

STAT40810 — Stochastic Models

Brendan Murphy

Week 5

Parametric Survival Analysis

Observation Types

- In typical survival analyses we have observations of the following types:
 - Fully observed: We observe T_i
 - Left censored: We observe that $T_i \leq c_i$.
 - Right censored: We observe that $T_i > c_i$.
 - Interval censored: We observed that $a_i < T_i \leq b_i$.
- The most common are fully observed and right censored, but the others can happen too.
- We can write out the contribution to the likelihood of each data type.

Likelihood Contribution

- When forming the likelihood the data types have the following contribution to the likelihood.
 - Fully observed: $f(t_i)$
 - Left censored: $\mathbb{P}\{T_i \leq c_i\} = F(c_i) = 1 - S(c_i)$
 - Right censored: $\mathbb{P}\{T_i > c_i\} = S(c_i) = 1 - F(c_i)$
 - Interval censored: $\mathbb{P}\{a_i < T_i \leq b_i\} = F(b_i) - F(a_i) = S(a_i) - S(b_i)$.
- The overall likelihood is formed by producting the relevant term for each observation.

Example: Leukemia

- The survival times of a number of leukemia patients who were on a maintained treatment were recorded.
- The times recorded are:
9 13 13+ 18 23 28+ 31 34 45+ 48 161+
- The times marked with a + are right censored.
- We want to fit an exponential and a Weibull model to the data and contrast them.

```
#Load survival package
library(survival)

#Load leukemia data
data(leukemia)

#Extract data for Maintained group
dat<-leukemia[leukemia$x=="Maintained",1:2]
dat

#Set up Surv() object
survdat <- Surv(time=dat[,1],event=dat[,2])
survdat

#Fit the model to the data
fit1 <- survreg(survdat~1,dist="exponential")
lambda <- 1/exp(fit1$coef)

#Study fit
summary(fit1)

#Plot the fit
tvec<-seq(0,200,length=201)
plot(tvec,dexp(tvec,lambda),type="l")
```

```
#Fit the model to the data
fit2 <- survreg(survdat~1,dist="weibull")

#Study fit
summary(fit2)

#Write parameters in original form
alpha <- 1/fit2$scale
beta <- exp(fit2$coef)

#Plot the fitted curves
tvec<-seq(0,200,length=201)
plot(tvec,dexp(tvec,lambda),type="l")
points(tvec,dweibull(tvec,alpha,beta),type="l",col="blue",lty=3)
```

Fit Comparison

- The shape parameter of the Weibull is not significantly different from 1.

Call:

```
survreg(formula = survdat ~ 1, dist = "weibull")
```

	Value	Std. Error	z	p
(Intercept)	4.0997	0.366	11.187	4.74e-29
Log(scale)	-0.0314	0.277	-0.113	9.10e-01

Scale= 0.969

Weibull distribution

Loglik(model)= -35.7 Loglik(intercept only)= -35.7

Number of Newton-Raphson Iterations: 5

n= 11

Fit Comparison

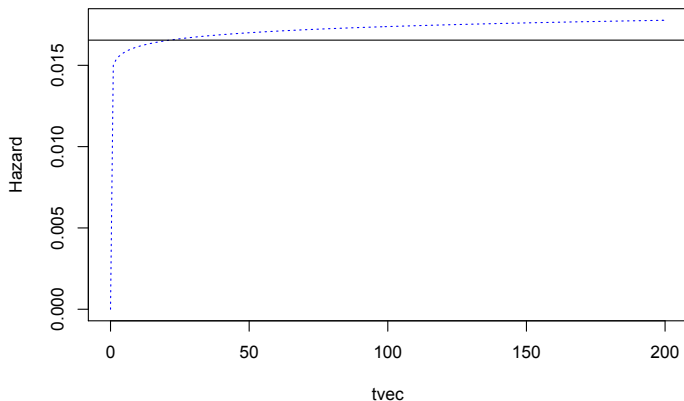


Figure : The hazard function for the Weibull is relatively flat for large times.