Asymptotic efficiency of estimation in the partial Koziol–Green model

By: Haimeng Zhang

Zhang, H. (2004). Asymptotic effciency for estimation in the partial Koziol-Green model. *Journal of Statistical Planning and Inference, 124(2)*, 399 - 408. doi: 10.1016/S0378-3758(03)00209-X

Made available courtesy of Elsevier: http://www.dx.doi.org/10.1016/S0378-3758(03)00209-X

Abstract:

In this paper, the convolution theorem and the minimax theorem for estimating the survival function in the partial Koziol–Green model (PKG) are presented. The result indicates that the partial Abdushukurov–Cheng–Lin (ACL) estimator in the PKG model is asymptotically efficient in the sense of being the least dispersed regular estimator. Consequently, the calculation shows that the ACL estimator in the KG model is also asymptotically efficient.

Keywords: Partial ACL estimator | Asymptotic efficiency | Semiparametric model | Censoring | Survival function | Convolution theorem

Article:

1. Introduction

Let T be a non-negative random variable defined on some probability space $(\Omega, \mathscr{A}, P)$. Assume that T is subject to being right censored by the minimum of two independent and non-negative random variables $TC$and TD, where TC is an informative censoring time satisfying the proportional hazards condition (1) below and TD is a non-informative censoring time. More explicitly, we assume that one observes the pair (Z, Δ), where $Z=\min(T,TC,TD)$ and $\Delta=1$ if $T\leqslant TC\wedge TD$, 0 if $TC\leqslant TD\wedge T$, and -1 if $TD\leqslant T\wedge TC$. Throughout this paper,T,TC and TD are assumed to be independent and continuous random variables. We write $G(t)$,$C(t)$ and $D(t)$ for the distribution functions, $g(t)$,$c(t)$ and $d(t)$ for the density functions, and $\overline{G}(t), \overline{C}(t)$ and $\overline{D}(t)$ for the survival functions of T,TC and TD, respectively. In addition, we let $H(t)$ and $\overline{H}(t)$ be the respective distribution function and survival function of Z. The proportional hazards assumption on the censoring variable TC can be expressed as follows:

equation(1)

$$\overline{C}(t) = \overline{G}(t)^{\theta}, \quad t \geqslant 0,$$

for some fixed but unknown parameter $\theta>0$. This is the model of censoring introduced by Gather and Pawlitschko (1998) under the name of the partial Koziol–Green (PKG) model.

Assume that $(Z_1, \Delta_1),(Z_2, \Delta_2),\cdots,(Z_n, \Delta_n)$ are n independent copies of (Z, Δ). Let $E_i=\min(TC_i,TD_i)$ and $\xi_i=I[\Delta_i=1]$, the indicator function of the event $\{\Delta_i=1\}$; then, $Z_i=\min(T_i,E_i)$. The PKG model reduces to a general model of random right censoring (GRC), which suggests the Kaplan–Meier estimator (Kaplan and Meier, 1958) for the estimation of the survival function $\overline{G} = 1 - G$,

equation(2)

$$\hat{\overline{G}}_{\mathrm{KM}}(t) = \prod_{i: Z_i \leqslant t} \left(\frac{n - R_i}{n - R_i + 1} \right)^{\xi_i}, \quad t \geqslant 0,$$

where R_i denotes the rank of Z_i within the Z-sample. For notational simplicity, the dependence of the estimator on n will be suppressed throughout the rest of the paper. Here the subscript KM indicates the method of estimation used. Notice that, as the Kaplan–Meier estimator $\hat{\overline{G}}_{\mathrm{KM}}(t)$ does not take advantage of (1), one can deduce that (2) may not be an efficient estimator under the PKG model.

As a special case of the PKG model, we consider the Koziol–Green model (KG), in which $TD=\infty$. Under the KG model, (1) implies

equation(3)

$$\overline{G}(t) = \overline{H}(t)^{p}, \quad t \geqslant 0,$$

since $\overline{H}(t) = \overline{C}(t)\overline{G}(t)$. Here $p=1/(1+\theta)=P(\Delta=1)$. It should also be noted that under the framework of the GRC model, (1) holds if and only if Z and Δ are independent provided that $0<P(T\leqslant TC)<1$. Therefore, an estimator of $\bar{G}$ can be exploited based on (3). Abdushukurov (1984) and Cheng and Lin (1987) independently proposed the so-called ACL estimator equation(4)

$$\hat{\bar{G}}_{\mathrm{ACL}}(t) = (\hat{\bar{H}}_n(t))^{\hat{p}},$$

where $\hat{\bar{H}}_n(t) = (1/n)\sum_{i=1}^{n} I[Z_i > t]$ and $\hat{p} = (1/n)\sum_{i=1}^{n} \Delta_i$ are the empirical counterparts of $\bar{H}(t)$ and p, respectively. The asymptotic consistency and normality of the ACL estimator has been fully investigated, for example, see Cheng and Lin (1987). Furthermore, under the KG model, it was shown that the ACL estimator is asymptotically more efficient than the Kaplan–Meier estimator (Cheng and Lin, 1987).

Under the PKG model, Gather and Pawlitschko (1998) proposed the so-called partial ACL (PACL) estimator analogue to the ACL estimator above, which is given by

equation(5)

$$\hat{\bar{G}}_{\mathrm{PACL}}(t) = (\hat{\bar{K}}_{\mathrm{KM}}(t))^{\tilde{p}}, \quad t \geqslant 0,$$

where $\tilde{p} = \sum_{i=1}^{n} I[\Delta_i = 1] / \sum_{i=1}^{n} I[\Delta_i \neq -1]$, with $\tilde{p} = 0$ if the denominator is zero, and $\hat{\bar{K}}_{\mathrm{KM}}(t)$ is the Kaplan–Meier estimator based on i.i.d. data $(Z_1, \eta_1),(Z_2, \eta_2),\cdots,(Z_n, \eta_n)$ with $\eta_i=I[\Delta_i\neq-1]$, i.e.,

equation(6)

$$\hat{\bar{K}}_{\mathrm{KM}}(t) = \prod_{i:Z_i \leqslant t} \left(\frac{n-R_i}{n-R_i+1}\right)^{\eta_i}, \quad t \geqslant 0,$$

where R_i denotes the rank of Z_i in the Z-sample. Under the PKG model, the PACL estimator was shown to be strongly consistent and asymptotically normally distributed (Gather and Pawlitschko, 1998). Further analytic result shows that the PACL estimator is asymptotically more efficient than the KM estimator given by (2), see Gather and Pawlitschko (1998) for more details.

Naturally, it is always interesting to know if a proposed estimator uses the available data in an efficient manner. In a regular parametric model, the Cramer–Rao lower bound provides the answer in terms of a variance lower bound for estimators of unknown finite dimensional parameters. Under regularity conditions it is well known that the maximum likelihood estimator achieves this lower bound and so is asymptotically efficient. The PKG model is a semiparametric model, however, where θ is the parametric component while $\bar{G}(t)$ is the nonparametric component of interest. Therefore, the asymptotic theory for investigating the efficiency property of estimates in semiparametric models, as originated by LeCam (1979), Hájek (1970)and later generalized, for example, by Begun et al. (1983), may be applied.

In this paper, we provide an analysis to determine the efficiency of the PACL estimator. In Section 2, we present our convolution theorem and minimax theorem analogous to those established by Begun et al. (1983) for general semiparametric models. Our theorems assert, roughly, that the limiting process for *any* sequence of *regular* estimators of $\overline{G}(t)$ must be as least as dispersed as the limiting process corresponding to the PACL estimator (5), and so the PACL estimator is asymptotically efficient. Furthermore, our calculation also shows that the ACL estimator under the KG model is asymptotically efficient. Some concluding remarks are given in Section 3.

We would like to mention that a simpler proof of the asymptotic efficiency of ACL estimator based on a result of Van der Vaart (1991) was given by Hollander et al. (2001). Such a proof relies on the fact that Hadamand differentiable functions of efficient estimators are efficient. The same approach may be applied to prove the asymptotic efficiency of PACL estimator. However, the calculation based on the approach of Begun et al. (1983) in this paper provides more general results such as local minimax theorem of the PACL estimator, and the asymptotic lower bounds calculated give us some indication of how close or far the estimators considered are from efficient if the asymptotic lower bounds are not achieved. An additional example of this application can be seen in Zhang and Goldstein (2003) in the consideration of the asymptotic efficiency of the relative risk parameter of case-cohort sampling design in Cox's regression model.

2. Efficiency of the PACL estimator

In this section, we obtain the asymptotic variance lower bound for estimating the survival function $\overline{G}(t)$ and the unknown parameter θ under the PKG model. We closely follow the treatment of Begun et al. (1983), referred to as BHHW in what follows. However, we only sketch the main idea and provide necessary notations used in this paper. The reader should refer to BHHW for more details such as definitions and theorems.

Recall from Section 1 that $\theta \in R^+$ is an unknown positive-valued parameter. Let $g(t)$, the density function of$T$, be an element of $\mathscr{G}$, a fixed subset of the set of all densities absolutely continuous with respect to Lebesgue measure ν on $R^+=[0,\infty)$; let τ be counting measure on $\{-1,0,1\}$. Then the i.i.d. vectors $\mathbf{X}_i=(Z_i,\Delta_i)$, $i=1,2,\cdots,n$, which take values in the space $\mathscr{X}=R^+\times\{-1,0,1\}$, have density $f(\mathbf{x})=f(\mathbf{x};\theta,g)$, with respect to the product measure $\mu=\nu\times\tau$, given by

equation(7)

$$f(\mathbf{x})=(g(t)\overline{C}(t)\overline{D}(t))^{I[\delta=1]}(c(t)\overline{G}(t)\overline{D}(t))^{I[\delta=0]}(d(t)\overline{G}(t)\overline{C}(t))^{I[\delta=-1]}$$

$$=(g(t)\overline{G}^{\theta}(t)\overline{D}(t))^{I[\delta=1]}(\theta g(t)\overline{G}^{\theta}(t)\overline{D}(t))^{I[\delta=0]}(d(t)\overline{G}^{\theta+1}(t))^{I[\delta=-1]}$$

from (1). Here $\mathbf{x} = (t, \delta)$. Let $L^2(\mu) = L^2(\mathscr{X}, \mu)$ and $L^2(\nu) = L^2(R^+, \nu)$ denote the usual L^2-spaces of square integrable functions and let $\langle\cdot,\cdot\rangle_\mu$ ($\|\cdot\|_\mu$) and $\langle\cdot,\cdot\rangle_\nu$ ($\|\cdot\|_\nu$) denote the usual inner products (and norms) in $L^2(\mu)$ and $L^2(\nu)$, respectively. To compute the effective information for θ in the presence of the unknown function g, we need to parametrize $\mathscr{G}$ locally by a subspace $\mathscr{B}$ of $L^2(\nu)$, where each $\beta \in \mathscr{B}$ is a possible "direction" in which to approach g. Explicitly, for every $g \in \mathscr{G}$, let

equation(8)

$$\mathscr{B} \equiv \{\beta \in L^2(\nu): \|n^{1/2}(g_n^{1/2} - g^{1/2}) - \beta\|_\nu \to 0 \text{ as } n \to \infty \text{ for some } \{g_n\} \subset \mathscr{G} \text{ and } g_n \text{ absolutely continuous with respect to } g\},$$

which implies $\mathscr{B} = \{\beta \in L^2 : \beta \perp g^{1/2}, \text{support}(\beta) \subset \text{support}(g)\}$. Therefore,

Proposition 1.

The set $\mathscr{B}$ is a subspace of $L^2(\nu)$.

In addition, we let $\mathscr{C}_0(g, \beta) \equiv \{\{g_n\} \in \mathscr{G}$; g_n absolutely continuous with respect to g and $\|n^{1/2}(g_n^{1/2} - g^{1/2}) - \beta\|_\nu \to 0$ as $n \to \infty\}$. Furthermore, we let $\mathscr{C}_0(g)$ be the union of all $\mathscr{C}_0(g, \beta)$ over $\beta \in \mathscr{B}$. Similarly, let $\Theta(\theta, h)$ denote all sequences $\{\theta_n\}_{n \geqslant 1}$ such that $|n^{1/2}(\theta_n - \theta) - h| \to 0$, as $n \to \infty$, and $\Theta(\theta) = \bigcup_{h \in R^1} \Theta(\theta, h)$.
Given $(\theta_n, g_n)_{n \geqslant 1} \in \Theta(\theta) \times \mathscr{C}_0(g)$ let $f_n \equiv f(\cdot\,; \theta_n, g_n)$ denote the corresponding sequence of densities.

The following proposition is required for the computation of the asymptotic lower bounds for regular estimators of θ and $\overline{G}(t)$.

Proposition 2.

Suppose $(\theta, g) \in R^+ \times \mathscr{G}$. If $\{(\theta_n, g_n)\}_{n \geqslant 1} \in \Theta(\theta, h) \times \mathscr{C}_0(g, \beta)$ for $h \in R^1$, $\beta \in L^2(\nu)$, and $f_n \equiv f(\cdot\,; \theta_n, g_n)$ and $f \equiv f(\cdot\,; \theta, g)$, then under the PKG model, we have

equation(9)

$$\|n^{1/2}(f_n^{1/2} - f^{1/2}) - \alpha\|_\mu \to 0 \quad as \ n \to \infty$$

with $\alpha \in L^2(\mu)$ given by $\alpha = h\rho + A\beta$, and $\rho \in L^2(\mu)$ and $A : L^2(\nu) \to L^2(\mu)$ are given by

$$\rho(t, \delta) = (1/2)[\log \overline{G}(t) + I[\delta = 0]/\theta] f^{1/2}(t, \delta),$$

$$A\beta(t,\delta) = \left[(I[\delta = 1] + I[\delta = 0])\beta(t)g^{-1/2}(t) + \theta \frac{\int_t^\infty \beta g^{1/2}\,\mathrm{d}\nu}{\overline{G}(t)} + I[\delta = -1] \frac{\int_t^\infty \beta g^{1/2}\,\mathrm{d}\nu}{\overline{G}(t)} \right] f^{1/2}(t,\delta),$$

for $\delta=-1,0$ *or* 1 *and* $t>0$.

Proof.

Under the PKG model, the verification of (9) and the determination of α, ρ and A parallel computations in Section 6 of BHHW for the Cox's model case, and Lemma 1 of Begun and Wellner (1982) for the two-sample case without censoring. □

Let $H \equiv \{\alpha \in L^2(\mu) : \alpha = h\rho + A\beta$ for some $h \in R^1, \beta \in \mathscr{B}\}$, and hence by Proposition 1, H is a subspace of $L^2(\mu)$. For $\alpha \in H$, we let $\mathscr{F}(f,\alpha)$ denote the collection of all sequences $\{f_n\}$ such that (9) holds for the given α and let $\mathscr{F}(f) \equiv \bigcup_{\alpha \in H} \mathscr{F}(f,\alpha)$.

We first calculate the effective information for θ in the presence of the unknown function g. We orthogonally project ρ onto the nuisance space $\{A\beta : \beta \in \mathscr{B}\}$ to yield the "effective score" for θ, $\rho - A\beta^*$, where $A\beta^*$, the orthogonal projection, is such that β^* satisfies the "normal equation":

equation(10)

$$A^*A\beta^* = A^*\rho,$$

where A^* is the adjoint operator of A. The effective asymptotic information then equals

equation(11)

$$I_*(\theta) = 4\|\rho - A\beta^*\|_\mu^2.$$

Now we present the main results for this section. The proofs of the theorems are deferred to the end of this section.

Theorem 1.

Suppose that $\widehat{\theta}_n$ *is any regular estimator of* θ *under the PKG model, its limit law is* $\mathscr{L} = \mathscr{L}(f)$*. Then* $\mathscr{L}$ *may be represented as the convolution of a* $N(0, 1/I_*)$ *distribution with* $\mathscr{L}_1 = \mathscr{L}_1(f)$*, a distribution depending only on* $f = f(\cdot\,;\theta, g)$*, where*

equation(12)

$$I_* = (p_0 + p_1)/(\theta(\theta+1)^2).$$

Here $p_i = P(\Delta = i),\ i = 0,1$.

To present our asymptotic minimax result, we introduce a subconvex loss function $l: R^1 \to R^+$, that is, $\{x: l(x) \leqslant y\}$ is closed, convex, and symmetric for every $y \geqslant 0$, and $l(x)$ satisfies $\int_{-\infty}^{\infty} l(z)\phi(sz)\,\mathrm{d}z < \infty$ for all $s>0$, where ϕ denotes the standard normal density function.

Theorem 2.

Let $l(x)$ be a subconvex function. Then under the PKG model and with $B_n(c) \equiv \{f_n \in \mathscr{F}(f): n^{1/2}\|f_n^{1/2} - f^{1/2}\|_\mu \leqslant c\}$,

$$\lim_{c\to\infty} \lim_{n\to\infty} \inf_{\widehat{\theta}_n} \sup_{f_n \in B_n(c)} \mathbf{E}_{f_n} l(n^{1/2}(\widehat{\theta}_n - \theta_n)) \geqslant \mathbf{E}\, l(Z_*),$$

where $Z_* \sim N(0, 1/I_*)$ for $I_* = (p_1 + p_0)/(\theta(\theta+1)^2)$. If $l(x)=x^2$, then $1/I_*$ is the asymptotic lower bound for the variance of any regular estimator.

Here the infimum over estimators $\widehat{\theta}_n$ is taken over the class of “generalized procedures,” the class of randomized (Markov kernel) procedures, as in BHHW.

Next we consider the asymptotic lower bound for the estimation of $\bar{G}$, the continuous survival function corresponding to the density $g \in \mathscr{G}$, over an interval $[0,T_0]$ with $P(T>T_0)>0$ and $P(TD>T_0)>0$. First, a (continuous) estimator $\widehat{\mathbb{G}}_n$ of $\bar{G}$ is said to be *regular* at $f=f(\cdot\,;\theta,g)$ if, for every sequence $\{f_n\}=\{f(\cdot\,;\theta_n,g_n)\}$ with $\{(\theta_n, g_n)\} \in \Theta(\theta) \times \mathscr{C}_0(g)$, the process $n^{1/2}(\widehat{\mathbb{G}}_n - \overline{G}_n)$, with $G_n \equiv \int_0^\cdot g_n\,\mathrm{d}\nu$, converges weakly on $C[0,T_0]$ to limit process $\mathbb{S}: n^{1/2}(\widehat{\mathbb{G}}_n - \overline{G}_n) \Rightarrow \mathbb{S}$ on $C[0,T_0]$ (under f_n) where the law of $\mathbb{S}$ on $C[0,T_0]$ does not depend on h or β.

Analogous to Theorem 1, we have

Theorem 3.

Suppose that $\widehat{\mathbb{G}}_n$ is any regular estimator of $\bar{G}$ with $G = \int_0^\cdot g\,\mathrm{d}\nu$ under the PKG model with limit process $\mathbb{S}$. Then $\mathbb{S} =_d \mathbb{Z}_* + \mathbb{W}$ where the process $\mathbb{W}$ is independent of $\mathbb{Z}_*$, and $\mathbb{Z}_*$ is the zero-mean Gaussian process on $[0,T_0]$ *given by*

equation(13)

$$\mathbb{Z}_* \equiv \mathbb{Z} - Z_* \overline{G}(t)\log(\overline{G}(t))/(\theta+1), \quad 0 \leqslant t \leqslant T_0$$

with covariance function $K_*(s,t)$ given by

equation(14)

$$K_*(s,t) = K(s,t) + \frac{\theta}{p_0 + p_1} \overline{G}(t) \log(\overline{G}(t)) \overline{G}(s) \log(\overline{G}(s)), \quad 0 \leqslant s,\ t \leqslant T_0.$$

Here $\mathbb{Z}$ *is a zero-mean Gaussian process on* $[0,T_0]$ *with covariance function* $K(s,t)$ *given by*

equation(15)

$$K(s,t) = \frac{\overline{G}(s)\overline{G}(t)}{\theta + 1} \int_0^{s \wedge t} \overline{G}^{-(1+\theta)} \overline{D}^{-1} \frac{\mathrm{d}G}{\overline{G}}, \quad 0 \leqslant s,\ t \leqslant T_0.$$

and $Z_* \sim N(0, 1/I_*)$ *is independent of* $\mathbb{Z}$.

To present our minimax theorem for the estimator of $\overline{G}(t)$, we let $l_1 : C[0, T_0] \to R^+$ be a subconvex loss function such as $l_1(x) \equiv ||x|| \equiv \sup_t |x(t)|,\ l_1(x) = \int |x(t)|^2 \, \mathrm{d}t$, or $l_1(x)=I[x:||x||\geqslant c]$.

Theorem 4.

Let $l_1(x)$ *be a subconvex loss function defined above. Then under the PKG model and with* $Bn(c)$ *defined in Theorem* 2,

$$\lim_{c \to \infty} \lim_{n \to \infty} \inf_{\widehat{\mathbb{G}}_n} \sup_{f_n \in B_n(c)} \mathbf{E}_{f_n} l_1(n^{1/2}(\widehat{\mathbb{G}}_n - \overline{G}_n)) \geqslant \mathbf{E}\, l_1(\mathbb{Z}_*)$$

where $\mathbb{Z}_*$ *is the zero-mean Gaussian process given in Theorem* 3.

Here again the infimum over estimators $\widehat{\mathbb{G}}_n$ is taken over the class of “generalized procedures,” the class of randomized (Markov kernel) procedures, as in BHHW. We say that a regular estimator $\widehat{\mathbb{G}}_n$ of $\bar{G}$ is *asymptotically efficient* if the asymptotic covariance function of $\widehat{\mathbb{G}}_n$ is $K_*(s,t)$. From Gather and Pawlitschko (1998), the PACL process $\{\widehat{\mathbb{U}}_n \equiv \sqrt{n}(\widehat{\overline{\mathbb{G}}}_{\mathrm{PACL}}(t) - \overline{G}(t))\}$ converges weakly to a zero-mean Gaussian process with covariance function $K_*(s,t)$ given by (14), in other words, $\widehat{\mathbb{U}}_n$ is a sequence of estimators for which $\mathbb{W} = 0$ in Theorem 3. Note that the PACL estimator is not continuous and thus not regular, but as in Wellner (1982), we can always construct a continuous estimator $\widehat{\mathbb{G}}_n^c$ from the continuous “lower linear interpolation” $\widehat{\overline{K}}_n^c$ of $\widehat{\overline{K}}_{\mathrm{KM}}$, yielding corresponding process $\widehat{\mathbb{U}}_n^c$. Therefore, as in Wellner (1982), since the jump sizes of $\widehat{\overline{K}}_{\mathrm{KM}}$ tend to zero in probability, we have $||\widehat{\mathbb{U}}_n - \widehat{\mathbb{U}}_n^c||_0^{T_0} = o_P(1)$, and $\widehat{\mathbb{U}}_n^c$ also converges weakly to $\mathbb{Z}_*$. We still use PACL for this continuous estimator. Hence, we have

Corollary 1.

The PACL estimator under the PKG model is asymptotically efficient.

Similar calculation and consideration apply to the derivation of the asymptotic lower bounds for estimating θ and $\overline{G}(t)$ under the KG model, and therefore, one can also conclude

Corollary 2.

The ACL estimator under the KG model is asymptotically efficient.

We are now back to the proofs of theorems.

Proofs of Theorem 1 **and** Theorem 2**.**

Our proofs parallel those of Theorems 3.1 and 3.2 of BHHW. We have verified the subspace condition of BHHW in Proposition 1, and the conclusion of Proposition 2.1 of BHHW in Proposition 2. Therefore, it remains only to compute I_*.

First, following the notations in Proposition 2, we compute $\beta^*(t)$, the solution of “normal equation” (10), and so the orthogonal projection $A\beta^*(\mathbf{x})$. Note that with classical functional analysis theory (cf. Luenberger, 1969) and straightforward calculation, we have

$$A^* A\beta(t) = (\theta+1)\left[R\beta(t)\overline{G}^{\theta}(t)\overline{D}(t) - \int_0^t R\beta(s)\overline{G}^{\theta}(s)\overline{D}(s)\frac{\mathrm{d}G}{\overline{G}}\right]g^{1/2}(t),$$

$$(A^* A)^{-1}\beta(t) = (1/(\theta+1))\left[R\beta(t)\overline{G}^{-\theta}(t)\overline{D}^{-1}(t) - \int_0^t R\beta(s)\overline{G}^{-\theta}(s)\overline{D}^{-1}(s)\frac{\mathrm{d}G}{\overline{G}}\right]g^{1/2}(t),$$

$$A^*\rho(t) = (1/2)g^{1/2}(t)\left[(\overline{G}^{\theta}(t)\overline{D}(t) - \int_0^t \overline{G}^{\theta-1}(s)\overline{D}(s)\,\mathrm{d}G\right],$$

where $R\beta(t) = \beta(t)g^{-1/2}(t) - \int_t^{\infty}\beta g^{1/2}\,\mathrm{d}\nu / \overline{G}(t)$. Hence we find that

equation(16)

$$\beta^*(t) = (A^* A)^{-1}A^*\rho(t) = \frac{1}{2(\theta+1)}(1+\log\overline{G}(t))g^{1/2}(t).$$

Therefore, with $\mathbf{x} = (t,\delta) \in \mathscr{X}$, the orthogonal projection $A\beta^*(\mathbf{x})$ of $\rho(\mathbf{x})$ onto the closed space $\{A\beta : \beta \in \mathscr{B}\}$ of $L^2(\nu)$ is given by

equation(17)

$$A\beta^*(t,\delta) = (1/2)[\log\overline{G}(t) + (1/(\theta+1))]f^{1/2}(t,\delta),$$

for δ=−1,0 or 1. Therefore, it is easy to get I_* from (11) after simple computation. Theorem 2 now follows by a direct application of Theorem 3.2 of BHHW with $I_* = (p_0 + p_1) / (\theta(\theta + 1)^2)$. □

Proofs of Theorem 3 **and** Theorem 4**.**

Our proofs parallel those of Theorem 4.1 and Theorem 4.2 of BHHW. We have verified the subspace condition of BHHW in Proposition 1, and the conclusion of Proposition 2.1 of BHHW in Proposition 2. In addition, from the proof of Theorem 1, we see that $(A^*A)^{-1}$ is bounded linear operator on $[0,T_0]$. Therefore, from Theorem 4.1 of BHHW, Theorem 3 holds with the covariance function $K_*(s,t)$ of $\mathbb{Z}_*$given by

$$K_*(s,t) = K(s,t) + (\theta+1)^{-2} I_*^{-1} \overline{G}(s)\overline{G}(t) \log \overline{G}(s) \log \overline{G}(t),$$

where

equation(18)

$$K(s,t) = \langle (I_{[0,s]} - G(s))g^{1/2}, (A^*A)^{-1}(I_{[0,t]} - G(t))g^{1/2} \rangle_\nu, \quad 0 \leqslant s, \ t \leqslant T_0.$$

Hence, it remains only to compute $K(s,t)$.

Notice that $R(I_{[0,t]}(\cdot) - G(t))g^{1/2}(\cdot) = I_{[0,t]}(\cdot)\overline{G}(t) / \overline{G}(\cdot)$ and

$$(A^*A)^{-1}(I_{[0,t]} - G(t))g^{1/2} = \frac{\overline{G}(t)}{\theta+1}\left[\frac{I_{[0,t]}}{\overline{G}^{\theta+1}\overline{D}} - \int_0^{\cdot} \frac{I_{[0,t]}}{\overline{G}^{\theta+1}\overline{D}}\frac{\mathrm{d}G}{\overline{G}}\right] g^{1/2}.$$

Therefore, for $0 \leqslant s, \ t \leqslant T_0$, $K(s,t)$ of (18) becomes

$$K(s,t) = \frac{\overline{G}(t)}{\theta+1}\left(\int_0^\infty I_{[0,t]} I_{[0,s]} \frac{\mathrm{d}G}{\overline{G}^{\theta+1}\overline{D}} - \int_0^\infty I_{[0,s]}\left\{\int_0^{\cdot} \frac{I_{[0,t]}}{\overline{G}^{\theta+1}\overline{D}}\frac{\mathrm{d}G}{\overline{G}}\right\}\mathrm{d}G\right)$$

$$= \frac{\overline{G}(t)}{\theta+1}\left(\int_0^{s\wedge t} \frac{\mathrm{d}G}{\overline{G}^{\theta+1}\overline{D}} - \int_0^{s\wedge t}(\overline{G} - \overline{G}(s))\frac{1}{\overline{G}^{\theta+1}\overline{D}}\frac{\mathrm{d}G}{\overline{G}}\right) = \frac{\overline{G}(s)\overline{G}(t)}{1+\theta}\int_0^{s\wedge t}\frac{1}{\overline{G}^{\theta+1}\overline{D}}\frac{\mathrm{d}G}{\overline{G}}.$$

Hence, noting that $p_0 = \theta p_1$ and I_* given in (12), we can get (14). Theorem 4 now follows by a direct application of Theorem 4.2 of BHHW. □

Proof of Corollary 2**.**

Under the KG model, one can easily see that Theorem 1, Theorem 2, Theorem 3 and Theorem 4 hold with $\overline{D}(t) \equiv 1$. Therefore, simple simplification gives, $0 \leqslant s < t \leqslant T_0$,

$$K(s,t) = (1+\theta)^{-2}\overline{G}(s)^{-\theta}\overline{G}(t)(1 - \overline{G}^{\theta+1}(s))$$

and

$$K_*(s,t) = K(s,t) + \theta \overline{G}(t)\log(\overline{G}(t))\overline{G}(s)\log(\overline{G}(s)),$$

since $p_0+p_1=1$ under the KG model. From Cheng and Lin (1987), notice that the asymptotic covariance of the ACL process $\{\sqrt{n}(\widehat{\overline{\mathbb{G}}}_{\mathrm{ACL}}(t) - \overline{G}(t))\}$ is $K_*(s,t)$, and therefore, the ACL estimator is asymptotically efficient. □

3. Concluding remarks

This paper provides an analysis of asymptotic efficiencies of the ACL estimator under the KG model and the PACL estimator under the PKG model. Our analysis shows that both estimators are asymptotically efficient in the sense of being the least dispersed regular estimators. Throughout this analysis we assume the existence of the density functions of the life time *T* and the non-informative censoring time *TD*. Our analysis follows closely the treatment of Begun et al. (1983) on asymptotic efficiency in semiparametric models.

In addition, one can easily verify that the estimator of the unknown parameter θ under either model is actually asymptotically efficient, i.e., achieves the corresponding asymptotic lower bound. Furthermore, the asymptotic lower bound for the estimation of joint parameter $(\theta, \overline{G}(t))$ can be easily obtained following Begun et al. (1983).

In addition to the ACL estimator considered in the Koziol‑Green model, Hollander et al. (2001) proposed a generalized maximum likelihood estimator (GMLE) which maximizes the likelihood of the data over the class of allowable distributions including all continuous and discrete distributions. This GMLE has been observed to be different from the ACL estimator, empirically (Hollander et al., 2001) the difference is small for large risk sets. Therefore, it is expected that the GMLE estimator is asymptotically equivalent to the ACL estimator. In particular it was recently proven in Zhang et al. (2003) that the GMLE and the ACL estimators have the same asymptotic distribution. Hence, from Corollary 2, we may expect that the GMLE estimator in Koziol‑Green model is also asymptotically efficient, and that an analogous result will hold in the PKG model.

Acknowledgements

The author appreciates the perceptive and illuminating comments from two anonymous referees. A substantial credit goes to the referees for an overall improvement of the paper. Special thanks to Dr. Larry Goldstein for discussions and editing the manuscript, leading to the great improvement of the presentation of the paper.

References

Abdushukurov, A., 1984. On some estimates of the distribution function under random censorship, Conference of Young Scientists, VINITI No. 8756-V, Math. Inst. Acad. Sci., Uzbek SSR, Tashkent.

Begun, J.M., Wellner, J.A., 1982. Asymptotic e&ciency of relative risk estimates. In: Sen, P.K. (Ed.), Contributions to Statistics: Essays in Honor of Norman L. Johnson, North-Holland Pub. Co., New York.

Begun, J.M., Hall, W.J., Huang, W., Wellner, J.A., 1983. Information and asymptotic e&ciency in parametric-nonparametric models. Ann. Statist. 11, 432–452.

Cheng, P.E., Lin, G.D., 1987. Maximum likelihood estimator of a survival function under the Koziol–Green proportional hazards model. Statist. Probab. Lett. 5, 75–80.

Gather, U., Pawlitschko, J., 1998. Estimating the survival function under a generalized Koziol–Green model with partially informative censoring. Metrika 48, 189–209.

Hájek, J., 1970. A characterization of limiting distributions of regular estimates. Z. Wahrsch. verw. Gebiete. 14, 323–330

Hollander, M., Laird, G., Song, K.-S., 2001. Maximum likelihood estimation in the proportional hazards model of random censorship. Statistics 35, 245–258.

Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457–481.

LeCam, L., 1979. On a theorem of J. HOajek. In: Jureckova, J. (Ed.), Contributions to Statistics; Jaroslan Hájek Memorial Volume. Reidel, Dordrecht, pp. 119–135.

Luenberger, D.G., 1969. Optimization by Vector Space Methods. Wiley, New York.

Van der Vaart, A., 1991. E&ciency and hadamard diQerentiability. Scand. J. Statist. 18, 63–75.

Wellner, J.A., 1982. Asymptotic optimality of the product limit estimation. Ann. Statist. 10, 595–802.

Zhang, H., Goldstein, L., 2003. Information and asymptotic e&ciency of case-cohort sampling design in Cox's regression model. J. Multivariate Anal. 85, 292–317.

Zhang, H., Rao, M.B., Mitra, R., 2003. On the generalized maximum likelihood estimator of survival function under proportional hazards model, submitted for publication.