Reliability Data Analysis Using S-PLUS

William Q. Meeker
Department of Statistics and
Iowa State University
Ames, IA 50011
32nd Symposium on the Interface:
Computing Science and Statistics
New Orleans, LA
April 5-8, 2000

Overview
- History, other software
- Specific user needs and GUI development goals
- Data objects and outline of menu structure
- Examples
- Numerical Methods
- Concluding remarks and future work

History and Development of SLIDA
- Motivated by GE's STATPAC (Nelson et al. 1972,...)
- CENSOR Fortran program (Meeker and Duke 1978-1980 at ISU)
- STAR written in Old S (Meeker and others at Bell Labs QAC 1985-1986) then translated into C to be a commercial product (Others at Bell Labs QAC 1986-1989)
- GENMAX Fortran (Meeker 1986-1992)
- SLIDA object-oriented S-PLUS commands (Meeker 1997-present)
- SLIDA S-PLUS GUI (Meeker 1998-present)

Other Reliability Data Analysis Software
- Terry Therneau's SURVIVAL package (in S-PLUS)
- SAS, JMP, MINITAB
- Special purpose packages:
  - WEIBULL++ and ALTA by Relisoft
  - WinSmith
- Various biomedical packages that handle censored data

General Needs for Reliability Data Analyses
- Complicated censoring and/or truncation; multiple failure modes.
- Wide range of standard and nonstandard models (e.g., non-normal distributions, nonlinear relationships)
- Estimates and statistical intervals for failure probabilities, distribution quantiles, failure rates, predictions for future number of failures, etc.
- Integration of analytical methods by graphically displaying data and fitted models together.
- Methods for planning reliability studies
- Use of simulation in inference and planning

SLIDA User Interface
- Most outputs given in graphical form.
- With numerous options, command argument specification is complicated.
- GUI simplifies option choice. The most important functionality in Meeker and Escobar (1998) plus recent developments (driven by courses).
SLIDA Data Objects

- Data objects contain important and useful (optional) information about a data set (defines response, censoring, truncation, weights, explanatory variables, title, units, notes, etc.)
  - Life data objects
    - Single distribution
    - Multiple failure modes
    - Single explanatory variable with a few levels
    - Comparison explanatory variable
    - General explanatory variables
  - Recurrence data (point process) objects
  - Repeated measures (degradation) data objects
- Multiple methods (different types of analyses) can be performed on particular data objects

Goals for the SLIDA GUI Design

- Easy for occasional new users.
- Organize according to how users want to do their work.
- Hide complexity for common users and everyday tasks; allow access for experts/experienced users.
- Develop structure to guide unfamiliar users through their work (without restricting the experts).
- Minimize required inputs; use defaults whenever possible.
- Minimize the need for typing/remembering. Present some or all choices whenever possible; eliminate in appropriate choices. Recall previous inputs for defaults, when appropriate.
- Signal bad (or questionable) input as soon as possible.

Result: Some complicated analyses much easier to do through the GUI than through commands; engineers like it.
SLIDA Top-Level Menu

- Make/summary/view/modify data object
- Plan single a distribution study
- Single distribution life data analyses
- Multiple failure mode life data analysis
- Comparison of distributions life data analysis
- Plan an accelerated life test (ALT)
- Simple regression (ALT) data analysis
- Multiple regression (ALT) life data analysis
- Regression residual analysis
- Recurrence (point process) data analysis
- Degradation (repeated measures) data analysis
- Preferences (change SLIDA default options)

SLIDA → Single distribution life data analyses

- Plot nonparametric estimate of cdf and confidence bands
- Probability plot with nonparametric confidence bands
- Probability plot with parametric ML fit
- Likelihood contour plot
- Compare distribution ML fits on probability plot
- Threshold parameter probability plot with parametric ML fit

Bearing Cage Failure-Time Data Lognormal Probability Plot and MLE

Bearing Cage Failure-Time Data Weibull-Lognormal Comparison

Bearing Cage Failure-Time Data Lognormal Probability Plot

Bearing Cage Failure-Time Data Weibull-Lognormal Comparison

Bearing Cage Failure-Time Data Lognormal Probability Plot

Bearing Cage Failure-Time Data Weibull-Lognormal Comparison

Bearing Cage Failure-Time Data Lognormal Probability Plot

Bearing Cage Failure-Time Data Weibull-Lognormal Comparison
**Bearing Cage Failure-Time Data**  
**Lognormal Joint Confidence Region**

<table>
<thead>
<tr>
<th>mu</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.5</td>
</tr>
<tr>
<td>8</td>
<td>1.0</td>
</tr>
<tr>
<td>10</td>
<td>1.5</td>
</tr>
<tr>
<td>12</td>
<td>2.0</td>
</tr>
<tr>
<td>14</td>
<td>2.5</td>
</tr>
<tr>
<td>16</td>
<td>3.0</td>
</tr>
<tr>
<td>18</td>
<td>3.5</td>
</tr>
<tr>
<td>20</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Thu Sep 30 18:02:51 CDT 1999

---

**Numerical Methods**

- **Algorithms**
  - Stable optimization
  - Accurate mathematical/statistical functions
  - Analytical and numerical derivatives
- **Maximum likelihood estimation**
  - Stable parameterization
  - Good starting values
  - Dealing with unbounded or flat likelihood functions.

---

**SLIDA → Plan single distribution study →**

- Specify life test planning information (planning values)
- Plot life test planning information (planning values)
- Plot of approximate required sample size
- Simulate a life test
- Probability of successful demonstration

---

**Life Test Planning Values**

Weibull Distribution with eta = 6464 and beta = 0.8037

**Needed Sample Size for a Life Test**

Needed sample size giving approximately a 50% chance of having a confidence interval factor for the 0.1 quantile that is less than R

Weibull Distribution with eta = 6464 and beta = 0.8045
Test censored at 1000 Hours with 20 percent failing
New Spring Experiment
Lognormal Regression Model Life vs Stress Plot

Fixed values of Temp=600, Method=New for the Spring Fatigue Data

---

New Spring Experiment
Sensitivity Analysis Plot

Spring Fatigue Data with Lognormal Strikes Log, Temp linear, Method class at 30,600, New Power Transformation Sensitivity Analysis on Stroke

---

New Spring Experiment
.1 Quantile Sensitivity Analysis Likelihood Profile

Profile Likelihood and 95% Confidence Interval for Box-Cox Transformation Power from the Lognormal Distribution

---

Reliability Data Analysis
Concluding Remarks and Future Work

Considerable progress has been made in this area in the past 10 years; much remains to be done.

- Improvements in the user interface. Make it easy to learn how to use.
- Better approximate methods for censored-data confidence intervals (e.g., likelihood ratio and/or bootstrap).
- Better and more methods for recurrence and degradation data.
- Methods for incorporating prior information (flexible, easy to use, Bayes methods).
- Better (i.e., easier to implement and more robust) methods for fitting user-specified models the system.