

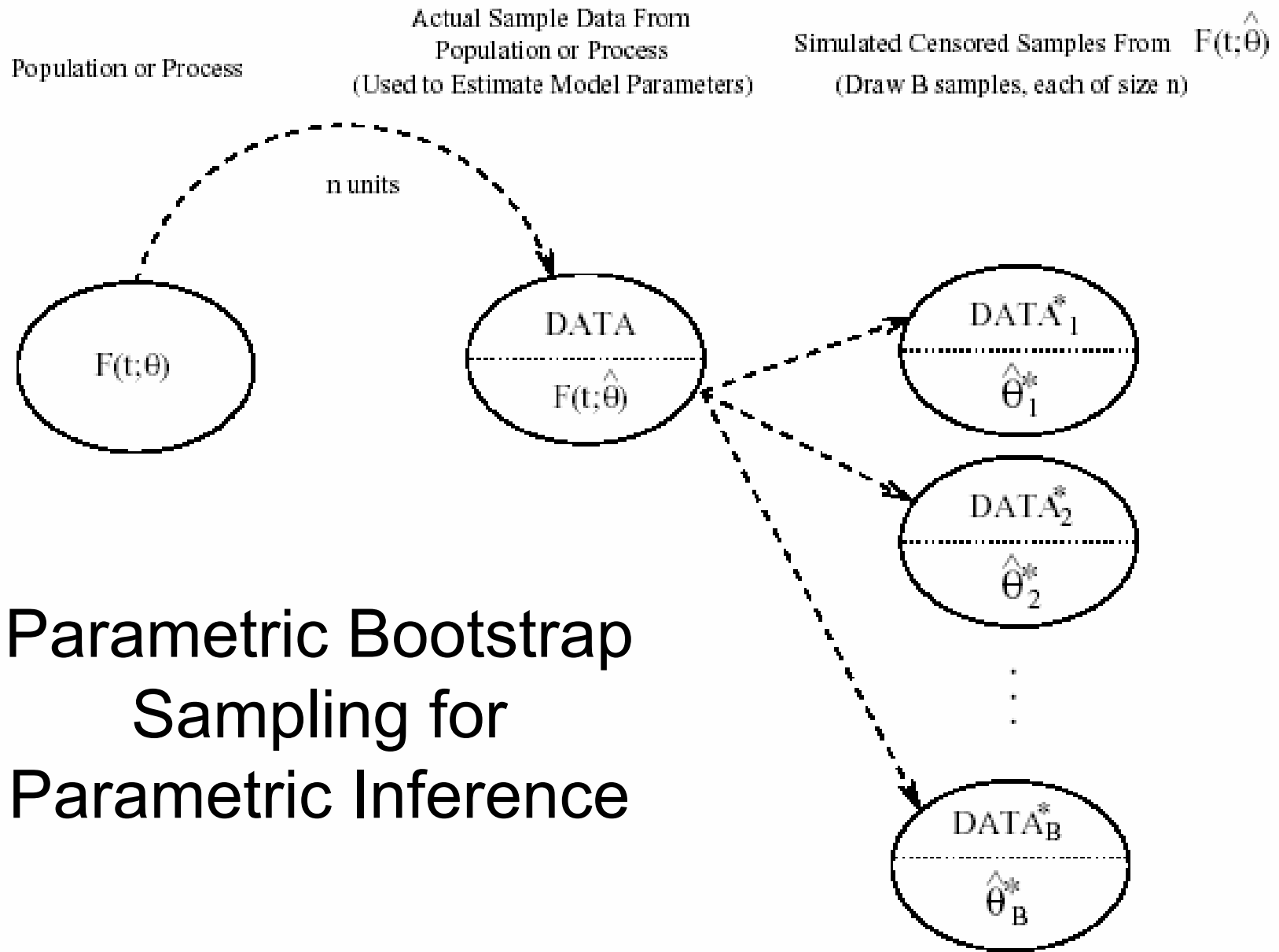
# Unit 9: Bootstrap Confidence Intervals

Ramón V. León

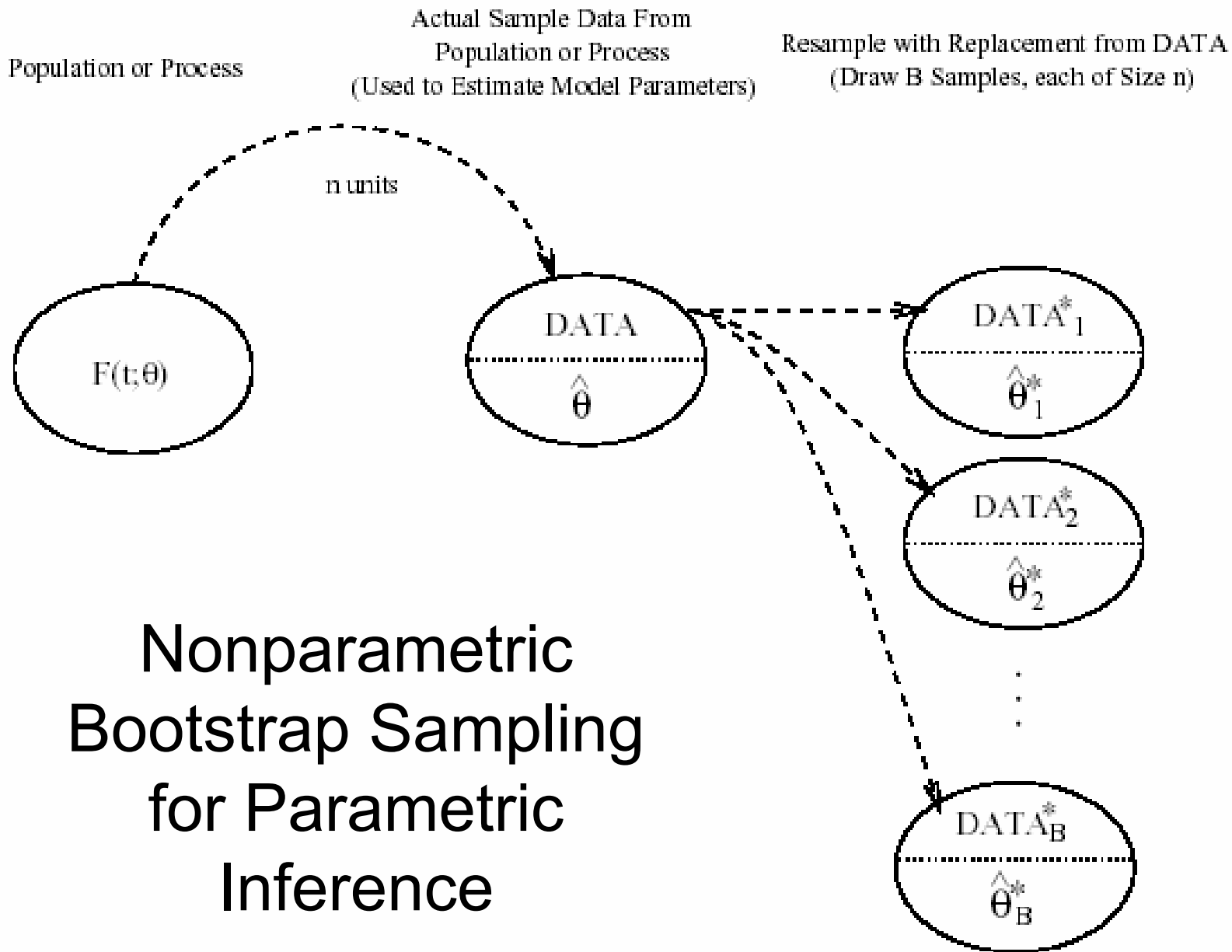
Notes largely based on  
“Statistical Methods for Reliability Data”  
by W.Q. Meeker and L. A. Escobar,  
Wiley, 1998 and on their class notes.

# Unit 9 Objectives

- Explain basic ideas behind the use of computer simulation to obtain bootstrap confidence intervals
- Explain different methods for generating bootstrap samples
- Obtain and interpret simulation-based pointwise parametric bootstrap confidence intervals

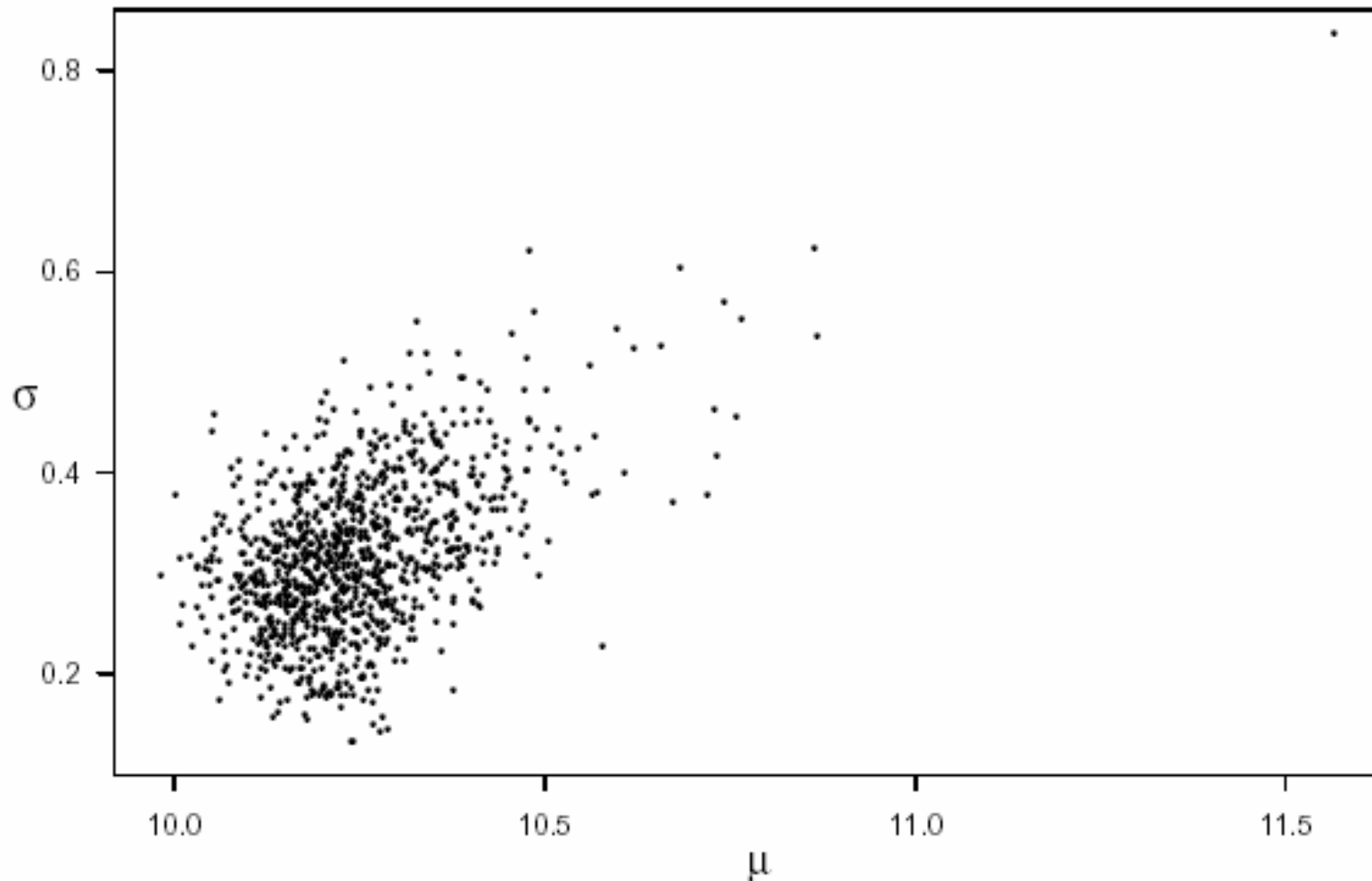


# Parametric Bootstrap Sampling for Parametric Inference

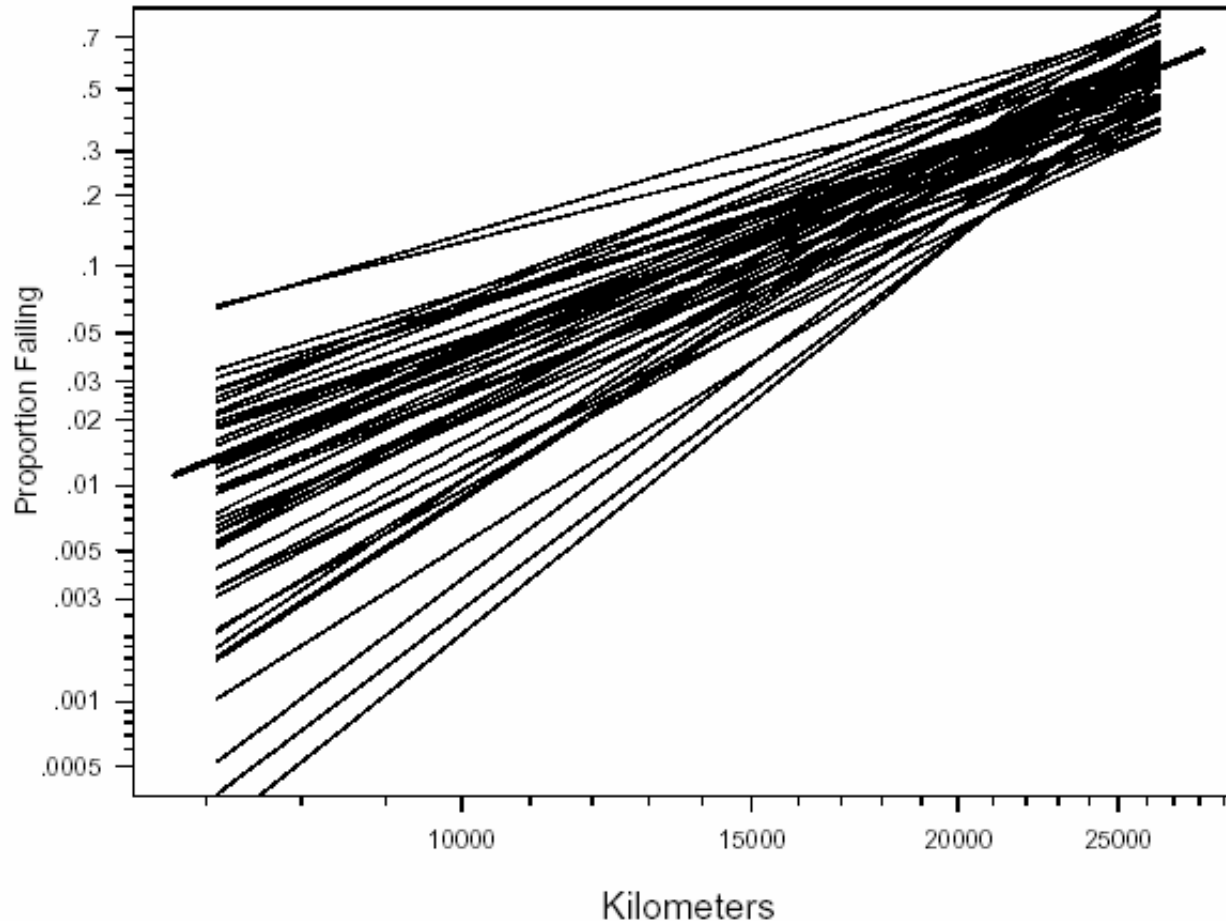


# Nonparametric Bootstrap Sampling for Parametric Inference

Scatterplot of 1,000 (Out of  $B = 10,000$ ) Bootstrap Estimates  $\hat{\mu}^*$  and  $\hat{\sigma}^*$  for Shock Absorber Weibull Model



**Weibull Plot of  $F(t; \hat{\mu}, \hat{\sigma})$  from the Original Sample  
(dark line) and 50 (Out of  $B = 10,000$ )  $F(t; \hat{\mu}^*, \hat{\sigma}^*)$   
Computed from Bootstrap Samples for the  
Shock Absorber**



# Bootstrap Sampling and Bootstrap Confidence Intervals

- Instead of assuming  $Z_{\hat{\mu}} = (\hat{\mu} - \mu) / \widehat{se}_{\hat{\mu}} \sim \text{NOR}(0, 1)$ , use Monte Carlo simulation to approximate the distribution of  $Z_{\hat{\mu}}$ .
- Simulate  $B = 4000$  values of  $Z_{\hat{\mu}^*} = (\hat{\mu}^* - \hat{\mu}) / \widehat{se}_{\hat{\mu}^*}$ .
- Some bootstrap approximations:
  - ▶  $Z_{\hat{\mu}} \sim Z_{\hat{\mu}^*}$
  - ▶  $Z_{\log(\hat{\sigma})} \sim Z_{\log(\hat{\sigma}^*)}$
  - ▶  $Z_{\text{logit}[\hat{F}(t)]} \sim Z_{\text{logit}[\hat{F}^*(t)]}$

when computing confidence intervals for  $\mu$ ,  $\sigma$ , and  $F$ .

# Bootstrap Confidence Interval for $\mu$

- With complete data or Type II censoring,

$$Z_{\hat{\mu}_j^*} = \frac{\hat{\mu}_j^* - \hat{\mu}}{\widehat{se}_{\hat{\mu}_j^*}}$$

has a distribution that does not depend on any unknown parameters. Such a statistic is called a **pivotal** statistic.

- By the definition of quantiles, then

$$\Pr\left(z_{\hat{\mu}^*(\alpha/2)} < Z_{\hat{\mu}_j^*} \leq z_{\hat{\mu}^*(1-\alpha/2)}\right) = 1 - \alpha$$

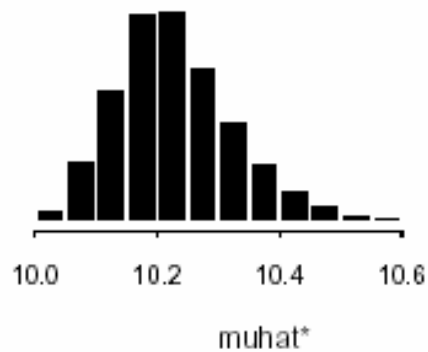
- Simple algebra shows that

$$[\underline{\mu}, \tilde{\mu}] = [\hat{\mu} - z_{\hat{\mu}^*(1-\alpha/2)} \widehat{se}_{\hat{\mu}}, \hat{\mu} - z_{\hat{\mu}^*(\alpha/2)} \widehat{se}_{\hat{\mu}}]$$

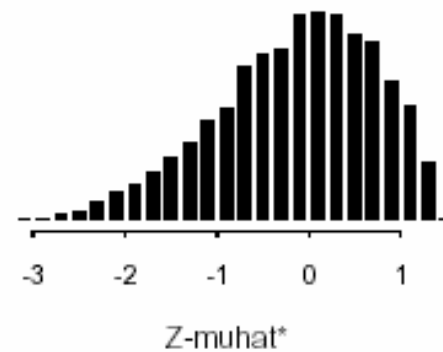
provides an exact 95% confidence interval for  $\mu$ . With other kinds of censoring, the interval is, in general, only **approximate**. (Assuming location scale or log location scale distribution.)

# Bootstrap Distributions of Weibull $\hat{\mu}^*$ and $Z_{\hat{\mu}^*}$ Based on $B=10,000$ Bootstrap Samples for the Shock Absorber

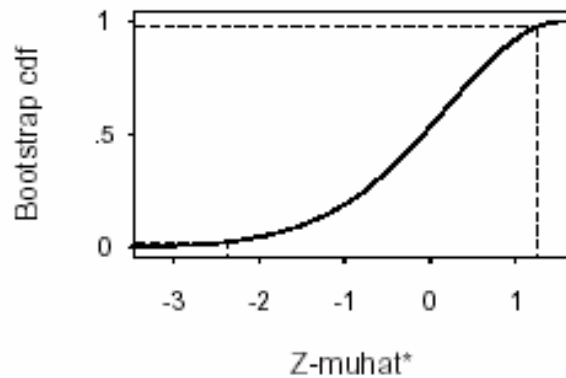
Bootstrap Estimates



Bootstrap-t Untransformed



Bootstrap-t Untransformed



# Bootstrap Confidence Interval for $\sigma$

- With complete data or Type II censoring,

$$Z_{\log(\hat{\sigma}^*)} = \frac{\log(\hat{\sigma}^*) - \log(\hat{\sigma})}{\widehat{\text{se}}_{\log(\hat{\sigma}^*)}}$$

has a distribution that does not depend on any unknown parameters. Such a statistics is called a **pivotal** statistic.

- By the definition of quantiles, then

$$\Pr\left(z_{\log(\hat{\sigma}^*)_{(\alpha/2)}} < Z_{\log(\hat{\sigma}^*)} \leq z_{\log(\hat{\sigma}^*)_{(1-\alpha/2)}}\right) = 1 - \alpha$$

- Simple algebra shows that

$$[\underline{\sigma}, \quad \tilde{\sigma}] = [\hat{\sigma}/\underline{w}, \quad \hat{\sigma}/\tilde{w}]$$

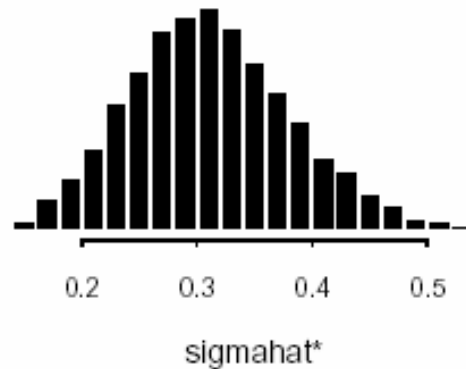
provides an exact 95% confidence interval for  $\sigma$ , where  $\underline{w} =$

$$\exp\left[z_{\log(\hat{\sigma}^*)_{(1-\alpha/2)}} \widehat{\text{se}}_{\log(\hat{\sigma})}\right] \text{ and } \tilde{w} = \exp\left[z_{\log(\hat{\sigma}^*)_{(\alpha/2)}} \widehat{\text{se}}_{\log(\hat{\sigma})}\right]$$

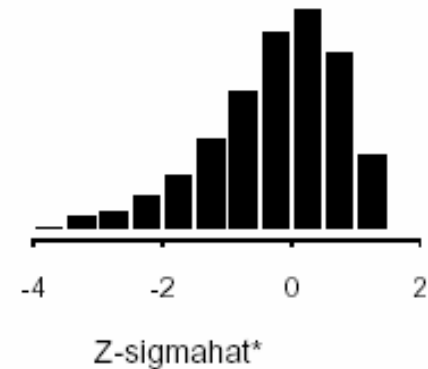
With other kinds of censoring, the interval is, in general, only **approximate**.

# Bootstrap Distributions of $\hat{\sigma}^*$ , $Z_{\hat{\sigma}^*}$ , and $Z_{\log(\hat{\sigma}^*)}$ Based on $B=10,000$ Bootstrap Samples

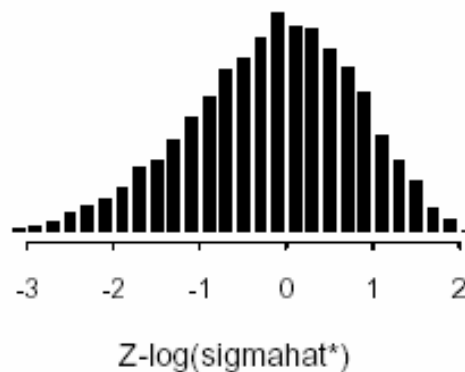
Bootstrap Estimates



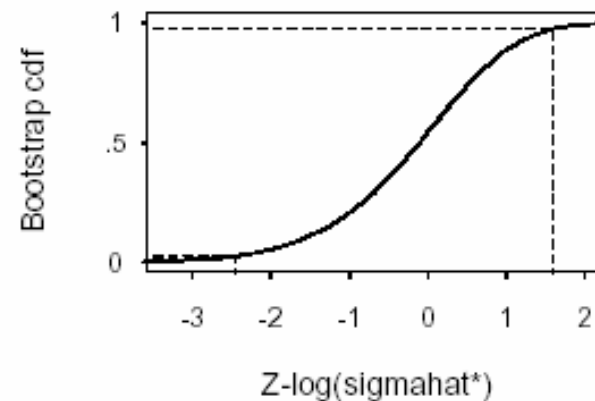
Bootstrap-t Untransformed



Bootstrap-t log-transform



Bootstrap-t log-transform



## Bootstrap Confidence Interval for $F(t_e)$

- With complete data or Type II censoring [using  $F = F(t_e)$ ],

$$Z_{\text{logit}(\hat{F}^*)} = \frac{\text{logit}(\hat{F}^*) - \text{logit}(\hat{F})}{\widehat{\text{se}}_{\text{logit}(\hat{F}^*)}}$$

has a distribution that does not depend on any unknown parameters. Such a statistics is called a **pivotal** statistic.

- By the definition of quantiles, then

$$\Pr \left( z_{\text{logit}(\hat{F}^*)_{(\alpha/2)}} < Z_{\text{logit}(\hat{F}^*)} \leq z_{\text{logit}(\hat{F}^*)_{(1-\alpha/2)}} \right) = 1 - \alpha$$

- Simple algebra shows that

$$[\underline{F}, \quad \tilde{F}] = [\hat{F}/\underline{w}, \quad \hat{F}/\tilde{w}]$$

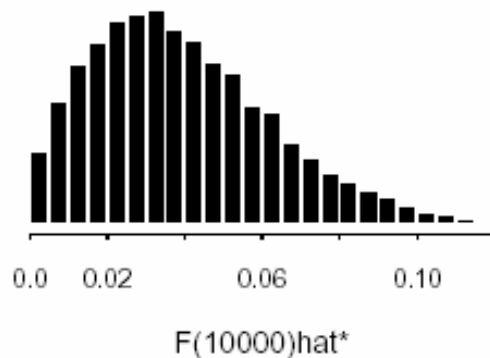
provides an exact 95% confidence interval for  $F$ , where  $\underline{w} =$

$$\exp \left[ z_{\text{logit}(\hat{F}^*)_{(1-\alpha/2)}} \widehat{\text{se}}_{\text{logit}(\hat{F})} \right] \text{ and } \tilde{w} = \exp \left[ z_{\text{logit}(\hat{F}^*)_{(\alpha/2)}} \widehat{\text{se}}_{\text{logit}(\hat{F})} \right]$$

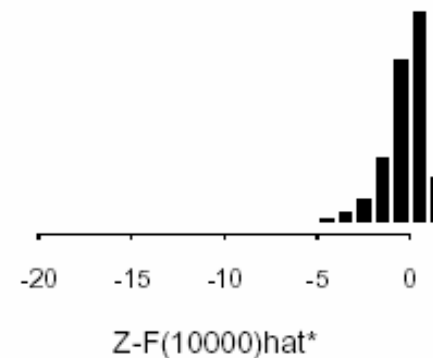
With other kinds of censoring, the interval is, in general, only **approximate**.

# Bootstrap Distributions of $\hat{F}(t_e)^*$ , $Z_{\hat{F}(t_e)^*}$ , and $Z_{\text{logit}[\hat{F}(t_e)^*]}$ for $t_e=10,000$ km Based on $B=10,000$ Bootstrap Samples

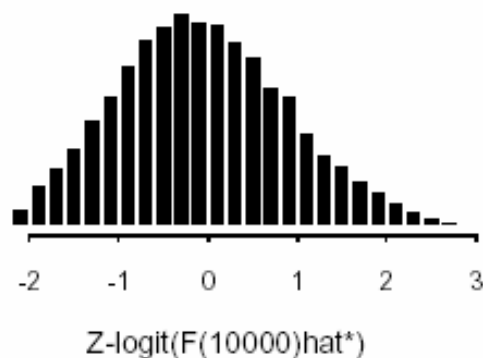
Bootstrap Estimates



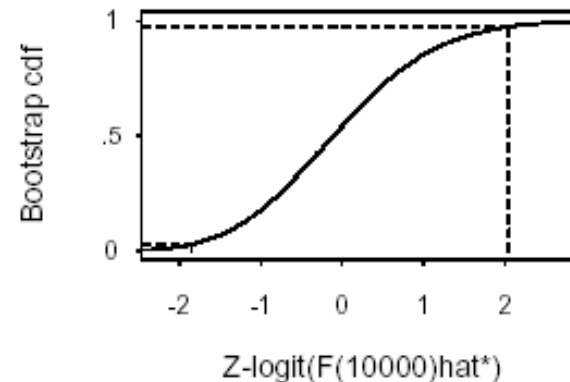
Bootstrap-t Untransformed



Bootstrap-t logit-transformed



Bootstrap-t logit-transformed



## Bootstrap Confidence Interval for $t_p$

- With complete data or Type II censoring,

$$Z_{\log(\hat{t}_p^*)} = \frac{\log(\hat{t}_p^*) - \log(\hat{t}_p)}{\widehat{\text{se}}_{\log(\hat{t}_p^*)}}$$

has a distribution that does not depend on any unknown parameters. Such a statistics is called a **pivotal** statistic.

- By the definition of quantiles, then

$$\Pr \left( z_{\log(\hat{t}_p^*)_{(\alpha/2)}} < Z_{\log(\hat{t}_p^*)} \leq z_{\log(\hat{t}_p^*)_{(1-\alpha/2)}} \right) = 1 - \alpha$$

- Simple algebra shows that

$$[t_p, \tilde{t}_p] = [\hat{t}_p/\underline{w}, \hat{t}_p/\tilde{w}]$$

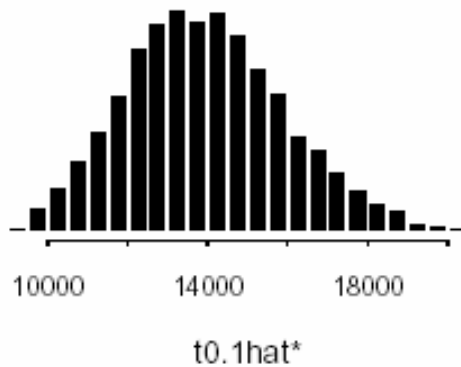
provides an exact 95% confidence interval for  $t_p$ , where  $\underline{w} =$

$\exp \left[ z_{\log(\hat{t}_p^*)_{(1-\alpha/2)}} \widehat{\text{se}}_{\log(\hat{t}_p)} \right]$  and  $\tilde{w} = \exp \left[ z_{\log(\hat{t}_p^*)_{(\alpha/2)}} \widehat{\text{se}}_{\log(\hat{t}_p)} \right]$

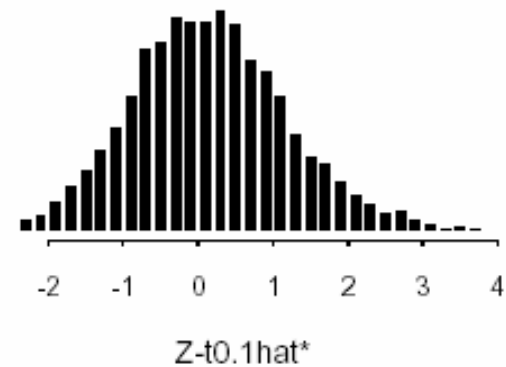
With other kinds of censoring, the interval is, in general, only **approximate**.

# Bootstrap Distributions of $\hat{t}_p^*$ , $Z_{\hat{t}_p^*}$ , and $Z_{\log[\hat{t}_p^*]}$ for $t_e=10,000$ km Based on $B=10,000$ Bootstrap Samples

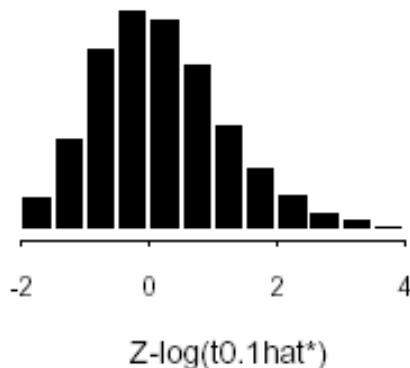
Bootstrap Estimates



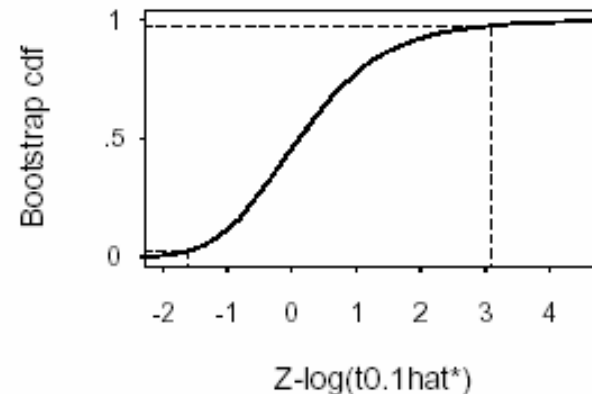
Bootstrap-t Untransformed



Bootstrap-t log-transform



Bootstrap-t log-transform



# Other Topics in Chapter 9

- Nonparametric bootstrap confidence Intervals (with nonparametric resampling)
- Bootstrap percentile method