Chapter 9

Parametric Bootstrap and Other Simulation-Based Confidence Interval Methods

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 55min

Chapter 9
Parametric Bootstrap and Other Simulation-Based Statistical Methods

Topics discussed in this chapter are:

- The basic concepts of using simulation and parametric bootstrap methods to obtain confidence intervals.
- Methods for generating parametric bootstrap samples and obtaining bootstrap estimates.
- How to obtain parametric or nonparametric confidence intervals from bootstrap samples.
- How to obtain parametric confidence intervals by using the simulated distribution of a pivotal quantity.
- How to obtain parametric confidence intervals by using the simulated distribution of a generalized pivotal quantity.

Chapter 9

Segment 1

Motivation Basic Ideas

Motivation

- Provide methods for constructing approximate or exact confidence intervals for distribution parameters and functions of distribution parameters like quantiles and probabilities.

  These methods usually outperform the **Wald** procedures.

- Wald confidence interval procedures may be adequate for initial informal analyses, particularly when the sample size (number of failures) is large.

- In general, likelihood-based methods for constructing confidence intervals outperform the Wald methods.

- Bootstrap methods provide useful alternatives to Wald and likelihood-based methods and may yield more accurate approximate confidence interval procedures.

- Sometimes bootstrap confidence intervals are used when there are not reasonable alternatives (e.g., likelihood-based confidence intervals are too demanding computationally).

Basic Ideas

- Replace mathematical approximations or intractable distribution theory with Monte Carlo simulation.
- For example, instead of assuming
$$Z_{\widehat{\mu}} = \frac{\widehat{\mu} - \mu}{\mathsf{se}_{\widehat{\mu}}} \,\dot{\sim}\, \mathsf{NORM}(0,1),$$
use a bootstrap approach to simulate $B = 100{,}000$ values of
$$Z_{\widehat{\mu}^*} = \frac{\widehat{\mu}^* - \widehat{\mu}}{\mathsf{se}_{\widehat{\mu}^*}}.$$
This provides an improved approximation to the **actual** distribution of $Z_{\widehat{\mu}}$ and a better confidence interval procedure for μ–especially with small data sets.
- Bootstrap methods provide **exact** distributions of pivotal quantities and generalized-pivotal quantities needed to obtain confidence intervals, sometimes leading to **exact** interval procedures.

General Concepts on Confidence Intervals

- An important criterion for judging an approximate procedure for constructing a statistical interval is how well the procedure would perform if it were repeated over and over.

- The coverage probability should be equal or close to the chosen nominal confidence level $100(1-\alpha)\%$.

- Prefer two-sided intervals for which the error probability α is split equally or approximately equally between the upper and lower interval bound (i.e., close to $\alpha/2$ for each side of the interval).

- In practice, one cannot actually repeat the sampling process over and over. Then simulate the sampling process to create bootstrap samples.

  The empirical sampling distribution of the appropriate statistics from the resulting bootstrap samples is used to compute the desired statistical interval, reducing the reliance on sometimes crude large-sample approximations.

Confidence Intervals and Bootstrap Interval Procedures

- Statistical intervals are computed as functions of the available data, consisting of n observations denoted by DATA.
- Bootstrap interval procedures employ, in addition, a set of B bootstrap samples, DATA_j^*, $j = 1, \ldots, B$, generated by Monte Carlo simulation that, in some sense, mimic the original sampling procedure.
- For each of the B bootstrap samples, one or more bootstrap statistics are computed.
- The bootstrap statistics are used to compute confidence intervals. There are several competing procedures to do this.

Chapter 9

Segment 2

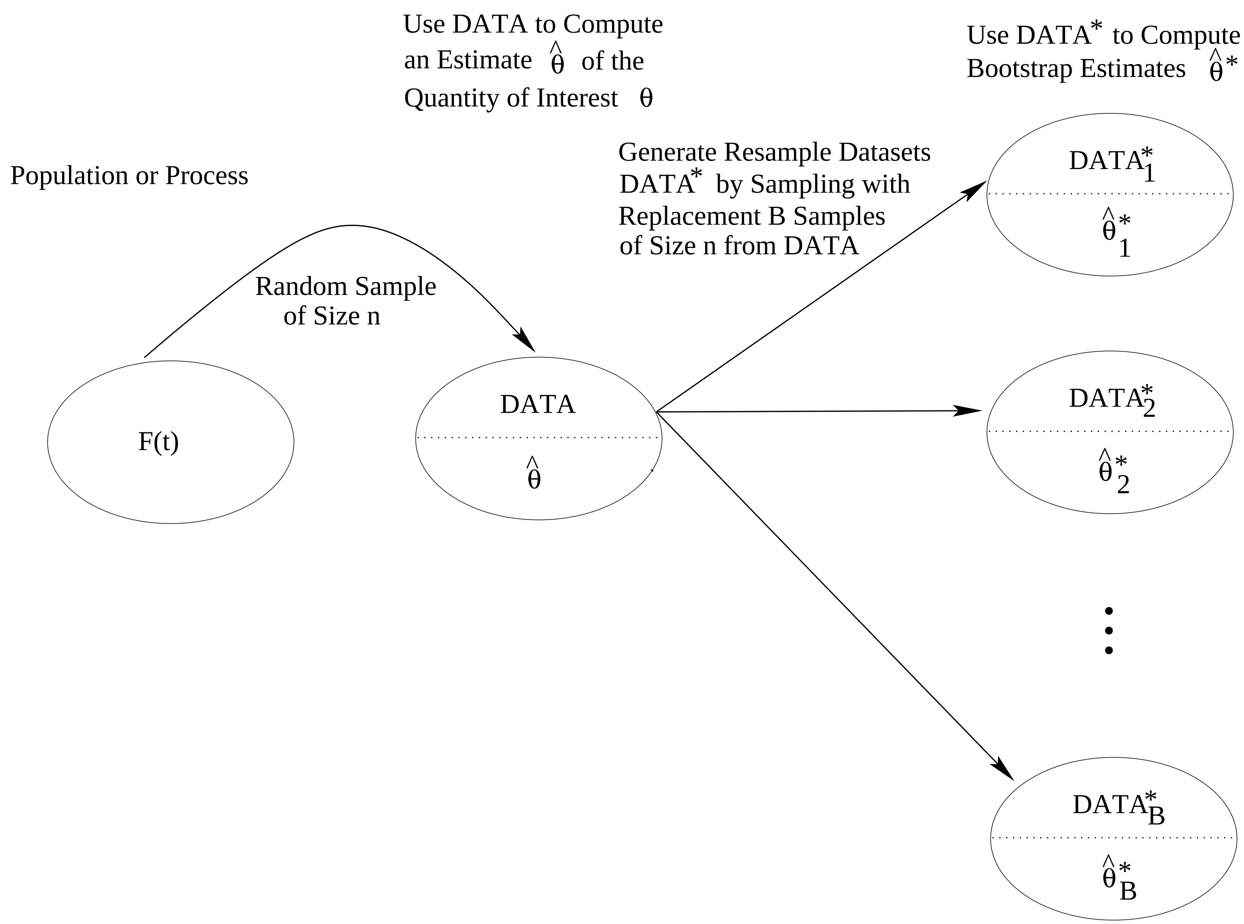
Methods for Generating Bootstrap Samples and Obtaining Bootstrap Estimates

Methods for Generating Bootstrap Samples and Obtaining Bootstrap Estimates

We use three main methods for generating bootstrap samples DATA_j^* and bootstrap estimates $\widehat{\theta}^*$.

- Nonparametric bootstrap **resampling**.
- **Fractional**-random-weight bootstrap sampling.
- Parametric bootstrap sampling.

Nonparametric Bootstrap Resampling for Obtaining Bootstrap Samples DATA$_j^*$ and Bootstrap Estimates $\widehat{\theta}_j^*$



Bootstrap Resampling

- In this method, a point estimate, $\widehat{\theta}$, of the scalar quantity of interest θ (or a particular function of interest computed from θ) is obtained initially directly from the data.
- Then B bootstrap samples (also called **resamples**), each of size n, are obtained by sampling, **with replacement**, from the n cases in the given data set.
- To obtain the j^{th} bootstrap sample DATA_j^*, we select with replacement a sample of size n from the n original observations in DATA.
- Each observation in DATA has an equal probability of being chosen on each draw.

Bootstrap Resampling (Continued)

- Because of the sampling with replacement, some observations in the original DATA may be selected more than once and others not at all in a single bootstrap sample (as will be illustrated later).
- For each bootstrap resample, an estimate of the desired distribution characteristic (or characteristics) of interest θ (or a function of interest related to θ) is computed from the n resample values, giving $\widehat{\theta}^*_j$, $j = 1, \ldots, B$.
- The resulting B values of $\widehat{\theta}^*$ can then be used to compute the desired statistical interval or intervals as described later.

Comments on the Bootstrap Resampling

- The **bootstrap resampling method** can also be viewed as a random-weight method of sampling where n integer weights $(\omega_1, \ldots, \omega_n)$, one for each observation in the data set, are a sample from an n-cell multinomial distribution with equal probability $1/n$ for each of the n cells.
- Some of the original observations will be resampled more than once (and thus have integer weights greater than 1) and others will be not be sampled at all (and thus will have weight 0).
- A **potential problem** with some of the bootstrap samples: Inability to estimate the quantity of interest (in the case of nonparametric bootstrap) or all of the model parameters (in the case of parametric bootstrap), even if the original data are able to do so.

Some Situations that May Lead to Estimability Issues with Some of the Bootstrap Samples

- When the data are censored and there are only a limited number of noncensored observations. In such cases, it is possible to obtain resamples with all observations censored.

- Even when data are not censored, we have encountered applications with small to moderate sample sizes where the resampling method resulted in noticeable instability in estimating parameters from resampled data.

Fractional-Random-Weight Bootstrap Sampling

- Random-weight bootstrap sampling is an appealing alternative to resampling that can be applied when the estimation method allows noninteger weights (e.g., ML or LS).
- Nonnegative weights can be generated from a **continuous** distribution of a positive random variable that has the same mean and standard deviation as the integer resampling weights (usually taken to be equal to 1).
- Weights generated independently from an exponential distribution with mean 1 is a common choice.
- Another alternative is to generate the weights from a uniform Dirichlet distribution, which can be achieved by standardizing the independent exponential weights to sum to n.
- In either of these cases, bootstrap estimates are obtained by applying an appropriate weighted estimation method, using the B sets of random weights. Parameters/quantities estimable with the original data, will usually be estimable for each set of random weights.

Comments on the Fractional-Random-Weight Bootstrap Sampling

- This method is nonparametric because generating the random weights does not require any assumptions about the underlying distribution of the data.
- To get the complete set of bootstrap estimates, the n random weights and the computation of the estimates from the weighted-estimation procedure is repeated B times.
- The method can be used in both nonparametric and parametric bootstrap applications.
- When there is little or no risk of estimation problems with resampling, the resampling (integer weight) method and the random continuous weight method will give similar bootstrap results.

Contrasting the Bootstrap Resampling and the Fractional-Random-Weight Bootstrap Sampling

- The important difference between the integer and continuous weight methods of generating bootstrap samples is that some of the integer weights are 0, indicating that the associated observations are completely ignored in computing the bootstrap statistics using this method.
- In contrast, when the continuous weights are used, each of the original observations has a contribution to the computation of the likelihood and resulting bootstrap estimates.
- Note that the data are constant and the random weights induce randomness in the computed bootstrap estimates.

The Fractional-Random-Weight Bootstrap and ML Estimation

- The fractional-random-weight bootstrap method should be used in situations with heavy censoring, complicated data and/or a complicated parametric model, and when maximum likelihood estimation is used.

- For maximum likelihood estimation, the weighted likelihood is

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}; \mathsf{DATA}) = \mathcal{C} \prod_{i=1}^{n} [L_i(\boldsymbol{\theta}; \mathsf{data}_i)]^{\omega_i}$$

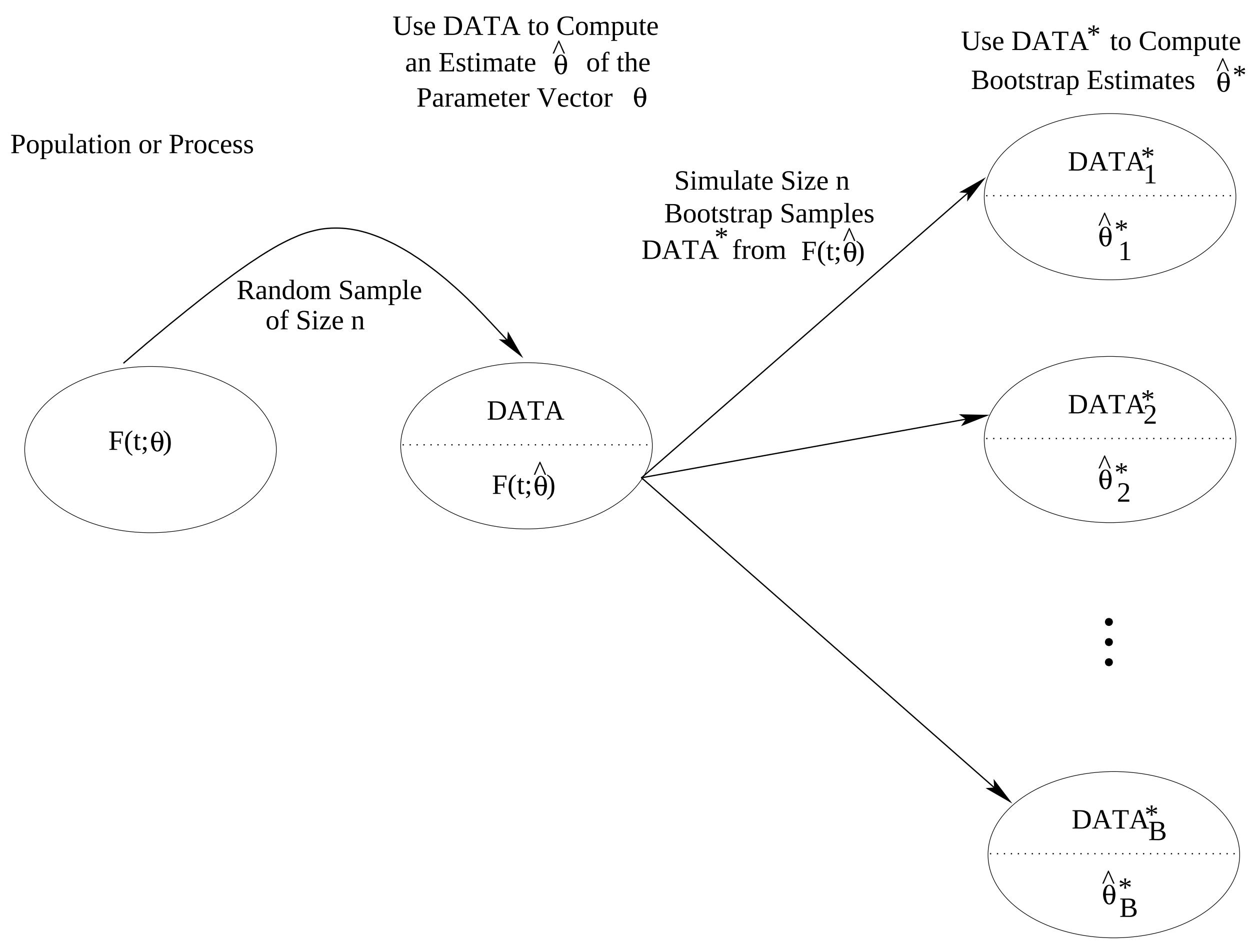
where $L_i(\boldsymbol{\theta}; \mathsf{data}_i)$ is the likelihood contribution for observation i.

- This weighted likelihood is maximized just like a regular likelihood to obtain bootstrap estimates $\widehat{\boldsymbol{\theta}}^*$ of the model parameters in $\boldsymbol{\theta}$.

Examples of Integer-Random-Weight and Fractional-Random-Weight Bootstrap Sampling From the Shock-Absorber Failure Data

			Combined uniform multinomial distribution integer random weights			Combined uniform Dirichlet distribution fractional random weights		
Kilometers	Status	Weight	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$
6700	Failed	1	0	1	3	0.32	0.85	3.73
6950	Censored	1	1	1	0	0.60	0.95	0.78
7820	Censored	1	2	0	2	2.78	0.95	1.16
8790	Censored	1	0	2	1	0.90	1.78	0.44
9120	Failed	1	2	1	1	0.22	0.43	2.24
9660	Censored	1	1	0	0	1.27	0.34	0.89
9820	Censored	1	0	1	0	1.43	1.86	0.81
11310	Censored	1	1	3	0	0.64	0.01	0.59
11690	Censored	1	1	1	1	0.04	1.60	0.75
11850	Censored	1	1	1	1	1.04	1.58	2.29
11880	Censored	1	2	0	0	1.38	0.83	0.15
...								
19410	Censored	1	1	0	0	0.34	1.19	0.46
20100	Failed	1	1	1	1	0.11	0.34	0.53
20100	Censored	1	2	0	1	0.41	0.57	0.38
20150	Censored	1	3	1	0	0.10	2.03	0.80
20320	Censored	1	1	0	0	2.49	0.31	2.01
20900	Failed	1	1	1	0	2.72	0.13	1.73
22700	Failed	1	3	0	4	1.63	0.77	0.84
23490	Censored	1	0	1	0	0.44	3.01	0.23
26510	Failed	1	0	1	1	1.41	0.79	0.54
27410	Censored	1	0	0	0	0.76	2.74	0.36
27490	Failed	1	0	3	0	0.00	1.04	0.10
27890	Censored	1	0	2	1	1.96	0.62	0.22
28100	Censored	1	0	2	2	1.35	0.29	0.31
Sum:		38	38	38	38	38.00	38.00	38.00
Number of failures:		11	10	11	17	11.45	6.98	14.76

Parametric Bootstrap Sampling for Obtaining Bootstrap Samples DATA$_j^*$ and Bootstrap Estimates $\widehat{\theta}_j^*$



Parametric Bootstrap Sampling and Bootstrap Estimates

This method is useful when there is no censoring; or the censoring is easy to simulate (e.g., simple time or failure censoring commonly used in life testing).

- First, use the n data cases to compute the ML estimate $\widehat{\boldsymbol{\theta}}$ and and estimate $F(t; \widehat{\boldsymbol{\theta}})$.
- Then B bootstrap samples of size n are simulated from $F(t; \widehat{\boldsymbol{\theta}})$ and these are denoted by DATA_j^*, $j = 1, \ldots, B$.
- For each of these B samples, obtain the ML bootstrap estimate of the parameter vector, $\widehat{\boldsymbol{\theta}}_j^*$. Similarly, for a quantity of interest $g(\boldsymbol{\theta})$, obtain the bootstrap estimates $g(\widehat{\boldsymbol{\theta}}_j^*)$, $j = 1, \ldots, B$.
- The values of $g(\widehat{\boldsymbol{\theta}}_j^*)$ or $\widehat{\boldsymbol{\theta}}_j^*$ can be used, in a variety of ways, to construct parametric bootstrap statistical intervals.

Using this method will, in some cases, lead to statistical interval procedures that are exact.

Which Bootstrap Sampling Method to Use?

- With complete data or simple censoring (failure or time censoring), it is straightforward to simulate bootstrap samples, that mimic the original data, from the fitted distribution. Then parametric bootstrap sampling is commonly used in these cases.

  Generating samples in this way can lead to exact confidence interval procedures.

- When the censoring is complicated, it is difficult to model the censoring in the simulation. In this case, either the resampling method or the fractional-random-weight sampling method would be easier to use and will provide an excellent approximation of the parametric method.

- In the presence of heavy censoring, the resampling method can perform poorly or fail altogether. In such cases, the fractional-random-weight bootstrap method should be used.

  We will use the fractional-random-weight bootstrap method with the shock-absorber failure data.

Chapter 9

Segment 3

Bootstrap Confidence Interval Methods

Bootstrap Confidence Interval Methods

- This presentation deals primarily with two-sided confidence intervals. A one-sided lower (or upper) confidence bound is obtained from the corresponding two-sided interval by substituting α for $\alpha/2$ in the expression for the lower (or upper) endpoint of the two-sided interval.
- A common approach for constructing parametric or non-parametric bootstrap confidence intervals for a quantity of interest is to use **appropriate** quantiles of the empirical bootstrap distribution of that quantity. There are a number of ways to select such quantiles. We present:
  - The **simple percentile method**.
  - The **bias-corrected (BC) percentile** method.
- The **bootstrap-t** method, based on the idea of an approximate pivotal quantity.
- **Pivotal** and **generalized pivotal** quantity methods.

Quantile of an Empirical Distribution

- Bootstrap inference usually requires one to calculate quantiles of the empirical distribution of bootstrap estimates of a function of interest.

- When the bootstrap estimates of a scalar function $g(\boldsymbol{\theta})$ are

$$g(\widehat{\boldsymbol{\theta}}_1^*), \ldots, g(\widehat{\boldsymbol{\theta}}_B^*),$$

  a common definition of the p quantile of their empirical distribution is the the k^{th} order statistic of the $g(\widehat{\boldsymbol{\theta}}_1^*)$, where

$$k = \begin{cases} pB & \text{if } pB \text{ is an integer} \\ \lfloor pB \rfloor + 1 & \text{if } pB \text{ is not an integer} \end{cases}$$

  and $\lfloor pB \rfloor$ denotes the integer part of pB.

- There are alternative definitions for the p quantile (e.g., rounding to the nearest integer).

  When B is large (as in our examples), the differences in the results obtained among the alternative definitions tend to be small.

The Simple Percentile Method

- The simple percentile method uses the $\alpha/2$ and $1-\alpha/2$ quantiles of the empirical bootstrap distribution of the estimates of the quantity of interest as the endpoints of the desired confidence interval. That is,

$$\left[\underset{\sim}{\theta},\ \widetilde{\theta}\right] = \left[\widehat{\theta}^*_{(\alpha/2)},\ \widehat{\theta}^*_{(1-\alpha/2)}\right],$$

  where $\widehat{\theta}^*_{(p)}$ is the p quantile of the empirical distribution of bootstrap estimates for θ, the quantity of interest.

- For example, a confidence interval for the mean of a distribution uses the $\alpha/2$ and $1-\alpha/2$ quantiles of the empirical distribution of bootstrap sample means as the end points of the desired confidence interval.

Example: The Simple Percentile Bootstrap Confidence Interval for the Weibull Shape Parameter of Shock-Absorber Life

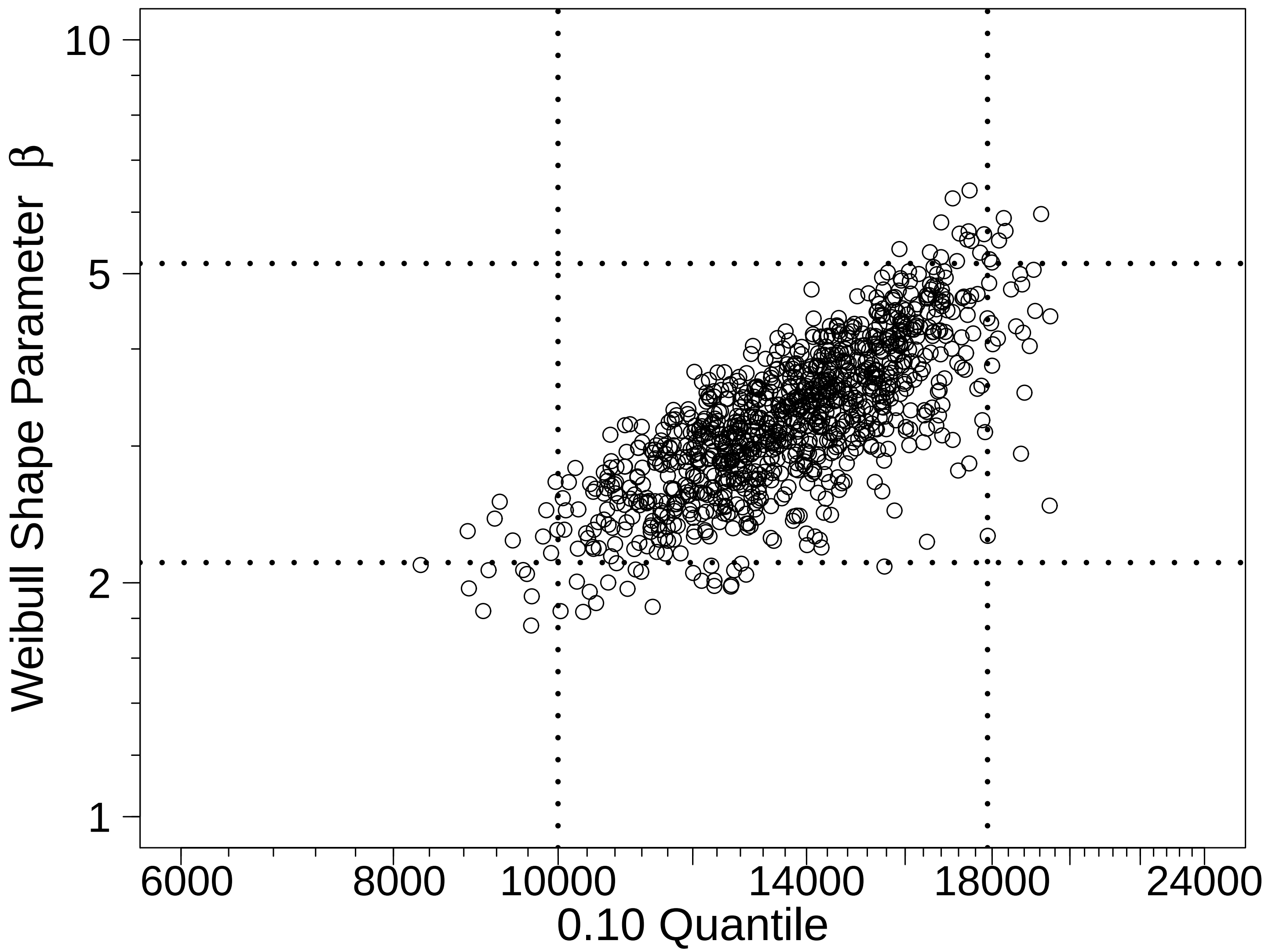
- To construct the simple percentile 95% confidence interval for the shock-absorber Weibull shape parameter, take the 0.025 and 0.975 quantiles of the empirical bootstrap distribution of the $\widehat{\beta}^{*}$ values as the lower and upper confidence bounds.

- The corresponding quantiles are: The 2,500$^{\text{th}}$ (i.e., $0.025 \times 100{,}000$) and 97,500$^{\text{th}}$ (i.e., $0.975 \times 100{,}000$) ordered observations.

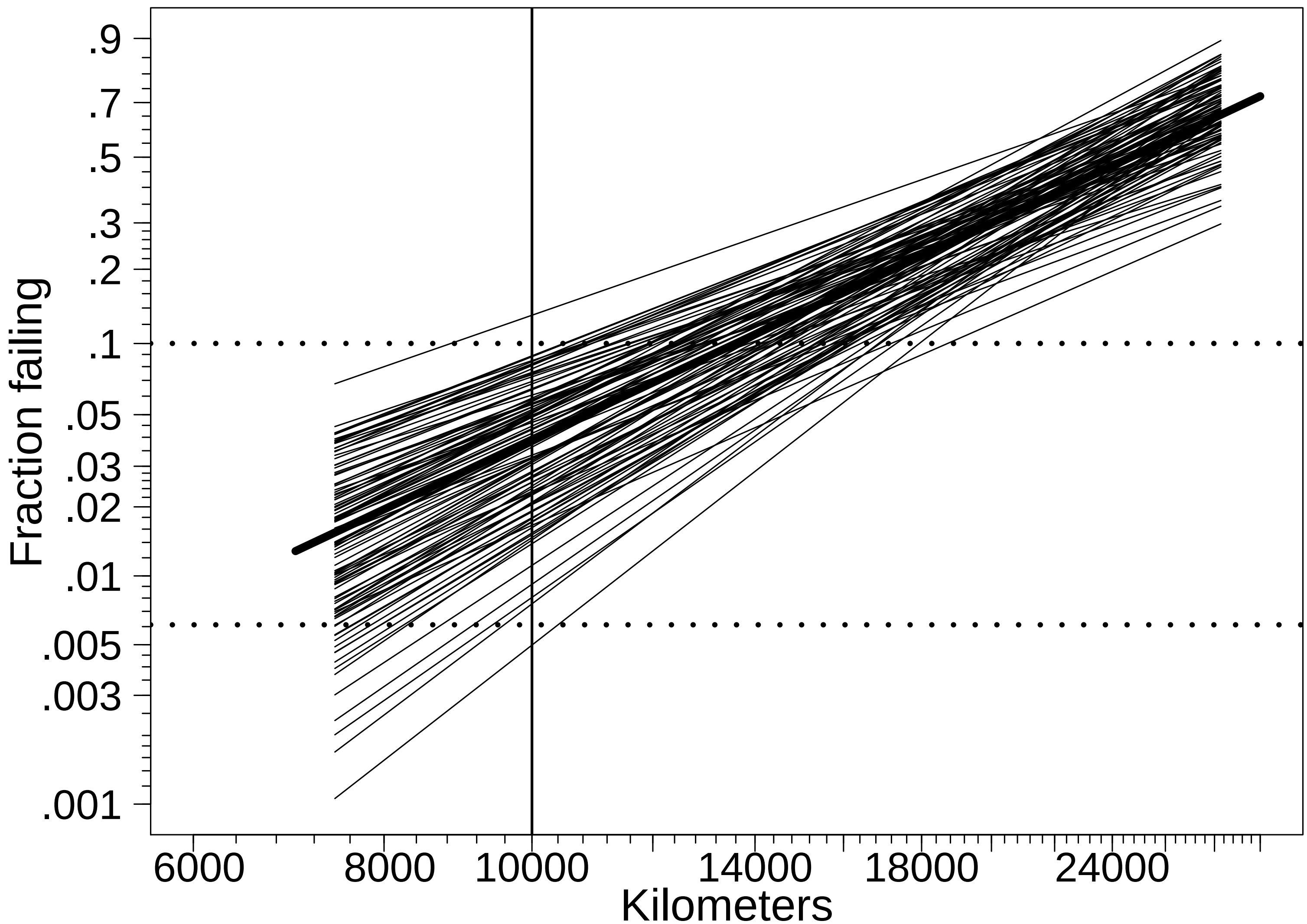
  This results in the 95% confidence interval

$$[\widehat{\beta}^{*}_{(0.025)}, \;\; \widehat{\beta}^{*}_{(0.975)}] = [2.12, \;\; 5.15].$$

Scatterplot of the First 1000 Bootstrap Samples $\widehat{\beta}^*$ and $\widehat{t}^*_{0.10}$ of β and $t_{0.10}$ for the Shock-Absorber Data with the 95% Simple Percentile Confidence Intervals for β and $t_{0.10}$



First 50 ML Bootstrap Estimates of $F(t)$. The Dotted Lines Indicate the Simple Percentile Bootstrap Method Confidence Interval for $F(10000)$.



The Simple Percentile Bootstrap Confidence Interval for the Weibull 0.10 Quantile of Shock-Absorber Life

- Using $\widehat{\mu} = 10.2299$ and $\widehat{\sigma} = 0.316409$, the ML estimate for $t_{0.10}$ is $\widehat{t}_{0.10} = \exp(\widehat{y}_{0.1}) = \exp(10.2299 - 2.2504 \times 0.316409) = 13600$.

- Compute the bootstrap sample estimates for $t_{0.10}$ for each of the 100,000 bootstrap samples. That is

$$\widehat{t}^{*}_{0.10_j} = \exp[\widehat{\mu}^{*}_{j} + \Phi^{-1}_{\text{sev}}(p)\, \widehat{\sigma}^{*}_{j}], \quad j = 1, \ldots, B.$$

- The 95% confidence interval for $t_{0.10}$, the shock-absorber Weibull 0.10 quantile, is obtained from the 2,500$^{\text{th}}$ and 97,500$^{\text{th}}$ ordered values of the empirical distribution of $\widehat{t}^{*}_{0.10}$. The interval is $[\widehat{t}^{*}_{0.10_{(0.025)}}, \ \widehat{t}^{*}_{0.10_{(0.975)}}] = [9998, \ 17889]$.

  Using `R` as a calculator gives

```
> library(StatInt)
> B10BootSamples <- exp(ShockAbsorberWeibullBootSamples[, "location"] +
+ ShockAbsorberWeibullBootSamples[, "scale"]*qsev(0.10))
> quantile(B10BootSamples[-1], p=c(0.025, 0.975))
     2.5%     97.5%
 9998.116 17888.869
```

The BC Percentile Method

- The bias-corrected percentile (BC) method can provide an improvement, relative to the simple percentile method.

- The BC method uses the adjusted quantiles given by

$$\alpha_1 = \Phi_{\text{norm}}\left[2z_{(\widehat{b})} - z_{(1-\alpha/2)}\right],$$
$$\alpha_2 = \Phi_{\text{norm}}\left[2z_{(\widehat{b})} + z_{(1-\alpha/2)}\right],$$

  where $z_{(\alpha)}$ is the α quantile of the standard normal distribution and

$$\widehat{b} = \text{fraction of the } B \text{ values of } \widehat{\theta}^* \text{ that are less than } \widehat{\theta}.$$

- In the above expressions, $z_{(\widehat{b})}$ is the bias-correction value that corrects for median bias in the distribution of $\widehat{\theta}^*$ (on the standard normal scale).

The BC Percentile Bootstrap Confidence Interval for $t_{0.10}$, the 0.10 Quantile of Shock-Absorber Life

- First, $\widehat{b}$ is computed as $43879/100000 = 0.43879$ and then $z_{(\widehat{b})} = \Phi^{-1}_{\text{norm}}(0.43879) = -0.1540377$, Direct computations give

$$\alpha_1 = \Phi_{\text{norm}}[2 \times (-0.1540377) - 1.959964] = 0.0116634,$$

$$\alpha_2 = \Phi_{\text{norm}}[2 \times (-0.1540377) + 1.959964] = 0.9507214.$$

Then the BC bootstrap approximate 95% confidence interval for $t_{0.10}$ is $[\widehat{t}^*_{0.10_{(0.0116634)}}, \;\; \widehat{t}^*_{0.10_{(0.9507214)}}] = [9400, \;\; 17269]$. Using `R` as a calculator gives

```
> bhat <- sum(B10BootSamples[-1] < B10BootSamples[1])/(length(B10BootSamples)-1)
> bhat
[1] 0.43879
> alpha1 <- pnorm(2*qnorm(bhat) - qnorm(0.975))
> alpha2 <- pnorm(2*qnorm(bhat) + qnorm(0.975))
> c(alpha1, alpha2)
[1] 0.0116634 0.9507214
> quantile(B10BootSamples[-1], c(alpha1, alpha2))
 1.16634% 95.07214%
 9399.664 17269.271
```

Chapter 9

Segment 4

Bootstrap Confidence Intervals Based on Pivotal Quantities

Bootstrap Confidence Intervals Based on Pivotal Quantities

- For data $\boldsymbol{X}$ with joint density $f(\boldsymbol{x};\boldsymbol{\theta})$, the scalar function $g(\boldsymbol{X},\boldsymbol{\theta})$ is a pivotal quantity (PQ) if the distribution of $g(\boldsymbol{X},\boldsymbol{\theta})$ does not depend on $\boldsymbol{\theta}$.

- If $\bar{X}$ is the sample mean and S is the sample standard deviation computed from a sample of size n from a NORM(μ,σ) distribution, then $(\mu-\bar{X})/(S/\sqrt{n})$ has a t-distribution with $n-1$ degrees of freedom. Then,
$$\Pr\left[t_{(\alpha/2;n-1)} \le \frac{\mu-\bar{X}}{S/\sqrt{n}} \le t_{(1-\alpha/2;n-1)}\right] = 1-\alpha.$$

- **Pivoting** on μ, we get
$$\Pr\left[\bar{X}+t_{(\alpha/2;n-1)}S/\sqrt{n} \le \mu \le \bar{X}+t_{(1-\alpha/2;n-1)}S/\sqrt{n}\right] = 1-\alpha.$$
This probability statement implies that
$$\left[\underset{\sim}{\mu},\ \ \widetilde{\mu}\right] = \left[\bar{x}+t_{(\alpha/2;n-1)}s/\sqrt{n},\ \ \bar{x}+t_{(1-\alpha/2;n-1)}s/\sqrt{n}\right]$$
is an exact $100(1-\alpha)\%$ confidence interval for μ, where $\bar{x}$ is the observed value of $\bar{X}$ and s is the observed value of S.

Quantiles for the Distributions of Some PQs in Samples from (Log-)Location-Scale Distributions

- Pivotal quantities exist for the parameters and quantiles of (log-)location-scale distributions when the data are complete or failure (Type 2) censored.

- For distributions other than the normal (and lognormal) or when censoring is involved, tables or computer functions for the needed quantiles of the distributions of the PQs are generally not available.

  In these cases, one can instead use parametric bootstrap simulation methods, as described next, to obtain the desired distribution quantiles. This yields exact confidence interval procedures.

- When the methods are not exact (e.g., because the data that have other than failure censoring), the coverage probability will generally be close to the nominal confidence level, with the approximation improving with larger sample sizes.

PQ Based Confidence Intervals for the Location Parameter of a Location-Scale Distribution or the Scale Parameter of a Log-Location-Scale Distribution

- When μ is a location parameter for a location-scale distribution then $\exp(\mu)$ is a scale parameter for a log-location-scale distribution.

- With no censoring or failure (Type 2) censoring with $r \geq 2$ failures, $\mu^{**} = (\mu - \widehat{\mu})/\widehat{\sigma}$ is a PQ. Thus an exact $100(1-\alpha)\%$ confidence interval for μ can be computed as

$$\left[\underset{\sim}{\mu}, \;\; \widetilde{\mu}\right] = \left[\widehat{\mu} + z_{\widehat{\mu}_{(\alpha/2;n,r)}}\widehat{\sigma}, \;\; \widehat{\mu} + z_{\widehat{\mu}_{(1-\alpha/2;n,r)}}\widehat{\sigma}\right],$$

  where $z_{\widehat{\mu}_{(\gamma;n,r)}}$ is the γ quantile of the distribution of μ^{**} for a sample of size n, censored at the point in time where the r^{th} $(2 \leq r \leq n)$ failure occurs.

- The corresponding $100(1-\alpha)\%$ confidence interval for the log-location-scale distribution scale parameter $\eta = \exp(\mu)$ is

$$\left[\underset{\sim}{\eta}, \;\; \widetilde{\eta}\right] = \left[\exp(\underset{\sim}{\mu}), \;\; \exp(\widetilde{\mu})\right].$$

The Distribution of $\mu^{**} = (\mu - \widehat{\mu})/\widehat{\sigma}$

The distribution of μ^{**} (and thus quantiles of the distribution) are obtained by using parametric bootstrap methods as follows.

- Obtain ML estimates $\widehat{\mu}$ and $\widehat{\sigma}$ of the assumed (log-)location-scale distribution parameters μ and σ using the available data.
- Then B simulated samples of size n are generated from the resulting fitted distribution (i.e., from the assumed distribution with parameters $\widehat{\mu}$ and $\widehat{\sigma}$).
- From each of the B samples, obtain bootstrap ML estimates $\widehat{\mu}_j^*$ and $\widehat{\sigma}_j^*$, $j = 1, \ldots, B$ and compute

  $$z_{\widehat{\mu},j}^* = \frac{\widehat{\mu} - \widehat{\mu}_j^*}{\widehat{\sigma}_j^*}.$$

  The desired quantiles of μ^{**} are then obtained from the ordered $z_{\widehat{\mu},j}^*$ values, as done before.

Approximate PQ Bootstrap Confidence Interval for the Shock-Absorber Weibull Distribution Scale Parameter

- Here, we use the bootstrap estimates computed using the fractional-random-weight method. The μ^{**} is only approximately pivotal in this example.

- The ML estimate for the Weibull distribution alternative parameters for the shock-absorber data are $\widehat{\mu} = 9.375$ and $\widehat{\sigma} = 0.491$.

- The parametric bootstrap ML estimates $\widehat{\mu}_j^*$ and $\widehat{\sigma}_j^*$ are used to compute $z^*_{\widehat{\mu},j} = (\widehat{\mu} - \widehat{\mu}_j^*)/\widehat{\sigma}_j^*$, $j = 1, \ldots, 100{,}000$. The quantiles of the distribution of $z^*_{\widehat{\mu}}$ are $z^*_{\widehat{\mu}(0.025)} = -0.68757$ and $z^*_{\widehat{\mu}(0.975)} = 0.622044$.

- Then the 95% confidence intervals for μ and $\eta = \exp(\mu)$ are:

$$
\begin{aligned}
\left[\underset{\sim}{\mu}, \ \widetilde{\mu}\right] &= [10.2299 - 0.68757 \times 0.316409, \ 10.2299 + 0.622044 \times 0.316409] \\
&= [10.0123, \ 10.4267], \\
\left[\underset{\sim}{\eta}, \ \widetilde{\eta}\right] &= \left[\exp(\underset{\sim}{\mu}), \ \exp(\widetilde{\mu})\right] = [\exp(10.0123), \ \exp(10.427)] \\
&= [22299, \ 33748].
\end{aligned}
$$

PQ Based Confidence Intervals for the Scale Parameter of a Location-Scale Distribution or the Shape Parameter of a Log-Location-Scale Distribution

- If σ is a scale parameter for a location-scale distribution, then it is also a shape parameter for a log-location-scale distribution. For the Weibull distribution, $\beta = 1/\sigma$ is more commonly used to represent the distribution shape parameter.

- With no censoring or failure (Type 2) censoring with $r \geq 2$ failures, $Z_{\widehat{\sigma}} = \sigma/\widehat{\sigma}$ is a PQ. Thus an exact $100(1-\alpha)\%$ confidence interval for σ can be computed as

$$\left[\underset{\sim}{\sigma}, \ \widetilde{\sigma}\right] = \left[z_{\widehat{\sigma}_{(\alpha/2)}}\widehat{\sigma}, \ z_{\widehat{\sigma}_{(1-\alpha/2)}}\widehat{\sigma}\right],$$

where $z_{\widehat{\sigma}_{(\gamma)}}$ is the γ quantile of the distribution of $Z_{\widehat{\sigma}}$, based on a sample of size n, censored at the point in time when the r^{th} $(2 \leq r \leq n)$ failure occurs.

- The corresponding, exact $100(1-\alpha)\%$ confidence interval for $\beta = 1/\sigma$ is

$$\left[\underset{\sim}{\beta}, \ \widetilde{\beta}\right] = \left[\frac{1}{\widetilde{\sigma}}, \ \frac{1}{\underset{\sim}{\sigma}}\right].$$

Approximate PQ Bootstrap Confidence Interval for Shock-Absorber Weibull Distribution Shape Parameter

- As done for the scale parameter, the parametric bootstrap estimates $\widehat{\sigma}_j^*$ are used to compute $z_{\widehat{\sigma},j}^* = \widehat{\sigma}/\widehat{\sigma}_j^*$, $j = 1, \ldots, 100{,}000$. The required quantiles of the distribution of $z_{\widehat{\sigma},j}^*$ are $z_{\widehat{\sigma}_{(0.025)}}^* = 0.67197$ and $z_{\widehat{\sigma}_{(0.975)}}^* = 1.63070$.

- The approximate 95% confidence interval for σ is

$$[\underset{\sim}{\sigma},\ \widetilde{\sigma}] = [0.67197 \times 0.316409,\ \ 1.63070 \times 0.316409] = [0.21262,\ \ 0.51597].$$

- The corresponding 95% confidence interval for the Weibull distribution shape parameter β is

$$[\underset{\sim}{\beta},\ \widetilde{\beta}] = \left[\frac{1}{\widetilde{\sigma}},\ \frac{1}{\underset{\sim}{\sigma}}\right] = [1/0.515968,\ \ 1/0.212618] = [1.94,\ \ 4.70].$$

Using `R` as a calculator gives

```
> 1/(quantile(ShockAbsorberWeibullBootSamples[1, "scale"]/
+     ShockAbsorberWeibullBootSamples[-1, "scale"], p=c(0.975, 0.025)) *
+     ShockAbsorberWeibullBootSamples[1, "scale"])
  97.5%    2.5%
1.93810 4.70328
```

PQ Confidence Intervals for the p Quantile of a Location-Scale or a Log-Location-Scale Distribution

- The p quantile of a location-scale distribution $F(y;\mu,\sigma) = \Phi[(y-\mu)/\sigma]$ is $y_p = \mu + \Phi^{-1}(p)\sigma$, and its ML estimator is $\widehat{y}_p = \widehat{\mu} + \Phi^{-1}(p)\widehat{\sigma}$.

- With complete data or failure (Type 2) censoring,

$$Z_{\widehat{y}_p} = \frac{y_p - \widehat{y}_p}{\widehat{\sigma}} = \left[\frac{\mu - \widehat{\mu}}{\widehat{\sigma}} + \left(\frac{\sigma}{\widehat{\sigma}} - 1\right)\Phi^{-1}(p)\right]$$

  is a PQ.

- The parametric ML bootstrap estimates $\widehat{\mu}_j^*$ and $\widehat{\sigma}_j^*$ for $j = 1,\ldots,B$ are used to compute

$$z^*_{\widehat{y}_p,j} = \left[\frac{\widehat{\mu} - \widehat{\mu}_j^*}{\widehat{\sigma}_j^*} + \left(\frac{\widehat{\sigma}}{\widehat{\sigma}_j^*} - 1\right)\Phi^{-1}(p)\right], \quad j = 1,\ldots,B.$$

  The quantiles $z_{\widehat{y}_p(\gamma;n,r)}$ of $Z_{\widehat{y}_p}$ are computed from the ordered values of $z^*_{\widehat{y}_p,j}$, $j = 1,\ldots,B$.

PQ Confidence Intervals for the p Quantile of a Location-Scale or a Log-Location-Scale Distribution (Continued)

- Then

$$\left[\underset{\sim}{y}_p,\ \ \widetilde{y}_p\right] = \left[\widehat{y}_p + z^*_{\widehat{y}_p(\alpha/2)}\widehat{\sigma},\ \ \widehat{y}_p + z^*_{\widehat{y}_p(1-\alpha/2)}\widehat{\sigma}\right]$$

is an approximate $100(1-\alpha)\%$ confidence interval for y_p.

The corresponding approximate $100(1-\alpha)\%$ confidence interval for $t_p = \exp(y_p)$ is

$$\begin{aligned}\left[\underset{\sim}{t}_p,\ \ \widetilde{t}_p\right] &= \left[\exp\left(\underset{\sim}{y}_p\right),\ \ \exp(\widetilde{y}_p)\right] \\ &= \left[\widehat{t}_p \exp\left(z^*_{\widehat{y}_p(\alpha/2)}\widehat{\sigma}\right),\ \ \widehat{t}_p \exp\left(z^*_{\widehat{y}_p(1-\alpha/2)}\widehat{\sigma}\right)\right].\end{aligned}$$

- The intervals above are exact when the data are complete (or Type 2 censored) and the parametric resampling is used to obtain the bootstrap samples.

Confidence Interval for $t_{0.10}$, the Shock-Absorber Lifetime Weibull 0.10 Quantile

- Because of the censoring in the data, we use the bootstrap estimates $\widehat{\mu}_j^*$ and $\widehat{\sigma}_j^*$ computed with the fractional-random-weight method. Thus the quantity $Z_{\widehat{y}_p}$ is only approximately pivotal in this example.

- Using $\widehat{\mu} = 10.2299$ and $\widehat{\sigma} = 0.316409$, the ML estimate for $t_{0.10}$ is $\widehat{t}_{0.10} = \exp(\widehat{y}_{0.1}) = \exp(10.2299 - 2.25037 \times 0.316409) = 13600$.

- The required quantiles of the empirical distribution are: $z^*_{\widehat{y}_{0.10}(0.025)} = -1.298539$ and $z^*_{\widehat{y}_{0.10}(0.975)} = 0.735484$. Then an approximate 95% confidence interval for $t_{0.10}$ is

$$
\begin{aligned}
\left[\underset{\sim}{t}_{0.10},\ \widetilde{t}_{0.10}\right] &= [13600\exp(-1.298539 \times 0.316409),\ \ 13600\exp(0.735484 \times 0.316409)] \\
&= [9018,\ \ 17164].
\end{aligned}
$$

Confidence Interval for $t_{0.10}$, the Shock-Absorber Lifetime Weibull 0.10 Quantile (Continued)

Using R as a calculator gives

```
> library(StatInt)
> the.quantiles <- quantile((ShockAbsorberWeibullBootSamples[1, "location"]-
+ ShockAbsorberWeibullBootSamples[-1, "location"])/
+ ShockAbsorberWeibullBootSamples[-1, "scale"] +
+ ((ShockAbsorberWeibullBootSamples[1, "scale"]/
+ ShockAbsorberWeibullBootSamples[-1, "scale"])-1)*qsev(0.10), p=c(0.025, 0.975))
> the.quantiles
     2.5%     97.5%
-1.298539  0.735484
> ShockAbsorberWeibullBootSamples[1, "location"] +
+          qsev(0.10)*ShockAbsorberWeibullBootSamples[1, "scale"] +
+          the.quantiles*ShockAbsorberWeibullBootSamples[1, "scale"]
   2.5%   97.5%
9.10696 9.75054
> exp(ShockAbsorberWeibullBootSamples[1, "location"] +
+          qsev(0.10)*ShockAbsorberWeibullBootSamples[1, "scale"] +
+          the.quantiles*ShockAbsorberWeibullBootSamples[1, "scale"])
    2.5%    97.5%
 9017.83 17163.51
```

Chapter 9

Segment 5

Confidence Intervals Based on Generalized Pivotal Quantities

Confidence Intervals Based on Generalized Pivotal Quantities (GPQs)

- Pivotal quantities are not available for all inferences of interest. In absence of a PQ, there may be a GPQ that can be used to construct a confidence interval for a function of parameters of interest.

- A GPQ is similar to a PQ in that it is a scalar function of the random parameter estimator or estimators (e.g., $\widehat{\mu}$ and $\widehat{\sigma}$) and the parameters to be estimated (e.g., μ and σ).

- A GPQ differs from a PQ in that the **unconditional** sampling distribution of the GPQ (i.e., the distribution that includes variability from repeated sampling) may depend on the unknown parameters (e.g., μ and σ).

Characterizing a GPQ

To be a GPQ, a function must have the following two properties:

- Conditional on the data (or on the observed value(s) of the parameter estimates calculated from the data, such as $\widehat{\mu}$ and $\widehat{\sigma}$ for a location-scale distribution), the distribution of a GPQ does not depend on any unknown parameters.
- If the random bootstrap parameter estimators (e.g., $\widehat{\mu}^*$ and $\widehat{\sigma}^*$) in a GPQ are replaced by the corresponding observed values of the parameter estimates (e.g., $\widehat{\mu}$ and $\widehat{\sigma}$), the GPQ must be equal to the actual value of the function of the parameters that is being estimated.

Properties of Intervals Based on GPQs

- Use of a GPQ-based procedure will, in general, lead to only an approximate confidence interval.

  GPQ methods tend to provide procedures with a coverage probability that is close to the nominal confidence level.

  It is possible to identify conditions under which a GPQ-based confidence interval procedure is exact.

- We illustrate that the computation of a GPQ-based confidence interval is similar to but simpler than the computation of a PQ interval.

- To obtain a $100(1-\alpha)\%$ GPQ confidence interval, one simulates a large number B of realizations of the GPQ. As with the simple percentile method, the GPQ-based confidence interval, is obtained from the $\alpha/2$ and $1-\alpha/2$ quantiles of the empirical GPQ distribution.

GPQs for μ and σ of a Location-Scale Distribution and for Functions of μ and σ

- There are GPQs for the location parameter μ and the scale parameter σ of a location-scale distribution.
- The GPQs for μ and σ could be used to compute confidence intervals for μ and σ. Such intervals, would **agree** with those obtained from the simpler pivotal quantity methods for these parameters.
- The GPQs for μ and σ can be used to obtain GPQs and corresponding confidence intervals for functions of μ and σ for which no PQ exists.
- Even if a PQ exists, one advantage of using the GPQ method instead is that once the empirical distribution of the GPQ distribution has been computed, the confidence interval procedure is similar to that used in the **simple percentile** method described earlier.

GPQs for the Location Parameter μ the Scale Parameter σ and Functions of μ and σ

- Easy to verify that a GPQ for μ is

$$\mu^{**} = \mu^{**}(\mu, \widehat{\mu}, \widehat{\sigma}, \widehat{\mu}^{*}, \widehat{\sigma}^{*}) = \widehat{\mu} + \left(\frac{\mu - \widehat{\mu}^{*}}{\widehat{\sigma}^{*}}\right)\widehat{\sigma}$$

- Easy to verify that a GPQ for σ is

$$\sigma^{**} = \sigma^{**}(\sigma, \widehat{\sigma}, \widehat{\sigma}^{*}) = \left(\frac{\sigma}{\widehat{\sigma}^{*}}\right)\widehat{\sigma}$$

- A GPQ for a function of interest $g(\mu, \sigma)$ is obtained by substituting the GPQs for μ and σ into the function $g(\mu, \sigma)$.

- Compute GPQ draws by substituting ML estimates $(\widehat{\mu}, \widehat{\sigma})$ for (μ, σ) and bootstrap estimates $\widehat{\mu}_j^{*}$ and $\widehat{\sigma}_j^{*}$ for $\widehat{\mu}^{*}$ and $\widehat{\sigma}^{*}$ above to get

$$\widehat{\mu}_j^{**} = \widehat{\mu} + \left(\frac{\widehat{\mu} - \widehat{\mu}_j^{*}}{\widehat{\sigma}_j^{*}}\right)\widehat{\sigma}, \quad \widehat{\sigma}_j^{**} = \left(\frac{\widehat{\sigma}}{\widehat{\sigma}_j^{*}}\right)\widehat{\sigma}, \text{ for } j = 1, \ldots, B.$$

Confidence Intervals for Tail Probabilities for Log-Location-Scale Distributions

- A lower tail probability for the log-location-scale distribution is given by

$$p = \Pr(T \leq t) = F(t; \mu, \sigma) = \Phi\left[\frac{\log(t) - \mu}{\sigma}\right].$$

- The ML estimate of p is $\widehat{p} = F(t; \widehat{\mu}, \widehat{\sigma})$, where $\widehat{\mu}$ and $\widehat{\sigma}$ are the ML estimates of the parameters obtained from the data.

  There does not exist a PQ that can be used directly to define a confidence interval procedure for the p.

- There is, however, a GPQ for this purpose. For example, substituting $\widehat{\mu}_j^{**}$ and $\widehat{\sigma}_j^{**}$ into the tail probability expression for $p = F(t; \mu, \sigma)$ and simplifying gives the GPQ for $p = F(t; \mu, \sigma)$,

$$\widehat{F}_j^{**} = \Phi\left[\frac{\log(t) - \widehat{\mu}_j^{**}}{\widehat{\sigma}_j^{**}}\right].$$

Confidence Intervals for Tail Probabilities for Location-Scale and Log-Location-Scale Distributions (Continued)

- The $\alpha/2$ and $1-\alpha/2$ quantiles of the empirical distribution of $\widehat{F}^{**}$ provide the endpoints of a $100(1-\alpha)\%$ confidence interval for $F = F(t)$.

- In this case, unlike the case for GPQs in general, this confidence interval procedure is exact if the data are complete or Type 2 censored and the parametric resampling is used to obtain the bootstrap samples. With Type 1 or multiple censoring, the procedure is approximate.

Confidence Intervals for Shock-Absorber Weibull Distribution Probabilities

- The ML estimate of $p = F(10000) = \Pr(T \le 10000)$ is obtained from the given data by substituting the ML estimates for μ and σ into $F(10000)$, giving

$$\widehat{p} = \widehat{F}(10000) = \Phi_{\text{sev}}\left[\frac{\log(10000) - 10.2299}{0.316409}\right] = 0.0391.$$

- Use the GPQ draws $\widehat{\mu}_j^{**}$ and $\widehat{\sigma}_j^{**}$ for $j = 1, \ldots, 100{,}000$ to obtain an empirical distribution of $\widehat{F}^{**}$.

- Then for a 95% confidence interval for $F(10000)$, the 0.025 and 0.975 quantiles of this distribution gives

$$\left[\underset{\sim}{F}(10000),\ \widetilde{F}(10000)\right] = \left[\widehat{F}^{**}_{(0.025)},\ \widehat{F}^{**}_{(0.975)}\right] = [0.0102,\ 0.122].$$

Confidence Intervals for Shock-Absorber Weibull Distribution Probabilities (Continued)

Using R as a calculator gives

```
> library(StatInt)
> psev((log(10000)-ShockAbsorberWeibullBootSamples[1, "location"])/
+        ShockAbsorberWeibullBootSamples[1, "scale"])
[1] 0.0390841
> mu.gpq <- ShockAbsorberWeibullBootSamples[1, "location"] +
+           ((ShockAbsorberWeibullBootSamples[1, "location"] -
+             ShockAbsorberWeibullBootSamples[-1, "location"])/
+             ShockAbsorberWeibullBootSamples[-1, "scale"]) *
+             ShockAbsorberWeibullBootSamples[1, "scale"]
> sigma.gpq <- (ShockAbsorberWeibullBootSamples[1, "scale"]/
+               ShockAbsorberWeibullBootSamples[-1, "scale"])*
+               ShockAbsorberWeibullBootSamples[1, "scale"]
> quantile(psev((log(10000)- mu.gpq)/sigma.gpq), p=c(0.025, 0.975))
      2.5%      97.5%
0.01015222 0.12182369
```

Confidence Intervals for the Mean of a Log-Location-Scale Distribution

- There are no known exact confidence interval procedures for the mean (expected value) of log-location-scale distributions, such as the lognormal and Weibull.
- Procedures based on GPQs have coverage probabilities that are close to the nominal confidence level.
- The mean of a lognormal distribution is

$$\mathrm{E}(T) = \exp\left(\mu + \sigma^2/2\right),$$

and the mean of a Weibull distribution is

$$\mathrm{E}(T) = \eta\Gamma\left(1 + \frac{1}{\beta}\right) = \exp(\mu)\Gamma(1+\sigma).$$

The GPQs for the Expectations of the Lognormal and Weibull Distributions

- Substituting GPQ draws $\widehat{\mu}_j^{**}$ for μ and $\widehat{\sigma}_j^{**}$ for σ gives

$$\widehat{\mathsf{E}(T)}_j^{**} = \widehat{\mathsf{E}(T)}^{**}(\widehat{\mu}, \widehat{\sigma}, \widehat{\mu}_j^{*}, \widehat{\sigma}_j^{*}) = \exp\left(\widehat{\mu}_j^{**} + \frac{(\widehat{\sigma}_j^{**})^2}{2}\right)$$

  for the lognormal mean and

$$\widehat{\mathsf{E}(T)}_j^{**} = \widehat{\mathsf{E}(T)}^{**}(\widehat{\mu}, \widehat{\sigma}, \widehat{\mu}_j^{*}, \widehat{\sigma}_j^{*}) = \exp(\widehat{\mu}_j^{**})\Gamma\Big(1 + \widehat{\sigma}_j^{**}\Big),$$

  for the Weibull mean.

- These GPQs are not pivotal because they have distributions that depend on the observed parameter estimates $\widehat{\mu}$ and $\widehat{\sigma}$.

The Approximate Confidence Intervals for the Expectations of the Lognormal and Weibull Distributions

- The empirical distributions of the GPQ, from Monte Carlo simulation, can be used to obtain approximate confidence intervals for $\mathsf{E}(T)$ for log-location-scale distributions.

- The $\alpha/2$ and $1-\alpha/2$ quantiles of the empirical distribution of $\widehat{\mathsf{E}(T)}^{**}$ provide the endpoints of the $100(1-\alpha)\%$ confidence interval for $\mathsf{E}(T)$.

- Similar methods can be applied for other log-location-scale distributions.

Confidence Interval for Shock-Absorber Mean Time to Failure Assuming a Weibull Distribution

- Use $\widehat{\mu} = 10.2299$, $\widehat{\sigma} = 0.316409$, and GPQ draws $\widehat{\mu}_j^{**}$ and $\widehat{\sigma}_j^{**}$ for $j = 1, \ldots, 100{,}000$ to obtain an empirical distribution of $\widehat{F}^{**}$.
- The endpoints of a 95% confidence interval for the Weibull distribution mean are given by the 0.025 and 0.975 quantiles of this empirical distribution. This results in

$$\left[\underset{\sim}{\mathsf{E}(T)},\ \widetilde{\mathsf{E}(T)}\right] = \left[\widehat{\mathsf{E}(T)}^{**}_{(0.025)},\ \widehat{\mathsf{E}(T)}^{**}_{(0.975)}\right] = [20{,}130,\ \ 30{,}063].$$

Using R as a calculator,

```
> quantile(exp(mu.gpq)*gamma(1+sigma.gpq), p=c(0.025, 0.975))
    2.5%    97.5%
20129.89 30063.15
```

Approximate 95% Confidence Intervals Shock-Absorber Weibull and Lognormal Mean Time to Failure

Method	95% confidence interval Weibull	95% confidence interval Lognormal
Wald	[20,202, 30,472]	[20,338, 42,203]
Simple percentile bootstrap	[21,029, 32,120]	[22,604, 45,387]
BC percentile bootstrap	[21,180, 32,645]	[23,209, 49,772]
GPQ	[20,130, 30,063]	[21,602, 44,120]

Bootstrap Conclusions

- Bootstrap methods provide easy-to-apply methods to compute trustworthy confidence (and prediction) intervals.
- There are many different methods to choose from.
- Completely nonparametric methods are available, but most reliability applications involve parametric assumptions (such as fitting a Weibull or lognormal distribution).
- The fractional-random-weight method of generating bootstrap estimates eliminates estimability problems that can arise with heavy censoring.
- If a PQ/GPQ method is available to construct confidence intervals, it will generally provide the best coverage properties (if not an exact method).

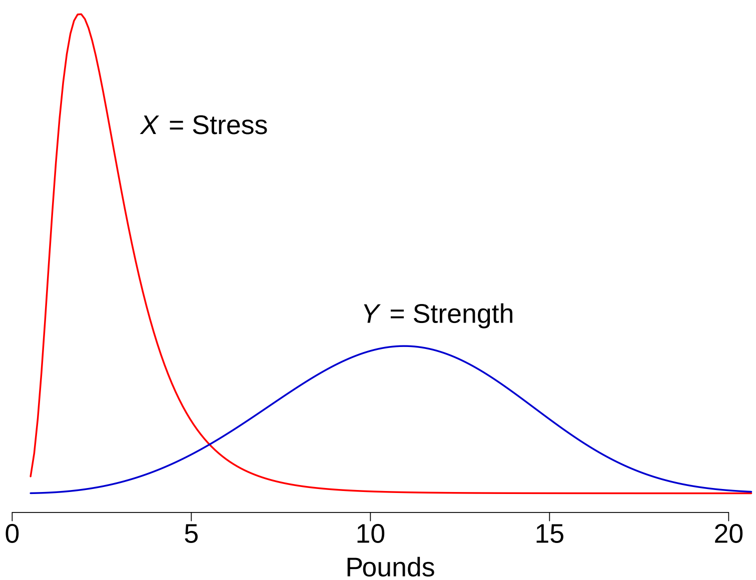
Chapter 9

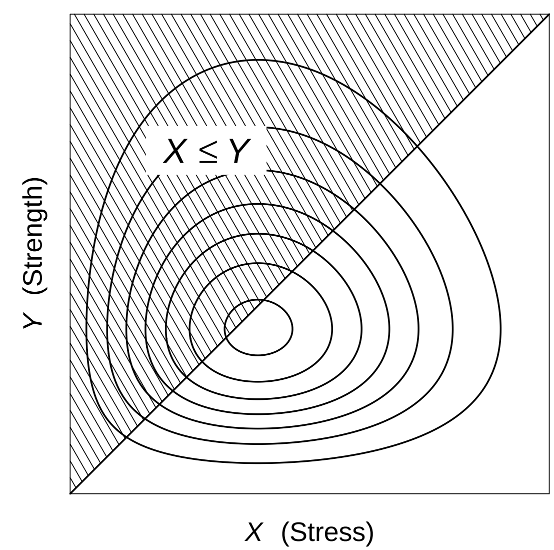
Segment 6

Using the GPQ Method for Stress-Strength Modeling

Stress-Strength Interference General Model

- Let the random variables X and Y denote stress and strength, respectively.
- Let $f(x, y)$ denote the joint density function of X and Y.





- The probability of not failing (a.k.a, reliability) is

$$\Pr(X \le Y) = \int_{-\infty}^{\infty} \int_{x}^{\infty} f(x, y)\, dy dx.$$

Stress-Strength for Independent X and Y

- Let $f_X(x)$ and $f_Y(y)$ denote the density functions of X and Y, respectively.
- The probability of not failing is

$$\Pr(X \le Y) = \int_{-\infty}^{\infty} \int_{x}^{\infty} f_X(x) f_Y(y)\, dy dx$$

$$= \int_{-\infty}^{\infty} f_X(x)[1 - F_Y(x)]\, dx.$$

- When X and Y are positive random variables,

$$\Pr(X \le Y) = \int_{0}^{\infty} f_X(x)[1 - F_Y(x)]\, dx.$$

- When X and Y are log-location-scale random variables,

$$\Pr(X \le Y) = \int_{0}^{\infty} f_X(x; \mu_X, \sigma_X)[1 - F_Y(x; \mu_Y, \sigma_Y)]\, dx.$$

  Estimate $\Pr(X \le Y)$ by evaluating at ML estimates $\widehat{\mu}_X$, $\widehat{\sigma}_X$, $\widehat{\mu}_Y$, and $\widehat{\sigma}_Y$ for the assumed distributions.

GPQ Confidence Interval for $\Pr(X \le Y)$ When X and Y are Log-Location-Scale Random Variables

- Compute the GPQs for μ_X, σ_X, μ_Y, and σ_Y:

$$\widehat{\mu}^{**}_{X_j} = \widehat{\mu}_X + \left(\frac{\widehat{\mu}_X - \widehat{\mu}^{*}_{X_j}}{\widehat{\sigma}^{*}_{X_j}}\right)\widehat{\sigma}_X, \quad \widehat{\sigma}^{**}_{X_j} = \left(\frac{\widehat{\sigma}_X}{\widehat{\sigma}^{*}_{X_j}}\right)\widehat{\sigma}_X$$

$$\widehat{\mu}^{**}_{Y_j} = \widehat{\mu}_Y + \left(\frac{\widehat{\mu}_Y - \widehat{\mu}^{*}_{Y_j}}{\widehat{\sigma}^{*}_{Y_j}}\right)\widehat{\sigma}_Y, \quad \widehat{\sigma}^{**}_{Y_j} = \left(\frac{\widehat{\sigma}_Y}{\widehat{\sigma}^{*}_{Y_j}}\right)\widehat{\sigma}_Y$$

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

- Compute the GPQ values for $\Pr(X \le Y)$:

  $$\widehat{R}^{**}_j = \Pr(X \le Y)^{**}_j = \int_0^{\infty} f_X(x; \widehat{\mu}^{**}_{X_j}, \widehat{\sigma}^{**}_{X_j})[1 - F_Y(x; \widehat{\mu}^{**}_{Y_j}, \widehat{\sigma}^{**}_{Y_j})]\, dx$$

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

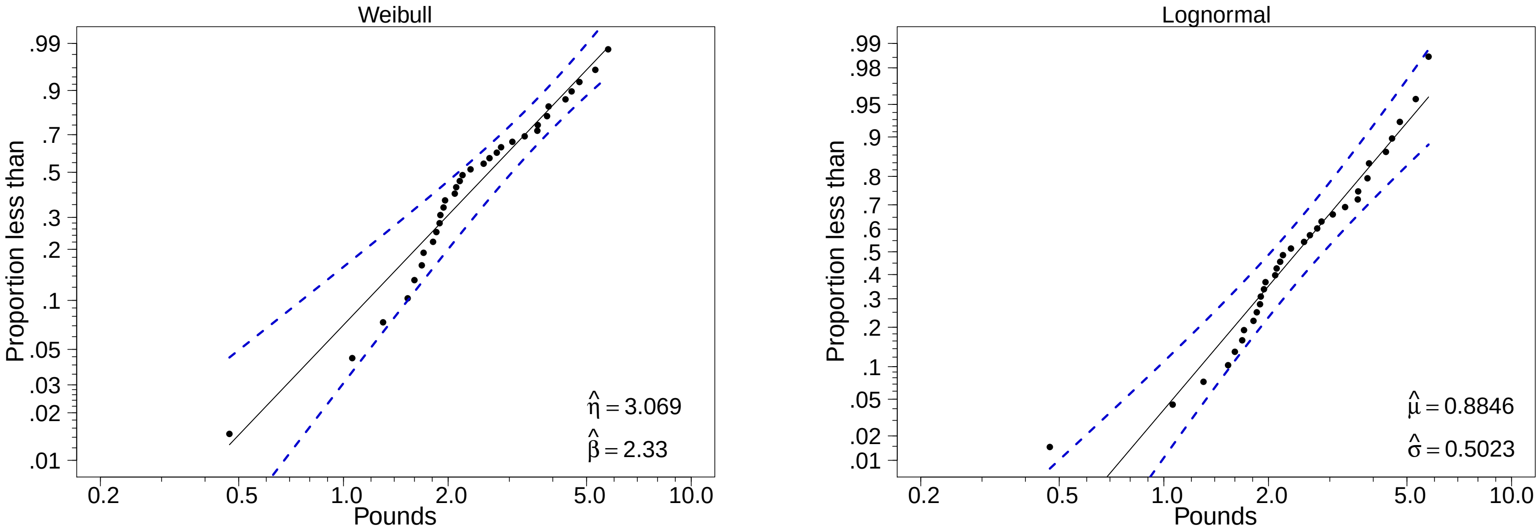
- A $100(1-\alpha)\%$ confidence interval for $\Pr(X \le Y)$ is obtained from the $\alpha/2$ and $(1 - \alpha/2)$ quantiles of the empirical distribution of $\widehat{R}^{**}$.

Example: Connector Reliability

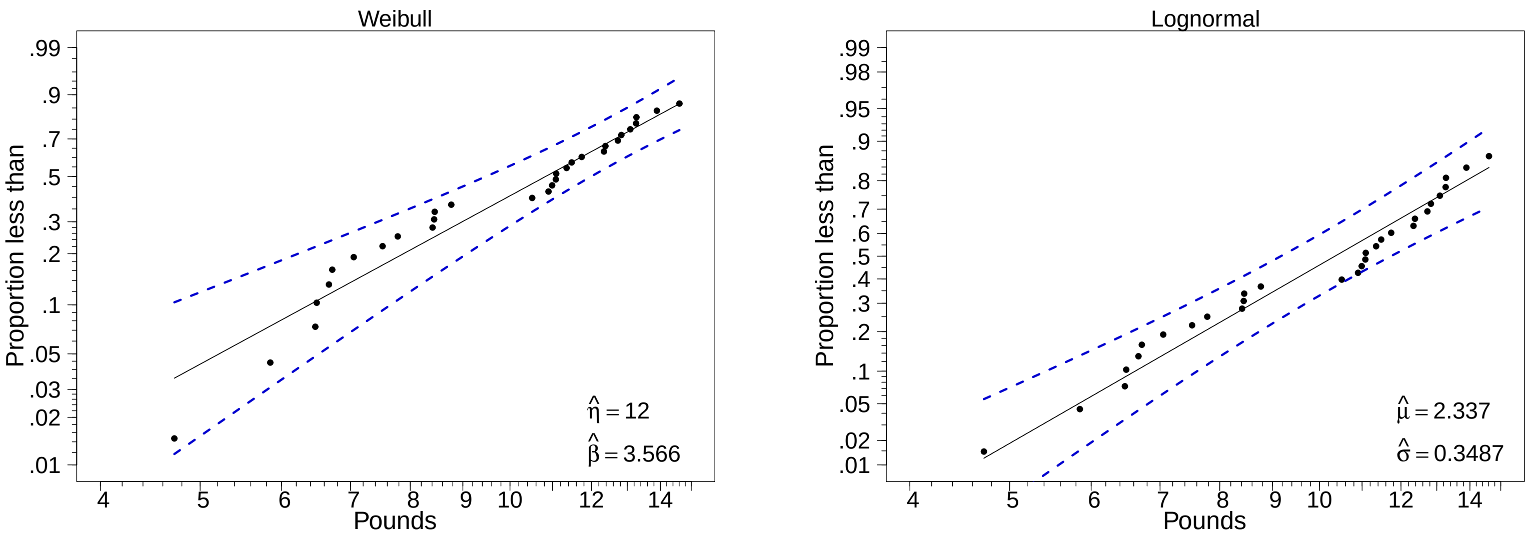
- Liu and Abeyratne (2019, Section 6.2) give data on connector **stress** and connector **strength**.
- Suppose a connector with strength Y is drawn from the population of connectors and the stress that it sees X is drawn from the population of stresses.
- The connector performs correctly if stress (X) is less than strength (Y).
- Thus connector reliability will be $\Pr(X \leq Y)$.
- Both lognormal and Weibull distributions were fit to both X and Y.
- A 95% confidence interval is needed for $R = \Pr(X \leq Y)$.
- For the stress variable, the AIC is slightly smaller for the Weibull distribution (112.5 versus 114.2).
- For strength, the lognormal was a little smaller (170.9 versus 171.26).

Which Distributions to Use?

Weibull and lognormal probability plots for Stress



Weibull and lognormal probability plots for Strength



Comparison of Evaluation of Pr($X \leq Y$) with Different Distributions

- Comparison of point estimates for Pr($X \leq Y$)

Stress Distribution	Strength Distribution Weibull	Lognormal
Weibull	0.9897	0.9969
Lognormal	0.9853	0.9912

- Comparison of approximate 95% confidence intervals for Pr($X \leq Y$)

Stress Distribution	Strength Distribution Weibull	Lognormal
Weibull	[0.9615, 0.9972]	[0.9799, 0.9994]
Lognormal	[0.9470, 0.9955]	[0.9597, 0.9978]

References

Liu, Y. and A. I. Abeyratne (2019). *Practical Applications of Bayesian Reliability*. Wiley. []

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