Cox PH Model

Inference

Introduction

- It is often of interest to consider hypothesis testing of the parameters in the Cox PH model.
- In Cox PH model
  $h(t|\boldsymbol{X}) = h_0(t)\exp(\beta_1 \mathrm{age} + \beta_2 \mathrm{Treatment})$
- For example,
  - $H_0: \beta_1$=0 v.s. $H_1: \beta_1 \neq$0 to test if age has a significant effect on survival.
  - $H_0: \beta_2$=0 v.s. $H_1: \beta_2 \neq$0 to test if there is significant difference in survival between treatment and placebo groups.

Partial likelihood

- The partial likelihood has similar properties as a regular likelihood function.
- Likelihood based testing methods can be defined similarly to make inference about $\boldsymbol{\beta}$.
  - Wald test
  - Score test
  - Likelihood ratio test

Properties of MPLE

- Let $\widehat{\boldsymbol{\beta}}$ be the Maximum Partial Likelihood Estimator.
- Let $U(\boldsymbol{\beta}) = \partial \log PL(\boldsymbol{\beta}) / \partial \boldsymbol{\beta}$ be the score vector of the partial likelihood.

  $\widehat{\boldsymbol{\beta}}$ satisfies $U(\widehat{\boldsymbol{\beta}}) = 0$
- Then

$$\widehat{\boldsymbol{\beta}} \sim N\left(\boldsymbol{\beta}, I_n^{-1}(\boldsymbol{\beta})\right) \quad asymptotically$$
$$U(\boldsymbol{\beta}) \sim N\left(\mathbf{0}, I_n(\boldsymbol{\beta})\right) \quad asymptotically$$

PMLE

Define the observed partial likelihood information matrix

$$\hat{\boldsymbol{I}}_n(\widehat{\boldsymbol{\beta}}) = -\frac{\partial^2 \log PL(\widehat{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T},$$

with its kh th element being

$$\hat{I}_{kh} = \hat{I}_{hk} = \sum_{i=1}^{D} \left\{ \sum_{j \in R(t_i)} (Z_{jk} - \bar{Z}_{(i)k})(Z_{jh} - \bar{Z}_{(i)h}) w_{ij} \right\}$$

Confidence intervals

- The variance of the PMLE $\widehat{\boldsymbol{\beta}}$ can be estimated by
$$Var(\widehat{\boldsymbol{\beta}}) \approx \hat{\boldsymbol{I}}_n^{-1}(\widehat{\boldsymbol{\beta}})$$
- In particular, the SE of $\hat{\beta}_k$ is the square root of the kth diagonal element of $\hat{\boldsymbol{I}}_n^{-1}(\widehat{\boldsymbol{\beta}})$.
- Wald $100(1-\alpha)\%$ CI of β_k is
$$\hat{\beta}_k \pm Z_{\alpha/2} SE(\hat{\beta}_k)$$
- The CI for $\exp(\beta_k)$, the hazard ratio associated with covariate Z_k, is given as
$$\left(e^{\widehat{\beta}_k - Z_{\alpha/2} SE(\widehat{\beta}_k)}, e^{\widehat{\beta}_k + Z_{\alpha/2} SE(\widehat{\beta}_k)}\right)$$

Hypothesis testing

- Consider a simple null hypothesis that
$$H0:\ \boldsymbol{\beta}=\boldsymbol{\beta}_0, H1:\ \boldsymbol{\beta}\neq\boldsymbol{\beta}_0.$$
- Let $\widehat{\boldsymbol{\beta}}$ be the MPLE.
- Three likelihood based testing methods can be defined as
  - The Wald's test
  $$Q_W=(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)^T\hat{\boldsymbol{I}}_n(\widehat{\boldsymbol{\beta}})(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)$$
  - Score test
  $$Q_S=\boldsymbol{U}^T(\boldsymbol{\beta}_0)\hat{\boldsymbol{I}}_n^{-1}(\boldsymbol{\beta}_0)U(\boldsymbol{\beta}_0)$$
  - Likelihood ratio test
  $$Q_{LR}=2\{\log PL(\widehat{\boldsymbol{\beta}})-\log PL(\boldsymbol{\beta}_0)\}$$
- Under null hypothesis, these test statistics all follows χ_p^2 distribution asymptotically. (p is the dimensionality of $\widehat{\boldsymbol{\beta}}$).

An example: Wald test

- For a single parameter, consider
$$H_0:\ \beta_k=0, vs\ H_1:\ \beta_k\neq 0.$$
- Then Wald test statistics is given as,
$$Q_W=\left(\frac{\widehat{\beta}_k}{SE(\widehat{\beta}_k)}\right)^2\sim\chi_1^2,\ \text{under } H_0.$$
- Reject null hypothesis, if
$$Q_W=\left(\frac{\widehat{\beta}_k}{SE(\widehat{\beta}_k)}\right)^2>\chi_{1,\alpha}^2.$$

An example: LR test

- LR test is commonly used to compare nested models.
- Multiple parameters:
$$H_0:\beta_{k+1}=\cdots=\beta_p=0$$
$$vs\ H_1:\ \beta_j\neq 0, \text{for some } j=k+1,\ldots,p.$$
- Consider full model
$$h(t\mid Z)=h_0(t)\exp(\beta_1 Z_1+\cdots+\beta_p Z_p)$$
- Reduced model
$$h(t\mid Z)=h_0(t)\exp(\beta_1 Z_1+\cdots+\beta_k Z_k)$$
- Partial likelihood ratio test statistic:
$$Q_{LR}=2\{\log PL_{full}-\log PL_{Reduced}\}$$
- Under null hypothesis, $Q_{LR}\sim\chi_{p-k}^2$ distribution asymptotically.
- df= # of parameters in full model − # of parameters in reduced model.
- Reject null hypothesis, if $Q_{LR}>\chi_{p-k,\alpha}^2$

An example: Score test

- Multiple parameters:
  $H_0: \beta_{k+1} = \cdots = \beta_p = 0$
  $vs\ H_1: \beta_j \neq 0$, for some $j = k+1, \ldots, p$.
- The reduced model
  $$h(t \mid Z) = h_0(t)\exp(\beta_1 Z_1 + \cdots + \beta_k Z_k)$$
  - Let $\widehat{\boldsymbol{\beta}}_0 = (\hat{\beta}_1, \ldots, \hat{\beta}_k, 0, \ldots, 0)^T$ be the MPLE under H_0.
- The Full model
  $$h(t \mid Z) = h_0(t)\exp(\beta_1 Z_1 + \cdots + \beta_p Z_p)$$
  - $\boldsymbol{U}(\boldsymbol{\beta})$ be score vector under the full model.
- Then the Score test statistics is defined as,
  $$Q_S = \boldsymbol{U}^T(\widehat{\boldsymbol{\beta}}_0)\hat{\boldsymbol{I}}_n^{-1}(\boldsymbol{\beta}_0)\boldsymbol{U}(\widehat{\boldsymbol{\beta}}_0)$$
- Under null hypothesis, $Q_S \sim \chi^2_{p-k}$ distribution asymptotically.
- Reject null hypothesis, if $Q_S > \chi^2_{p-k,\alpha}$

SAS Example

Wald test statistics

Analysis of Maximum Likelihood Estimates					
Parameter	Estimate	SE	Chi-Square	Pvalue	Hazard Ratio
Kps	-0.03300	0.00554	35.5051	<.0001	0.968
Duration	0.00323	0.00949	0.1159	0.7335	1.003
Age	-0.01353	0.00962	1.9772	0.1597	0.987
Cell:adeno	0.78356	0.3038	6.6512	0.0099	2.189
Cell:small	0.48230	0.26537	3.3032	0.0691	1.620
Cell:squamous	-0.40770	0.28363	2.0663	0.1506	0.665
Prior: yes	0.45914	0.28868	2.5296	0.1117	.
Therapy:test	0.56662	0.24765	5.2349	0.0221	.
Prior*Therpy	-0.87579	0.42976	4.1528	0.0416	.

SAS Example

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	26.5239	4	<.0001
Score	31.2621	4	<.0001
Wald	24.3090	4	<.0001

Summary

- Hypothesis testing based on partial likelihood function
  - Wald test
  - Score test
  - LR test
- Three tests are asymptotically equivalent. No theoretical basis for preferring one test than anther for large samples.