Combining Information for Reliability Assessment

Alyson Wilson, Ph.D.

Department of Statistics
North Carolina State University
agwilso2@ncsu.edu

April 12, 2016

Acknowledgments

My research is funded through the

Science of Test Consortium
Director, Operational Test and Evaluation
Department of Defense

Acknowledgments

- Christine Anderson-Cook, Los Alamos National Laboratory
- Caleb Browning, North Carolina State University
- Rebecca Dickinson, Institute for Defense Analyses
- Laura Freeman, Institute for Defense Analyses
- Kassie Fronczyk, Institute for Defense Analyses
- Randy Griffiths, Army Test and Evaluation Command
- Jiqiang Guo, NPD Group
- Michael Hamada, Los Alamos National Laboratory
- Aparna Huzurbazar, Los Alamos National Laboratory
- Harry Martz, Los Alamos National Laboratory (retired)
- Shane Reese, Brigham Young University
- Brian Weaver, Los Alamos National Laboratory
- Roger Zoh, Texas A&M University

Overview

- Motivation and Challenges
- Bayesian Binomial Assurance Test
- Combining Data
- Future Challenges

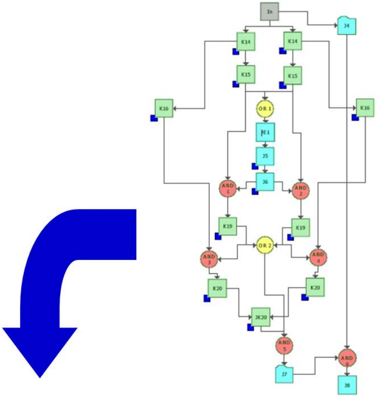
System Reliability

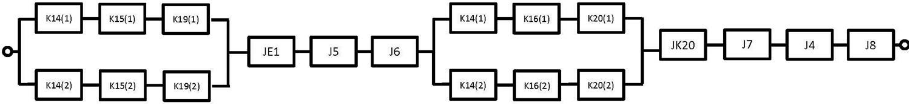
Heterogeneous Information

- DoD systems experience system design, contractor testing, developmental testing, and operational testing.
- Later in the lifecycle, we see new variants of systems, life extension programs, ...
- The goal of science-based stockpile stewardship is the assessment of safety and reliability in aging warheads in the absence of nuclear testing.

Statistical Challenges

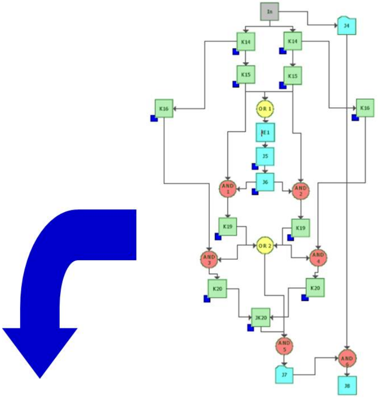
- Data
  - Multilevel
  - Multiple types: binary, lifetime, degradation, expert judgement, computer model
- Systems
  - Representation
  - Assessment: model checking and diagnostics, model fit
- Planning data collection

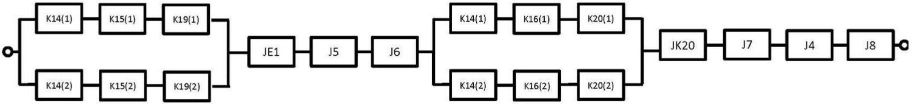




Statistical Challenges

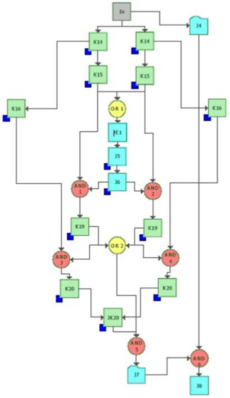
- Data
  - Multilevel
  - Multiple types: binary, lifetime, degradation, expert judgement, computer model
- Systems
  - Representation
  - Assessment: model checking and diagnostics, model fit
- Planning data collection

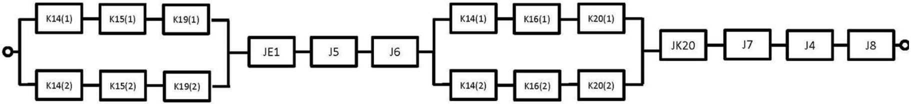




Statistical Challenges

- Data
  - Multilevel
  - Multiple types: binary, lifetime, degradation, expert judgement, computer model
- Systems
  - Representation
  - Assessment: model checking and diagnostics, model fit
- Planning data collection





Reliability Demonstration and Assurance

- Example: Using minimal assumptions, to demonstrate that reliability at time t_0 hours is 0.99, with 90% confidence, requires testing at least 230 units for t_0 hours with zero failures. To have a 80% chance of passing the test, requires that the true reliability be approximately 0.999.
- For complicated, expensive systems, traditional reliability demonstration is usually not practical
- Reliability assurance is the alternative: Use whatever relevant knowledge you have in a principled Bayesian approach to plan the test

Bayesian Binomial Test Plan

- Suppose we want to develop a *Bayesian binomial test plan*.
- We want to determine (n, c) where n is the test sample size and c is the number of systems allowed to fail before the "test is failed."
- What criteria do we use to choose n and c?

Test Criteria

There are two errors we could make:

- We could decide the "test is failed" when the system reliability π is higher than a specified π_P

Posterior Producer's Risk: Choose a test plan so that if the test is failed, there is a small probability that the reliability at t_I (the time of interest) is high

Test Criteria

There are two errors we could make:

- We could decide the "test is failed" when the system reliability π is higher than a specified π_P

Posterior Producer's Risk: Choose a test plan so that if the test is failed, there is a small probability that the reliability at t_I (the time of interest) is high

Test Criteria

- We could decide the "test is passed" when the system reliability is lower than a specified π_C.

Posterior Consumer's Risk: Choose a test plan so that if the test is passed, there is a small probability that the reliability at t_l is low

- Reliable Life Criterion: Choose a test plan so that if the test is passed, there is a high probability that the $1 - \alpha$ quantile of the distribution is greater than $t = 1$

Test Criteria

- We could decide the "test is passed" when the system reliability is lower than a specified π_C.

Posterior Consumer's Risk: Choose a test plan so that if the test is passed, there is a small probability that the reliability at t_l is low

- Reliable Life Criterion: Choose a test plan so that if the test is passed, there is a high probability that the $1 - \alpha$ quantile of the distribution is greater than $t = 1$

Posterior Producer's Risk

$$
\begin{aligned}
\textit{Posterior Producer's Risk} &= \mathbf{P}(\pi \geq \pi_P \,|\, \textit{Test Is Failed}, \mathbf{x}) \\
&= \int_{\pi_P}^{1} p(\pi \,|\, y > c, \mathbf{x}) d\pi \\
&= \int_{\pi_P}^{1} \frac{f(y > c \,|\, \pi) p(\pi \,|\, \mathbf{x})}{\int_0^1 f(y > c \,|\, \pi) p(\pi \,|\, \mathbf{x}) d\pi} \, d\pi \\
&= \frac{\int_{\pi_P}^{1} \left[\sum_{y=c+1}^{n} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) d\pi}{\int_0^1 \left[\sum_{y=c+1}^{n} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) d\pi} \\
&= \frac{\int_{\pi_P}^{1} \left[1 - \sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) d\pi}{1 - \int_0^1 \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y}\right] p(\pi \,|\, \mathbf{x}) d\pi}
\end{aligned}
$$

Posterior Consumer's Risk

$$
\begin{aligned}
\textit{Posterior Consumer's Risk} &= \mathbf{P}(\pi \leq \pi_C \,|\, \textit{Test Is Passed}, \mathbf{x}) \\
&= \int_0^{\pi_C} p(\pi \,|\, y \leq c, \mathbf{x}) d\pi \\
&= \int_0^{\pi_C} \frac{f(y \leq c \,|\, \pi) p(\pi \,|\, \mathbf{x})}{\int_0^1 f(y \leq c \,|\, \pi) p(\pi \,|\, \mathbf{x}) d\pi} \, d\pi \\
&= \frac{\int_0^{\pi_C} \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi \,|\, \mathbf{x}) d\pi}{\int_0^1 \left[\sum_{y=0}^{c} \binom{n}{y} (1-\pi)^y \pi^{n-y} \right] p(\pi \,|\, \mathbf{x}) d\pi}.
\end{aligned}
$$

We evaluate these integrals using Monte Carlo integration and draws from the posterior distribution of $p(\pi \,|\, \mathbf{x})$.

When be Bayesian?

From Meeker and Hong (2014), "Many of the applications described in this article and particularly in this concluding section will require combining information from different sources (e.g., data, inexact physics-based knowledge, and certain kinds of expert opinion). Additionally, statistical models being used will often contain multiple sources of variability. *Bayesian statistical methods provide a natural approach for combining such information*."

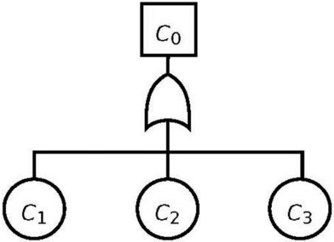
Why be Bayesian

- Allows incorporation of prior information
  - May supplement limited data
  - May provide improvements in cost or precision
  - Provides a formal framework to think about how to combine information
- Computational simplifications
  - Censored data
  - When framed as a Bayesian problem, complex models can often be fit relatively easily using Markov chain Monte Carlo or other computational algorithms.
- Straightforward to produce estimates and credible intervals for complicated functions of model parameters (e.g., predictions, probability of failure, quantiles of lifetime distribution)
- ~~Philosophical~~ Pragmatic

Multilevel Data

Multilevel Pass/Fail Data, Series System

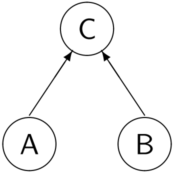
- Informaton collected at C_0, C_1, C_2, and C_3
- Information at C_0 provides partial information about C_1, C_2, and C_3
- Goal: simultaneous inference about system and component reliabilities



	Successes	Failures	Trials
Component 1	8	2	10
Component 2	7	2	9
Component 3	3	1	4
System	10	2	12

Example

Multilevel Pass/Fail Data, Bayesian Network, Reliability Changing with Time



Age	A	B	C
1	19/19	35/35	15/16
2	-	47/48	14/14
3	16/19	37/38	12/14
4	12/12	-	-
5	-	44/45	13/14
6	-	35/37	11/12
7	9/13	-	-
8	-	33/42	5/16
9	-	-	12/19
10	3/10	30/39	8/14

Model

- "Component" probabilities, logistic regression,

$$p_A(t) = \frac{\exp(\alpha_A + \beta_A t)}{1 + \exp(\alpha_A + \beta_A t)}$$
$$p_B(t) = \frac{\exp(\alpha_B + \beta_B t)}{1 + \exp(\alpha_B + \beta_B t)}$$

- Conditional probabilities

$$\tau_{11} = \mathbf{P}(C = 1 \mid A = 1, B = 1)$$
$$\tau_{10} = \mathbf{P}(C = 1 \mid A = 1, B = 0)$$
$$\tau_{01} = \mathbf{P}(C = 1 \mid A = 0, B = 1)$$
$$\tau_{00} = \mathbf{P}(C = 1 \mid A = 0, B = 0)$$

Model

- "Component" probabilities, logistic regression,

$$p_A(t) = \frac{\exp(\alpha_A + \beta_A t)}{1 + \exp(\alpha_A + \beta_A t)}$$

$$p_B(t) = \frac{\exp(\alpha_B + \beta_B t)}{1 + \exp(\alpha_B + \beta_B t)}$$

- Conditional probabilities

$$\tau_{11} = \mathbf{P}(C = 1 \mid A = 1, B = 1)$$

$$\tau_{10} = \mathbf{P}(C = 1 \mid A = 1, B = 0)$$

$$\tau_{01} = \mathbf{P}(C = 1 \mid A = 0, B = 1)$$

$$\tau_{00} = \mathbf{P}(C = 1 \mid A = 0, B = 0)$$

Model

- "Component" probabilities, logistic regression,

$$p_A(t) = \frac{\exp(\alpha_A + \beta_A t)}{1 + \exp(\alpha_A + \beta_A t)}$$
$$p_B(t) = \frac{\exp(\alpha_B + \beta_B t)}{1 + \exp(\alpha_B + \beta_B t)}$$

- Conditional probabilities

$$\tau_{11} = \mathbf{P}(C = 1 \,|\, A = 1, B = 1)$$
$$\tau_{10} = \mathbf{P}(C = 1 \,|\, A = 1, B = 0)$$
$$\tau_{01} = \mathbf{P}(C = 1 \,|\, A = 0, B = 1)$$
$$\tau_{00} = \mathbf{P}(C = 1 \,|\, A = 0, B = 0)$$

Model

- "System" probability

$$p_C(t) = \tau_{11}\frac{\exp(\alpha_A+\alpha_B+(\beta_A+\beta_B)t)}{(1+\exp(\alpha_A+\beta_A t))(1+\exp(\alpha_B+\beta_B t))} + \tau_{10}\frac{\exp(\alpha_A+\beta_A t)}{(1+\exp(\alpha_A+\beta_A t))(1+\exp(\alpha_B+\beta_B t))} + \tau_{01}\frac{\exp(\alpha_B+\beta_B t)}{(1+\exp(\alpha_A+\beta_A t))(1+\exp(\alpha_B+\beta_B t))} + \tau_{00}\frac{1}{(1+\exp(\alpha_A+\beta_A t))(1+\exp(\alpha_B+\beta_B t))}$$

Likelihood and Prior

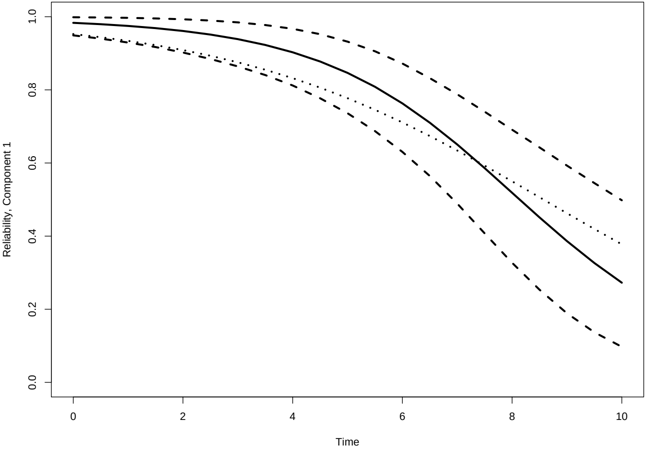
Data at $t = 3$

Age	A	B	C
3	16/19	37/38	12/14

Likelihood at $t = 3$

$$L(\alpha_A, \alpha_B, \beta_A, \beta_B, \tau_{11}, \tau_{10}, \tau_{01}, \tau_{00}) = p_A(3)^{16}(1 - p_A(3))^3 p_B(3)^{37}(1 - p_B(3)) p_C(3)^{12}(1 - p_C(3))^2$$

Posterior Distributions



Extensions

This basic approach has been extended in many directions.

- Multiple diagnostics measured at the components (Anderson-Cook et al. (2008))
- Binary, lifetime, or degradation data at components and system (Guo and Wilson (2013))
- Parallel line of research focused on developing system models: elicitation, software, representations (e.g., Wilson et al. (2007), Anderson-Cook (2008))
- Prior distributions (to capture knowledge about parameters "before" this experiment)
  - "Naive" specifications can lead to surprisingly bad results
  - Variable selection priors

Assurance Testing Example

- Missile system consisting of 10 components.
- The 10 components are labeled A – K, and each component has two or three versions as denoted by the numbers following the identifying letters. For example, component A1 is component A, version 1.
- The missile is a series system.
- Over time, seven variants of the missile have been tested.
- For some systems we do not have information about which variant of the component was tested.

Components Used in Variants of System

System Variants

I	II	III	IV	V	VI	VII
A1	A1	A1	A1	A2	A1	A2
B1	B1	B1	B1	B1	B2	B2
C1	C1	C2	C3	C3	–	C3
D1	D2	D2	–	–	–	–
E1	E2	E2	E1	E2	E2	E2
F1	F2	F3	F3	F3	F3	F2
G1	G2	G3	–	–	–	–
H1	H2	H2	H2	H2	H2	H2
J1	J2	J3	J3	J3	J3	J3
K1	K1	K1	K1	K1	K1	K2

Previous Tests

Component	Tests	Successes	Component	Tests	Successes
A1	638	633	F2	338	336
A2	70	65	F3	196	195
B1	662	651	G1	174	174
B2	46	43	G2	296	294
C1	470	468	G3	144	141
C2	144	141	H1	174	172
C3	90	89	H2	534	529
D1	174	174	J1	174	172
D2	440	435	J2	296	296
E1	194	192	J3	238	237
E2	514	512	K1	664	650
F1	174	174	K2	42	42

Model

We assume that the versions of a component are similar so that we can model the reliabilities hierarchically. In particular, we model the successes of the *j*th version of the *i*th component as:

$$
\begin{aligned}
X_{ij} \mid \pi_{ij} &\sim \text{Binomial}(n_{ij}, \pi_{ij}), \\
\pi_{ij} \mid \delta_i, \gamma_i &\sim \text{Beta}(\delta_i, \gamma_i) \\
\delta_i &\sim \text{Uniform}(0, 5000) \\
\gamma_i &\sim \text{Uniform}(0, 5000)
\end{aligned}
$$

Designing a Test

Consider the situation where a new version of component A is under consideration for the system, while the other components remain at the most recent version. Draws for π for this new system that we can use in our Monte Carlo evaluation of consumer and posterior risk are

$$\pi^{(k)} = \pi_A^{(k)} \, \pi_{B2}^{(k)} \, \pi_{C3}^{(k)} \, \pi_{D2}^{(k)} \, \pi_{E2}^{(k)} \, \pi_{F3}^{(k)} \, \pi_{G3}^{(k)} \, \pi_{H2}^{(k)} \, \pi_{J3}^{(k)} \, \pi_{K2}^{(k)}$$

where we use the predictive distribution for π_1 (component A) and posterior distributions for π_{ij} (specific versions of the rest of the components).

We set values for π_P, π_C, posterior producer's risk, and posterior consumer's risk and then solve for n and c that satisfy the conditions.

Moving Forward

There has been substantial progress in the last 20 years with models, methods, and tools for Bayesian reliability.

- How do we design tests and optimize data collection when we acknowledge that we have many kinds of information and considerable relevant historical information?
- How do we embed reliability growth models in a Bayesian framework?
- How do we think about the idea of mission reliability?