Reliability-HW4

ID : 111024517   Name：鄭家豪

due on 05/17

Problem 1

(b)

Fit Expoential distribution with θ to this data,the likelihood:

$$L(\theta) = \prod_{i=1}^{n}(f(t_i;\theta))^{d_i}(1 - F(t_i;\theta))^{r_i}, \text{failure time } t_i > 0,$$

$$\text{where } f(t_i;\theta) = \frac{1}{\theta}\exp(-\frac{t_i}{\theta}), F(t_i;\theta) = 1 - \exp(-\frac{t_i}{\theta})$$

$$d_i : \text{the failed count at time } t_i,$$
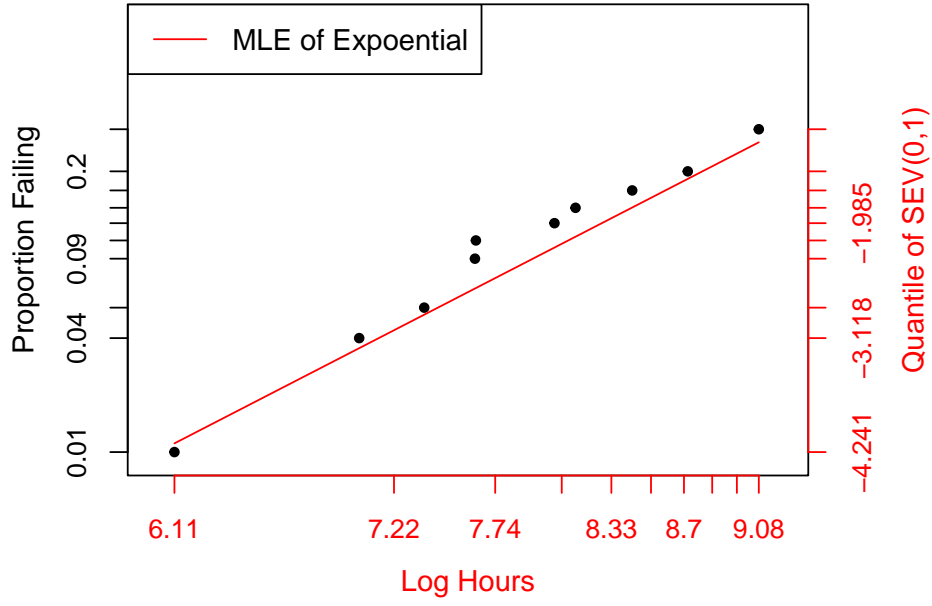
$$r_i : \text{the censored count at time } t_i.$$

Solve MLE:

$$\frac{\partial \log L}{\partial \theta} = \sum_{i=1}^{n}(d_i(\frac{-1}{\theta} + \frac{t_i}{\theta^2}) + r_i\frac{t_i}{\theta^2}) = \frac{1}{\theta^2}\sum_{i=1}^{n}(d_i(-\theta + t_i) + r_it_i) = 0$$

$$\Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n}t_i(d_i + r_i)}{\sum_{i=1}^{n}d_i}.\text{By MLE invariant,the MLE of } F(t_i;\theta) = F(t_i;\hat{\theta}).$$

Then ,draw Weibull plot(the scatter points are estimated by KM):

$$\log(-\log(1 - \tilde{F}(t_i))) = a + b \times \log t_i, \text{where } a : \text{intecept}, b : \text{slope},$$

$$\tilde{F}(t_i) = 1 - \prod_{j=1}^{i}(d_i/n_i)$$

(c)

By delta method and Wald confidence,

$$\text{The 95\% CI for } \theta : \hat{\theta} \pm 1.96 \times \text{s.e.}(\hat{\theta}) = (L_{\hat{\theta}}, U_{\hat{\theta}}),$$

$$\text{where s.e.}(\hat{\theta}) = \left(\frac{\mathrm{d}\log L}{\mathrm{d}\theta}\Big|_{\theta=\hat{\theta}}\right)^{-1/2} = \left(\frac{\hat{\theta}^2}{\sum_{i=1}^{n} d_i}\right)^{1/2}.$$

$$\Rightarrow \text{The 95\% CI for F}(1250) : (F(1250; U_{\hat{\theta}}), F(1250; L_{\hat{\theta}})).$$

$$(\because F(t;\theta) : \text{nonincreasing in } \theta)$$

```
The MLE of F(1250): 0.04261431
```

```
The 95% CI of F(1250): ( 0.0274295 , 0.09542977 )
```

(d)

By delta method and Wald confidence,

$$\text{The 95\% CI for } t_{0.1} : (-\log(1-0.1))\hat{\theta} \pm 1.96 \times \text{s.e.}(\hat{\theta}) \times \log(1-0.1).$$

```
The MLE of 0.1 quantile: 3024.198
```

```
The 95% CI of 0.1 quantile: ( 1313.128 , 4735.268 )
```

2

(e)

Based on Weibull distribution is log-location family,

By delta method(two parameter),

$$F(T \leq t) = \Phi_{\text{sev}}(\frac{\log(t) - \mu}{\sigma}), \text{where } \theta = (\mu, \sigma).$$

$$B = \frac{\partial F}{\partial \theta} = [\frac{\partial F}{\partial \mu}, \frac{\partial F}{\partial \sigma}] = \phi_{\text{sev}}(z)[-1/\sigma, z/\sigma], \text{where } \frac{\log(t) - \mu}{\sigma} = z.$$
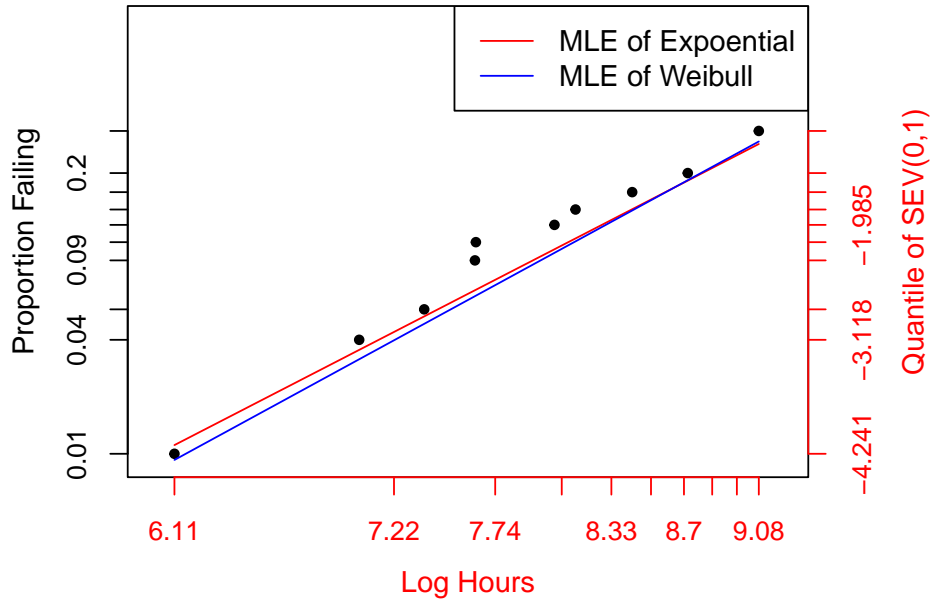
$$\text{The 95\% CI for F}(1250): F(1250; \hat{\mu}, \hat{\sigma}) \pm 1.96 \times (B[\text{COV}(\hat{\theta})]B^T)$$

Similar to part (b),draw the Weibull plot:

```
The MLE of mu and sigma on Weibull: 10.17678 0.9455166
```

```
The MLE of F(1250) on Weibull: 0.03911405
```

```
The 95% CI of F(1250) on Weibull: ( 0.0009697705 , 0.07725832 )
```


Based on the probability plot shown above, it can be observed that the Weibull and expoential distribution fit similarly.

Moreover ,from the both confidence interval,the inference of Weibull distribution is more conservative than expoential distribution.

(f)

By variable transformation,

$$T \sim \text{Wei}(\theta, k : \text{shape}) \Rightarrow \log(T) \sim \text{SEV}(\mu, \sigma),$$

$$\text{with } \sigma = 1/k, \mu = \log(\theta).$$

From part (e),we know the MLE of σ.By MLE invariant,$\hat{k} = 1/\hat{\sigma}$,and its CI can be obtained by delta method.

```
The MLE of Weibull shape parameter: 1.057623
```

```
The 95% CI of Weibull shape parameter: ( 0.7893928 , 1.325853 )
```

Because the 95% confidence interval of Weibull shape parameter includes 1, it means that the exponential distribution can also be used to fit. In other words, the Hazard function of the diesel generator fan is a constant and does not depend on time.

Problem 2

(b)

The relative likelihood of $\theta = (\mu, \sigma)$:$R(\theta) = \dfrac{L(\theta)}{L(\hat{\theta}_{\text{MLE}})}$,where $\hat{\theta}_{\text{MLE}}$ is obtained by optim().

```
The MLE of mu and sigma: ( 9.374678 , 0.4912114 )
```


Contour Plot of R(μ, σ)
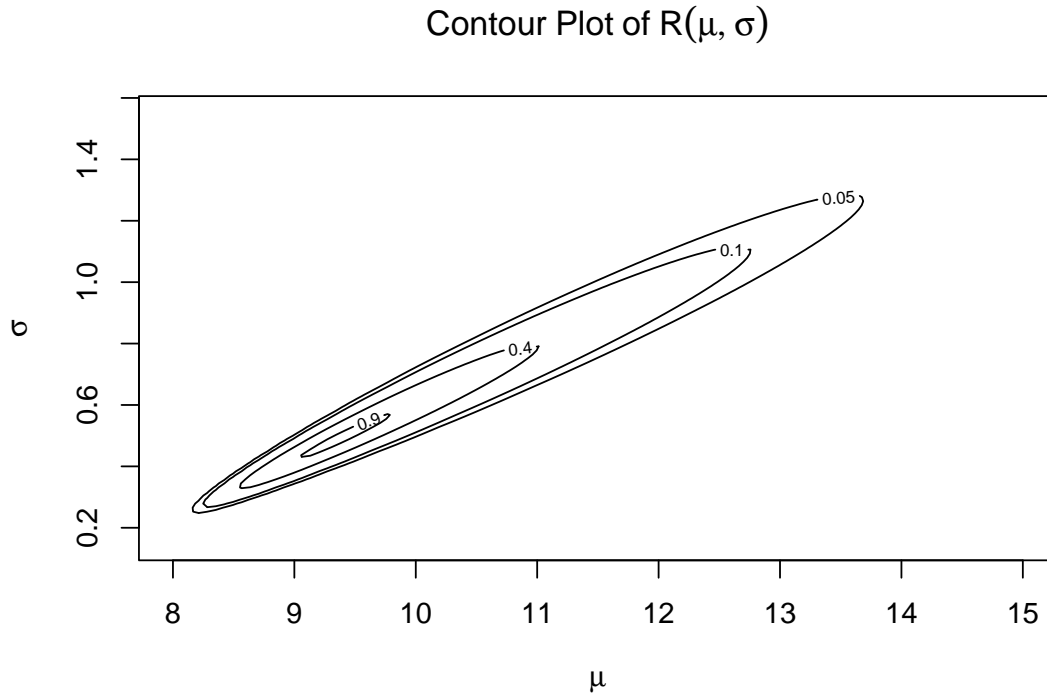
From the above,μ and σ are positive correlation.

(c)

By delta method and Wald CI,

$$\because \Phi_{\text{sev}}(\frac{\log(t_p) - \mu}{\sigma}) = p, \therefore \log(t_p) = \mu + \sigma \times \Phi_{\text{sev}}^{-1}(p)$$

The standard error of $\log(t_{0.1})$: $s_t = ([1, \Phi_{\text{sev}}^{-1}(0.1)][\text{COV}(\hat{\theta})][1, \Phi_{\text{sev}}^{-1}(0.1)]^T)^{1/2}$
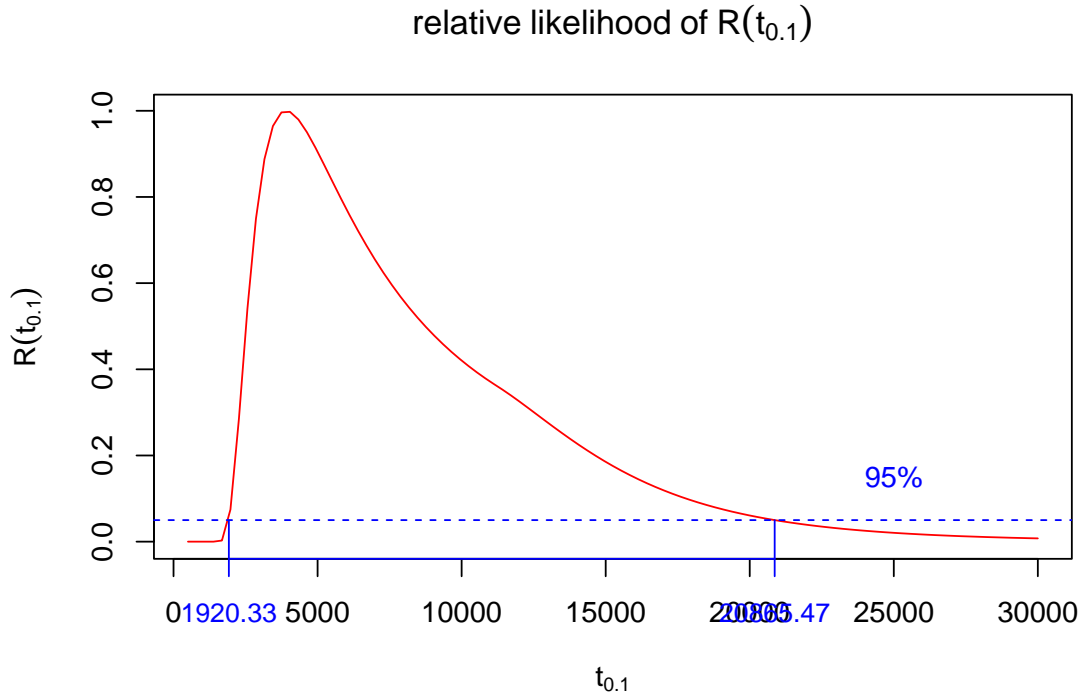
The 95% CI of $\log(t_{0.1})$: $\log(t_p)|_{\theta=\hat{\theta}} \pm 1.96 \times s_t = (L_{0.1}, U_{0.1})$

The 95% CI of $t_{0.1}$: $(\exp(L_{0.1}), \exp(U_{0.1}))$.

```
The Wald 95% CI for 0.1 quantile: ( 1488.772 , 10227.52 )
```

(d)

Draw the relative likelihood w.r.t $t_{0.1}$ plot(from part (b)):

relative likelihood of R(t_{0.1})

```
The LR 95% CI : ( 1920.335 , 20865.47 )
```

(e)

I think Wald CI is trustworthy based on the informative range and large sample size because its width of 95% CI for $t_{0.1}$ is less than LR 95% CI for $t_{0.1}$.

In this case,LR CI reflects the uncertainty in the estimation of true parameter,and indicates that the LR method is more conservative and has a lower level of precision.

(f)

For $\log(T) \sim \mathrm{SEV}(\mu, \sigma)$,by MLE from part(b),the estimated hazard rate of T:

$$h(t; \hat{\mu}, \hat{\sigma}) = \frac{f(t; \hat{\mu}, \hat{\sigma})}{1 - F(t; \hat{\mu}, \hat{\sigma})} = \frac{1}{e^{\hat{\mu}}\hat{\sigma}}(\frac{t}{e^{\hat{\mu}}})^{1/\hat{\sigma}-1}$$
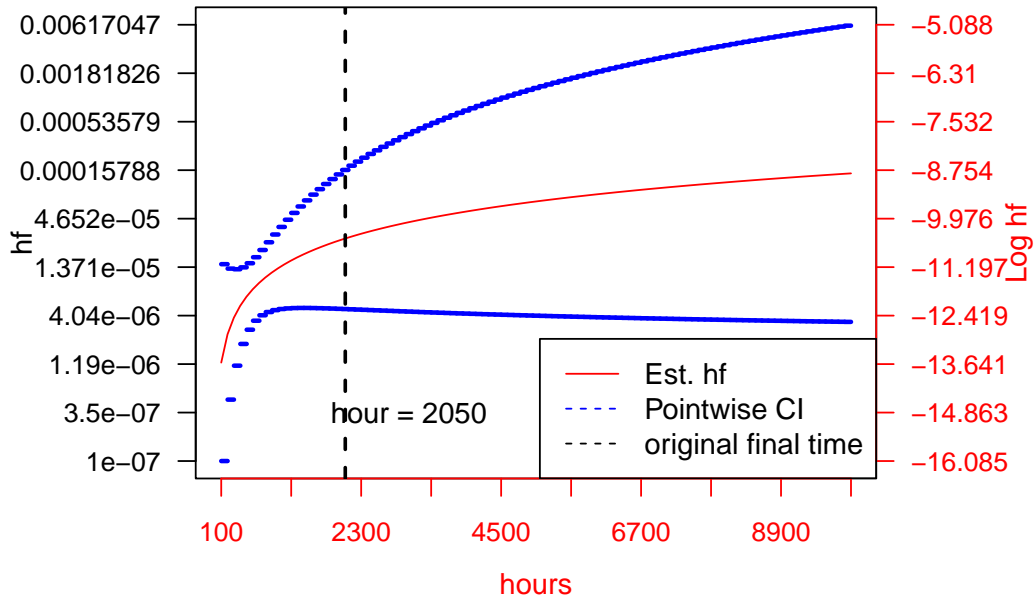
By delta method and Wald CI,

$$B = [\frac{\partial \log(h(t; \mu, \sigma))}{\partial \mu}, \frac{\partial \log(h(t; \mu, \sigma))}{\partial \sigma}]|_{\mu=\hat{\mu}, \sigma=\hat{\sigma}}, \text{where } \log(h(t; \mu, \sigma)) = -\mu - \log \sigma + (1/\sigma - 1)(\log(t) - \mu)$$

$$= [-1/\hat{\sigma}, -1/\hat{\sigma} - 1/\hat{\sigma}^2(\log(t) - \hat{\mu})] = 1/\hat{\sigma}^2[-\hat{\sigma}, -\hat{\sigma} - \log(t) + \hat{\mu}]$$

$$\mathrm{Var}(\log(h(t))) \approx B[\mathrm{COV}(\mu, \sigma)]B^T$$

The 95% CI of $\log(h(t))$: $\log(h(t; \hat{\mu}, \hat{\sigma})) \pm 1.96 \times \sqrt{\mathrm{Var}(\log(h(t)))}$

(Since the value of hf is too small, it is presented in the follolwing as an estimate of log(hf) and its pointwise CIs)

Appendix(code)

```
# 1
## pre-process
Fan <- read.csv("Fan.csv",header = T)
t.index <- unique(Fan$Hours)
di <- rep(0,length(t.index))
```

6

```r
label=0
for (i in 1:length(t.index)){
  label <- which(Fan$Hours == t.index[i] & Fan$Censoring.Indicator=="Fail")
  di[i] <- ifelse(length(label)==0,0,Fan$Count[label])
}
ri <- rep(0,length(t.index))
for (i in 1:length(t.index)){
  label <- which(Fan$Hours == t.index[i] & Fan$Censoring.Indicator=="Censored")
  ri[i] <- ifelse(length(label)==0,0,Fan$Count[label])
}
ni <- 0
ni[1] <- sum(Fan$Count)
for(i in 2:length(di)){
  ni[i] <- ni[1] - sum(ri[1:i-1])-sum(di[1:i-1])
}
org.Fan <- data.frame("ti"=t.index,"di"=di,
                      "ri" = ri,
                      "ni"=ni)


##(c)
ti <- org.Fan$ti
theta.hat <- sum(ti*(di+ri))/sum(di)
F.1250.hat <- pexp(1250,rate=1/theta.hat) # F(T<1250) MLE
se.hat.theta.hat <- theta.hat/sqrt(sum(di))
theta.CI <- theta.hat+c(-1,1)*qnorm(0.975)*se.hat.theta.hat
F.1250.CI <- pexp(1250,rate=1/c(theta.CI[2],theta.CI[1])) #F(T<1250) CI


## (b)
t = ti[di>0] #fail time
S.hat <- cumprod(1-org.Fan$di/org.Fan$ni)
prob <- (1-S.hat)[di>0] #KM cdf
par(plt=c(0.2,0.8,0.2,0.8))
plot(log(t), log(-log(1-prob)), type="p",
     col="black",yaxt="n",ylab="",xaxt="n",
     xlab="",ylim=c(-4.3,0),xlim=c(6,9.2),
     pch=20)
xl=log(seq(min(t),max(t),length=length(prob)))
yl=log(-log(1-prob))
axis(1,at=xl,lab=round(xl,2),cex.axis=0.9,col = "red",col.axis="red")
axis(2,at=yl,lab=round(prob,2),cex.axis=0.9)
```

```r
axis(4,at=yl,round(yl,3),cex.axis=0.9,
     col = "red",col.axis="red")
mtext("Log Hours", side=1, line=2.5,col="red")
mtext("Proportion Failing", side=2, line=2.5)
mtext("Quantile of SEV(0,1)", side=4, line=2.5,col = "red")
curve(log(-log(1-pexp(exp(x),rate = 1/theta.hat))), from = log(min(t)), to = log(max(t)),
      add = T, col='red')
legend('topleft', legend ='MLE of Expoential', col = 'red', lty=1)
##(d)
t_0.1.hat=-theta.hat*log(1-0.1)
t_0.1.CI=-theta.CI*log(1-0.1)


##(e)
logL.Weibull=function(theta,ti,di,ri){
  mu=theta[1];sig=theta[2]
  if(sig<0){sig=0.001}
  l=0
  z=(log(ti)-mu)/sig
  for(i in 1:length(ti)){
    l=l+di[i]*(-log(sig)-log(ti[i])+z[i]-exp(z[i]))+
      ri[i]*(-exp(z[i]))
  }
  return(-l)
} # Weibull -log L
op.Weibull=optim(c(1,2),logL.Weibull,ti=ti,di=di,ri=ri,hessian=T)
#P(W <1250)
mle.Weibull=op.Weibull$par
COV.Weibull=solve(op.Weibull$hessian)
psev=function(x){
  1-exp(-exp(x))
}
dsev=function(x){
  exp(x-exp(x))
}
z = (log(1250)-mle.Weibull[1])/mle.Weibull[2]
F.1250.hat.Wei <- psev(z)
phi=dsev(z)
B=phi*c(-1,-z)/mle.Weibull[2]
se_F.1250.hat.Wei=sqrt(B%*%COV.Weibull%*%B)
F.1250.CI.Wei = F.1250.hat.Wei+c(-1,1)*qnorm(0.975)*as.numeric(se_F.1250.hat.Wei)
```

```r
par(plt=c(0.2,0.8,0.2,0.8))
plot(log(t), log(-log(1-prob)), type="p",
     col="black",yaxt="n",ylab="",xaxt="n",
     xlab="",ylim=c(-4.3,0),xlim=c(6,9.2),
     pch=20)
xl=log(seq(min(t),max(t),length=length(prob)))
yl=log(-log(1-prob))
axis(1,at=xl,lab=round(xl,2),cex.axis=0.9,col = "red",col.axis="red")
axis(2,at=yl,lab=round(prob,2),cex.axis=0.9)
axis(4,at=yl,round(yl,3),cex.axis=0.9,
     col = "red",col.axis="red")
mtext("Log Hours", side=1, line=2.5,col="red")
mtext("Proportion Failing", side=2, line=2.5)
mtext("Quantile of SEV(0,1)", side=4, line=2.5,col = "red")
curve(log(-log(1-pexp(exp(x),rate = 1/theta.hat))), from = log(min(t)), to = log(max(t)),
      add = T, col='red')
curve(log(-log(1-psev((x-mle.Weibull[1])/mle.Weibull[2]))), from = log(min(t)), to = log(max(t)),
      add = T, col='blue')
legend('topright', legend =c('MLE of Expoential','MLE of Weibull'), col = c('red',"blue"), lty=1)


## (f)
beta.CI = 1/mle.Weibull[2] + c(-1,1)*sqrt(COV.Weibull[2,2])/mle.Weibull[2]^2



# 2
## Pre-process
BearingCage <- read.csv("BearingCage.csv",header = T)
t.index <- unique(BearingCage$Hours)
di <- rep(0,length(t.index))
label=0
for (i in 1:length(t.index)){
  label <- which(BearingCage$Hours == t.index[i] & BearingCage$Censoring.Indicator=="Failed")
  di[i] <- ifelse(length(label)==0,0,BearingCage$Count[label])
}
ri <- rep(0,length(t.index))
for (i in 1:length(t.index)){
  label <- which(BearingCage$Hours == t.index[i] & BearingCage$Censoring.Indicator=="Censored")
  ri[i] <- ifelse(length(label)==0,0,BearingCage$Count[label])
}
ni <- 0
```

```r
ni[1] <- sum(BearingCage$Count)
for(i in 2:length(di)){
  ni[i] <- ni[1] - sum(ri[1:i-1])-sum(di[1:i-1])
}
org.BearingCage <- data.frame("ti"=t.index,"di"=di,
                              "ri" = ri,
                              "ni"=ni)
## (b)
ti <- t.index
mle <- optim(par = c(10,0.5),logL.Weibull,di=di,ri=ri,ti=ti,hessian = T)


x <- seq(8,15,length.out=100)
y <- seq(0.2,1.5,length.out=100)
z=matrix(0,100,100)
for(i in 1:100){
  for(j in 1:100){
    z[i,j]=exp(-logL.Weibull(c(x[i],y[j]),ti,di,ri)+mle$value)
  }
}
contour(x,y,z,xlab=expression(mu),ylab=expression(sigma),
        levels=c(0.9,0.4,0.1,0.05),
        main=expression(Contour~Plot~of~R(mu,sigma)),
        xlim = c(min(x),max(x)),ylim = c(min(y)-0.05,max(y)+0.05))


## (c)
qsev=function(p){
  log(-log(1-p))
}
tp <- exp(mle$par[1]+mle$par[2]*qsev(0.1))
B <- tp*c(1,qsev(0.1))
sd.tp <- (B%*%solve(mle$hessian)%*%B)^(1/2)
w <- exp(qnorm(0.975)*sd.tp/tp)
CI.Wald <- c(tp/w,tp*w)


## (d)
ng.relative.L.tp <- function(tp,sigma){
  fn = exp(-logL.Weibull(c(log(tp)-qsev(0.1)*sigma,sigma),ti,di,ri)+mle$value)
  -fn
}
LR <- function(time){
```

```r
    rt <- c()
    for (t in 1:length(time)){
      max <- optim(mle$par[2],ng.relative.L.tp,lower=0.1,
                   upper=0.8,method="Brent",tp=time[t])
      value <- ng.relative.L.tp(time[t],max$par)
      rt <- c(rt,-value)
    }
    rt
}


curve(LR(x),500,30000,xlab = expression(t[0.1]),ylab=expression(R(t[0.1])),
      main=expression(relative~likelihood~of~R(t[0.1])),col="red")
abline(h=exp(-qchisq(0.95,2)/2),lty=2,col="blue")
text(25000,exp(-qchisq(0.95,2)/2)+0.1,"95%",col="blue")
LR.CI <- function(t){(LR(t)-exp(-qchisq(0.95,2)/2))^2}
L.95 <- optim(tp,LR.CI,lower=1500,upper=tp,method="Brent")$par
U.95 <- optim(tp,LR.CI,lower=tp,upper=25000,method="Brent")$par
lines(x=c(L.95,L.95),y=c(-1,0.05),lty=1,col="blue")
lines(x=c(U.95,U.95),y=c(-1,0.05),lty=1,col="blue")
axis(1,at=c(L.95,U.95),lab=round(c(L.95,U.95),2),cex.axis=0.9,col = "blue",col.axis="blue")
cat("The LR 95% CI :","(",L.95,",",U.95,")","\n")


## (f)
d.ti <- seq(100,10000,length=100) #used in plot
hf.hat.fun <- function(t){
  t^(1/mle$par[2]-1)*exp(-mle$par[1]/mle$par[2])*1/mle$par[2]
}
hf.hat <- hf.hat.fun(d.ti)
var.loghf <- function(t){
  mu = mle$par[1];sigma = mle$par[2]
  B1 = sigma
  B2 = -sigma+mu-log(t)
  B = c(B1,B2)/sigma^2
  B %*% solve(mle$hessian) %*% B
} #delta method for calculate var.log(hf)
l.log.hf <- c()
u.log.hf <- c()
for ( i in 1:length(d.ti)){
  l.log.hf[i] <- log(hf.hat[i])-qnorm(0.975)*sqrt(var.loghf(d.ti[i]))
  u.log.hf[i] <- log(hf.hat[i])+qnorm(0.975)*sqrt(var.loghf(d.ti[i]))
```

```r
} #build pointwise CI
par(plt=c(0.2,0.8,0.2,0.8))
curve(log(hf.hat.fun(x)),100,10000,
      yaxt="n",ylab="",xaxt="n",xlab="",col="red",
      ylim=c(min(l.log.hf),max(u.log.hf)),xlim=c(95,10001))
xaxis <- seq(100,10000,length=10)
yaxis <- seq(min(l.log.hf),max(u.log.hf),length=10)
axis(1,at=xaxis,lab=round(xaxis,0),cex.axis=0.9,
     col = "red",col.axis="red")
axis(2,at=yaxis,lab=round(exp(yaxis),8),cex.axis=0.9,las=1)
#axis(3,at=xaxis,round(xaxis,3),las=1,cex.axis=0.9,
#     col = "red",col.axis="red")
axis(4,at=yaxis,round(yaxis,3),cex.axis=0.9,las=1,
     col = "red",col.axis="red")
mtext("hours", side=1, line=2.5,col = "red")
mtext("hf", side=2, line=3.5)
#mtext("Log hours", side=3, line=2.5,col = "red")
mtext("Log hf", side=4, line=3.5,col = "red")
for(i in 1:99){
  lines(d.ti[i:(i+1)],rep(l.log.hf[i],2),lty=2,col="blue",lwd=2.5)
  lines(d.ti[i:(i+1)],rep(u.log.hf[i],2),lty=2,col="blue",lwd=2.5)
}
abline(v = max(ti),col="black",lty=2,lwd=2)
text(max(ti)+1000,yaxis[2],"hour = 2050",col="black")
legend('bottomright', legend =c('Est. hf','Pointwise CI','original final time'),
       col = c('red','blue','black'), lty=c(1,2,2))
```