HW2

```
transplant <- read.table("~/Desktop/stat5703/HW/HW2/transplant.txt", quote="\"", comment.char="")
colnames(transplant)<-c("time","type","s")
```

####1 Answer: We could assume the right censoring in this data set is random because the death of the plant will occur after the recorded time.

####2

```
library(ggfortify)
```

```
## Loading required package: ggplot2
```

```
library(survival)
```

```
##
## Attaching package: 'survival'
```

```
## The following object is masked _by_ '.GlobalEnv':
##
##     transplant
```
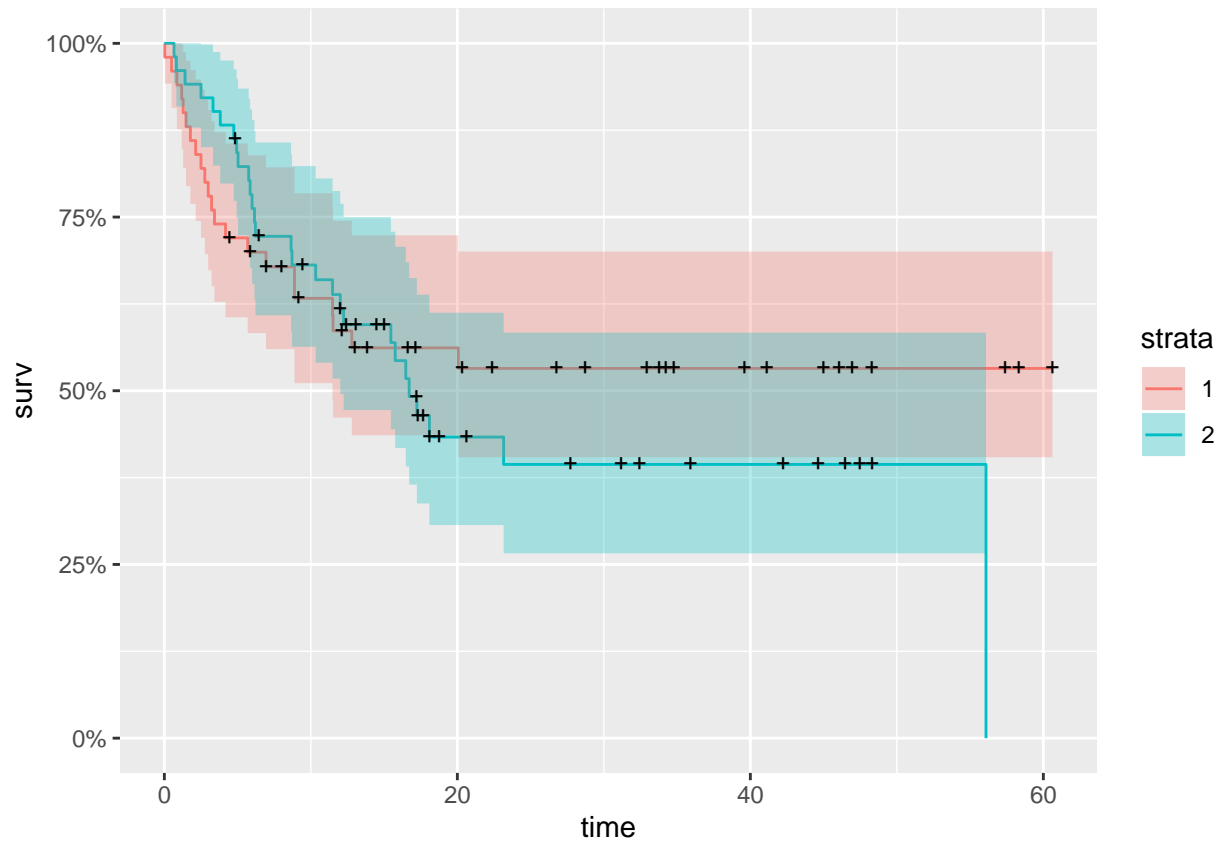
```
library(dplyr)
```

```
##
## 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
```

```
library(ggplot2)
autoplot(survfit(Surv(time, s) ~ type,data=transplant))
```

The plot seem to indicate a difference between two groups. Transplant allogeneic seems to be more efficient.

3

```
fit_exp <- survreg(Surv(time,s)~type, dist="exp",data=transplant)
summary(fit_exp)
```

```
##
## Call:
## survreg(formula = Surv(time, s) ~ type, data = transplant, dist = "exp")
##              Value Std. Error    z      p
## (Intercept)  4.066      0.466  8.72 <2e-16
## type        -0.325      0.285 -1.14   0.25
##
## Scale fixed at 1
##
## Exponential distribution
## Loglik(model)= -228   Loglik(intercept only)= -228.6
##  Chisq= 1.31 on 1 degrees of freedom, p= 0.25
## Number of Newton-Raphson Iterations: 5
## n= 101
```

From the summary table, we can see that the estimated value of autologous is a negative value (-0.325). Which means Transplant autologous would cause less survival time. This result is consistant with our previous observation.

####4 From summary table, we get the p value is 0.25 which is larger than normal thresholds 0.05. So there is no strong evidence against the null hypothesis :there is no difference between the transplant type. These conclusion depend on the assumption of exponential parametric model and asymptotic approximation to the distribution of the likelihood ratio statistic.

####5

```r
fit<- survfit(Surv(time,s)~type,data=transplant)
par(mfrow=c(1,1))
plot(fit,xlab="Days",conf.int=TRUE,main="Survival times",col=c("red","blue"))

x <- seq(from=0,to=2300,by=1)
lines(x,1-pexp(x,exp(-coef(fit_exp)[1])),col="red")
lines(x,1-pexp(x,exp(-sum(coef(fit_exp)))),col="blue")
```

Survival times
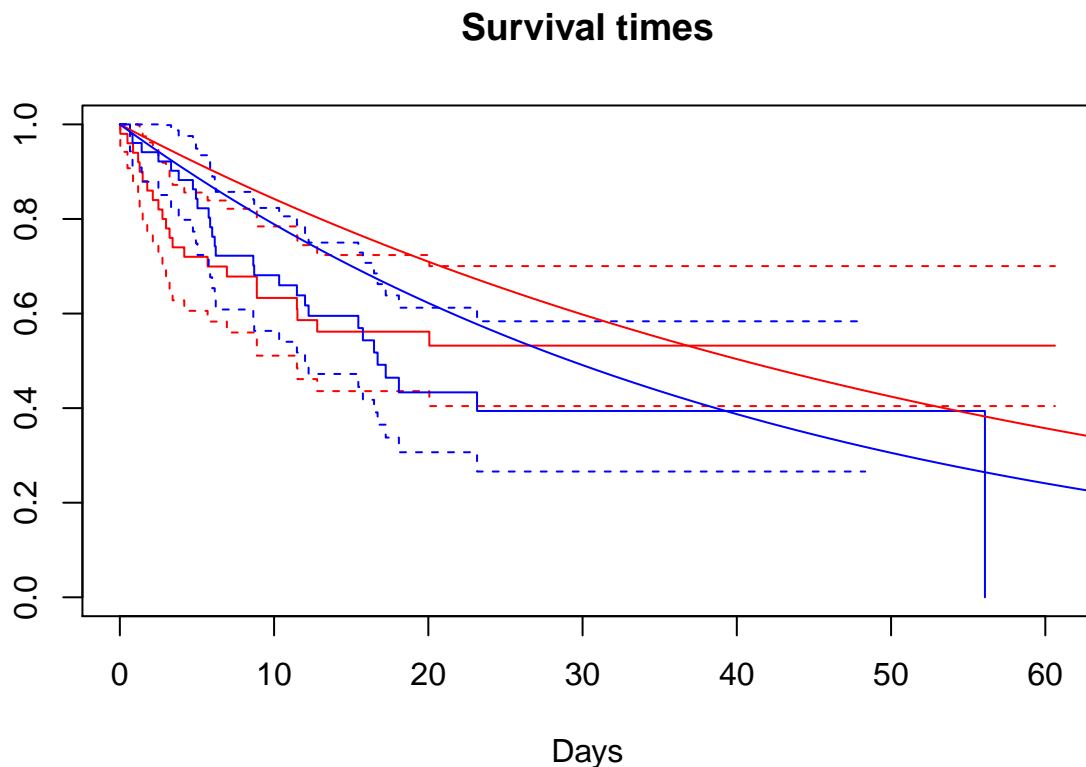
From the graph, we can see that exponential model does great to fit. Most time the lines represent the exponential model lie inside the corresponding types' confidence interval. It does better job for fitting type 2 than fitting type 1. But it is not really good when fitting type 1 as it the fitted line goes outside of confidence interval after around time=55.

```r
fit_wei <- survreg(Surv(time,s)~type,data=transplant)
summary(fit_wei)
```

```
##
## Call:
## survreg(formula = Surv(time, s) ~ type, data = transplant)
```

3

```
##               Value Std. Error     z       p
## (Intercept)   4.342      0.698   6.22   5e-10
## type         -0.374      0.420  -0.89  0.3742
## Log(scale)    0.388      0.123   3.15  0.0016
##
## Scale= 1.47
##
## Weibull distribution
## Loglik(model)= -222   Loglik(intercept only)= -222.4
##   Chisq= 0.8 on 1 degrees of freedom, p= 0.37
## Number of Newton-Raphson Iterations: 5
## n= 101
```

```r
par(mfrow=c(1,1))
plot(fit,col=c("red","blue"),xlab="Days",conf.int=TRUE,main="Survival times")

x <- seq(from=0,to=2300,by=1)
lines(x,1-pexp(x,exp(-coef(fit_exp)[1])),col="red")
lines(x,1-pexp(x,exp(-sum(coef(fit_exp)))),col="blue")

lines(x,1-pweibull(x,fit_wei$scale,exp(coef(fit_wei)[1])),col="green")
lines(x,1-pweibull(x,fit_wei$scale,exp(sum(coef(fit_wei)))),col="darkgreen")
```

Survival times
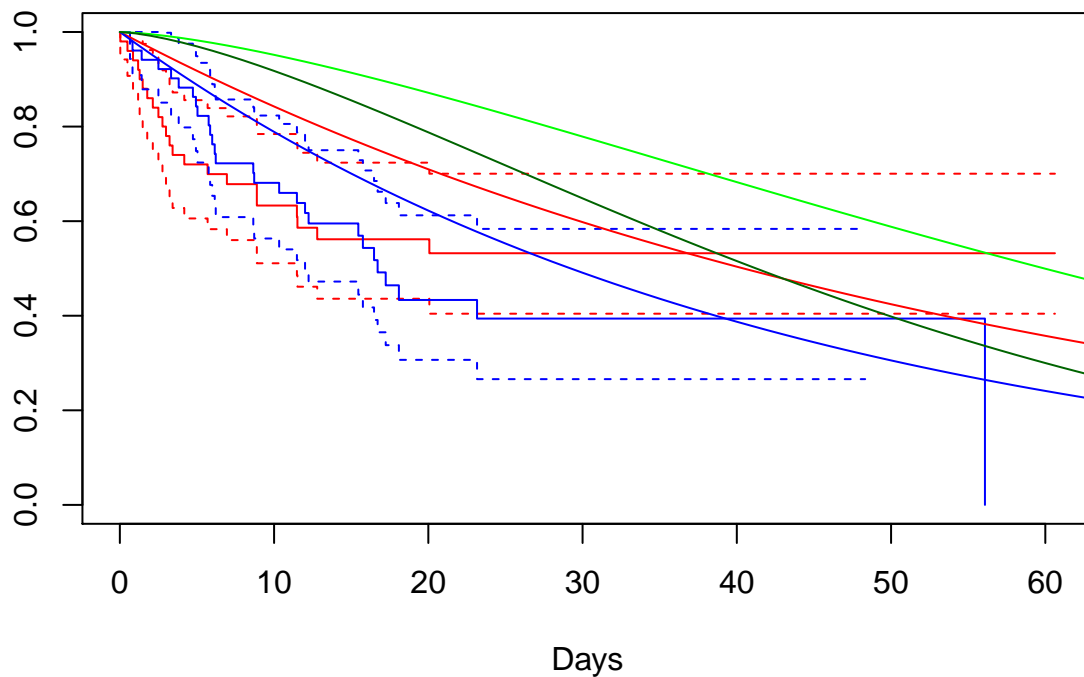
We see that the Weibull model does not fit the data better. The new parameter log(scale) is significant since the pvalue is small, but the type still seems not significant.From the plot we cannot see any improvement in the estimated Weibull survival functions.