

Assignment 4

Eric Mittman

February 16, 2016

| t | $\widehat{F}(t)$ |
|--------|------------------|
| 0.00 | 0.00 |
| 17.88 | 0.04 |
| 28.92 | 0.09 |
| 33.00 | 0.13 |
| 41.52 | 0.17 |
| 42.12 | 0.22 |
| 45.60 | 0.26 |
| 48.40 | 0.30 |
| 51.84 | 0.35 |
| 51.96 | 0.39 |
| 54.12 | 0.43 |
| 55.56 | 0.48 |
| 67.80 | 0.52 |
| 68.64 | 0.61 |
| 68.88 | 0.65 |
| 84.12 | 0.70 |
| 93.12 | 0.74 |
| 98.64 | 0.78 |
| 105.12 | 0.83 |
| 105.84 | 0.87 |
| 127.92 | 0.91 |
| 128.04 | 0.96 |
| 173.40 | 1.00 |

Table 1: Nonparametric estimate of $F(t)$ (ball bearing data).

- 6.5 (a) See Figure 1.
- (b) See Figure 1.
- (c) See Figure 1.
- (d) Both the Weibull and lognormal probability plot transform the data to lie approximately on a straight line, suggesting that both might be adequate models. The lognormal seems to fit the data best.

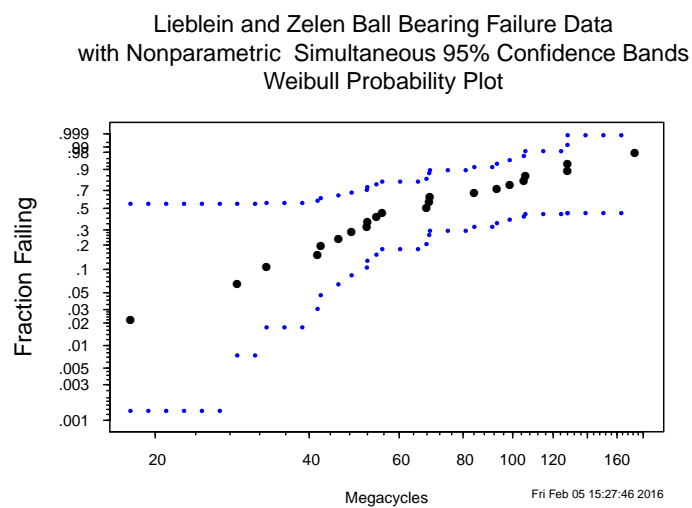
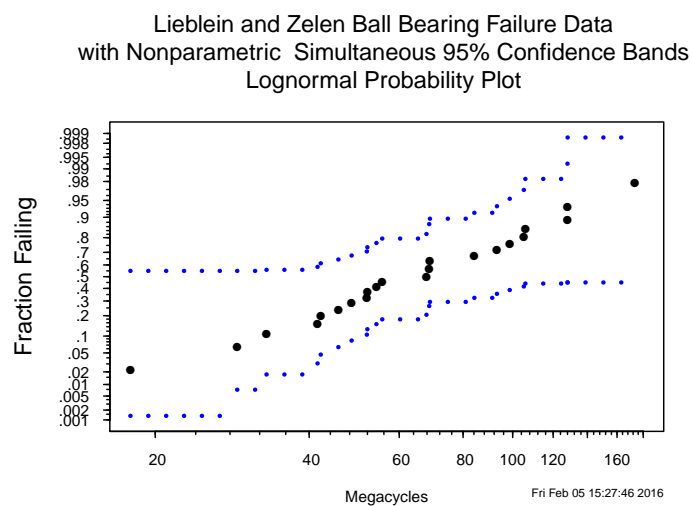
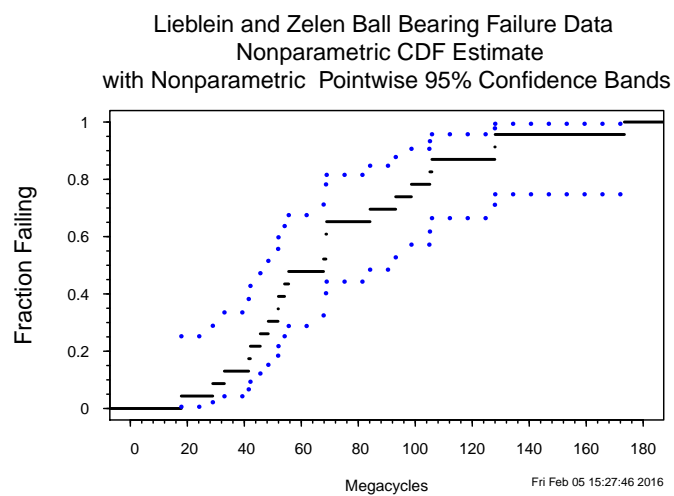


Figure 1: Empirical cdf and probability plots for the ball bearing data.

6.7 (a) See Table 2.

| t | $\widehat{F}(t)$ |
|--------|------------------|
| 0.00 | 0.00 |
| 17.88 | 0.04 |
| 28.92 | 0.09 |
| 33.00 | 0.13 |
| 41.52 | 0.17 |
| 42.12 | 0.22 |
| 45.60 | 0.26 |
| 48.40 | 0.30 |
| 51.84 | 0.35 |
| 51.96 | 0.39 |
| 54.12 | 0.43 |
| 55.56 | 0.48 |
| 67.80 | 0.52 |
| 68.64 | 0.61 |
| 68.88 | 0.65 |
| 84.12 | 0.70 |
| 93.12 | 0.74 |
| 98.64 | 0.78 |
| 105.12 | 0.83 |
| 105.84 | 0.87 |
| 127.92 | 0.91 |
| 128.04 | 0.96 |
| 173.40 | 1.00 |

Table 2: Nonparametric estimate of $F(t)$ (alloy data.)

(b) See Figure 2.

(c) See Figure 3.

(d) The fitted Weibull cdf given by the line in Figure 3 follows the data quite closely and thus seems adequate.

(e) $t_{.1}$ is not within the range of observed data, so is extrapolation. However, it is not an extreme case of extrapolation and the Weibull distribution is conservative by nature, so I would be comfortable predicting it with these data.

6.9 It appears that modeling photo-detector failure times as Weibull, exponential, lognormal and normal are all reasonable choices as they linearize the data. Of these the exponential appears the poorest fit. However, the sample size is small enough that we might expect the data to be fit many location-scale models reasonably well. The normal distribution fits well to the data that we have collected, but it is not an ideal choice since it has support on the negative numbers.

6.10 (a) Because the mean is a parameter of interest and it is also the median ($t_{.5}$)

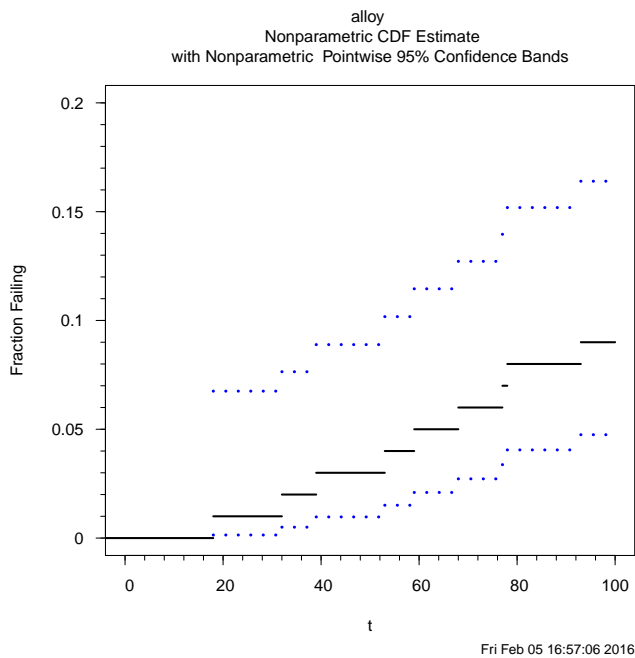


Figure 2: Empirical distribution of alloy failure data.

(b)

$$\log(-\log(1-p)) = 0 \iff 1-p = e^{-1} \iff p = 1 - e^{-1} \approx 0.6321$$

```
6.11 alloy_df <- readRDS("alloy_df.rds")

dfw <- data.frame(x = log(alloy_df[2:56,1]), y = qnorm(alloy_df[2:56,3]))
fit2 <- lm(x ~ y, dfw)

pnorm( ( log(200) - fit2$coef[1])/fit2$coef[2])

## (Intercept)
## 0.7238951
```

6.12 All of the graphical methods are going to give similar results because they are showing the same data. The empirical estimate that they give is about 0.75, which is a little higher than in 6.11. That reflects the discrepancy between the lognormal model fit and the data.

C6.16 (a) A logistic distribution. See Figure 5.

(b) A graphical estimate of the median is 1037. This is equivalent to the sample median of failure times.

(c) A numerical estimate of the mean time to failure is 1067.

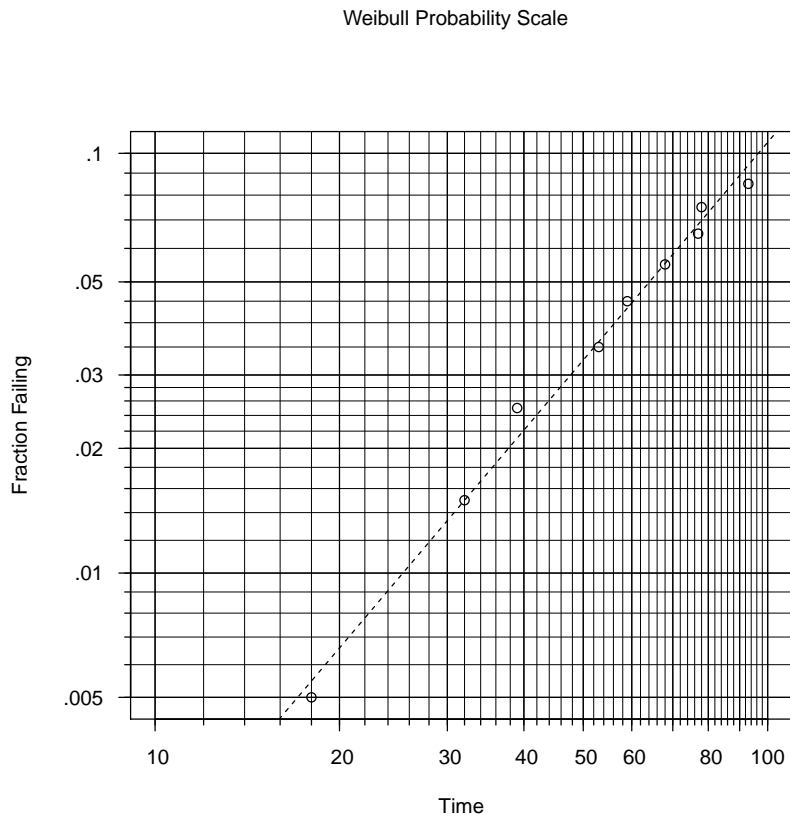


Figure 3: The fitted slope $\hat{\sigma} = 1.759$, $\hat{\beta} = \frac{1}{\widehat{\sigma}} = .569$.

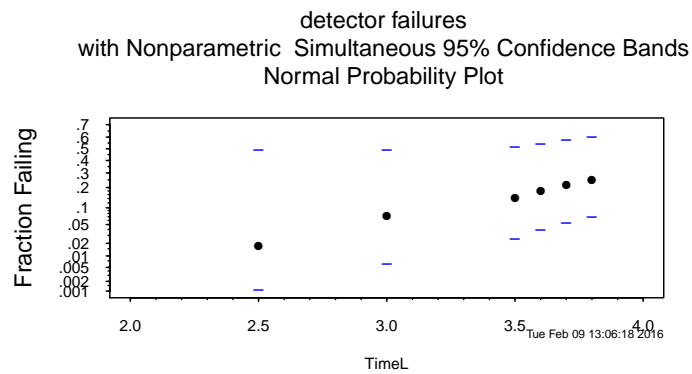
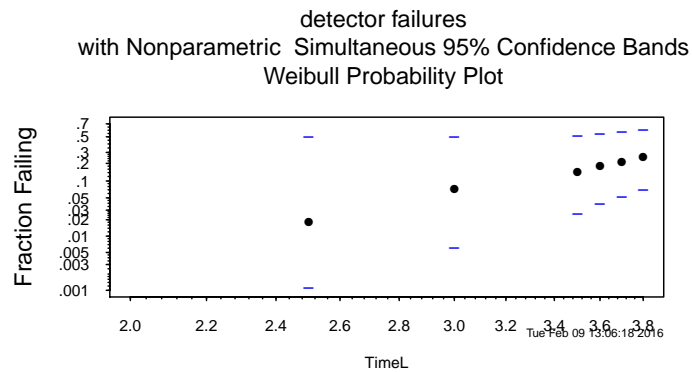
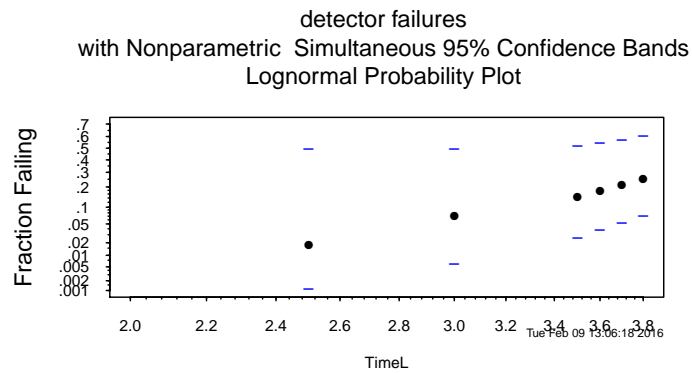
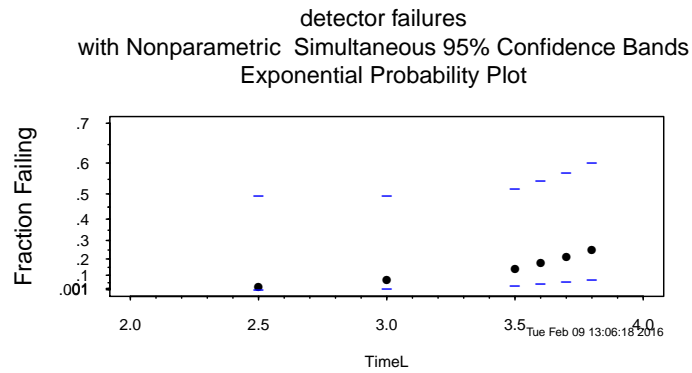


Figure 4: Probability plots for detector failure times.

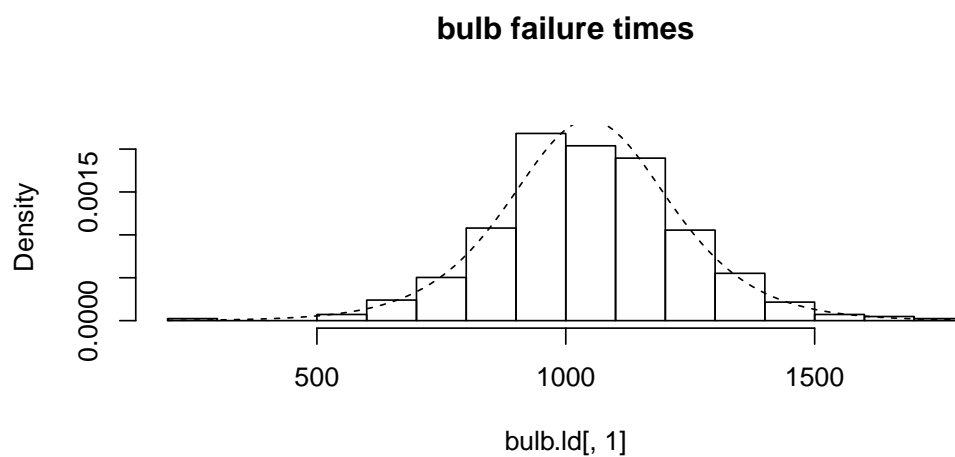
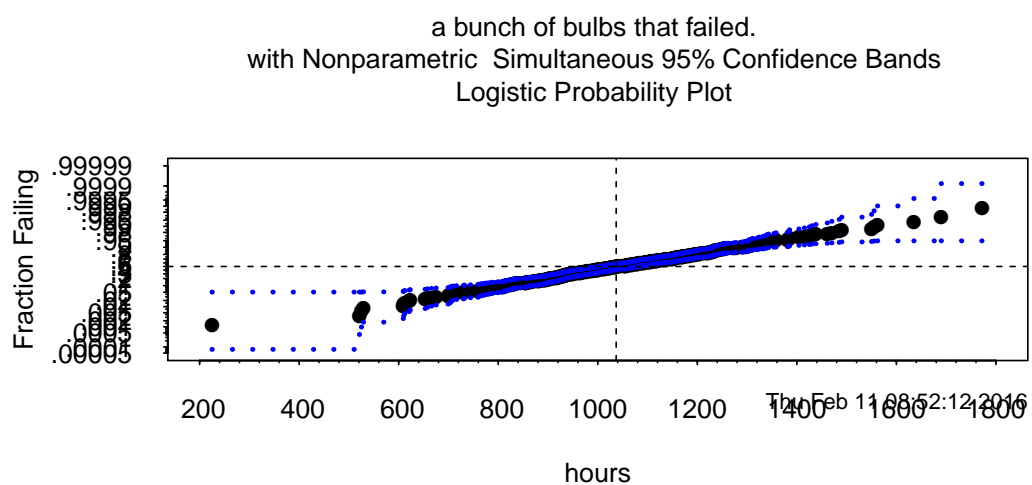


Figure 5: Logistic probability plot and histogram of bulb failure times with estimated logistic distribution function. Median time to failure is 1037.

- (d) The mean and median will be noticeably different if the distribution is asymmetric.
- (e) It appears it would be, since it is a good estimate of the typical lifespan of a bulb, and this seems like a useful number.