Kaplan-Meier estimate of survival function

- Non-parametric
- Not well-suited to answering questions about effects of covariates
- Particular issue with covariates which are continuous, e.g. age

Cox regression model (a.k.a. "proportional hazards")

- Ideal material for exam questions!
- D. R. Cox **1972** *Regression models and life tables*
- Basic idea is that hazard function is proportional to a base hazard
- Explanatory variables assumed to act *multiplicatively*…
- …or *additively* on a logarithmic scale

Single-covariate case

- Pick case to act as *baseline* with baseline hazard function, $\lambda_0(t)$
- The hazard function of the i^{th} observation is a proportion, A_i, of the baseline hazard, i.e. $\lambda_i(t) = A_i\lambda_0(t), A_i > 0$
- Note that this proportion is more usually written in exponential form related to the covariate being studied, i.e. $A_i = e^{\beta_i z_i}$, where z_i is the regression covariate (say age) and β_i is the regression parameter. This gives:

$$\lambda_i(t) = \lambda_0(t)e^{\beta_i z_i}$$

- Under Cox model, hazards are in constant proportion at all times, so:

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{e^{\beta_i z_i}}{e^{\beta_j z_j}}$$

- General shape of hazard function set by the baseline, $\lambda_0(t)$…
- …while exponential term accounts for differences between observations (lives)
- The baseline hazard, $\lambda_0(t)$, is a *nuisance parameter*, but is elimated by the proportional-hazards model

Multi-covariate case

- In general there might be p different pieces of information per observation, say age, gender, drug treatment type, etc
- Data for each observation is a vector, $z_i = (z_{i1}, z_{i2}, \ldots, z_{ip})'$
- …and regression parameter is also a vector, $\beta = (\beta_1, \beta_2, \ldots, \beta_p)'$
- In general the hazard becomes:

$$\begin{aligned}\lambda(t; z_i) &= \lambda_0(t) \exp(\beta_1 z_{i1} + \beta_2 z_{i2} + \ldots + \beta_p z_{ip}) \\ &= \lambda_0(t) \exp(\beta' z_i)\end{aligned}$$

- …and the proportional-hazards model gives:

$$\begin{aligned}\frac{\lambda(t; z_i)}{\lambda(t; z_j)} &= \frac{\exp(\beta_1 z_{i1} + \beta_2 z_{i2} + \ldots + \beta_p z_{ip})}{\exp(\beta_1 z_{i1} + \beta_2 z_{i2} + \ldots + \beta_p z_{ip})} \\ &= \frac{\exp(\beta' z_i)}{\exp(\beta' z_j)}\end{aligned}$$

Partial likelihood function

- n lives
- r ordered distinct event (death) times $t_{(1)}, t_{(2)}, \ldots, t_{(r)}$, i.e. *no ties*
- $R(t_{(j)})$ is the *risk set* just prior to time $t_{(j)}$, i.e. the set of individuals at risk of the event (death) and uncensored at time $t_{(j)}$
- $R(t_{(j)})$ is directly analogous to the n_j in the Kaplan-Meier approach
- In Kaplan-Meier, individuals are implicitly identical, hence simple count of n_j
- In Cox model, each individual i has a vector of different characteristics, z_i, hence $R(t_{(j)})$ instead of n_j as individuals are no longer assumed identical
- Partial likelihood, $L(\beta)$, is defined as:

$$L(\beta) = \prod_{j=1}^{r} \frac{\exp(\beta' z_{(j)})}{\displaystyle\sum_{l \in R(t_{(j)})} \exp(\beta' z_l)}$$

Derivation of partial likelihood function

- Since we assume no ties in event (death) times, there is exactly one event (death) at each $t_{(j)}$

- Likelihood is then product of conditional probabilities at the event (death) times

- First consider one of the conditional probabilities

- Recall $\Pr(A|B) = \dfrac{\Pr(A)}{\Pr(B)}, \quad A \subset B$

- Event A is the probability that the life with $z_{(t_j)}$ dies in the interval $[t, t+\delta)$, i.e. $\Pr(A) = {}_{\delta}q_{z_{(j)},t}$

- Event B is the probability that exactly one life in the risk set $R(t_{(j)})$ dies in the interval $[t, t+\delta)$, and that all others survive

- The probability that life l dies is ${}_{\delta}q_{z_l,t}$ and the probability all others survive is $\prod\limits_{k \neq l} (1 - {}_{\delta}q_{z_k,t})$

- Thus, $\Pr(B)$, the probability that exactly one life dies is the sum over all possible members $l \in R(t_{(j)})$, i.e.

$$\Pr(B) = \sum_{l \in R(t_{(j)})} {}_{\delta}q_{z_l,t} \prod_{k \neq l} (1 - {}_{\delta}q_{k,t})$$

- So, the conditional probability for a single event is:

$$\begin{aligned} \Pr(A|B) &= \frac{\Pr(A)}{\Pr(B)} \\ &= \frac{{}_{\delta}q_{z_{(j)},t}}{\sum\limits_{l \in R(t_{(j)})} {}_{\delta}q_{z_l,t} \prod\limits_{k \neq l} (1 - {}_{\delta}q_{k,t})} \end{aligned}$$

$$
\begin{aligned}
\Pr(A|B) &= \frac{\Pr(A)}{\Pr(B)} \\
&= \frac{\delta q_{z_{(j)},t}}{\sum\limits_{l \in R(t_{(j)})} \delta q_{z_l,t} \prod\limits_{k \neq l} (1 - \delta q_{k,t})} \\
&= \frac{\delta\lambda(t; z_{(j)}) + o(\delta)}{\sum\limits_{l \in R(t_{(j)})} (\delta\lambda(t; z_l) + o(\delta)) \prod\limits_{k \neq l} (1 - \delta\lambda(t; z_k) + o(\delta))} \\
&= \frac{\delta\lambda(t; z_{(j)}) + o(\delta)}{\sum\limits_{l \in R(t_{(j)})} \delta\lambda(t; z_l) + o(\delta)}, \qquad \text{where} \lim_{\delta \to 0^+} \frac{o(\delta)}{\delta} \to 0 \\
&= \frac{\lambda(t; z_{(j)}) + \frac{o(\delta)}{\delta}}{\sum\limits_{l \in R(t_{(j)})} \lambda(t; z_l) + \frac{o(\delta)}{\delta}} \\
&= \frac{\lambda(t; z_{(j)})}{\sum\limits_{l \in R(t_{(j)})} \lambda(t; z_l)}, \qquad \text{as} \quad \delta \to 0^+ \\
&= \frac{\lambda_0 \exp(\beta' z_{(j)})}{\sum\limits_{l \in R(t_{(j)})} \lambda_0 \exp(\beta' z_l)} \\
&= \frac{\exp(\beta' z_{(j)})}{\sum\limits_{l \in R(t_{(j)})} \exp(\beta' z_l)}
\end{aligned}
$$

- The above is the result for a single conditional probability, so the partial likelihood is the product of all of them:

$$
L(\beta) = \prod_{j=1}^{r} \frac{\exp(\beta' z_{(j)})}{\sum\limits_{l \in R(t_{(j)})} \exp(\beta' z_l)}
$$

Recipe for creating Cox (partial) likelihood

- Strong parallels to recipe for Kaplan-Meier estimate of $\hat{S}(t)$
- Order the data by event (death) times
- Identify the r distinct event (death) times
- Decide what the baseline is going to be
- There will be r product terms in the partial likelihood function, one for each of the event (death) cases $z_{(j)}$

$$L(\beta) = \overbrace{— \times — \times \ldots \times —}^{r \text{ terms}}$$

- For each product term, the numerator is $\exp(\beta' z_{(j)})$

$$L(\beta) = \frac{\exp(\beta' z_{(1)})}{\quad} \times \frac{\exp(\beta' z_{(2)})}{\quad} \times \ldots \times \frac{\exp(\beta' z_{(r)})}{\quad}$$

- Note that for cases where $z_{(j)}$ matches the baseline, the numerator is simply $e^0 = 1$
- In example 2 on page 11 of the Red Book, there are five claim events so there are five product terms. Since all claims except the second and fifth are from males (the baseline), the numerators are:

$$L(\beta) = \frac{1}{—} \times \frac{e^\beta}{—} \times \frac{1}{—} \times \frac{1}{—} \times \frac{e^\beta}{—}$$

- At the first claim at time 2, there are five males and five females in the risk set:

$$L(\beta) = \frac{1}{5 + 5e^\beta} \times \frac{e^\beta}{—} \times \frac{1}{—} \times \frac{1}{—} \times \frac{e^\beta}{—}$$

where the first 5 is $5e^0$ as the five males are baseline.

- At the second claim at time 4, there are four males (5+, 9, 11+, 14) and five females (4, 7+, 10, 12+, 15) in the risk set:

$$L(\beta) = \frac{1}{5 + 5e^\beta} \times \frac{e^\beta}{4 + 5e^\beta} \times \frac{1}{—} \times \frac{1}{—} \times \frac{e^\beta}{—}$$

- At the third claim at time 9, there are three males (9, 11+, 14) and three females (10+, 12+, 15) in the risk set:

$$L(\beta) = \frac{1}{5+5e^{\beta}} \times \frac{e^{\beta}}{4+5e^{\beta}} \times \frac{1}{3+3e^{\beta}} \times \frac{1}{} \times \frac{e^{\beta}}{}$$

- At the fourth claim at time 14, there is only one male (14) and one female (15) in the risk set:

$$L(\beta) = \frac{1}{5+5e^{\beta}} \times \frac{e^{\beta}}{4+5e^{\beta}} \times \frac{1}{3+3e^{\beta}} \times \frac{1}{1+e^{\beta}} \times \frac{e^{\beta}}{}$$

- Finally, at the fifth and final claim at time 15, there is only one female (15) in the risk set:

$$L(\beta) = \frac{1}{5+5e^{\beta}} \times \frac{e^{\beta}}{4+5e^{\beta}} \times \frac{1}{3+3e^{\beta}} \times \frac{1}{1+e^{\beta}} \times \frac{e^{\beta}}{0+e^{\beta}}$$

- The partial likelihood in example 2 on page 11 thus reduces to:

$$L(\beta) = \frac{e^{\beta}}{(1+e^{\beta})(4+5e^{\beta})}$$

Using the partial likelihood

- Similar to a proper likelihood function
- Can take natural logarithms to get log partial likelihood, $\ell = \log L(\beta)$
- Solve $\dfrac{\partial \ell}{\partial \beta} = 0$ to find estimate of β, $\hat{\beta}$
- $\mathrm{E}(\hat{\beta}) = \beta_0$, where β_0 is the true unknown value of β

$$\begin{aligned}\mathrm{Var}(\hat{\beta}) &= -1/\mathrm{E}\left(\frac{\partial^2 \ell}{\partial \beta^2}\right)\Big]_{\beta=\beta_0} \\ &\approx -1/\left(\frac{\partial^2 \ell}{\partial \beta^2}\right)\Big]_{\beta=\hat{\beta}}\end{aligned}$$

- In example 2 on page 12 of the Red Book:

$$\frac{\partial \ell}{\partial \beta} = 1 - \frac{3e^{\beta}}{1+e^{\beta}} - \frac{5e^{\beta}}{4+5e^{\beta}}$$

and this is maximised at $\dfrac{\partial \ell}{\partial \beta} = 0$ for $\beta = -1.1567$. This is $\hat{\beta}$.

- $\mathrm{Var}(\hat{\beta}) \approx 1/0.7486$

$\Rightarrow \mathrm{StdErr}(\hat{\beta}) = 1.1558$

Hypothesis testing

- Ideal material for exam questions!
- Define the *score function*, $U(\beta) = \dfrac{\partial \ell}{\partial \beta}$
- Define the *observed information function*, $I(\beta) = -\dfrac{\partial^2 \ell}{\partial \beta^2}$
- z test: compare either $z = \dfrac{\hat{\beta}}{\sqrt{\mathrm{Var}(\hat{\beta})}}$ with $N(0,1)$ or z^2 with χ^2_1
- Score test: compare $S = \dfrac{U(\beta_0)^2}{I(\beta_0)}$ with χ^2_1 if β is scalar
- Likelihood ratio test: compare $-2\log\Lambda = -2(\ell_0 - \ell_1)$ with χ^2_1, where ℓ_i is the maximised log-likelihood value of ℓ under $H_i : i = 0, 1$

Fitting a proportional-hazards model in R