Problem 1

a)

$$f(t) = \frac{\partial S(t)}{\partial t} = \frac{\partial e^{-(at)^{\beta}}}{\partial t} = a\beta(at)^{\beta-1}e^{-(at)^{\beta}}$$

b)

If T is (absolutely) continuous.

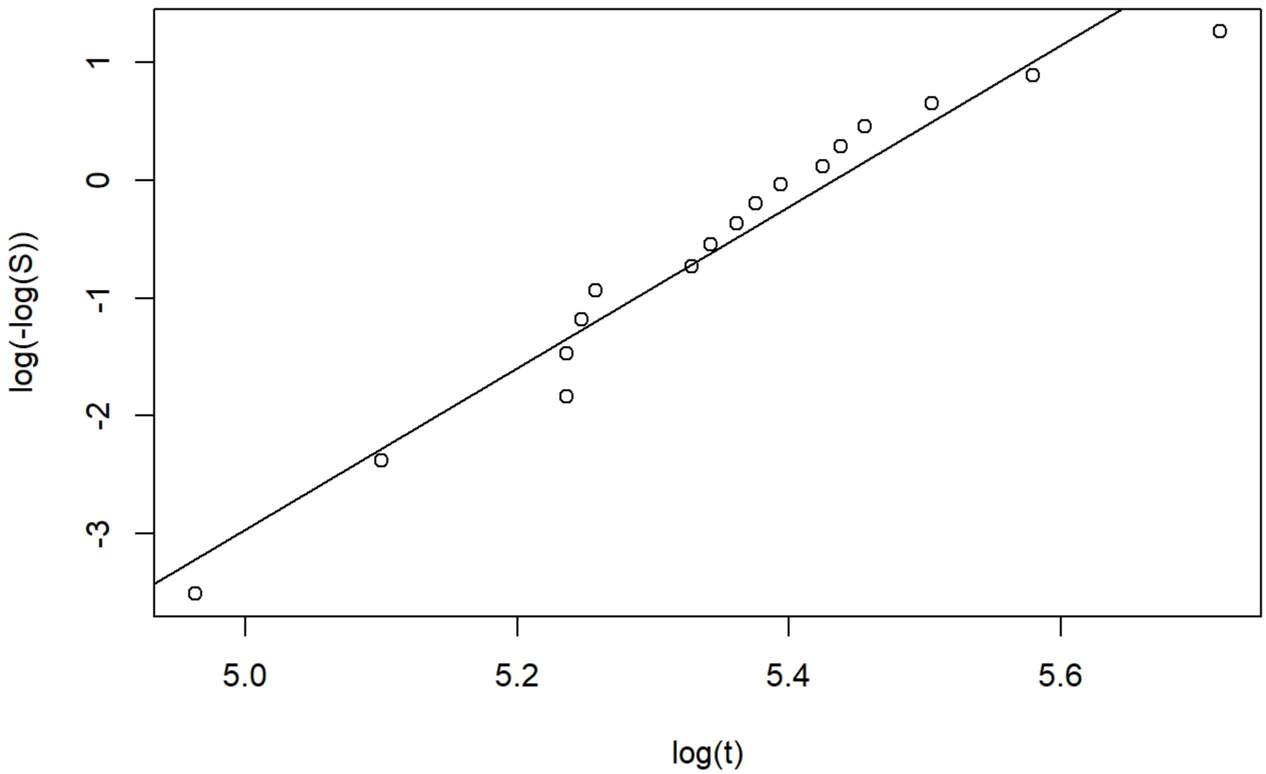
$$\lambda(t) = \frac{f(t)}{S(t)} = a\beta(at)^{\beta-1}$$

c)

$$log(-log(S(t))) = log[-log(e^{-(at)^{\beta}})] = \beta log(at) = \beta log(a) + \beta log(t)$$

d)

```
data = c(143, 164, 188, 188, 190, 192, 206, 209, 213, 216, 220, 227, 230, 234, 246, 265, 304)
n = length(data)

s = 1 - (1:n-0.5)/n

x = log(data)
y = log(-log(s))

model = lm(y ~ x)

plot(x, y, xlab = "log(t)", ylab = "log(-log(S))")
abline(model)
```



[Solution]. Since there is a strong linear association between `log(t)` and `log(-log(S))` , based on the pervious question, it shows the Weibull distribution is appropriate for these data.

e)

```
model
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Coefficients:
## (Intercept)            x
##     -37.233        6.854
```

From least square approach, I got $\hat{\beta} = 6.854$ and $\beta \hat{log(a)} = -37.233$, then $a = 0.0044$

Problem 2

a)

```
library(survival)
```

```
## Warning: package 'survival' was built under R version 3.5.3
```

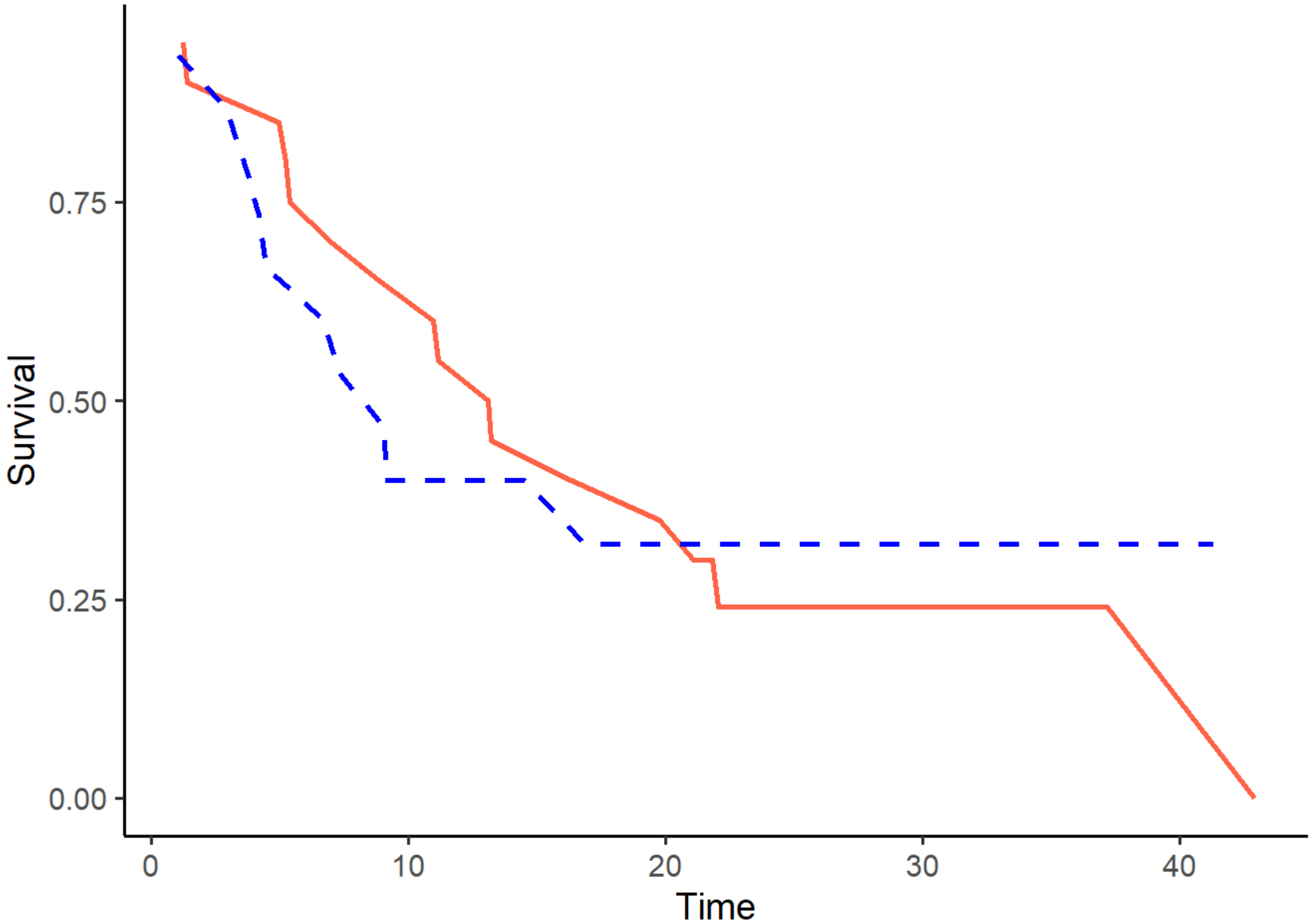
```
library(ggplot2)

A = c(1.25, 1.41, 4.98, 5.25, 5.38, 6.92, 8.89, 10.98, 11.18, 13.11, 13.21, 16.33, 19.77, 21.08,
21.84, 22.07, 31.38, 32.61, 37.18, 42.92)

B = c(1.05, 2.92, 3.61, 4.20, 4.49, 6.72, 7.31, 9.08, 9.11, 14.49, 16.85, 18.82, 26.59, 30.26, 41.34)

deltaA = c(rep(1,14),0,1,0,0,0,1)
deltaB = c(rep(1,9),0,1,0,0,0,0)

kmA = survfit(Surv(A,deltaA)~1, type="kaplan-meier")
cat("Kaplan Meier estimates of SA: \n", kmA$surv, "\n")
```

```
## Kaplan Meier estimates of SA:
##  0.95 0.9 0.85 0.8 0.75 0.7 0.65 0.6 0.55 0.5 0.45 0.4 0.35 0.3 0.3 0.24 0.24 0.24 0.24 0
```

```
kmB = survfit(Surv(B,deltaB)~1, type="kaplan-meier")
cat("Kaplan Meier estimates of SB: \n", kmB$surv)
```

```
## Kaplan Meier estimates of SB:
##  0.9333333 0.8666667 0.8 0.7333333 0.6666667 0.6 0.5333333 0.4666667 0.4 0.4 0.32 0.32 0.32 0.32 0.32
```

```
ggplot()+
  geom_line(aes(kmA$time, kmA$surv), col = "tomato", size = 1)+
  geom_line(aes(kmB$time, kmB$surv), col = "blue", size = 1, linetype = "dashed")+
  xlab("Time")+
  ylab("Survival")+
  theme_classic(14)
```



b)

```
summary(kmA)
```

```
## Call: survfit(formula = Surv(A, deltaA) ~ 1, type = "kaplan-meier")
##
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   1.25     20       1     0.95  0.0487        0.859        1.000
##   1.41     19       1     0.90  0.0671        0.778        1.000
##   4.98     18       1     0.85  0.0798        0.707        1.000
##   5.25     17       1     0.80  0.0894        0.643        0.996
##   5.38     16       1     0.75  0.0968        0.582        0.966
##   6.92     15       1     0.70  0.1025        0.525        0.933
##   8.89     14       1     0.65  0.1067        0.471        0.897
##  10.98     13       1     0.60  0.1095        0.420        0.858
##  11.18     12       1     0.55  0.1112        0.370        0.818
##  13.11     11       1     0.50  0.1118        0.323        0.775
##  13.21     10       1     0.45  0.1112        0.277        0.731
##  16.33      9       1     0.40  0.1095        0.234        0.684
##  19.77      8       1     0.35  0.1067        0.193        0.636
##  21.08      7       1     0.30  0.1025        0.154        0.586
##  22.07      5       1     0.24  0.0980        0.108        0.534
##  42.92      1       1     0.00     NaN           NA           NA
```

```
summary(kmB)
```

```
## Call: survfit(formula = Surv(B, deltaB) ~ 1, type = "kaplan-meier")
##
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   1.05     15       1    0.933  0.0644        0.815        1.000
##   2.92     14       1    0.867  0.0878        0.711        1.000
##   3.61     13       1    0.800  0.1033        0.621        1.000
##   4.20     12       1    0.733  0.1142        0.540        0.995
##   4.49     11       1    0.667  0.1217        0.466        0.953
##   6.72     10       1    0.600  0.1265        0.397        0.907
##   7.31      9       1    0.533  0.1288        0.332        0.856
##   9.08      8       1    0.467  0.1288        0.272        0.802
##   9.11      7       1    0.400  0.1265        0.215        0.743
##  16.85      5       1    0.320  0.1239        0.150        0.684
```

From tables above, it shows the 95% confidence interval for SA(10) is [0.471, 0.897], and the 95% confidence interval for SB(10) is [0.215, 0.743]