

# STAT 687 - Homework # 3 - Solution

*Due 0800, 21 July 2014*

## From the text problem # 6.5

Consider the ball bearing fatigue data given in Example 1.1 and Table 1.1

- (a) Compute a nonparametric estimate of  $F(t)$ , the proportion of units failing as a function of time. Plot your estimate on paper with linear scales

Since the data are uncensored,  $\hat{F}(t)$  can be found using the binomial approach (Eq. 3.1) where

$$\hat{F}(t_i) = \frac{\sum_{j=1}^i d_j}{n}.$$

For this data set,  $\sum_{j=1}^i d_j$  is updated after each failure. Since the failure times are exact,  $\hat{F}(t_i)$  is a constant for  $[t_{i-1} - t_i)$ ,  $i = 1, \dots, 23$ . Thus, the plot will appear as a step function. Using RSplida, we can find the values of  $\hat{F}(t_i)$  and plot them on linear scales using the following code.

```
lzbearing.ld<-frame.to.ld(lzbearing,
                           response.column = 1,
                           time.units = "Megacycles")
```

- (b) Make a lognormal probability plot of the data. This is accomplished by ordering the failure times in increasing order,  $T_{(1)} \leq \dots \leq T_{(23)}$ . Then plot  $t_{(i)}$  versus  $(i - .5)/n$  on lognormal probability paper.

The requested plot is obtained by plotting the ordered failure times against  $\Phi_{NOR}^{-1}(\hat{p})$  where  $\hat{p}$  is computed using the Hazen plotting position formula  $\frac{i-0.5}{n}$ :

These values are then plotted on semi-log paper as shown in the R code below

- (c) Do the same as in part (b) but on Weibull probability paper.

The Weibull probability plot is obtained by plotting time against  $\Phi_{SEV}^{-1}(\hat{p})$  where  $\hat{p}$  is again computed using the Hazen plotting position formula

These values are then plotted on semi-log paper as shown in the R code below

- (d) Comment on the adequacy of the lognormal and Weibull models to describe these data.

Upon observing the plots in parts (b) and (c) above, both the Weibull and lognormal distributions appear to fit these data equally well.

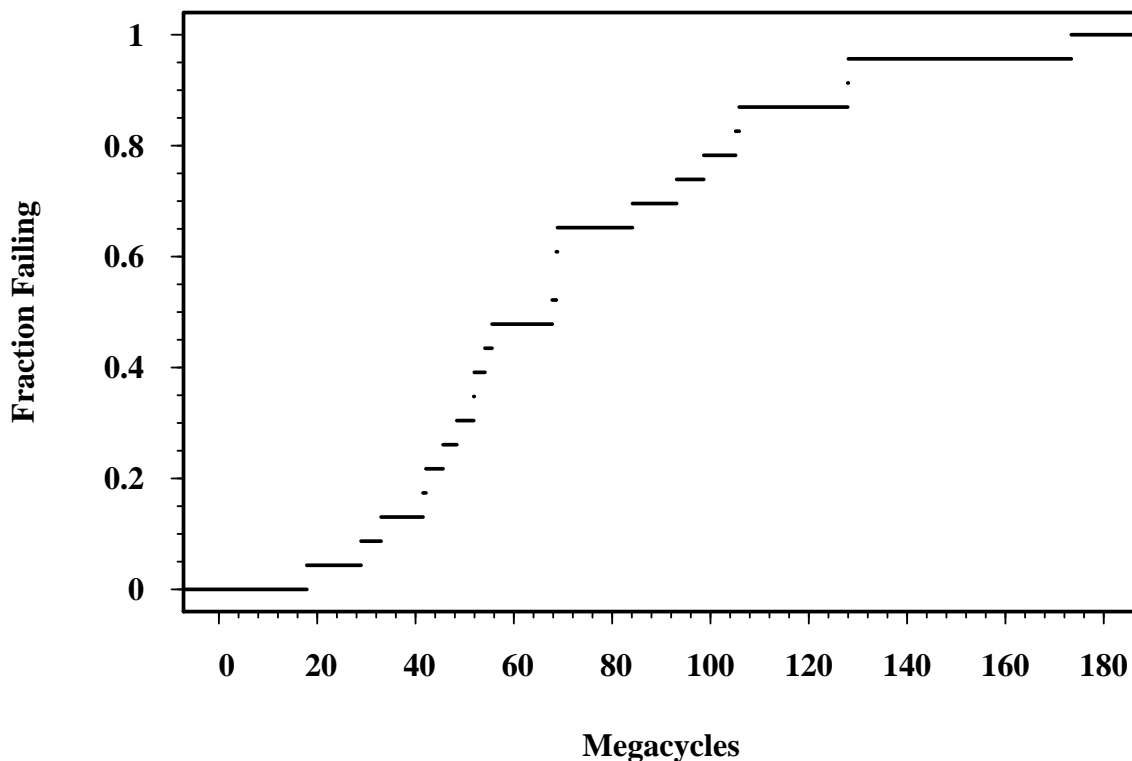


Figure 1: Nonparametric probability plot of the LZ Bearing data set

### From the text problem # 6.7

A sample of 100 specimens of a titanium alloy were subjected to a fatigue test to determine the time to crack initiation. The test was run up to a limit of 100,000 cycles. The observed time of crack initiation (in units of 1,000's of cycles) were: 18, 32, 39, 53, 59, 68, 77, 78, 93. No crack had initiated in any of the other 91 specimens.

- (a) Compute a nonparametric estimate,  $\hat{F}(t)$  of the cdf  $F(t)$  using both the simple binomial method and the Kaplan-Meier method (in this case the two methods provide the same answer)

Recall that the Kaplan-Meier estimator for the survivor function can be expressed as

$$\hat{S}(t_i) = \prod_{j=1}^i \left[ 1 - \frac{d_j}{n_j} \right]$$

where both  $d_j$  and  $n_j$  change by a value of one after each failure because the failure times are exact and because this data set is singly right censored at 100 kilocycles. The nonparametric estimate of the cdf at 100 kilocycles, for example, is then

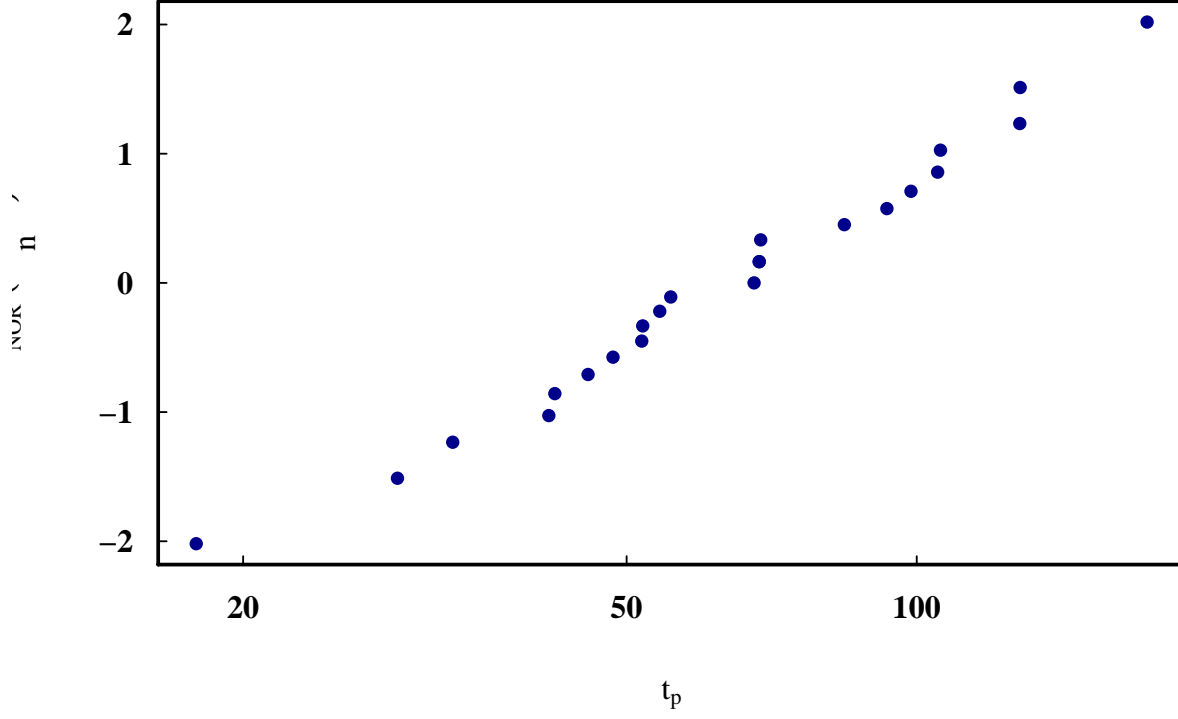


Figure 2: Lognormal probability plot of the LZ Bearing data set

$$\hat{F}(100) = 1 - \hat{S}(100) = 1 - \frac{99}{100} \times \frac{98}{99} \times \frac{97}{98} \times \frac{96}{97} \times \frac{95}{96} \times \frac{94}{95} \times \frac{93}{94} \times \frac{92}{93} \times \frac{91}{92} = 1 - \frac{91}{100} = \frac{9}{100}$$

The simple binomial method produces an estimate for the cdf as

$$\hat{F}(t_j) = \frac{\sum_{j=1}^i d_j}{n}$$

where only the numerator changes by one after each failure. The estimate of the cdf at 100 kilocycles using the binomial method is

$$\hat{F}(100) = \frac{9}{100}.$$

For each failure time, the value of  $\hat{F}(t_j), j = 1, 2, \dots, i$  computed using both approaches will be equivalent. These values are: (.01, .02, .03, .04, .05, .06, .07, .08, .09). For the censored observations we are unable to accurately estimate the value of the cdf, thus for these 91 non-failures we assign a value of .09.

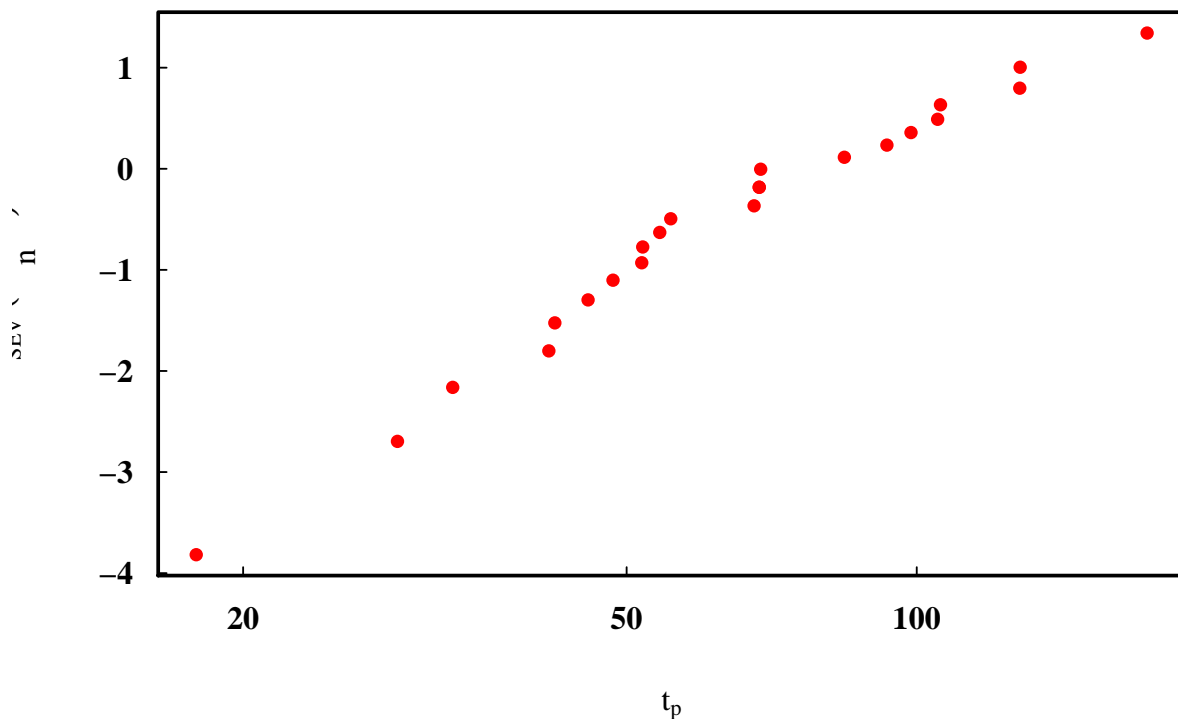


Figure 3: Weibull probability plot of the LZ Bearing data set

(b) Plot  $\hat{F}(t)$  on linear axes.

The SMRD package can be used to produce the requested plot of  $\hat{F}(t)$  with the following code. Although the data set for this problem is loaded in the SMRD package as `titanium01`, we'll produce the plot by creating the data frame from scratch.

```
obs <- c(18,32,39,53,59,68,77,78,93,rep(100,91))
cen <- c(rep("failed",9), rep("right",91))

titanium.ld <- frame.to.ld(data.frame(obs,cen),
                           response.column = 1,
                           censor.column = 2,
                           time.units = "Kilocycles")

plot(titanium.ld)
```

(c) Use  $\hat{F}(t)$  to compute plotting positions and plot the data on Weibull paper. Use the plot to obtain an estimate of the Weibull shape parameter  $\beta$

The SMRD package can compute the plotting positions for this singly right-censored data set and plot these points on Weibull paper using the following R code

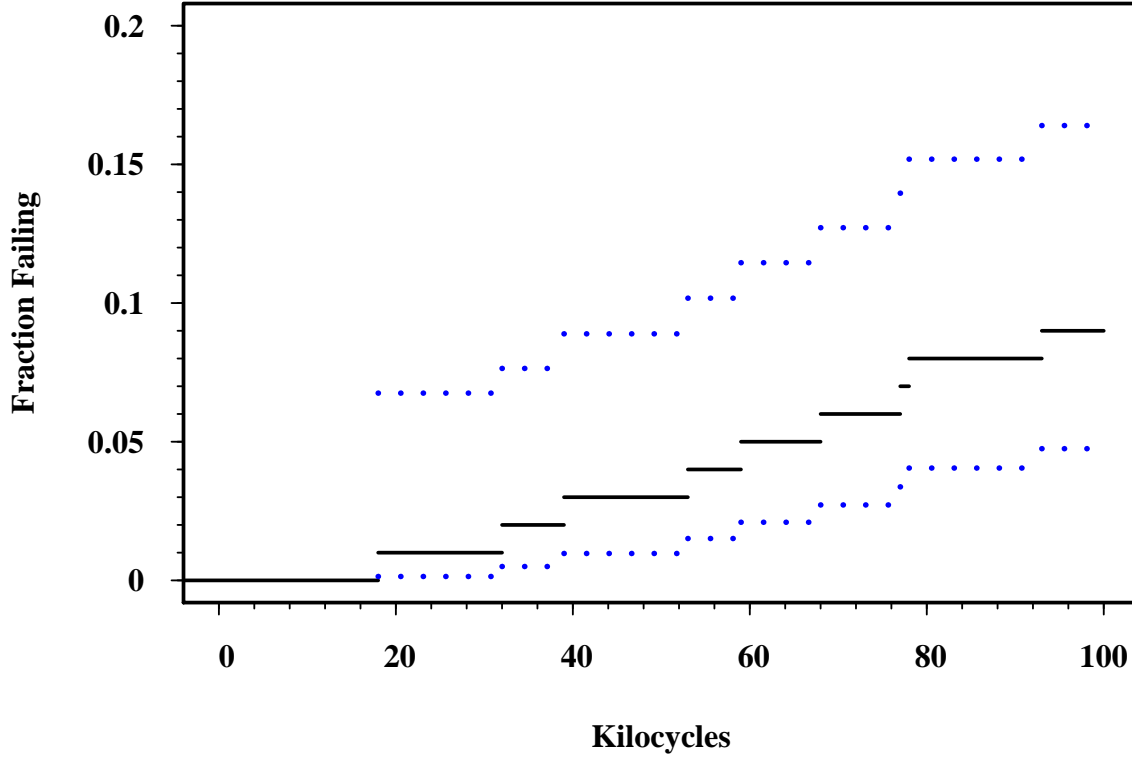


Figure 4: Nonparametric CDF plot for the titanium fatigue crack initiation data

Estimating  $\beta = \sigma^{-1}$  from the plot requires that the plotting position values are known. The SMRD package does not provide these values, but we can calculate them ourselves. In § 6.4.2 we see that for singly right-censored data Meeker recommends using Hazen plotting positions. These values can be created with the following code

A graphical estimate for the Weibull shape parameter  $\beta$  may be found from the slope of the best fit line through the data. From Figure 5, an estimate of the slope is

$$\frac{\log[-\log[1 - .09]] - \log[-\log[1 - .01]]}{\log[93] - \log[18]} = \frac{\log\left[\frac{\log[.91]}{\log[.99]}\right]}{\log\left[\frac{93}{18}\right]} = 1.36$$

- (d) Comment on the adequacy of the Weibull distribution\_\_

The transformed data appears to fall close to a straight line. Thus, the assumption that the data were generated from a Weibull process seems reasonable.

- (e) Comment on the adequacy of the available data if the purpose of the experiment was to estimate  $t_{.1}$ .

Since only 9% of the systems under test failed, estimating  $t_{.1}$  will require extrapolating outside of the data. However, since none of the data fall far from the best fit line there may be little risk in this short of an extrapolation.

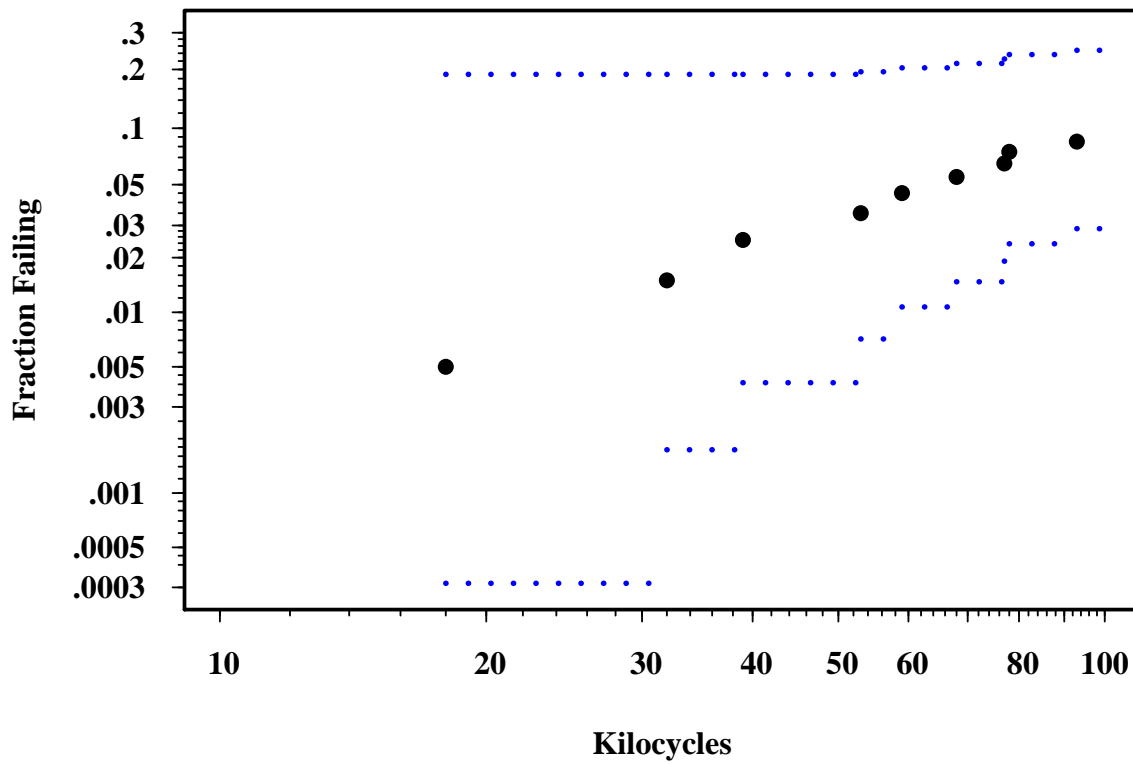


Figure 5: Nonparametric CDF plot for the titanium fatigue crack initiation data on Weibull paper

### From the text problem # 6.9

Using the life test data on silicon photodiode detectors from Exercise 3.12, construct probability plots for the exponential, Weibull, and lognormal distributions. Which distributions look like they might provide an adequate model for photodiode detector life?

This problem does not indicate that we must use the Hazen plotting position. Therefore, using the SMRD package, the requested plots can be produced with the following code:

```
spd.ld<-frame.to.ld(photodetector,
                    response.column = c(1,2),
                    censor.column = 3,
                    case.weight.column = 4,
                    time.units = "Hours (in Hundreds)")
par(family = "serif", mfrow = c(3,1))
plot(spd.ld, distribution = "exponential")
plot(spd.ld, distribution = "weibull")
plot(spd.ld, distribution = "lognormal")
```

```
par(mfrow = c(1,1))
```

Looking at these plots it appears that both the lognormal and Weibull distributions adequately fit the photodiode detector life data, while the exponential distribution does not.

## From the text problem # 7.7

A life test for a new insulating material used 50 specimens. The specimens were tested simultaneously at 40kV (considerably higher than the rated voltage of 20kV). The test was run until 10 of the specimens (this is known as “failure” or Type II censoring). The failure times were recorded as 8, 11, 12, 13, 19, 21, 28, 34, 36, and 44 hours. The engineers responsible for the reliability believe, based on previous experience with similar materials tested under similar conditions, that the failure-time distribution at 40kV can be described by an  $EXP(\theta)$  distribution.

- (a) Construct an exponential probability plot of the data. Does the plot provide any evidence that the exponential distribution is inadequate?

This problem involves Type II censoring, as the test was terminated after observing a specified number of failures (10). The 40 specimens that did not fail should then be considered as right censored observations where  $t_{11}, \dots, t_{50} \in (44, \infty)$ .

The simplest method for constructing this exponential probability plot is using the SMRD package. This requires that we first put the event times and event designators in a data frame for the “frame.to.ld” function to use to create a life data object. The following code can be used to accomplish this.

```
obs7.7<-c(8,11,12,13,19,21,28,34,36,rep(44,41))
cen7.7<-c(rep("failed",10), rep("censored",40))
dat7.7<-data.frame(obs7.7,cen7.7)

insulate.ld<-frame.to.ld(dat7.7,
                          response.column = 1,
                          censor.column = 2,
                          time.units = "Hours")

par(family="serif")
plot(insulate.ld, distribution="Exponential", plot.censored.ticks="top")
```

This may also be found by inverting the cdf  $F(\theta|t) = 1 - \exp(-\frac{t}{\theta})$  to obtain the exponential distribution quantile function  $\log\left(\frac{1}{1-p}\right) = \frac{t}{\theta}$ . Since this data set is singly censored we replace the quantile value  $p$  with the Hazen plotting position  $\frac{i-0.5}{n}$  where  $i$  denotes the ordered index of failure for each item under test and  $n$  is the total sample size. These transformed quantile values are then plotted against the observation times on a standard axis plot.

- (b) Compute TTT, the total time on test, and  $\hat{\theta}$ , the ML estimate for  $\theta$

The total time on test,  $TTT$  can be calculated by summing the elements in the data vector defined above in part (a)

```
obs7.7 <- c(8,11,12,13,19,21,28,34,36,rep(44,41))
TTT    <- sum(obs7.7)
```

Thus,

$$TTT = \sum_{i=1}^n t_i = 1986.$$



The ML estimate  $\hat{\theta}_{MLE}$  is computed by noting that the likelihood function for exponentially distributed data with exact failures and right censored observations may be expressed as

$$L(\theta|\underline{X}) = \prod_{i=1}^n \left(\frac{1}{\theta}\right)^{\delta_i} \exp\left(-\frac{t_i}{\theta}\right)$$

The log-likelihood function is then

$$\begin{aligned} \mathcal{L}(\theta|\underline{X}) &= \sum_{i=1}^n -\delta_i \log(\theta) - \left(\frac{t_i}{\theta}\right) \\ &= \sum_{i=1}^n -\delta_i \log(\theta) - \frac{\sum_{i=1}^n t_i}{\theta} \end{aligned}$$

Taking the derivative of  $\mathcal{L}(\theta|\underline{X})$  with respect to  $\theta$  and setting equal to 0 results in

$$\hat{\theta}_{MLE} = \frac{\sum_{i=1}^n t_i}{\sum_{i=1}^n \delta_i}$$

We note that the numerator in the above equation is the total time on test,  $TTT$ , while the denominator is the number of observed failures, thus

$$\hat{\theta}_{MLE} = \frac{1986 \text{ hours}}{10 \text{ failures}} = 198.6.$$

- (c) Compute an estimate of the standard error of  $\theta$

The standard error for  $\hat{\theta}$  may be found from

$$\hat{se}_{\hat{\theta}} = \sqrt{\left[-\frac{d^2 \mathcal{L}(\theta|\underline{X})}{d\theta^2}\right]^{-1}}.$$

Returning to (??), the derivative of this function (the second derivative of the log-likelihood function) taken with respect to  $\theta$  at the value  $\hat{\theta}_{MLE}$  results in

$$\begin{aligned} \frac{d^2 \mathcal{L}(\theta|\underline{X})}{d\theta^2} &= \frac{\sum_{i=1}^n \delta_i}{(\hat{\theta}_{MLE})^2} - \frac{2 \sum_{i=1}^n t_i}{(\hat{\theta}_{MLE})^3} \\ &= \frac{r}{(\hat{\theta}_{MLE})^2} - \frac{2(TTT)}{\hat{\theta}_{MLE}^3}. \end{aligned}$$

Multiplying this function by  $-1$ , taking the reciprocal and simplifying results in

$$\frac{d^2 \mathcal{L}(\theta|\underline{X})}{d\theta^2} = \frac{(\hat{\theta}_{MLE})^2}{r}$$

and

$$\hat{se}_{\hat{\theta}_{MLE}} = \frac{\hat{\theta}_{MLE}}{\sqrt{r}} = \frac{198.6}{\sqrt{10}} = 62.8$$

- (d) Compute 95% confidence intervals for  $\theta$  based on  $Z_{\hat{\theta}} \sim NOR(0, 1)$ ,  $Z_{\log(\hat{\theta})} \sim NOR(0, 1)$ , and the exact distribution of  $2(TTT/\theta)$ . Which of these intervals would you feel comfortable using? Why?

The requested confidence intervals are shown below and may be computed using equations 7.10, 7.11 and the top equation on page 167, respectively.

$$\begin{aligned} CI(Z_{\hat{\theta}}) &= \hat{\theta} \pm Z_{(1-\frac{\alpha}{2})} \hat{se}_{\hat{\theta}} = [75.51, 321.69] \\ CI(Z_{\log(\hat{\theta})}) &= \hat{\theta} \times \exp(\pm Z_{(1-\frac{\alpha}{2})} \hat{se}_{\hat{\theta}}/\hat{\theta}) = [106.86, 369.11] \\ CI(\text{Exact}) &= \frac{2(TTT)}{\chi^2_{(1-\alpha/2; 2r)}}, \frac{2(TTT)}{\chi^2_{(\alpha/2; 2r)}} = [116.24, 414.15] \end{aligned}$$

Knowing that the exponential distribution is positively skewed, I would feel least comfortable using the  $Z_{\hat{\theta}}$  intervals. I would feel most comfortable using the exact confidence interval.

- (e) For this problem, is there any extrapolation involved in estimating  $\theta$ ? Explain.

Extrapolation exists in assuming that the data follow an  $\text{Exp}(\theta)$  distribution as this belief leads us to conclude that the ML estimate for  $\theta = \frac{TTT}{r}$

- (f) Compute 95% confidence intervals for  $t_{.1}$ ,  $h(50; \theta)$ , and  $F(50; \theta)$  based on the exact distribution of  $2(TTT/\theta)$ . Is there any extrapolation involved in these intervals? Explain.

As discussed in class and in the text an exact  $100(1 - \alpha)\%$  confidence interval for  $\theta$  may be expressed as

$$[\underline{\theta}, \bar{\theta}] = \left[ \frac{2(TTT)}{\chi^2_{(1-\alpha/2; 2r)}}, \frac{2(TTT)}{\chi^2_{(\alpha/2; 2r)}} \right]$$

Since  $t_{0.1}$  is a monotone function of  $\theta$  where  $t_{0.1} = -\theta \log(0.9)$  (via the exponential quantile function) it follows that an exact 95% confidence interval for  $t_{0.1}$  under Type II censoring can be expressed as

$$\left[ \frac{-2(TTT) \times \log(0.9)}{\chi^2_{(0.975; 18)}}, \frac{-2(TTT) \times \log(0.9)}{\chi^2_{(0.025; 18)}} \right] = [12.25, 43.63]$$

Similarly, the exponential hazard function  $h(t) = \theta^{-1}$  is a monotone function of  $\theta$  and an exact 95% confidence interval for  $h(t)$  under Type II censoring can be expressed as

$$\left[ \frac{\chi^2_{(0.025; 18)}}{2(TTT)}, \frac{\chi^2_{(0.975; 18)}}{2(TTT)} \right] = [.00024, .00086]$$

Finally, the exponential cdf  $F(50|\theta)$  is also a monotone function of  $\theta$ . Therefore, an exact 95% confidence interval for  $F(50|\theta)$  can be expressed as

$$\left[ 1 - \exp \left( -50 \frac{\chi^2_{(0.025;18)}}{2(TTT)} \right), 1 - \exp \left( -50 \frac{\chi^2_{(0.975;18)}}{2(TTT)} \right) \right] = [.114, .350]$$

## From the text problem # 7.8

Use the results from Exercise 7.7 to compute and plot (use an exponential plot) the ML estimate of  $F(t; \theta)$  and 95% simultaneous parametric confidence bands for  $F(t; \theta)$ , based on the exact distribution of  $2(TTT/\theta)$ .

From problem 7.7(b)  $\hat{\theta}_{MLE}$  was determined to be 198.6. Thus,

$$\hat{F}(t)_{MLE} = F(t; \hat{\theta}_{MLE}) = 1 - \exp\left(-\frac{t}{198.6}\right)$$

From problem 7.7(f) the upper and lower limits of the 95% confidence interval on  $\hat{F}(t)_{MLE}$  based on the exact distribution of  $\frac{2(TTT)}{\theta}$  were determined to be

$$\left[1 - \exp\left(-t \frac{\chi^2_{(0.025;18)}}{2(TTT)}\right), 1 - \exp\left(-t \frac{\chi^2_{(0.975;18)}}{2(TTT)}\right)\right]$$

Inserting the value of the Chi-square quantile function and the  $TTT$  these limits become

$$\left[1 - \exp\left(\frac{-t}{482.58}\right), 1 - \exp\left(\frac{-t}{125.99}\right)\right]$$

The plot of  $\hat{F}(t)_{MLE}$  along with these pointwise confidence limits is shown below

Converting these pointwise limits into simultaneous limits of a 95% confidence interval is accomplished by using equation (3.19, pg 62) where  $\hat{se}_{\hat{F}}$  is expressed as shown in equation 3.4 (since the data is singly censored) and  $e_{(a,b,1-\alpha/2)}$  is determined from Table 3.5. The values of  $a$  and  $b$  chosen as the minimum and maximum quantile values that contain the data. For the data in problem 7.7 the minimum and maximum quantile values computed from the Hazen plotting position formula are 0.01 and 0.17. Thus, selecting  $a = 0.01$  and  $b = 0.90$  from Table 3.5 is sufficient for these data and results in  $e_{(.01,.90,.975)} = 3.21$  Inserting  $e$  into the equation (3.19) results in the simultaneous confidence intervals shown in the following plot.

### From the text problem # 7.9

A life test was conducted for the same insulating material described in Exercise 7.7. Again, 50 specimens were tested, but at 25 kV. The test ran for 20 hours without any failures. The test had to be terminated at this time so that the test equipment could be used for other experiments.

- (a) Compute the  $TTT$  and show why the ML estimate for  $\theta$  is equal to  $\infty$ .

For this test the  $TTT$  is  $50 \times 20 = 1000$  hours. Because no failures were observed the maximum likelihood estimate is undefined as

$$\lim_{r \rightarrow \infty} \hat{\theta}_{MLE} = \frac{1000}{0} \rightarrow \infty$$

- (b) Compute a conservative 95% lower confidence bound for  $\theta$ .

As shown in equation (7.17), a conservative  $100(1 - \alpha)\%$  lower confidence bound for  $\theta$

$$\underline{\theta} = \frac{TTT}{-\log(\alpha)} = \frac{1000}{-\log(0.05)} = 333.81 \text{ hours.}$$

- (c) For this problem, is there any extrapolation involved in computing the lower confidence bound for  $\theta$ ? Explain.

There is extrapolation in this problem since we are attempting to estimate  $\theta$  (the time at which 63.2% of a population has failed) from a test that includes no failures.

- (d) Use the result in part (b) to compute a conservative 95% lower confidence bound for  $t_{.1}$ , and conservative 95% upper confidence bounds for  $h(50; \theta)$  and  $F(50; \theta)$ .

A conservative 95% lower confidence bound on  $t_{0.1}$  can be found from inverting the exponential cdf where

$$t_{0.1} = -\theta \times \log(1 - 0.1) = -333.8 \times \log(1 - 0.1) = 35.17.$$

A conservative 95% upper confidence bound on  $F(50; \theta)$ , the probability of failing before 50 hours, can be expressed as

$$F(50; \theta) = 1 - \exp\left(\frac{-50}{333.8}\right) = .139.$$

A conservative 95% upper confidence bound on  $h(50; \theta)$ , the conditional probability of failure after 50 hours, given that failure has not occurred at 50 hours, is expressed as

$$h(50; \theta) = \frac{1}{\theta} = \frac{1}{333.8} = 0.0029.$$

## From the text problem # 8.5

Return to the fatigue crack-initiation test in Exercise 6.7. Fit a Weibull distribution with a given shape parameter  $\beta = 2$ .

- (a) Compute the ML estimate of  $\eta$ . What is the practical interpretation of this estimate?

If neither  $\eta$  nor  $\beta$  are known, a closed-form analytical solution for the MLE's does not exist. In this problem, the value of  $\beta$  is known, thus  $F(t|\eta, \beta = 2) = 1 - e^{-\left(\frac{t}{\eta}\right)^2}$  and  $f(t) = \frac{2}{\eta} \left(\frac{t}{\eta}\right) e^{-\left(\frac{t}{\eta}\right)^2}$ . The likelihood function for the data is

$$L(\eta) = \prod_{i=1}^9 f(t_i) \prod_{i=10}^{100} 1 - F(t_i) = \frac{2^9}{\eta^{18}} \left( \prod_{i=1}^9 t_i \right) \exp \left( -\frac{\sum_{i=1}^9 t_i^2}{\eta^2} \right) \exp \left( -\frac{\sum_{i=10}^{100} t_i^2}{\eta^2} \right)$$

After combining terms, the log-likelihood function is

$$\mathcal{L}(\eta) = C - 18 \times \log(\eta) - \eta^{-2} \sum_{i=1}^{100} t_i^2.$$

Taking the derivative of the log-likelihood function wrt  $\eta$  gives

$$\frac{d\mathcal{L}}{d\eta} = \frac{-18}{\eta} + \frac{2 \sum_{i=1}^{100} t_i^2}{\eta^3} \rightarrow \eta^2 = \frac{\sum_{i=1}^{100} t_i^2}{9} \rightarrow \hat{\eta} = \frac{\sqrt{\sum_{i=1}^{100} t_i^2}}{3}$$

$$\frac{d\mathcal{L}^2}{d\eta^2} = \frac{18}{\eta^2} + \frac{6 \sum_{i=1}^{100} t_i^2}{\eta^4}$$

So,  $\hat{\eta}_{MLE} = \sqrt{314815} = 323.94$  kilocycles, the practical implication of this result is that the Weibull shape parameter, or characteristic life (the time at which 63.2% of a population of units has failed) is equal to the  $\hat{\beta}_{MLE}^{-1}$  root of the ratio of the sum of the individual items times on test, raised to the  $\hat{\beta}_{MLE}$  power, divided by the number of failures.

- (b) Obtain the estimate  $\hat{se}_{\hat{\eta}}$ .

The value for  $\hat{se}_{\hat{\eta}}$  is obtained from taking the square root of the negative-inverse of the Observed Fisher Information matrix. For this problem the Observed Fisher Information matrix is a single term, thus the inverse is found by taking the reciprocal.

$$\hat{se}_{\hat{\eta}} = \sqrt{\left[ -\frac{d^2 \mathcal{L}(\eta|\underline{X})}{d\eta^2} \right]^{-1}} = \sqrt{\left( -\frac{18}{(\hat{\eta}_{MLE})^2} + \frac{6 \sum_{i=1}^{100} t_i^2}{(\hat{\eta}_{MLE})^4} \right)^{-1}} = \sqrt{2914.87} = 53.99$$

- (c) Compute a conservative 95% confidence interval for  $\eta$ .

To find the confidence intervals for  $\hat{\eta}$  we substitute  $\hat{\mu}_{MLE} = \log[\hat{\eta}_{MLE}]$  into Equation (8.7) where

$$[\underline{\mu}, \bar{\mu}] = \hat{\mu} \pm Z_{(1-\alpha/2)} \hat{se}_{\hat{\mu}}$$

and

$$\hat{se}_{\hat{\mu}} = \sqrt{\left( -\frac{18}{(\log[\hat{\eta}_{MLE}])^2} + \frac{6 \sum_{i=1}^{100} t_i^2}{(\log[\hat{\eta}_{MLE}])^4} \right)^{-1}} = \sqrt{.000197} = 0.01404$$

The confidence intervals for  $\eta$  may then be expressed as

$$[\underline{\eta}, \bar{\eta}] = \exp(\log[\hat{\eta}_{MLE}] \pm 1.96 \times 0.01404) = [332.97, 315.15]$$

- (d) Plot the Weibull estimate of  $F(t)$  along with the nonparametric estimate. Comment on the adequacy of the Weibull distribution to describe the data.

In Problem 6.7(b) a plot of the nonparametric estimate of the cdf was produced for the titanium alloy fatigue test data. Here, we are asked to plot a Weibull  $(\hat{\eta}_{MLE}, \hat{\beta}_{MLE})$  cdf on top of the nonparametric estimate to verify that the maximum likelihood distribution fits the data. A poor fit could result from (1) calculating  $\hat{\beta}_{MLE}$  or  $\hat{\eta}_{MLE}$  incorrectly, (2) assuming the Weibull distribution is correct when it is not, or (3) insufficient data to accurately estimate the parameter values. The plot, shown below, includes the nonparametric estimate of the cdf as well as three different Weibull models.

For the first Weibull model  $\hat{\beta}_{MLE} = 2$  as was directed in part (a) of this problem. The corresponding scale parameter  $\hat{\eta}_{MLE}$  was found to be 323.94. For the second Weibull model the shape parameter was found by using the Kaplan-Meier plotting position values calculated in problem 6.7(a), resulting in  $\hat{\beta}_{MLE} = 1.36$ . The shape parameter for the third Weibull model was found using the Hazen plotting position values, for this model  $\hat{\beta}_{MLE} = 1.75$ . Using the procedure shown in part (a) to calculate  $\hat{\eta}_{MLE}$  resulted in the scale parameter values shown in the legend.

From observing these curves, the plotting position values obtained from the Kaplan-Meier estimate result in a Weibull model that fits the data better than the model resulting from the Hazen plotting position values or whatever method was used to arrive at  $\hat{\beta}_{MLE} = 2$ .

- (e) What is an estimate of the .1 quantile of the time-to-initiation distribution?

The maximum likelihood estimate for  $t_{0.1}$  is found from the Weibull quantile function where  $\hat{\beta}_{MLE} = 2$  and  $\hat{\eta}_{MLE} = 323.94$ . Substituting these values into the quantile function gives

$$(\hat{t}_{0.1})_{MLE} = \hat{\eta}_{MLE} \sqrt{-\log[1 - 0.1]} = 105.15 \text{ kilocycles}$$

- (f) Compute a conservative 95% confidence interval for the .1 quantile of the time-to-initiation distribution.

Because the quantile function is a monotone function of  $\eta$  a 95% confidence interval on  $\left(\hat{t}_{0.1}\right)_{MLE}$  is found from

$$[t_{0.1}, \bar{t}_{0.1}] = \left[ \underline{\eta} \times \sqrt{-\log[0.9]}, \bar{\eta} \times \sqrt{-\log[0.9]} \right] = [75.85, 145.77]$$



### From the text problem # 8.23

Analysis of field data has suggested that a particular engine bearing is an important life-limiting component. The responsible engineers believe that improving the bearing's reliability would have an important effect on overall engine reliability. A redesigned bearing was tested extensively in a bench life-test with simulated loads. Ten bearings were tested, each for 500 hours, with no reported failures. On the basis of previous field failure data for this class of bearing, a Weibull distribution with a shape parameter of  $\beta = 2.3$  had provided an adequate description for the life of this bearing. Find a 95% lower confidence bound for the .01 quantile of the bearing life distribution.

Using a fixed value for  $\beta = 2.3$ , an approximate 95% lower confidence bound on  $\eta$  is

$$\underline{\eta} = \left( \frac{2 \sum_{i=1}^n [t_i]^{2.3}}{\chi_{(0.95, 2)}^2} \right)^{1/2.3} = \left( \frac{2 \times 10 \times [500]^{2.3}}{5.99} \right)^{1/2.3} = \left( \frac{2 \times 2692801}{5.99} \right)^{1/2.3} = 844.457 \text{ hours.}$$

From this, an approximate 95% lower confidence bound on  $t_{0.01}$  is expressed as

$$\underline{t_{0.01}} = \underline{\eta} \times (-\log[1 - 0.01])^{1/2.3} = 114.28 \text{ hours.}$$

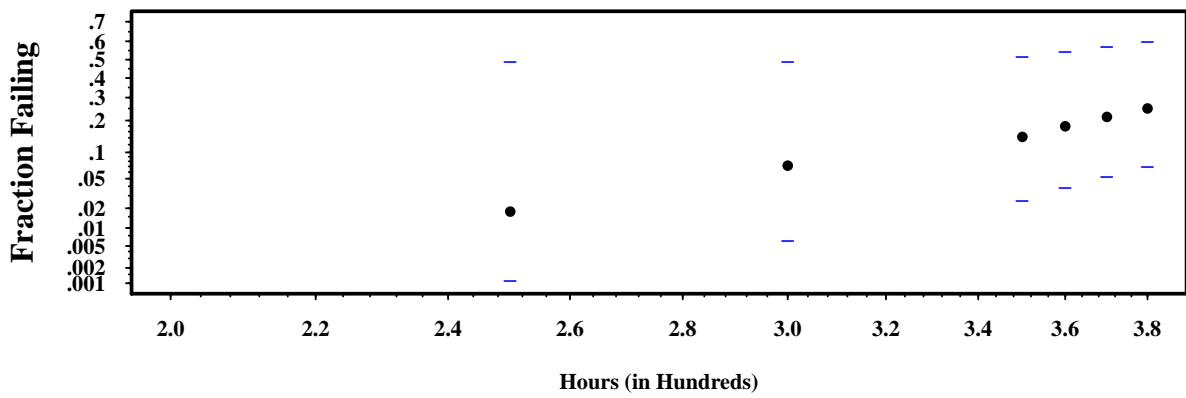
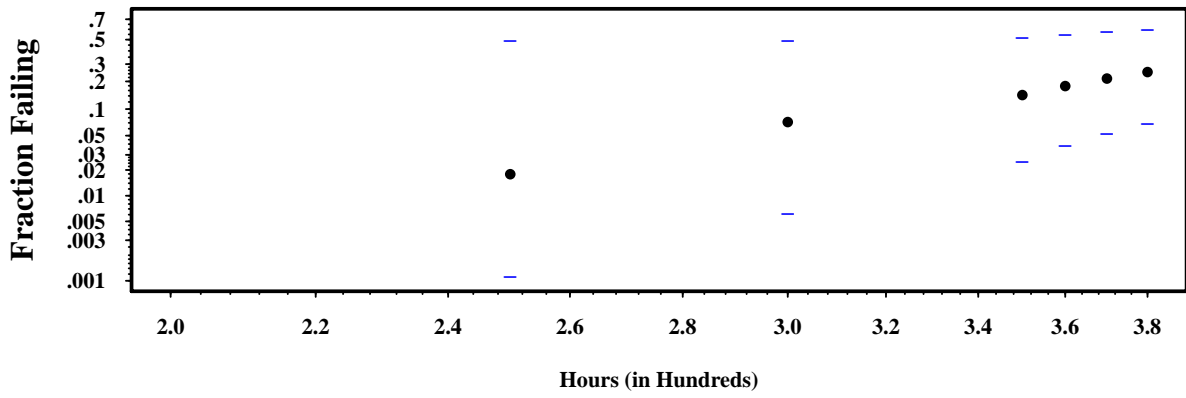
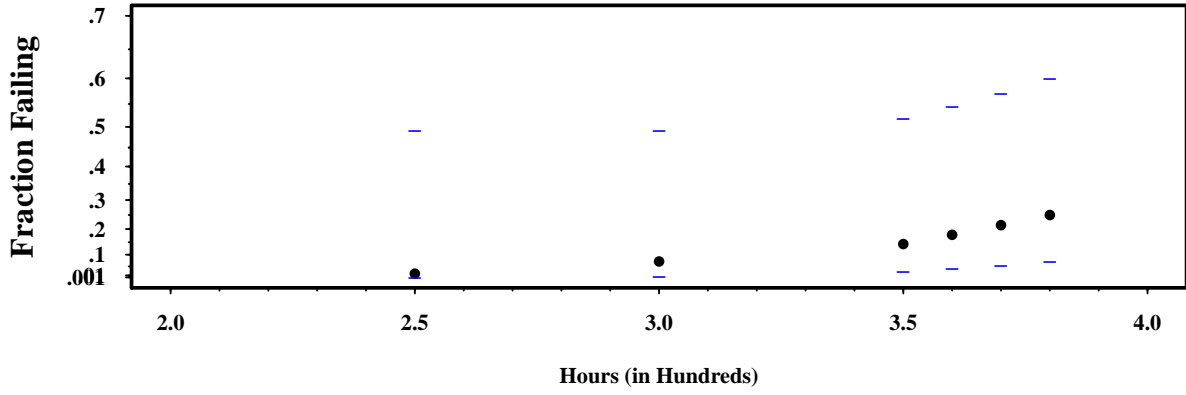


Figure 6: Probability plots of the silicon photodiode detector failure data using the exponential (top), Weibull (middle), and lognormal (bottom) distributions

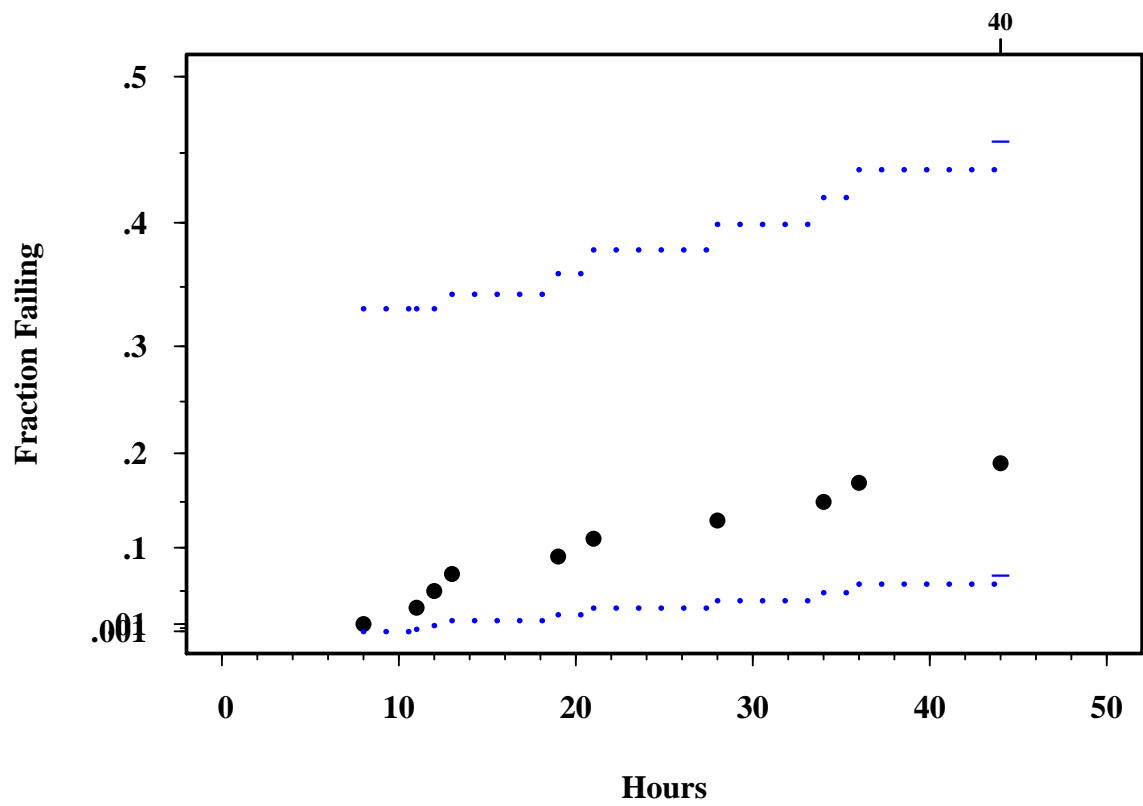


Figure 7: Exponential probability plot for the insulating material accelerated life test data (40kV)

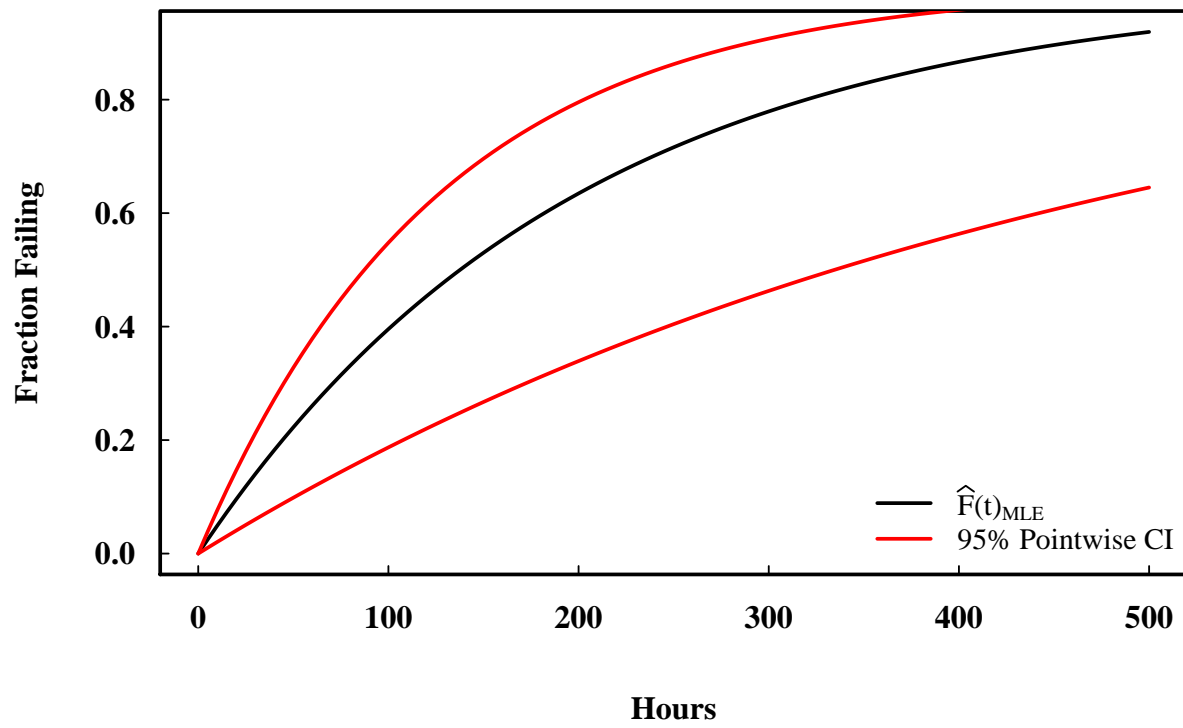


Figure 8: Maximum Likelihood Estimate of  $F(t)$  for the insulating material test data with 95% Pointwise CI

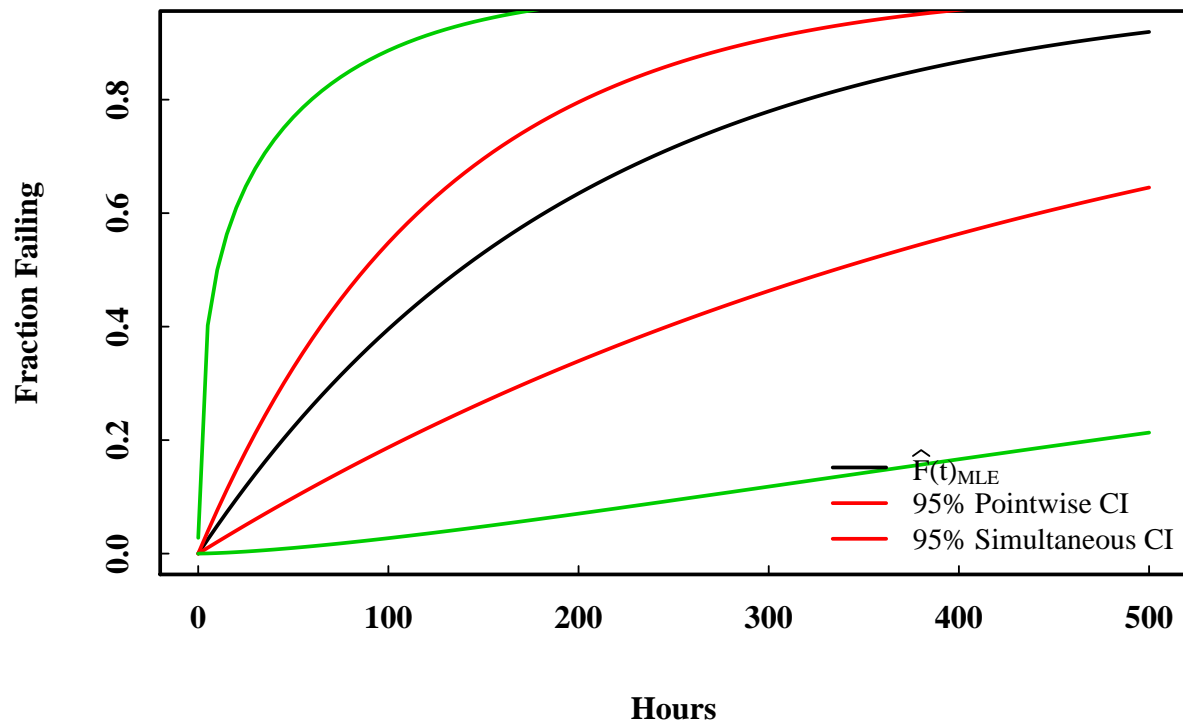


Figure 9: Maximum Likelihood Estimate of  $F(t)$  for the insulating material test data with 95% Pointwise and 95% Simultaneous Confidence Intervals

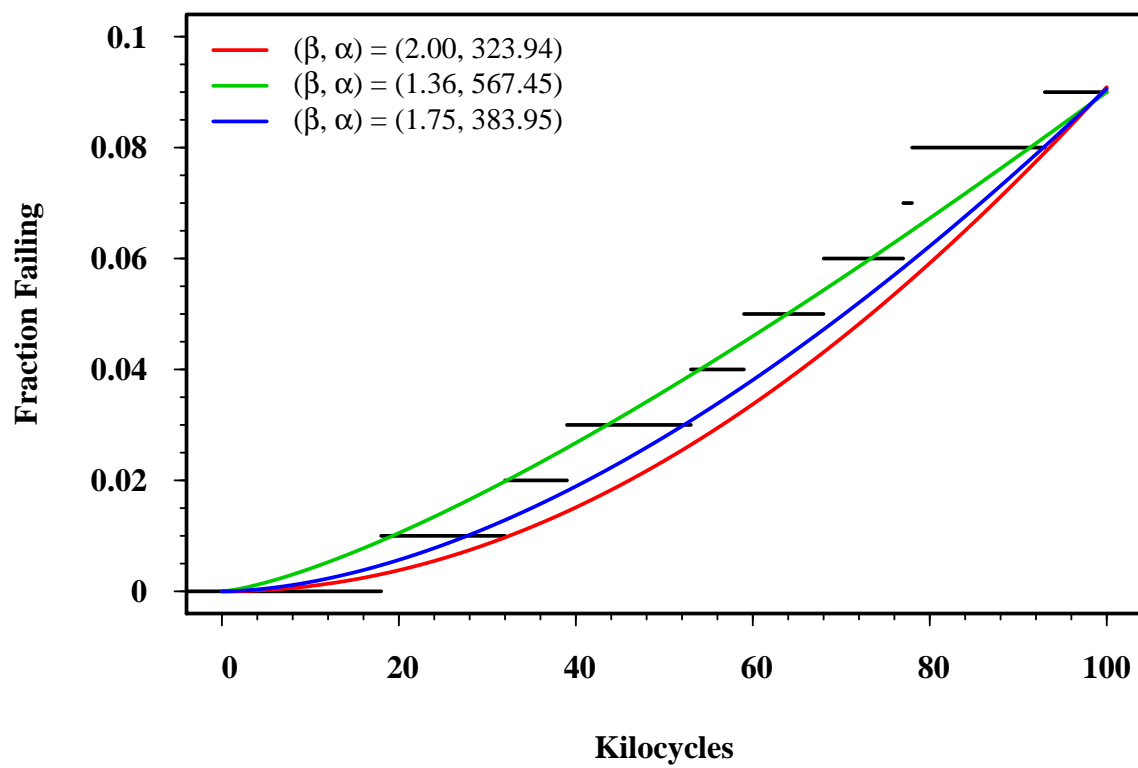


Figure 10: Fit comparison of three Weibull distributions to the titanium fatigue crack data