

Computer Lab 8: Interaction and Survival Data

1. Interaction

Example, log-linear models (interaction):

The data given in Splus build in data frame *Insurance* (in the library MASS) consist of the numbers of policy-holders, *Holders*, the numbers of car insurance claims made by those policyholders, *Claims*. There are three explanatory variables, *District* (four levels), *Group* (of car, four levels), and *Age* (four ordered levels). Please analyze the data up to the two way interaction with offset *log(Holders)*.

```
[Splus: interaction]
library(MASS)
data=Insurance
age=factor(data$Age)
group=factor(data$Group)
dis=factor(data$District)
hol=log(data$Holders)
cla=data$Claims
data3=data.frame(hol,dis,group,age,cla)
glm1=glm(cla~offset(hol)+dis+group+age,family=poisson,data=data3)
anova(glm1,test="Chisq")
summary(glm1,cor=F)
glm2=glm(cla~offset(hol)+(dis+group+age)^2,family=poisson,data=data3)
anova(glm2,test="Chisq")
summary(glm2,cor=F)
```

2. Estimation of survival function

Example:

For the data set “leukemia” in Splus,

- Assume the survival time is exponentially distributed, please find the estimated survival function and plot the survival function. Please compare the estimated survival function with the Kaplan-Meier estimates and by plot the two survival functions in the same figure.
- Using the Kaplan-Meier estimates for two groups to find the survival functions evaluated at 32 and plot the two survival functions in the same Figure. Please make conclusions.

[Splus code]:

```
##### (a)
lambda=sum(leukemia$status)/sum(leukemia$time)
lambda
k.fit0 = survfit(Surv(time, status), data=leukemia)

### Compare the estimated survival functions
summary(k.fit0)      ### Kaplan-Meier Estimates
dtime=summary(k.fit0)$time
pfit=exp(-lambda*dtime)
pfit                  ### parametric estimates

plot(k.fit0)
dtime2=seq(0,160,by=1)
pfit2=exp(-lambda*dtime2)
lines(dtime2,pfit2)

##### (b)
k.fit = survfit(Surv(time, status) ~ group, data=leukemia)
summary(k.fit)
plot(k.fit, lty=2:3,col=2:3)
legend(100, 1, c("Maintained", "Nonmaintained"),lty=2:3,col=2:3)
```

Example: Kaplan-Meier Estimate:

For the following data,

Patient	Time	Cens	Treat	Age	LBR
1	281	1	0	46	3.2
2	604	0	0	57	3.1
3	457	1	0	56	2.2
4	384	1	0	65	3.9
5	341	0	0	73	2.8
6	842	1	0	64	2.4
7	1514	1	1	69	2.4
8	182	0	1	62	2.4
9	1121	1	1	71	2.5
10	1411	0	1	69	2.3
11	814	1	1	77	3.8
12	1071	1	1	58	3.1

Cens: censor indicator; Treat: treatment.

Calculate the Kaplan-Meier estimates for two treatment groups and plot the survival functions in the same Figure.

[Splus Code:]

```
time=c(281,604,457,384,341,842,1514,182,1121,1411,814,1071)
status=c(1,0,1,1,0,1,1,0,1,0,1,1)
treat=c(0,0,0,0,0,1,1,1,1,1,1)
age=c(46,57,56,65,73,64,69,62,71,69,77,58)
lbr=c(3.2,3.1,2.2,3.9,2.8,2.4,2.4,2.4,2.5,2.3,3.8,3.1)
data=data.frame(time,status,treat,age,lbr)

k.fit0 = survfit(Surv(time, status), data=data)
summary(k.fit0)
plot(k.fit0)

k.fit = survfit(Surv(time, status) ~ treat, data=data)
summary(k.fit)
plot(k.fit, lty=2:3,col=2:3)
legend(200, 0.3, c("Control", "Treatment"),lty=2:3,col=2:3)
```