An estimation of the treatment effect for the right censored data

Hyo-Il Park1 · Ju Sung Kim2

1Department of Statistics, Chong-ju University
2Department of Informational Statistics, Chungbuk National University
Received 22 January 2011, revised 15 March 2011, accepted 28 March 2011

Abstract
In this article, we propose an estimation procedure for the treatment effect for the right censored data. We apply the least square method for deriving the estimation equation and obtain an explicit formula for an estimation. Then we consider some asymptotic properties with derivation of the asymptotic normality for the estimate. Finally we illustrate our procedure with an example and discuss some interesting aspects for the estimation procedure.

Keywords: Convolution, least square method, nonparametric method, two-sample problem.

1. Introduction
For the comparison study for a newly developed treatment over a control or between two treatments, one may take a test whether there exists any difference between them assuming a suitable model. If any, then one may be interested in measuring the difference. In the survival analysis, the most famous and widely used model is the proportional hazards model (Cox, 1972). Since the proportional hazards model defines the proportionality between two hazard functions, it would be inappropriate to describe the difference between two treatments directly. Then a well-known and useful one for this purpose would be the location translation model. In short, the location translation model assumes that the difference between two quantile points keeps constant over time. Then the difference between two quantile points can be considered as the difference between two treatments and has been called as the treatment effect.

The study on the estimation of the treatment effect for the two-sample problem based on the right censored data has been one of the research topics for a long time with the estimation for the mean-life in the survival analysis. Many authors have reported their research results with some drawback or drawbacks which may be inevitable because of the possibility of
censoring of the largest observation. We review some important results in the sequel and note that they are all based on the location translation model. Akritas (1986) proposed an estimate as a median of the convolution (Feller, 1971) of two Kaplan-Meier estimates (Kaplan and Meier, 1958). However a simple median of the convolution may lack the consistency since the Kaplan-Meier estimate should be incomplete when the largest observation is right censored. In order to overcome this inconsistency, Meng et al. (1991) and Bassiakos et al. (1991) have strived to modify Akritas’ estimate and proposed new estimates by introducing some artificial auxiliary variable, which made the procedures and forms of estimates very complicated and required additional information for the censoring distributions, which are of no interest in our concern. Tsiatis (1990) considered to use the linear rank statistics which may be used for testing the equality between two distributions or survival functions under the two-sample problem setting as the estimation functions. The resulting estimating procedure requires an iterative computation process and so the estimate does not have a closed form. Also Park and Park (1995) considered an estimate by integrating the difference between two quantiles. However the estimate may incur some efficiency loss by deleting or omitting some observations of the data. Zhou and Liang (2005) considered a procedure with assuming the distribution of control group is known but that of treatment group, unknown. However it is difficult to assume a specific distribution or survival function for the survival data. Thus the Zhou and Liang’s procedure has some intrinsic drawback for the real application.

In this study, we consider to propose a new nonparametric estimate procedure which is simple in calculation and easy to use. In the next section, we obtain a nonparametric estimate by applying the least square method. In section 3, we discuss some asymptotic properties for the estimate. In section 4, we illustrate our estimate with an example and discuss some interesting aspects for the estimation procedure as concluding remarks.

2. Least square estimate

We consider the following linear model.

\[ Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \cdots, n. \]  

(2.1)

The covariate \( x_i \) takes value 0 or 1 according as the nonnegative response variable \( Y_i \) comes from the control or treatment group. Without loss of generality, we assume that the first \( n_1, n_1 < n \) number of observations are assigned to the control group and the rest \( n_2 = n - n_1 \) number of observations, assigned to the treatment group. Also we assume that the distribution \( F \) of the error term \( \epsilon_i \) is unknown but continuous with mean 0 and finite variance, which is also unknown. Then we have the following results:

\[ E(Y_i) = \beta_0, \quad i = 1, \cdots, n_1 \]
\[ E(Y_i) = \beta_0 + \beta_1, \quad i = n_1 + 1, \cdots, n. \]

Therefore \( \beta_1 \) can be considered as the location translation parameter and should take on the role of the treatment effect. Based on this two-sample problem setting, a lot of research results for the estimation of \( \beta_1 \) have been proposed yet using the parametric and nonparametric methods when no censoring is involved. However in this study, we will consider the possibility of right censoring for \( Y_i \). For this let \( C_1, \cdots, C_n \) be the censoring random variables independent of \( Y_1, \cdots, Y_n \) with censoring distributions \( G_1 \) for \( i = 1, \cdots, n_1 \) and \( G_2 \)
for \( n_1 + 1 \leq i \leq n \). Then we only can observe that for each \( i, i = 1, \ldots, n \)

\[
T_i = \min \{ Y_i, C_i \} \quad \text{and} \quad \delta_i = I (Y_i \leq C_i), \tag{2.2}
\]

where \( I(\cdot) \) is an indicator function. Thus if \( T_i \) is censored then \( \delta_i = 0 \) and uncensored, \( \delta_i = 1 \). Under this random censoring scheme, by considering \( \beta_0 \) as a nuisance parameter, one may obtain an estimate of \( \beta_1 \) by finding \( \beta_1 \) which minimizes the following equation

\[
Q_M (\beta_0, \beta_1) = \sum_{i=1}^{n} (T_i - \beta_0 - \beta_1 x_i)^2 dF_n(T_i - \beta_0 - \beta_1 x_i). \tag{2.3}
\]

In (2.3), \( F_n \) is the Kaplan-Meier estimate of \( F \) based on

\[
T_1 - \beta_0 - \beta_1 x_1, \ldots, T_n - \beta_0 - \beta_1 x_n
\]

and \( dF_n(t) = F_n(t) - F_n(t-) \) is the jump size of \( F_n \) at \( t \). Therefore when \( \delta_i = 0 \), \( dF_n(T_i - \beta_0 - \beta_1 x_i) = 0 \) and \( dF_n(T_i - \beta_0 - \beta_1 x_i) > 0 \) if \( \delta_i = 1 \). This version of the application of the least square method was initiated by Miller (1976) and requires an iterative procedure for calculation since the value of \( F_n(T_i - \beta_0 - \beta_1 x_i) \) can vary with the values of \( \beta_0 \) and \( \beta_1 \). Also one may estimate \( \beta_1 \) using the method of Buckley and James (1979) by introducing the pseudo random variables for the censored observations and minimizing a similar equation with (2.3). However the use of an iterative procedure for calculation has been inevitable.

Now we note the following fact about the model (2.1). Under the model (2.1), for any two observations \( Y_i \) and \( Y_j, \ i = 1, \ldots, n_1 \) and \( j = n_1 + 1, \ldots, n \), we see that with the notation that \( \epsilon_{ij} = \epsilon_j - \epsilon_i \)

\[
Y_j - Y_i = \beta_1 + \epsilon_{ij}.
\]

We note that the distribution \( H \) of the difference \( Y_j - Y_i \) takes \( \beta_1 \) as its mean or median since the distribution of \( \epsilon_{ij} \) should be symmetric about 0. In order to proceed our discussions for \( H \) more concretely, let \( F_1 \) and \( F_2 \) be the distribution functions of \( Y_i \) and \( Y_j \), respectively. From (2.1) and the assumptions introduced up to now, the location translation model holds for \( F_1 \) and \( F_2 \) such as for any real number \( y \),

\[
F_2(y) = F_1(y - \beta_1).
\]

Also we note that \( H \) must be the convolution of \( F_1 \) and \( F_2 \) such that for any real number \( t, t \in (-\infty, \infty) \) we have

\[
H(t) = \Pr \{ Y_j - Y_i \leq t \} = \int_{0}^{\infty} F_2(t+y) dF_1(y).
\]

Let \( F_{1n_1} \) and \( F_{2n_2} \) be the Kaplan-Meier estimates of \( F_1 \) and \( F_2 \) based on \( (T_1, \delta_1), \ldots, (T_{n_1}, \delta_{n_1}) \) and \( (T_{n_1+1}, \delta_{n_1+1}), \ldots, (T_n, \delta_n) \), respectively. Then an estimate \( H_n \) of \( H \) can be obtained by using the convolution as follows. For any real number \( t \), we have that

\[
H_n(t) = \int_{0}^{\infty} F_{2n_2}(t+y) dF_{1n_1}(y), -\infty < t < \infty. \tag{2.4}
\]
As we already mentioned in the introduction, Akritas (1986) proposed a Hodges-Lehmann type of estimate by noting that \( \beta_1 \) is a median and \( H_n \) is an estimate of \( H \). However one may construct an estimating equation \( Q \) which is similar to \( Q_M \) in (2.3) using (2.4) based on the approach of Miller (1976) as follows.

\[
Q(\beta_1) = \sum_{i=1}^{n_1} \sum_{j=n_i+1}^{n} (T_j - T_i - \beta_1)^2 dH_n(T_j - T_i - \beta_1).
\]

By differentiating \( Q(\beta_1) \) with respect to \( \beta_1 \), we obtain that

\[
\frac{\partial Q(\beta_1)}{\partial \beta_1} = -2 \sum_{i=1}^{n_1} \sum_{j=n_i+1}^{n} (T_j - T_i - \beta_1) dH_n(T_j - T_i - \beta_1).
\]

Then by taking \( \partial Q(\beta_1)/\partial \beta_1 = 0 \), solving it with respect to \( \beta_1 \) and putting it \( \hat{\beta}_1 \), we may obtain a least square estimate \( \hat{\beta}_1 \) of \( \beta_1 \). At a first glance, an iterative procedure for calculation would be also required because the treatment effect \( \beta_1 \) is contained in the expression of \( H_n \).

However we will see that any iterative procedure would be unnecessary in the sequel. First of all, in order to obtain \( H_n(T_j - T_i - \beta_1) \) using (2.4), we have to have the Kaplan-Meier estimates \( F_{1n_1} \) and \( F_{2n_2} \) based on

\[
(T_1, \delta_1), \ldots, (T_{n_1}, \delta_{n_1}) \text{ and } (T_{n_1+1} - \beta_1, \delta_{n_1+1}), \ldots, (T_n - \beta_1, \delta_n) \tag{2.5}
\]

or

\[
(T_1 + \beta_1, \delta_1), \ldots, (T_{n_1} + \beta_1, \delta_{n_1}) \text{ and } (T_{n_1+1}, \delta_{n_1+1}), \ldots, (T_n, \delta_n). \tag{2.6}
\]

For any given \( \beta_1 \), let \( f_{1n_1}(T_i; \delta_i) \) and \( f_{2n_2}(T_j - \beta_1; \delta_j) \) be the jumps at \( T_i \) and \( T_j - \beta_1 \) for \( F_{1n_1} \) and \( F_{2n_2} \), which are obtained from (2.5). Also let \( f_{1n_1}(T_i + \beta_1; \delta_i) \) and \( f_{2n_2}(T_j; \delta_j) \) be the jumps at \( T_i + \beta_1 \) and \( T_j \) for \( F_{1n_1} \) and \( F_{2n_2} \), which are obtained from (2.6). Then from (2.4), we have that

\[
H_n(T_j - T_i - \beta_1) = \sum_{T_i - T_k - \beta_1 \leq T_j - T_i} f_{1n_1}(T_k; \delta_k) f_{2n_2}(T_i - \beta_1; \delta_i)
\]

or

\[
H_n(T_j - T_i - \beta_1) = \sum_{T_i - T_k - \beta_1 \leq T_j - T_i} f_{1n_1}(T_k + \beta_1; \delta_k) f_{2n_2}(T_i; \delta_i).
\]

Now we note for all \( i, i = 1, \ldots, n_1 \) that for any given \( \beta_1 \)

\[
f_{1n_1}(T_i + \beta_1; \delta_i) = f_{1n_1}(T_i; \delta_i)
\]

since the orders of \( T_i \) and \( T_i + \beta_1 \) are the same among \( T_1, \ldots, T_{n_1} \) and \( T_1 + \beta_1, \ldots, T_{n_1} + \beta_1 \). Also we may conclude the same for \( T_j - \beta_1 \) and \( T_j, j = n_1+1, \ldots, n \) such that for any given \( \beta_1 \)

\[
f_{2n_2}(T_j - \beta_1; \delta_j) = f_{2n_2}(T_j; \delta_j).
\]
Furthermore this tells us that for any \( i \) and \( j \) and for any given \( \beta_1 \)
\[
H_n(T_j - T_i) = H_n(T_j - T_i - \beta_1).
\]

Therefore the least square estimate, \( \hat{\beta}_1 \) can be proposed as follow.
\[
\hat{\beta}_1 = \frac{1}{H_n(D_n)} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} (T_j - T_i) dH_n(T_j - T_i), \tag{2.7}
\]
where \( D_n = \max \{ T_j - T_i : 1 \leq i \leq n_1, n_1 + 1 \leq j \leq n \} \). We note that any iterative procedure for calculation is not required. This is an advantage of our estimate and the calculation is also relatively simple.

For obtaining more explicit formular for \( \hat{\beta}_1 \) from (2.7), first of all, we note that for any pair of \( i \) and \( j \) such that \( f_{1n_1}(T_i; \delta_i) > 0 \) and \( f_{2n_2}(T_j; \delta_j) > 0 \), since
\[
H_n(T_j - T_i) = \sum \sum_{T_i - T_k \leq T_j - T_i} f_{1n_1}(T_k; \delta_k) f_{2n_2}(T_i; \delta_i),
\]
we have that
\[
dH_n(T_j - T_i) = f_{1n_1}(T_i) f_{2n_2}(T_j).
\]

Therefore \( \hat{\beta}_1 \) can be expressed as
\[
\hat{\beta}_1 = \frac{1}{H_n(D_n)} \sum_{i=1}^{n_1} \sum_{j=n_1+1}^{n} (T_j - T_i) f_{1n_1}(T_i) f_{2n_2}(T_j),
\]
\[
= \frac{1}{H_n(D_n)} \left\{ \sum_{i=1}^{n_1} f_{1n_1}(T_i) \sum_{j=n_1+1}^{n} T_j f_{2n_2}(T_j) - \sum_{i=1}^{n_1} T_i f_{1n_1}(T_i) \sum_{j=n_1+1}^{n} f_{2n_2}(T_j) \right\}
\]
\[
= \frac{F_{1n_1}(T_{(n_1)}^{1})}{H_n(D_n)} \sum_{j=n_1+1}^{n} T_j dF_{2n_2}(T_j) - \frac{F_{2n_2}(T_{(n_2)}^{2})}{H_n(D_n)} \sum_{i=1}^{n_1} T_i dF_{1n_1}(T_i), \tag{2.8}
\]
where \( T_{(n_1)}^{i} \) is the largest observation for each \( i, i = 1, 2 \). If the largest observations, \( T_{(n_1)}^{1} \) and \( T_{(n_2)}^{2} \) are not censored in both the samples, the expression in (2.8) can be much simpler.

For this case, since \( F_{1n_1}(T_{(n_1)}^{1}) = F_{2n_2}(T_{(n_2)}^{2}) = H_n(D_n) = 1 \),
\[
\hat{\beta}_1 = \sum_{j=n_1+1}^{n} T_j dF_{2n_2}(T_j) - \sum_{i=1}^{n_1} T_i dF_{1n_1}(T_i)
\]
\[
= T_{2n_2} - T_{1n_1}, \text{ say}.
\]
In other words, \( \hat{\beta}_1 \) implies the difference between the two sample means when both the largest observations are not censored. Therefore our estimate can be considered as an extension of the difference between two sample means for the right censored case. Also we note that the jump size of the Kaplan-Meier estimate is 0 for the censored observation. This means that
the largest observation when it is censored can not contribute its amount to \((2.8)\). For this
matter, let \(X\) be a non-negative random variable with the distribution function \(F\). Then we
note that
\[
E(X) = \int_0^\infty x dF(x) = \int_0^\infty (1 - F(x)) \, dx. \tag{2.9}
\]
For (2.9), you may refer to Chung (1974). Then in the light of (2.9), we may modify (2.8)
as
\[
\hat{\beta}_1 = \frac{F_{1n_1}(T_{(n_1)^i})}{H_n(D_n)} \int_0^{T_{(n_2)^i}} \{1 - F_{2n_2}(x)\} \, dx - \frac{F_{2n_2}(T_{(n_2)^i})}{H_n(D_n)} \int_0^{T_{(n_1)^i}} \{1 - F_{1n_1}(x)\} \, dx. \tag{2.10}
\]
Then we note that \(\hat{\beta}_1\) in (2.10) may include the effect of the largest observation even though
it may be censored.

3. Some asymptotic properties for the estimate

In this section, we discuss some asymptotic properties for \(\hat{\beta}_1\). First of all, by assuming
that for each \(i, i = 1, 2\)
\[
F_{1n_1}(T_{(n_1)^i}) \to^P 1 \tag{3.1}
\]
we see that \(\hat{\beta}_1\) is a consistent estimate of \(\beta_1\) since (3.1) guarantees that \(H_n(D_n) \to^P 1\),
where \(\to^P\) means the convergence in probability. However when either one of the conditions
in (3.1) is not satisfied, the contamination with any bias in \(\hat{\beta}_1\) would be inevitable. Now we
consider the asymptotic normality for
\[
\sqrt{n} \left( \hat{\beta}_1 - \beta_1 \right).
\]
For this, for each \(i, i = 1, 2\), let
\[
\tau_i = \inf \{ t : F_i(t) = 1 \}.
\]
Now we assume that
\[
\int_0^{\tau_i} \left[ \int_t^\infty \{1 - F_i(s)\} \, ds \right]^2 \frac{dF_i(t)}{\{1 - F_i(t-)^i\}^2 \{1 - G_i(t-)\}} < \infty, \tag{3.2}
\]
\[
\sqrt{n_i} \int_{T_{(n_1)^i}}^{\infty} \{1 - F_i(t)\} \, dt \to^P 0 \text{ as } n_i \to \infty \tag{3.3}
\]
and
\[
\sqrt{n_i} \left\{ \frac{F_{1n_1}(T_{(n_1)^i})}{H_n(D_n)} - 1 \right\} \to^P 0 \tag{3.4}
\]
Also we assume that
\[ n_1 / n_2 \to \lambda > 0 \text{ as } n \to \infty. \] (3.5)

**Theorem 3.1** Under the assumptions (3.2)-(3.5), the distribution of \( \sqrt{n} \left( \hat{\beta}_1 - \beta_1 \right) \) converges in distribution to a normal random variable with mean 0 and variance \( \sigma^2 \),

\[
\sigma^2 = (1 + \lambda^{-1}) \int_0^T \left[ \int_t^\infty \left\{ 1 - F_1(s) \right\} ds \right]^2 \frac{dF_1(t)}{\{1 - F_1(t-)\}^2 \{1 - G_1(t-)\}} + (1 + \lambda) \int_0^T \left[ \int_t^\infty \left\{ 1 - F_2(s) \right\} ds \right]^2 \frac{dF_2(t)}{\{1 - F_2(t-)\}^2 \{1 - G_2(t-)\}}.
\]

**Proof.** In order to prove this theorem, first of all, we note that
\[
\beta_1 = \int_0^\infty t dF_2(t) - \int_0^\infty t dF_1(t) = \int_0^\infty \{1 - F_2(t)\} dt - \int_0^\infty \{1 - F_1(t)\} dt.
\]
Then from (2.10), we have that
\[
\sqrt{n} \left( \hat{\beta}_1 - \beta_1 \right) = \sqrt{n} \left[ \frac{F_{1n_1}(T_{n_1}^1)}{H_n(D_n)} \int_0^{T_{n_2}^2} \left\{ 1 - F_{2n_2}(t) \right\} dt - \int_0^\infty \left\{ 1 - F_2(t) \right\} dt \right] - \sqrt{n} \left[ \frac{F_{2n_2}(T_{n_2}^2)}{H_n(D_n)} \int_0^{T_{n_1}^1} \left\{ 1 - F_{1n_1}(t) \right\} dt - \int_0^\infty \left\{ 1 - F_1(t) \right\} dt \right].
\]
Also we have that
\[
\sqrt{n_1} \left[ \frac{F_{2n_2}(T_{n_2}^2)}{H_n(D_n)} \int_0^{T_{n_1}^1} \left\{ 1 - F_{1n_1}(t) \right\} dt - \int_0^\infty \left\{ 1 - F_1(t) \right\} dt \right] = - \frac{F_{2n_2}(T_{n_2}^2)}{H_n(D_n)} \int_0^{T_{n_1}^1} \sqrt{n_1} \left\{ 1 - F_{1n_1}(t) - F_1(t) \right\} dt + \sqrt{n_1} \left\{ \frac{F_{2n_2}(T_{n_2}^2)}{H_n(D_n)} - 1 \right\} \int_0^{T_{n_1}^1} \left\{ 1 - F_1(t) \right\} dt - \sqrt{n_1} \int_0^{T_{n_1}^1} \left\{ 1 - F_1(t) \right\} dt.
\]
Then using the integration by parts (Shorack and Wellner, 1986), we have that
\[
- \int_0^{T_{n_1}^1} \sqrt{n_1} \left\{ F_{1n_1}(t) - F_1(t) \right\} dt = - \int_0^{T_{n_1}^1} \left\{ \sqrt{n_1} \frac{F_{1n_1}(t) - F_1(t)}{1 - F_1(t)} \right\} \left\{ 1 - F_1(t) \right\} dt
\]
\[
= \int_0^{T_{n_1}^1} \left[ \int_t^\infty \left\{ 1 - F_1(s) \right\} ds \right] d \left\{ \sqrt{n_1} \frac{F_{1n_1}(t) - F_1(t)}{1 - F_1(t)} \right\}.
\]
Thus we see from Gill (1983) and the assumption (3.1) that
\[- \int_0^{T_{(n_1)}} \sqrt{n_1} \{ F_{1n_1}(t) - F_1(t) \} \, dt \]

converges in distribution to a normal random variable with 0 mean variance
\[
\int_0^{T_1} \left[ \int_t^{\infty} \{ 1 - F_1(s) \} \, ds \right]^2 \frac{dF_1(t)}{(1 - F_1(t))^2 \{ 1 - G_1(t^-) \}}.
\]

Thus from the assumptions (3.2)-(3.5) and the Slutsky’s Theorem (Bickel and Doksum, 1977), we have that
\[
\sqrt{n} \left[ \frac{F_{2n_2}(T_{(n_2)})}{H_n(D_n)} \int_0^{T_{(n_1)}} \{ 1 - F_{1n_1}(t) \} \, dt - \int_0^{\infty} \{ 1 - F_1(t) \} \, dt \right]
\]

converge in distribution to a normal random variable $Z_1$ with 0 mean and variance $\sigma_1^2$,
\[
\sigma_1^2 = (1 + \lambda^{-1}) \int_0^{T_1} \left[ \int_t^{\infty} \{ 1 - F_1(s) \} \, ds \right]^2 \frac{dF_1(t)}{(1 - F_1(t))^2 \{ 1 - G_1(t^-) \}}.
\]

With the same arguments and assumptions, we see that
\[
\sqrt{n} \left[ \frac{F_{1n_1}(T_{(n_1)})}{H_n(D_n)} \int_0^{T_{(n_2)}} \{ 1 - F_{2n_2}(t) \} \, dt - \int_0^{\infty} \{ 1 - F_2(t) \} \, dt \right]
\]

converge in distribution to the normal random variable $Z_2$, which are independent of $Z_1$ with 0 mean and variance $\sigma_2^2$,
\[
\sigma_2^2 = (1 + \lambda) \int_0^{\infty} \left[ \int_t^{T_2} \{ 1 - F_2(s) \} \, ds \right]^2 \frac{dF_2(t)}{(1 - F_2(t))^2 \{ 1 - G_2(t^-) \}},
\]

which completes the proof of Theorem with the fact that $Z_1$ and $Z_2$ are independent.

Then in order to apply the result of Theorem 3.1 for the inference for $\beta_1$, we need a consistent estimate of $\sigma^2$. For this, let $T_{(i)}$ be the $i$ th largest observation among $T_{n_1}, \cdots, T_{n_1}$ and $\delta_{(i)}^1$, its concomitant variable. Also let $T_{(j)}$ be the $j$ th largest observation among $T_{n_1+1}, \cdots, T_n$ and $\delta_{(j)}^2$, its concomitant variable. Then an estimate $\hat{\sigma}_n^2$ of $\sigma^2$ can be obtained as follows (Gill, 1983):
\[
\hat{\sigma}_n^2 = n \frac{F_{2n_2}(T_{(n_2)})}{H_n(D_n)} \sum_{i=1}^{n_1} \left[ \int_{T_{(i)}}^{T_{(n_1)}} \{ 1 - F_{1n_1}(s) \} \, ds \right]^2 \frac{\delta_{(i)}^1}{(n_1 - i)(n_1 - i + 1)}
\]
\[
+ n \frac{F_{1n_1}(T_{(n_1)})}{H_n(D_n)} \sum_{j=1}^{n_2} \left[ \int_{T_{(j)}}^{T_{(n_2)}} \{ 1 - F_{2n_2}(s) \} \, ds \right]^2 \frac{\delta_{(j)}^2}{(n_2 - j)(n_2 - j + 1)}.
\]

We note that we cannot say that $\hat{\sigma}_n^2$ is a consistent estimate of $\sigma^2$ if the conditions (3.1)-(3.5) are not satisfied.
4. A numerical example and some concluding remarks

In order to illustrate our procedure, we use the following data in Table 4.1 from Pike (1966), which gave the times from insult with carcinogen DMBA to mortality from vaginal cancer in rats. According to a pre-treatment regime, 40 rats were divided into two groups as Group 1 and 2 with sample sizes 19 and 21, respectively. In Table 4.1, + implies censored observation. Then under the location translation model, in order to estimate the treatment effect, \( \beta_1 \), for the pre-treatment regimens, we obtained the following relevant quantities.

\[
F_{1,19}(304) = 1, \quad F_{2,21}(344) = 0.95, \quad H_{40}(201) = 0.95,
\]

\[
\int_0^{304} \{1 - F_{1,n_1}(t)\} \, dt = 218.76 \quad \text{and} \quad \int_0^{344} \{1 - F_{2,n_2}(t)\} \, dt = 241.85.
\]

Thus we have that

\[
\hat{\beta}_1 = \frac{241.85 - 0.95 \times 218.76}{0.95} = 35.82.
\]

Table 4.1 Days to vaginal cancer mortality in rats

<table>
<thead>
<tr>
<th>Group</th>
<th>Survival times</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>143, 164, 188, 188, 190, 192, 206, 209, 213, 216, 216+, 220, 227, 230, 234, 244+, 246, 265, 304</td>
</tr>
</tbody>
</table>

Also we obtained

\[
\sum_{i=1}^{19} \left[ \int_{T_{1,i}}^{T_{1,19}} \{1 - F_{1,19}(s)\} \, ds \right]^2 \frac{\delta_{1,i}^1}{(19 - i)(19 - i + 1)} = 82.61
\]

and

\[
\sum_{j=1}^{21} \left[ \int_{T_{2,j}}^{T_{2,21}} \{1 - F_{2,21}(s)\} \, ds \right]^2 \frac{\delta_{2,j}^2}{(21 - j)(21 - j + 1)} = 121.62.
\]

Thus we have that

\[
\hat{\gamma}_{40}^2 = 8694.76.
\]

Then for testing \( H_0 : \beta_1 = 0 \), since

\[
\sqrt{40}\frac{\hat{\beta}_1}{\sqrt{\hat{\gamma}_{40}^2}} = 2.430,
\]

we may see that the asymptotic \( p \)-value is 0.0075 from the standard normal distribution for the one-sided alternative \( H_1 : \beta_1 > 0 \). In passing, we note that Kalbfleisch and Prentice (1980) obtained the 2.75 and 2.72 as the values of the approximate chi-square statistics.
for the linear rank and log-rank statistics, respectively. Therefore the test based on our proposed estimate suggest stronger evidence for a difference between the two populations.

That either condition in (3.1) is violated or omitted would not guarantee the unbiasedness for the estimate $\hat{\beta}_1$. Therefore our estimation procedure can not completely prevent the bias from the censoring mechanism but provide a tool with easy application to obtain the treatment effect. Also the explicit form for the estimate may be another advantage of our method.

As a by-product of our procedure, we may obtain a simple estimate of the mean-life for the censored data using the same arguments. For this suppose that we have a sample $(T_i, \delta_i), i = 1, \cdots, n$ from a population having the unknown distribution function $F$ with mean $\mu$, where $T_i$ and $\delta_i$ are defined as (2.2) under the iid setting. Let $F_n$ be the corresponding Kaplan-Meier estimate of $F$. Then the least square estimate $\hat{\mu}$ can be obtained with the same modification in (2.10) as $\hat{\mu} = \frac{1}{F_n(T_{(n)})} \int_0^{T_{(n)}} \{1 - F_n(t)\} dt,$

where $T_{(n)}$ is the largest observation from $T_1, \cdots, T_n$. Thus we note that $\hat{\mu}$ becomes the usual sample mean-life when the largest observation is not censored. For the fixed censoring case, especially the type I censoring, the least square estimate $\hat{\mu}$ can be useful since the censoring time $T$ is pre-assigned. Then in this case, $\hat{\mu}$ can be considered an estimate of the mean for the truncated distribution at $T$.

One may try to apply the least square method to the models which may include some covariate or covariates. Then it would be inevitable to use the iterative procedure for the calculation to obtain the estimate since (2.3) is one of such models.

Finally we note that the two expressions for $\hat{\beta}_1$ in (2.8) and (2.10) coincide if the largest observation from each sample is not censored. Since the jump size for the Kaplan-Meier estimate is 0 for the censored observation, if the largest observation is censored then the contribution of the largest observation can be included for (2.10) but not for (2.8), which may incur some serious underestimation of for each component. This is why we have proposed (2.10) instead of (2.8).

Acknowledgment

The authors wish to express their sincere appreciation to the three anonymous referees for pointing out errors and suggesting constructive ideas.

References


An estimation of the treatment effect for the right censored data


