Statistics for Biology and Health
Chapter 9 Additive Hazards Regression Models

Qi Guo

July 25,2019

THE UNIVERSITY OF TEXAS AT DALLAS
School of Natural Sciences and Mathematics

Contents

Introduction

Introduction

- The regression models for survival data based on a proportional hazards model before show the effect of covariates act multiplicatively on some unknown baseline hazard rate, and covariates which do not act on the baseline hazard rate in this fashion were modeled either by the inclusion of a time-dependent covariate or by stratification.
- Now consider an alternative to the semiparametric multiplicative hazard model, namely, the additive hazard model, which is expressed by:

$$h[t|Z(t)] = \beta_0(t) + \sum_{k=1}^{p} \beta_k(t) Z_k(t)$$

where the $\beta_k(t)$'s are covariate functions to be estimated from the data.

Aalen’s Nonparametric, Additive Hazard Model

Introduction

- The first additive hazard model is Aalen's Nonparametric model, the unknown risk coefficients in this model are allowed to be functions of time so that the effect of a covariate may vary over time.
- And the risk coefficients can be estimated by a least-squares technique.

Notations

- The data consists of a sample $[T_j, \delta_j, Z_j(t)],\ j = 1, \cdots, n$, T_j is the on study time, δ_j the event indicator, and $Z_j(t) = [Z_{j1}(t), \cdots, Z_{jp}(t)]$ is a $p-$vector of, possibly, time-dependent covariates. For the *j*th individual define:

$$Y_j(t) = \begin{cases} 1 \textit{ if individual } j \textit{ is under observation } (\textit{at risk}) \textit{at time } t \\ 0 \textit{ otherwise} \end{cases}$$

- If the data is left-truncated, then, $Y_j(t)$ is 1 only between an individual's entry time into the study and exit time from the study. For right-censored data $Y_j(t)$ is 1 if $t \leq T_j$.

Aalen's Nonparametric additive hazard model

- For individual j, the conditional hazard rate at time t, given $Z_j(t)$, by

$$h[t|Z_j(t)] = \beta_0(t) + \sum_{k=1}^{p} \beta_k(t) Z_{jk}(t)$$

  where $\beta_k(t),\ k = 1, \cdots, p$ are unknown parametric functions to be estimated.
- Directly estimate the cumulative risk function like before $B_k(t)$, defined by:

$$B_k(t) = \int_0^t \beta_k(u) du,\ k = 0, 1, \cdots, p$$

- Crude estimates of $\beta_k(t)$ are given by the slope of estimate of $B_k(t)$. Better estimates of $\beta_k(t)$ can be obtained by using a kernel-smoothing technique.

Least-squares estimate

- To find the estimates of $B_k(t)$ a least-squares method is used, define an $n \times (p+1)$ design matrix, $X(t)$, as follows:
- For the ith row of $X(t)$, set $X_i(t) = Y_i(t)(1, Z_j(t))$. That's $X_i(t) = (1, Z_{j1}(t), \cdots, Z_{jp}(t))$ if the ith subject is at risk at time t, and a $p+1$ vector of zeros if this subject is not at risk. Let $I(t)$ be the $n \times 1$ vector with ith element equal to 1 if subject i dies at t and 0 otherwise.
- The least-squares estimate of the vector $B(t) = (B_0(t), B_1(t), \cdots, B_p(t))'$ is

$$\hat{B}(t) = \sum_{T_i \leq t} [X'(T_i)X(T_i)]^{-1} X'(T_i) I(T_i)$$

Least-squares estimate

- The variance-covariance matrix of $B(t)$ is:

$$\hat{Var}(\hat{B}(t)) = \sum_{T_i \leq t} [X'(T_i)X(T_i)]^{-1} X'(T_i) I^D(T_i) X(T_i) \{[X'(T_i)X(T_i)]^{-1}\}'$$

- $I^D(t)$ is the diagonal matrix with diagonal elements equal to $I(t)$. The estimator $B(t)$ only exists up tp the time t, which is the smallest time at which $X'(T_i)X(T_i)$ becomes singular.
- The estimators $\hat{B}_k(t)$ estimate the integral of the regression function b_k in the same fashion as the NA estimator discussed in Chapter 3.

Example

- Here we have a single covariate Z_{j1} equal to 1 if the jth observation is from sample 1 and 0 otherwise. The design matrix is:

$$X(t) = \begin{pmatrix} Y_1(t) & Y_1(t)Z_{11} \\ \cdots & \cdots \\ Y_n(t) & Y_n(t)Z_{1n} \end{pmatrix}$$

- and

$$X'(t)X(t) = \begin{pmatrix} N_1(t) + N_2(t) & N_1(t) \\ N_1(t) & N_1(t) \end{pmatrix}$$

- Here $N_k(t)$ is the number at risk in the kth group at time t. The $X'(t)X(t)$ is nonsingular as long as there is at least one subject still at risk in each group.

Example

- From the formula before find that:

$$\hat{B}_0(t) = \sum_{T_i \le t} d_i \{ \frac{1}{N_2(T_i)} - \frac{Z_{i1}}{N_2(T_i)} \}$$
$$= \sum_{T_i \le t, i \in sample\ 2} \frac{d_i}{N_2(T_i)}.$$

- The NA estimator of the hazard rate using the data in sample 2 only. Also:

$$\hat{B}_1(t) = \sum_{T_i \le t} d_i \{ \frac{-1}{N_2(T_i)} + Z_{i1} [\frac{1}{N_1(T_i)} + \frac{1}{N_2(T_i)}] \}$$
$$= \sum_{T_i \le t, i \in sample\ 1} \frac{d_i}{Y_1(T_i)} - \sum_{T_i \le t, i \in sample\ 2} \frac{d_i}{Y_2(T_i)}.$$

- The difference of the NA estimators of the hazard rate in the two samples.

The Weighted additive hazard model

- To make it more flexible, when test the hypothesis $H_0 : \beta_k(t) = 0$ for all $t \leq \tau$ and all k in some set k, it based on a weighted stochastic integral of the estimated value of $\beta_k(t)$ as compared to its expected value, zero, under H_0.
- To perform the test we need a matrix of weights to use in constructing the test, this weight matrix is a diagonal matrix $W(t)$ with diagonal elements $W_j(t), j = 1, \cdots, p+1$. The test statistic:

$$U = \sum_{T_i} W(T_i)[X'(T_i)X(T_i)]^{-1}X'(T_i)I(T_i)$$

- The $(j+1)$st element of U is the test statistic for testing the hypothesis $H_j : \beta_j(t) = 0$

The Weighted additive hazard model

- The elements of U are weighted sums of the increments of $\hat{B}_k(t)$ and elements of V are also obtained from elements of $\hat{Var}(\hat{B}(t))$.
- Using the statistics a simultaneous test of the hypothesis that $\beta_j(t) = 0$ for all $j \in J$ where J is a subset of $\{0, 1, \cdots, p + 1, \}$ is

$$X = U_J' V_J^{-1} I U_J$$

- Here U_J is the subvector of U corresponding to elements in J and V_J the corresponding subcovariance matrix.
- And we can choose the weight function like before.

Lin and Ying's Additive Hazards Model

Lin and Ying's Additive Hazards Model

- The Lin and Ying additive model for the conditional hazard rate for individual j with covariate vector $Z_j(t)$ is:

$$h(t|Z_j(t)) = \alpha_0(t) + \sum_{k=1}^{p} \alpha_k Z_{jk}(t)$$

  where $\alpha_k, k = 1, \cdots, p$ are unknown parameters and $\alpha_0(t)$ is an arbitrary baseline function.
- As usual our data consists of a sample (T_i, δ_j, Z_j), and here skip the introduction of all notations.

Lin and Ying's Additive Hazards Model

- To construct the estimates of $\alpha_k, k = 1, \cdots, p$, first construct the vector $\bar{Z}(t)$, which is the average value of the covariates at time t. That is,

  $$\bar{Z}(t) = \frac{\sum_{i=1}^{n} Z_i Y_i(t)}{\sum_{i=1}^{n} Y_i(t)}$$

  the numerator is the sum of the covariates for all individuals at risk at time t and the denominator is the number at risk at time t, and then construct the $p \times p$ matrix A given by:

  $$A = \sum_{i=1}^{n} \sum_{j=1}^{i} (T_j - T_{j-1})[Z_i - \bar{Z}(T_j)]'[Z_i - \bar{Z}(T_j)]$$

Lin and Ying's Additive Hazards Model

- The $p-$vector B given by:

$$B' = \sum_{i=1}^{n} \delta_i [Z_i - \bar{Z}(T_i)]$$

and the $p \times p$ matrix C by:

$$C = \sum_{i=1}^{n} \delta_i [Z_i - \bar{Z}(T_i)]' [Z_i - \bar{Z}(T_i)]$$

- The estimate of $\alpha = (\alpha_1, \cdots, \alpha_p)$ is:

$$\hat{\alpha} = A^{-1} B'$$

- and the estimated variance of α is given by:

$$\hat{V} = \hat{Var}(\hat{\alpha}) = A^{-1} C A^{-1}$$

Lin and Ying's Additive Hazards Model

- To test the hypothesis $H_j : \alpha_k = 0$, use the statistic:

$$\frac{\hat{\alpha}_j}{\sqrt{\hat{V}_{jj}}}$$

which has a $N(0, 1)$ distribution for large n under the H_0, the test of the hypothesis that $\alpha_j = 0$ for all $j \in J$ is based on the quadratic form:

$$\chi^2 = [\hat{\alpha}_J - 0]' \hat{V}_J^{-1} [\hat{\alpha}_J - 0]$$

- Under the H_0 the statistic has a chi squared distribution with degrees of freedom equal to the dimension of J.