Survival HW#5

Tonghao Zhang

2016/11/14

12.4

```r
library(survival)
library(KMsurv)
library(ggplot2)
data("kidney");attach(kidney)
logfit1<-survreg(Surv(time[type==1],delta[type==1])~1,dist = "loglogistic")
logfit2<-survreg(Surv(time[type==2],delta[type==2])~1,dist = "loglogistic")

mu1<--2.964
sigma1<--0.539
lambda1<-exp(-mu1/sigma1)
alpha1<-1/sigma1
covhatstar1<-logfit1$var
covhat1<-covhatstar1
covhat1[,2]<-covhat1[,2]*logfit1$scale
covhat1[2,]<-covhat1[2,]*logfit1$scale
var.lambda1<-exp(-2*mu1/sigma1)*(covhat1[1,1]/sigma1^2+mu1^2*covhat1[2,2]/sigma1^4-2*mu1*covhat1[1,2]/si
var.alpha1<-covhat1[2,2]/sigma1^4
sd.lambda1<-sqrt(var.lambda1)
sd.alpha1<-sqrt(var.alpha1)

mu2<--5.005
sigma2<--1.74
lambda2<-exp(-mu2/sigma2)
alpha2<-1/sigma2
covhatstar2<-logfit2$var
covhat2<-covhatstar2
covhat2[,2]<-covhat2[,2]*logfit2$scale
covhat2[2,]<-covhat2[2,]*logfit2$scale
var.lambda2<-exp(-2*mu2/sigma2)*(covhat2[1,1]/sigma2^2+mu2^2*covhat2[2,2]/sigma2^4-2*mu2*covhat2[1,2]/si
var.alpha2<-covhat2[2,2]/sigma2^4
sd.lambda2<-sqrt(var.lambda2)
sd.alpha2<-sqrt(var.alpha2)
```

a)

group 1: $\hat{(\alpha)} = 1.8552876, \hat{(\lambda)} = 0.0040906, var(\hat{(\alpha)}) = 0.38119, var(\hat{(\lambda)}) = 0.0042164$

group 2: $\hat{(\alpha)} = 0.5747126, \hat{(\lambda)} = 0.0563351, var(\hat{(\alpha)}) = 0.1496979, var(\hat{(\lambda)}) = 0.0249043$

b)

```r
expfit1<-survreg(Surv(time,delta)~1,data=kidney[type==1,],dist = "exponential")
expfit2<-survreg(Surv(time,delta)~1,data=kidney[type==2,],dist = "exponential")
LR1<--2*(expfit1$loglik[1]-logfit1$loglik[1])
```

```
LR2<--2*(expfit2$loglik[1]-logfit2$loglik[1])
p1L<-1-pchisq(LR1,1)
p2L<-1-pchisq(LR2,1)
wald1<-sigma1*(1/covhat1[2,2])*sigma1
wald2<-sigma2*(1/covhat2[2,2])*sigma1
p1W<-1-pchisq(wald1,1)
p2W<-1-pchisq(wald2,1)
```

group 1: likelihood ratio test statistic is 3.8707427, p-value is 0.0491348. Wald test statistic is 23.6886076, p-value 1.1325096×10^{-6}.

group 2: likelihood ratio test statistic is 7.4943038, p-value is 0.0061894. Wald test statistic is 4.5657264, p-value 0.0326178.

c)

```
S1<-1/(1+lambda1*5^alpha1)
S2<-1/(1+lambda2*5^alpha2)
kpfit1<-survfit(Surv(time,delta)~1,data = kidney[type==1,])
kpfit1.est<-kpfit1$surv[3];CI.1<-c(kpfit1$lower[3],kpfit1$upper[3])
kpfit2<-survfit(Surv(time,delta)~1,data = kidney[type==2,])
kpfit2.est<-kpfit2$surv[3];CI.2<-c(kpfit2$lower[3],kpfit2$upper[3])
```

The point estimate of first group is 0.9250552 [0.5285844,1], while KP estimate is 0.9523256, [0.8899461, 1].

The point estimate of group type 2 is 0.8756071 [0.727708,1], while KP estimate is 0.8881579, [0.8174954, 0.9649283].

d)

```
logfit.all<-survreg(Surv(time,delta)~as.factor(type),data=kidney,dist = "loglogistic")
mu.all<-2.9590
sigma.all<-1.06
gamma.all<-0.4588
alpha.all<-1/sigma.all
lambda.all<-exp(-mu.all/sigma.all)
beta.all<--gamma.all/sigma.all

afactor<-exp(beta.all)^(1/alpha.all)
podd<-exp(-beta.all)
```

H_0: no effect. p-value = 0.357, fail to reject the null hypothesis.

The point estimate of relative odds is 1.5416144, point estimate of acceleration factor is 0.6320416.

12.10

```
library(ggplot2)
library(gridExtra)
cumhaz.1<--log(kpfit1$surv);cumhaz.2<--log(kpfit2$surv)
expfit1<-survreg(Surv(time,delta)~1,data = kidney[type==1,],dist = "exp")
expfit2<-survreg(Surv(time,delta)~1,data = kidney[type==2,],dist = "exp")
lambda1.exp<-exp(-3.48)
lambda2.exp<-exp(-4.01)
```

```r
p1<-qplot(kpfit1$time,cumhaz.1)+labs(x="",y="")+ggtitle("exponetial")+geom_line()
p2<-qplot(kpfit2$time,cumhaz.2)+labs(x="",y="")+ggtitle("exponetial")+geom_line()

weibfit1<-survreg(Surv(time,delta)~1,data = kidney[type==1,],dist = "weibull")
mu1.weib<-3.194;sigma1.weib<-0.627
lambda1.weib<-exp(-mu1.weib/sigma1.weib);alpha1.weib<-1/sigma1.weib
p3<-qplot(log(kpfit1$time),log(cumhaz.1))+labs(x="",y="")+ggtitle("weibull")+geom_line()
weibfit2<-survreg(Surv(time,delta)~1,data = kidney[type==2,],dist = "weibull")
mu2.weib<-5.411;sigma2.weib<-1.85
lambda2.weib<-exp(-mu2.weib/sigma2.weib);alpha2.weib2<-1/sigma2.weib
p4<-qplot(log(kpfit2$time),log(cumhaz.2))+labs(x="",y="")+ggtitle("weibull")+geom_line()

lognormal1<-survreg(Surv(time,delta)~1,data = kidney[type==1,],dist = "loggaussian")
lognormal2<-survreg(Surv(time,delta)~1,data = kidney[type==2,],dist = "loggaussian")
p5<-qplot(log(kpfit1$time),qnorm(1-exp(-cumhaz.1)))+labs(x="",y="")+ggtitle("log-normal")+geom_line()
p6<-qplot(log(kpfit2$time),qnorm(1-exp(-cumhaz.2)))+labs(x="",y="")+ggtitle("log-normal")+geom_line()
p7<-qplot(log(kpfit1$time),log(exp(-log(kpfit1$surv))-1))+labs(x="",y="")+ggtitle("log-logistic")+geom_
p8<-qplot(log(kpfit2$time),log(exp(-log(kpfit2$surv))-1))+labs(x="",y="")+ggtitle("log-logistic")+geom_
grid.arrange(p1,p2,p3,p4,ncol=2)
```
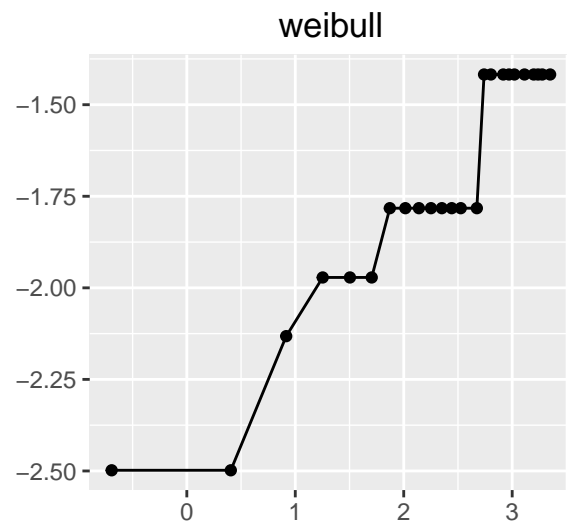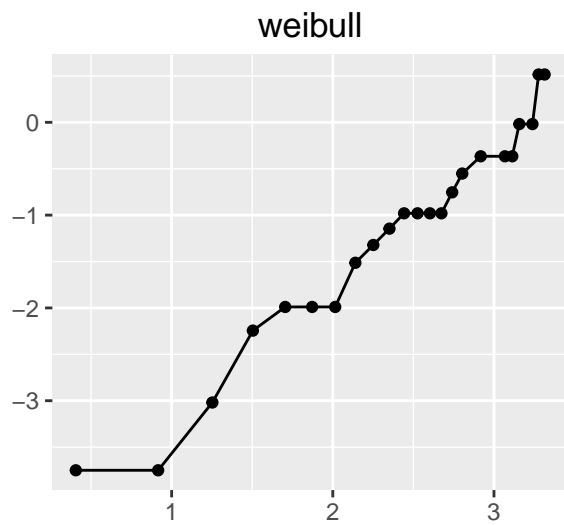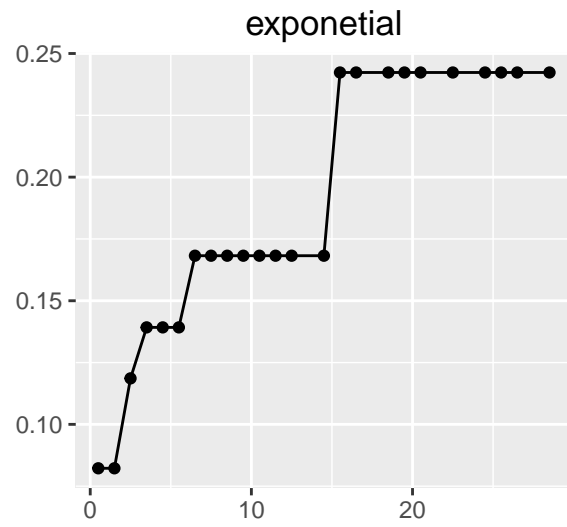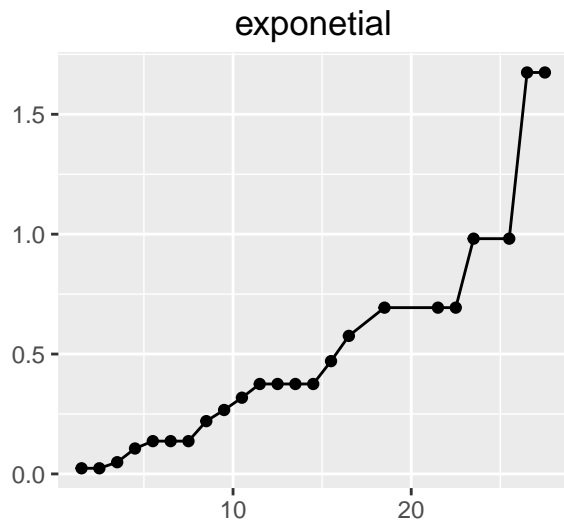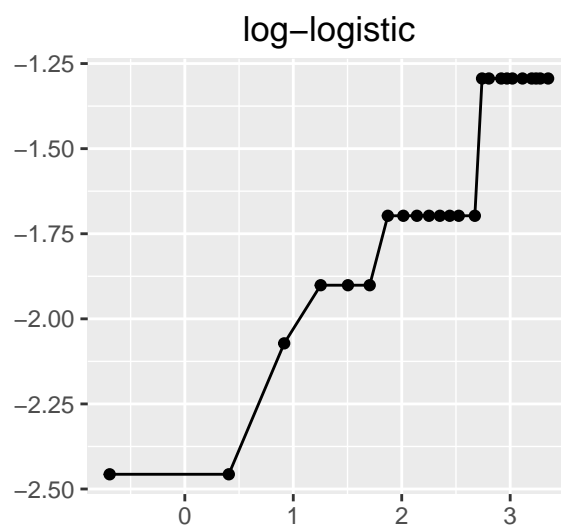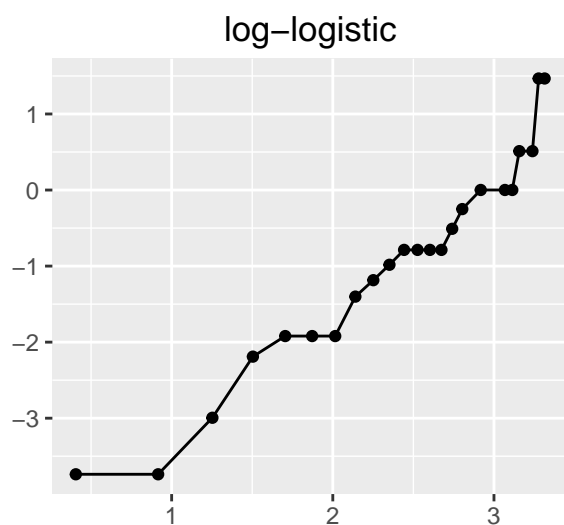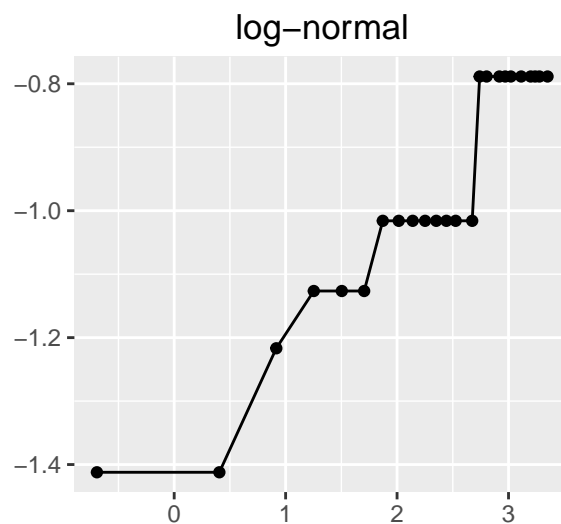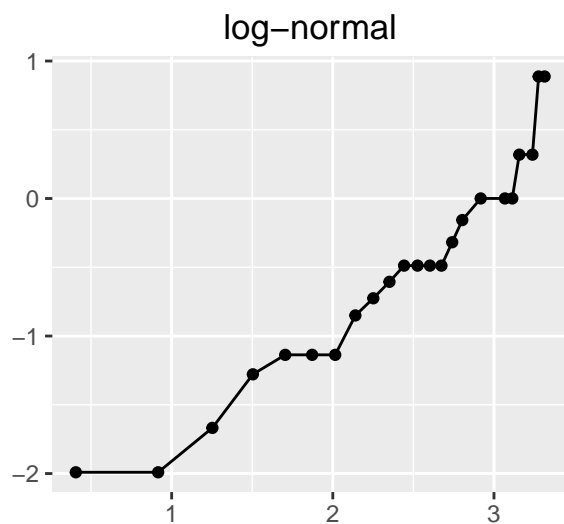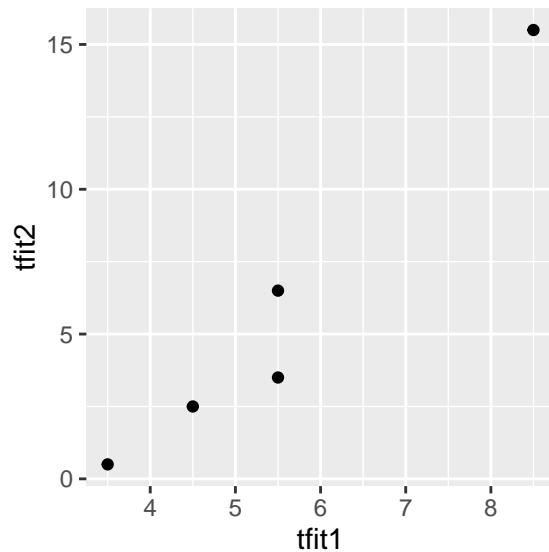
```r
grid.arrange(p5,p6,p7,p8,ncol=2)
```

As there not much observed event in type 2 group, its estimate is more like stepwise, not smooth. From the plots above, it seems to me weibull and log-logistic are better models.

12.12

```
tfit2<-c(0.5,2.5,3.5,6.5,15.5)
tfit1<-c(3.5,4.5,5.5,5.5,8.5)
qplot(tfit1,tfit2)
```

Again number of observed events in group 2 is too small, I can only find 5 quantiles. From this five points qqplot log-logistic model is ok.

12.14

a)

```
data("hodg")
G<-1*(hodg$gtype==1);D<-1*(hodg$dtype==1)
weibfit<-survreg(Surv(time,delta)~G+D+G*D,data = hodg,dist = "weibull")
summary(weibfit)
```

```
##
## Call:
## survreg(formula = Surv(time, delta) ~ G + D + G * D, data = hodg,
##     dist = "weibull")
##              Value Std. Error    z       p
## (Intercept)  6.952      0.641 10.85 2.03e-27
## G           -3.319      0.991 -3.35 8.13e-04
## D           -1.160      0.842 -1.38 1.68e-01
## G:D          5.358      1.377  3.89 9.98e-05
## Log(scale)   0.503      0.167  3.01 2.63e-03
##
## Scale= 1.65
##
## Weibull distribution
## Loglik(model)= -176.5   Loglik(intercept only)= -183.3
##  Chisq= 13.54 on 3 degrees of freedom, p= 0.0036
## Number of Newton-Raphson Iterations: 5
## n= 43
```

```
mu<-6.952;sigma<-1.65;gamma<-c(-3.319,-1.160,5.358)
beta<--gamma/sigma
lambda<-exp(-mu/sigma);alpha<-1/sigma

coxsnell.weib<-lambda*exp(beta[1]*G+beta[2]*D+beta[3]*G*D)*hodg$time^alpha
```
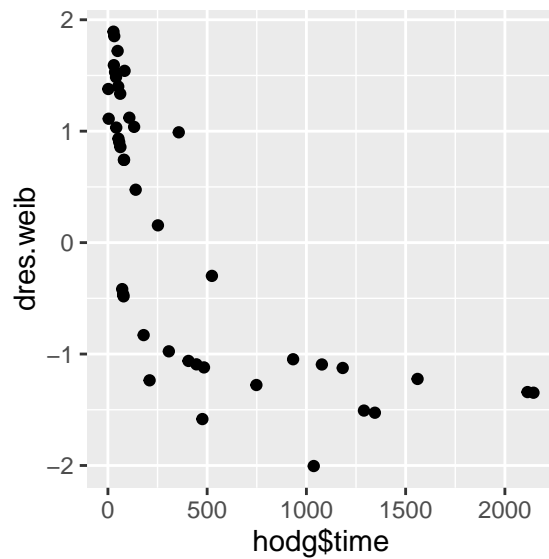
```
kp.cs<-survfit(Surv(coxsnell.weib,hodg$delta)~1)
qplot(kp.cs$time,-log(kp.cs$surv))
```
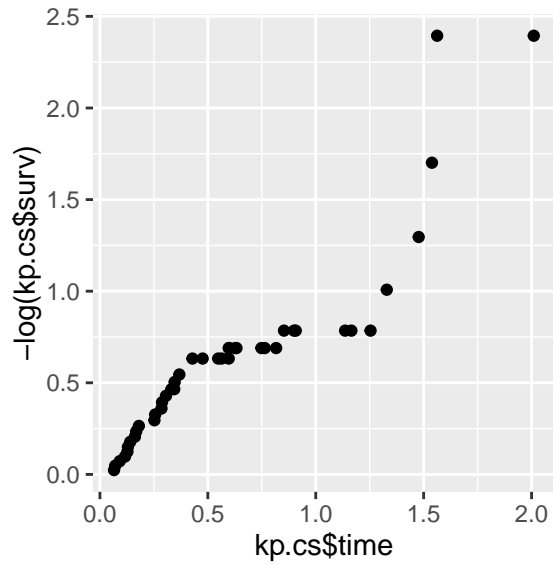

There is a severe departure from straight line, which indicates Weibull is not a suitable model here.

b)

```
mres.weib<-hodg$delta-coxsnell.weib
dres.weib<-sign(mres.weib)*(-2*(mres.weib+hodg$delta*log(hodg$delta-mres.weib)))^0.5
qplot(hodg$time,dres.weib)
```


Deviance plot indicates Weibull model is not suitable, it is very far from what we expect from a random noise plot.

c)

```
loglogistic<-survreg(Surv(time,delta)~G+D+G*D,data = hodg,dist = "loglogistic")
mu<--5.991;sigma<-1.26;gamma<-c(-2.874,-0.786,4.411)
```

```
beta<--gamma/sigma
lambda<-exp(-mu/sigma);alpha<-1/sigma

#coxsnell.ll<-log(1+exp(-loglogistic$linear.predictors*alpha)*hodg$time^alpha)
coxsnell.ll<-log(1/(1+lambda*exp(beta[1]*G+beta[2]*D+beta[3]*G*D)*hodg$time^alpha))
kp.cs.ll<-survfit(Surv(coxsnell.ll,hodg$delta)~1)
qplot(kp.cs.ll$time,-log(kp.cs.ll$surv))
```
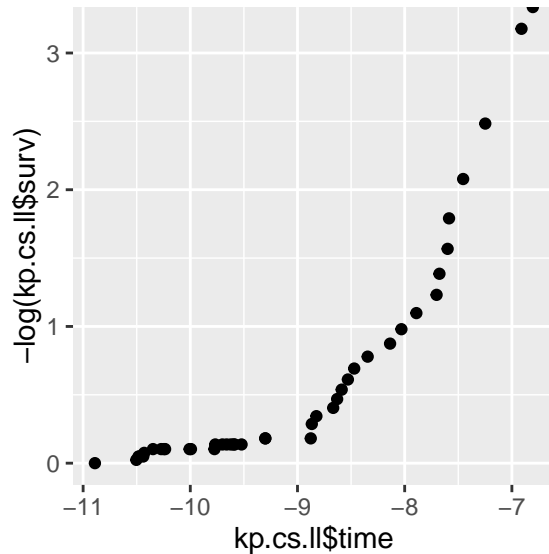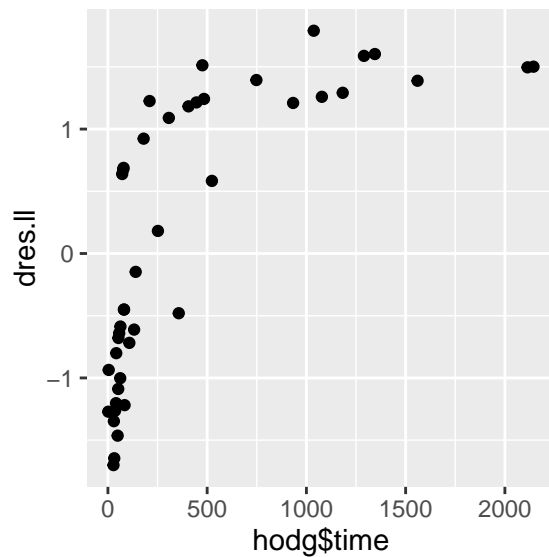


```
mres.ll<-hodg$delta-coxsnell.ll
dres.ll<-resid(loglogistic,type = "deviance")
qplot(hodg$time,dres.ll)
```


Log-logistic model is not suitable neither.