Journal of Statistial
Theary & Methods
1994, Vol. 5, pp45 ~ 57

임의 중단모형하에서의 평균잔여수명함수의 추정

이 인 석 , 이 우 동[1)]

요 약 이 연구에서는 Hjort(1991)에의해 제안된 누적위험률함수의 비모수적 추정량을 이용하여 무한인 구간까지 정의된 평균잔여수명함수의 추정량을 제안하고 제안된 추정량의 일치성과 점근적 정규성을 밝히고, 모의실험을 통하여 다른 추정량들과 비교하고자 한다.

1. 서 론

임의의 기계가 t시간까지 살았다는 가정하에서 그 기계의 앞으로 남아있는 평균수명시간, 즉, 평균잔여수명(mean residual life)은 의학분야나 공학분야등, 수명시간분석에 관련된 분야에서 관심이 되는 중요한 측도 이다. 이러한 평균잔여수명은 생존함수(survival function)의 함수로, 몇몇 연구자들에 의해 생존함수의 비모수적 추정량을 이용한 추정량들이 제안되어 왔다.

임의중단모형(random censoring model)하에서, Yang(1977,1978)은 평균잔여수명에대한 비모수적 추정량을 제안했으며, 제안된 추정량이 미리 정의된 유한인 구간 상에서 가우시안(Gaussian) 과정으로 수렴함을 밝혔다. 그리고 베이지안 측면에서, Ghorai, Susarla, Susarla 와 Van Ryzin(1980)은 제곱손실오차(squared error loss)를 고려하여 추정량을 제안하였다. Kumazawa(1987)는 Kaplan-Meier추정량을 이용하여 평균잔여수명함수에대한 추정량을 제안하고 Yang의 연구를 Gill(1983)의 결과를 이용하여 무한인 구간 상으로 확장하여 제안된 추정량의 정규성을 밝혔다.

이 연구에서는 Hjort에의해 제안된 누적위험률함수의 비모수적 추정량을 이용하여 무한인 구간까지 정의된 평균잔여수명함수의 추정량을 제안하고 제안된 추정량의 점근적 일치성과 점근적 정규성을 밝히고, 모의실험을 통하여 기존의 다른 추정량들과

1경북대학교 통계학과

비교하고자 한다.

2. 평균잔여수명함수의 추정

$T_1, T_2, \cdots, T_n$을 $[\ 0, \infty)$상에서 정의된 분포함수(distribution function), F로부터 추출된 확률표본(random samples)이라 두자. 그리고 수명함수는

$$\begin{aligned} S(t) &= P\{T_i > t\} \\ &= 1-F(t) \\ &= \exp(-\Lambda(s)) \end{aligned}$$

라 하자. 여기서 $\Lambda(s)$는 누적위험률함수(cumulative hazard function)로서,

$$\Lambda(s) = \int_0^s \lambda(t)dt$$

이고

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr\{T \in (t, t+\Delta t) | T > t\}}{\Delta t}$$

이며 $\lambda(t)$는 어떤 기계가 t시간이상 살았다는 가정(조건)하에서 아주 짧은 시간 혹은 순간에 죽을 확률 혹은 위험률(hazard rate)이다. 또, $C_1, C_2, \cdots, C_n$ 을 분포함수로 $1-G(t)$를 갖는 중단확률표본이라 하고 T_i와 C_i는 독립이라 가정하자. 그리고

$$X_i \equiv (T_i \wedge C_i) = \min(T_i, C_i),\ i=1,2,\cdots,n$$

라두고 그 분포함수를 $H(t)$라 하면,

$$\begin{aligned} 1-H(t) &= P\{X_i > t\} \\ &= S(t)G(t) \end{aligned}$$

인 관계가 성립한다.

임의중단 모형에서 관측되는 자료의 형태는 $(X_1, \delta_1), (X_2, \delta_2), \cdots, (X_n, \delta_n)$이며 여기서

$$\begin{aligned} \delta_i &= I\{T_i < C_i\} \\ &= \begin{cases} 1 & \text{if } T_i < C_i \\ 0 & \text{if } T_i \geq C_i \end{cases} \end{aligned}$$

로서 중단 지시자(indicator function)이다. $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$을 $X_1, X_2, \cdots, X_n$의 순서통계량이라하고 $X_{(i)}$의 중단지시자를 $\delta_{(i)}$라두자.

위의 임의 중단모형에서 널리 이용되는 수명함수의 비모수적 추정량을 소개하면 아래와 같다. 먼저, Kaplan과 Meier(1958)는 생존함수의 추정량(K-M 추정량)을 아래와 같이 제안하였고

$$\hat{S}_{KM}(t) = \begin{cases} \prod_{i : X_{(i)} \le t} \left\{ \frac{n-i}{n-i+1} \right\}^{\delta_{(i)}}, & \text{for } t \le X_{(n)} \\ 0 \quad \text{if } \delta_{(n)} = 0, & \text{for } t > X_{(n)} \\ \textit{undefined} \quad \text{if } \delta_{(n)} = 1, & \text{for } t > X_{(n)} \end{cases}$$

그후, 여러학자들에의하여 K-M 추정량에 대한 통계적 성질이 밝혀져 왔는데, 그 중 대표적인 연구자와 연구내용을 간략히 소개해보면, Breslow와 Crowley(1974)는 K-M추정량의 극한분포가 정규분포임을 증명하였고, Peterson(1977)은 K-M 추정량을 두개의 부생존함수로 표시하였으며, Gill(1983)은 K-M 추정량의 점근적성질을 밝혔다. 또, K-M 추정량에 비교될만한 생존함수의 추정량으로는 Nelson(1972)과 Aalen(1978)에 의해 제안된 누적위험률함수에 대한 경험적(empirical) 추정량으로 그들의 추정량은

$$\begin{aligned} \hat{\Lambda}(t) &= \int_0^t \frac{dN(s)}{Y(s)} \\ &= \sum_{i : X_{(n)} \le t} \frac{\delta_{(i)}}{n-i+1}, \end{aligned}$$

이며, 여기서 $N(t)$는 t보다 작거나 같은 임의중단되지않은 관찰치(uncensored observation)의 개수이며 $Y(t)$는 t보다 작거나 같은 관찰치(observation)의 개수이다. 즉

$$N(t) = \sum_{i=1}^{n} I_{\{X_{(i)} \le t,\, \delta_{(i)} = 1\}}$$

$$Y(t) = \sum_{i=1}^{n} I_{\{X_{(i)} \ge t\}}$$

이다.

Nelson과 Aalen의 누적위험률함수에 대한 추정량을 이용한 생존함수의 추정량(N-A 추정량)은 분포함수와 누적위험률함수의 관계를 이용하여 다음과 같이 제안할 수 있다.

$$\hat{S}_{NA}(t) = e^{-\hat{\Lambda}(t)}.$$

N-A 추정량의 점근적성질은 셈과정이론(counting process theory)을 이용하여 Aalen(1978)과 Anderson과 Borgan(1985)등에 의해 밝혀졌다.

한편, 비모수적 베이지안 추정량들은 손실함수를 제곱손실오차로 가정했을 때 생존함수 $S(t)$에대한 추정량으로 Susarla 와 Van Ryzin(1976)이 아래의 Susarla 와 Van Ryzin추정량(S-V 추정량), 누적위험률함수 $\Lambda(t)$의 추정량으로 Hjort(1991)가 제안한 추정량이 있다. 먼저 S-V 추정량은 아래와 같다.

$$\hat{S}_{SV}(t) = \frac{\alpha(t,\infty)+N^+(t)}{n+\alpha(R^+)} \prod_{i=1}^{n} \left\{ \frac{N^+(X_i)+\alpha(X_i,\infty)+1}{N^+(X_i)+\alpha(X_i,\infty)} \right\}^{[X_i > t,\ \delta_i=0]}$$

여기서

$$N^+(t) = \sum_{i=1}^{n} I_{(X_{(i)} \geq t)}$$

이고 α는 R^+상에서 임의의 유한한 측도(finite measure)이며 과거나 사전지식의 정보를 이용한 $S(t)$의 사전 정보이다. S-V 추정량에 대한 자료의 개수가 무한인 경우 점근 분포가 정규분포함을 Susarla 와 Van Ryzin(1978)에 의해 증명되었다.

그리고 최근에 Hjort에의해 제안된 누적위험률함수에대한 추정량은

$$\hat{\Lambda}(t) = \int_0^t \frac{c(s)d\Lambda_0(s)+dN(s)}{c(s)+Y(s)},$$

이다. 여기서 $N(s)$와 $Y(s)$는 N-A 추정량에 정의된 실험적 과정(empirical process)이며 $c(s)$와 $\Lambda_0(s)$는 각 각 s시간에서의 위험(risk)의 수와 누적위험률함수에대한 사전정보나 지식이다. 이 Hjort의 추정량을 이용한 생존함수의 추정량(H 추정량)을 아래와 같이 생각할 수 있다.

$$\hat{S}(t) = e^{-\hat{\Lambda}(t)}.$$

Hjort는 만약 $c(s)$와 $\Lambda_0(s)$가 0으로 수렴한다면 N-A 추정량으로 수렴한다는 것과 $c(s)$는 음이 아닌 조각연속(piecewise continuous)인 함수로 제안했다.

앞으로 $D[0,T]$ 를 $[0,T]$ 상에서 정의된 오른쪽으로부터 연속(right continuous)이고 왼쪽극한(lefthand limit)을 가지며 Skorohod metric을 가진 함수들의 집합이라 하자.(Pollard(1984)와 Billingsley(1968)참조)

Lee와 Lee(1993)는 위의 주어진 H 추정량에 대한 점근적 성질을 아래의 보조정리들에 밝혔다.

보조정리 2. 1 (Lee and Lee(1993)) T는 $H(T)<1$ 이고, 모든 $s \in [0,T]$ 에대해,

(1) $Y(s) \rightarrow^{p} \infty$

(2) $c(s)$는 $\frac{c(s)}{Y(s)} \rightarrow^{p} 0$인 예측가능(predictable)하고 유계(bounded)이면

$$\lim_{n\to\infty} P\{\sup_{s\in[0,T]} | \tilde{S}(s)-S(s) | > \varepsilon\} = 0.$$

보조정리 2. 2 (Lee and Lee(1993)) T는 $H(T)<1$ 이고, 모든 $s \in [0, T]$ 에대해,

(1) $Y(s) \rightarrow^{p} \infty$

(2) $c(s)$는 $\frac{c(s)}{Y(s)} \rightarrow^{p} 0$인 예측가능(predictable)하고 유계(bounded)이고

(3) $\inf_{0\le s\le T} c(s) \neq 0$ 이면

$n \rightarrow \infty$일 때, $D[0, T]$ 상에서

$\sqrt{n}(\tilde{S}(t)-S(t)) \rightarrow^{D} Z^{*}(t)$,

여기서, $Z^{*}(t)$는 평균은 0이고 공분산함수(covariance function)가

$$Cov(Z^{*}(s), Z^{*}(t)) = S(s)S(t)\int_{0}^{s\wedge t} \frac{dF(u)}{(1-H(u))S(u)}, \quad 0\le s, t \le T$$

인 가우시안 확률과정(Gaussian process)이다.

보조정리2.1과 2.2를 이용하여 이 연구에서 제안된 평균잔여수명함수의 추정문제와 그 점근적성질을 알아보자.

먼저, 평균잔여수명함수의 정의는 아래와 같다.

$$e_t = E(T-t \mid T > t)$$

$$= \frac{\int_t^{\infty} S(s)ds}{S(t)}.$$

위의 정의로부터 $e_0 = \mu$임을 알 수있다.

Yang(1977)과 Susarla와 Van Ryzin(1980)등은 아래와 같이 절단된 평균잔여수명함수

$$e_t^T = \frac{\int_t^T S(s)ds}{S(t)}, \quad H(T) < 1$$

의 추정문제를 연구하였다. 특히 Yang(1978)은 임의 중단모형에서 e_t^T에 포함된 생존함수를생존함수의 비모수적 추정량인 K-M추정량을 대입하여 구간 $[0, T]$ 상에서 일

치성과 극한분포를 구하였다.

최근에 Kumazawa(1987)는 임의중단모형에서, $\tau_F = \sup_t\{t \mid F(t)<1\}$로 정의하고,

$$e_t^{\tau_F} = \frac{\int_t^{\tau_F} S(s)ds}{S(t)}$$

의 추정문제를 고려하였고 여기서 τ_F는 ∞일 수 있다. 그는 K-M추정량을 $e_t^{\tau_F}$에 포함된 생존함수에 대입하여 추정량을 제안하였고 제안된 추정량의 점근분포가 정규분포임을 증명하였다.

본 연구에서는 $e_t^{\tau_F}$ 에 H 추정량을 이용하여 추정량을 제안하고, 점근적 일치성과 정규성을 밝힌다. 먼저

$T^* = \max_{1\le i\le n}\{X_i\}$, $\tau_G = \sup_t\{t \mid G(t)<1\}$

그리고

$\tau_H = \sup_t\{t \mid H(t)<1\}$

이라 정의하자.

만약, $\tau_G<\tau_F$인 관계가 성립한다면 τ_G시점을 넘어서는 중단되지 않은 자료의 관측은 불가능할 것이며, τ_G를 넘어서는 시점에서 생존함수 S의 추정은 불가능할 것이다. 그러므로 $\tau_F\le\tau_G$로 가정하면 $\tau_F=\tau_H$인 관계가 성립되고, 다음과 같이 $e_t^{\tau_F}$의 추정량을 제안하자.

$$\hat{e}_t = \frac{\int_t^{T^*} \hat{S}(s)ds}{\hat{S}(t)}.$$

유사한 방법으로 생존함수의 추정량 $\hat{S}_{KM}(t)$, $\hat{S}_{NA}(t)$, $\hat{S}_{SV}(t)$를 이용하여 $e_t^{\tau_F}$의 추정량을 각각 $\hat{e}_{t^{KM}}$, $\hat{e}_{t^{NA}}$, $\hat{e}_{t^{SV}}$라두자. 지금부터 $\hat{e}_t$의 점근적 성질을 밝혀 보자.

아래의 정리 2.1에서 제안된 추정량의 약 일치성(consistency)을 보였다.

정리 2. 1 만약

$$\sqrt{n}\int_{T^*}^{\tau_F} S(s)ds \to^p 0,\ as\ n\to\infty.$$

그러면

$$\sup_{0\le t\le T^*}|\ \hat{e}_t - e_t^{\tau_F}| \to^p 0,\ as\ n\to\infty.$$

제안된 추정량의 점근분포를 알아보기 위하여, $Z_n(t) \equiv \sqrt{n}(S(t) - \tilde{S}(t)) / S(t)$라 정의하면 확률과정 $Z_n(t)$는 보조정리 2.2에 의해서, 확률과정 $S(t)Z_n(t)$ 는 평균이 0이고 공분산함수가 $Cov(Z^*(s), Z^*(t))$인 가우시안 확률과정으로 수렴하기 때문에, $[0, T]$, $H(T)<1$상에서 브라우니안 모션 확률과정(Brownian motion process), $Z \equiv B(V)$로 수렴하고 여기서 $V(t) = \int_0^t \frac{dF(s)}{(1-F(s))S(s)}$이다. 이 결과와 Gill(1983)의 정리2.2를 이용하여 제안된 추정량의 점근분포를 아래의 정리 2.2에 밝혔다.

정리 2. 2 $h(t) = \int_t^{\tau_F} S(s)ds$라두고, 만약

(1) $\sqrt{n}\ h(T^*) \rightarrow^p 0.$

(2) $\int_0^{\tau_H} h^2(s)dV(s) < \infty$

가 성립한다면, $0 \le t \le T^*$에 대하여, $D[0, \tau_F]$ 상에서

$$\sqrt{n}(\hat{e}_t - e_t^{\tau_F}) \rightarrow^D B^*(t) \equiv - \frac{\int_t^{\tau_F} h(s)dZ(s)}{S(t)}, \quad n \rightarrow \infty$$

여기서 B^*는 평균이 0이고 공분산함수가

$$Cov(B^*(s), B^*(t)) = (S(s)S(t))^{-1} \int_{s \wedge t}^{\tau_F} h^2(u)dV(u) .$$

인 가우시안 확률과정이다

증명 함수 $h(t)$와 조건 (2)에의해 Gill(1983)의 정리 2.2의 조건을 만족해주고, Fleming과 Harrington(1991)의 부록에 있는 정리A.1.2를 이용하면,

$$\begin{aligned}\sqrt{n}(\hat{e}_t - e_t^{\tau_F}) = \ & (\hat{S}(t)S(t))^{-1}\{ -S(t)\int_t^{T^*} h(s)dZ_n(s) \\ & + S(t)Z_n(T^*)h(T^*) - S(t)\sqrt{n}h(T^*) \}\end{aligned}$$

이다. 그리고 조건 (1)에의해 위의 정리가 성립함을 보였다.■

위의 정리의 결과를 이용하여 $e_t^{\tau_F}$에대한 점근적 신뢰구간을 구해보자. 먼저 정리2.2

에서 $\hat{e}_t$의 점근적 분산 $\sigma^2(t)$는 아래와 같다.

$$\sigma^2(t) = (S(t))^{-2} \int_t^{\tau_H} h^2(s) dV(s)$$

그리고 $\sigma^2(t)$의 추정량을 다음과 같이 제안하였다.

$$\hat{\sigma}^2(t) = n(\hat{S}(t))^{-2} \int_t^{T^*} (\int_s^{T^*} \hat{S}(u) du)^2 d\hat{V}(s)$$

<u>따름정리 2. 1</u> 분포함수 F가 연속이고, 정리 2.2의 가정이 만족한다면, $0 \le t \le T^*$인 각 t에대해

$$\lim_{n \to \infty} \Pr\{ \hat{e}_t - l \le e_t^{\tau_F} \le \hat{e}_t + l \} = 1 - \alpha$$

여기서 $l = z_{\alpha/2} \sqrt{\hat{\sigma}^2(t)}$이고, $z_{\alpha/2}$는 표준정규분포의 100α분위수이다.

3. 모의실험및 그 결과

지금부터 위에 제안된 추정량들 $\hat{e}_t^{KM}$, $\hat{e}_t^{NA}$, $\hat{e}_t^{SV}$, $\hat{e}_t$에대한 소표본 특성을 컴퓨터를 이용한 모의실험을 통하여 추정된 편의와 평균제곱오차 측면을 고려하여 비교하고자 한다.

모의실험에서 수명분포(F)는 모수 α(척도모수), β(형상모수)에따라 다양한 형태를 갖는 와이블(Weibull)분포를 가정하고 위험률함수가 증가함수인 경우(IFR, $\alpha = 2.0$, $\beta = 1.5$), 상수함수인경우(Constant FR, $\alpha = 1.0, \beta = 1.0$), 감소함수인 경우(DFR, $\alpha = 1.0, \beta = 0.9$)를 가정하였다. 그리고 중단분포(G)는 지수분포(Exponential)와 일양분포(Uniform)분포로 가정하여 중단비율(censoring rate)이 약 10%와 30%가 되도록 중단분포의 모수를 가정하였다. 가정된 자료의 수는 20, 30, 50으로 1000번을 반복하였고 사용된 난수는 IMSL의 부 프로그램인 "GGUBS"로부터 얻었고 IBM 386pc를 이용하였다.

모의실험의 결과는 표 3-1에 보였다. 표 3-1로부터 다음의 결과를 얻을 수 있었다.

(1)수명함수 F의 분포가 Weibull(1.0, 0.9)일 때 $\hat{e}_t^{KM}$와 $\hat{e}_t^{NA}$이 $\hat{e}_t^{SV}$, $\hat{e}_t$보다 더 작은 편의를 가졌다. 특히 중단비율이 30%인 경우 $\hat{e}_t^{NA}$의 편의가 가장 우수하였다. 그 경우에 $\hat{e}_t^{SV}$와 $\hat{e}_t$를 비교해볼때 편의의 측면에서는 $\hat{e}_t$이 $\hat{e}_t^{SV}$보다 더 적은 편의를 가졌다. 중단비율이 10%인 경우에는 $S(t)$가 0.3 에서 0.1 까지의 영역에서는

$\hat{e}_{t^{NA}}$의 편의가 가장작고 그 외의 부분에서는 $\hat{e}_{t^{KM}}$의 편의가 더 적었다.

평균제곱오차의 측면에서 볼 때, $\hat{e}_{t^{SV}}$가 가장 우수하였고 $\hat{e}_t$는 $\hat{e}_{t^{KM}}$, $\hat{e}_{t^{NA}}$보다 더 우수하였다.

(2)수명함수 F의 분포가 Weibull(1.0, 1.0)인 경우, 평균제곱오차의 측면에서 $\hat{e}_{t^{SV}}$와 $\hat{e}_t$는 $\hat{e}_{t^{KM}}$, $\hat{e}_{t^{NA}}$보다 더욱 우수하였고, 특히, 중단비율이 10%인경우 표본의 크기에 상관없이 $\hat{e}_t$는 $\hat{e}_{t^{SV}}$보다 수명분포의 끝부분에서 더 작은 평균제곱오차를 가졌다.

편의의 측면에서볼때 $\hat{e}_{t^{KM}}$, $\hat{e}_{t^{NA}}$이 $\hat{e}_{t^{SV}}$와 $\hat{e}_t$보다 더 작은 편의를 가졌다. $\hat{e}_{t^{KM}}$은 $\hat{e}_{t^{NA}}$보다 작은 편의를 가졌다. $\hat{e}_{t^{SV}}$와 $\hat{e}_t$를 비교해볼때 중단 비율이 30%인 경우, $\hat{e}_t$의 편의가 $\hat{e}_{t^{SV}}$보다 작고 중단비율이 10%인 경우는 $\hat{e}_{t^{SV}}$의 편의가 $\hat{e}_t$보다 작았다.

(3)수명함수 F의 분포가 Weibull(2.0, 1.5)일 때, 표본의 크기와 중단비율에 관계없이 $\hat{e}_t$와 $\hat{e}_{t^{SV}}$의 평균제곱오차는 $\hat{e}_{t^{KM}}$나 $\hat{e}_{t^{NA}}$보다 작았다. 그러나 편의의 측면에서 볼 때 $\hat{e}_{t^{KM}}$, $\hat{e}_{t^{SV}}$의 편의가 $\hat{e}_{t^{NA}}$와 $\hat{e}_t$보다 작고 특히 $\hat{e}_{t^{KM}}$의 편의가 가장 작은 것을 알 수있었다.

결론적으로 편의나 평균제곱오차의 측면에서는 중단비율이 높다고 생각되는 자료에서는 $\hat{e}_t$의 사용이 더 바람직하다고 할 수있다.

표 3-1 추정된 편의와 평균제곱오차

F : *Weibull*(1.0, 1.0)

G : Exponential(0.111) *Censoring rate* : 10%

G : Exponential(0.429) *Censoring rate* : 30%

n = 20

$S(t)$		$\hat{e}_t^{NA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$	$\hat{e}_t^{NA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$
0.1	bias	-.2486	-.3271	-.2896	-.2951	-.6314	-.6490	-.6404	-.6509
	mse	.6950	.6537	.6673	.6637	.7227	.7180	.7192	.7143
0.2	bias	-.0115	-.1712	-.0959	-.1197	-.2839	-.3533	-.3194	-.3679
	mse	.4687	.4035	.4229	.4014	.4706	.4592	.4612	.4353
0.3	bias	.0550	-.1277	-.0427	-.0764	-.1128	-.2296	-.1730	-.2560
	mse	.2939	.2298	.2474	.2140	.3627	.3450	.3488	.2767
0.4	bias	.0714	-.1047	-.0232	-.0568	-.0403	-.1751	-.1102	-.2077
	mse	.2237	.1749	.1873	.1603	.2747	.2539	.2586	.2041
0.5	bias	.0711	-.0961	-.0188	-.0493	-.0142	-.1526	-.0861	-.1838
	mse	.1747	.1308	.1421	.1280	.2066	.1842	.1891	.1548
0.6	bias	.0685	-.0871	-.0150	-.0439	.0027	-.1311	-.0667	-.1625
	mse	.1401	.1053	.1138	.1022	.1695	.1501	.1547	.1240
0.7	bias	.0652	-.0787	-.0117	-.0383	.0106	-.1170	-.0553	-.1465
	mse	.1173	.0909	.0970	.0874	.1411	.1235	.1277	.1034
0.8	bias	.0606	-.0738	-.0109	-.0359	.0139	-.1073	-.0484	-.1346
	mse	.0941	.0737	.0779	.0709	.1149	.1008	.1038	.0839
0.9	bias	.0583	-.0679	-.0085	-.0320	.0164	-.0984	-.0422	-.1235
	mse	.0820	.0652	.0685	.0626	.0980	.0865	.0887	.0739

n = 30

$S(t)$		$\hat{e}_t^{NA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$	$\hat{e}_t^{NA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$
0.1	bias	-.0467	-.1688	-.1107	-.1339	-.4458	-.4821	-.4641	-.4925
	mse	.6921	.6282	.6489	.6238	.6166	.6143	.6136	.5913
0.2	bias	.0815	-.1092	-.0203	-.0501	-.1124	-.2201	-.1677	-.2504
	mse	.3799	.2714	.3069	.2913	.3822	.3613	.3665	.3113
0.3	bias	.0905	-.0934	-.0083	-.0427	.0013	-.1424	-.0731	-.1756
	mse	.2119	.1416	.1617	.1414	.2818	.2437	.2555	.2040
0.4	bias	.0800	-.0852	-.0087	-.0404	.0166	-.1332	-.0612	-.1642
	mse	.1530	.1044	.1175	.0987	.1964	.1640	.1728	.1353
0.5	bias	.0728	-.0746	-.0060	-.0344	.0226	-.1178	-.0502	-.1466
	mse	.1056	.0784	.0841	.0733	.1396	.1143	.1206	.0970
0.6	bias	.0662	-.0676	-.0049	-.0304	.0269	-.1034	-.0402	-.1297
	mse	.0840	.0638	.0677	.0606	.1081	.0891	.0933	.0765
0.7	bias	.0606	-.0619	-.0040	-.0274	.0279	-.0934	-.0342	-.1168
	mse	.0709	.0558	.0584	.0527	.0914	.0750	.0786	.0657
0.8	bias	.0556	-.0578	-.0038	-.0252	.0287	-.0843	-.0286	-.1057
	mse	.0565	.0448	.0465	.0426	.0762	.0626	.0655	.0549
0.9	bias	.0518	-.0538	-.0031	-.0231	.0269	-.0787	-.0262	-.0980
	mse	.0484	.0392	.0403	.0372	.0612	.0514	.0531	.0455

n = 50

$S(t)$		$\hat{e}_t^{NA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$	$\hat{e}_t^{NA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$
0.1	bias	.0678	-.1228	-.0335	-.0645	-.3416	-.4193	-.3812	-.4377
	mse	.4526	.3372	.3765	.3515	.5213	.5215	.5180	.4642
0.2	bias	.1112	-.0753	.0110	-.0220	-.0498	-.1789	-.1168	-.2065
	mse	.2255	.1493	.1731	.1501	.2786	.2545	.2605	.2141
0.3	bias	.0982	-.0596	.0138	-.0162	-.0065	-.1397	-.0755	-.1718
	mse	.1554	.1052	.1207	.1049	.1886	.1664	.1720	.1391
0.4	bias	.0867	-.0488	.0149	-.0112	.0214	-.1033	-.0429	-.1297
	mse	.1155	.0811	.0917	.0806	.1382	.1170	.1228	.1000
0.5	bias	.0711	-.0475	.0089	-.0135	.0197	-.0933	-.0379	-.1163
	mse	.0711	.0518	.0568	.0515	.1011	.0865	.0900	.0750
0.6	bias	.0632	-.0431	.0081	-.0118	.0221	-.0809	-.0297	-.1008
	mse	.0559	.0420	.0454	.0420	.0744	.0643	.0664	.0564
0.7	bias	.0548	-.0415	.0054	-.0124	.0190	-.0754	-.0278	-.0925
	mse	.0441	.0343	.0364	.0343	.0617	.0546	.0558	.0484
0.8	bias	.0509	-.0378	.0059	-.0105	.0210	-.0667	-.0219	-.0814
	mse	.0374	.0294	.0311	.0292	.0549	.0482	.0496	.0428
0.9	bias	.0465	-.0358	.0052	-.0099	.0199	-.0619	-.0196	-.0744
	mse	.0319	.0256	.0268	.0251	.0454	.0402	.0411	.0363

F : *Weibull*(1.0, 0.9)

G : Exponential(0.105) *Censoring rate* : 10%

G : Exponential(0.423) *Censoring rate* : 30%

n = 20

$S(t)$		$\hat{e}_{t^{KM}}$	$\hat{e}_t$	$\hat{e}_{t^{KM}}$	$\hat{e}_{t^{SV}}$	$\hat{e}_{t^{KM}}$	$\hat{e}_t$	$\hat{e}_{t^{KM}}$	$\hat{e}_{t^{SV}}$
0.1	bias	-.3455	-.4412	-.3954	-.4039	-.8517	-.8683	-.8602	-.8731
	mse	1.0881	1.0289	1.0480	1.0395	1.1561	1.1563	1.1551	1.1551
0.2	bias	-.0359	-.2297	-.1384	-.1692	-.4002	-.4744	-.4379	-.5048
	mse	.6937	.5991	.6268	.5973	.7135	.7134	.7097	.6730
0.3	bias	.0513	-.1697	-.0669	-.1090	-.1787	-.3087	-.2453	-.3536
	mse	.4408	.3432	.3704	.3246	.5221	.5163	.5139	.4150
0.4	bias	.0744	-.1371	-.0392	-.0814	-.0830	-.2346	-.1611	-.2873
	mse	.3308	.2571	.2762	.2399	.3820	.3664	.3679	.2949
0.5	bias	.0754	-.1242	-.0319	-.0694	-.0475	-.2036	-.1283	-.2507
	mse	.2551	.1905	.2072	.1872	.2808	.2621	.2646	.2215
0.6	bias	.0725	-.1122	-.0266	-.0618	-.0250	-.1753	-.1026	-.2204
	mse	.1999	.1491	.1616	.1459	.2241	.2073	.2098	.1723
0.7	bias	.0695	-.1006	-.0215	-.0535	-.0121	-.1550	-.0857	-.1964
	mse	.1638	.1259	.1346	.1219	.1836	.1679	.1703	.1416
0.8	bias	.0645	-.0937	-.0198	-.0495	-.0063	-.1414	-.0755	-.1784
	mse	.1303	.1011	.1071	.0977	.1492	.1370	.1383	.1171
0.9	bias	.0619	-.0858	-.0165	-.0441	-.0014	-.1288	-.0663	-.1623
	mse	.1111	.0875	.0922	.0845	.1264	.1165	.1173	.0992

n = 30

$S(t)$		$\hat{e}_{t^{KM}}$	$\hat{e}_t$	$\hat{e}_{t^{KM}}$	$\hat{e}_{t^{SV}}$	$\hat{e}_{t^{KM}}$	$\hat{e}_t$	$\hat{e}_{t^{KM}}$	$\hat{e}_{t^{SV}}$
0.1	bias	-.0747	-.2270	-.1543	-.1908	-.6441	-.6784	-.6613	-.7003
	mse	1.1127	1.0018	1.0395	.9938	.9784	.9854	.9800	.9516
0.2	bias	.0936	-.1432	-.0326	-.0788	-.2005	-.3186	-.2609	-.3661
	mse	.6067	.4322	.4904	.4436	.5892	.5934	.5857	.4985
0.3	bias	.1050	-.1217	-.0167	-.0631	-.0426	-.2051	-.1263	-.2576
	mse	.3305	.2183	.2510	.2177	.4150	.3825	.3903	.3127
0.4	bias	.0930	-.1092	-.0155	-.0582	-.0151	-.1867	-.1039	-.2345
	mse	.2383	.1593	.1813	.1501	.2771	.2433	.2511	.2003
0.5	bias	.0841	-.0952	-.0118	-.0494	-.0007	-.1608	-.0834	-.2056
	mse	.1580	.1147	.1243	.1070	.1974	.1720	.1769	.1430
0.6	bias	.0770	-.0850	-.0091	-.0428	.0072	-.1407	-.0688	-.1801
	mse	.1242	.0920	.0987	.0865	.1485	.1303	.1329	.1108
0.7	bias	.0706	-.0771	-.0074	-.0378	.0135	-.1240	-.0566	-.1592
	mse	.1034	.0791	.0838	.0742	.1269	.1099	.1128	.0939
0.8	bias	.0644	-.0717	-.0070	-.0347	.0164	-.1111	-.0481	-.1428
	mse	.0824	.0633	.0666	.0597	.1060	.0921	.0945	.0777
0.9	bias	.0594	-.0667	-.0064	-.0317	.0155	-.1034	-.0442	-.1310
	mse	.0682	.0534	.0557	.0505	.0821	.0723	.0735	.0631

n = 50

$S(t)$		$\hat{e}_{t^{KM}}$	$\hat{e}_t$	$\hat{e}_{t^{KM}}$	$\hat{e}_{t^{SV}}$	$\hat{e}_{t^{KM}}$	$\hat{e}_t$	$\hat{e}_{t^{KM}}$	$\hat{e}_{t^{SV}}$
0.1	bias	.0651	-.1721	-.0608	-.1054	-.5035	-.5869	-.5456	-.6214
	mse	.7235	.5518	.6097	.5562	.8655	.8721	.8649	.7805
0.2	bias	.1301	-.1015	.0058	-.0411	-.0997	-.2439	-.1738	-.3018
	mse	.3520	.2340	.2710	.2314	.4554	.4239	.4329	.3423
0.3	bias	.1181	-.0767	.0140	-.0287	-.0348	-.1897	-.1145	-.2456
	mse	.2397	.1620	.1863	.1581	.2932	.2628	.2708	.2208
0.4	bias	.1049	-.0615	.0167	-.0194	.0052	-.1399	-.0690	-.1874
	mse	.1689	.1177	.1333	.1157	.2094	.1812	.1891	.1522
0.5	bias	.0868	-.0585	.0106	-.0206	.0077	-.1235	-.0587	-.1643
	mse	.1086	.0778	.0861	.0763	.1501	.1305	.1354	.1119
0.6	bias	.0770	-.0526	.0097	-.0174	.0124	-.1069	-.0472	-.1417
	mse	.0829	.0609	.0665	.0607	.1088	.0957	.0984	.0833
0.7	bias	.0662	-.0506	.0062	-.0178	.0114	-.0978	-.0424	-.1273
	mse	.0640	.0488	.0522	.0486	.0895	.0797	.0815	.0705
0.8	bias	.0612	-.0459	.0067	-.0151	.0149	-.0862	-.0344	-.1116
	mse	.0534	.0411	.0439	.0406	.0781	.0688	.0708	.0609
0.9	bias	.0552	-.0437	.0054	-.0145	.0144	-.0795	-.0308	-.1012
	mse	.0450	.0351	.0372	.0346	.0640	.0568	.0582	.0511

F : *Weibull*(2.0, 1.5)

G : *Uniform*(18.05) *Censoring rate* : 10%

G : *Uniform*(6.005) *Censoring rate* : 30%

n = 20

$S(t)$		$\hat{e}_t^{KA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$	$\hat{e}_t^{KA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$
0.1	bias	-.1578	-.2380	-.1996	-.2057	-.4252	-.4517	-.4388	-.4505
	rmse	.5312	.4888	.5032	.4973	.5997	.5965	.5969	.5833
0.2	bias	.0331	-.1295	-.0531	-.0709	-.1426	-.2354	-.1906	-.2343
	rmse	.4216	.3638	.3803	.3505	.4541	.4317	.4380	.3815
0.3	bias	.0876	-.1002	-.0130	-.0406	-.0087	-.1536	-.0842	-.1534
	rmse	.3124	.2512	.2683	.2150	.3925	.3546	.3660	.2756
0.4	bias	.0990	-.0855	-.0001	-.0281	.0433	-.1236	-.0442	-.1219
	rmse	.2625	.2132	.2262	.1795	.3303	.2877	.3003	.2294
0.5	bias	.0901	-.0883	-.0058	-.0299	.0551	-.1177	-.0357	-.1163
	rmse	.1850	.1435	.1531	.1385	.2522	.2135	.2237	.1764
0.6	bias	.0867	-.0821	-.0037	-.0274	.0647	-.1039	-.0237	-.1069
	rmse	.1582	.1248	.1321	.1173	.2215	.1869	.1961	.1528
0.7	bias	.0824	-.0765	-.0023	-.0250	.0644	-.0979	-.0204	-.1024
	rmse	.1354	.1112	.1154	.1063	.1871	.1602	.1664	.1342
0.8	bias	.0778	-.0733	-.0022	-.0244	.0613	-.0949	-.0197	-.1013
	rmse	.1141	.0957	.0981	.0912	.1549	.1336	.1377	.1156
0.9	bias	.0747	-.0696	-.0012	-.0233	.0620	-.0884	-.0153	-.0979
	rmse	.1048	.0902	.0915	.0863	.1425	.1243	.1275	.1081

n = 30

$S(t)$		$\hat{e}_t^{KA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$	$\hat{e}_t^{KA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$
0.1	bias	-.0094	-.1280	-.0716	-.0876	-.2362	-.2957	-.2669	-.2878
	rmse	.4790	.4346	.4481	.4341	.5419	.5236	.5287	.5111
0.2	bias	.0946	-.0863	-.0020	-.0245	.0085	-.1298	-.0637	-.1158
	rmse	.2936	.2189	.2423	.2179	.3876	.3356	.3525	.3062
0.3	bias	.0999	-.0774	.0046	-.0193	.0785	-.0921	-.0113	-.0792
	rmse	.1880	.1361	.1502	.1315	.2957	.2371	.2554	.1989
0.4	bias	.0896	-.0731	.0024	-.0202	.0803	-.0928	-.0109	-.0863
	rmse	.1444	.1098	.1180	.1008	.2409	.1913	.2062	.1524
0.5	bias	.0831	-.0657	.0038	-.0169	.0739	-.0900	-.0124	-.0846
	rmse	.1087	.0869	.0908	.0826	.1552	.1210	.1292	.1043
0.6	bias	.0766	-.0610	.0038	-.0154	.0727	-.0812	-.0077	-.0779
	rmse	.0938	.0770	.0797	.0735	.1290	.1029	.1084	.0895
0.7	bias	.0713	-.0573	.0039	-.0146	.0688	-.0761	-.0062	-.0744
	rmse	.0842	.0715	.0731	.0682	.1126	.0919	.0959	.0812
0.8	bias	.0654	-.0558	.0025	-.0153	.0659	-.0712	-.0044	-.0709
	rmse	.0692	.0596	.0603	.0569	.0968	.0797	.0828	.0714
0.9	bias	.0620	-.0532	.0029	-.0149	.0610	-.0692	-.0049	-.0707
	rmse	.0647	.0570	.0573	.0542	.0800	.0682	.0694	.0617

n = 50

$S(t)$		$\hat{e}_t^{KA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$	$\hat{e}_t^{KA}$	$\hat{e}_t$	$\hat{e}_t^{KM}$	$\hat{e}_t^{SV}$
0.1	bias	.0755	-.0929	-.0145	-.0234	-.1384	-.2449	-.1937	-.2301
	rmse	.3085	.2365	.2588	.2531	.3755	.3547	.3588	.3385
0.2	bias	.1016	-.0678	.0103	-.0082	.0546	-.0998	-.0268	-.0879
	rmse	.1652	.1170	.1304	.1118	.2553	.2072	.2223	.1903
0.3	bias	.0907	-.0554	.0124	-.0052	.0593	-.0912	-.0201	-.0790
	rmse	.1328	.0982	.1081	.0938	.1642	.1323	.1405	.1228
0.4	bias	.0798	-.0482	.0121	-.0034	.0721	-.0669	-.0007	-.0565
	rmse	.1005	.0765	.0831	.0754	.1307	.1056	.1119	.0975
0.5	bias	.0651	-.0493	.0054	-.0080	.0579	-.0687	-.0076	-.0599
	rmse	.0633	.0512	.0533	.0504	.0948	.0795	.0824	.0743
0.6	bias	.0595	-.0449	.0057	-.0067	.0569	-.0604	-.0030	-.0529
	rmse	.0533	.0447	.0459	.0442	.0748	.0629	.0649	.0590
0.7	bias	.0521	-.0445	.0030	-.0088	.0468	-.0619	-.0078	-.0552
	rmse	.0461	.0402	.0406	.0394	.0650	.0578	.0582	.0545
0.8	bias	.0492	-.0414	.0039	-.0075	.0475	-.0551	-.0032	-.0491
	rmse	.0414	.0366	.0368	.0359	.0605	.0533	.0541	.0510
0.9	bias	.0456	-.0401	.0034	-.0079	.0427	-.0545	-.0044	-.0492
	rmse	.0 86	.0349	.0349	.0339	.0525	.0474	.0475	.0459

참 고 문 헌

(1) Aalen, O. (1978), Nonparametric inference for a family of counting processes. Annals of Statistics, 6, 701-726.

(2) Anderson, P.K. and Borgan, ϕ. (1985), Counting process models for life history data : A review, Scandinavian Journal of Statistics, 12, 97-158.

(3) Bréslow,N.E. and Crowley, J. (1974), A large sample study of the life table and product limit estimators under random censorship. Annals of Statistics, 2, 435-453.

(4) Fleming, T.R. and Harrington, D.P. (1991), Counting Process and Survival Analysis, John Wiley & Sons Ins., New York.

(5) Ghorai, J.K., Susarla, A., Susarla, V. and Van Ryzin, J.(1980), Nonparametric estimation of mean residual life time with censored data. Colloquia Mathetica Societatis Janos Bolyai, Nonparametric Statistical Inference, 32, 269-291.

(6) Gill, R.D. (1983), Large sample behavior of the product-limit estimator on the whole line. Annals of Statistics, 11, 49-58.

(7) Hjort, N.L. (1991), Nonparametric Bayes estimators based on beta processes in models for life history data. Annals of Statistics, 3, 1259-1294.

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

(9) Kumazawa, Y.(1987), A note on an estimator of life expectancy with random censorship. Biometrika, 74, 655-658.

(10) Lee I.S. and Lee, W.D. (1993), Nonparametric Bayes estimation of survival function, Unpublished ph. D. thesis.

(11) Nelson ,W.B.(1972), Theory and applications of hazard plotting for censored failure data. Technometrics, 14, 945-996.

(12) Peterson, A.V. (1977), Expressing the Kaplan-Meier estimator as a function of empirical survival functions. Journal of the American Statistical Association, 72, 854-858.

(13) Susarla, V. and Van Ryzin, J. (1980), Large sample theory for an estimator of the mean survival time from censored samples. Annals of Statistics, 8, 1002-1016.

(14) Yang, G.L. (1977), Life expectancy under random censorship. Stochastic process and their applications, 6, 33-39.

(15) Yang, G.L. (1978), Estimation of biometric function. Annals of Statistics,6,112-116.