

Chapter 14
Introduction to the Use of Bayesian Methods for
Reliability Data

William Q. Meeker and Luis A. Escobar
Iowa State University and Louisiana State University

Copyright 1998-2008 W. Q. Meeker and L. A. Escobar.
Based on the authors' text *Statistical Methods for Reliability
Data*, John Wiley & Sons Inc. 1998.

July 30, 2009
8h 4min

14 - 1

Introduction to the Use of Bayesian Methods for
Reliability Data
Chapter 14 Objectives

- Describe the use of Bayesian statistical methods to combine **prior** information with data to make inferences.
- Explain the relationship between Bayesian methods and likelihood methods used in earlier chapters.
- Discuss sources of prior information.
- Describe useful computing methods for Bayesian methods.
- Illustrate Bayesian methods for estimating reliability.
- Illustrate Bayesian methods for prediction.
- Compare Bayesian and likelihood methods under different assumptions about prior information.
- Explain the dangers of using wishful thinking or expectations as prior information.

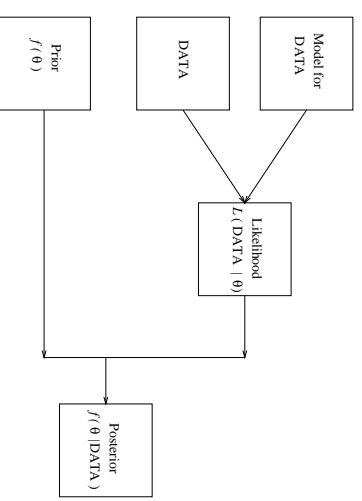
14 - 2

Introduction

- Bayes 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: prior information subjective.
 - Empirical Bayes: prior information from past data.
- Bayesian methods are closely related to likelihood methods.

14 - 3

Bayes Method for Inference



14 - 4

Updating Prior Information Using Bayes Theorem

Bayes Theorem provides a mechanism for combining *prior* information with sample data to make inferences on model parameters.

For a vector parameter θ the procedure is as follows:

- Prior information on θ is expressed in terms of a pdf $f(\theta)$.
- We observe some data which for the specified model has likelihood $L(\text{DATA}|\theta) \equiv L(\theta; \text{DATA})$.

- Using Bayes Theorem, the conditional distribution of θ given the data (also known as the **posterior** of θ) is

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

where $R(\theta) = L(\theta)/L(\theta)$ is the relative likelihood and the multiple integral is computed over the region $f(\theta) > 0$.

14 - 5

Some Comments on Posterior Distributions

- The posterior $f(\theta|\text{DATA})$ is function of the prior, the model, and the data.
- In general, it is impossible to compute the multiple integral $\int L(\text{DATA}|\theta)f(\theta)d\theta$ in closed form.
- New statistical and numerical methods that take advantage of modern computing power are facilitating the computation of the posterior.

14 - 6

Differences Between Bayesian and Frequentist Inference

- Nuisance parameters
- ▶ Bayes methods use marginals.
- ▶ Large-sample likelihood theory suggest maximization.
- There are not important differences in large samples.
- Interpretation
- ▶ Bayes methods justified in terms of probabilities.
- ▶ Frequentist methods justified on repeated sampling and asymptotic theory.

14 - 7

Sources of Prior Information

- Informative.
 - ▶ Past data.
 - ▶ Physical, chemical, and mechanical theory.
 - ▶ Expert knowledge.
- Diffuse (or approximately non-informative).
 - ▶ Uniform over finite range of parameter (or function of parameter).
 - ▶ Uniform over infinite range of parameter (improper prior).
 - ▶ Other vague or diffuse priors.

14 - 8

Proper Prior Distributions

Any positive function defined on the parameter space that integrates to a finite value (usually 1).

- **Uniform prior:** $f(\theta) = 1/(b - a)$ for $a \leq \theta \leq b$. This prior does not express strong preference for specific values of θ in the interval.
- **Examples of non-uniform prior distributions:**
 - ▶ Normal with mean at a and standard deviation b .
 - ▶ Beta between specified a and b with specified shape parameters (allows for a more general shape).
 - ▶ Isosceles triangle with base (range between) a and b .

For a positive parameter θ , may want to specify the prior in terms of $\log(\theta)$.

14 - 9

Improper Prior Distributions

Positive function $f(\theta)$ over parameter space for which

$$\int f(\theta)d\theta = \infty,$$

- **Uniform** in an interval of infinite length: $f(\theta) = c$ for all θ .
 - For a positive parameter θ the corresponding choice is $f[\log(\theta)] = c$ and $f(\theta) = (c/\theta)$, $\theta > 0$.
- To use an improper prior, one must have
- $$\int f(\theta)L(\theta|DATA)d\theta < \infty$$
- (a condition on the form of the likelihood and the DATA).
- These prior distributions can be made to be proper by specification of a finite interval for θ and choosing c such that the total probability is 1.

14 - 10

Effect of Using Vague (or Diffuse) Prior Distributions

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

$$f(\theta|DATA) = \frac{R(\theta)f(\theta)}{\int R(\theta)f(\theta)d\theta} = \frac{R(\theta)}{\int R(\theta)d\theta}$$
 which indicates that the posterior $f(\theta|DATA)$ is proportional to the likelihood.

- The posterior is approximately proportional to the likelihood for a proper (finite range) uniform if the range is large enough so that $R(\theta) \approx 0$ where $f(\theta) = 0$.
- Other diffuse priors also result in a posterior that is approximately proportional to the likelihood if $R(\theta)$ is large relative to $f(\theta)$.

14 - 11

Eliciting or Specifying a Prior Distribution

- The elicitation of a meaningful joint prior distribution for vector parameters may be difficult
 - ▶ The marginals may not completely determine the joint distribution.
 - ▶ Difficult to express/elicit dependences among parameters through a joint distribution.
 - ▶ The standard parameterization may not have practical meaning.

- General approach: choose an appropriate parameterization in which the priors for the parameters are approximately independent.

14 - 12

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 is **no** useful informative prior information, determine the form and range of the vague prior (e.g., uniform over a wide interval).
- For parameters for which there is useful informative prior information, specify the form and range of the distribution (e.g., lognormal with 99.7% content between two specified points).

14 - 13

Example of Eliciting Prior Information: Bearing-Cage Time to Fracture Distribution

With appropriate questioning, engineers provided the following information:

- Time to fracture data can often be described by a Weibull distribution.
- From previous similar studies involving heavily censored data, (μ, σ) tend to be correlated (making it difficult to specify a joint prior for them).
- For small p (near the proportion failing in previous studies), (t_p, σ) are approximately independent (which allows for specification of approximately independent priors).

14 - 14

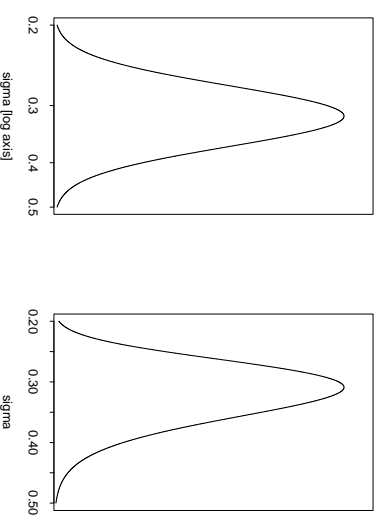
Example of Eliciting Prior Information: Bearing-Cage Fracture Field Data (Continued)

- 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.
- The engineers did not have strong prior information on possible values for the distribution quantiles.
- For the Weibull shape parameter $\log(\sigma) \sim \text{NOR}(a_0, b_0)$, where a_0 and b_0 are obtained from the specification of two quantiles $\sigma_{\gamma/2}$ and $\sigma(1-\gamma/2)$ of the prior distribution for σ . Then

$$a_0 = \log \left[\sqrt{\sigma_{\gamma/2} \times \sigma(1-\gamma/2)} \right], \quad b_0 = \log \left[\frac{\sqrt{\sigma(1-\gamma/2)/\sigma_{\gamma/2}}}{\gamma(1-\gamma/2)} \right]$$
- Uncertainty in the Weibull 0.01 quantile will be described by UNIFORM[log(a_1), log(b_1)] distribution where $a_1 = 100$ and $b_1 = 5000$ (wide range—not very informative).

14 - 15

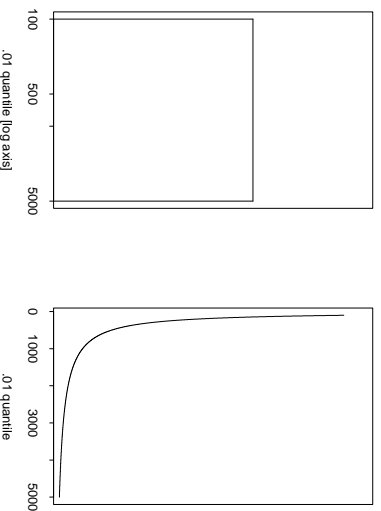
Prior pdfs for log(σ) and σ when $\sigma_{0.005} = 0.2, \sigma_{0.995} = 0.5$



14 - 16

Prior pdfs for log($t_{0.01}$) and $t_{0.01}$ when

$$a_1 = 100, b_1 = 5000$$



14 - 17

Joint Lognormal-Uniform Prior Distributions

- The prior for log(σ) is normal

$$f[\log(\sigma)] = \frac{1}{b_0} \phi_{\text{nor}} \left[\frac{\log(\sigma) - a_0}{b_0} \right], \quad \sigma > 0.$$

The corresponding density for σ is $f(\sigma) = (1/\sigma) f[\log(\sigma)]$.

- The prior for log(t_p) is uniform

$$f[\log(t_p)] = \frac{1}{\log(b_1/a_1)}, \quad a_1 \leq t_p \leq b_1.$$

The corresponding density for t_p is $f(t_p) = (1/t_p) f[\log(t_p)]$.

- Consequently, the joint prior distribution for (t_p, σ) is

$$f(t_p, \sigma) = \frac{f[\log(t_p)] f[\log(\sigma)]}{t_p \sigma} \quad a_1 \leq t_p \leq b_1, \quad \sigma > 0.$$

14 - 18

Joint Prior Distribution for (μ, σ)

- The transformation $\mu = \log(t_p) - \Phi_{\text{sev}}^{-1}(p)\sigma, \sigma = \sigma$ yields the prior for (μ, σ)

$$\begin{aligned} f(\mu, \sigma) &= \frac{f[\log(t_p)]}{t_p} \times \frac{f[\log(\sigma)]}{\sigma} \times t_p \\ &= f[\log(t_p)] \times \frac{f[\log(\sigma)]}{\sigma} \\ &= \frac{1}{\log(b_1/a_1)} \times \frac{\phi_{\text{nor}}\{[\log(\sigma) - a_0]/b_0\}}{\sigma b_0} \end{aligned}$$

where $\log(a_1) - \Phi_{\text{sev}}^{-1}(p)\sigma \leq \mu \leq \log(b_1) - \Phi_{\text{sev}}^{-1}(p)\sigma, \sigma > 0$.

- The region in which $f(\mu, \sigma) > 0$ is South-West to North-East oriented because $\text{Cov}(\mu, \sigma) = -\Phi_{\text{sev}}^{-1}(p)\text{Var}(\sigma) > 0$.

14 - 19

Joint Posterior Distribution for (μ, σ)

- The likelihood is

$$L(\mu, \sigma) = \prod_{i=1}^{2003} \left\{ \frac{1}{\sigma t_i} \phi_{\text{sev}} \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^4 \times \left\{ 1 - \Phi_{\text{sev}} \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^{1-\delta_i}$$

where δ_i indicates whether the observation i is a failure or a right censored observation.

- The posterior distribution is

$$f(\mu, \sigma | \text{DATA}) = \frac{L(\mu, \sigma) f(\mu, \sigma)}{\iint L(v, w) f(v, w) dv dw} = \frac{R(\mu, \sigma) f(\mu, \sigma)}{\iint R(v, w) f(v, w) dv dw}.$$

14 - 20

Computing the Posterior Using Simulation

Using simulation, one can draw a sample from the posterior using only the likelihood and the prior. The procedure for a general parameter θ and prior distribution $f(\theta)$ is as follows:

- Let $\theta_{i1}, i = 1, \dots, M$ be a random sample from $f(\theta)$.

- The i th observation, θ_{i1} , is retained with probability $R(\theta_{i1})$.

Then if U_{i1} is a random observation from a uniform $(0, 1)$, θ_{i1} is retained if

$$U_{i1} \leq R(\theta_{i1}).$$

- It can be shown that the retained observations, say $\theta_{11}^*, \dots, \theta_{M^*1}^*$ ($M^* \leq M$) are observations from the posterior $f(\theta | \text{DATA})$.

14 - 22

Methods to Compute the Posterior

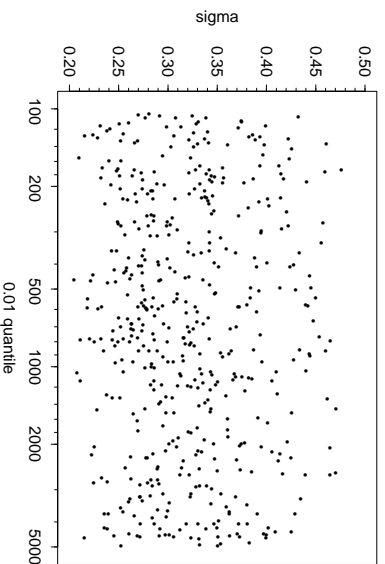
- Numerical integration:** to obtain the posterior, one needs to evaluate the integral $f(\theta | \text{DATA}) = \int R(\theta) f(\theta) d\theta$ over the region on which $f(\theta) > 0$.

In general there is not a closed form for the integral and the computation has to done numerically using fixed quadrature or adaptive integration algorithms.

- Simulation methods:** the posterior can be approximated using Monte Carlo simulation resampling methods.

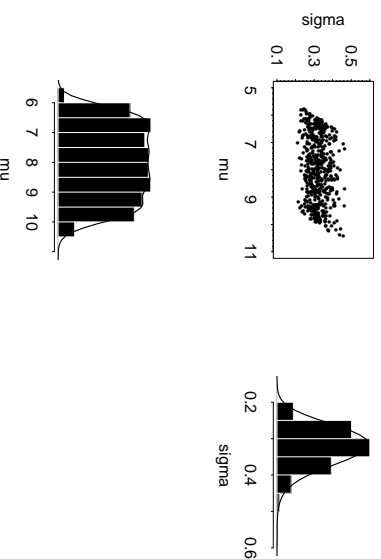
14 - 21

Simulated Joint Prior for $t_{0.01}$ and σ



14 - 23

Simulated Joint and Marginal Prior Distributions for μ and σ



14 - 24

Sampling from the Prior

The joint prior for $\theta = (\mu, \sigma)$, is generated as follows:

- Use the inverse cdf method (see Chapter 4) to obtain a pseudorandom sample for t_p , say

$$(t_p)_i = a_1 \times b_1^{U_{1i}}, \quad i = 1, \dots, M$$

where U_{11}, \dots, U_{1M} are a pseudorandom sample from a uniform $(0, 1)$.

- Similarly, obtain a pseudorandom sample for σ , say

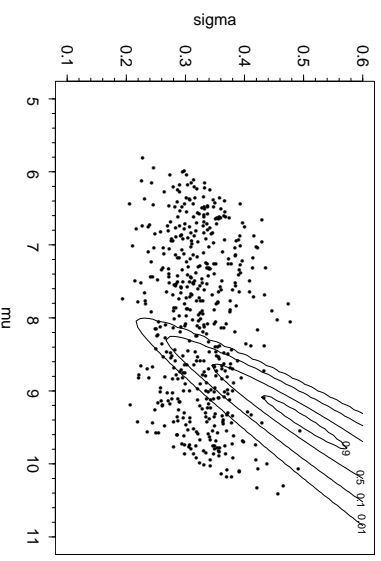
$$\sigma_i = \exp[a_0 + b_0 \Phi_{\text{nor}}^{-1}(U_{2i})]$$

where U_{21}, \dots, U_{2M} are another independent pseudorandom sample from a uniform $(0, 1)$.

- Then $\theta_i = (\mu_i, \sigma_i)$ with $\mu_i = \log[(t_p)_i] - \Phi_{\text{sev}}^{-1}(p)\sigma_i$ is a pseudorandom sample from the (μ, σ) prior.

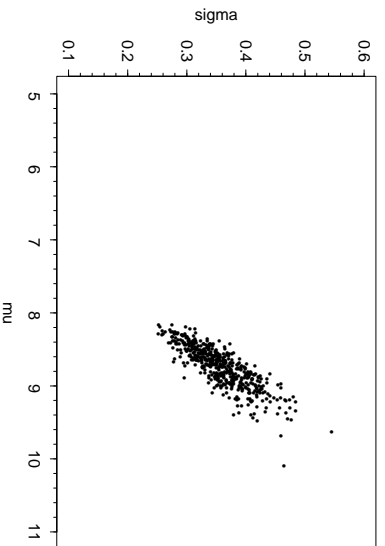
14-25

Simulated Joint Prior Distribution with μ and σ Relative Likelihood



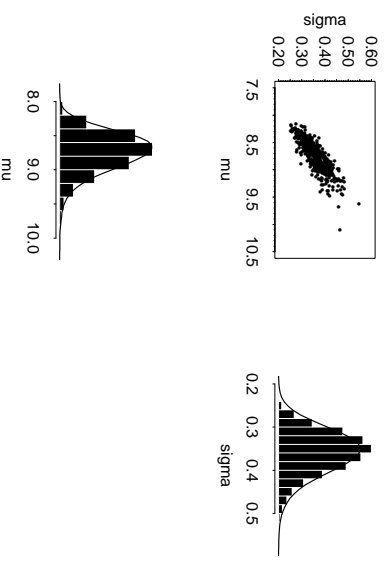
14-26

Joint Posterior for μ and σ



14-27

Joint Posterior and Marginals for μ and σ for the Bearing Cage Data



14-28

Comments on Computing Posteriors Using Resampling

The number of observations M^* from the posterior is random with an expected value of

$$E(M^*) = M \int f(\theta)R(\theta)d\theta$$

Consequently,

- When the prior and the data do not agree well, $M^* \ll M$ otherwise and a larger prior sample will be required.
- Can add to the posterior by sequentially filtering groups of prior points until a sufficient number is available in the posterior.

14-29

Posterior and Marginal Posterior Distributions for the Model Parameters

- Inferences on individual parameters are obtained by using the marginal posterior distribution of the parameter of interest. The marginal posterior of θ_j is

$$f[\theta_j | \text{DATA}] = \int f(\theta | \text{DATA})d\theta'$$

where θ' is the subset of the parameters excluding θ_j .

- Using the general resampling method described above, one gets a sample for the posterior for θ , say $\theta_i^* = (\mu_i^*, \sigma_i^*)$, $i = 1, \dots, M^*$.
- Inferences for μ or σ alone are based on the corresponding marginal distributions μ_i^* and σ_i^* , respectively.

14-30

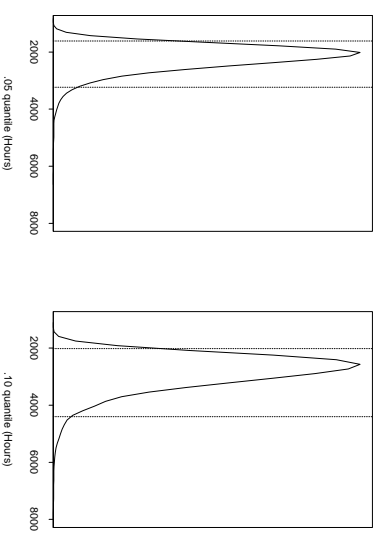
Posterior and Marginal Posterior Distributions for the Functions of Model Parameters

- Inferences on a scalar function of the parameters $g(\theta)$ are obtained by using the marginal posterior distribution of the functions of the parameters of interest, $f[g(\theta)|\text{DATA}]$.
- Using the simulation method, inferences are based on the simulated posterior marginal distributions. For example:
 - ▶ The marginal posterior distribution of $f(t_p|\text{DATA})$ for inference on quantiles is obtained from the empirical distribution of $\mu_i^* + \Phi_{\text{seV}^{-1}}^{-1}(p)\sigma_i^*$.
 - ▶ The marginal posterior distribution of $f[F(t_c)|\text{DATA}]$ for inference for failure probabilities at t_c is obtained from the empirical distribution of $\Phi_{\text{seV}} \left[\frac{\log(t_c) - \mu_i^*}{\sigma_i^*} \right]$.

14 - 31

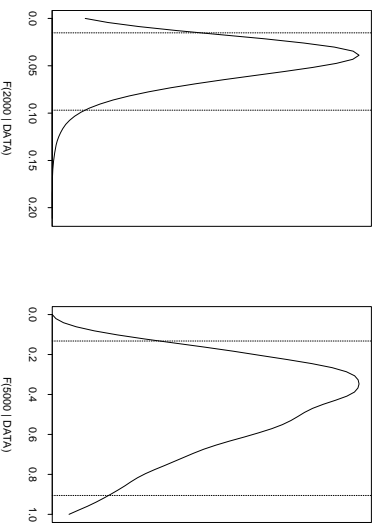
Simulated Marginal Posterior Distributions

for $t_{0,05}$ and $t_{0,10}$



14 - 32

Simulated Marginal Posterior Distributions for $F(2000)$ and $F(5000)$



14 - 33

Bayes Point Estimation

Bayesian inference for θ and functions of the parameters $g(\theta)$ are entirely based on their posterior distributions $f(\theta|\text{DATA})$ and $f[g(\theta)|\text{DATA}]$.

Point Estimation:

- If $g(\theta)$ is a scalar, a common Bayesian estimate of $g(\theta)$ is its posterior mean, which is given by

$$\hat{g}(\theta) = E[g(\theta)|\text{DATA}] = \int g(\theta)f(\theta|\text{DATA})d\theta.$$
 In particular, for the i th component of θ , $\hat{\theta}_i$ is the posterior mean of θ_i . This estimate is the the Bayes estimate that minimizes the square error loss.

- Other possible choices to estimate $g(\theta)$ include (a) the posterior mode, which is very similar to the ML estimate and (b) the posterior median.

14 - 34

One-Sided Bayes Confidence Bounds

- A $100(1 - \alpha)\%$ Bayes lower confidence bound (or credible bound) for a scalar function $g(\theta)$ is value \tilde{g} satisfying

$$\int_{\tilde{g}}^{\infty} f[g(\theta)|\text{DATA}]d_g(\theta) = 1 - \alpha$$

- A $100(1 - \alpha)\%$ Bayes upper confidence bound (or credible bound) for a scalar function $g(\theta)$ is value \tilde{g} satisfying

$$\int_{-\infty}^{\tilde{g}} f[g(\theta)|\text{DATA}]d_g(\theta) = 1 - \alpha$$

14 - 35

Two-Sided Bayes Confidence Intervals

- A $100(1 - \alpha)\%$ Bayes confidence interval (or credible interval) for a scalar function $g(\theta)$ is any interval $[\tilde{g}_l, \tilde{g}_h]$ satisfying

$$\int_{\tilde{g}_l}^{\tilde{g}_h} f[g(\theta)|\text{DATA}]d_g(\theta) = 1 - \alpha \quad (1)$$

- The interval $[\tilde{g}_l, \tilde{g}_h]$ can be chosen in different ways
 - ▶ Combining two $100(1 - \alpha/2)\%$ intervals puts equal probability in each tail (preferable when there is more concern for being incorrect in one direction than the other).
 - ▶ A $100(1 - \alpha)\%$ Highest Posterior Density (HPD) confidence interval chooses $[\tilde{g}_l, \tilde{g}_h]$ to consist of all values of g with $f(g|\text{DATA}) > c$ where c is chosen such that (1) holds. HPD intervals are similar to likelihood-based confidence intervals. Also, when $f[g(\theta)|\text{DATA}]$ is unimodal the HPD is the narrowest Bayes interval.

14 - 36

Bayesian Joint Confidence Regions

The same procedure generalizes to confidence regions for vector functions $g(\theta)$ of θ .

- A $100(1-\alpha)\%$ Bayes confidence region (or credible region) for a vector valued function $g(\theta)$ is defined as

$$CR_B = \{g(\theta) | f[g|\text{DATA}] \geq c\}$$

where c is chosen such that

$$\int_{CR_B} f[g(\theta)|\text{DATA}]dg(\theta) = 1 - \alpha$$

- In this case the presentation of the confidence region is difficult when θ has more than 2 components.

14 - 37

Bayes Versus Likelihood

- Summary table or plots to compare the Likelihood versus the Bayes Methods to compare confidence intervals for μ , σ , and $t_{0.1}$ for the Bearing-cage data example.

14 - 38

Prediction of Future Events

- Future events can be predicted by using the Bayes predictive distribution.

- If X [with pdf $f(\cdot|\theta)$] represents a future random variable

▶ the posterior predictive pdf of X is

$$f(x|\text{DATA}) = \int f(x|\theta)f(\theta|\text{DATA})d\theta \\ = E_{\theta|\text{DATA}}[f(x|\theta)]$$

▶ the posterior predictive cdf of X is

$$F(x|\text{DATA}) = \int_{-\infty}^x f(u|\theta)du = \int F(x|\theta)f(\theta|\text{DATA})d\theta \\ = E_{\theta|\text{DATA}}[F(x|\theta)]$$

where the expectations are computed with respect to the posterior distribution of θ .

14 - 39

Approximating Predictive Distributions

- $f(x|\text{DATA})$ can be approximated by the average of the posterior pdfs $f(x|\theta_i^*)$. Then

$$f(x|\text{DATA}) \approx \frac{1}{M^*} \sum_{i=1}^{M^*} f(x|\theta_i^*).$$

- Similarly, $F(x|\text{DATA})$ can be approximated by the average of the posterior cdfs $F(x|\theta_i^*)$. Then

$$F(x|\text{DATA}) \approx \frac{1}{M^*} \sum_{i=1}^{M^*} F(x|\theta_i^*).$$

- A two-sided $100(1-\alpha)\%$ Bayesian prediction interval for a new observation is given by the $\alpha/2$ and $(1-\alpha/2)$ quantiles of $F(x|\text{DATA})$.

14 - 40

Location-Scale Based Prediction Problems

Here we consider prediction problems when $\log(T^*)$ has a location-scale distribution.

- Predicting a future value of T . In this case, $X = T$ and $x = t$, then

$$f(t|\theta) = \frac{1}{\sigma t} \phi(\zeta), \quad F(t|\theta) = \Phi(\zeta)$$

where $\zeta = [\log(t) - \mu]/\sigma$.

- Thus, for the Bearing-cage fracture data, approximations of the predictive pdf and cdf for a **new** observation are:

$$f(t|\text{DATA}) \approx \frac{1}{M^*} \sum_{i=1}^{M^*} \frac{1}{\sigma t_i^*} \phi_{\text{sev}}(\zeta_i^*) \\ F(t|\text{DATA}) \approx \frac{1}{M^*} \sum_{i=1}^{M^*} \Phi_{\text{sev}}(\zeta_i^*)$$

where $\zeta_i^* = [\log(t) - \mu_i^*]/\sigma_i^*$.

14 - 41

Prediction of an Order Statistic

Here we consider prediction of the k th order statistic in a future sample of size m from the distribution of T when $\log(T^*)$ has a location-scale distribution.

- In this case, $X = T_{(k)}$ and $x = t_{(k)}$, then

$$f[t_{(k)}|\theta] = \frac{m!}{(k-1)!(m-k)!} \times [\Phi(\zeta)]^{k-1} \times \frac{1}{\sigma t_{(k)}} \phi(\zeta) \\ \times [1 - \Phi(\zeta)]^{m-k} \\ F[t_{(k)}|\theta] = \sum_{j=k}^m \frac{m!}{j!(m-j)!} [\Phi(\zeta)]^j \times [1 - \Phi(\zeta)]^{m-j}$$

where $\zeta = [\log(t_{(k)}) - \mu]/\sigma$.

14 - 42

Predicting the 1st Order Statistic

When $k = 1$ (predicting the 1st order statistic), the formulas simplify to

- Predictive pdf

$$f(t_{(1)}|\theta) = m \times \frac{1}{\sigma t_{(1)}} \phi(\zeta) \times [1 - \Phi(\zeta)]^{m-1}$$

- Predictive cdf

$$F[t_{(1)}|\theta] = 1 - [1 - \Phi(\zeta)]^m$$

where $\zeta = [\log(t_{(1)}) - \mu]/\sigma$.

14 - 43

Predicting the 1st Order Statistic for the Bearing-Cage Fracture Data

For the Bearing-cage fracture data:

- An approximation for the predictive pdf for the 1st order statistic is

$$f(t_{(1)}|\text{DATA}) \approx \frac{1}{M^*} \sum_{i=1}^{M^*} \left\{ m \times \frac{1}{\sigma_i^* t} \phi(\zeta_i^*) \times [1 - \Phi(\zeta_i^*)]^{m-1} \right\}$$

- The corresponding predictive cdf is

$$F[t_{(1)}|\text{DATA}] \approx \frac{1}{M^*} \sum_{i=1}^{M^*} \left\{ 1 - [1 - \Phi(\zeta_i^*)]^m \right\}$$

where $\zeta_i^* = [\log(t) - \mu_i^*]/\sigma_i^*$.

14 - 44

Predicting a New Observation

- $F(t|\text{DATA})$ can be approximated by the average of the posterior probabilities $F(t|\theta_i^*)$, $i = 1, \dots, M^*$.

- Similarly, $f(t|\text{DATA})$ can be approximated by the average of the posterior densities $f(t|\theta_i^*)$, $i = 1, \dots, M^*$.

- In particular for the Bearing-cage fracture data, an approximation for the predictive pdf and cdf are

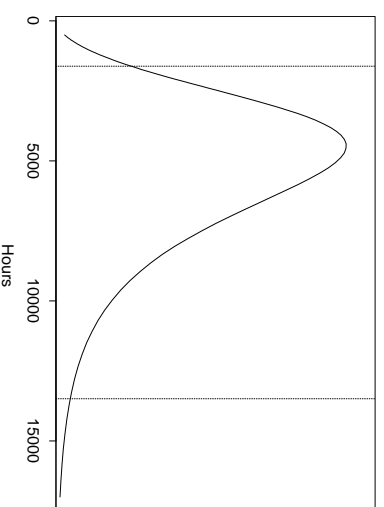
$$f(t|\text{DATA}) \approx \frac{1}{M^*} \sum_{i=1}^{M^*} \frac{1}{\sigma_i^*} \phi_{\text{sev}} \left[\frac{\log(t) - \mu_i^*}{\sigma_i^*} \right]$$

$$F(t|\text{DATA}) \approx \frac{1}{M^*} \sum_{i=1}^{M^*} \Phi_{\text{sev}} \left[\frac{\log(t) - \mu_i^*}{\sigma_i^*} \right].$$

- A 100(1 - α)% Bayesian prediction interval for a new observation is given by the percentiles of this distribution.

14 - 45

Predictive Density and Prediction Intervals for a Future Observation from the Bearing Cage Population



14 - 46

Caution 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.
- We must beware of the use of **wishful thinking** 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 ones model (including the inputted prior information)

14 - 47