# 1. Introduction.

Let $x_1^0, \cdots, x_n^0$ be a sequence of positive random variables, usually representing survival times of $n$ patients in a clinical trial. Let $z_1, \cdots, z_n$ be their corresponding covariates. For a given $z$, assume that the hazard function of $x^0$ satisfy the proportional hazard model

$$\lambda(t|z) = \lambda_0(t)\exp(\beta_0 z), \tag{1.1}$$

where $\lambda_0$ is called the baseline hazard function. In practice, $x^0$'s are not directly observable. We observe $(x_i, \delta_i)$, $i = 1, \cdots, n$, where $x_i = \min(x_i^0, c_i)$, $\delta_i = 1_{[x_i = x_i^0]}$ and $c_i$ is the censoring variable. Assume that the censoring is non-informative, Cox (1972,1975) proposed to consider the partial likelihood

$$L_n(\beta) = \prod_{i=1,\cdots,n;\, x_i \le T} \left[\frac{\exp(\beta z_i)}{\sum_{j=1}^n \exp(\beta z_j) 1_{[x_j \ge x_i]}}\right]^{\delta_i}, \tag{1.2}$$

for inference on $\beta_0$, where $T$ satisfy $\Pr\{x_i \ge T\} > 0$. Let $\hat\beta$ be a number such that $L_n(\hat\beta) = \sup_\beta L_n(\beta)$, then we shall expect $\hat\beta$ to behave like that of a maximum likelihood estimate.

Assuming that $y_i = (z_i, x_i, \delta_i)$, $i = 1, \cdots, n$ are i.i.d. and some other regularity conditions, Tsiatis (1981) has shown that there exits a sequence $\hat\beta$ such that $\hat\beta$ converge to $\beta_0$ almost surely and that $\sqrt{n}(\hat\beta - \beta_0)$ converges in distribution to a normal variable with mean 0 and finite variance. The variance can be estimated by $(I_n(\hat\beta))^{-1}$, where $I_n$ is minus the second derivative of $\log L_n$. Let $U_n(\beta_0)$, the first derivative of $\log L_n$ at value $\beta_0$ (The score statistics). Then Gu (1991) has show that under some regularity conditions

$$\begin{aligned} U_n(\beta_0) &= \sum_{i=1}^n g(y_i) + \frac{1}{n}\sum_{i<j}\psi(y_i, y_j) + R_{1,n}, && (1.3)\\ I_n(\beta_0) &= n\sigma^2 + \sum_{i=1}^n h(y_i) + R_{2,n}, && (1.4)\\ I_n(\hat\beta) &= n\sigma^2 + \sum_{i=1}^n f(y_i) + R_{8,n}, && (1.5)\end{aligned}$$

where for any $\epsilon > 0$, $\Pr\left(|R_{1,n}| > n^{-1/2+\epsilon}\right) = o(n^{-1/2})$, $\Pr\left(|R_{j,n}| > n^{\epsilon}\right) = o(n^{-1/2})$, $j = 2, 8$ and

$$\begin{aligned} g(y_i) &= \int_0^T \left(z_i - \frac{\alpha_1(t)}{\alpha_0(t)}\right) dM_i(t), \qquad h(y_i) = h_1(y_i) + h_2(y_i),\\ h_1(y_i) &= \int_0^T \left(\alpha_2(t) - \frac{\alpha_1^2(t)}{\alpha_0(t)}\right)\frac{dM_i(t)}{\alpha_0(t)}, \qquad h_2(y_i) = \int_0^T \rho_i(t)\lambda_0(t)dt,\end{aligned}$$

$$
\begin{aligned}
\rho_i(t) &= \xi_{i,2}(t)-\alpha_2(t)-2\frac{\alpha_1(t)}{\alpha_0(t)}(\xi_{i,1}(t)-\alpha_1(t))+\frac{\alpha_1^2(t)}{\alpha_0^2(t)}(\xi_{i,0}(t)-\alpha_0(t)),\\
\psi(y_i,y_j) &= \psi_0(y_i,y_j)+\psi_0(y_j,y_i), \qquad \psi_0(y_i,y_j)=-\int_0^T \pi_i(t)dM_j(t), \qquad (1.6)\\
\pi_i(t) &= \frac{1}{\alpha_0(t)}(\xi_{i,1}(t)-\alpha_1(t))-\frac{\alpha_1(t)}{\alpha_0^2(t)}(\xi_{i,0}(t)-\alpha_0(t)).\\
f_1(y_i) &= \int_0^T\left(\frac{\alpha_2(t)}{\alpha_0(t)}-\frac{\alpha_1^2(t)}{\alpha_0^2(t)}\right)(\xi_{i,0}(t)-\alpha_0(t))\,\lambda_0(t)dt,
\end{aligned}
$$

where $M_i(t)=1_{[x_i\le t,\delta_i=1]}-\int_0^t\xi_{i,0}(s)\lambda_0(s)ds$, and for $k=0,1,2$, $\xi_{i,k}(t)=z_i^k 1_{[x_i\ge t]}e^{\beta_0 z_i}$. It is well known that $M_i(t)$, $i=1,\cdots,n$ are independent mean zero martingales with $\langle M_i\rangle(t)=\int_0^t\xi_{i,0}(s)\lambda_0(s)ds$ (see, for example, Gill, 1980).

In developing Edgeworth expansion for the statistics related to Cox regression model, we need to have explicit expressions for the quantities $\zeta$, $\mu$ and $\kappa_3$ which are defined by

$$
\begin{aligned}
\zeta &= E[g(y_1)h(y_1)],\\
\mu &= E[g(y_1)f_1(y_1)] \qquad (1.7)\\
\kappa_3 &= \frac{1}{\sigma^3}\left\{Eg^3(y_1)+3E[g(y_1)g(y_2)\psi(y_1,y_2)]\right\},
\end{aligned}
$$

where $\sigma^2=E[g^2(y_1)]$.

## 2. Result and Proof.

In this Section, we prove

**Theorem 2.1** *With the assumption discussed in section 1, we have*

$$
\begin{aligned}
\zeta &= \Delta_A-\Delta_B, \qquad (2.1)\\
\mu &= -\sigma_A^2+\sigma_B^2, \qquad (2.2)\\
\kappa_3 &= \frac{1}{\sigma^3}\left(-\Delta+3\zeta\right), \qquad (2.3)
\end{aligned}
$$

*where*

$$
\sigma^2 = \int_0^T\left(\alpha_2(t)-\frac{\alpha_1^2(t)}{\alpha_0(t)}\right)\lambda_0(t)dt
$$

$$
\begin{aligned}
\sigma_A^2 &= \int_0^T \left(\alpha_2(t) - \frac{\alpha_1^2(t)}{\alpha_0(t)}\right) \frac{\tilde{\alpha}_1(t)}{\alpha_0(t)} \Lambda_0(t)\lambda_0(t)dt, \\
\sigma_B^2 &= \int_0^T \left(\alpha_2(t) - \frac{\alpha_1^2(t)}{\alpha_0(t)}\right) \frac{\tilde{\alpha}_0(t)}{\alpha_0(t)} \Lambda_1(t)\lambda_0(t)dt, \qquad (2.4) \\
\Delta &= -\int_0^T \left(\alpha_3(t) - 3\frac{\alpha_2(t)\alpha_1(t)}{\alpha_0(t)} + 2\frac{\alpha_1^3(t)}{\alpha_0^2(t)}\right) \lambda_0(t)dt, \\
\Delta_A &= -\int_0^T \left(\tilde{\alpha}_3(t) - 2\frac{\tilde{\alpha}_2(t)\alpha_1(t)}{\alpha_0(t)} + \frac{\tilde{\alpha}_1(t)\alpha_1^2(t)}{\alpha_0^2(t)}\right) \Lambda_0(t)\lambda_0(t)dt. \\
\Delta_B &= -\int_0^T \left(\tilde{\alpha}_2(t) - 2\frac{\tilde{\alpha}_1(t)\alpha_1(t)}{\alpha_0(t)} + \frac{\tilde{\alpha}_0(t)\alpha_1^2(t)}{\alpha_0^2(t)}\right) \Lambda_1(t)\lambda_0(t)dt,
\end{aligned}
$$

*where for $k = 1, 2, 3$, $\alpha_k(t)$, $\tilde{\alpha}_k(t)$ and $\Lambda_k(t)$ are defined by*

$$
\begin{aligned}
\alpha_k(t) &= E\left[z^k e^{\beta_0 z} 1[x \geq t]\right], \\
\tilde{\alpha}_k(t) &= E\left[z^k e^{2\beta_0 z} 1[x \geq t]\right], \\
\Lambda_k(t) &= \int_0^t \frac{\alpha_k(t)}{\alpha_0(t)} \lambda_0(t)dt.
\end{aligned}
$$

**Lemma 2.1** *Let $\{H(t), H_1(t), H_2(t), 0 \leq t \leq T\}$ be left continuous processes $\{N_1(t), N_2(t), 0 \leq t \leq T\}$ be two independent counting processes with continuous compensators $A_1$ and $A_2$ and $\{\mathcal{F}_t, 0 \leq t \leq T\}$ be the filtration. Let $M_1 = N_1 - A_1$ and $M_2 = N_2 - A_2$. Then*

$$
\begin{aligned}
&E\left[\int_0^T H(t)dM_1(t)\right]^3 \\
&\quad = E\left[\int_0^T H^3(t)dA_1(t)\right] + 3E\left[\int_0^T H(t)dM_1(t) \int_t^T H^2(s)dA_1(s)\right] \qquad (2.5) \\
&E\left[\int_0^T H(t)dM_1(t) \int_0^T H_1(s)dM_2(s) \int_0^T H_2(u)dM_2(u)\right] \\
&\quad = E\left[\int_0^T H(t)dM_1(t) \int_t^T H_1(s)H_2(s)dA_2(s)\right] \qquad (2.6)
\end{aligned}
$$

*Proof of Lemma 2.1.* Let $D_1$ $D_2$ $D_3$ and $D_4$ be regions of $R^3$ defined by

$$
\begin{aligned}
D_1 &= \{(t, s, u) : t < s < u\} \\
D_2 &= \{(t, s, u) : t < s = u\} \\
D_3 &= \{(t, s, u) : t = s < u\} \qquad (2.7) \\
D_4 &= \{(t, s, u) : t = s = u\}
\end{aligned}
$$

It is easy to see that the left side of (2.5) can be broken into

$$6E[\int_{D_1} f] + 3E[\int_{D_2} f] + 3E[\int_{D_3} f] + E[\int_{D_4} f] \equiv I_1 + I_2 + I_3 + I_4$$

where

$$f = f(t,s,u) = H(t)H(s)H(u)dM_1(t)dM_1(s)dM_1(u).$$

First we consider the integration on $D_1$.

$$I_1 = E\left[\int_{t<s} H(t)H(s)dM_1(t)dM_1(s)E\{\int_s^T H(u)dM_1(u)|\mathcal{F}_s\}\right] = 0, \tag{2.8}$$

Use the same method, $I_2$ can be reduced to the second term on the right of (2.5) since

$$\begin{aligned} & E\left[\int_{s<t=u<T} H(t)H(u)dM_1(t)dM_1(u)|\mathcal{F}_s\right] \\ = \; & E\left[\int_{s<t<T} H^2(t)(dN_1(t) - 2dN_1(t)dA_1(t) + (dA_1(t))^2)|\mathcal{F}_s\right] \\ = \; & E\left[\int_{s<t<T} H^2(t)dN_1(t)|\mathcal{F}_s\right] \\ = \; & E\left[\int_{s<t<T} H^2(t)dA_1(t)|\mathcal{F}_s\right], \end{aligned} \tag{2.9}$$

where the second equality follows by the continuity of $A_1$ and the last equality holds since $A_1$ is the compensator of $N_1$. Similarly, $I_3 = 0$ and $I_4$ reduced to the first term on the right of (2.5).

To show (2.6), we use a similar argument. We note that the integration over the region $\{t > s \text{ and } t > u\}$, $\{s > t \text{ and } s > u\}$ and $\{u > t \text{ and } u > s\}$ are zero. Integration over the region $\{t = s > u\}$ and $\{t = u > s\}$ are also zero since $M_1$ and $M_2$ are independent. It remains to show that the integration over the region $\{t < s = u\}$ equals to the right side of (2.6). This is true by a derivation similar to (2.9). We only have to replace $H^2$ by $H_1H_2$.

*Proof of Theorem 2.1.* First of all, it is easy from standard martingale theory (Gill, 1980) that

$$\begin{aligned} \sigma^2 &= E[g(y_1)^2] = E\left[\int_0^T \left(z_1 - \frac{\alpha_1(t)}{\alpha_0(t)}\right)^2 \xi_{1,0}(t)\lambda_0(t)dt\right] \\ &= \int_0^T E\left[\xi_{1,2}(t) - 2\xi_{1,1}(t)\frac{\alpha_1(t)}{\alpha_0(t)} + \xi_{1,0}(t)\frac{\alpha_1^2(t)}{\alpha_0^2(t)}\right]\lambda_0(t)dt \end{aligned}$$

which is equivalent to the first expression of (2.4).

To show (2.1), first we have $\zeta = E[g(y_1)h(y_1)] = E[g(y_1)h_2(y_1)]$ since

$$\begin{aligned} E[g(y_1)h_1(y_1)] &= E\left[\int_0^T \left(z_1 - \frac{\alpha_1(t)}{\alpha_0(t)}\right)\left(\alpha_2(t) - \frac{\alpha_1^2(t)}{\alpha_0(t)}\right)\xi_{1,0}(t)\lambda_0(t)dt\right] \\ &= \int_0^T (\alpha_1(t) - \alpha_1(t))\left(\alpha_2(t) - \frac{\alpha_1^2(t)}{\alpha_0(t)}\right)\lambda_0(t)dt = 0. \end{aligned}$$

Let $H(s) = z_1 - \alpha_1(s)/\alpha_0(s)$. For $k = 0, 1, 2$, we have

$$\begin{aligned} E\left[(\xi_{1,k}(t) - \alpha_k(t))g(y_1)\right] &= E\left[\xi_{1,k}(t)\int_0^T H(s)dM_1(s)\right] \\ &= E\left[\xi_{1,k}(t)\int_0^{t-} H(s)dM_1(s)\right] \\ &= -E\left[\xi_{1,k}(t)\int_0^{t-} H(s)dA_1(s)\right] \qquad (2.10) \\ &= -E\left[\int_0^{t-}\left(z_1 - \frac{\alpha_1(s)}{\alpha_0(s)}\right)\xi_{1,k}(t)\xi_{1,0}(s)\lambda_0(s)ds\right] \\ &= -\int_0^t \left(\tilde{\alpha}_{k+1}(t) - \frac{\alpha_1(s)}{\alpha_0(s)}\tilde{\alpha}_k(t)\right)\lambda_0(s)ds \\ &= -\left(\tilde{\alpha}_{k+1}(t)\Lambda_0(t) - \tilde{\alpha}_k(t)\Lambda_1(t)\right). \end{aligned}$$

By the definition of $h_2$,

$$\begin{aligned} \zeta &= -\int_0^T \left(\tilde{\alpha}_3(t)\Lambda_0(t) - \tilde{\alpha}_2(t)\Lambda_1(t)\right)\lambda_0(t)dt \\ &\quad +2\int_0^T \frac{\alpha_1(t)}{\alpha_0(t)}\left(\tilde{\alpha}_2(t)\Lambda_0(t) - \tilde{\alpha}_1(t)\Lambda_1(t)\right)\lambda_0(t)dt \qquad (2.11) \\ &\quad -\int_0^T \frac{\alpha_1^2(t)}{\alpha_0^2(t)}\left(\tilde{\alpha}_1(t)\Lambda_0(t) - \tilde{\alpha}_0(t)\Lambda_1(t)\right)\lambda_0(t)dt \end{aligned}$$

which reduces to (2.1).

To show (2.2), we note that by definition

$$\begin{aligned} \mu &= E[g(y_1)f_1(y)] \\ &= \int_0^T \left(\frac{\alpha_2(t)}{\alpha_0(t)} - \frac{\alpha_1^2(t)}{\alpha_0^2(t)}\right)E[(\xi_{i,0}(t) - \alpha_0(t))\,g(y_1)]\lambda_0(t)dt, \end{aligned}$$

which is equivalent to (2.2).

To show (2.3), we first note that by (2.6) of Lemma 2.1

$$\begin{aligned} E[g(y_1)g(y_2)\psi(y_1,y_2)] &= 2E[g(y_1)g(y_2)\psi_0(y_1,y_2)] \\ &= -2E[\int_0^T H(t)dM_1(t)\int_0^T H_1(s)dM_2(s)\int_0^T H_2(u)dM_2(u)] \\ &= -2E\left[\int_0^T H(t)dM_1(t)\int_t^T H_1(s)H_2(s)\xi_{2,0}(s)\lambda_0(s)ds\right] \end{aligned} \tag{2.12}$$

where

$$H(t) = z_1 - \frac{\alpha_1(t)}{\alpha_0(t)}, \qquad H_1(t) = z_2 - \frac{\alpha_1(t)}{\alpha_0(t)}, \qquad H_2(t) = \pi_1(t)$$

The right side of (2.12) is zero since

$$E[H_1(s)H_2(s)\xi_{2,0}(s)|y_1] = H_2(s)E\left[\alpha_1(s) - \alpha_1(s)\right] = 0$$

and $H$ and $M_1$ are functions of $y_1$ only. Therefore $\kappa_3 = \sigma^{-3}E[g^3(y_1)]$. Applying (2.5), we have

$$\begin{aligned} E[g^3(y_1)] &= \int_0^T E\left[H^3(t)\xi_{1,0}(t)\right]d\lambda_0(t) + 3E\left[\int_0^T H^2(t)\xi_{1,0}(t)d\lambda_0(t)\int_0^t H(s)dM_1(t)\right] \\ &= -\Delta + 3\zeta, \end{aligned}$$

where the second term after the first equality is by changing the order of integration and the last equality is by similar argument in (2.11) and (2.10).

## References


Cox, D. R. (1972). Regression model and life tables (with discussion). J. Roy. Statist. Soc. Ser. B, 34, 187-220.

Cox, D. R. (1975). Partial likelihood. Biometrika, 62, 269-276.

Tsiatis, A. A. (1981). A large sample study of the Cox's regression model. Ann. Statist., 9, 93-108.

Gill, R.D. (1980). *Censoring and Stochastic Integrals.* Mathematisch Centrum Tract 124, Amsterdam.

Gu, M. (1991). On the Edgeworth expansion and bootstrap approximation for the Cox regression model under random censorship. Report from the Dept. of Math. and Stat., McGill University 91-09.

Technical Report 92-01

# Some Moments Concerning the Cox Regression Model

Minggao Gu[1]
Department of Mathematics and Statistics
McGill University

## Abstract

We calculate some moments concerning the score statistics related to the Cox regression model. A formular was develop to calculate the third moment of the martingale associated with a one jump counting process, which is essential in developing Edgeworth expansions for the statistics in the Cox regression model.