Published in final edited form as:

J Stat Plan Inference. 2009 April 1; 139(4): 1381–1393. doi:10.1016/j.jspi.2008.08.006.

Semiparametric inference of proportional odds model based on randomly truncated data

Rajeshwari Sundaram¹

Biometry and Mathematical Statistics Branch, Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, NIH, DHHS, 6100 Executive Boulevard, Rockville, MD 20878, USA

Abstract

This paper studies the estimation in the proportional odds model based on randomly truncated data. The proposed estimators for the regression coefficients include a class of minimum distance estimators defined through weighted empirical odds function. We have investigated the asymptotic properties like the consistency and the limiting distribution of the proposed estimators under mild conditions. The finite sample properties were investigated through simulation study making comparison of some of the estimators in the class. We conclude with an illustration of our proposed method to a well-known AIDS data.

Key words and Phrases

Proportional odds model; Weighted empirical processes; Minimum Distance; Truncation; Survival

1 Introduction

In many prospective and retrospective studies, survival data encountered are often incomplete due to censoring and/or truncation. Examples of randomly truncated data can be found in various areas like astronomy, economics (e.g. Woodroffe (1985), Fiegelson and Babu (1992)), epidemiology and biometry (e.g. Keiding, Bayer and Watt-Boolsen (1987), Lagakos, Barraj and DeGruttola (1988)). For instance, random truncation occurs naturally in retrospective studies of a disease when only data available are those pertaining to individuals who have experienced some ascertaining event (E_2), prior to some time point t^* . So, individuals are enrolled into the study only if their time of occurrence t_2 of the event E_2 is less than or equal to t^* . For these individuals, the time t_1 of an initiating event (E_1) is determined retrospectively. Note that the survival time of interest $T = t_2 - t_1$ is observed only if $T \leq Y$, with $Y = t^* - t_1$ denoting the right truncating variable. An example of such a data was encountered in AIDS study (Kalbfleisch and Lawless (1989)), where an individual is infected with human immunodeficiency virus (HIV) at time t_1 and diagnosed with acquired immune deficiency syndrome (AIDS) at time t_2 . If the observation period is terminated at t^* , then only those individuals for whom the incubation time $T = t_2 - t_1$ is less than or equal to $Y = t^* - t_1$ can be observed and right truncation occurs. Another example can be found in AIDS cohort studies, where a group of patients who are infected with HIV at time s but have

¹This research was supported by the Intramural Research Program of the NIH, NICHD

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

not yet developed AIDS (at some time $t > s$) are selected. If the recruitment starts at Y_0 then only those individuals for whom $T = t - s \geq Y = Y_0 - s$ are observed and random left truncation occurs. So, under random left truncation, one observes a pair (T, Y) only if $T \geq Y$, where T is the true variable of interest and Y is the truncating random variable. One can view random right truncation as random left truncation by just multiplying the survival time T and the truncating time Y by -1 .

Nonparametric inference of distribution function based on randomly truncated data has been studied by Woodroffe (1985), Wang, Jewell and Tsai (1986), Keiding and Gill (1990), Lai and Ying (1991) among others; regression analysis of such data has been studied by Gross and Lai (1996), Grigoletto and Akritas (1999) and He and Yang (2003) among others. In this paper, our goal is to develop methodology to analyze the semiparametric proportional odds model based on randomly truncated data. The proportional odds model has been suggested as an alternative to the Cox's proportional hazards model in analyzing the effect of covariates on event times. Bennett (1983a,b) and Pettitt (1982) were among the first to generalize proportional odds model to the context of survival analysis. In the two sample setup, the proportional odds model is useful for fitting the data whose hazard rates converge asymptotically. Bennett (1983a) has suggested using proportional odds structure to model effective cure as the mortality rate of a disease group would converge to the mortality rate of a control group as time progresses under this model. Moreover, as pointed out by Murphy et. al. (1997), this model can be used to demonstrate that the morbidity rate for a group halting a deleterious habit, such as smoking, converges to the morbidity rate of similar never-smokers as time progresses. Other examples of application of this model can be found in the analysis of environmental health data and in analysis of nonlethal tumors. In studying harmful effects of chemical exposure, experiments are performed in which groups of animals are exposed to different levels of chemicals. The resulting tumor distributions between groups are often compared using the proportional odds structure.

The proportional odds model specifies that the distribution function F_Z for failure time T associated with a p -dimensional covariate Z is parameterized by

$$\frac{1 - F_Z(t)}{F_Z(t)} = \exp(Z' \beta) \frac{1 - F(t)}{F(t)}. \quad (1.1)$$

Inference procedures under the above model has been studied by Pettitt (1982), Bennett (1983a,b), Bickel (1986), Dabrowska and Doksum (1988a) among others for complete data. Cheng et. al. (1995), Dabrowska and Doksum (1988b), Murphy et. al. (1997), and Yang and Prentice (1999), Sundaram (2006) have studied the inference for the above model based on right censored data. Huang and Rossini (1997) studied the above model for interval censored data and Rossini and Tsiatis (1996) for current status data. Grigoletto and Akritas (1999) have studied inference when the data is possibly left truncated. However, their method is only applicable for data with just one continuous covariate and the rest of the covariates have to be categorical. This may be too restrictive in practice. In this paper, we propose a methodology for regression analysis of semiparametric proportional odds model based on randomly truncated data with no such restriction on the covariates.

Motivated by robustness properties and ease of computation, Yang and Prentice (1999) proposed a weighted empiricals approach for analyzing the proportional odds model based on right censored data. They showed that their estimators had comparable performance in terms of efficiency, to the profile likelihood estimator of Murphy et. al. (1997). In this paper, we generalize the weighted empiricals approach to study estimation of model (1.1) under random truncation. The organization of the paper is as follows: in Section 2, we discuss the

inference procedure, the large sample properties of the proposed estimators are discussed in Section 3 and in Section 4, we present a numerical study and analyze the AIDS incubation time data presented in Kalbfleisch and Lawless (1989) to illustrate the proposed methodology.

2 Preliminaries

Let (T, Z) denote the survival time and the corresponding p -covariate. Let $F_Z(S_Z)$ denote the distribution (survival) function of T given Z , and $F(S)$ denote the distribution (survival) function of T given $Z = 0$. Let $R_Z = F_Z/(1 - F_Z)$ denote the odds function for failure given Z and let $R = F/(1 - F)$ denote the baseline odds function for failure, i.e., the odds function for failure given $Z = 0$. Under proportional odds model, the distribution of T given the covariate Z is modeled by

$$R_Z(t) = \exp(-Z'\beta)R(t). \quad (2.2)$$

The unknown parameters are R , the baseline odds function and β , a p -dimensional vector of regression coefficients. This model (2.2) has been studied by Yang and Prentice (1999) and Murphy et. al. (1997) when data is possibly right censored. This paper studies inference of the above model (2.2) when the variable of interest T is possibly randomly left truncated. Under random truncation, the observation (T, Z) is interfered by another independent variable Y such that all three quantities are observed only if $T \geq Y$. In absence of random truncation, one could think of the observations as an infinite sequence of independent random vectors $\{(T_k, Y_k, Z_k); k = 1, 2, \dots\}$. Under random truncation, however, some of these vectors would be missing and only a subsequence $\{(T_{k_i}, Y_{k_i}, Z_{k_i}); i = 1, 2, \dots\}$ can be observed. We will denote the observable subsequence by

$$(U_i, V_i, W_i), i=1, 2, \dots \text{ subject to } U_i \geq V_i, \quad (2.3)$$

i.e., $U_i = T_{k_i}, V_i = Y_{k_i}, W_i = Z_{k_i}$. So, the observed randomly truncated sample of size n will be denoted by $\{(U_i, V_i, W_i), i = 1, 2, \dots, n\}$.

Denote the cumulative distribution (survival) function of the survival time T and the random truncating variable Y by $F_z(S_z)$ and $G_z(\bar{G}_z)$ respectively, conditional on $Z = z$. Typically, under random truncation T and Y are assumed independent, conditional on the covariate $Z = z$. Let H denote the cumulative distribution function of Z . Let (a_F, b_F) denote the support of F defined by

$$a_F = \inf\{x: F(x) > 0\} \text{ and } b_F = \sup\{x: F(x) < 1\}.$$

Let a_G and b_G be similarly defined for the distribution G . Under random truncation, only the conditional distributions

$$P(T \leq x | T \geq a_G) \text{ and } P(Y \leq x | Y \leq b_F)$$

can be estimated nonparametrically as random truncation restricts the observation range of T and Y . Furthermore, one can show that under the restriction $a_G \leq a_F$, the model (2.2) is

identifiable. As a consequence, throughout we will assume $a_G = a_F = 0$ and $b_F = \infty$. Under this assumption, the conditional distributions mentioned above will be the actual distributions of T and Y . In addition, we also assume that

$$P(T > Y) = \alpha > 0.$$

Obviously, if $\alpha = 0$, no data can be observed. Note that Lemma 2.1 of He and Yang (2003) shows that the i.i.d. property of the observations is preserved under random truncation. As shown in their Lemma 2.1, the random vectors (U_i, V_i, W_i) defined in (2.3) are i.i.d. with a common distribution

$$\begin{aligned} F^*(u, v, w) &= P(U_1 \leq u, V_1 \leq v, W_1 \leq w) \\ &= P(T \leq u, Y \leq v, Z \leq w | T \geq Y) \\ &= \frac{1}{\alpha} \int \int_{x \leq u, z \leq w} G_z(x \wedge v) F_z(dx) H(dz), \end{aligned} \quad (2.4)$$

where $x \wedge v = \min(x, v)$. We will denote the marginal distribution of U and V by F^* and G^* respectively. We are now ready to present our estimation method for the regression coefficient β and the unknown baseline function R for randomly truncated data modeled by the proportional odds model (2.2). This estimation procedure has been successfully used by Yang and Prentice (1999) and Sundaram (2006a) for right censored data.

The proportional odds model (2.2) can be expressed in terms of the cumulative hazard $\Lambda_Z(t)$ of T given Z as follows:

$$d\Lambda_Z(t) = \frac{dR(t)}{R(t) + \exp(Z'\beta)},$$

which can be rewritten as follows

$$dR(t) = \exp(Z'\beta) d\Lambda_Z(t) + R(t) d\Lambda_Z(t).$$

Multiplying both sides of the above equation with subject specific non-negative weights

$a_i(t)$, $i = 1, \dots, n$, with the property that $\sum_{i=1}^n a_i(t) \equiv 1$ and integrating throughout from 0 to t yields

$$R(t) = \int_0^t \sum_{i=1}^n a_i(s) \exp(Z_i'\beta) d\Lambda_{Z_i}(s) + \int_0^t R(s) \sum_{i=1}^n a_i(s) d\Lambda_{Z_i}(s). \quad (2.5)$$

The above equation is a Volterra integral equation of the first kind. Using Theorem II.6.3 of Andersen, Borgan, Gill, Keiding (1992), the above equation has a unique solution, which can be explicitly expressed. So, for appropriate choice of weights a_i , R can be estimated and easily computed. In particular, we consider the following $(p+1)$ weights for the n subjects

$$a_i(t) = \frac{1}{\alpha_i} Z_{ij} \bar{F}_{Z_i}(t) G_{Z_i}(t) / \sum_{k=1}^n \frac{1}{\alpha_k} Z_{kj} \bar{F}_{Z_k}(t) G_{Z_k}(t), i=1, \dots, n, \quad (2.6)$$

where $j = 0, 1, \dots, p$; $Z_{i0} = 1$ and Z_{ij} , $j = 1, \dots, p$ are the j -th component for the i -th subject and $\alpha_i = P(T > Y | Z = Z_i)$. Denote by

$$\Lambda_{nj}(t) = \frac{\int_0^t \sum_{i=1}^n \frac{1}{\alpha_i} Z_{ij} G_{Z_i}(s) dF_{Z_i}(s)}{\sum_{i=1}^n \frac{1}{\alpha_i} Z_{ij} \bar{F}_{Z_i}(s) G_{Z_i}(s)}$$

and

$$\Lambda_{nj}(t; b) = \frac{\int_0^t \sum_{i=1}^n \frac{1}{\alpha_i} Z_{ij} e^{Z_i' b} G_{Z_i}(s) dF_{Z_i}(s)}{\sum_{i=1}^n \frac{1}{\alpha_i} Z_{ij} \bar{F}_{Z_i}(s) G_{Z_i}(s)}, b \in \mathbb{R}^p.$$

With this choice of weight functions, the equation (2.5) can be expressed as

$$R(t) = \Lambda_{nj}(t; \beta) + \int_0^t R(s) \Lambda_{nj}(ds), \quad j=0, 1, \dots, p. \quad (2.7)$$

By replacing β by $b \in \mathbb{R}^p$, we define a class of equations (in R_{nj})

$$R_{nj}(t; b) = \Lambda_{nj}(t; b) + \int_0^t R_{nj}(s; b) \Lambda_{nj}(ds), \quad j=0, 1, \dots, p. \quad (2.8)$$

Again, using Theorem II.6.3 of Andersen, Borgan, Gill, Keiding (1992), the above equation has a unique solution, which can be explicitly expressed as follows:

$$R_{nj}(t; b) = \int_0^t \prod_{(s, t]} (1 + d\Lambda_{nj}(du)) d\Lambda_{nj}(ds; b),$$

for $j = 0, 1, \dots, p$ and $b \in \mathbb{R}^p$. As a consequence of uniqueness of the equation (2.8), which is a Volterra equation of the first kind, $R_{nj}(t, b) = R(t)$, if $b = \beta$ for $j = 0, 1, \dots, p$. Now, to define estimators for R , some notation and few definitions need to be introduced.

For $0 \leq t < \infty$, $j = 0, 1, \dots, p$ and fixed b , define the following weighted empirical processes

$$\widehat{K}_{nj}(t) = \sum_{i=1}^n Z_{ij} I\{V_i \leq t < U_i\}, \quad (2.9)$$

$$\widehat{L}_{nj}(t; b) = \sum_{i=1}^n Z_{ij} \exp(Z_i' b) I\{U_i \leq t\}, \quad (2.10)$$

and by

$$\widehat{L}_{nj}(t) = \sum_{i=1}^n Z_{ij} I\{U_i \leq t\}. \quad (2.11)$$

Note that the above empirical processes are weighted versions of the commonly used counting processes, counting the number of events that occur by certain time t , and the risk set size by time t in case of random truncation. In presence of random truncation, the risk set $\widehat{K}_{nj}(\cdot)$ is non-monotonic and may be zero for k -th order statistics $U_{(k)}$ because of empty risk sets. In order to avoid instability caused by this property, Lai and Ying (1991) have suggested using sufficiently large risk sets while defining the estimate for the hazards function. Throughout, we will assume the covariates to be non-negative for technical reasons, which happens quite naturally in survival analysis. Furthermore, observe that conditional on the covariates Z_i , using (2.5) the expectations of the weighted empirical processes defined above are

$$\begin{aligned} E\widehat{L}_{nj}(t) &= \sum_{i=1}^n \frac{1}{\alpha_i} Z_{ij} \int_0^t \overline{F}_{Z_i}(s) G_{Z_i}(s) d\Lambda_{Z_i}(s), \\ E\widehat{L}_{nj}(t; b) &= \sum_{i=1}^n \frac{1}{\alpha_i} Z_{ij} \exp(Z_i' b) \int_0^t \overline{F}_{Z_i}(s) G_{Z_i}(s) d\Lambda_{Z_i}(s) \end{aligned}$$

and

$$E\widehat{K}_{nj}(t) = \sum_{i=1}^n \frac{1}{\alpha_i} Z_{ij} \overline{F}_{Z_i}(t) G_{Z_i}(t).$$

So, using these weighted empirical processes, for each b , $0 \leq j \leq p$, one can estimate Λ_{nj} and Λ_{nj} by

$$\widehat{\Lambda}_{nj}(t; b) = \int_0^t I\{\widehat{K}_{nj}(s) > cn^\gamma\} \frac{\widehat{L}_{nj}(ds)}{\widehat{K}_{nj}(s)}$$

and

$$\widehat{\Lambda}_{nj}(t) = \int_0^t I\{\widehat{K}_{nj}(s) > cn^\gamma\} \frac{\widehat{L}_{nj}(ds)}{\widehat{K}_{nj}(s)},$$

where the integrand is weighted by the $I\{\widehat{K}_{nj}(s) > cn^\gamma\}$ for $c > 0$ and $0 \leq \gamma < 1=2$ to avoid instability caused by non-monotonic weighted risk sets $\widehat{K}_{nj}$.

Hence, the first order Volterra integral equation (2.7) defined for R with $a_i(t)$ as in (2.6) and each b , $0 \leq j \leq p$, can be estimated by the following equation in $\widehat{R}_{nj}(t; b)$:

$$\widehat{R}_{nj}(t; b) = \int_0^t \widehat{R}_{nj}(s; b) \widehat{\Lambda}_{nj}(ds) + \widehat{\Lambda}_{nj}(t) \quad (2.12)$$

which is solely based on quantities that are observable. As before, this equation is a first order Volterra integral equation in $\widehat{R}_{nj}$. Hence, using Theorem II.6.3 of Andersen, Borgan, Gill, Keiding (1992), the above equation has an explicit solution given by:

$$\widehat{R}_{nj}(t; b) = \frac{1}{\widehat{P}_{nj}(t)} \int_0^t \widehat{P}_{nj-}(s) \widehat{\Lambda}_{nj}(ds; b), \quad (2.13)$$

where

$$\widehat{P}_{nj}(v) = \prod_{s \leq v} (1 + \widehat{\Lambda}_{nj}(ds)), \quad v > 0. \quad (2.14)$$

for each $j = 0, 1, \dots, p$. This in turn yields a class of estimators for the baseline distribution function F as follows

$$\widehat{F}_{nj}(t; b) = \frac{\widehat{R}_{nj}(t; b)}{1 + \widehat{R}_{nj}(t; b)}, \quad j = 0, \dots, p.$$

Using strong law of large numbers and the continuity of the product integral transformation, it follows that

$$\widehat{R}_{nj}(t; b) \rightarrow R_j(t; b) \text{ as } n \rightarrow \infty$$

with probability 1, uniformly in t and b , each in compact sets. Also, under certain minimal conditions (to be specified later) $R_j(t; b) = R(t)$ if and only if $b = \beta$. This motivates our estimation procedure for the regression parameter β . We choose that value of b as an estimator for β which minimizes an L_2 -discrepancy between $\widehat{F}_{nj}$ functions and $\widehat{F}_{n0}$. In particular, we can use the distance

$$D_n(b) = \int_0^\infty \sum_{j=1}^p (\widehat{F}_{nj}(t; b) - \widehat{F}_{n0}(t; b))^2 \Phi_n(dt; b), \quad (2.15)$$

with some appropriate (possibly) data-dependent integrating measure Φ_n . In fact, by varying the integrating measure Φ_n , we have defined a class of estimators $\hat{\beta}$ for β , i.e.

$$\widehat{\beta} = \operatorname{argmin}_b D_n(b). \quad (2.16)$$

These type of minimum distance (MD) estimators have been shown to have nice robustness properties in parametric setup with no censoring. When the underlying distribution has some known parametric form and when there is no censoring, such minimum L_2 distance estimators are usually asymptotically robust (Beran 1977, 1982). Also, the minimum L_2 distance estimators are highly efficient at some familiar parametric models for proper choices of the integrating measures. (Koul and De Wet 1983; Parr and De Wet 1981). Recently Yang and Prentice (1999) and Sundaram (2006b) studied these estimators for various semi-parametric models with censored or truncated data.

3 Large Sample Properties

In this section we establish the strong consistency and weak convergence of the proposed estimators $\hat{\beta}$ for the regression parameter β and for $\hat{R}_{nj}$ the estimators of the baseline odds function R under sufficient conditions listed below:

- A0** (T_i, Y_i, Z_i) are independent replicates of (T_1, Y_1, Z_1) , and conditional on $Z = z$, T and Y are independent and have distribution functions F_z and G_z . The support of each of the distributions F_z and G_z is $(0, \infty)$. The distribution F is continuous with continuous density.
- A1** Z is nonnegative and bounded with distribution H , $\beta \in N(\beta)$, a compact neighborhood in R^p .
- A2** $\Phi(t, b) = \lim_n \Phi_n(t; b)$ exists uniformly for $b \in N(\beta)$ and $t \in [0, \tau]$, where $\tau < \infty$ and Φ, Φ_n are finite measures on $[0, \tau]$.
- A3** $\alpha_Z = P_Z(T \geq Y) > 0$ almost everywhere Z .

The assumption A0 is the standard i.i.d. assumption about the underlying processes. The assumption A1 concerning non-negativity of the covariates can in fact handle all bounded covariates by just adding sufficiently large value. Assumption A2 allows the integrating measures Φ_n in the definition (2.15) to be data dependent. The last assumption is required as otherwise data may be unobservable for certain values of covariates.

For t in $[0, \tau]$ and b in $N(\beta)$, define

$$\begin{aligned} L_j(t) &= \int_0^t \frac{1}{\alpha_z} z_j G_z(s) dF_z(s) dH(z) \\ L_{\gamma j}(t; b) &= \int_0^t \frac{1}{\alpha_z} z_j e^{z'b} G_z(s) dF_z(s) dH(z) \\ K_j(t) &= \int_{\alpha_z} \frac{1}{z_j} G_z(t) \bar{F}_z(t) dH(z) \end{aligned}$$

Using strong law of large numbers and the assumption A1 that the covariates are bounded, the weighted empirical processes $\hat{K}_{nj}$, $\hat{L}_{nj}$ and $\hat{L}_{nyj}$ defined in (2.9)–(2.11) converge uniformly in t and $b \in N(\beta)$. Denote by

$$\Lambda_j(t) = \int_0^t \frac{dL_j(s)}{K_j(s)} \text{ and } \Lambda_{yj}(t; b) = \int_0^t \frac{dL_{yj}(s; b)}{K_j(s)}.$$

Using Lemma 1, 3 and Theorem 1 of Lai and Ying (1991), we have

$$\sup_{0 \leq t \leq \tau} |\hat{\Lambda}_{nj}(t) - \Lambda_j(t)| = 0, \quad \text{a.s.} \quad (3.17)$$

$$\sup_{0 \leq t \leq \tau, b \in N(\beta)} |\hat{\Lambda}_{nyj}(t; b) - \Lambda_{yj}(t; b)| = 0, \quad \text{a.s.} \quad (3.18)$$

Furthermore, using the continuity of the product integral transformation, it follows that

$$\sup_{0 \leq t \leq \tau} |\hat{P}_{nj}(t) - P_j(t)| = 0, \quad \text{a.s.}, \quad (3.19)$$

where $P_j(t) = \exp\{-\Lambda_j(t)\}$. Hence,

$$\sup_{0 \leq t \leq \tau, b \in N(\beta)} |\hat{R}_{nj}(t; b) - R_j(t; b)| = 0, \quad \text{a.s.} \quad (3.20)$$

This together with assumption A2 implies that the distance (2.15) converges a.s. to $D(b)$ uniformly in $b \in N(\beta)$, where

$$D(b) = \sum_{j=1}^p \int_0^{\tau} (F_j(t; b) - F_0(t; b))^2 \Phi(dt; b).$$

Hence, assuming

$$D(b) \neq 0, \text{ for } b \notin N(\beta), \quad (3.21)$$

the proposed minimum distance estimator $\hat{\beta} \rightarrow \beta$, a.s. as $n \rightarrow \infty$. This condition is satisfied under k -sample setup or when the distribution of the truncating variable does not depend on the covariates. Hence, the following theorem holds.

Theorem 3.1—Suppose that assumptions A0-A3 and (3.21) hold. Let $\hat{\beta}$ be the minimum L_2 -distance estimator of β defined in (2.16). Then, with probability 1,

$$\widehat{\beta} \rightarrow \beta, \text{ as } n \rightarrow \infty.$$

Next, we will establish the asymptotic normality of $\sqrt{n}(\widehat{\beta} - \beta)$ in several steps.

3.1 Linear approximation of $\widehat{F}_n(t; b)$ near β

We will first establish a linear approximation of $\widehat{R}_{nj}(t; b)$ at $b = \beta$. Denote by

$$\widetilde{L}_{nyj}(t; b) = \sum_{i=1}^n Z_i Z_{ij} \exp(Z_i' b) I\{V_i \leq t \leq U_i\}$$

and by

$$\widehat{\Lambda}_{nyj}(t; b) = \int_0^t \frac{d\widetilde{L}_{nyj}(ds; b)}{\widehat{K}_{nj}(s)}.$$

Using Mean Value Theorem

$$\widehat{R}_{nj}(t; b) - \widehat{R}_{nj}(t; \beta) = \widetilde{R}_{nj}'(t; b^*)'(b - \beta), \quad j=0, 1, \dots, p, \quad (3.22)$$

where

$$\widetilde{R}_{nj}'(t; b^*) = \frac{1}{\widehat{P}_{nj}(t)} \int_0^t \widehat{P}_{nj-}(s) \widehat{\Lambda}_{nyj}'(ds; b^*)$$

and b^* is an element in the line segment joining b and β . Again, with minor modification of Lemma 1, 3 and Theorem 1 of Lai and Ying (1991), it can be shown that

$$\sup_{0 \leq t \leq \tau, b \in N(\beta)} |\widehat{\Lambda}_{nyj}'(t; b) - \Lambda_{yj}'(t; b)| = 0, \text{ a.s.}$$

This in turn with (3.19) implies that

$$\sup_{0 \leq t \leq \tau, b \in N(\beta)} |\widetilde{R}_{nj}'(t; b) - R_j'(t; b)| = 0, \text{ a.s.} \quad (3.23)$$

Furthermore, using $F_j'(t; b) = \overline{F}_j^2(t; b) R_j'(t; b)$ and combining (3.22) and (3.23), we have almost surely

$$\sup_{0 \leq t \leq \tau, \|b-\beta\| \rightarrow 0} \left\| (\widehat{F}_{nj}(t; b) - \widehat{F}_{nj}(t; \beta)) - (\widehat{F}_{n0}(t; b) - \widehat{F}_{n0}(t; \beta)) \right\| = \overline{F}_j^2(t; \beta) (R'_j(t; \beta) - R'_0(t; \beta) + o(1))(b - \beta), \quad (3.24)$$

3.2 Martingale representation for $\widehat{R}_{nj}$

Next, we will present a martingale representation for the process $\widehat{R}_{nj}(t; b)$. Under random truncation,

$$M_i(t) = I\{U_i \leq t\} - \int_0^t I\{V_i \leq s \leq U_i\} d\Lambda_i(s), i=1, \dots, n \quad (3.25)$$

are martingales (Keiding and Gill (1990)). Denote by

$$\tilde{\Lambda}_{nj}(t) = \int_0^t I\{\widehat{K}_{nj}(s) \geq cn^\gamma\} d\Lambda_{nj}(s), \quad \tilde{\Lambda}_{nj\gamma}(t) = \int_0^t I\{\widehat{K}_{nj}(s) \geq cn^\gamma\} \Lambda_{nj\gamma}(ds; b). \quad (3.26)$$

Further define by

$$M_{nj}(t) = \int_0^t R(s) I\{\widehat{K}_{nj}(s) \geq cn^\gamma\} d\widehat{\Lambda}_{nj}(s) - \int_0^t R(s) I\{\widehat{K}_{nj}(s) \geq cn^\gamma\} d\Lambda_{nj}(s) + \int_0^t I\{\widehat{K}_{nj}(s) \geq cn^\gamma\} d\widehat{\Lambda}_{nj\gamma}(s) - \int_0^t I\{\widehat{K}_{nj}(s) \geq cn^\gamma\} d\Lambda_{nj\gamma}(s).$$

This can be rewritten as

$$M_{nj}(t) = \int_0^t \frac{I\{\widehat{K}_{nj}(s) \geq cn^\gamma\}}{\widehat{K}_{nj}(s)} \sum_{i=1}^n Z_{ij}(R(s) + \exp(Z_i' \beta)) dM_i(s).$$

Thus $M_{nj}(t)$ is also a martingale. Recall that

$$\widehat{R}_{nj}(t; \beta) = \int_0^t \widehat{R}_{nj}(s; \beta) d\widehat{\Lambda}_{nj}(s) + \widehat{\Lambda}_{nj\gamma}(t; \beta).$$

Then,

$$\widehat{R}_{nj}(t; \beta) - R(t) = \int_0^t (\widehat{R}_{nj}(s; \beta) - \widehat{R}_{nj}(s)) d\widehat{\Lambda}_{nj}(s) + M_{nj}(t) + \int_0^t R(s; \beta) d(\tilde{\Lambda}_{nj}(s) - \Lambda_{nj}(s)) + (\tilde{\Lambda}_{nj\gamma}(t) - \Lambda_{nj\gamma}(t)).$$

Using arguments similar to Theorem 5(i) of Lai and Ying (1991), the last two terms in the above equation is each $o_P(n^{-1/2})$, uniformly in $t \in [0, \tau]$. Hence,

$$\widehat{R}_{nj}(t; \beta) - R(t) = \int_0^t (\widehat{R}_{nj}(s; \beta) - \tilde{R}_{nj}(s)) d\widehat{\Lambda}_{nj}(s) + M_{nj}(t) + o_p(n^{-1/2}).$$

Hence, $\widehat{R}_{nj}(t; \beta) - R(t)$ can be viewed as solution of a Volterra integral equation of first kind in probability and has an explicit solution

$$\widehat{R}_{nj}(t; \beta) - R(t) = \frac{1}{\widehat{P}_{nj}(t)} \int_0^t P_{nj-}(s) dM_{nj}(s) + o_p(n^{-1/2}).$$

Furthermore, using standard martingale analysis

$$\sqrt{n}(\widehat{R}_{nj}(t; \beta) - R(t)) \Rightarrow \frac{1}{\alpha^{1/2} P_j(t)} W_j(t), j=1, \dots, p, \quad (3.28)$$

where $\alpha = \int P_z(T > Y) dH(z)$ is the truncation probability and $W_j(t)$ are zero-mean Gaussian process with mean 0 and covariance process

$$\langle W_j, W_k \rangle(t) = \int_{s \leq t} P_j(s) P_k(s) \frac{1}{K_j(s)} \frac{1}{K_k(s)} (R(s) K_{jk}(s) + K_{yjk}(s)) dR(s).$$

3.3 Asymptotic quadratic approximation of the distance $D_n(b)$ near β

Let

$$U_{nj}(t) = \overline{F}^2(t) (\widehat{R}_{nj}(t; \beta) - \widehat{R}_{n0}(t; \beta)), j=1, \dots, p,$$

and let $\Psi = (\psi_{jk})_{p \times p}$, where

$$\psi_{jk}(t) = \overline{F}^2(t) \frac{\partial}{\partial b_k} (R_j(t; \beta) - R_0(t; \beta)).$$

For any p -vector Δ ,

$$\tilde{D}(\Delta) = \sum_{j=1}^p \int_0^\tau (U_{nj}(t) + n^{-1/2} \psi_j \Delta)^2 \Phi(dt; \beta), \quad (3.29)$$

where ψ_j is the j -th row of Ψ and the integrating measure Φ is as in assumption A2. Then, using the asymptotic linearity of $\widehat{F}_{nj}$ (3.24) and the assumption A2 concerning the integrating measure Φ_n , it follows that

$$\sup_{\|\Delta\| \leq B} |D_n(\beta + n^{-1/2}\Delta) - \tilde{D}(\Delta)| = o_p(1).$$

Hence the minimizer of $D_n, \sqrt{n}(\hat{\beta} - \beta)$ is asymptotically equivalent to the minimizer of $\tilde{D}(\Delta)$.

Denote by $A = \left(\int_0^\tau \Psi'(t) \Psi(t) \Phi(dt; \beta) \right)$. Then,

$$\sqrt{n}(\hat{\beta} - \beta) = A^{-1} \int_0^\tau \sum_{j=1}^P \psi'_{j\cdot}(t) \sqrt{n} U_{nj}(t) \Phi(dt; \beta) + o_p(1). \quad (3.30)$$

Note that

$$\sqrt{n} U_{nj}(t) = \sqrt{n}(\hat{R}_{nj}(t; \beta) - R(t)) - \sqrt{n}(\hat{R}_{n0}(t; \beta) - R(t)).$$

Combining (3.28) with (3.30), the following holds:

$$\sqrt{n}(\hat{\beta} - \beta) \Rightarrow N(0, \Sigma)$$

where

$$\Sigma = A^{-1} B A^{-1}, \quad (3.31)$$

and $B = (b_{uv})_{p \times p}$ is given by

$$\begin{aligned} b_{uv} = & \int_0^\tau \int_0^\tau \sum_{j=1}^P \sum_{k=1}^P \psi_{ju}(s) \bar{F}^2(s) \psi_{kv}(t) \bar{F}^2(t) \left(\frac{\langle W_j, W_k \rangle(s \wedge t)}{\alpha P_j(s) P_k(t)} \right. \\ & - \frac{\langle W_j, W_0 \rangle(s \wedge t)}{\alpha P_j(s) P_0(t)} \\ & - \frac{\langle W_0, W_k \rangle(s \wedge t)}{\alpha P_0(s) P_k(t)} \\ & \left. + \frac{\langle W_0, W_0 \rangle(s \wedge t)}{\alpha P_0(s) P_0(t)} \Phi(ds; \beta) \Phi(dt; \beta) \right). \end{aligned}$$

Hence, we have the following result.

Theorem 3.2—Assume that A0-A3 and (3.21) are satisfied, and let $\hat{\beta}$ be the estimator of β defined in (2.16). Then

$$\sqrt{n}(\hat{\beta} - \beta) \Rightarrow N(0, \Sigma),$$

where Σ is as in (3.31).

Estimation of the Covariance—The covariance matrix Σ can be estimated consistently by replacing the unknown quantities by using the MD estimator $\hat{\beta}$ and the weighted empiricals defined in (2.9)–(2.11) and functions thereof, as well as the consistent estimate for the truncating probability α . This is explained in more detail below. Recall from Section 2.1,

$$\widetilde{L}_{nyj}(t;b) = \sum_{i=1}^n Z_i Z_{ij} \exp(Z_i' b) I\{V_i \leq t \leq U_i\}$$

and

$$\widehat{\Lambda}_{nyj}(t;b) = \int_0^t \frac{d\widetilde{L}_{nyj}(ds;b)}{\widehat{K}_{nj}(s)}, \quad \widetilde{R}_{nj}(t;b) = \frac{1}{\widehat{P}_{nj}(t)} \int_0^t \widehat{P}_{nj-}(s) \widehat{\Lambda}_{nyj}(ds;b).$$

Define:

$$\widehat{K}_{nj}(t) = \sum_{i=1}^n Z_{ij} Z_{ik} I\{V_i \leq t \leq U_i\}$$

and

$$\widehat{K}_{nyjk}(t) = \sum_{i=1}^n Z_{ij} Z_{ik} \exp(Z_i' b) I\{V_i \leq t \leq U_i\}.$$

Using these, we can define the consistent estimator $\widehat{\Psi}(t)$ for the $p \times p$ matrix $\Psi(t)$ as follows:

$$\widehat{\psi}_j(t;\widehat{\beta}) = (1 - \widehat{F}_{n0}(t;\widehat{\beta}))^2 \left(\widetilde{R}_{nj}(t;\widehat{\beta}) - \widetilde{R}_{n0}(t;\widehat{\beta}) \right),$$

where $\widehat{\psi}_j$ is the j -th column of $\widehat{\Psi}$. Furthermore, the truncation probability $\alpha = P(T \geq Y) = \int P_z(T \geq Y) dH(z)$ can be consistently estimated as follows:

$$\widehat{\alpha} = \sum_{i=1}^n \int_0^{\tau} \widehat{G}_{zi}(s) d\widehat{F}_{n0}(ds;\widehat{\beta}).$$

In case the distribution of the truncating variable Y is independent of the covariate process Z , $\widehat{G}_{zi}$ is simply the Lynden-Bell estimator for the distribution of the variable Y (which is right truncated by the variable T). However, in situations where the truncating variable depends on z , $\widehat{G}_z$ can be evaluated by the Lynden-Bell estimator from the observations in the window around z or its asymptotically equivalent version (Akritas and LaValley, 1997). Combining these, the matrix A can be consistently estimated by

$$\widehat{A} = \left(\int_0^\tau \widehat{\Psi}'(t) \widehat{\Psi}(t) \Phi_n(dt; \beta) \right)$$

and

$$\langle \widehat{W}_j, \widehat{W}_k \rangle(t) = \int_{s \leq t} \widehat{P}_{nj}(s) \widehat{P}_{nk}(s) \frac{1}{\widehat{K}_{nj}(s)} \frac{1}{\widehat{K}_{nk}(s)} (\widehat{R}_{n0}(s; \widehat{\beta}) \widehat{K}_{nj}(s) + \widehat{K}_{nj}(s; \widehat{\beta})) d\widehat{R}_{n0}(s; \widehat{\beta}).$$

Combining the estimators of α , A and the above, we have a consistent estimator for the asymptotic covariance matrix Σ .

Optimal Choice of Weight Functions—The performance of the minimum distance estimator depends on the integrating measure Φ_n . The role of the integrating measure Φ_n is similar to that of the weight function for the weighted log-rank statistics. Ideally Φ_n should be chosen in an optimal way, for example by minimizing the variance of $\widehat{\beta}$. However, in general it is difficult to find such an optimal integrating measure analytically as the asymptotic variance is a function of the covariates, as well as the survival function and the truncating distributions. In fact, in the special case of $\beta = 0$ and when the truncating variable is independent of the covariate process, analogous to Appendix E of Yang and Prentice (1999) one can obtain the optimal weight Φ_n as follows:

$$\Phi(dt) = \frac{d(K_1(t)(1 - F(t)))}{F(t)(1 - F(t))^2}$$

In practice, a simple choice of weight function is $\widehat{F}_{n0}(dt; b)$. This often results in highly efficient estimators in parametric setup based on complete data (Koul (2002)). Several other choices are considered and investigated in the simulation study in Section 4.

4 Numerical Study

In this section, we first present a simulation study investigating the performance of the proposed estimation procedure under various choice of integrating measures and for varying sample sizes and truncation probability. We generated the proportional odds model from the following model:

$$R_z(t) = te^{\beta z}$$

where the baseline odds function $R_0(t)$ was chosen to equal t , the regression coefficient was taken to be $\beta = -1, 0$ or 1 and the covariate $Z \sim U(0, 1)$. The truncating variable was generated from an exponential distribution with mean $= \mu$ and the constant μ was chosen appropriately to result in 10% or 30% truncation. The table below represents the bias and the sampling standard deviation of the estimate $\widehat{\beta}$ based on 5000 replications for sample size $n = 100, 200$ respectively. As we have defined a class of minimum distance estimators, we compared the performance of proposed estimators for various choices of integrating function Φ . Motivated by the optimal weight choice mentioned in previous section, we considered the optimal weight

$$d\widehat{\Phi}_n(t;b) = \frac{d(K_{nj}(t)(1 - \widehat{F}_{n0}(t;b)))}{\widehat{F}_{n0}(t;b)^d(1 - \widehat{F}_{n0}(t;b))^e}$$

Our numerical results showed that a slight modification of the optimal weight

$$d\widehat{\Phi}_n(t;b) = K_{n1}(t) \frac{d\widehat{F}_{n0}(t;b)}{\widehat{F}_{n0}(t;b)^d(1 - \widehat{F}_{n0}(t;b))^e}$$

performed better than the optimal weight. In fact, this finding is similar to the findings in Yang and Prentice (1999) for right censored data. Hence, motivated by this, we considered the following class of weights:

$$d\widehat{\Phi}_n(t;b) = K_{nj}(t)^c \frac{d\widehat{F}_{n0}(t;b)}{\widehat{F}_{n0}(t;b)^d(1 - \widehat{F}_{n0}(t;b))^e}, j=0, 1; c=0, 1; d, e=0, 1, 2.$$

We present below the comparison of the three best in the class of weights mentioned above and $\widehat{F}_{n0}(t)$. They are

$$\begin{aligned} \text{MDF: } d\widehat{\Phi}_n(t) &= d\widehat{F}_{n0}(t;b), \\ \text{MD1: } d\widehat{\Phi}_n(t) &= \frac{d\widehat{F}_{n0}(t;b)}{(1 - \widehat{F}_{n0}(t;b))}, \\ \text{MD2: } d\widehat{\Phi}_n(t) &= \frac{d\widehat{F}_{n0}(t;b)}{\widehat{F}_{n0}(t;b)(1 - \widehat{F}_{n0}(t;b))^2} \end{aligned}$$

and

$$\text{MD3: } d\widehat{\Phi}_n(t) = \widehat{K}_{n1}(t) \frac{d\widehat{F}_{n0}(t;b)}{\widehat{F}_{n0}(t;b)(1 - \widehat{F}_{n0}(t;b))^2}.$$

Note that the integrating measure MDF corresponds to using the distribution function $\widehat{F}_{n0}$, and hence gives weight to observed time points based on the best possible distribution estimate based on current estimate for the regression parameter b . The integrating measure for MD1 puts more weight on the right tail of the observed time values, same is true about the other two but they assign weight in a more complicated manner with the weight corresponding to MD3 being also weighted by the size of the risk sets. We investigated the performance of the proposed estimators based on 5000 replications. In the table below, we present the bias, sampling standard deviation (denoted as Samp. SD), estimate of the asymptotic standard deviation (Asym. SD) and the coverage probability (CP) based 5000 replications. The finite sample performance of the regression parameter estimators $\hat{\beta}$ seems to be reasonable for various integrating measures. As the sample size increases, the bias and the sampling standard deviation reduces for all the four estimators presented here. Furthermore, their performance seems reasonable even when the truncating percentage increases from 10% to 30%. Among the four estimators, MD3 seems to have the smallest bias and the sampling standard deviation for MD1 seems to be the smallest. However, the bias of MD1 seems to be larger in general than MD3. A good compromise between them is

MD2 which has reasonably small bias and sampling standard deviation. The estimate for the asymptotic standard deviation is reasonable, becoming closer to the sampling standard deviation as the sample size increases from $n = 100$ to $n = 200$. This indicates nice finite sample performance of the approximations for the asymptotic formula for the variance of the proposed regression estimators. The coverage probabilities become closer to the nominal level of 95% confidence level as the sample size increases from $n = 100$ to $n = 200$. Overall, the minimum distance approach seems to have very good finite sample property. The computational burden of this approach also is reasonable. In order to calculate the minimum distance estimators, we can use optimization procedures such as grid search, random search methods, simulated annealing or quasi-Newton approach which are easily available in commonly used statistical software like S-Plus/R or Matlab.

In order to demonstrate the proposed methodology, we applied our method to analyze a well-known data set on AIDS. This data set has been studied in Lagakos et. al. (1988) and in Kalbfleisch and Lawless (1989). This data set has information on 295 subjects who were diagnosed with AIDS before July 1, 1986 and reported before January 1, 1987 to Centers for Disease Control (CDC), Atlanta, Georgia. For these cases infection was attributed to a single transfusion or short series of transfusions. The data consists of time to infection with HIV virus, time to diagnosis of AIDS and the age of the subject. The question of interest is whether age at transfusion influences the induction time or incubation time, i.e., the lag time between the infection time and the diagnosis time. Since, only the subjects with diagnosis time before July 1, 1986 are included, the induction time is right truncated by difference between July 1, 1986 and the infection time. As had been done in Kalbfleisch and Lawless (1989), we treated age as categorical variable with three categories of age: 0–4, 5–59 and ≥ 60 . We fitted the proportional odds model with two covariates: $Z_1 = 1$ for individuals of age group 0–4 and $Z_1 = 0$ otherwise, and $Z_2 = 1$ for individuals of age group 5–59 and $Z_2 = 0$ otherwise. We fitted the following proportional odds model:

$$R_z(t) = R_0(t) e^{-Z_1 \beta_1 - Z_2 \beta_2}.$$

Note that the baseline odds function corresponds to the odds function for individuals who are at least 60 years of age. The table below presents the MD2 estimates, i.e. the MD estimators for the regression coefficients based on the integrating measure $\Phi_n(t) = \hat{F}_{n0}(dt; b)/\hat{F}_{n0}(t; b)(1 - \hat{F}_{n0}(t; b))^2$.

Age Group	Point Estimate	Standard Error	P-Value
0 – 4	2.7037	.5303	< .0001
5 – 59	–0.9307	.1931	< .0001

The coefficients indicate that the incubation times are smaller for children compared to the elderly subjects and that the incubation times for adults are larger than that of the elderly subjects. The Figure 1 shows the plot of the estimated distribution function for the incubation times for each of the three age groups based on the proportional odds model fitted above. Note the distribution looks similar to that given in Figure 3 of Kalbfleisch and Lawless (1989) showing agreement with their analysis.

References

1. Akritas, MG.; LaValley, MP. Statistical analysis with incomplete data: A selective review. In: Maddala, GS.; Rao, CR., editors. Handbook of Statistics. Amsterdam: Elsevier Science B. V; 1997. p. 551-632.
2. Andersen, PK.; Borgan, O.; Gill, RD.; Keiding, N. Statistical Models Based on Counting Processes. Springer; New York: 1991.
3. Bennett S. Analysis of Survival Data by the Proportional Odds Model. Statistics in Medicine. 1983a; 2:273–277. [PubMed: 6648142]
4. Bennett S. Log-Logistic Regression Models for Survival Data. Applied Statistics. 1983b; 32:165–171.
5. Beran RJ. Minimum Hellinger Distance Estimates for Parametric Models. The Annals of Statistics. 1977; 5:445–463.
6. Beran RJ. Robust Estimation in Models for Independent NonIdentically Distributed Data. The Annals of Statistics. 1982; 10:415–428.
7. Bickel, PJ. Papers on Semiparametric Models at the ISI Centenary Series. Amsterdam: 1986. Efficient Testing in a Class of Transformation Models; p. 63-81.
8. Cheng SC, Wei LJ, Ying Z. Analysis of Transformation Models with Censored Data. Biometrika. 1995; 82:835–845.
9. Dabrowska DM, Doksum KA. Estimation and Testing in a Two-Sample Generalized Odds-Rate Model. Journal of the American Statistical Association. 1988a; 83:744–749.
10. Dabrowska DM, Doksum KA. Partial Likelihood in Transformation Models With Censored Data. Scandinavian Journal of Statistics. 1988b; 15:1–23.
11. Fiegelson, ED.; Babu, GJ., editors. Statistical Challenges in Modern Astronomy. Springer; New York: 1992.
12. Grigoletto M, Akritas MG. Analysis of covariance with incomplete data via semiparametric model transformation. Biomterics. 1999; 55:1177–1187.
13. Gross ST, Lai TL. Nonparametric estimation and regression analysis with left truncated and right-censored data. Journal of the American Statistical Association. 1996; 91:1166–1180.
14. He S, Yang GL. Estimation of regression parameters with left truncated data. Journal of Statistical Planning and Inference. 2003; 117:99–122.
15. Huang J, Rossini AJ. Sieve estimation for the proportional-odds failure-time regression model with interval censoring. Journal of the American Statistical Association. 1997; 92:960–967.
16. Kalbfleisch JD, Lawless JF. Inference based on restrospective ascertainment: analysis of the data on transfusion-related aids. Journal of the American Statistical Association. 1989; 84:360–372.
17. Keiding N, Bayer T, Watt-Boolsen S. Confirmatory analysis of survival data using left truncation of the life times of primary survivors. Statistics in Medicine. 1987; 6:939–944. [PubMed: 3438619]
18. Keiding N, Gill R. Random Truncation and Markov Processes. The Annals of Statistics. 1990; 18:582–602.
19. Koul HL, DeWet T. Minimum distance estimation in a linear regression model. The Annals of Statistics. 1983; 11:921–932.
20. Koul, HL. Lecture notes in Statistics. Vol. 166. Springer-Verlag; New York: 2002. Weighted empirical processes in dynamic nonlinear models.
21. Lagakos SW, Barraj LM, DeGruttola V. Nonparametric analysis of truncated survival data, with application to AIDS. Biometrika. 1988; 75:515–523.
22. Lai TZ, Ying Z. Estimating a distribution function with truncated and censored data. The Annals of Statistics. 1991; 19:417–442.
23. Murphy SA, Rossinini AJ, Van der Waart AW. Maximum Likelihood Estimation in the Proportional Odds Model. Journal of the American Statistical Association. 1997; 92:968–976.
24. Parr WC, DeWet T. On minimum Cramér-von Mises-norm parameter estimation. Commuinactions in Statistics A – Theory and Methods. 1981; 10:1149–1166.

25. Pettitt AN. Proportional odds model for survival data and estimates using ranks. *Applied Statistics*. 1982; 33:169–175.
26. Rossini AJ, Tsiatis AA. A Semiparametric Proportional Odds Regression Model for the Analysis of Current Status Data. *Journal of the American Statistical Association*. 1996; 91:713–721.
27. Sundaram R. Semiparametric inference for the proportional odds model with time-dependent covariates. *Journal of Statistical Planning and Inference*. 2006a; 136:320–334.
28. Sundaram R. Estimation in two sample randomly truncated scale model. *Journal of Statistical Planning and Inference*. 2006b; 136:2983–3006.
29. Yang S, Prentice RL. Semiparametric Inference in the Proportional Odds Regression Model. *Journal of the American Statistical Association*. 1999; 94:125–136.
30. Wang M, Jewell N, Tsai W. Asymptotic properties of the product limit estimate under random truncation. *The Annals of Statistics*. 1986; 14:1597–1605.
31. Woodroffe M. Estimating a distribution function with truncated data. *The Annals of Statistics*. 1985; 13:163–177.

Figure 1.
Cumulative distribution function for incubation time for the three age groups.

Table 1

Simulation results based on 5000 replicates

$n = 100$		Truncation = 10%			Truncation = 30%		
		$\beta = 0$	$\beta = 1$	$\beta = -1$	$\beta = 0$	$\beta = 1$	$\beta = -1$
MDF: Bias		-.0371	-.0573	.0263	-.0317	-.0479	-.0367
Samp. SD		.5212	.5099	.4956	.5281	.5290	.5210
Asym. SD		.5618	.5504	.5209	.5920	.5725	.5913
CP		.9686	.9645	.9660	.9678	.9674	.9614
MD1: Bias		.0223	-.0261	.0253	.0106	-.0327	.0201
Samp. SD		.5002	.5031	.4967	.5230	.5259	.5266
Asym. SD		.5088	.5402	.5373	.5735	.5990	.5903
CP		.9629	.9658	.9636	.9560	.9615	.9750
MD2: Bias		.0230	-.0041	.0209	.0147	-.0244	.0120
Samp. SD		.5065	.5032	.5093	.5215	.5234	.5262
Asym. SD		.5266	.5282	.5400	.5786	.5684	.5469
CP		.9728	.9542	.9662	.9695	.9611	.9582
MD3: Bias		-.0097	-.0296	.0239	-.0182	-.0288	-.0100
Samp. SD		.5049	.5117	.4961	.5304	.5334	.5245
Asym. SD		.5310	.5662	.5122	.5735	.5797	.5664
CP		.9732	.9639	.9616	.9727	.9779	.9740
$n = 200$		Truncation = 10%			Truncation = 30%		
		$\beta = 0$	$\beta = 1$	$\beta = -1$	$\beta = 0$	$\beta = 1$	$\beta = -1$
MDF: Bias		-.0202	-.0349	.008	-.0280	-.0247	-.0249
Samp. SD		.4417	.4831	.4385	.4962	.5149	.4820
Asym. SD		.4337	.4940	.4747	.5267	.5146	.5056
CP		.9409	.9440	.9593	.9610	.9555	.9566
MD1: Bias		.0117	-.0167	.0174	-.0112	-.0305	.0138
Samp. SD		.4303	.4603	.4478	.4849	.5032	.4794
Asym. SD		.4201	.4549	.4759	.5070	.5197	.4781

$n = 100$	Truncation = 10%				Truncation = 30%			
	$\beta = 0$	$\beta = 1$	$\beta = -1$	$\beta = 0$	$\beta = 1$	$\beta = -1$	$\beta = 0$	$\beta = -1$
CP	.9523	.9494	.9549	.9412	.9436	.9441		
MD2: Bias	.0218	.0032	.0164	-.0011	-.0105	.0021		
Samp. SD	.4436	.4609	.4616	.4833	.4959	.4835		
Asym. SD	.4255	.4512	.4648	.4864	.5189	.5057		
CP	.9484	.9477	.9548	.9456	.9543	.9519		
MD3: Bias	-.0031	-.0158	.0063	-.0002	-.0156	-.0008		
Samp. SD	.4436	.4710	.4428	.4909	.5044	.4835		
Asym. SD	.4314	.4524	.4420	.5162	.5012	.4974		
CP	.9423	.9497	.9442	.9510	.9547	.9441		