

ICRSA, USC, Columbia

May 21-24, 2003

Heterogeneity Modelling for Recurrent Events

Bo H. Lindqvist

*Department of Mathematical Sciences,
Norwegian University of Science and Technology,*

Trondheim

bo@math.ntnu.no

<http://www.math.ntnu.no/~bo/>

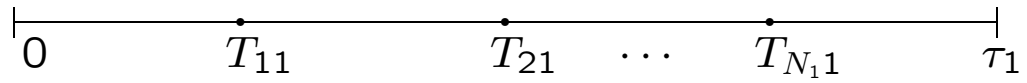
RECURRENT EVENTS DATA

- Observe times for one or more events for each individual.
- Modelling as a point process for each individual
- Applications: Medicine, epidemiology, reliability analysis, finance, etc.

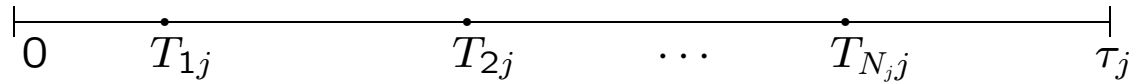
IMPORTANT ASPECTS FOR MODELLING

- Trend in times between events?
- Renewal or Poisson behavior at events?
- Dependence of covariates?
- *Unobserved heterogeneity between individual processes?*

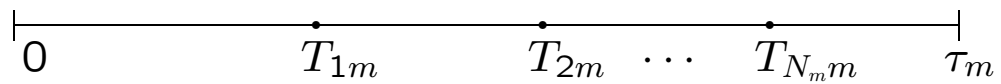
TYPICAL DATA



\vdots



\vdots



(possibly with a vector of covariates for each process)

APPLICATIONS IN RELIABILITY:

- Repairable systems – events correspond to failures and repairs
- Manufactured products – events are warranty claims
- Examples:
 - Proschan (Technometrics 1963) – failures of airconditioners in Boeing airplanes
 - Nelson and Doganaksoy (1989) – valve seat data

APPLICATIONS IN MEDICAL STUDIES:

- Repeated events, e.g. recurrence of infections, epileptic seizures, cancer tumors
- Examples:
 - Aalen and Husebye (Statistics in Medicine 1991) – gastroenterological study
 - Byar and Blackard (Urology 1977), Wei, Lin and Weissfeld (JASA 1989) – recurrence of bladder cancer

WHAT IS HETEROGENEITY?

- Individual variation between systems not explained by observed covariates
- Biostatistics: Heterogeneity = *Frailty*

STANDARD MODELING OF HETEROGENEITY/FRAILTY

- Intensity of events is proportional to unobserved random effect ("frailty"), specific to each individual or system
- Frailties are independent *unobserved* realisations from a distribution (often Gamma)
- Frailties are integrated out in likelihood function (since unobserved)

TREATMENT OF HETEROGENEITY IN RELIABILITY

1955 – Royal Statistical Society, London –

David R. Cox: *“Some Statistical Methods Connected with Series of Events”*

Motivated from the clothing industry.

Introduces concepts and methods for the analysis of repairable systems, among them *heterogeneity*, which he calls *variance components*.

1963 – Technometrics Journal –

Frank Proschan: *“Theoretical Explanation of Observed Decreasing Failure Rate”*

Aircondition data: Failures of aircondition system on 13 Boeing 720 airplanes.

Conclusions:

- HPP appropriate for each system
- *heterogeneity* across systems

1987 – IEEE Transactions on Reliability –

Engelhardt and Bain: *Statistical Analysis of a Compound Power Law Model for Repairable Systems*

TREATMENT OF FRAILITY IN BIOSTATISTICS

1979 – Demography –

Vaupel, Manton & Stallard: *"The impact of heterogeneity in individual frailty on the dynamics of mortality"*.

1984 – Biometrika –

Philip Hougaard: *"Life table methods for heterogeneous populations: distributions describing the heterogeneity"*.

1991 – Statistics in Medicine –

Aalen & Huseby: *"Statistical analysis of repeated events forming renewal processes"*.

STANDARD MODELLING OF RECURRENT EVENTS DATA

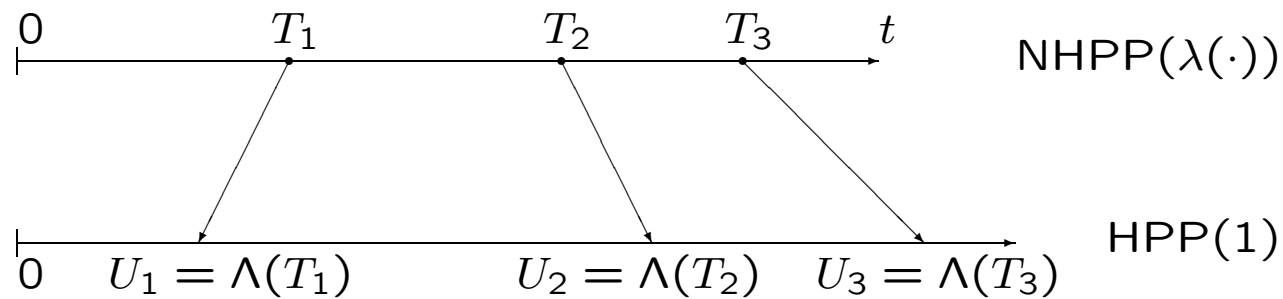
- Homogeneous Poisson-process (HPP)
- Nonhomogeneous Poisson-process (NHPP)
- Renewal process (RP)
- Superimposed renewal process
- Imperfect repair process
- Effective age process
- *Trend-Renewal Process*

TOWARDS TREND-RENEWAL PROCESS

(Lindqvist, Elvebakk, Heggland 2003)

Characterizing property of NHPP($\lambda(\cdot)$):

Cumulative intensity $\Lambda(t) = \int_0^t \lambda(u) du$



INHOMOGENEOUS GAMMA PROCESS

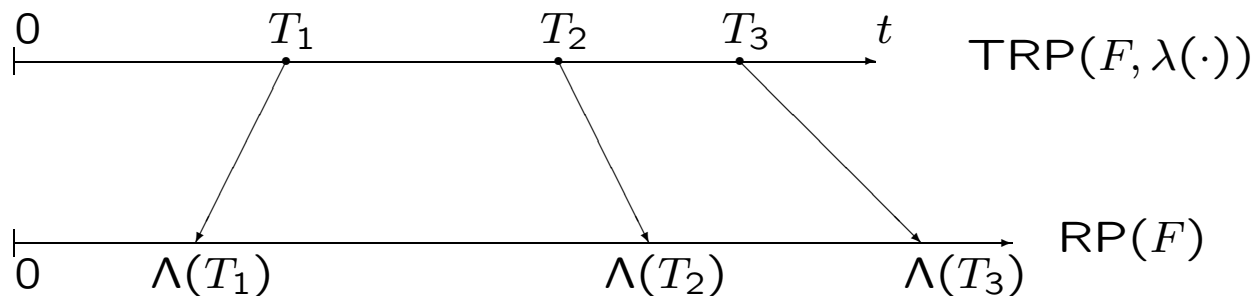
Berman (1981):

Consider only every κ th event of an NHPP($\lambda(\cdot)$) (κ positive integer)

TREND RENEWAL PROCESS –

Defining property of $\text{TRP}(F, \lambda(\cdot))$:

- Trend function: $\lambda(t)$
(cumulative $\Lambda(t) = \int_0^t \lambda(u) du$)
- Renewal distribution: F with expected value 1

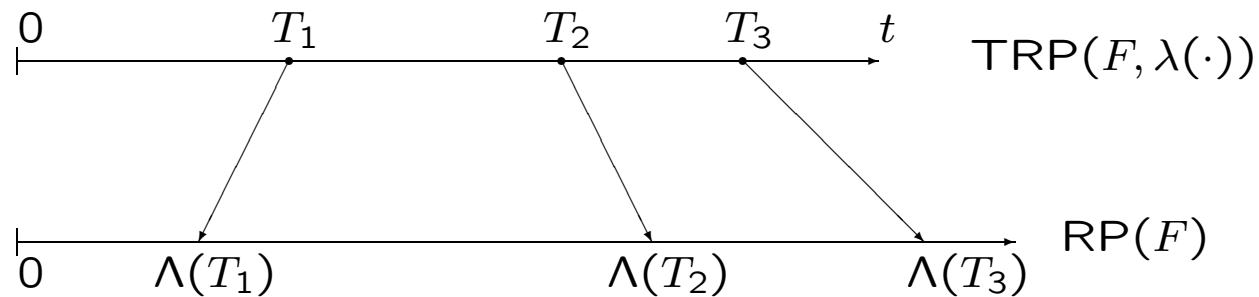


SPECIAL CASES:

- NHPP: If F is standard exponential distribution
- RP: If $\lambda(t)$ is constant in t

STATISTICAL INFERENCE FOR TRP

SINGLE SYSTEM OBSERVED ON $[0, \tau]$



Events occur at $T_1, T_2, \dots, T_{N(\tau)}$

LIKELIHOOD FOR $\text{TRP}(F, \lambda(\cdot))$:

$$L = \left\{ \prod_{i=1}^{N(\tau)} f(\Lambda(T_i) - \Lambda(T_{i-1})) \lambda(T_i) \right\} \\ \times \{1 - F(\Lambda(\tau) - \Lambda(T_{N(\tau)}))\}$$

where $f = F'$

SEVERAL INDEPENDENT SYSTEMS

$$L = \prod L_j$$

INCLUDING COVARIATES IN THE TRP MODEL

- m systems are observed
- Covariate vector x_j observed for j th process
- Process j is TRP($F, \lambda_j(t)$) where

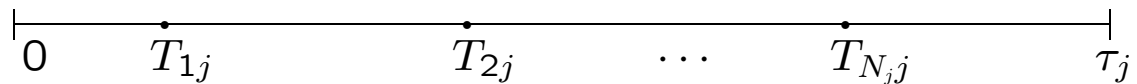
$$\lambda_j(t) = \lambda(t)g(x_j; \beta)$$

(for example of Cox-type).

UNOBSERVED HETEROGENEITY – THE HETEROGENEOUS TRP (HTRP)

DEFINITION OF HTRP($F, \lambda(\cdot), H$)

- m systems are observed
- j th system observed in $[0, \tau_j]$, with N_j observed failures



- Process j is TRP($F, \lambda_j(t)$) where

$$\lambda_j(t) = a_j \lambda(t)$$

- The a_j are i.i.d. unobservable random variables with d.f. H , expected value 1.

LIKELIHOOD:

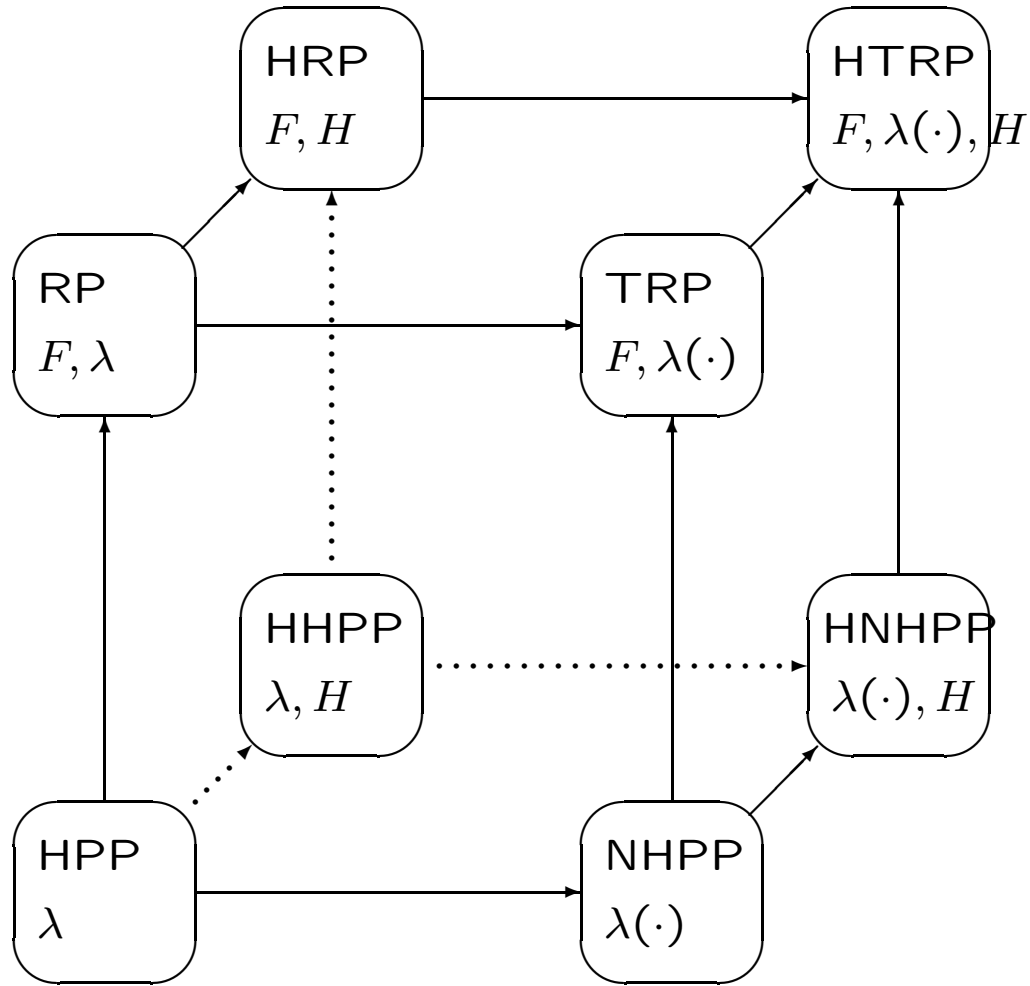
For given value of a_j the likelihood for the j th system is

$$L_j(a_j) = \text{likelihood for TRP}(F, a_j \lambda(\cdot))$$

Full likelihood of HTRP($F, \lambda(\cdot), H$):

$$L = \prod_{j=1}^m \int L_j(a_j) dH(a_j)$$

SUBMODELS OF HTRP – THE MODEL CUBE



SPECIAL PARAMETRIC CASE:

THE INHOMOGENEOUS GAMMA PROCESS WITH UNOBSERVED HETEROGENEITY – HTRP($F_g, \lambda(\cdot), H_g$)

- F_g is gamma-distribution with expectation 1 and variance γ , $\lambda(\cdot)$ is a given parametric trend function, H_g is gamma-distribution with expectation 1 and variance δ
- m processes, j th process censored at the n_j th event ($j = 1, \dots, m$).

Likelihood for j th process, given a_j , is

$$\left\{ \prod_{i=1}^{n_j} K_{ij}^{1/\gamma-1} \lambda(T_{ij}) \right\} \frac{a_j^{n_j/\gamma} \gamma^{-n_j/\gamma}}{(\Gamma(1/\gamma))^{n_j}} \exp\{-a_j \Lambda(T_{n_jj})/\gamma\}$$

where $K_{ij} = \Lambda(T_{ij}) - \Lambda(T_{i-1,j})$ and $T_{0j} = 0$.

Unconditional likelihood L_j found by taking expected value w.r.t. a_j ,

$$L_j = \left\{ \prod_{i=1}^{n_j} K_{ij}^{1/\gamma-1} \lambda(T_{ij}) \right\} \frac{\gamma^{-n_j/\gamma}}{(\Gamma(1/\gamma))^{n_j}} \frac{\delta^{-1/\delta}}{\Gamma(1/\delta)} \frac{\Gamma(n_j/\gamma + 1/\delta)}{(1/\delta + (1/\gamma)\Lambda(T_{n_jj}))^{n_j/\gamma+1/\delta}}$$

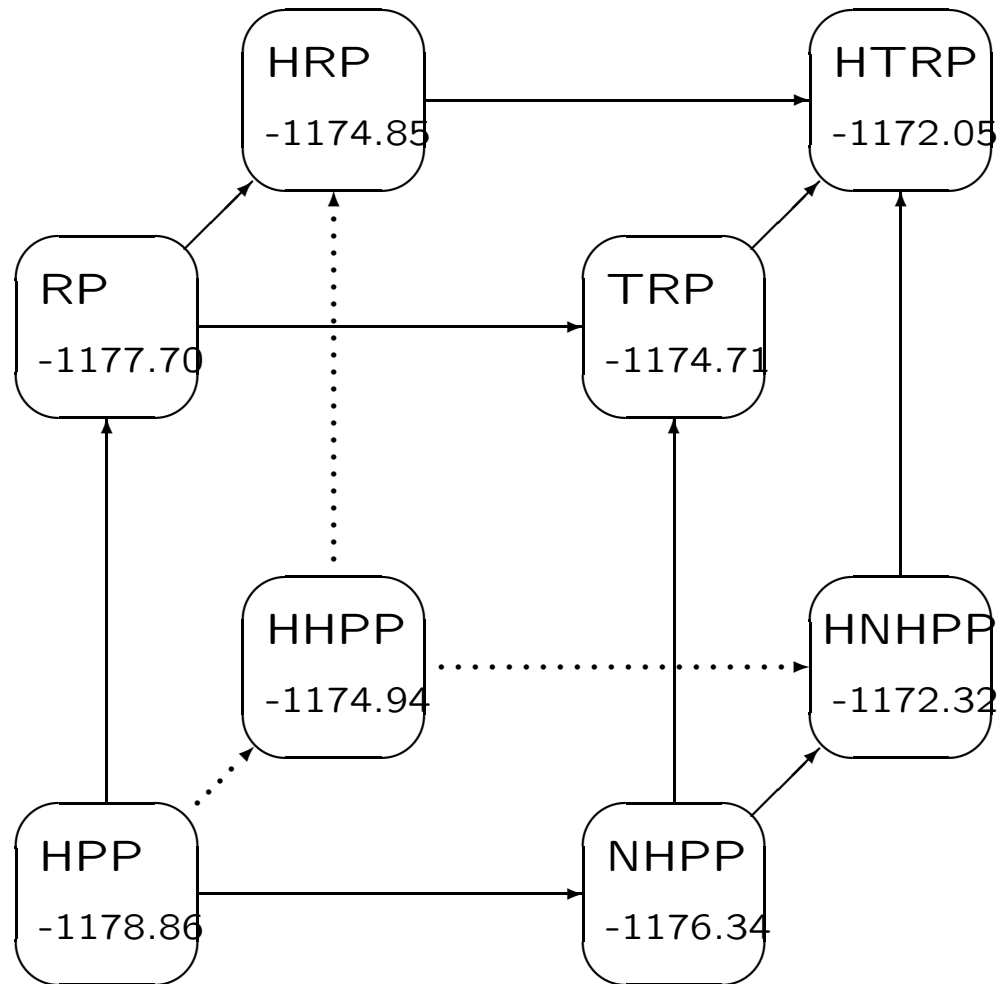
Full likelihood: $\prod L_j$.

PROSCHAN'S AIRCONDITION DATA

13 planes, 17 systems due to Major Overhaul

Plane	Interfailure times						
7907	194	15	41	29	33	181	
7908	413	14	58	37	100	65	9
	169	447	184	36	201	118	
7908 MO	34	31	18	18	67	57	62
	7	22	34				
7909	90	10	60	186	61	49	14
	24	56	20	79	84	44	59
	29	118	25	156	310	76	26
	44	23	62				
7909 MO	130	208	70	101	208		
⋮			⋮				

LIKELIHOOD CUBE FOR PROSCHAN DATA



PROSCHAN'S CONCLUSION: HHPP

OUR CONCLUSION: HNHPP (slightly better ?)

Left face: Follmann and Goldberg (1988), Bottom face: Lawless (1987)

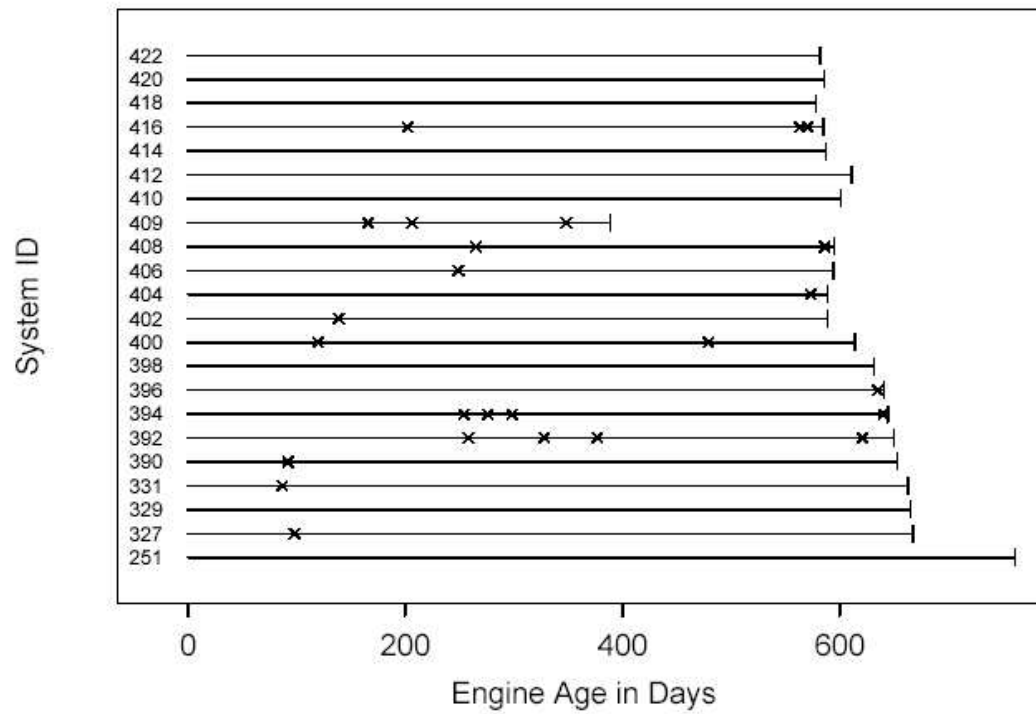
BAYESIAN INFERENCE IN HTRP MODEL

- Introduction of heterogeneity can be viewed as a problem of EMPIRICAL BAYES
- For prediction of particular system we need to take into account the unobservable factor a_j
- Parameter δ can be estimated from ensemble of systems

NEXT, FULL BAYES ANALYSIS MEANS PUTTING A PRIOR ON δ (AND OTHER PARAMETERS)

EXAMPLE – BAYESIAN ANALYSIS OF VALVE SEAT DATA (41 UNITS)

Valve Seat Replacement Times Event Plot (Nelson and Doganaksoy 1989)



VALVE SEAT DATA

Model: $\text{HTRP}(F, \lambda(\cdot), H)$

- F (renewal distribution): Weibull distribution
- $\lambda(\cdot)$ (ageing function): Power law,
i.e. $\lambda(t) = abt^{b-1}$
- H (heterogeneity): Gamma distribution

ESTIMATION BY MCMC –
REQUIRES (ESSENTIALLY ONLY) COMPUTATION OF LIKELIHOOD:

Recall that for given value of a_j the likelihood for the j th system is

$L_j(a_j) = \text{likelihood for TRP}(F, a_j\lambda(\cdot))$

Full likelihood of HTRP($F, \lambda(\cdot), H$):

$$L = \prod_{j=1}^m \int L_j(a_j) dH(a_j)$$

This expectation easily gets messy.

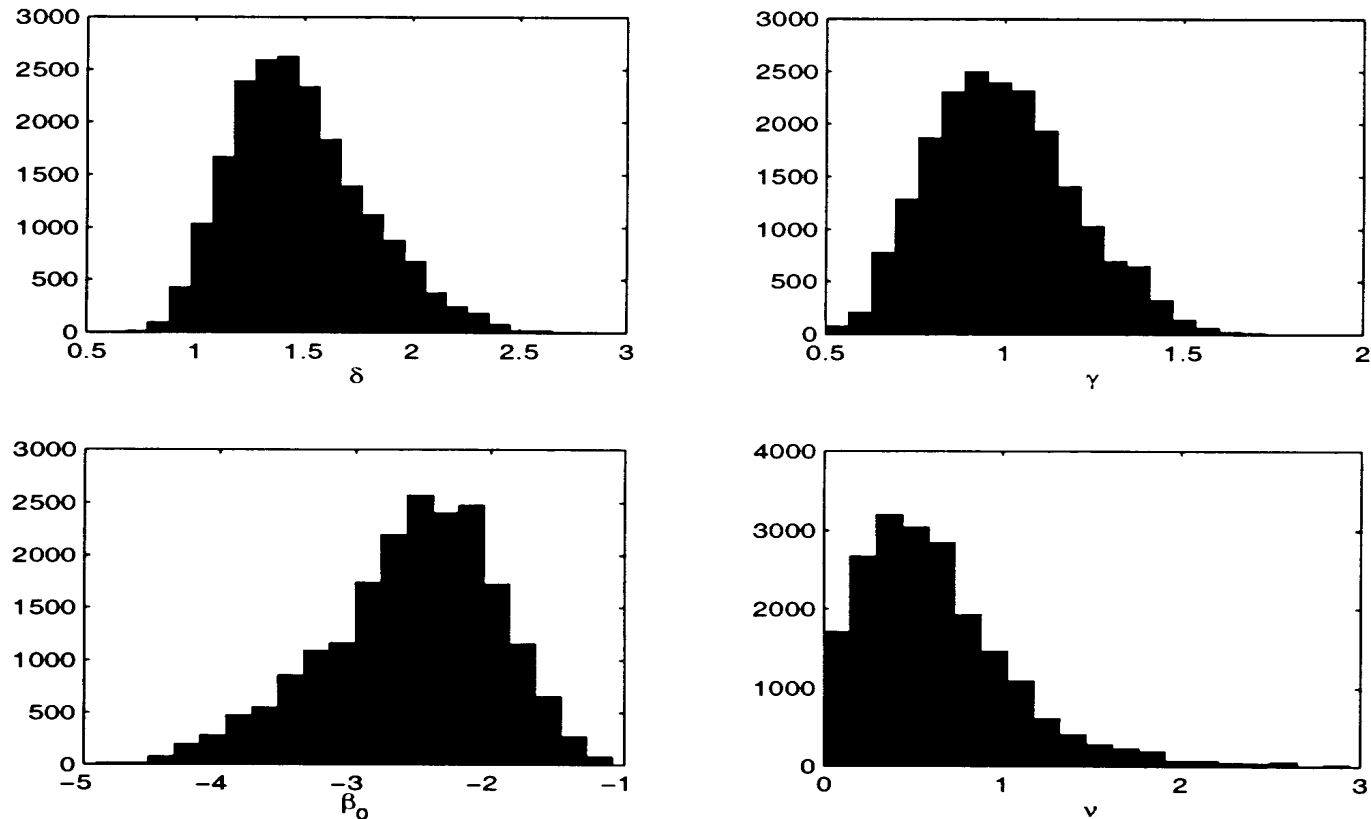
Suggestion: Compute by simulation, simulating k gamma(δ)-distributed a , and taking the average.

(For "clever" ways, see Balakrishnan 2003)

POSTERIOR DISTRIBUTIONS FOR VAGUE PRIORS – COMPUTED BY MCMC (METROPOLIS + RANDOM WALK SAMPLER)

Upper left: Trend (1 = no trend), *Upper right:* Weibull shape (1 = expon),
Lower left: log(scale), *Lower right:* Heterogeneity (variance)

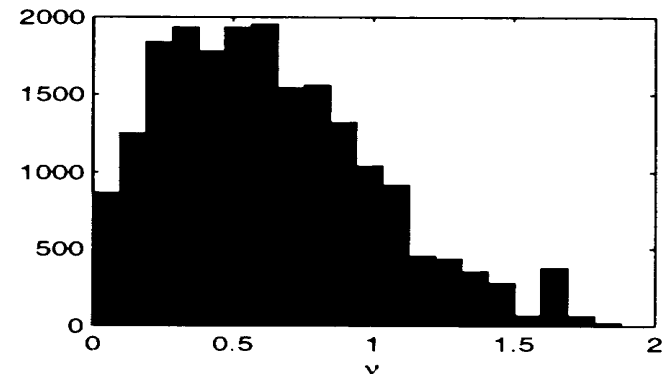
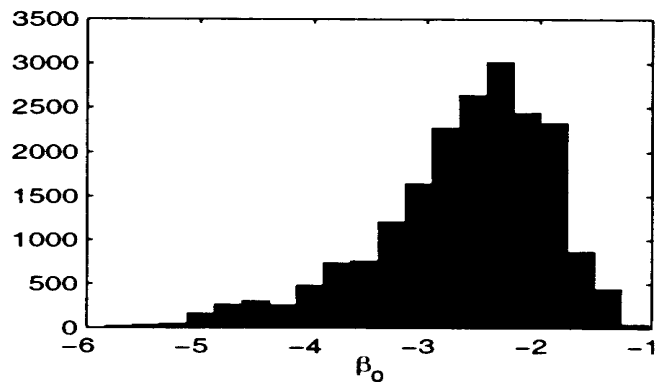
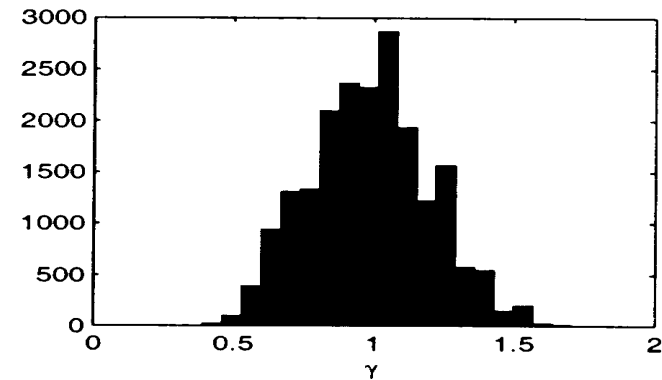
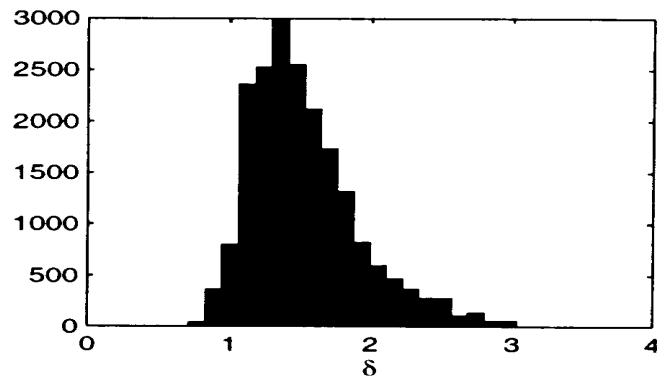
$k = 200$



POSTERIOR DISTRIBUTIONS FOR VAGUE PRIORS – COMPUTED BY MCMC (METROPOLIS + RANDOM WALK SAMPLER)

Upper left: Trend (1 = no trend), *Upper right:* Weibull shape (1 = expon),
Lower left: log(scale), *Lower right:* Heterogeneity (variance)

$$k = 20$$



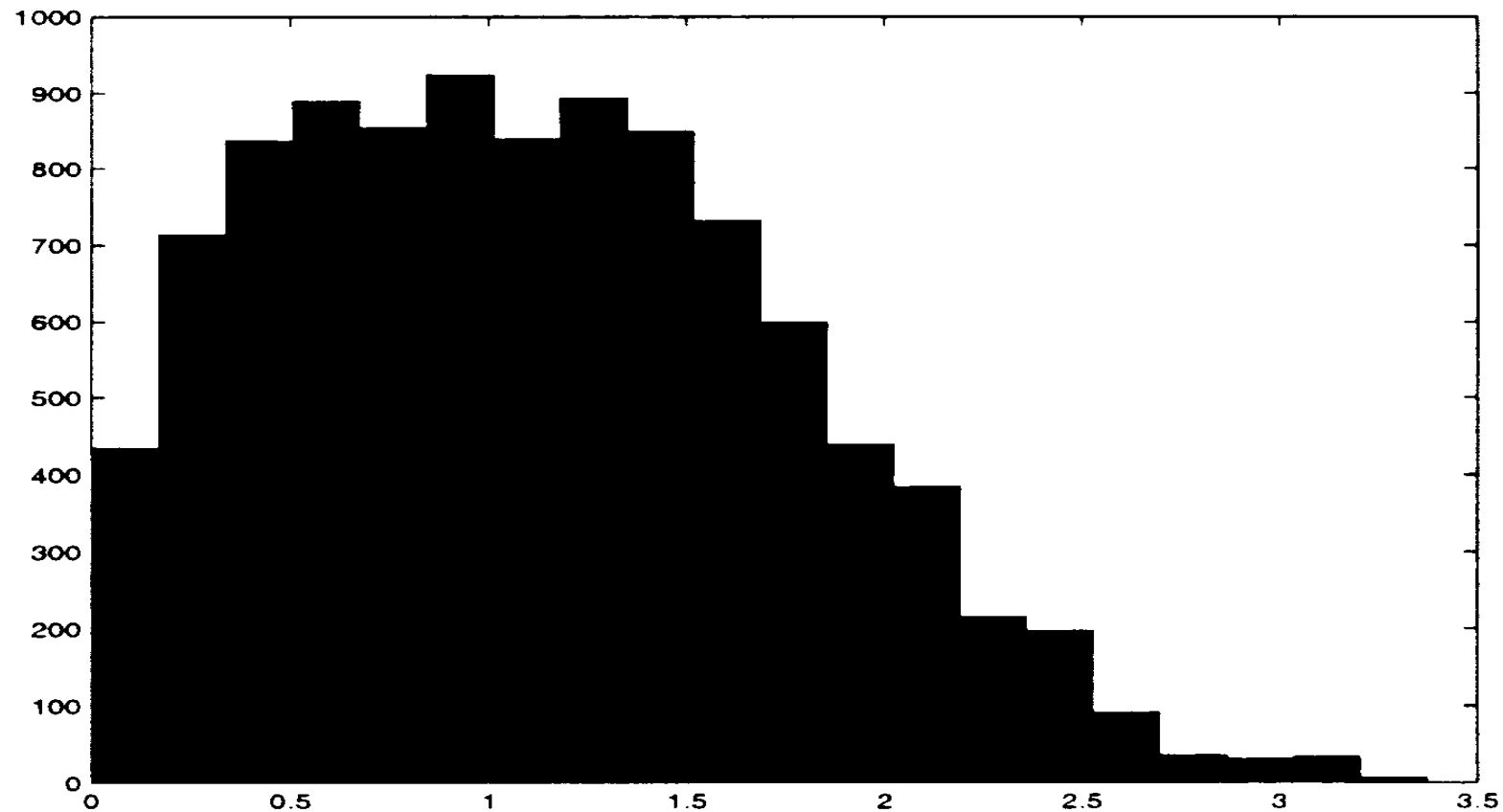
SPECIAL INTEREST: HETEROGENEITY PARAMETER

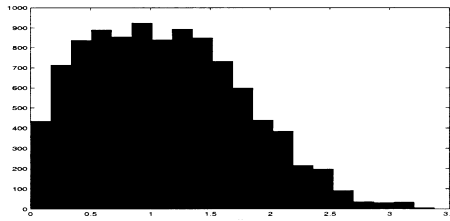
Simpler model with independence should be preferred
IF COMPATIBLE WITH DATA

Leads to study of

TESTS FOR $H_0 : \delta = 0$

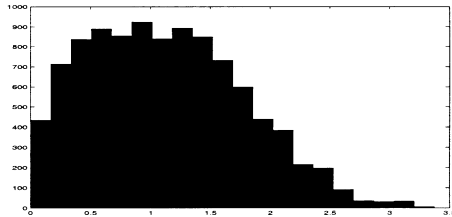
TYPICAL LOOK OF POSTERIOR FOR δ :





PROBLEMS WITH TESTING $H_0 : \delta = 0$:

1. $\delta = 0$ is on the *border* of the parameter set (frequentistic asymptotic theory invalid!)
2. $\delta = 0$ is a *sharp* hypothesis (Lindley's paradox in Bayesian testing!)
 - FREQUENTISTIC SOLUTION: Show that expression for likelihood is valid also for small and negative δ . Then *augment* the parameter set so that $\delta = 0$ becomes an *inner point*. Motivation: See figure.
 - BAYESIAN SOLUTIONS:
 1. Standard Bayesian hypothesis testing via Bayes factors (Lindley's paradox may apply). Simulations have shown that this is true, i.e. test favors H_0 also when $\delta = 0$ seems very unreasonable. Problem gets bigger with increasing number of systems.
 2. Using Bayesian reference criterion (e.g. Bernardo and Rueda, IS-Review, 2002, Bernardo 1999). Uses vague (reference) prior, no sharp hypothesis, uses objective discrepancy between models based on Kullback-Leibler, computes the posterior discrepancy.



POSSIBLE APPROACH: BAYESIAN ANALYSIS ALLOWING $\delta < 0$.

- Figure indicates how posterior density can be extended "by eye" to negative δ .
- Likelihood suggests that admissible negative δ are in absolute value less than approximately $\min_j (n_j \gamma)^{-1}$, where $1/\gamma$ is shape parameter of Weibull renewal distribution.
- Enables one to compute posterior probabilities of $\delta \leq 0$ and $\delta > 0$ and thus avoids sharp null hypothesis. Problems: What is an appropriate (noninformative) prior for negative values? What is an appropriate loss function?

CONCLUSIONS

- Heterogeneity between systems is an example of empirical Bayes
- Standard modelling of heterogeneity can be extended to trend-renewal processes
- Bayes formulation and solution by MCMC is efficient in case of complex likelihoods
- Question of homogeneity or heterogeneity is important for prediction
- Classical frequentistic testing is invalid, but can be modified
- Bayes testing: needs to be careful about purpose.
Open questions remain to be resolved!