www.inl.gov

Idaho National Laboratory

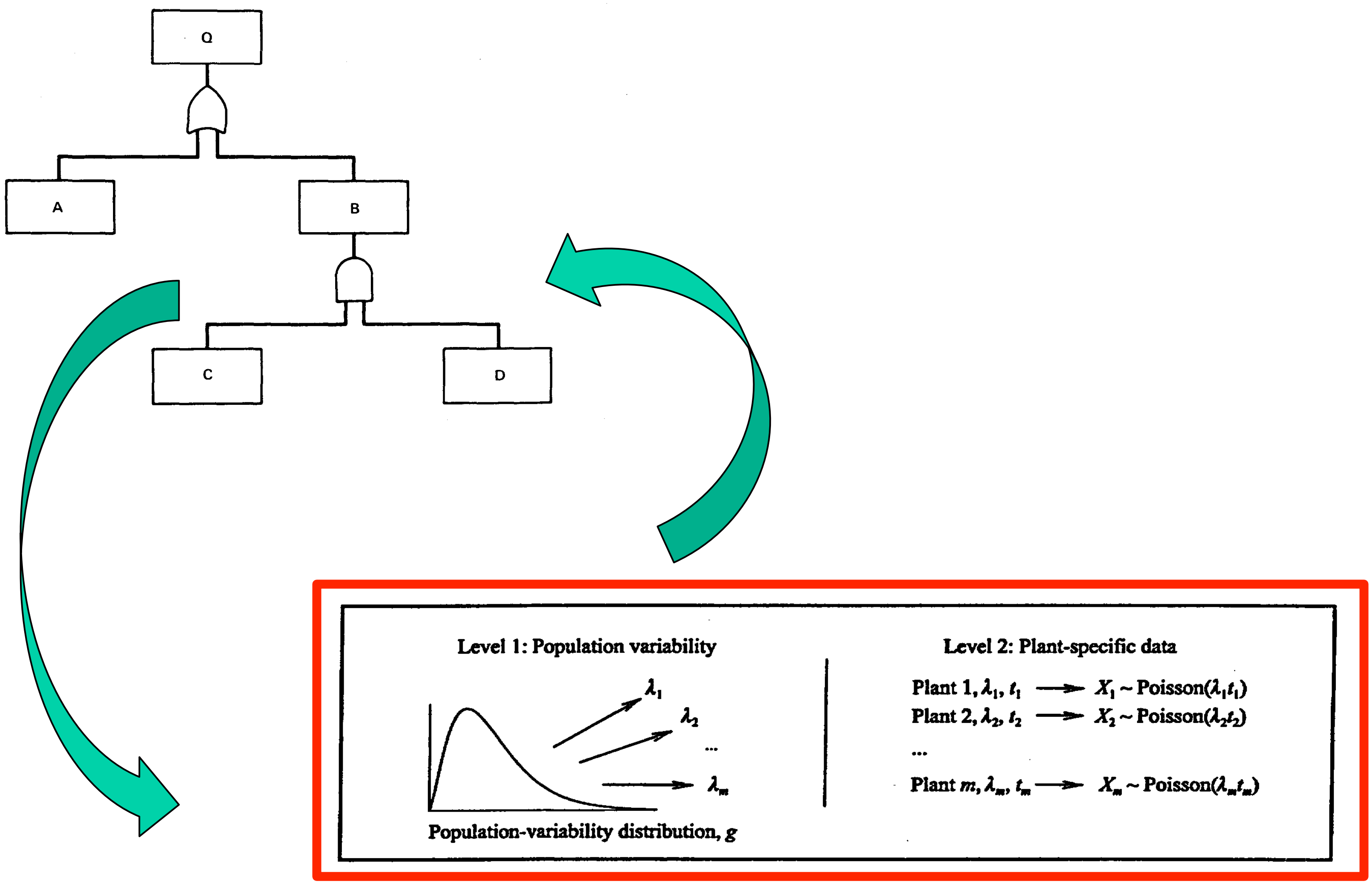
Performance of Empirical Bayes Estimation Techniques Used in Probabilistic Risk Assessment on Failure Data collected in U.S NRC Reactor Operating Experience Database

Andrei Gribok, Vivek Agarwal, Vaibhav Yadav

2018 Probabilistic Safety Assessment & Management Conference (PSAM 14)

Los Angeles , CA, USA September 16-21, 2018

Two Parts of PRA Analysis



SECY-17-0112. PLANS FOR INCREASING STAFF CAPABILITIES TO USE RISK INFORMATION IN DECISION-MAKING ACTIVITIES

Challenge \ Strategy	Strategy I *Evaluate and Update Guidance*	Strategy II *Develop Graded Approach*	Strategy III *Enhance Staff Training*	Strategy IV *Advance Risk Initiatives*	Strategy V *Increase Communication*
(1) Knowledge and Support	X		X		X
(2) Review Integration	X	X	X		X
(3) Guidance Development	X	X	X		X
(4) PRA Acceptability				X	X
(5) PRA Realism	X			X	X
(6) Risk Aggregation	X			X	X

- Bayesian analysis depends on prior elicitation/selection
- Prior elicitation/selection is a fundamental problem of Bayesian inference
- To produce a prior we need to convert information about observable values (number of failures, time) into information about unobservable variables (probability of failure, failure rate)
- Is there the best prior? Is there a best method to elicit prior? Is there the best method for parameter estimation?
- The historical data is our best, most objective, and reliable source of information about true probability of failure or failure rate
- The credibility of PRA analysis depends on how closely it reflects industry-wide historical data*

*ALI MOSLEH, "PRA: A PERSPECTIVE ON STRENGTHS, CURRENT LIMITATIONS, AND POSSIBLE IMPROVEMENTS", NUCLEAR ENGINEERING AND TECHNOLOGY, VOL.46 NO.1 FEBRUARY 2014

Bayesian Parameter Estimation

Bayesian inference

$$\pi(\lambda/x,\alpha) = \frac{L(x/\lambda)\cdot\pi(\lambda/\alpha)}{\int L(x/\lambda)\cdot\pi(\lambda/\alpha)d\lambda}$$

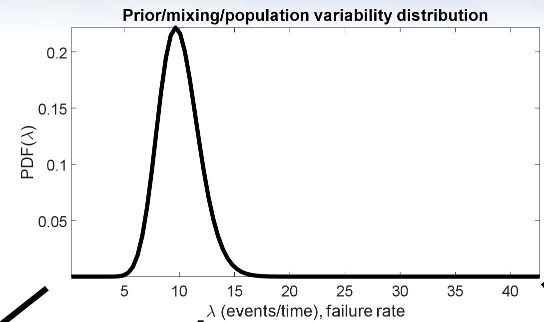
Prior predictive distribution

$$\pi(x/\alpha) = \int_0^\infty L(x/\lambda)\cdot\pi(\lambda/\alpha)d\lambda$$

Empirical Bayes

$$\pi(\lambda/x,\alpha) = \frac{L(x/\lambda)\cdot\pi(\lambda/\hat{\alpha})}{\int L(x/\lambda)\cdot\pi(\lambda/\hat{\alpha})d\lambda}$$

Gamma-Poisson Compound Distribution Model



Gamma distribution is a continuous distribution

λ_1 λ_2 λ_k

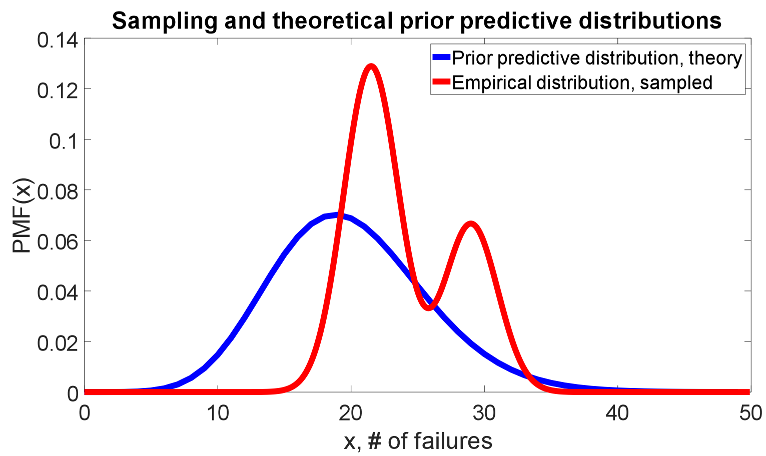
Poisson distribution is a discrete distribution

$$P(X=x_1 \text{ in } [0,t_1]/\lambda_1) = \frac{(\lambda_1 t_1)^{x_1}}{x_1!} e^{-\lambda_1 t_1}$$

$$P(X=x_2 \text{ in } [0,t_2]/\lambda_2) = \frac{(\lambda_2 t_2)^{x_2}}{x_2!} e^{-\lambda_2 t_2}$$

$$P(X=x_k \text{ in } [0,t_k]/\lambda_k) = \frac{(\lambda_k t_k)^{x_k}}{x_k!} e^{-\lambda_k t_k}$$

$\boldsymbol{x_1, t_1}$ $\boldsymbol{x_2, t_2}$ $\boldsymbol{x_k, t_k}$



Negative binomial (NB) distribution is a discrete distribution

Prior Predictive Distribution for Gamma-Poisson Compound Distribution Model

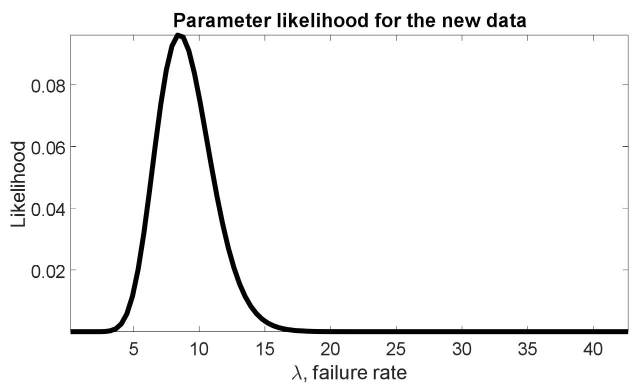
$$P(\mathrm{X} = x/\alpha\,, \beta, T\,) = \mathrm{NB}\left(\mathrm{x};\, \alpha, \frac{T}{\beta + T}\right)$$

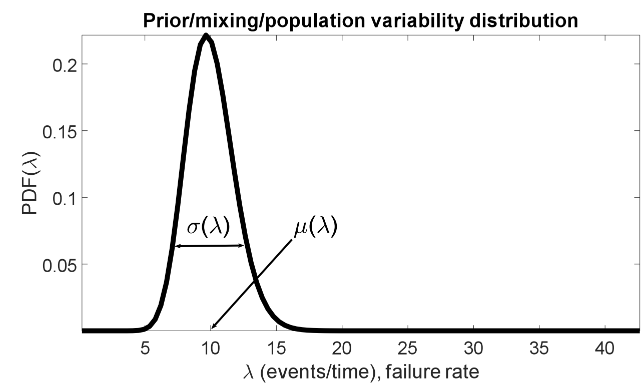
$$= \frac{\Gamma(x+\alpha)}{x!\,\Gamma(\alpha)} \cdot \left[\frac{T}{\beta + T}\right]^{x} \cdot \left[\frac{\beta}{\beta + T}\right]^{\alpha}$$

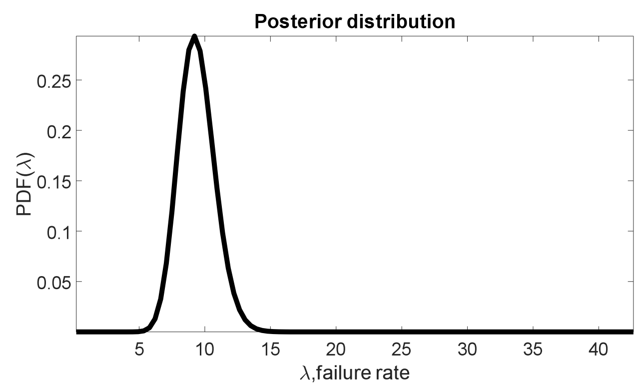
$$= \int_0^{\infty} \underbrace{\frac{(\lambda \cdot T)^x}{x!} e^{-\lambda \cdot T}}_{Likelihood} \cdot \underbrace{\frac{\beta^{\alpha} \cdot \lambda^{\alpha - 1} \cdot e^{-\lambda \cdot \beta}}{\Gamma(\alpha)}}_{Prior} d\lambda$$

Why Do We Need Nonparametric Bayes?

$$\text{Gamma}(\lambda/\alpha,\beta) = \frac{\beta^{\alpha}\cdot\lambda^{\alpha-1}\cdot e^{-\lambda\cdot\beta}}{\Gamma(\alpha)},\ \mu(\lambda)=\alpha/\beta,\ \sigma(\lambda)=\sqrt{\alpha}/\beta$$







Performance of Different Estimation Techniques on NROD MDP FR Data

- $\text{RMSE} = \sqrt{\frac{1}{95}\sum_{i=1}^{95}\left(\lambda_i^{est} - \lambda_i^{"true"}\right)^2}$
- Posterior mean is posterior summary statistics
- Plant-specific MLE
- Industry average MLE
- The James–Stein (JS) Estimator
- Jeffreys noninformative prior, Gamma(0.5,0)
- Constrained noninformative prior (CNI), Gamma (0.5,0.5·Industry average)
- Method of moments
- Maximum likelihood fitting of Gamma (α,β)
- Maximum Likelihood II (MLE2)
- Nonparametric empirical Bayes (NPEB)

U.S. NRC Reactor Operating Experience Database (NROD)

- NROD contains failure data for thousands of components
- The data typically span the period from 1998 to 2015 (2016,2017 is now available)
- Binomial data are presented as number of failures and number of demands
- Poisson data are presented as number of failures and time period

1998-2000
Estimate

2001-2015
“Truth”

NROD Data for Motor-Driven Pump (MDP) Failure to Run (FR)

Plant Name	Device count	Failure Count (1998-2000)	Run hours (1998-2000)	Failure Count (2001-2015)	Run hours (2001-2015)
Plant 1	3	2	54938.82	2	268467.12
Plant 2	2	0	26280.00	1	131400.00
Plant 3	6	0	88100.52	0	346328.61
…..					
…..					
Plant 94	9	5	169077.24	4	918448.20
Plant 95	12	3	139221.72	4	816829.16
Total	624	55	9312985.75	168	50313464.88

λ_i^{est}

$$\lambda_i^{"true"} = \frac{(Failure\ Count)_i}{(Run\ hours)_i}$$

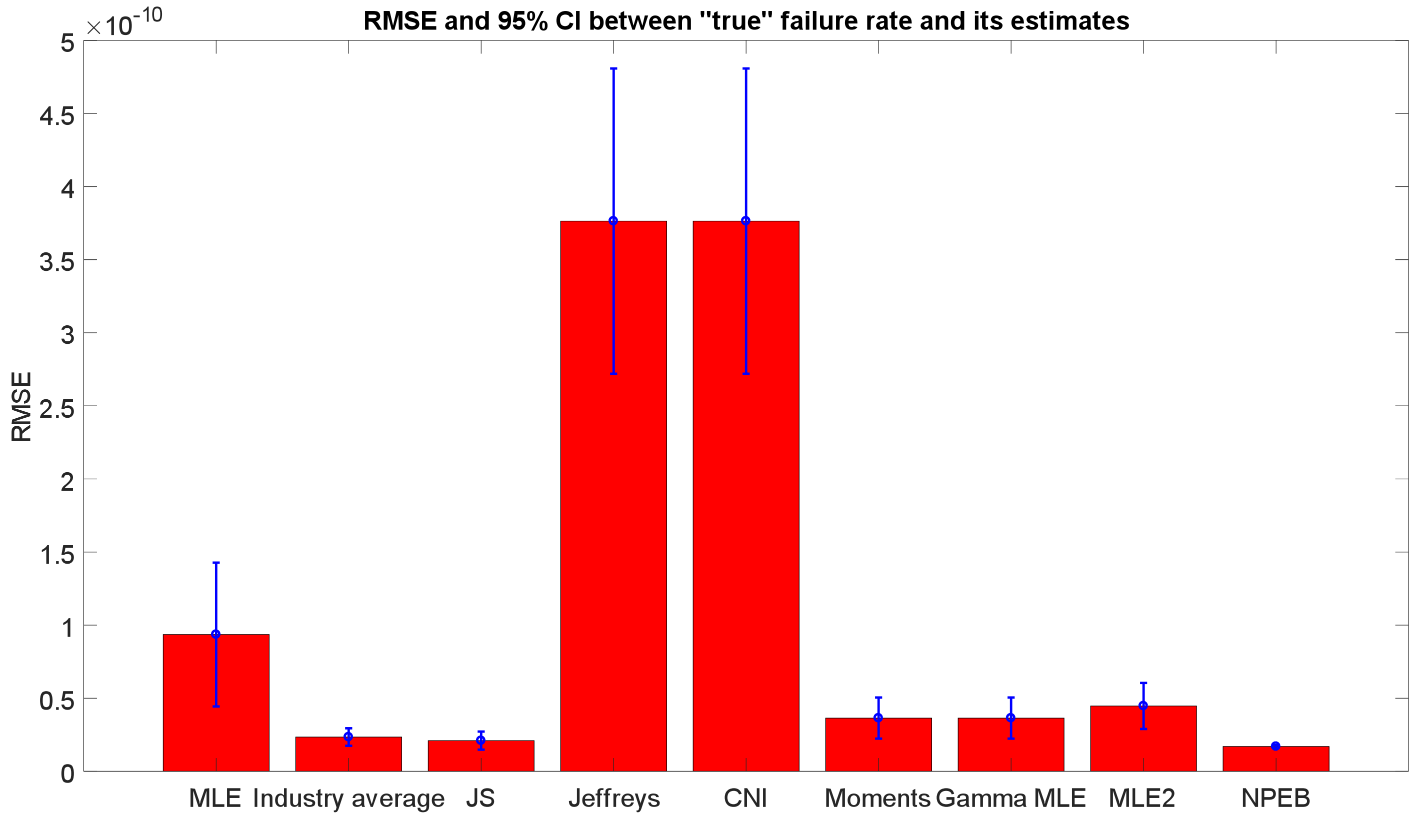
The chi-square test of homogeneity (data pullability)

- Test if population variability distribution is not a degenerate one

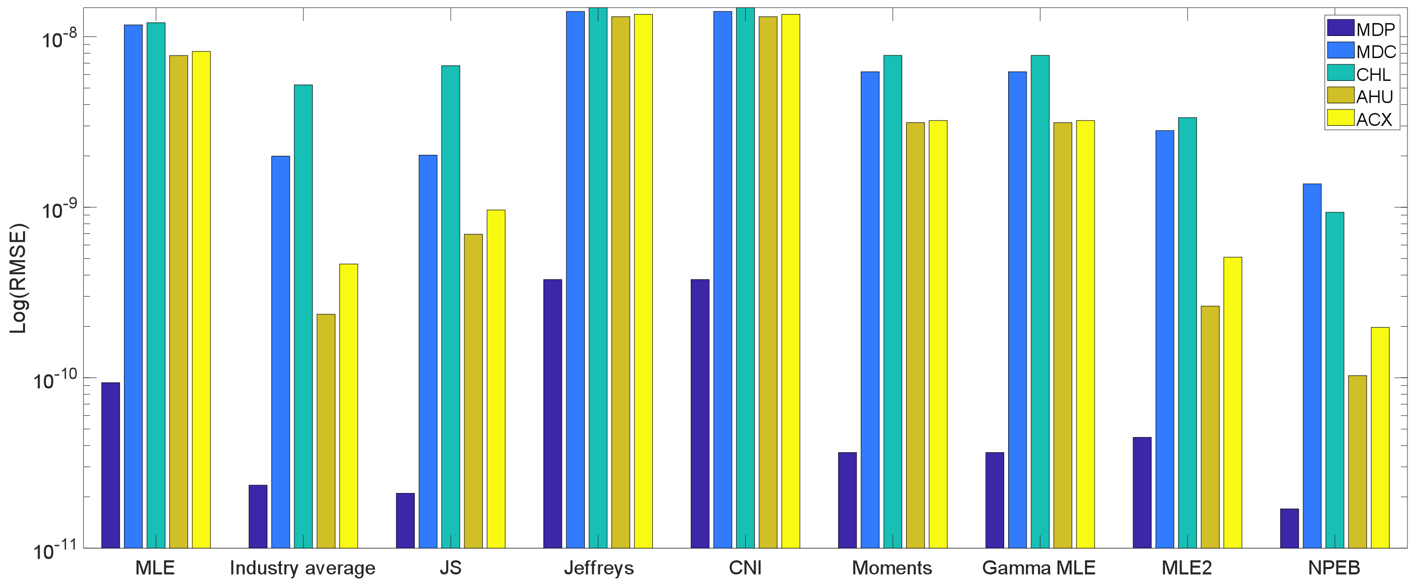
- $H_0: \lambda_1 = \lambda_2 = \ldots = \lambda_k$, all failure rates are the same
- $H_1: \lambda_1 \neq \lambda_2 = \ldots = \lambda_k$, at least one λ_i is different from others

- $\chi^2 = \sum_{i=1}^{k} \frac{\left[\lambda_i^{obs} - \lambda_i^{exp}\right]^2}{\lambda_i^{exp}}$, if $\chi^2 > \chi^2_{0.95}$(df=k-1) then H_0 is rejected and H_1 is accepted
- For MDP FR data, χ^2=143.5877, $\chi^2_{0.95}$(94)=117.6317

Performance of Different Estimation Techniques on NROD MDP FR Data

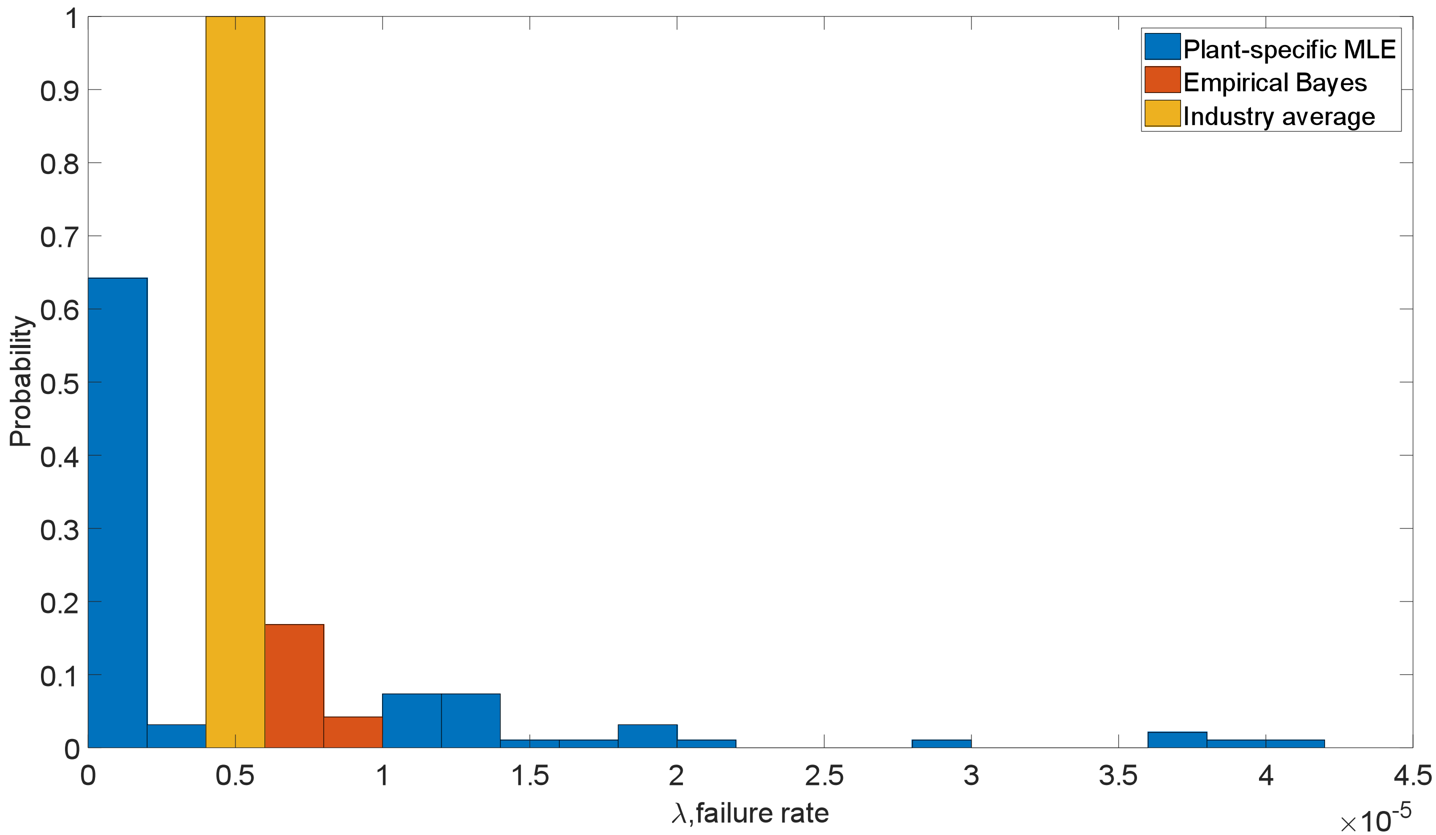


Performance of Different Parameter Estimation Methods for Different Components on NROD Data Set.

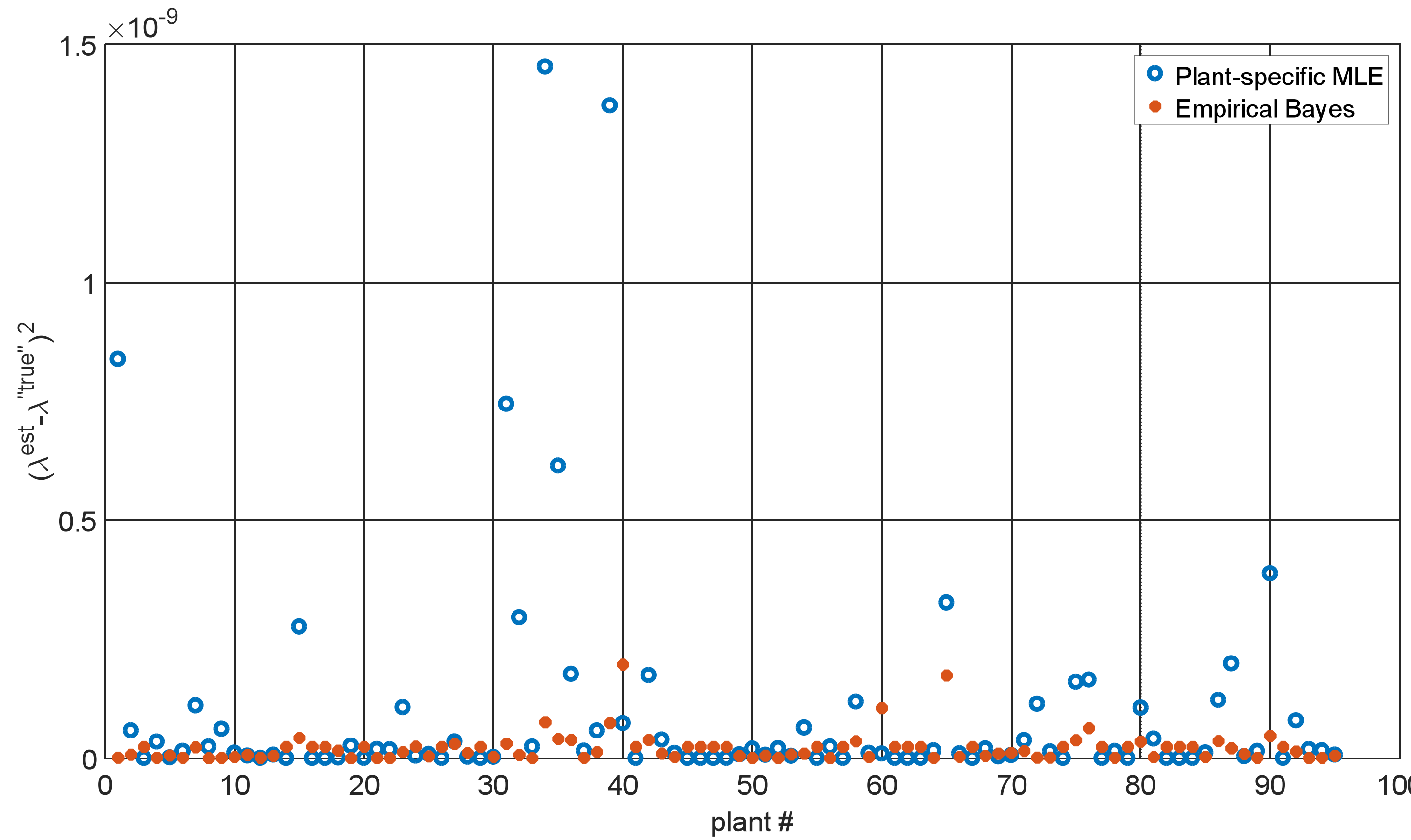


- ACX_FR: Accumulator Fail to Run (ACX)
- AHU_NR_FTR: Air Handling Unit Normally Running Fails to Run (AHU)
- CHL_FR: Chiller Unit Fails to Run (CHL)
- MDC_FR: Motor Driven Compressor Fail to Run (MDC)
- MDP_FR: Motor Driven Pump Fail to Run (MDP)

Empirical Bayes is a shrinkage estimator. It is biased.



Why do not we always use EB in lieu of MLE?



MLE is better for 38 plants, including 13 nonzero λs

Number of plants for which Plant-Specific MLE Outperforms NPEB

Component	# of plants in NROD database	# plants for which NPEB has larger RMSE than plant-specific MLE
ACX	34	13
AHU	34	14
CHL	22	3
MDC	29	12
MDP	99	33

Conclusions

- Nonparametric EB method is more flexible, allowing to change prior's mean and variance independently
- Nonparametric EB method requires selection of two parameters: kernel width for kernel density estimation and regularization parameter for Bayesian deconvolution
- EB and hierarchical Bayes improve industry-wide performance with respect to plant-specific MLE, however, they do not guarantee an improvement for a specific plant
- Single-stage noninformative prior is likely to degrade the accuracy of the estimate in comparison to MLE
- Bayesian methods are shrinkage (biased) estimators. The variance of the estimate is shrunk toward zero and the value of the estimator is shrunk towards the mean value of prior distribution
- If we need industry-wide improvement for the accuracy of parameter estimates, EB is the way to go, however, if we need to improve estimate for a specific plant, all options should be considered

INL

Idaho National Laboratory

Fredholm Integral Equation of the First Kind
Bayesian Deconvolution

$$g(x) = \int_a^b K(\lambda, x) f(\lambda) d\lambda, \qquad c \leq x \leq d$$

$$\underbrace{\hat{g}(x)}_{predictive} = \int_0^\infty \underbrace{\frac{(\lambda \cdot T)^x}{x!} e^{-\lambda \cdot T}}_{Kernel} \cdot \underbrace{f(\lambda)}_{prior} \, d\lambda, 0 \leq \lambda \leq \infty, T \; specified, x = 0,1,2 \ldots$$

$$\underbrace{argmin}_{f(\lambda)} \left\{ \left\| \underbrace{\int_0^\infty \underbrace{\frac{(\lambda \cdot T)^x}{x!} e^{-\lambda \cdot T}}_{Kernel} \cdot \underbrace{f(\lambda)}_{prior} \, d\lambda - \underbrace{\hat{g}(x)}_{predictive}}_{Residuals} \right\|_2^2 + \underset{weight}{\epsilon^2} \underbrace{\int_0^\infty \left(\frac{d^2 f(\lambda)}{d\lambda^2} \right)^2 d\lambda}_{smoothing\ norm} \right\}$$

Discretization of the Integral Equation

$$\underbrace{\begin{pmatrix}\hat{g}(x_1)\\ \hat{g}(x_2)\\ \vdots \\ \hat{g}(x_m)\end{pmatrix}}_{predictive} = \underbrace{\begin{pmatrix}\frac{(\lambda_1 T)^{x_1}}{x_1!}e^{-\lambda_1 T} & \frac{(\lambda_2 T)^{x_1}}{x_1!}e^{-\lambda_2 T} & \cdots & \frac{(\lambda_n T)^{x_1}}{x_1!}e^{-\lambda_n T}\\ \frac{(\lambda_1 T)^{x_2}}{x_2!}e^{-\lambda_1 T} & \frac{(\lambda_2 T)^{x_2}}{x_2!}e^{-\lambda_2 T} & \cdots & \frac{(\lambda_n T)^{x_2}}{x_2!}e^{-\lambda_n T}\\ \vdots & \vdots & \cdots & \vdots \\ \frac{(\lambda_1 T)^{x_m}}{x_m!}e^{-\lambda_1 T} & \frac{(\lambda_2 T)^{x_m}}{x_m!}e^{-\lambda_2 T} & \cdots & \frac{(\lambda_n T)^{x_m}}{x_m!}e^{-\lambda_n T}\end{pmatrix}}_{kernel\ matrix} \cdot \underbrace{\begin{pmatrix}\tilde{f}_1\\ \tilde{f}_2\\ \vdots \\ \tilde{f}_n\end{pmatrix}}_{prior}$$

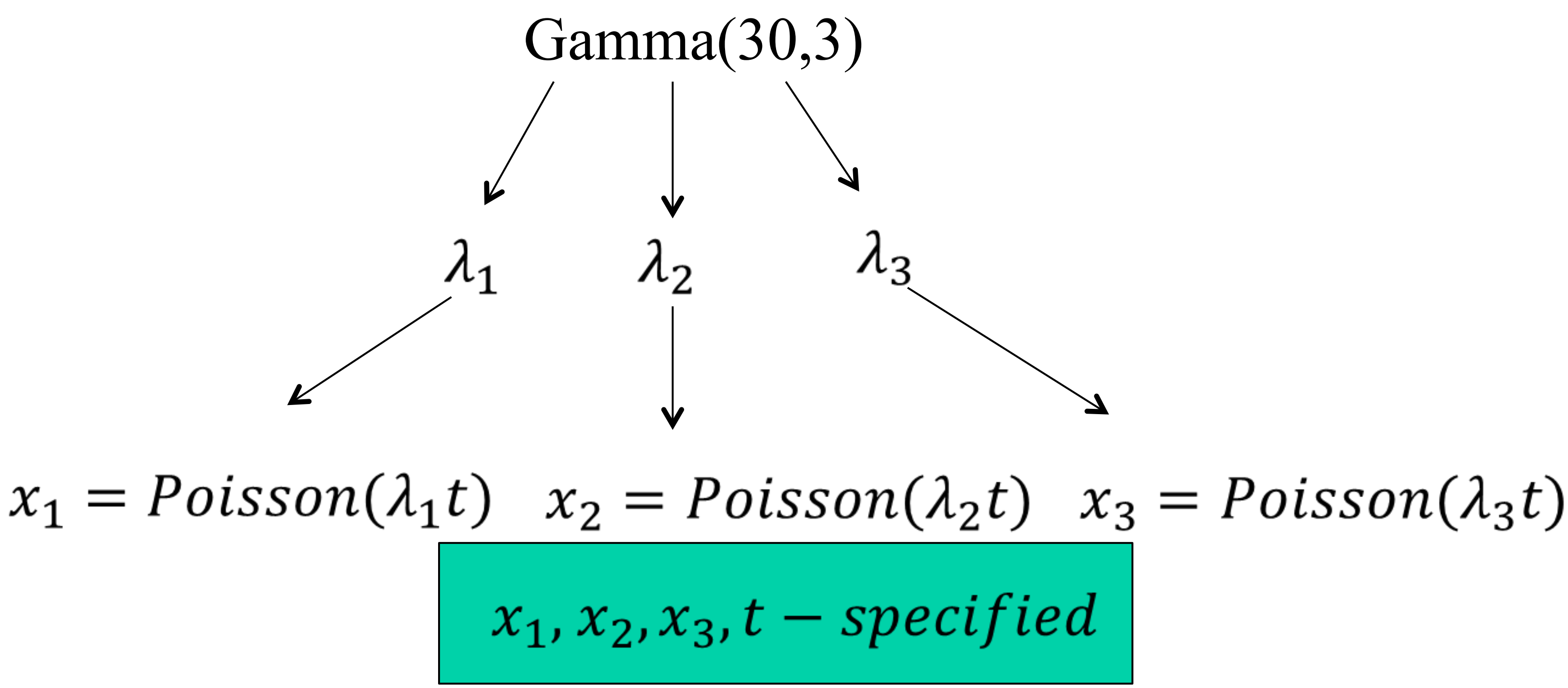
$$\underbrace{\hat{g}}_{m\times 1} = \underbrace{P}_{m\times n} \cdot \underbrace{\tilde{f}}_{n\times 1}$$

$$\underbrace{argmin}_{\tilde{f}}\left\{\left\|P\cdot\tilde{f}-\hat{g}\right\|_2^2 + \lambda^2\cdot\left\|L\cdot\tilde{f}\right\|_2^2\right\},\ L = \begin{pmatrix}1 & -2 & 1 & \cdots & 0\\ \ddots & \ddots & \ddots & \ddots & \ddots\\ 0 & \cdots & 1 & -2 & 1\end{pmatrix} \in R^{(n-2)\times n}$$

$$\underbrace{argmin}_{\tilde{f}}\left\|\underbrace{\begin{pmatrix}P\\ \lambda L\end{pmatrix}}_{m+n-2\times n}\cdot\underbrace{\tilde{f}}_{n\times 1} - \underbrace{\begin{pmatrix}\hat{g}\\ 0\end{pmatrix}}_{m+n-2\times 1}\right\|_2^2, \qquad subject\ to\ \tilde{f}\geq 0$$

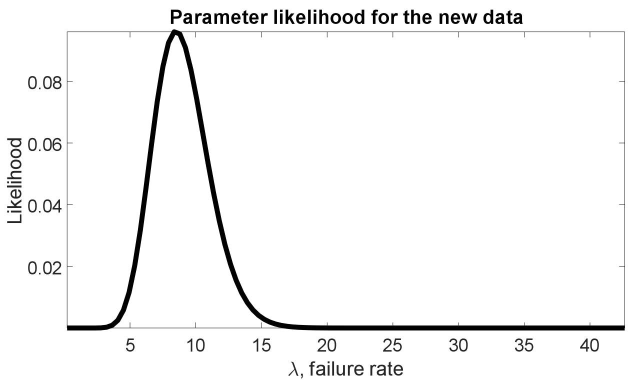
Prior Selection Methods

- Method of moments
- Maximum likelihood fitting of Gamma (α,β)
- Maximum Likelihood II (Evidence maximization)
- Nonparametric empirical Bayes (NPEB)

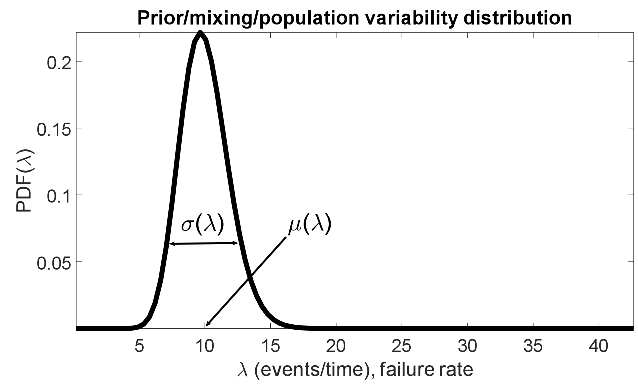


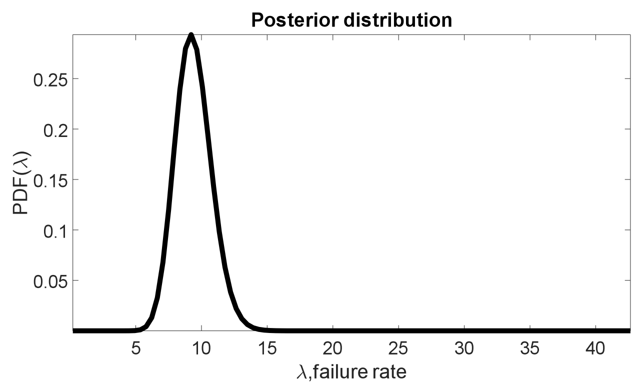
Inference for a New Plant

$$x_{new} = Poisson(\lambda_{true}^{new} \cdot t) \longleftarrow \lambda_{true}^{new} \longleftarrow \text{Gamma}(30,3)$$



+





$$\lambda_{est}^{new} = \int_0^{\infty} \lambda \cdot p(\lambda) d\lambda$$

Performance Measures

- Root Mean Squared Error (RMSE) between true prior and estimated one

$$RMSE = \frac{\|f_{true} - f_{est}\|_2}{\sqrt{n}}$$

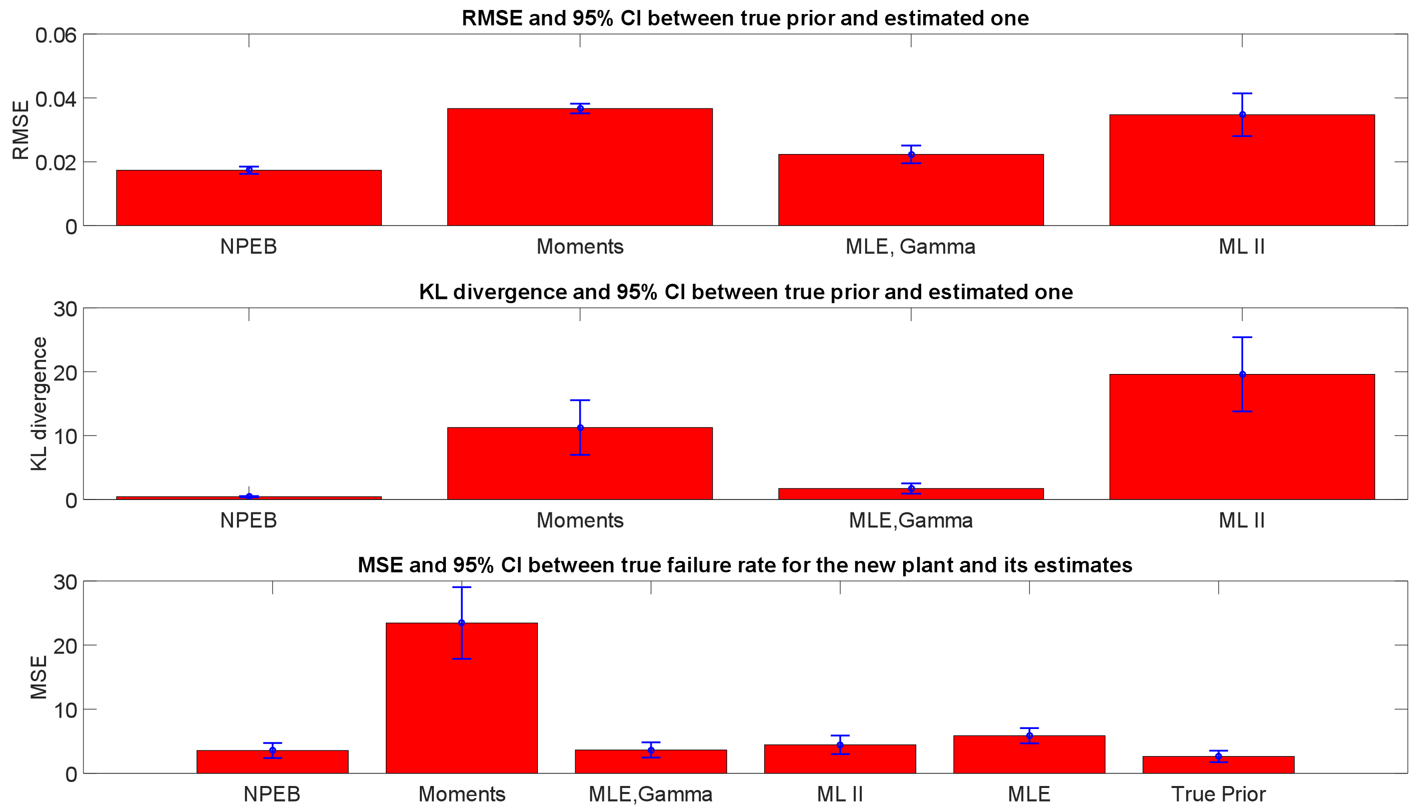
Kullback–Leibler (KL) divergence

$$KL = \sum_{i=1}^{n} f_{true}^{i} \cdot log\left[\frac{f_{true}^{i}}{f_{est}^{i}}\right]$$

MSE between λ_{true}^{new} and λ_{est}^{new}

$$\text{MSE} = \left[\lambda_{true}^{new} - \lambda_{est}^{new}\right]^2$$

Comparison of Different Prior Estimation Techniques



Applying NPEB to MDP FR Data