Lecture Nine: Cox Propotional Hazards Models (IV)

The expected survival for individuals can be estimated after the baseline (cumulative) hazard or baseline survival function is estimated.

1. Estimating the hazard and survival functions

   (a) **Prodiuct-limit estimates**: Cox model (Kalbfleisch and Prentice, *Biometrika*, 60, 267-278 (1973)).

   Assume the hazard is zero except at event times, then

   $$\hat{h}_0(t_{(j)}) = 1 - \hat{\xi}_j,$$

   where $\hat{\xi}_j$ is the solution of equation

   $$\sum_{l \in D(t_{(j)})} \frac{exp(\hat{\beta}' x_l)}{1 - \hat{\xi}_j^{exp(\hat{\beta}' x_l)}} = \sum_{l \in R(t_{(j)})} exp(\hat{\beta}' x_l), \tag{1}$$

   for $j = 1, 2, ..., r$. Where $D(t_{(j)})$ is the set of all d_j individuals who die at the j'th ordered death time $t_{(j)}$.

   If we make another assumption that the hazard of death is constant between adjacent death times, then, $\hat{\xi}_j$ can be regarded as the probability that an individual survives through from $t_{(j)}$ to $t_{(j+1)}$. The baseline survival function can be estimated by

   $$\hat{S}_0(t) = \prod_{j=1}^{k} \hat{\xi}_j,$$

   for $t_{(k)} \leq t < t_{(k+1)}, k = 1, 2, ..., r-1$, and

   $$\hat{H}_0(t) = -log\, \hat{S}_0(t) = -\sum_{J=1}^{k} log\, \hat{\xi}_j,$$

   therefore

   $$\hat{S}_i(t) = [\hat{S}_0(t)]^{exp(\hat{\beta}' x_i)},$$

(b) **The Nelson-Aalen or Breslow estimates**: Empirical $H(t)$

When there are tied survival times, iterative method is used to solve equation (1). The iterative process can be avoid by using an approximation to the summation on the left hand side of equation (1).

The **Breslow estimator** (JRSS, series B, vol. 34, 216-217) of the baseline cumulative hazard is

$$\tilde{H}_0(t) = \sum_{j=1}^{k} \frac{d_j}{\sum_{l \in R(t_{(j)})} exp(\hat{\beta}' x_l)},$$

for $t_{(k)} \leq t < t_{(k+1)}, k = 1, 2, ..., r-1$. The corresponding estimated baseline survival function is given by

$$\tilde{S}_0(t) = \prod_{j=1}^{k} (1 - \frac{d_j}{\sum_{l \in R(t_{(j)})} exp(\hat{\beta}' x_l)}).$$

(Recall the Tylor expansion of $log(1-x)$ used in Nelson-Aalen estimation of cumulative hazard estimation in chapter 2)

2. Example 3.14: Treatment of hypernephroma

(a) SAS program

```
/* For Figure 3.5 */
options linesize=80 nodate nonumber;
libname fu '../../sdata';
data work1;
        set fu.kidney;
if age=2 then a2 = 1; else a2 = 0;
if age=3 then a3 = 1; else a3 = 0;
data nset1;
        input a2 a3 neph;
        datalines;
        0 0 0
        0 0 1
;
proc phreg data=work1;
        model survt*censor(0)=a2 a3 neph / covb risklimits;
```

```
        baseline covariates=nset1 out=pred1 survival=S
               lower=S_lower upper=S_upper /nomean;
data temp1;
        set pred1;
        if a2 = 0 and a3 = 0 and neph = 0 then pattern=1;
        else if a2 = 0 and a3 = 0 and neph = 1 then pattern=2;

filename gsasfile 'kidney1.gsf';
goptions reset = all gunit = pct border ftext=swissb htitle=6
             htext=2.5 gaccess=gsasfile ROTATE=LANDSCAPE
                     gsfmode=replace device=ps;

legend1 label=none shape=line(4)
      value=(f=swiss h=.8 'age < 60 without a nephrectomy'
                'age < 60 with a nephrectomy');
axis1 label=(h=1 f=swiss a=90) minor=(n=1);
axis2 label=(h=1 f=swiss 'Survival Time') minor=(n=4);
title 'Expected survival by nephrectomy';
proc gplot data=temp1;
      plot S*survt=Pattern / legend=legend1 vaxis=axis1
                           haxis=axis2    cframe=white;
      symbol1 interpol=stepLJ h=1 l=1 v=square  c=blue;
      symbol2 interpol=stepLJ h=1 l=2 v=diamond c=black;
      note f=swiss h=1.5 j=c 'Hypernephroma Study';
run;
```

(b) Expected survival: Figure 3.5 and Figure 3.6

3. Concordance index (C-index), predictive ability, and explained variation

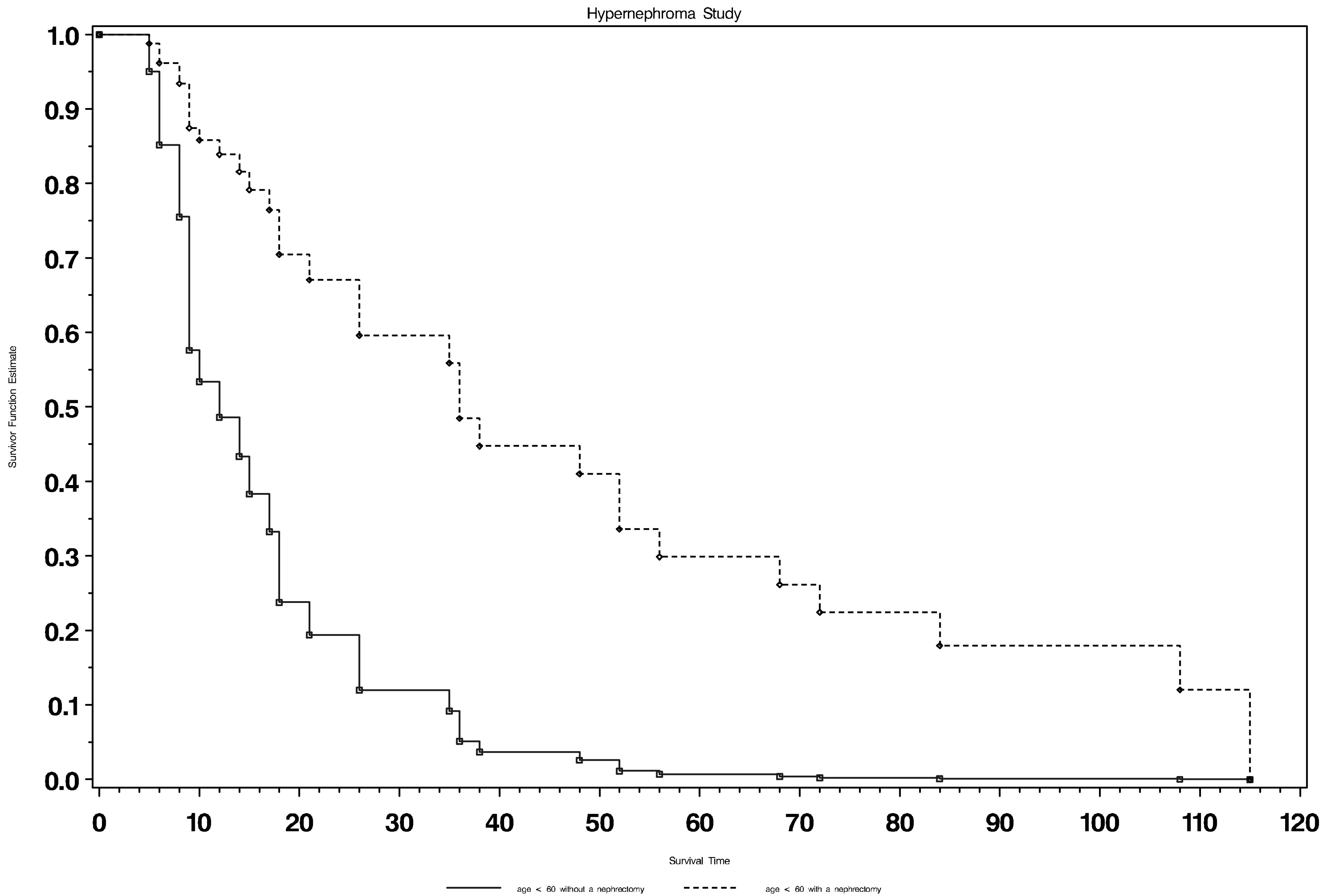
4. Receiver operating characteristics (ROC) and time-dependent ROC

5. Cox model and the log-rank test

   Follow the derivation at page 110 - 112

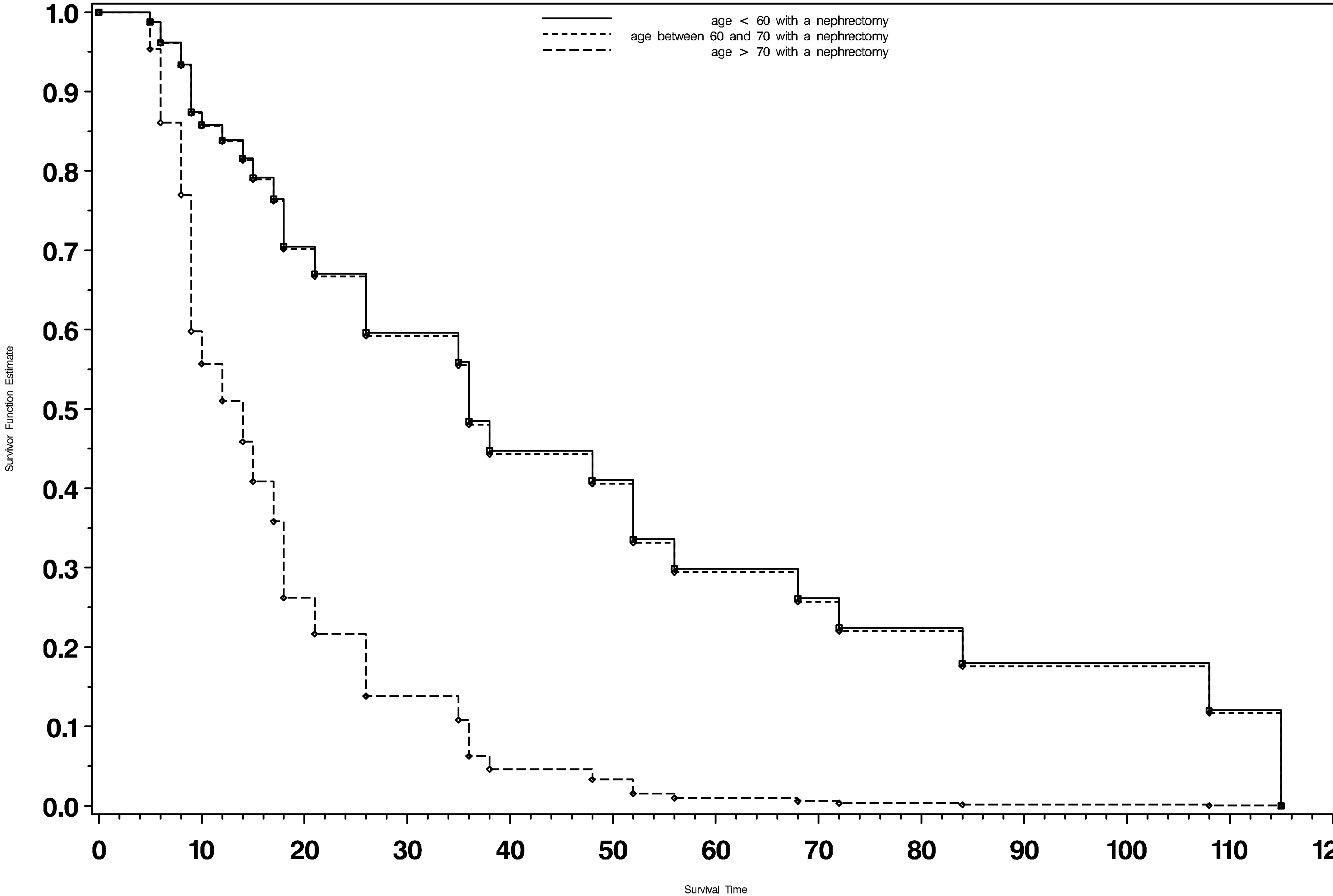
Assignment five: Assume the important prognostic factors for the hypernephroma study (see Exampe 3.4 and Table 3.6) were age group of a patient and nephrectomy. Based on the estimates of Cox model, generate the estimated survival curves for the following two cohorts: $Age < 60$ without nephrectomy and $60 \leq age \leq 70$ with nephrectomy.

Expected survival by nephrectomy



Expected survival by age group

Hypernephroma Study



Expected survival by age group