Chapter 10

Introduction to Bayesian Methods for Reliability Data Analysis

W. Q. Meeker, L. A. Escobar, and F. G. Pascual
Iowa State University, Louisiana State University, and Washington State University.

Based on Meeker, Escobar, and Pascual (2021): *Statistical Methods for Reliability Data, Second Edition*, John Wiley & Sons Inc.

May 24, 2021
10h 56min

Chapter 10
Introduction to the Use of Bayesian Methods for Reliability Data

Topics discussed in this chapter are:

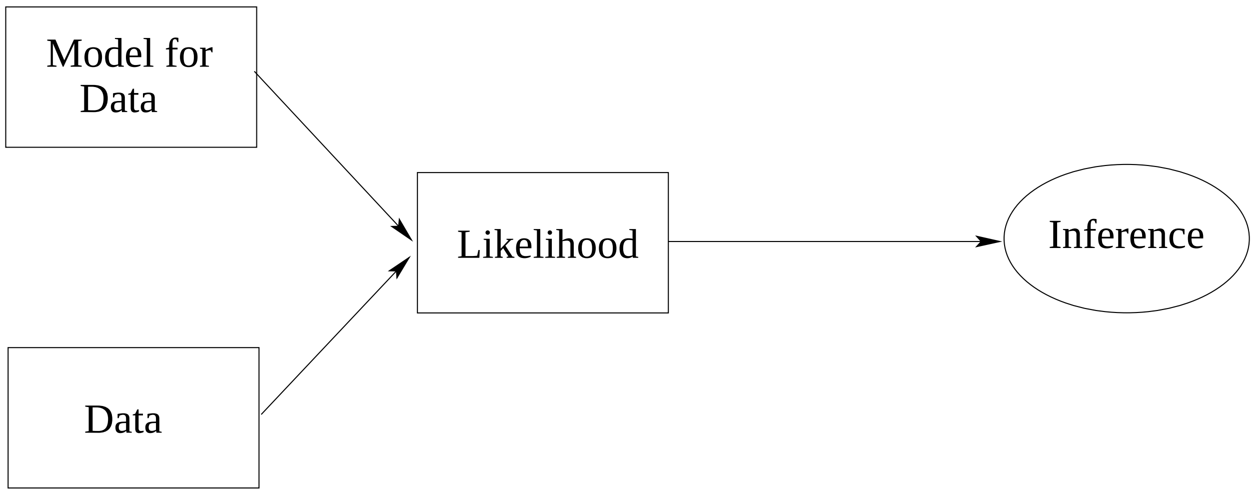
- The relationship between likelihood and Bayesian inference methods, background, and motivation.
- Bayes' Theorem, prior information, and the importance of a good parameterization
- An example using the simple accept/reject Monte Carlo method.
- Bayesian point estimation and credible interval methods.
- The basic ideas of using Markov Chain Monte Carlo (MCMC) simulation to draw from a joint posterior distribution.
- A second example that illustrates the power of Bayesian methods to answer important reliability-related questions.

Chapter 10

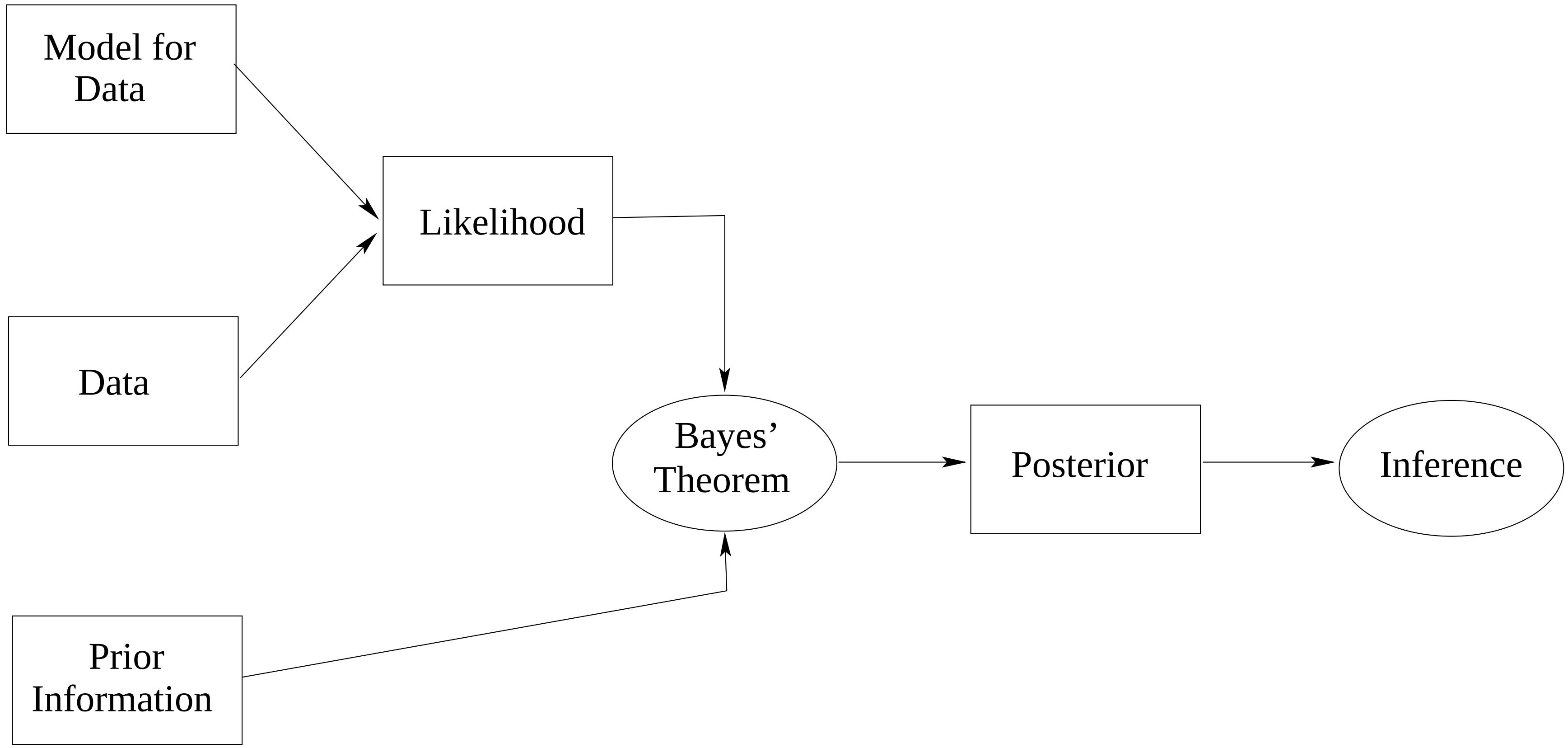
Segment 1

Background and Motivation

Likelihood Inference



Bayesian Inference



Background

- Bayesian methods augment likelihood with **prior** information.
- A probability distribution is used to describe our **prior** beliefs about a parameter or set of parameters.
- Sources of prior information:
  - Subjective Bayes: subjective prior information, usually informative.
  - Objective Bayes: diffuse prior (approximately noninformative) or matching.
- In many practical applications, there is informative prior information only in one dimension (e.g., for one parameter).

Motivation for Using Bayesian Methods

- Bayesian methods allow an analyst to incorporate prior information into a data analysis/modeling problem to supplement limited data, often providing important improvements in precision (or cost savings).

- Bayesian methods can handle, with relative ease, complicated data-model combinations for which no ML software exists (e.g., combinations of nonlinear relationships, random effects, and censored data).

- For complicated models with **weakly informative prior distributions**, Bayesian methods provide trustworthy confidence (credible) intervals.

- In applications with multiple subpopulations, hierarchical models allow borrowing strength across the subpopulations.

Chapter 10

Segment 2

Bayes' Theorem, Prior Information, and Parameterization

Combining Data and Prior Information Using Bayes' Theorem

- Prior information on a vector of parameters $\boldsymbol{\theta}$ is expressed in terms of a joint prior distribution (pdf) $f(\boldsymbol{\theta})$.

- We observe data which for the specified model has likelihood $L(\text{DATA}|\boldsymbol{\theta}) \equiv L(\boldsymbol{\theta}; \text{DATA})$.

- Using Bayes' Theorem, the conditional distribution of $\boldsymbol{\theta}$ given the data (also known as the **joint posterior** distribution of $\boldsymbol{\theta}$) is

$$f(\boldsymbol{\theta}|\text{DATA}) = \frac{L(\text{DATA}|\boldsymbol{\theta}) f(\boldsymbol{\theta})}{\int L(\text{DATA}|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}}$$

  where the multiple integral is computed over the region $f(\boldsymbol{\theta}) > 0$.

- Doing the integration directly is usually intractable.

- Simulation methods can often be used to generate "draws" from $f(\boldsymbol{\theta}|\text{DATA})$, the joint posterior distribution of $\boldsymbol{\theta}$.

Differences Between Bayesian and Non-Bayesian Inference

- Nuisance parameters:
  - Large-sample likelihood theory suggests maximizing out nuisance parameters, giving profile likelihood for parameters or functions of parameters that are of interest.
  - Bayesian methods integrate out nuisance parameters to get marginal distributions of parameters or functions of parameters that are of interest.
- There are not important differences between integrating and maximizing out nuisance parameters when the joint posterior distribution is symmetric (large samples).
- Justification:
  - Non-Bayesian methods are justified on the basis of large-sample approximations and related asymptotic theory.
  - Bayesian methods are justified on the basis of probability theory.

Sources of Prior Information

- Informative:
  - Previous experience with the same kind of failure mechanism.
  - Physical, chemical, and mechanical theory.
  - Expert knowledge.
- Weakly informative, vague (or approximately non-informative):
  - Uniform over infinite range of parameter (improper prior distribution).
  - Uniform over finite range of a parameter (or function of the parameters).
  - Normal (or truncated normal) distribution with a large variance.
- Often we have an informative prior distribution in one dimension and weakly informative in other dimensions (i.e., for other parameters).

Effect of Using Weakly Informative (or Vague) Prior Distributions

- For a uniform prior distribution $f(\theta)$ (possibly improper) across all possible values of θ

$$f(\boldsymbol{\theta}|\text{DATA}) = \frac{L(\boldsymbol{\theta})f(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta})f(\boldsymbol{\theta})d\boldsymbol{\theta}} = \frac{L(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta})d\boldsymbol{\theta}}$$

  which indicates that the joint posterior distribution $f(\boldsymbol{\theta}|\text{DATA})$ is proportional to the likelihood.

- Other weakly informative prior distributions also result in a joint posterior distribution that is approximately proportional to the likelihood if $L(\boldsymbol{\theta})$ is large relative to $f(\boldsymbol{\theta})$.

Expert Opinion and Eliciting Prior Information

- Identify parameters that, from past experience (or data), can be specified approximately independently (e.g., for high reliability applications a small quantile and the Weibull shape parameter).

- Determine for which parameters there is useful informative prior information.

- For parameters for which there in **no** useful informative prior information, determine the form and range of the weakly informative prior distribution (e.g., normal or truncated normal with a large standard deviation).

- For parameters for which there is useful informative prior information, specify the form and range of the distribution (e.g., lognormal with 99% content between two specified points).

Importance of a Good Parameterization

- One needs to think carefully about how to define parameters in a model.

- We would like to have stable parameters.

- Stable parameters generally correspond to things that one can effectively identify in a plot of the data.

- Use of Stable parameters generally leads to
  - ▶ More sure and faster convergence of estimation algorithms.
  - ▶ Parameters for which it is easier to elicit prior information.

A Useful Reparameterization of the Weibull Distribution

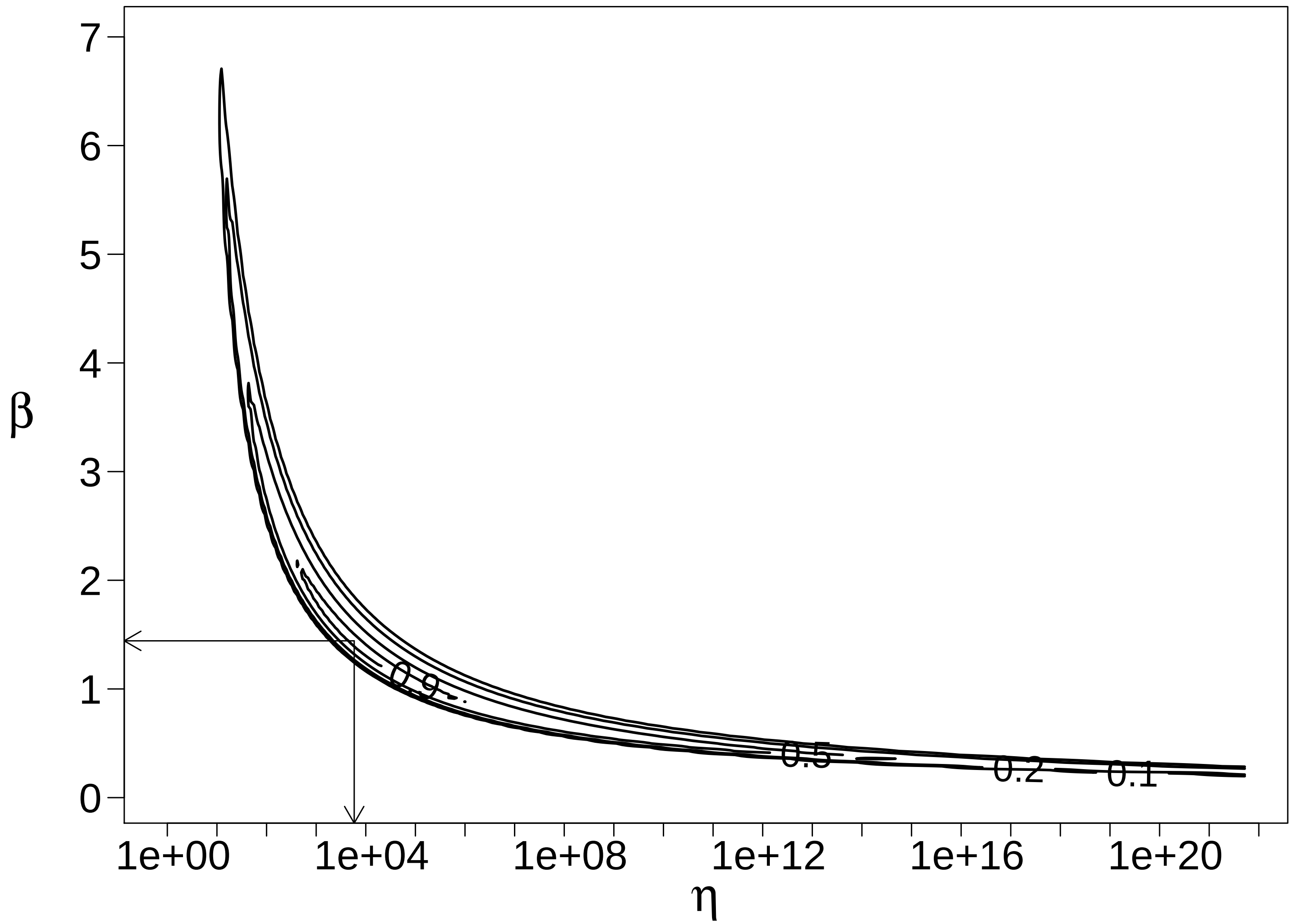
From Chapter 4, the Weibull distribution is usually written as

$$\Pr(T \le t; \eta, \beta) = F(t; \eta, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \quad t > 0.$$

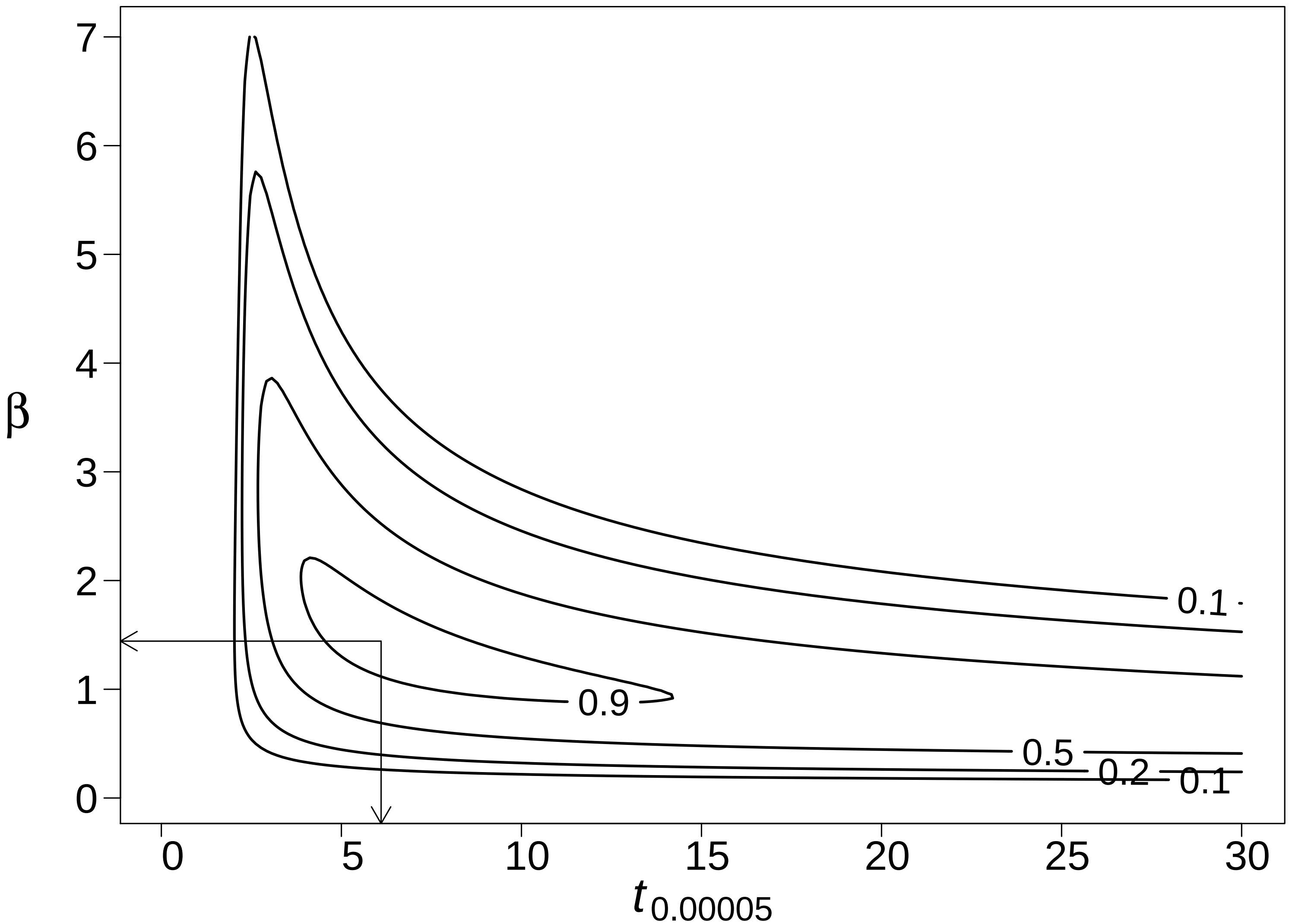
Replacing η with $\eta = t_p/[-\log(1-p)]^{1/\beta}$ gives

$$\begin{aligned}\Pr(T \le t; t_p, \beta) = F(t; t_p, \beta) &= 1 - \exp\left[-\left(\frac{t}{t_p/[-\log(1-p)]^{1/\beta}}\right)^{\beta}\right] \\ &= 1 - \exp\left[\log(1-p)\left(\frac{t}{t_p}\right)^{\beta}\right], \quad t > 0.\end{aligned}$$

Contour Plots of the Weibull Distribution Relative Likelihood for the Extreme Heavy-Censoring Example Using the Traditional (η, β) Parameterization



Contour Plots of the Weibull Distribution Relative Likelihood for the Extreme Heavy-Censoring Example Using the stable $(t_{0.00005}, \beta)$ Parameterization



Chapter 10

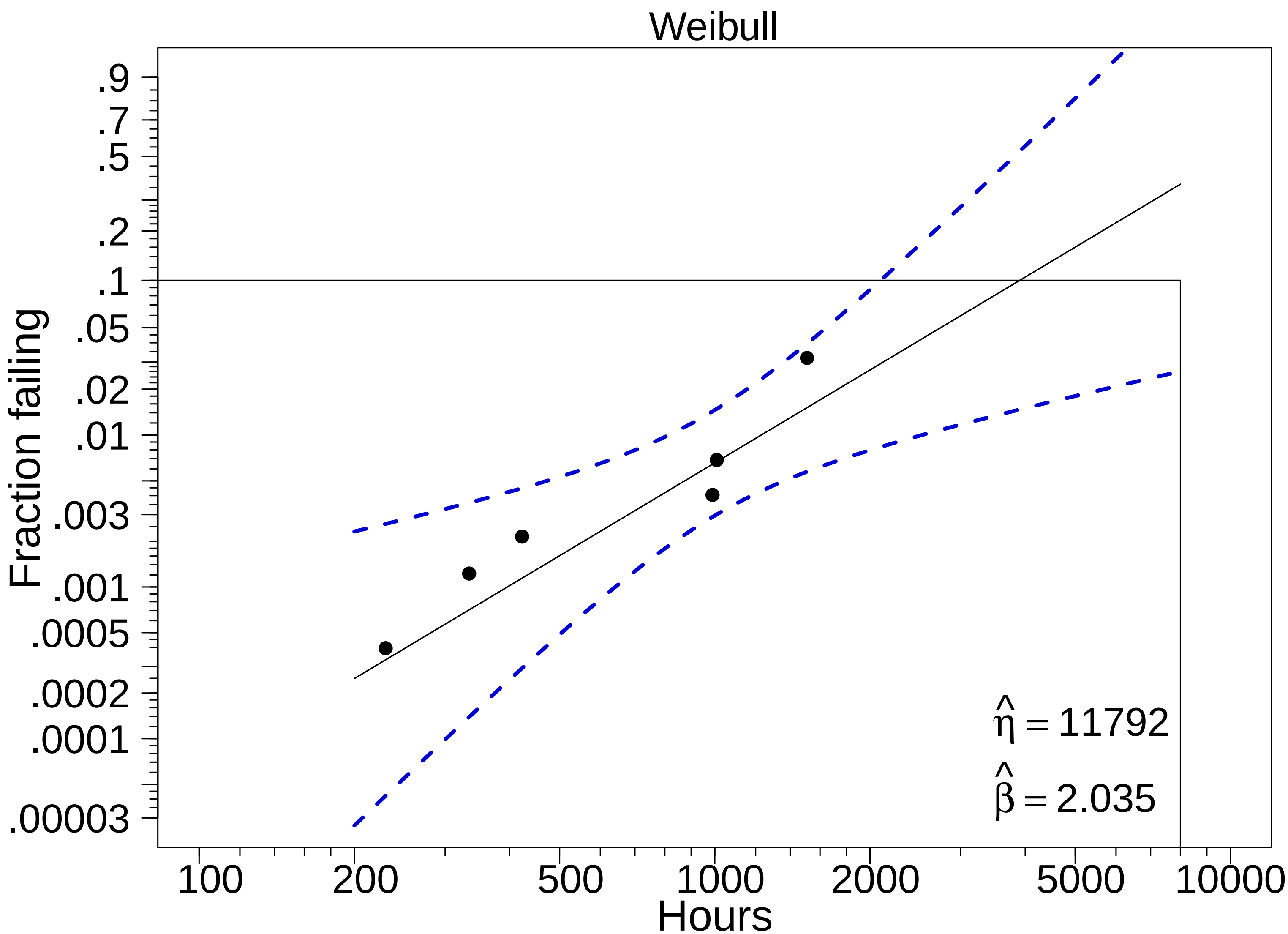
Segment 3

The Simple Accept/Reject Monte Carlo Method Applied to the Bearing Cage Field Failure Data

Bearing Cage Field Failure Data, Revisited

- Data from the Weibull Handbook Abernethy et al. (1983)
- 1703 units had been introduced into the field over time
- Oldest unit at 2220 hours of operation
- 6 units had failed
- Design life specification was B10 = 8000 hours of operation
- Do we have a serious problem? Re-design needed?

Weibull Probability Plot of the Bearing Cage Failure Data and the Maximum Likelihood Estimates of $F(t)$



Example of Eliciting Prior Information: Bearing-Cage Fracture Field Data

- Based on experience with previous products of the same material and knowledge of the failure mechanism, there is strong prior information about the Weibull shape parameter.
- For the Weibull shape parameter **weakly informative** prior distribution, we will use a truncated normal distribution with probability 0.99 between 0.10 and 8.0, denoted by <TNORM>(0.10, 8.0).
- For the Weibull shape parameter **informative** prior distribution, we will use a truncated normal distribution with probability 0.99 between 1.5 and 3.0, denoted by <TNORM>(1.5, 3.0).
- The engineers did not have strong prior information on possible values for the distribution quantiles.
- Uncertainty in the Weibull 0.10 quantile will be described by a lognormal <LNORM>(100, 500,000) distribution (wide range–not very informative).

Bearing Cage Prior Distributions Summary

	Weibull distribution stable parameters	
Analysis	$t_{0.10}$	β
Weakly informative prior	<LNORM>(100, 500,000)	<TNORM>(0.10, 8.0)
Informative prior	<LNORM>(100, 500,000)	<TNORM>(1.5, 3.0)

Evaluating the Joint Posterior Distribution Using Accept/Reject Monte Carlo Method

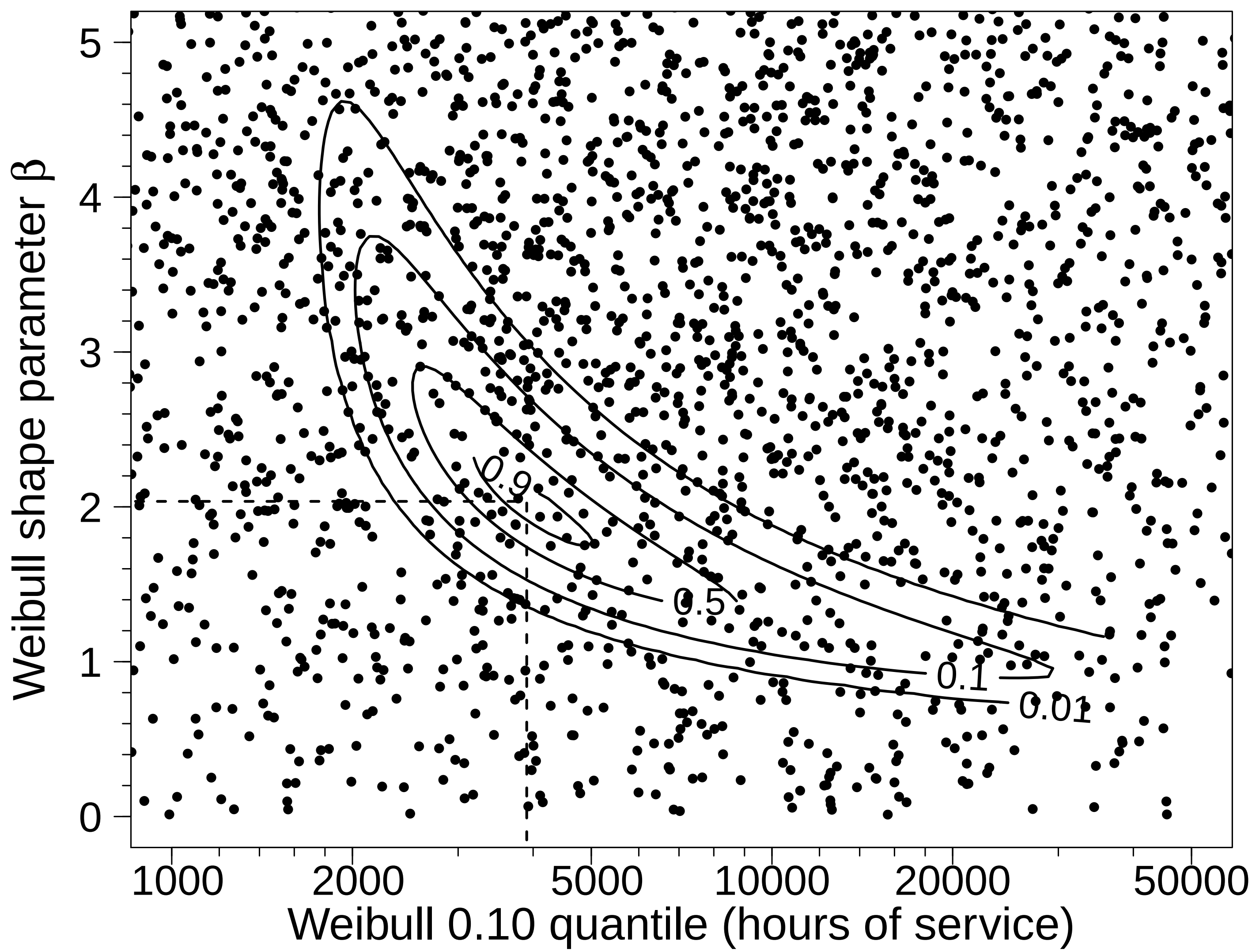
Using simulation, one can obtain sample draws from the joint posterior distribution using only the likelihood and the prior distribution.

- The procedure for a general parameter $\boldsymbol{\theta}$ and joint prior distribution $f(\boldsymbol{\theta})$ is as follows:
- Let $\boldsymbol{\theta}_i$, $i = 1, \ldots$ be sample draws from $f(\boldsymbol{\theta})$.
- Then $\boldsymbol{\theta}_i$, prior draw i, is accepted with probability $R(\boldsymbol{\theta}_i)$, the relative likelihood at the point. Thus, if U_i is a random observation from a uniform UNIFORM$[0, 1]$, $\boldsymbol{\theta}_i$ is accepted if

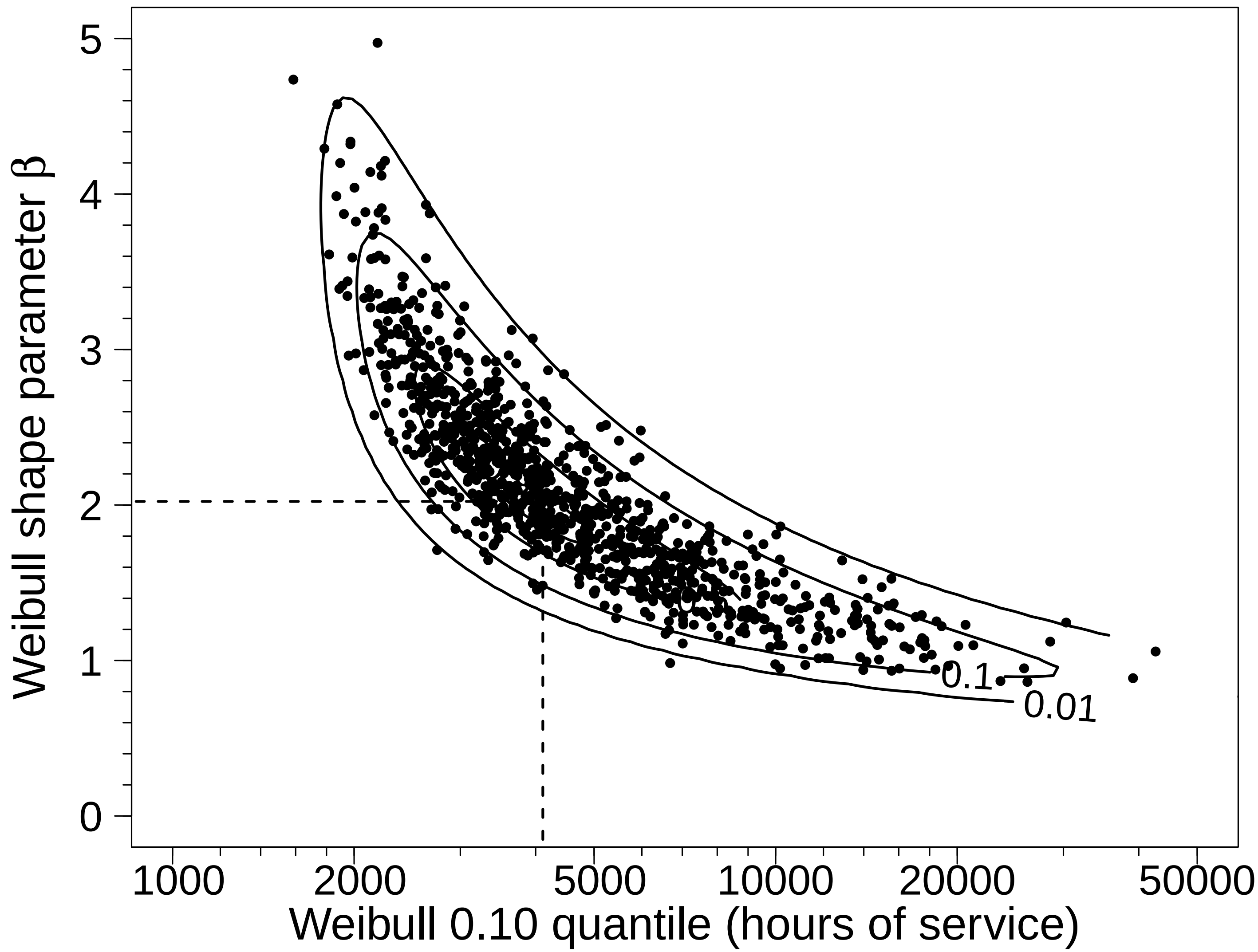
$$U_i \leq R(\boldsymbol{\theta}_i).$$

- It can be shown that the retained observations, say $\boldsymbol{\theta}_1^*, \ldots \boldsymbol{\theta}_{B^*}^*$ are draws from the joint posterior distribution $f(\boldsymbol{\theta}|\text{DATA})$.

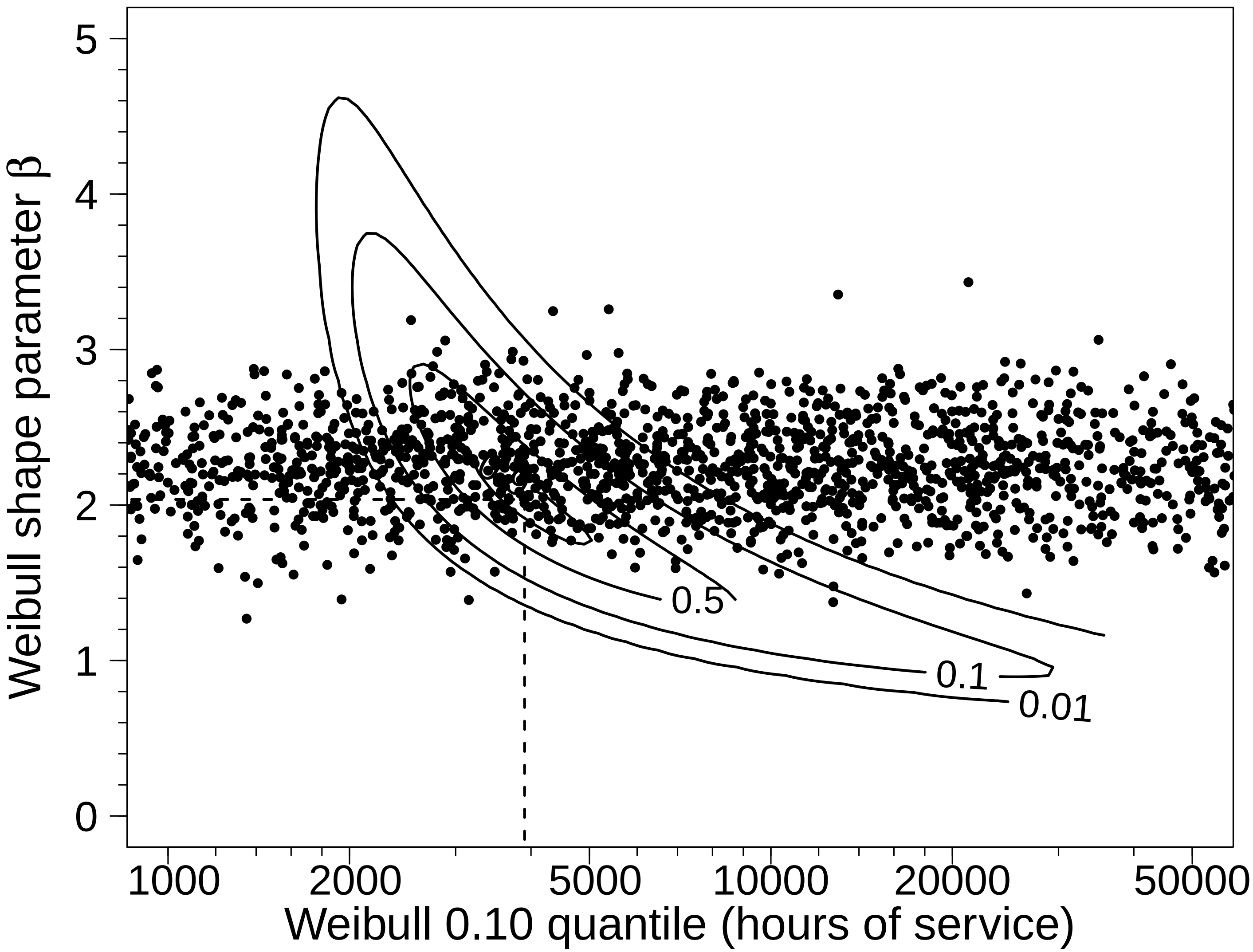
Likelihood Contours and Draws from the Weakly Informative Joint Prior Distribution of $t_{0.10}$ and β



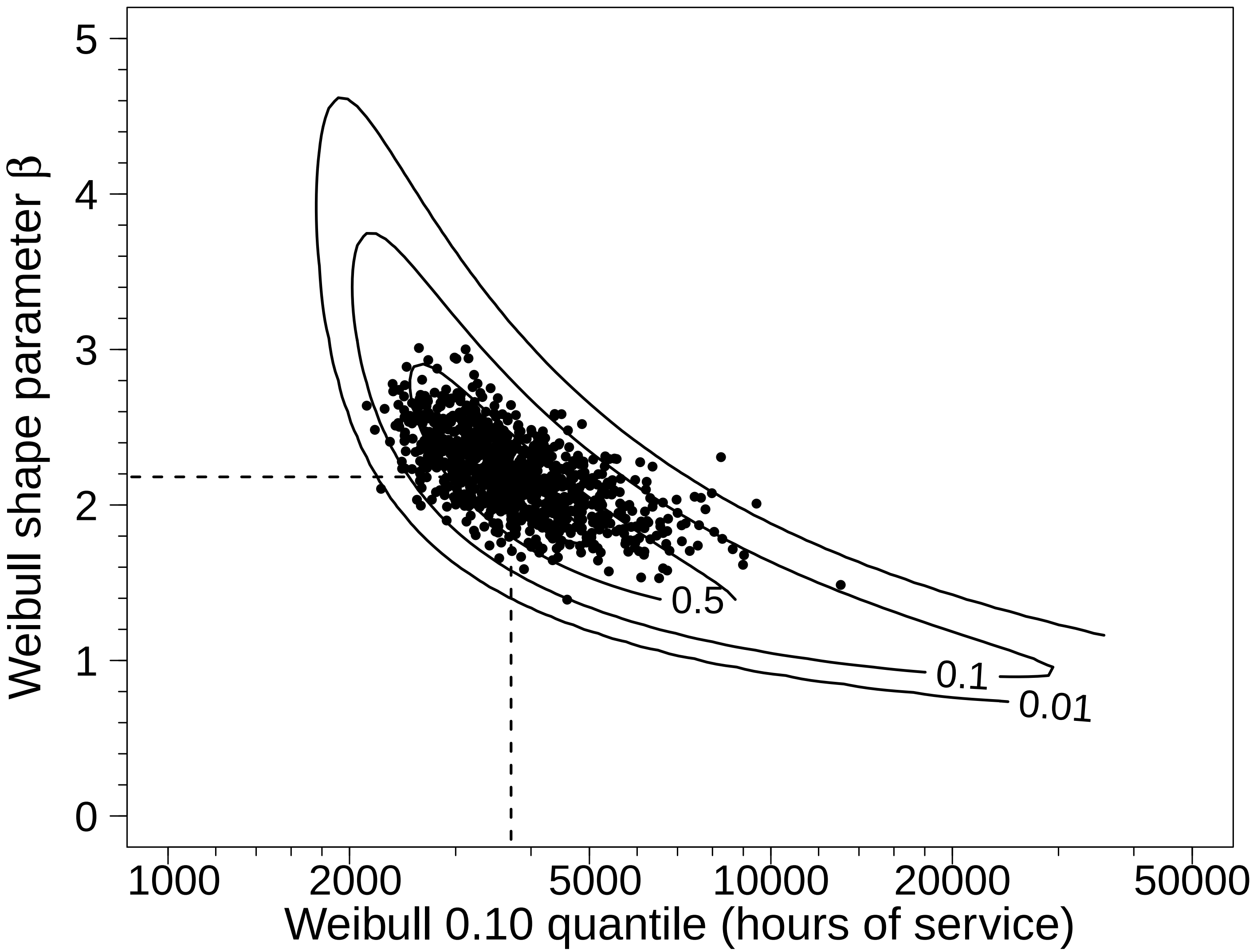
Likelihood Contours and Draws from the <u>Joint Posterior</u> of $t_{0.10}$ and β Based on the Weakly Informative a Joint Prior



Likelihood Contours and Draws from the Informative <u>Joint Prior</u> Distribution of $t_{0.10}$ and β



Likelihood Contours and Draws from the <u>Joint Posterior</u> of $t_{0.10}$ and β Based on the Informative a Joint Prior



Comments on Computing Draws from the Joint Posterior Distribution Using the Simple Accept/Reject Monte Carlo Method

The probability that a draw from the prior distribution is accepted is

$$\int f(\boldsymbol{\theta})R(\boldsymbol{\theta})d\boldsymbol{\theta}.$$

- When the prior distribution and the data do not agree well, the acceptance probability will be small and more draws from the joint prior will be required, **increasing computational effort**.

- The easiest way to evaluate this probability is to run the algorithm.

Chapter 10

Segment 4

Bayesian Point Estimation and Credible Interval Methods

Bayesian Point Estimation for Parameters and Functions of Parameters

- Inferences on a scalar function of the parameters $g(\boldsymbol{\theta})$ are obtained by using the marginal posterior distribution of the functions of the parameters of interest, $f[g(\boldsymbol{\theta})|\text{DATA}]$.

- For point estimation, one could use the posterior **mean** of $g(\boldsymbol{\theta})$

$$\widehat{g}(\boldsymbol{\theta}) = \text{E}[g(\boldsymbol{\theta})|\text{DATA}] = \int g(\boldsymbol{\theta}) f(\boldsymbol{\theta}|\text{DATA}) d\boldsymbol{\theta}.$$

- A better alternative is to use the posterior **median**.

Bayesian Inference Based on Draws from the Joint Posterior Distribution

$$\begin{pmatrix} \boldsymbol{\theta}_1^* \\ \boldsymbol{\theta}_2^* \\ \vdots \\ \boldsymbol{\theta}_B^* \end{pmatrix} = \begin{pmatrix} \theta_{11}^* & \theta_{12}^* & \cdots & \theta_{1q}^* \\ \theta_{21}^* & \theta_{22}^* & \cdots & \theta_{2q}^* \\ \vdots & \vdots & \vdots & \vdots \\ \theta_{B1}^* & \theta_{B2}^* & \cdots & \theta_{Bq}^* \end{pmatrix}.$$

Bayesian Inference Based on Draws from the Joint Posterior Distribution and Marginal Distributions for Other Quantities of Interest

Often our interest is on particular functions of the parameters such as $g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), \ldots$. Then we can compute, directly, additional columns in the matrix of draws, corresponding to draws from the marginal posterior distributions of those particular functions.

$$\begin{pmatrix} \theta^*_{11} & \theta^*_{12} & \cdots & \theta^*_{1q} & g_1(\boldsymbol{\theta}^*_1) & g_2(\boldsymbol{\theta}^*_1) \cdots \\ \theta^*_{21} & \theta^*_{22} & \cdots & \theta^*_{2q} & g_1(\boldsymbol{\theta}^*_2) & g_2(\boldsymbol{\theta}^*_2) \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \theta^*_{B1} & \theta^*_{B2} & \cdots & \theta^*_{Bq} & g_1(\boldsymbol{\theta}^*_B) & g_2(\boldsymbol{\theta}^*_B) \cdots \end{pmatrix}.$$

Using Sample Draws from the Joint Posterior Distribution to Compute Point Estimates and Credible Intervals

- For inferences on a quantile t_p, draws from the marginal posterior distribution $f(t_p|\text{DATA})$ are obtained by computing the Weibull p quantile (Chapter 4), row-wise

$$t_p^* = \eta_i^*[-\log(1-p)]^{1/\beta_i^*}, \quad 0 < 0 < 1,$$

for $i = 1, \ldots, B$.

- For inferences on a failure probability $F(t_e)$ at a time t_e, draws from the marginal posterior distribution $f[F(t_e)|\text{DATA}]$ are obtained by computing the Weibull cdf (Chapter 4), row-wise

$$F^*(t_e) = 1 - \exp\left[-\left(\frac{t_e}{\eta_i^*}\right)^{\beta_i^*}\right], \quad t > 0,$$

for $i = 1, \ldots, B$.

Example Sample Draws from the Joint Posterior Distribution of the Weibull Distribution Parameters $t_{0.10}$ and β and the additional quantities of interest η, $F(5000)$, and $F(8000)$

$t^*_{0.10}$	β^*	η^*	$F^*(5000)$	$F^*(8000)$
4280	1.84	14500	0.131	0.284
5460	2.16	15500	0.083	0.214
4570	2.24	12500	0.121	0.309
3330	2.21	9240	0.227	0.517
3350	2.22	9230	0.226	0.517
3560	2.18	10000	0.198	0.458
2990	2.16	8460	0.274	0.588
3220	2.38	8280	0.259	0.602
5710	1.82	19700	0.079	0.177
4950	1.79	17400	0.102	0.221

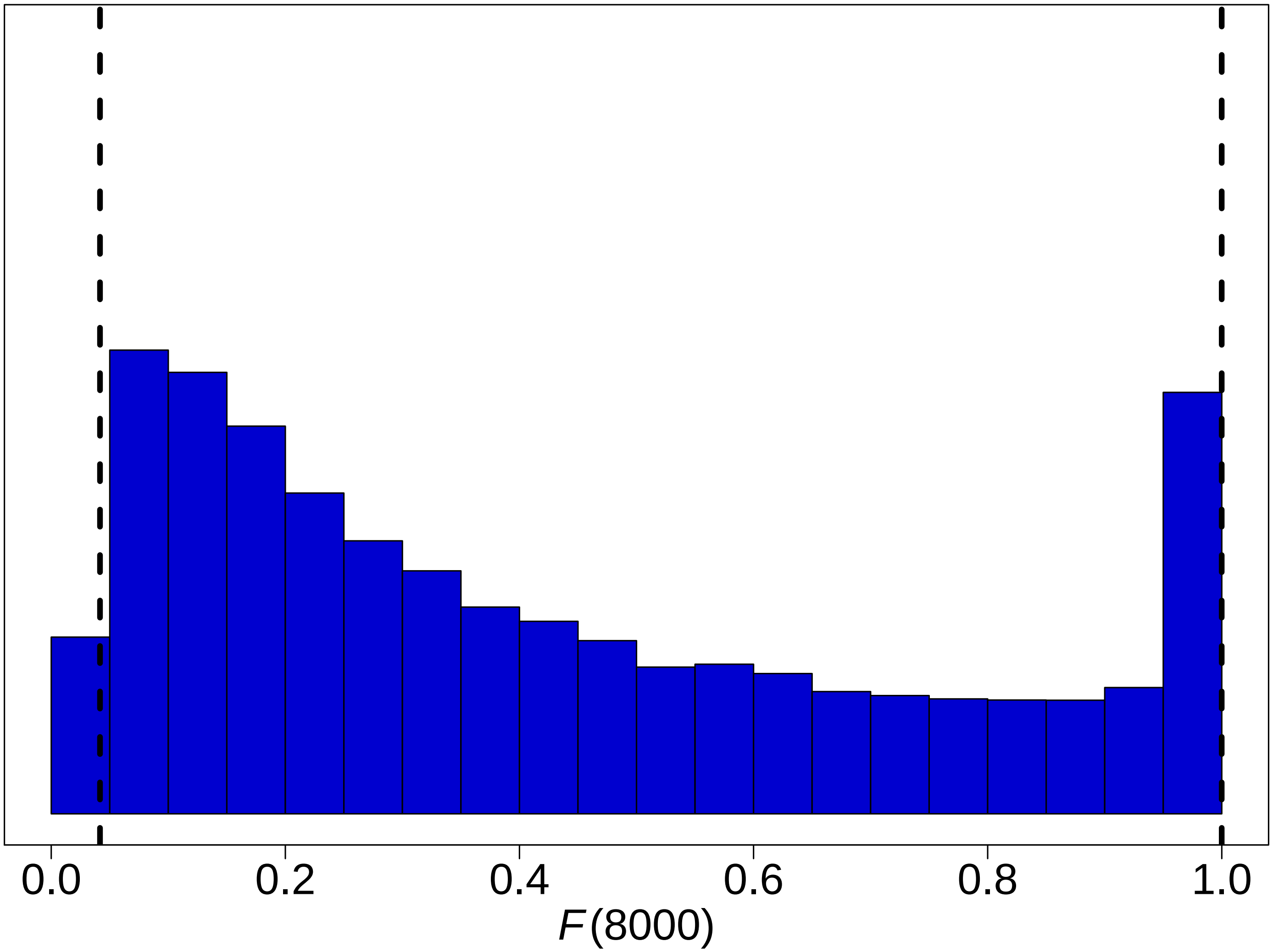
Bayesian Inference Using Sample Draws from Marginal Posterior Distributions

- Bayesian point estimates and credible intervals are **easy to compute** given draws from the marginal posterior distributions of the quantities of interest.

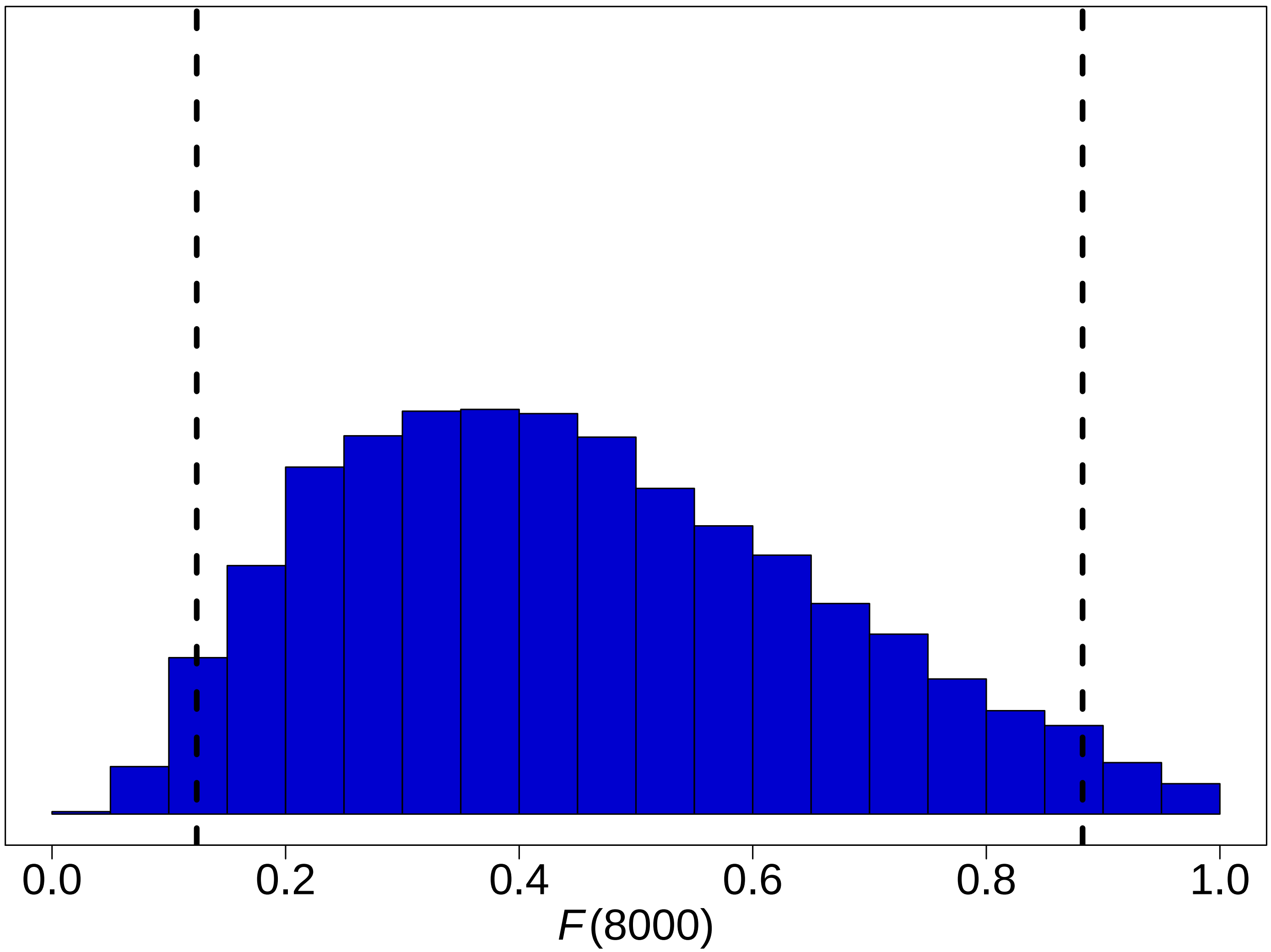
- The following `R` commands, applied to the matrix of 100,000 sample draws, similar to those shown in the previous slide, give:

```
> quantile(drawsBearingCageWeaklyInformative[,"F5000"], probs=c(0.025, 0.975))
  2.5%  97.5%
0.0243 0.8446
> quantile(drawsBearingCageInformative[,"F5000"], probs=c(0.025, 0.975))
  2.5%  97.5%
0.0556 0.4668
> quantile(drawsBearingCageWeaklyInformative[,"F8000"], probs=c(0.025, 0.975))
  2.5%  97.5%
0.0417 0.9999
> quantile(drawsBearingCageInformative[,"F8000"], probs=c(0.025, 0.975))
 2.5% 97.5%
0.124 0.882
```

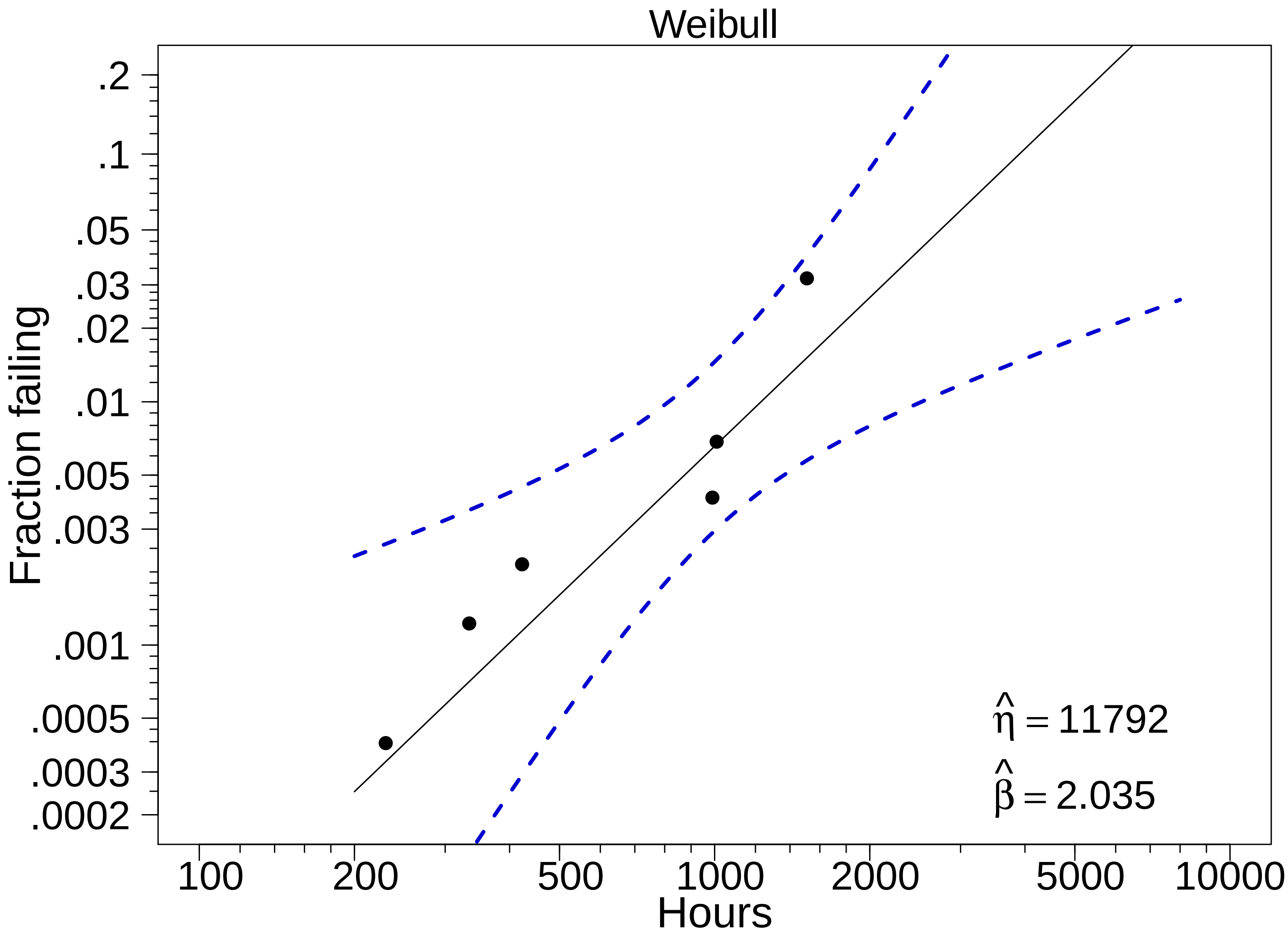
Marginal Posterior Distributions and Credible Intervals for Bearing Cage Failure Probabilities Weakly Informative Prior Distributions



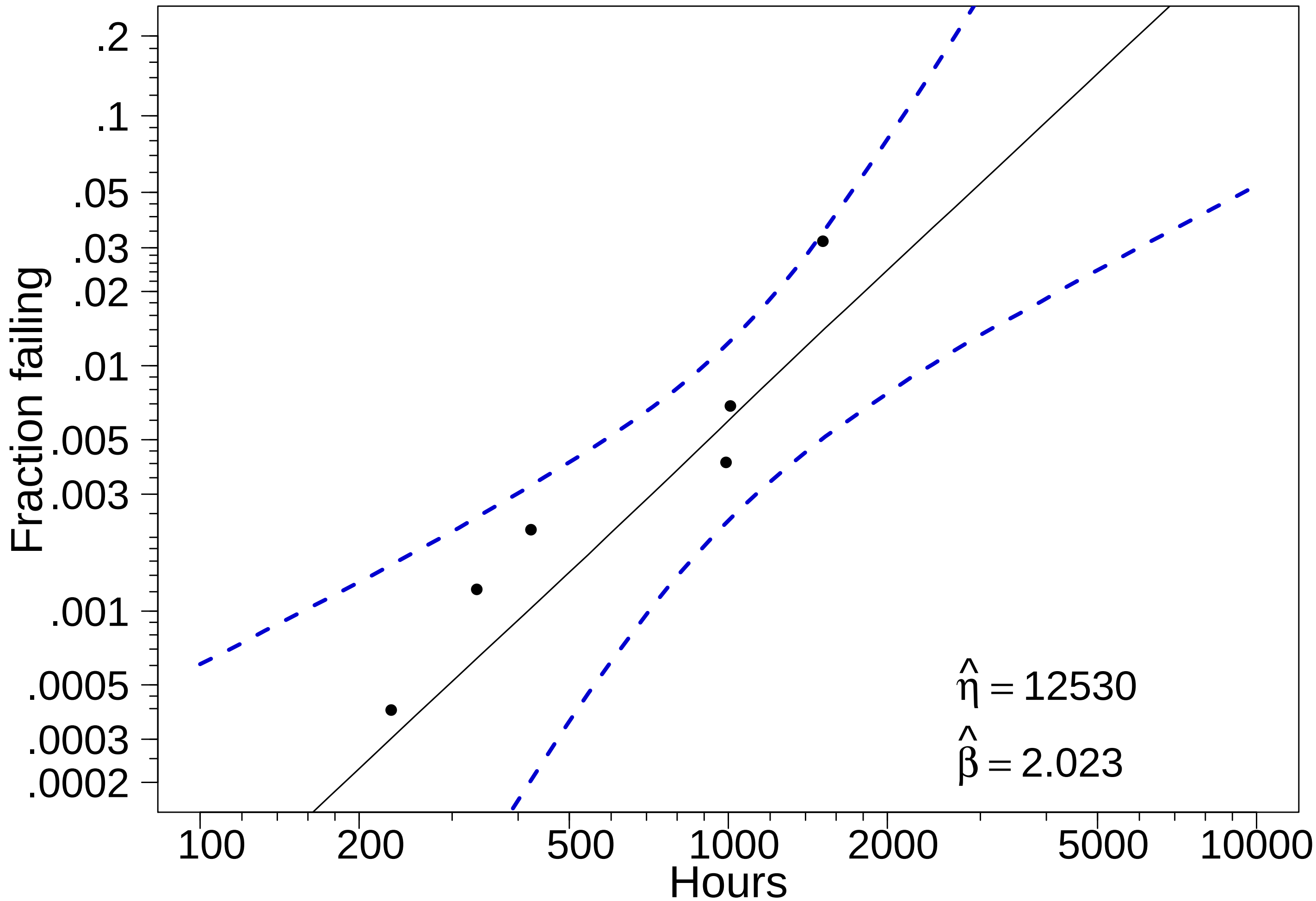
Marginal Posterior Distributions and Credible Intervals for Bearing Cage Failure Probabilities Informative Prior Distributions



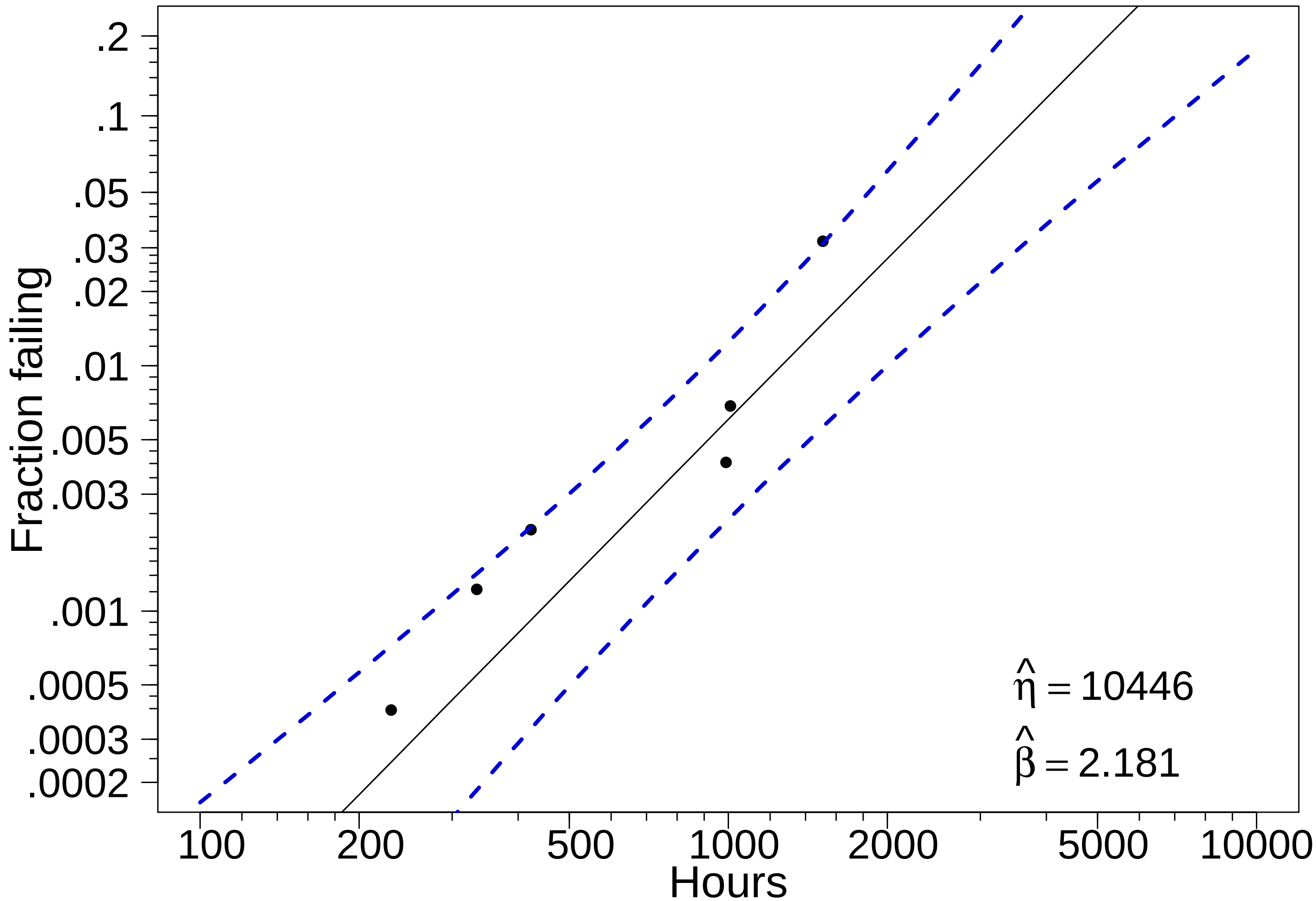
Weibull Probability Plot of the Bearing Cage Failure Data and the Maximum Likelihood Estimates of $F(t)$ with 95% Confidence Intervals



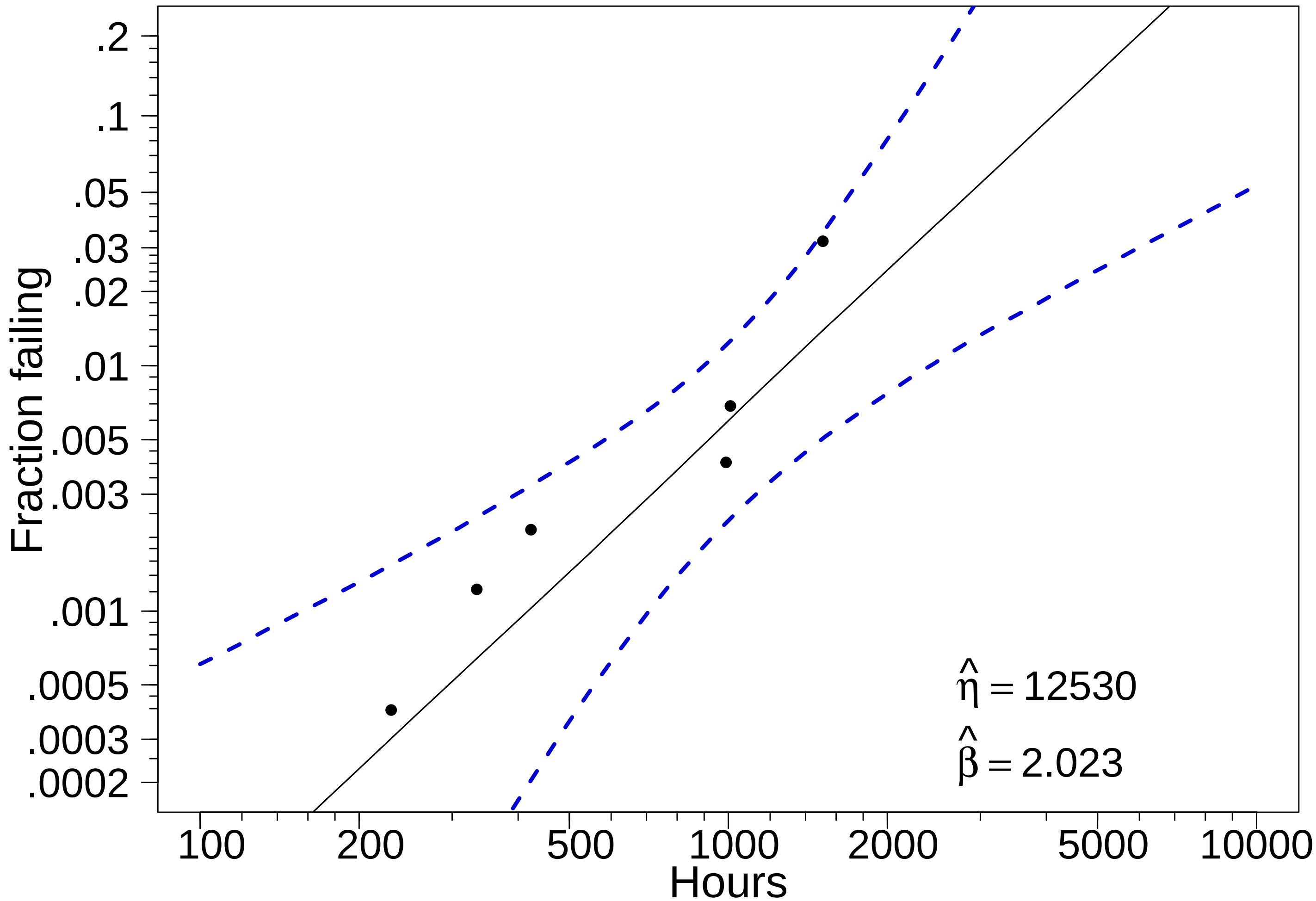
Weibull Probability Plot of the Bearing Cage Failure Data Bayesian Estimates of $F(t)$ Based on the Weakly Informative Prior Distribution



Weibull Probability Plots of the Bearing Cage Failure Data and the Bayesian Estimates of $F(t)$ Based on the Informative Prior Distribution



Weibull Probability Plot of the Bearing Cage Failure Data and the Bayesian Estimates of $F(t)$ Based on the Weakly Informative Prior Distribution



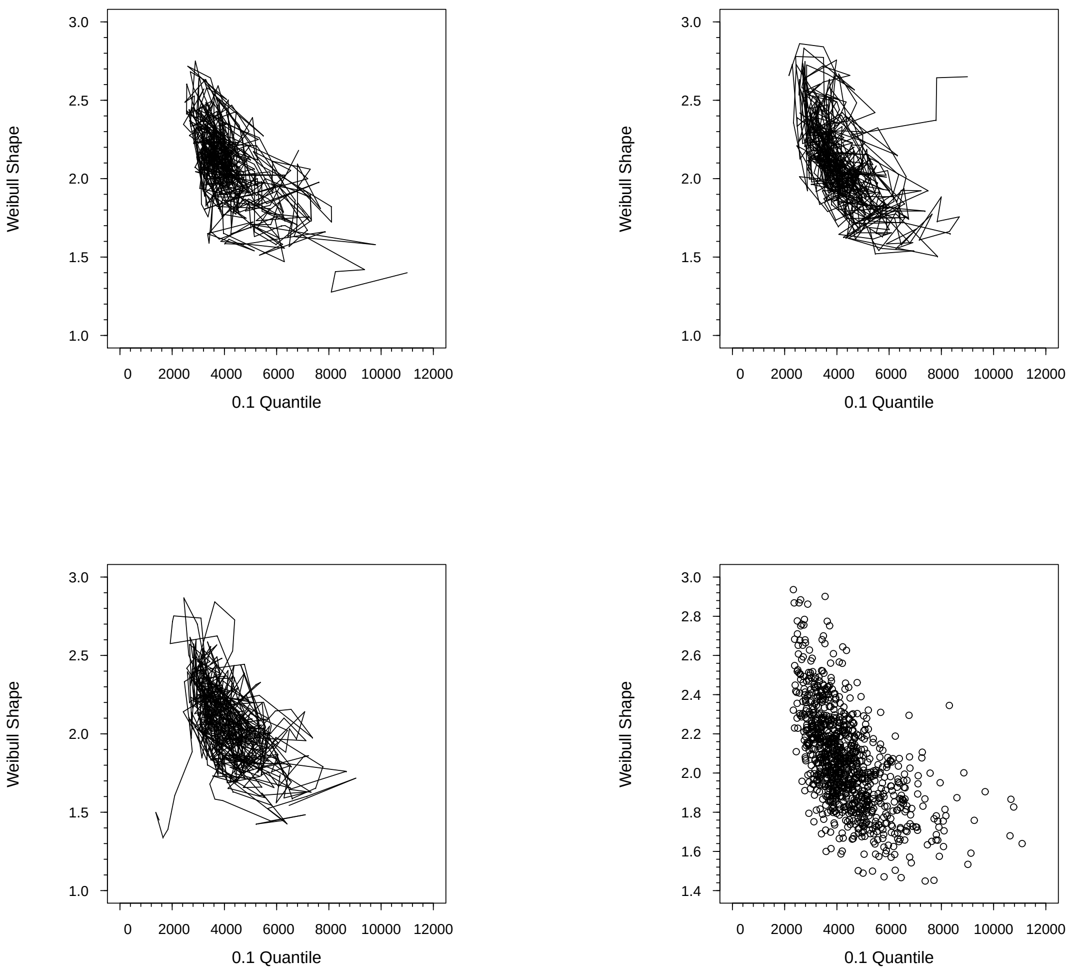
Simple Accept/Reject Monte Carlo Method for Obtaining Draws from the Joint Posterior Distribution

- Is easy to implement.
- Results in independent draws (no autocorrelation).
- Easy to assure that one is covering the entire posterior distribution space.
- When used with weakly informative prior distributions, especially with multiple parameters, the acceptance probability tends to be very small (e.g. 0.001) and obtaining an adequate number of draws will take a considerable amount of computational effort.
- Will work effectively only if a small number of parameters is being estimated (e.g., fewer than 4 or 5).

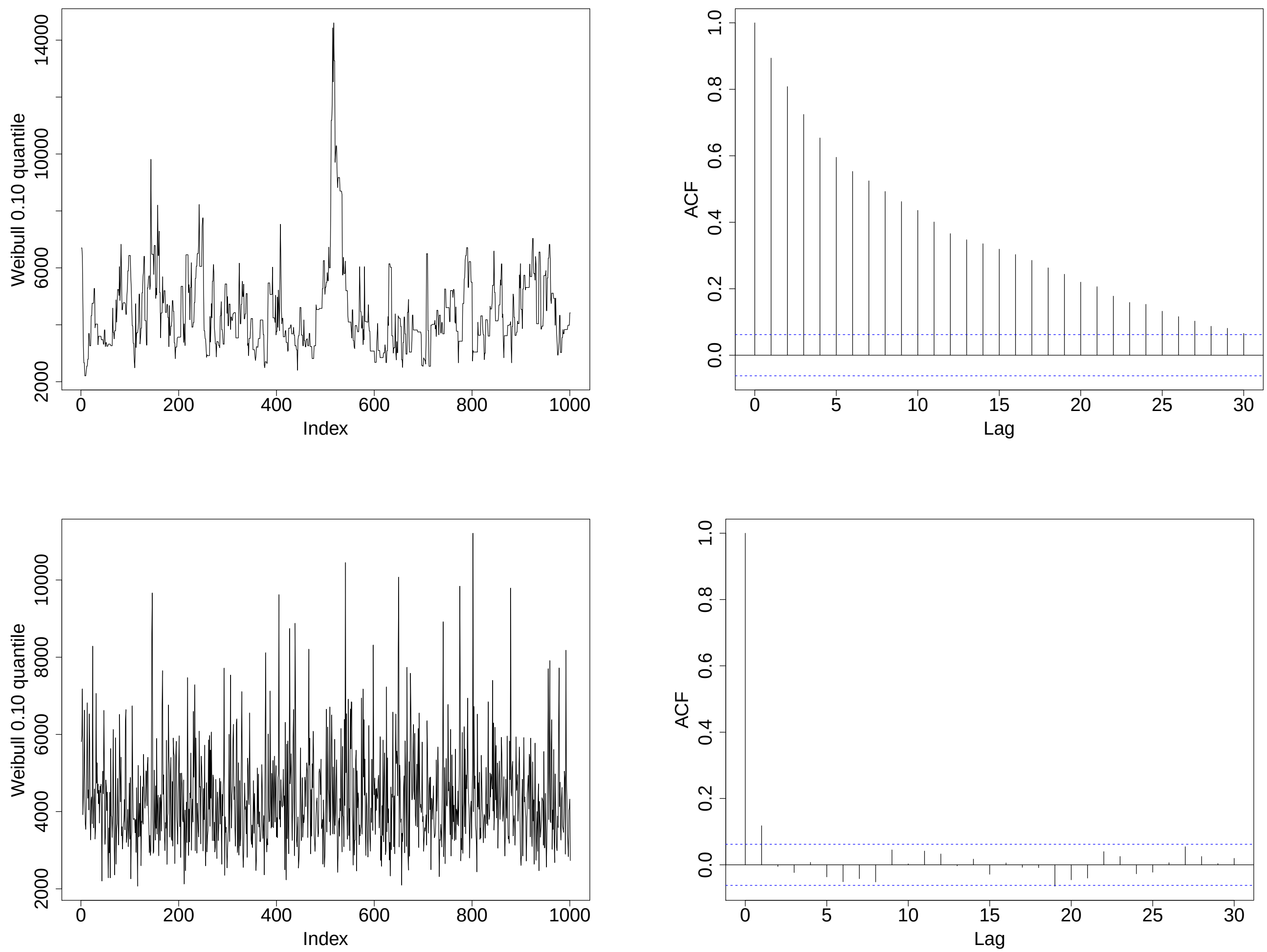
Markov Chain Monte Carlo (MCMC) Methods for Obtaining Draws from the Joint Posterior Distribution

- Can handle complicated hierarchical models with a large number of parameters.
- Many different MCMC methods to choose from (e.g., Gibbs sampling and the Metropolis-Hastings algorithms).
- MCMC methods tend to be more complicated to implement and tune.
- Tradeoffs between complexity and performance.
- Draws have autocorrelation implying that more draws are generally needed.
- Must use MCMC diagnostics to assure that the draws are representative of the target joint posterior distribution.

Sample Paths from a Markov Chain Generating Samples From a Joint Posterior Distribution for the Bearing Cage Data and a Weibull Distribution



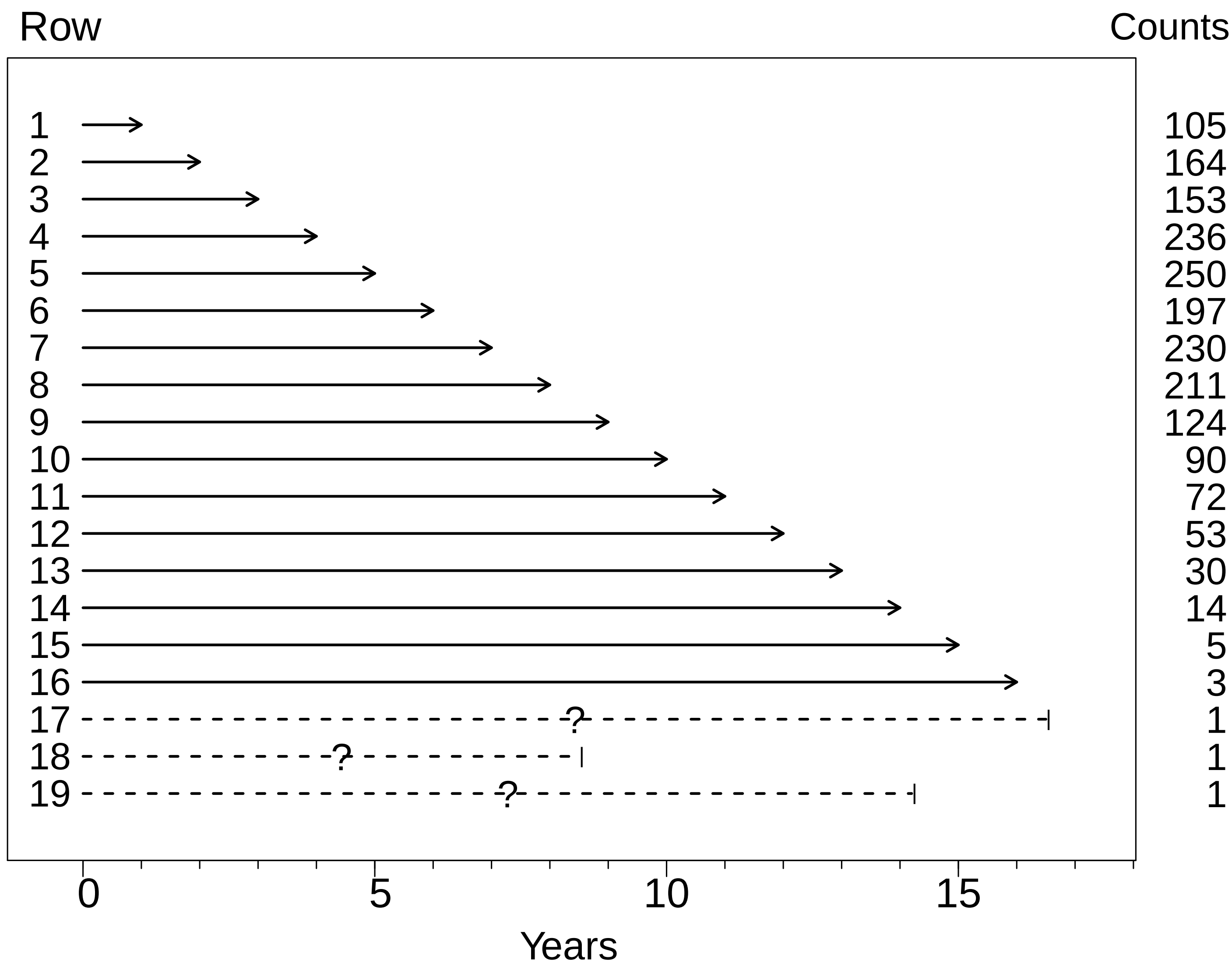
Trace Plots (on the left) and ACF Plots (on the right) Comparing Unthinned Draws (on the top) and Thinned Draws (on the bottom) for the Bearing Cage 0.10 Quantile



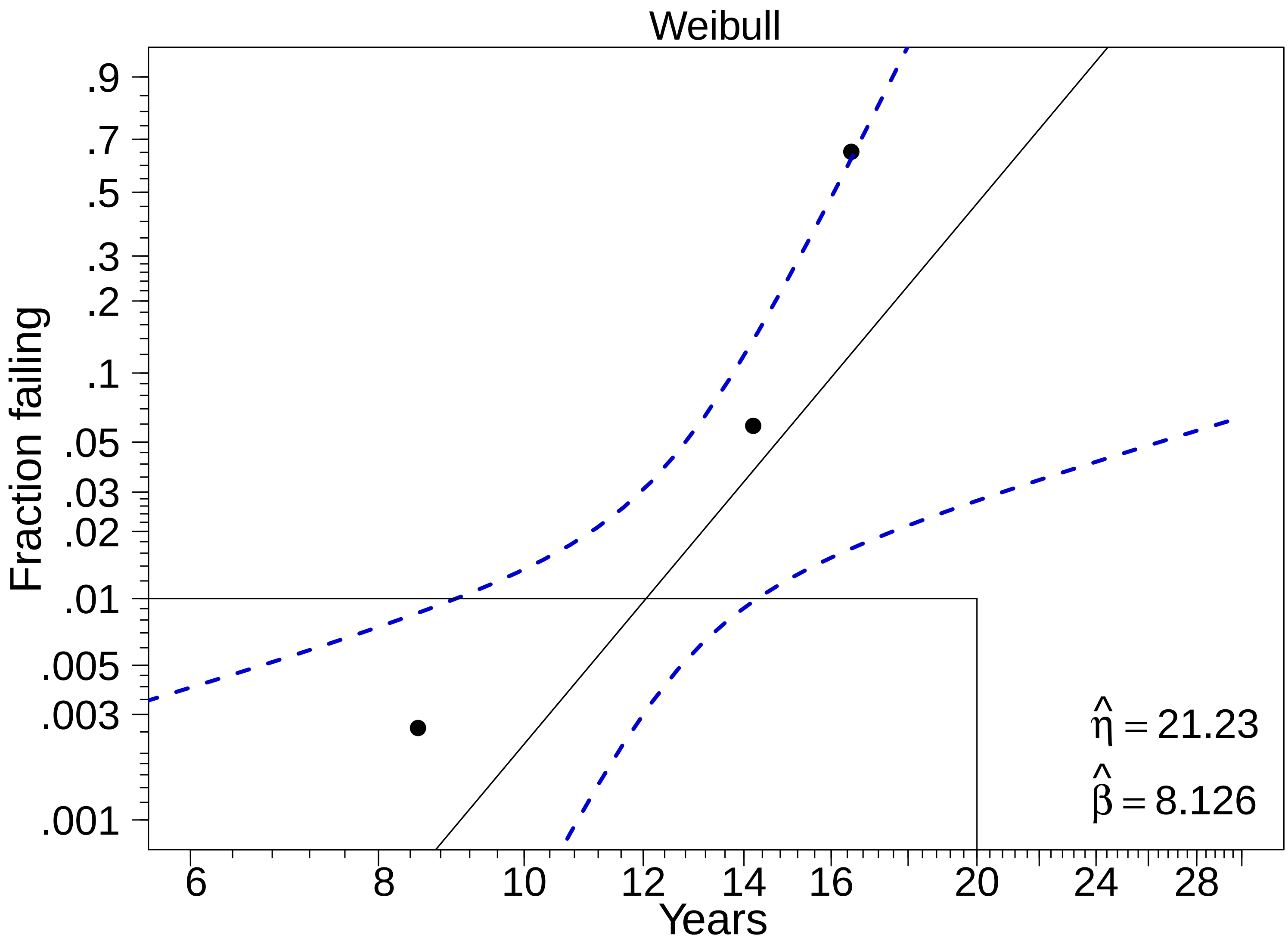
Rocket Motor Field Data Analysis

- Data from Olwell and Sorell (2001).
- Rocket motor is one of five critical missile components.
- Approximately 20,000 missiles in inventory.
- 1,940 firings over the life of the missile over 18 years; catastrophic motor failures for **3 older missiles**.
- Concern about a possible wearout failure mode and the health of the stockpile.

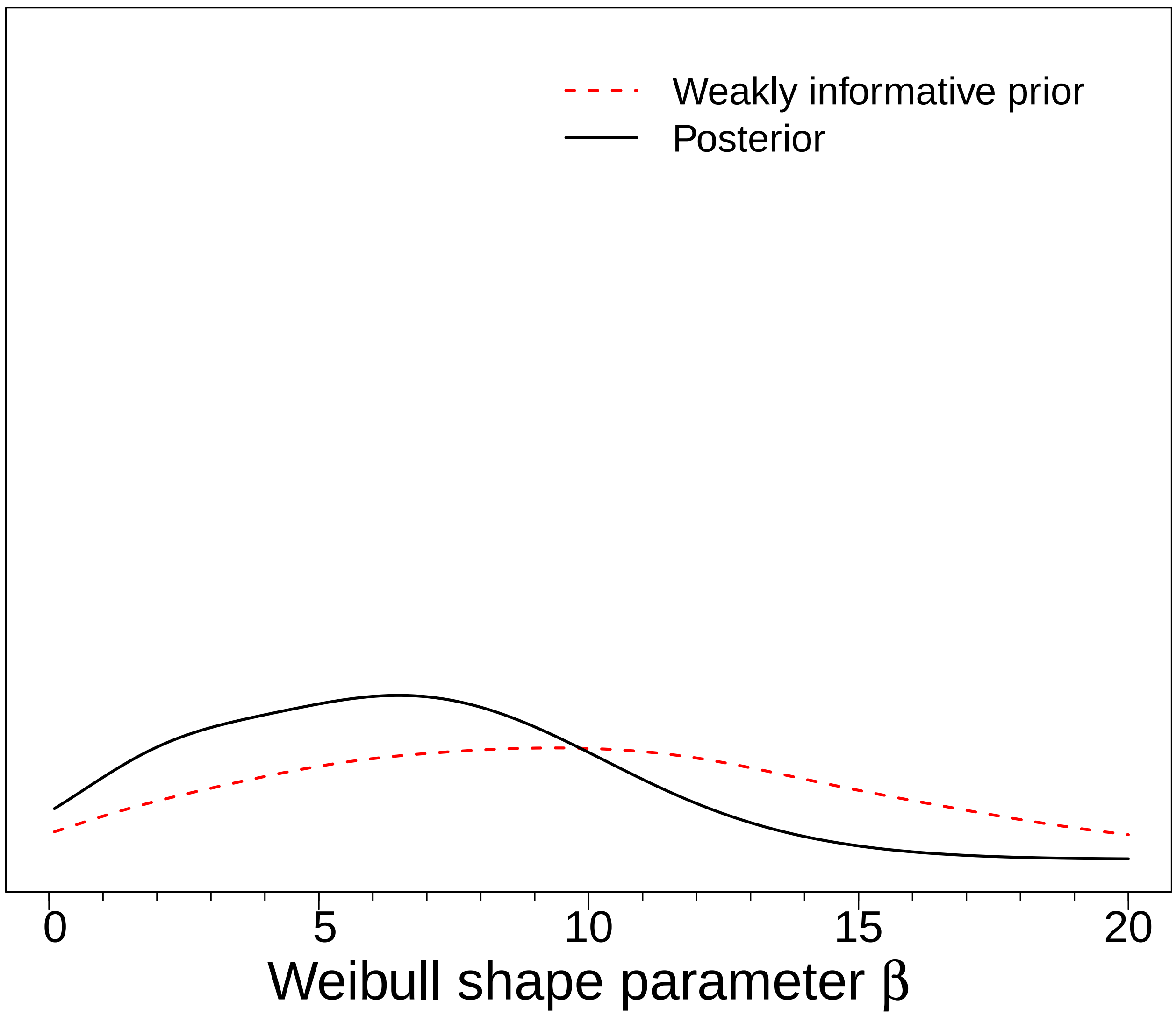
Event Plot of the Rocket Motor Field Data



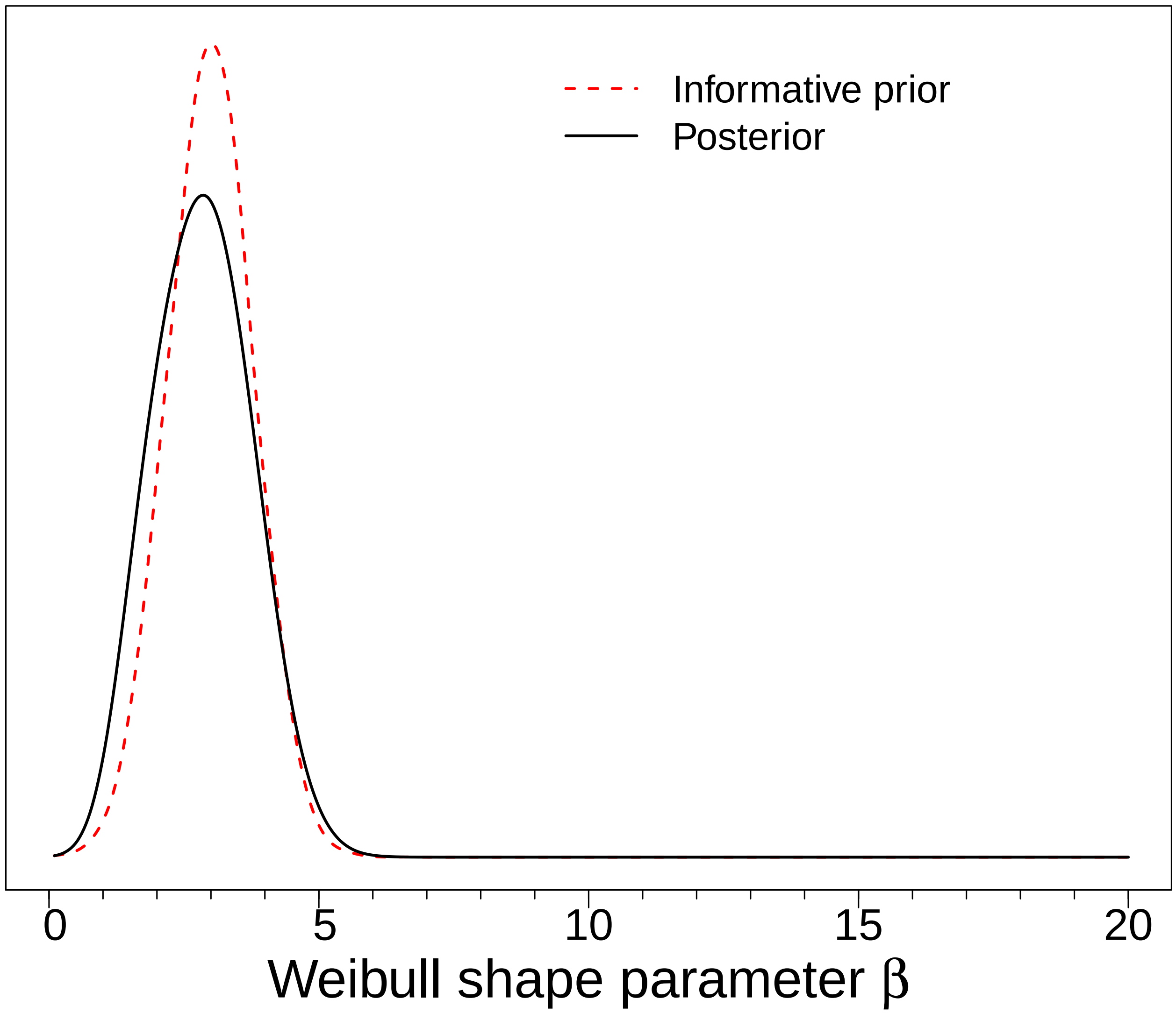
Weibull Probability Plot of the Rocket Motor Current-Status Data



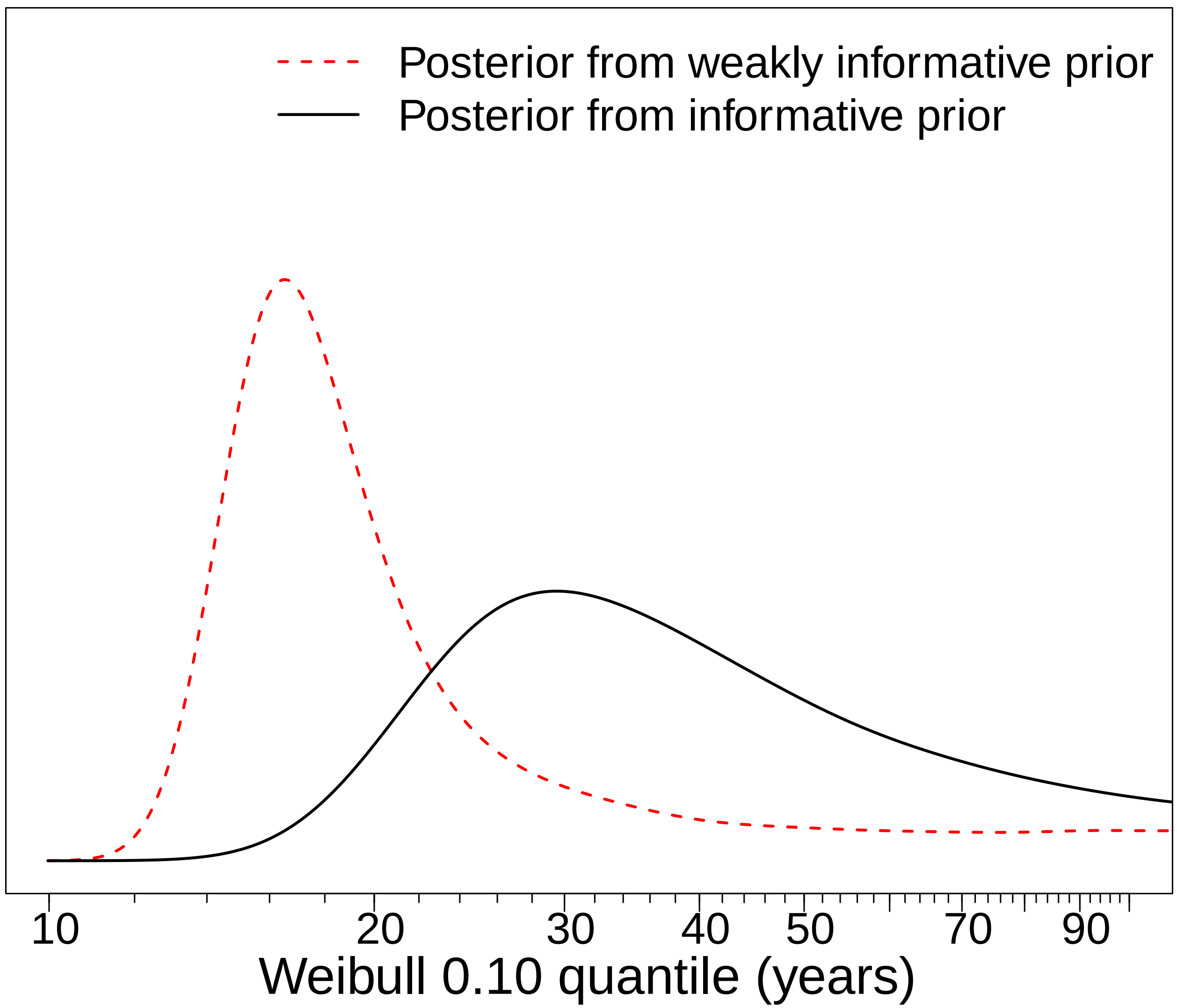
Posterior and Prior Distributions for the Rocket Motor Weibull Distribution Shape Parameter β with a Weakly Informative Prior Distribution



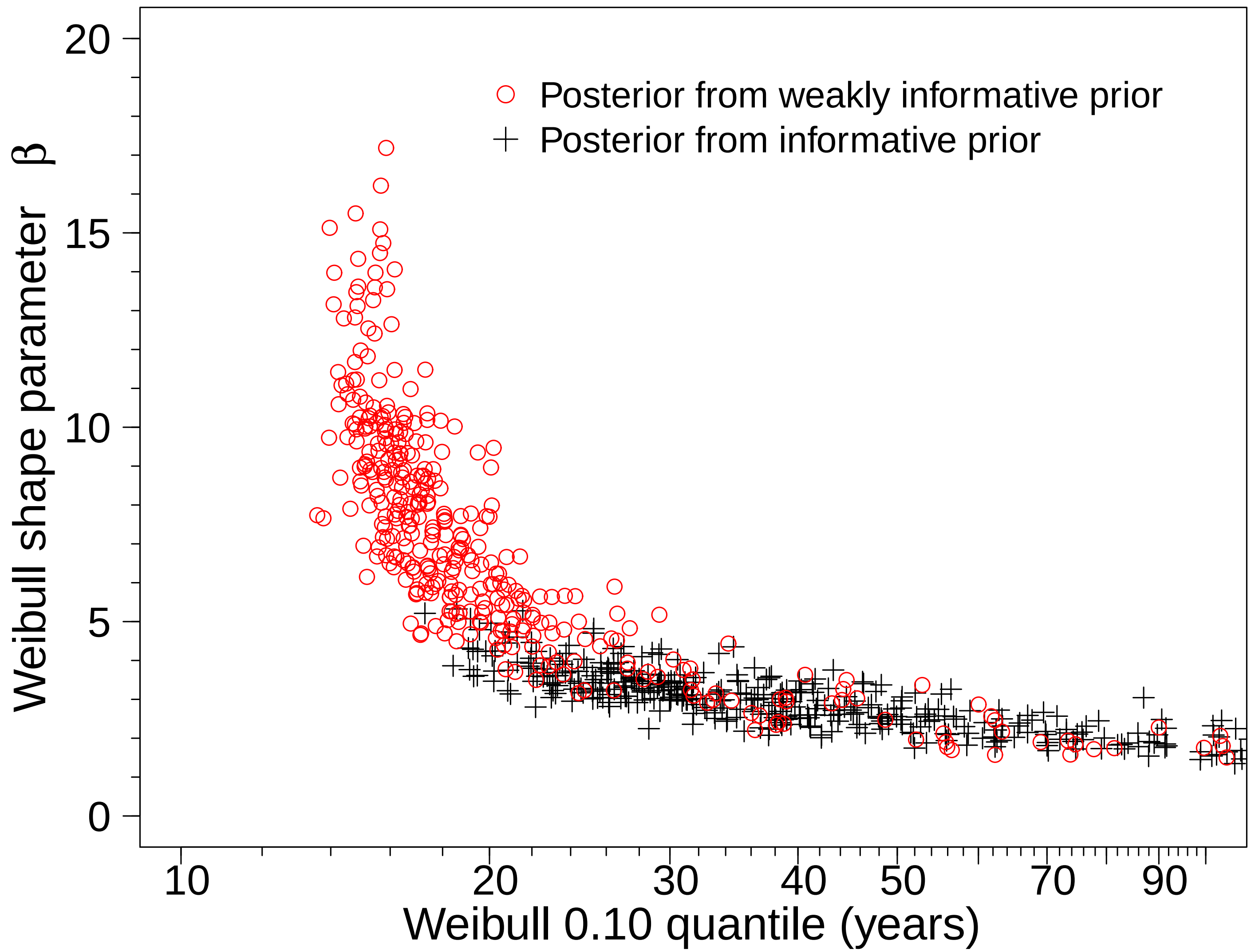
Posterior and Prior Distributions for the Rocket Motor Weibull Distribution Shape Parameter β with an Informative Prior Distribution



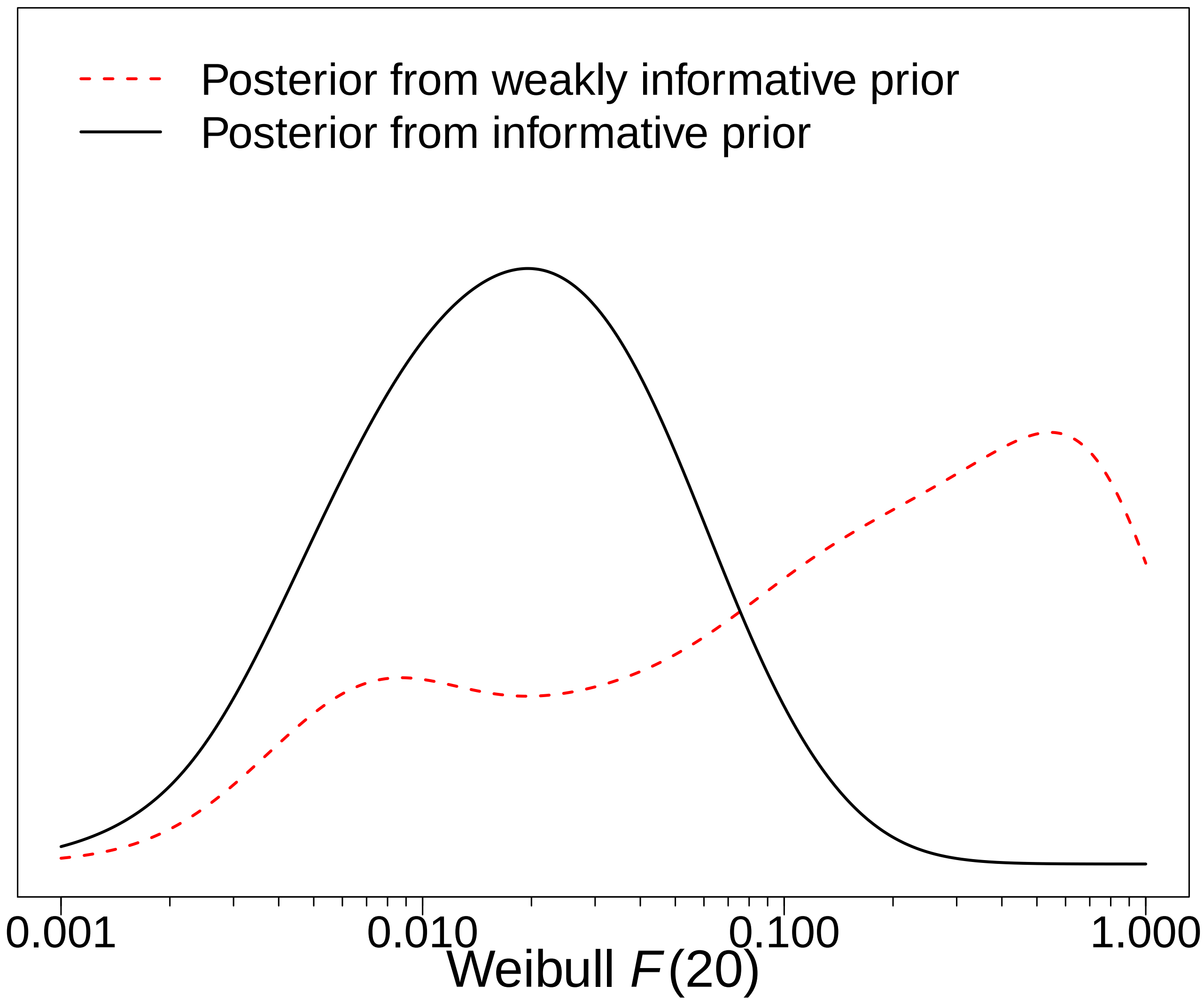
Marginal Posterior Distributions for the Rocket Motor Failure-Time Distribution 0.10 Quantile Using Weakly Informative and Informative Priors



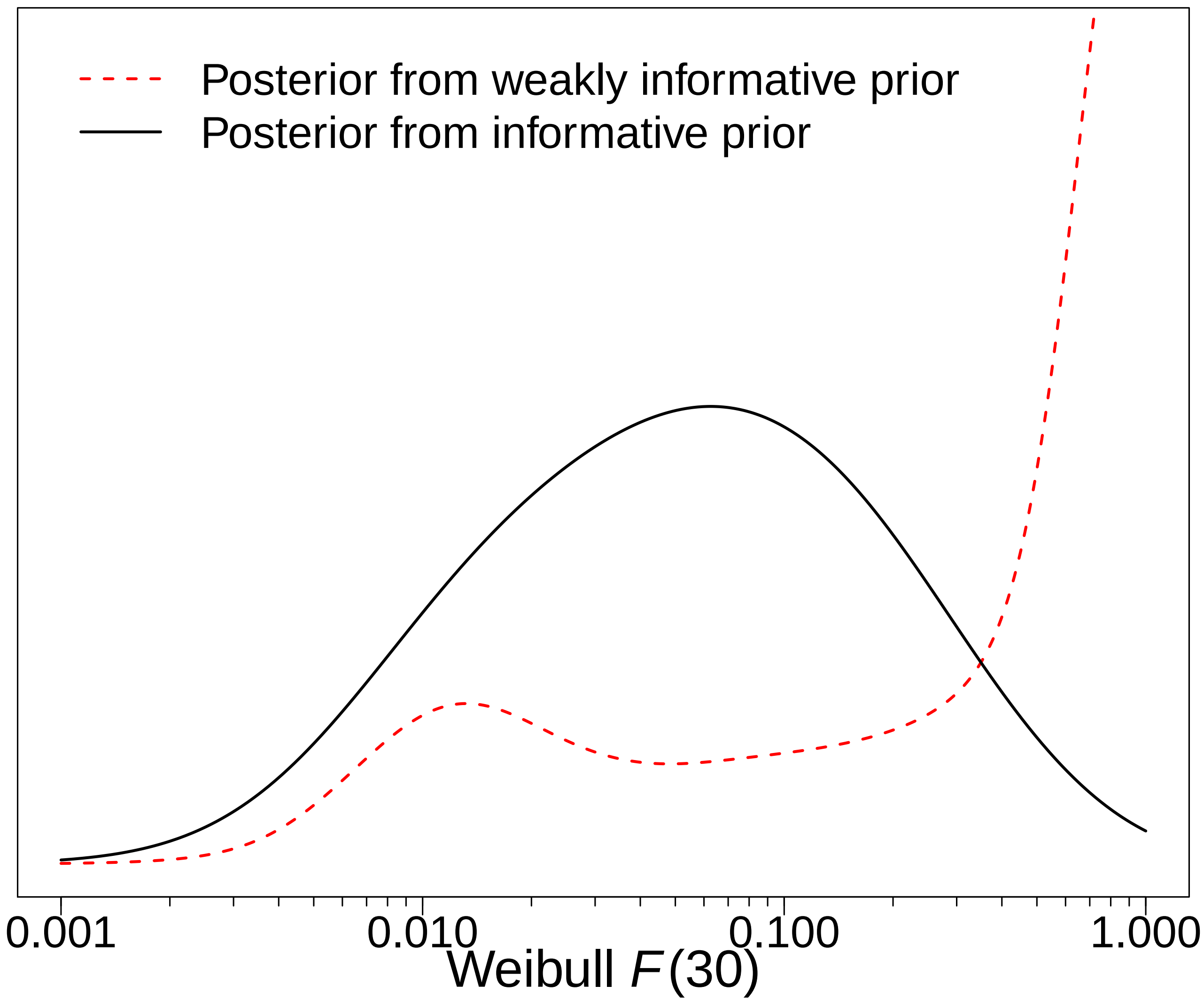
Joint Posterior Distributions Based on Weakly Informative and Informative Prior Distributions



Marginal Posterior Distributions of the Rocket Motor $F(20)$ Using Weakly Informative and Informative Priors



Marginal Posterior Distributions of the Rocket Motor $F(30)$ Using Weakly Informative and Informative Priors



Lessons Learned from the Rocket Motor Example

- Even with very limited data, useful inferences on reliability may be possible.

- It is important to use external (prior) information, when available.

- The shapes of the likelihood and posterior distribution are important in determining posterior inferences.

Whose Prior Distribution Should We Use?

- It is possible to specify a weakly informative prior distribution for parameters that are unknown.

- With limited data, the choice of a prior distribution (even if weakly informative) can be highly influential.

- In some applications, solid prior information, based on a combination of physics of failure and previous empirical experience, is available.

- In general, whoever is assuming the **risk** associated with decision making should be allowed to choose the prior distribution.

- If different people or groups have difference risk functions, we have a difficult conflict.

Final Cautions on the Use of Prior Information

- In many applications, engineers really have useful, indisputable prior information. In such cases, the information should be integrated into the analysis.

- Beware of **wishful thinking** masquerading as prior information. The potential for generating seriously misleading conclusions is high.

- As with other inferential methods, when using Bayesian methods, it is important to do sensitivity analyses with respect to uncertain inputs to one's model (including the inputted prior information).

References

Abernethy, R. B., J. E. Breneman, C. H. Medlin, and G. L. Reinman (1983). *Weibull Analysis Handbook*. Air Force Wright Aeronautical Laboratories Technical Report AFWAL-TR-83-2079. Available from: http://apps.dtic.mil/dtic/tr/fulltext/u2/a143100.pdf. []

Meeker, W. Q., L. A. Escobar, and F. G. Pascual (2021). *Statistical Methods for Reliability Data* (Second Edition). Wiley. [1]

Olwell, D. H. and A. A. Sorell (2001). Warranty calculations for missiles with only current-status data, using Bayesian methods. In *2001 Proceedings of the Annual Reliability and Maintainability Symposium (RAMS)*, 133–138. IEEE. []