Bayesian Inference in Risk Assessment (P-102)

Curtis Smith
&
Tony Koonce

December 2015

United States Nuclear Regulatory Commission

Idaho National Laboratory

DISCLAIMER NOTICE

Idaho National Laboratory

Idaho National Laboratory

Bayesian Inference in Risk Assessment (P-102)

Course Presented by

Curtis Smith
Curtis.Smith@inl.gov

Tony Koonce
Tony.Koonce@inl.gov

November 30 – December 4, 2015
NRC Professional Development Center
Rockville, MD

Dedication

- The “probability and stats” course was developed by Dr. Dana Kelly in the 1990s

- We dedicate this course to his memory

- One of Dana’s mantras (taken from the Copenhagen interpretation of quantum physics) was “Shut up and calculate!” We embrace this approach in this class…

Bayesian Inference in Risk Assessment (P-102)

- Section 1: Course Topics
- Section 2: Review of Basic Probability Calculations
- Section 3: Introduction to Bayesian Inference
- Section 4: Introduction to Monte Carlo Sampling
- Section 5: Uncertainty Propagation in Risk Assessment

Let us discuss the information in these sections…

Section 1: Course Topics

- The "stats" course, P-102, comprises three sections
  - Review of basic probability calculations
    - Things you should already know, so we will just remind you of them
  - Basic Bayesian statistical inference
    - We will use Excel to do the math
      - Will demo another tool for harder problems
  - Uncertainty propagation in risk assessment
    - Simple Monte Carlo sampling
      - Propagation of parameter uncertainty through risk model
      - We will illustrate this in Excel

Section 2: Review of Probability

- Purpose
  - Students will review probability axioms and operations
- Objectives
  - Students will be able to calculate results involving
    - "AND", "OR", "NOT" operations
    - Conditional probabilities
    - Bayes' theorem
    - Discrete and continuous probability distributions
  - Students will understand the terms mean, variance, standard deviation, percentile, and be able to relate these to particular distributions used in the course

Section 3: Bayesian Statistical Inference

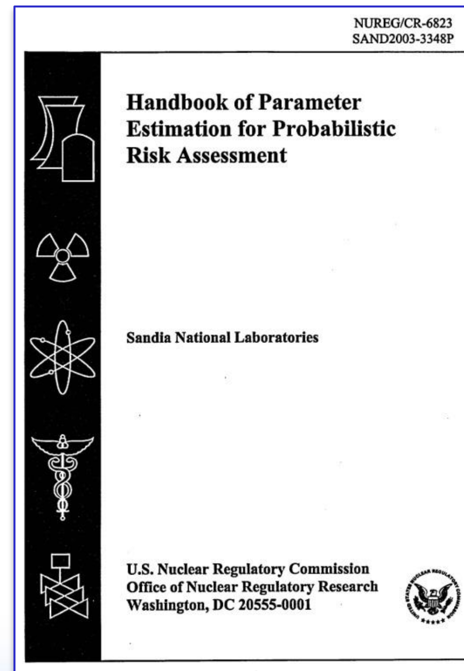
- Purpose
  - Students will learn subjectivist interpretation of probability, concept of Bayesian updating, and applications to commonly encountered kinds of stochastic models
- Objectives
  - Students will learn
    - Probability interpreted as a quantification of degree of plausibility
    - Bayesian inference using Excel for
      - Discrete priors
      - Conjugate priors for Poisson, binomial, and exponential models
      - Formal priors for Poisson, binomial, and exponential models
    - RADS Calculator for conjugate and non-conjugate priors

Sections 4 and 5: Uncertainty Propagation in Risk Assessment

- Purpose
  - Students will see an overview of how Bayesian estimates of risk metrics (e.g., core damage frequency) are obtained
- Objectives
  - Through examples using Excel, students will learn about
    - Monte Carlo sampling of distributions
    - Estimation of a “top event” probability by propagation of distributions through a logic model
    - Simple Monte Carlo sampling and Latin hypercube sampling

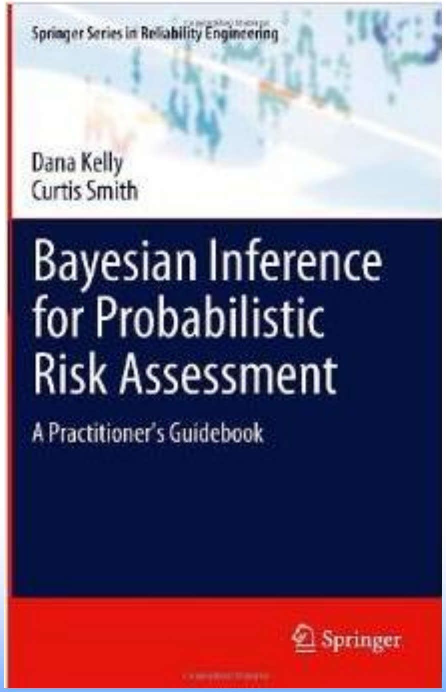
Course Reference

- *Handbook of Parameter Estimation for Probabilistic Risk Assessment*, NUREG/CR-6823, September 2003.
  - Available on NRC web site at
  - www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6823



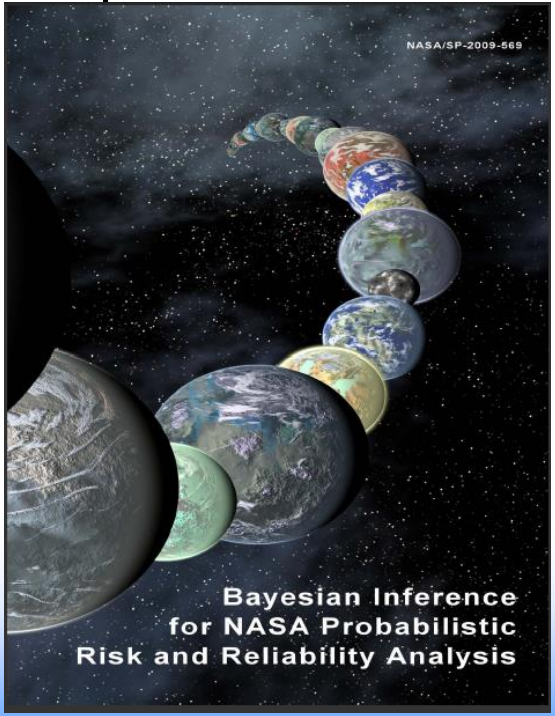
Supplemental Reference

- *Bayesian Inference for PRA: A Practitioner's Guidebook*, 2011
  - Available at
  - www.amazon.com/Bayesian-Inference-Probabilistic-Risk-Assessment/dp/1849961867
  - The text for P-502



Supplemental Reference

- *Bayesian Inference for NASA Probabilistic Risk and Reliability Analysis*, NASA/SP-2009-569, 2009
  – Available at
  – www.hq.nasa.gov/office/codeq/doctree/SP2009569.pdf



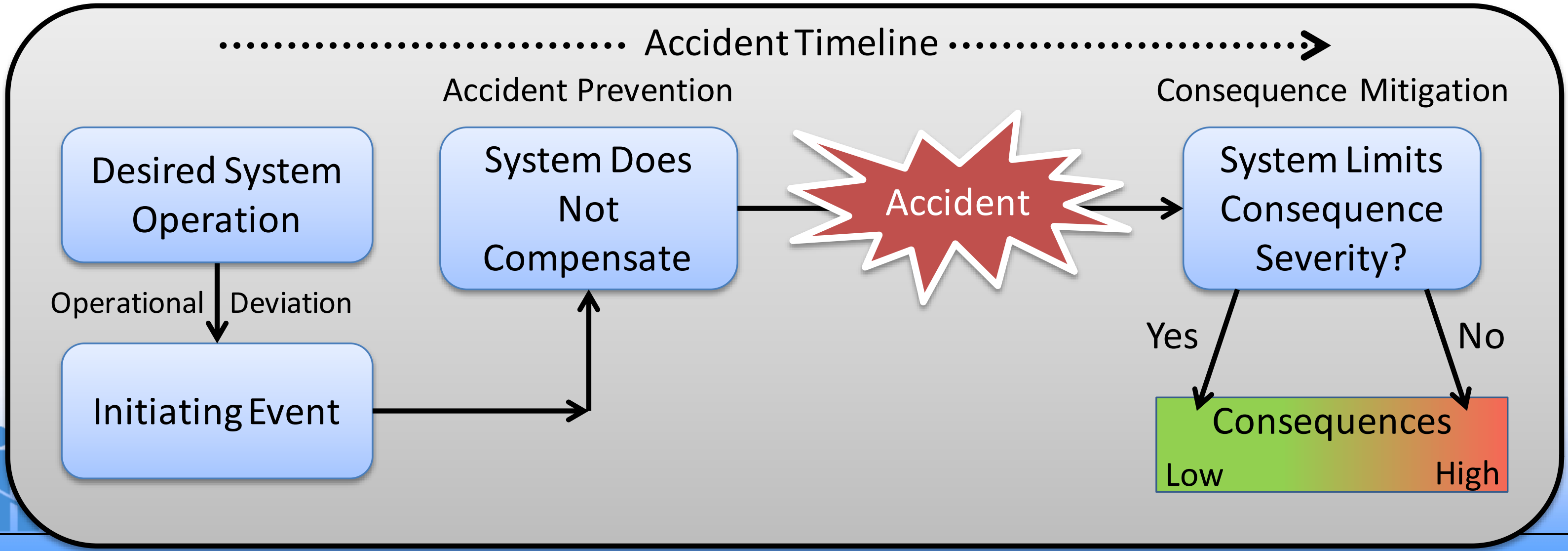
Idaho National Laboratory

Loss of Offsite Power (LOSP) Example

- The “LOSP example” will be used as a central example throughout most of the course
- A system uses offsite power, but has two standby emergency diesel generators (EDGs)
- Occasionally offsite power is lost (an “initiating event”)
  - When this happens the EDGs are demanded to start and run
- The system
  - Succeeds if either EDG starts and runs for six-hour mission time
  - Fails otherwise

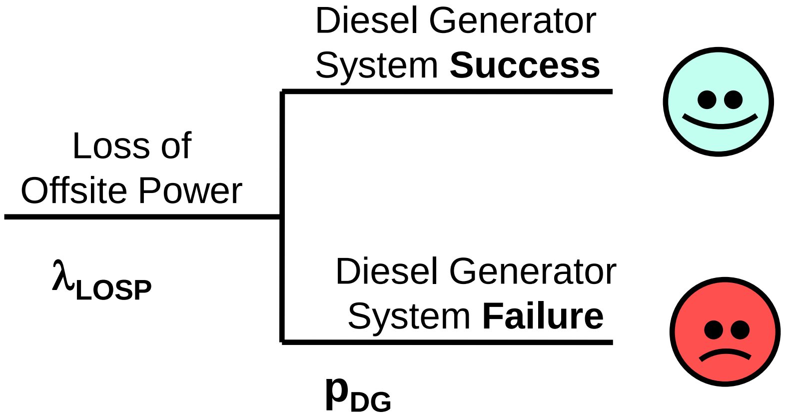
The Concept of a Scenario

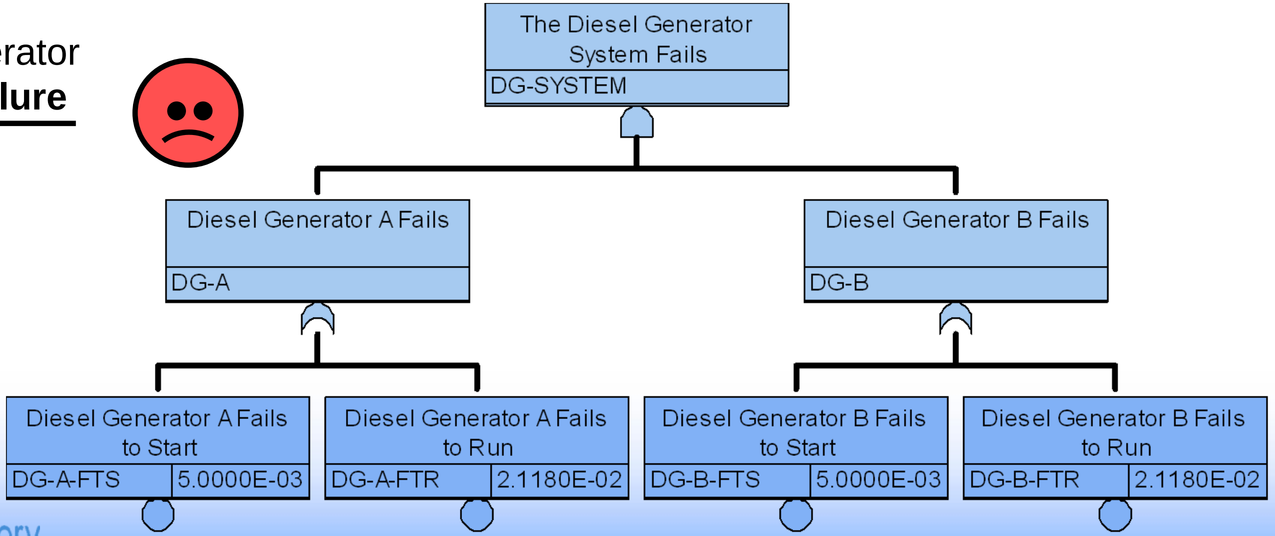
- Scenario modeling
  - For each hazard, identify an initiating event and necessary enabling conditions that result in undesired consequences
- Enabling conditions often involve failure to recognize a hazard or failure to implement controls such as protective barriers
- Accident scenario is the sequence of events comprised of:
  - Initiating event + enabling conditions + events that lead to adverse consequences



LOSP Example

- A PRA will have an event tree representing the scenario
  - Fault trees will represent the diesel generator failures





The Minimal Cut Sets

LOSP * DG-A-FTS * DG-B-FTS   or

LOSP * DG-A-FTS * DG-B-FTR   or

LOSP * DG-A-FTR * DG-B-FTS   or

LOSP * DG-A-FTR * DG-B-FTR

Recovery of Offsite Power

- Core damage can be averted if offsite power is recovered
- Assume traditional engineering analysis shows…
  - Recovery must occur by six hours to avert core damage
- Append nonrecovery event to minimal cut sets
  - This represents probability that offsite power is **not** recovered within six hours

Recovered Cut Sets

LOSP*DG-A-FTS*DG-B-FTS***OSP-NONREC** or

LOSP*DG-A-FTS*DG-B-FTR***OSP-NONREC** or

LOSP*DG-A-FTR*DG-B-FTS***OSP-NONREC** or

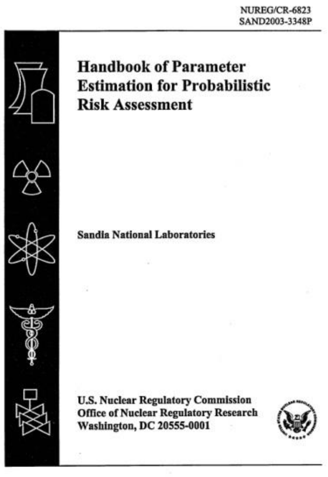
LOSP*DG-A-FTR*DG-B-FTR***OSP-NONREC**

“Real” PRA Cut Sets

- A “real” PRA may have additional terms not considered in this class
  - **Common Cause Failure**
  - **Unavailability** of the component (e.g., out for test or maintenance activities)
  - **Human reliability**
  - **Component recoveries**

Section 2: Review of Basic Probability Calculations

- Purpose
  - Students will review fundamentals of probability
- Objectives
  - Students will be able to perform simple calculations involving
    - "AND", "OR", "NOT" operations
    - Conditional probabilities, independent events
    - Bayes' theorem
  - Students will understand
    - Discrete and continuous probability distributions
    - Moments and percentiles of distributions



NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

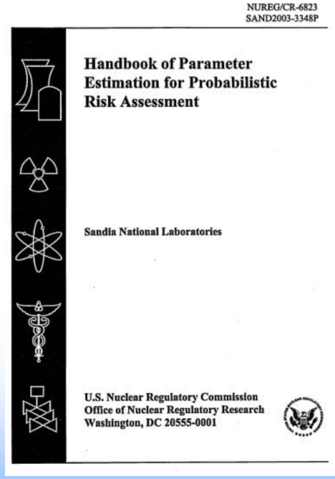
Appendix A

Outline

- Topics to be covered include
  - Basic framework for probabilistic models
  - Rules for manipulating probabilities
  - Discrete probability distributions
  - Continuous probability distributions
  - Moments and percentiles of distributions

Basic Framework

- An **experiment** can result in a number of **outcomes**. Experiment may be “trial,” “test,” “demand,” etc.
- Sample space **S** is the set of all possible outcomes on any one experiment
- An **event is a set of outcomes**
  - Its probability is the sum of the probability of each constituent outcome



Pages A-1 through A-4

Example 1

- Experiment: Rolling six-sided die
- The possible outcomes (i.e. the sample space, S)
  - One of the six faces of the die
- Some possible events
  - A particular number
  - Even number
  - Odd number
  - Etc.

Example 2

- Experiment: Try to start EDG-A
- The possible outcomes (i.e. the sample space, S)
  - Failure to start (FTS_A)
  - Start but failure to run (FTR_A)
  - Start and run to end of mission ($Success_A$)
- Some possible events
  - EDG-A fails somehow
  - EDG-A starts
  - Etc.

Example 3

- Experiment: Try to start two EDGs, EDG-A and EDG-B
- The outcomes (i.e. the sample space):

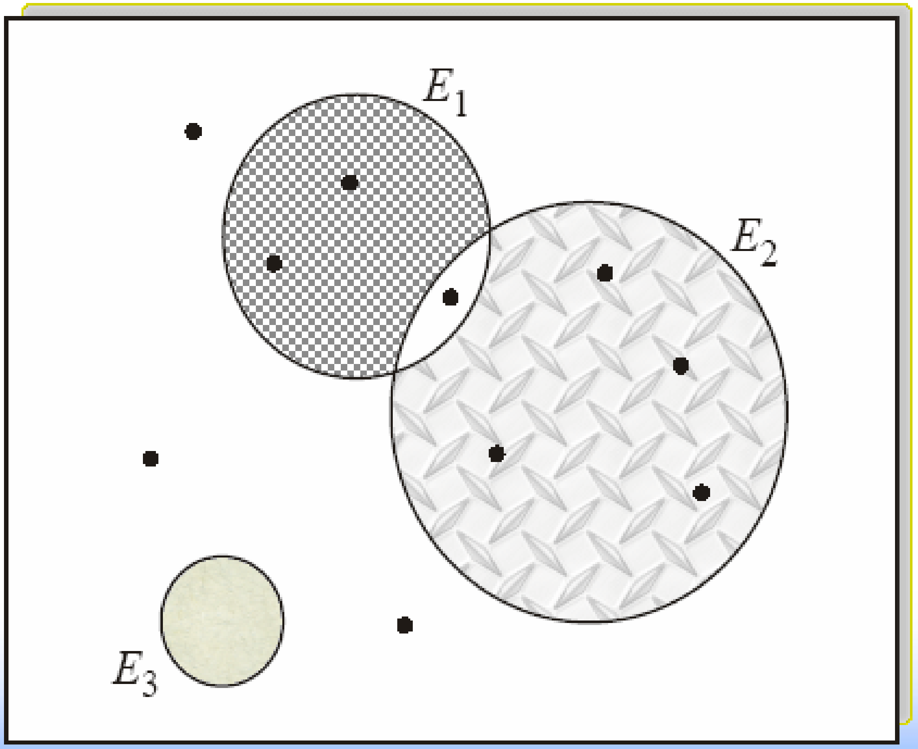
FTS_A & FTS_B	FTS_A & FTR_B	FTS_A & $Success_B$
FTR_A & FTS_B	FTR_A & FTR_B	FTR_A & $Success_B$
$Success_A$ & FTS_B	$Success_A$ & FTR_B	$Success_A$ & $Success_B$

- Some possible events
  - At least one EDG succeeds
  - Both EDGs fail somehow
  - At least one EDG fails to start
  - Exactly one EDG fails
  - Etc.

Example 4

John Venn

- It is sometimes helpful to show **events** and **outcomes** via a Venn diagram
  - Three events, 10 outcomes



Building Events from Other Events or Outcomes — OR

- A OR B = combined event containing all events that are in A or in B
  - Also written A ∪ B, the **union** of A and/or B
  - The union symbol, ∪, is easy to remember since symbol looks like the letter "U"
- In a PRA, minimal cut sets are "ORed" together to obtain overall results of the analysis

Building Events from Other Events or Outcomes — AND

- A AND B = combined event containing all events that are both in A and in B
  - Also called **intersection** of A and B, written $A \cap B$
  - The intersection symbol $\cap$ can be remembered as the opposite of the union symbol, or n in and
- In a PRA, the events **within a single minimal cut set** are “ANDed” together to obtain the cut set value
- A and B are **disjoint or mutually exclusive** if they have no events in common
  - I.e. A AND B is empty (denoted by $\varnothing$)

Building Events from Other Events or Outcomes — NOT

- The **complement** of A, or NOT A, is the event containing all the events (in the sample space) that are not in A.
- Written $\overline{A}$ or /A or A* or $\neg$A
  - Example: In SAPHIRE (will see later), successfully starting DG-B denoted as /DG-B-FTS

Elementary “Rules” of Probability

1. Probability of an event A, “Pr(A)” or “P(A),” is a **nonnegative** real number
2. Probability of the **union** of non-overlapping (disjoint) events is the sum of the event probabilities
3. Probability of **all** possible outcomes (i.e., the sample space) equals 1.0

- Can show from above axioms that $0 \leq Pr(A) \leq 1$

Andrei Kolmogorov

Rules for Manipulating Probabilities - Complements

- The NOT (or complement) operation
  - Subtract probability from 1.0
  - Example, Pr(not A) = 1 - Pr(A)
- A probability problem tip
  - With messy problems using terms such as “at least” or “at most,” first calculate probability of complement of event:
    - Pr(A) = 1 - Pr(not A)
    - For example, Pr(at least one failure) = 1 – Pr(zero failures)
  - “At least” $\rightarrow$ >
    - For example, Pr(at least one failure) = Pr(# failures > 0)
  - “At most” $\rightarrow$ <

Rules for Manipulating Probabilities – OR (Union)

- For the **OR** (or union) operation, we consider two cases
  1. **If** A, B are **disjoint**
     - Pr(A or B) = Pr(A) + Pr(B)
     - Examples
       – With a die, Pr(1 or 2) = Pr(1) + Pr(2) because outcomes are disjoint
       – With a coin toss, Pr(H or T) = Pr(H) + Pr(T)
  2. **In general**, even if A, B are not disjoint
     - Pr(A or B) = Pr(A) + Pr(B) - Pr(A AND B)
     - Can extend to three or more events by using the inclusion-exclusion rule
       – http://en.wikipedia.org/wiki/Inclusion-exclusion_principle

Rules for Manipulating Probabilities – AND (Intersection)

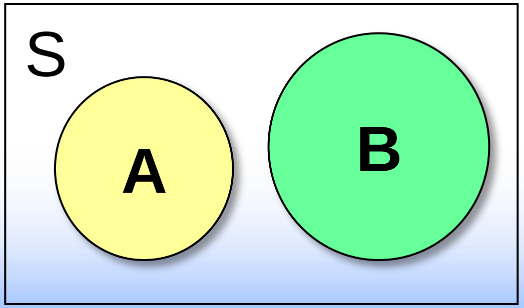
- For the **AND** (or intersection) operation, we consider two cases
  1. If A, B are **independent**
     - Pr(A AND B) = Pr(A) • Pr(B) (this is definition of statistical independence)
  2. If A, B are **not independent** (i.e., dependent)
     - Pr(A AND B) = Pr(A) • Pr(B | A)
       = Pr(B) • Pr(A | B)
     - Pr(B | A) read as "probability of B occurring, given that A occurs," or more simply, "probability of B, given A"
       – By conditioning on A, we are "renormalizing" the sample space to be just A
       – Pr(B | A) is the fraction of B that is found within A

Definition of "Conditional Probability"

- **Conditional** probability definition
  - We said that in general
    - Pr(A AND B) = Pr(A) • Pr(B | A)
  - The conditional probability is last term, **Pr(B | A)**, so
    - Pr(B | A) = Pr(A AND B) / Pr(A) , Pr(A) $\neq$ 0
    - Pr(A | B) = Pr(A AND B) / Pr(B) , Pr(B) $\neq$ 0
  - These last equations define "conditional probability"
- We will see (later) that this product rule of conditional probabilities leads us to "Bayes' Theorem"

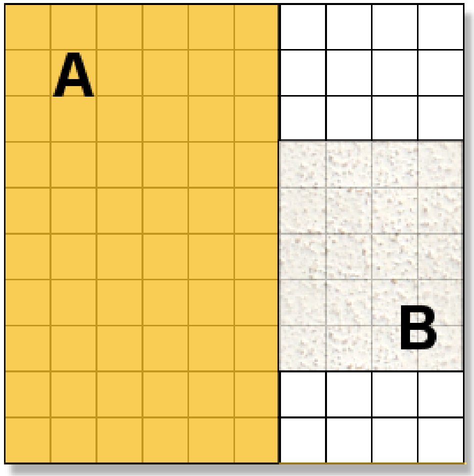
Do Not Confuse Independent and Disjoint

- If A, B are **mutually exclusive** (i.e., disjoint), then
  - Pr(A AND B) = 0
- If Pr(A AND B) = 0 then Pr(A AND B) ≠ Pr(A) • Pr(B) unless either Pr(A) or Pr(B) = 0
- When mutually exclusive, A and B are **not** independent
  - In fact, they are very **strongly dependent**
    - If one event occurs, other event cannot occur
      - If heads occurs on a coin toss, tails cannot occur
    - They simply are disjoint
    - On a Venn diagram, they do not overlap

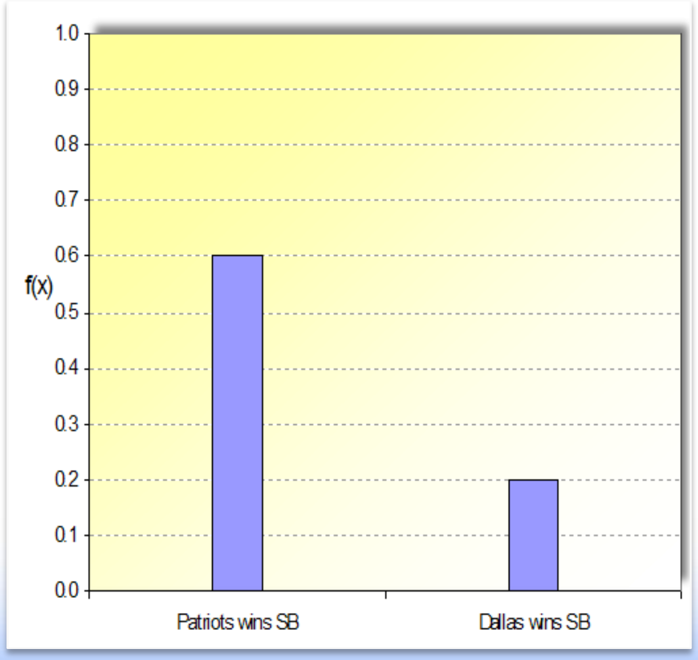


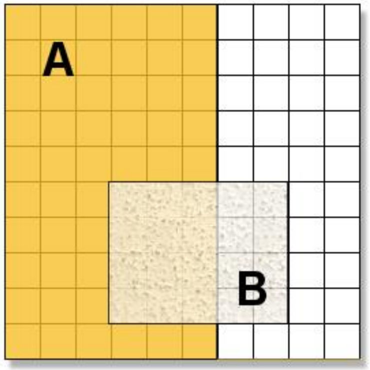
Independent versus Disjoint

- An example using disjoint events
  - If two events A and B are disjoint (mutually exclusive)
    - Pr(A AND B) = 0
    - If Pr(A) = 0.6 while Pr(B) = 0.2 then the Venn diagram is



Disjoint

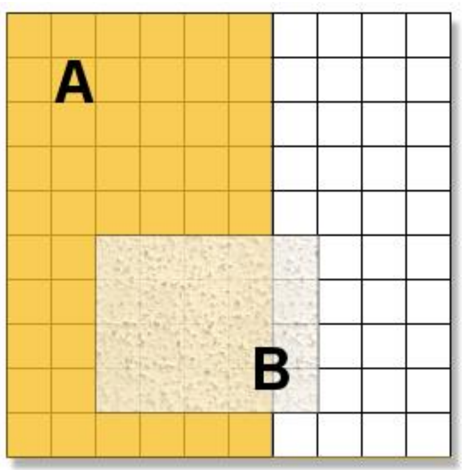




Pr(A AND B) = 0.12 *if* A, B were independent…

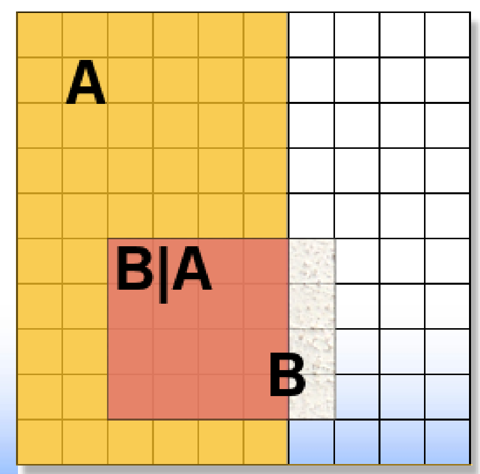
Independent versus Dependent

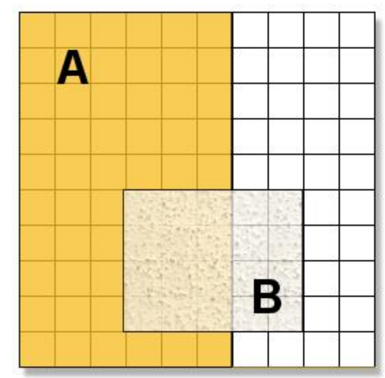
- An example using **dependent** events
  - If Pr(A) = 0.6, P(B) = 0.2, and Pr(A AND B) = 0.16
    - Then Pr(B | A) = 0.26667 since
    - Pr(A AND B) = Pr(A) • Pr(B | A)



A and B are *dependent*

Where is Pr(B|A) on the Venn diagram??
16 blocks/60 blocks = 0.26667

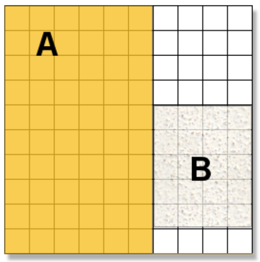


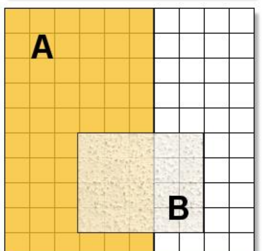


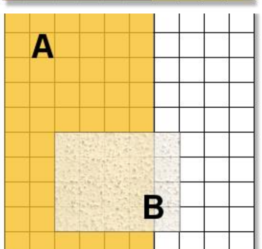
Pr(A AND B) = 0.12
***if* A, B were independent…**

Disjoint, Independent, Dependent Summary

- The table below summarized the probability rules when quantifying multiple events



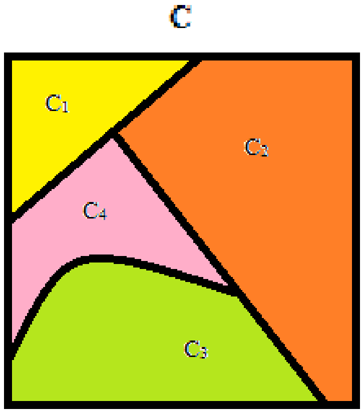




Case	Operation	Rule
Disjoint	OR	p(A OR B) = p(A)+p(B)
	AND	p(A AND B) = 0
Independent	OR	p(A OR B) = p(A)+p(B) - p(A AND B) ≈ p(A)+p(B) (rare event approx.)
	AND	p(A AND B) = p(A)p(B)
Dependent	OR	p(A OR B) = p(A)+p(B) - p(A AND B) ≈ p(A)+p(B) (rare event approx.)
	AND	p(A AND B) = p(A)p(B | A) = p(B)p(A | B)

Bayes' Theorem

Thomas Bayes



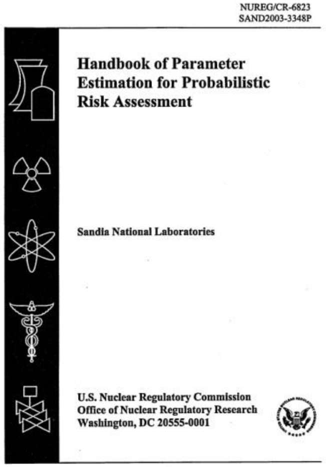
- A set of events $\{C_i\}$ is a **partition** of the sample space **C**
  - If all $\{C_i\}$s in **C** are mutually exclusive
    - Each pair is mutually exclusive…no overlap
  - And if union of $\{C_i\}$s is the entire sample space **C**
- Bayes' Theorem: If $\{C_i\}$ is a partition of the sample space,

$$\Pr(C_i \mid E) = \frac{\Pr(E \mid C_i)\Pr(C_i)}{\sum_j \Pr(E \mid C_j)\Pr(C_j)}$$

- Bottom term is Pr(E) (where E is the "evidence")

$$\Pr(E) = \sum_j \Pr(E \mid C_j)\Pr(C_j)$$

is called "Law of Total Probability"


NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages A-4 through A-12

Bayes' Theorem

- If we are calculating probability of **event C** where **evidence E** is available

  Pr(C | E) = Pr(C) Pr(E | C) / Pr(E)

- Terms in equation above have specified names

  Pr(C | E): Posterior probability (or posterior distribution)

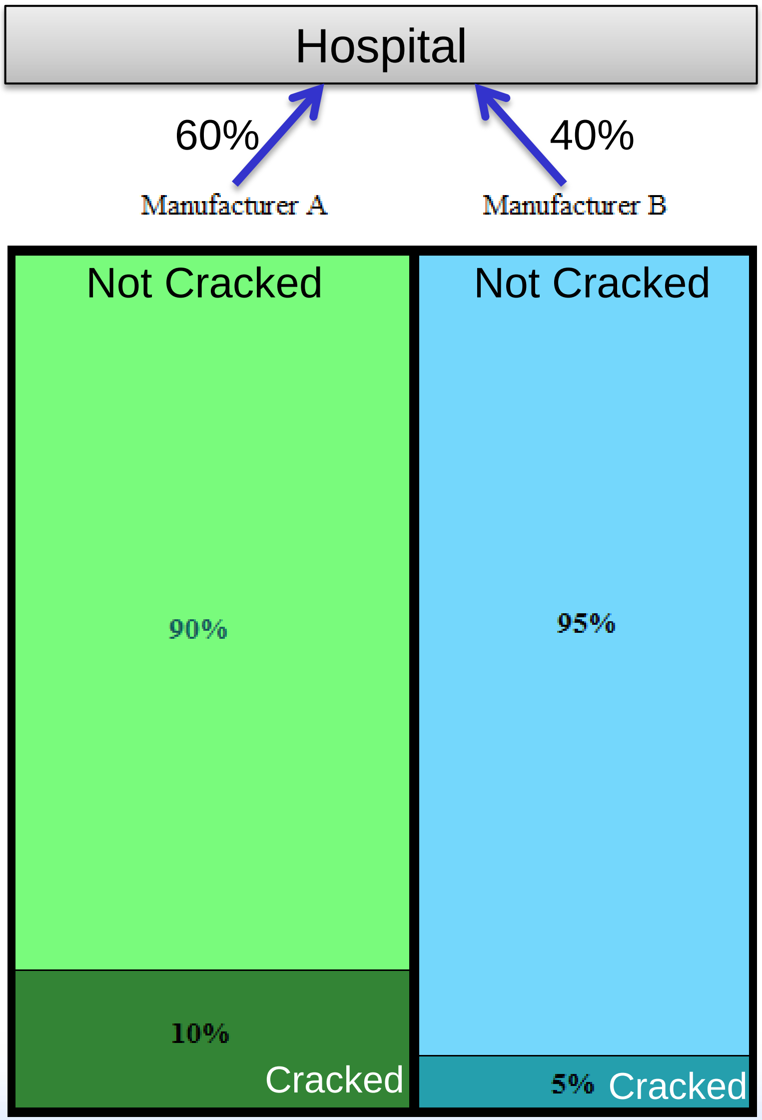
  Pr(C): Prior probability (or prior distribution)

  Pr(E | C): Probabilistic model, likelihood, or aleatory model

  Pr(E): Unconditional (marginal) probability of evidence

Bayes Example

- Tests for integrity are carried out on radiation sources by the manufacturer
- Hospital gets 60% of its sources from manufacturer A, the rest from manufacturer B
  - Manufacturer A results from its tests: 10% cracked
  - Manufacturer B results from its tests: 5% cracked



Bayes Example

- Incident report is later sent to the NRC regarding leak from cracked source at the hospital
  - What is the probability that cracked source came from manufacturer B?

→Pr(Manufacturer B | crack) =

Pr(Manufacturer B) Pr(crack | Manufacturer B) **/** Pr(crack)

= (0.4)(0.05) **/** [(0.6)(0.10) +(0.4)(0.05)]

= 0.02 / 0.08

= 0.25

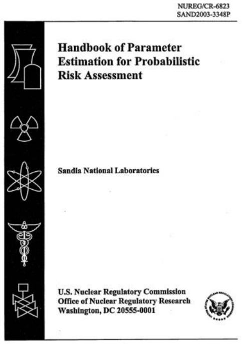
25% chance it came from Manufacturer B

75% chance it came from Manufacturer A

Discrete Probability Distributions

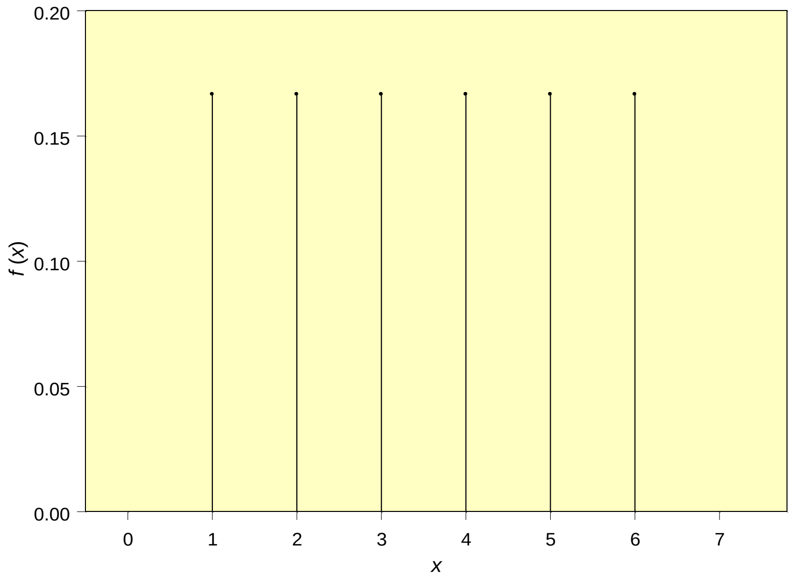
- Outcomes can be summarized by a **random variable** X, which takes possible real values x
- An "event" is then a set of possible values assumed by X
- Probabilities of events are calculated using X's distribution function (sometimes called probability mass function)
  - $f(x) = \Pr(X = x)$
  - Cumulative distribution is: $F(X \leq x_j) = \sum_{i=0 \, to \, j} f(x_i)$
- Facts about a discrete distribution:

$f(x_i) \geq 0$ and $\sum_{all \, i} f(x_i) = 1$

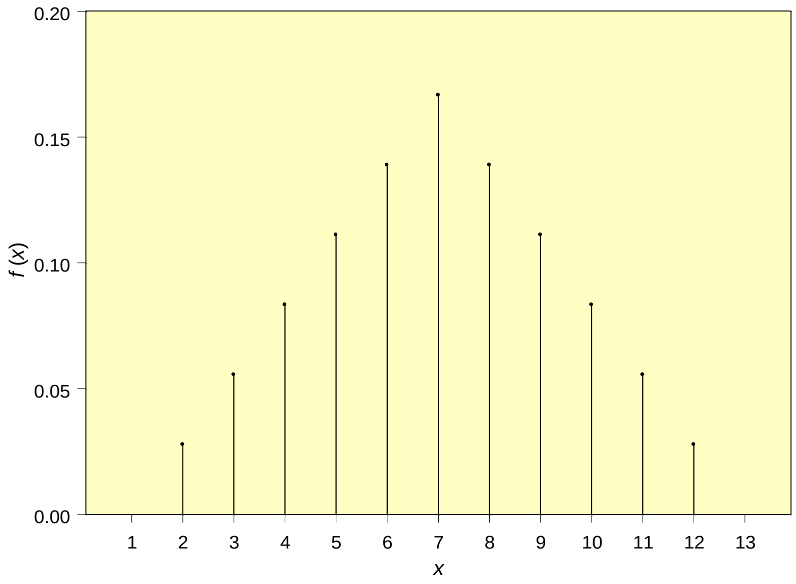


Page A-5

Examples: Number of Spots on Dice



Spots on One Die



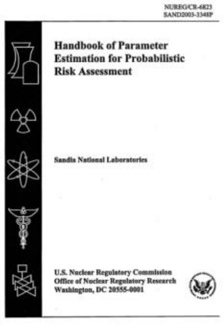
Total Spots on Two Dice

Idaho National Laboratory

Continuous Probability Distributions

- Random variable X takes on values in a continuous range, such as from 0 to ∞
- For any random variable X, $\Pr(a \le X \le b) = F(b) - F(a)$
  - where F is the cumulative distribution function (cdf)
- In most cases, can write this in terms of a **probability density function**, f(x), which is the derivative of F(x):

$$\Pr(a \le X \le b) = \int_a^b f(x)dx$$


NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

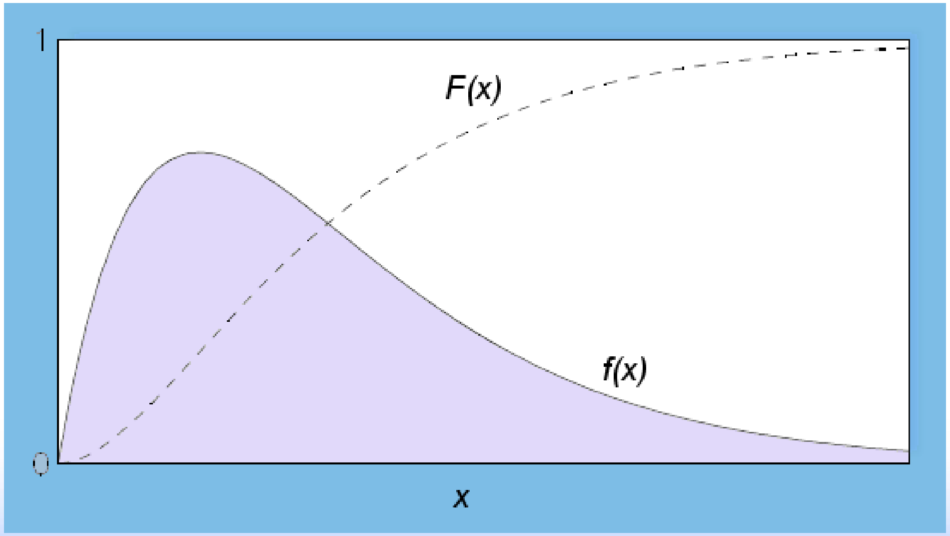
U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages A-5 through A-6

Continuous Probability Distributions

- Relations between pdf [f(x)] and CDF [F(x)]
  - $F(x) \equiv \Pr(X \le x) = \int_{-\infty}^{x} f(x')dx'$, has no units
  - $f(x) = dF(x)/dx$, has units x^{-1}
- Note, $\Pr(X = x) = 0$ for any specific value of x
  - But probability that X is in an **interval** is typically nonzero

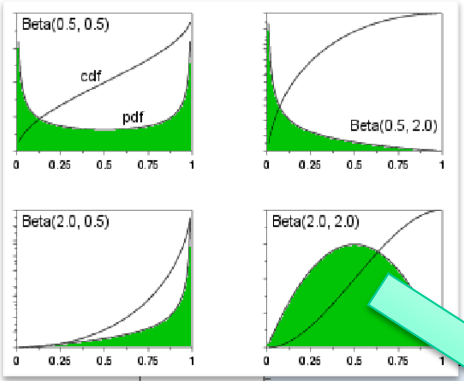
Note that graph scale is for F(x)



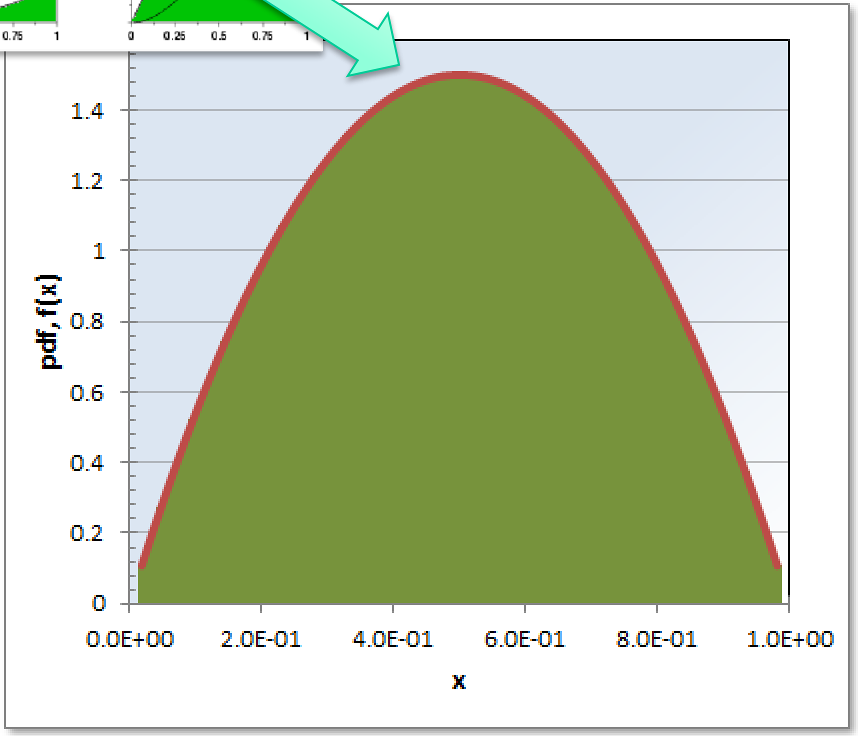
Continuous Probability Distributions

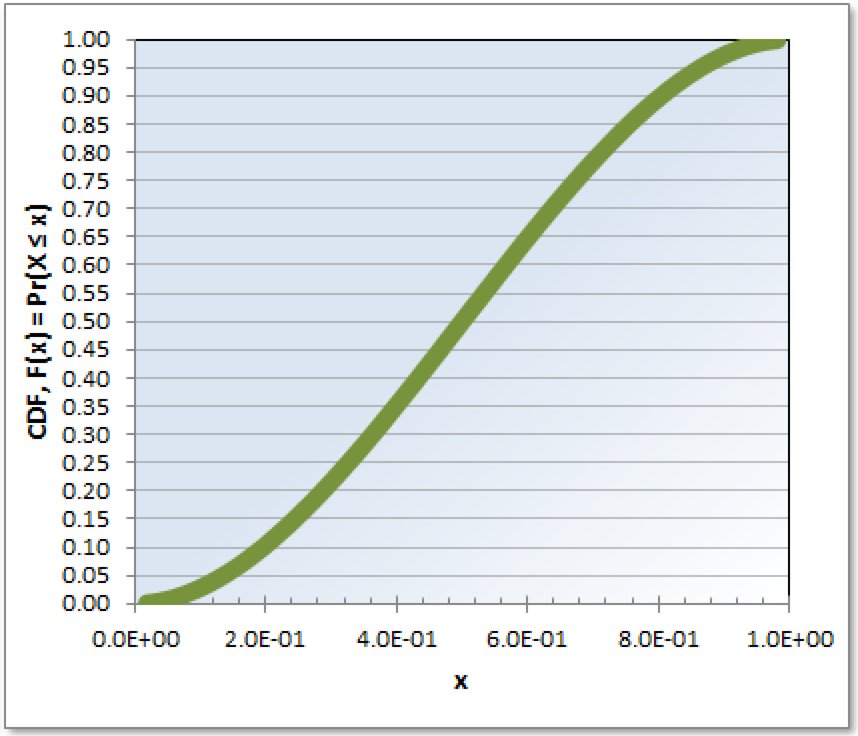
- Properties of probability density function, f(x)
  - $f(x) \geq 0$ for all x
  - $\int_{-\infty}^{\infty} f(x)dx = 1$
    - Sometimes we will use *improper* distributions in Bayesian inference, where this integral diverges

Continuous Probability Distributions

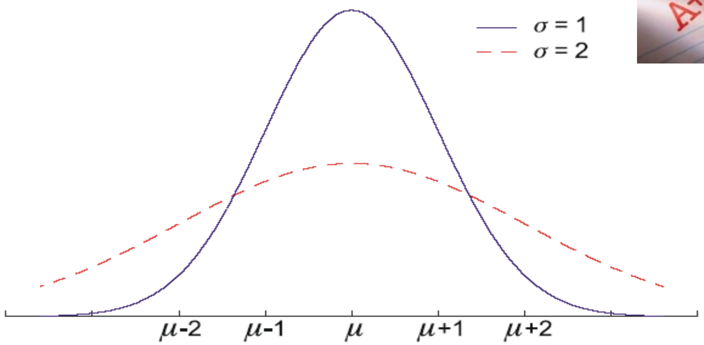


Beta(2, 2)





Normal Distribution



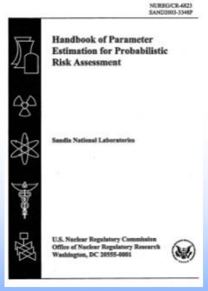
- Arises in many settings
  - Primary application in this course is as a “link” to the lognormal distribution
  - Density function in HOPE, page A-15
- If X has a **normal**(μ, σ^2) distribution, then

$$\Pr(X \le x) = \Pr\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Carl Friedrich Gauss

Φ is tabulated in many books, for example HOPE Table C-1

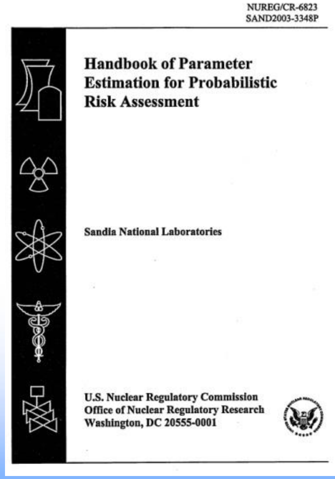
Can also use =NORMDIST(x, μ, σ, TRUE) in Excel



Pages A-15 through A-16

Moments and Percentiles

- The **mean**, or **expected value**, or expectation, of X is weighted average of the values of X
  - $E(X) = \sum_{x} x \Pr(X = x) = \sum_{x} x f(x)$ if X discrete
  - $E(X) = \int_{-\infty}^{\infty} x f(x) dx$ if X continuous


NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages A-8 through A-10

Idaho National Laboratory

Moments and Percentiles

- The **variance** is the weighted average of $[X - E(X)]^2$

$$\rightarrow \operatorname{var}(X) = \sum_x [x - E(X)]^2 f(x) \quad \text{if X discrete}$$

$$\rightarrow \operatorname{var}(X) = \int_{-\infty}^{\infty} [x - E(X)]^2 f(x)dx \quad \text{if X continuous}$$

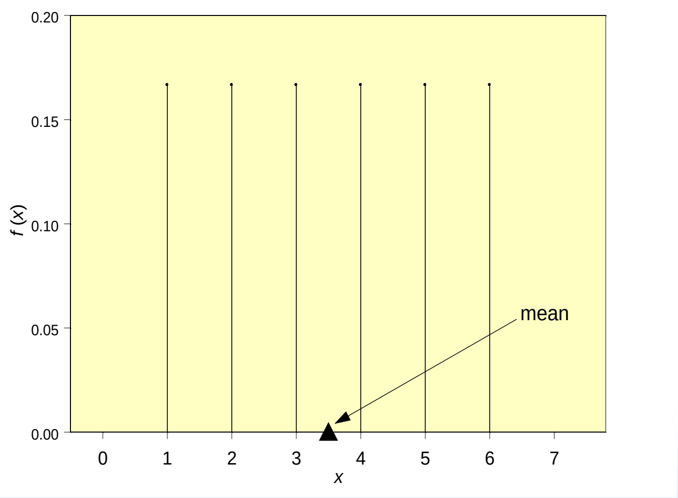
- The **standard deviation** is the square root of the variance (same units as x)

  σ = sqrt(Variance)

Idaho National Laboratory

Mean Example

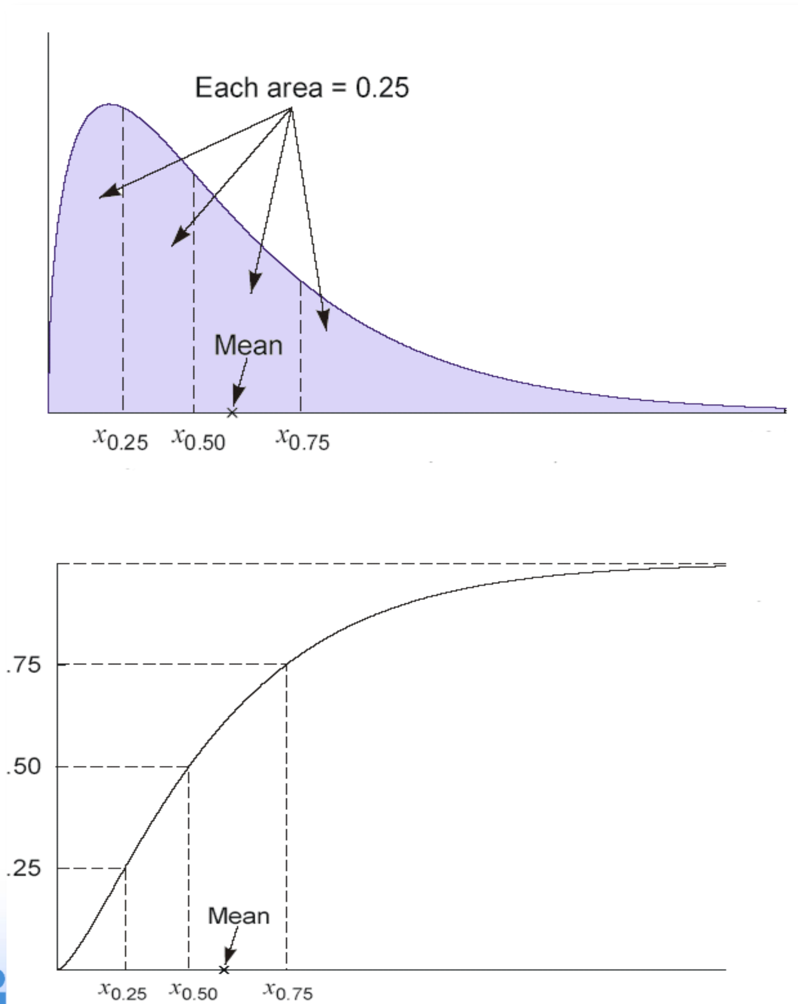
- We have the discrete distribution for a single die
  - What is the expected value?
  - Pr(X = x) = 1/6, x = 1, 2,…, 6
  - E[X] = 1(1/6) + 2(1/6) + 3(1/6) + 4(1/6) + 5(1/6) + 6(1/6)
    = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 1 = 3.5
  - Since discrete, we can not really get an outcome of 3.5. Possible outcomes are 1, 2, 3, 4, 5, or 6
  - In general, mean can be any value

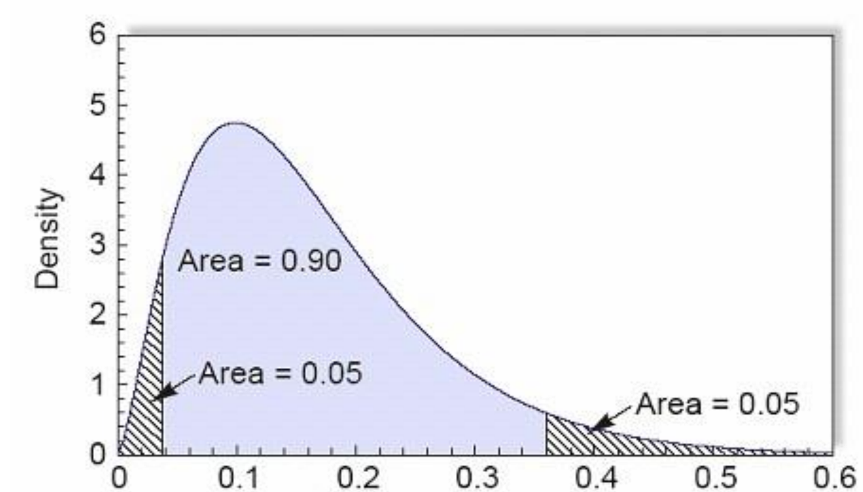


Moments and Percentiles

- The 95th **percentile**, denoted $x_{0.95}$, is the value such that
  $F(x_{0.95}) = 0.95$
- Similar definition for any number from 0 to 100 percent
- Special cases common in PRA include
  - **Median** = 50th percentile
  - **Upper bound** = 95^{th}
    - Should properly be called 95% upper bound
  - **Lower bound** = 5^{th}
    - Should properly be called the 5% lower bound
- For discrete distributions, exact percentile may not be observable value, as was the case for the mean

Moments and Percentiles





Moments and **Percentiles**

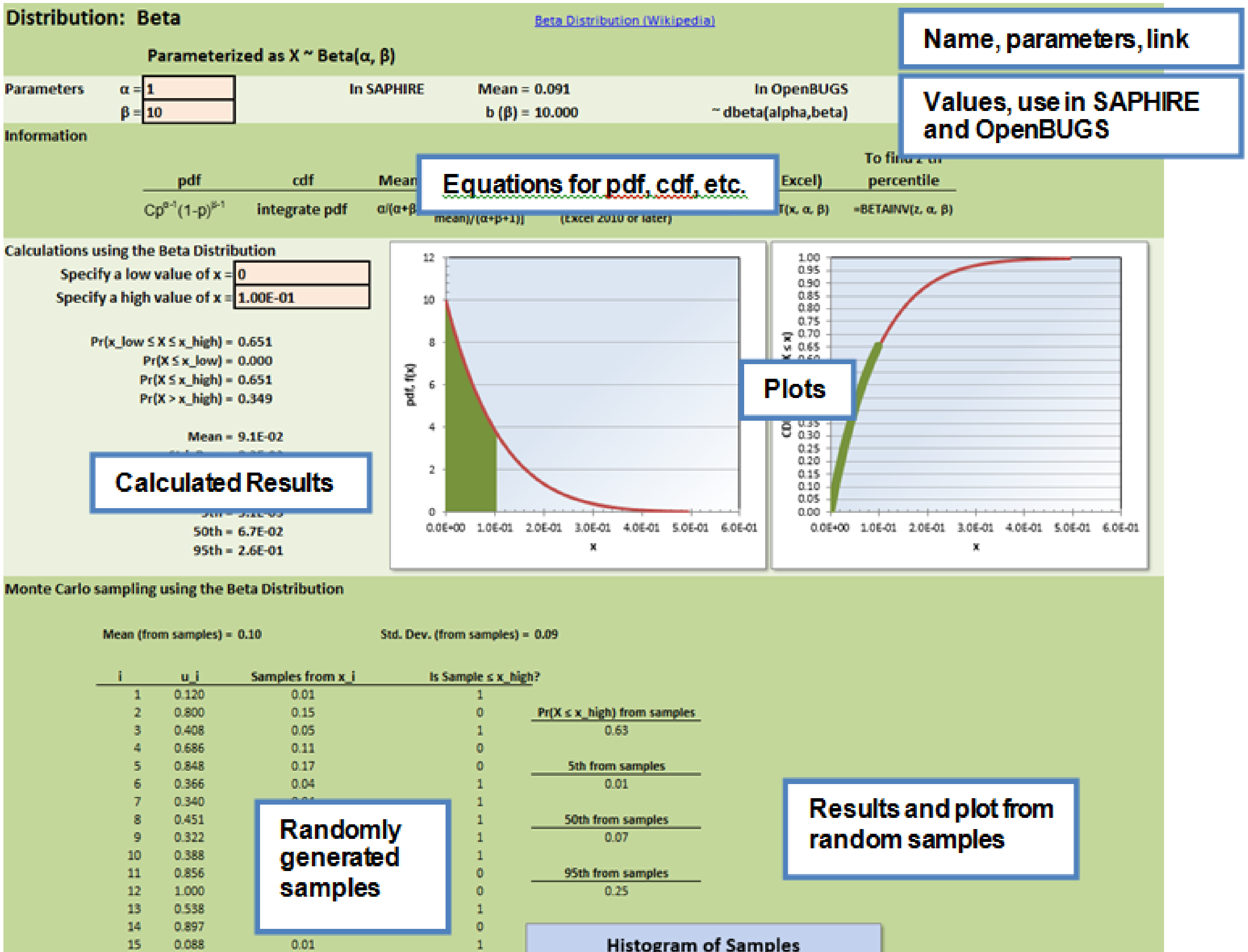
- Alternative language
  - The q **quantile** is the 100q percentile
- If a distribution is **positively skewed** (longer tail on the right), then mean is greater than median
  - E[X] > 50th percentile
  - Also, the mode (highest point on the pdf) is less than the 50th percentile for positively skewed distributions
    - Mode < Median < Mean

Distribution Summary Worksheet

- A tool for this course is the Excel spreadsheet titled "Distribution Summary Worksheets" (DSW)
- DSW is divided into two different types of worksheets
  - Bayesian inference (for conjugate cases)
  - Probability distributions
- The probability distributions include:

– Beta	Binomial	Exponential
– Gamma	Lognormal	Normal
– Poisson	Weibull	

DSW Distribution Example



Tips for Solving Problems

- Write what you have
  - Can you list the outcomes?
  - What events are relevant?
  - What is “fixed” and what is “random”?
  - What is the problem asking for?
  - What formulas relate to this question?
- Do not try to do everything in your head.
  - Use pencil and paper, and proceed step by step through the problem

Section 3: Introduction to Bayesian Inference

- Purpose
  - Present the subjectivist interpretation of probability, Bayesian inference for single-parameter problems, use of Excel functions, and applications to commonly encountered probabilistic models
- Objectives: Students will learn
  - Probability interpreted as a quantification of state of knowledge
  - Bayes' Theorem, Bayesian inference for parameter values in following:
    - Binomial, Poisson, and exponential (aleatory) models
    - Relation of these models to likelihood function in Bayes' Theorem
  - Use of discrete priors
  - Conjugate priors for Poisson, binomial, and exponential likelihoods
  - Formal priors for Poisson, binomial, and exponential data
  - Use of spreadsheets to update conjugate priors
  - Use of online RADS calculator for updating conjugate priors and nonconjugate lognormal priors

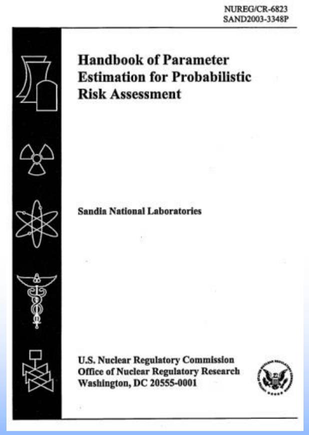
Elementary Bayesian Statistical Inference

- Topics to be covered
  - Subjective interpretation of probability
  - Bayes' Theorem as mechanism for Bayesian inference
  - Likelihood functions (aleatory models)
    - Binomial distribution
    - Poisson distribution
    - Exponential distribution
  - Prior distributions (epistemic uncertainty)
    - Discrete
    - Conjugate
    - Formal
    - Nonconjugate

George Apostolakis

Bayesian Statistical Inference

- General framework is covered in HOPE…
  - Page 6-2 (one-page introduction)
  - Section 6.2.2 for initiating events and running failures
    - Failure to run is also covered in Section 6.5
  - Section 6.3.2 for failures on demand
  - Section B.5 for summary of Bayesian estimation

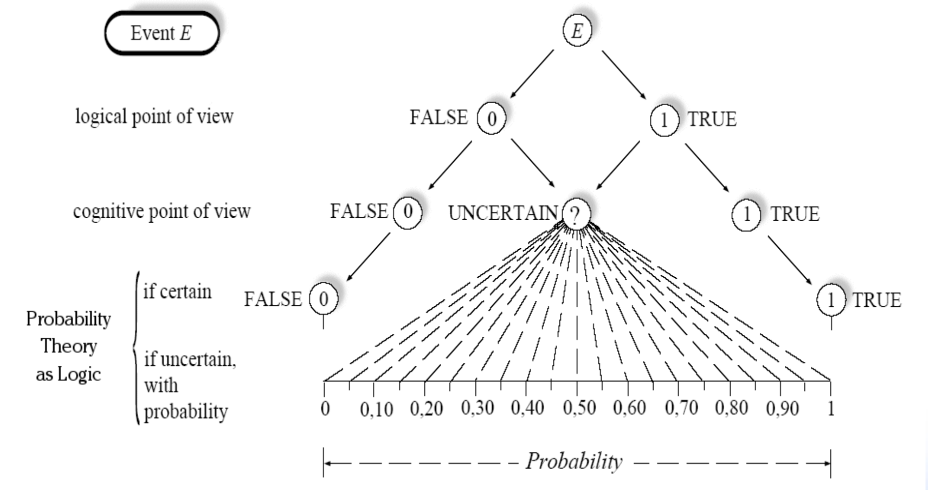


Motivation for Bayesian Inference

- Problems with frequentist inference
  - If data are sparse, estimates can be unrealistic (0 events in some cases)
  - No way to incorporate nonempirical "data"
    - For example, expert judgment
  - Difficult to propagate uncertainties (i.e., confidence intervals) through logic models
- Solution: A different interpretation of "probability"
  - Information about the parameter, beyond what is in the empirical data, is included in the estimate
  - Use Monte Carlo sampling to propagate uncertainties (expressed as probability distributions) through logic models

Subjective Probability

- In the **Bayesian**, or "subjectivist," approach, probability is a **quantification of state of knowledge**
  - It is used to describe the plausibility of an event
    - Plausibility – "Apparent validity"
  - A mechanism to encode **information**
- Note that, for "Bayes' Theorem,"
  - Thomas Bayes never wrote it
  - Laplace first used it in real problems



Adapted from "Probability and Measurement Uncertainty in Physics" by Giulio D'Agostini, December 1995

Bayesian Parameter Estimation

- The general procedure is:
  1. Begin with an **aleatory** model for the process of interest
  2. Specify a **prior distribution** for parameter(s) in this model, quantifying epistemic uncertainty, i.e., quantifying state of knowledge about the possible parameter values
  3. Collect **data**
  4. Obtain the posterior (i.e., updated) distribution for the parameter(s) of interest
  5. Check validity of model (P-501 and P-502 courses)
- We follow this process to make inferences, that is, to determine the probability that a model or hypothesis is reasonable, conditional on **all** available evidence

Common Aleatory Models in PRA

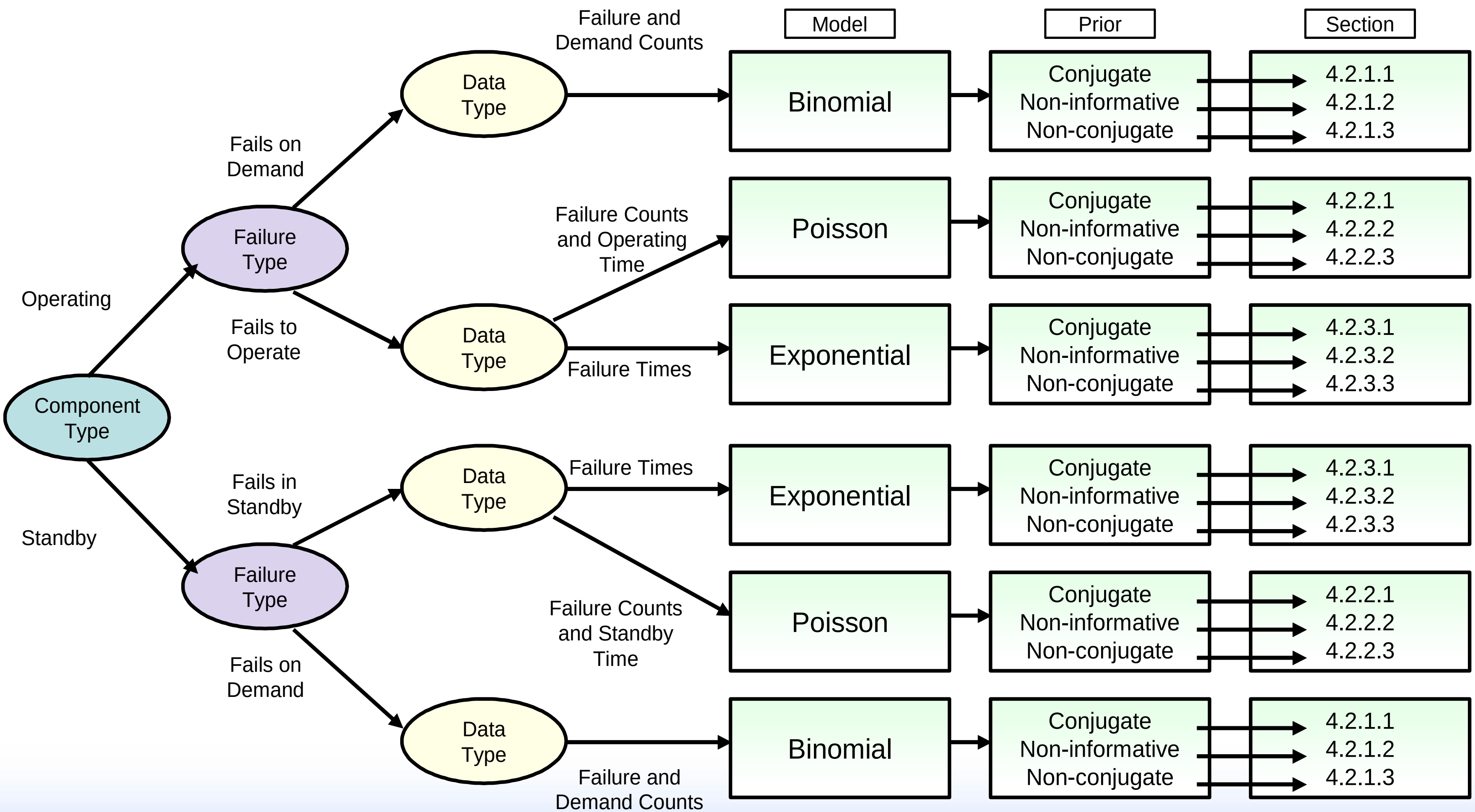
- Binomial
- Poisson
- Exponential
- We will use these models to “count” failures

What can we count?

- Examples of Poisson processes
  - Counting particles such as neutrons or photons
  - Number of (lit) lights failing
  - Arrival of customers
  - Large earthquakes
  - HTTP requests on a server
  - Loss of Offsite Power

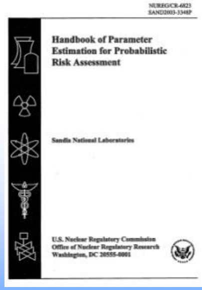
- Examples of Bernoulli (binomial) processes
  - Tossing a coin
  - Starting a car
  - Discrete random walk
  - Turning on a light
  - Birth of a child
  - Launching a rocket
  - Failures of a EDG

A “roadmap” (from NASA/SP-2009-569)



Binomial Distribution

- Commonly used model for **failure to change state**.
- Assumptions about the physical process
  1. On each demand, outcome is a failure with probability p (alternatively, a success with probability q=1 – p)
     - This p is the same on every demand
     - Called a Bernoulli trial
  2. Occurrences of failures on different demands are independent
- Form of the data
  - We observe a **random number of failures**, X, in a fixed or specified number of demands n


NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages 2-7 through 2-10, A-12 through A-13

Binomial Distribution: Functional Form

- Then the random variable X has a **binomial**(n, p) distribution:

$$f(x) = Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

for x = 0, 1, …, n

(x = number of failures)

Distribution parameters are p (unknown) and n (specified)

- For Bayesian inference, we write f(x) as f(x|p), called the likelihood function, sometimes denoted L(p)
  - Leads to frequentist maximum likelihood estimate (MLE) for p of x/n
- X is observed (failures) and p is unknown (focus of inference)

Idaho National Laboratory

Binomial Coefficient

- The binomial coefficient is defined as

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

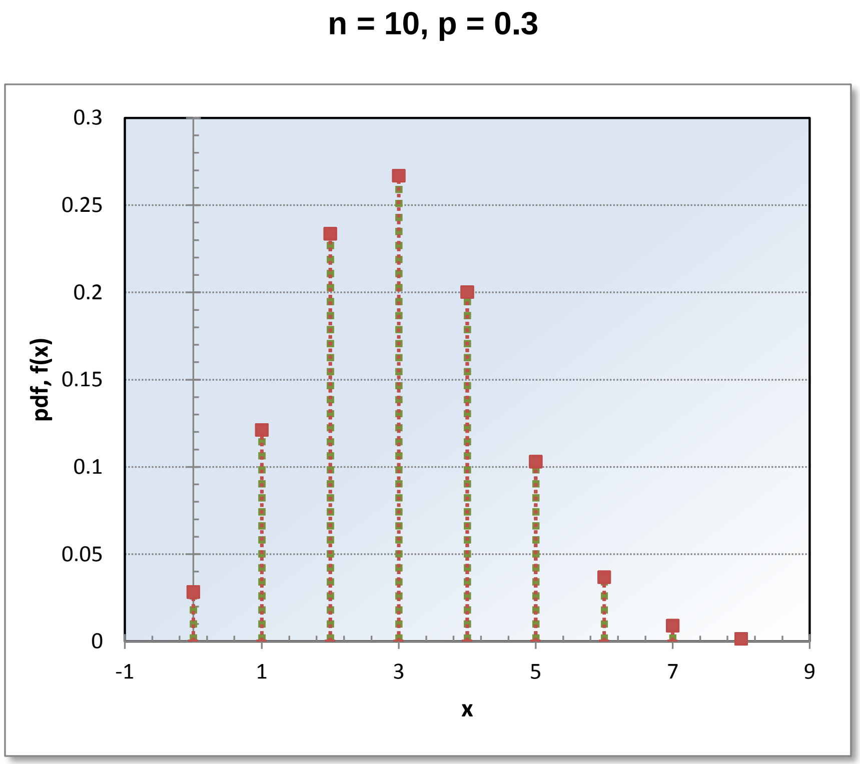
- Example

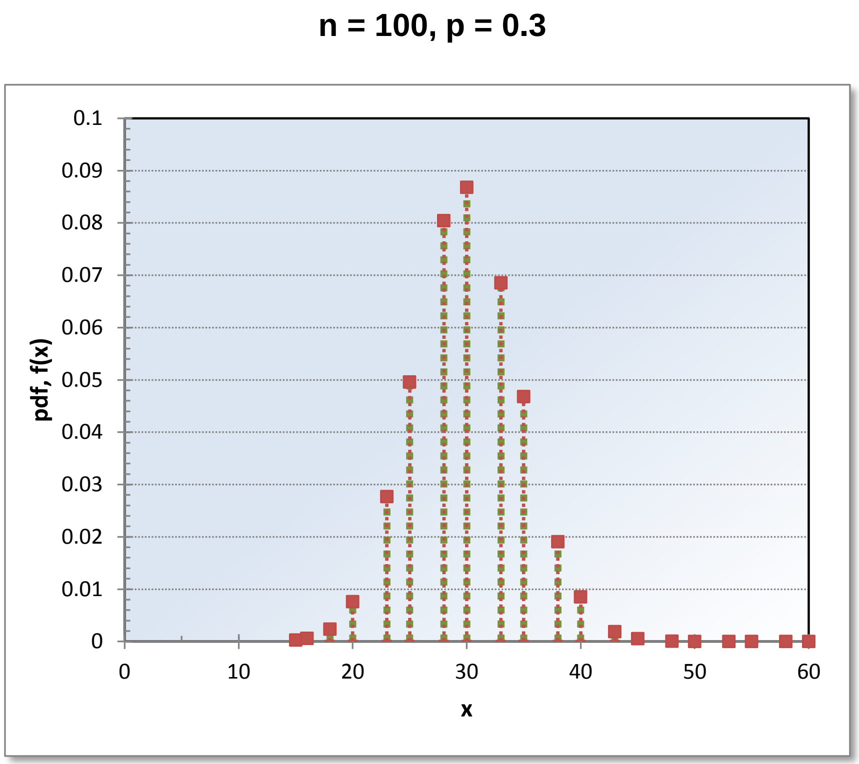
$$\binom{1}{1} = \frac{1!}{1!(1-1)!} = \frac{1}{1(0!)} = 1$$

  - Note that 0! = 1

$$\binom{3}{2} = \frac{3!}{2!(3-2)!} = \frac{6}{2(1!)} = 3$$

Binomial Distribution: Examples





Idaho National Laboratory

Binomial Distribution: Summary Measures

- Moments
  - Mean = np
  - Variance = np(1-p)
- Probability
  - To find probability of seeing **exactly x outcomes** in n number of trials [or Pr(X=x | n, p)] use

    =BINOMDIST(x, n, p, FALSE) in Excel

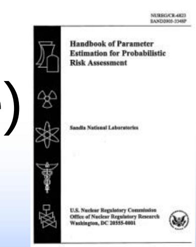
    - To find the **cumulative** of this use [Pr(0≤X≤x)]

    =BINOMDIST(x, n, p, TRUE) in Excel
  - To find approximate (100×z)th percentile of X use
    - **=CRITBINOM(n, p, z)** in Excel
    - Example, to find 95th

      =CRITBINOM(n, p, 0.95)

Poisson Distribution

- Most commonly used aleatory model for initiating events and failure to operate for specified time period
- Assumptions on the physical process
  1. Probability of event in short time period Δt is approximately $\lambda \times \Delta t$, for a constant λ
  2. Simultaneous events do not occur
  3. Occurrences of events in disjoint time periods are independent
- Form of the data
  – We observe a random number of events, X, in a fixed or specified time period t
- X is observed and λ is unknown (focus of inference)


Handbook of Parameter Estimation for Probabilistic Risk Assessment

Pages 2-2 through 2-7, A-13 through A-14

Poisson Distribution: Functional Form

- Then the random variable X has a **Poisson**(λt) distribution:

$$f(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

for x = 0, 1, 2, …
(x = number of events)

The distribution depends on one quantity, λt, (λ unknown, t specified)

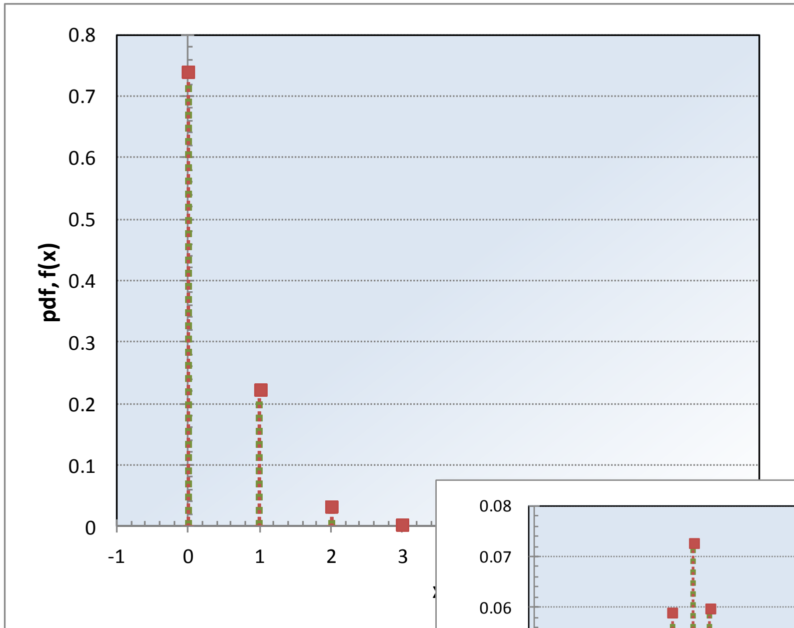
  – Therefore, product λt is sometimes written as μ (or even λ), and the distribution is called Poisson(μ)

- For Bayesian inference, we write f(x) as f(x|λ), called the <u>likelihood function</u>, sometimes denoted L(λ)

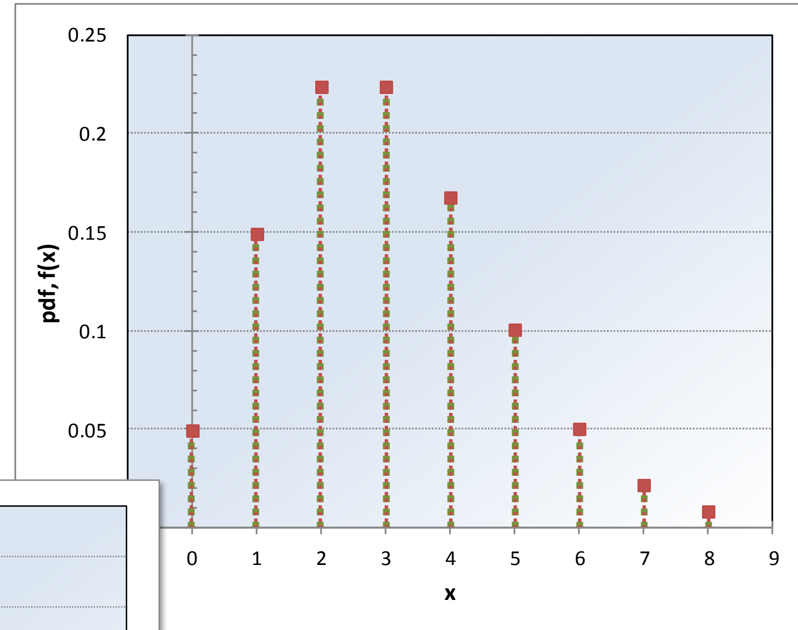
  – Leads to frequentist MLE for λ of x/t

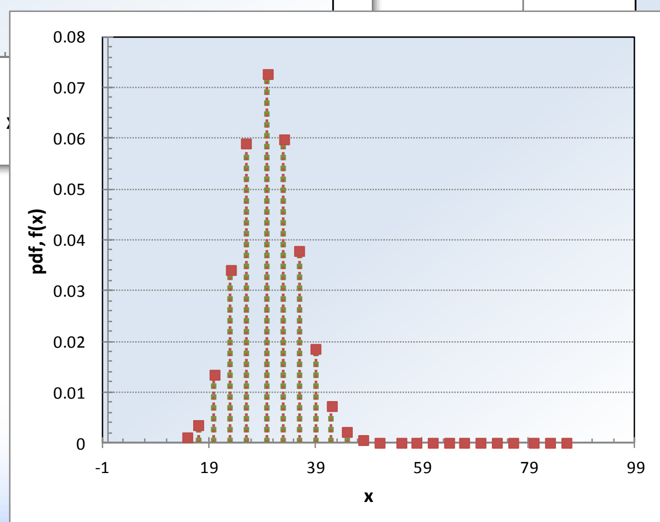
Poisson Distribution: Examples

mu = 0.3



mu = 3





mu = 30

Idaho National Laboratory

Poisson Distribution: Summary Measures

- Moments
  - Mean = $\lambda t = \mu$
  - Variance = $\lambda t = \mu$
- Probability
  - To find probability of seeing **exactly x** outcomes in time t [or $Pr(X=x \mid t, \lambda)$] use

    =POISSON(x, mean, FALSE) in Excel

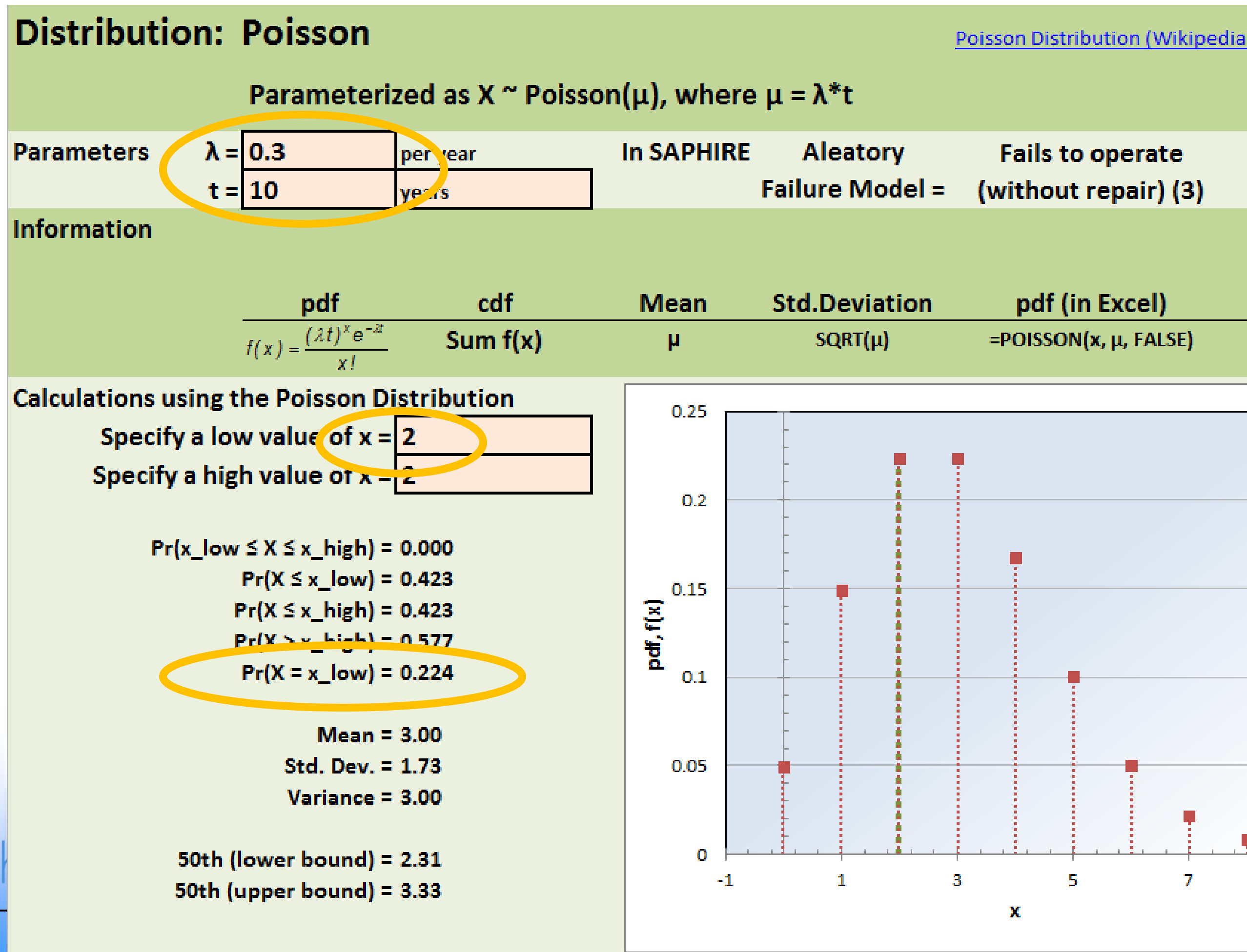
    - To find the cumulative of this use [$Pr(0 \le X \le x)$]

    =POISSON(x, mean, TRUE) in Excel

  - To find approximate ($100 \times z$)th percentile of X?
    - There is no “**CRITPOISSON**” in Excel, so need to look at the cumulative distribution to determine this

DSW Poisson Example #1

- To find Pr(**X=2** | t=10 yr, λ=0.3/yr)



DSW Poisson Example #2

- To find Pr(**X ≤ 2** | t=10 yr, λ=0.3/yr)

Distribution: Poisson

Poisson Distribution (Wikipedia)

Parameterized as X ~ Poisson(μ), where μ = λ*t

Parameters	λ =	0.3	per year	In SAPHIRE	Aleatory	Fails to operate
	t =	10	years		Failure Model =	(without repair) (3)

Information

pdf	cdf	Mean	Std.Deviation	pdf (in Excel)
$f(x)=\frac{(\lambda t)^x e^{-\lambda t}}{x!}$	Sum f(x)	μ	SQRT(μ)	=POISSON(x, μ, FALSE)

Calculations using the Poisson Distribution

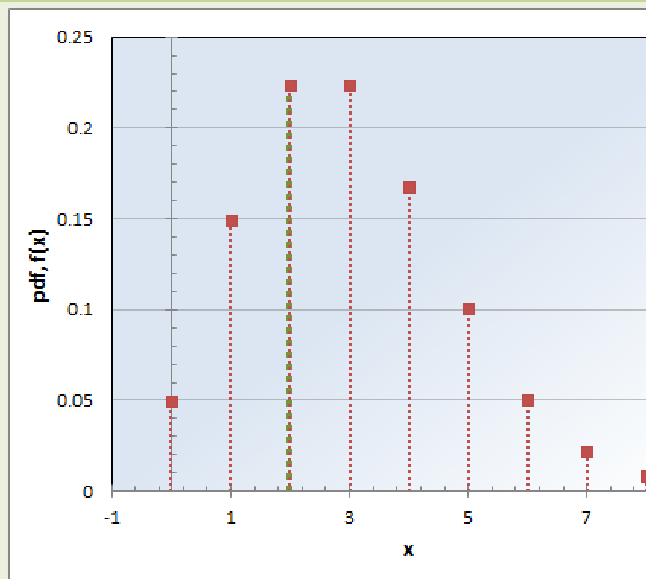
Specify a low value of x = 2

Specify a high value of x = 2

Pr(x_low ≤ X ≤ x_high) = 0.000
Pr(X ≤ x_low) = 0.423
Pr(X ≤ x_high) = 0.423
Pr(X > x_high) = 0.577
Pr(X = x_low) = 0.224

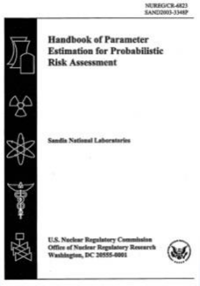
Mean = 3.00
Std. Dev. = 1.73
Variance = 3.00

50th (lower bound) = 2.31
50th (upper bound) = 3.33



Exponential Distribution

- A commonly used aleatory model for a time duration
  - Time to repair component, time to suppress fire, etc.
- Very simple (sometimes too simple)
- Setting: Watch something until an event of interest occurs, for example
  - Failure to run
  - Restoration of power
  - Suppression of fire, etc.
- Let **T** be a random variable representing time when event occurs



Pages 2-17, A-17 through A-18

Exponential Distribution: Genesis

- Assumption on the physical process

  1. For $t \geq 0$ and small Δt

  $$\Pr(T \leq t + \Delta t \mid T > t) \approx \lambda \times \Delta t \quad \text{(for a constant } \lambda\text{)}$$

- Interpretation
  – If the system is running at time t, probability that system will fail in next small time interval Δt is $\lambda \times \Delta t$, regardless of what t is.
  – That is, the system does not improve or degrade (i.e., age) as a function of time

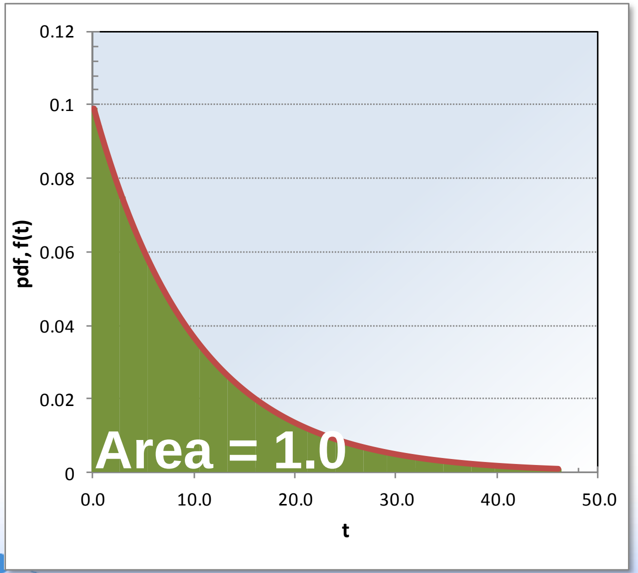
- Form of the data
  – We observe the **event times**, T_i, $i = 1, 2, \ldots, n$

- T is <u>observed</u>, λ is <u>unknown</u> (the focus of inference)

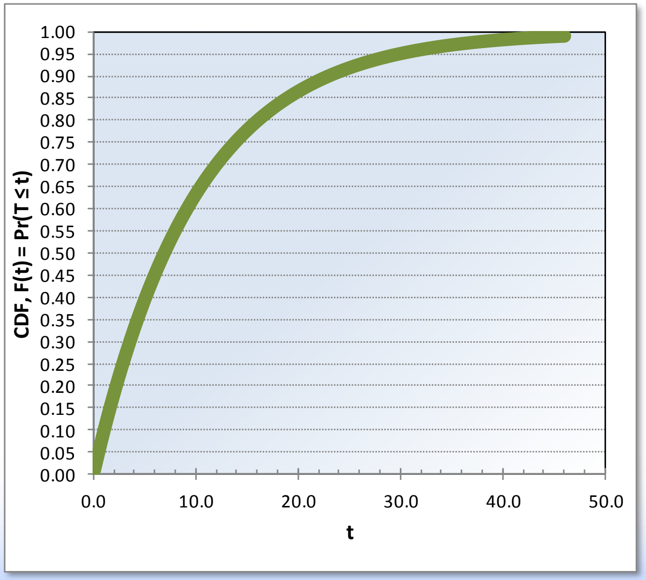
Exponential Distribution: Graphs

- Under the assumptions from the previous page
  - T has an **exponential**(λ) distribution

$$f(t) = \lambda e^{-\lambda t}$$ for t ≥ 0



$$F(t) = 1 - e^{-\lambda t}$$ for t ≥ 0



Idaho National Laboratory

Exponential Distribution

- Units
  - “λ t” is unitless
  - λ has units of 1/t (in PRA, usually per hour)
    - Initiating events are often per year
- Alternative parameterization in terms of $\mu = 1/\lambda$.
  - Just rewrite formulas in obvious way
  - Units of μ are units of t
    - Also known as “mean time to failure” (MTTF)
- Moments
  - Mean = $1/\lambda = \mu$
  - Variance = $1/\lambda^2 = \mu^2$

Exponential Distribution: Likelihood Function

- Likelihood function for n observed times, t_i

$$f(\lambda \mid t_1 \cdots t_n) = \prod_{i=1}^{n} f(t_i) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i)$$

- Leads to frequentist MLE for λ of $n / \Sigma t_i$

Idaho National Laboratory

Bayes' Theorem and Bayesian Parameter Estimation – Discrete Case

- Consider the unknown parameter λ (same idea if the parameter is p)
- For now, assume X (observed variable) is discrete, with $f(x|\lambda) = \Pr(X=x|\lambda)$
- Also assume that the unknown parameter λ can only take discrete values, $\lambda_1, \lambda_2, \ldots$
- Define discrete prior distribution, $\pi_{prior}(\lambda_i) = \Pr(\lambda = \lambda_i)$.
- By Bayes' Theorem,

$$\Pr(\lambda = \lambda_i \mid X = x) = \frac{\Pr(X = x \mid \lambda = \lambda_i)\Pr(\lambda = \lambda_i)}{\sum_j \Pr(X = x \mid \lambda = \lambda_j)\Pr(\lambda = \lambda_j)}$$

or

$$\pi_{post}(\lambda_i \mid x) = \frac{f(x \mid \lambda_i)\pi_{prior}(\lambda_i)}{\sum_j f(x \mid \lambda_j)\pi_{prior}(\lambda_j)}$$

- Denominator is a normalizing constant

Idaho National Laboratory

Bayes' Theorem and Bayesian Parameter Estimation – General Case

- Define $\pi_{prior}(\lambda)$, the prior pdf of λ
  - Discrete, continuous, or mixed
- Let $f(x \mid \lambda)$ be the pdf of X, dependent on λ
  - This is the **likelihood or aleatory model**
- The posterior pdf of λ is

$$\pi_{post}(\lambda \mid x) \propto f(x \mid \lambda)\pi_{prior}(\lambda)$$

- π_{post} **is proportional to** the product of the prior distribution and the likelihood function
  - π_{post} is what we put into our PRA basic events

Bayes' Theorem is Basis for Bayesian Updating of Data

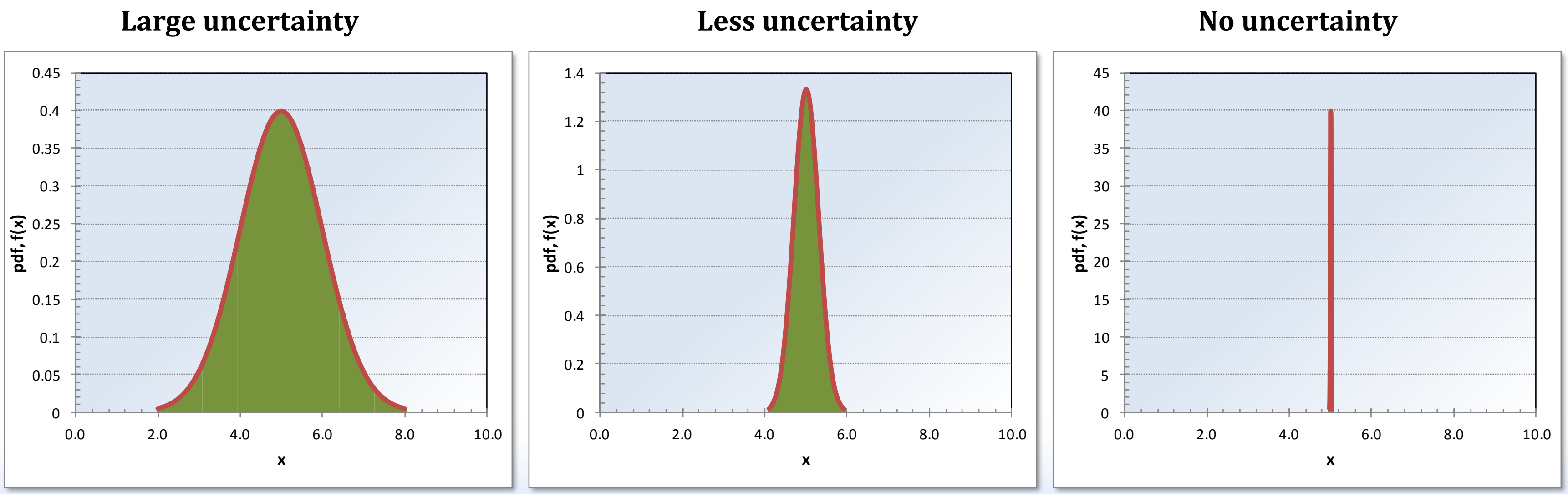
- Bayes' Theorem:

$$\pi_1(\theta | E) = \frac{L(E | \theta)\pi_0(\theta)}{\int L(E | \theta)\pi_0(\theta)\, d\theta}$$

- Where:
  - θ is parameter of interest
  - $\pi_o(\theta)$ is prior distribution
  - $L(E|\theta)$ is likelihood function
  - $\pi_1(\theta|E)$ is posterior distribution (updated estimate)

Probability Distributions Represent Uncertainty

- Usually used to represent state of knowledge of parameter values
  - Model assumptions typically addressed via sensitivity studies
- Probability distribution $\pi(\lambda)$ represents analyst's uncertainty about unknown value of λ
  - Note that λ may *not* be observable (for example, if a failure rate)



Bayes' Theorem and Bayesian Parameter Estimation – General Case

- If prior distribution is continuous
  - Parameter (e.g., λ) has values over a continuous range (and a continuum of possible values).
- Even though our **goal** is to obtain posterior distribution [$\pi_{post}(\lambda \mid x)$] for a parameter λ, need to remember
  - λ is assigned a prior distribution (representing information about possible values of λ)
    - Often convenient to summarize distribution by metrics such as mean, variance, or percentiles
  - Note that the distribution (of PRA parameters) is usually subjective, not a real, physical or empirical distribution
    - We do not "see" probabilities

Historical Use of Bayes Theorem

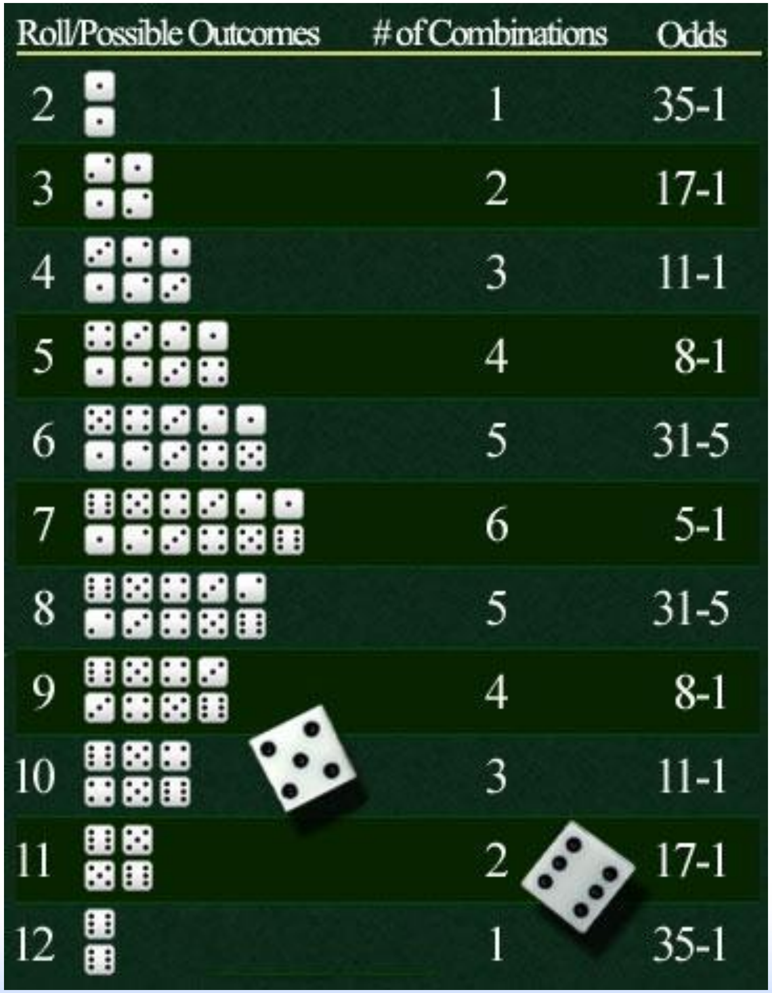
- Laplace, in 1774, used Bayesian methods to estimate the mass of Saturn
  - Assumed uniform prior density (what was known at the time)
  - Data consisted of mutual perturbations between Jupiter and Saturn
- His result was that he gave **odds** of 11,000 to 1 that his mass estimate* is not in error by more than 1%
  - What do odds of 11,000 to 1 imply?
    - That the point estimate $\pm$ 1% is the 99.99% credible interval
- 200 years of science increased his estimate by about 0.6%
  - Laplace would have won his bet (so far!)

Pierre Laplace

*1/3512th of solar mass = 5.7×10^{26} kg

Odds?

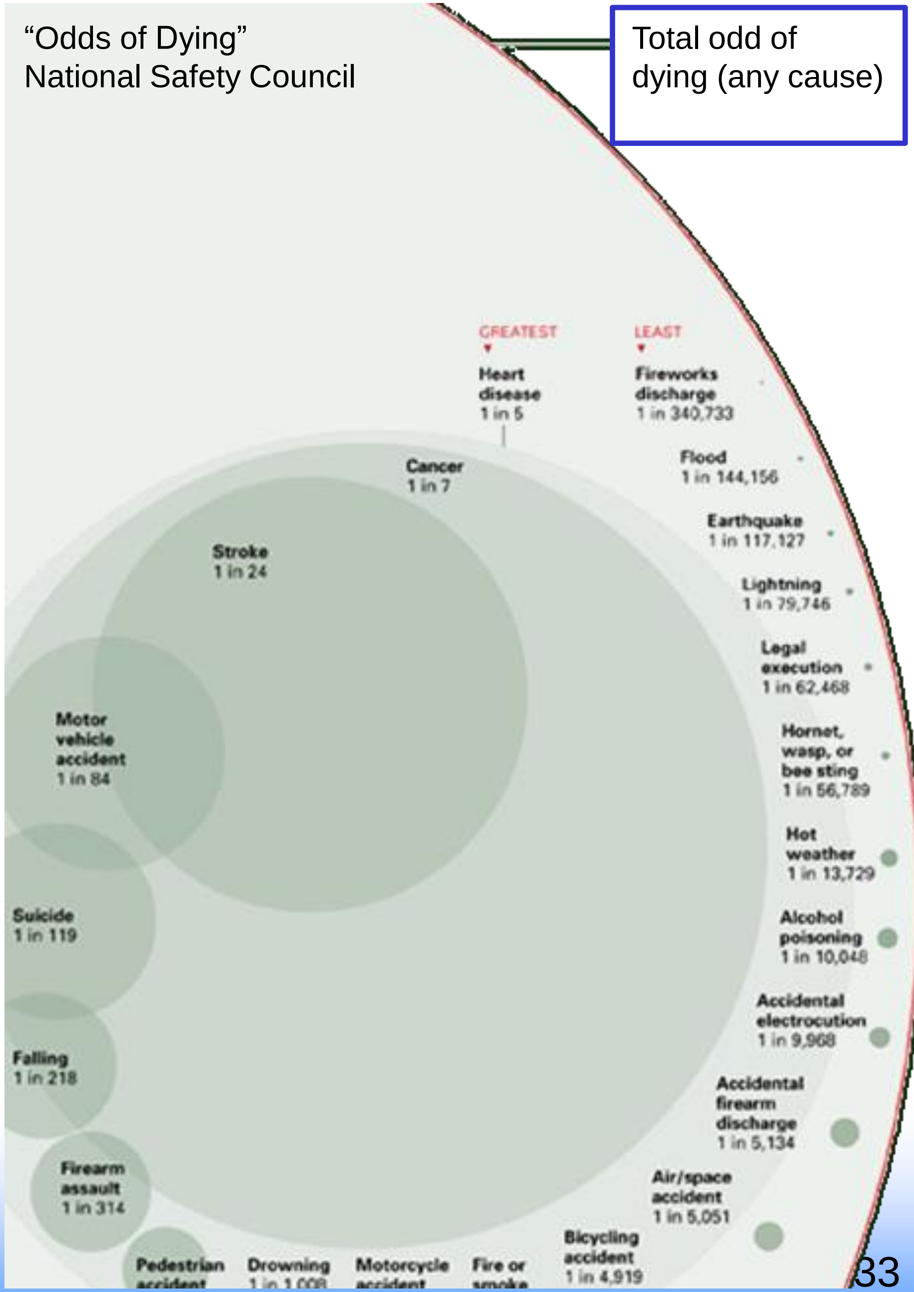
- Odds typically thought of as a “betting” term
  - Really a way to sneak-in probability to a discussion!
- Odds = $\dfrac{P(event)}{[1-P(event)]}$.
  - This is the odds for something



Roll/Possible Outcomes		# of Combinations	Odds
2		1	35-1
3		2	17-1
4		3	11-1
5		4	8-1
6		5	31-5
7		6	5-1
8		5	31-5
9		4	8-1
10		3	11-1
11		2	17-1
12		1	35-1

Odds?

- Odds typically thought of as a "betting" term
  - Really a way to sneak-in probability to a discussion!
- Odds = $\frac{P(event)}{[1-P(event)]}$.
  - This is the odds for something



Uses of Posterior Distribution

- For presentation purposes
  - Plot the posterior pdf
  - Give the posterior mean
  - Give a **Bayes credible interval**, an interval that contains most of the posterior probability (e.g. 90% or 95%)
    - 90% interval → <5^{th}, 95^{th}>
    - 95% interval → <2.5^{th}, 97.5^{th}>
- For risk assessment
  - Sample from the distribution of each parameter
  - Combine the results to obtain sample from Bayes distribution of end-state frequency

Prior Distributions

- We are going to examine three different situations related to **different types** of prior information
  - Discrete priors
  - Conjugate priors
    - Informative
    - Noninformative (or formal)
  - Nonconjugate priors

Discrete Prior Distributions

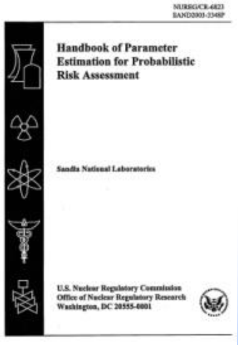
- These priors are easy to update with a spreadsheet (e.g., Excel)
  - Follows directly from Bayes' Theorem
    - For example, see "discrete prior.xls" in Excel folder

$$Pr(\lambda = \lambda_i \mid X = x) = \frac{Pr(X = x \mid \lambda = \lambda_i) Pr(\lambda = \lambda_i)}{\sum Pr(X = x \mid \lambda = \lambda_j) Pr(\lambda = \lambda_j)}$$

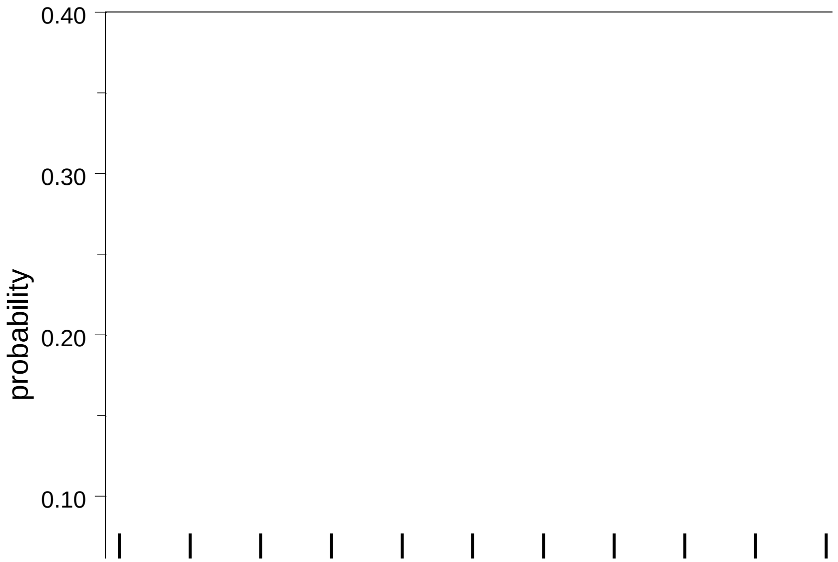
- Numerator in Bayes' Theorem is product of likelihood and prior probability of λ_i
  - To obtain full posterior probability, divide every such product by the sum of all such products
  - This makes the posterior probabilities (for all possible values of λ_i) sum to 1.0
- Discrete priors were once common in risk assessment
  - **Not used much these days**

Microsoft Excel

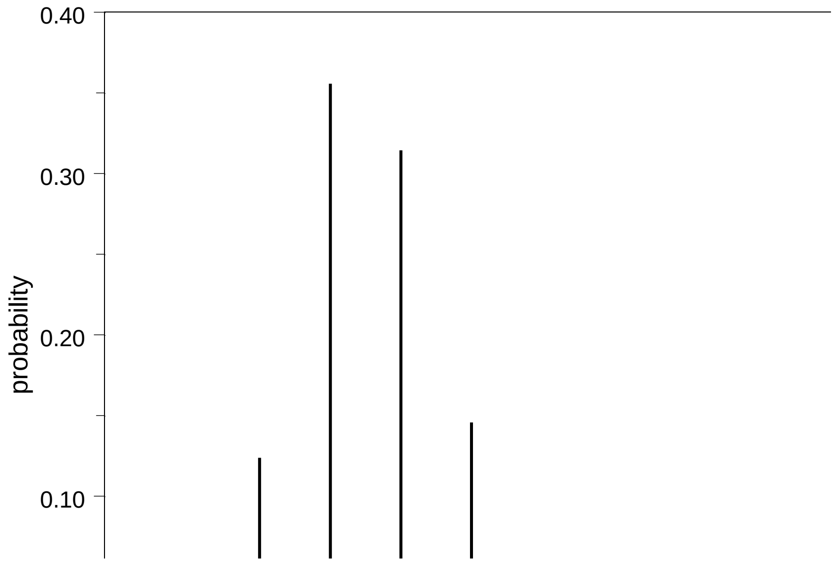
discrete prior.xls



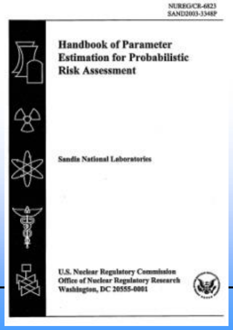
Example of Discrete Prior and Posterior



Coarse discrete prior for λ (events per year)



Posterior for λ, based on 10 observed events in 6 years



Page 6-10

CONJUGATE PRIORS

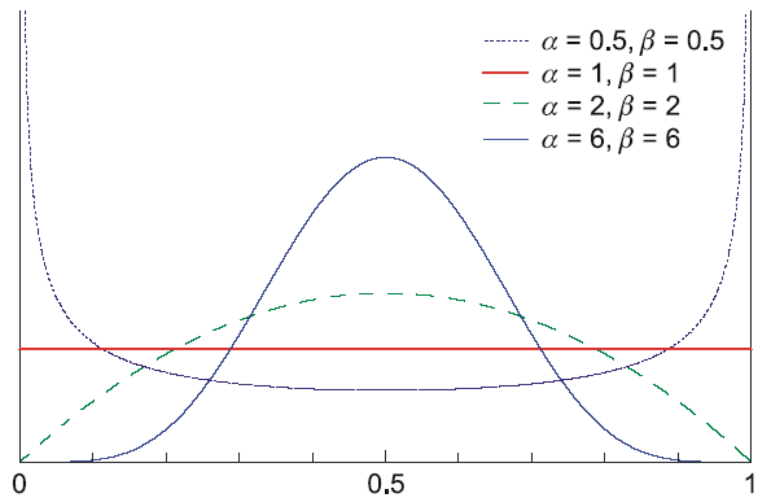
Conjugate Priors

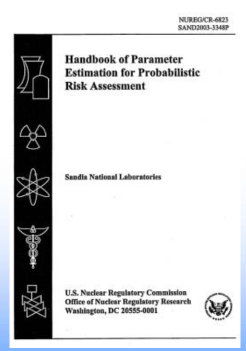
- Prior and posterior distribution have same functional form
  - Only distribution parameters change to reflect data incorporated via likelihood function
    - This means you can write down the posterior distribution with just arithmetic
    - Mathematically convenient (no integration)
  - Widely used in PRA (perhaps too widely)
- In this section, we will address conjugate priors for **three** aleatory models commonly used in PRA
  - Binomial distribution
  - Poisson distribution
  - Exponential distribution

Binomial Likelihood – Beta Conjugate Prior

- Facts about **beta(α, β)** distribution
  - beta(α, β) density
    - $f(p) = C\,p^{\alpha-1}(1-p)^{\beta-1}$
  - Mean: $\alpha / (\alpha + \beta)$
  - Variance: mean(1 – mean)/$(\alpha + \beta + 1)$
  - Percentiles from tables in HOPE, App. C
  - Easier and more accurate to use BETAINV in Excel:

  100p percentile = BETAINV(p, α, β)
  - SAPHIRE uses mean and β (called "b" by SAPHIRE)




NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

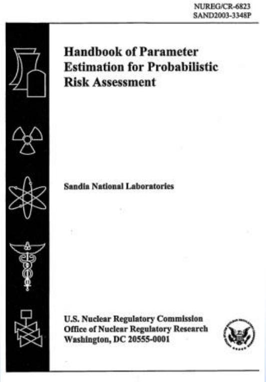
Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages A-21 through A-22, C-8 through C-13

Binomial Likelihood – Beta Conjugate Prior

- If X is **binomial**(n, p) and $g_{prior}(p)$ is **beta**(α_{prior}, β_{prior})
  - Then posterior distribution of p is
    - **beta**(α_{post}, β_{post})
      - $\alpha_{post} = \alpha_{prior} + x$ (x = # events)
      - $\beta_{post} = \beta_{prior} + n - x$ (n = total # trials)
    - α_{prior} is like prior number of events
    - β_{prior} is like prior number of successes
    - Posterior mean is $(\alpha_{prior} + x)/(\alpha_{prior} + \beta_{prior} + n)$
    - This is a weighted average of MLE, x/n, and prior mean, $\alpha_{prior} /(\alpha_{prior} + \beta_{prior})$



DSW Binomial-Beta Bayesian Example

- Assume our prior is ~Beta(1, 1)
- Assume we see 15 failures in 87 demands

Step 1 – Specify the Prior (Beta Tab)

Distribution: Beta — Beta Distribution (Wikipedia)

Parameterized as X ~ Beta(α, β)

Parameters α = 1, β = 1 — In SAPHIRE: Mean = 0.500, b (β) = 1.000 — In ~ dbeta(

Information

pdf	cdf	Mean	Std.Deviation	pdf (in Excel)	CDF (i
$Cp^{\alpha-1}(1-p)^{\beta-1}$	integrate pdf	α/(α+β)	sqrt[mean(1-mean)/(α+β+1)]	=BETA.DIST(x, α, β, FALSE) (Excel 2010 or later)	=BETAD

Calculations using the Beta Distribution

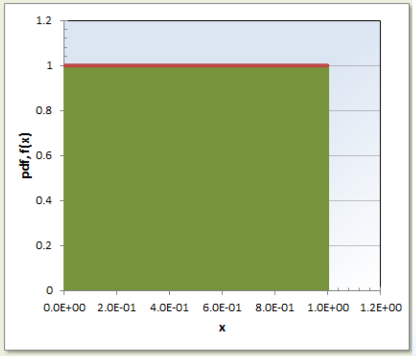
Specify a low value of x = 0

Specify a high value of x = 1.00E+00

Pr(x_low ≤ X ≤ x_high) = 1.000
Pr(X ≤ x_low) = 0.000
Pr(X ≤ x_high) = 1.000
Pr(X > x_high) = 0.000

Mean = 5.0E-01
Std. Dev. = 2.9E-01
Variance = 8.3E-02

5th = 5.0E-02
50th = 5.0E-01
95th = 9.5E-01



Step 2 – Specify Data (Bayes-Binomial)

Binomial Data

This worksheet performs conjugate Bayesian updating calcul

For Binomal Data, the conjugate prior is Beta(α, β)

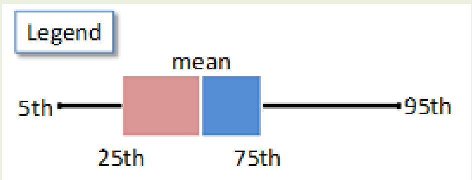
		Parameters		
Informed Prior		α =	1	Beta(α, β), where α and β are specif
		β =	1	
Jeffreys Prior		α =	0.5	Beta(½, ½)
		β =	0.5	
Constrained Non-informative (CNI)	Mean	α =	0.5	Beta(α, β)
	1.00E-01	β =	4.5	
Data Observed		x =	15	(number of failures)
		n =	87	(number of demands)

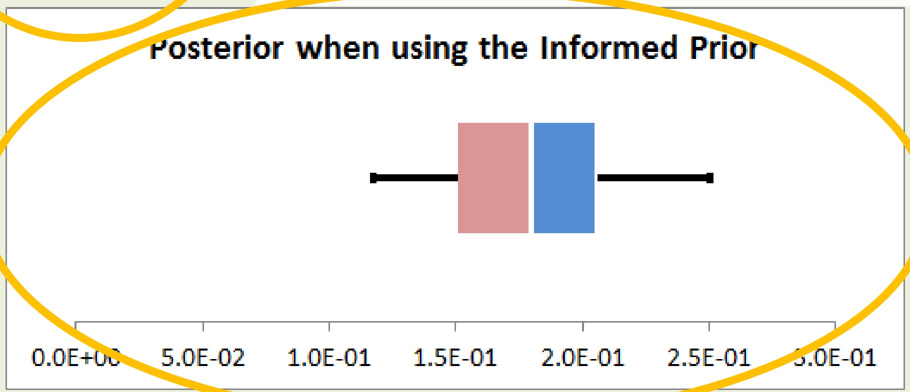
DSW Binomial-Beta Bayesian Example

- Read the results

Posterior Results

	Informed	Jeffreys	CNI
α =	16	15.5	15.5
β =	73	72.5	76.5
Posterior Distribution =	Beta(16, 73)	Beta(15.5, 72.5)	Beta(15.5, 76.5)
5th =	1.2E-01	1.1E-01	1.1E-01
25th =	1.5E-01	1.5E-01	1.4E-01
mean =	1.8E-01	1.8E-01	1.7E-01
75th =	2.1E-01	2.0E-01	1.9E-01
95th =	2.5E-01	2.5E-01	2.4E-01





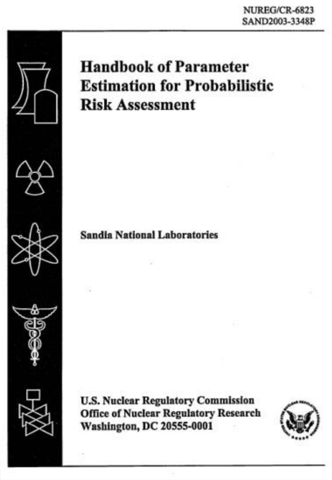
Poisson Likelihood – Gamma Conjugate Prior

- Facts about **gamma(α, β)** distribution, see HOPE
  - gamma(α, β) density is $g(\lambda) = C\,\lambda^{\alpha-1}e^{-\lambda\beta}$
  - mean = α / β
  - variance = α / β^2
  - 100p percentile = GAMMAINV(p, α, 1/β) or GAMMAINV(p, α, 1)/β
  - Excel uses 1/ β instead of β as the second parameter

Idaho National Laboratory

Poisson Likelihood – Gamma Conjugate Prior

- SAPHIRE uses mean and α (called “r” by SAPHIRE)
- Example gamma distributions:

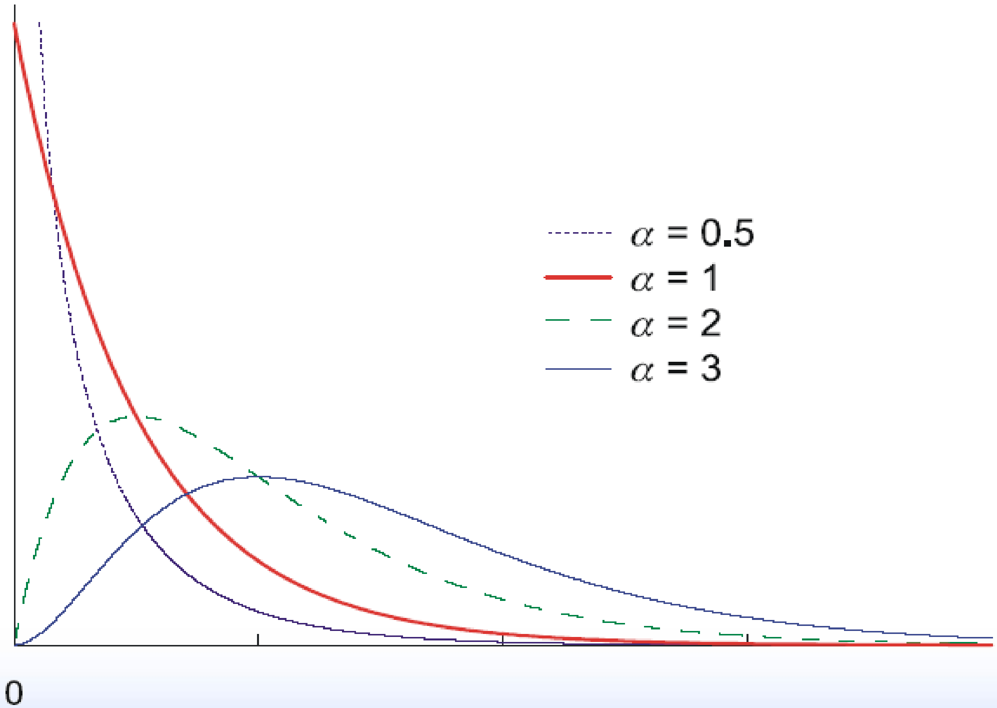

NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

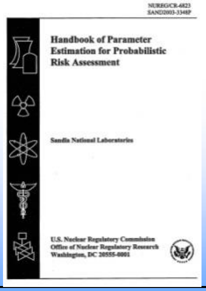
U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages A-18 through A-20



Poisson Likelihood – Gamma Conjugate Prior

- If X is **Poisson**(λ t) and $g_{prior}(\lambda)$ is **gamma**(α_{prior} , β_{prior}), then posterior distribution of λ is
    - **gamma**(α_{post} , β_{post})
        - $\alpha_{post} = \alpha_{prior} + x$ (x = # events)
        - $\beta_{post} = \beta_{prior} + t$ (t = observation time)
    - α_{prior} is like prior number of events
    - β_{prior} is like prior observation time
    - Therefore, posterior mean = $(\alpha_{prior} + x)/(\beta_{prior} + t)$
        - Again, a weighted average of MLE, x/t, and prior mean, $\alpha_{prior} / \beta_{prior}$


Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

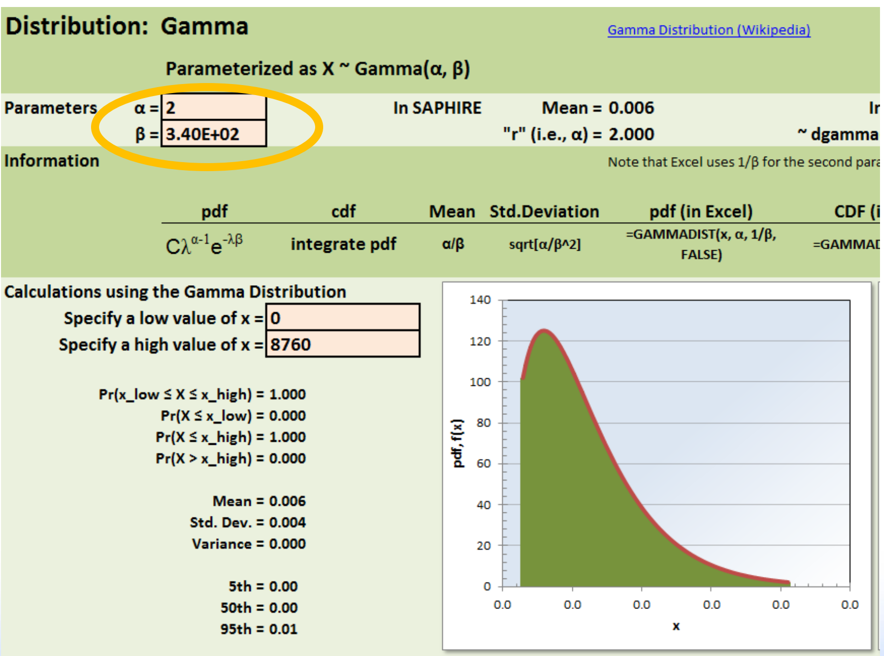
U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research

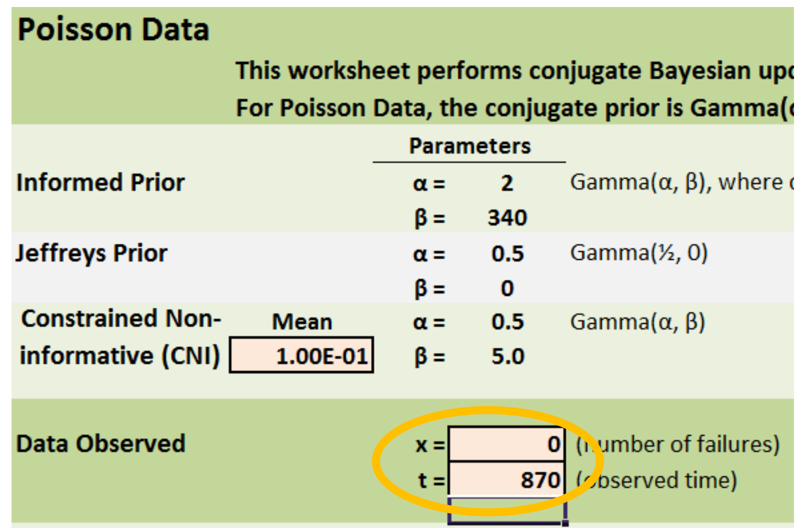
Pages 6-12, 6-13

DSW Poisson-Gamma Bayesian Example

- Assume our prior is ~Gamma(2, 340 hr)
- Assume we see 0 failures in 870 hours

Step 1 – Specify the Prior (Gamma Tab) Step 2 – Specify Data (Bayes-Poisson)

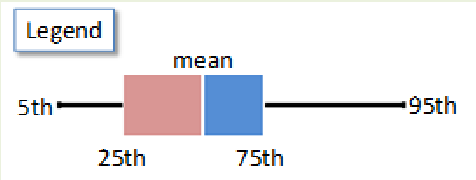


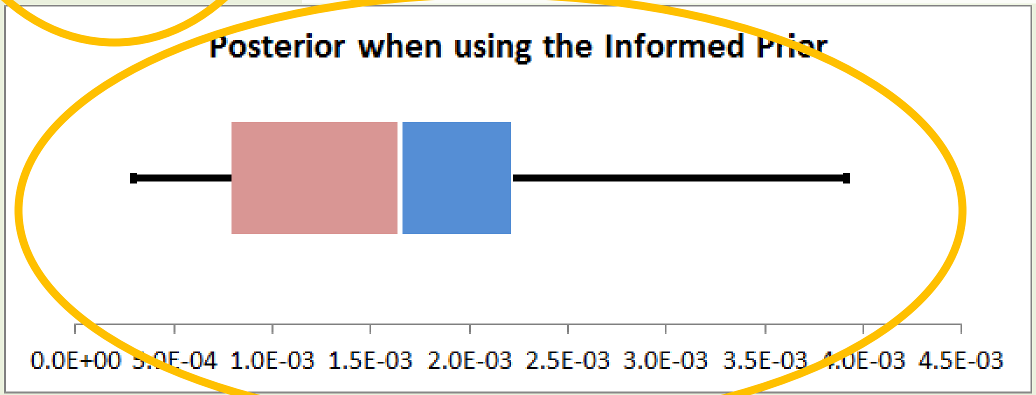


DSW Poisson-Gamma Bayesian Example

- Read the results

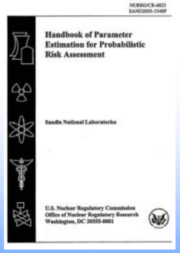
Posterior Results	Informed	Jeffreys	CNI
α =	2	0.5	0.5
β =	1210	870	875.0
Posterior Distribution =	Gamma(2, 1210)	Gamma(0.5, 870)	Gamma(0.5, 875)
5th =	2.9E-04	2.3E-06	2.2E-06
25th =	7.9E-04	5.8E-05	5.8E-05
mean =	1.7E-03	5.7E-04	5.7E-04
75th =	2.2E-03	7.6E-04	7.6E-04
95th =	3.9E-03	2.2E-03	2.2E-03





Exponential Likelihood – Gamma Conjugate Prior

- If $T_1, \ldots, T_n$ are independent observations from **exponential**(λ) distribution and $g_{prior}(\lambda)$ is **gamma**(α_{prior}, β_{prior}), then posterior distribution of λ is
    - **gamma**(α_{post}, β_{post})
        - $\alpha_{post} = \alpha_{prior} + n$ (n = # events)
        - $\beta_{post} = \beta_{prior} + \Sigma t_i$ (t_i = observed times of n events)
- Again, for a **gamma(α, β)** distribution
    - mean = α / β
    - variance = α / β^2
    - 100p percentile = GAMMAINV(p, α, 1/β)


Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research

Page 6-61

DSW Exponential-Gamma Bayesian Example

- Assume our prior is ~Gamma(2, 340 hr)
- Assume we tested five components and saw times-to-failure of: 205, 100, 760, 450, 1100 hours

Step 1 – Specify the Prior (Gamma Tab)

Distribution: Gamma

Gamma Distribution (Wikipedia)

Parameterized as X ~ Gamma(α, β)

Parameters α = 2
β = 3.40E+02

In SAPHIRE Mean = 0.006
"r" (i.e., α) = 2.000

~ dgamma

Information

Note that Excel uses 1/β for the second para

pdf	cdf	Mean	Std.Deviation	pdf (in Excel)	CDF (i
$C\lambda^{\alpha-1}e^{-\lambda\beta}$	integrate pdf	α/β	sqrt[α/β^2]	=GAMMADIST(x, α, 1/β, FALSE)	=GAMMAD

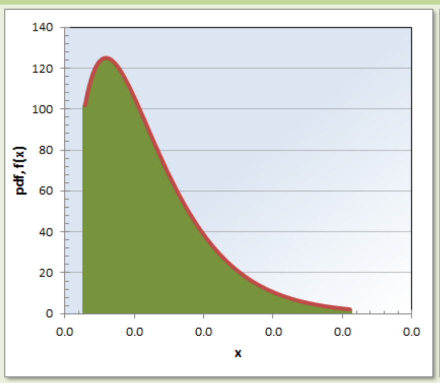
Calculations using the Gamma Distribution

Specify a low value of x = 0
Specify a high value of x = 8760

Pr(x_low ≤ X ≤ x_high) = 1.000
Pr(X ≤ x_low) = 0.000
Pr(X ≤ x_high) = 1.000
Pr(X > x_high) = 0.000

Mean = 0.006
Std. Dev. = 0.004
Variance = 0.000

5th = 0.00
50th = 0.00
95th = 0.01



Step 2 – Specify Data (Bayes-Exponential)

Exponential Data

This worksheet performs conjugate Bayesian updating calculation
For Exponential Data, the conjugate prior is Gamma(α, β)

	Parameters		
Informed Prior	α =	2	Gamma(α, β), where α and β are specifi
	β =	340	
Jeffreys Prior	α =	0	Gamma(0, 0)
	β =	0	

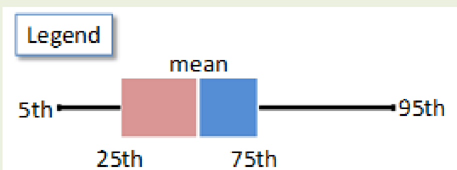
Data Observed number of times = 5

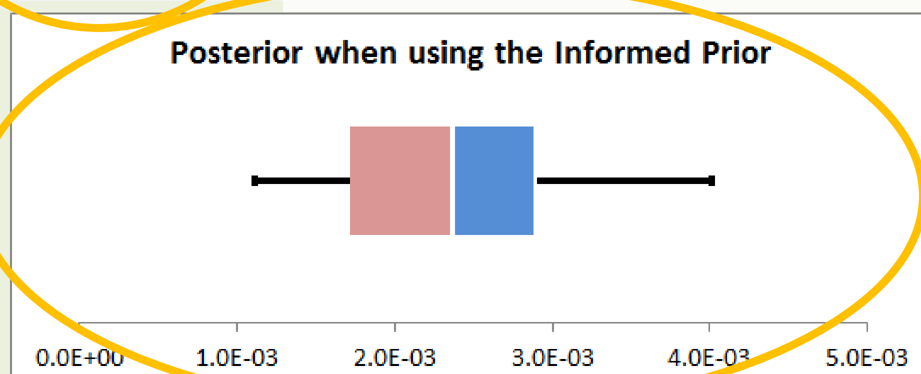
times =	205	100	760	450	1100

DSW Exponential-Gamma Bayesian Example

- Read the results

Posterior Results	Informed	Jeffreys	
α =	7	5	
β =	2955	2615	
Posterior Distribution =	Gamma(7, 2955)	Gamma(5, 2615)	
5th =	1.1E-03	7.5E-04	
25th =	1.7E-03	1.3E-03	
mean =	2.4E-03	1.9E-03	
75th =	2.9E-03	2.4E-03	
95th =	4.0E-03	3.5E-03	
MTT_ =	4.2E+02	5.2E+02	mean time to ___ (e.g., r





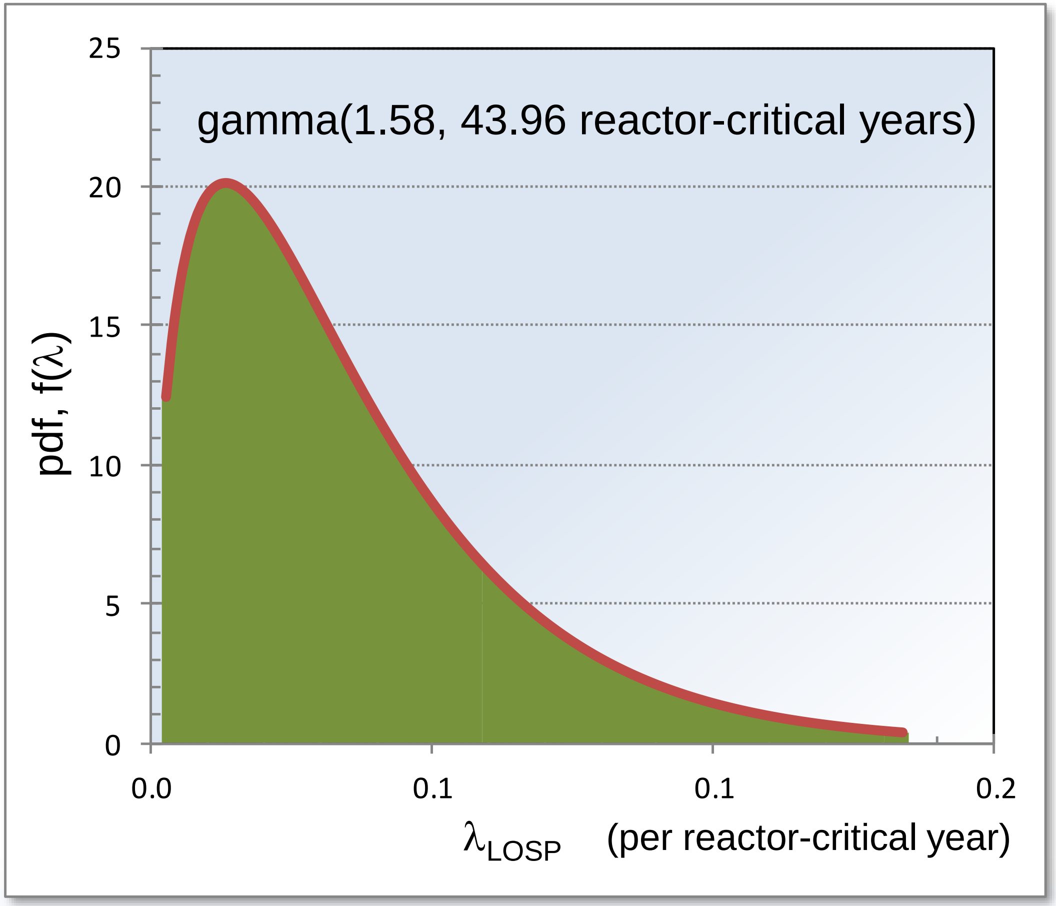
LOSP EXAMPLE

Prior Distributions for LOSP Example (For Later Reference)

- λ_{LOSP} ~ gamma(1.58, 43.96 reactor-critical years)
  - From "Reevaluation of Station Blackout Risk at Nuclear Power Plants: NUREG/CR-6890, December 2005
  - This is the composite from several subtypes of LOSP event
- p_{FTS} ~ beta(0.957, 190)
  - From S. A. Eide, "Historical Perspective On Failure Rates for US Commercial Reactor Components," Reliability Engineering and System Safety, **80** (2003), pp. 123-132
- λ_{FTR} ~ gamma(1.32,1137 hrs)
  - From Eide (2003)
  - This is the composite of two rates

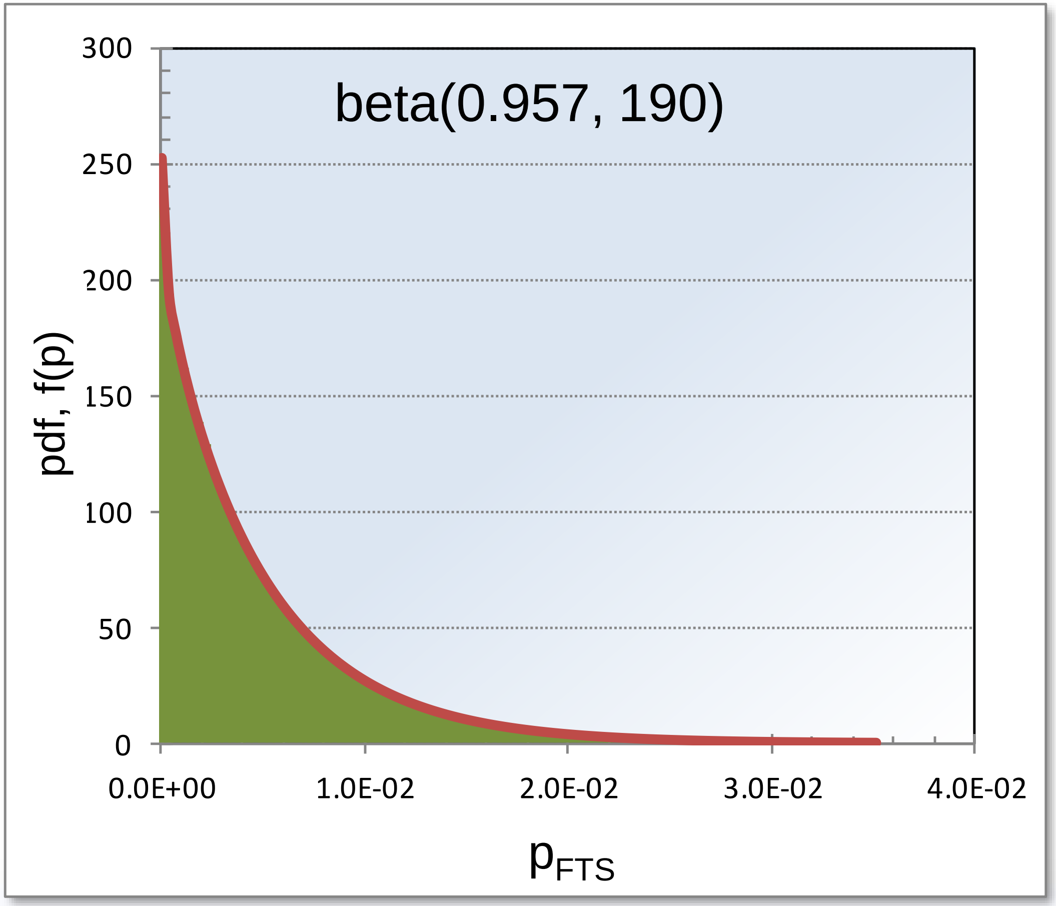
Prior Density Plots: LOSP

- λ_{LOSP}



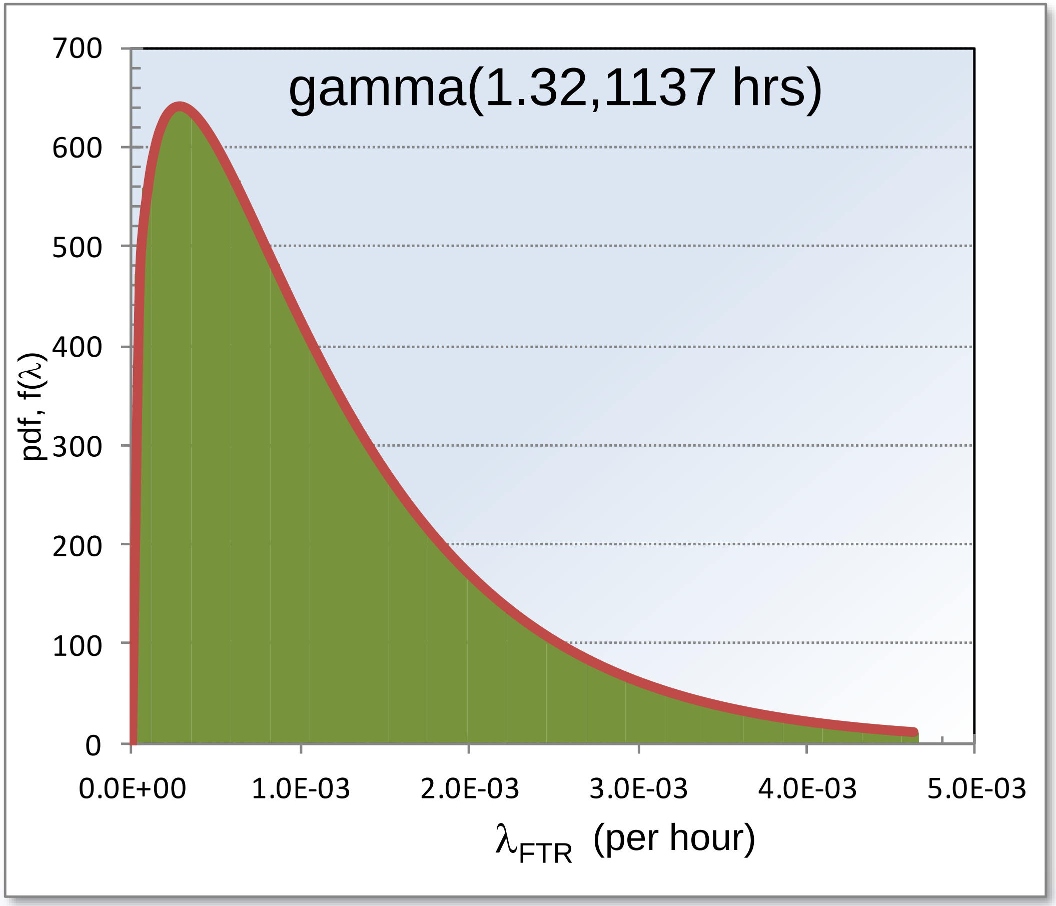
Prior Density Plots: FTS

- p_{FTS}



Prior Density Plots: FTR

- λ_{FTR}



LOSP Example Data

- The observed number of LOSP events over a period of time
  - **1 initiating event in 9.2** operating years
- The observed number of failures out of a number of demands
  - **1 failure to start in 75** demands
- The observed number of failures in an observed total operating time
  - **0 failures to run in 146** running hours

LOSP Frequency Update with DSW

- For LOSP frequency, aleatory model is Poisson
  1. Specify prior (Gamma tab)~gamma(1.58, 43.96 rcy)
  2. Bayesian update, so select “Bayes-Poisson” tab

Poisson Data

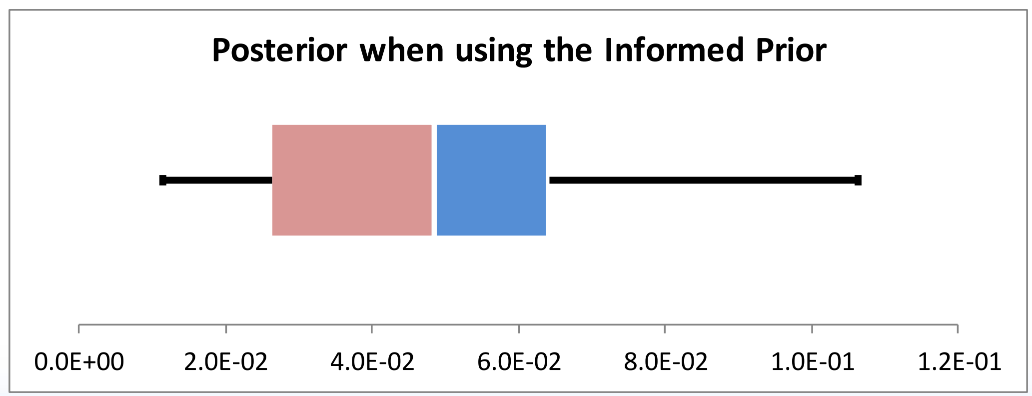
This worksheet performs conjugate Ba
For Poisson Data, the conjugate prior i

			Parameters	
Informed Prior		α =	1.58	Gamma(α,
		β =	43.96	
Jeffreys Prior		α =	0.5	Gamma(½
		β =	0	
Constrained Non-	Mean	α =	0.5	Gamma(α,
informative (CNI)	1.00E-01	β =	5.0	
Data Observed		x =	1	(number c
		t =	9.2	(observed

		Informed
Posterior Results	α =	2.58
	β =	53.16
	Posterior Distribution =	Gamma(2.58, 53.16)
	5th =	1.1E-02
	25th =	2.6E-02
	mean =	4.9E-02
	75th =	6.4E-02
	95th =	1.1E-01

Microsoft Excel

Distribution Summary Worksheet – Ver 1e.xlsx



EDG FTS Update with DSW

- For EDF fails to start, aleatory model is Binomial
  1. Specify prior (Beta tab)~beta(0.957, 190)
  2. Bayesian update, so select "Bayes-Binomial" tab

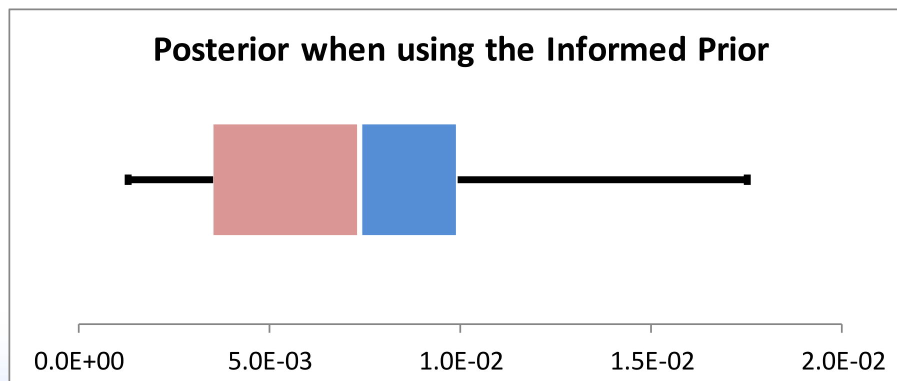
Binomial Data

This worksheet performs conjugat
For Binomal Data, the conjugate pr

			Parameters	
Informed Prior		α =	0.957	Beta(α
		β =	190	
Jeffreys Prior		α =	0.5	Beta(½
		β =	0.5	
Constrained Non-informative (CNI)	Mean	α =	0.5	Beta(α
	1.00E-01	β =	4.5	
Data Observed		x =	1	(numb
		n =	75	(numb
			Informed	
Posterior Results		α =	1.957	
		β =	264	
		Posterior Distribution =	Beta(1.957, 264)	
		5th =	1.3E-03	
		25th =	3.5E-03	
		mean =	7.4E-03	
		75th =	9.9E-03	
		95th =	1.8E-02	

Microsoft Excel

Distribution Summary Worksheet – Ver 1e.xlsx



EDG FTR Update with DSW

- EDG fails to run, aleatory model is Poisson
  1. Specify prior (Gamma tab) ~gamma(1.32,1137 hrs)
  2. Bayesian update, so select “Bayes-Poisson” tab

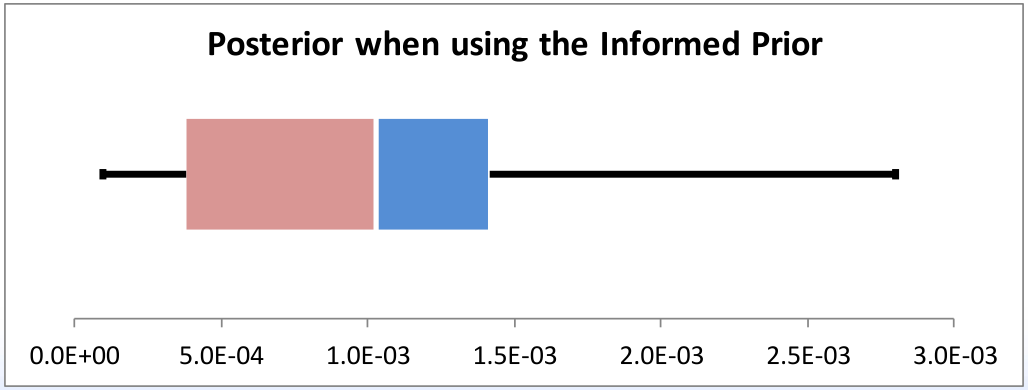
Poisson Data

This worksheet performs conjugate Bay
For Poisson Data, the conjugate prior is

			Parameters	
Informed Prior		α =	1.32	Gamma(α,
		β =	1137	
Jeffreys Prior		α =	0.5	Gamma(½,
		β =	0	
Constrained Non-	Mean	α =	0.5	Gamma(α,
informative (CNI)	1.00E-01	β =	5.0	
Data Observed		x =	0	(number o
		t =	146	(observed
			Informed	
Posterior Results		α =	1.32	
		β =	1283	
	Posterior Distribution =		Gamma(1.32, 1283)	
		5th =	9.6E-05	
		25th =	3.8E-04	
		mean =	1.0E-03	
		75th =	1.4E-03	
		95th =	2.8E-03	

Microsoft Excel

Distribution Summary Worksheet – Ver 1e.xlsx



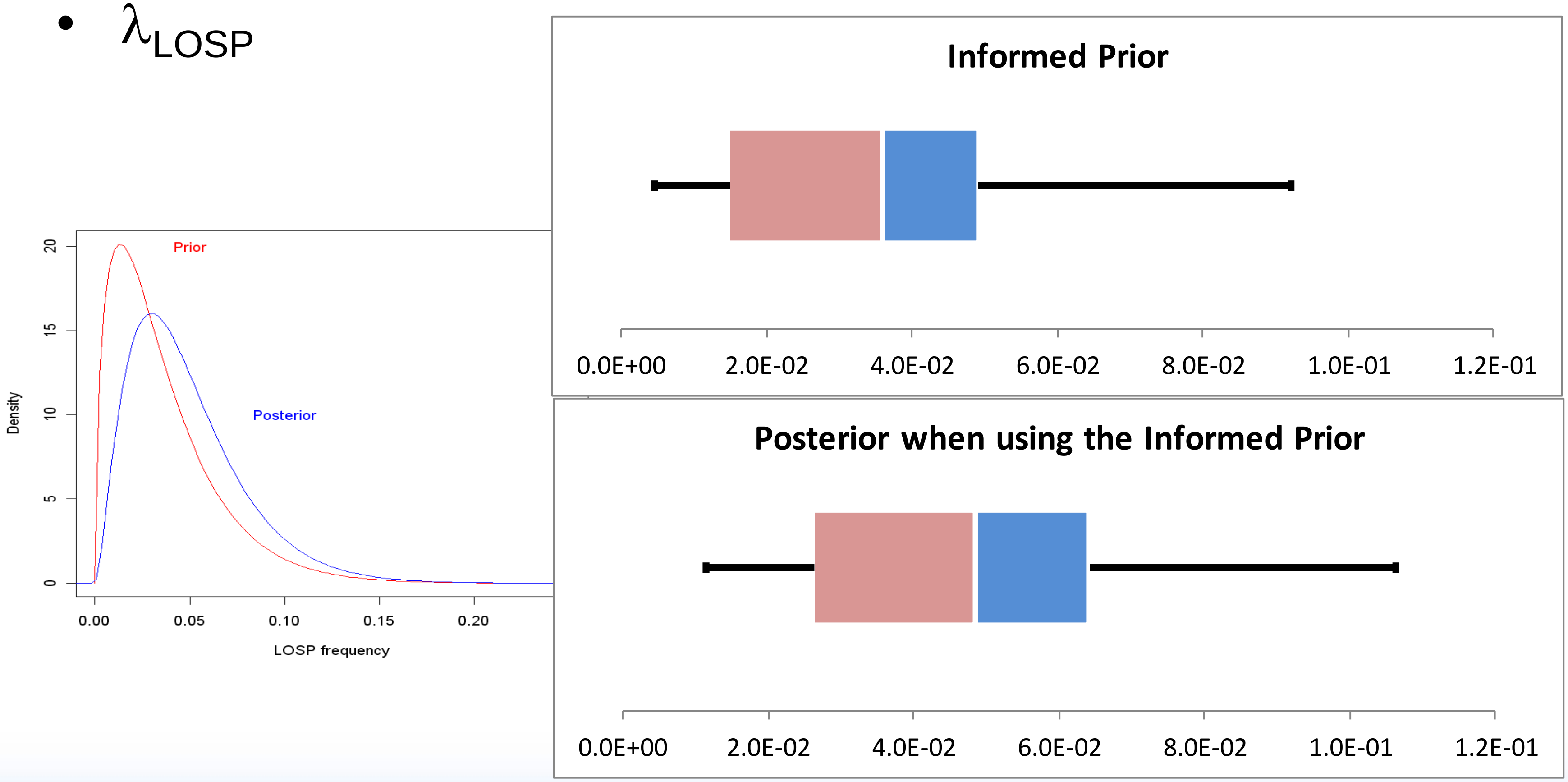
Summary of Bayesian Estimates for LOSP Example

Parameter	Distribution	Point Est. (Mean)	90% Interval	Distribution
λ_{LOSP}	*Industry Prior*	*3.6E-2 yr^{-1}*	*(4.6E-3, 9.2E-2) yr^{-1}*	*Gamma(1.58, 43.96)*
	Posterior	***4.9E-2 yr^{-1}***	***(1.1E-2, 1.1E-1) yr^{-1}***	***Gamma(2.58, 53.16)***
p_{FTS}	*Industry Prior*	*5.0E-3*	*(2.3E-4, 1.5E-2)*	*Beta(0.957, 190)*
	Posterior	***7.4E-3***	***(1.3E-3, 1.8E-2)***	***Beta(1.957, 264)***
λ_{FTR}	*Industry Prior*	*1.2E-3 hr^{-1}*	*(1.1E-4, 3.2E-3) hr^{-1}*	*Gamma(1.32, 1137)*
	Posterior	***1.0E-3 hr^{-1}***	***(9.6E-5, 2.8E-3) hr^{-1}***	***Gamma(1.32, 1283)***

Posterior credible intervals generally are shorter than those from data alone (i.e., confidence interval) or prior alone (prior credible interval)

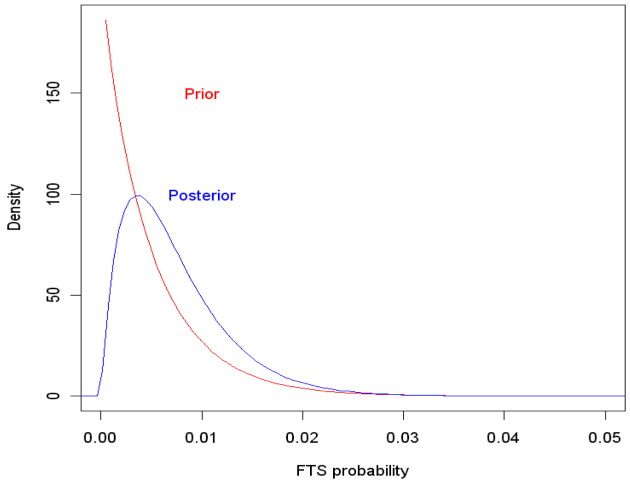
Comparison of Posterior and Prior: LOSP

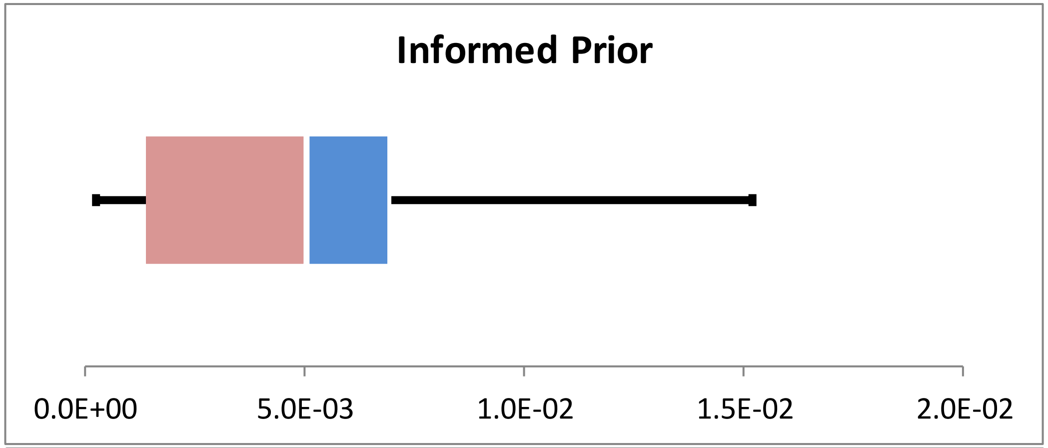
- λ_{LOSP}

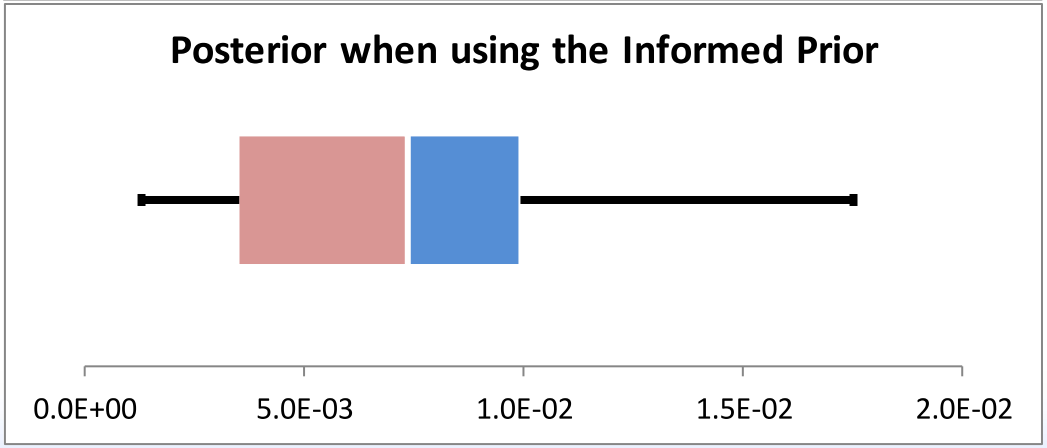


Comparison of Posterior and Prior: FTS

- p_{FTS}



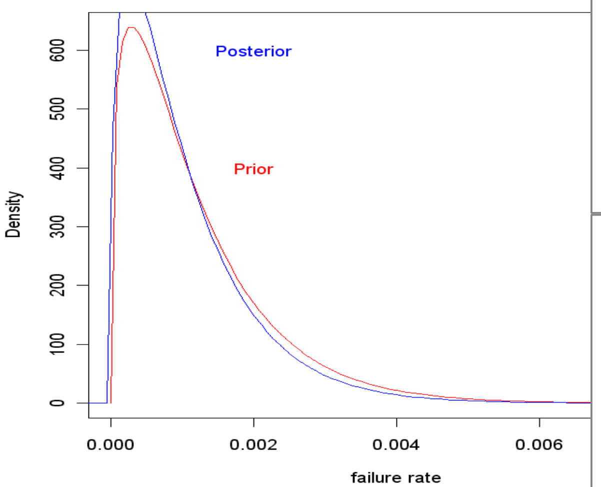


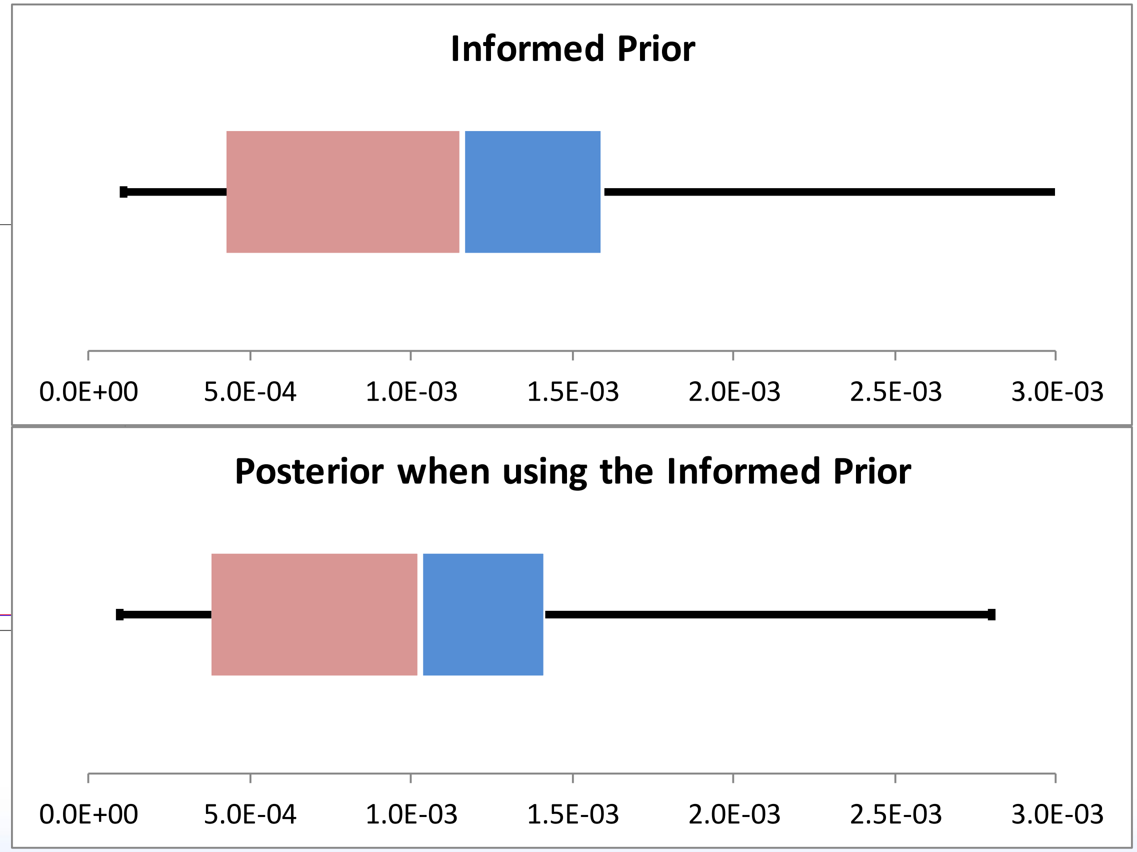


Idaho National Laboratory

Comparison of Posterior and Prior: FTR

- λ_{FTR}





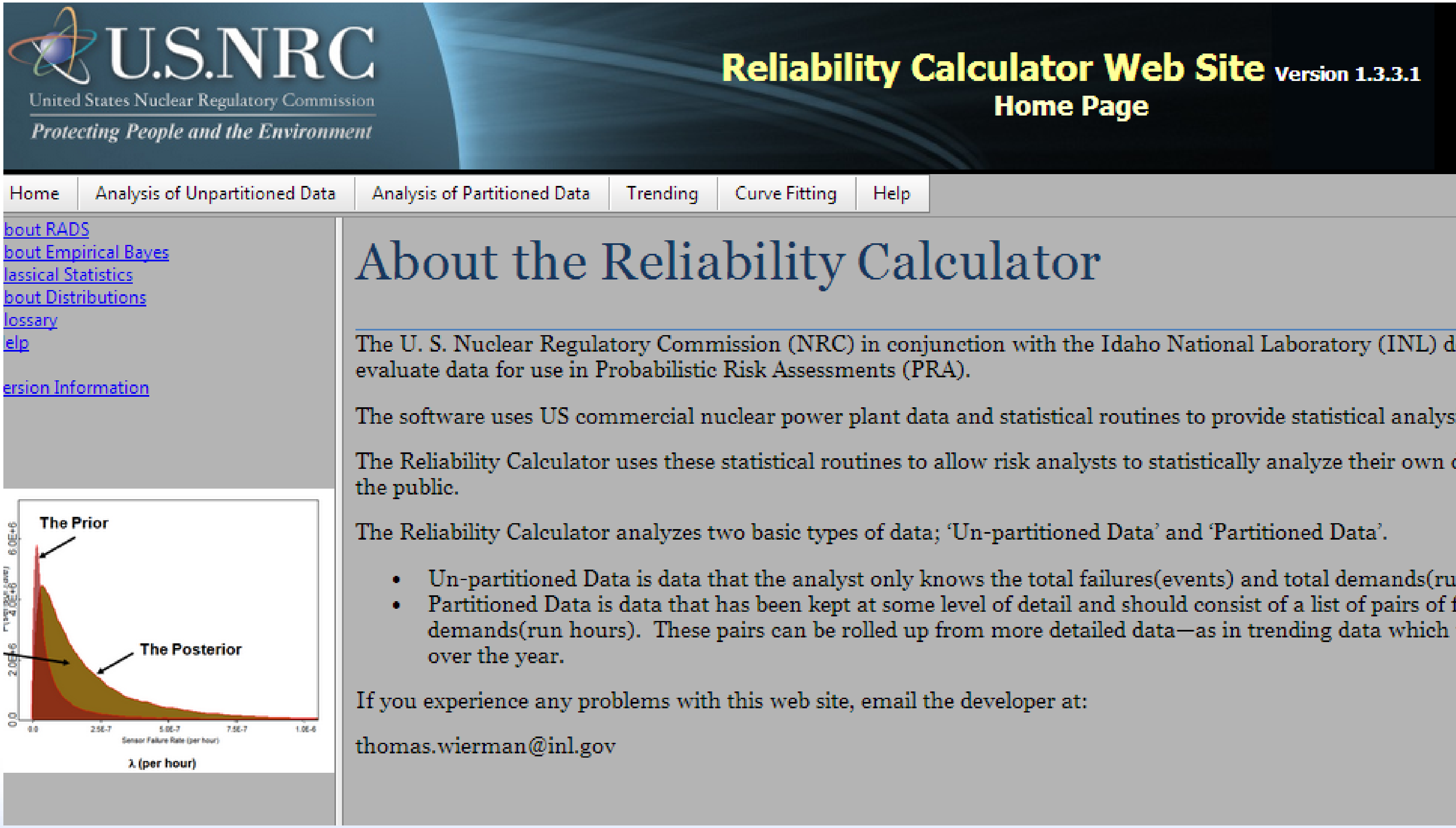
LOSP Updates with RADS Calculator

- RADS (Reliability and Availability Data System) was developed for NRC by INL
- Both stand-alone version and web-based calculator
- Will show web-based calculator in this course
- Access calculator at

  https://nrcoe.inel.gov/radscalc/Default.aspx

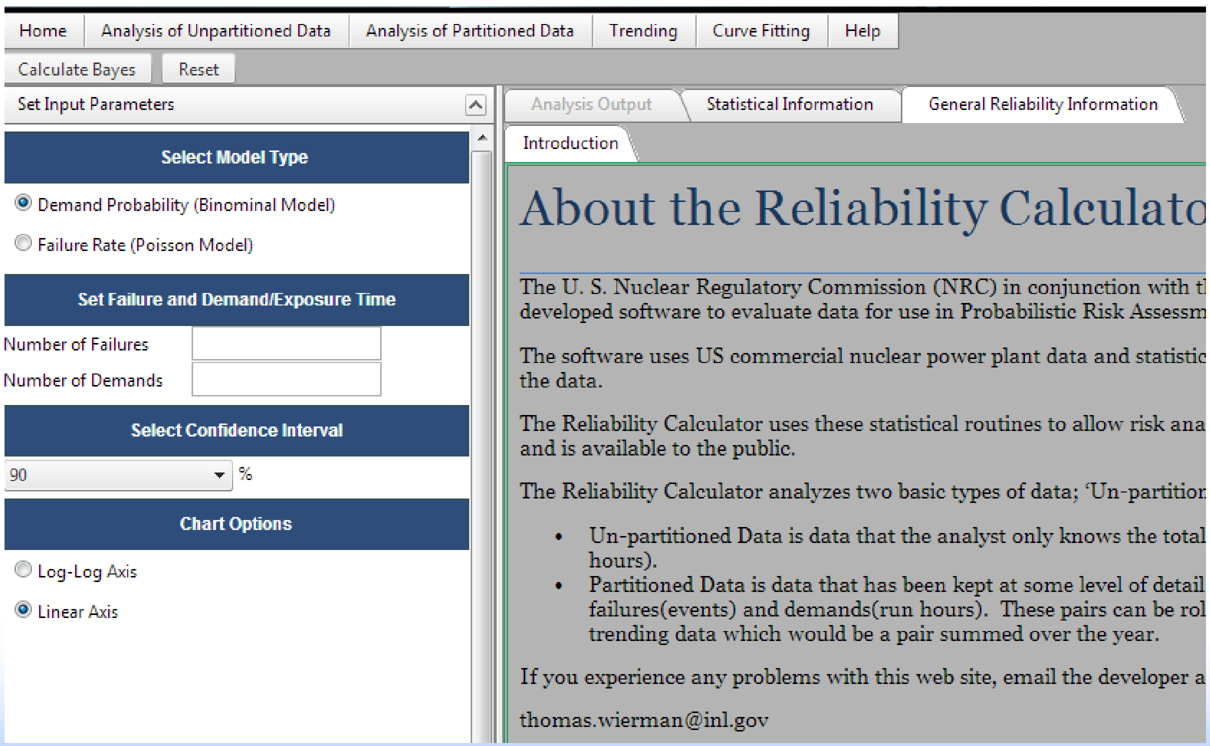
Idaho National Laboratory

LOSP Updates with RADS Calculator



LOSP Update with RADS Calculator

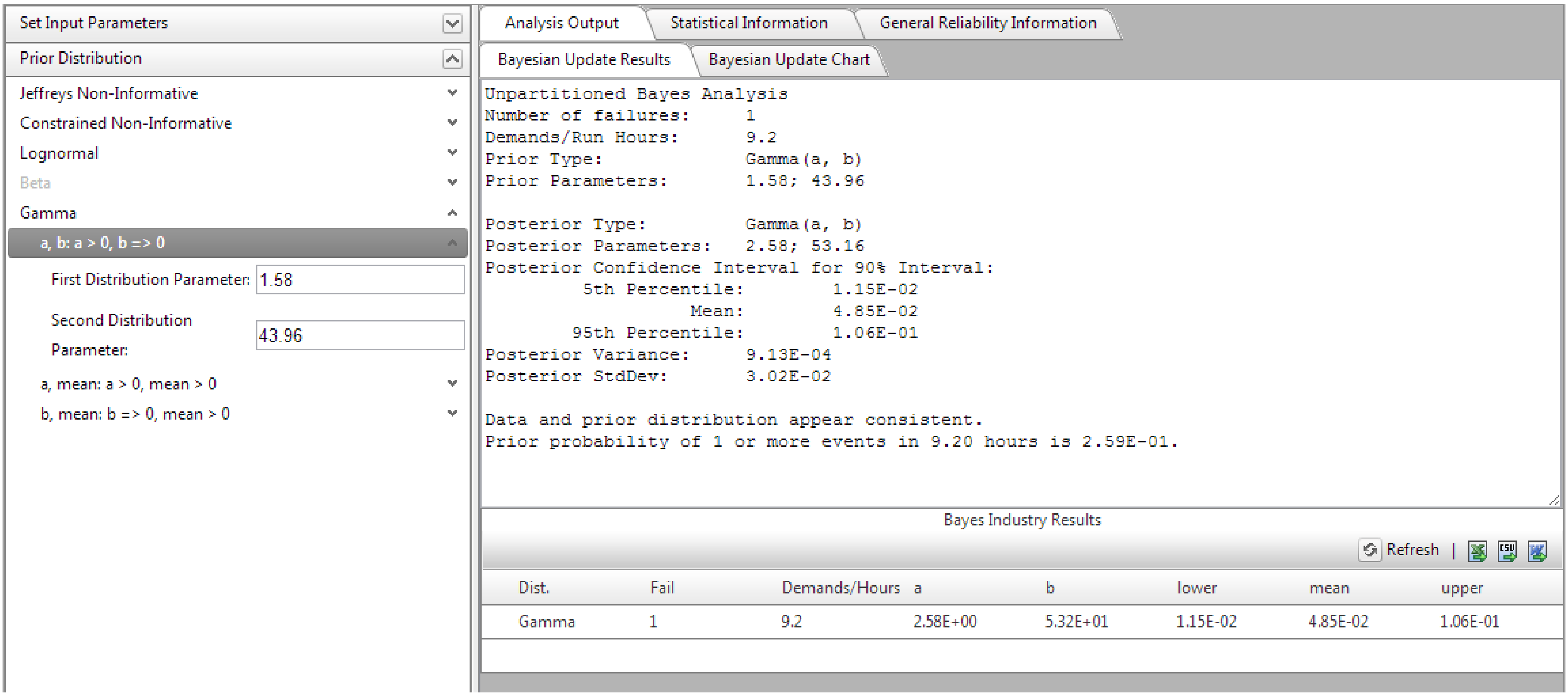
- Menu options across the screen
  - Select Analysis of Unpartitioned Data
    - Bayes Analysis



LOSP Update with RADS Calculator

- For LOSP frequency, select "Failure Rate"
- Enter data
  - "Number of Failures"
  - "Exposure Time"
    - Units are not specified
- Click the "Prior Distribution" bar towards bottom-left
  - Select gamma prior ~gamma(1.58, 43.96)
  - Enter "a" (alpha=1.58) and "b" (beta=43.96)
- Press "Calculate Bayes"

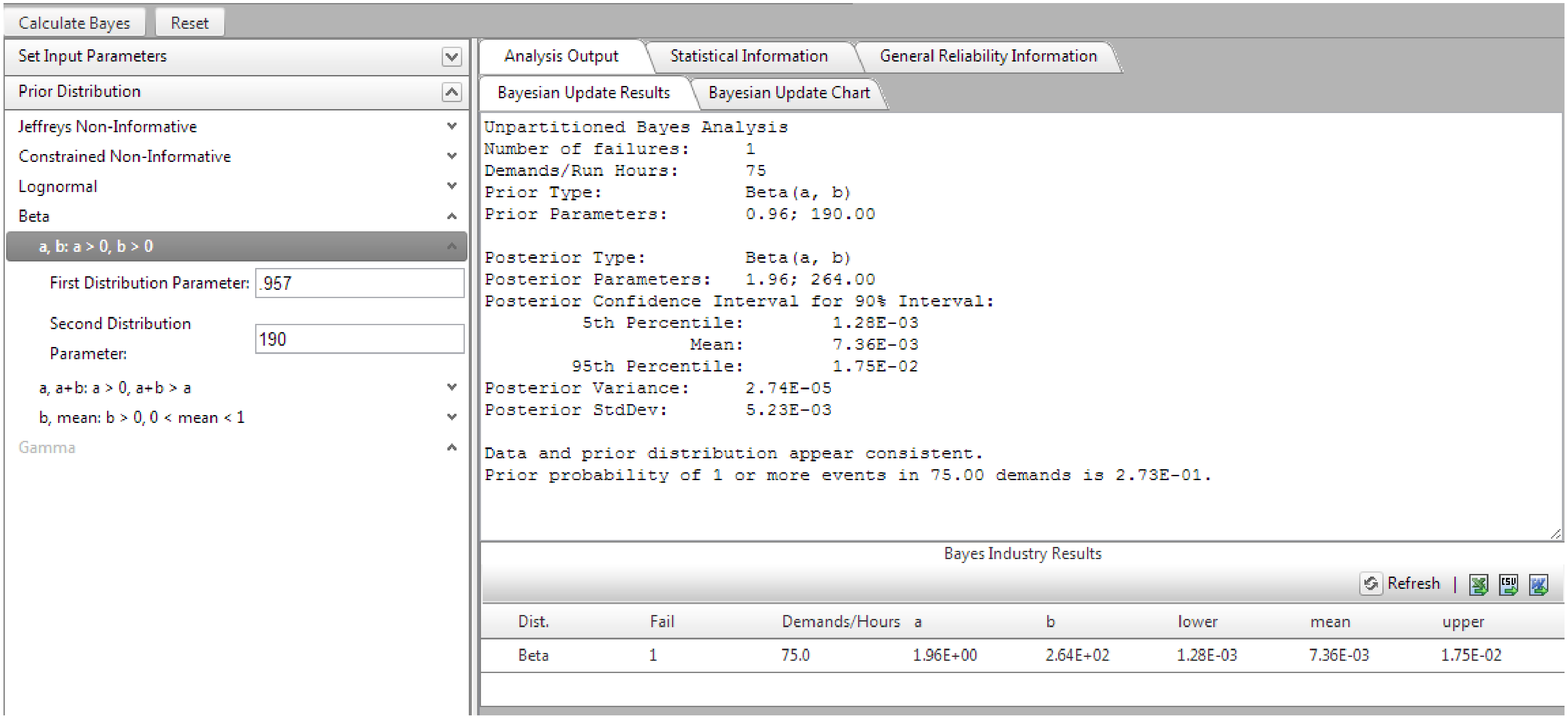
RADS Calculator results



LOSP Update with RADS Calculator

- For EDG failure to start, select "Demand Probability"
- Enter data
  - "Number of Failures"
  - "Number of Demands"
- Select beta prior and enter "a" (alpha) and "b" (beta)
- Press "Calculate Bayes"

EDG FTS Update with RADS Calculator



Idaho National Laboratory

NONINFORMATIVE PRIORS

Noninformative (Formal) Prior Distributions

- The original intent of "**noninformative**" priors was to answer question
  - How do we find a prior representing complete ignorance?
- Rev. Bayes suggested a **uniform** prior
  - Laplace used this in his activities with great success
  - But, there are philosophical/mathematical problems with this

Noninformative (Formal) Prior Distributions

- Sir Harold Jeffreys suggested a prior that was invariant to changes in
  - Scale
  - Location
- Consequently, so-called "noninformative" prior is typically not uniform for the parameter of interest
  - It is uniform (or approximately so) for some transformed parameter
  - Instead, it depends on the process generating the data
  - Has property that the Bayes posterior intervals are approximately equal to frequentist confidence intervals (exactly equal for continuous data)
    - Formal priors "let the empirical data speak for themselves"

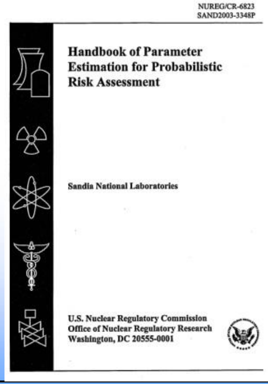
Noninformative (Formal) Prior Distributions

- Other formal priors have been developed
- Often used as “objective” or reference prior
  – Also called “vague” or “diffuse” prior
  – Jeffreys prior is most common choice of formal prior for single-parameter problems

Sir Harold Jeffreys

Jeffreys Prior Distributions

- For **binomial**(n, p) aleatory model
  – Jeffreys noninformative prior for p is **beta**(1/2, 1/2), which is a proper prior (integral equals 1)
- For **Poisson**(λt) aleatory model
  – Jeffreys noninformative prior for λ can be thought of as **gamma**(1/2, 0) (improper prior, integral diverges)
- For **exponential**(λ) aleatory model
  – Jeffreys noninformative prior for λ can be thought of as **gamma**(0, 0) (improper prior, integral diverges)


NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

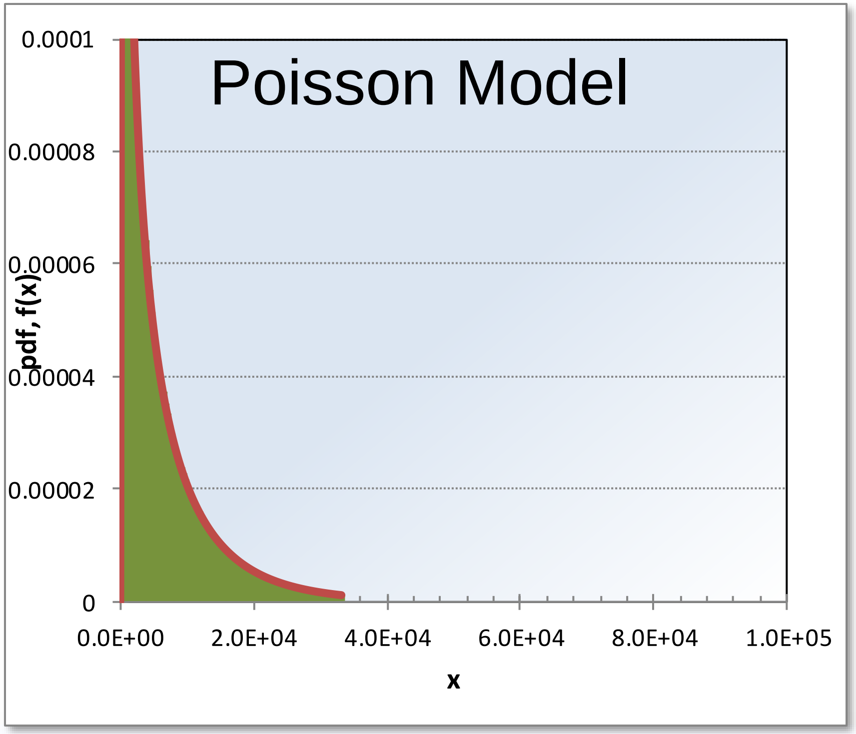
U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages B-12 through B-14, 6-14, 6-37, 6-61, 6-62

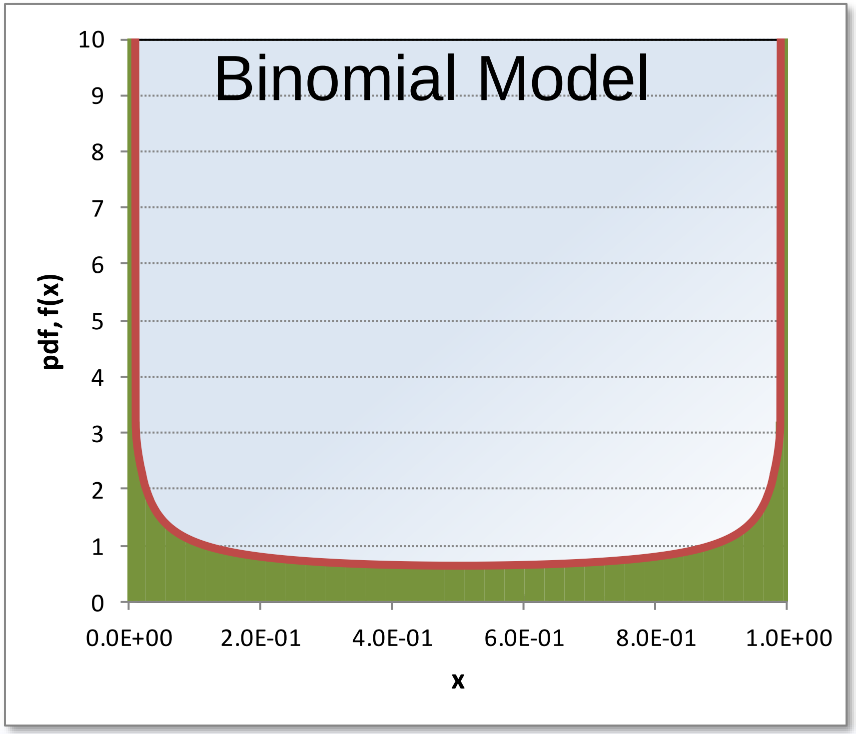
Idaho National Laboratory

Jeffreys Prior Distributions

- Two of these priors look like



gamma(1/2, 0)



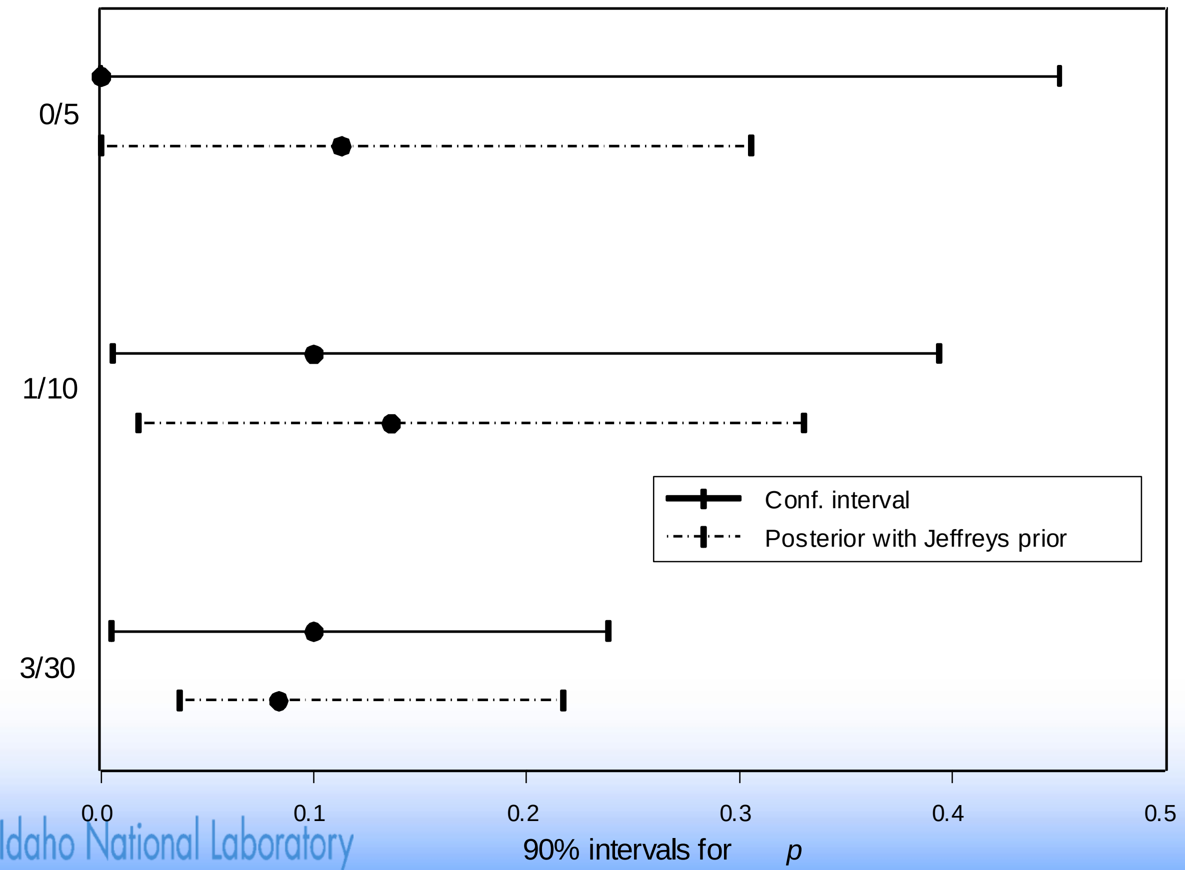
beta(1/2, 1/2)

Updating Jeffreys Prior Distributions

- For **binomial**(n, p) aleatory model
  - Jeffreys noninformative prior for p is **beta**(½, ½)
  - Posterior is beta(x + ½, n – x + ½)
- For **Poisson**(λt) aleatory model
  - Jeffreys noninformative prior for λ is **gamma**(½, 0)
  - Posterior is gamma(x + ½, t)
- For **exponential**(λ) aleatory model
  - Jeffreys noninformative prior for λ is **gamma**(0, 0)
  - Posterior is gamma(n, $\Sigma\ t_i$)

Jeffreys Prior Distributions

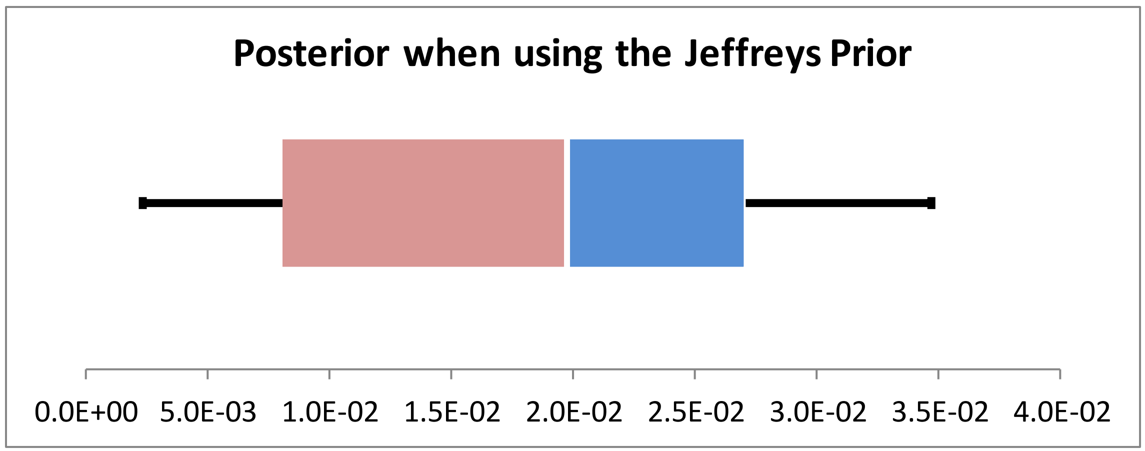
- Let us compare results from these priors to frequentist confidence intervals



Using DSW to Update Jeffreys Prior Distributions

- Select appropriate tab (e.g., Bayes-Binomial)
- Enter observed data (e.g., 1 failure in 75 demands)
- Read off posterior from column labeled "Jeffreys"

Jeffreys
1.5
74.5
Beta(1.5, 74.5)
2.4E-03
8.1E-03
2.0E-02
2.7E-02
5.1E-02

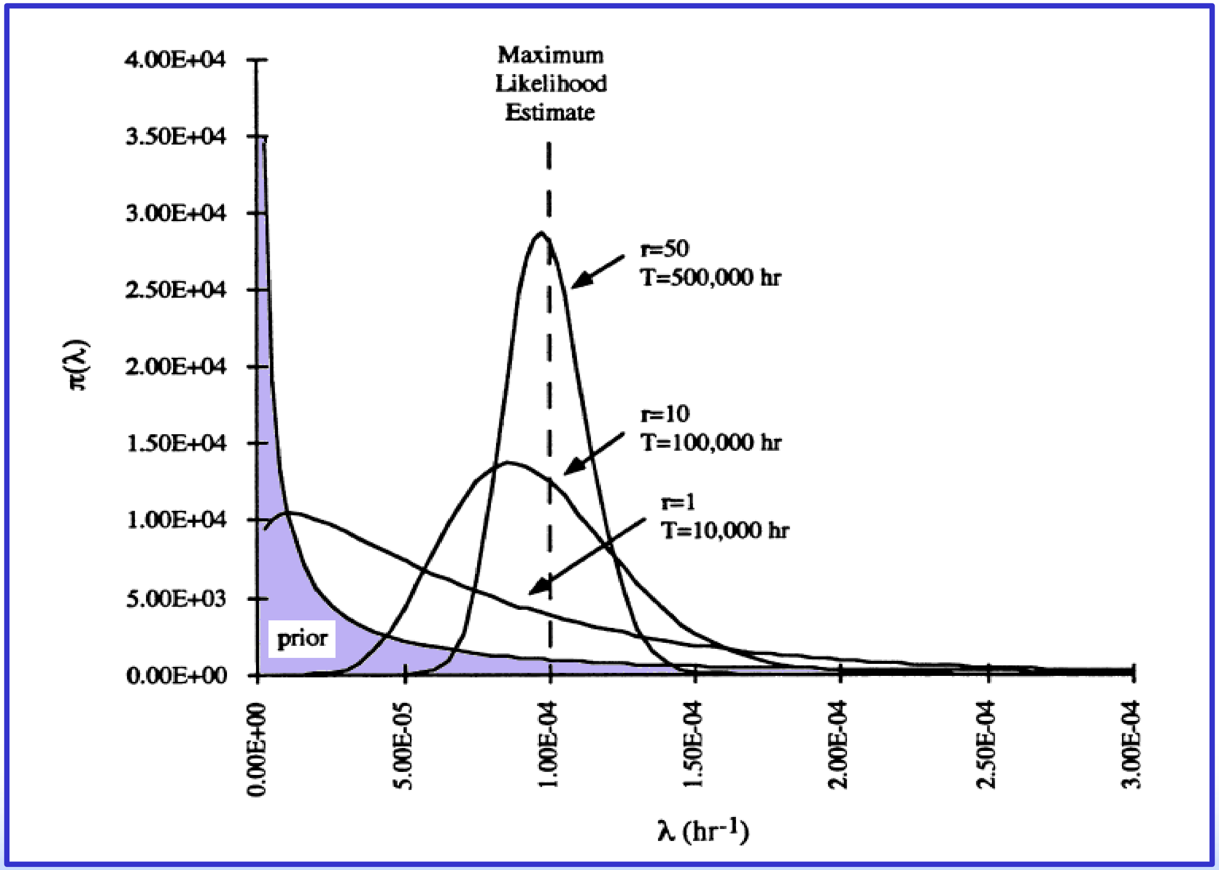


Using RADS Calculator to Update Jeffreys Prior Distributions

- Select “Jeffreys noninformative” as prior distribution
- Enter observed data
- Press “Calculate Bayes”

Data Versus Prior Distribution

- As amount of data increases, prior becomes less important



Siu and Kelly, 1998

Constrained Noninformative (CNI) Prior

- Recall that Jeffreys prior for binomial likelihood is beta(½, ½)
  – The prior mean is 0.5 (quite large if a failure prob.)
- With sparse data, prior mean can influence results too much
- To overcome this, can use prior which has specified mean, but is close to Jeffreys prior otherwise
  – Specified mean might be industry average value
- Result is called “constrained noninformative” (CNI) prior

Left to right: Claude Shannon, Edward Jaynes, Corwin Atwood

Idaho National Laboratory

Constrained Noninformative Prior

- Cannot be written in form of standard distribution for case of binomial likelihood
  - Approximated well by beta distribution with α = 0.5
    - See Table C.8 in HOPE for precise values of α
  - For the beta, the mean = α / (α + β), β can be found to be

    β = α(1-mean)/mean
- For Poisson likelihood, CNI prior is gamma(½, 1/(2*mean))
- For exponential likelihood, cannot define CNI prior, as it is not proper, therefore cannot have finite mean
  - Alternative is maximum entropy prior, which is
    - Gamma(1, 1/mean)
    - This is an exponential distribution

Idaho National Laboratory

Updating CNI Prior with DSW

- Is conjugate for Poisson and binomial data
- Enter specified mean value
- Enter data observed
- Read off posterior results in column labeled “CNI”
- Example: assume mean probability of EDG failure to start is thought to be 0.01
  - Take this as mean of CNI prior and update with 1 failure in 75 demands

Idaho National Laboratory

Updating CNI Prior with DSW

Binomial Data

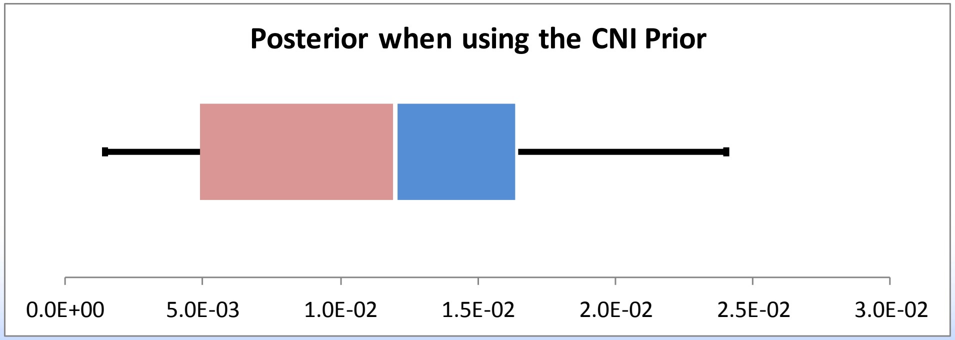
This worksheet performs conjugate Bayesian upd
For Binomal Data, the conjugate prior is Beta(α, β

	Mean	Parameters		
Informed Prior		α =	0.957	Beta(α, β), where α and
		β =	190	
Jeffreys Prior		α =	0.5	Beta(½, ½)
		β =	0.5	
Constrained Non-informative (CNI)		α =	0.5	Beta(α, β)
	1.00E-02	β =	49.5	

Data Observed			
	x =	1	(number of failures)
	n =	75	(number of demands)

CNI
1.5
123.5
Beta(1.5, 123.5)
1.4E-03
4.9E-03
1.2E-02
1.6E-02
3.1E-02

Results



Updating CNI Prior with RADS Calculator

- Input parameters of beta or gamma CNI prior, as appropriate
- Press "Calculate Bayes"

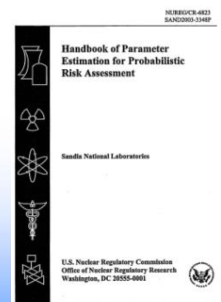
NONCONJUGATE PRIORS

Nonconjugate Priors

- For each aleatory model, there is **at most** one conjugate prior type
  - All other distributional forms are nonconjugate
  - Integral in denominator of Bayes' Theorem must be done numerically
- Most common nonconjugate prior in PRA is lognormal distribution
  - Originally used in WASH-1400 to represent plant-to-plant variability in parameter values
  - Very convenient when uncertainty spans several orders of magnitude
  - Useful in expert elicitation (e.g., seismic PRA, NUREG-1829)
    - Elicit median and upper bound, 5^{th} and 95^{th} percentiles, etc.

Nonconjugate Prior Distributions – The Concept

- Generic databases often express uncertainty in terms of lognormal distribution
- Experts often provide order-of-magnitude estimates, represented well by lognormal distribution
- For these or other reasons, we may prefer a nonconjugate prior
- When prior is not conjugate
  - Posterior distribution must be found by numerical integration. Will use online RADS calculator.



NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

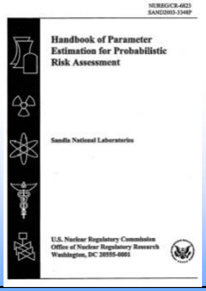
Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Pages 6-16 through 6-20, 6-39 through 6-43

Lognormal Distribution

- Definition of a lognormal distribution:
  - X is **lognormal**(μ, σ^2) if ln(X) is **normal**(μ, σ^2)
- Will encounter lognormal distribution in various areas of risk assessment
  - **Often used** as a prior distribution in PRA, even though it is **not conjugate**
  - Sometimes used as likelihood function (e.g., LOSP recovery time)
    - Covered in P-501 and P-502 courses
  - Often used to model hazard (earthquake frequency) and fragility (probability of seismic failure) in seismic PRA



Pages A-16, A-17

Facts About the Lognormal Distribution

- Median of X is e^{μ} Mean of X is $\exp[\mu + (½)\sigma^2]$
- Variance of X is $(\text{mean})^2[\exp(\sigma^2) - 1]$
- Error factor (EF) is defined as $e^{1.645\sigma}$
- Other ways to write EF (applies only to lognormal)
  - EF = $95^{th}/50^{th} = 50^{th}/5^{th} = (95^{th}/5^{th})^{1/2}$
- Probability

$$\Pr(X \le x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$

where Φ is tabulated in HOPE, Appendix C

- Can also use =LOGNORMDIST(x, μ, σ) in Excel
- or =NORMDIST(LN(x), μ, σ, TRUE)

- Percentile
  - Can use =LOGINV(p, μ, σ) in Excel

Idaho National Laboratory

Lognormal Distribution

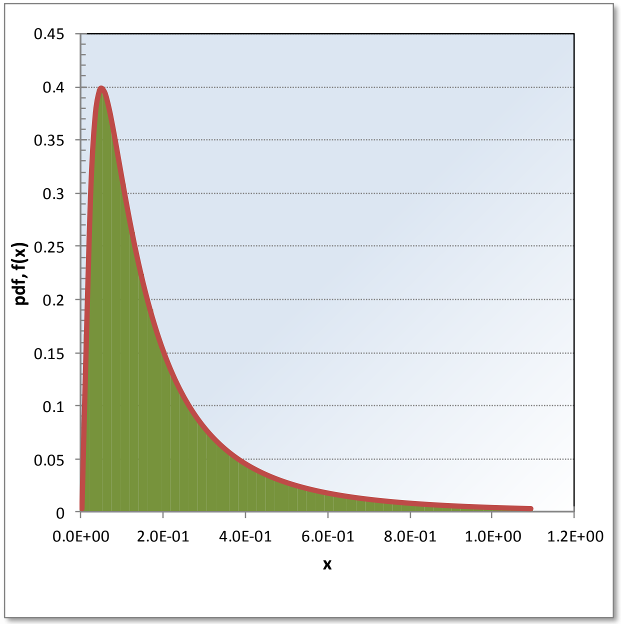
- Lognormal distribution is determined (in general) by **any two** of
  - μ
  - σ^2
  - median
  - mean
  - variance
  - EF or upper percentile
- SAPHIRE uses mean and EF

Distribution: Lognormal				Lognormal Dis
	Parameterized as X ~ Lognormal(μ, σ)			
Parameters	μ = -3	In SAPHIRE	LN Mean =	8.21E-02
	σ = 1		EF (error factor) =	5.18
Information				

Lognormal Distribution in DSW

- Parameterized using μ and σ
- Sheet also includes “lognormal calculator” when given
  - Mean & EF
  - Median & EF
  - Mean and standard deviation
    - These return μ and σ

29	Parameter Conversion	
30	Mean = 8.21E-02	μ = -3.00
31	EF = 5.18	σ = 1.000
32		
33	Median = 4.98E-02	μ = -3.00
34	EF = 5.18	σ = 1.000
35		
36	Mean 8.21E-02	μ = -3.00
37	Std. Dev. 0.11	σ = 1.000



Updating Lognormal Prior with RADS Calculator

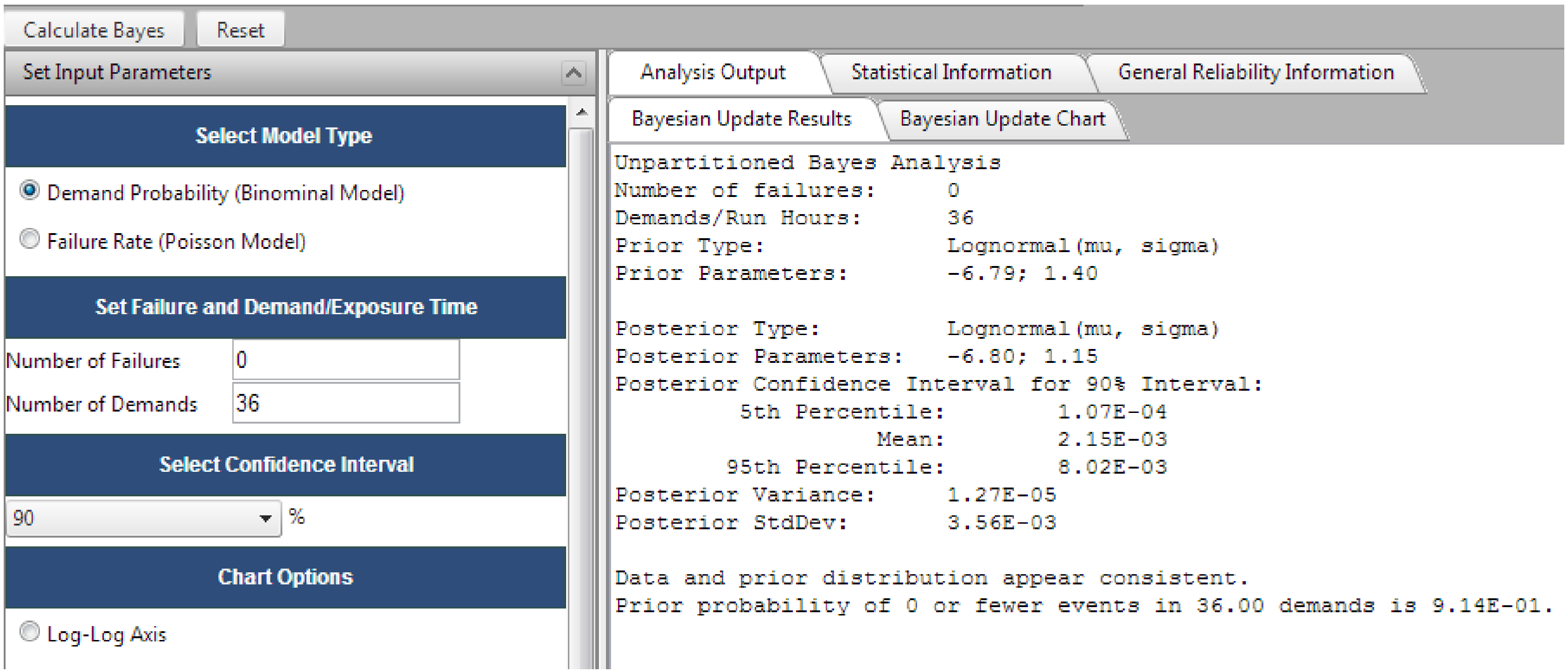
- Example: Interested in failure on demand for standby pump
- Generic database shows p is **lognormal** with
  - mean of 0.003
  - error factor of 10
- Observe 0 failures in 36 demands
- What is posterior mean of p?

Idaho National Laboratory

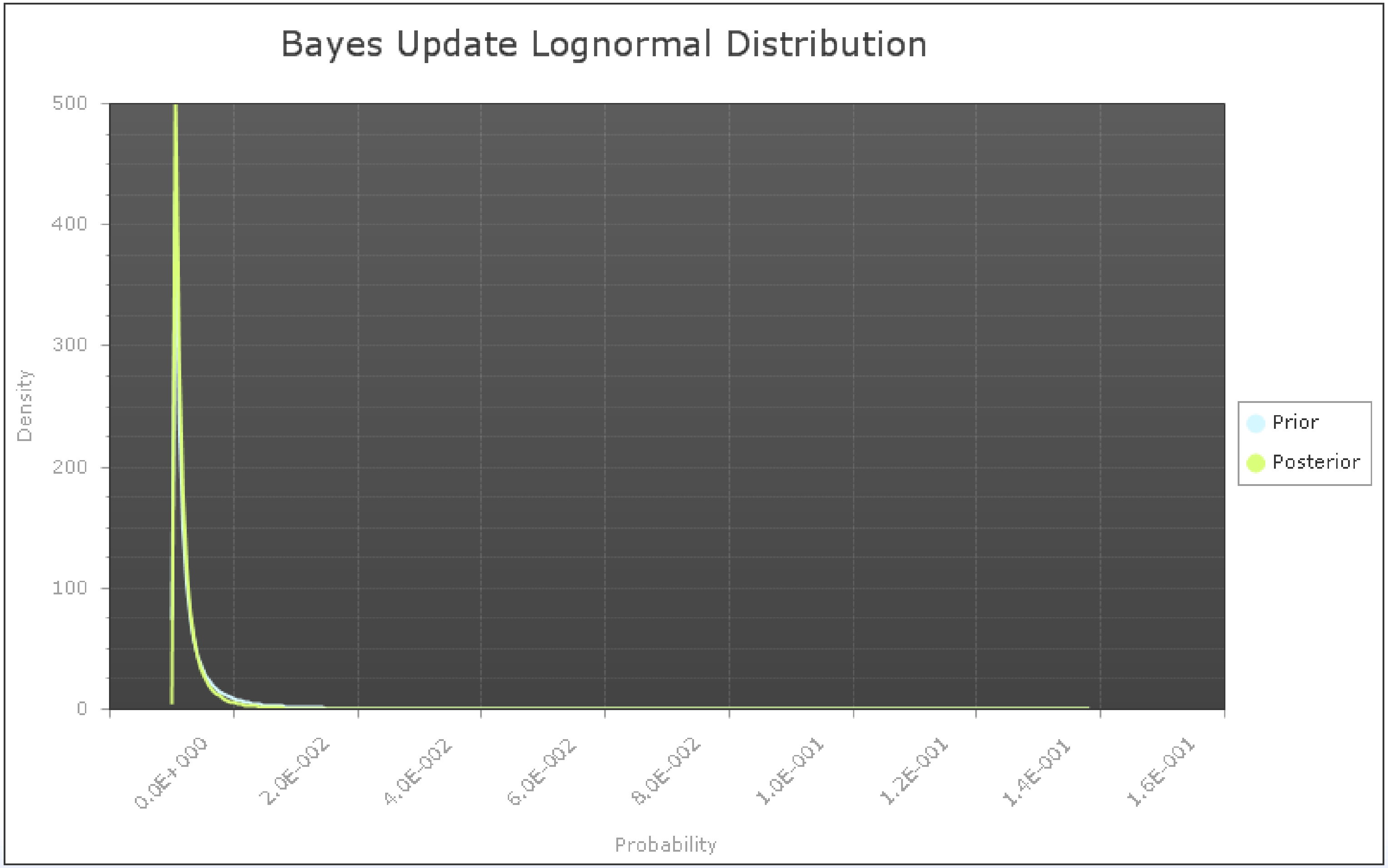
Updating Lognormal Prior with RADS Calculator

- Select "Demand Probability" and enter input data in usual way
- Select "Lognormal" as the prior distribution and enter mean and error factor
- Push "Calculate Bayes"
- Note that posterior distribution is **not** lognormal

Updating Lognormal Prior with RADS Calculator



Lognormal (cont.)



Section 4: Introduction to Monte Carlo Sampling

- Purpose
  - Concept of simulating a random variable via Monte Carlo sampling will be introduced and illustrated using Excel
- Objectives
  - Students will learn
    - **How to generate uniform random numbers in Excel**
    - How to generate a binomial random variable
    - **Concept of using inverse c.d.f. to generate random samples from a specified distribution**
    - Use of transformations to generate random samples
    - Determining sample size

Monte Carlo Sampling – Purpose in PRA

- Approximate a distribution by generating a large random sample from the distribution
- Useful for
  - Propagating uncertainties through logic model (e.g. fault tree or event tree)
  - Approximating posterior distribution when it does not have simple form (e.g. when prior is not conjugate)

Stanislaw Ulam

Sampling from a Uniform(0,1) Distribution

- Many software packages can sample from uniform distribution
  - Excel, R, Visual Basic, FORTRAN, SAPHIRE, etc.
- Completely deterministic, not random
  - “Looks” random, thus called “pseudorandom”
  - Really, the output is a long (e.g. ~2^{31}) sequence of distinct numbers
    - Order of numbers is unpredictable unless algorithm used to generate them is known
  - User inputs a “seed”, or computer uses the clock time
    - Seed determines where in the sequence we start

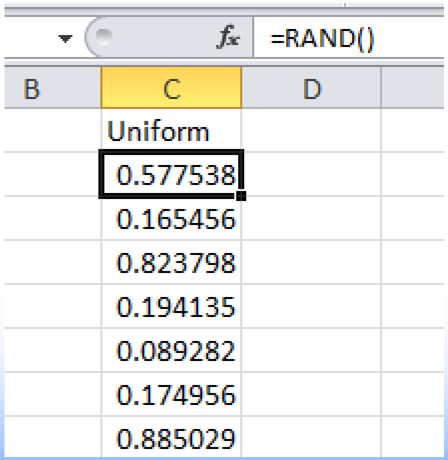
Pseudorandom Numbers

"Dilbert" Scott Adams

Idaho National Laboratory

Generating Uniform Random Numbers in Excel

- Use RAND() to generate from uniform(0, 1)
  - Can use F9 function key to recalculate (generate new random number)
- Use (b – a)*RAND() + a to generate random numbers from uniform(a, b)
- Use RANDBETWEEN(a, b) to generate uniformly distributed <u>integers</u> between a and b



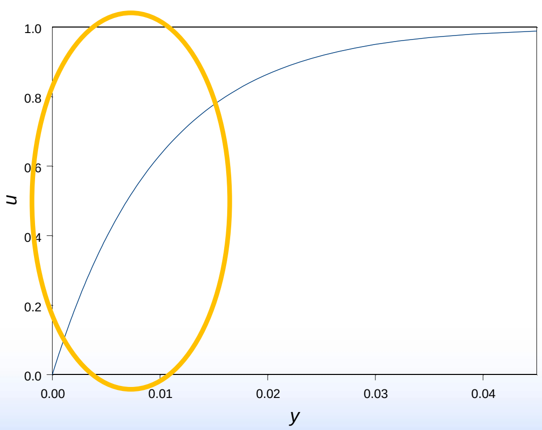
Idaho National Laboratory

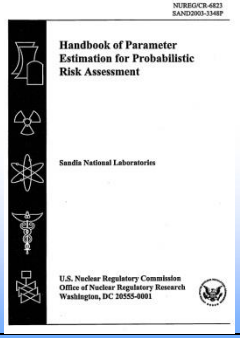
Sampling from a Binomial Random Variable (aleatory model)

- To simulate a **binomial**(n,p) random variable, iterate the following:
  - Generate n random numbers u_1 through u_n from a **uniform**(0,1) distribution
  - If $u_i < p$ define $x_i = 1$. Otherwise define $x_i = 0$.
  - Set $y = x_1 + ... + x_n$
  - Repeat
- The values of y are a sample from a **binomial**(n,p) distribution

Use of “Inverse CDF method”

- Iterate the following:
  - Generate u from a uniform(0,1) distribution
  - Set $y = F^{-1}(u)$, where
    - F is the CDF of Y, $F(y) = Pr(Y < y)$
    - F^{-1} is inverse function, $F(y) = u$ $F^{-1}(u)=y$
- Values of y are a random sample from the distribution of Y
- Idea…
  - Choose most values where F is steep




NUREG/CR-6823
SAND2003-3348P

Handbook of Parameter Estimation for Probabilistic Risk Assessment

Sandia National Laboratories

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

Page 6-41

Idaho National Laboratory

Example: Sampling from Exponential Variable via Inverse CDF

- Recall CDF for exponential

  $F(t) = 1 - e^{-\lambda t}$

- $t_i = F^{-1}(u_i) = -1/\lambda[\ln(1 - u_i)]$
- The t_is are a sample from an exponential(λ) distribution
- Thus, if we know
  - The rate λ
  - And can generate a uniform random number
  - We can generate exponential times, t_i

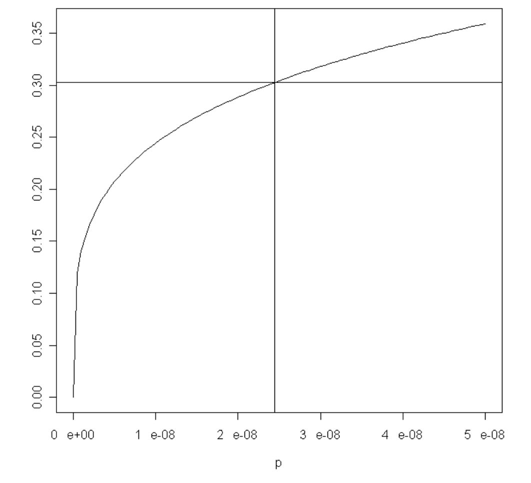
Idaho National Laboratory

Example: Sampling from Beta(α, β) Distribution in Excel

- Generate u_i from uniform(0, 1) distribution
  - RAND()
- Obtain beta-distributed values from BETAINV(u_i, α, β)
  - These “inverse functions” in Excel allow us to easily generate random samples from the distribution

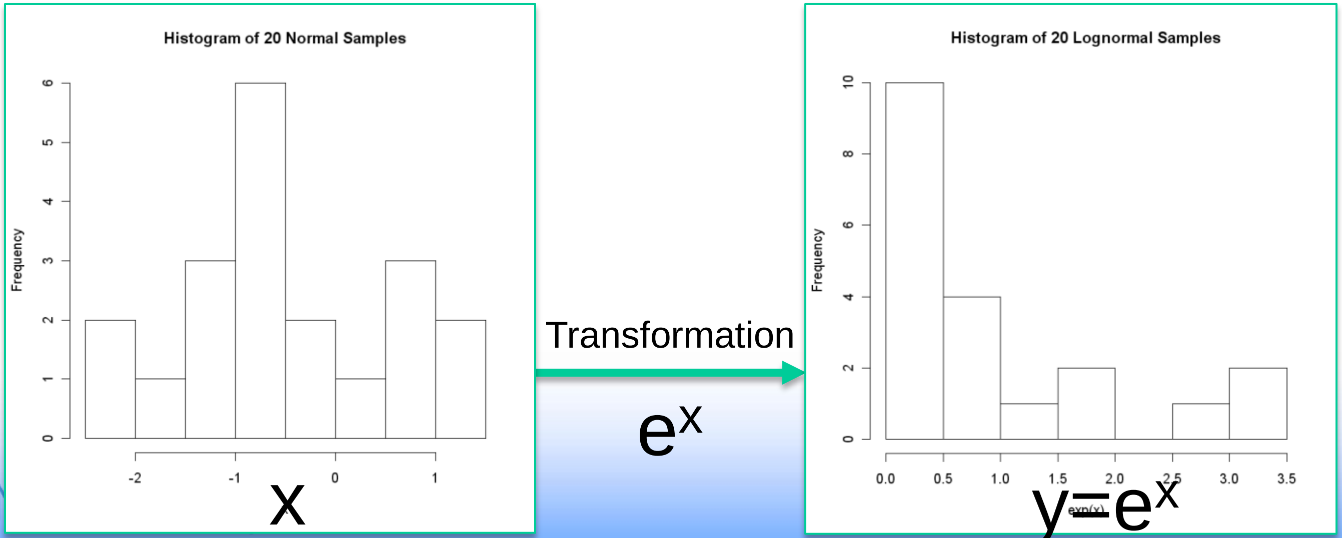
B6 =BETAINV(A6,0.24,189360)

	A	B	C
1	uniform	beta(0.24, 189360)	
2	0.82839	2.21096E-06	
3	0.30491	2.52048E-08	
4	0.23713	8.82039E-09	
5	0.60933	4.83265E-07	
6	0.30256	2.44033E-08	
7			
8			
9			



Use of Transformation

- For example, to generate lognormal Y
  - First generate n values from a normal distribution, call them x_1 through x_n
    - Set $y_i = \exp(x_i)$, so that $\ln(y_i) = x_i$
- The y_i values are a random sample from a lognormal distribution



Example: Lognormal Sampling with mean of 2E-4 and EF=7

- Generate u_i as before
  - RAND()
- Generate normal distribution values via =**norminv**(x,μ,σ)
  - First need to calculate μ and σ
- Obtain lognormal distribution values by taking e^y
- NOTE: We could sample directly using

  =LOGINV(x,μ,σ)

Random sampling from a lognormal distribution

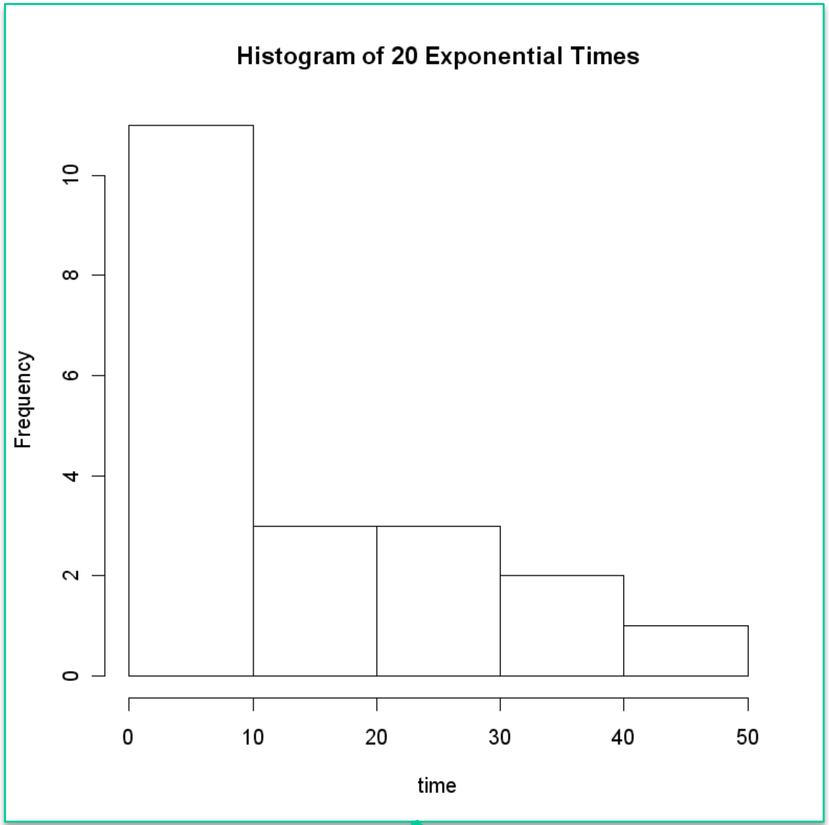
i	RAND	Normal(μ,σ)	Lognormal (via transformation)	Lognormal (via LOGINV)
1	0.216	-10.1	3.92E-05	3.92E-05
2	0.125	-10.6	2.55E-05	2.55E-05
3	0.877	-7.9	3.89E-04	3.89E-04
4	0.212	-10.2	3.85E-05	3.85E-05
5	0.656	-8.7	1.59E-04	1.59E-04
6	0.466	-9.3	8.97E-05	8.97E-05
7	0.500	-9.2	9.90E-05	9.90E-05
8	0.107	-10.7	2.28E-05	2.28E-05
9	0.439	-9.4	8.27E-05	8.27E-05
10	0.475	-9.3	9.21E-05	9.21E-05
Mean=	**0.407**	**-9.54**	**1.04E-04**	**1.04E-04**
Exact Mean=	0.5	-9.22	2.00E-04	2.00E-04

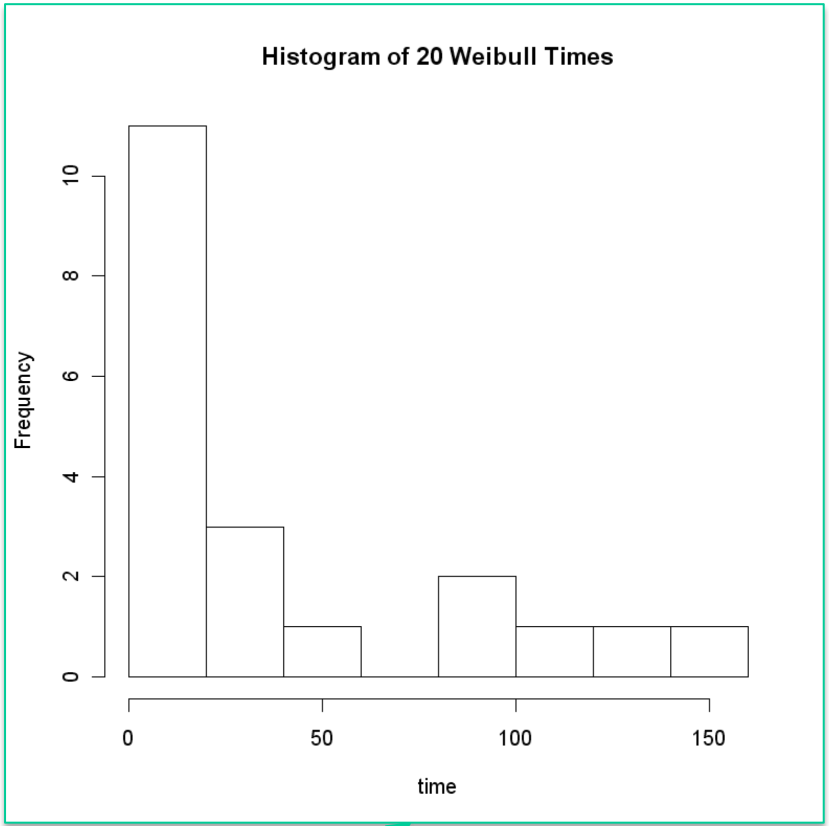
Use of Transformation: Exponential To Weibull

- Can generate Weibull samples from exponential samples
  - First generate n values from an exponential(λ) distribution, call them x_1 through x_n
  - Set each $t_i = (x_i)^{1/\alpha}$
  - The t_is will have a Weibull(α, λ) distribution
    - $f(t) = \alpha \lambda t^{\alpha - 1} \exp(-\lambda t^{\alpha})$

Waloddi Weibull

Use of Transformation: Exponential To Weibull





$$t = t^{1/\alpha}$$

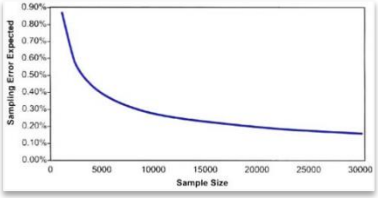
Determining Sample Size

- Let true distribution of Y have mean μ and variance σ^2
- Generate (large) sample, $y_1, \ldots, y_n$
- Estimate μ by sample mean, i.e. average of sample values, $\bar{y}$
- Approximate 95% confidence interval for μ is

$$\bar{y} \pm \frac{2s}{\sqrt{n}}$$

  – Here s is sample standard deviation, an estimate of σ

  – $\frac{s}{\sqrt{n}}$ is called the **standard error**



- So to estimate μ and cut the "error" by a factor of 2, n must be increased by a factor of 4
- Pragmatically, just keep track of the mean and 95th percentile
  – If relatively stable, you have enough samples

Sampling in DSW

- Each of the probability distribution pages has a section demonstrating Monte Carlo sampling
- For example, for Lognormal

D44 =LOGINV(C44,C5,C6)

Monte Carlo sampling using the Lognormal Distribution

Mean (from samples) = 9.54E-02 **Std. Dev. (from samples) = 1.25E-01**

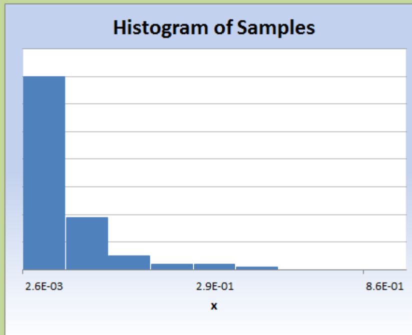
i	u_i	Samples from x_i	Is Sample ≤ x_high?
1	0.468	4.60E-02	1
2	0.057	1.02E-02	1
3	0.130	1.62E-02	1
4	0.146	1.74E-02	1
5	0.523	5.28E-02	1
6	0.353	3.41E-02	1
7	0.458	4.48E-02	1
8	0.276	2.75E-02	1
9	0.584	6.15E-02	1
10	0.898	1.78E-01	1
11	0.998	8.61E-01	1
12	0.006	4.06E-03	1
13	0.564	5.85E-02	1
14	0.560	5.78E-02	1
15	0.571	5.95E-02	1
16	0.828	1.28E-01	1
17	0.439	4.27E-02	1
18	0.599	6.40E-02	1
19	0.840	1.35E-01	1
20	0.020	6.42E-03	1
21	0.877	1.59E-01	1
22	0.886	1.66E-01	1
23	0.641	7.15E-02	1
24	0.869	1.53E-01	1
25	0.739	9.44E-02	1
26	0.849	1.40E-01	1
27	0.482	4.76E-02	1
28	0.474	4.67E-02	1
29	0.930	2.18E-01	1
30	0.845	1.38E-01	1
31	0.988	4.73E-01	1
32	0.567	5.90E-02	1

Pr(X ≤ x_high) from samples
1

5th from samples
0.01

50th from samples
0.05

95th from samples
0.31



Sampling in DSW (cont.)

Monte Carlo sampling using the Lognormal Distribution

Mean (from samples) = 1.69E-04 ← **=Average**

Std. Dev. (from samples) = 2.28E-04 ← **=Stdev**

i	u_i	amples from x_i	Is Sample ≤ x_high?
1	0.2[illegible]	[illegible]-05	1
2	0.9[illegible]	[illegible]-04	1
3	0.7[illegible]	[illegible]-04	1
4	0.135	2.68E-05	1
5	0.249	4.45E-05	1
6	0.941	6.28E-04	1
7	0.454	8.64E-05	1
8	0.840	3.21E-04	1
9	0.347	6.23E-05	1
10	0.700	1.84E-04	1
11	0.735	2.08E-04	1
12	0.[illegible]16	2.8[illegible]-04	[illegible]
13	[illegible]	2.[illegible]-04	[illegible]

[illegible]gh) from samples
1

5th from samples
0.00

50th from samples
0.00

95th from samples
0.00

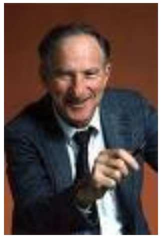
=RAND **=Loginv** **=IF** **=Percentile**

Section 5: Uncertainty Propagation in Risk Assessment

- Purpose
  - Illustrate, using Excel, how epistemic uncertainties in parameters are propagated through PRA models to obtain Bayesian estimates of risk metrics
- Objectives
  - Through examples using Excel, students will learn about
    - Monte Carlo sampling from distributions
    - Estimation of a “top event” probability or sequence frequency by propagation of distributions through a logic model
    - Simple Monte Carlo sampling and Latin Hypercube sampling

Risk

Stan Kaplan



John Garrick

- Recall the three questions a risk analysis attempts to answer:
  - What undesired things could happen?
  - What are their probabilities or frequencies?
  - What are their consequences?
- Must quantify answers, and assess uncertainty in these answers
- In LOSP example
  - Events
    - Initiating event could occur
    - Then EDG power system could successfully operate or it could fail
  - Consequences
    - Plant trip likely, perhaps worse if EDGs fail
  - Frequency of bad consequence is subject of this section

Overall Approach to Uncertainty Propagation

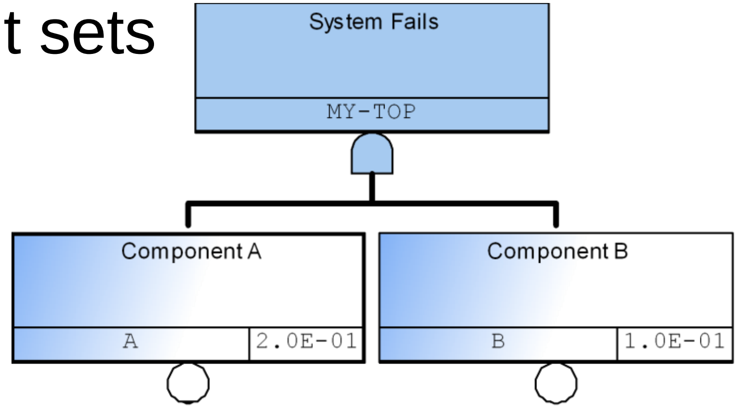
- In risk assessment, we estimate
  - Probability of “top event” (if looking at fault trees)
  - Frequency of “end state” (if looking at event trees)
- These estimates are typically based upon “minimal cut sets”
  - Minimal cut sets contain parameters such as
    - Failure rates
    - Probabilities of failure on demand
  - In the LOSP example, we develop λ_{SBO} as a (fairly complicated) function of λ_{LOSP}, p_{FTS}, and λ_{FTR}
  - We approximate the Bayes distribution of the end-state frequency as follows

Overall Approach to Uncertainty Propagation

1. Randomly sample a value of each basic parameter
   - This sample comes from the parameter's posterior distribution
2. Samples are used to quantify a desired
   - Top-event probability
   - End-state frequency
3. This process is repeated many times
   - Use new sampled values of the basic parameters on each iteration
   - Obtain many calculated values of desired result
   - Resulting values are a (pseudo)random sample from the Bayesian distribution of the top-event probability or end-state frequency
     - Together, they approximate the resulting distribution

Example

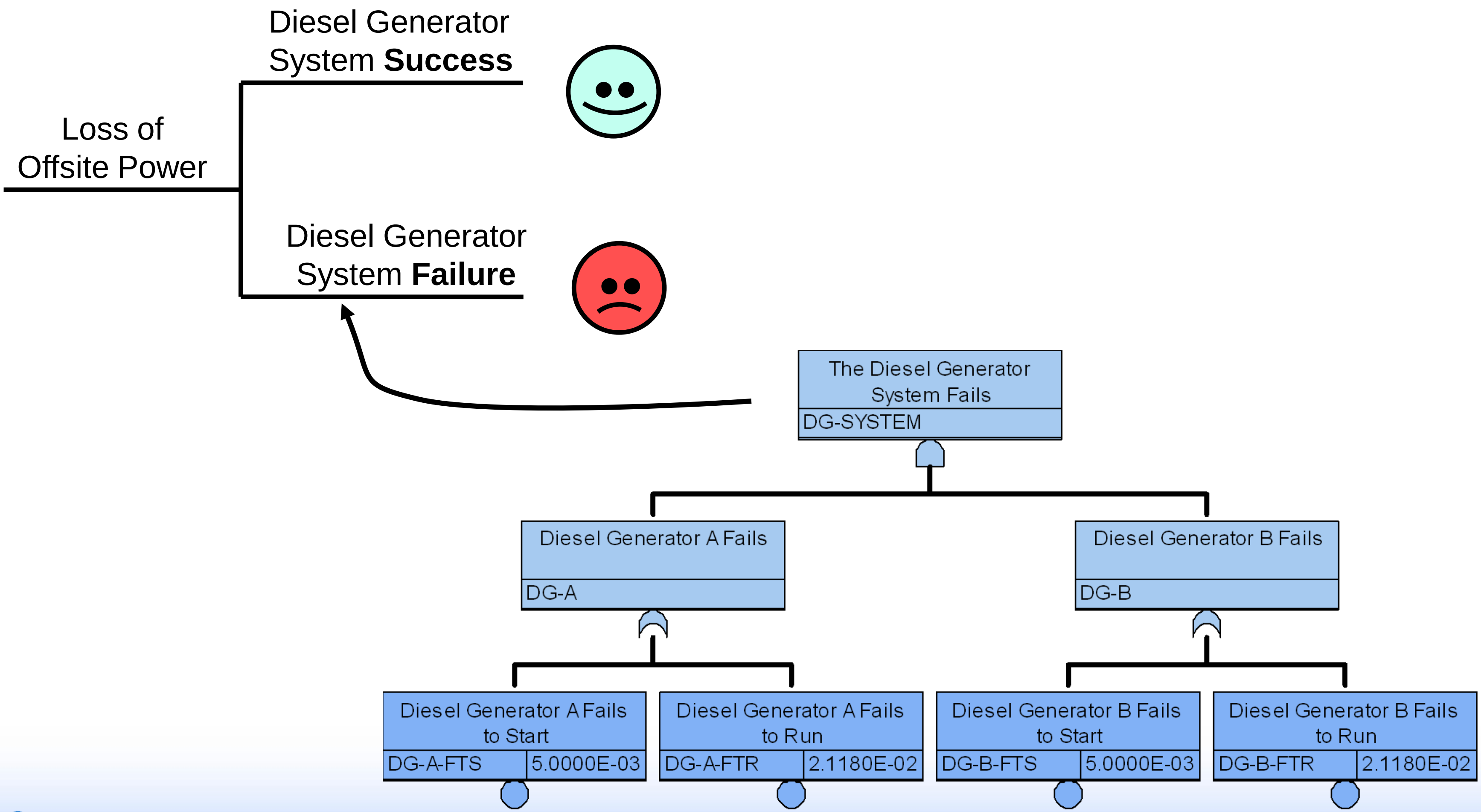
- Evaluate fault tree to determine cut sets
  - Single cut set → A*B
- In Excel, create samples for A, B
- Use samples to sample top event
  - MY-TOP = A × B
- Use MY-TOP samples to produce
  - E[ ]
  - Percentiles
  - Distribution
  - …



Random sampling for a cut set A*B

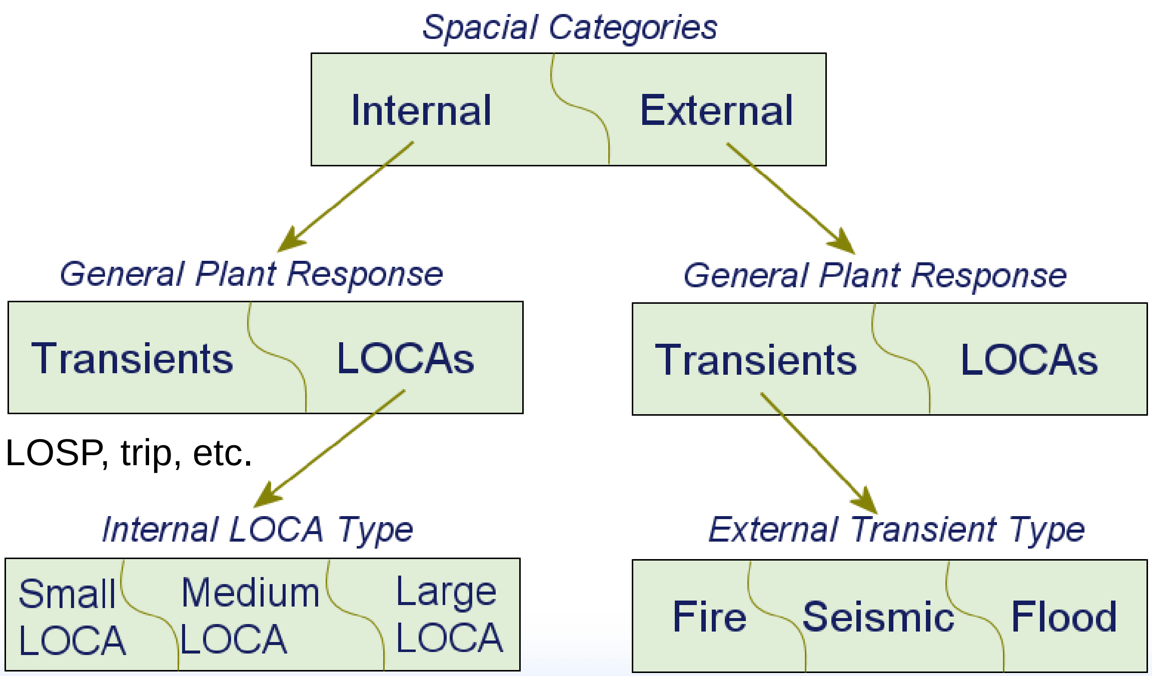
i	RAND	A	RAND	B	Cut Set
1	0.750	0.3	0.410	0.082	0.024605
2	0.612	0.2	0.398	0.080	0.019494
3	0.553	0.2	0.670	0.134	0.029653
4	0.331	0.1	0.994	0.199	0.026323
5	0.117	0.0	0.775	0.155	0.007236
6	0.729	0.3	0.081	0.016	0.004705
7	0.603	0.2	0.408	0.082	0.019696
8	0.359	0.1	0.900	0.180	0.025822
9	0.780	0.3	0.988	0.198	0.061584
10	0.622	0.2	0.784	0.157	0.038993
Mean=	**0.545**	**0.22**	**0.641**	**0.128**	**0.026**
Exact Mean=	0.5	0.2	0.5	0.1	

LOSP Example



LOSP Example

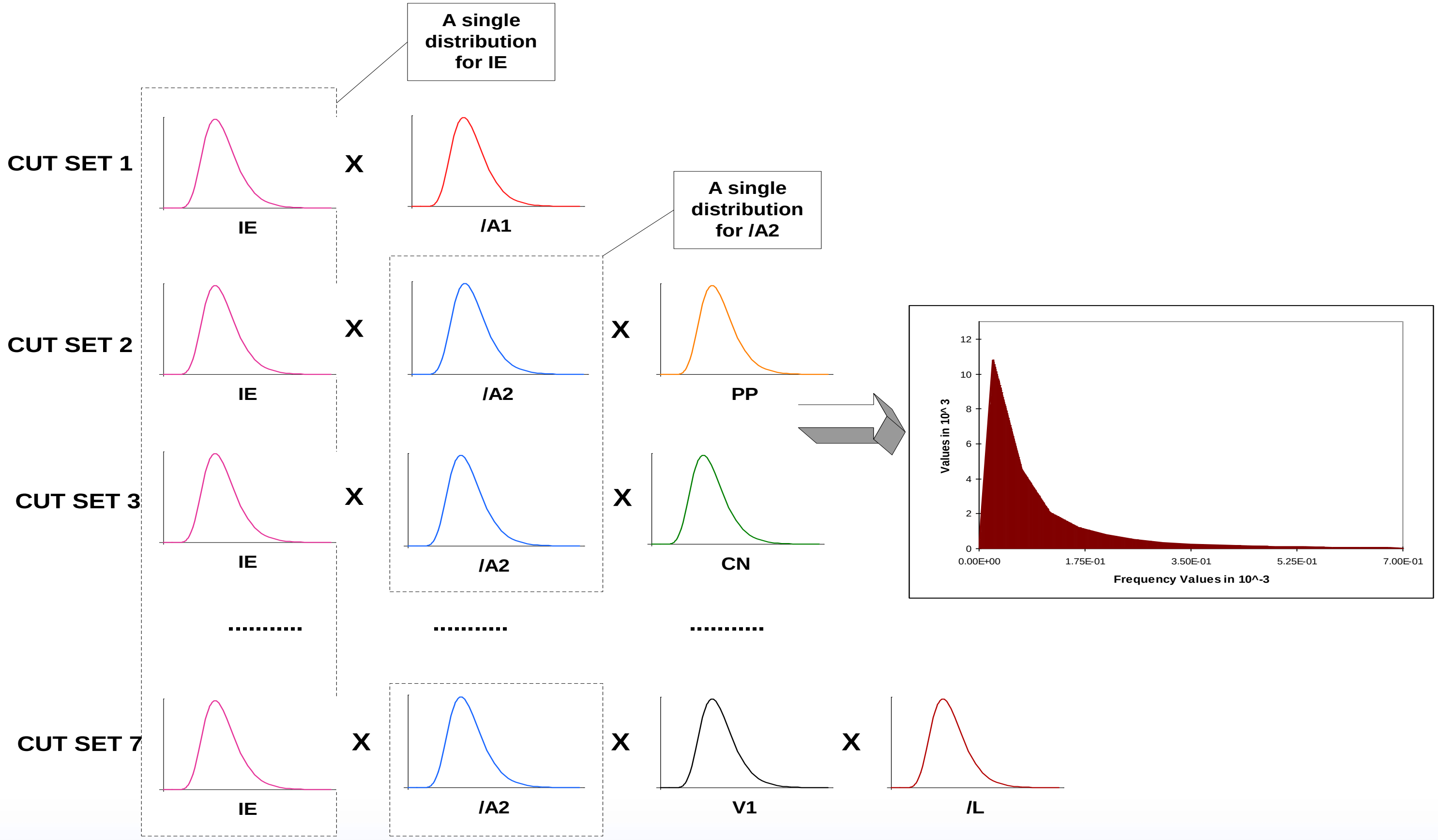
- Note that LOSP is just one initiating event…this analysis process is carried out for **all** results from **all** initiating events



PRA Minimal Cut Sets

- In every minimal cut set there are “basic events”
- Every basic event stored in PRA database typically has some epistemic uncertainty about the value used for the event
  - The **propagation** of this uncertainty through cut sets must be performed in order to understand the uncertainty in the overall result (e.g., CDF)
  - This uncertainty can be characterized via summary measures, such as mean value and 95th percentile

Schematic View of Uncertainty Propagation



Minimal Cut Sets in LOSP Example

- $\lambda_{SBO} = \lambda_{LOSP} \times$ Pr[EDG system fails]
- Pr[EDG system fails]

$$= Pr[(FTS_A \text{ and } FTS_B) \text{ or } (FTS_A \text{ and } FTR_B) \text{ or } (FTS_B \text{ and } FTR_A) \text{ or } (FTR_A \text{ and } FTR_B)]$$

$$\approx Pr(FTS_A \text{ and } FTS_B) + Pr(FTS_A \text{ and } FTR_B) + Pr(FTS_B \text{ and } FTR_A) + Pr(FTR_A \text{ and } FTR_B)$$

  - Using rare event approximation, Pr[EDG system fails]:

$$= Pr(FTS_A) \times Pr(FTS_B) + Pr(FTS_A) \times Pr(FTR_B) + Pr(FTS_B) \times Pr(FTR_A) + Pr(FTR_A) \times Pr(FTR_B)$$

assuming EDGs A and B fail independently

Minimal Cut Sets in LOSP Example

- Generic forms for basic event probabilities
  - $\Pr(FTS) = p_{FTS}$
  - $\Pr(FTR) = 1 - e^{-\lambda_{FTR} t_{mission}} \approx \lambda_{FTR} t_{mission}$
- $\Pr(FTS_A) \times \Pr(FTS_B) = ?$
  - p_{FTS}^2 ? (one estimated parameter)
  - $p_{FTS\text{-}A} \times p_{FTS\text{-}B}$? (two estimated parameters)

How Many Distinct Parameters in Example?

- If we distinguish between p_{FTS-A} and p_{FTS-B}
  - Assumes that the two p's differ significantly
  - Use only data from i^{th} EDG to estimate p_{FTS-i}
    - Have relatively more uncertainty in each estimate
    - Same prior for each p_{FTS-i} ?
- If we model only a single p_{FTS}
  - Assumes that the two p's are nearly equal
  - Uses data from both EDGs, and generic prior, to estimate the one p
    - Have relatively less uncertainty in the one estimate
    - Use generic prior

How Many Distinct Parameters in Example?

- If we assign independent Bayes distributions to $p_{FTS\text{-}A}$ and $p_{FTS\text{-}B}$
  - $E(p_{FTS\text{-}A} \times p_{FTS\text{-}B}) = E(p_{FTS\text{-}A}) \times E(p_{FTS\text{-}B})$
- If we assign Bayes distribution to p_{FTS}
  - $E(p_{FTS}^2) > E(p_{FTS}) \times E(p_{FTS})$
- So if the two parameters are really the same
  - Modeling them with independent distributions (uncorrelated sampling) yields too small a mean.
- In SAPHIRE, to force $p_{FTS\text{-}A}$ and $p_{FTS\text{-}B}$ to equal each other, i.e. to equal p_{FTS}
  - Assign them to a single **correlation class.**
- For additional information, see (Apostolakis and Kaplan, 1981)

Idaho National Laboratory

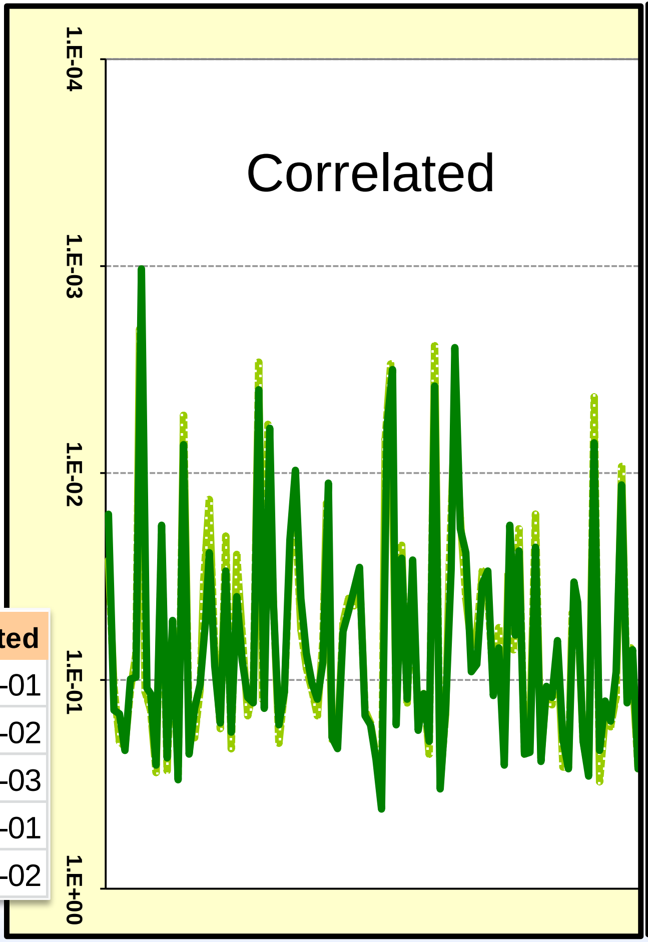
Un- versus Correlated Example

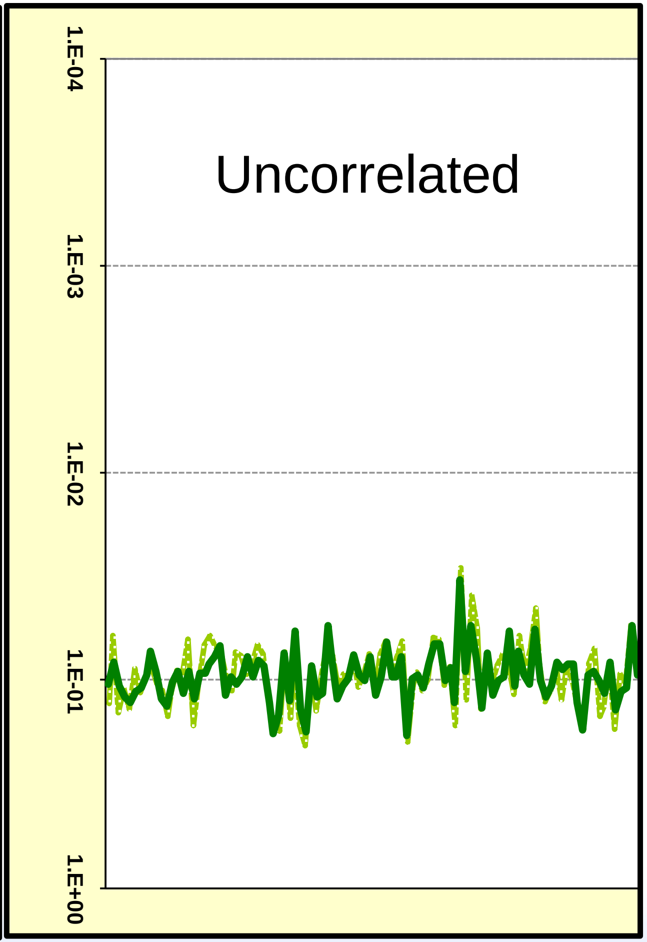
OR gate with 10 inputs

Each inputs is ~beta(1, 95)

Mean=0.01

Metric	Uncorrelated	Correlated
AVERAGE	9.9E-02	1.1E-01
MEDIAN	9.6E-02	9.8E-02
5th	5.8E-02	5.8E-03
95th	1.4E-01	2.7E-01
STD.DEV.	2.7E-02	8.9E-02



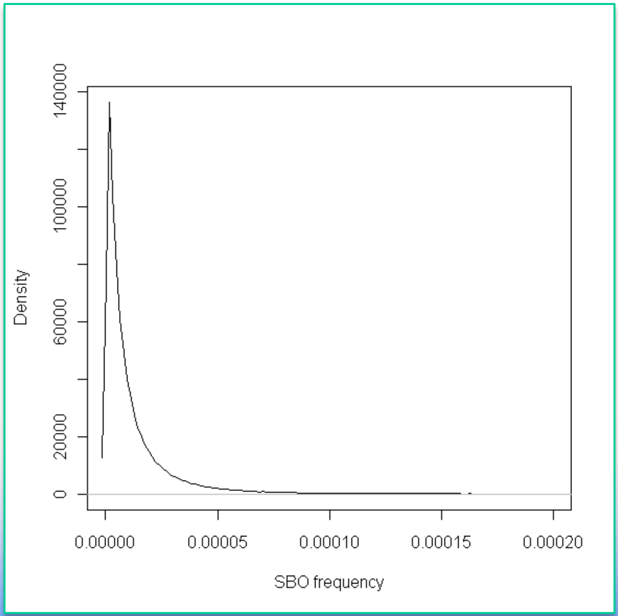


SBO Frequency in LOSP Example

- Assume single p_{FTS}, single λ_{FTR}
- $\lambda_{SBO} \approx \lambda_{LOSP} \times [p_{FTS}^2 + 2p_{FTS}\,\lambda_{FTR}t_{mission} + (\lambda_{FTR}t_{mission})^2]$
- Approximate the Bayes distribution of λ_{SBO} by a (large) random sample from the distribution

Microsoft Excel

losp demo problem.xls



Uncertainty Analysis for Other Applications

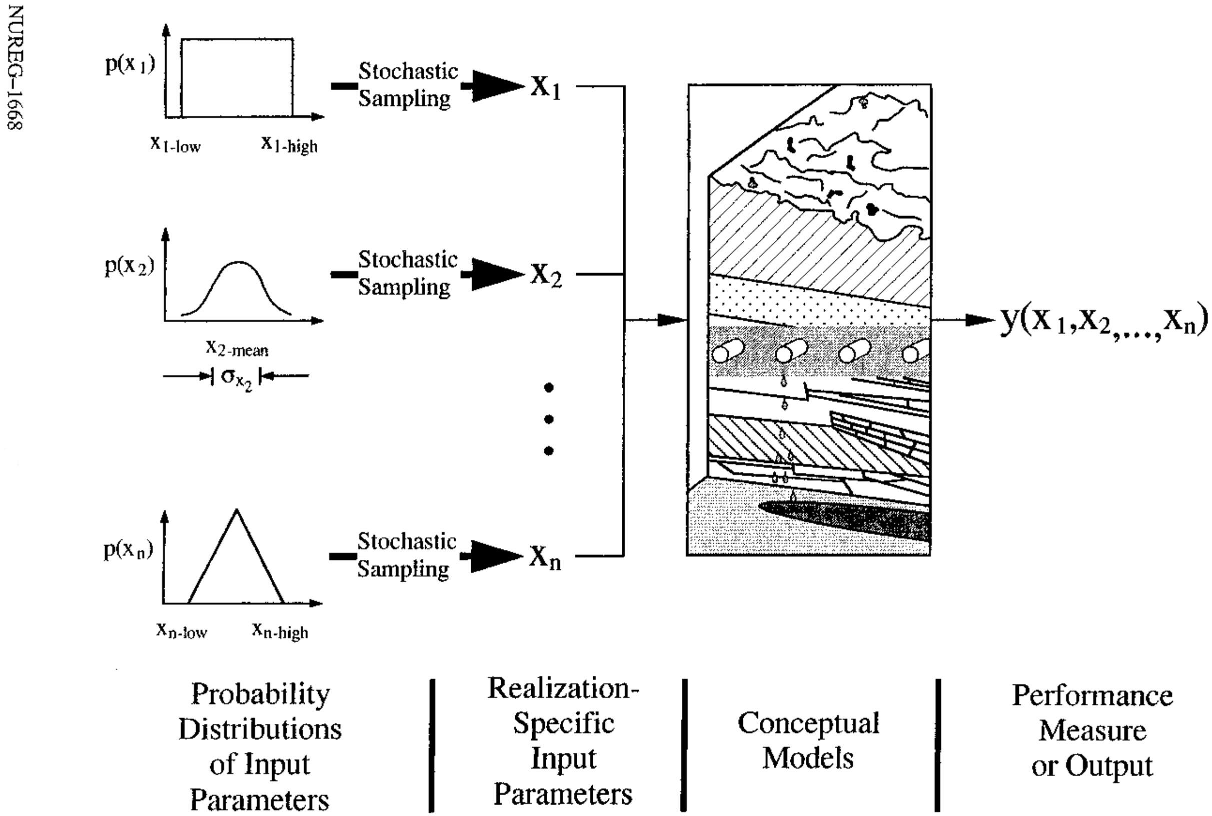


Figure 4-1. A diagram illustrating the use of the Monte Carlo method in performance assessment.

Propagation of Uncertainty

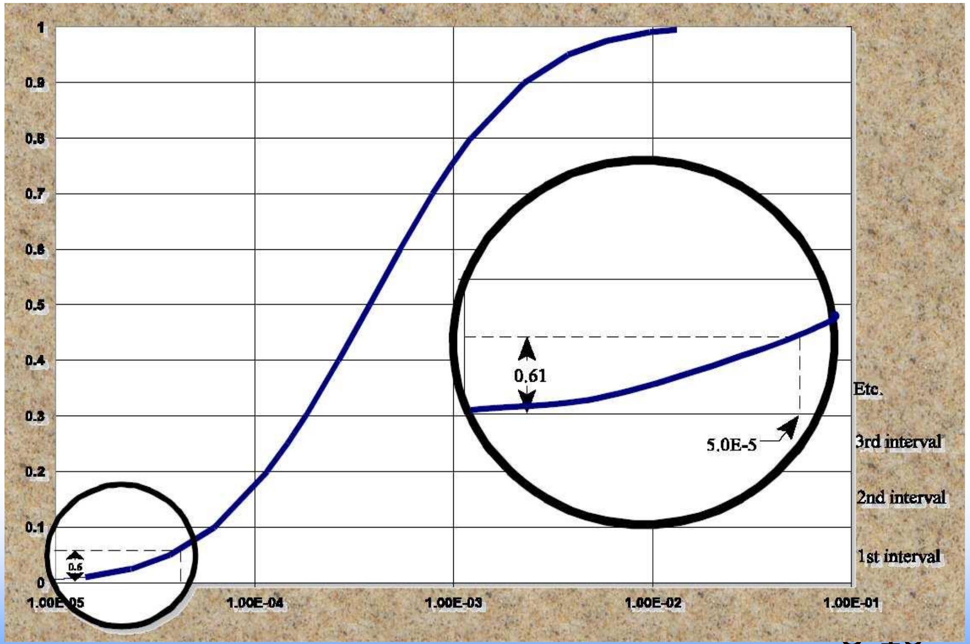
- To perform uncertainty analysis, the analyst must specify:
  1. The type of sampling
     - Simple Monte Carlo sampling (SMCS)
       - Called “Monte Carlo” in SAPHIRE
     - Latin hypercube sampling (LHS)
     - Grid sampling (not discussed in this course)
  2. The number of iterations (i.e., samples)
     - For example, if we specify 2,000 samples and there are 10 unique basic events, we generate 20,000 random numbers
  3. The random number generator seed value

Two Kinds of Sampling

- **Simple Monte Carlo Sampling (SMCS)**
- In simple Monte Carlo sampling, each parameter is sampled (pseudo)randomly from its specified distribution
  1. Each set of sampled values is entered into the minimal cut sets
  2. The frequency/probability of the top event is calculated for each set of sampled values
  3. This process is repeated many times (up to the number of samples specified)

Two Kinds of Sampling

- **Latin Hypercube Sampling (LHS)**
- In Latin hypercube sampling, each parameter is sampled in a stratified way, to guarantee that **each** portion of the range of the distribution is represented
- An example with 10 stratifications is shown
  - Within each portion, we randomly sample



Latin Hypercube Sampling

- For example, let us denote one parameter by p
  - Bayesian distribution of p is known: the posterior distribution of p based on prior information and relevant data
  - If 10 samples were to be taken
    1. p would be sampled randomly from interval ($p_{0.0}$, $p_{0.10}$), giving a value that we denote as p_1
    2. Again, sample randomly from interval ($p_{0.10}$, $p_{0.20}$), giving a value that we denote at p_2
    3. Repeat process until we have p_{10} [from interval ($p_{0.90}$, $p_{1.0}$)]
  - This is **stratified sampling**, in which the sampled points are forced to cover entire range of the distribution

Latin Hypercube Sampling

- After all parameters in the model have been sampled in this stratified way, they are randomly matched to each other
  - In example with λ_{LOSP}, p_{FTS}, and λ_{FTR}, one of the sampled values of each parameter would be chosen
  - However, they would be chosen so that the largest value of one parameter is **not** necessarily matched with largest or smallest values of other parameters
  - Instead, the choice of each pairing is random
  - For the chosen values, the top-event is calculated
  - Then another set of sampled parameter values is chosen, using values that have not been chosen yet
- In this way, a number (10 in this example) of values are calculated for the end-state frequency

Differences Between Sampling Types

- While there are computational differences between the two techniques (SMCS and LHS):
  - One should not be too concerned about which technique is selected for a particular analysis
  - Instead, one should be concerned about **convergence** of the numeric calculation
  - Convergence may be checked by noting change (or lack thereof) of uncertainty results as the number of samples is varied
- The samples from either method converge to the Bayes distribution of the end-state frequency or top-event probability

The Seed Value

- A seed value tells software where, in sequence of possible random numbers, to **start selecting** random numbers
  - The random number generator gives a sequence of "random" integers (which are typically converted to real numbers)
  - A seed of "51" may tell us to start at the i^{th} random integer
  - A seed of "1,236" may tell us to start at the j^{th} random integer
  - etc.
- Again, checking for **convergence** should make seed selection irrelevant
  - But, to reproduce analysis results, one must use the same seed and same number of samples

Accuracy of Sampling

- Accuracy of a simple random sample is roughly **proportional** to square root of sample size
  - For example, if λ_{SBO} is sampled from its distribution n times
    - Mean of the distribution is estimated by **average** of n sampled values (the sample mean), and this average has standard deviation proportional to $1/\sqrt{n}$
    - Estimate of this quantity is the standard error
      - A confidence interval equals the sample mean $\pm$ a multiple of the standard error
- LHS is more **complicated** than simple random sampling
  - But requires **fewer samples** for comparable accuracy
  - Therefore, it is justified if each calculation of top-event is expensive or time-consuming

Uncertainty Analysis Results

- Every result from the PRA is uncertain
  - Parameters used to quantify basic events (p for FTS, λ for FTR, etc.)
  - Initiating event frequency
  - System failure probability
  - Overall results such as core damage frequency and importance measures