Lecture 16: Parametric Survival Models

Lin Chen

Department of Public Health Sciences
The University of Chicago

slide 31-46: NOT IN FINAL

Regression models for survival data

- When analyzing survival data, one may want to assess the associations between predictors (characteristics, factors such as treatment) to survival times. One possible and common way is to associate predictors with the hazard function
- There are some different constructs for how survival times are influenced by predictors. We will focus on two of these:
  1. Proportional hazards model
  2. Accelerated failure time model (a non-PH model)

  We begin with the proportional hazards model, as it is (by far) the most common approach

Proportional hazards model

- This model is based on the assumption of proportional hazards.
- Consider the situation where the hazard at a particular time depends on the values of p explanatory variables, $X_1, X_2, ..., X_p$. Denote i^{th} subject's values of the explanatory variable $\boldsymbol{x_i} = (x_{1i}, x_{2i}, ..., x_{pi})$
- Let $h_0(t)$ be the *baseline hazard function*: the hazard function for an individual for whom $\boldsymbol{x_i} = \mathbf{0}$
- The hazard function for the i^{th} individual is given by

$$h_i(t) = \psi(\boldsymbol{x_i}) h_0(t), \tag{1}$$

where the *relative hazard*, $\psi(\boldsymbol{x_i})$, is a function of the values of the explanatory variables for the i^{th} individual. $\psi(\cdot)$ can be interpreted as the hazard at time t for an individual with explanatory variables $\boldsymbol{x_i}$ relative to the hazard for an individual with explanatory variables $\boldsymbol{x} = \mathbf{0}$

Proportional hazards (continued)

- Since the relative hazard, $\psi(\boldsymbol{x_i})$, must be non-negative, it is convenient to model it as $\exp(\eta_i)$, where η_i is a linear combination of the p explanatory variables:

$$\eta_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} \tag{2}$$

so that $\sum_{i=1}^{p} \beta_j x_{ji}$. In matrix notation, $\eta_i = \boldsymbol{\beta}^T \boldsymbol{x_i}$.

- The general proportional hazards model is given by

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}) h_0(t) \tag{3}$$

equivalently,

$$\log\left(\frac{h_i(t)}{h_0(t)}\right) = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} \tag{4}$$

the proportional hazards model may be regarded as a linear model for the (natural) logarithm of the hazard ratio.

Proportional hazards (continued)

- There are two components in the proportional hazards model, the baseline hazard function $h_0(t)$ and the relative hazard function $\exp(\boldsymbol{\beta}^T \boldsymbol{x_i})$.
- If we specify **a parametric functional form** for the baseline hazard function $h_0(t)$, i.e., the distribution of survival time for individuals with $\boldsymbol{x_i} = \mathbf{0}$, we will obtain **a parametric proportional hazards model.**
- Later, we will see that $h_0(t)$ can be left completely unspecified. That gives us the semi-parametric proportional hazards model (Cox proportional hazards model).

Proportional hazards model for the comparison of two groups

- Suppose that we want to compare two groups of survival times: Group I vs. Group II. Let X be the group indicator, 1 if Group II, 0 if Group I.
- Under the proportional hazards model, the hazard of death at time t is given by

$$h_i(t) = e^{\beta x_i} h_0(t) \tag{5}$$

- Consequently, the hazard at time t for an individual in Group I ($x_i = 0$) is $h_0(t)$, and that for an individual in Group II ($x_i = 1$) is $e^{\beta} h_0(t)$.
- The **hazard ratio** is $\frac{h_i(t)}{h_0(t)} = e^{\beta}$

similar concept as e^{β} = odds Ratio (OR)

A parametric survival distribution: Weibull

- If the survival time t follows a **Weibull distribution** $W(\lambda, \gamma)$:

$$\begin{aligned} f(t) &= \lambda\gamma t^{\gamma-1}\exp(-\lambda t^{\gamma}), \gamma > 0, \lambda > 0 \\ h(t) &= \lambda\gamma t^{\gamma-1}, \gamma > 0, \lambda > 0 \\ H(t) &= \lambda t^{\gamma} \\ \log\{H(t)\} &= \log(h) + \gamma \times \log(t) \\ S(t) &= \exp(-\lambda t^{\gamma}) \end{aligned}$$

- The Weibull is a generalized distribution of the exponential model. Exponential model has only **a scale parameter** λ. The Weibull has an additional **shape parameter** γ. Some notes:
  - If $\gamma = 1$, the distribution is exponential. The model is reduced to **the exponential survival model** with a constant hazard λ.
  - If $\gamma > 1$, the hazard is increasing over time, if $\gamma < 1$, decreasing
  - If the Weibull model holds, $H(t)$ should be linear in $\log(t)$ with intercept $\log(\lambda)$ and slope γ

Weibull proportional hazards model for the comparison of two groups

- We now make the additional assumption that the survival times for the individual in Group I have a Weibull distribution $W(\lambda, \gamma)$,

$$h_0(t) = \lambda\gamma t^{\gamma-1} \tag{6}$$

then the hazard function for those in Group II is $e^{\beta} h_0(t)$, that is,

$$e^{\beta}\lambda\gamma t^{\gamma-1} \tag{7}$$

which is a hazard function for a Weibull distribution $W(e^{\beta}\lambda, \gamma)$

- The hazard of a Weibull can be multiplied by a constant and the new hazard is also Weibull. The Weibull distribution has the *proportional hazards property* (not all survival distributions do).
- The Weibull proportional hazards model is given by

① $h_i(t) = \exp(\beta \cdot x_i) h_0(t)$,

Relative Hazard • Baseline Hazard

$e^{\beta x_i}$

② where $h_0(t) = \lambda\gamma t^{\gamma-1}$.

(parametric form)

★ Specification !!! ★

① Generalized Hazard Model

① + ② Weibull Prop. Hazard Model

The log-cumulative hazard plot

- When a single sample of survival times has a Weibull distribution $W(\lambda, \gamma)$, then

$$\log(H_0(t)) = \log\lambda + \gamma \cdot \log(t)$$

- It then follows that if the survival times in a second group have a $W(e^{\beta}\lambda, \gamma)$, then

$$\log(H_1(t)) = (\beta + \log\lambda) + \gamma \cdot \log(t)$$

- Thus, if the assumption of Weibull survival times holds for the data, when we plot the estimated log-cumulative hazard function against the log of the survival time for the two groups, we expect to see two approximately parallel straight lines.

log(t)

log(H)

same slope (constant rate)

Need to check for BOTH!!

- Check Weibull Distribution Assumption: Plot (Straight Line!) → parallel, no changes with time
- If lines curved: Proportional Hazard Assumption → parallel, but changes with time

The log-cumulative hazard plot (continued)

- If the two lines in a log-cumulative hazard plot are essentially straight, but not parallel, this means that the shape parameter γ (which governs how the hazard changes over time) is different in the two groups, and the hazards are no longer proportional.

- If the lines are not particularly straight, the Weibull model may not be appropriate. However, if the two curves are parallel, the proportional hazards assumption is valid, and we can use a model with a different hazard or even without specifying $h_0(t)$.

Example: *Prognosis in women with breast cancer by tumor marker staining*

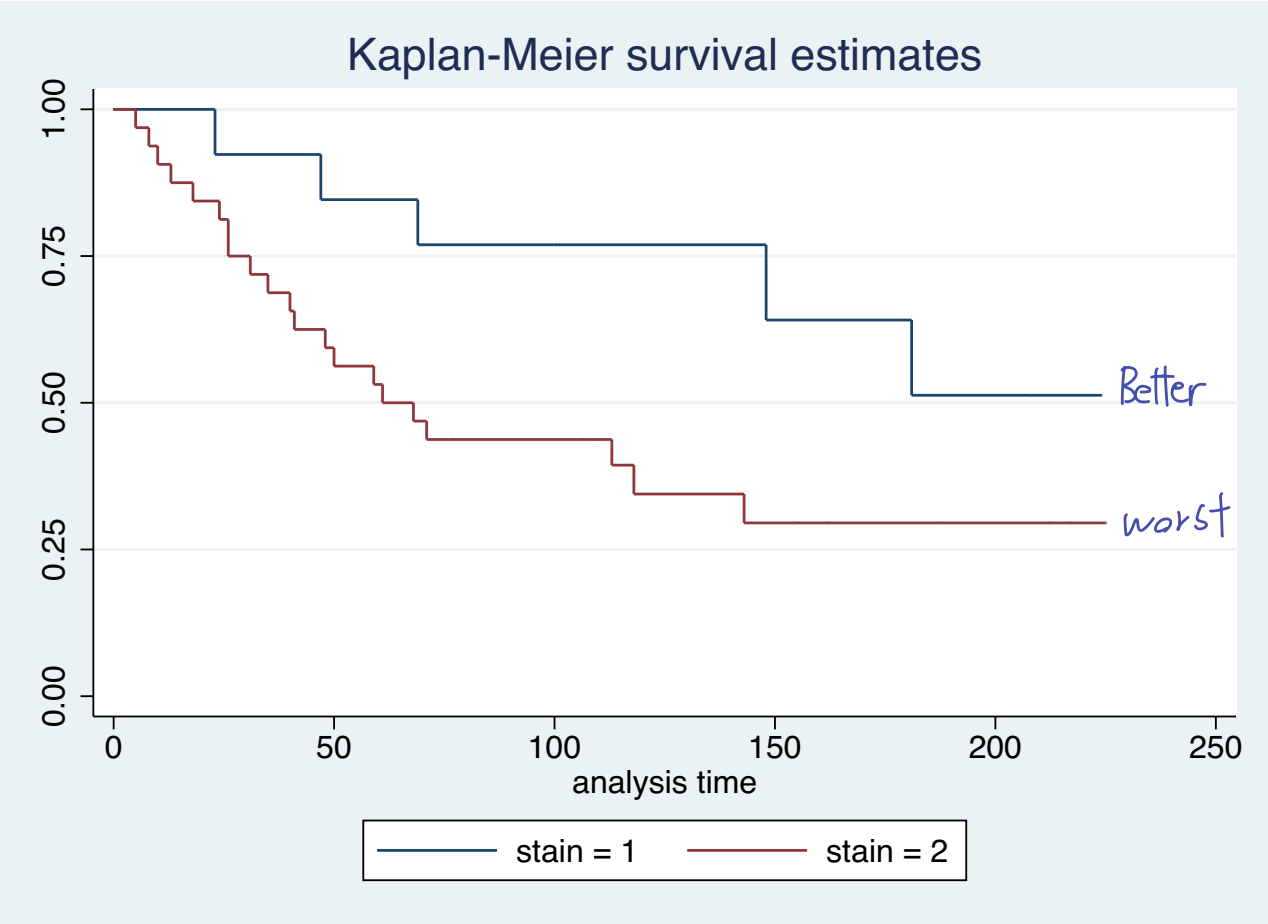
Table 1: Survival times (in months) of women with tumors that were negatively or positively stained for *Helix promatia* HPA. Censored times are labeled with an asterisk. From Collett book (Leatham & Brooks Lancet 1987)

Negative staining	Positive staining	
23	5	68
47	8	71
69	10	76*
70*	13	105*
71*	18	107*
100*	24	109*
101*	26	113
148	26	116*
181	31	118
198*	35	143
208*	40	154*
212*	41	162*
224*	48	188*
	50	212*
	59	217*
	61	225*

Example: *Prognosis for women with breast cancer* (continued)

- stain codes: 1 = negative staining, 2= positive staining Note that the variable is not 0-1 coded.

```
. use prognosis_breast_cancer.dta
. stset time status
......
. sts graph, by (stain)
```



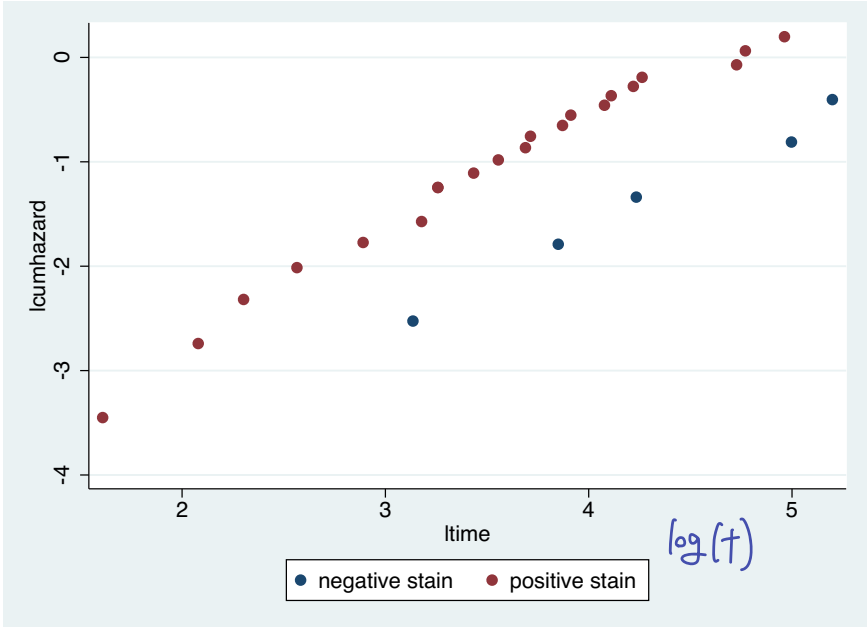
Example: *Prognosis for women with breast cancer* (continued)

Next, estimate the survival function and the cumulative hazard using KM method. Plot log of cumulative hazard versus log of time.

H(t)=-log(S)

```
. sts generate survf = s, by(stain)
. generate cumhazard = -log(survf)
. generate lcumhazard = log(cumhazard)
. generate ltime = log(time)
. graph twoway scatter lcumhazard ltime if status==1 & stain==1    ||
    scatter lcumhazard ltime if status==1 & stain==2,
  legend(order(1 "negative stain" 2 "positive stain"))
```

Check 2 Assumption:
- Weibull
- Prop.



log(H)

log(t)

Meti
- Weibull ✓
- Prop. ✓

Prognosis for women with breast cancer: default form

To make baseline rate parameter "_cons" interpretable, use `i.stain` instead of `stain`

```
. streg i.stain, dist(weibull) nolog
        failure _d:  status
  analysis time _t:  time
Weibull PH regression

No. of subjects =           45                 Number of obs    =        45
No. of failures =           26
Time at risk    =         4331
                                               LR chi2(1)       =      4.14
Log likelihood  =   -60.883962                 Prob > chi2      =    0.0418

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     2.stain |   2.545372   1.271665     1.87   0.061     .9560751    6.776579
       _cons |   .0041365   .0037257    -6.09   0.000     .0007079    .0241707
-------------+----------------------------------------------------------------
       /ln_p |  -.0646417   .1673746    -0.39   0.699    -.3926898    .2634064
-------------+----------------------------------------------------------------
           p |   .9374033   .1568975                      .6752382    1.301355
         1/p |   1.066777   .1785513                      .7684296    1.480959
------------------------------------------------------------------------------
Note: _cons estimates baseline hazard.
```

- By default, STATA outputs the hazard ratios (e^{β}) as the coefficient; and _cons = λ, p = γ in our notation.
- Option "nohr" can provide coefficients on the log relative hazard scale, and gives estimates for $\beta, \log(\lambda), \gamma$

Prognosis for women with breast cancer (continued)

Interpretations of the parameter estimates:

- log hazard ratio $\hat{\beta} = \log(2.545372) = .93427681$, baseline rate parameter $\hat{\lambda} = 0.0041365$, and shape parameter $\hat{\gamma} = 0.9374033$
- The hazard function for the group with negative staining is estimated to be $\hat{h}_0(t) = \hat{\lambda}\hat{\gamma}t^{\hat{\gamma}-1}$, and the hazard function for the group with positive staining is estimated to be $\hat{h}_i(t) = e^{\hat{\beta}}\hat{\lambda}\hat{\gamma}t^{\hat{\gamma}-1}$
- Since $e^{\hat{\beta}} = 2.55$, a woman in the positive HLA group has about two and a half times the risk of death at any given time, compared to a woman whose tumor was HLA negative.
- $e^{\hat{\beta}} = 2.55 > 1$ and the 95% CI for e^{β} is $(0.96,\ 6.77)$ only just includes unity (borderline significant), suggests that women with positively stained tumors have a poorer prognosis than those whose tumors were negatively stained.

Prognosis for women with breast cancer (continued)

Quantiles

- Recall for individuals in the negative stain group, their survival time has a Weibull distribution $W(\lambda, \gamma)$, and the survival function is $S_0(t) = \exp(-\lambda t^\gamma)$. For individuals in the positive stain group, their survival time has a Weibull distribution $W(e^\beta \lambda, \gamma)$, and $S_i(t) = \exp(-e^\beta \lambda t^\gamma)$
- The median and other percentiles of the survival time distributions in the two groups can be estimated from the values of $\hat{\beta}, \hat{\lambda}$ and $\hat{\gamma}$.
  - The estimated p*th* percentile for those with negative staining is given by

$$\hat{t}_{\frac{p}{100}} = \left\{ \frac{1}{\hat{\lambda}} \log\left(\frac{100}{100 - p} \right) \right\}^{1/\hat{\gamma}} \tag{8}$$

  - The estimated p*th* percentile for those with positive staining is given by

$$\hat{t}_{\frac{p}{100}} = \left\{ \frac{1}{e^{\hat{\beta}} \hat{\lambda}} \log\left(\frac{100}{100 - p} \right) \right\}^{1/\hat{\gamma}} \tag{9}$$

The Weibull proportional hazards model

But wait . . . if γ not different from 1.0, ($\hat{\gamma} = 0.9374033$) then we have an exponential survival model

- Simpler model may be preferred if it can be fit
- Exponential model is *nested* within Weibull - can use LR test in addition to CI on output.

Fit the Exponential Survival Model for Prognosis of women with breast cancer

```
. streg i.stain, dist(exponential) nolog

        failure _d:  status
  analysis time _t:  time

Exponential PH regression

No. of subjects =           45                 Number of obs    =        45
No. of failures =           26
Time at risk    =         4331
                                               LR chi2(1)       =      4.36
Log likelihood  =   -60.960708                 Prob > chi2      =    0.0369

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     2.stain |   2.589922   1.28878     1.91   0.056     .9766015    6.868405
       _cons |   .0030266   .0013536   -12.97   0.000     .0012598    .0072716
------------------------------------------------------------------------------
Note: _cons estimates baseline hazard.
```

$LR = -2(-60.960708 + 60.883962) = 0.153492$. The LR test statistics is far away from significance. The two model has a similar fit and the exponential survival model is preferred.

Fit Exponential Model for Breast Cancer Data

- Note that $\exp(\beta)$ for stain effect ($HR = 2.59$) is not much different from Weibull ($HR = 2.55$)
- The scale parameter (rate of failure, here 0.0030266) is not much different from the scale parameter in the Weibull model. What is this number and how does it relate to staining?

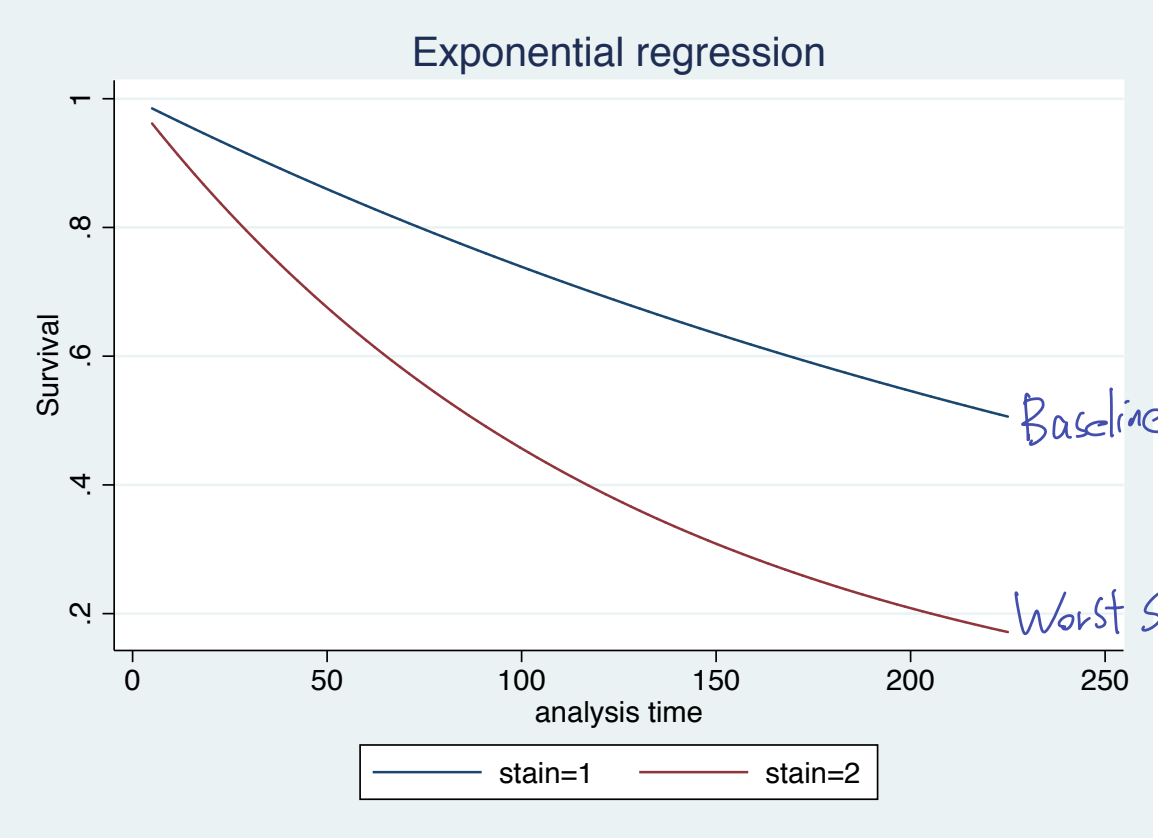
```
. stsum, by(stain)

         failure _d:  status
   analysis time _t:  time
                |                  Incidence      Number of   |------ Survival time -----|
      stain     | Time at risk          rate       subjects         25%         50%         75%
----------------+-------------------------------------------------------------------------
          1     |        1,652     .0030266             13         148           .           .
          2     |        2,679     .0078387             32          26          61           .
----------------+-------------------------------------------------------------------------
      Total     |        4,331     .0060032             45          40         113

.* Note _cons term in model is failure rate at stain = 1
. display .0030266*2.59
.00783889
```

Answer: $\hat{\lambda}$ from the exponential survival model is the baseline incidence rates and $e^{\hat{\beta}}$ is the hazard ratio.

Fit Exponential Model for Breast Cancer Data

Plotting the curves

```
. stcurve, survival at1(stain=1) at2(stain=2)
```



The Weibull Proportional Hazards Model

- More generally, when there are p explanatory variables $X_1, ..., X_p$, under the proportional hazards model, the hazard of death at time t for the i^{th} individual is

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}) h_0(t) \tag{10}$$

- If baseline hazard function $h_0(t)$ is specified as a Weibull model with scale parameter λ and shape parameter γ, i.e., $h_0(t) = \lambda\gamma t^{\gamma-1}$, the hazard function for the ith individual in the study is then given by

$$h_i(t) = \exp(\boldsymbol{\beta}^T \boldsymbol{x_i}) \lambda\gamma t^{\gamma-1} \tag{11}$$

- From the form of (11), the survival time of the ith individual in the study has a Weibull distribution $W(\exp(\boldsymbol{\beta}^T \boldsymbol{x_i})\lambda, \gamma)$
- The survivor function corresponding to the hazard function in (11) is

$$S_i(t) = \exp\{-\exp(\boldsymbol{\beta}^T \boldsymbol{x_i}) \lambda t^{\gamma}\} \tag{12}$$

Example: Weibull PH model with multiple covariates

Treatment of hypernephroma

In a study carried out at the University of Oklahoma Health Sciences Center, data were obtained on the survival times of 36 patients with a maglignant tumor in the kidney, or hypernephroma.

Of particular interest is whether the survival time of the patients depends on their age at the time of diagnosis and on whether or not they had received a nephrectomy, or surgical removal of the kidney.

Treatment of hypernephroma

```
. use treatment_of_hypernephroma.dta
. list in 1/10
```

	nephre~y	age	time	status
1.	0	1	9	1
2.	0	1	6	1
3.	0	1	21	1
4.	0	2	15	1
5.	0	2	8	1
6.	0	2	17	1
7.	0	3	12	1
8.	1	1	104	0
9.	1	1	9	1
10.	1	1	56	1

- time is recorded in months.
- nephrectomy codes: 0 = no nephrectomy, 1 = nephrectomy
- age codes: 1 for age<60, 2 for age 60-70, 3 for age >70

Treatment of hypernephroma

- Let Age_2 be the indicator for age being in the range 60-70, and Age_3 be the indicator for age being in the range >70. Then a Weibull proportional hazard model could be specified to be

$$h(t) = \exp(\beta_1 \cdot nephrectomy + \beta_2 \cdot Age\_2 + \beta_3 \cdot Age\_3) h_0(t) \qquad (13)$$

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

- $h_0(t)$ is the hazard for a subject that didn't receive a nephrectomy and of an age <60 at the time of diagnosis.
- $\exp(\beta_1)$ is the hazard ratio comparing receiving a nephrectomy to receiving no nephrectomy adjusting for age. (Equivalently, β_1 is the increase in log hazard for 1 unit increase in nephrectomy (1 vs. 0) adjusting for age.)
- $\exp(\beta_2)$ is the hazard ratio comparing age group 60-70 to age group <60 adjusting for nephrectomy status.

Treatment of hypernephroma

```
. stset time status
. streg nephrectomy i.age, dist(weibull) nolog
        failure _d:  status
  analysis time _t:  time
Weibull PH regression
No. of subjects =           36                 Number of obs    =         36
No. of failures =           32
Time at risk    =         1340
                                               LR chi2(3)       =      17.13
Log likelihood  =    -43.87881                 Prob > chi2      =     0.0007
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 nephrectomy |   .1919802   .1006891    -3.15   0.002     .0686785    .5366509
             |
         age |
          2  |   1.085593   .4363889     0.20   0.838     .4937408    2.386903
          3  |   5.218136   3.088109     2.79   0.005     1.635956    16.64406
             |
       _cons |   .0170522   .0131496    -5.28   0.000     .0037617    .0772992
-------------+----------------------------------------------------------------
        /ln_p |   .3438972   .1411602     2.44   0.015     .0672284     .620566
-------------+----------------------------------------------------------------
           p |   1.410434    .199097                      1.06954    1.859981
         1/p |   .7090018   .1000828                       .53764    .9349817
------------------------------------------------------------------------------
Note: _cons estimates baseline hazard.
```

"omit status"

dist(weibull)

λ _cons

√ p

Keep weibull

CI outside of 1, no need check exp.

Treatment of hypernephroma

- Interpretation of coefficients:
  - $\exp(\hat{\beta}_1) = .1919802$, with a 95% CI for $\exp(\beta_1)$ being $(0.0686785, 0.5366509)$. The CI does not contain 1 and the upper bound is smaller than 1, meaning that the nephrectomy substantially reduces the hazard of death at any given time controlling for age.
  - $\exp(\hat{\beta}_2) = 1.085593$, with a 95% CI for $\exp(\beta_2)$ being $(.4937408, 2.386903)$ and containing 1. There is no evidence that hazard of death is different for age group 60-70 compared to age group < 60 controlling for nephrectomy status.
  - $\exp(\hat{\beta}_3) = 5.218136$, with a 95% CI for $\exp(\beta_3)$ being $(1.635956, 16.64406)$. There is strong evidence that mortality hazard at any given time is higher for age group >70 compared to age group 60 controlling for nephrectomy status.
  - $\hat{\lambda} = .0170522$, $\hat{\gamma} = 1.410434$, thus the hazard for a subject that didn't receive a nephrectomy and of an age <60 at the time of diagnosis at time t is estimated to be $\hat{\lambda}\hat{\gamma}t^{\hat{\gamma}-1}$.

Comparing alternative Weibull models

- We can still use the likelihood ratio test (difference in deviance) to compare two nested models.
- Suppose the smaller model contains p explanatory variables, and the larger model contains k *extra* explanatory variables; the maximized likelihoods under smaller model and larger model are $\hat{L}_s$ and $\hat{L}_l$, respectively.
- Under H_0 the smaller model is correct, $-2(\log\hat{L}_s - \log\hat{L}_l)$ has an approximate χ^2_k distribution.

Treatment of hypernephroma: with interaction terms

In Stata, use "`xi:`" function to build an interaction model

```
. xi: streg i.age*nephrectomy, dist(weibull) nolog
i.age             _lage_1-3           (naturally coded; _lage_1 omitted)
i.age*nephrec~y   _lageXnephr_#       (coded as above)

        failure _d:  status
  analysis time _t:  time

Weibull regression -- log relative-hazard form

No. of subjects =           36                 Number of obs   =        36
No. of failures =           32
Time at risk    =         1340
                                               LR chi2(5)      =     21.82
Log likelihood  =   -41.532133                 Prob > chi2     =    0.0006

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      _lage_2 |   .9183851   .7500492    -0.10   0.917     .1852869    4.552028
      _lage_3 |   1.121983   1.297394     0.10   0.921     .1163344    10.82093
  nephrectomy |   .0875388   .0632624    -3.37   0.001     .0212351    .3608657
_lageXnephr_2 |   1.128947   1.061025     0.13   0.897     .1789303    7.123004
_lageXnephr_3 |   12.65381   16.80464     1.91   0.056     .9371311    170.8609
        _cons |   .0187577   .0159487    -4.68   0.000     .0035436    .0992918
-------------+----------------------------------------------------------------
        /ln_p |   .4407051   .1457169     3.02   0.002     .1551052    .7263049
-------------+----------------------------------------------------------------
            p |   1.553802   .2264153                      1.167781    2.067427
          1/p |   .6435825   .0937809                       .483693    .8563251
------------------------------------------------------------------------------
```

Treatment of hypernephroma

- The model is $h(t) = \exp(\beta_1 \text{neph} + \beta_2 \text{age2} + \beta_3 \text{age3}$
  interaction interaction
  $+\beta_4 \text{neph} \times \text{age2} + \beta_5 \text{neph} \times \text{age3}) \cdot h_0(t)$,

  where baseline hazard for patients with age < 60 and no nephrectomy is $h_0(t) = \lambda\gamma t^{\gamma-1}$.
- Comparing the two nested models: without interaction between age and nephrectomy vs. with interaction between age and nephrectomy:

$$-2(\log\hat{L}_s - \log\hat{L}_l) = -2*(-43.87881-(-41.532133)) = 4.7 \sim \chi^2_2 \text{ under } H_0$$

The p-value is not significant, therefore we choose simpler model (no interaction).

```
. di chi2tail(2, 4.693354)
.0956866
```

Treatment of hypernephroma

- Based on the interaction model, the hazard ratio comparing a patient with nephrectomy and aged 70+ versus another patient without nephrectory in 60-70 is: *Step by Step*

  The hazard function for a patient A with nephrectomy and aged 70+ (`age3=1`) is $h(t)_A = \exp(\hat{\beta}_1 + \hat{\beta}_3 + \hat{\beta}_5) \cdot h_0(t)$.

  The hazard function for a patient B without nephrectomy and aged 60-70 (`age2=1`) is $h(t)_B = \exp(\hat{\beta}_2) \cdot h_0(t)$.

- The Hazard Ratio is

$$\exp(\hat{\beta}_1 + \hat{\beta}_3 + \hat{\beta}_5 - \hat{\beta}_2).$$

- Note that the default Stata output provides $\exp(\hat{\beta})$'s. In this example, HR=$\exp(\hat{\beta}_1) \times \exp(\hat{\beta}_3) \times \exp(\hat{\beta}_5) / \exp(\hat{\beta}_2)$
  $= .0875388 \times 1.121983 \times 12.65381 / .9183851 = 1.35$ ✓

A non-PH model: The Accelerated Failure Time (AFT) Regression Model

- The **accelerated failure time model** is a parametric model form that provides an alternative to the commonly used proportional hazards model.
- Unlike a proportional hazards model which assumes that the effect of a covariate is to multiply the hazard by some constant, an accelerated failure time model assumes that the effect of a covariate is to accelerate or decelerate the time to event by some constant.
- Compared with proportional hazards model, the accelerated failure time model may provide a more straightforward way to think about covariates in relation to survival time, although it is less used, and fewer are familiar with it.

AFT model for two group comparison

- Suppose that we want to compare two treatments, a standard treatment, S, or a new treatment, N.
- Under the **accelerated failure time model**, the survival time (t) of an individual on the new treatment is taken to be a multiple of the survival time for an individual on the standard treatment. Thus the effect of the new treatment is to "speed up" or "slow down" the passage of time.
- Under this assumption, the probability that an individual on the new treatment survives beyond time t will be equal to the probability that an individual on the standard treatment survives beyond time t/ϕ

AFT model for two group comparison

- Let $S_S(t)$ and $S_N(t)$ be the survivor functions for individuals in the two treatment groups. Then, the accelerated failure time model specifies that

$$S_N(t) = S_S(t/\phi), \tag{14}$$

  where the quantity $1/\phi$ is often called the *acceleration factor*.
- Interpretation:
  - Under AFT model, $S_N[t_N(50)] = S_s[t_N(50)/\phi] = 0.5 = S_S[t_S(50)]$. The median survival time for the new treatment is ϕ times the median survival time for the standard treatment, $t_N(50) = \phi t_S(50)$.
  - The pth percentile of survival time for the new treatment is ϕ times the pth percentile of survival time for the standard treatment, $t_N(p) = \phi t_S(p)$.
  - $\phi < 1$ corresponds to an acceleration in the time to event for the new treatment, relative to an individual on the standard treatment. The standard treatment has longer survival time.
  - $\phi > 1$ corresponds to an deceleration in the time to death of an individual assigned to the new treatment, relative to an individual on the standard treatment.

Accelerated failure time model for the comparison of two groups (continued)

- Let X be an indicator variable that takes the value 0 for an individual in the standard treatment group, and 1 for an individual in the new treatment group. Let $S_0(t)$ be the baseline survivor function (survivor function for subjects with $x = 0$), under the accelerated failure time model,

$$S_i(t) = S_0(t/\phi) \tag{15}$$

- Since ϕ must be positive, it is convenient to set $\phi = e^{\beta}$.
- Formally, the accelerated failure time model is given by

$$S_i(t) = S_0(t/e^{\beta x_i}) \tag{16}$$

and also it can be written as

$$h_i(t) = e^{-\beta x_i} h_0(t/e^{\beta x_i}) \tag{17}$$

Assessing the validity of an accelerated failure time model

- Consider the accelerated failure time model for individuals in the two treatment groups, as we showed before, the model says that the relationship between any pth percentiles of survival times of patients on the new and standard treatment is given by

$$t^N_{\frac{p}{100}} = e^{\beta} t^S_{\frac{p}{100}} \tag{18}$$

- This means that a plot of the quantity $\hat{t}^N_{\frac{p}{100}}$ against $\hat{t}^S_{\frac{p}{100}}$, for suitably chosen values of p, should given a straight line through the origin if the accelerated failure time model is appropriate. The slope of this line will be an estimate of e^{β}.
- This type of plot is called *percentile-percentile plot*, also known as the *quantile - quantile plot* or the ***Q-Q plot***.

Example: *Prognosis for women with breast cancer* (continued)

- Let X be the indicator for staining status, 1 for positive staining, and 0 for negative staining.
- Let $S_0(t)$ be the baseline survivor function, in this example, it's the survivor function for negative staining group.
- Under the accelerated failure time model,

$$S_i(t) = S_0(te^{-\beta x_i}) \tag{19}$$

Prognosis for women with breast cancer

```
. use prognosis_breast_cancer.dta

. stset time status

. sts list, by(stain)

        failure _d:  status
  analysis time _t:  time
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf.	Int.]
stain=1							
23	13	1	0	0.9231	0.0739	0.5664	0.9888
47	12	1	0	0.8462	0.1001	0.5122	0.9591
.....							
212	2	0	1	0.5128	0.1673	0.1756	0.7738
224	1	0	1	0.5128	0.1673	0.1756	0.7738
stain=2							
5	32	1	0	0.9688	0.0308	0.7982	0.9955
8	31	1	0	0.9375	0.0428	0.7725	0.9840
.....							
217	2	0	1	0.2953	0.0900	0.1366	0.4736
225	1	0	1	0.2953	0.0900	0.1366	0.4736

```
. * recode for modeling later
. gen stain1 = stain-1
(45 real changes made)
```

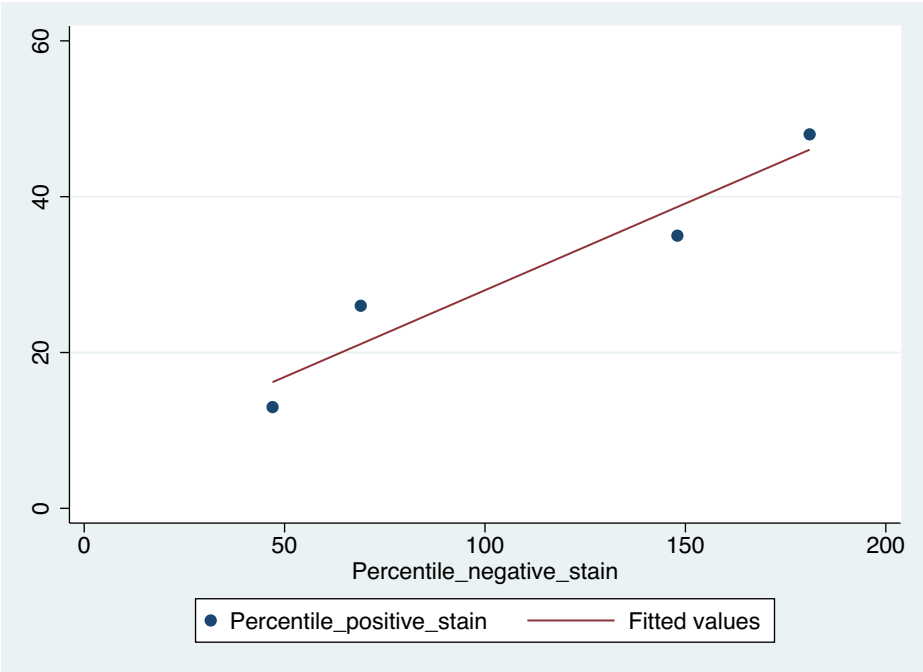
Prognosis for women with breast cancer

Table 2: Estimated percentiles of the distributions of survival times for women with tumors that were positively or negatively stained

Percentile	Negative staining	Positive staining
10	47	13
20	69	26
30	148	35
40	181	48
50	-	61
60	-	113
70	-	143
80	-	-

Prognosis for women with breast cancer

```
. clear

. input Percentile_negative_stain Percentile_positive_stain
47 13
69 26
148 35
181 48
end

. graph twoway (scatter Percentile_positive_stain Percentile_negative_stain)
(lfit Percentile_positive_stain Percentile_negative_stain)
          , xlabel(0(50)200) ylabel(0(20)60)
```



Prognosis for women with breast cancer: the percentile-percentile plot

- The points fall on a reasonably straight line roughly through the origin, suggesting that the accelerated failure time model would be appropriate. However, it should also be noted that there are only a limited number of points in the graph.
- The slope of the line (a rough estimate for e^{β} is smaller than 1, suggesting that for women whose tumors were positively stained, the disease process is **speeded up** relative to those whose tumors were negatively stained. In other words, the survival time for those HLA+ tumors is shorter compared to the survival time for women with HLA- tumors.

The Weibull accelerated failure time model

- We have examined the AFT assumption and now let us further assume that the survival times in the negative staining group have a Weibull distribution $W(\lambda, \gamma)$, so that

$$S_0(t) = \exp(-\lambda t^\gamma) \tag{20}$$

- Under the accelerated failure time model,

$$S_i(t) = S_0(t/e^{\beta x_i}) = \exp(-\lambda(e^{-\beta x_i})^\gamma t^\gamma), \tag{21}$$

  which is a survivor function for a Weibull distribution $W(\lambda(e^{-\beta x_i})^\gamma, \gamma)$. The Weibull distribution has the accelerated failure time property.
- Weibull distribution is the only distribution that has both proportional hazard property and accelerated failure time property. A Weibull AFT model can be written as

$$S_i(t) = S_0(t/\exp(\beta \cdot x_i)),$$

  where $S_0(t) = \exp(-\lambda t^\gamma)$.

Prognosis for women with breast cancer

```
. streg stain1, distribution(weibull) nolog time

        failure _d:  status
  analysis time _t:  time

Weibull AFT regression

No. of subjects =           45                 Number of obs    =        45
No. of failures =           26
Time at risk    =         4331
                                               LR chi2(1)       =      4.14
Log likelihood  =   -60.883962                 Prob > chi2      =    0.0418

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      stain1 |  -.9966647   .5440936    -1.83   0.067    -2.063069    .0697391
       _cons |   5.854364   .4988778    11.74   0.000     4.876581    6.832146
-------------+----------------------------------------------------------------
      /ln_p |  -.0646417   .1673746    -0.39   0.699    -.3926898    .2634064
-------------+----------------------------------------------------------------
          p |   .9374033   .1568975                      .6752382    1.301355
        1/p |   1.066777   .1785513                      .7684296    1.480959
------------------------------------------------------------------------------
```

- p gives estimate for γ
- *stain* gives estimate for β
- *_cons* gives estimate for $-\log(\lambda)/\gamma$

Prognosis for women with breast cancer

- $\hat{\beta} = -0.9966647$, $\hat{\gamma} = 0.9374033$, and $\hat{\lambda} = \exp(-\hat{\gamma} \cdot 5.854364) = 0.0041365$
- The estimated $S(t)$ for the negative staining group (baseline):

$$\hat{S}_0(t) = \exp(-\hat{\lambda} t^{\hat{\gamma}}) \tag{22}$$

- $\hat{S}_i(t) = \hat{S}_0(t/\exp(\hat{\beta} x_i))$, thus the estimated survivor function for positive staining group ($x_i = 1$) is given by

$$\hat{S}_{pos}(t) = \hat{S}_0(t/e^{-.9966647}) = \hat{S}_0(t/0.3691085) \tag{23}$$

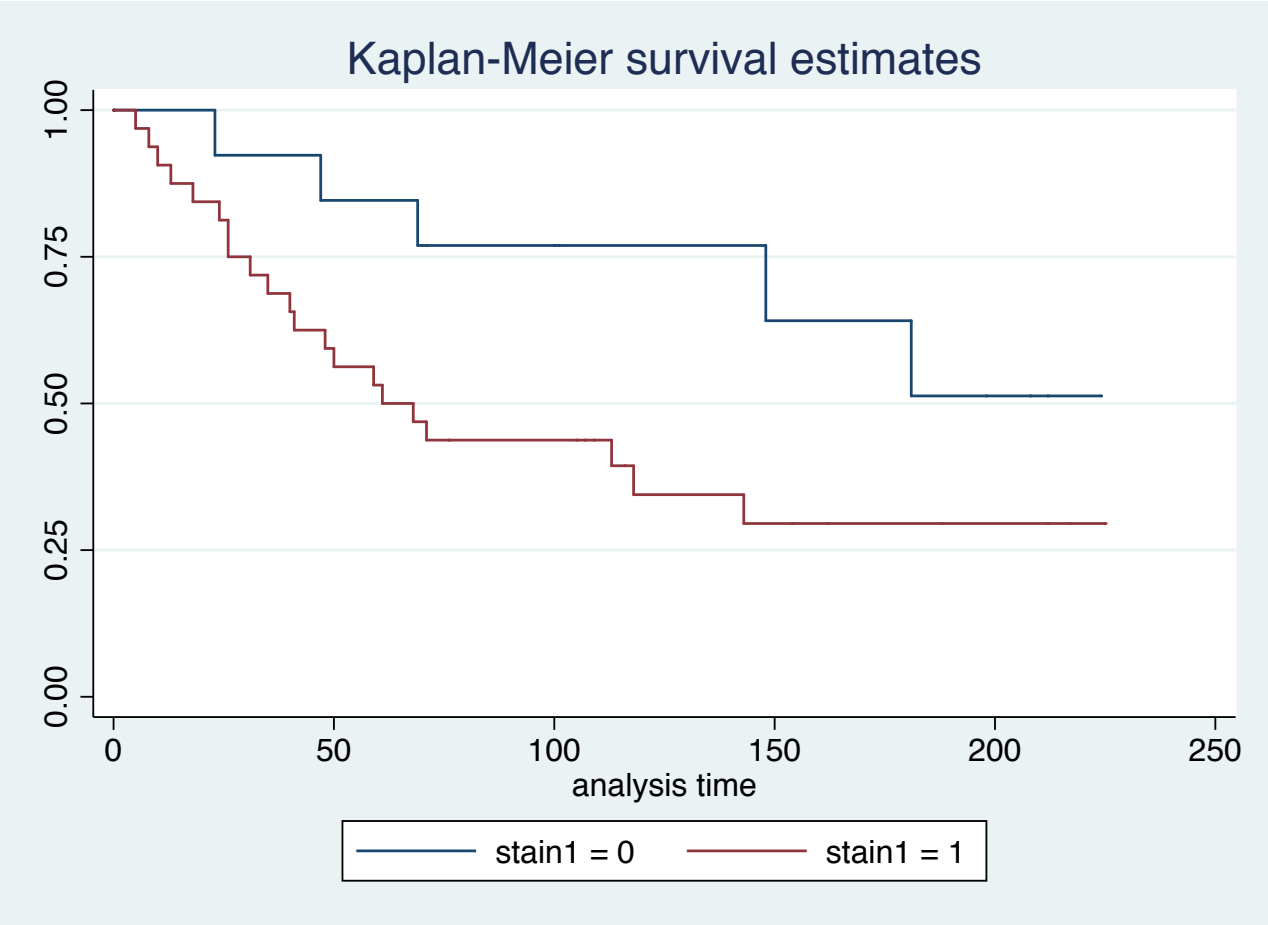
- The time ratio is $e^{\hat{\beta}} = 0.37$, with a 95% CI $\left(e^{-2.063069},\ e^{0.0697391}\right) = (0.13, 1.07)$. $\hat{t}^{pos}_{p/100} = 0.37 \hat{t}^{neg}_{p/100}$. Comparing a positive staining subject to a negative staining subject (baseline), the median (or any other percentile) survival time for a positive staining subject will be about 37% of that for a negative staining subject. Positive stain has worse survival but the reduction in survival time is not significant.

Prognosis for women with breast cancer

- We can also output directly in 'time ratio' scale in Stata, using the option `tr`.

```
. streg stain1, distribution(weibull) nolog time tr

        failure _d:  status
  analysis time _t:  time

Weibull regression -- accelerated failure-time form

No. of subjects =           45                 Number of obs    =        45
No. of failures =           26
Time at risk    =         4331
                                               LR chi2(1)       =      4.14
Log likelihood  =   -60.883962                 Prob > chi2      =    0.0418
------------------------------------------------------------------------------
          _t | Time Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      stain1 |   .3691085   .2008296    -1.83   0.067     .1270635    1.072228
       _cons |    348.753   173.9851    11.74   0.000     131.1814    927.1788
-------------+----------------------------------------------------------------
       /ln_p |  -.0646417   .1673746    -0.39   0.699    -.3926898    .2634064
-------------+----------------------------------------------------------------
           p |   .9374033   .1568975                      .6752382    1.301355
         1/p |   1.066777   .1785513                      .7684296    1.480959
------------------------------------------------------------------------------
```

- Are relevant times at different percentiles t of HLA positive survival curve shifted down by .37 times t from HLA negative survival curve ? Check this on KM plot



The General Accelerated Failure Time Model

- The accelerated failure time model for comparison between two groups can be generalized to the situation where there are p explanatory variables $X_1, ..., X_p$,

$$S_i(t) = S_0(t/\exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi})) \tag{24}$$

  where $S_0(t)$ is the baseline survivor function (for an individual for whom $\boldsymbol{x_i} = 0$)
- Different models for $S_0(t)$ will give different parametric accelerated failure time models. Parametric models commonly used are the Weibull, log-logistic, and lognormal model
- Positive β for an explanatory variable X means that increasing X values are associated with longer survival time controlling for the other variables in the model; negative β (or $e^{\beta} < 1$) means that increasing X value are associated with shorter survival time controlling for the other variables in the model.

Summary

Parametric Survival Regression Models

- The proportional hazards model assumes that the effect of a covariate is to multiply the hazard by some constant; and an AFT model assumes that the effect of a covariate is to accelerate or decelerate the life course of a disease by some constant.
- These models provide a way to incorporate covariates, as well as efficiently estimate survival quantities, provided that the model fit is adequate
- Next: Flexible semi-parametric model (happens to be PH form, but flexible extensions available)