The Korean Communications
in Statistics Vol. 10 No. 3, 2003
pp. 1017-1024

Estimation of Treatment Effect for Bivariate Censored Survival Data[1]

Choon-Mo Ahn[2], Sang-Gue Park[3]

Abstract

An estimation problem of treatment effect for bivariate censored survival data is considered under location shift model between two sample. The proposed estimator is very intuitive and can be obtained in a closed form. Asymptotic results of the proposed estimator are discussed and simulation studies are performed to show the strength of the proposed estimator.

Keywords : Bivariate censored survival data; Kaplan-Meier estimator; Location shift model; Paired failure times; Quantile function; Treatment effect.

1. Introduction

Consider the problem that there are two failure times for each observational unit like a paired experiment. Let $T=(T_1, T_2)$ be a pair of nonnegative random variables. The variables T_1 and T_2 may represent failure times of paired subjects, times from individuals of a treatment until first response in two successive courses of a same patients, etc. Under bivariate censoring, the observable variables are given by $Y=(Y_1, Y_2)$ and $\delta=(\delta_1, \delta_2)$, where $Y_i=\min(T_i, Z_i)$ and $\delta_i=I(T_i=Z_i)$ $(i=1,2)$. Here $Z=(Z_1, Z_2)$ is a pair of censoring times thought to represent times to withdrawals from the study. This type of censoring mechanism can be found in Clayton(1978), Campbell(1981), Clayton and Cuzick(1985) and Dabrowska(1988, 1989) among others.

Some methods related linear regression like Lee, Wei and Ying (1993) in the existing literature address this problem. These regression methods can be used here in the two-sample problem by setting the covariate 0 for the first sample and 1 for the second. However, so-called direct procedures for the estimation of two-sample problem would be preferred

---

1) This is supported by Chung-Ang University Research Grant in 2002.
2) Senior researcher, Information Technology Management Research Group, ETRI, Daejon, 305-350
E-mail : cmahn@etri.re.kr
3) Professor, Department of Statistics, Chung-Ang University, Seoul, 156-756, Korea
E-mail : spark@cau.ac.kr

because it imposes fewer assumptions than the regression methods. Meng, Bassiakos and Lo (1991) and Park and Park (1995) proposed direct procedures of the estimation of treatment effect for the two-sample problem. They are all inspired Akritas (1986)'s method for quantile estimation in the shift version of the two-sample problem.

In present paper, we consider the estimation of treatment effect of bivariate censored data under bivariate censoring. Under location shift model, the treatment effect is defined as the difference between medians of two failure times. Suppose $T=(T_1, T_2)$ has the continuous distribution function F and let F_i be the marginal distribution functions of T_i with density f_i $(i=1,2)$. For some $\Delta>0$,

$$F_2(t)=F_1(t-\Delta) \quad \text{for all} \quad t>0. \tag{1.1}$$

We can easily see that Δ is the median of the distribution T_2-T_1 and the treatment effect Δ is of interest. Another approach for estimating Δ is as follows. Let us denote the quantile function $F_i^{-1}(y)$ $(i=1,2)$ for $0\le y\le 1$, $F_i^{-1}(y)=\inf\{t: F_i(t)\ge y\}$.

It is easy to show that

$$\frac{1}{\beta}\int_0^{\beta}(F_2^{-1}(t)-F_1^{-1}(t))\,dt=\Delta \tag{1.2}$$

under the model (1.1). Park and Park (1995) proposed the quantile estimator of treatment effect from the two independent failure times by using the relation (1.2).

We propose the estimator of treatment effect by generalizing (1.2) to bivariate censored failure times and establish the asymptotic normality by using some well-known results in section 2. Simulation studies are presented in section 3.

2. Estimator and its asymptotic properties

Under the model (1.1), for $0<\beta\le 1$, Δ can be estimated by

$$\hat{\Delta}=\frac{1}{\beta}\int_0^{\beta}(\hat{F}_2^{-1}(t)-\hat{F}_1^{-1}(t))\,dt \tag{2.1}$$

where $\hat{F}_i$ $(i=1,2)$ is the Kaplan and Meier (1958) estimator of F_i $(i=1,2)$ using marginal sample (Y_{ij}, δ_{ij}) $(i=1,2; j=1,\cdots,n)$ respectively.

One reasonable choice of β is $\min\{\hat{F}_1(\tau), \hat{F}_2(\tau)\}$, where τ is the minimum of two largest uncensored observation obtained in each sample. This estimator looks same as Park and Park (1995)'s estimator for the case of independence, but it's not since the asymptotic variance of the estimator is different. Major advantage of this estimator is to avoid the estimation of marginal densities of failure times in estimating the asymptotic variance of the estimator.

Now we show the asymptotic normality of $\sqrt{n}(\hat{\Delta}-\Delta)$. This result will be used for estimation of treatment effect. Suppose $Z=(Z_1, Z_2)$ has the continuous distribution function G and let G_i be the marginal distribution functions of Z_i $(i=1,2)$. We also let the distribution function of Y be H and the survival function of Y be $\overline{H}$. $\overline{F}$, $\overline{F}_i$, $\overline{G}$ and $\overline{G}_i$ can be defined similarly. Then we have $\overline{H}(u,v)=\overline{F}(u,v)\overline{G}(u,v)$ for $u,v>0$. Using (1.2) and (2.1) we have

$$\sqrt{n}(\widehat{\Delta}-\Delta)=\frac{1}{\beta}\int_0^{\beta}\{\sqrt{n}(\widehat{F}_2^{-1}(t)-F_2^{-1}(t))-\sqrt{n}(\widehat{F}_1^{-1}(t)-F_1^{-1}(t))\}dt. \tag{2.2}$$

Theorem 2.1 Suppose that β satisfies $\overline{H}(F_1^{-1}(\beta), F_2^{-1}(\beta))>0$ and $\inf_{0\le t\le\beta} f_i(F_i^{-1}(t))>0$ ($i=1,2$). Then $\widehat{\Delta}$ is a strongly consistent estimate of Δ.

Proof. This follows from lemma 1 and theorem 1 in Park and Park (1995).

Theorem 2.2 Suppose that β satisfies $\overline{H}(F_1^{-1}(\beta), F_2^{-1}(\beta))>0$ and F_i $(i=1,2)$ has a bounded second derivative F_i'' on $[0, F_i^{-1}(\beta)+\delta)$ for some $\delta>0$ and $\inf_{0\le t\le\beta} f_i(F_i^{-1}(t))>0$. Then $\sqrt{n}(\widehat{\Delta}-\Delta)$ is asymptotically normal distribution.

Proof. By Bahadur representation of Kaplan-Meier estimator represented by Cheng (1984), (2.2) can be transformed as (2.3) under assumed conditions;

$$\sqrt{n}(\widehat{\Delta}-\Delta)=\sum_{i=1}^{2}(-1)^{i+1}\frac{1}{\beta}\int_0^{\beta}\frac{\sqrt{n}(\widehat{F}_i(F_i^{-1}(t))-t)}{f_i(F_i^{-1}(t))}dt+\int_0^{\beta}R_n(t)\,dt \tag{2.3}$$

where $R_n(t)$ is the remainder term with $\sup_{0\le t\le\beta}|R_n(t)|=O(n^{-1/4}(\log n)^{3/4})$.

If $\tau=(\tau_1,\tau_2)$ be a point such that $\overline{H}(\tau_1,\tau_2)>0$. Then we know that $\sqrt{n}(\widehat{F}_1(s)-F_1(s), \widehat{F}_2(t)-F_2(t))$ converge weakly on $D[[0,\tau_1]\times[0,\tau_2]]$ to mean zero bivariate normal process (Wang and Wells (1997)). So by continuity theorem (Theorem 5.1, Billingsley (1968)) and the property of normal process, $\sqrt{n}(\widehat{\Delta}-\Delta)$ converges weakly to normal distribution. ■

Asymptotic variance σ^2 of $\sqrt{n}(\widehat{\Delta}-\Delta)$ can be computed using (2.3) and the change of variable technique. Note that "acov" means asymptotic covariance.

$$\begin{aligned}\beta^2\sigma^2 &= \sum_{i=1}^{2}\int_0^{F_i^{-1}(\beta)}\int_0^{F_i^{-1}(\beta)}\mathrm{acov}(\sqrt{n}(\widehat{F}_i(s)-F_i(s)), \sqrt{n}(\widehat{F}_i(t)-F_i(t)))dsdt\\ &\quad -2\int_0^{F_2^{-1}(\beta)}\int_0^{F_1^{-1}(\beta)}(\sqrt{n}(\widehat{F}_1(s)-F_1(s)), \sqrt{n}(\widehat{F}_2(t)-F_2(t)))dsdt\end{aligned}$$

where

$$\operatorname{acov}(\sqrt{n}(\widehat{F}_i(s)-F_i(s)), \sqrt{n}(\widehat{F}_i(t)-F_i(t))) = \overline{F}_i(s)\overline{F}_i(t)\int_0^{s\wedge t}\frac{d\Lambda_i(x)}{\overline{H}_i(x)} \tag{2.4}$$

(refer section 6.3 of Fleming & Harrington (1992)) and

$$\begin{aligned}&\operatorname{acov}(\sqrt{n}(\widehat{F}_1(s)-F_{1(s)}), \sqrt{n}(\widehat{F}_2(t)-F_2(t)))\\ &= \overline{F}_1(s)\overline{F}_2(t)\int_0^t\int_0^s \frac{\overline{H}(u,v)}{\overline{H}_1(u)\overline{H}_2(v)}\{\Lambda_{11}(du,dv)+\frac{\overline{H}_{10}(du,v)}{\overline{H}(u,v)}\Lambda_2(dv)\\ &+\frac{\overline{H}_{01}(u,dv)}{\overline{H}(u,v)}\Lambda_1(du)+\Lambda_1(du)\Lambda_2(dv)\}\end{aligned} \tag{2.5}$$

where $\overline{H}_{10}(u,v)=\Pr(Y_1>u,\delta_1=1,Y_2>v)$, $\overline{H}_{01}(u,v)=\Pr(Y_1>_{u,}Y_2>v,\delta_2=1)$, Λ_i is hazard function of T_i, Λ_{11} is hazard function which fail T_1 and T_2 simultaneously. The formula given in (2.5) can be derived using Lo and Singh's (1986) representation of Kaplan - Meier estimator (Theorem 1, p.456). This is also derived in Wang and Wells (1997).

For estimating σ^2, a consistent estimator can be obtained by replacing $\overline{F}_i$, $\overline{H}_i$, Λ_{11}, Λ_i, $\overline{H}_{10}$, $\overline{H}_{01}$ by their empirical distribution and Kaplan-Meier estimator, respectively. Explicitly, assume that $Y_{1(j)}$ and $Y_{2(j)}$ $(j=1,\cdots,n)$ are the order statistic of Y_{1j} and Y_{2j}, respectively. Then second term of asymptotic covariance given in (2.5) can be estimated by product of $\overline{\widehat{F}_1}(s)\overline{\widehat{F}_2}(t)$ and the following formula;

$$\begin{aligned}&\sum_{Y_{2(l)}\le t}\sum_{Y_{1(k)}\le s}\{n\times\frac{\sum_{j=1}^n\{I(Y_{2j}\le Y_{2(l)},\delta_{2j}=1)-I(Y_{2j}<Y_{2(l)},\delta_{2j}=1)\}}{\sum_{j=1}^n I(Y_{2j}\ge Y_{2(l)})}\\ &\times\frac{\sum_{i=1}^n\{I(Y_{1i}>Y_{1(k)},\delta_{1i}=1,Y_{2i}>Y_{2(l)})-(Y_{1i}\ge Y_{1(k)},\delta_{1i}=1,Y_{2i}>Y_{2(l)})\}}{\sum_{i=1}^n I(Y_{1i}>Y_{1(k)})\times\sum_{j=1}^n I(Y_{2j}>Y_{2(l)})}\\ &\times(Y_{2(l)}-Y_{2(l-1)})(Y_{1(k)}-Y_{1(k-1)})\}.\end{aligned}$$

Asymptotic variance given in (2.4) and the remaining formulas given in (2.5) can be estimated similarly.

3. An illustrated example

We present a numerical example to show the application of the proposed estimator and compare its result with others. We consider the well-known matched pair data of Holt and Prentice (1974), which is considered by many authors. We assume the logarithm of survival times of closely Y_{2i} and poorly Y_{1i} matched skin graft on the same burn patient follows the model (1.1). The log survival times were recorded like Table 1.

[Table 1] log(Days) of survival of skin grafts on burn patients

	1	2	3	4	5	6	7	8	9	10	11
log Y2	1.57	1.28	1.76+	1.97	1.20	1.34	1.30	1.26	1.80	1.46	1.78+
log Y1	1.46	1.11	1.18	1.41	1.04	1.23	1.41	1.32	1.63	1.17	1.60

Jung and Su (1995) and Lee et al. (1993) proposed the estimators of treatment effect in bivariate censored data. Both methods are based on so-called minimum dispersion statistic. Lee et al. showed that the Wilcoxon score is generally acceptable. So we choose it for comparison. Assuming that $(\log Y_1, \log Y_2)$ follows model (1.1), 95% confidence intervals of $\varDelta$ by Jung and Su and Lee et al. are (-0.20, 1.24) and (0.12, 0.80), respectively.

Now we calculate the proposed interval estimator for $\varDelta$. First we have $\hat{\tau}=1.63$ and $\beta=7/11$. Therefore we have $\hat{F}_1^{-1}(\beta)=1.41$ and $\hat{F}_2^{-1}(\beta)=1.57$. So the point estimate of treatment effect is $\hat{\varDelta}=(3.626-3.429)/(7/11)=0.310$. Using (2.4) and (2.5), we get $\hat{\sigma}^2=(0.108+0.100-2\times 0.050)/(7/11)^2=0.267$. So the 95% confidence interval of $\varDelta$ is (0.005, 0.615).

We now compare these methods by examining three constructed approximate 95% confidence intervals. The interval length of Jung and Su's method is much wider than intervals based on Lee et al.'s and the proposed one. Also Lee et al.'s method and the proposed method detect significant treatment effect at 5% level, but Jung and Su do not detect the difference.

This example suggests the further studies of small sample behaviors. We discuss this problem in the next section.

4. Simulation and Conclusions

Some simulation studies with 1000 repetitions were carried out to compare the finite sample (n=20, 40) performance of several estimators of $\varDelta$, i.e. Lee et al. (1993), Jung and Su (1995) and the proposed method. Lee et al.'s statistics are based on minimum dispersion statistic and considered the several types of weights and other methods as mentioned earlier. They showed rank method with Wilcoxon score performs the best in normal or non-normal case, so we choose it for comparison.

In the three simulations, we follow the method used by Lee et al. with the different bivariate distribution. That is, the pairs of failure times (T_1, T_2) were distributed according to the bivariate normal distribution with marginal mean 0, marginal variance 1 and the correlation 0.5, to the bivariate exponential distribution,

$$\overline{F}(t_1, t_2)=e^{-(t_1+t_2)}\{1+\theta(1-e^{-t_1})(1-e^{-t_2})\}$$

with $\theta=1$ and

$$\overline{F}(t_1, t_2) = (e^{t_1/\theta} + e^{t_2/\theta} - 1)^{-\theta}$$

with $\theta=0.25$, by Gumbel (1960) and Clayton (1978), respectively. The Gumbel model with $\theta=1$ represents fairly weak positive dependence while the Clayton model with $\theta=0.25$ represents fairly strong positive dependence. Gumbel and Clayton model failure times are obtained by the method given in Prentice and Cai (1992). And the censoring times are the same distribution function to failure times with $\theta=1$ for Gumbel model and $\theta=0.5$ for Clayton model. In this setting, we choose $\Delta=0$ for Gumbel and Clayton model and $\Delta=0.1$ for the bivariate normal distribution. All the results are summarized in Table 2, 3 and 4.

In table 2, we can find that the proposed method shows better performance in estimating location-shift parameter with respect to coverage probability and average interval length when the marginal distribution is normal. The method proposed by Jung and Su(1995) is very conservative compared with other methods and the method proposed by Lee et al.(1993) does not maintain the nominal level.

In table 3 and 4, the marginal distribution is exponential distribution. The proposed method still works better comparing other two methods. The method given by Lee et al. seems to have shorter average interval length, but it did not maintain the nominal level.

In overall, the proposed method is recommendable with respect to coverage probability and the average interval length. However, one should be cautious when the sample size is less than 40.

All the calculations in this article were programmed in FORTRAN 77 with double arithmetic precision on a IBM personal computer. Random numbers were generated using IMSL subroutine. The simulation program is available from the first author.

[Table 2] Empirical coverage probability from bivariate normal distribution with $\Delta=.1$

Sample size	20	40
censoring rate of Y1	50.5%	49.9%
censoring rate of Y2	50.2%	50.0%
double censoring	33.5%	33.3%
Jung and Su	0.992 (3.586)	0.991 (1.849)
Lee et al.	0.917 (1.240)	0.938 (0.826)
Proposed	0.899 (1.107)	0.950 (0.794)

* ( ) means average length of confidence interval

[Table 3] Empirical coverage probability from Gumbel distribution with $\Delta=0$

Sample size	20	40
censoring rate of Y1	49.7%	49.7%
censoring rate of Y2	50.5%	50.0%
double censoring	30.5%	30.2%
Jung and Su	0.991 (6.424)	0.993 (2.533)
Lee et al.	0.932 (0.779)	0.948 (0.464)
Proposed	0.932 (0.599)	0.955 (0.550)

* ( ) means average length of confidence interval

[Table 4] Empirical coverage probability from Clayton distribution with $\Delta=0$

Sample size	20	40
censoring rate of Y1	49.8%	49.9%
censoring rate of Y2	50.5%	50.0%
double censoring	39.3%	39.3%
Jung and Su	0.996 (4.002)	0.999 (1.973)
Lee et al.	0.931 (0.463)	0.938 (0.284)
Proposed	0.931 (0.476)	0.955 (0.419)

* ( ) means average length of confidence interval

References

[1] Akritas, M. G. (1986). Empirical processes associated with V-statistics and a class of estimators under random censoring, *The Annals of Statistics*, 14, 619-637

[2] Billingsley, P. (1968). Convergence of probability measures, Wiley, New York.

[3] Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data, *Biometrika*, 68, 417-422

[4] Cheng, K. F. (1984). On almost sure representation for quantiles of the product limit estimator with applications, *Sankhya, Series A, The Indian Journal of Statistics*, 46, 426-443.

[5] Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, 65, 141-151.

[6] Clayton, D. G. and Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model (with discussion), *Journal of the Royal Statistical Society Series A*, 148, 82-117.

[7] Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane, *The Annals of*

Statistics, 16, 1475-1489.

[8] Dabrowska, D. M. (1989). Kaplan-Meier estimate on the plane: weak convergence, LIL, and the bootstrap, *Journal of Multivariate Analysis*, 29, 308-325.

[9] Fleming, T. R. and Harrington, D. P. (1991). Counting processes and survival analysis, Wiley, New York.

[10] Gumbel, E. J. (1960). Bivariate exponential distributions, *Journal of American Statistical Association*, 55, 698-707 (1960).

[11] Holt, J. D. and Prentice, R. L. (1974). Survival analysis in twin studies and matched pairs experiments, *Biometrika*, 61, 17-30.

[12] Jung, S. H. and Su, J. Q. (1995). Nonparametric estimation for the difference or ratio of median failure times for paired observations, *Statistics in Medicine*, 14, 275-281.

[13] Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53, 457-481.

[14] Lee, E. W., Wei, L. J. and Ying, Z. (1993). Linear regression analysis for highly stratified failure time data, *Journal of American Statistical Association*, 88, 557-565.

[15] Lo, S. H. and Singh, K. (1985). The product-limit estimator and the bootstrap: some asymptotic representations, *Probability Theory and Related Fields*, 71, 455-465..

[16] Meng, X. L., Bassiakos, Y. and S. H. Lo. (1991). Large-sample properties for a general estimator of the treatment effect in the two-sample problem with right censoring, *The Annals of Statistics*, 19, 1786-1812.

[17] Park, H. I. and Park, S. G. (1995). Quantile estimation of treatment effect for the two sample problem with right censored data, *Statistics and Probability Letters*, 24, 139-145.

[18] Prentice, R. L. and Cai, J. (1992). Covariance and survivor function estimation using censored multivariate failure time data, *Biometrika*, 79, 495-512.

[19] Wang, W. and Wells, M. T. (1997). Nonparametric estimators of the bivariate survival function under simplified censoring conditions, *Biometrika*, 84, 863-880.

[Received July 2003, Accepted October 2003]