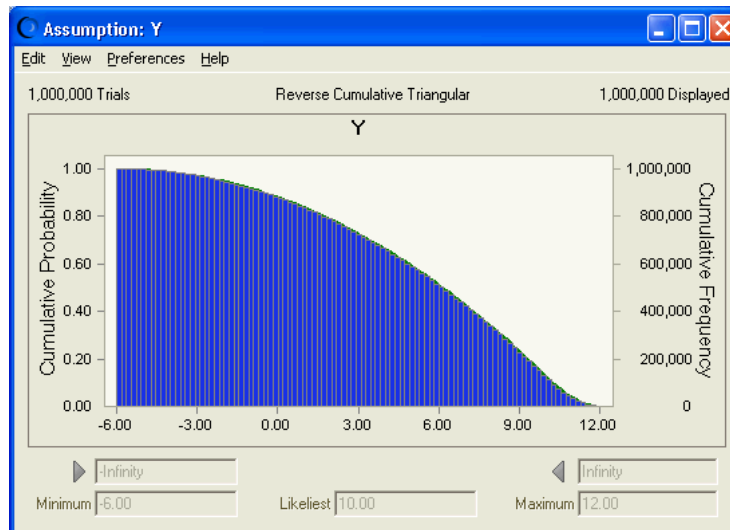
---

- **Belief/Plausibility Distributions**
  - Evidence Over Fuzzy Sets for Each Variable
- **Convolute Distributions** per the Rule Base
  - Mathematics of Belief/Plausibility
- **Same Concept as Convolution of Probability Distributions**
  - Mathematics of Probability

# Convolute Probability Distributions: Crystal Ball Software

$$Z = X + Y$$

**X and Y Independent**



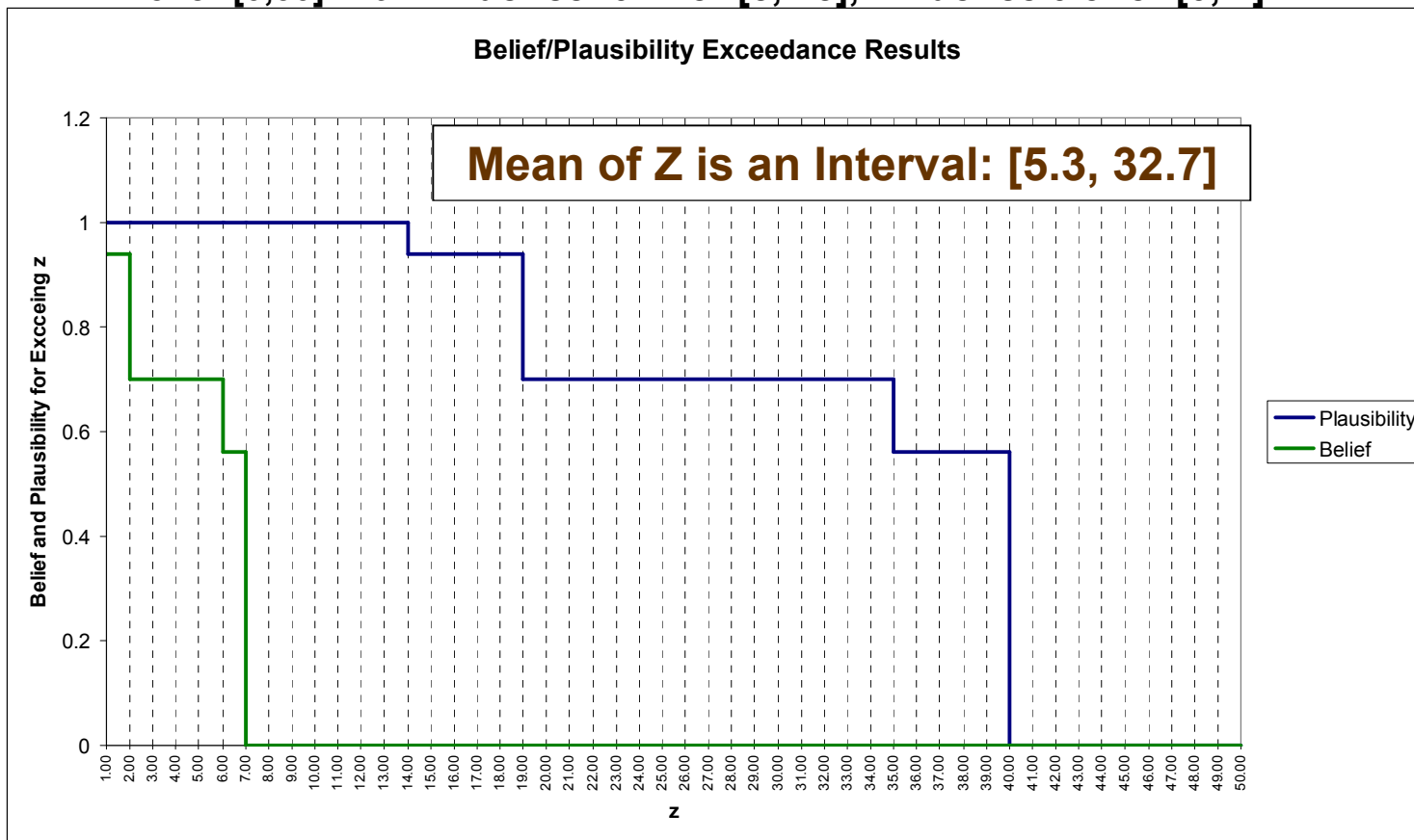
**Mean of Z is a Point value: 8.34**

# Convolute Belief/Plausibility Distributions for Numeric Variables: BeliefConvolution Software

$Z = X + Y$ ,  $X$  and  $Y$  Non-Interactive

$X$  over  $[1, 20]$  with Evidence: 0.8 for  $[2, 15]$ , 0.2 for  $[1, 10]$

$Y$  over  $[0, 30]$  with Evidence: 0.7 for  $[5, 25]$ , Evidence 0.3 for  $[0, 4]$





# Convolute Belief/Plausibility Distributions for Linguistic Variables: LinguisticBelief Software

---

- **Example Follows**



# Combining Qualitative Variables: Approximate Reasoning

---

- Mathematics for **Combining Words**
- If we use Words instead of numbers we need a way of combining the Words for Different Variables
- Implemented as A **Rule Base for Combining Fuzzy Sets from Different Variables**



# Combination of Linguistic Variables: Example

---

**Develop the Model: Happiness for Any Individual**  
**Define the Variables and their Fuzzy Sets**

- **Basic Variables**
  - **Health**
    - Bad, Moderate, Excellent
  - **Wealth**
    - Poor, Middle Class, Rich
  - **Outlook on Life**
    - Pessimist, Optimist
- **Rule Based Variables**
  - **Quality of Life = Health x Wealth (x per rule base)**
    - Not so Good, Good
  - **Happiness = Outlook on Life x Quality of Life**
    - Depressed, Accepting, Very Happy

# Combination of Linguistic Variables: Example

**Develop the Approximate Reasoning Rule Base for  
Rule Based Variables**

**Quality of Life**

Rules for selected RuleLinguistic

**Rules for RuleLinguistic: Quality of Life**

Fuzzy Set for Input Linguistic: Health	Fuzzy Set for Input Linguistic: Wealth	Output Fuzzy Set for Rule (blank if rule not set)
Bad	Poor	Not so Good
Bad	Middle Class	Not so Good
Bad	Rich	Not so Good
Moderate	Poor	Not so Good
Moderate	Middle Class	Not so Good
Moderate	Rich	Good
Excellent	Poor	Good
Excellent	Middle Class	Good
Excellent	Rich	Good

Specify Output Fuzzy Set for Selected Rule Choices Are: ▼

Accept Rules as Shown Cancel



# Combination of Linguistic Variables: Example

## Happiness

Rules for selected RuleLinguistic

**Rules for RuleLinguistic: Happiness**

Fuzzy Set for Input Linguistic: Outlook on Life	Fuzzy Set for Input Linguistic: Quality of Life	Output Fuzzy Set for Rule (blank if rule not set)
Pessimist	Good	Accepting
Pessimist	Not so Good	Depressed
Optimist	Good	Very Happy
Optimist	Not so Good	Accepting

Specify Output Fuzzy Set for Selected Rule Choices Are: ▼

Accept Rules as Shown Cancel

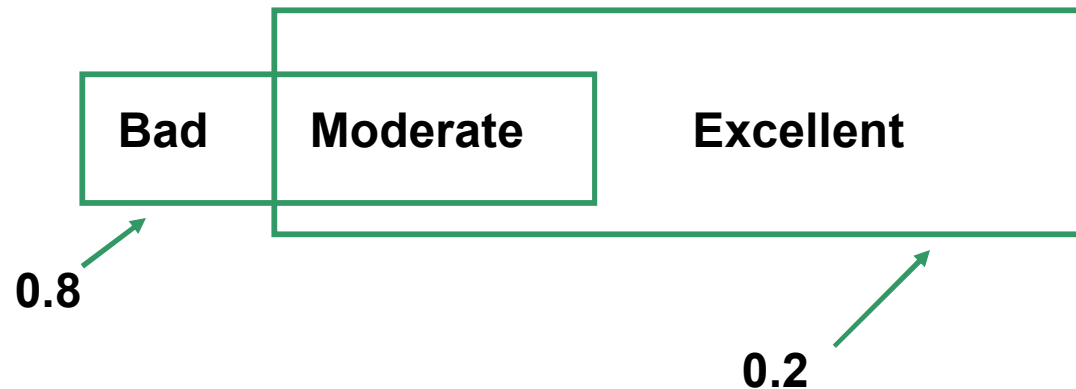


# Combination of Linguistic Variables: Example

---

**Evaluate the Model for Specific Individual:  
Happiness for “John”**

**Assign Evidence to Fuzzy Sets for Basic Variables**  
**Health**

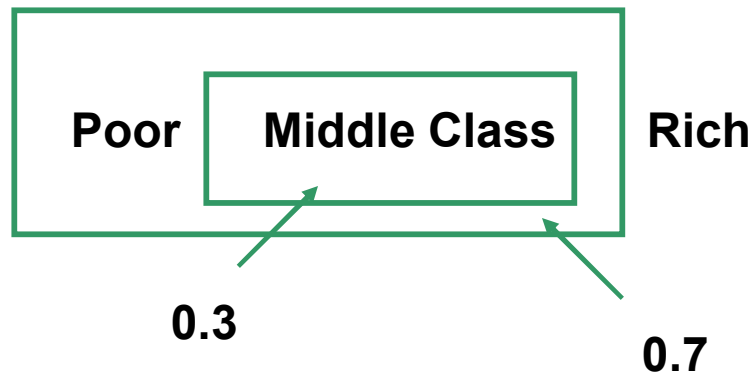




# Combination of Linguistic Variables: Example

---

**Wealth**





# Combination of Linguistic Variables: Example

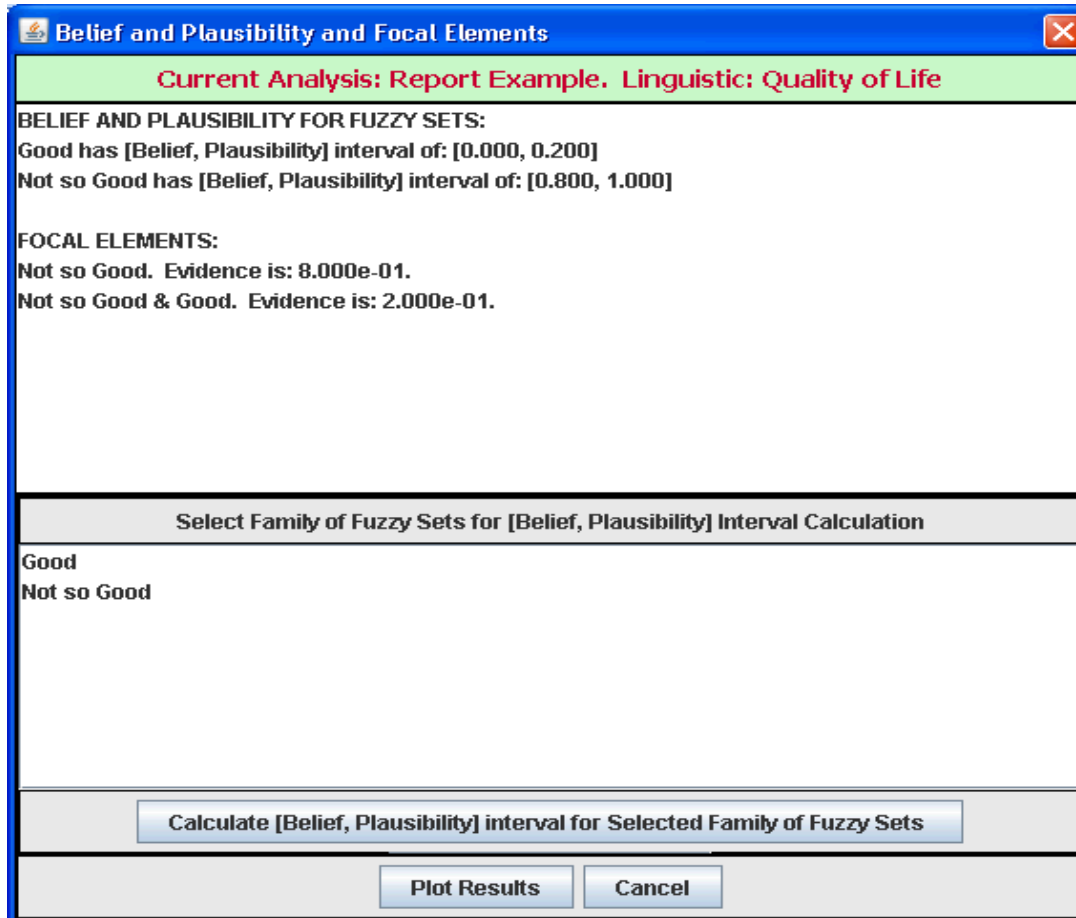
---

## Outlook on Life



# Combination of Linguistic Variables: Example

## Evaluate Variable: Quality of Life for John



**Belief and Plausibility and Focal Elements**

**Current Analysis: Report Example. Linguistic: Quality of Life**

**BELIEF AND PLAUSIBILITY FOR FUZZY SETS:**  
Good has [Belief, Plausibility] interval of: [0.000, 0.200]  
Not so Good has [Belief, Plausibility] interval of: [0.800, 1.000]

**FOCAL ELEMENTS:**  
Not so Good. Evidence is: 8.000e-01.  
Not so Good & Good. Evidence is: 2.000e-01.

**Select Family of Fuzzy Sets for [Belief, Plausibility] Interval Calculation**

Good  
Not so Good

**Calculate [Belief, Plausibility] interval for Selected Family of Fuzzy Sets**

**Plot Results** **Cancel**

# Combination of Linguistic Variables: Example

## Evaluate Variable Happiness for John

**Belief and Plausibility and Focal Elements**

**Current Analysis: Report Example. Linguistic: Happiness**

**BELIEF AND PLAUSIBILITY FOR FUZZY SETS:**  
Depressed has [Belief, Plausibility] interval of: [0.016, 1.000]  
Accepting has [Belief, Plausibility] interval of: [0.000, 0.984]  
Very Happy has [Belief, Plausibility] interval of: [0.000, 0.196]

**FOCAL ELEMENTS:**  
Depressed. Evidence is: 1.600e-02.  
Depressed & Accepting. Evidence is: 7.880e-01.  
Depressed & Accepting & Very Happy. Evidence is: 1.960e-01.

**Select Family of Fuzzy Sets for [Belief, Plausibility] Interval Calculation**

Depressed  
Accepting  
Very Happy

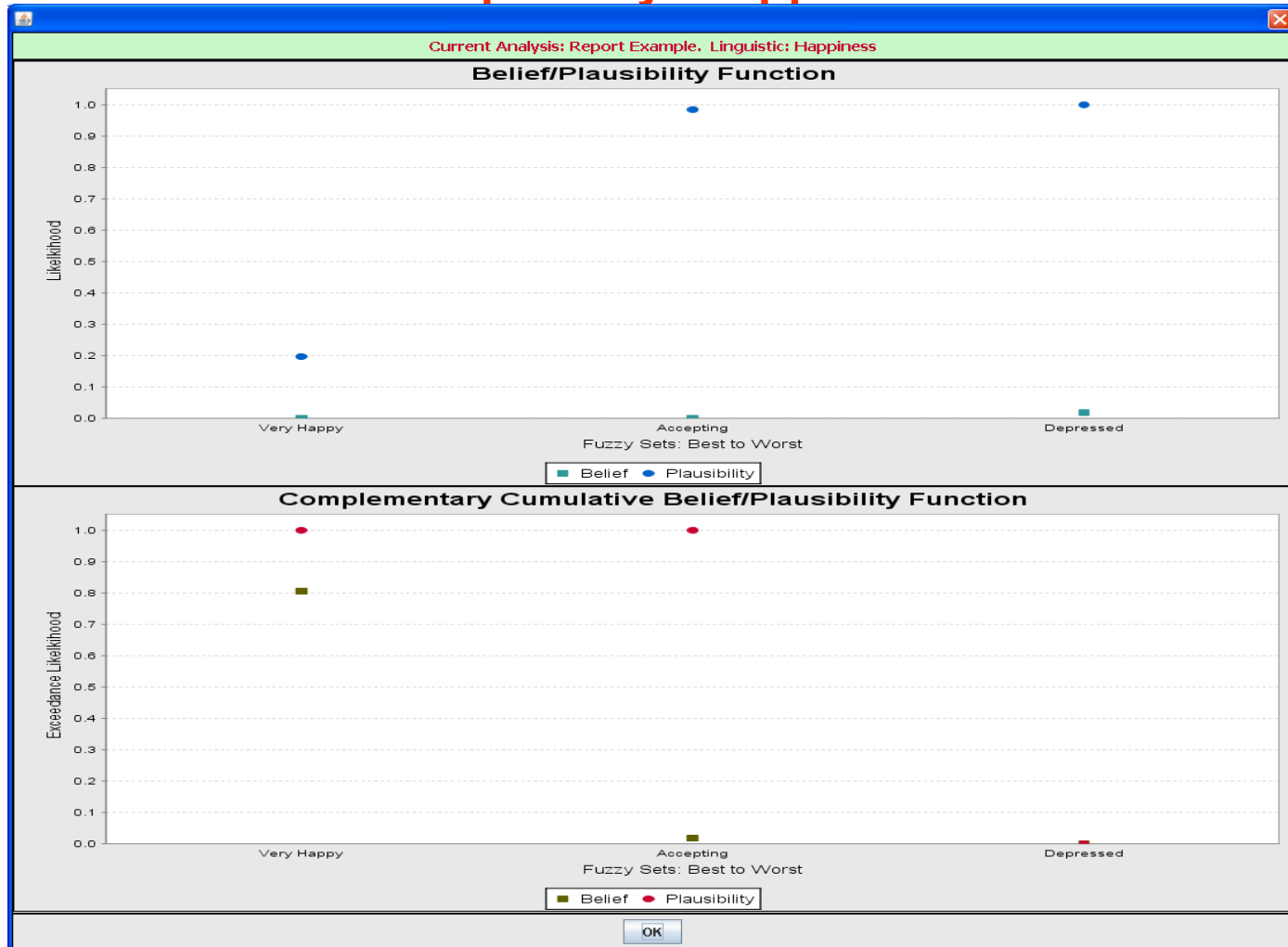
**Calculate [Belief, Plausibility] interval for Selected Family of Fuzzy Sets** [0.000e+00, 9.840e-01]

**Plot Results** **Cancel**

jldarby@sandia.gov

# Combination of Linguistic Variables: Example

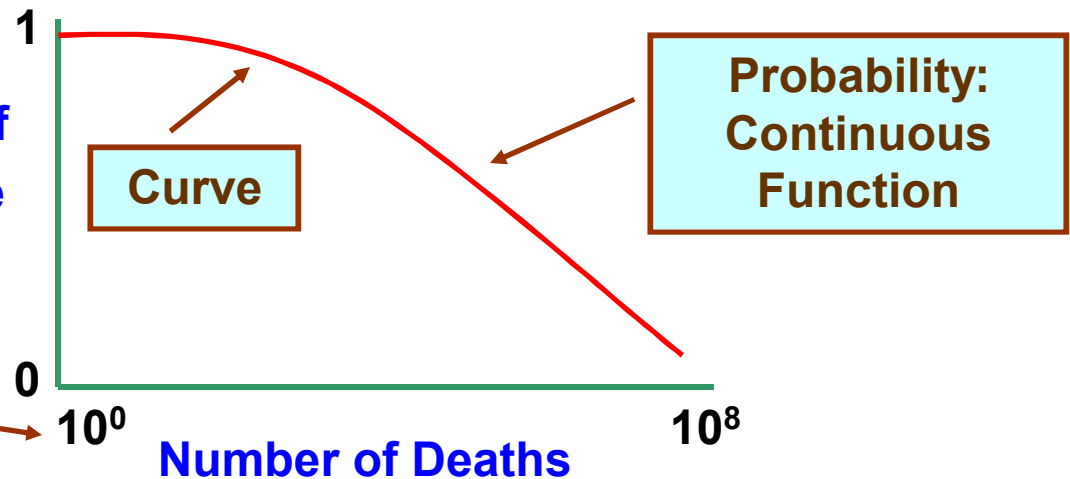
## Summarize Results Graphically: Happiness for “John”



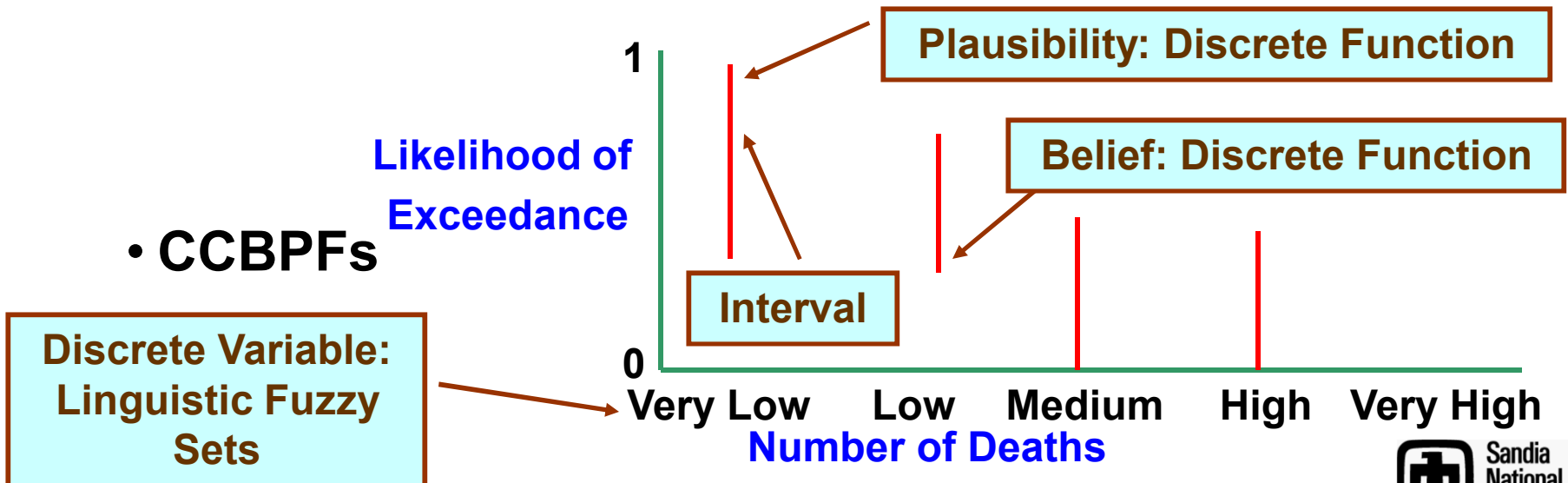


# CCDF and CCBPFs

- **CCDF** Likelihood of Exceedance



- **CCBPFs** Likelihood of Exceedance







# Custom SNL Software Tools: Java

---

- **BeliefConvolution**
  - Convolution of **Numeric** Variables with Belief/Plausibility
- **LinguisticBelief**
  - Evaluation of **Linguistic** Variables
    - Linguistic Fuzzy Sets
    - Approximate Reasoning
    - Belief/Plausibility
- **PoolEvidence**
  - Multiple Experts provide Evidence for Variables
    - Linguistic Fuzzy Sets
  - Combine Evidence
    - Pooled Evidence for Variables
  - Input for LinguisticBelief

# LinguisticBelief: Example Application

**LinguisticBelief Application**

File Utilities Help

New Open Save EXIT! Exit

Current Analysis: National Planning Scenario

- Rule Linguistics
  - Long Term Consequence
  - Casualties (Equivalent Dead)
  - Adversary Estimate Consequence
  - Adversary Estimate Information Required
  - Detect Adversary Gathering Information
  - Adversary Estimate Scientific/Engineering Attributes Required
  - Detect Adversary Gathering Scientific/Engineering Attributes
  - Adversary Estimate Attributes Required
  - Adversary Estimate Resources Required
  - Detect Adversary Gathering Attributes
  - Immediate Consequence
  - Detect Adversary Gathering Resources
  - Vulnerability
  - Consequence
  - Adversary Estimate Vulnerability
  - Threat
  - Risk**
- Basic Linguistics
  - Dead
  - Injured/III
  - Infrastructure Damage
  - Evacuation
  - Contamination
  - National Economic Impact

**Information for Selected Tree Node**

fuzzy sets: credible, Very Low, medium, -- Rule output fuzzy set: Negligible\*, \* Rule input fuzzy sets: credible, Very Low, high, -- Rule output fuzzy set: Negligible\*, \* Rule input fuzzy sets: credible, Very Low, very high, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Low, very low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Low, low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Low, medium, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Low, high, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Low, very high, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Medium, very low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Medium, low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Medium, medium, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, Medium, high, -- Rule output fuzzy set: Medium\*, \* Rule input fuzzy sets: credible, Medium, very high, -- Rule output fuzzy set: High\*, \* Rule input fuzzy sets: credible, High, very low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, High, low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: credible, High, medium, -- Rule output fuzzy set: Medium\*, \* Rule input fuzzy sets: credible, High, high, -- Rule output fuzzy set: High\*, \* Rule input fuzzy sets: credible, High, very high, -- Rule output fuzzy set: Extreme\*, \* Rule input fuzzy sets: likely, Very Low, very low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Very Low, low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Very Low, medium, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Very Low, high, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Very Low, very high, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Low, very low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Low, low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Low, medium, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Low, high, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Low, very high, -- Rule output fuzzy set: Medium\*, \* Rule input fuzzy sets: likely, Medium, very low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, Medium, low, -- Rule output fuzzy set: Medium\*, \* Rule input fuzzy sets: likely, Medium, medium, -- Rule output fuzzy set: High\*, \* Rule input fuzzy sets: likely, Medium, high, -- Rule output fuzzy set: High\*, \* Rule input fuzzy sets: likely, Medium, very high, -- Rule output fuzzy set: High\*, \* Rule input fuzzy sets: likely, High, very low, -- Rule output fuzzy set: Low\*, \* Rule input fuzzy sets: likely, High, low, -- Rule output fuzzy set: Medium\*, \* Rule input fuzzy sets: likely, High, medium, -- Rule output fuzzy set: High\*, \* Rule input fuzzy sets: likely, High, high, -- Rule output fuzzy set: Extreme\*, \* Rule input fuzzy sets: likely, High, very high, -- Rule output fuzzy set: Extreme\*, . [Belief, Plausibility] Intervals for fuzzy sets are as follows : Fuzzy Set: Negligible [0.000, 0.384], Fuzzy Set: Low [0.000, 0.349], Fuzzy Set: Medium [0.155, 1.000], Fuzzy Set: High [0.000, 0.720], Fuzzy Set: Extreme [0.000, 0.000], .



## **Conclusion: We Covered**

---

- **Risk for evaluation of safety**
- **Probabilistic uncertainty**
  - **Classical approach**
  - **Bayesian approach**
- **Epistemic uncertainty**
- **Belief / Plausibility measure**
- **Fuzzy Sets**
- **Approximate reasoning for purely linguistic variables**



# Suggestions

---

- **Never: create your own “new” approach for risk and uncertainty on the fly**
- **Good: Select an existing technique best suited to the fidelity of the information you have**
- **Better: Ask for help from an expert**
  - You are the subject matter expert
  - Get help from experts on risk and uncertainty



# Backup Information Some More Details

---



# Probability Concepts

---

- **Convention: upper case letter denotes a Random Variable, lower case letter denotes a specific value of the random variable**
  - **Random Variable:  $R$** 
    - e.g.  $R = \{-2, -1, 0, 1, 2\}$
  - **$r$ : specific value for  $R$** 
    - e.g.  $r = 2$
- **Random Variables can be Discrete or Continuous**
  - $R = \{-2, -1, 0, 1, 2\}$  is Discrete,  $r$  cannot be 0.2
  - $T = \{x \mid x \text{ in } [-2, 2]\}$  is Continuous,  $t$  can be 0.2
    - $[a, b]$  denotes the **interval** of all real numbers between  $a$  and  $b$  inclusive
    - $(a, b)$  denotes the interval of all real numbers between  $a$  and  $b$  excluding  $a$  including  $b$
- **Combinations of Random Variables**
  - Random Vector
  - $R$  and  $T$  random variables
  - Cartesian product  $R \times T = \{<r, t>\}$  is a Random Vector
  - $R$  and  $T$  are independent random variables  
if  $P(<r, t>) = P(r) * P(t)$



# Probability Concepts

---

- **Functions of a Random Variable**
  - R a random variable, r a specific value of r
  - $f(r)$  a function for R
    - $z = f(r)$
    - $P(z) = \sum P(r) \mid f(r) = z$ 
      - Add since **mutually exclusive**
    - **Example**
      - R is  $\{-2, 1, 0, 1, 2\}$
      - Assume each r has same probability: 0.2
      - $f(r) = r^2$
      - $P(z = 4) = P(r = -2) + P(r = 2) = 0.4$
      - $P(z = 1) = P(r = -1) + P(r = 1) = 0.4$
      - $P(z = 0) = P(r = 0) = 0.2$
      - $P(\text{any } z) = 0.4 + 0.4 + 0.2 = 1.0$



# Probability Concepts

---

- Functions of a Random Vector
  - R and T random variables
  - function  $z = f(r, t)$
  - $P(z) = \sum P(\langle r, t \rangle \mid f(r, t) = z)$ 
    - Add since **mutually exclusive**
    - This is convolution: folding two probability distributions
  - Example
    - $R = \{-2, -1, 0, 1, 2\}$  each outcome prob 0.2
    - $T = \{0, 1\}$  each outcome prob 0.5
    - $R \times T = \{\langle -2, 0 \rangle, \langle -1, 0 \rangle, \langle 0, 0 \rangle, \langle 1, 0 \rangle, \langle 2, 0 \rangle, \langle -2, 1 \rangle, \langle -1, 1 \rangle, \langle 0, 1 \rangle, \langle 1, 1 \rangle, \langle 2, 1 \rangle\}$
    - Assume R and T **independent**:  $P(\langle r, t \rangle) = P(r) * P(t)$
    - Define  $f(r, t) = r + t$
    - $P(z = 2) = P(\langle 2, 0 \rangle) + P(\langle 1, 1 \rangle) = 0.2 * 0.5 + 0.2 * 0.5 = 0.2$
    - $R + T = \{-2, -1, 0, 1, 2, 3\}$ 
      - $P(-2) = 0.1, P(-1) = 0.2, P(0) = 0.2, P(1) = 0.2, P(2) = 0.2, P(3) = 0.1$





# Probability Concepts

---

- **Probability for Function of a Continuous Random Vector**
  - **Convolute (faltung or fold) the probability distributions for the constituent random variables under the operation specified by the function**
  - **$z = f(r, t)$ , R and T continuous random variables**
  - **PDF(z) =  $\int \text{PDF}(r, t) \mid f(r, t) = z$      $z = x + y$      $P(z) = \int_x P(x) * P(z - x) dx$**
  - **If R and T are independent random variables**  
**PDF(r, t) = PDF(r) \* PDF(t)**
- **End digression on Probability; back to Uncertainty for Risk**



# Classical Statistical Inference

---

- Upper Confidence Level (UCL)
- “p is in the interval [0, UCL] to confidence level C%” means C% of a large number of [0, UCL] confidence intervals constructed from repeated samples contains P.
- [0, UCL] is an upper one sided confidence interval
- From [Martz and Waller Bayesian Reliability Analysis](#)
  - $\alpha$  specifies  $(1 - \alpha)$  confidence interval

Binomial dist.

$$C\% \equiv (1-\alpha)100\%$$

$$UCL(x) = \frac{(x+1)F_{1-\alpha}(2x+2, 2n-2x)}{(n-x) + (x+1)F_{1-\alpha}(2x+2, 2n-2x)}$$

- For  $x = 0$ , 50% UCL is:  $1 - 0.5^{1/n}$



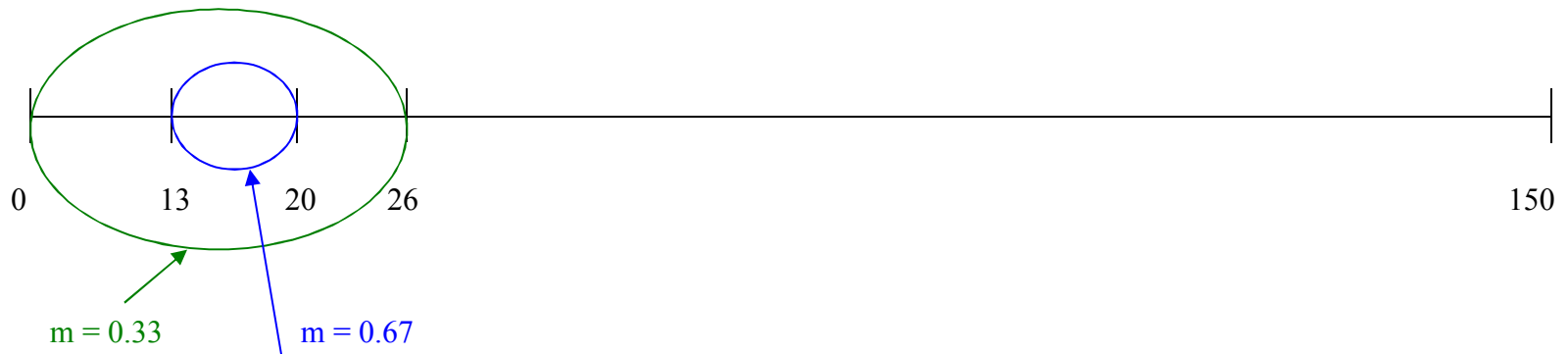
# Evidence is NOT The Measure of Uncertainty

---

- Evidence for any subset (interval) R is: “Likelihood” that outcome is **exactly** in R (in R and nowhere else)
- Belief for any subset (interval) R is: “Likelihood” that outcome is in R or **any subset** of R
- Plausibility for any subset (interval) R is: “Likelihood” that outcome is in R or any subset that **overlaps** (is not disjoint with) R
- Probability thinkers have trouble understanding how T a **subset** of R can have **more** evidence than R
  - $T \subseteq R$ 
    - Since R contains T,  $\text{Probability}(R) \geq \text{Probability}(T)$
  - They Confuse Evidence with Belief/Plausibility
    - Even If  $\text{Evidence}(R) < \text{Evidence}(T)$ 
      - $\text{Belief}(R) \geq \text{Belief}(T)$  and  $\text{Plausibility}(R) \geq \text{Plausibility}(T)$
  - Example Follows

# Example of Evidence

- $(0, 26)$  contains  $[13, 20)$ , but the evidence 0.33 for  $(0, 26)$  is **less** than the evidence 0.67 for  $[13, 20)$ , because the evidence for an interval is the “likelihood” of being **exactly** in that interval and *not* localized within any subinterval.



# Example of Evidence

- $(0, 26)$  contains  $[13, 20)$ . The belief for  $(0, 26)$  will be greater than or equal to than the belief for  $[13, 20)$ . The belief for an interval is the total evidence of being *in that interval or any other interval within that interval*.
- $(0, 26)$  contains  $[13, 20)$ . The plausibility for  $(0, 26)$  will be greater than or equal to than the plausibility for  $[13, 20)$ . The plausibility for an interval is the total evidence of being *in any interval that overlaps that interval (any interval not disjoint with that interval)*.





# Belief/Plausibility can be Viewed as Evidence that Supports/Does not Contradict

---

- Example: X a continuous random variable over  $[0, 1]$

- Event is any Interval

- Consider Event  $[0, 5 \times 10^{-6}]$

- Experts Assign Evidence as Follows

$[0]$  has evidence 0.01

$[0, 10^{-6}]$  has evidence 0.45

$[10^{-6}, 10^{-5}]$  has evidence 0.32

$[10^{-6}, 10^{-4}]$  has evidence 0.20

$[10^{-4}, 10^{-2}]$  has evidence 0.01

$[10^{-3}]$  has evidence 0.01

Evidence  
Supports (within)  
 $[0, 5 \times 10^{-6}]$

Evidence Does Not  
Contradict  
(overlaps)  
 $[0, 5 \times 10^{-6}]$

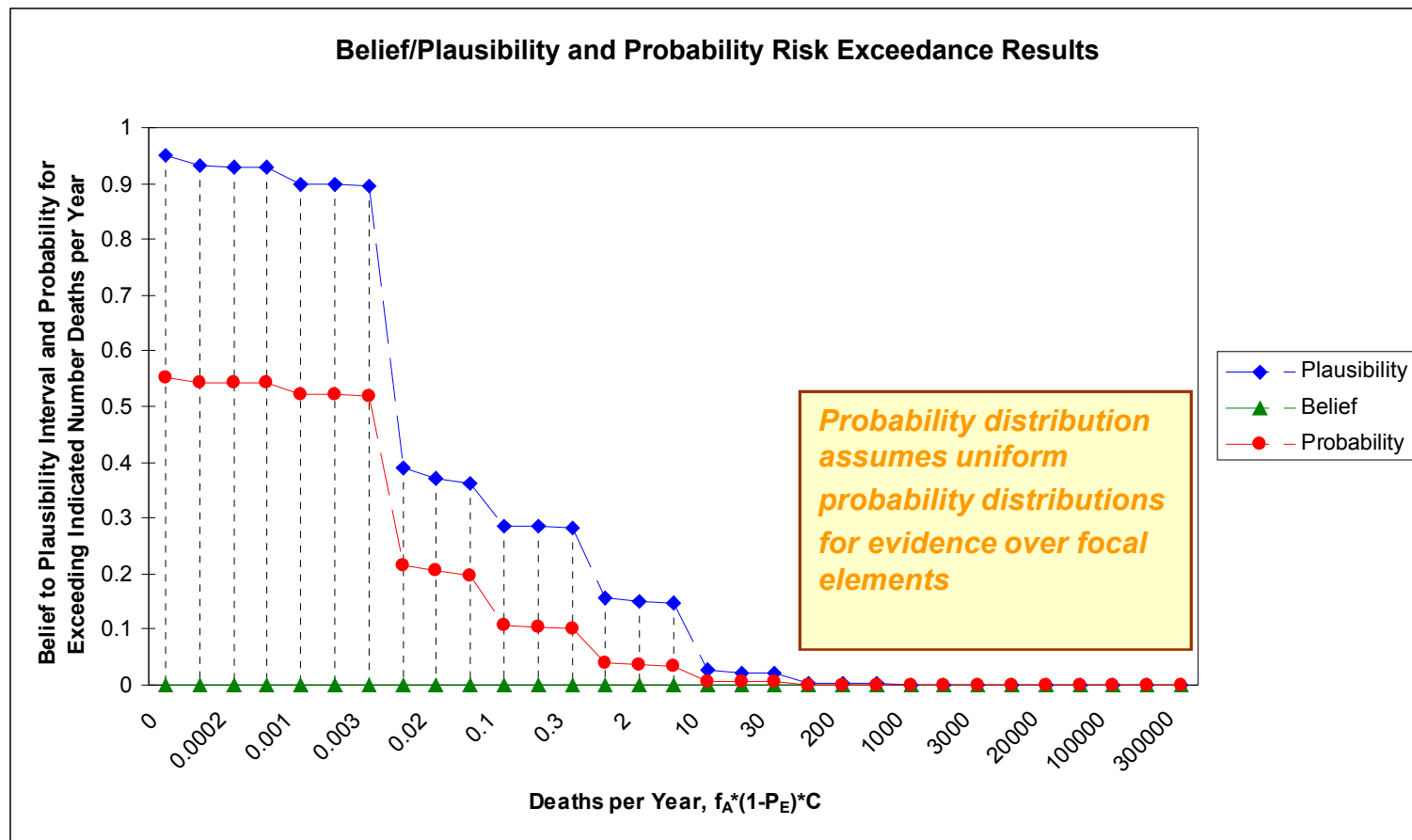
Evidence Contradicts  $[0, 5 \times 10^{-6}]$

- Belief for  $[0, 5 \times 10^{-6}] = 0.01 + 0.45 = 0.46$

- Plausibility for  $[0, 5 \times 10^{-6}] = 0.01 + 0.45 + 0.32 + 0.20 = 0.98$

# Belief / Plausibility

- Risk as Exceedance of Consequence
  - Calculations with BeliefConvolution SNL custom Java software





# Belief / Plausibility

---

- Mean for random variable  $X$  is an interval  $[E_*(X), E^*(X)]$ 
  - Inf (infimum) means greatest lower bound
  - sup (supremum) means least upper bound
  - $A_i$  is a focal element, interval of real numbers

$$E_*(X) = \sum_{all\ A_i \subseteq X} \inf(A_i) * m(A_i)$$

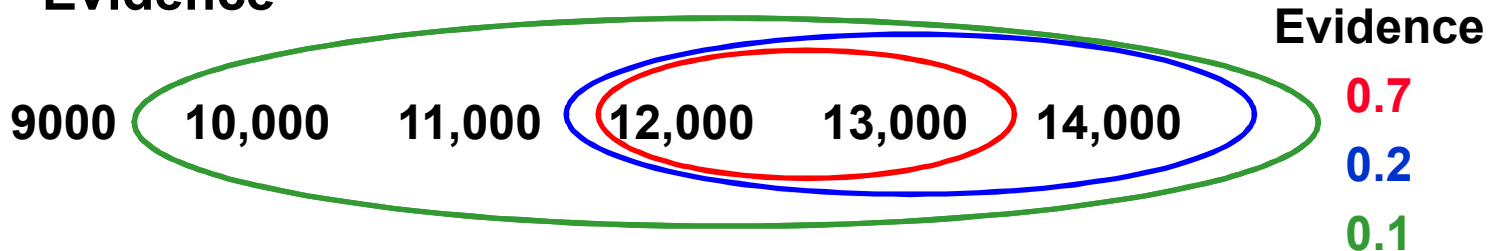
$$E^*(X) = \sum_{all\ A_i \subseteq X} \sup(A_i) * m(A_i)$$



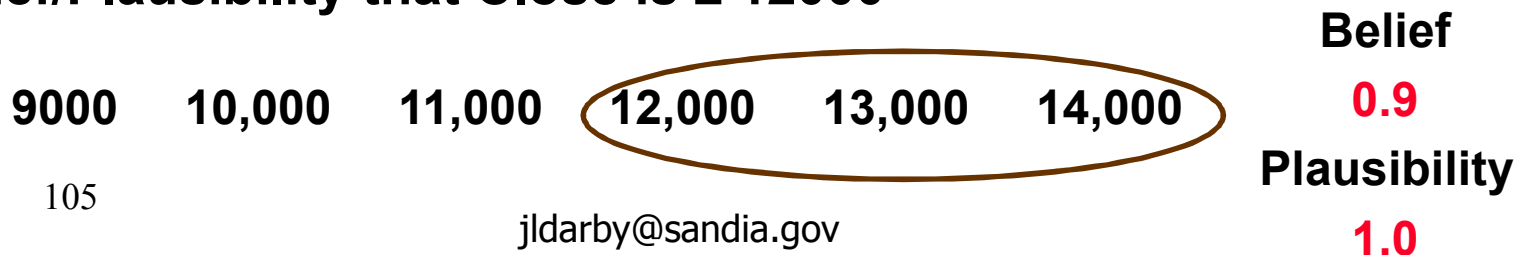
# Evidence for Variable X

- Evidence for any A a subset of X is: “Likelihood” that X is **exactly** in A (in A and nowhere else)
- Belief for any A a subset of X is: “Likelihood” that X is in A or **any subset** of A
- Plausibility for any A a subset of X is: “Likelihood” that X is in A or any subset that **overlaps** (is not disjoint with) A
- Example: Stock Market Close Dec. 31, 2007

– Evidence



- Belief/Plausibility that Close is  $\geq 12000$





# Belief / Plausibility for Function of Random Variables

---

- For a function of random variables
  - Random vector is Cartesian product  $X \times Y$
  - Function  $z = f(x, y)$
  - Evidence is binary relation,  $R$ , on  $X \times Y$  ( $R$  is a subset of  $X \times Y$ )
  - $R_X$  is projection of  $R$  on  $X$   $R_X = \{x \in X \mid \langle x, y \rangle \in R \text{ for some } y \in Y\}$
  - $R_Y$  is projection of  $R$  on  $Y$   $R_Y = \{y \in Y \mid \langle x, y \rangle \in R \text{ for some } x \in X\}$
  - For any subset  $A$  the marginal evidence is  $m_X(A)$  and  $m_Y(B)$ 
    - $\text{Pow}(A)$  denotes power set of  $A$  (set of all subsets of  $A$ )
    - $R|A=R_X$  means all relations  $R$  such that the projection of  $R$  onto  $X$  ( $R_X$ ) is equal to  $A$

$$m_X(A) = \sum_{R|A=R_X} m(R) \text{ for all } A \in \text{Pow}(X)$$

$$m_Y(B) = \sum_{R|B=R_Y} m(R) \text{ for all } B \in \text{Pow}(Y)$$

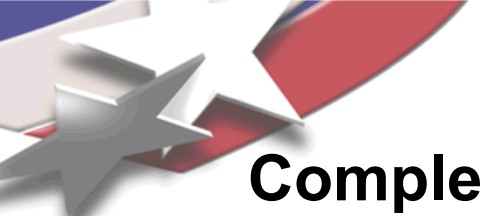
- If subsets  $A$  and  $B$  are non-interactive (extension of independence for probability)
  - $m(A \times B) = m_X(A) * m_Y(B)$ 
    - Like  $P(\langle x, y \rangle) = P(x) * P(y)$  if probabilistic independence
  - $m(R) = 0$  for all  $R \neq A \times B$



# Quantification of Margins and Uncertainty

---

- **Techniques for evaluating data on aging concerns in nuclear weapons**
- **QMU is the “math” to evaluate predictive / diagnostic “data”**



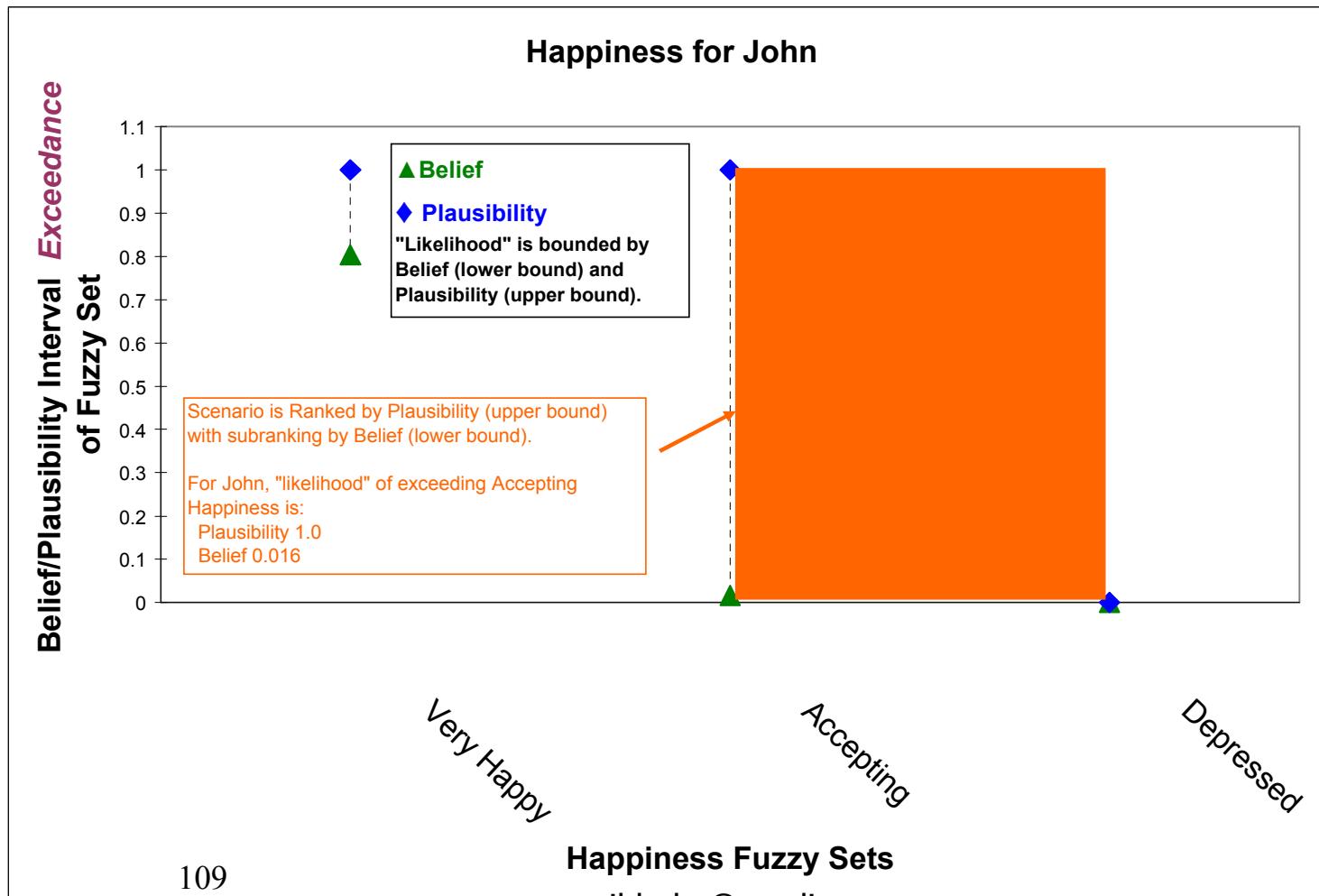
## Results Expressed as Complementary Cumulative Belief/Plausibility Functions (CCBPFs) for Linguistic Variable

---

- Linguistic Fuzzy Sets **Ordered** from “Best” to “Worst”
  - CCBPFs are **Non-Increasing**
- “Likelihood” of **Exceeding** Fuzzy Set
- “Likelihood” is **Belief/Plausibility Interval**
- Analogous to Complementary Cumulative Distribution Function (CCDF) for Probability
  - CCDF Random Variable is a real number
    - discrete or continuous
  - CCBPFs Variable has linguistic fuzzy sets
    - Discrete
  - CCDF is a One Function: a Curve
  - CCBPFs are Two Functions: an Interval

# Combination of Linguistic Variables: Example

## Graphical Summary for Ranking: "John"



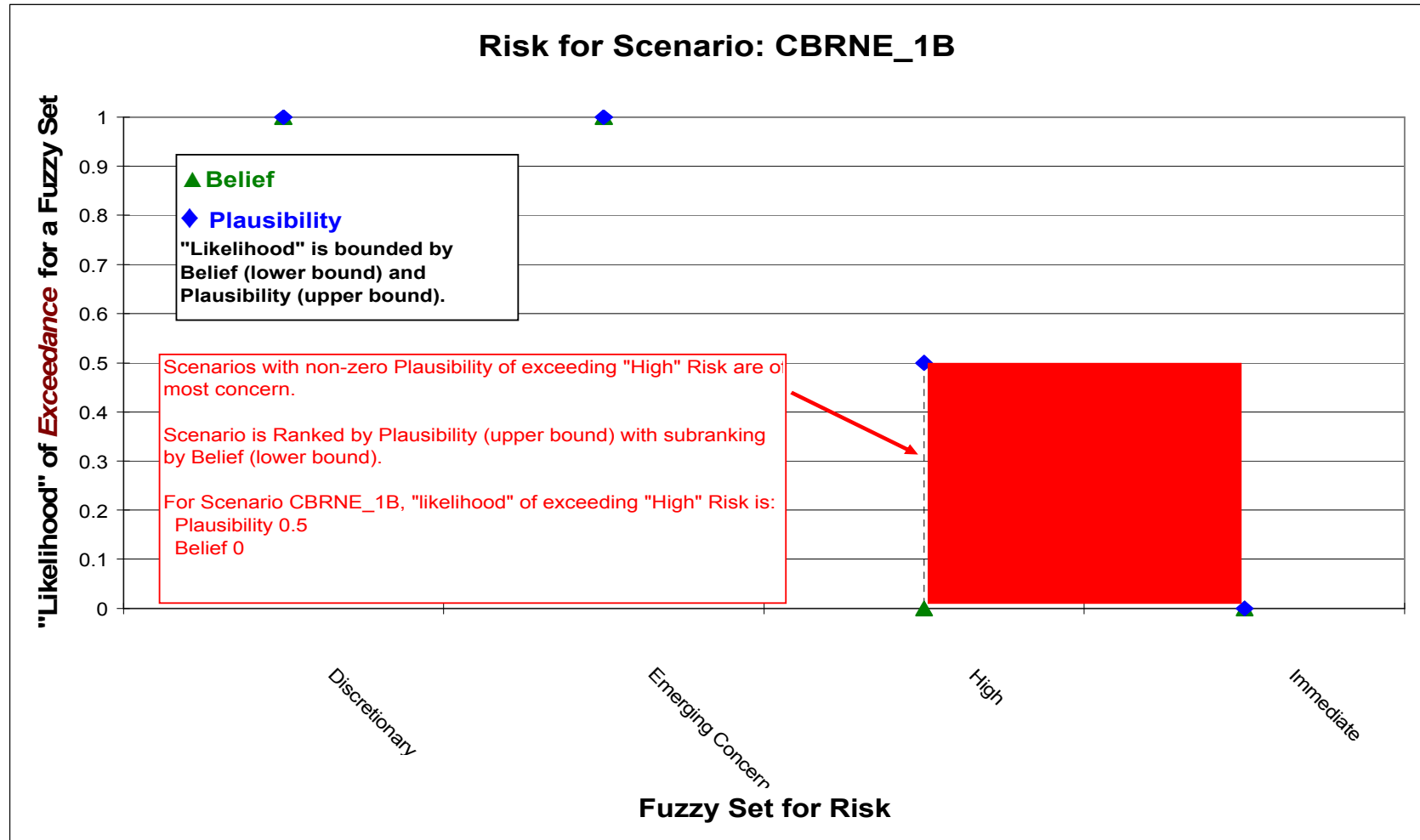


# Rank Order Scenarios by Risk

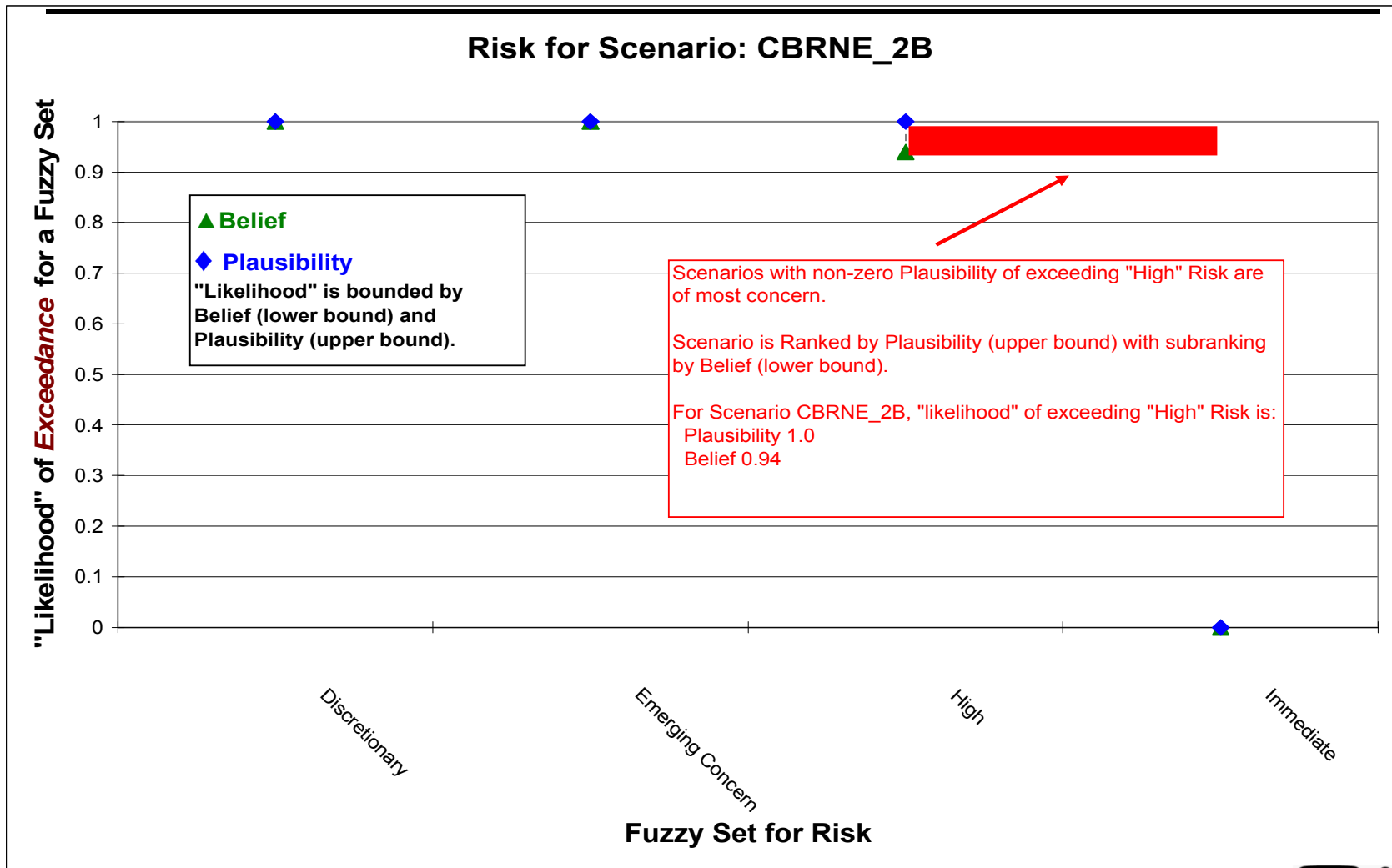
---

- **Rank a scenario by the Highest Non-Zero Plausibility of Exceeding the “Worst” Fuzzy Set**
  - **For Scenarios with Equal Plausibility, Subrank by Highest Belief**
- **Extension of “Probability of Exceedance” approach**
  - **Uses Fuzzy Sets instead of Numbers**
  - **Uses Belief/Plausibility Interval instead of Probability**
- **Can be “Color Coded”**
  - **Shown for 3 of 5 scenarios in Following from SAND2007-1301**

# Rank Order Results

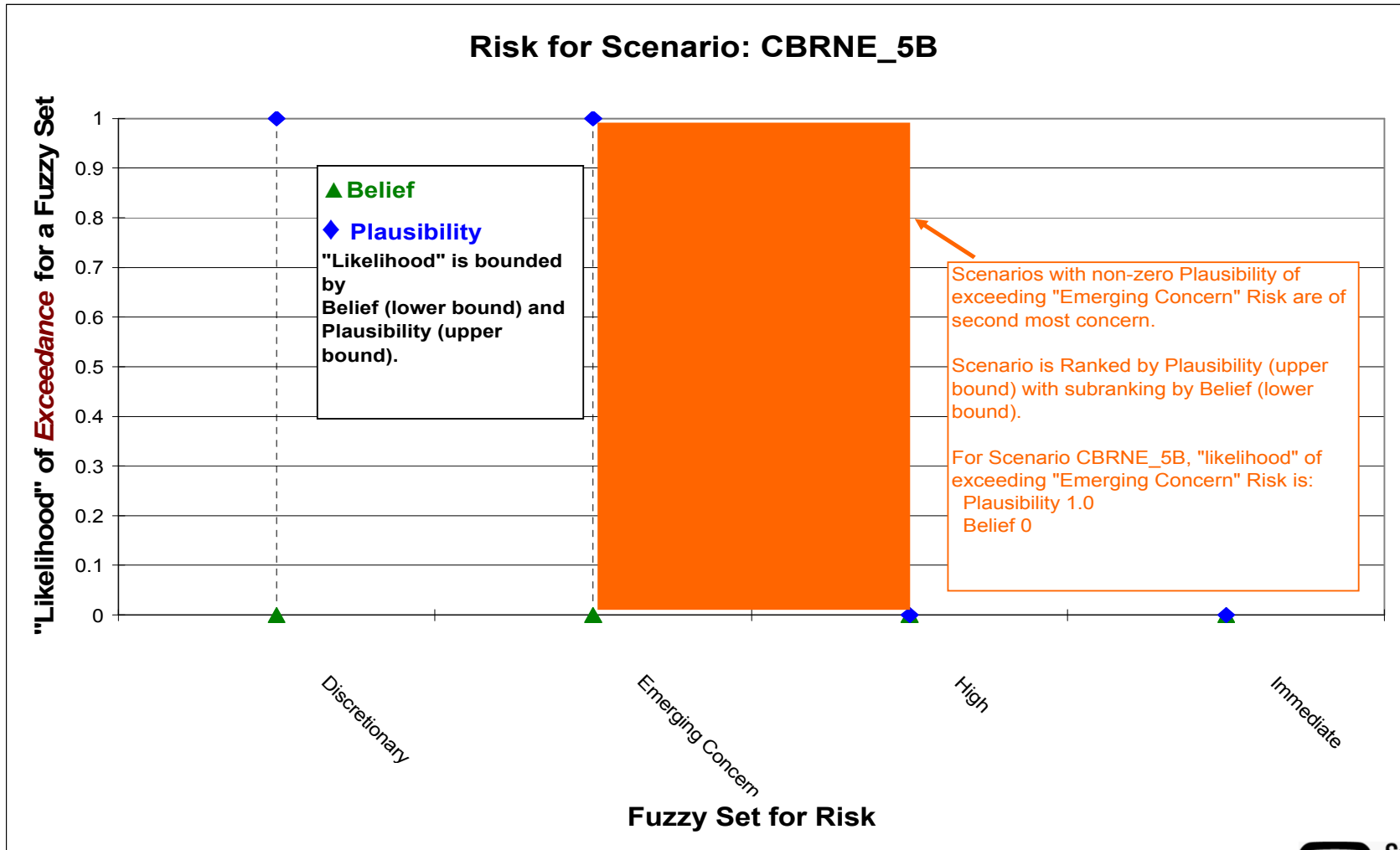


# Rank Order Results





# Rank Order Results





# Results of Ranking of All Five Scenarios

---

## RANKING FOR SCENARIOS CBRNE\_1B through CBRNE\_5B

For **Exceeding** Fuzzy Set “**High**” the Scenarios rank ordered (decreasing) are:

CBRNE\_2B has plausibility of exceedance of 1.0 and belief of exceedance of 0.94

CBRNE\_3B has plausibility of exceedance of 1.0 and belief of exceedance of 0.77

CBRNE\_4B has plausibility of exceedance of 1.0 and belief of exceedance of 0.64

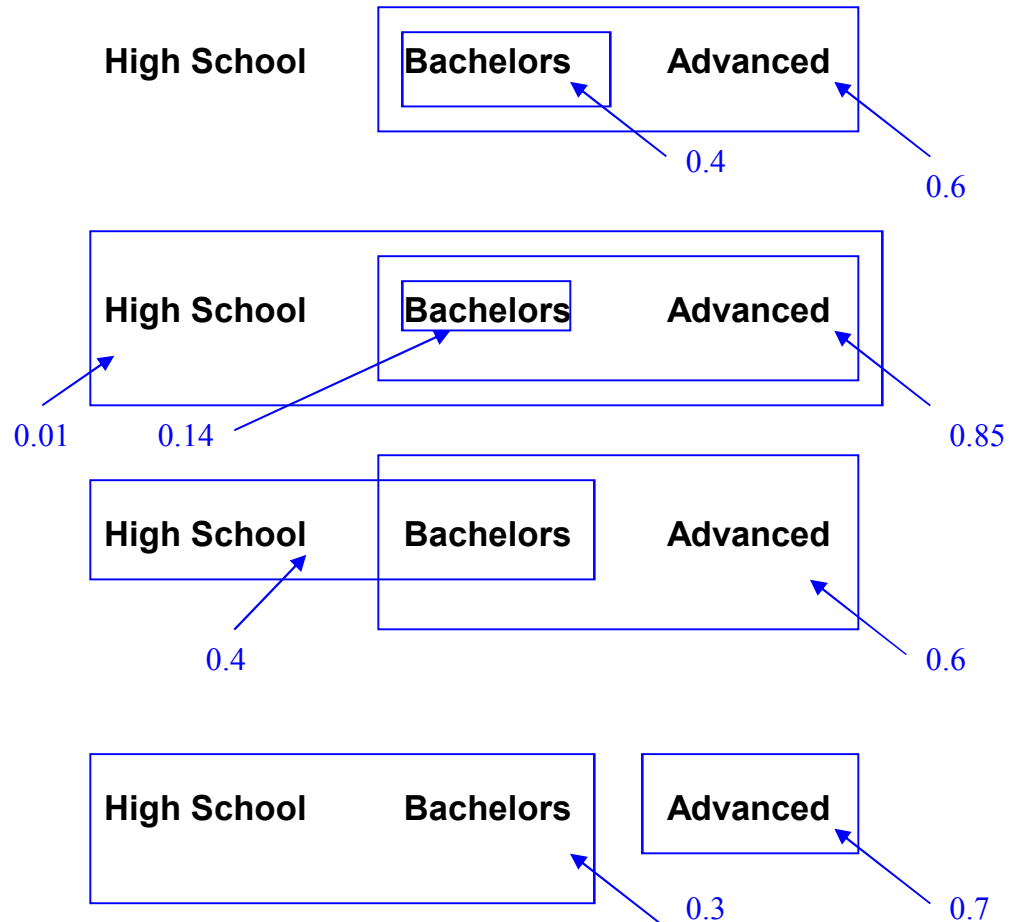
CBRNE\_1B has plausibility of exceedance of 0.5 and belief of exceedance of 0.0

For **Exceeding** Fuzzy Set “**Emerging Concern**” the Scenarios rank ordered (decreasing) (not already ranked for a worse fuzzy set) are:

CBRNE\_5B has plausibility of exceedance of 1.0 and belief of exceedance of 0.0

# Pool Evidence from Many Experts

**Adversary Level of  
Technical Training:  
4 Experts Assign  
Evidence**



# PoolEvidence: Example Application

**Pooled Evidence Application**

File Help

New Open Save EXIT! Exit

Current Analysis: None

Variables for Analysis	Fuzzy Sets for Selected Variable
<p>Variables for Current Analysis: Actual Test</p> <ul style="list-style-type: none"><li>Adversary Motivation</li><li>Defender Resources</li><li>Result</li></ul>	<p>Fuzzy Sets for Variable: Adversary Motivation</p> <ul style="list-style-type: none"><li>Very Low</li><li>Low</li><li>Medium</li><li>High</li><li>Very High</li></ul>

Show Pooled Focal Elements for Selected Variable

Experts for Selected Variable	Focal Elements for Selected Expert
<p>Experts for Variable: Adversary Motivation</p> <ul style="list-style-type: none"><li>Pepper</li><li>Charlie</li><li>Heather</li><li>John</li></ul>	<p>Focal Elements for Expert: John</p> <ul style="list-style-type: none"><li>Low, Medium, High, with Evidence: 2.30000e-01</li><li>Very Low, Medium, with Evidence: 3.20000e-01</li><li>Low, High, with Evidence: 4.50000e-01</li></ul>

Close

# PoolEvidence

