Problem 2.

```
library(dplyr)
library(latex2exp)
library(knitr)
fan.data = read.csv("Fan.csv")
fan = fan.data %>%
    mutate(di = ifelse(Censoring.Indicator=="Fail",Count,0)) %>%
    mutate(ri = ifelse(Censoring.Indicator=="Censored",Count,0)) %>%
    group_by(Hours) %>%
    summarise(di = sum(di), ri = sum(ri)) %>%
    ungroup()
kable(fan)
```

Hours	di	ri
450	1	0
460	0	1
1150	2	0
1560	0	1
1600	1	0
1660	0	1
1850	0	5
2030	0	3
2070	2	0
2080	1	0
2200	0	1
3000	0	4
3100	1	0
3200	0	1
3450	1	0
3750	0	2
4150	0	4
4300	0	4
4600	1	0
4850	0	4

Hours	di	ri
5000	0	3
6100	1	3
6300	0	1
6450	0	2
6700	0	1
7450	0	1
7800	0	2
8100	0	2
8200	0	1
8500	0	3
8750	1	2
9400	0	1
9900	0	1
10100	0	3
11500	0	1

(b)

Likelihood function

$$L(\theta\ ;\ t)\ =\ \prod_{i=1}^{n}\left(\frac{1}{\theta}e^{-\frac{t_i}{\theta}}\right)^{d_i}\left(e^{-\frac{t_i}{\theta}}\right)^{r_i}$$

log likelihood function

$$l(\theta\ ;\ t)\ =\ \sum_{i=1}^{n} d_i\left(-\log(\theta)-\frac{t_i}{\theta}\right)+r_i\left(-\frac{t_i}{\theta}\right)$$

Then MLE of θ

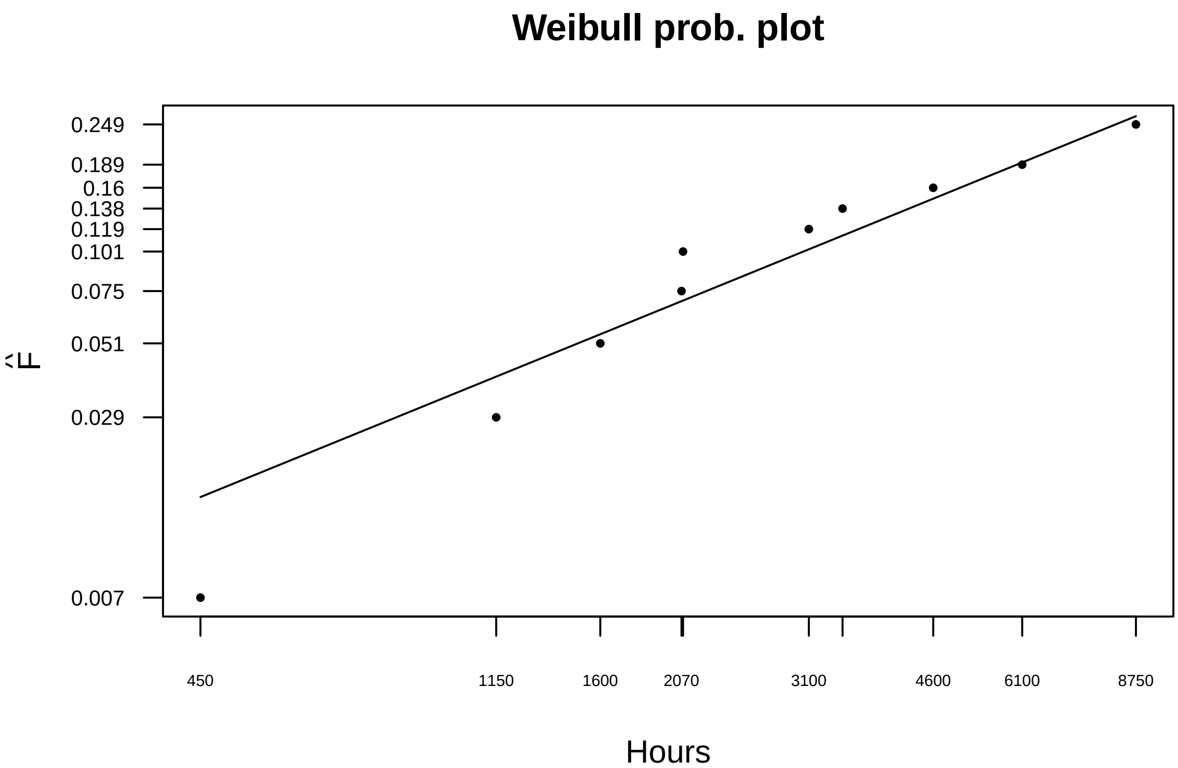
$$\hat{\theta}\ =\ \arg\max\ l(\theta\ ;\ t)\ =\ 28676.5$$

Compute the MLE of $F(t)$

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

```
neg_logL = function(theta, t=fan$Hours, di=fan$di, ri=fan$ri) {
    l = sum(di*(-log(theta)-t/theta)+ri*(-t/theta))
    return(-l)
}
op = optim(5, neg_logL, hessian = T)
theta_mle = op$par
```

```
fan = fan %>%
    mutate(ni = c(70,70-cumsum(di+ri))[-36], pi = di/ni,
           Si = Reduce("*",1-pi,acc=T), F.hat = 1-Si) %>% # non parametric
    mutate(F.hat_exp = pexp(Hours,1/theta_mle)) # parametric
qsev=function(p){
  log(qweibull(p,1,1))
}
psev=function(x){
  pweibull(exp(x),1,1)
}
tab = fan %>% select(Hours,di,F.hat, F.hat_exp) %>%
    filter(di>0)
yi=(tab$F.hat[-1]+tab$F.hat[-10])/2;yi=c(tab$F.hat[1]/2,yi) # plotting position
plot(log(tab$Hours), qsev(yi), pch = 16, cex = 0.6,
     yaxt = "n", xaxt = "n", xlab = "Hours", ylab = TeX("$\\hat{F}$"), main = "Weibull prob. plot")
points(log(tab$Hours), qsev(tab$F.hat_exp), type = "l")
axis(1,log(tab$Hours),round(tab$Hours,2),cex.axis=0.5)
axis(2,qsev(yi),round(yi,3),las=1,cex.axis=0.7)
```



(c)

$$\hat{F}(1250) \;=\; 1 - \exp\left(-\frac{1250}{\hat{\theta}}\right) \;=\; 0.04265332$$

```
F_hat.1250 = pexp(1250,1/theta_mle)
F_hat.1250
```

```
## [1] 0.04265332
```

The 95% confidence interval

$$[F_L(1250)\ ,\ F_U(1250)] \;=\; \left[\frac{\hat{F}}{\hat{F} + \left(1-\hat{F}\right) \times w}\ ,\ \frac{\hat{F}}{\hat{F} + \left(1-\hat{F}\right)/w}\right]$$

where

$$w \;=\; \exp\left[\frac{Z_{0.975}\ se\left(\hat{F}\right)}{\hat{F}\left(1-\hat{F}\right)}\right]$$

and by Delta method

$$se\left(\hat{F}\right) \;=\; \sqrt{\left[\frac{\partial F}{\partial \theta}\right]^2_{\theta=\hat{\theta}}}\ se\left(\hat{\theta}\right) \;=\; \left[\frac{1250}{\hat{\theta}^2 \exp\left(\frac{-1250}{\hat{\theta}}\right)}\right] se\left(\hat{\theta}\right)$$

```r
se.theta = sqrt(1/op$hessian[1,1])
se.F = 1250/(theta_mle^2*exp(-1250/theta_mle))*se.theta
w = exp(qnorm(0.975)*se.F/(F_hat.1250*(1-F_hat.1250)))
c(F_hat.1250/(F_hat.1250+(1-F_hat.1250)*w),F_hat.1250/(F_hat.1250+(1-F_hat.1250)/w))
```

```
## [1] 0.02296929 0.07786171
```

(d)

$$\hat{t}_{0.1} \;=\; \hat{F}^{-1}(0.1) \;=\; -\hat{\theta}\log(1-0.1) \;=\; 3021.371$$

```r
t_0.1 = qexp(0.1, 1/theta_mle)
t_0.1
```

```
## [1] 3021.371
```

The 95% confidence interval

$$[t_L \;,\; t_U] \;=\; [\hat{t}_{0.1}/w \;,\; \hat{t}_{0.1} \times w]$$

where

$$w \;=\; \exp\left[\frac{Z_{0.975}\; se\left(\hat{t}_{0.1}\right)}{\hat{t}_{0.1}}\right]$$

and by Delta method

$$se\left(\hat{t}_{0.1}\right) \;=\; \sqrt{\left[\frac{\partial t_{0.1}}{\partial \theta}\right]^2_{\theta=\hat{\theta}}}\; se\left(\hat{\theta}\right) \;=\; \sqrt{[\log(0.9)]^2}\; se\left(\hat{\theta}\right)$$

```r
se.t = sqrt((log(0.9))^2)*se.theta
w = exp(qnorm(0.975)*se.t/t_0.1)
c(t_0.1/w, t_0.1*w)
```

```
## [1] 1702.966 5360.462
```

(e)

Likelihood funciton

$$L\left(\theta = (\mu,\sigma)\;;\; t\right) \;=\; \prod_{i=1}^{n}\left\{\frac{1}{\sigma t_i}\phi_{sev}\left[\frac{\log(t_i)-\mu}{\sigma}\right]\right\}^{d_i}\;\left\{1-\Phi_{sev}\left[\frac{\log(t_i)-\mu}{\sigma}\right]\right\}^{r_i}$$

Log likelihood function

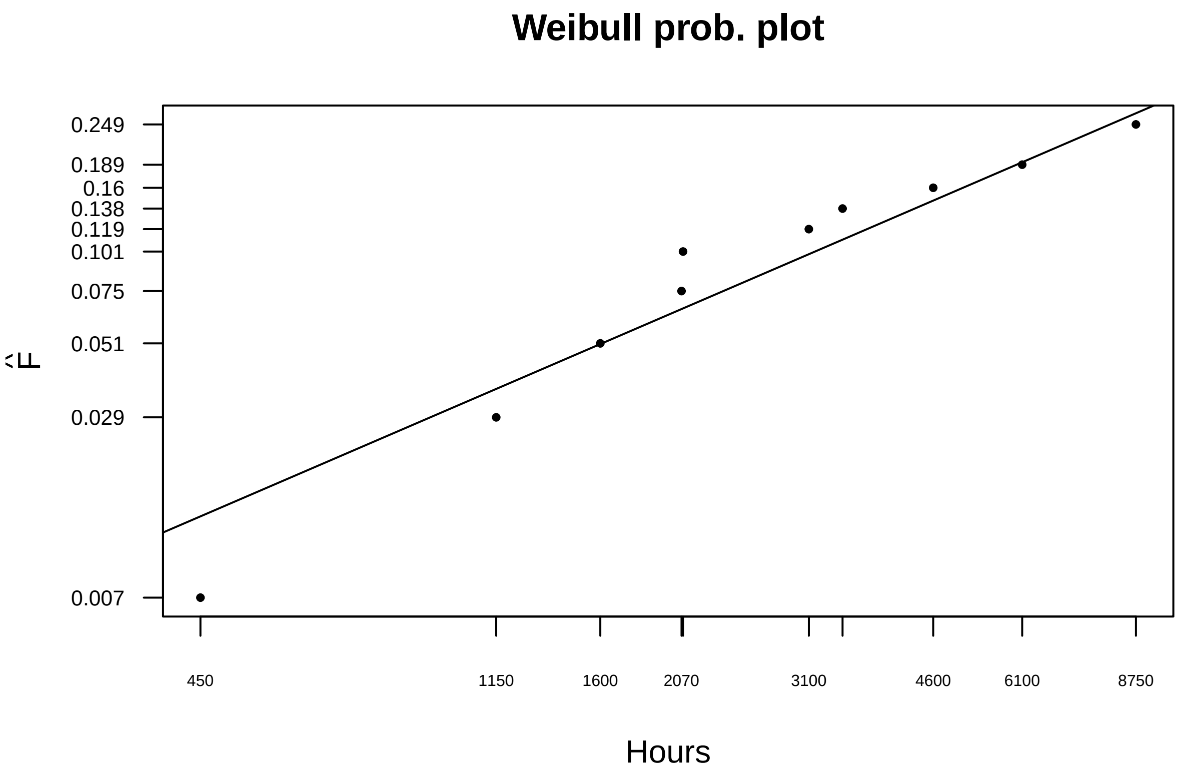
$$l\left(\theta=(\mu,\sigma)\ ;\ t\right)\ =\ \sum_{i=1}^{n} d_i\left[\frac{\log(t_i)-\mu}{\sigma}-\exp\left(\frac{\log(t_i)-\mu}{\sigma}\right)-\log(\sigma)-\log(t_i)\right]\ +\ r_i\left[-\exp\left(\frac{\log(t_i)-\mu}{\sigma}\right)\right]$$

Then MLE of $\theta\ =\ (\eta\ ,\ \beta)$

$$\hat{\theta}\ =\ (\hat{\mu}\ ,\ \hat{\sigma})\ =\ \arg\max l(\mu\ ,\ \sigma\ ;\ t)\ =\ (10.1770769\ ,\ 0.9448977)$$

```
ld=function(x,mu,sig){ #log density function
  z=(log(x)-mu)/sig
  (z-exp(z))-log(sig)-log(x)
}
lS=function(x,mu,sig){ #log survival function
  z=(log(x)-mu)/sig
  -exp(z)
}
logL=function(theta,ti=fan$Hours,di=fan$di,ri=fan$ri){
  mu=theta[1];sig=abs(theta[2])
  l=0
  l = sum(di*ld(ti,mu,sig)+ri*lS(ti,mu,sig))
  -l
}
op.wei = optim(c(2,5),logL,hessian = T)
mu.mle = op.wei$par[1] ; sig.mle = op.wei$par[2]
eta.mle = exp(mu.mle) ; beta.mle = 1/sig.mle
fan = fan %>% mutate(F.hat_wei = 1-exp(-exp((log(Hours)-mu.mle)/sig.mle)))
```

```
plot(log(tab$Hours), qsev(yi), pch = 16, cex = 0.6,
     yaxt = "n", xaxt = "n", xlab = "Hours", ylab = TeX("$\\hat{F}$"),
     main = "Weibull prob. plot")
#points(log(fan$Hours),qsev(fan$F.hat_wei), type = "l")
abline(-mu.mle/sig.mle, 1/sig.mle)
axis(1,log(tab$Hours),round(tab$Hours,2),cex.axis=0.5)
axis(2,qsev(yi),round(yi,3),las=1,cex.axis=0.7)
```



$$\hat{F}(1250) \;=\; \Phi_{sev}\left(\frac{\log(1250)-\hat{\mu}}{\hat{\sigma}}\right) \;=\; 0.03902108$$

```
F1250_wei = 1-exp(-exp((log(1250)-mu.mle)/sig.mle))
F1250_wei
```

```
## [1] 0.03902108
```

The 95% confidence interval

$$[F_L(1250)\;,\;F_U(1250)] \;=\; \left[\frac{\hat{F}}{\hat{F}+\left(1-\hat{F}\right)\times w}\;\;,\;\;\frac{\hat{F}}{\hat{F}+\left(1-\hat{F}\right)/w}\right]$$

where

$$w \;=\; \exp\left[\frac{Z_{0.975}\;se\left(\hat{F}\right)}{\hat{F}\left(1-\hat{F}\right)}\right]$$

and by Delta method

$$se\left(\hat{F}\right) \;=\; \sqrt{\left[\frac{\partial F(1250)}{\partial\theta}\right]^T_{\theta=\hat{\theta}}\hat{Var}\left(\hat{\theta}\right)\left[\frac{\partial F(1250)}{\partial\theta}\right]_{\theta=\hat{\theta}}}$$

```r
z = (log(1250)-mu.mle)/sig.mle
partial = c(-1/sig.mle*exp(z-exp(z)), -z/sig.mle*exp(z-exp(z)))
se.F = sqrt(partial %*% solve(op.wei$hessian) %*% partial)[1,1]
w = exp(qnorm(0.975)*se.F/(F1250_wei*(1-F1250_wei)))
c(F1250_wei/(F1250_wei+(1-F1250_wei)*w),F1250_wei/(F1250_wei+(1-F1250_wei)/w))
```

```
## [1] 0.01450743 0.10072267
```

(f)

The 95% confidence interval for σ

$$[\sigma_L \ , \ \sigma_U] \ = \ \hat{\sigma} \ \pm \ Z_{0.975} \ se\,(\hat{\sigma})$$

```r
se.sig = sqrt(diag(solve(op.wei$hessian)))[2]
c(sig.mle-qnorm(0.975)*se.sig, sig.mle+qnorm(0.975)*se.sig)
```

```
## [1] 0.4754929 1.4143024
```

Then the 95% confidence interval for shape parameter β

$$[\beta_L \ , \ \beta_U] \ = \ [1/\sigma_U \ , \ 1/\sigma_L]$$

```r
c(1/(sig.mle+qnorm(0.975)*se.sig) , 1/(sig.mle-qnorm(0.975)*se.sig))
```

```
## [1] 0.7070624 2.1030808
```

We can see that $\beta \ = \ 1$ is fall in the 95% confidence interval, so it may be proper to fit an exponential distribution to describe the diesel generator fan data. The hazard function of the fan may be a constant. Thus we regard the old one and the young one equally.

(g)

$$\hat{t}_{0.1} \ = \ \hat{F}^{-1}(0.1) \ = \ \exp\left[\log(-\log(1-0.1)) \ \hat{\sigma} \ + \ \hat{\mu}\right] \ = \ 3136.021$$

```r
t_0.1 = exp(log(-log(0.9))*sig.mle+mu.mle)
t_0.1
```

```
## [1] 3136.021
```

By Delta method

$$se\left(\hat{t}_{0.1}\right) \;=\; \sqrt{\left[\frac{\partial t_{0.1}}{\partial \theta}\right]^{T}_{\theta=\hat{\theta}} \hat{Var}\left(\hat{\theta}\right) \left[\frac{\partial t_{0.1}}{\partial \theta}\right]_{\theta=\hat{\theta}}}$$

where

$$\left[\frac{\partial t_{0.1}}{\partial \theta}\right] \;=\; \begin{bmatrix} \exp\left[\log(-\log(0.9))\ \sigma \ + \ \mu\right] \\ \log(-\log(0.9)) \exp\left[\log(-\log(0.9))\ \sigma \ + \ \mu\right] \end{bmatrix}$$

```r
partial = c(exp(log(-log(0.9))*sig.mle+mu.mle),
            log(-log(0.9))*exp(log(-log(0.9))*sig.mle+mu.mle))
se.t = sqrt(partial %*% solve(op.wei$hessian) %*% partial)[1,1]
se.t
```

```
## [1] 993.6667
```

The 95% confidence interval based on $Z_{\hat{t}_{0.1}} \ \dot{\sim} \ N(0,1)$

$$[t_L\ ,\ t_U] \;=\; \hat{t}_{0.1} \ \pm\ Z_{0.975}\ se(\hat{t}_{0.1})$$

```r
c(t_0.1-qnorm(0.975)*se.t, t_0.1+qnorm(0.975)*se.t)
```

```
## [1] 1188.470 5083.572
```

The 95% confidence interval based on $Z_{\log(\hat{t}_{0.1})} \ \dot{\sim} \ N(0,1)$

$$[t_L\ ,\ t_U] \;=\; \left[\hat{t}_{0.1}/w\ ,\ \hat{t}_{0.1} \times w\right]$$

where

$$w \;=\; \exp\left[\frac{Z_{0.975}\ se\left(\hat{t}_{0.1}\right)}{\hat{t}_{0.1}}\right]$$

```r
w = exp(qnorm(0.975)*se.t/t_0.1)
c(t_0.1/w, t_0.1*w)
```

```
## [1] 1685.275 5835.622
```

(h)

It just happened. The smallest observation does not have to fall into the CI for $t_{0.1}$.

Problem 3.

(a)

```
bearing = read.csv("Bearingcage.csv")
bearing = bearing %>% mutate(di = ifelse(Censoring.Indicator=="Failed",Count,0),
                             ri = ifelse(Censoring.Indicator=="Censored",Count,0),
                             ni = c(1703,1703-cumsum(Count))[-26],
                             pi = di/ni, Si = Reduce("*",1-pi,acc=T), Fi = 1-Si)
```

```
logL=function(theta,ti=bearing$Hours,di=bearing$di, ri = bearing$ri){
  mu=theta[1];sig=abs(theta[2])
  l = sum(di*ld(ti,mu,sig)+ri*lS(ti,mu,sig))
  -l
}
op.bear=optim(c(20,1),logL,hessian=T)
mu.hat = op.bear$par[1] ; sig.hat = op.bear$par[2]
se.mu = sqrt(diag(solve(op.bear$hessian)))[1] ; se.sig = sqrt(diag(solve(op.bear$hessian)))[2]
```

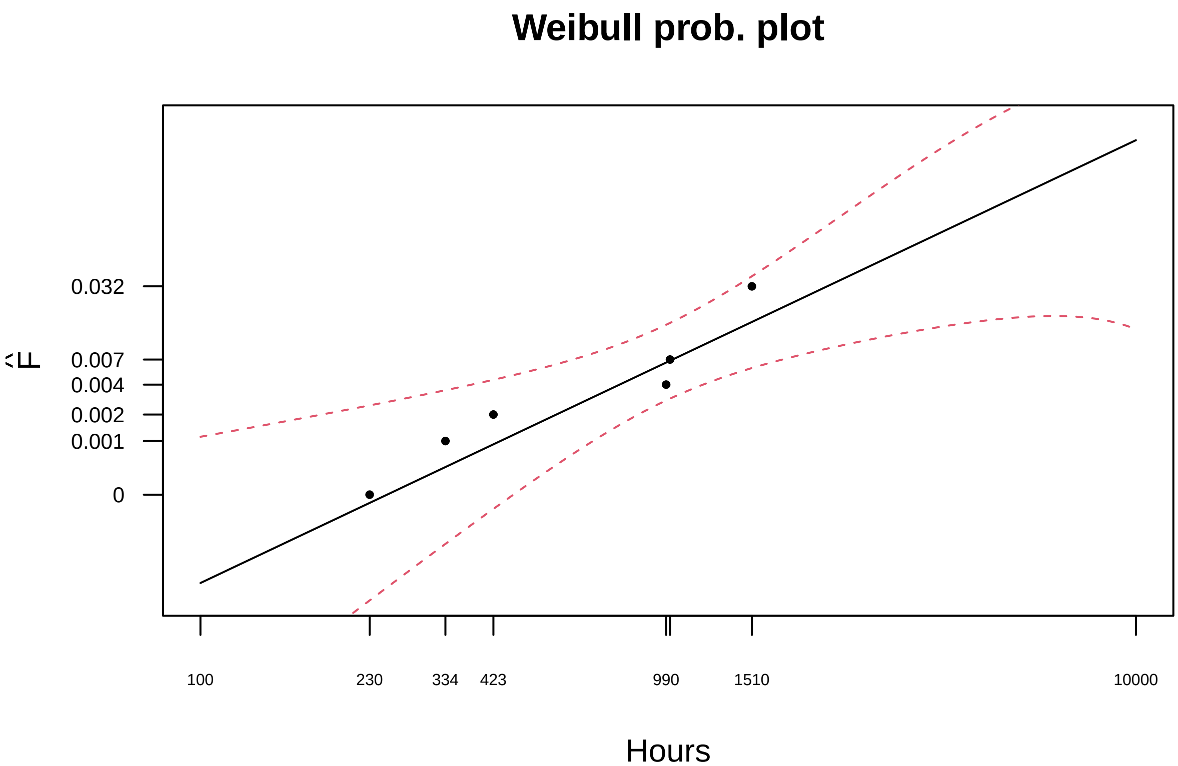
The MLE for $(\mu\ ,\ \sigma)$

```
c(mu.hat,sig.hat)
```

```
## [1] 9.3759872 0.4914871
```

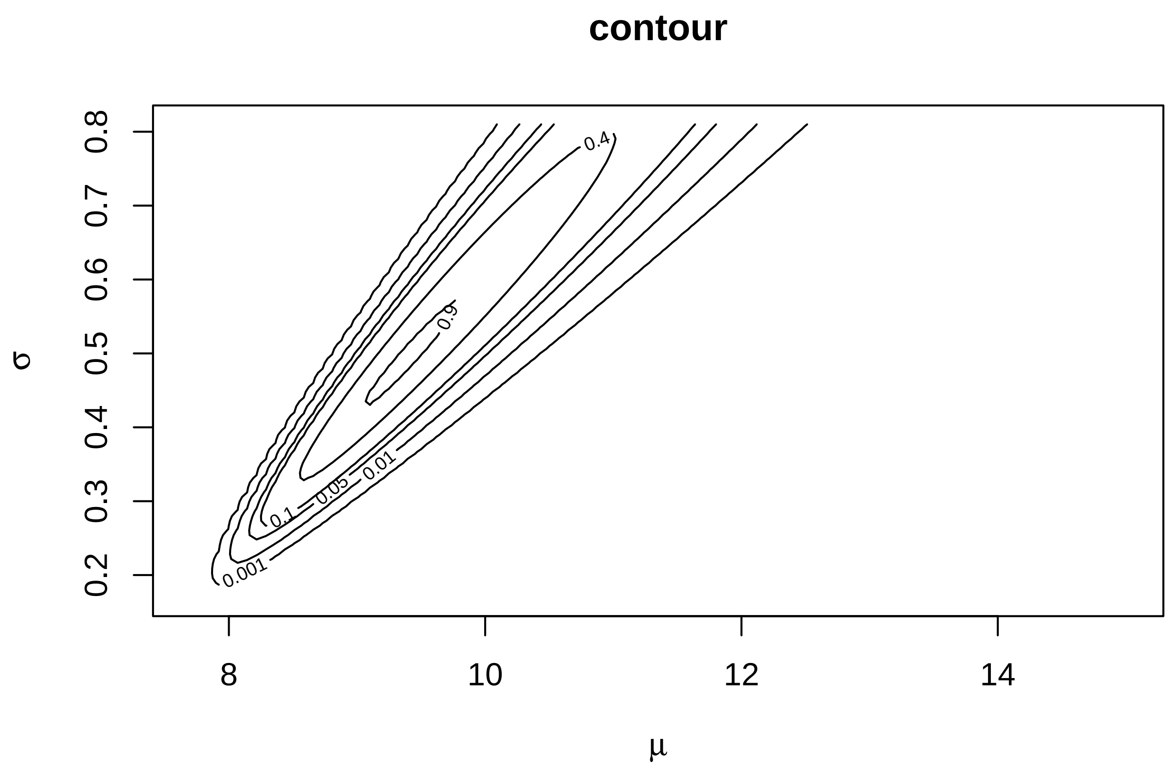
```
bearing = bearing %>% mutate(F.hat_wei = pweibull(Hours,1/sig.hat,exp(mu.hat)))
tab = bearing %>% select(Hours,di,Fi,F.hat_wei) %>%
    filter(di > 0)
yi = (tab$Fi[-1]+tab$Fi[-6])/2 ; yi = c(tab$Fi[1]/2, yi)
plot(log(tab$Hours),qsev(yi), xlim = c(log(100),log(10000)),ylim = c(-10,0),
     pch=16,cex=0.6,yaxt = "n", xaxt = "n", xlab = "Hours", ylab = TeX("$\\hat{F}$"),
     main = "Weibull prob. plot")
curve((x-mu.hat)/sig.hat, log(100),log(10000), add=T)
axis(1,log(tab$Hours),round(tab$Hours,2),cex.axis=0.5)
axis(1,c(log(100),log(10000)),c(100,10000),cex.axis=0.5)
axis(2,qsev(yi),round(yi,3),las=1,cex.axis=0.7)
dsev = function(z) {
```

```
    exp(z-exp(z))
}
CI = function(x,mu=mu.hat,sig=sig.hat,hessian=op.bear$hessian) {
    z = (log(x)-mu)/sig
    Ft = psev(z)
    B = c(-dsev(z)/sig, -z*dsev(z)/sig)
    F.se = sqrt(B%*%solve(hessian)%*%B)[1,1]
    w = exp(qnorm(0.975)*F.se/(Ft*(1-Ft)))
    return(c(Ft/(Ft+(1-Ft)*w), Ft/(Ft+(1-Ft)/w)))
}
xl = exp(seq(log(100),log(10000),len=100))
y.low = c()
y.up = c()
for (i in 1:100) {
    y.low[i] = CI(xl[i])[1]
    y.up[i] = CI(xl[i])[2]
}
points(log(xl), qsev(y.low), col=2, lty=2, type="l")
points(log(xl), qsev(y.up), col=2, lty=2, type="l")
```



(b)

```
x = seq(7.7,15,len=100)
y = y=seq(0.17,0.81,len=100)
z = matrix(0,100,100)
for(i in 1:100){
  for(j in 1:100){
    exp(-logL(c(x[i],y[j]))+op.bear$value)
    z[i,j]=exp(-logL(c(x[i],y[j]))+op.bear$value)
  }
}
contour(x,y,z,levels=c(0.9,0.4,0.1,0.05,0.01,0.001),main="contour", xlab = TeX("$\\mu$"),ylab=TeX("$\\s
```



We can see that μ and σ have positive correlation.

(c)

```
raw=read.csv("BearingCage.csv")
ti=raw[,1]
di=(raw[,2]=="Failed")
wi=raw[,3]
qsev=function(p){
  log(qweibull(p,1,1))
}
psev=function(x){
  pweibull(exp(x),1,1)
}
ld=function(x,mu,sig){ #log density function
  z=(log(x)-mu)/sig
  (z-exp(z))-log(sig)-log(x)
```

```
}
lS=function(x,mu,sig){ #log survival function
  z=(log(x)-mu)/sig
  -exp(z)
}
logL=function(theta,ti,di,wi){
  mu=theta[1];tp=theta[2];
  inv=log(qexp(0.1))
  sig=(log(tp)-mu)/inv
  if(sig<0.01)sig=0.01
  l=0
  for(i in 1:length(ti)){
    l=l+di[i]*wi[i]*ld(ti[i],mu,sig)+(1-di[i])*wi[i]*lS(ti[i],mu,sig)
  }
  -l
}
op=optim(c(10,500),ti=ti,di=di,wi=wi,logL,hessian=T)
mle=op$p
se=diag(solve(op$h))^0.5
```

The Wald 95% confidence interval for $t_{0.1}$

$$[t_L \ , \ t_U] \ = \ \hat{t}_{0.1} \ \pm \ Z_{0.975} \ se\left(\hat{t}_{0.1}\right)$$

```
op$p[2]+c(-1,1)*qnorm(0.975)*se[2]
```

```
## [1]  160.4161 7641.4650
```

(d)

Profile likelihood ratio

$$R(t_{0.1}) \ = \ \max_{\mu} \left[\frac{L(t_{0.1},\mu)}{L(\hat{t}_{0.1},\hat{\mu})}\right]$$

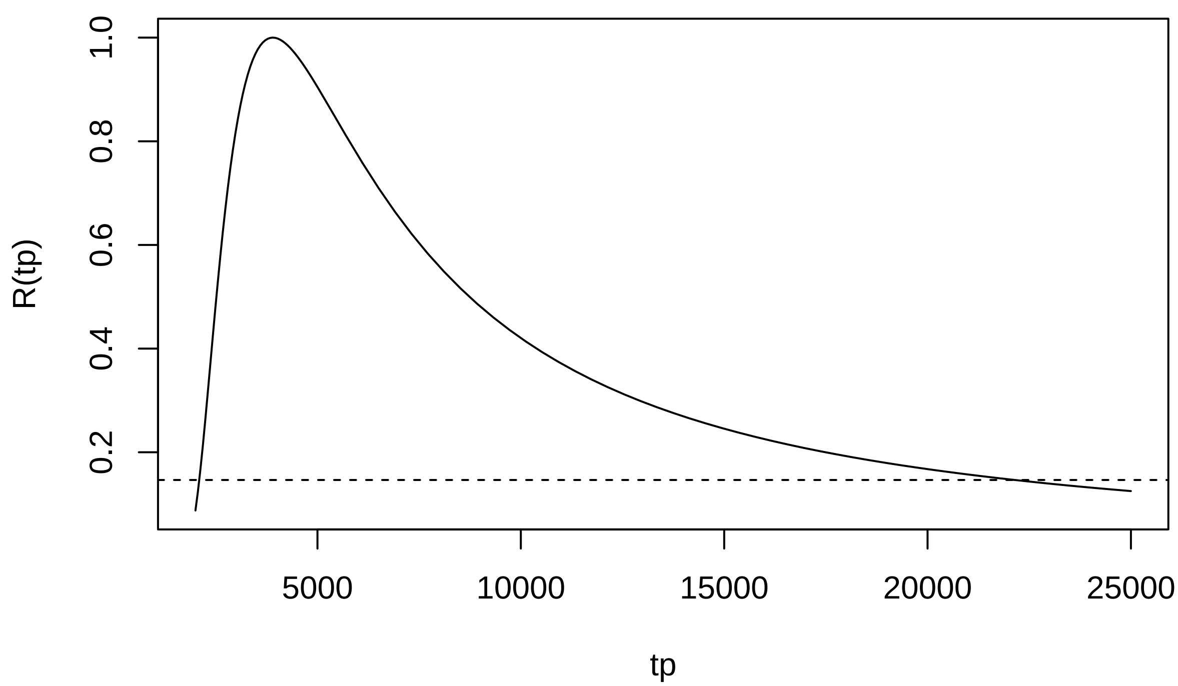
```
logL1=function(mu,tp,ti,di,wi){
  logL(c(mu,tp),ti,di,wi)
}
x=seq(7.5,15,len=100)
```

```r
y=c(seq(2000,5000,len=50),seq(5300,25000,len=50))
Rx=Ry=c()
for(i in 1:100){
  Rx[i]=exp(-optim(500,mu=x[i],ti=ti,di=di,wi=wi,logL1)$value+op$value)
  Ry[i]=exp(-optim(10,tp=y[i],ti=ti,di=di,wi=wi,logL1)$value+op$value)
}
```

The LR-based 95% confidence interval for $t_{0.1}$

$$\left\{t_{0.1} \ : \ R(t_{0.1}) \ > \ \exp\left(-\frac{1}{2}\chi^2_{0.95,1}\right)\right\}$$

```r
plot(y,Ry,type="l", xlab = "tp", ylab = "R(tp)")
abline(h=exp(-qchisq(0.95,df=1)/2),lty=2)
```




```r
y[c(3,93)]
```

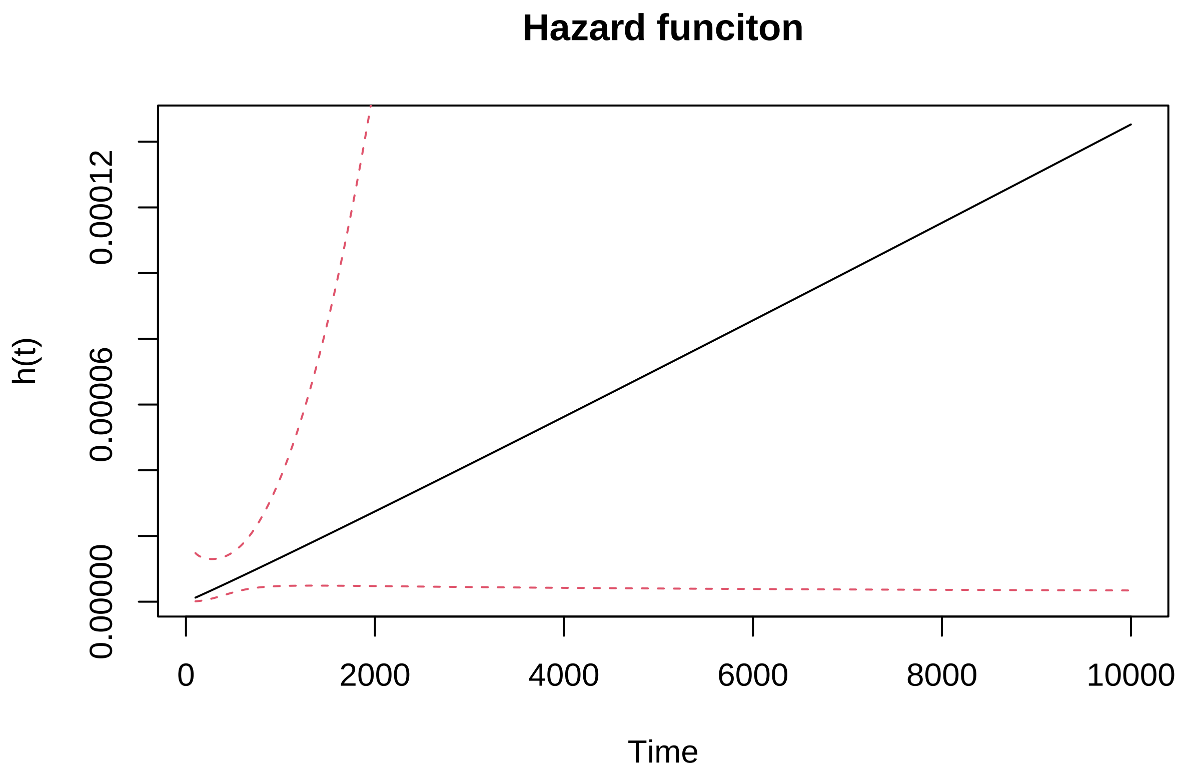
```
## [1]  2122.449 22185.714
```

(e)

LR-based confidence interval. However, the Wald confidence interval procedures are quick, useful, and adequate for exploratory work. When more accurate confidence interval approximations are required, one should use likelihood procedures.

(f)

```
hazard = function(t,mu=mu.hat,sig=sig.hat) {
    z = (log(t)-mu)/sig
    return(exp(z)/(sig*t))
}
xl = seq(100,10000,len=1000)
plot(xl,hazard(xl), type = "l", xlab = "Time", ylab="h(t)", main="Hazard funciton")
CI = function(t, mu=mu.hat,sig=sig.hat) {
    h.hat = hazard(t)
    z = (log(t)-mu)/sig
    B = c(-exp(z)/(sig^2*t) , -exp(z)/(sig^2*t)*(1+z))
    h.se = sqrt(B%*%solve(op.bear$hessian)%*%B)[1,1]
    w = exp(qnorm(0.975)*h.se/h.hat)
    return(c(h.hat/w , h.hat*w))
}
h.low = c()
h.up = c()
for (i in 1:1000) {
    h.low[i] = CI(xl[i])[1]
    h.up[i] = CI(xl[i])[2]
}
points(xl,h.low, col = 2, lty = 2, type = "l", ylim = c(0,0.001))
points(xl,h.up, col = 2, lty = 2, type = "l")
```



(g)

Likelihood ratio

$$R(\mu\ ,\ t_{0.1})\ =\ \frac{L(\mu\ ,\ t_{0.1})}{L(\hat{\mu}\ ,\ \hat{t}_{0.1})}$$

The 95% joint confidence region of $(\mu\ ,\ t_{0.1})$

$$\left\{(\mu\ ,\ t_{0.1})\ :\ R(\mu\ ,\ t_{0.1})\ >\ \exp\left(-\frac{1}{2}\chi^2_{0.95,2}\right)\right\}$$

```
x=seq(7.7,15,len=100)
y=seq(1000,54000,len=100)
y=c(seq(1000,5000,len=100),seq(5100,54000,len=100))
z=matrix(0,100,200)
for(i in 1:100){
  for(j in 1:200){
    exp(-logL(c(x[i],y[j]),ti,di,wi)+op$value)
    z[i,j]=exp(-logL(c(x[i],y[j]),ti,di,wi)+op$value)
  }
```

```r
}
par(mfrow = c(1,2))
contour(x,y,z,levels=c(0.9,0.4,0.1,0.05,0.01,0.001),main="contour",
        xlab=TeX("$\\mu$"),ylab=TeX("$t_{0.1}$"))
contour(x,y,z,levels=c(0.05),main="contour",xlab=TeX("$\\mu$"),ylab=TeX("$t_{0.1}$"))
```