BASTION TECHNOLOGIES

Expectation vs. Reality

The Sunrise Problem Applied to Probabilistic Risk Assessment

November 2022

Joseph Osowski, QD35, Bastion Technologies





Introduction

- **The "problem": *What is the probability that the sun will rise tomorrow?***
- **Evaluation of the sunrise problem shows the difficulty of using statistics for certain applications.**
- **When applied to Probabilistic Risk Assessment (PRA), the sunrise problem illustrates a pitfall that analysts should recognize and avoid, when possible.**

Agenda

- Introduction to the Sunrise Problem
- Further Statistical Analysis
- Expectation vs Reality
- Application to PRA
- Other Applications
- Avoiding the Pitfalls
- Summary

Introduction to the Sunrise Problem

- First introduced by Pierre-Simon Laplace.
- Given no prior information on the frequency of sunrises *p*, Laplace noted that a uniform probability distribution should be used and informed using observations.
- Applying Bayes' rule, a conditional probability can be shown as:

$$\Pr(\text{Sun rises tomorrow} \mid \text{It has risen } k \text{ times previously}) = \frac{\int_0^1 p^{k+1}\,dp}{\int_0^1 p^k\,dp} = \frac{k+1}{k+2}$$

Observations (years)	Observations (days)	Pr(Sunrise)	Pr(No sunrise)	"1 in" Days	"1 in" Years
28	10,220	0.999902	9.78E-05	10,222	28.01
80	29,200	0.999966	3.42E-05	29,202	80.01
1000	365,000	0.999997	2.74E-06	365,002	1,000.01
5000	1,825,000	0.999999	5.48E-07	1,825,002	5,000.01

Further Statistical Analysis

- Alternatively, could update a Jeffreys Prior beta distribution with observations.
  - Common method used to estimate failure probability of a binomial (pass/fail) variable.
  - Essentially adds ½ of a failure and ½ of a success to observed data.
    - Alpha= Failures + ½ , Beta = Successes + ½
- Pr(Sun rises tomorrow | It has risen *k* times previously) described by $\mathrm{Beta}\left(\frac{1}{2}, \frac{1}{2} + k\right)$

Observations (years)	Observations (days)	Pr(Sunrise) (median)	Pr(No sunrise)	"1 in" Days	"1 in" Years
28	10,220	0.999978	2.23E-05	44,931	123
80	29,200	0.999992	7.79E-06	128,371	351
1000	365,000	0.999999	6.23E-07	1,604,621	4,396
5000	1,825,000	1.000000	1.25E-07	8,023,101	21,981

Expectation vs Reality

- The sunrise problem, and attempts to “solve” it, show a dichotomy between what raw statistics predict and what we “know” to be true: the sun will continue to rise.

- Some factors contributing to this dichotomy include:
  - Reference class problem.
  - Attempts to predict a low-probability event far in the future.
  - Improper assumptions.
  - Statistical solutions do not include the bulk of our prior knowledge.

- Laplace recognized these shortcomings:

  But this number [the probability of the sun coming up tomorrow] is far greater for him who, seeing in the ***totality of phenomena*** *the principle regulating the days and seasons, realizes that nothing at the present moment can arrest the course of it.*

Application to PRA

Similar problems can arise in the reliability/PRA realm when analyzing failure risk.

Example:

A small boat manufacturer wants to know the probability that a rivet fails, creating a hole that could sink their boats.

Each boat uses 600 rivets, and the manufacturer has already built 6 boats with no occurrences of this failure.

____

Use Jeffrey's prior & Beta distribution:
Alpha = 0 + ½ = ½
Beta = 600*6 + ½ =3600.5

Pr(single rivet failure)= 6.32E-05
Pr(at least 1 rivet failure on a boat) = **0.037**

Application to PRA (cont.)

- Is this high (1 in 26) failure probability reasonable? Does the probabilistic expectation match reality?

- Factors contributing to the analysis' shortcomings:
  - Reference class problem.
    - Should more data be used from similar designs outside of the manufacturer history?
  - Extrapolating a low-probability event.
    - Low risk when looking at a single rivet, but risk increases due to the large number of rivets on each boat.
  - Improper assumptions.
    - Can we treat rivet failures as independent/random?
  - Statistical solution does not include the bulk of our prior knowledge.
    - Is the failure even possible barring external events? Has material strength/failure physics been analyzed?

Other Applications

- The shortcomings that the sunrise problem illustrates can be found in various other types of statistical analyses:

- Probabilistic Risk Analysis.
  - Shown in earlier example.
  - Similar outcomes may arise when analyzing long-duration reliability.
- Population estimates.
  - Extrapolating small local populations across large areas can be highly inaccurate.
- Predictions of weather or geological events.
  - Limited accuracy and obvious flaws of predictions across different locations or long time periods.

Avoiding the Pitfalls

- **Failing to recognize (or explain) a sunrise problem can lead to faulty assessments and loss of trust in PRA methodologies.**

- **Risk analysts should notice the “warning signs” of a sunrise problem:**
  - Low probability failure events evaluated over long durations or many points of failure.
  - Risk estimations of failures that have no history of occurrence.
  - Unexpectedly high risk of system failure compared to failure history.
  - Simplified failure criteria used to streamline statistical analysis (pass/fail).

- **Most shortcomings can be addressed by taking a wider view of the analysis.**
  - Does the result pass the “sanity check”?
  - What *actually* causes the failure to occur?
  - Can more/better data be used as an input?
  - Can extrapolation be limited?

Summary

- **At its core, the sunrise problem illustrates a notable misapplication of statistical prediction methods.**

- **PRA analyses similar to the sunrise problem often result in unusually high risk and typically occur when there is a combination of:**
  - Limited data.
  - Many potential points of failure.
  - Low failure probabilities.
  - Misunderstood failure requirements.

- **Occurrences of such problems in a risk assessment should be reviewed for these shortcomings and overall plausibility. Assessments may be improved with the addition of:**
  - Additional physics/material properties analysis.
  - Additional historical data.
  - Limited analysis scope and acknowledgement of shortcomings.

POC & References

- **Joseph.J.Osowski@nasa.gov**
  - 256-961-0751

- **E.T. Jaynes, *Probability Theory: The Logic of Science,* 2003.**

- **Additional information can be found on Wikipedia:**
  - https://en.wikipedia.org/wiki/Sunrise_problem

- **Images from Excel stock image library**