Chapter 16

Analysis of Data with More than One Failure Mode

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
11h 2min

Chapter 16
Analysis of Data with More than One Failure Mode

Topics discussed in this chapter are:

- Examples, model, data and likelihood for describing multiple-failure-mode data.
- Maximum likelihood estimation, illustrated by estimating the MTTF of Device-G.
- Estimating the Shock Absorber failure-time distribution and when it is important to use failure-mode information.
- Other topics related to multiple failure modes.

Chapter 16

Analysis of Data with More than One Failure Mode

Segment 1

Examples and Model

Data and Likelihood

Examples of Products With Two or More Causes of Failure (or Multiple Failure Modes)

Many units, systems, subsystems, or components have more than one cause of failure. For example:

- Laptop computer failure modes include keyboard, screen, processor, main memory, solid-state drive, and power supply.
- A shock absorber may fail from a broken spring or a crack in the casing.
- A semiconductor device can fail at a junction or at a lead.
- A hard drive can fail because a manufacturing defect (infant mortality) or because of mechanical wearout.
- For an automobile tire, tread can wearout or the tire may suffer a puncture.

Basic Ideas Behind the Analysis of Data with Multiple Failure Modes

- There is a joint distribution of failure times for all failure modes.
- Usually, only the first failure is observed.
- If the failure times for the individual failure modes are independent, then analysis is relatively simple.
- When only the first failure is observed and **no parametric distribution assumption** is made, there is no information in the data about the nature of the dependency.
- When only the first failure is observed and **a parametric distribution is assumed**, there is some information in the data about the nature of the dependency (but not much).

Multiple Failure Modes Data and the Series-System Model

- Let T_j be the lifetime from failure mode j, $j = 1, \ldots J$.
- The failure-time of the system is
$$T = \min\{T_1, \ldots T_J\}.$$
- When the failure modes are **independent** and have a continuous distribution, the cdf for the failure-time is
$$F_T(t) = 1 - \prod_{j=1}^{J} [1 - F_j(t)]$$
where $F_j(t)$ is the marginal cdf for mode j and $f_j(t) = dF_j(t)/d\,t$ is the marginal pdf for mode j.
- The observed data are $(t_i, \kappa_i), i = 1, \ldots, n$ where t_i is a given time for observation i and
$$\kappa_i = \begin{cases} j & \text{if the cause of failure } i \text{ was mode } j \\ 0 & \text{if the unit } i \text{ was censored} \end{cases}.$$

Analysis of Data with Multiple Failure Modes Assuming Independence

- When failure times for the different failure modes are independent, it is possible to analyze the individual failure modes separately.
- Create a separate data set, with n cases, for each failure mode.
- For failure mode j, the likelihood is

$$L(\mu_j, \sigma_j) = \prod_{i=1}^{n} \left[f_j(t_i; \mu_j, \sigma_j)\right]^{\delta_i} \left[1 - F_j(t_i; \mu_j, \sigma_j)\right]^{1-\delta_i}$$

where

$$\delta_i = \begin{cases} 1 & \text{if } \kappa = j \\ 0 & \text{otherwise} \end{cases}$$

Chapter 16

Analysis of Data with More than One Failure Mode

Segment 2

Estimating the MTTF of Device-G

Device-G Background

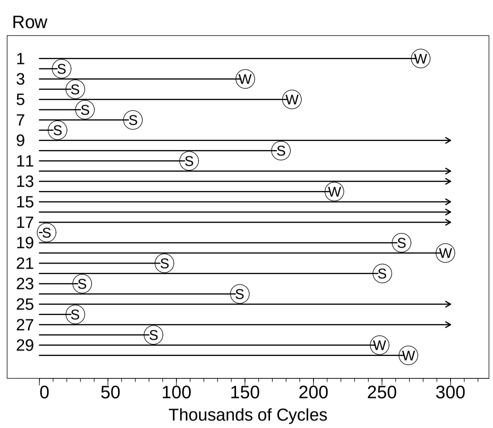
- Design life of 300 thousand cycles.
- Units were failing in the field more rapidly than had been expected.
- **Needed:**
  - ▶ An estimate of device MTTF
  - ▶ Information about how to improve reliability
- Mode W failures, caused by normal product wear, began to appear after 100 thousand cycles of use.
- Mode S failures were caused by failures on an electronic component due to electrical surge. These failures predominated early in life.

Device-G Data

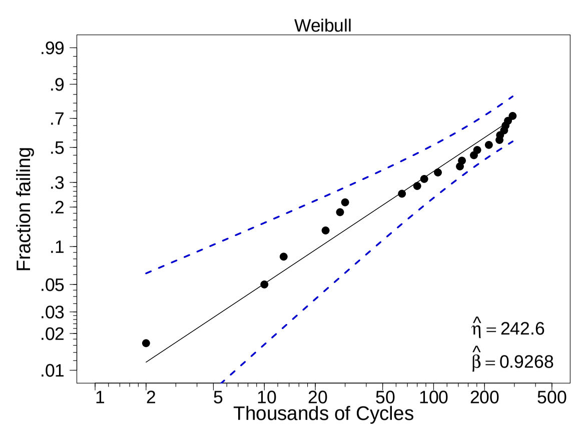
Thousands of Cycles	Failure Mode	Thousands of Cycles	Failure Mode	Thousands of Cycles	Failure Mode
275	W	106	S	88	S
13	S	300	–	247	S
147	W	300	–	28	S
23	S	212	W	143	S
181	W	300	–	300	–
30	S	300	–	23	S
65	S	300	–	300	–
10	S	2	S	80	S
300	–	261	S	245	W
173	S	293	W	266	W

W indicates a wearout failure, S indicates an electrical surge failure, and – indicates a unit still operating after 300 thousand cycles.

Event Plot of Device-G Data



Weibull Analyses of Device-G Data Estimating Time to Failure Ignoring the Cause of Failure



Series-System Model for Device-G

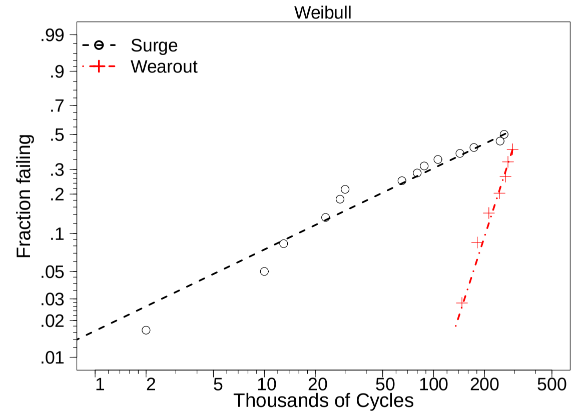
- Let T_{S} be the lifetime from the electrical surge failure mode (S) and T_{W} be the lifetime from the wearout failure mode (W).
- The failure-time of the product is

$$T = \min(T_{\mathrm{S}}, T_{\mathrm{W}}).$$

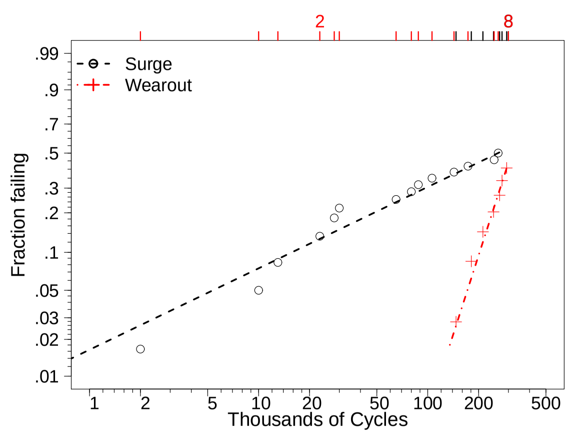
- When the failure modes are independent, the cdf for the failure-time is

$$\begin{aligned} F_T(t) &= \Pr(T \le t) = 1 - \Pr(T > t) = 1 - \Pr(T_{\mathrm{S}} > t \cap T_{\mathrm{W}} > t) \\ &= 1 - \Pr(T_{\mathrm{S}} > t)\Pr(T_{\mathrm{W}} > t) \\ &= 1 - [1 - F_{\mathrm{S}}(t)][1 - F_{\mathrm{W}}(t)]. \end{aligned}$$

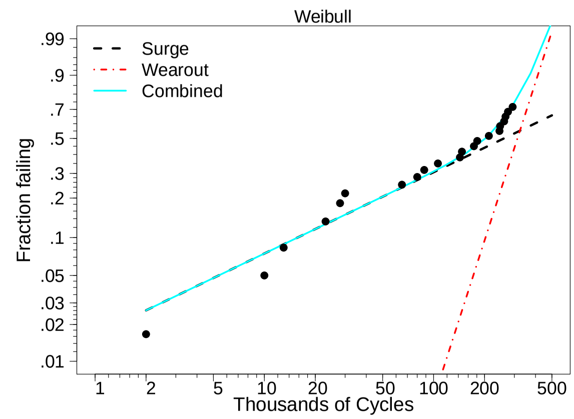
Weibull Analyses of Device-G Data Individual Failure Modes



Weibull Analyses of Device-G Data Individual Failure Modes



Weibull Analyses of Device-G Data Estimating Time to Failure Using Series System Model



Weibull Distribution Models for the Device-G Data

- Failure times for each of the two failure modes modeled with a separate Weibull distribution. That is,

$$F_j(t) = \Phi_{\mathrm{sev}}\left(\frac{\log(t) - \mu_j}{\sigma_j}\right), \quad j = \mathrm{S,\ W}.$$

- The **series-system model** for two independent failure modes acting together is

$$F_T(t) = 1 - [1 - F_{\mathrm{S}}(t)] \times [1 - F_{\mathrm{W}}(t)].$$

- The model ignoring the cause of failure information is

$$F_T(t) = \Phi_{\mathrm{sev}}\left(\frac{\log(t) - \mu_{\mathrm{SW}}}{\sigma_{\mathrm{SW}}}\right).$$

Device-G Field-Tracking Data ML Weibull Distribution Estimation Results for the Electric Surge (S) and Wearout (W) Failure Modes

Mode	Parameter	ML Estimate	Standard Error	95% Approximate Confidence Interval Lower	Upper
S	μ_{S}	6.11	0.43	5.27	6.95
	σ_{S}	1.49	0.35	0.94	2.36
W	μ_{W}	5.83	0.11	5.62	6.04
	σ_{W}	0.23	0.08	0.12	0.44
S & W	μ_{SW}	5.49	0.23	5.04	5.94
	σ_{SW}	1.08	0.21	0.74	1.57

For Mode S alone, $\mathcal{L}_{\mathrm{S}} = -101.36$ for Mode W alone, $\mathcal{L}_{\mathrm{W}} = -47.16$, and for both modes together, $\mathcal{L}_{\mathrm{SW}} = -142.62$.

Some Comments on the Weibull Analyses of Device-G Data

- The Weibull distribution provides a good fit to both the S failure mode and the W failure mode data.
- Weibull analysis ignoring the cause of failure information shows evidence of a change in the slope of the plotted points, indicating a gradual shift from one failure mode to another.
- The Weibull cdf estimate obtained from ignoring the cause of failure and the series-system cdf estimate for the two failure modes acting together diverge rapidly after 200 thousand cycles.
- Estimates of the mean time to failure computed from $\widehat{\mathrm{MTTF}} = \int_0^\infty [1 - \widehat{F}_T(t)]dt$ were 251.3 and 196.0 thousands of cycles, for the models ignoring and using the failure mode information, respectively.

Chapter 16

Analysis of Data with More than One Failure Mode

Segment 3

Estimating the Shock Absorber Failure-Time Distribution

Confidence Interval Methods

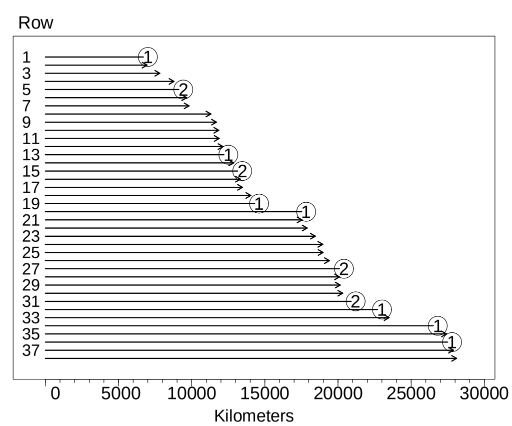
When it is Important to Use Failure-Mode Information?

Shock Absorber Failure Data

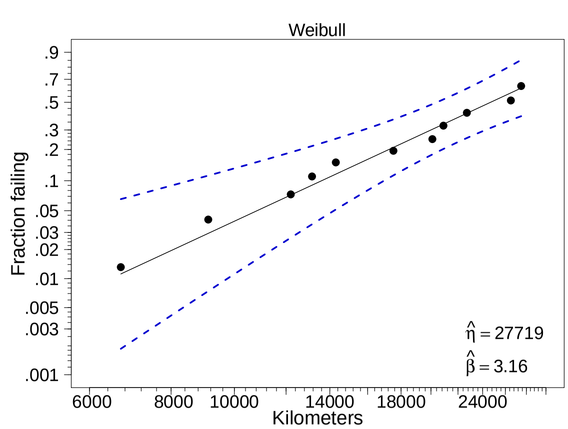
First reported in O'Connor (1985).

- Failure times, in number of kilometers of use, of vehicle shock absorbers.
- Two failure modes, denoted by Mode 1 and Mode 2.
- Need to estimate the failure-time distribution for the shock absorbers.
- One might be interested in the failure-time distribution for
  - Mode 1 (e.g., after Mode 2 has been eliminated)
  - Mode 2 (e.g., after Mode 1 has been eliminated)

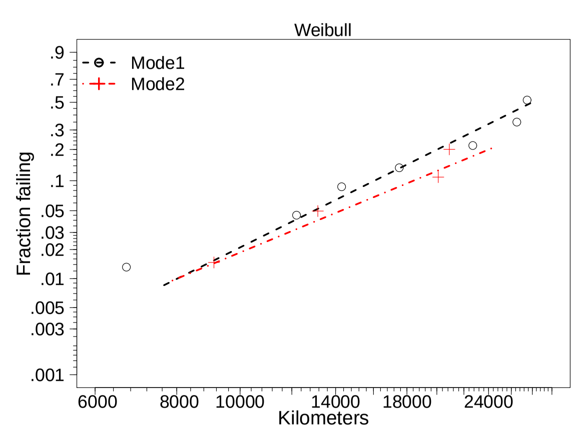
Event Plot of Shock Absorber Data



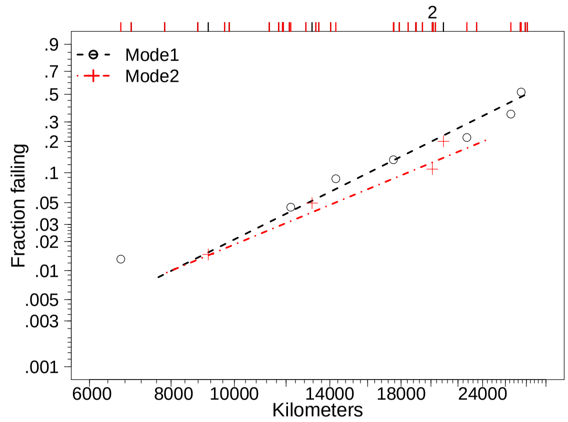
Weibull Analyses of Shock Absorber Data Estimating Time to Failure Ignoring the Cause of Failure



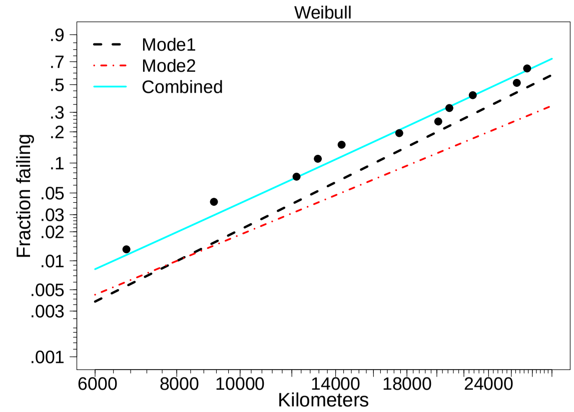
Weibull Analyses of Shock Absorber Data Individual Failure Modes



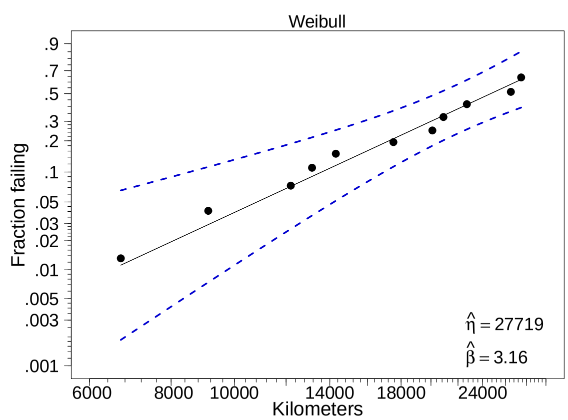
Weibull Analyses of Shock Absorber Data Individual Failure Modes



Weibull Analyses of Shock Absorber Data Estimating Time to Failure Using Series System Model



Weibull Analyses of Shock Absorber Data Estimating Time to Failure Ignoring the Cause of Failure



Confidence Intervals for Failure Probabilities and Quantiles

- Let $\widehat{\boldsymbol{\theta}}$ be the ML estimate of $\boldsymbol{\theta}$, and $\widehat{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ the ML estimate of $\Sigma_{\widehat{\boldsymbol{\theta}}}$ obtained from the component data. Then using same methods as in previous chapters

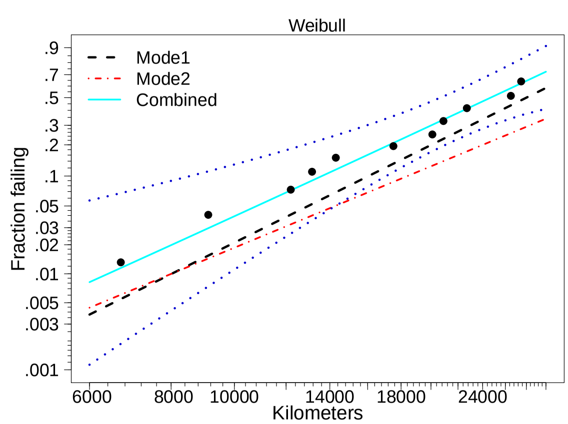
$$\widehat{F}_T = F_T(\widehat{\boldsymbol{\theta}}) = g[F_1(\widehat{\boldsymbol{\theta}}_1), \ldots, F_m(\widehat{\boldsymbol{\theta}}_m)]$$

$$\widehat{\text{Var}}(\widehat{F}_T) = \left(\frac{\partial F_T}{\partial \boldsymbol{\theta}}\right)' \widehat{\Sigma}_{\widehat{\boldsymbol{\theta}}} \left(\frac{\partial F_T}{\partial \boldsymbol{\theta}}\right)$$

where the derivatives are evaluated at $\widehat{\boldsymbol{\theta}}$.

- Confidence interval can be computed using
  - The delta method and a Wald approximation.
  - Inverting a likelihood ratio test
  - Bootstrap/GPQ methods

Weibull Analyses of Shock Absorber Data Estimating Time to Failure Using Series System Model



When is it Important to Take Account of Failure-Mode Information?

- If there is more than one failure mode, it is important to separate and analyze failure modes separately when:
  - Shape parameters are importantly different (e.g., when there is both infant mortality and wearout).
  - There is need to assess the impact of eliminating a failure mode.
  - There is a need to predict causes of future failures to assure that a sufficient number of appropriate replacement parts are available.
- If interest is on the failure time distribution of the entire system and that distribution can be adequately described by a simple distribution, then failure-mode information can be ignored.

Chapter 16

Analysis of Data with More than One Failure Mode

Segment 4

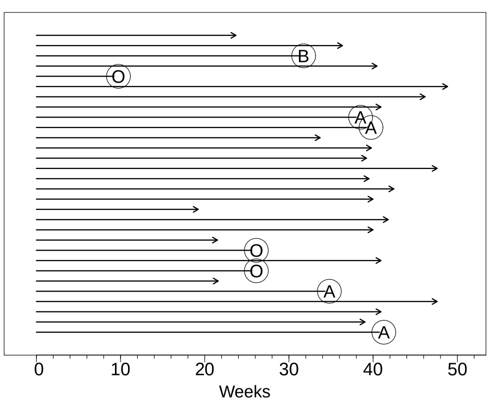
The Effect on System Reliability of Eliminating a Failue Mode

Product E Warranty Returns

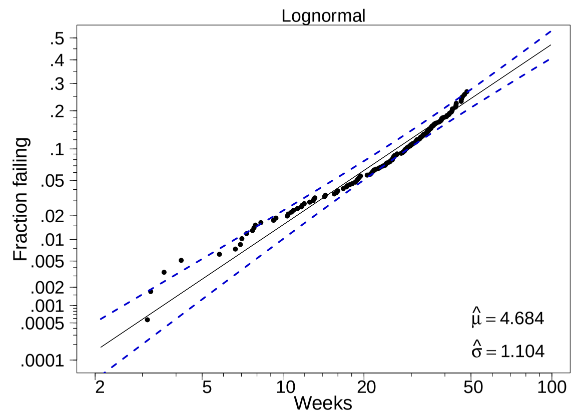
Product E Warranty Returns

- Failure times, in weeks of service.
- There were 19 different failure modes that appeared in the ErrorCode field of the data files. Failure modes with the same mechanism were combined.
- In the final data set there are three failure modes, denoted by Mode A and Mode B, and Other.
- Pareto analysis showed that the dominant failure mode was Failure Mode A with Failure Mode B also causing problems. Mode Other was mostly infant mortality.
- Need to estimate the failure-time distribution for the individual failure modes and make predictions of the number of failures of each type for the next year.

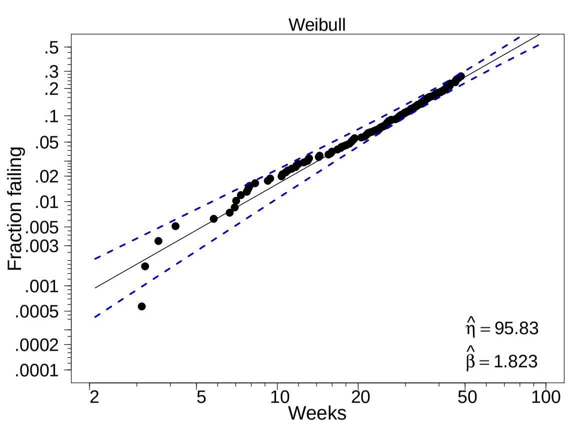
Event Plot of a Random Sample of Records from the Product-E Data



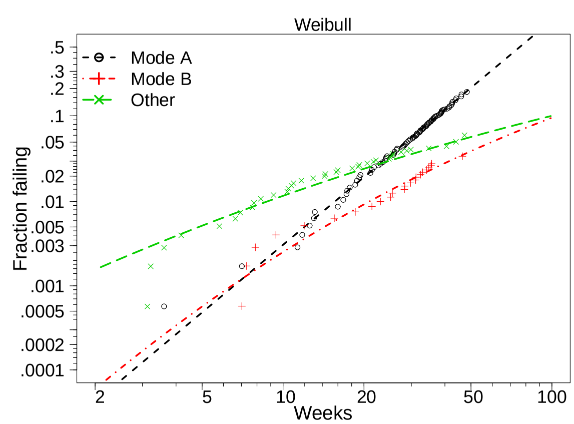
Lognormal Analyses of Product-E Data Estimating Time to Failure Ignoring the Cause of Failure



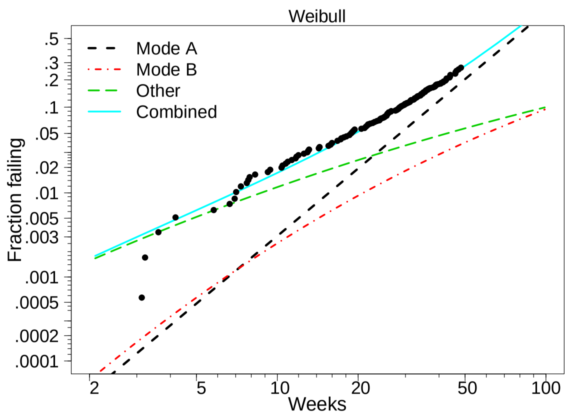
Weibull Analyses of Product-E Data Estimating Time to Failure Ignoring the Cause of Failure



Analyses of Product-E Data Individual Failure Modes Mode A is Weibull; Modes B and Other are Lognormal



Analyses of Product-E Data Estimating Time to Failure Using Series System Model Mode A is Weibull; Modes B and Other are Lognormal



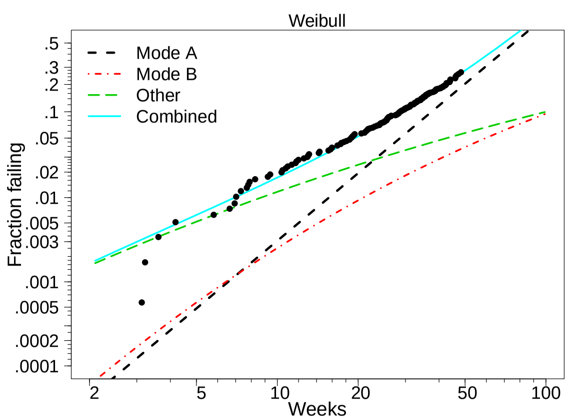
Estimating the Effect of Eliminating a Failure Mode

- Suppose that failure mode k can be eliminated without affecting the behavior of the other modes. Then under the series-system model with independent failure times for the different modes, the cdf of product life becomes

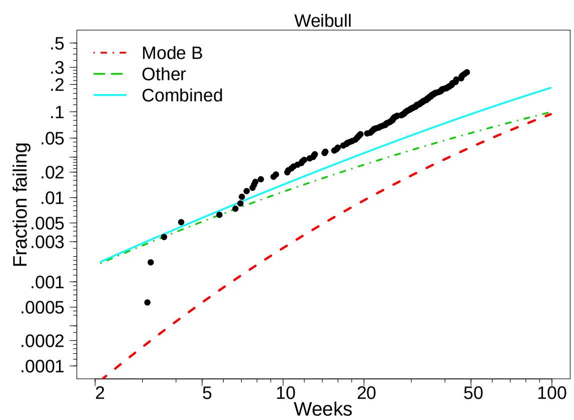
$$F_T(t) = 1 - \prod_{j \neq k} [1 - F_j(t)].$$

- If there are only $J = 2$ failure modes and one mode is eliminated, the failure-time distribution is exactly the marginal distribution of the remaining mode.
- Mode A was the dominant mode for Product E. Management wanted to know how much improvement would be seen after that mode would be eliminated.

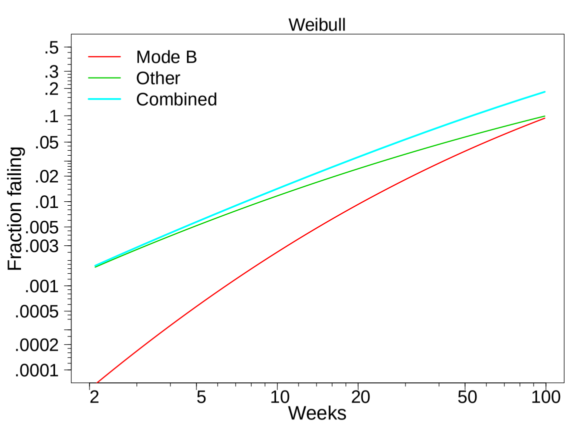
Analyses of Product-E Data Estimating Time to Failure Using Series System Model Mode A is Weibull; Modes B and Other are Lognormal



Weibull Analyses of Product-E Data Estimating Time to Failure Using Series System Model After Eliminating Failure Mode A



Weibull Analyses of Product-E Data Estimating Time to Failure Using Series System Model After Eliminating Failure Mode A



Chapter 16

Analysis of Data with More than One Failure Mode

Segment 5

Subdistribution Functions and Prediction for Product-E Individual Failure Modes

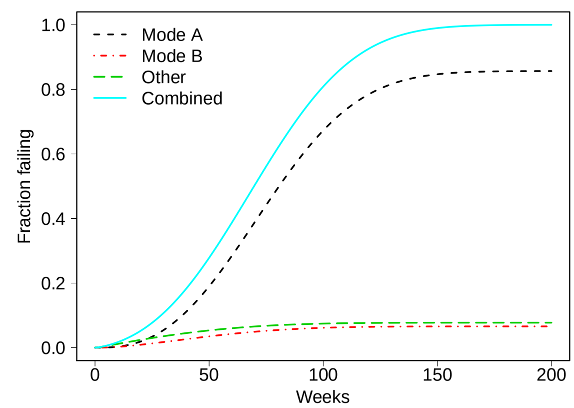
Subdistribution Functions of the Individual Failure Modes

- The subdistribution function of failure mode j is defined as

$$F_j^*(t) = \Pr(T \le t \cap \text{cause} = j) = \int_0^t f_j(w) \prod_{l \ne j} [1 - F_l(w)]\, dw.$$

- The subdistribution density of mode j is $f_j^*(t) = dF_j^*(t)/dt$.
- The collection of subdistribution functions for failure modes $j = 1, \ldots, J$ can be viewed as the joint distribution of **system** time to failure and failure mode.
- The system cdf is $F_T(t) = \sum_{j=1}^{J} F_j^*(t)$.
- $F_j^*(t)/F_T(t)$ is the fraction of mode j failures at time t.
- As $t \to \infty$, $F_j^*(t) \to p_j$, the overall proportion of units that will eventually fail due to mode j.
- Subdistributions are identifiable (and thus estimable), even when individual failure modes are dependent.

ML Estimates of the Product-E Subdistribution Functions



Predictions for Product-E Individual Failure Modes

	Δt (weeks)												
Failure mode	4	8	12	16	20	24	28	32	36	40	44	48	52
Mode A	25	27	30	32	33	35	35	36	36	35	34	33	31
Mode B	3	3	3	3	3	2	2	2	2	2	2	1	1
Mode Other	3	3	3	3	3	2	2	2	2	2	2	1	1
Total	31	33	36	38	39	39	39	40	40	39	38	35	33

Chapter 16

Analysis of Data with More than One Failure Mode

Segment 6

Non-Identifiability of Dependence Structure

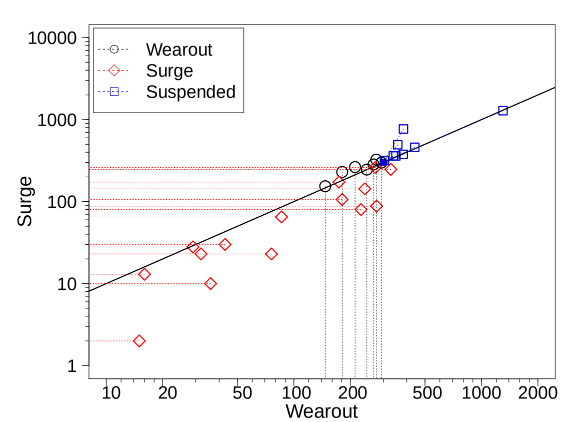
Visualizing that Multiple Failure Mode Data Have Little or No Information About Dependency

Device-G Pseudo Failure-Time Joint Distributions

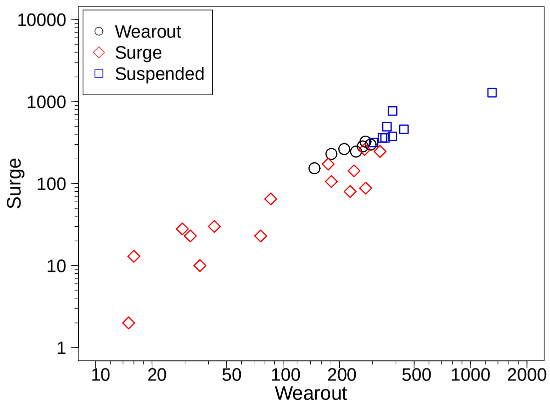
More About the Non-Identifiability of the Dependence Structure

- **Result:** Given a joint failure-time distribution of J competing risks (failure modes) with an arbitrary dependence structure, there exists a corresponding model with independent risks that have the same J subdistribution functions. (Tsiatis, 1975)
- The result implies that given usual multiple failure mode data, one cannot identify the competing-risk dependence structure.
- Any given multiple failure mode data set could have arisen from a distribution that has independent risks or dependent risks.
- These points can be illustrated with pseudo-data joint distributions derived from the Device-G data.

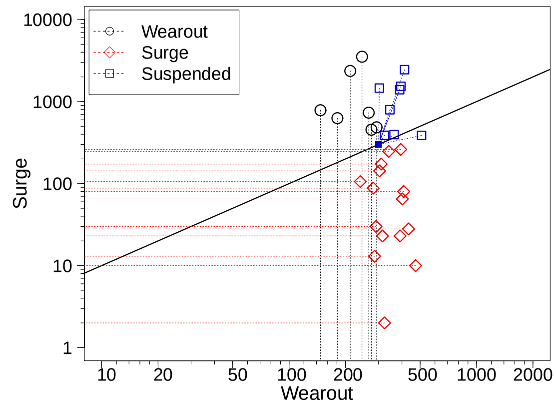
Device-G Dependent Pseudo Data



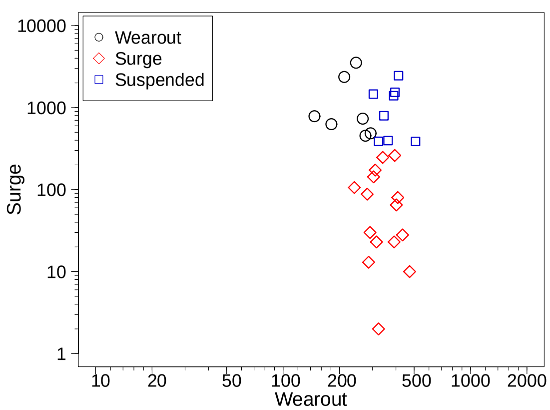
Device-G Dependent Pseudo Data



Device-G Independent Pseudo Data



Device-G Independent Pseudo Data



Chapter 16

Analysis of Data with More than One Failure Mode

Segment 7

Other Topics Related to Multiple Failure Modes

Estimation When Failure Mode is Identified for Only Some Failures

When failure modes are not identified or are only partially identified for some units, it is still possible to estimate the individual $F_i(t)$ distributions by using maximum likelihood.

- Known as **masking** of failure modes.
- More difficult because the analysis is not separable.
- Parameter estimates for the distribution for one mode will be correlated with those of the other modes.
- In practice, one is likely to analyze the data as if there were only a single mode. Potentially misleading if extrapolating outside the range of the data.

Effect of Dependency Among Failure Modes

- The common assumption of independent failure modes is sometimes unrealistic.

  When there is dependence, still one can use the relationship

  $$F_T(t) = \Pr(T \le t) = 1 - \Pr(T_1 > t \cap T_2 > t)$$

  but the evaluation has to be done with respect to the bivariate joint distribution of T_1 and T_2.
- Usually, when there is dependence, the dependence is positive. Then long (short) failure times of one mode tend to go with long (short) failure times of another.
- If the failure modes are positively dependent, then attempting to predict the effect of eliminating one of the failure modes, using the independent failure mode model, can give incorrect and overly optimistic predictions.

Modeling Dependency Among Failure Modes

- Dependent multiple failure modes can be modeled by specifying marginal distributions for the different failures modes along with a "copula" function to describe the dependence.

- There is no information in the data to help specify the form of the copula.

- A combination of external information and sensitivity analysis can be used to help specify a copula.

- Bayesian methods would be useful to describe uncertainty in copula parameters.

References

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

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceeding of the National Academy of Sciences of the United States of America 72*, 20–22. []