

BIOS 622 Homework 4

Instructor: Sijin Wen

February 20, 2025

Please answer each question precisely and completely. It is permissible to discuss each other, but all of your work must be your own.

1. When we use Wald statistic to find estimation and its 95%CI, we assume that sample size is large, thus $\hat{\lambda}$ will approximately have Chi-square or normal distribution (after taking square of root). If the sample size is small, the normal distribution may not be a good approximation. Two alternative approaches are (1) bootstrapping and (2) Bayesian method using MCMC algorithm to get a lot of samples from the true distribution of λ , then we can find $\hat{\lambda}$ using sample-mean, sample-standard-deviation, sample-quantiles for 95%CI. Also you can visualize the true distribution of $\hat{\lambda}$ using histogram on those MCMC samples. A modification of example from the course notes (Page 60). T_1, \dots, T_n iid exponential(λ). The outcome is time to death from severe viral disease. The modified data are as follows (the same as HW3 Q2):
 - Steroid: 1(2), 1+(2), 4+, 5, 7, 8, 10, 10+, 12+, 16+(3)
 - Control: 1, 2, 3(2), 3+, 5+(2), 16+(8)
 - (a) Find 95% CI of $\hat{\lambda}_i$ ($i = 1, 2$) for each treatment group, using bootstrapping approach
 - (b) Find $\hat{\lambda}_i$ ($i = 1, 2$) and its 95% CI for each treatment group, using Bayesian approach (you may use the R code below)
 - (c) Based on (a) and (b), estimate hazard ratio between two groups ($HR = \lambda_1/\lambda_2$)
 - (d) Estimate a 95% CI of hazard ratio using either bootstrapping or Bayesian approach.

```
##### steroid group/data:
t1 <- c(1,1,1,1,4,5,7,8,10,10,12,16,16,16)
cen1 <- c(1,1,0,0,0,1,1,1,1,0,0,0,0,0)
sum(cen1)/sum(t1)  ##MLE of lambda
```

```

sum(t1)/sum(cen1)  ##MLE of mu

##### control group:
t2 <- c(1,2,3,3,3,5,5,rep(16,8))
cen2 <- c(1,1,1,1,0,0,0,rep(0,8))
sum(cen2)/sum(t2)  ## MLE of lambda
sum(t2)/sum(cen2)  ## MLe of mu

#####log-likelihood

loglik <- function(lamda,y,ind){
res.u <- sum(log(lamda*exp(-lamda*y[ind==1])))
res.cen <- sum(log(exp(-lamda*y[ind==0])))
return((res.u+res.cen))
}

#####MCMC

mcmc.s <- function(y,ind,lam, e0, N)
{
set.seed(99)
acc <- 0
opt <- rep(NA, N)
for(i in 1:N) {
print(i)
old <- loglik(lam,y,ind)
cand <- lam + e0 * rnorm(1)
while(cand <0){cand <- lam + e0 * rnorm(1)}
new <- loglik(cand,y,ind)
if(log(runif(1)) < new - old) {
lam <- cand
acc <- acc + 1
}
opt[i] <- lam
}

```

```

}
par(mfrow = c(2, 1))
plot(1:N, opt, type = "l", xlab = "", ylab = "lambda")
plot(1:N, 1/opt, type = "l", xlab = "", ylab = "mu")
print(acc/N)
return(opt)  ##lambda
}

junk<-mcmc.s(y=t1,ind=cen1,lam=sum(cen1)/sum(t1) ,e0=0.09,10000)
mean(junk[-(1:5000)])  ##lambda
sd(junk[-(1:5000)])  ## sd of lambda
sum(cen1)/sum(t1)  ## comparing MLE

junk2<-mcmc.s(y=t2,ind=cen2,lam=sum(cen2)/sum(t2),e0=0.03,10000)
mean(junk2[-(1:5000)])  ## lambda2
sd(junk2[-(1:5000)])  ## sd
sum(cen2)/sum(t2)  ## comparing MLE

hist(junk[-(1:5000)]/junk2[-(1:5000)])  ## histogram of hazrd ratio?

```