

마팅게일잔차에 기초한 가산위험모형의 적합도검정법¹⁾

김진흠²⁾, 이승연³⁾

요약

본 논문에서는 하나의 이진공변량(binary covariate)을 갖는 가산위험모형(additive risk model)의 적합도를 검정하기 위한 검정법을 제안하였다. 제안된 검정법은 마팅게일잔차(martingale residual)에 기초한 Wei(1984)의 비례위험모형(proportional hazards model)의 적합도검정법을 가산위험모형으로 확장한 형태이다. 제안된 검정통계량의 점근성질을 유도하였고 실제자료에 적용하여 보았다.

1. 서론

일반적으로 생존분석분야에서 공변량과 생존시간 사이의 관계를 연구하기 위하여 비례위험모형과 가산위험모형을 가정하는 경우가 많다. 비례위험모형은 위험함수가 공변량에 대하여 지수적으로 증가(또는 감소)하는 모형이고 가산위험모형은 위험함수가 공변량에 대하여 가산적으로 증가(또는 감소)하는 모형이다. 따라서, 두 모형은 다음과 같이 표현될 수 있다.

$$\lambda(t; \mathbf{Z}) = \lambda_0(t) + \boldsymbol{\beta}_0' \mathbf{Z}(t), \quad (1.1)$$

$$\lambda(t; \mathbf{Z}) = \lambda_0(t) \exp[\boldsymbol{\gamma}_0' \mathbf{Z}(t)]. \quad (1.2)$$

여기서, $\lambda_0(\cdot)$ 는 기저위험함수(baseline hazard function)이고 $\mathbf{Z}(\cdot)$ 는 $p \times 1$ 벡터 공변량이며 $\boldsymbol{\beta}_0$ 와 $\boldsymbol{\gamma}_0$ 는 $p \times 1$ 벡터 회귀모수들이다.

Cox(1972)가 제안한 이후 모형 (1.2)에서 가정된 비례위험모형의 타당성을 검토하기 위하여 많은 검정법들이 제안되어 왔다. 그 중에서 마팅게일잔차에 기초한 검정법들을 중심으로 몇 가지 소개하고자 한다. Wei(1984)는 모형 (1.2)에서 $p=1$ 이고 $\mathbf{Z}(t)$ 가 0과 1의 값만 취하는 특별한 경우에 대하여 비례위험가정의 타당성을 검정하기 위한 검정법을 제안하였다. 다시 말해서, 두 그룹의 위험함수가 비례위험가정을 만족하는지 아닌지에 대한 검정법을 제안하였다. Therneau, Grambsch and Fleming(1990)은 Wei(1984)의 검정법을 $p \times 1$ 벡터 시간불변공변량(time-invariant covariate)을 갖는 모형으로 확장하였다. 그러나, Therneau *et al.*(1990)의 검정법은 공변량들 사이의 독립성을 요구하고 있기 때문에 실제로 그 요구조건이 만족되지 않는 경우에 적용시키기가 어렵다. Lin, Wei and Ying(1993)은 Therneau *et al.*(1990)이 제안한 검정

1) 본 연구는 1994년 한국과학재단 연구과제 KOSEF 94-0701-01-3의 지원에 의하여 수행되었음.
2) (445-743) 경기도 화성군 봉담면 와우리 산 2-2, 수원대학교 응용통계학과 전임강사.
3) (133-747) 서울특별시 광진구 군자동 98, 세종대학교 응용통계학과 조교수.

통계량의 분포를 시뮬레이션을 통해서 근사시키는 방법을 제안하였고 그 방법으로 얻어진 검정 통계량의 유의확률값은 공변량들의 공분산 구조에 관계없이 점근적으로 타당함을 보였다. 또한, Lin *et al.*(1993)은 그래픽 방법으로도 비례위험가정의 타당성을 살펴보았다.

본 논문에서는 가산위험모형 (1.1)에서 $p=1$ 이고 이진공변량을 갖는 다음과 같은 단순한 모형을 고려하였다.

$$\lambda(t; Z) = \lambda_0(t) + \beta_0 Z. \quad (1.3)$$

여기서, β_0 는 회귀모수이며 Z 는 한 개체가 그룹 1에 속하면 1이고 그룹 2에 속하면 0인 이진 공변량이다. 이 모형의 적합도검정법을 전개하기 위하여 필요한 가정과 확률과정들을 2절에서 소개하고, 3절에서 마팅계일관차에 기초한 검정법을 제안하였다. 이 검정통계량은 비례위험모형에 대하여 전개한 Wei(1984)의 적합도검정법을 Lin and Ying(1994)의 가산위험모형으로 확장한 형태이다. 또한, 제안된 검정통계량의 점근성질을 유도하였고 4절에서 두 가지 실제자료에 적용하여 그 결과를 토론하였다.

2. 기호와 가정

다음과 같이 기호를 정의하자.

$$X_{ij} = \min(T_{ij}, C_{ij}) \quad (i=1, 2; j=1, \dots, n_i),$$

$$\Delta_{ij} = I(T_{ij} \leq C_{ij}) \quad (i=1, 2; j=1, \dots, n_i),$$

$$Z_{1j} = 1, \quad (j=1, \dots, n_1); \quad Z_{2j} = 0, \quad (j=1, \dots, n_2).$$

여기서, $I(\cdot)$ 는 어떤 집합 A 에 대하여

$$I(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

로 정의되는 지표함수(indicator function)이고 T_{ij} 는 절대연속인(absolutely continuous)분포함수를 갖는 생존시간이며 C_{ij} 는 T_{ij} 에 대응하는 중도절단시간(censoring time)이다. T_{ij} 와 C_{ij} 는 서로 독립이라고 가정하자. 더불어 몇 가지 확률과정들을 정의하자.

$$N_{ij}(t) = I(X_{ij} \leq t, \Delta_{ij} = 1), \quad Y_{ij}(t) = I(X_{ij} \geq t) \quad (i=1, 2; j=1, \dots, n_i),$$

$$N_i(t) = \sum_{j=1}^{n_i} N_{ij}(t), \quad Y_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t) \quad (i=1, 2).$$

그러므로, $N_i(t)$ ($i=1, 2$)는 그룹 i 에서 시간 t 까지 사망한 개체수를 의미하고 $Y_i(t)$ ($i=1, 2$)는 그룹 i 에서 시간 t 를 포함해서 그 이후에 생존하고 있는 개체수를 의미한다. 한편, 모형 (1.3)하에서 계수과정(counting process) $N_{ij}(t)$ ($i=1, 2; j=1, \dots, n_i$)의 강도

함수(intensity function)는

$$Y_{ij}(t)d\Lambda_i(t) = Y_{ij}(t)[d\Lambda_0(t) + \beta_0 Z_{ij}dt]$$

로 주어진다. 여기서, $\Lambda_i(\cdot)$ ($i=1,2$)는 그룹 i 에 대응하는 누적위험함수(cumulative hazard function)이고 $\Lambda_0(\cdot) = \int_0^\cdot \lambda_0(u)du$ 이다. 또한, Doob-Meyer의 분해(decomposition)에 의하면 $N_{ij}(t)$ 는

$$M_{ij}(t) = N_{ij}(t) - \int_0^t Y_{ij}(u)d\Lambda_i(u) \quad (2.1)$$

로 유일하게 표현될 수 있다. 여기서, $M_{ij}(\cdot)$ 는 제곱적분가능(square integrable) 마팅계일이다 (Andersen and Gill(1982)).

3. 적합도검정통계량과 점근성질

모형 (1.3)에서 β_0 를 추정하기 위하여 Lin and Ying(1994)은 다음과 같은 추정함수(estimated function)를 제안하였다.

$$U(\beta, t) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \int_0^t [Z_{ij} - \bar{Z}(u)] [dN_{ij}(u) - Y_{ij}(u)\beta Z_{ij}du] \quad (3.1)$$

여기서, $\bar{Z}(t) = Y_1(t)/[Y_1(t) + Y_2(t)]$ 이다. 한편, $Z_{1j} = 1$ ($j=1, \dots, n_1$) 이고 $Z_{2j} = 0$ ($j=1, \dots, n_2$)이기 때문에 (3.1)은 다음과 같이 간단히 표현될 수 있다.

$$U(\beta, t) = \sum_{j=1}^{n_1} \int_0^t \frac{Y_2(u)}{Y_1(u) + Y_2(u)} dM_{1j}(u) - \sum_{j=1}^{n_2} \int_0^t \frac{Y_1(u)}{Y_1(u) + Y_2(u)} dM_{2j}(u) \quad (3.2)$$

(3.2)로부터 $E[U(\beta, t)] = 0$ 임을 알 수 있으며 이 결과로부터 가산위험가정이 타당할 때 (3.2)는 그룹 1에서 관찰된 사망수와 기대되는 사망수의 차로 해석될 수 있다. 따라서, 비례위험 모형 하에서 제안된 Wei(1984)의 적합도검정법과 같은 관점에서 $\sup_{0 \leq t < \infty} |U(\beta, t)|$ 을 가산위험모형에 대한 적합도검정통계량으로 고려할 수 있다.

$\hat{\beta}$ 를 $U(\beta, \infty) = 0$ 의 해라고 하자. 적절한 조건과 모형 (1.3) 가정하에서 $\hat{\beta}$ 는 β_0 의 일치추정량이며 $n^{1/2}(\hat{\beta} - \beta_0)$ 는 점근적으로 정규분포를 따른다 (Lin and Ying(1994)). 우선, (2.1)로부터 마팅계일잔차를

$$\hat{M}_{ij}(t) = N_{ij}(t) - \int_0^t Y_{ij}(u)[d\hat{\Lambda}_0(\hat{\beta}, u) + \hat{\beta}Z_{ij}du] \quad (i=1,2; j=1, \dots, n_i)$$

로 정의하자. 여기서, $\hat{\Lambda}_0(\hat{\beta}, t)$ 는 $\Lambda_0(t)$ 의 일치추정량이며

$$\widehat{\Lambda}_0(\widehat{\beta}, t) = \int_0^t \frac{dN_1(u) + dN_2(u) - \widehat{\beta} Y_1(u) du}{Y_1(u) + Y_2(u)}$$

로 주어진다 (Lin and Ying(1994)). 이제, (3.2)를 이용하여 관찰과정(observed process)을 다음과 같이 정의하자.

$$n^{-1/2} U(\widehat{\beta}, t) = n^{-1/2} \left\{ \sum_{j=1}^{n_1} \int_0^t \frac{Y_2(u)}{Y_1(u) + Y_2(u)} d\widehat{M}_{1j}(u) - \sum_{j=1}^{n_2} \int_0^t \frac{Y_1(u)}{Y_1(u) + Y_2(u)} d\widehat{M}_{2j}(u) \right\}. \quad (3.3)$$

한편, (3.3)의 점근분포를 유도하기 위하여 필요한 조건들은 다음과 같다.

조건 (1): $n_i/n \rightarrow \rho_i \in (0, 1)$ ($i=1, 2$), $n = n_1 + n_2$.

조건 (2): 모형 (1.3) 가정하에서 $\sup_{0 \leq t < \infty} |Y_i(t)/n_i - y_i(t)| \xrightarrow{p} 0$ ($i=1, 2$)를 만족하는 함수들 $y_i(\cdot) \in (0, 1)$ ($i=1, 2$)가 존재한다.

아래 모든 정리의 증명들은 부록 A에 수록하였다.

정리 3.1 조건 (1)과 (2)를 만족한다고 가정하면, 모형 (1.3) 가정하에서 확률과정 $n^{-1/2} U(\widehat{\beta}, t)$ 과 $n^{-1/2} \widehat{U}(\beta, t)$ 은 점근적으로 동치이다. 여기서,

$$n^{-1/2} \widehat{U}(\beta, t) = n^{-1/2} \left\{ \sum_{j=1}^{n_1} \int_0^\infty \frac{\rho_2 y_2(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] dM_{1j}(u) - \sum_{j=1}^{n_2} \int_0^\infty \frac{\rho_1 y_1(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] dM_{2j}(u) \right\} \quad (3.4)$$

이고 $\eta(t) = \int_0^t \frac{\rho_1 y_1(u) \rho_2 y_2(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} du$ 이다.

정리 3.2 조건 (1)과 (2)를 만족한다고 가정하면, 모형 (1.3) 하에서 $n^{-1/2} \widehat{U}(\beta, \cdot)$ 는 평균 0이고 공분산 $\alpha(s, t)$ ($s \leq t$)를 갖는 가우스과정(Gaussian process)으로 수렴한다. 여기서,

$$\alpha(s, t) = \xi(s) - \frac{\eta(t)}{\eta(\infty)} \xi(s) - \frac{\eta(s)}{\eta(\infty)} \xi(t) + \frac{\eta(s)\eta(t)}{\eta(\infty)^2} \xi(\infty) \quad (3.5)$$

이고 $\xi(t) = \int_0^t \frac{\rho_1 y_1(u) \rho_2^2 y_2^2(u)}{[\rho_1 y_1(u) + \rho_2 y_2(u)]^2} d\Lambda_1(u) + \int_0^t \frac{\rho_1^2 y_1^2(u) \rho_2 y_2(u)}{[\rho_1 y_1(u) + \rho_2 y_2(u)]^2} d\Lambda_2(u)$ 이다.

정리 3.3 $\hat{\eta}(t) = n^{-1} \int_0^t \frac{Y_1(u)Y_2(u)}{Y_1(u)+Y_2(u)} du$ 이고

$$\hat{\xi}(t) = n^{-1} \left\{ \int_0^t \left[\frac{Y_2(u)}{Y_1(u)+Y_2(u)} \right]^2 dN_1(u) + \int_0^t \left[\frac{Y_1(u)}{Y_1(u)+Y_2(u)} \right]^2 dN_2(u) \right\}$$

라고 하자. 정리 3.2와 동일한 조건하에서 $\sigma(s, t)$ ($s \leq t$)의 일치추정량 $\hat{\sigma}(s, t)$ 는

$$\hat{\sigma}(s, t) = \hat{\xi}(s) - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \hat{\xi}(s) - \frac{\hat{\eta}(s)}{\hat{\eta}(\infty)} \hat{\xi}(t) + \frac{\hat{\eta}(s)\hat{\eta}(t)}{\hat{\eta}(\infty)^2} \hat{\xi}(\infty) \quad (3.6)$$

로 주어진다.

정리 3.1-정리 3.3으로부터 관찰과정 $n^{-1/2}U(\hat{\beta}, \cdot)$ 는 평균 0이고 식 (3.5)와 같은 공분산을 갖는 가우스과정으로 수렴한다. 또한, 그 공분산의 일치추정량은 식 (3.6)으로부터 구할 수 있다. 이 관찰과정을 기초로 하여 Wei(1984)의 검정통계량과 같은 형태의 적합도검정통계량을 다음과 같이 제안한다.

$$S = \sup_{0 \leq t < \infty} A^{-1/2} |n^{-1/2}U(\hat{\beta}, t)|.$$

여기서, $A = n^{-1} \left\{ \int_0^\infty \left[\frac{Y_2(u)}{Y_1(u)+Y_2(u)} \right]^2 dN_1(u) + \int_0^\infty \left[\frac{Y_1(u)}{Y_1(u)+Y_2(u)} \right]^2 dN_2(u) \right\}$ 이다.

검정통계량 S 는 모형 부적합(model misspecification) 하에서도 일치성이 만족됨을 보일 수 있다 (Wei(1984), Struthers and Kalbfleisch(1986)). 그러나, 식 (3.5)의 공분산의 형태로부터 알 수 있듯이 $n^{-1/2}U(\hat{\beta}, \cdot)$ 는 점근적으로 독립증분성(independent increments)을 갖고 있지 않기 때문에 $n^{-1/2}U(\hat{\beta}, \cdot)$ 의 점근분포를 수치적으로 계산하기는 어렵다. 관찰과정 $n^{-1/2}U(\hat{\beta}, \cdot)$ 의 분포를 근사시킬 수 있는 방법으로 Lin et al.(1993)이 제안한 시뮬레이션과정을 이용할 수 있다. 우선, 식 (3.4)에 있는 $M_{ij}(t)$ ($i=1, 2; j=1, \dots, n_i$)를 $N_{ij}(t)G_{ij}$ ($i=1, 2; j=1, \dots, n_i$)로 대체하는데, 여기서 G_{ij} 는 표준정규확률변수이다. 그리고, (3.4)에 있는 나머지 미지의 양들은 모두 각각의 일치추정량으로 대체한다. 그러면, 시뮬레이션과정(simulated process)을 다음과 같이 정의할 수 있다.

$$n^{-1/2} \hat{U}(\hat{\beta}, t) = n^{-1/2} \left\{ \sum_{j=1}^{n_1} \int_0^\infty \frac{Y_2(u)}{Y_1(u)+Y_2(u)} \left[I(u \leq t) - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \right] G_{1j} dN_{1j}(u) - \sum_{j=1}^{n_2} \int_0^\infty \frac{Y_1(u)}{Y_1(u)+Y_2(u)} \left[I(u \leq t) - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \right] G_{2j} dN_{2j}(u) \right\}. \quad (3.7)$$

시뮬레이션과정(simulated process) (3.7)에서 G_{ij} ($i=1, 2; j=1, \dots, n_i$)는 확률변수로 간주되고

$(X_{ij}, \Delta_{ij}, Z_{ij}) (i=1, 2; j=1, \dots, n_i)$ 는 고정된 것으로 간주되기 때문에 (3.7)은 시간 t 에서 서로 독립인 확률변수들의 합이라고 볼 수 있다.

정리 3.4 모형 (1.3) 가정하에서 조건 (1)과 (2)가 만족된다고 하자. $(X_{ij}, \Delta_{ij}, Z_{ij}) (i=1, 2; j=1, \dots, n_i)$ 가 주어졌을 때, $n^{-1/2} O(\hat{\beta}, t)$ 는 평균 0이고 식 (3.5)와 같은 공분산을 갖는 가우스과정으로 수렴한다.

정리 3.4에 의하면 $(X_{ij}, \Delta_{ij}, Z_{ij}) (i=1, 2; j=1, \dots, n_i)$ 이 주어졌을 때 관찰과정 (3.3)과 시뮬레이션과정 (3.7)은 모형 (1.3) 가정하에서 같은 점근분포를 갖는다. 따라서, 시뮬레이션과정을 이용하여 적합도검정통계량 S 의 점근분포를 구할 수 있다. 구체적으로, 시뮬레이션과정은 다음과 같이 반복적으로 생성할 수 있다.

단계 (1): 시뮬레이션을 통해서 표준정규확률변수 $G_{ij} (i=1, 2; j=1, \dots, n_i)$ 를 생성한다.

단계 (2): 주어진 $(X_{ij}, \Delta_{ij}, Z_{ij}) (i=1, 2; j=1, \dots, n_i)$ 와 단계 (1)에서 얻어진 G_{ij} 를 이용하여 (3.7)를 구한다.

단계 (3): 단계 (1)과 (2)를 반복수행한다.

실제적으로, 시뮬레이션과정을 이용하여 적합도검정을 할 수 있는 방법은 두 가지가 있다. 먼저, 그래픽 방법으로 관찰과정 (3.3)과 시뮬레이션과정 (3.7)을 그려보았을 때 만약 관찰과정이 시뮬레이션과정으로부터 많이 벗어나면 가산위험모형이 타당하지 않다고 결론짓는다. 그러나, 이와같은 그래픽 방법은 객관적인 기준이 애매모호하기 때문에, 대안책으로 수치적 방법을 제안한다. 이 방법은 유의확률 $P(S \geq s)$ 을 $P(\hat{S} \geq s)$ 로 근사시켜 만약 근사된 유의확률이 고정된 유의수준 α 보다 작으면 가산위험모형의 가정을 기각시킨다. 여기서, s 는 검정통계량 S 의 관찰값이고 $\hat{S} = \sup_{0 \leq t < \infty} A^{-1/2} |n^{-1/2} O(\hat{\beta}, t)|$ 이다.

4. 실제 예

3절에서 제안한 적합도검정법을 두 가지 실제자료에 적용하여 각 자료가 가산위험모형에 적합한지를 그래픽 방법과 수치적 방법으로 살펴보고자 한다.

검정통계량 S 의 경험적 유의확률값(empirical p-values)들은 1,000번의 반복을 통해서 얻었고 부록B에 수록된 그림 1과 그림 2에서 관찰과정(observed process)은 실선으로 표시하였으며 15개의 시뮬레이션과정(simulated processes)들은 점선으로 표시하였다.

자료 1은 Fleming *et al.*(Gill and Schumacher(1987))에서 연구된 것으로 자궁암 Stage II 또는 Stage IIA의 증상을 갖는 35명의 환자들에게 약물을 투여한 이후 병이 진행되기 시작하는 시간까지를 관측한 자료이다. 이 자료가 가산위험모형으로 적합한 지를 검정하기 위하여 구한 검

정통계량 S 의 경험적 유의확률값은 0.019이었다. 따라서, 유의수준 5% 하에서 가산위험모형이 적합하지 않다는 결론을 내릴 수 있으며, 이 결론은 그림 1에서 보여주는 것과 같이 관찰과정이 시뮬레이션과정을 벗어나는 것이 뚜렷하므로 그래픽방법과도 일치한다.

자료 2는 Ying, Jung and Wei(1995)에서 분석된 것으로 121명의 SCLC(small cell lung cancer)환자들의 생존시간을 측정된 것이다. SCLC 환자에 대한 대표적인 치료법은 etoposide와 cisplatin을 조합하여 투여하는 것이다. 그러나, 아직까지는 최적의 조합순서 뿐만 아니라 서로 다른 두 조합순서에 따른 치료효과의 관계도 알려져 있지 않다. 121명의 환자들 중에서 59명의 환자들에게는 Arm A의 치료법을, 나머지 62명의 환자들에게는 Arm B의 치료법을 수행하였다. 여기서, Arm A의 치료법은 etoposide 투여한 후 cisplatin 투여하는 방법이고 Arm B의 치료법은 cisplatin 투여한 후 etoposide 투여하는 방법이다. 이 자료에서 두 치료법의 위험함수가 가산적인지 아닌지를 검정하기 위하여 구한 검정통계량 S 의 경험적 유의확률값은 0.181이었다. 유의수준 5% 하에서 검정통계량의 값이 유의하지 않으므로 이 자료는 가산위험모형이 타당하다고 볼 수 있으며 그림 2에서 보여주는 것과 같이 관찰과정이 시뮬레이션과정을 거의 벗어나지 않으므로 일치된 결론을 내릴 수 있다. 한편, 같은 자료에 대하여 Wei(1984)의 검정법을 이용하여 비례위험가정의 타당성을 검정해 본 결과 유의확률값이 0.003으로 매우 유의하므로 비례위험모형이 적합하지 않다는 결론을 내릴 수 있다.

5. 결론

본 논문에서는 하나의 이진공변량을 갖는 가산위험모형의 적합도검정법을 두 가지 방법으로 제안하고 실제적인 자료에 적용한 결과를 보여 주었다. 4절에서 소개된 자료 2의 경우처럼 비례위험모형가정이 타당하지 않다고 판단될 때 이미 연구된 비례위험모형의 통계적 방법들을 적용하는 것은 적합하지 못하다. 왜냐하면, Cox(1972) 이후 연구된 대부분의 통계적 방법들은 비례위험가정을 기초하여 개발된 것이므로 이 가정이 타당하지 않을 경우에 많은 오류를 초래할 수 있기 때문이다. 그러므로, 가정한 모형에서 관심의 대상이 되는 모수를 통계적으로 추론하기에 앞서 가정한 모형이 적합한지를 검정하는 것이 선행되어야 한다. 이 논문에서는 비례위험모형의 적합도검정법을 가산위험모형으로 확장시킴으로서 타당한 모형을 찾을 수 있는 구체적인 하나의 방법을 제시하고 있다. 앞으로는 제안한 검정법을 k -표본 ($k > 2$) 문제 뿐만 아니라 p -차원 시간불변공변량(time-invariant covariate)을 포함하는 일반적인 가산위험모형으로까지 확장하고자 한다.

참 고 문 헌

- [1] Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting process : a large sample study, *The Annals of Statistics*, Vol. 10, 1100-1120.
- [2] Cox, D. R. (1972). Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society*, Ser. 34, 187-220.
- [3] Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*, John Wiley and Sons, New York.
- [4] Gill, R. D. and Schumacher, M. (1987). A simple test of the proportional hazards assumption, *Biometrika*, Vol. 74, 289-300.
- [5] Lin, D. Y., Wei, L. J. and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals, *Biometrika*, Vol. 80, 557-572.
- [6] Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model, *Biometrika*, Vol. 81, 61-71.
- [7] Struthers, C. A. and Kalbfleisch, J. D. (1986). Misspecified proportional hazard models, *Biometrika*, Vol. 73, 363-369.
- [8] Therneau, T. M., Grambsch, P. M. and Fleming, T. R. (1990). Martingale-based residuals for survival models, *Biometrika*, Vol. 77, 147-160.
- [9] 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.
- [10] Ying, Z., Jung, S. H. and Wei, L. J. (1995). Survival analysis with median regression model, *Journal of the American Statistical Association*, Vol. 90, 178-184.

부록 A

(정리 3.1의 증명) β_0 근방에서 $U(\hat{\beta}, t)$ 와 $U(\hat{\beta}, \infty)$ 의 Taylor 급수전개와 Fleming and Harrington(1991)를 이용하여 $n^{-1/2}U(\hat{\beta}, t)$ 과 $n^{-1/2}\tilde{U}(\beta, t)$ 이 점근적으로 동치임을 보일 수 있다. 먼저, 다음을 주목하자.

$$\begin{aligned} n^{-1/2}U(\hat{\beta}, t) &= n^{-1/2}U(\beta, t) + n^{-1/2}(\hat{\beta} - \beta) \left\{ n^{-1} \int_0^t \frac{Y_1(u)Y_2(u)}{Y_1(u) + Y_2(u)} du \right\} \\ &= n^{-1/2}U(\beta, t) - n^{-1/2}U(\beta, \infty) \left\{ \frac{n^{-1} \int_0^t \frac{Y_1(u)Y_2(u)}{Y_1(u) + Y_2(u)} du}{n^{-1} \int_0^\infty \frac{Y_1(u)Y_2(u)}{Y_1(u) + Y_2(u)} du} \right\} \\ &= n^{-1/2}U(\beta, t) - n^{-1/2}U(\beta, \infty) \left\{ \frac{\eta(t)}{\eta(\infty)} + o_p(1) \right\}. \end{aligned}$$

또한, Fleming and Harrington(1991)의 Theorem 6.2.1으로부터

$$n^{-1/2} \left\{ \int_0^t \frac{Y_2(u)}{Y_1(u) + Y_2(u)} dM_1(u) - \int_0^t \frac{\rho_2 y_2(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} dM_1(u) \right\} \xrightarrow{d} 0$$

$$n^{-1/2} \left\{ \int_0^t \frac{Y_1(u)}{Y_1(u) + Y_2(u)} dM_2(u) - \int_0^t \frac{\rho_1 y_1(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} dM_2(u) \right\} \xrightarrow{d} 0$$

임을 알 수 있다. 그러므로, 위의 결과로부터

$$\begin{aligned} n^{-1/2}U(\hat{\beta}, t) &= n^{-1/2} \left\{ \int_0^\infty \frac{\rho_2 y_2(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] dM_1(u) \right. \\ &\quad \left. - \int_0^\infty \frac{\rho_1 y_1(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] dM_2(u) \right\} \\ &\quad + o_p(1) \end{aligned}$$

임으로 $n^{-1/2}U(\hat{\beta}, t)$ 과 $n^{-1/2}\tilde{U}(\beta, t)$ 은 점근적으로 동치이다. ■

(정리 3.2의 증명) 먼저,

$$H_1(t, u) = \frac{\rho_1 y_1(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \left\{ I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right\},$$

$$H_2(t, u) = \frac{\rho_2 y_2(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \left\{ I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right\}$$

라 하자. 그러면, $n^{-1/2}\tilde{U}(\beta, t)$ 는 다음과 같이 표현될 수 있다.

$$n^{-1/2} \tilde{U}(\beta, t) = \int_0^\infty H_2(t, u) dM_1(u) - \int_0^\infty H_1(t, u) dM_2(u).$$

$n^{-1/2} \tilde{U}(\beta, \cdot)$ 의 타이트성(tightness)은 $H_1(t, \cdot)$ 과 $H_2(t, \cdot)$ 의 유계성(boundedness)과 Fleming and Harrington (1991)의 Corollary 3.4.1로부터 만족된다. 또한, $n^{-1/2} \tilde{U}(\beta, \cdot)$ 은 서로 독립인 확률변수들의 간단한 합의 형태이기 때문에 Linderberg-Feller의 중심극한정리와 타이트성으로부터 $n^{-1/2} \tilde{U}(\beta, \cdot)$ 는 평균이 0인 가우스과정으로 수렴한다. 한편, $s \leq t$ 에 대하여 $n^{-1/2} \tilde{U}(\beta, \cdot)$ 의 공분산 함수를 다음과 같이 유도할 수 있다.

$$\begin{aligned} & E\{n^{-1/2} \tilde{U}(\beta, s) \cdot n^{-1/2} \tilde{U}(\beta, t)\} \\ &= n^{-1} E\left\{ \int_0^\infty H_2(s, u) H_2(t, u) d\langle M_1, M_1 \rangle(u) + \int_0^\infty H_1(s, u) H_1(t, u) d\langle M_2, M_2 \rangle(u) \right. \\ &\quad \left. - \int_0^\infty H_2(s, u) H_1(t, u) d\langle M_1, M_2 \rangle(u) - \int_0^\infty H_2(t, u) H_1(s, u) d\langle M_1, M_2 \rangle(u) \right\} \quad (\text{A.1}) \\ &= n^{-1} E\left\{ \int_0^\infty H_2(s, u) H_2(t, u) d\langle M_1, M_1 \rangle(u) + \int_0^\infty H_1(s, u) H_1(t, u) d\langle M_2, M_2 \rangle(u) \right\} \\ &= n^{-1} E\left\{ \int_0^\infty H_2(s, u) H_2(t, u) Y_1(u) d\Lambda_1(u) + \int_0^\infty H_1(s, u) H_1(t, u) Y_2(u) d\Lambda_2(u) \right\}. \end{aligned}$$

여기서, $M_1(t)$ 과 $M_2(t)$ 은 서로 수직이며 $\langle M_i, M_i \rangle(t) = \int_0^t Y_i(u) d\Lambda_i(u)$ ($i=1, 2$) 이다. 또

$$\text{한, } A_1(u) = \left\{ \frac{\rho_2 y_2(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \right\}^2, \quad A_2(u) = \left\{ \frac{\rho_1 y_1(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} \right\}^2 \text{라하고 Tonelli의 정리를}$$

(A.1)의 마지막 식에 적용하면

$$\begin{aligned} & E\{n^{-1/2} \tilde{U}(\beta, s) \cdot n^{-1/2} \tilde{U}(\beta, t)\} \\ &= E\left\{ \int_0^\infty A_1(u) \left[I(u \leq s) - \frac{\eta(s)}{\eta(\infty)} \right] \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] \frac{Y_1(u)}{n} d\Lambda_1(u) \right. \\ &\quad \left. + \int_0^\infty A_2(u) \left[I(u \leq s) - \frac{\eta(s)}{\eta(\infty)} \right] \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] \frac{Y_2(u)}{n} d\Lambda_2(u) \right\} \\ &\xrightarrow{p} \int_0^\infty \left\{ \left[I(u \leq s) - \frac{\eta(s)}{\eta(\infty)} \right] \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] \frac{\eta(s)\eta(t)}{\eta(\infty)^2} \right. \\ &\quad \left. \times \left[A_1(u) \rho_1 y_1(u) + A_2(u) \rho_2 y_2(u) \right] d\Lambda_2(u) + \beta A_1(u) \rho_1 y_1(u) du \right\} \\ &= \int_0^\infty \left\{ \left[I(u \leq s) - \frac{\eta(s)}{\eta(\infty)} \right] \left[I(u \leq t) - \frac{\eta(t)}{\eta(\infty)} \right] \frac{\eta(s)\eta(t)}{\eta(y)^2} \right. \\ &\quad \left. \times \left[\frac{\rho_1 \rho_2 y_1(u) y_2(u)}{\rho_1 y_1(u) + \rho_2 y_2(u)} d\Lambda_2(u) + \frac{\beta \rho_1 \rho_2^2 y_1(u) y_2(u)^2}{[\rho_1 y_1(u) + \rho_2 y_2(u)]^2} du \right] \right\} \end{aligned}$$

와 같이 유도된다. ■

(정리 3.3의 증명) 주어진 조건 (1)과 (2) 그리고 $N_i(t)/n \xrightarrow{p} \int_0^t \rho_i y_i(u) d\Lambda_i(u)$ ($i=1, 2$)로

부터 쉽게 유도할 수 있다. ■

(정리 3.4의 증명) 먼저, 타이트성을 검토하면 다음과 같다. $t_1 \leq t \leq t_2$ 에 대하여,

$$C_{1j}(t_1, t) = \frac{\Delta_{1j} Y_2(X_{1j})}{Y_1(X_{1j}) + Y_2(X_{1j})} \left\{ [I(X_{1j} \leq t) - I(X_{1j} \leq t_1)] - \left[\frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} - \frac{\hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right] \right\}, \quad (\text{A.2})$$

$$C_{1j}(t, t_2) = \frac{\Delta_{1j} Y_2(X_{1j})}{Y_1(X_{1j}) + Y_2(X_{1j})} \left\{ [I(X_{1j} \leq t_2) - I(X_{1j} \leq t)] - \left[\frac{\hat{\eta}(t_2)}{\hat{\eta}(\infty)} - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \right] \right\}, \quad (\text{A.3})$$

$$C_{2j}(t_1, t) = \frac{\Delta_{2j} Y_1(X_{2j})}{Y_1(X_{2j}) + Y_2(X_{2j})} \left\{ [I(X_{2j} \leq t) - I(X_{2j} \leq t_1)] - \left[\frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} - \frac{\hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right] \right\}, \quad (\text{A.4})$$

$$C_{2j}(t, t_2) = \frac{\Delta_{2j} Y_1(X_{2j})}{Y_1(X_{2j}) + Y_2(X_{2j})} \left\{ [I(X_{2j} \leq t_2) - I(X_{2j} \leq t)] - \left[\frac{\hat{\eta}(t_2)}{\hat{\eta}(\infty)} - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \right] \right\} \quad (\text{A.5})$$

라 하자. 여기서, $j=1, \dots, n_1$; $j'=1, \dots, n_2$ 이다. (A.2)-(A.5)로부터

$$|n^{-1/2} \hat{U}(\hat{\beta}, t) - n^{-1/2} \hat{U}(\hat{\beta}, t_1)|^2 = |n^{-1/2} \sum_{j=1}^{n_1} C_{1j}(t_1, t) G_{1j} - n^{-1/2} \sum_{j'=1}^{n_2} C_{2j'}(t_1, t) G_{2j'}|^2, \quad (\text{A.6})$$

$$|n^{-1/2} \hat{U}(\hat{\beta}, t_2) - n^{-1/2} \hat{U}(\hat{\beta}, t)|^2 = |n^{-1/2} \sum_{j=1}^{n_1} C_{1j}(t, t_2) G_{1j} - n^{-1/2} \sum_{j'=1}^{n_2} C_{2j'}(t, t_2) G_{2j'}|^2 \quad (\text{A.7})$$

와 같이 표현할 수 있다.

표준정규확률변수 G_{1j} 과 $G_{2j'}$ 에 대하여 $E(G_{1j}^4) = 3$ 이고 $E(G_{2j'}^4) = 3$ 임을 주목하자. 또

한, $\frac{Y_2(X_{1j})}{Y_1(X_{1j}) + Y_2(X_{1j})} \leq K_2(\ll \infty)$ 와 $\frac{Y_1(X_{2j'})}{Y_1(X_{2j'}) + Y_2(X_{2j'})} \leq K_1(\ll \infty)$ 라 하자. 그러면,

$C_{1j}(t_1, t)$, $C_{1j}(t, t_2)$, $C_{2j'}(t_1, t)$ 과 $C_{2j'}(t, t_2)$ 의 유계성으로부터 다음을 알 수 있다.

$$\begin{aligned} C_{1j}(t_1, t)^2 &\leq K_2^2 \left\{ I(t_1 < X_{1j} \leq t) - \left[\frac{\hat{\eta}(t) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right] \right\}^2 \\ &\leq K_2^2 \left\{ \frac{\hat{\eta}(t) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^2 \\ &\leq K_2^2 \left\{ \frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^2, \end{aligned} \quad (\text{A.8})$$

$$C_{1j}(t, t_2)^2 \leq K_2^2 \left\{ \frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^2, \quad (\text{A.9})$$

$$C_{2j'}(t_1, t)^2 \leq K_1^2 \left\{ \frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^2, \quad (\text{A.10})$$

$$C_{2j'}(t, t_2)^2 \leq K_1^2 \left\{ \frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^2, \quad (\text{A.11})$$

$$\begin{aligned} & C_{1j}(t_1, t) C_{1k}(t, t_2) \\ & \leq K_2^2 \left\{ -I(t < X_{1k} \leq t_2) \left[\frac{\hat{\eta}(t) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right] - I(t_1 < X_{1j} \leq t) \left[\frac{\hat{\eta}(t_2) - \hat{\eta}(t)}{\hat{\eta}(\infty)} \right] \right. \\ & \quad \left. + \frac{[\hat{\eta}(t) - \hat{\eta}(t_1)] [\hat{\eta}(t_2) - \hat{\eta}(t)]}{\hat{\eta}(\infty)^2} \right\} \\ & \leq K_2^2 \left\{ \frac{[\hat{\eta}(t) - \hat{\eta}(t_1)] [\hat{\eta}(t_2) - \hat{\eta}(t)]}{\hat{\eta}(\infty)^2} \right\} \\ & \leq K_2^2 \left\{ \frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^2, \end{aligned} \quad (\text{A.12})$$

$$C_{2j'}(t_1, t) C_{2k}(t, t_2) \leq K_1^2 \left\{ \frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^2. \quad (\text{A.13})$$

여기서, $j, k = 1, \dots, n_1$; $j', k' = 1, \dots, n_2$ 이다.

이제, (A.6)-(A.13)으로부터 타이트성을 쉽게 검토할 수 있다.

$$\begin{aligned} & n^2 E \{ |n^{-1/2} \hat{U}(\hat{\beta}, t) - n^{-1/2} \hat{U}(\hat{\beta}, t_1)|^2 |n^{-1/2} \hat{U}(\hat{\beta}, t_2) - n^{-1/2} \hat{U}(\hat{\beta}, t)|^2 \} \\ & = \sum_{j=1}^{n_1} \sum_{k=1}^{n_1} C_{1j}(t_1, t)^2 C_{1k}(t, t_2)^2 E(G_{1j}^2 G_{1k}^2) + \sum_{j=1}^{n_1} \sum_{j'=1}^{n_2} C_{1j}(t_1, t)^2 C_{2j'}(t, t_2)^2 E(G_{1j}^2 G_{2j'}^2) \\ & \quad + \sum_{j=1}^{n_2} \sum_{k=1}^{n_2} C_{2j'}(t_1, t)^2 C_{2k}(t, t_2)^2 E(G_{2j'}^2 G_{2k}^2) + \sum_{j=1}^{n_1} \sum_{j'=1}^{n_2} C_{1j}(t_1, t)^2 C_{2j'}(t, t_2)^2 E(G_{1j}^2 G_{2j'}^2) \\ & \quad + 4 \sum_{j=1}^{n_1} \sum_{\substack{k=1 \\ j < k}}^{n_1} C_{1j}(t_1, t) C_{1k}(t_1, t) C_{1j}(t, t_2) C_{1k}(t, t_2) E(G_{1j}^2 G_{1k}^2) \\ & \quad + 4 \sum_{j=1}^{n_2} \sum_{\substack{k=1 \\ j < k}}^{n_2} C_{2j'}(t_1, t) C_{2k}(t_1, t) C_{2j'}(t, t_2) C_{2k}(t, t_2) E(G_{2j'}^2 G_{2k}^2) \\ & = 3 \sum_{j=1}^{n_1} C_{1j}(t_1, t)^2 C_{1j}(t, t_2)^2 + \sum_{j=1}^{n_1} \sum_{\substack{k=1 \\ j \neq k}}^{n_1} C_{1j}(t_1, t)^2 C_{1k}(t, t_2)^2 \\ & \quad + 3 \sum_{j=1}^{n_2} C_{2j'}(t_1, t)^2 C_{2j'}(t, t_2)^2 + \sum_{j=1}^{n_2} \sum_{\substack{k=1 \\ j \neq k}}^{n_2} C_{2j'}(t_1, t)^2 C_{2k}(t, t_2)^2 \\ & \quad + 2 \sum_{j=1}^{n_1} \sum_{j'=1}^{n_2} C_{1j}(t_1, t)^2 C_{2j'}(t, t_2)^2 \\ & \quad + 4 \sum_{j=1}^{n_1} \sum_{\substack{k=1 \\ j < k}}^{n_1} C_{1j}(t_1, t) C_{1k}(t_1, t) C_{1j}(t, t_2) C_{1k}(t, t_2) \\ & \quad + 4 \sum_{j=1}^{n_2} \sum_{\substack{k=1 \\ j < k}}^{n_2} C_{2j'}(t_1, t) C_{2k}(t_1, t) C_{2j'}(t, t_2) C_{2k}(t, t_2) \\ & \leq [3n_1^2 K_2^4 + 2n_1 n_2 K_1^2 K_2^2 + 3n_2^2 K_1^4] \left[\frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right]^4. \end{aligned}$$

그러므로, 어떤 $K(\infty)$ 에 대하여,

$$E\{|n^{-1/2} \hat{U}(\hat{\beta}, t) - n^{-1/2} \hat{U}(\hat{\beta}, t_1)|^2 | n^{-1/2} \hat{U}(\hat{\beta}, t_2) - n^{-1/2} \hat{U}(\hat{\beta}, t)|^2\} \\ \leq K \left\{ \frac{\hat{\eta}(t_2) - \hat{\eta}(t_1)}{\hat{\eta}(\infty)} \right\}^4$$

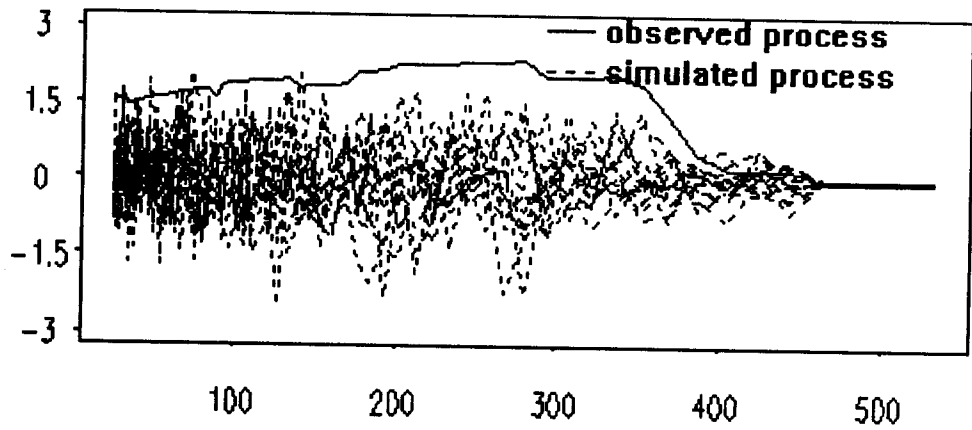
임으로 타이트성을 만족한다.

Lindeberg-Feller의 중심극한정리와 타이트성으로부터 $n^{-1/2} \hat{U}(\hat{\beta}, \cdot)$ 는 평균이 0인 가우스과정으로 수렴한다. 한편, $s \leq t$ 에 대하여,

$$E\{n^{-1/2} \hat{U}(\hat{\beta}, s) \cdot n^{-1/2} \hat{U}(\hat{\beta}, t) | (X_{ij}, \Delta_{ij}, Z_{ij})\} \\ = n^{-1} \left\{ \sum_{j=1}^{n_1} \left[\frac{Y_2(X_{1j})}{Y_1(X_{1j}) + Y_2(X_{1j})} \right]^2 \left[I(X_{1j} \leq s) - \frac{\hat{\eta}(s)}{\hat{\eta}(\infty)} \right] \left[I(X_{1j} \leq t) - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \right] \Delta_{1j} \right. \\ \left. + \sum_{j=1}^{n_2} \left[\frac{Y_1(X_{2j})}{Y_1(X_{2j}) + Y_2(X_{2j})} \right]^2 \left[I(X_{2j} \leq s) - \frac{\hat{\eta}(s)}{\hat{\eta}(\infty)} \right] \left[I(X_{2j} \leq t) - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \right] \Delta_{2j} \right\} \\ = n^{-1} \int_0^\infty \left\{ \left[I(u \leq s) - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} I(u \leq s) - \frac{\hat{\eta}(s)}{\hat{\eta}(\infty)} I(u \leq t) + \frac{\hat{\eta}(s) \hat{\eta}(t)}{\hat{\eta}(\infty)^2} \right] \right. \\ \left. \times \left\{ \left[\frac{Y_2(u)}{Y_1(u) + Y_2(u)} \right]^2 dN_1(u) + \left[\frac{Y_1(u)}{Y_1(u) + Y_2(u)} \right]^2 dN_2(u) \right\} \right\} \\ = \hat{\xi}(s) - \frac{\hat{\eta}(t)}{\hat{\eta}(\infty)} \hat{\xi}(s) - \frac{\hat{\eta}(s)}{\hat{\eta}(\infty)} \hat{\xi}(t) + \frac{\hat{\eta}(s) \hat{\eta}(t)}{\hat{\eta}(\infty)^2} \hat{\xi}(\infty) \\ = \hat{\sigma}(s, t) \\ \xrightarrow{p} \sigma(s, t)$$

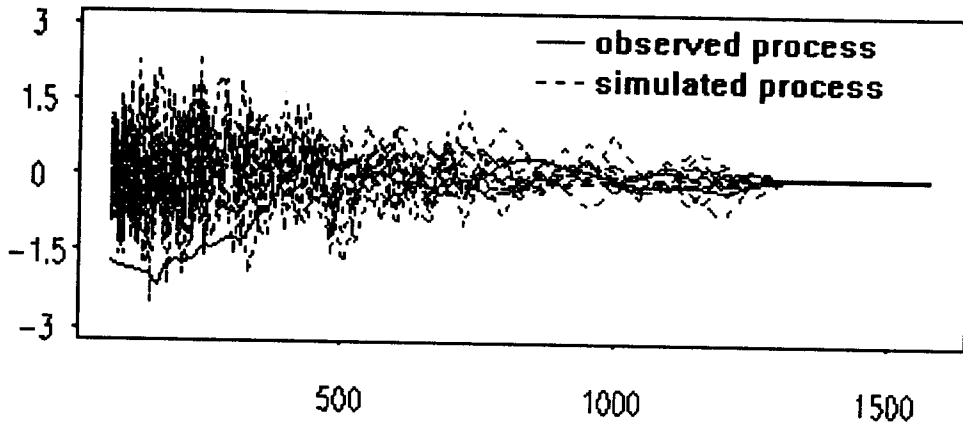
임을 알 수 있다. 따라서, $\{(X_{ij}, \Delta_{ij}, Z_{ij})\}$ 이 주어졌을 때 $n^{-1/2} \hat{U}(\hat{\beta}, \cdot)$ 은 $n^{-1/2} \tilde{U}(\beta, \cdot)$ 동일한 공분산 구조를 갖고 평균이 0인 가우스과정으로 수렴한다. ■

부록 B



Times(days) from treatment to disease progression of patients with ovarian cancer

그림 1 자궁암환자들의 병의 진행시간에 대한 관찰과정과 15개 시뮬레이션과정의 그림



Survival times(days) of patients with SCLC

그림 2 SCLC를 가진 환자들의 생존시간에 대한 관찰과정과 시뮬레이션과정의 그림

A Goodness-of-Fit Test Based on Martingale Residuals for the Additive Risk Model⁴⁾

Jinheum Kim⁵⁾, Seungyeoun Lee⁶⁾

Abstract

This paper proposes a goodness-of-fit test for checking the adequacy of the additive risk model with a binary covariate. The test statistic is based on martingale residuals, which is the extended form of Wei(1984)'s test. The proposed test is shown to be consistent and asymptotically normally distributed under the regularity conditions. Furthermore, the test procedure is illustrated with two set of real data and the results are discussed.

4) This research is supported by Korean Science and Engineering Foundation KOSEF #94-0701-01-3.

5) College Instructor, Department of Applied Statistics, The University of Suwon, Kyonggido, 445-743, Korea.

6) Assistant Professor, Department of Applied Statistics, Sejong University, Seoul, 133-747, Korea.