Bayesian Statistics

Fabrizio Ruggeri

Istituto di Matematica Applicata e Tecnologie Informatiche
Consiglio Nazionale delle Ricerche
Via Alfonso Corti 12, I-20133, Milano, Italy, European Union
fabrizio@mi.imati.cnr.it
www.mi.imati.cnr.it/fabrizio/

RELIABILITY: BASIC DEFINITIONS

- Failure time T, with pdf $f(t)$ and cdf $F(t)$, $t \geq 0$
- Failures $T_1, \ldots, T_n \Rightarrow W_i = T_i - T_{i-1}$ interfailure (interarrival) times
- Reliability function: $S(t) = \mathbb{P}(T > t) = \int_t^\infty f(x)dx$
- Mean time to failure $MTTF = \int_0^\infty tf(t)dt = \int_0^\infty S(t)dt$
- Hazard function (hazard rate, failure rate):

$$h(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$

  - $h(t)$ not a density but $h(t)\Delta t \approx \mathbb{P}(t \leq T < t + \Delta t | T \geq t)$
  - $h(t) = \frac{f(t)}{S(t)}$
  - $f(t) = h(t)S(t) = h(t)e^{-\int_0^t h(x)dx}$

SOFTWARE RELIABILITY

- Software reliability can be defined as *the probability of failure-free operation of a computer code for a specified mission time in a specified input environment*

- Here we are interested in the evolution of the reliability during the phase of software testing, where the goal is to remove as many bugs as possible before marketing but there is the risk of introducing new bugs and worsening the reliability

- Software testing is an important phase before marketing a product
  - Software released too early might have many bugs (errors), causing losses to the producers, due to repair costs and worsened reputation
  - Software released too late might have very few or no bugs but it could be obsolete or marketed when competitors have already sold their own software

- Optimal release time is a target, possibly achievable by using (Bayesian) statistics and decision analysis in combining available data (and experts' opinions) and utility functions based on costs and revenues

HIDDEN MARKOV MODEL

- Failure times $t_1 < t_2 < \ldots < t_n$ in $(0, y]$
- Y_t latent process describing *reliability status* of software at time t (e.g. growing, decreasing and constant)
- Y_t changing only after a failure $\Rightarrow Y_t = Y_m$ for $t \in (t_{m-1}, t_m]$, $m = 1, \ldots, n+1$, with $t_0 = 0$, $t_{n+1} = y$ and $Y_{t_0} = Y_0$
- $\{Y_n\}_{n \in \mathbb{N}}$ Markov chain with
  - discrete state space E
  - transition matrix $\mathbb{P}$ with rows $\mathbb{P}_i = (P_{i1}, \ldots, P_{ik})$, $i = 1, \ldots, k$
- k might be fixed and known or to be estimated

HIDDEN MARKOV MODEL

- Interarrival times of m-th failure $X_m|Y_m = i \sim \mathcal{E}(\lambda(i)), i = 1, \ldots, k, m = 1, \ldots, n$
- X_m's independent given $Y \Rightarrow f(X_1, \ldots, X_n|Y) = \prod_{m=1}^{n} f(X_m|Y)$
- $\mathbb{P}_i \sim \mathcal{D}ir(\alpha_{i1}, \ldots, \alpha_{ik}), \forall i \in E$, i.e. $\pi(\mathbb{P}_i) \propto \prod_{j=1}^{k} P_{ij}^{\alpha_{ij}-1}$
- Independent $\lambda(i) \sim \mathcal{G}(a(i), b(i)), \forall i \in E$
- For fixed k (number of states), available full conditionals $\Rightarrow$ Gibbs sampling
  - $\lambda(i) \sim \mathcal{G}, \forall i \in E$
  - $\mathbb{P}_i \sim \mathcal{D}, \forall i \in E$,
  - Y_m given by a discrete distribution in closed form, $m = 1, \ldots, n$

HIDDEN MARKOV MODEL: SOME ISSUES

- We dealt earlier with many independent r.v.'s and computed the likelihood as the product of their density

- Here we have interarrival times X_m which are not independent but we are still able to consider the product of their conditional densities since they are independent given the hidden states Y_m

  $\mathbb{P}\left(X_1, \ldots, X_n | Y\right) = \int \prod_{m=1}^{n} \mathbb{P}(X_m | Y_m) \mathbb{P}(Y_1, \ldots, Y_n) dY$

- In the Bayesian framework the hidden states Y_m are treated as if they were parameters and they are drawn at each iteration of the MCMC from an adequate posterior conditional distribution

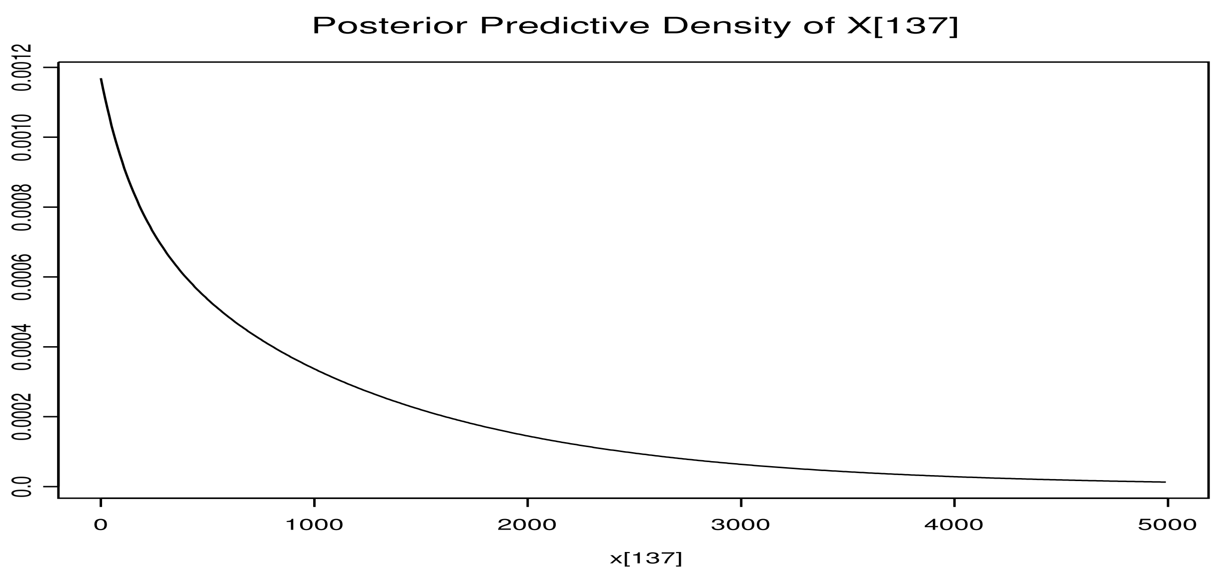
- At the end it is possible to estimate the state of the hidden chain at each time m by considering the frequency of each state in the sample drawn from the MCMC

- Estimation of the number of states K was done either through Reversible jump MCMC or Bayes factor out of MCMC (Chib's method)

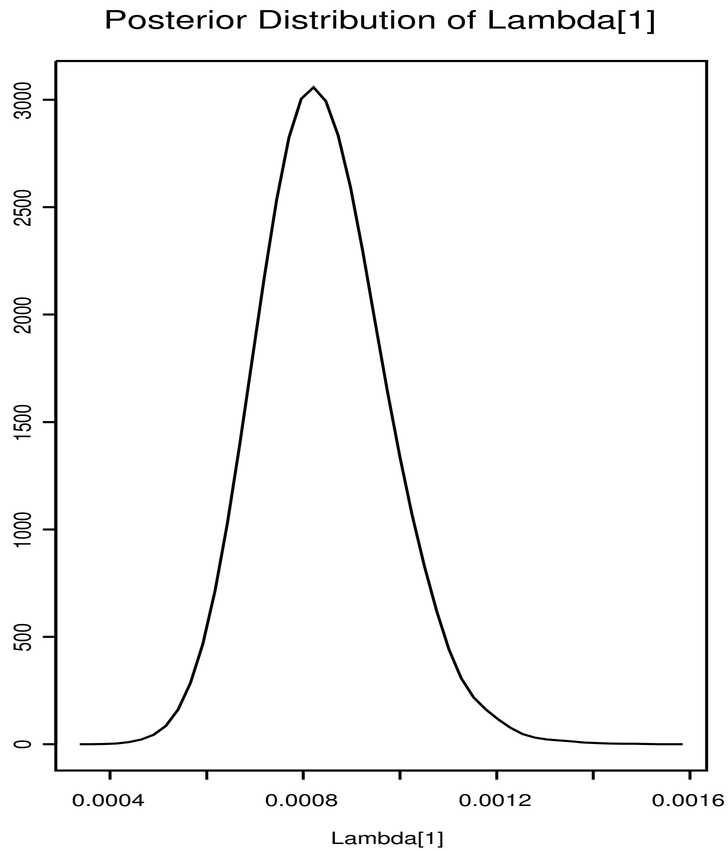
SOFTWARE RELIABILITY - MUSA DATA

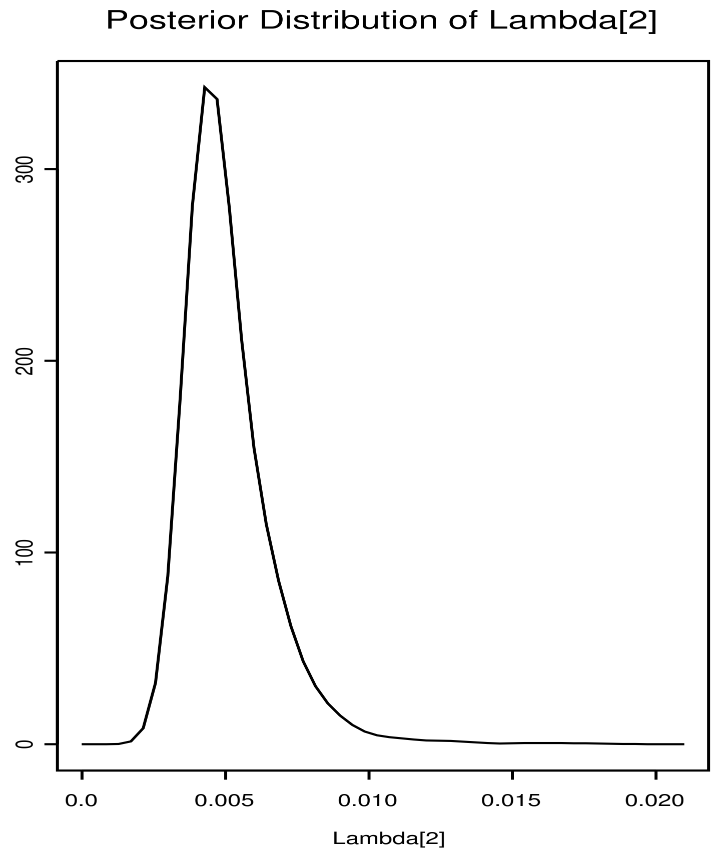
- Musa System 1 data: 136 software failure times

- Hidden Markov model with 2 unknown states

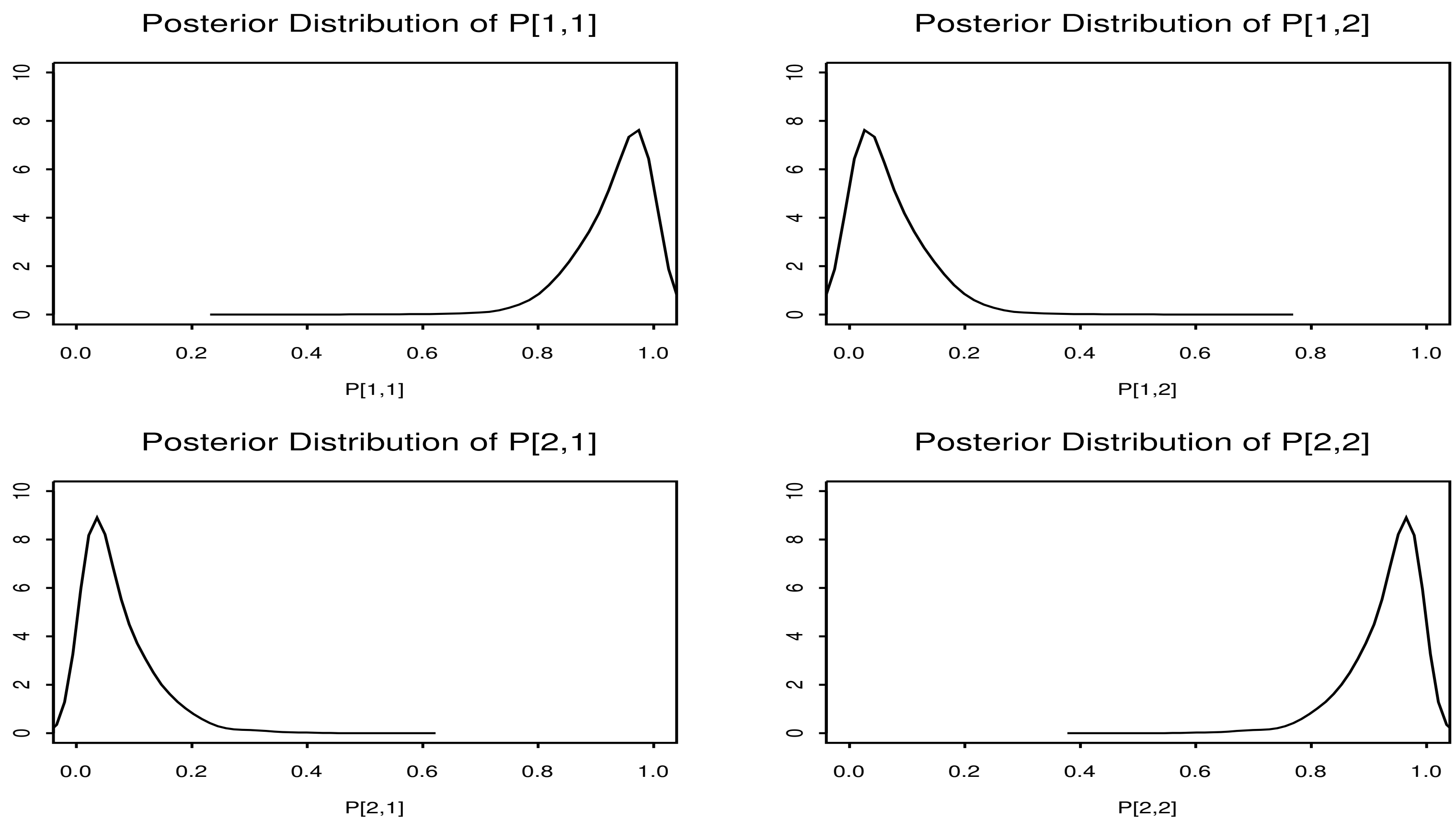


SOFTWARE RELIABILITY - MUSA DATA

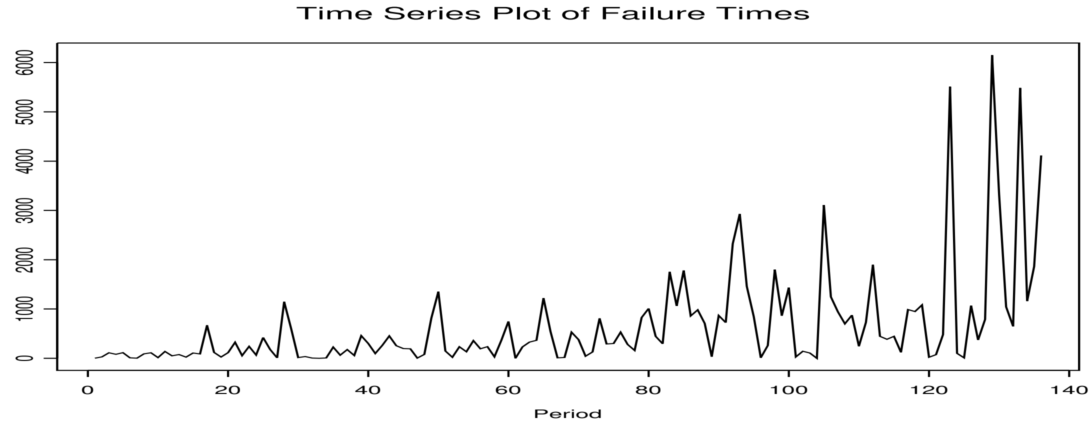


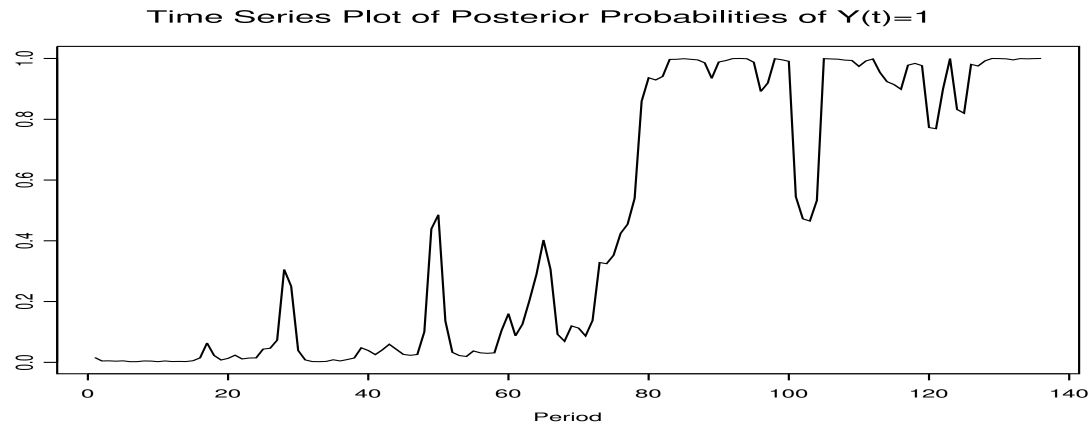


SOFTWARE RELIABILITY - MUSA DATA



SOFTWARE RELIABILITY - MUSA DATA





Longer failure times $\Rightarrow$ higher Bayes estimator of probability of ”good” state

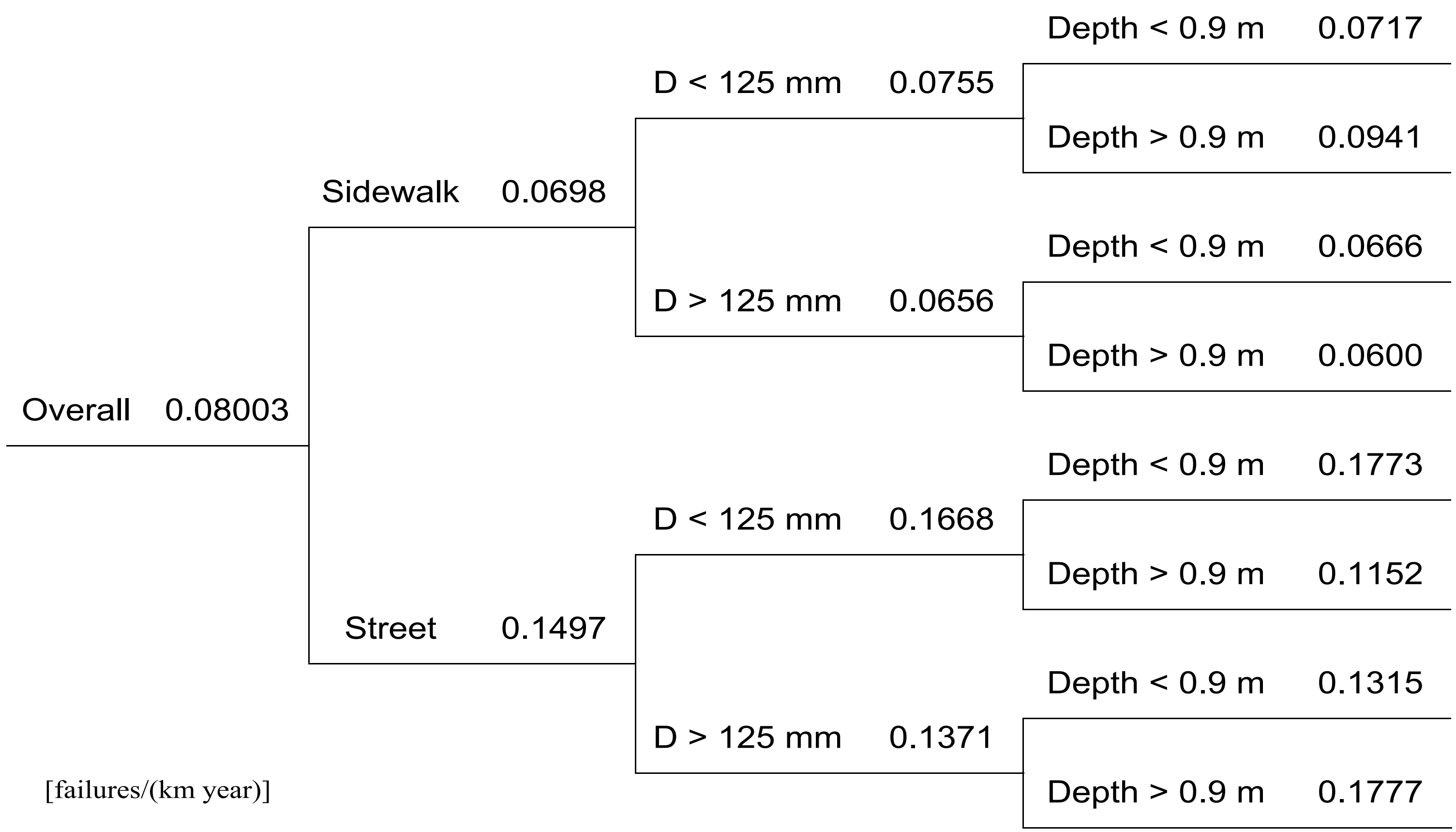
GAS ESCAPES IN A CITY NETWORK

- Setting up an efficient replacement policy in a large metropolitan gas distribution network developed in the last century
- Assessment of failure rate of the pipelines, different for materials (cast iron, steel, polyethylene, etc.) and conditions (diameter, laying depth, etc.)
- Change of pipelines with highest failure rate
- (Traditional) cast iron pipelines with higher failure rate than other materials
  - not subject to corrosion (aging)
  - propensity-to-failure in a unit time period or unit length does not vary significantly with time and space
  - rare events occurring *randomly* and not affecting the next ones
  - $\Rightarrow$ HPP
- EDA identified diameter, laying depth and location as the most significant factors

FAILURES IN CAST-IRON PIPES

- HPP with parameter λ (unit failure rate in time and space)
- n failures in $[0, T] \times \mathcal{S}, \Rightarrow L(\lambda | n, T, \mathcal{S}) = (\lambda s T)^n e^{-\lambda s T}$, with $s = meas(\mathcal{S})$
- Data: $n = 150$ failures in $T = 6$ years on a net $\approx s = 312$ Km long
- $\Rightarrow L(\lambda | n, T, \mathcal{S}) = (1872\lambda)^{150} e^{-1872\lambda}$
- MLE $\widehat{\lambda} = n/(sT) = 150/1872 = 0.080$
- $\lambda \sim \mathcal{G}(\alpha, \beta) \Rightarrow \lambda | n, T, \mathcal{S} \sim \mathcal{G}(\alpha + n, \beta + sT)$
- Consider 8 classes determined by two levels of relevant covariates: diameter, location and depth

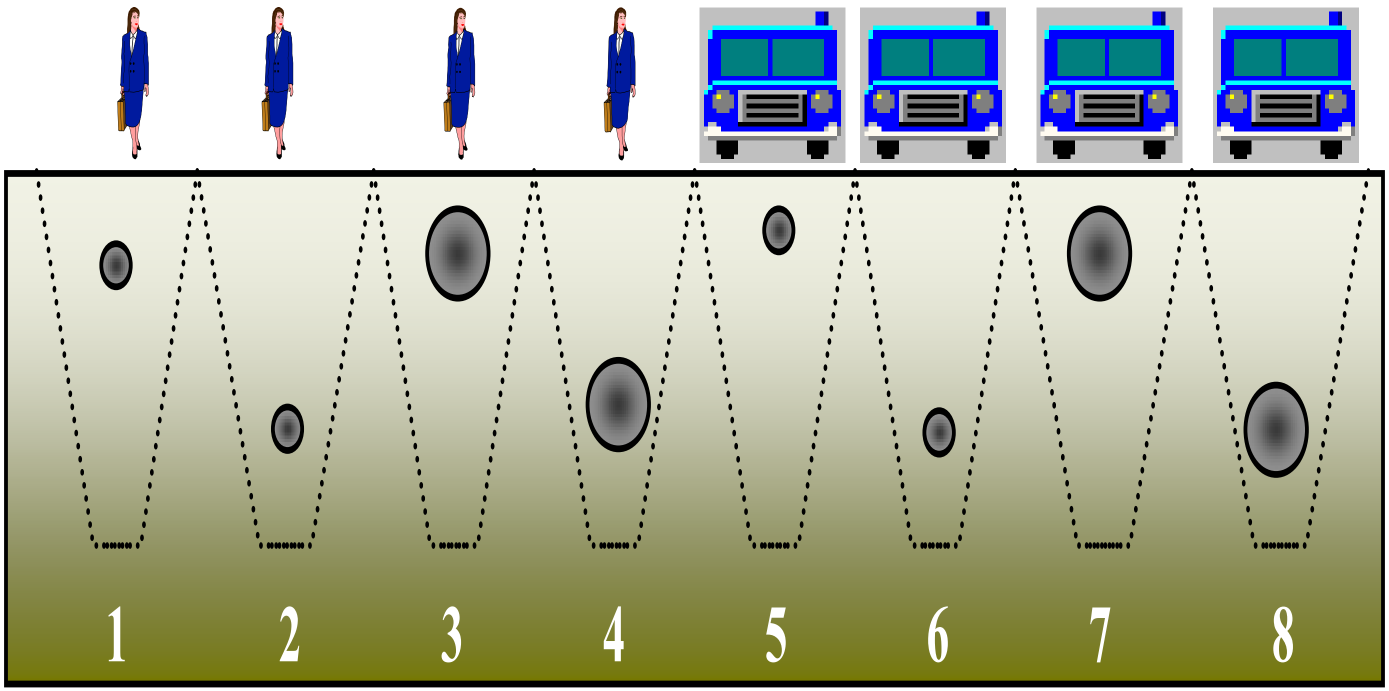
FAILURES IN CAST-IRON PIPE



ELICITATION OF EXPERTS' OPINIONS

- A questionnaire was given to 26 experts from different areas within the company
- Interviewees were unable to say how many failures they expected to see on a kilometer of a given kind of pipe in a year (even upper and lower bounds on them!)
- The experts had great difficulty in saying how and how much a factor influenced the failure and expressing opinions directly on the model parameters while they were able to compare the performance against failure of different pipeline classes
- To obtain such a propensity-to-failure index, each expert was asked to compare the pipeline classes pairwise. In a pairwise comparison the judgement is the expression of the relation between two elements that is given, for greater simplicity, in a linguistic shape
- The linguistic judgement scale is referred to a numerical scale (Saaty's proposal: Analytic Hierarchy Process) and the numerical judgements can be reported in a single matrix of pairwise comparisons

ELICITATION OF EXPERTS' OPINIONS



ANALYTIC HIERARCHY PROCESS

- Two alternatives A and B

B	"equally likely as"	A $\rightarrow$ 1
B	"a little more likely than"	A $\rightarrow$ 3
B	"much more likely than"	A $\rightarrow$ 5
B	"clearly more likely than"	A $\rightarrow$ 7
B	"definitely more likely than"	A $\rightarrow$ 9

- Pairwise comparison for alternatives $A_1, \ldots, A_n$
- $\Rightarrow$ square matrix of size n
- $\Rightarrow$ (normalized) eigenvector associated with the largest eigenvalue
- $\Rightarrow (P(A_1), \ldots, P(A_n))$
- **Question**: if a gas escape occurs, where do think it will occur if you have to choose between subnetwork A and subnetwork B?

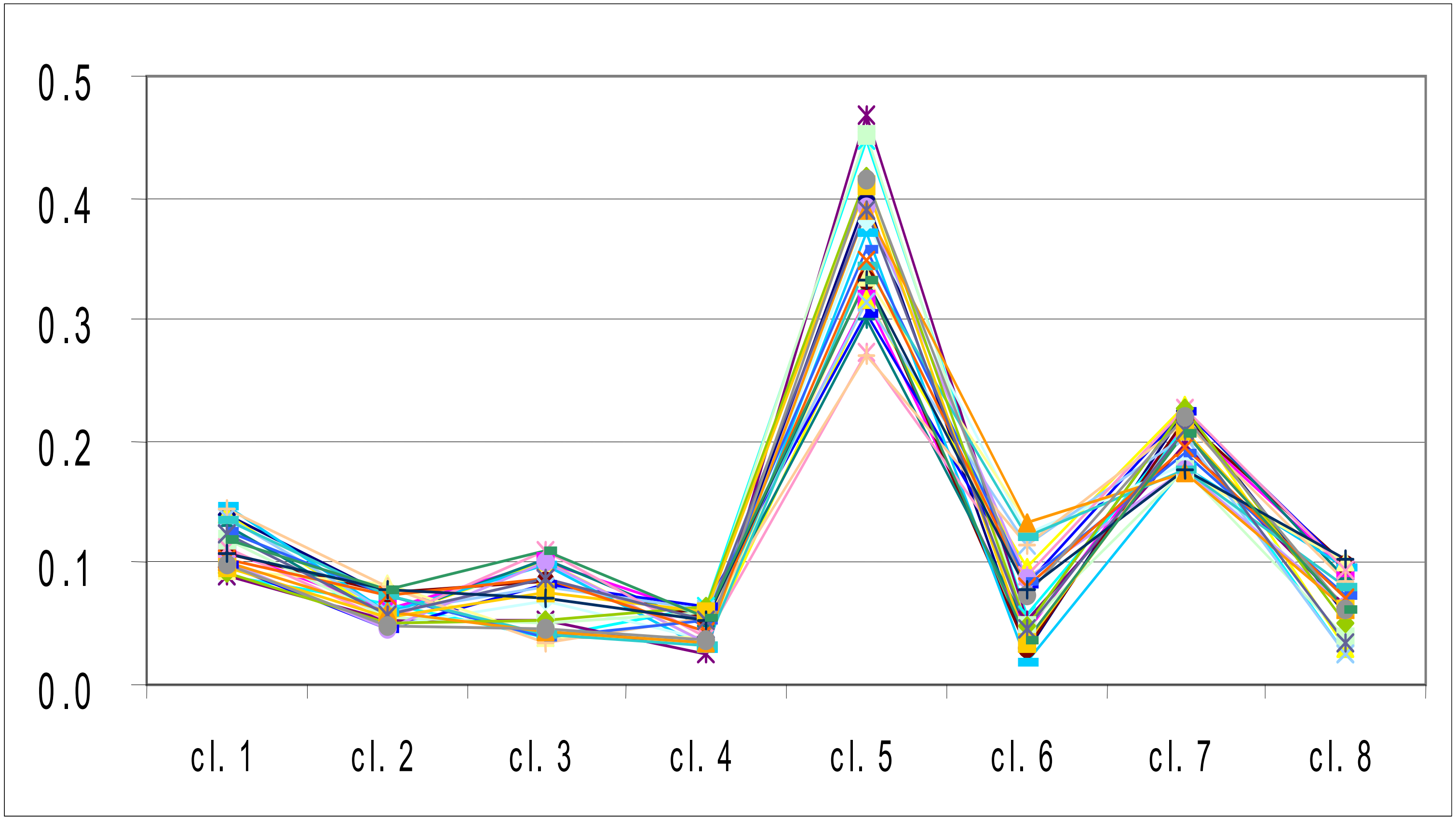
ANALYTIC HIERARCHY PROCESS

An expert's opinion on propensity to failure of cast-iron pipes

Class	1	2	3	4	5	6	7	8
1	1	3	3	3	1/6	1	1/6	3
2	1/3	1	1/4	2	1/6	1/2	1/5	1
3	1/3	4	1	1	1/4	1	1/6	2
4	1/3	1/2	1	1	1/5	1	1/5	1
5	6	6	4	5	1	4	4	5
6	1	2	1	1	1/4	1	1/6	1
7	6	5	6	5	1/4	6	1	4
8	1/3	1	1/2	1	1/5	1	1/4	1

ELICITATION OF EXPERTS' OPINIONS

Values elicited by experts $\Rightarrow$ similar opinions



MODELS FOR CAST-IRON PIPES

Independent classes $A_i, i = 1, 8$, given by 3 covariates (diameter, location and depth) $\Rightarrow$ find the "most risky" class

- Failures in the network occur at rate λ and allocated to class A_i with probability $P(A_i) \Rightarrow$ failures in class A_i occur at rate $\lambda_i = \lambda P(A_i)$ (*Coloring Theorem*)
- $P(A_i)$ given by AHP for any expert
- Choice of $\lambda \Rightarrow$ *critical*
  - Proper way to proceed:
    * Use experts' opinions through AHP to get a Dirichlet prior on $p_i = P(A_i)$
    * Ask the experts about the expected number of gas escapes for given period and length of network $\Rightarrow$ statements on λ, unit failure rate for entire network, and get a gamma prior on it
  - What we did
    * Estimate λ by MLE $\widehat{\lambda}$ with a unique HPP for the network
    * Use experts' opinions through AHP to get a prior on $\lambda_i = \widehat{\lambda} P(A_i)$

MODELS FOR CAST-IRON PIPES

- Choice of priors
  - Gamma vs. Lognormal [informal sensitivity]
  - For each expert, eigenvector from AHP multiplied by $\widehat{\lambda} \Rightarrow$ *sample* about $(\lambda_1, \ldots, \lambda_8)$
  - Mean and variance of priors on λ_i's estimated from the *sample* of size 26 (number of experts) using the method of moments
- Posterior mean of failure rate λ_i for each class
- Classes ranked according to posterior means (largest $\Rightarrow$ most keen to gas escapes)
- Sensitivity
  - Classes of Gamma priors on λ with mean and/or variance in intervals
  - Quantile class on λ

ESTIMATES' COMPARISON

- Location: **W** (under walkway) or **T** (under traffic)
- Diameter: **S** (small, $<$ 125 mm) or **L** (large, $\geq$ 125 mm)
- Depth: **N** (not deep, $<$ 0.9 m) or **D** (deep , $\geq$ 0.9 m)

Class	MLE	Bayes ($\mathcal{LN}$)	Bayes ($\mathcal{G}$)
TSN	*.177*	**.217**	**.231**
TSD	*.115*	*.102*	*.104*
TLN	*.131*	*.158*	*.143*
TLD	**.178**	*.092*	*.094*
WSN	.072	.074	.075
WSD	.094	.082	.081
WLN	.066	.069	.066
WLD	.060	.049	.051

Highest value; 2^{nd}-4^{th} *values*

- Location is the most relevant covariate
- TLD: 3 failures along 2.8 Km but quite unlikely to fail according to the experts
- $\mathcal{LN}$ and $\mathcal{G} \Rightarrow$ similar answers

RENEWAL PROCESS

- Sequence of failure times $X_0 = 0 \leq X_1 \leq X_2 \leq \cdots$
- Interfailure times $T_i = X_i - X_{i-1}$ for $i = 1, 2, \ldots$
- If $T_1, T_2, \ldots$ sequence of i.i.d. random variables
- $\Rightarrow \{T_i\}$ stochastic process called *renewal process*
- The previous HPP $N(t)$, with $N(t) \sim \mathcal{P}(\lambda t)$, is an example of renewal process since the interfailure times $T_1, \ldots$ are i.i.d. exponential random variables
- We check only for T_1, but it is possible to prove for all the other interfailure times:
  $\mathbb{P}(N(t) = 0) = e^{-\lambda t} = \mathbb{P}(T_1 > t) \Rightarrow$ survival/reliability function of $\mathcal{E}(\lambda)$

RENEWAL PROCESS

- Number of cycles run by washing machines before failure (same model, operated under same conditions)

- $\Rightarrow$ Sequence of integer-valued random variables $N_1, \ldots, N_n$

- Assume N_i's, $i = 1, \ldots, n$, are i.i.d. Poisson $\mathcal{P}(\lambda)$ random variables

- $\Rightarrow$ likelihood given by $l(\lambda|data) = \dfrac{\lambda^{\sum_{i=1}^n N_i}}{\prod_{i=1}^n N_i!} e^{-n\lambda}$

- Conjugate gamma $\mathcal{G}(\alpha, \beta)$ prior for λ

- $\Rightarrow$ Posterior $\mathcal{G}\left(\alpha + \sum_{i=1}^n N_i, \beta + n\right)$

- Posterior mean $\dfrac{\alpha + \sum_{i=1}^n N_i}{\beta + n}$

NONHOMOGENEOUS POISSON PROCESS

- NHPP are used in many fields but we will concentrate here on reliability
- NHPP used to model reliability growth/decay
- NHPP good for
  - prototype testing
  - repair of small components in complex systems
- Repair strategies in a NHPP:
  - instantaneous
  - minimal repair ($\Rightarrow$ back to previous reliability)

Repairs could worsen the reliability

NONHOMOGENEOUS POISSON PROCESS

- NHPPs characterized by intensity function $\lambda(t)$ varying over time

- $\Rightarrow$ NHPPs useful to describe (*rare*) events whose rate of occurrence evolves over time (e.g. gas escapes in steel pipelines)

  – Life cycle of a new product

    * initial elevated number of failures (*infant mortality*)

    * almost steady rate of failures (*useful life*)

    * increasing number of failures (*obsolescence*)

    $\Rightarrow$ NHPP with a *bathtub* intensity function

- NHPP has no stationary increments, unlike the HPP

- Superposition and Coloring Theorems can be applied to NHPPs

- Elicitation of priors raises similar issues as before

NONHOMOGENEOUS POISSON PROCESS

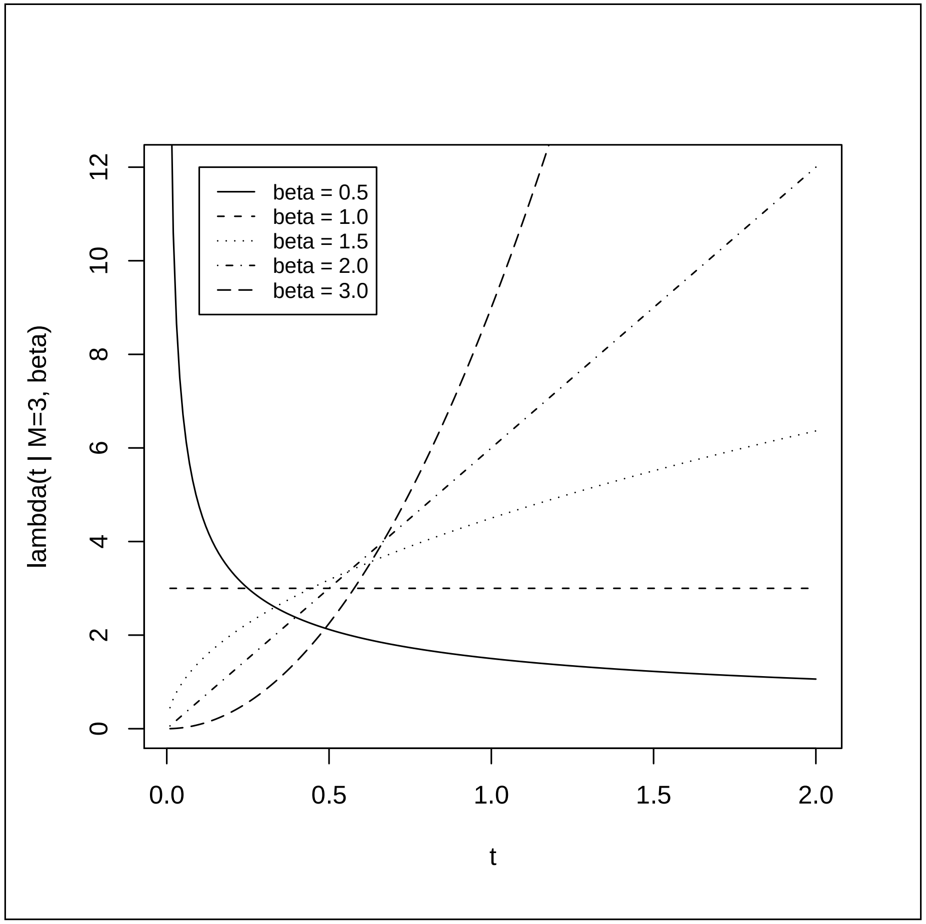
$N(t)$ Power Law process (PLP) (or Weibull process)

- $l(M, \beta \mid T_1, \ldots, T_n) = M^n \beta^n \prod_{i=1}^{n} T_i^{\beta-1} e^{-MT^\beta}$
- Independent priors $M \sim \mathcal{G}(\alpha, \delta)$ and $\beta \sim \mathcal{G}(\mu, \nu)$
- Possible dependent prior: $M|\beta \sim \mathcal{G}\left(\alpha, \delta^\beta\right)$
- $\Rightarrow$ posterior conditionals (in red changes for dependent prior)

$$
\begin{aligned}
M|T_1, \ldots, T_n\beta &\sim \mathcal{G}\left(\alpha + n, \delta^{\beta} + T^\beta\right) \\
\beta|T_1, \ldots, T_n M &\propto \beta^{\mu+n-1} \exp\{\beta(\sum_{i=1}^{n} \log T_i - \nu) - MT^\beta - M\delta^\beta\}
\end{aligned}
$$

- Sample from posterior applying Metropolis step within Gibbs sampler

POWER LAW PROCESS



RELIABILITY MEASURES

- System reliability (for a PLP)
  - Data on the same system (observed up to y):

    $$R((y,s]) = P\left(N(y,s) = 0|M,\beta\right) = e^{-M\left(s^\beta - y^\beta\right)}$$

  - Data on equivalent system:

    $$R(s) = P\left(N(s) = 0|M,\beta\right) = e^{-Ms^\beta}$$

- Expected number of failures in future intervals
  - Same system: $E[N(y,s]|M,\beta] = M\left(s^\beta - y^\beta\right)$
  - Equivalent system: $E[N(s)|M,\beta] = Ms^\beta$

- Intensity function at y:

  Reliability growth models without further improvements $\Rightarrow$ constant intensity $\lambda(y)$

RELIABILITY MEASURES

- Posterior on (M, β) given data $\mathbf{T} \Rightarrow$ sample $\left\{M^{(i)}, \beta^{(i)}\right\}_{i=1}^{N}$
- Reliability measures (now to be integrated w.r.t. posteriors):
  - $R((y,s] = E[R((y,s]|M,\beta)] = \int e^{-M(s^\beta - y^\beta)} f(M,\beta|\mathbf{T}) dM d\beta$
  - $E[N(y,s] = E[E[N(y,s]|M,\beta]] = \int M(s^\beta - y^\beta) f(M,\beta|\mathbf{T}) dM d\beta$
- Estimates:
  - $R\widehat{((y,s}]) = \sum_{i=1}^{N} e^{-M^{(i)}(s^{\beta^{(i)}} - y^{\beta^{(i)}})}$
  - $E[\widehat{N(y,}s]] = \sum_{i=1}^{N} M^{(i)}(s^{\beta^{(i)}} - y^{\beta^{(i)}})$

GAS ESCAPES IN STEEL PIPES

- Replacements not affecting network reliability $\Rightarrow$ repairable system
- Steel pipes subject to corrosion (aging) $\Rightarrow$ NHPP and relevance of installation date
- Network split into subnetworks based upon year of installation, *as if* pipes were installed on July, 1st each year
- Independent PLP's $N_i(t)$ for each subnetwork, with $\lambda_i(t) = M_i \beta_i t^{\beta_i - 1}$
- *Superposition Theorem*: Sum of independent NHPPs $N_i(t)$ with intensity functions $\lambda_i(t)$ is still a NHPP $N(t)$ with intensity function $\lambda(t) = \sum \lambda_i(t)$
- Characteristics of pipes installed vs. PLP parameters
  - Equal pipes $\Rightarrow$ same M and β for each $N_i(t)$
  - Completely different pipes $\Rightarrow$ different, independent M_i and β_i for each $N_i(t)$
  - Similar pipes $\Rightarrow$ M_i and β_i for each $N_i(t)$ coming from a common distribution (exchangeability)

GAS ESCAPES IN STEEL PIPES

- Experts asked about interval of first gas escape X_1
  - Choice of section of the network (e.g. length l)
  - Choice of time intervals in a list (e.g. $[T_0, T_1]$)
  - Degree of belief on each interval (choice among 95%, 85% and 75%); e.g. for PLP (M, β)
    $\Rightarrow P(X_1 \in [T_0, T_1]) = \text{exp}\{-lMT_0^\beta\} - \text{exp}\{-lMT_1^\beta\} = 0.95$
  - Check for consistency, e.g. $A \subset B \not\Rightarrow P(A) > P(B)$

- Pooling of experts' opinions
  - $\Rightarrow$ *sample* from priors
  - $\Rightarrow$ hyperparameters in priors, matching moments

GAS ESCAPES IN STEEL PIPES

- *Known* length l_s of network installed in year $s = 1, \ldots, r$
- *Known* installation date δ_k of k-th failed pipe
- Likelihood $L(\underline{M}, \underline{\beta}; \underline{t}, \underline{\delta}) = \prod_{k=1}^{n} \beta_{\delta_k} l_{\delta_k} M_{\delta_k} (t_k - \delta_k)^{\beta_{\delta_k}-1} e^{-\sum_{s=1}^{r} l_s M_s [(T_1-s)^{\beta_s} - (s \vee T_0 - s)^{\beta_s}]}$
- $M_s \sim \mathcal{E}(\theta_M) \perp \beta_s \sim \mathcal{E}(\theta_\beta)$, $s = 1, \ldots, r$, but exchangeable among themselves
- $\theta_M \sim \mathcal{E}(\tau_M)$ and $\theta_\beta \sim \mathcal{E}(\tau_\beta)$
- Posterior $\pi(\underline{M}, \underline{\beta} | \underline{t}, \underline{\delta})$ obtained integrating out θ_M and θ_β
- $\pi(\underline{M}, \underline{\beta} | \underline{t}, \underline{\delta}) \propto \left(\prod_{s=1}^{r} (l_s M_s \beta_s)^{|I_s|} \right) \left(\prod_{k=1}^{n} (t_k - \delta_k)^{\beta_{\delta_k}-1} \right) e^{-\sum_{s=1}^{r} l_s M_s [(T_1-s)^{\beta_s} - (s \vee T_0 - s)^{\beta_s}]}$.

$$\cdot \tau_M \tau_\beta \frac{r!}{\left[\sum_{s=1}^{r} (M_s + \tau_M / r)\right]^{r+1}} \frac{r!}{\left[\sum_{s=1}^{r} (\beta_s + \tau_\beta / r)\right]^{r+1}}$$

MODELS FOR STEEL PIPES

Exchangeable M and β; known installation dates

95% credible intervals for reliability measures:

- System reliability over 5 years: $P\{N(1998, 2002) = 0\} \Rightarrow [0.0000964, 0.01]$
- Expected number of failures in 5 years: $EN(1998, 2002) \Rightarrow [4.59, 9.25]$
- Mean value function (solid) vs. cumulative # failures (points)