1 Aalen's Additive Risk Model

Ian W. McKeague

1.1 The Model

In medical statistics and survival analysis, it is important to assess the association between risk factors and disease occurrence or mortality. Underlying disease mechanisms are invariably complex, so the idea is to simplify the relationship between survival patterns and covariates in such a way that only essential features are brought out. Aalen's additive risk model [1] is one of three well-developed approaches to this problem; the others are the popular proportional hazards model introduced by D. R. Cox in 1972 [5], and the accelerated failure-time model, which is a linear regression model with unknown error distribution, introduced in the context of right-censored survival data by R. G. Miller in 1976 [18].

Aalen's model expresses the conditional hazard function $\lambda(t|\mathbf{z})$ of a survival time T as a linear function of a p-dimensional covariate vector $\mathbf{z}$

$$\lambda(t|\mathbf{z}) = \boldsymbol{\alpha}(t)'\mathbf{z} = \sum_{j=1}^{p} \alpha_j(t) z_j, \qquad (1)$$

where $\boldsymbol{\alpha}(t)$ is a nonparametric p vector of regression functions [constrained by $\lambda(t|\mathbf{z}) \geq 0$] and $\mathbf{z} = (z_1, \ldots, z_p)'$. Some authors refer to (1) as the *linear hazard model*.

As a function of the covariates $z_1, \ldots, z_p$, the *additive* form of Aalen's model contrasts with the *multiplicative* form of Cox's model

$$\begin{aligned}\lambda(t|\mathbf{z}) &= \lambda_0(t) \exp\{\boldsymbol{\beta}'\mathbf{z}\} \\ &= \lambda_0(t) \prod_{j=1}^{p} \exp\{\beta_j z_j\},\end{aligned}$$

where $\lambda_0(t)$ is a nonparametric baseline hazard function and $\boldsymbol{\beta}$ is a vector of regression parameters. Aalen's model has the feature that the influence of each covariate can vary separately and nonparametrically through time, unlike Cox's model or the accelerated failure-time model. This feature can be desirable in some applications, especially when there are a small number of covariates.

Consider the following simple example with three covariates: T is the age at which an individual contracts melanoma (if at all), z_1 = indicator male, z_2 = indicator female, and z_3 = number of serious sunburns as a child. Then the corresponding regression functions, α_1, α_2, and α_3, can be interpreted as the (age-specific) background rates of melanoma for males and females and as the excess rate of melanoma due to serious sunburns is childhood, respectively.

Aalen's model is expected to provide a reasonable fit to data, since the first step of

a Taylor series expansion of a general conditional hazard function about the zero of the covariate vector can be expressed in the form (1). It is somewhat more flexible than Cox's model and can be especially helpful for exploratory data analysis. A rough justification for the additive form can be given in terms of p independent competing risks, since the hazard function of the minimum of p independent random variables is the sum of their individual hazard functions.

It is generally sensible to include a nonparametric baseline function in the model, by augmenting $\mathbf{z}$ with a component that is set to 1. Also, it is often natural to center the covariates in some fashion, so the baseline can be interpreted as the "hazard" function for an "average" individual. In some cases, however, a baseline hazard is already implicit in the model and it is not necessary to center the covariates, as in the melanoma example above.

Aalen originally proposed his model in a counting process setting, which allows time-dependent covariates and general patterns of censorship, and which can be studied using powerful continuous-time martingale techniques. In a typical application the observed survival times are subject to right censorship, and it is customary to assume that the censoring time, C, say, is conditionally independent of T given $\mathbf{z}$. One observes $(X, \delta, \mathbf{z})$, where $X = T \wedge C$ and $\delta = I\{X = T\}$. Aalen's model (1) is now equivalent to specifying that the counting process $N(t) = I(X \leq t, \delta = 1)$, which indicates an uncensored failure by time t, has intensity process

$$\lambda(t) = \boldsymbol{\alpha}(t)'\mathbf{y}(t),$$

where $\mathbf{y}(t) = \mathbf{z}I\{X \geq t\}$ is a covariate process.

1.2 Model Fitting

To fit Aalen's model one first estimates the p vector of integrated regression functions $\mathbf{A}(t) = \int_0^t \boldsymbol{\alpha}(s)\,ds$. Denote by $(t_i, \delta_i, \mathbf{z}_i)$ the possibly right-censored failure time t_i, indicator of noncensorship δ_i, and covariate vector $\mathbf{z}_i$ for n individuals. Let $\mathbf{N} = (N_1, \ldots, N_n)'$ and $\mathbf{Z} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)'$, where N_i is the counting process and $\mathbf{y}_i$ is the associated covariate process for individual i.

Aalen [1] introduced an ordinary least squares (OLS) estimator of $\mathbf{A}(t)$ given by

$$\hat{\mathbf{A}}(t) = \int_0^t (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\,d\mathbf{N},$$

where the matrix inverse is assumed to exist; $\hat{\mathbf{A}}$ is a step function, constant between uncensored failures, and with jump

$$\boldsymbol{\Delta}_i = \left(\sum_{t_k \geq t_i} \mathbf{z}_k\mathbf{z}_k'\right)^{-1} \mathbf{z}_i \qquad (2)$$

at an uncensored failure time t_i. The matrix inverse exists unless there is collinearity between the covariates or there are insufficiently many individuals at risk at time t_i. A heuristic motivation for $\hat{\mathbf{A}}$ comes from applying the method of least squares to increments of the multivariate counting process $\mathbf{N}$. The estimator is consistent and asymptotically normal [14,9]. The covariance matrix of $\hat{\mathbf{A}}(t)$ can be estimated [1,2] by $\hat{\mathbf{V}}(t) = \sum_{t_i \leq t} \delta_i \boldsymbol{\Delta}_i \boldsymbol{\Delta}_i'$.

Plots of the components of $\hat{\mathbf{A}}(t)$ against t, known as *Aalen plots*, are a useful graphical diagnostic tool for studying time-varying covariate effects [2–4,7,9,12,13]. Mau [12] coined the term Aalen plots and made a strong case for their importance in survival analysis. Roughly constant slopes in the plots indicate periods when a covariate has a non-time-dependent regression coefficient; plateaus indicate times at which a covariate has no effect on the hazard. Interpretation of the plots is helped by the inclusion of pointwise or simultaneous confidence limits. An approximate pointwise $100(1-\alpha)\%$ confidence interval

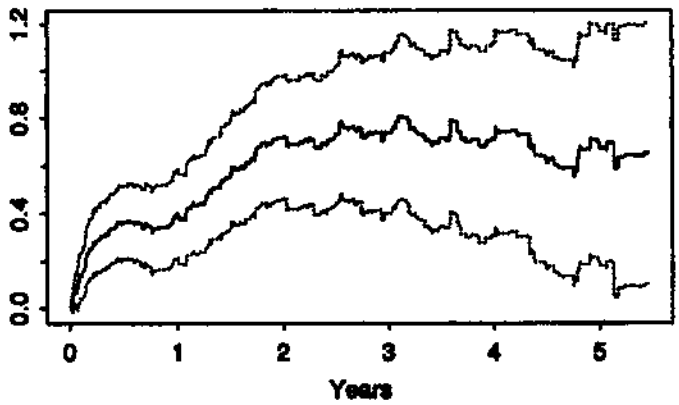

Figure 1: An Aalen plot with 95% confidence limits for the myelamatosis data.

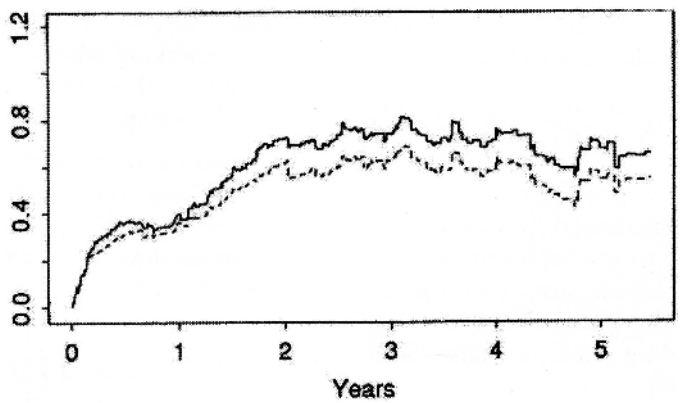

Figure 2: Comparison of Aalen plots of the WLS estimates (dashed line) and OLS (solid line) estimates for the myelamatosis data.

for the jth component of $\mathbf{A}(t)$ is given by

$$\hat{\mathbf{A}}_j(t) \pm z_{\alpha/2}\hat{\mathbf{V}}_{jj}(t)^{1/2},$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ quantile of the standard normal distribution and $\hat{\mathbf{V}}_{jj}(t)$ is the jth entry on the diagonal of $\hat{\mathbf{V}}(t)$. To avoid wild fluctuations in the plots (which occur when the size of the risk set is small), estimation should be restricted to time intervals over which the matrix inverse in (2) is numerically stable.

Figure 1 shows an Aalen plot based on survival data for 495 myelamatosis patients [17]. The plot gives the estimated integrated regression function for one particular covariate, serum β_2-microglobulin, which was log-transformed to adjust for skewness. Pointwise 95% confidence limits are also shown. Serum β_2-microglobulin is seen to have a strong effect on survival during the first 2 years of follow-up.

The vector of regression functions $\boldsymbol{\alpha}$ can be estimated by smoothing the increments of $\hat{\mathbf{A}}$. One approach is to extend Ramlau-Hansen's kernel estimator [19] to the additive-risk-model setting [3,9,14]. For a kernel function K that integrates to 1 and some bandwidth $b > 0$,

$$\hat{\boldsymbol{\alpha}}(t) = b^{-1}\sum_{i=1}^{n} K\left(\frac{t-t_i}{b}\right)\boldsymbol{\Delta}_i$$

consistently estimates $\boldsymbol{\alpha}$ provided the bandwidth tends to zero at a suitable rate with increasing sample size. Plots of the regression function estimates in some real and simulated data examples have been given by Aalen [3].

Huffer and McKeague [9] introduced a weighted least-squares (WLS) estimator of $\mathbf{A}$; see Figure 2 for a comparison with the OLS estimator. The weights consistently estimate $[\lambda(t|\mathbf{z}_i)]^{-1}$ and are obtained by plugging $\hat{\alpha}(t)$ and $\mathbf{z} = \mathbf{z}_i$ into (1). The WLS estimator is an approximate maximum-likelihood estimator and an approximate solution to the score equations [20]. It is consistent and asymptotically normal provided $\lambda(t|\mathbf{z})$ is bounded away from zero [9,14]. Furthermore, the WLS estimator is asymptotically efficient in the sense of having minimal asymptotic variance [6,20,4]. Simulation studies [9] show that significant variance reductions are possible using WLS compared with OLS estimators, especially in large samples where the weights are more stable. When there are no covariates ($p = 1$ and $\mathbf{z}_i = 1$), the OLS and WLS estimators reduce to the Nelson–Aalen estimator of the baseline cumulative hazard function. Simultaneous confidence bands for $\mathbf{A}$ based on OLS and WLS estimators, for continuous or grouped data, can be found in References 9, 15.

Tests of whether a specific covariate (say, the jth component of $\mathbf{z}$) has any effect on survival can be carried out within the Aalen model setting. The idea is to test the null hypothesis $H_0 : \mathbf{A}_j(t) = 0$ over the follow-up period. This can be done [2] using a test statistic of the form $\sum_{i=1}^{n} w(t_i)\boldsymbol{\Delta}_{ij}$ for a suitable weight function w. Kolmogorov–Smirnov-type tests are also available [9]; such tests are equivalent to checking whether the confidence band for the jth component of $\mathbf{A}$ contains the zero function.

To predict survival under Aalen's model, one estimates the conditional survival probability $P(T > t|\mathbf{z}) = \exp\{-\mathbf{A}(t)'\mathbf{z}\}$. This can be done using the product-limit estimator

$$\hat{P}(T > t|\mathbf{z}) = \prod_{t_i \le t}(1 - \boldsymbol{\Delta}_i'\mathbf{z}),$$

or by plugging $\hat{\mathbf{A}}(t)$ into $P(T > t|\mathbf{z})$ in place of the unknown $\mathbf{A}(\mathbf{t})$. When there are no covariates, $\hat{P}(T > t|\mathbf{z})$ reduces to the Kaplan–Meier estimator of the survival function corresponding to the baseline hazard.

1.3 Model Diagnostics

Some goodness-of-fit checking procedures are available for additive risk models. Aalen [2,3] suggested making plots against t of sums of the martingale residual processes $\hat{M}_i(t) = \delta_i I(t_i \leq t) - \hat{\mathbf{A}}(t_i \wedge t)'\mathbf{z}_i$ over groups of individuals. If the model fits well, then the plots would be expected to fluctuate around the zero line. McKeague and Utikal [16] suggested the use of a standardized residual process plotted against t and $\mathbf{z}$, and developed a formal goodness-of-fit test for Aalen's model.

Outlier detection has been studied by Henderson and Oman [8], who considered the effects on $\hat{\mathbf{A}}(t)$ of deletion of an observation from a dataset. They show that unusual or influential observations can be detected quickly and easily. They note that Aalen's model has an advantage over Cox's model in this regard, because closed-form expressions for the estimators are available, leading to exact measures of the effects of case deletion.

Mau [12] noticed that Aalen plots are useful for diagnosing time-dependent covariate effects in the Cox model. To aid interpretation of the plots in that case, Henderson and Milner [7] suggested that an estimate of the shape of the curve expected under proportional hazards be included.

1.4 Related Models

More recently a number of variations on the additive structure of Aalen's model have been introduced. McKeague and Sasieni [17] considered a partly parametric additive risk model in which the influence of only a subset of the covariates varies nonparametrically over time, and that of the remaining covariates is constant

$$\lambda(t|\mathbf{x},\mathbf{z}) = \boldsymbol{\alpha}(t)'\mathbf{x} + \boldsymbol{\beta}'\mathbf{z}, \qquad (3)$$

where $\mathbf{x}$ and $\mathbf{z}$ are covariate vectors and $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\beta}$ are unknown. This model may be more appropriate than (1) when there are a large number of covariates and it is known that the influence of only a few of the covariates is time-dependent. Lin and Ying [10] studied an additive analog of Cox's proportional hazards model that arises as a special case of (3):

$$\lambda(t|\mathbf{z}) = \alpha_0(t) + \boldsymbol{\beta}'\mathbf{z}. \qquad (4)$$

Efficient WLS-type estimators for fitting (3) and (4) have been developed.

A variation in the direction of Cox's proportional hazards model [5] has been studied by Sasieni [21,22]: the *proportional excess hazards* model

$$\lambda(t|\mathbf{x},\mathbf{z}) = \alpha_0(t|\mathbf{x}) + \lambda_0(t)\exp\{\boldsymbol{\beta}'\mathbf{z}\}, \quad (5)$$

where $\alpha_0(t|\mathbf{x})$ is a known background hazard (available from national mortality statistics, say) and $\lambda_0(t)$ and $\boldsymbol{\beta}$ are unknown. A further variation in this direction is due to Lin and Ying [11], who considered an *additive–multiplicative hazards* model that includes

$$\lambda(t|\mathbf{x},\mathbf{z}) = \boldsymbol{\gamma}'\mathbf{x} + \lambda_0(t)\exp\{\boldsymbol{\beta}'\mathbf{z}\}, \quad (6)$$

where $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$, and $\lambda_0(t)$ are unknown. Finding efficient procedures for fitting the models (5) and (6) involves a combination of Cox partial likelihood techniques and the estimation of efficient weights similar to those needed for the standard additive risk model (1).

1.5 Conclusion

Despite the attractive features of Aalen's model as an alternative to Cox's model in many applications, it has received relatively little attention from practitioners or researchers. Cox's model has been perceived to be adequate for most applications, but it can lead to serious bias when the influence of covariates is time-dependent. Fitting separate Cox models over disjoint time intervals (years, say)

is an ad hoc way around this problem. Aalen's model, however, provides a more effective approach. Interest in it, and especially in Aalen plots, is expected to increase in the future.

Acknowledgments. This research was partially supported by NSF Grant ATM-9417528.

References

1. Aalen, O. O. (1980). A model for nonparametric regression analysis of counting processes. *Lecture Notes Statist.*, **2**, 1–25. Springer-Verlag, New York. (Aalen originally proposed his model at a conference in Poland; this paper appeared in the conference proceedings.)

2. Aalen, O. O. (1989). A linear regression model for the analysis of life times. *Statist. Med.*, **8**, 907–925. (A readable introduction to additive risk models with an emphasis on graphical techniques. Results based on the Aalen and Cox models are compared.)

3. Aalen, O. O. (1993). Further results on the nonparametric linear regression model in survival analysis. *Statist. Med.*, **12**, 1569–1588. (Studies the use of martingale residuals for assessing goodness of fit of additive risk models.)

4. Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York. (A comprehensive and in-depth survey of the counting process approach to survival analysis, including the additive risk model and its associated estimators.)

5. Cox, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. B*, **34**, 187–220.

6. Greenwood, P. E. and Wefelmeyer, W. (1991). Efficient estimating equations for nonparametric filtered models. In *Statistical Inference in Stochastic Processes*, N. U. Prabhu and I. V. Basawa, eds., pp. 107–141. Marcel Dekker, New York. (Shows that the weighted least-squares estimator is an approximate maximum-likelihood estimator and that this implies asymptotic efficiency.)

7. Henderson, R. and Milner, A. (1991). Aalen plots under proportional hazards. *Appl. Statist.*, **40**, 401–410. (Introduces a modification to Aalen plots designed to detect time-dependent covariate effects in Cox's proportional hazards model.)

8. Henderson, R. and Oman, P. (1993). Influence in linear hazard models. *Scand. J. Statist.*, **20**, 195–212. (Shows how to detect unusual or influential observations under Aalen's model.)

9. Huffer, F. W. and McKeague, I. W. (1991). Weighted least squares estimation for Aalen's additive risk model. *J. Am. Statist. Assoc.*, **86**, 114–129.

10. Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika*, **81**, 61–71.

11. Lin, D. Y. and Ying, Z. (1995). Semiparametric analysis of general additive–multiplicative hazard models for counting processes. *Ann. Statist.*, **23**, 1712–1734.

12. Mau, J. (1986). On a graphical method for the detection of time-dependent effects of covariates in survival data. *Appl. Statist.*, **35**, 245–255. (Makes a strong case that Aalen plots can provide important information that might be missed when only Cox's proportional hazards model is applied.)

13. Mau, J. (1988). A comparison of counting process models for complicated life histories. *Appl. Stoch. Models Data Anal.*, **4**, 283–298. (Aalen's model is applied in the context of intermittent exposure in which individuals alternate between being at risk and not.)

14. McKeague, I. W. (1988). Asymptotic theory for weighted least squares estimators in Aalen's additive risk model.

In *Statistical Inference from Stochastic Processes*, N. U. Prabhu, ed.; *Contemp. Math.*, **80**, 139–152. American Mathematical Society, Providence. (Studies the asymptotic properties of estimators in the counting process version of Aalen's model.)

15. McKeague, I. W. (1988). A counting process approach to the regression analysis of grouped survival data. *Stoch. Process. Appl.*, **28**, 221–239. (Studies the asymptotic properties of grouped data based estimators for Aalen's model.)

16. McKeague, I. W. and Utikal, K. J. (1991). Goodness-of-fit tests for additive hazards and proportional hazards models. *Scand. J. Statist.*, **18**, 177–195.

17. McKeague, I. W. and Sasieni, P. D. (1994). A partly parametric additive risk model. *Biometrika*, **81**, 501–514.

18. Miller, R. G. (1976). Least squares regression with censored data. *Biometrika*, **63**, 449–464.

19. Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *Ann. Statist.*, 11, 453–466.

20. Sasieni, P. D. (1992). Information bounds for the additive and multiplicative intensity models. In *Survival Analysis: State of the Art*, J. P. Klein and P. K. Goel, eds., pp. 249–265. Kluwer, Dordrecht. (Shows asymptotic efficiency of the weighted least-squares estimators in the survival analysis setting.)

21. Sasieni, P. D. (1995). Efficiently weighted estimating equations with application to proportional excess hazards. *Lifetime Data Analysis*, 1, 49–57.

22. Sasieni, P. D. (1996). Proportional excess hazards. *Biometrika*, **83**, 127–141.