한국통계학회논문집 제3권 1호
1996년 4월 pp. 61-71

모의실험을 통한 가산위험모형에 대한 적합도검정법들의 비교

김 진 흠[1]

요 약

Kim and Song(1995)과 Kim and Lee(1996)는 하나의 이진공변량(binary covariate)을 갖는 가산위험모형(additive risk model)의 적합도검정법(goodness-of-fit test)을 제안했다. 전자는 모수의 가중추정량들의 차에 기초한 검정법이며 후자는 마팅게일잔차(martingale residual)에 기초한 검정법이다. 본 논문에서는 모의실험을 통하여 두 검정법을 비교하였다.

1. 서 론

Lin and Ying(1994)은 생존시간과 공변량 사이의 관계를 연구하기 위하여 위험함수가 공변량에 대하여 가산적으로 증가 또는 감소하는 다음과 같은 형태의 가산위험모형을 고려하였다.

$$\lambda(t; Z) = \lambda_0(t) + \beta_0' Z(t). \tag{1.1}$$

여기서, $\lambda_0(t)$는 미지의 기저위험함수(baseline hazard function)이고, β_0는 $p\times 1$벡터 회귀모수이며, $Z(t)$는 $p\times 1$벡터 시간가변공변량(time-varying covariate)이다. 모형 (1.1)은 생존분석학에서 널리 적용되고 있는 Cox의 비례위험모형이 타당하지 않을 때 하나의 대안으로서 뿐만 아니라 역학(epidemiology) 등의 분야에서 흔히 사용되고 있는 모형이다 (Buckley(1984)).

Kim and Song(1995)과 Kim and Lee(1996)는 하나의 이진공변량 만을 갖는 모형 (1.1)의 특별한 경우에 대하여 모형의 적합성을 검정할 수 있는 검정법을 제안하였다. 다시 말해서, $p=1$이고 $Z(t)$가 0과 1의 값만을 취하는 경우에 대하여 가산위험가정의 타당성을 검정하기 위한 검정법을 제안하였다. Kim and Song(1995)과 Kim and Lee(1996)가 고려한 모형은 다음과 같다.

$$\lambda(t;Z) = \lambda_0(t) + \beta_0 Z. \tag{1.2}$$

Kim and Song(1995)이 제안한 검정법은 모수 β_0에 대한 서로 다른 가중추정량들의 차에 기초한 것이며, Kim and Lee(1996)가 제안한 검정법은 마팅게일잔차에 기초한 것이다. 전자는 비례위험모형의 적합도검정을 위하여 제안된 Gill and Schumacher(1987)의 방법을 가산위험모형으로 확장한 형태이며, 후자는 Wei(1984)의 방법을 가산위험모형으로 확장한 형태이다.

---

1) (445-743) 경기도 화성군 봉담면 와우리 산 2-2 수원대학교 응용통계학과 전임강사.

본 논문에서는 모의실험을 통하여 모수 β_0에 대한 Lin and Ying(1994)의 추정량과 Kim and Song(1995)의 가중추정량을 비교하였고, 표본의 크기가 50으로 소표본인 경우에 대하여 Kim and Song(1995)과 Kim and Lee(1996)가 제안한 적합도검정법들을 비교하였다.

2. 적합도 검정통계량의 비교

각 i $(i=1,\dots,n)$에 대하여 $X_i=\min(T_i, C_i)$, $\Delta_i=I(T_i\le C_i)$, $Z_i=I(i\in$ 그룹 $1)$라 하자. 여기서, T_i는 절대연속인 분포함수 $F(\cdot)$를 갖는 생존시간이고, C_i는 T_i에 대응하는 중도절단시간(censoring time)이다. T_i와 C_i는 서로 독립이라고 가정하자.

2-1. 마팅게일잔차에 기초한 검정법

Lin and Ying(1994)은 모수 β_0에 대한 일치추정량으로 다음을 제안하였다.

$$\hat{\beta}=\hat{\Omega}(\infty)^{-1}\sum_{i=1}^{n}\int_0^{\infty}[Z_i-\bar{Z}(u)]dN_i(u). \tag{2.1}$$

여기서, $N_i(t)=\Delta_i I(X_i\le t)$ 이고 $Y_i(t)=I(X_i\ge t)$ 이다. 또한, $\bar{Z}(t)=\sum_{i=1}^{n}[Y_i(t)/\sum_{j=1}^{n}Y_j(t)]Z_i$ 이고 $\hat{\Omega}(t)=n^{-1}\sum_{i=1}^{n}\int_0^t Y_i(u)[Z_i-\bar{Z}(u)]^2du$ 이다.

먼저, Kim and Lee(1996)가 제안한 검정통계량 계산의 편의를 위하여 Kim and Lee(1996)가 제안한 관찰과정(observed process)과 시뮬레이션과정(simulated process)의 순환식(recursive formula)을 살펴보고자 한다.

Kim and Lee(1996)가 제안한 관찰과정 $U(\hat{\beta},t)$는

$$U(\hat{\beta},t)=\sum_{i=1}^{n}\Delta_i I(X_i\le t)[Z_i-\bar{Z}(X_i)]-\hat{\beta}\sum_{i=1}^{n}\int_0^t Y_i(u)[Z_i-\bar{Z}(u)]^2du$$

와 같이 표현할 수 있다. 만약 X_k $(k=1,\dots,n)$들이 순서화 되었다고 가정하면 $t=X_k$ 에서 $U(\hat{\beta},t)$는 다음과 같다.

$$U(\hat{\beta},X_k)=\sum_{i=1}^{k}\Delta_i[Z_i-\bar{Z}(X_i)]-n\hat{\beta}\,\hat{\Omega}(X_k)$$

이고

$$\hat{\Omega}(X_k)=\hat{\Omega}(X_{k-1})+n^{-1}(X_k-X_{k-1})\sum_{i=k}^{n}[Z_i-\bar{Z}(X_k)]^2$$

이다. 따라서,

$$\begin{aligned} U(\hat{\beta}, X_k) &= U(\hat{\beta}, X_{k-1}) + \Delta_k[Z_k - \overline{Z}(X_k)] - n\hat{\beta}[\hat{\Omega}(X_k) - \hat{\Omega}(X_{k-1})] \\ &= U(\hat{\beta}, X_{k-1}) + \Delta_k[Z_k - \overline{Z}(X_k)] - \hat{\beta}(X_k - X_{k-1}) \sum_{i=k}^{n}[Z_i - \overline{Z}(X_k)]^2 \end{aligned}$$

이다. 여기서, $X_0 \equiv 0$, $U(\hat{\beta}, X_0) \equiv 0$ 라 하자.

또한, Kim and Lee(1996)가 제안한 시뮬레이션과정(simulated process) $\hat{U}(\hat{\beta}, t)$는 다음과 같다.

$$\hat{U}(\hat{\beta}, t) = \sum_{i=1}^{n} \Delta_i G_i [I(X_i \le t) - \hat{\Omega}(t)\hat{\Omega}(\infty)^{-1}][Z_i - \overline{Z}(X_i)].$$

여기서, G_i $(i=1, \dots, n)$는 표준정규확률변수이다. $t = X_k$에 대하여, 위 식은

$$\hat{U}(\hat{\beta}, X_k) = \sum_{i=1}^{k} \Delta_i G_i [Z_i - \overline{Z}(X_k)] - \hat{\Omega}(X_k)\hat{\Omega}(X_n)^{-1} \sum_{i=1}^{n} \Delta_i G_i [Z_i - \overline{Z}(X_i)]$$

이 된다. 그러므로, 우리는 다음의 순환식을 이용하여 $\hat{U}(\hat{\beta}, X_k)$ $(k=1, \dots, n)$을 쉽게 계산할 수 있다.

$$\begin{aligned} &\hat{U}(\hat{\beta}, X_k) \\ &= \hat{U}(\hat{\beta}, X_{k-1}) + \Delta_k G_k [Z_k - \overline{Z}(X_k)] - [\hat{\Omega}(X_k) - \hat{\Omega}(X_{k-1})]\hat{\Omega}(X_n)^{-1} \\ &\qquad \times \sum_{i=1}^{n} \Delta_i G_i [Z_i - \overline{Z}(X_i)] \\ &= \hat{U}(\hat{\beta}, X_{k-1}) + \Delta_k G_k [Z_k - \overline{Z}(X_k)] - n^{-1}(X_k - X_{k-1}) \sum_{i=k}^{n}[Z_i - \overline{Z}(X_k)]^2 \hat{\Omega}(X_n)^{-1} \\ &\qquad \times \sum_{i=1}^{n} \Delta_i G_i [Z_i - \overline{Z}(X_i)]. \end{aligned}$$

여기서, $X_0 \equiv 0$, $\hat{U}(\hat{\beta}, X_0) \equiv 0$ 라 하자.

Kim and Lee(1996)는 모형 (1.2)의 적합도검정을 위하여 다음과 같은 검정통계량을 제안했다.

$$S = \max_{1 \le k \le n} \Sigma^{-1/2} |n^{-1/2} U(\hat{\beta}, X_k)|. \tag{2.2}$$

여기서, $\Sigma = n^{-1} \sum_{i=1}^{n} \Delta_i [Z_i - \overline{Z}(X_i)]^2$ 이다. 또한, Kim and Lee(1996)는 Lin, Ying and Wei(1993)의 방법을 사용하여 검정통계량 S의 관측값, s의 유의확률(p-value) $P(S \ge s)$을 $P(\hat{S} \ge s)$로 근사시키는 방법을 제안했다. 여기서, $\hat{S} = \max_{1 \le k \le n} \Sigma^{-1/2} |n^{-1/2} \hat{U}(\hat{\beta}, X_k)|$ 이다. 다시 말해서, (X_i, Z_i, Δ_i) $(i=1, \dots, n)$이 주어진 조건하에서 표준정규확률변수 G_i를 표본의 크

기만큼 n개 생성한 후 $\hat{S}$를 계산하여 그것을 $\hat{s}_1$라 하자. 동일한 시행을 $B(>0)$번 반복하여 $\hat{s}_1, \cdots, \hat{s}_B$을 얻었을 때 $P(\hat{S} \geq s)$을 B번의 시행 중에서 " $\hat{s}_b \geq s$ $(b=1, \cdots, B)$ "을 만족하는 상대도수로서 정의하는 방법이다. 따라서, 만약 근사시켜 구한 p값이 가정한 유의수준 보다 작으면 모형 (1.2)의 가정은 타당하지 않다고 결론 내릴 수 있다.

2-2. 가중추정량들의 차에 기초한 검정법

한편, Kim and Song(1995)은 모수 β_0에 대한 가중추정량

$$\hat{\beta}_K = \int_0^\infty \frac{K(u)}{\int_0^\infty K(v)dv} \left[\frac{d\,\overline{N}_1(u)}{d\,\overline{Y}_1(u)} - \frac{d\,\overline{N}_2(u)}{d\,\overline{Y}_2(u)} \right] du \tag{2.3}$$

을 제안하였다. 여기서, $\overline{Y}_1(t) = \sum_{i=1}^n I(Z_i=1)Y_i(t)$, $\overline{Y}_2(t) = \sum_{i=1}^n I(Z_i=0)Y_i(t)$, $\overline{N}_1(t) = \sum_{i=1}^n I(Z_i=1)N_i(t)$, $\overline{N}_2(t) = \sum_{i=1}^n I(Z_i=0)N_i(t)$ 이다. 또한, $K(t)$는 예측 가능(predictable) 가중함수이고 $\overline{Y}_1(t)\ \overline{Y}_2(t)=0$ 이면 0인 함수이다. Kim and Song(1995)은 서로 다른 가중함수 K_1, K_2를 사용하여 얻은 가중추정량들의 차에 기초한 검정법을 제안하였다. 그 검정통계량은 다음과 같다.

$$T_{K_1K_2} = n^{1/2}\ \hat{\sigma}_{K_1K_2}{}^{-1} D_{K_1K_2}. \tag{2.4}$$

여기서,

$$D_{K_1K_2} = \hat{\beta}_{K_1} - \hat{\beta}_{K_2},$$

$$\hat{\sigma}_{K_1K_2}{}^2 = n(\hat{V}_{11} - \hat{V}_{12} - \hat{V}_{21} + \hat{V}_{22}),$$

$$\hat{V}_{jj'} = \int_0^\infty \frac{K_j(u)K_{j'}(u)}{\int_0^\infty K_j(v)dv \int_0^\infty K_{j'}(v)dv} \left[\frac{d\,\overline{N}_1(u)}{\overline{Y}_1(u)^2} + \frac{d\,\overline{N}_2(u)}{\overline{Y}_2(u)^2} \right] \ (j, j'=1,2).$$

따라서, 모수 β_0의 두 가중추정량 $\hat{\beta}_{K_1}$, $\hat{\beta}_{K_2}$ 모두가 일치추정량이기 때문에 가정한 모형이 타당하지 않으면 검정통계량 $T_{K_1K_2}$의 값이 매우 큰 값을 가질 것이다 (Kim and Song(1995)). 한편, 검정통계량 $T_{K_1K_2}$의 점근정규성을 이용하면 모형 (1.2)의 적합도검정을 쉽게 할 수 있다 (Kim and Song(1995)).

3. 모의실험

본 절에서는 모수 β_0에 대한 Lin and Ying(1994)의 추정량과 Kim and Song(1995)의 가중추정량을 비교하고, $n=50$ 인 소표본인 경우에 대하여 Kim and Song(1995)과 Kim and Lee(1996)가 제안한 적합도검정법들을 비교하기 위하여 모의실험을 하였다. 본 모의실험에서는 가중함수 $K(\cdot)$로 초기 사망에 가중값을 많이 부여하는 Gehan의 가중함수, $K_G(t)=[\overline{Y}_1(t)+\overline{Y}_2(t)]\cdot L(t)$, 시간에 관계없이 동일한 가중값을 주는 Log-rank 가중함수, $K_L(t)=1\cdot L(t)$, 그리고 Gehan의 가중함수에 비하여 중도절단시간의 분포가 고려된 Prentice-Wilcoxon의 가중함수, $K_P(t)=\widehat{F}(t)\cdot L(t)$, 세 가지를 고려하였다. 여기서, $L(t)=[\overline{Y}_1(t)+\overline{Y}_2(t)]/[\overline{Y}_1(t)\ \overline{Y}_2(t)]$ 이고 $\widehat{F}(t)$는 혼합표본에 기초한 시각 t 에서의 생존함수 $\overline{F}(t)=1-F(t)$에 대한 Kaplan-Meier(1958) 추정량이다.

3-1. β_0에 대한 추정량들의 비교

2-1절에서 식 (2.1)과 (2.3)과 같이 주어진 모수 β_0의 두 추정량들을 비교하기 위하여 모의실험을 하였다. 그러나, Log-rank 가중함수의 경우는 Lin and Ying(1994)의 추정량과 동일하기 때문에 <표 1>에 결과를 나타내지 않았다. 모의실험계획은 다음과 같다.

모형 : $\lambda(t;Z)=\lambda_0(t)+\beta_0 Z,\ \lambda_0(t)=1,\ 2t;\ \beta_0=0.5,\ 2.0.$

표본의 크기(n) : 50, 100.

중도절단 비율(cp) : 25%, 50%, (n=100인 경우)75%.

중도절단시간 분포 : $Uni(0,c)$.

여기서, $Uni(0,c)$ 은 $[0,c]$의 값을 취하는 균등분포를 의미하고 c는 원하는 중도절단 비율을 만족하는 값이다. 한편, 모의실험으로 생성된 n개의 표본들이 각 그룹에 속할 확률은 동일하다. <표 1> 에서 *mean* 는 1,000 번의 반복 실험에 의해서 얻어진 추정량들의 평균이며, *mse* 는 편의제곱평균이다. <표 1> 의 *mse* 측면에서 다음과 같은 결과들을 얻을 수 있다.

첫째, Lin and Ying(1994)에 의하면 β_0의 값이 0에 가까울수록, $\lambda_0(t)$의 값이 1 에 가까울수록 식 (2.1)에 주어진 추정량은 우수하다고 알려져 있다. <표 1> 로부터 $\beta_0=0.5$ 이고 $\lambda_0(t)=1$인 경우와 $\beta_0=2.0$ 이고 $\lambda_0(t)=1$인 경우에 각각 이 점을 확인할 수 있다.

둘째, $\lambda_0(t)=2t$ 일 때는 시각에 따라 위험률이 변하기 때문에 어떤 가중함수를 고려한 경우보다 Log-rank 가중함수를 고려한 경우에 가장 좋지 못한 추정이 될 것이다. 또한, $\beta_0=0.5$일

때보다 $\beta_0=2.0$일 때 두 그룹의 위험의 차가 크기 때문에 중도절단시간의 분포를 고려한 $\hat{\beta}_P$가 $\hat{\beta}_G$보다 우수할 것이다. 이 두 사실 모두를 <표 1> 에서 확인할 수 있다.

3-2. 검정통계량 S와 $T_{K_1K_2}$의 비교

2-2절에서 제안한 두 적합도 검정통계량 (2.2)와 (2.4)를 비교하기 위하여 다음과 같은 모의실험을 설계하였다.

$$\text{모형}: \quad \lambda(t;Z)=\lambda_0(t)+\beta_0(t)Z,\ \lambda_0(t)=1;\ \beta_0=0.5,\ 3t,\ 7t.$$

표본의 크기(n) : 50.

중도절단 비율(cp) : 25%, 50%.

중도절단시간 분포 : $Uni(0,c)$, $Exp(c)$.

여기서, $Exp(c)$은 모수의 값이 c인 지수분포를 의미하고 c는 각 중도절단 비율을 만족하도록 하는 값이다. 특히, 검정통계량 S가 귀무가설하에서 얼마나 크기(size)를 잘 제어하는지를 알아보고 또한, 모형의 단조이탈(monotone departure)에 대하여 얼마만큼 검정력(power)이 있는 지를 알아보기 위하여 가정한 모형으로부터 $n=50$개씩 자료를 1,000번 반복 생성하여 $S_1,\cdots,S_{1000}$를 얻었다. 또한, 매번 얻어진 자료에 기초하여 $\hat{S}_{rep,1},\cdots,\hat{S}_{rep,1000}$ $(rep=1,\cdots,1000)$ 을 계산한 후 이들의 표본분포로부터 상위 $\alpha\%$에 해당하는 기각역의 값 $\hat{s}_{rep}(\alpha)$ $(rep=1,\cdots,1000)$를 구한다. 만약 S_{rep}가 $\hat{s}_{rep}(\alpha)$보다 크면 가정한 모형이 타당하지 않다고 결정한다. 결국, <표 2> 와 <표 3> 에 나타난 수값들은 1,000번의 반복실험 중에서 기각역에 속하는 상대도수를 의미한다. 한편, 검정통계량 $T_{K_1K_2}$의 계산을 위하여 Gehan의 가중함수와 logrank 가중함수의 쌍, Gehan의 가중함수와 Prentice-Wilcoxon 가중함수의 쌍, logrank의 가중함수와 Prentice-Wilcoxon 가중함수의 쌍을 고려하였다. 또한, 귀무가설하에서 $T_{K_1K_2}$는 점근정규성을 만족하므로 기각역의 값으로 $z_{\alpha/2}$를 사용하였다. <표 2> 와 <표 3> 의 결과로부터 다음과 같은 몇 가지 결론을 내릴 수 있다.

첫째, 중도절단이 연구기간동안 균등하게 일어나는 균등분포의 경우 혹은 연구초기에 중도절단이 집중적으로 발생하는 지수분포의 경우 모두 제안한 두 검정통계량 S와 $T_{K_1K_2}$는 검정법의 크기를 비교적 잘 제어한다. 또한, 제안한 검정통계량 모두 중도절단 비율에 영향을 받지 않고 검정법의 크기를 잘 제어한다고 할 수 있다.

둘째, 제안한 검정통계량들의 검정력이 전체적으로 우수하지 못한 것은 표본의 크기에 기인한 것이므로 본 논문의 관심사인 검정법들의 비교에 크게 우려되는 결과는 아니라고 생각된다. 한편, 검정력은 중도절단시간의 분포에 관계없이 거의 유사한 경향을 띄고 있음을 몇 가지 결과로부터 확인할 수 있다. 먼저, 단조이탈이 커질수록 그에 상응하여 검정력도 커진다. 중도절단 비율이 증

가할수록 검정력의 감소가 나타나는 데 그것은 cp가 50%일 때 생존시간이 완전하게 얻어지는 자료의 수는 겨우 25개 정도로 그 수가 매우 적기 때문에 발생한 결과라고 할 수 있다. 제안한 검정법들 사이의 검정력을 살펴보면 검정통계량 T_{GP}는 T_{GL}, T_{LP}에 비하여 검정력이 약간 떨어지며, 검정통계량 S는 T_{GL}, T_{LP} 보다는 검정력이 약하지만 T_{GP} 보다는 우수함을 알 수 있다. 그러나, 검정법들 사이의 그와 같은 검정력의 차이가 크게 두드러진다고 볼 수는 없다.

4. 결 론

본 논문에서는 하나의 이진공변량을 갖는 가산위험모형의 적합도검정을 위하여 Kim and Song(1996)과 Kim and Lee(1996)가 제안한 두 검정법을 모의실험을 통하여 살펴보았다. 기저위험함수를 지수함수로 가정했을 때 두 검정법 모두 중도절단시간과 중도절단 비율에 관계없이 비교적 검정력의 크기를 잘 제어함을 알 수 있었다. 또한, Kim and Lee(1996)가 제안한 검정법보다 Kim and Song(1995)이 제안한 검정법의 검정력이 약간 우수하게 나타났지만 이런 경향은 어떤 가중함수의 쌍을 형성하느냐에 따라 조금씩 변화될 수 있음을 알 수 있었다. Kim and Song(1995)의 검정법이 지닌 또 다른 장점은 Kim and Lee(1996)의 검정법 보다 계산이 편리함은 물론 방대한 자료에 대해서도 계산속도가 빠르다는 것이다. 그러나, Kim and Song(1995)의 검정법이 지닌 단점은 가장 우수한 검정력을 얻기 위해서 어떤 쌍의 가중함수를 고려해야만 하는가라는 문제점이 남아 있다고 하겠다. 따라서, 제안한 두 검정법을 실제자료에 적용하여 모형의 적합성을 검정하고자 할 때 Kim and Lee(1996)의 검정법이 계산속도 측면에서는 약점을 지니고 있지만 Kim and Song(1995)의 검정법이 지닌 단점을 극복할 수 있다는 점에서 Kim and Lee(1996)의 검정법을 이용하는 것이 다소 유리하다고 할 수 있다.

< 표 1> 1,000번 반복에 기초한 $\hat{\beta}$, $\hat{\beta}_G$, $\hat{\beta}_P$의 평균과 평균제곱오차에 대한 실험적 추정량

β_0	$\lambda_0(t)$	n	cp	$\hat{\beta}$ mean	$\hat{\beta}$ mse	$\hat{\beta}_G$ mean	$\hat{\beta}_G$ mse	$\hat{\beta}_P$ mean	$\hat{\beta}_P$ mse
0.5	1	50	25%[1]	0.5460	0.2366	0.5236	0.2749	0.5221	0.2600
			50%[2]	0.5707	0.2967	0.5634	0.3414	0.5560	0.3027
		100	25%[1]	0.4985	0.0989	0.4982	0.1187	0.4959	0.1117
			50%[2]	0.4938	0.1378	0.4838	0.1742	0.4832	0.1512
			75%[3]	0.4787	0.2634	0.4796	0.3004	0.4733	0.2610
	2t	50	25%[4]	0.5131	0.1647	0.5167	0.1258	0.5113	0.1270
			50%[5]	0.4962	0.1583	0.5046	0.1246	0.4934	0.1280
		100	25%[4]	0.5037	0.0810	0.5051	0.0607	0.5033	0.0620
			50%[5]	0.5062	0.0818	0.5095	0.0632	0.5040	0.0660
			75%[6]	0.5059	0.0912	0.5111	0.0742	0.5014	0.0791
2.0	1	50	25%[7]	2.1087	0.7070	2.0801	0.7253	2.0682	0.7016
			50%[8]	2.0787	0.8114	2.0404	0.9378	2.0226	0.8385
		100	25%[7]	2.0674	0.2893	2.0555	0.3178	2.0495	0.3048
			50%[8]	2.0166	0.3751	2.0106	0.4205	1.9951	0.3832
			75%[9]	1.9693	0.6406	1.9606	0.7166	1.9432	0.6319
	2t	50	25%[10]	2.0644	0.4215	2.0547	0.3770	2.0424	0.3708
			50%[11]	2.0639	0.4847	2.0666	0.4415	2.0350	0.4285
		100	25%[10]	2.0255	0.1954	2.0271	0.1744	2.0196	0.1725
			50%[11]	2.0261	0.2420	2.0125	0.2134	2.0023	0.2131
			75%[12]	2.0180	0.2990	2.0139	0.2897	1.9952	0.2795

(단, 각 경우 중도절단시간의 분포는 다음과 같다. [1]의 경우 $Uni(0, 3.245)$, [2]의 경우 $Uni(0, 1.299)$, [3]의 경우 $Uni(0, 0.489)$, [4]의 경우 $Uni(0, 2.503)$, [5]의 경우 $Uni(0, 0.908)$, [6]의 경우 $Uni(0, 0.319)$, [7]의 경우 $Uni(0, 3.138)$, [8]의 경우 $Uni(0, 1.531)$, [9]의 경우 $Uni(0, 0.817)$, [10]의 경우 $Uni(0, 2.530)$, [11]의 경우 $Uni(0, 1.174)$, [12]의 경우 $Uni(0, 0.519)$ 이다.)

< 표 2> 1,000번 반복에 기초한 검정통계량 T_{GL}, T_{GP}, T_{LP}, S의 실험적 크기(size)와 검정력(power) (중도절단시간의 분포가 균등분포인 경우)

α	*hd*	0.5		3*t*		7*t*	
	cp	25%[1]	50%[2]	25%[3]	50%[4]	25%[5]	50%[6]
0.01	T_{GL}	0.005	0.013	0.049	0.021	0.106	0.046
	T_{GP}	0.006	0.009	0.042	0.025	0.081	0.055
	T_{LP}	0.006	0.012	0.043	0.015	0.100	0.036
	S	0.005	0.010	0.042	0.022	0.077	0.049
0.05	T_{GL}	0.034	0.055	0.253	0.136	0.397	0.254
	T_{GP}	0.041	0.049	0.171	0.119	0.288	0.232
	T_{LP}	0.034	0.055	0.251	0.126	0.386	0.245
	S	0.044	0.054	0.203	0.126	0.327	0.219
0.10	T_{GL}	0.106	0.098	0.407	0.233	0.575	0.405
	T_{GP}	0.101	0.102	0.268	0.204	0.421	0.360
	T_{LP}	0.103	0.091	0.407	0.240	0.598	0.416
	S	0.100	0.097	0.336	0.217	0.512	0.352

(단, 각 경우 중도절단시간의 분포는 다음과 같다. [1]의 경우 $Uni(0, 3.245)$, [2]의 경우 $Uni(0, 1.299)$, [3]의 경우 $Uni(0, 2.848)$, [4]의 경우 $Uni(0, 1.160)$, [5]의 경우 $Uni(0, 2.564)$, [6]의 경우 $Uni(0, 0.9834)$ 이다.)

< 표 3> 1,000번 반복에 기초한 검정통계량 T_{GL}, T_{GP}, T_{LP}, S의 실험적 크기(size)와 검정력(power) (중도절단시간의 분포가 지수분포인 경우)

α	hd	0.5		$3t$		$7t$	
	cp	25%[1]	50%[2]	25%[3]	50%[4]	25%[5]	50%[6]
0.01	T_{GL}	0.005	0.003	0.057	0.030	0.085	0.052
	T_{GP}	0.003	0.004	0.060	0.025	0.075	0.058
	T_{LP}	0.004	0.003	0.050	0.015	0.075	0.031
	S	0.002	0.004	0.047	0.017	0.057	0.031
0.05	T_{GL}	0.048	0.041	0.259	0.143	0.378	0.242
	T_{GP}	0.043	0.040	0.190	0.126	0.292	0.224
	T_{LP}	0.045	0.037	0.249	0.126	0.371	0.224
	S	0.041	0.039	0.204	0.115	0.295	0.199
0.10	T_{GL}	0.097	0.098	0.392	0.268	0.582	0.412
	T_{GP}	0.090	0.092	0.284	0.222	0.422	0.344
	T_{LP}	0.093	0.098	0.394	0.267	0.588	0.421
	S	0.085	0.088	0.342	0.229	0.480	0.331

(단, 각 경우 중도절단시간의 분포는 다음과 같다. [1]의 경우 $Exp(0.404)$, [2]의 경우 $Exp(1.225)$, [3]의 경우 $Exp(0.454)$, [4]의 경우 $Exp(1.342)$, [5]의 경우 $Exp(0.511)$, [6]의 경우 $Exp(1.552)$ 이다.)

참 고 문 헌

[1] Buckley, J. D. (1984). Additive and Multiplicative Models for Survival Rates, *Biometrics*, Vol. 40, 51-62.

[2] Gill, R. D. and Schumacher, M. (1987). A Simple Test of the Proportional Hazards Assumption, *Biometrika*, Vol. 74, 289-300.

[3] Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations, *Journal of the American Statistical Association*, Vol. 53, 457-481.

[4] Kim, J. H. and Lee, S. Y. (1996). 마팅게일잔차에 기초한 가산위험모형의 적합도검정법, 「응용통계연구」, 제9권 1호, 75-89.

[5] Kim, J. H. and Song, M. S. (1995). A Goodness-of-fit Test for the Additive Risk Model with a Binary Covariate, 「統計學 硏究」, 제24권 2호, 537-549.

[6] Lin, D. Y. and Ying, Z. (1994). Semiparametric Analysis of the Additive Risk Model, *Biometrika*, Vol. 81, 61-71.

[7] Wei, L. J. (1984). Testing Goodness-of-fit for Proportional Hazards Model with Censored Observations, *Journal of the American Statistical Association*, Vol. 79, 649-652.