

Lecture fourteen: Parametric PH Models: The Weibull Model(I)

1. Population Survivor Distribution
 - (a) The distribution of a population survival is often unknown.
 - (b) Any theoretical distribution to approximate the true one?
 - (c) Only few theoretical distributions available. Its use depends on
 - i. How good the approximation is.
 - ii. Mathematical tractability
 - (d) Empirical distribution solely based on data, and it's non-parametric.
2. The strengths and limitations of parametric models (Therneau and Grambsch, 2000, pp. 28-31; JASA, 2004, vol. 99: 890-896)
 - (a) Inferences will be more precise if the parametric assumption is valid.
 - (b) Estimates tend to have smaller standard error (thus more efficient).
 - (c) Few distributions available for parametric analysis, and difficulties in justification.
 - (d) Bias when a wrong distribution is chosen.
3. The exponential and Weibull distributions

(a) **Exponential**

Hazard function:

$$h(t) = \lambda,$$

for $0 \leq t < \infty$, $\lambda > 0$.

Survival and density functions:

$$S(t) = \exp(-\lambda t),$$

and

$$f(t) = \lambda \exp(-\lambda t),$$

Mean and the p th percentile ($S\{t(p)\} = 1 - (p/100)$):

$$E(T) = \frac{1}{\lambda}.$$

$$t(p) = \frac{1}{\lambda} \log\left(\frac{100}{100-p}\right).$$

Memoryless property:

$$P(T > t_1 + t_0 | T > t_0) = \exp\{-\lambda t_1\}.$$

The result means that conditional on survival to time t_0 , the excess survival time beyond t_0 still has an exponential distribution with same parameter λ .

(b) **Weibull**

Hazard function:

$$h(t) = \lambda \gamma t^{\gamma-1},$$

for $0 \leq t < \infty$, $\lambda > 0$ and $\gamma > 0$.

Survival and density functions:

$$S(t) = \exp(-\lambda t^\gamma),$$

$$f(t) = \lambda \gamma t^{\gamma-1} \exp(-\lambda t^\gamma),$$

Mean and the p th percentile:

$$E(T) = \lambda^{-1/\gamma} \Gamma(\gamma^{-1} + 1).$$

$$t(p) = \left[\frac{1}{\lambda} \log\left(\frac{100}{100-p}\right) \right]^{1/\gamma}.$$

(c) Figures: hazard and density for Weibull distribution

```

# Splus program generating the figures
weibull.s<-function(){
  sh <-c(3,1.5,0.5)
  med <- 20
  sc <- log(2)/med^sh
  t<- seq(0,45, by=0.5)
  f1<-sc[1]*sh[1]*t^(sh[1]-1)*exp(-sc[1]*t^sh[1])
  f2<-sc[2]*sh[2]*t^(sh[2]-1)*exp(-sc[2]*t^sh[2])
  f3<-sc[3]*sh[3]*t^(sh[3]-1)*exp(-sc[3]*t^sh[3])
  h1<-sc[1]*sh[1]*t^(sh[1]-1)
  h2<-sc[2]*sh[2]*t^(sh[2]-1)
  h3<-sc[3]*sh[3]*t^(sh[3]-1)
  motif()
  par(mfrow=c(2,2))
  plot(t, h1, xlab="Time", ylab="Hazard function",
        type="l",lty=1,lwd=1)
  lines(t, h2,lty=2,type="l",lwd=2)
  lines(t, h3,lty=3,type="l",lwd=2)
  legend(5,0.45,c("shape = 3","shape=1.5","shape=0.5"),
        cex=0.6,lty=1:3)
  plot(t, f1, xlab="Time", lty=1,lwd=1,type="l",
        ylab="probability density function")
  lines(t, f2,lty=2,type="l",lwd=2)
  lines(t, f3,lty=3,type="l",lwd=2)
  text(t[50]+5, f1[50]-0.0025, "shape=3",cex=0.6)
  text(t[74]+4, f2[74]+0.0025, "shape=1.5",cex=0.6)
  text(t[34]+3, f3[34]+0.003, "shape=0.5",cex=0.6)
}

```

4. Assessing the suitability of a parametric model

Recall that exponential distribution is a special case of Weibull. For Weibull distribution,

$$\log\{-\log S(t)\} = \log \lambda + \gamma \log t.$$

- (a) Example 5.1: IUD data
- (b) The SAS program

```

options ls=80;
libname fu '.././sdata';
data w;
    set fu.iud;
filename gsasfile 'ex51a.gsf';
goptions reset=all gunit=pct border ftext=swissb htitle=6 htext=2.5
gaccess=gsasfile ROTATE=LANDSCAPE gsfmode=replace device=ps;
title1 'Log-cumulative hazard plot for IUD data';
symbol1 c=black i=join v=triangle height=3;

proc lifetest notable plots=(lls) method=pl outsurv=a;
    time survt*censor(0);
data x1;
    set a;
    if survival=0 then delete;
    if survt = 0 then delete;
    ls=-log(survival);
    lls=log(-log(survival));
    logt=log(survt);
proc reg;
    model ls = survt/noint;
proc reg;
    model lls = logt;

filename gsasfile 'ex51b.gsf';
goptions reset=all gunit=pct border ftext=swissb htitle=6 htext=2.5
gaccess=gsasfile ROTATE=LANDSCAPE gsfmode=replace device=ps;
axis1 label=('survival time')
    width=2
    order=(10 to 107 by 15);
axis2 label=(a=90 'survival functions')
    width=2;
title1 'survival functions for IUD data';
footnote1 c=black f=special h=4 'L L'
    f=swiss h=4 ' exp'
    c=blue f=special h=3 ' m m '
    f=swissb h=4 ' Weibull';
symbol1 c=black i=join v=triangle height=3;

```

```

symbol2 c=blue i=join v=star height=3;
/* plug in parameters estimated from above regressions */
data x2;
    set x1;
    se=exp(-0.01021*survt);
    sw=exp(-exp(-6.3582)*(survt**(1.377252)));
proc gplot ;
    plot se*survt sw*survt /overlay haxis=axis1 vaxis=axis2;
run;

```

(c) *The log-cumulative hazard plot*

5. Fitting a parametric model to a single sample

Likelihood function, log-likelihood function, score equations, and information matrix, MLE (see the second lecture notes for chapter 3)

(a) Exponential distribution case

- i. without censoring
- ii. with censoring (what if no events?)

follow the steps in the book at page 182

(b) Example 5.2: IUD data (by hand and by sas program)

```

/* SAS program */
options ls=80;
libname fu '../..sdata';
data w;
    set fu.iud;
proc lifereg;
    model survt*censor(0) = / covb dist=exponential;
run;
/* SAS output */
Analysis of Parameter Estimates

```

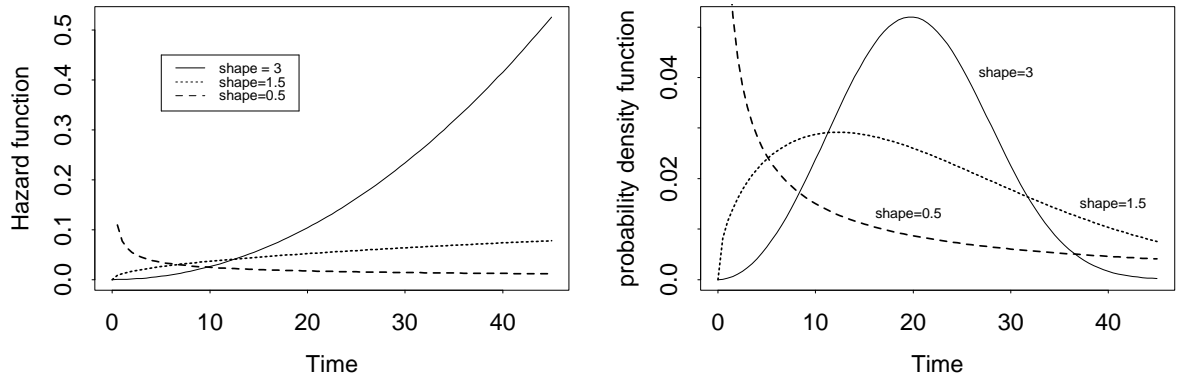
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr>ChiSq
Intercept	1	4.7555	0.3333	4.1022	5.4088	203.53	<.0001

Scale	0	1.0000	0.0000	1.0000	1.0000
Weibull Scale	1	116.2221	38.7407	60.4721	223.3688
Weibull Shape	0	1.0000	0.0000	1.0000	1.0000

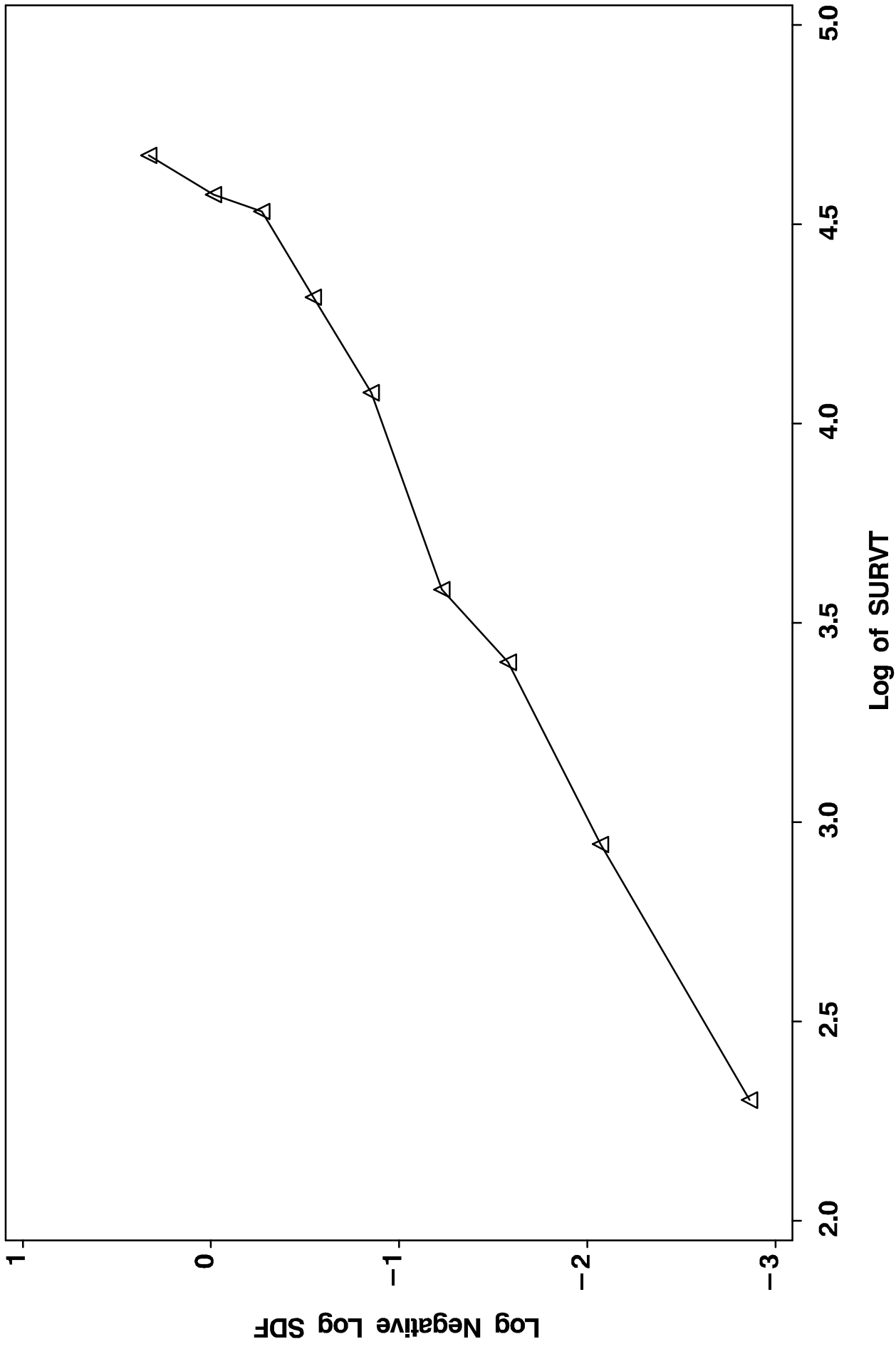
(c) Weibull distribution case

(d) Example 5.3: IUD data

Figure 1:



Log – cumulative hazard plot for IUD data



survival functions for IUD data

