likelihood

陳寧寬

2019-03-06

wooldridge\

標準的 right censored 和 failure time 問題

$$t_i = min(t_i^*, c_i)$$

t_i 代表觀察到的期間\ t_i^* 代表真實存活時間\ c_i 代表 censored time\

我們去建構 likilihood 時應該使用 t_i, 因為這才是觀察的到的

在此處 di=1 代表 right censored di=0 代表 interval censored\

站在 102 年的角度下

對 0-5 歲的樣本

$$\frac{[F(10) - F(0)]^{d_i}[1 - F(5)]^{1-d_i}}{1 - F(0)}$$

對 5-10 歲的樣本

$$\frac{[F(15) - F(5)]^{d_i}[1 - F(10)]^{1-d_i}}{1 - F(5)}$$

對 10-15 歲的樣本

$$\frac{[F(20) - F(10)]^{d_i}[1 - F(15)]^{1-d_i}}{1 - F(10)}$$

對 15-20 歲的樣本

$$\frac{[F(25) - F(15)]^{d_i}[1 - F(20)]^{1-d_i}}{1 - F(15)}$$

對 20-25 歲的樣本

$$\frac{[F(30) - F(20)]^{d_i}[1 - F(25)]^{1-d_i}}{1 - F(20)}$$

對 25 以上歲的樣本

$$\frac{[F(100) - F(25)]^{d_i}[1 - F(30)]^{1-d_i}}{1 - F(25)}$$

simulation 的設定

從民國 52 年開始生成資料，每年都蓋 100 戶房子生到民國 102 年

觀察那些民國 102 年還健在的房子，然後去紀錄他們民國 107 年的情形

民國 107 年還健在的就是 right cenosred 不在的就是 interval censored

step1 先判斷房子在 102 年還有沒有活著

民國 52 年蓋的房子去扣掉 50 民國 53 年蓋的房子去扣掉 49

```
lamda=0.05
n=100
y=50
A <- rexp(n*y,rate = lamda)
#hist(A,probability = T,breaks =15 )
data <- matrix(A,nrow=y,ncol=n)
data102 <- matrix(0,nrow=y,ncol=n  )
for(i in 1:y){
  c=51-i
 data102[i,] <- data[i,]-c

}
ind<- which(data102>=0)
datanew <- matrix(0,nrow=y,ncol=n  )
data102alivelife=matrix
datanew[ind] <- data[ind]
data102alivelife <- datanew # 在 102 年活著的樣本的壽命
#data102aliveyear <-          # 在 102 年活著的樣本的年紀
```

```
data107 <- matrix(0,nrow=y,ncol=n  )

for(i in 1:y ){
  c=56-i
 data107[i,] <- data[i,]-c

}

ind2<- which(data107>=0)
datanew2 <- matrix(0,nrow=y,ncol=n  )
data107alivelife=matrix
datanew2[ind2] <- data[ind2]
data107alivelife <- datanew2
```

```
ind102<- ind
ind107<- ind2
ind_dead <- setdiff(ind,ind2)
```

```
data102year=matrix(0,nrow=y,ncol=n  )
for(i in 1:y){
c=51-i
for(j in 1:n)
{
if(data102alivelife[i,j]>0){
data102year[i,j]=c
}
}
}
```

```
data107year=matrix(0,nrow=y,ncol=n  )
for(i in 1:y){
c=56-i
for(j in 1:n)
{
if(data107alivelife[i,j]>0){
data107year[i,j]=c
}
}
}
#data107year 為 107 年尚健在的建物的年齡
#data102year 為 102 年尚健在的建物的年齡
```

```
data107v <- data107year[which(data107year>0)]

data102v <- data102year[which(data102year>0)]
```

```
require(dplyr)
```

```
## Loading required package: dplyr
```

```
## Warning: package 'dplyr' was built under R version 3.4.4
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
require(reshape2)
```

```
## Loading required package: reshape2
```

```
## Warning: package 'reshape2' was built under R version 3.4.4
```

```
gentable=function(a){
a1<-melt(table(cut(a,breaks=seq(0,100,5) )))
a2<-data.frame(sapply(a1,function(x) gsub("\\(|\\]","",gsub("\\,","-",x))))
colnames(a2)<-c("numbers","Freq")
return(a2)
}
data102table <- gentable(data102v)
data107talbe <- gentable(data107v)
A <- rbind(  t(data102table),t(data107talbe))
B <- A[-3,]
rownames(B)<-c("year","102","107  ")
```

```
require(data.table)
```

```
## Loading required package: data.table
```

```
## Warning: package 'data.table' was built under R version 3.4.4
```

```
##
```

```
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:reshape2':
##
##     dcast, melt
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

```
a <- as.numeric(B[2,])
b <-as.numeric(B[3,])
c<- B[1,]
class(a)
```

```
## [1] "numeric"
```

```
d <- data.table(cbind(  B[1,],a,b   ))
```