Lecture fifteen: Parametric PH Models: The Weibull Model(II)

1. The Weibull proportional hazards model

   (a) The proportional hazards property

      i. One dummy variable (two groups)

      $$h_i(t) = e^{\beta x_i} h_0(t),$$

      where x_i is dummy variable, and the baseline hazard is from Weibull distribution (known instead of arbitrary in Cox model).

      $$h_0(t) = \lambda \gamma t^{\gamma - 1}. \quad (1)$$

      Let $\psi = exp(\beta)$, then $\psi h_0(t)$ is hazard function for a Weibull distribution with scale parameter $\psi\lambda$ and the same shape parameter γ (*the proportional hazards property*).

      ii. More than one covariates

      $$h_i(t) = exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}) h_0(t),$$

      where $h_0(t)$ is from (1). The corresponding survival function and density function are still from Weibull distribution with scale parameter $\lambda exp(\beta' x_i)$ and shape parameter γ. Namely

      $$S_i(t) = exp\{-\lambda exp(\beta' x_i) t^{\gamma}\}. \quad (2)$$

      $$f_i(t) = \lambda \gamma exp(\beta' x_i) t^{\gamma - 1} exp\{-\lambda exp(\beta' x_i) t^{\gamma}\}.$$

      iii. Maximum Likelihood Estimation (**Not MPLE as in Cox model**)

      iv. Formulas for median and pth percentile of survival time.

      See the ones in lecture one of this chapter.

   (b) An alternative form of the Weibull PH model

   Consider following transformation of r.v. T_i of survival time:

   $$log\, T_i = \mu + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \ldots + \alpha_p x_{ip} + \sigma \epsilon_i.$$

If ϵ has the *Gumbel* distribution, also known as *extreme value* distribution: see Kalbfleisch and Prentice's book (at page 33) for details, with pdf:

$$f(\epsilon) = exp(\epsilon - e^{\epsilon}),$$

then, $\xi = e^{\epsilon}$ has exponential distribution Exp(1).

Follow steps at page 237- 238, then we have

$$S_i(t) = exp[-exp\{\frac{log\, t - \mu - \alpha' x_i}{\sigma}\}]. \tag{3}$$

In model (3), μ is *intercept* and σ is *scale parameter.*
Let (2) and (3) equal, then we have

$$\lambda = exp(-\mu/\sigma), \qquad \gamma = \sigma^{-1}, \qquad \beta_j = -\alpha_j/\sigma,$$

for $j = 1, 2, \ldots, p$ (page 203, and pages 237-238).

In SAS, the output from **PROC LIFEREG** is expressed in terms of $\mu, \sigma, \alpha_1, \alpha_2, \ldots, \alpha_p$ in equation (3) (still use maximum likelihood estimation, but different parameterization).

2. Example 5.3: IUD data (one sample, intercept only)

   For exponetial distribution, the score equation has closed form solution, but for Weibull distribution, the score equations are non-linear in γ, thus, iterative procedures, such as the Newton-Raphson algorithm, are required.

```
/* SAS program */
options ls=80;
libname fu '../../sdata';
data w;
        set fu.iud;
proc lifereg;
        model survt*censor(0) = / covb dist=weibull;
run;
```

```
/* SAS output */
          Analysis of Parameter Estimates

                             Standard   95% Confidence   Chi-
Parameter     DF Estimate     Error        Limits      Square  Pr>ChiSq

Intercept      1   4.5915    0.2044   4.1910   4.9921 504.74   <.0001
Scale          1   0.5965    0.1638   0.3482   1.0218
Weibull Scale  1  98.6440   20.1601  66.0859 147.2425
Weibull Shape  1   1.6764    0.4604   0.9786   2.8717
```

3. The log-cumulative plot: Example 5.4: Breast cancer study

   Is the Weibull PH model appropriate for the data?

   (a) Log-cumulative hazard plot

   (b) SAS program

```
options ls=80;
libname fu '../../sdata';
data w;
        set fu.hpa;
filename gsasfile 'ex54.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 HPA data';
footnote1 c=black f=special h=4 'L L'
        f=swiss h=4 ' Negative'
        c=blue f=special h=3 '   m m '
        f=swissb h=4 ' Positive';
symbol1 c=black i=join v=triangle height=3;
symbol2 c=blue i=join v=star height=3;

proc lifetest  notable plot = (lls) method=pl;
        time survt*censor(0);
        strata group;
run;
```

4. Example 5.5: Breast cancer study

   Fitting exponential model; calculate the risk ratio and its variance. We use output from sas instead of formulas at page 183.

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

                              Standard   95% Confidence   Chi-
Parameter     DF Estimate      Error        Limits      Square Pr>ChiSq

Intercept      1   5.8003     0.4472   4.9238   6.6768 168.22   <.0001
GROUP          1  -0.9516     0.4976  -1.9269   0.0237   3.66   0.0558
Scale          0   1.0000     0.0000   1.0000   1.0000
Weibull Shape  0   1.0000     0.0000   1.0000   1.0000
```

5. Example 5.6: Breast cancer study

   Fitting Weibull model; calculate the risk ratio and its variance using output from sas program.

```
/* SAS program */
options ls=80;
libname fu '../../sdata';
data w;
        set fu.hpa;
proc lifereg;
        model survt*censor(0) = group / covb dist = weibull;
run;
```

```
/* SAS output */
              Analysis of Parameter Estimates

                            Standard   95% Confidence   Chi-
Parameter      DF Estimate     Error       Limits      Square Pr>ChiSq

Intercept       1    5.8544    0.4989   4.8766  6.8321 137.71    <.0001
GROUP           1   -0.9967    0.5441  -2.0631  0.0697   3.36    0.0670
Scale           1    1.0668    0.1786   0.7684  1.4810
Weibull Shape   1    0.9374    0.1569   0.6752  1.3014


                    Estimated Covariance Matrix

                      Intercept          GROUP           Scale


        Intercept      0.248878      -0.245010        0.026044
        GROUP         -0.245010       0.296036       -0.021310
        Scale          0.026044      -0.021310        0.031880
```

What's the variance of $\hat{t}(50)$? Recall

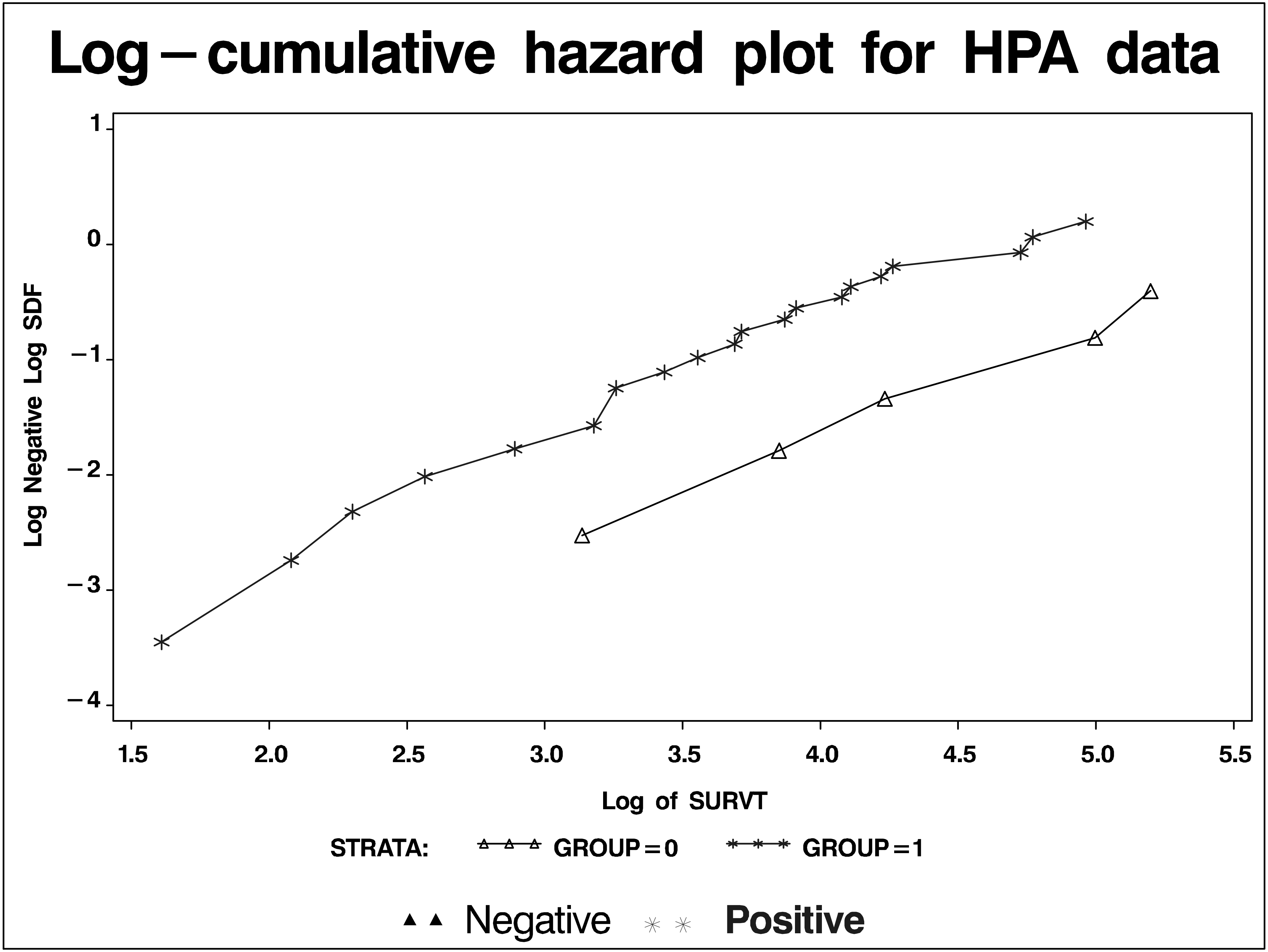
$$\hat{t}(p) = [\frac{1}{\hat{\lambda}} log(\frac{100}{100-p})]^{1/\hat{\gamma}},$$

6. Example 5.7: Breast cancer study: CI for hazard ratio

   (a) Delta method: A function of two estimates (two random variables)

   The general formula for $g(\hat{\theta}_1, \hat{\theta}_2)$ (see the eq. (5.24) at p. 188)

   (b) Illustration: CI for $\hat{\beta}$, i.e.
   $g(\hat{\alpha}, \hat{\sigma}) = -\frac{\hat{\alpha}}{\hat{\sigma}}$.


Log – cumulative hazard plot for HPA data
Log Negative Log SDF
1
0
–1
–2
–3
–4
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
Log of SURVT
STRATA: GROUP = 0 GROUP = 1
Negative
Positive