Bayesian inference for reliability of systems and networks using the survival signature

Louis J. M. Aslett[1], Frank Coolen[2] and Simon P. Wilson[3]

[1]University of Oxford, [2]Durham University, [3]Trinity College Dublin

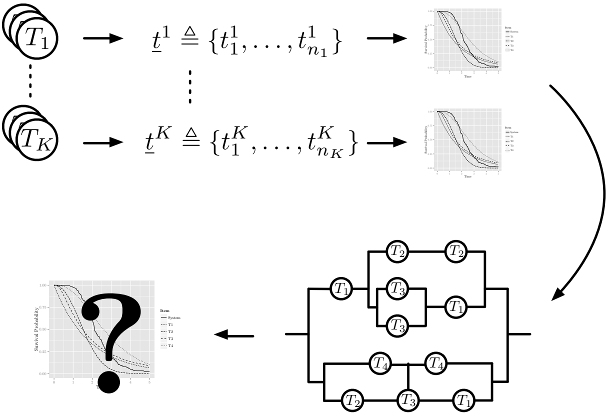
Durham Asset Management Seminar
12th March 2014

i-like.org.uk

Problem setting

Test data available on a components to be used in a system.

Objective: inference on system/network reliability given component test data.



A nonparametric model of components

At a fixed time t, probability component of type k functions is Bernoulli(p_t^k) for some unknown p_t^k.

$\implies$ number functioning at time t in iid batch of n_k is Binomial(n_k, p_t^k).

A nonparametric model of components

At a fixed time t, probability component of type k functions is Bernoulli(p_t^k) for some unknown p_t^k.

$\implies$ number functioning at time t in iid batch of n_k is Binomial(n_k, p_t^k).

Let $S_t^k \in \{0, 1, \dots, n_k\}$ be number of working components in test batch of n_k components of type k. Then,

$$S_t^k \sim \text{Binomial}(n_k, p_t^k) \;\; \forall \, t > 0$$

A nonparametric model of components

At a fixed time t, probability component of type k functions is Bernoulli(p_t^k) for some unknown p_t^k.

$\implies$ number functioning at time t in iid batch of n_k is Binomial(n_k, p_t^k).

Let $S_t^k \in \{0, 1, \dots, n_k\}$ be number of working components in test batch of n_k components of type k. Then,

$$S_t^k \sim \text{Binomial}(n_k, p_t^k) \;\; \forall\, t > 0$$

Given test data $\underline{t}^k = \{t_1^k, \dots, t_{n_k}^k\}$, for each t we can form corresponding observation from Binomial model

$$s_t^k \triangleq \sum_{i=1}^{n_k} \mathbb{I}(t_i^k > t)$$

Bayesian inference for nonparametric model

Taking prior $p_t^k \sim \text{Beta}(\alpha_t^k, \beta_t^k)$, exploit conjugacy result

$$p_t^k \mid s_t^k \sim \text{Beta}(\alpha_t^k + s_t^k, \beta_t^k + n_k - s_t^k)$$

Then, posterior predicitive for number of components surviving in a new batch of m_k components is

$$C_t^k \mid s_t^k \sim \text{Beta-binomial}(m_k, \alpha_t^k + s_t^k, \beta_t^k + n_k - s_t^k)$$

Bayesian inference for nonparametric model

Taking prior $p_t^k \sim \text{Beta}(\alpha_t^k, \beta_t^k)$, exploit conjugacy result

$$p_t^k \mid s_t^k \sim \text{Beta}(\alpha_t^k + s_t^k, \beta_t^k + n_k - s_t^k)$$

Then, posterior predicitive for number of components surviving in a new batch of m_k components is

$$C_t^k \mid s_t^k \sim \text{Beta-binomial}(m_k, \alpha_t^k + s_t^k, \beta_t^k + n_k - s_t^k)$$

Summary: for any fixed t, s_t^k provides a minimal sufficient statistic for computing posterior predictive distribution of the number of components surviving to t in a new batch, without any parametric model for component lifetime being assumed.

Propagating uncertainty to the system

Now take collection of component types $k \in \{1, \ldots, K\}$, each with test data $\underline{t} = \{\underline{t}^1, \ldots, \underline{t}^k\}$, and corresponding collection of minimal sufficient statistics for a fixed t, $\{s_t^1, \ldots s_t^K\}$.

Propagating uncertainty to the system

Now take collection of component types $k \in \{1, \ldots, K\}$, each with test data $\underline{t} = \{\underline{t}^1, \ldots, \underline{t}^k\}$, and corresponding collection of minimal sufficient statistics for a fixed t, $\{s_t^1, \ldots s_t^K\}$.

Survival probability for a new system S^* comprising these component types follows naturally via posterior predictive and surival signature:

$$P(T_{S^*} > t \,|\, s_t^1, \ldots s_t^K)$$

$$= \int \cdots \int P(T_{S^*} > t \,|\, p_t^1, \ldots p_t^K) P(p_t^1 \mid s_t^1) \ldots P(p_t^K \mid s_t^K) \, dp_t^1 \ldots dp_t^K$$

Propagating uncertainty to the system

Now take collection of component types $k \in \{1, \dots, K\}$, each with test data $\underline{t} = \{\underline{t}^1, \dots, \underline{t}^k\}$, and corresponding collection of minimal sufficient statistics for a fixed t, $\{s_t^1, \dots s_t^K\}$.

Survival probability for a new system S^* comprising these component types follows naturally via posterior predictive and surival signature:

$$
\begin{aligned}
&P(T_{S^*} > t \,|\, s_t^1, \dots s_t^K) \\
&= \int \cdots \int P(T_{S^*} > t \,|\, p_t^1, \dots p_t^K) P(p_t^1 \,|\, s_t^1) \dots P(p_t^K \,|\, s_t^K)\, dp_t^1 \dots dp_t^K \\
&= \int \cdots \int \left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K) P\left(\bigcap_{k=1}^{K} \{ C_t^k = l_k \,|\, p_t^k \} \right) \right] \\
&\qquad \times P(p_t^1 \,|\, s_t^1) \dots P(p_t^K \,|\, s_t^K)\, dp_t^1 \dots dp_t^K
\end{aligned}
$$

$$P(\,T_{S^*} > t \,|\, s_t^1, \dots s_t^K)$$

$$= \int \cdots \int P(\,T_{S^*} > t \,|\, p_t^1, \dots p_t^K) P(p_t^1 \mid s_t^1) \dots P(p_t^K \mid s_t^K)\, dp_t^1 \dots dp_t^K$$

$$= \int \cdots \int \left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K) P\left(\bigcap_{k=1}^{K} \{\, C_t^k = l_k \,|\, p_t^k \} \right) \right] \times P(p_t^1 \mid s_t^1) \dots P(p_t^K \mid s_t^K)\, dp_t^1 \dots dp_t^K$$

$$= \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K) \prod_{k=1}^{K} \int P(C_t^k = l_k \,|\, p_t^k) P(p_t^k \mid s_t^k)\, dp_t^k$$

Final integral is simply the posterior predictive (Beta-binomial).

System survival probability

$$P(T_{S^*} > t \mid s_t^1, \dots s_t^K)$$
$$= \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K)$$
$$\times \prod_{k=1}^{K} \binom{m_k}{l_k} \frac{\mathrm{B}(l_k + \alpha_t^k + s_t^k, m_k - l_k + \beta_t^k + n_k - s_t^k)}{\mathrm{B}(\alpha_t^k + s_t^k, \beta_t^k + n_k - s_t^k)}$$

Incredibly easy to implement this algorithmically since survival signature has factorised the survival function by component type!

Why not structure function?

$$\phi(\underline{x}) = \prod_{j=1}^{s} \left(1 - \prod_{i \in C_j} (1 - x_i) \right)$$

where $\{C_1, \ldots, C_s\}$ is the collection of minimal cut sets of the system. Recall don't need $x \in \{0, 1\}$ — we can plug in probabilities. So why not?

Why not structure function?

$$\phi(\underline{x}) = \prod_{j=1}^{s} \left(1 - \prod_{i \in C_j} (1 - x_i) \right)$$

where $\{C_1, \ldots, C_s\}$ is the collection of minimal cut sets of the system. Recall don't need $x \in \{0, 1\}$ — we can plug in probabilities. So why not?

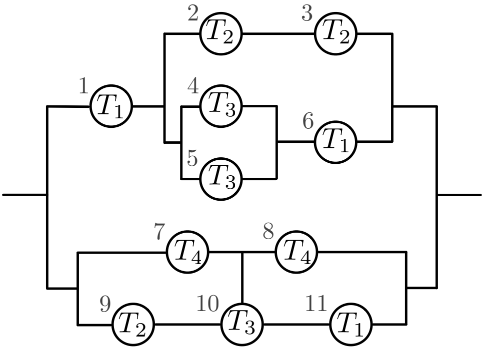
$$P(\, T_{S^*} > t \,|\, s_t^1, \ldots s_t^K) = \int \cdots \int \phi(p_t^{x_1}, \ldots, p_t^{x_n}) P(p_t^1 \mid s_t^1) \ldots P(p_t^K \mid s_t^K) \, dp_t^1 \ldots dp_t^K$$

where $p_t^{x_i}$ is the element of $\{p_t^1, \ldots, p_t^K\}$ corresponding to component i.

Have fun with that integral for large $K \ldots$!

Example system layout, $K = 4$, $n = 11$

Example system:



$T_1 \sim \text{Exp}(\lambda_1 = 0.55)$

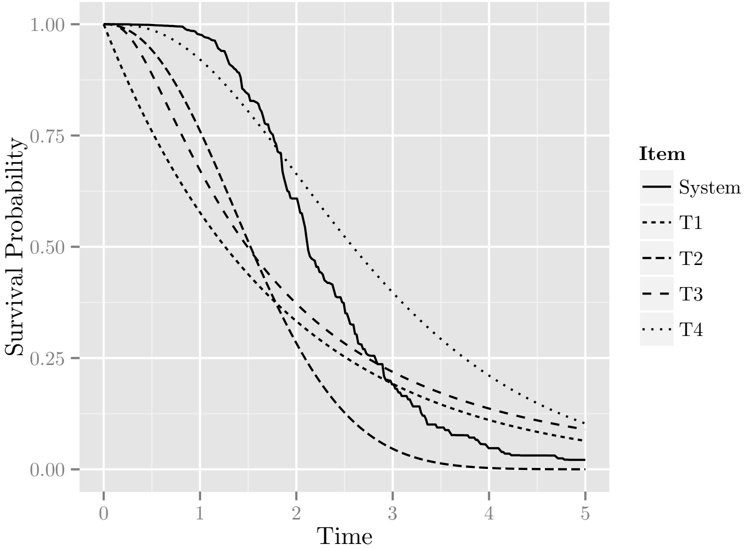
$T_2 \sim \text{Wei}(\lambda_2 = 1.8, \gamma_1 = 2.2)$

$T_3 \sim \text{Log-N}(\mu = 0.4, \tau = 1.234)$

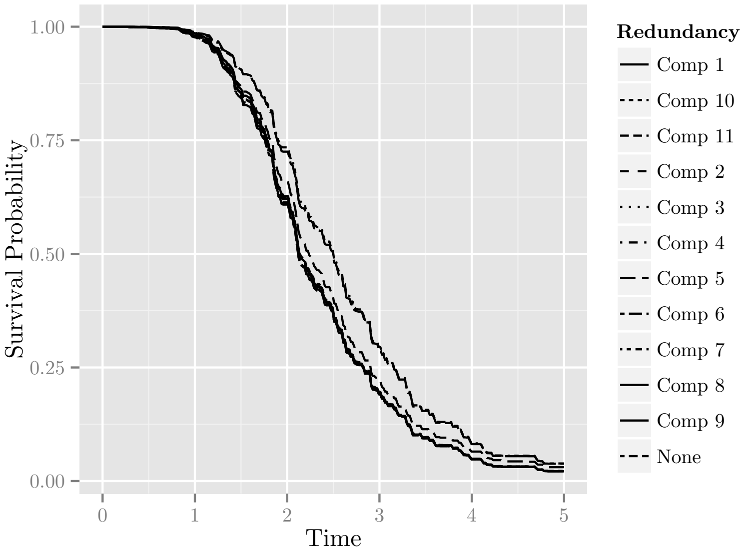
$T_4 \sim \text{Gam}(\lambda_3 = 0.9, \gamma_2 = 3.2)$

Simulated test data with $n_k = 100 \; \forall \, k$

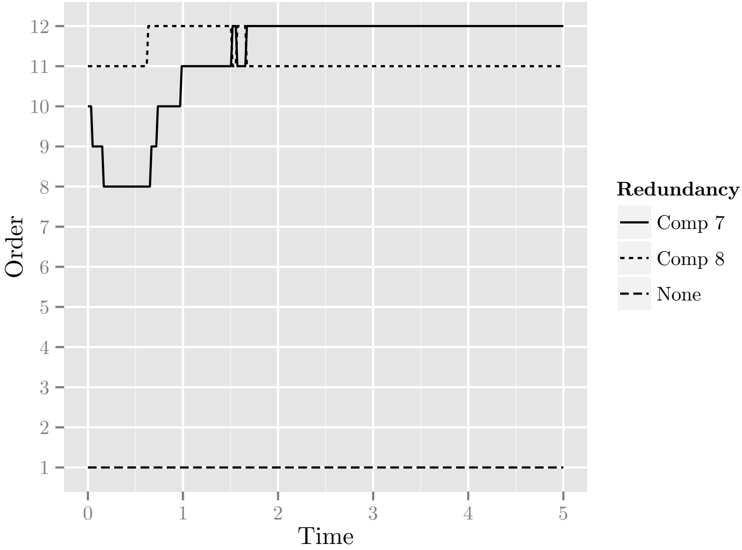
Posterior predictive survival curves



Optimal redundancy?



Optimal redundancy?



Parametric models of components

The survival signature achieves the same factorisation of system lifetime when using parametric models for the components.

Model the lifetime of component k directly via likelihood function f_k

$$T_k \sim f_k(t; \psi_k)$$

As before, given test data $\underline{t}^k = \{t_1^k, \dots, t_{n_k}^k\}$ for component k, posterior density is:

$$f_{\Psi_k \,|\, \underline{T}^k}(\psi_k \,|\, \underline{t}^k) \propto f_{\Psi_k}(\psi_k) \prod_{i=1}^{n_k} f_k(t_i^k; \psi_k)$$

$$P(T_{S^*} > t \,|\, \underline{t}^1, \dots \underline{t}^K)$$
$$= \int \cdots \int P(T_{S^*} > t \,|\, \psi_1, \dots \psi_K) f_{\Psi_1 \,|\, \underline{T}^1}(d\psi_1 \,|\, \underline{t}^1) \dots f_{\Psi_K \,|\, \underline{T}^K}(d\psi_K \,|\, \underline{t}^K)$$

$$
\begin{aligned}
&P(T_{S^*} > t \,|\, \underline{t}^1, \ldots \underline{t}^K) \\
&= \int \cdots \int P(T_{S^*} > t \,|\, \psi_1, \ldots \psi_K) f_{\Psi_1 \,|\, \underline{T}^1}(d\psi_1 \,|\, \underline{t}^1) \ldots f_{\Psi_K \,|\, \underline{T}^K}(d\psi_K \,|\, \underline{t}^K) \\
&= \int \cdots \int \left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \ldots, l_K) P\left(\bigcap_{k=1}^{K} \{ C_t^k = l_k \,|\, \psi_k \} \right) \right] \\
&\qquad \times f_{\Psi_1 \,|\, \underline{T}^1}(d\psi_1 \,|\, \underline{t}^1) \ldots f_{\Psi_K \,|\, \underline{T}^K}(d\psi_K \,|\, \underline{t}^K)
\end{aligned}
$$

$$P(T_{S^*} > t \,|\, \underline{t}^1, \dots \underline{t}^K)$$

$$= \int \cdots \int P(T_{S^*} > t \,|\, \psi_1, \dots \psi_K) f_{\Psi_1 \,|\, \underline{T}^1}(d\psi_1 \,|\, \underline{t}^1) \dots f_{\Psi_K \,|\, \underline{T}^K}(d\psi_K \,|\, \underline{t}^K)$$

$$= \int \cdots \int \left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K) P\left(\bigcap_{k=1}^{K} \{ C_t^k = l_k \,|\, \psi_k \} \right) \right]$$
$$\times f_{\Psi_1 \,|\, \underline{T}^1}(d\psi_1 \,|\, \underline{t}^1) \dots f_{\Psi_K \,|\, \underline{T}^K}(d\psi_K \,|\, \underline{t}^K)$$

$$= \int \cdots \int \left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K) \right.$$
$$\left. \times \prod_{k=1}^{K} \binom{m_k}{l_k} [F_k(t; \psi_k)]^{m_k - l_k} [1 - F_k(t; \psi_k)]^{l_k} \right]$$
$$\times f_{\Psi_1 \,|\, \underline{T}^1}(d\psi_1 \,|\, \underline{t}^1) \dots f_{\Psi_K \,|\, \underline{T}^K}(d\psi_K \,|\, \underline{t}^K)$$

$$
\begin{aligned}
&P(T_{S^*} > t \,|\, \underline{t}^1, \dots \underline{t}^K) \\
&= \int \cdots \int \left[\sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K) \right. \\
&\qquad \left. \times \prod_{k=1}^{K} \binom{m_k}{l_k} [F_k(t; \psi_k)]^{m_k - l_k} [1 - F_k(t; \psi_k)]^{l_k} \right] \\
&\qquad \times f_{\Psi_1 \,|\, \underline{T}^1}(d\psi_1 \,|\, \underline{t}^1) \dots f_{\Psi_K \,|\, \underline{T}^K}(d\psi_K \,|\, \underline{t}^K) \\
&= \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \dots, l_K) \\
&\qquad \times \prod_{k=1}^{K} \binom{m_k}{l_k} \int [F_k(t; \psi_k)]^{m_k - l_k} [1 - F_k(t; \psi_k)]^{l_k} f_{\Psi_k \,|\, \underline{T}^k}(d\psi_k \,|\, \underline{t}^k)
\end{aligned}
$$

Final term posterior predictive of l_k components of type k surviving to t.

Computing the integral for arbitrary models

Three possibilities. The posterior, $f_{\Psi_k \mid \underline{T}^k}(d\psi_k \mid \underline{t}^k)$, is:

1. in closed form and integral tractable;
2. known distribution, but the integral is intractable;
3. not in closed form.

Also, note the integral is just:

$$\mathbb{E}_{\Psi_k \mid \underline{T}^k}\left[[F_k(t;\psi_k)]^{m_k - l_k} [1 - F_k(t;\psi_k)]^{l_k} \right]$$

Computing the integral for arbitrary models

Three possibilities. The posterior, $f_{\Psi_k \,|\, \underline{T}^k}(d\psi_k \,|\, \underline{t}^k)$, is:

1. in closed form and integral tractable;
2. known distribution, but the integral is intractable;
3. not in closed form.

Also, note the integral is just:

$$\mathbb{E}_{\Psi_k \,|\, \underline{T}^k}\left[[F_k(t;\psi_k)]^{m_k - l_k}[1 - F_k(t;\psi_k)]^{l_k} \right]$$

Thus, for samples $\psi_k^{(1)}, \dots, \psi_k^{(N)} \sim \Psi_k \,|\, \underline{T}^k$ we can always fall back to evaluating:

$$\frac{1}{N}\sum_{i=1}^{N}[F_k(t;\psi_k^{(i)})]^{m_k - l_k}[1 - F_k(t;\psi_k^{(i)})]^{l_k}$$

$$\xrightarrow{N\to\infty} \mathbb{E}_{\Psi_k \,|\, \underline{T}^k}\left[[F_k(t;\psi_k)]^{m_k - l_k}[1 - F_k(t;\psi_k)]^{l_k} \right]$$

Posterior predictive survival curves for both methods