TMA4275 LIFETIME ANALYSIS

Slides 17: Unobserved heterogeneity in NHPP

Bo Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

https://www.ntnu.edu/employees/bo
bo.lindqvist@ntnu.no

NTNU, Spring 2020

CONTENTS OF SLIDES 17

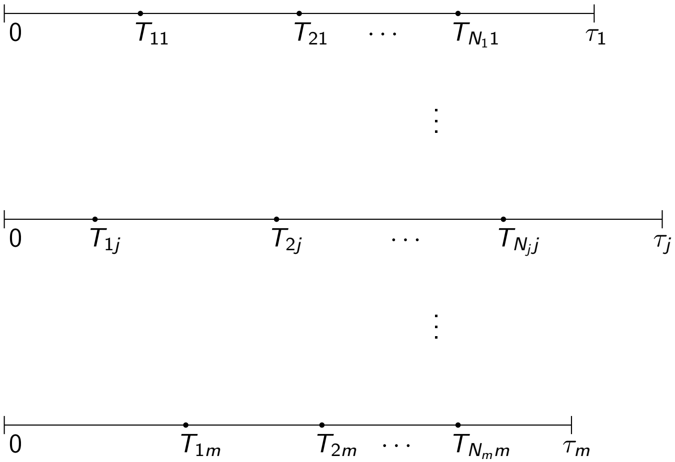
These slides are based on the paper:

Asfaw, Z. G., & Lindqvist, B. H. (2015). Unobserved heterogeneity in the power law nonhomogeneous Poisson process. Reliability Engineering & System Safety, 134, 59-65.

Minimal repair models (NHPP)

- Reliability analyses of repairable systems are commonly made under the assumption of *minimal repairs*, leading to nonhomogeneous Poisson process (NHPP) models.
- Minimal repair intuitively means that a failed system is restored just back to a functioning state, which is commonly named an "as bad as old" condition.
- an NHPP is characterized by the
  - intensity function (ROCOF): $w(t)$
  - or, equivalently, its cumulative version: $W(t) = \int_0^t w(u)du$

Typical data format



Example: Failure times for 104 closing valves at boiling water reactor plants in Finland; follow-up time 3286 hours

System #	Failure times *(External Leakage)*							
1	610	614	943	2024	2087	2104	2399	2525
2	126	323	943	1132	2087	2399	2426	
3	860	915	1606	3181				
4	10	19	104	2352				
5	293	2567						
6	2434	2676						
7	1963							
8	1262							
9	2501							
10	1963							
11	132							
12	1623							
13	3127							
14	3211							
15	1225							
16	1222							
17-104	–							

(From M. Bhattacharjee, E. Arjas and U. Pulkkinen, MMR 2002)

Problem with closing valve failure data? Unexpected large variation in failure behaviour

- The data show an indication of a larger variation in the failure behaviour of the systems than expected if they were identically distributed.
- If overlooked, such variations may lead to wrong predictions for future behaviour or predictions for new systems.

Possible reason for variability: *Heterogeneous systems*

Suppose that basically the systems follow an NHPP with intensity function $w(t)$.

Heterogeneity is introduced in the model by modifying the intensity for a particular system to

$$w_h(t) = a\, w(t),$$

where a is an unobserved positive constant ("frailty") specific to the system under study, and which may vary from system to system.

Thus, if m systems are observed, there are frailties $a_1, a_2, ..., a_m$, which are modeled as independent realizations from a probability distribution with mean 1 and variance $\delta \geq 0$, say.

A standard assumption is to let the a_j be gamma-distributed, which facilitates the derivation of likelihood functions.

Potential sources for heterogeneity

The variability among systems may, for example, be caused by differences in

- environmental conditions
- maintenance strategies
- maintenance philosophy or attitude.
- variation in the quality of the production of the systems themselves ("Monday cars")

or, the omission of certain covariates in the model.

Gamma distribution with expected value 1

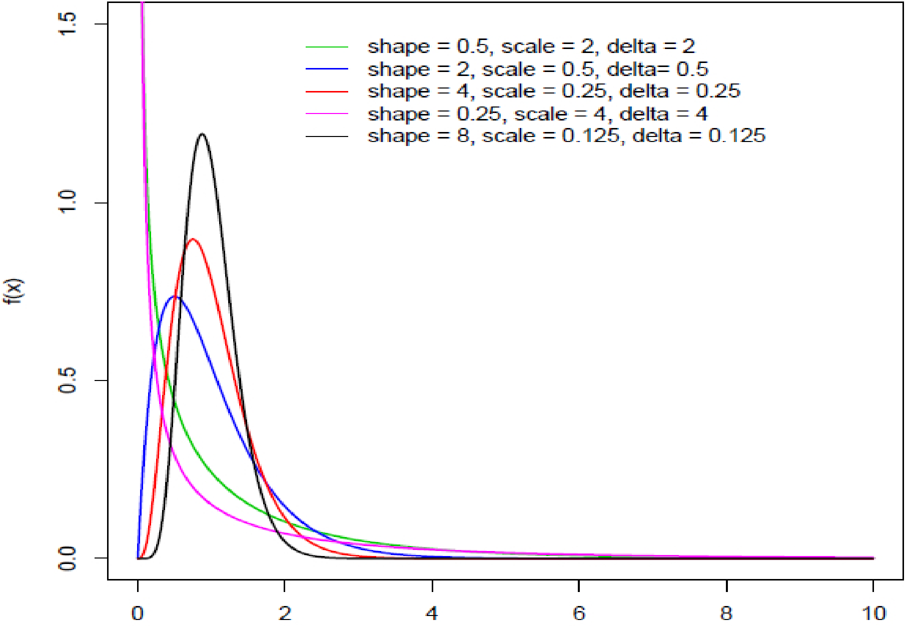
The density of the two-parameter gamma distribution is generally given as

$$f(a) = \frac{a^{k-1} e^{-\frac{a}{\theta}}}{\theta^k \Gamma(k)}$$

for $a > 0$, where $k > 0$ is the shape parameter and $\theta > 0$ is the scale parameter. The corresponding expected value and variance are, respectively, $k\theta$ and $k\theta^2$. Since we require $E(a) = 1$ and $Var(a) = \delta$, we use $k = 1/\delta$ and $\theta = \delta$. The density of a hence becomes

$$h(a) = \frac{a^{\frac{1}{\delta}-1} e^{-\frac{a}{\delta}}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}}$$

Graphs of gamma densities with expected value 1



Likelihood function

Assume that the jth system is observed on the time interval $(0, \tau_j]$. Let n_j be the total number of failures in $(0, \tau_j]$, and let the failure times be $t_{1j}, t_{2j}, \ldots, t_{n_j j}$.

Generally, the likelihood function for an NHPP with intensity $\lambda(t)$ is $\{\prod_{i=1}^{n} \lambda(t_i)\} e^{-\Lambda(\tau)}$, where $\Lambda(t) = \int_0^t \lambda(u)du$.

Thus, for the jth system, where the intensity is $a_j w(t)$, the likelihood for given value of a_j is

$$L_j(a_j) = \left\{ \prod_{i=1}^{n_j} a_j w_j(t_{ij}) \right\} e^{-a_j W(\tau_j)}$$

Likelihood function (cont.)

$$L_j(a_j) = \left\{ \prod_{i=1}^{n_j} a_j w(t_{ij}) \right\} e^{-a_j W(\tau_j)}$$

Since a_j is unobservable, the contribution to the full likelihood from this system is obtained by unconditioning with respect to a_j. This gives the likelihood for the jth system:

$$\begin{aligned} L_j &= E\left[L_j(a_j)\right] = \int L_j(a_j) h(a_j) da_j \\ &= \int \left\{ \prod_{i=1}^{n_j} a_j w(t_{ij}) \right\} e^{-a_j W(\tau_j)} \frac{a_j^{\frac{1}{\delta}-1} e^{\frac{-a_j}{\delta}}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}} da_j \\ &= \frac{\prod_{i=1}^{n_j} w(t_{ij})}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}} \int_0^\infty a_j^{n_j + \frac{1}{\delta} - 1} e^{-a_j (W(\tau_j) + \frac{1}{\delta})} da_j \end{aligned}$$

Likelihood function (cont.)

$$\frac{\prod_{i=1}^{n_j} w(t_{ij})}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}} \int_0^\infty a_j^{n_j+\frac{1}{\delta}-1} e^{-a_j(W(\tau_j)+\frac{1}{\delta})} da_j$$

Now it is easy to show that $\int_0^\infty a^{r-1} e^{-sa} da = \Gamma(r)/s^r$ for all $r, s > 0$, so we get

$$L_j = \frac{\prod_{i=1}^{n_j} w(t_{ij})}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}} \frac{\Gamma(n_j + \frac{1}{\delta})}{\left[W(\tau_j) + \frac{1}{\delta}\right]^{n_j+\frac{1}{\delta}}}.$$

Assuming that the processes have the same baseline $w(t)$, and that the a_j are drawn independently from the given gamma-distribution (parametrized by δ), the full likelihood will be $L = \prod_{j=1}^m L_j$.

Specializing the above to the power law, $w(t) = \lambda\beta t^{\beta-1}$, we get

$$L_j = \frac{\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1}\Gamma(n_j + \frac{1}{\delta})}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}\left[\lambda\tau_j^\beta + \frac{1}{\delta}\right]^{n_j+\frac{1}{\delta}}}$$

Maximum likelihood estimators

Asfaw and B.L. (2015) specialize to the case where $\tau_j = \tau$ for $j = 1, \ldots, m$ and obtain

$$\hat{\lambda} = \frac{n}{m\tau^{\hat{\beta}}}$$

$$\hat{\beta} = \frac{n}{n \ln \tau - \sum_{j=1}^{m} \sum_{i=1}^{n_j} \ln t_{ij}},$$

where $n = \sum_{j=1}^{n} n_j$.

These are in fact exactly the same functions of the data as for the power law case without heterogeneity.

(**Note:** *This would not be the case if the observation time intervals were not all equal for all the m processes.*)

The above assumption (equal τ_j), simplifies the calculation of the maximum likelihood estimator of δ, since we can put the $\hat{\lambda}$ and $\hat{\beta}$ into the log likelihood which will then contain only one parameter to solve for.

Estimation with data from Bhattacharjee et al. 2002

We fitted the following model,

$w_h(t) = a\,\lambda\beta t^{\beta-1}$ with $a \sim \text{gamma}(1/\delta, 1/\delta)$

The maximum likelihood estimates were found to be:

$\hat{\lambda}$	$\hat{\beta}$	$\hat{\delta}$	$\log L$
$4.59 \cdot 10^{-4}$	0.822	8.34	-340.79

As a partial check of model fit, we computed the estimated expected number of systems which will have no failures, i.e., $104 \cdot P(T_1 > 3286)$.

$$\begin{aligned} P(T_1 > 3286) &= P(\text{no failures in } (0{,}3286)) \\ &= E\left[P(\text{no failures in } (0{,}3286)|a)\right] \\ &= E(e^{-\lambda \cdot 3286^{\beta} \cdot a}) \end{aligned}$$

Substituting $\hat{\lambda}$ and $\hat{\beta}$ for λ and β, and letting $a \sim \text{gamma}(1/\hat{\delta}, 1/\hat{\delta})$, we get 0.848 which by multiplication with 104 gives 88.2 (while observed is 88).

Estimation with Proschan's airconditioner data (1963)

We fitted the same model as for the Bhattacharjee et al. data,

$w_h(t) = a\,\lambda\beta t^{\beta-1}$ with $a \sim \text{gamma}(1/\delta, 1/\delta)$

The maximum likelihood estimates are given in the following table, where the first row is for the model with heterogeneity, while the second line is for the model with no heterogeneity, i.e., $\delta = 0$.

$\hat{\lambda}$	$\hat{\beta}$	$\hat{\delta}$	$\log L$
$3.42 \cdot 10^{-3}$	1.163	0.114	-1172.23
$3.57 \cdot 10^{-3}$	1.152	0	-1176.23

By calculating 2 times the difference in log likelihood, it is seen that the p-value is

$$P(\chi^2_1 > 8.0) = 0.0047$$

There is, though, estimated a certain trend ($\hat{\beta} > 1$) which turns out to be significant, thus partly violating the final conclusion by Proschan.

10 000 simulations of $m = 20$ systems for each parameter combination.

Data	m	True Value			n		Estimates					
10000	20	λ	β	δ	Ave.	St.D	$\hat{\lambda}$		$\hat{\beta}$		$\hat{\delta}$	
							Ave.	St.D	Ave.	St.D	Ave.	St.D
		2	1.5	0	63.1451	8.0541	2.0059	0.2057	1.5008	0.0430	0.0021	0.0036
				0.1	63.0648	21.4421	2.0075	0.2456	1.5005	0.0421	0.0942	0.0352
				0.2	63.2006	29.6606	2.0042	0.2853	1.5013	0.0425	0.1897	0.0651
				0.4	63.0469	40.1756	2.0033	0.3450	1.5009	0.0426	0.3795	0.1208
				0.6	63.3880	49.3983	2.0063	0.4033	1.5014	0.0431	0.5718	0.1750
				0.8	63.0149	56.1749	2.0050	0.4452	1.5013	0.0430	0.7511	0.2223
				1	63.4131	64.1655	2.0033	0.5023	1.5021	0.0432	0.9309	0.2610
			1	0	20.0288	4.4658	2.0051	0.2526	1.0019	0.0509	0.0064	0.0119
				0.2	19.9885	9.9238	1.9987	0.3231	1.0023	0.0509	0.1879	0.0774
				0.4	20.2335	13.4745	2.0052	0.3802	1.0026	0.0508	0.3742	0.1305
				0.6	19.8679	15.8587	1.9968	0.4285	1.0043	0.0514	0.5520	0.1769
				0.8	20.1814	18.2127	1.9923	0.4691	1.0051	0.0518	0.7158	0.2107
				1	20.4585	20.8251	1.9843	0.5149	1.0083	0.0524	0.8679	0.2421

Scenario 1: An ordinary power law is anticipated, while there may be an unrecognized heterogeneity

- For prediction of number of failures of a new system, by an erroneous assumption of *no* heterogeneity, one gets too short predicted intervals for the number of failures in a given time period. In fact, the expected value is correctly estimated, but the variation could be much bigger than expected if heterogeneity is not accounted for.
- Neither the expected value nor the standard error of $\hat{\beta}$ are much influenced by the heterogeneity. Thus, the inference for β is not much influenced by a wrong assumption of no heterogeneity.
- For the estimation of λ, one would get a too optimistic estimate for the standard error (and hence too short confidence intervals) by assuming no heterogeneity, and also a downward bias will be present in the estimate.

Scenario 2: The correct model, a power law model with heterogeneity, is used for statistical inference

- The estimator $\hat{\delta}$ seems to behave quite satisfactorily. Still, $\hat{\delta}$ slightly underestimates the true value of δ. Its standard error increases with δ as should be expected.
- As already noted in Scenario 1, the estimator $\hat{\beta}$ is approximately unbiased, with a standard error which does not depend on δ.
- The estimator $\hat{\lambda}$ appear to be approximately unbiased, but has a standard error which increases with δ.