War-Winning Capabilities … On Time, On Cost

U.S. AIR FORCE

Combining Simulation with Flight Test through Bayesian Inference

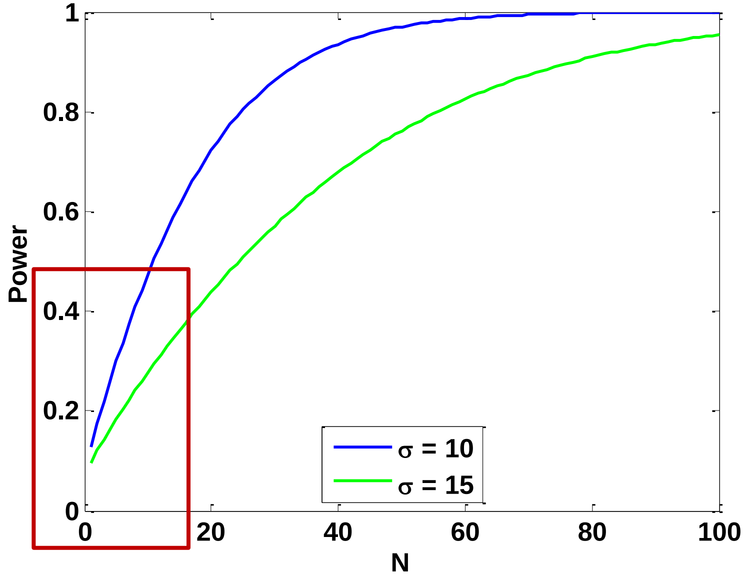
Daniel Carlson
771st TS
661-275-9976

Overview

- Motivation
- Bayesian Inference
- Bayesian vs. Frequentist Statistics
- Radar Warning Receiver (RWR) Example
- Flight Test Example
- Conclusion

Motivation

- **Flight test is limited due to various constraints such as cost and schedule**
- **Drawing conclusions with a high statistical confidence is difficult when number of test points is small**



Power = 1-β

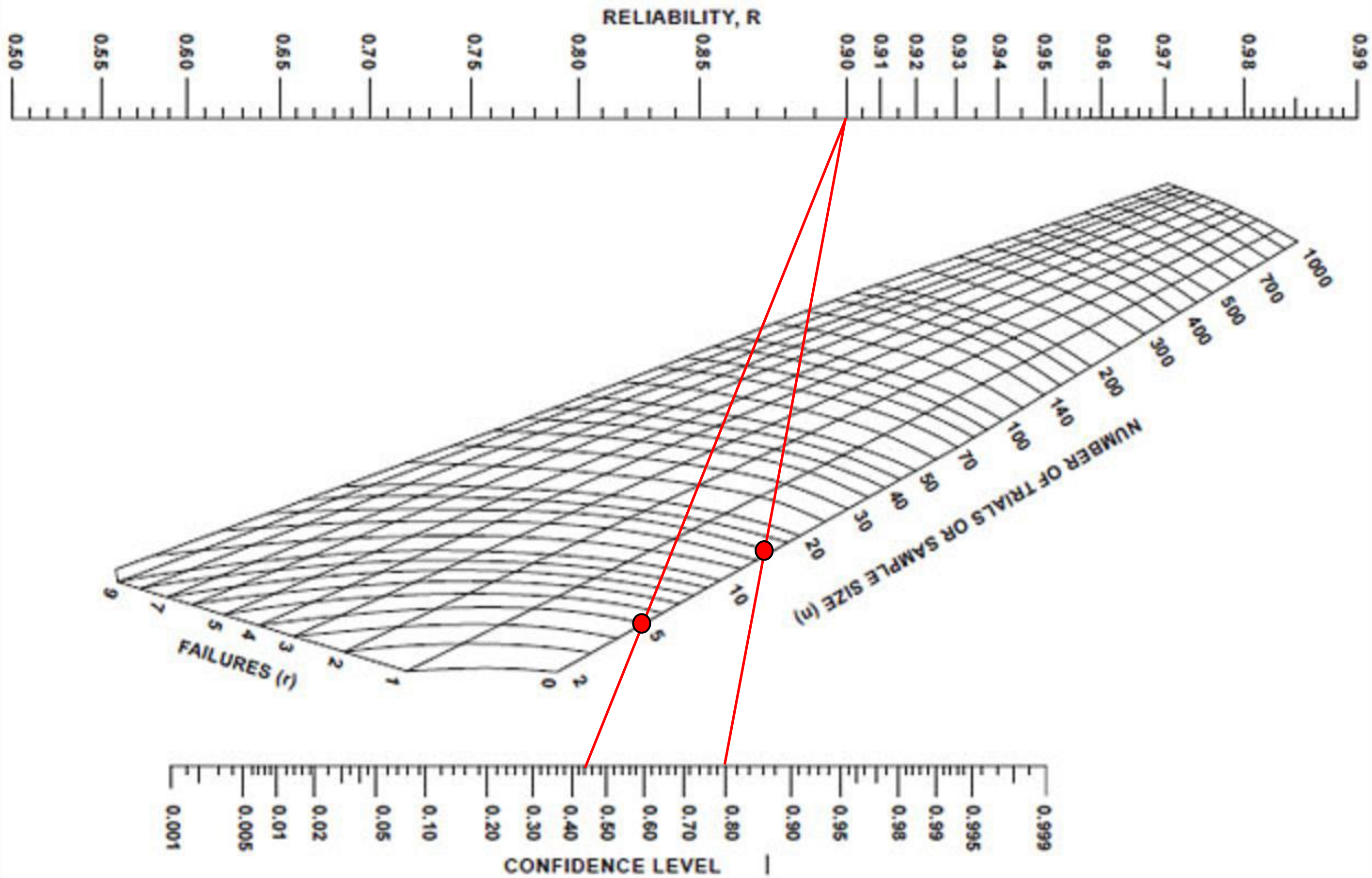
β = Probability of making a Type II error (false negative)

N = number of samples

α = .05 (significance level), μ_0=0 (null hypothesis), μ_{true}=5

Motivation

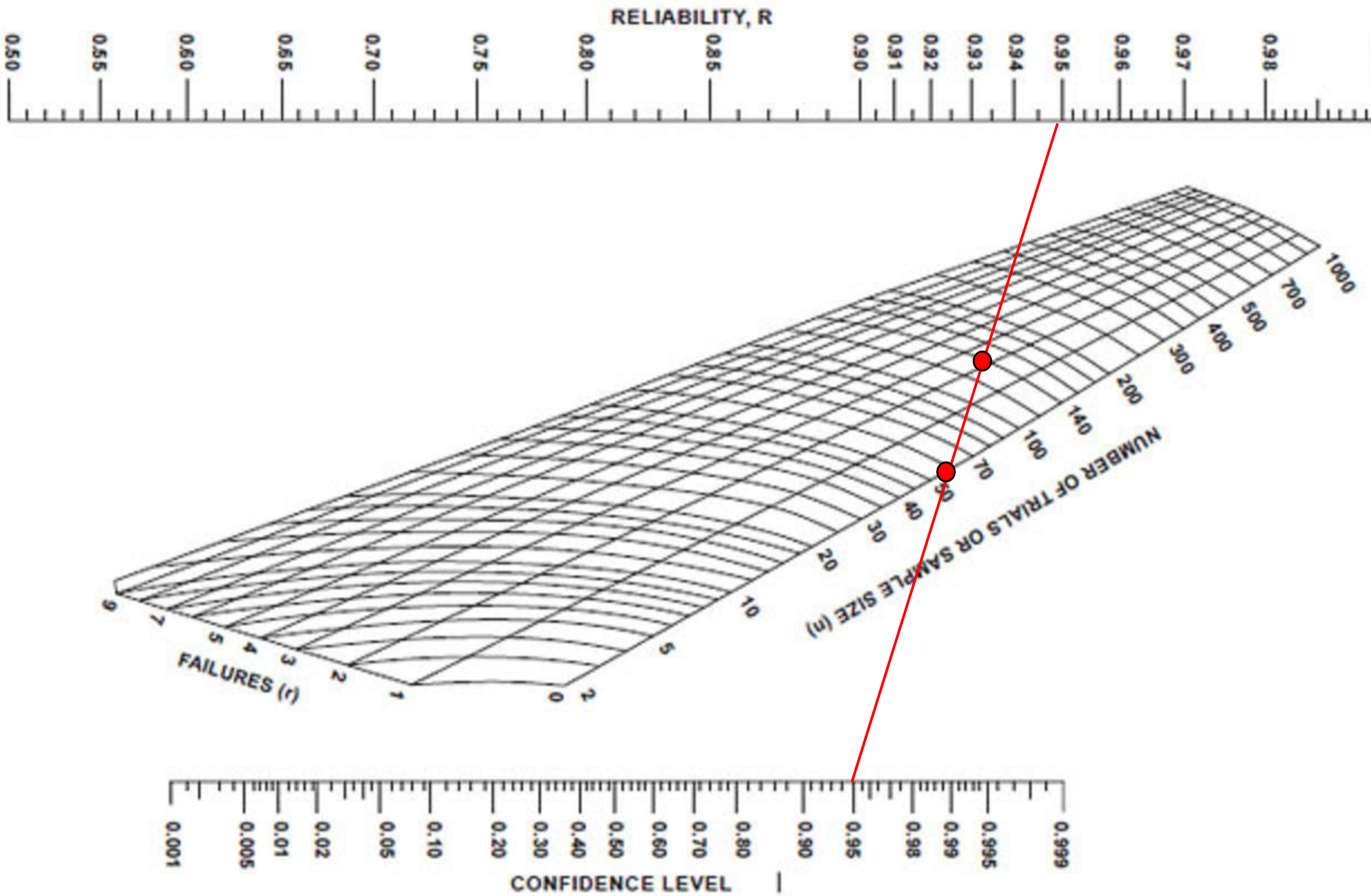
Binomial Nomograph



https://acc.dau.mil/CommunityBrowser.aspx?id=294528

Motivation

Binomial Nomograph



https://acc.dau.mil/CommunityBrowser.aspx?id=294528

Bayesian Inference

- **f(A|B) – posterior distribution of the variable A given the observed measurements of the variable B**
- **f(A) – assumed distribution of A (prior)**
- **f(B|A) – the distribution of the observed measurements of B assuming A is true**
- **f(B) – the distribution of the observed measurements B (likelihood)**

$$f(A|B) = \frac{f(A)\,f(B|A)}{f(B)}$$

Bayesian Inference

- **Calculate the following integral…**

$$\iint \lambda^n \beta^n \exp(-\lambda T^{*\beta}) \prod_{i-1}^{n} T_i^{\beta-1} \frac{\lambda^{k-1}\exp(-\lambda/\theta)}{\Gamma(k)\theta^k} \frac{\beta^{k'-1}\exp(-\beta/\theta')}{\Gamma(k')\theta'^{k'}} d\lambda d\beta$$

- **This is a Weibull likelihood for the time between failures, T^* is the termination time, gamma priors on λ, β**
- **No closed form solution**

Bayesian vs. Frequentist

- **Frequentist:**
  - Throw away any previous information
  - Estimate a parameter assumed to be an unknown constant
  - Hypothesis test, p-value, confidence intervals, power, type I and type II errors
- **Bayesian:**
  - Use prior information
  - Estimate a distribution for a parameter as a random variable
    - Admits hierarchical models
  - Conclusions driven by posterior density function

Bayesian vs. Frequentist

412th TW

- **Frequentist:**
  - 95% confidence interval: with a large number of samples 95% of the confidence intervals would contain the true mean value
- **Bayesian:**
  - 95% credible interval: there is an 95% probability that the true mean is within the interval

RWR Example

- **New RWR being tested for Response Time against an SA-X**
- **Only 10 data points from limited range blocks**
- **RWR manufacturer has provided these specs from historical data and modeling and simulation:**
  - **Predicted mean: μ=3s**
  - **Predicted standard deviation: σ=1s**
  - **Predicted distribution: log-normal**

Response Time (s)
2.98
2.70
2.79
2.35
3.86
2.28
3.87
4.52
2.52
2.31

RWR Example

Frequentist Method

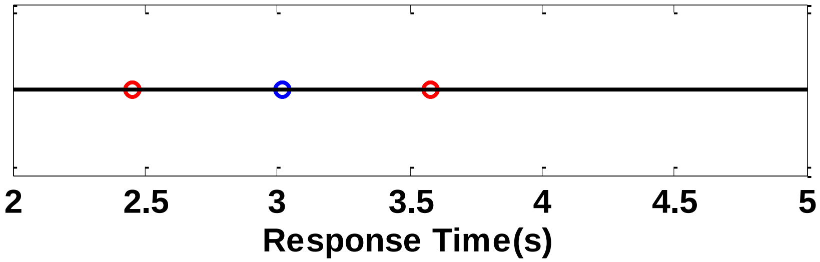
- **Calculate mean**

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}$$

- **Calculate 95% Confidence Interval**

$$CI(95\%) = \bar{x} \pm \frac{t_c s}{\sqrt{N}}$$

Response Time (s)
2.98
2.70
2.79
2.35
3.86
2.28
3.87
4.52
2.52
2.31

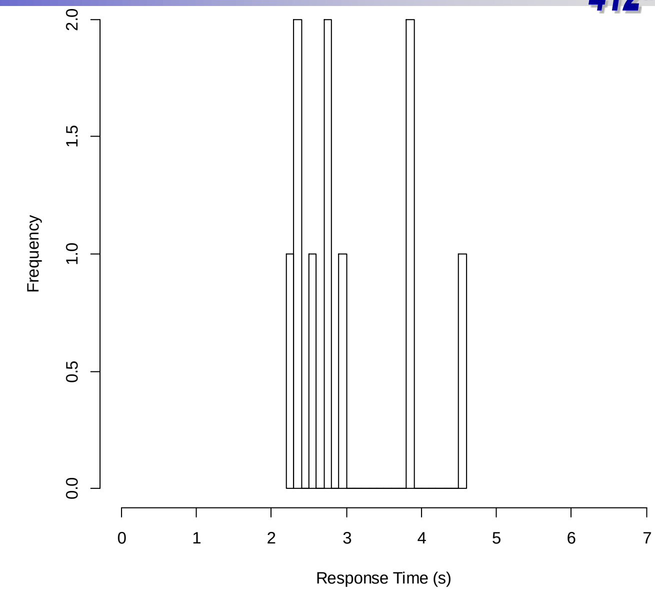


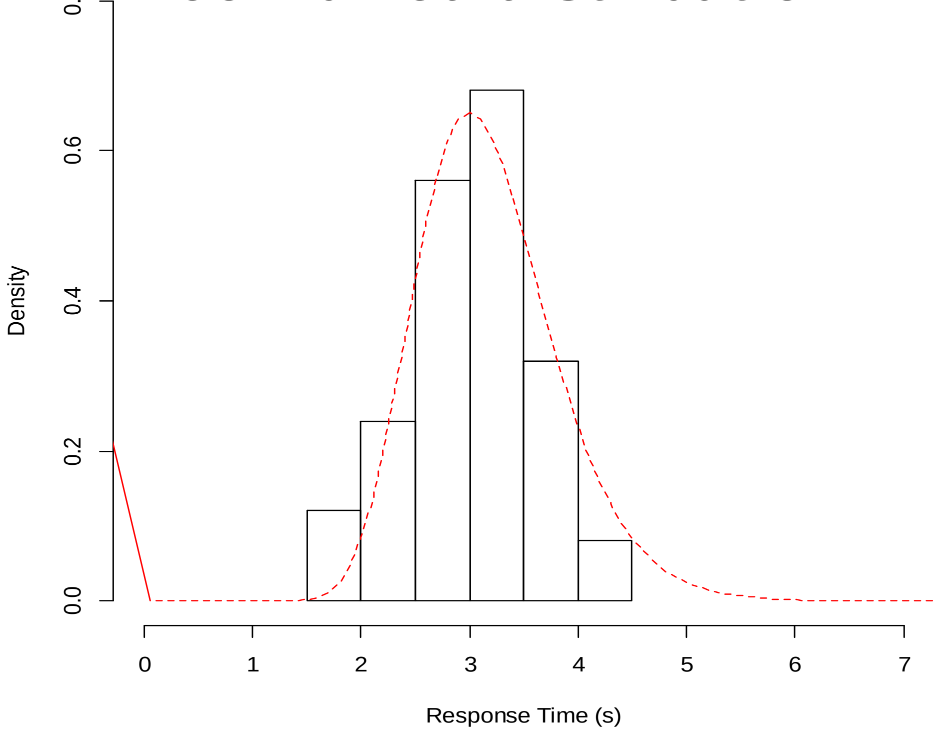
***t=upper 2.5% critical value for t distribution with n-1 degrees of freedom**

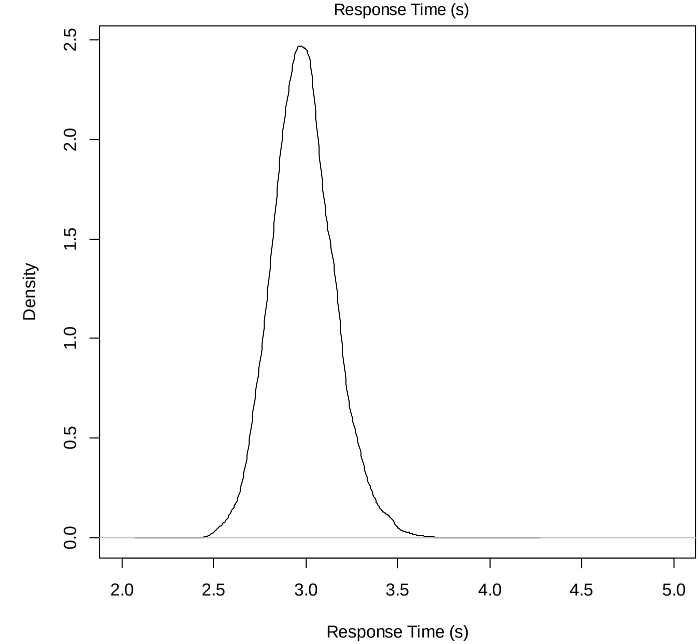
RWR Example

Bayesian Method

- Predicted distribution = prior
- 10 data points = likelihood
- Use Gibbs sampling to create combined distribution

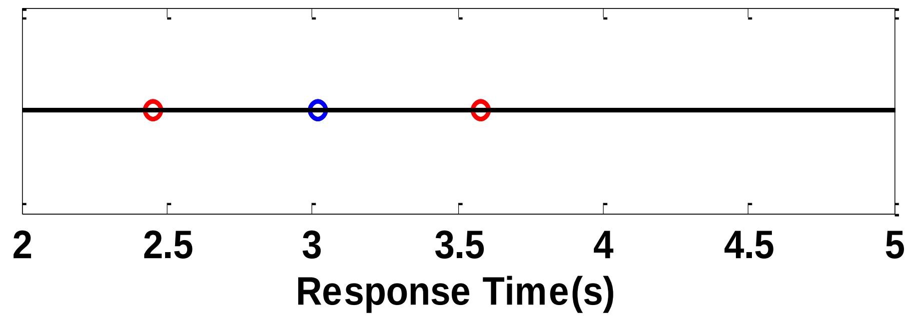




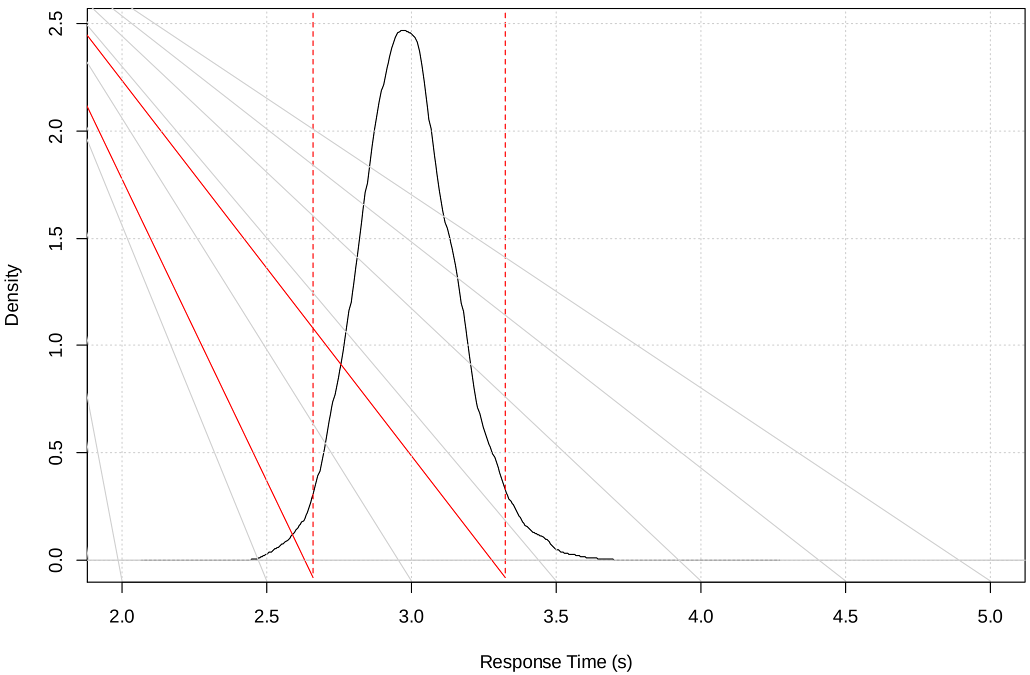


RWR Example

- **Frequentist Method:**
  - **Mean and 95% Confidence Interval**



- **Bayesian Method:**
  - **Posterior Distribution with 95% Credible Interval**



Frequentist 95% Confidence Interval	Bayesian 95% Credible Interval
2.45-3.58	2.66-3.32

Flight Test Example

- **Flight test program at the Air Force Test Center (AFTC) is measuring reduction in engagement range (RER) due to electronic attack**

$$\frac{R_{Dry} - R_{Wet}}{R_{Dry}}$$

- **Flight test has limited test points**
- **Additional runs are performed at hardware in the loop and other simulation facilities**

Flight Test Example

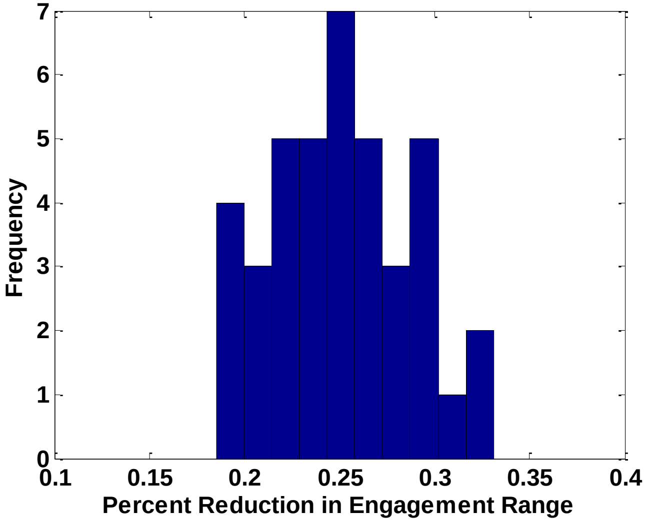
- **f(A|B) – Combined distribution to estimate RER**
- **f(A) – Simulation Data**
- **f(B|A) – Flight Test Data**
- **f(B) – Normalizing Factor (not needed)**

$$f(A|B) = \frac{f(A)f(B|A)}{f(B)}$$

Flight Test Example

- Flight test data
- 4 dry runs and 10 wet runs
- RER distribution created by all possible combinations of dry and wet runs

Dry				Wet				
100.21	98.18	104.26	97.24	78.26	69.75	73.41	74.22	77.94
				79.23	75.23	76.31	74.16	71.16

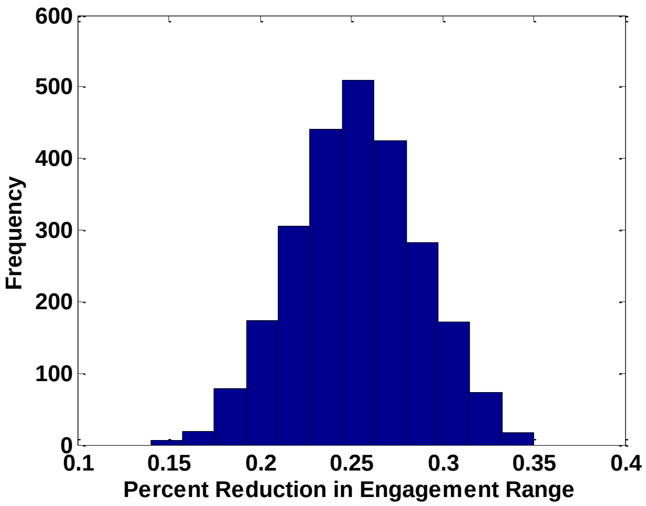


All data presented is fictitious

Flight Test Example

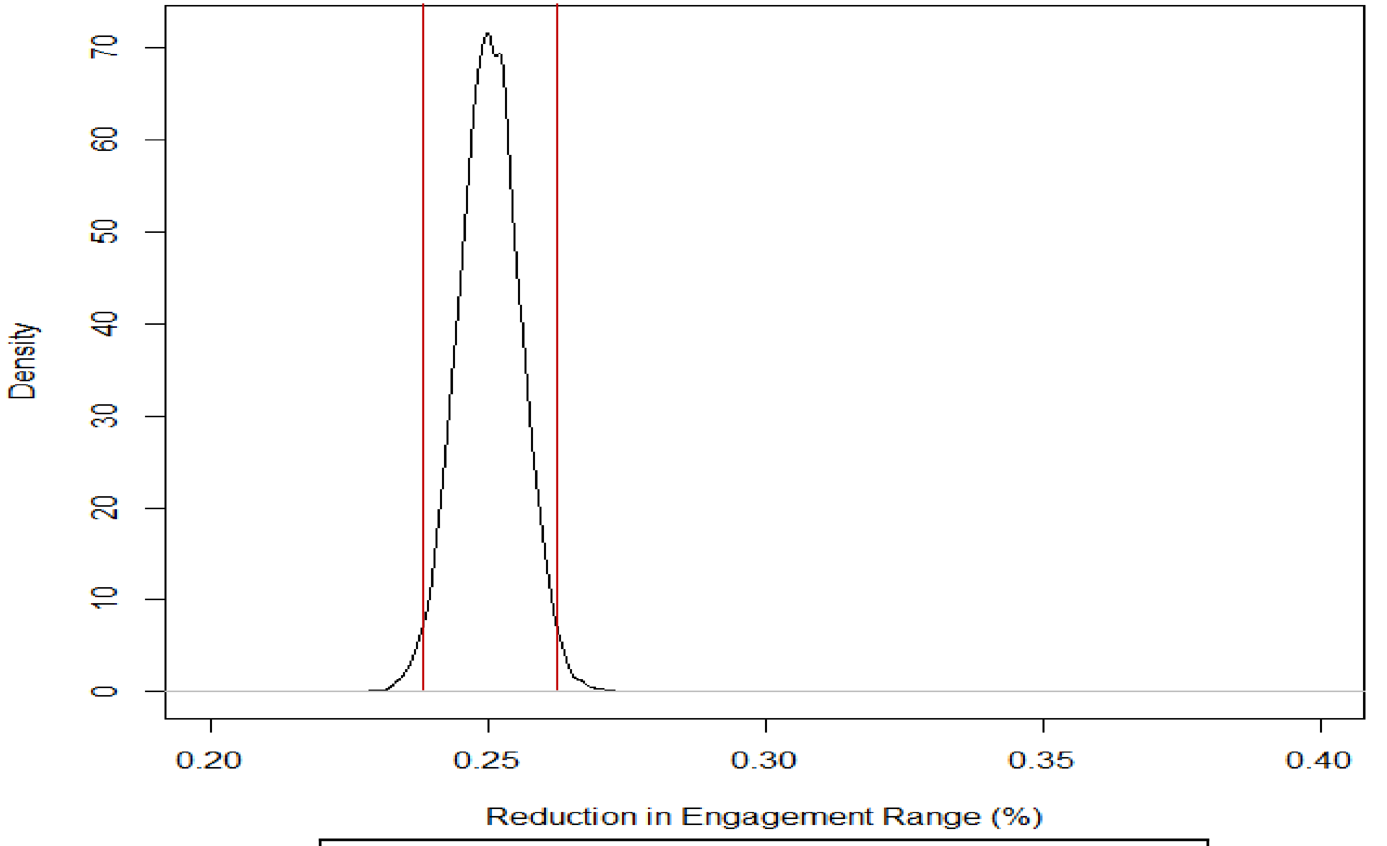
- **Simulation Data**
- **50 dry runs and 50 wet runs**
- **RER distribution created by all possible combinations of dry and wet runs**

Dry Simulation					Wet Simulation				
106.14	100.76	102.47	101.69	99.79	74.87	77.31	75.61	74.89	73.99
103.20	100.93	97.15	101.01	96.34	73.18	75.11	76.64	75.86	70.13
100.30	104.80	105.71	98.84	102.37	79.87	78.02	74.28	74.31	73.18
103.61	97.16	100.72	99.14	98.52	81.42	79.57	72.08	76.26	78.40
99.18	96.63	101.75	100.35	100.57	74.13	76.64	77.17	74.59	76.63
104.07	99.41	102.91	101.74	97.82	74.05	71.89	72.76	70.22	71.72
103.61	103.36	100.10	105.13	95.65	72.52	72.18	76.85	78.80	68.98
104.79	95.28	102.39	100.08	102.43	78.53	76.38	75.91	72.63	70.88
100.10	98.23	101.04	95.58	102.89	78.51	74.70	77.33	74.90	77.21
102.63	96.44	99.48	100.68	97.06	75.29	77.21	70.29	76.12	76.50



Flight Test Example

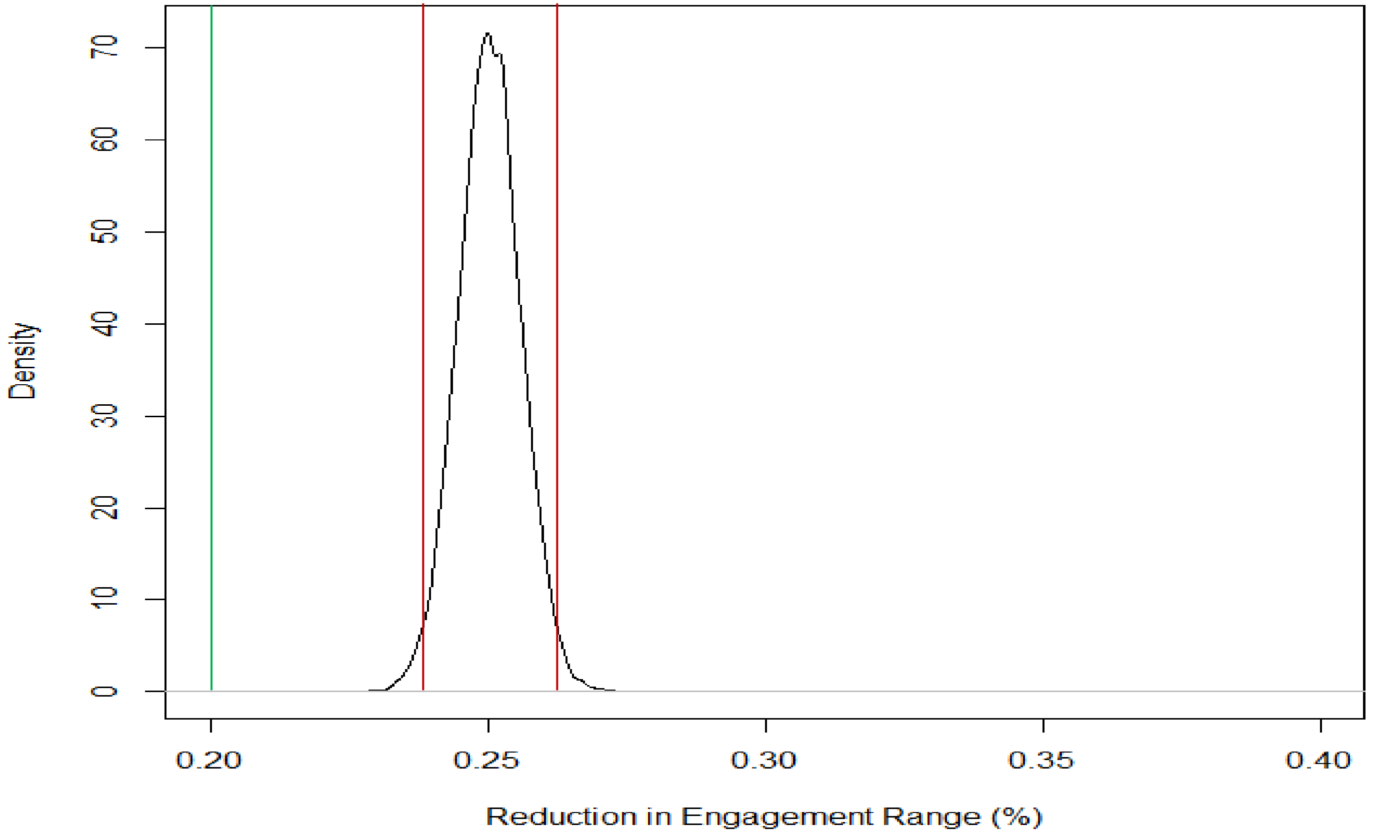
- **Combined distribution created using Bayesian Inference**



Bayesian 95% Credible Interval
0.239-.261

Flight Test Example

- **Combined distribution created using Bayesian Inference**



Bayesian 95% Credible Interval
0.239-.261

Conclusion

412th TW

- **Bayesian analysis in not the perfect solution for every problem**
- **Bayesian analysis has the unique ability to incorporate prior information in statistical analysis**
- **Bayesian methods provide techniques for calculating statistical confidence values from 2 data sets**
- **Can provide a way of making statistically significant calculations with small data sets**
- **May be applied to situations like flight test where resources are limited and number of data points is small**

References

- Agresti, Alan, and David B. Hitchcock. "Bayesian Inference for Categorical Data Analysis." *Statistical Methods and Application (Journal of the Italian Statistical Society)*, 2005.
- Bauwens, Luc, and Michel Lubrano. "Bayesian Inference on GARCH Models Using the Gibbs Sampler." *Econometrics Journal*, 1998: C23-C46.
- Bayes, Thomas. "An Essay Toward Solving a Problem in the Doctrine of Chances." *Philosophical Transactionsof the Royal Society* 53 (1763): 370-418.
- Brownlow, James. "Stastically Defensible Test & Evaluation: Bayesian Reliability Analysis." *Air Force Systems Engineering Conference.* AFFTC-PA-11600, 2011.
- Carrier, Richard C. "Bayes' Theorem for Beginners: Formal Logic and Its Relevance to Historical Method." *Sources of the Jesus Tradition: An Inquiry.* Amherst, 2008.
- Friston, J. K., W. Penny, C. Phillips, S. Kiebel, G. Hinton, and J. Ashburner. "Classical and Bayesian Inference in Neuroimaging: Theory." *NeuroImage* 16 (2002): 465-483.
- Geweke, John. "Posterior Simulators in Econometrics." Working Paper 555, Research Department, Federal Reserve Band ok Minneapolis, 1995.
- Kececioglu, Dimitri. *Reliability & Life Testing Handbook.* Lancaster, PA: DEStech Publications, 2002.
- Laplace, Pierre Simon. "Mémoire sur la Probabilité des Causes par les Evénements." *Mémoires de l'Académie Royale des Sciences*, 1774: 621-656.
- O'Neill, Philip D. "A Tutorial Introduction to Bayesian Inference for Stochastic Epidemic Models Using Markov Chain Monte Carlo Methods." *Mathematical Boisciences*, no. 180 (2002): 103-114.
- Tipping, Michael E. "Bayesian Inference: An Introduction to Principles and Practice in Machine Learning." In *Advanced Lectures on Machine Learning*, by O. Bousquet, U. von Luxburg and G. Ratsch, 41-62. Springer, 2004.
- Vallverdu, Jordi. "The False Dilemma: Bayesia vs. Frequentist." *Electronic Journal for Philosophy*, 2008.
- Yang, Yang. "Bayesian Inference for Heirarchical Age-Period_Cohort Models of Repeated Cross-Section Survey Data." *Sociological Methodology* 36, no. 1 (2006).

Questions?