

# Review for time-dependent ROC analysis under diverse survival models

Yang-Jin Kim <sup>1,a</sup>

<sup>a</sup>Department of Statistics, Sookmyung Women's University

## Abstract

The receiver operating characteristic (ROC) curve was developed to quantify the classification ability of marker values (covariates) on the response variable and has been extended to survival data with diverse missing data structure. When survival data is understood as binary data (status of being alive or dead) at each time point, the ROC curve expressed at every time point results in time-dependent ROC curve and time-dependent area under curve (AUC). In particular, a follow-up study brings the change of cohort and incomplete data structures such as censoring and competing risk. In this paper, we review time-dependent ROC estimators under several contexts and perform simulation to check the performance of each estimators. We analyzed a dementia dataset to compare the prognostic power of markers.

Keywords: area under curve (AUC), inverse probability censoring weighting (IPCW), prediction, survival analysis, time-dependent receiver operating curve (ROC)

## 1. 서론

Receiver operating characteristic (ROC) 곡선은 이항 반응 변수에 대한 연속형 마커(marker)의 예측력을 판단하기 위해 적용될 수 있다. 환자의 상태를 진단해야 하는 의학 연구에서 질병 진단을 위한 마커의 역할은 매우 중요하다. 질병의 빠른 진단은 환자의 고통을 줄일 수 있으며 생존 가능성을 높일 수 있다. 한 예로 보건복지부 자료에 의하면 우리 나라 치매환자수는 2014년 416,309명에서 2018년 712,386명으로 5년간 71% 증가했으며 2018년 치매환자는 인구대비 1.37% 까지 증가하게 되었다. 또한, 통계청이 발표한 2019년 사망 원인 통계에 따르면 알츠하이머 치매로 사망한 인구는 전년 대비 약 10% 가까이 증가해 한국인의 10대 사망 원인 9위로 꼽혔다. 급속한 고령화 사회에서 이러한 치매율의 증가는 사회적, 경제적 문제를 가져올 수 있다. 특히 치매는 장기간 천천히 진행되는 경향이 있어 치매 발병을 인식하기 어려울 뿐만 아니라 일단 중증단계에 이르면 호전될 가능성이 그리 높지 않다. 따라서 초기 단계의 치매 진단을 통해 약물 치료와 운동 요법 등을 적용함으로써 중증 단계로 악화되는 것을 막을 수 있다. 질병 마커의 판별력을 검증하기 위해 여러 가지 통계 방법론이 적용될 수 있다. 이들 방법 중 ROC 곡선은 마커의 임계치를 변화시켜 가면서 구한 민감도(sensitivity)와 특이도(specificity)를 통해 구하게 된다. 예를 들어, 마커를  $X$ 로 표시하고 질병 여부를  $D = \{0, 1\}$ 로 표시하자. 여기서 마커 값이 높을수록 질병 발병가능성이 높다고 가정한다(만약 낮은 마커 값이 질병의 지표라면 아래 공식에서  $X$  대신에  $-X$ 가 사용되어야 한다). 임계치  $c$ 에 대해서 민감도와 특이도는 각각  $se(c) = \Pr(X > c | D = 1)$ 와

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2020R1A2C1A01100755)

<sup>1</sup> Department of Statistics, Sookmyung Women's University, Cheongpa-ro 47-gil 100, Yongsan-Gu, Seoul 04310, Korea.  
E-mail: yjin@sookmyung.ac.kr

$sp(c) = \Pr(X < c | D = 0)$ 로 표현되는데 임계치  $c$ 는  $X$ 가 취할 수 있는 모든 가능한 값이 적용될 수 있다.  $c$  값을 바꿔가며 계산된 민감도와 특이도를 이용하여 구한 ROC 곡선은  $(1 - sp(c), se(c)) = (FPR(c), TPR(c))$ 로 표시된다. 여기서 false positive rate (FPR)은 위양성이라하고 true positive rate (TPR)로 진양성이라 한다. 환자의 상태를 보여주는 마커와 질병 유무가 이미 알려져 있는 진단학 연구(diagnostic study)에서 로지스틱 모형의 진단력을 평가하기 위해 ROC 곡선과 곡선 아래 면적인 area under curve (AUC)가 적용되어 왔으며 AUC 값이 클수록 그 모형 예측력이 더 우수함을 의미한다. 또, 다른 표현으로 해석하면 AUC 값은 랜덤하게 선택된 두 사람(한 명은 환자, 또 다른 사람은 건강인)에 대해, 환자의 마커 값이 건강인의 마커 값보다 클 확률을 의미한다. 하지만 생존 분석에서는 반응 변수인 사건 발생 여부는 시간이 경과함에 따라 관측 가능하기 때문에 관련된 예측 연구(prognostic study)에서는 시간 가변(time dependent) 반응 변수의 상태(예를 들어, 사망 여부)를 분석에 반영할 필요가 있으며 이를 위해 민감도와 특이도는 시간의 함수로 정의되어야 한다. Heagerty와 Zheng (2005)은 이러한 시간 가변 ROC 곡선을 다음과 같이 세 가지 유형으로 정의하였다. 여기서  $T$ 는 사건 발생 시간을 의미한다.

- 누적 민감도(cumulative sensitivity), 다이내믹 특이도(dynamic specificity) :  $Se^C(c, t), Sp^D(c, t)$

$$Se^C(c, t) = P(X > c | T \leq t) \quad Sp^D(c, t) = P(X \leq c | T > t) \quad (1.1)$$

- 발생 민감도(incident sensitivity), 다이내믹 특이도(dynamic specificity):  $Se^I(c, t), Sp^D(c, t)$

$$Se^I(c, t) = P(X > c | T = t) \quad Sp^D(c, t) = P(X \leq c | T > t) \quad (1.2)$$

- 발생 민감도(incident sensitivity), 고정 특이도(static specificity):  $Se^I(c, t), Sp^S(c, t)$

$$Se^I(c, t) = P(X > c | T = t) \quad Sp^S(c, t) = P(X \leq c | T > t^*) \quad (1.3)$$

본 논문의 목적은 우중도 생존 분석과 경쟁 위험 모형에서 제안된 여러가지 시간 가변 ROC 곡선과 AUC 추정량을 정리하고 의미를 해석하는 것이다. 먼저 2장에서는 위 세 가지 유형 중 가장 많이 사용되는 누적 민감도와 다이내믹 특이도를 구하기 위한 여러 가지 추정량을 소개한다. 3장에서는 경쟁 위험 모형에 대해 2장의 추정량들이 어떻게 확장되었는지를 설명한다. 4장에서는 간단한 모의 실험을 통해 추정량의 적합도를 검토하며 5장에서는 치매 자료 분석을 통해 위 추정량들을 비교한다. 6장에서는 관련된 향후 연구내용을 제시함으로써 마무리한다.

## 2. 우중도 절단 자료에서 ROC 곡선 추정

전향 연구를 통해 구해지는 생존 분석 자료에서 가장 흔히 발생하는 우중도 절단 자료를 고려해보자.  $n$ 명의 관측 개체 중  $i$  번째 관측 개체에 대해 분석할 자료는 다음과 같다.

$$\{(T_i^*, \delta_i, X_i), \quad i = 1, \dots, n\}$$

여기서  $T_i^* = \min(C_i, T_i)$ 이며  $T_i$ 는 관심 있는 사건의 발생 시간을  $C_i$ 는 중도 절단 시간을 의미한다.  $\delta_i = I(T_i < C_i)$ 는 중도 절단 지시 함수이며 마커 값은  $X_i$ 로 표시된다. 우중도 절단은 관심 있는 사건과 독립적으로 발생한다고 가정하거나 공변량이 주어진 경우, 조건부 독립을 가정한다. 생존 분석의 특성 상, 위에서 정의된 세 가지 유형의 민감도와 특이도 정의에서 누적 민감도와 다이내믹 특이도가 가장 많이 적용된다. 즉,  $t$  시점까지 사건을 경험한 사례군(case)은 아직 사건을 경험하지 못한 대조군(control)보다 더 높은 마커값을 가질

것이라 기대된다. 식(1.1)에 정의된 조건부 확률을 이용하여 누적 민감도, 다이나믹 특이도와 해당 시간 가변 AUC에 대한 Naive 추정량은 다음과 같다.

$$\begin{aligned}\widehat{\text{Se}}_N^C(c, t) &= \frac{\sum_{i=1}^n \delta_i I(X_i > c, T_i^* \leq t)}{\sum_{i=1}^n \delta_i I(T_i^* \leq t)}, \\ \widehat{\text{Sp}}_N^D(c, t) &= \frac{\sum_{i=1}^n I(X_i \leq c, T_i^* > t)}{\sum_{i=1}^n I(T_i^* > t)}, \\ \text{AUC}^{\text{Naive}}(t) &= \frac{\sum_{i=1}^n \sum_{j=1}^n \delta_i I(X_i > X_j) I(T_i^* \leq t, T_j^* > t)}{\sum_{i=1}^n \sum_{j=1}^n \delta_i I(T_i^* \leq t) I(T_j^* > t)}.\end{aligned}\quad (2.1)$$

식 (2.1)에서 제시된 누적 민감도 추정량은 중도 절단률이 증가할수록 편이된 추정량을 가져올 수 있다. 예를 들어,  $t$  시점의 누적 민감도 계산에서,  $t$  시점 전에 중도 절단된 관측개체  $i$  ( $C_i < t$ )의 사건 발생 시간  $T_i$ 는 알려져 있지 않으며 대신  $C_i < T_i$  라는 정보만을 가지게 된다. 따라서  $t$ 가 증가할수록 무시되어지는 중도 절단 자료의 비율은 점점 더 커지게 된다. 하지만 중도 절단 자료를 누적 민감도 계산에서 완전히 제외시키는 것은 표본의 편이성(biased sampling)을 가져오게 된다. 따라서 이러한 관측 자료의 불완전성을 보정하기 위해 여러 가지 방법들이 제안되었다. Heagerty 등 (2000)이 제시한 Kaplan-Meier를 이용한 방법과 비모수 이변량 분포함수를 이용한 방법, Chambles와 Diao (2006)가 제시한 재귀(recursive)방법에 의해 재계산된 위험 그룹을 이용한 방법, Uno 등 (2007)과 Hung과 Chiang (2010)이 제안한 편이 추출을 보정하기 위한 역확률 중도 절단 가중(inverse probability censoring weighting, IPCW)방법이 있다. 다음 절에서 이들 방법들을 간략하게 요약하고자 한다.

## 2.1. Kaplan-Meier based estimator: Heagerty *et al.* (2000)

중도 절단 자료의 영향력을 반영하기 위해 카플란 마이어 추정량을 이용하여 구한 누적 민감도와 다이나믹 특이도의 추정량은 다음과 같다.

$$\widehat{\text{Se}}_{H,1}^C(c, t) = \frac{[1 - \hat{S}(t|X > c)](1 - \widehat{F}_X(c))}{1 - \hat{S}(t)}, \quad \widehat{\text{Sp}}_{H,1}^C(c, t) = \frac{\hat{S}(t|X \leq c) \widehat{F}_X(c)}{\hat{S}(t)}.\quad (2.2)$$

여기서  $\widehat{F}_X$ 는 마커  $X$ 의 분포함수이며  $\hat{S}(t)$ 는 주변(marginal) 생존 함수의 Kaplan-Meier 추정량이다. 마커 값의 영향력을 반영하기 위해 주어진  $c$ 를 이용하여  $X > c$  또는  $X \leq c$ 를 만족하는 두 그룹에 대해 조건부(conditional) 생존 함수의 Kaplan-Meier 추정량  $\hat{S}(t|X > c)$ 와  $\hat{S}(t|X \leq c)$ 을 각각 구하게 된다. 하지만 식 (2.2)는 중도 절단이 마커와 연관된 경우에 편이 추정량을 가져오게 되며 이러한 마커 종속 중도 절단(marker dependent censoring)은 임상 연구에서 매우 흔히 발생된다. 위 추정량의 더욱 심각한 단점은 고정된 시점  $t$ 에서 변화되어지는  $c$ 값에 대해  $\widehat{\text{Se}}_{H,1}^C(c, t)$ 와  $\widehat{\text{Sp}}_{H,1}^D(c, t)$ 가 단조성을 보장하지 못하는 경우가 발생하게 되며 이는 ROC(t)가  $[0, 1] \times [0, 1]$  범위를 벗어나는 결과를 가져오게 된다. 이러한 비단조성(non-monotonicity)은 조건부 생존 함수 추정량  $\hat{S}(t|X > c)$ 와  $\hat{S}(t|X \leq c)$  계산에 적용될 관측 개체( $X > c$  또는  $X \leq c$ )들이  $c$  값에 따라 변하기 때문이다. Heagerty 등 (2000)은 이러한 단점을 극복하기 위해 Akritas (1987)에 의해 제안된 비모수 스무딩 기법을 이용한 가중 카플란 마이어 함수(weighted Kaplan-Meier)를 제안하였다.

$$\hat{S}_\lambda(t|x) = \prod_{s \leq t} \left( 1 - \frac{\sum_j K_\lambda(x, x_j) I(T_j = s) \delta_j}{\sum_j K_\lambda(x, x_j) I(T_j \geq s)} \right),$$

여기서  $K_\lambda(x, x_j)$  는 스무딩 모수  $\lambda$  에 의해 조정되는 커널 함수  $K_\lambda(x, x_j) = I(-\lambda \leq \widehat{F}_X(x) - \widehat{F}_X(x_j) < \lambda)$ 이며 이를 이용한 누적 민감도와 다이내믹 특이도의 추정량은 다음과 같았다.

$$\widehat{\text{Se}}_{H,2}^C(c, t) = \frac{(1 - \widehat{F}_X(c)) - \widehat{S}_\lambda(c, t)}{1 - \widehat{S}_\lambda(t)}, \quad \widehat{\text{Sp}}_{H,2}^D(c, t) = 1 - \frac{\widehat{S}_\lambda(c, t)}{\widehat{S}_\lambda(t)}, \quad (2.3)$$

$\widehat{S}_\lambda(t) = \widehat{S}_\lambda(-\infty, t)$ 이며 식 (2.2)와 식 (2.3)과 관련된 AUC를 각각  $\text{AUC}^{\text{KM}}(t)$ 와  $\text{AUC}^{\text{NNE}}(t)$ 라 하며 *survivalROC* 패키지를 이용하여 구할 수 있다.

## 2.2. Chamless and Diao (2006) estimator

Chamless와 Diao (2006)는 우중도 절단의 효과를 보정하기 위해 순서 통계량과 베이즈 정리를 이용하였다.  $\delta = 1$  을 가지는 사건 발생 시간을 순서대로 다음과 같이 나열한다  $0 = \tilde{s}_0 < \tilde{s}_1 < \dots < \tilde{s}_m$ . 따라서  $\tilde{s}_m$  은 사건 발생 시간의 최대값을 의미한다. 특정 시점  $t$ 에서의 누적 민감도와 다이내믹 특이도를 구하기 위해,  $0 = \tilde{s}_0 < \tilde{s}_1 < \dots < \tilde{s}_{m(t)} < t$ 에서 관측된 정보가 사용되는데 여기서  $\tilde{s}_{m(t)}$ 은  $t$  시점보다 빠른 관측 시점 중에서  $t$ 와 가장 가까운 시점을 의미한다. 베이즈 정리를 이용하여 누적 민감도  $P(X > c | T \leq t) = P(X > c, T \leq t) / P(T \leq t)$ 의 분자를  $\sum_{k=1}^{m(t)} P(X > c | \tilde{s}_{k-1} < T \leq \tilde{s}_k) \times P(\tilde{s}_{k-1} < T \leq \tilde{s}_k)$ 로 유도하고  $P(X > c | \tilde{s}_{k-1} < T \leq \tilde{s}_k)$ 에 대한 추정량  $I(X_{(k)} > c)$ 과  $P(\tilde{s}_{k-1} < T \leq \tilde{s}_k)$ 의 추정량  $\widehat{S}(\tilde{s}_{k-1}) - \widehat{S}(\tilde{s}_k)$ 을 이용하여 다음의 누적 민감도 추정량을 구하게 된다. 비슷하게 다이내믹 특이도  $P(X \leq c | T > t)$ 의 추정량을 위해  $P(X \leq c | \tilde{s}_{k-1} < T \leq \tilde{s}_k)$ 의 추정량  $I(X_{(k)} \leq c)$ 과  $X$ 의 분포함수 ( $F_X(x) = P(X \leq x)$ ) 추정량  $\widehat{F}_X(c)$ 을 다음과 같이 적용한다.

$$\widehat{\text{Se}}_{\text{CD}}^C(c, t) = \frac{\sum_{k=1}^{m(t)} I(X_k > c) (\widehat{S}(\tilde{s}_{k-1}) - \widehat{S}(\tilde{s}_k))}{1 - \widehat{S}(\tilde{s}_{m(t)})}, \quad \widehat{\text{Sp}}_{\text{CD}}^D(c, t) = \frac{\widehat{F}_X(c) - \sum_{k=1}^{m(t)} I(X_k \leq c) (\widehat{S}(\tilde{s}_{k-1}) - \widehat{S}(\tilde{s}_k))}{\widehat{S}(\tilde{s}_{m(t)})}. \quad (2.4)$$

위의  $\widehat{\text{Se}}_{\text{CD}}^C(c, t)$ 는 단조성을 만족하는 반면,  $\widehat{\text{Sp}}_{\text{CD}}^D(c, t)$  추정량은 여전히 단조성 성질을 만족하지 않아 관련 ROC 곡선도  $[0, 1] \times [0, 1]$  범위 밖에서 정의될 수 있다. 식(2.4)와 관련된  $\text{AUC}^{\text{CD}}(t)$ 는 *survAUC* 패키지를 이용하여 구할 수 있다.

## 2.3. Uno *et al.* (2007) and Hung and Chiang (2010) estimator

Uno 등 (2007)과 Hung과 Chiang (2010)이 제안한 편이 추출을 보정하기 위한 역확률 중도 절단 가중(inverse probability censoring weighting, IPCW)방법은 생존 함수에서 편이 표본(biased sampling) 또는 중도 절단과 관심 사건의 독립성이 위배되는 경우, 예를 들어, 경쟁 위험 모형과 재발 사건 자료의 종속 절단(induced dependent censoring)에 적용되어 왔다. IPCW의 주요 개념은 관측 확률로 중도 절단 되지 않을 확률을 가중치로 사용함으로써 표본 편이를 보정하고자 하는 것이다. 식 (2.1)의 Naive 추정량을 확장하여 누적 민감도와 다이내믹 특이도를 다음과 같이 유도할 수 있다.

$$\widehat{\text{Se}}_{\text{IPCW}}^C(c, t) = \frac{\sum_{i=1}^n \frac{\delta_i}{n\widehat{S}_C(T_i^*)} I(X_i > c, T_i^* \leq t)}{\sum_{i=1}^n \frac{\delta_i}{n\widehat{S}_C(T_i^*)} I(T_i^* \leq t)}, \quad \widehat{\text{Sp}}_{\text{IPCW}}^D(c, t) = \frac{\sum_{i=1}^n I(X_i \leq c, T_i^* > t) \frac{1}{n\widehat{S}_C(t)}}{\sum_{i=1}^n I(T_i^* > t) \frac{1}{n\widehat{S}_C(t)}}, \quad (2.5)$$

여기서  $\widehat{S}_C(\cdot)$ 는 중도 절단 시간  $C$ 의 생존 함수 추정량으로  $(T_i^*, 1 - \delta_i)$ 를 이용하여 Kaplan-Meier 추정량을 적용된다. 위에서 정의된 누적 민감도와 다이내믹 특이도를 이용한 ROC 곡선은  $[0, 1] \times [0, 1]$ 의 범위에서 정의된다. Blanche 등 (2013)은  $\widehat{\text{Se}}_{\text{IPCW}}^C(c, t)$ 가  $\widehat{\text{Se}}_{\text{CD}}^C(c, t)$ 와 거의 유사함을 증명하였으며 다음의 시간 가변

Table 1: Summary of R package for a time dependent ROC curve

R package	Description	Estimator
<b>survivalROC</b>	$AUC^{KM}, AUC^{NNE}$	(2.2), (2.3)
<b>survAUC</b>	$AUC^{CD}, AUC^{IPCW}$	(2.4), (2.5)
<b>timeROC</b>	$AUC^{IPCW}, AUC^{CIPCW}, AUC^{CR,1}$	(2.5), (2.6), (3.3), (3.4)
<b>CompRisksROC, SurvCompetingRisk</b>	$AUC^{CR,1}, AUC^{CR,2}$	(3.5), (3.6)
<b>&amp; risksetROC</b>	$AUC^{*,CRC}, AUC^{CRC}$	& (3.7), (3.8)

AUC 추정량을 유도하였다.

$$\widehat{AUC}^{IPCW}(t) = \frac{\sum_{i=1}^n \sum_{j=1}^n \delta_i I(X_i > X_j), I(T_i^* \leq t, T_j^* > t) \frac{\delta_i}{n^2 \hat{s}_C(T_i^*) \hat{s}_C(t)}}{\sum_{i=1}^n \sum_{j=1}^n \delta_i I(T_i^* \leq t) I(T_j^* > t) \frac{\delta_i}{n^2 \hat{s}_C(T_i^*) \hat{s}_C(t)}}.$$

하지만 지금까지 설명한 어떤 AUC 추정량도 마커 종속 중도 절단(marker dependent censoring)을 고려하지는 못하였다. 이를 위해 Blanche 등 (2013)은 중도 절단 조건부 생존 함수 추정량을 구하기 위해 마커값을 공변량으로 하는 회귀모형을 이용하였다. Cox의 비례 위험 모형, Aalen의 가법 모형이 적용되거나 비모수 스무딩 기법을 적용하여 추정될 수 있다. 이렇게 구한  $\hat{S}_C(t|X)$ 는 민감도, 특이도 그리고 AUC에 대해 다음과 같이 적용된다.

$$\begin{aligned} \widehat{Se}_{CIPCW}^C(c, t) &= \frac{\sum_{i=1}^n \frac{\delta_i}{n \hat{s}_C(T_i^*|X_i)} I(X_i > c, T_i^* \leq t)}{\sum_{i=1}^n \frac{\delta_i}{n \hat{s}_C(T_i^*|X_i)} I(T_i^* \leq t)}, \quad \widehat{Sp}_{CIPCW}^D(c, t) = \frac{\sum_{i=1}^n I(X_i \leq c, T_i^* > t) \frac{1}{n \hat{s}_C(t|X_i)}}{\sum_{i=1}^n I(T_i^* > t) \frac{1}{n \hat{s}_C(t|X_i)}}, \\ \widehat{AUC}^{CIPCW}(t) &= \frac{\sum_{i=1}^n \sum_{j=1}^n \delta_i I(X_i > X_j), I(T_i^* \leq t, T_j^* > t) \frac{\delta_i}{n^2 \hat{s}_C(T_i^*|X_i) \hat{s}_C(t|X_j)}}{\left( \sum_{i=1}^n \frac{\delta_i}{n \hat{s}_C(T_i^*|X_i)} I(T_i^* \leq t) \right) \left( \sum_{j=1}^n I(T_j^* > t) \frac{1}{n \hat{s}_C(t|X_j)} \right)}. \end{aligned} \quad (2.6)$$

추정량 (2.5)와 (2.6)은 *survAUC* 와 *timeROC* 를 이용하여 구할 수 있다. Table 1은 위에서 설명한 여러 가지 추정량을 포함한 패키지를 보여준다.

### 3. 경쟁 위험 모형에서 receiver operating characteristics (ROC) 곡선 추정

일반적인 우중도 절도 생존 자료는 이항 자료(사건 발생 여부)와 사건 발생 시간의 결합 자료로 이해할 수 있다. 이항 자료의 확장으로 경쟁 위험 모형(competitive risk model)은 사건 발생 시간과 그 사건 발생의 원인을 다항 자료로 표현할 수 있다. 만약 관심 사건 외의 모든 사건을 중도 절단으로 간주할 경우 중도 절단 사건과 관심 사건의 독립성이 위배되며 이는 편이된 추정 결과를 가져올 수 있다. 예를 들어, 노년층을 대상으로 치매 발병이 관심 사건인 경우 사망 사건은 중도 절단 사건으로 간주하기 보다는 경쟁 사건으로 분석되어야 한다. 즉, 이른 시점에 사망할 경우 치매를 경험할 수 있는 상태에 이르지 못하는 반면 오랫동안 생존하는 사람은 치매 위험에 상대적으로 노출될 수 기간이 더 길어지게 된다. 따라서 위험 노출에 대한 편이를 보정하기 위한 방법이 적절하게 고려될 필요가 있으며 역확률 중도 절단 가중치(IPCW)를 이용한 회귀 모형 (Fine과 Gray, 1999)이 널리 적용되고 있다. 경쟁 위험 모형에서도 마커 또는 모형의 예측력을 평가하기 위한 여러 가지 방법론이 제안되었다. 경쟁 위험 모형은 다단계 모형의 한 가지 유형으로 간주할 수 있는데 일반적인 우중도 절단 생존모형이 0(생존)  $\rightarrow$  1(사망)의 모형이라면 경쟁위험 모형은 마지막 흡수 상태(absorbing state)가 여러 가지 원인을 가지는 사망으로 간주될 수 있을 것이다. 이러한 개념을 확장하여 Foucher 등 (2010)은 두 가지 경쟁 사건(사망, 치매)에 대해 3개의 그룹으로 분류한 후 ROC 곡선을 적용하였다. Blanche 등 (2013b)

은 Uno 등 (2000)과 Hung과 Chiang (2010)이 제안한 IPCW에 기반한 ROC 곡선 추정량을 경쟁 위험 모형으로 확장하였다.

경쟁 위험 모형의 관측 자료는  $\{(\tilde{T}_i, \delta_i, \eta_i, X_i), i = 1, \dots, n\}$  로 구성된다. 여기서  $\eta_i$  는 경쟁 위험의 원인에 대한 지시함수로  $K$  개의 경쟁 사건에 대해  $\{1, 2, \dots, K\}$ 의 값을 가진다. 여기서 주 관심사건이  $\eta_i = 1$  일 때, 누적 민감도 (1.1) 은 다음과 같이 재정의된다.

$$Se_k^{CR}(c, t) = Pr(X > c | T \leq t, \eta = 1).$$

위 통계량의 추정량을 구하기 위해 다음의 원인 별 누적 발생 함수(cumulative incidence function)가 적용된다.

$$C_j(t) = Pr(T \leq t, \eta = j), \quad j = 1, 2, \dots, K, \quad \lim_{t \rightarrow \infty} \sum_{j=1}^K C_j(t) = 1.$$

관심 사건이  $\eta = 1$  로 표시될 때, Sara와 Heagerty (2010)는 사건 별 누적 민감도를 다음과 같이 유도하였다.

$$\widehat{Se}_1^{CR}(c, t) = P(X > c | T \leq t, \eta = 1) = \frac{\int_c^\infty \hat{C}_1(t|X = u) g_X(u) du}{\int_{-\infty}^\infty \hat{C}_1(t|X = u) g_X(u) du}, \quad (3.1)$$

여기서  $g_X(\cdot)$  는 마커  $X$  의 확률 밀도 함수(pdf) 이며  $\hat{C}_j(t|X = m)$  는 다음과 같이 추정된다.

$$\hat{C}_j(t|X = x_i) = \sum_{s \leq t} \hat{S}(s|X = x_i) \hat{\lambda}_j(s|X = x_i)$$

여기서  $\hat{\lambda}_j(s|x)$  는 원인별 위험 함수(cause-specific hazard function)를 의미한다.

### 3.1. Three-class 진단방법의 확장

Foucher 등 (2010)는 원인별 위험 함수  $\lambda_j(t|x)$ 를 직접적으로 이용하여 누적 민감도와 다이내믹 특이도를 유도하였다. 마커에 대한 비례 회귀모형을 적용하여, 원인  $j$  에 대한 생존 함수는  $S_j(t|x) = S_{0j}(t)^{\exp(\alpha_j' X)}$ 로 정의된다. 여기서 회귀계수  $\alpha_j$  를 추정하기 위해 다른 원인에 의한 모든 사건 발생은 중도 절단으로 간주하게 된다. 이에 통해 구한  $z_j = \alpha_j' X$  를 마커로 이용한 누적 민감도와 다이내믹 특이도를 다음과 같이 유도하였다.

$$Se_2^{CR}(c, t) = \frac{\int_c^\infty (1 - S_j(t|z_j)) g_j(z_j) dz_j}{\int_{-\infty}^\infty (1 - S_j(t|z_j)) g_k(z_j) dz_j}, \quad Sp_2^{CR}(c, t) = \frac{\int_c^\infty (S_j(t|z_j)) g_j(z_j) dz_j}{\int_{-\infty}^\infty (S_j(t|z_j)) g_j(z_j) dz_j}. \quad (3.2)$$

### 3.2. Inverse probability censoring weighting (IPCW) 방법의 적용

Blanche 등 (2013b)은 경쟁 사건에 대한 표본 편이를 고려하기 위해 다음 두 가지 시간 가변 다이내믹 특이도를 정의한다.

$$Sp_{IPCW}^{CR,*}(c, t) = Pr(X \leq c | T > t), \quad (3.3)$$

$$Sp_{IPCW}^{CR}(c, t) = Pr(X \leq c | \{T > t \cup \{T \leq t, \eta \neq 1\}\}). \quad (3.4)$$

여기서 식 (3.3)과 (3.4)의 차이는 경쟁 사건에 대한 처리에 있다.  $t$  시간 전에 발생된 경쟁 사건을 식 (3.3)에 서는 대조군(control)으로 고려하지 않는 반면에, 식 (3.4)에서는 위험그룹으로 인식함으로써 대조군 그룹에

포함시키게 된다. 다만 경쟁 사건 발생시점에 따라 가중치를 부여함으로써 위험그룹의 기여도가 시간에 따라 변화도록 한다. 위 두 가지 특이도에 따라 AUC 추정량은 다음과 같이 두 가지로 표현된다.

$$\text{AUC}_{\text{IPCW}}^{\text{CR},*}(t) = \Pr(X_i > X_l | T_i \leq t, \eta_i = 1, T_l > t), \quad (3.5)$$

$$\text{AUC}_{\text{IPCW}}^{\text{CR}}(t) = \Pr(X_i > X_l | T_i \leq t, \eta_i = j, \{T_l > t \cup \{T_l \leq t, \eta \neq j\}\}). \quad (3.6)$$

식 (3.2)에 기반하여 구한  $\text{AUC}^{\text{CR}}(t)$  의 의미는  $t$  시점에 원인  $j$ 의 사건을 가진 환자와  $t$  시점까지 원인  $j$ 의 사건을 가지지 않은 환자를 구분하는 확률로 해석된다. 반면에, 식 (3.3)과 (3.4)을 이용한  $\text{AUC}_{\text{IPCW}}^{\text{CR}}(t)$  와  $\text{AUC}_{\text{IPCW}}^{\text{CR},*}(t)$ 에 대한 해석은  $t$ 시점에서 원인  $j$  를 가진 환자와 어떤 사건도 경험하지 않은 환자를 구분하는 확률로 해석된다.

만약 마커가 경쟁 사건 발생 확률에 영향을 주는 경우, 위에서 정의된 식 (3.3)과 (3.4)는 매우 다른 결과를 가지게 된다. 마커와 관련된 중도절단의 영향력을 반영하기 위해 Blanche 등 (2013b)은 누적 민감도와 두 가지 유형의 다이나믹 특이도를 제안하였다.

$$\begin{aligned} \widehat{\text{Se}}(c, t) &= \frac{\sum_{i=1}^n \delta_i \widehat{W}(T_i^*) I(X_i > c, T_i^* \leq t, \eta_i = j)}{\sum_{i=1}^n \delta_i \widehat{W}(T_i^*) I(T_i^* \leq t, \eta_i = j)}, \\ \widehat{\text{Sp}}^*(c, t) &= \frac{\sum_{i=1}^n I(X_i \leq c, T_i^* > t)}{\sum_{i=1}^n I(T_i^* > t)}, \\ \widehat{\text{Sp}}(c, t) &= \frac{\sum_{i=1}^n I(X_i \leq c) (I(T_i^* > t) \widehat{W}(t) + I(T_i^* \leq t, \eta_i \neq \{0, j\}) \widehat{W}(T_i^*))}{\sum_{i=1}^n I(T_i^* > t) \widehat{W}(t) + I(T_i^* \leq t, \eta_i \neq \{0, j\}) \widehat{W}(T_i^*)}. \end{aligned}$$

해당 AUC 추정량을 다음과 같이 유도할 수 있다.

$$\begin{aligned} \widehat{\text{AUC}}^{*,\text{CR}_C}(t) &= \frac{\sum_{i=1}^n \sum_{j=1}^n I(X_i > X_j), I(T_i^* \leq t, T_j^* > t, \eta_i = j) \frac{\delta_i}{n^2 \widehat{S}_C(T_i^*) \widehat{S}_C(t)}}{\left\{ \sum_{i=1}^n I(T_i^* \leq t, \eta_i = j) \widehat{W}(T_i^*) \right\} \left\{ \sum_{j=1}^n I(T_j^* > t) \widehat{W}(t) \right\}}, \quad (3.7) \\ \widehat{\text{AUC}}^{\text{CR}_C}(t) &= \frac{\sum_{i=1}^n \sum_{j=1}^n I(X_i > X_j), I(T_i^* \leq t, \eta_i = j) \widehat{W}(T_i^*) (I(T_j^* > t) \widehat{W}(t) + I(T_j^* < t, \eta_j \neq \{0, j\}) \widehat{W}(T_j^*))}{\left\{ \sum_{i=1}^n I(T_i^* \leq t, \eta_i = j) \widehat{W}(T_i^*) \right\} \left\{ \sum_{j=1}^n I(T_j^* > t) \widehat{W}(t) + I(T_j^* \leq t, \eta_j \neq \{0, j\}) \widehat{W}(T_j^*) \right\}}. \quad (3.8) \end{aligned}$$

여기서  $\widehat{W}(t) = 1/\widehat{S}_C(t)$  로 정의된다.

#### 4. 모의 실험

앞에서 설명된 여러 가지 시간 가변 AUC(t) 추정량의 특성을 검토하기 위해 간단한 모의 실험을 실시한다. 총 300번의 반복 시행을 하였으며 표본 크기는  $n = 300$  이 적용되었다. 마커값을 유도하기 위해 공변량  $X \sim N(0, 1)$  에서 추출되었으며 생존 시간  $T$ 와 중도 절단 시간  $C$  의 위험 함수는  $\lambda_T(t) = \lambda_0(t) \exp(\beta_1 X)$ ,  $\lambda_C(c) = \lambda_1(c) \exp(\beta_2 X)$ 로 정의되며  $\lambda_0(t)$ 와  $\lambda_1(t)$ 는 와이블 위험 함수  $\lambda(t) = \alpha/\theta^\alpha t^{\alpha-1}$ 로  $\lambda_0$ 에 대해서는  $(\alpha, \theta) = (2, 2.5)$ 가 적용되며  $\lambda_1(t)$ 의  $(\alpha, \theta)$ 에 대해서는 AUC와 중도 절단율에 따라 서로 다른 값이 적용되었다. 한 시점  $t = 2$ 에 대해  $\text{AUC}(2) = 0.92(\beta_1 = 2.0)$ 와  $\text{AUC}(2) = 0.70(\beta_1 = 0.5)$ 의 추정량을 구하였다. 또한 마커와 중도 절단과의 연관 관계 정도를 모형화하기 위해 중도 절단 위험률의  $\beta_2$ 에 대해 세 가지 값이 적용되었다.  $\beta_2 = (0, 0.3, 0.9)$ 에서  $\beta_2 = 0$ 은 중도 절단이 마커와 무관함을  $\beta_2 = 0.3$ 은 마커와 중간 정도의 연관 관계를  $\beta_2 = 0.9$ 는 중도 절단과 마커가 강한 연관관계를 의미한다. Table 2는 중도 절단율이 20%와 45% 하에서 AUC(2)의 결과 값을

Table 2: Comparison of biases and RMSEs under several situation( $\times 1000$ )

			Censoring rate=20%					
AUC(2)	$\beta_2$		AUC <sup>naive</sup>	AUC <sup>KM</sup>	AUC <sup>NNE</sup>	AUC <sup>CD</sup>	AUC <sup>IPCW</sup>	AUC <sup>CIPCW</sup>
0.70	0.0	Bias	-6.890	0.383	-11.053	0.476	0.195	0.362
		RMSE	1.014	1.098	0.288	0.301	0.285	0.241
	0.3	Bias	-16.12	-3.48	-13.54	-0.981	6.287	0.163
		RMSE	0.272	0.233	0.302	0.267	0.279	0.223
	0.9	Bias	-36.20	-12.18	-15.41	-6.268	17.32	-0.121
		RMSE	0.328	0.232	0.293	0.277	0.279	0.213
0.92	0.0	Bias	-3.759	0.775	-5.755	-1.609	0.059	0.108
		RMSE	1.014	1.098	0.323	3.309	1.387	1.182
	0.3	Bias	-5.756	-1.436	-5.724	-3.180	1.614	0.001
		RMSE	1.179	0.911	1.225	2.844	1.247	0.927
	0.9	Bias	-10.065	-4.516	-6.130	-2.730	3.381	-0.047
		RMSE	2.487	1.236	1.673	3.333	1.621	1.350
			Censoring rate=45%					
AUC(2)	$\beta_2$		AUC <sup>naive</sup>	AUC <sup>KM</sup>	AUC <sup>NNE</sup>	AUC <sup>CD</sup>	AUC <sup>IPCW</sup>	AUC <sup>CIPCW</sup>
0.70	0.0	Bias	-42.137	-0.435	-14.52	-3.168	0.156	-0.204
		RMSE	1.847	1.773	2.089	4.727	4.071	2.024
	0.3	Bias	-73.35	-13.91	-12.93	-5.39	18.96	-1.119
		RMSE	1.900	1.563	1.886	4.548	2.447	1.858
	0.9	Bias	-157.83	-59.23	-32.21	-10.54	30.62	-0.983
		RMSE	16.934	3.645	3.461	4.756	9.251	5.652
0.92	0.0	Bias	18.97	0.498	-5.887	-2.992	0.603	0.883
		RMSE	0.748	0.424	0.424	0.376	0.646	0.419
	0.3	Bias	-25.30	-6.870	-6.540	-2.357	6.097	-0.136
		RMSE	0.978	0.414	0.404	0.323	0.565	0.394
	0.9	Bias	-41.47	-20.83	-8.14	-2.46	12.73	1.046
		RMSE	2.168	0.919	0.517	0.349	0.752	0.456

이용하여 구한 편이(bias)(( $\widehat{AUC}(2) - AUC(2)$ )  $\times 10^3$ )의 300번 개의 평균값과 RMSE를 보여준다. 먼저 중도 절단율이 20% 인 경우,  $AUC(2)^{Naive}$  추정량의 편이가 가장 큼을 알 수 있다.  $\beta_2 = 0$  인 경우에 대부분의 추정량은 비슷한 값을 가진다. 중도 절단과 마커값의 연관 관계가 커질수록 ( $\beta_2 = 0.3, \beta_2 = 0.9$ ),  $AUC(2) = 0.70$ 에 대한 추정량들의 편이가 조금씩 증가함을 볼 수 있다. 그러나  $AUC(2)^{CIPCW}$ 는 모든 경우에서 가장 작은 편이를 가지면 참값과 가장 가까운 추정치를 보여준다.  $AUC(2) = 0.70$ 와  $AUC(2) = 0.92$ 의 추정값을 비교해보면  $AUC(2) = 0.7$ 일 때 0.92보다 편이가 더 큼을 알 수 있으며 중도 절단율 45%에서 중도 절단율 20%보다 편이가 더 증가함을 알 수 있다. 또한 여전히  $AUC(2) = 0.7$ 에서 추정량의 편이가  $AUC(2) = 0.92$ 의 추정량의 편이보다 더 크며 마커와 중도 절단과의 연관관계가 증가할수록 편이가 증가하였다. 하지만  $\beta_2 = 0.9$ 일 때  $AUC^{CIPCW}(2)$ 는 여전히 가장 작은 편이를 가짐을 알 수 있다. 하지만 RMSE에 대해서는 중도 절단율이 45%이고  $AUC^{KM}$ 과  $AUC^{NNE}$ 와 비슷하거나 더 큰 값을 준다.

## 5. 자료분석

앞에서 설명한 우중도 절단 자료의 누적 민감도와 다이내믹 특이도를 적용하기 위해 프랑스에서 수집된 치매자료를 분석하고자 한다 (Letenneur 등, 1994). 치매 관련 자료의 주요 목적은 노인의 향후 치매 발병 위험 정도를 예측하고 예방하는 것이므로 치매 위험 노인들을 대상으로 꾸준하게 세 가지 인지 테스트 mini-



Table 3: PH model application: Estimates and standard errors of predictors of Paquid data

		MMRT	Age	BVRT	IST	Score of physical dependency
Marker 1	$\beta$	-0.191	0.116			
	se(p-value)	0.024(<0.001)	0.017(<0.001)			
Marker 2	$\beta$	-0.135	0.117	-0.108	-0.034	-0.278
	se(p-value)	0.036(0.0002)	0.0187(<0.001)	0.044(0.014)	0.020(0.014)	0.132(0.036)

mental state examination (MMSE), Benton visual retention test (BVRT), Isaacs set test (IST)와 신체 자립도 점수(Score of physical dependency)를 포함한 예측모형을 고려해본다. 연구에 참여한 487명의 노인들의 연구 참여 시 연령 범위는(65.22, 92.14)이었으며 연구 참여 이후 연구가 종료할 때까지 위 테스트 점수는 방문 할 때마다 기록되며 환자마다 관측 횟수는 적게는 한번 많게는 9번을 면담하였다. 이들 중에 치매 진단을 받은 노인은 126명(26%)이었으며 평균 발병 나이는 85.15세이고 범위는(67.51, 99.49)이었다. 치매 여부는 정신과 전문의의 진단과 신경학과 전문의의 뇌 초음파 자료에 의해 진단된다. 자료 분석의 목적은 연구 참여 이후 치매 발병을 예측하기 위한 마커의 예측력을 평가하기 위해 앞에서 설명한 여러가지 우중도 절단 자료에 대한 여러가지 AUC 통계량을 적용하고 그 결과값을 비교하고자 한다. 관심있는 반응 변수는 연구 시작에서 치매 발병까지 걸린 시간이며 참여자의 연구 참여 나이가 동일하지 못한 점을 반영하기 위해 공변량으로 참여 당시의 나이를 포함한다. 마커값을 구하기 위해 두 가지 예측 모형이 고려되었다. 세 가지 인지 테스트 중 가장 많이 사용되는 MMSE와 연구 참여 당시 나이를 비례 위험 모형에 적용하여 얻은 회귀계수 값을 이용하여  $M_{i1} = -0.191\text{MMSE}_i + 0.116\text{Age}_i$  이었다. 여기서 MMSE 점수가 낮을수록 나이가 많을수록 치매 발병이 빨리 진행됨을 알 수 있다. 또 다른 마커에서는 세 가지 인지 테스트 값(MMSE, BVRT, IST), 신체 자립도(Score of physical dependency) 그리고 나이를 이용하여 구한 회귀 계수값이 적용된다.  $M_{i2} = -0.135\text{MMSE}_i - 0.10\text{BVRT}_i - 0.034\text{IST}_i - 0.278\text{Score of physical dependency}_i + 0.117\text{Age}_i$ . 즉, 인지 관련 점수가 낮고 신체 자립도가 낮은 나이가 많은 노인이 치매 발병 시점이 빠름을 알 수 있다(Table 3). 또 다른 예측 모형의 판별 측도(discrimination measure)로 자주 사용되는 c-index 를 확장한 시간 가변 C-index  $C(t)$ 도 적용해 보았다.

$$\widehat{C}(t) = \frac{\sum_{i=1}^n \sum_{j=1}^n I(T_i < T_j) N_i(t) I(M_n(t, Z_i) > M_n(t, Z_j))}{\sum_{i=1}^n \sum_{j=1}^n I(T_i < T_j) N_i(t)}, \quad N_i(t) = I(T_i \leq t) \delta_i,$$

여기서  $C(t)$  는  $t$  시점에서 랜덤하게 선택된 두 사람의 마커값의 순위와 사건 발생 시점의 순위의 일치쌍(concordance)의 비율을 의미한다 (Harrell 등, 1996). 따라서  $AUC(t)$ 와 마찬가지로 1에 가까울수록 높은 마커 값을 가진 환자가 더 빠른 발생 시점을 가짐을 의미하게 된다. 중도 절단 독립성이 위배되는 경우를 반영하기 위해  $C^{\text{IPCW}}(t)$  와  $C^{\text{CIPCW}}(t)$  가 각각 다음과 같이 유도되었다 (Gerds 등, 2013).

$$\widehat{C}^{\text{IPCW}}(t) = \frac{\sum_{i=1}^n \sum_{j=1}^n I(T_i < T_j) N_i(t) I(M_n(t, Z_i) > M_n(t, Z_j)) \frac{1}{S_C(T_i)S_C(T_j)}}{\sum_{i=1}^n \sum_{j=1}^n I(T_i < T_j) N_i(t) \frac{1}{S_C(T_i)S_C(T_j)}},$$

$$\widehat{C}^{\text{CIPCW}}(t) = \frac{\sum_{i=1}^n \sum_{j=1}^n I(T_i < T_j) N_i(t) I(M_n(t, Z_i) > M_n(t, Z_j)) \frac{1}{S_C(T_i|X_i)S_C(T_j|X_j)}}{\sum_{i=1}^n \sum_{j=1}^n I(T_i < T_j) N_i(t) \frac{1}{S_C(T_i|X_i)S_C(T_j|X_j)}}.$$

Table 4 은 두 마커 값을 이용하여  $t = 5, 10, 15, 18$  시점에서 구한 여러가지  $AUC(t)$  값과  $C(t)$ 값을 보여준다. 위 결과를 통해, 여러가지 인지 테스트값을 포함한 마커값(Marker 2)이 MMRT 변수만을 사용하는 경우보다 조금 더 큰 값을 보여준다. AUC 값들간의 비교를 고려하였을 때,  $AUC^{\text{IPCW}}(t)$ 와  $AUC^{\text{CIPCW}}(t)$ 가 다른 값들에

Table 4: Application of several AUCs using two markers to Paquid dataset

t	Marker 1: $M_1 = -0.191\text{MMRT} + 0.116\text{Age}$							
	$\text{AUC}^{\text{KM}}(t)$	$\text{AUC}^{\text{NNE}}(t)$	$\text{AUC}^{\text{CD}}(t)$	$\text{AUC}^{\text{IPCW}}(t)$	$\text{AUC}^{\text{CIPCW}}(t)$	$\hat{C}(t)$	$\hat{C}^{\text{IPCW}}(t)$	$\hat{C}^{\text{CIPCW}}(t)$
5	0.862	0.823	0.829	0.885	0.879	0.713	0.688	0.684
10	0.816	0.820	0.757	0.842	0.830	0.685	0.651	0.644
15	0.808	0.820	0.802	0.830	0.827	0.697	0.678	0.672
18	0.813	0.823	0.800	0.831	0.836	0.689	0.690	0.690
Time	Marker 2: $M_2 = -0.135\text{MMRT} - 0.10\text{BVRT} - 0.03\text{IST} - 0.278\text{Score of physical dependency} + 0.117\text{Age}$							
	$\text{AUC}^{\text{KM}}(t)$	$\text{AUC}^{\text{NNE}}(t)$	$\text{AUC}^{\text{CD}}(t)$	$\text{AUC}^{\text{IPCW}}(t)$	$\text{AUC}^{\text{CIPCW}}(t)$	$\hat{C}(t)$	$\hat{C}^{\text{IPCW}}(t)$	$\hat{C}^{\text{CIPCW}}(t)$
5	0.870	0.839	0.856	0.902	0.897	0.741	0.723	0.720
10	0.801	0.835	0.748	0.845	0.835	0.712	0.688	0.684
15	0.813	0.842	0.813	0.842	0.845	0.732	0.714	0.709
18	0.827	0.848	0.816	0.846	0.847	0.727	0.714	0.711

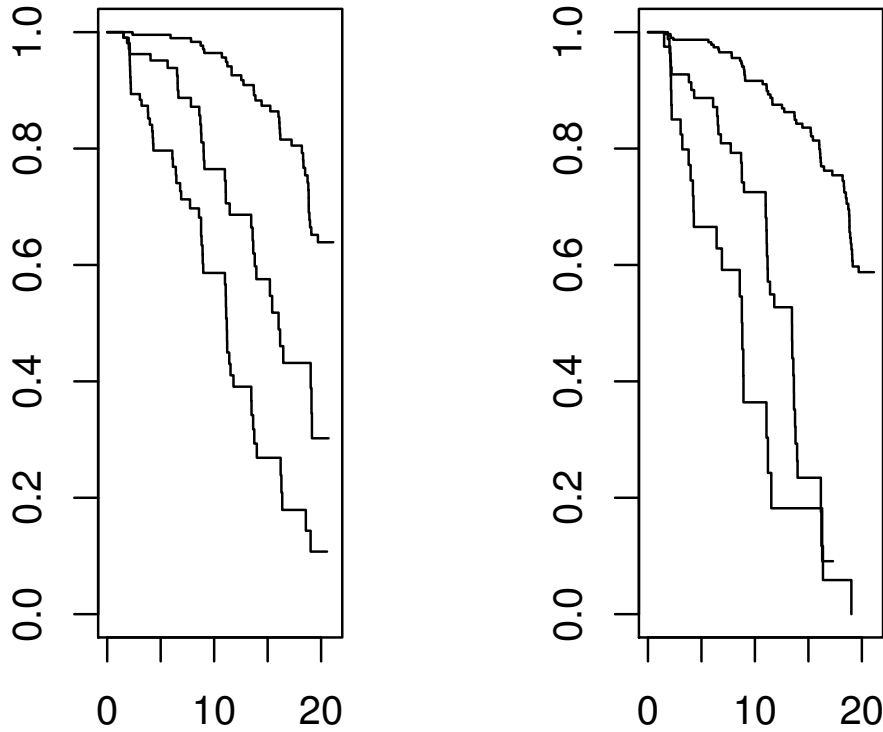


Figure 1: Survival curves of three group according to markers value: left (Marker 1),right (Marker 2).

비해 약간 높은 값을 가진다. 하지만 시간이 경과함에 따라 이들값간에는 큰 차이를 보이지 않았다. Figure 1은 마커값의 25, 50, 75 백분위수를 경계값으로 하여 487명을 3개의 그룹으로 나눈 후 각 그룹의 생존 함수 추정량을 보여준다. 왼쪽은 Marker 1을 오른쪽은 Marker 2의 값을 이용한 결과이다. 이 그림을 통해 Marker 2에 기반한 그룹간의 생존 함수에 더 유의한 차이를 보여주었다. 로그 순위 검정을 적용한 결과도 Marker 1에 대해서  $\hat{p} = 0.04$ 이고 Marker 2에 대해서는  $\hat{p} = 0.01$ 이었다. 세가지 c-index 값은 AUC 보다는 더 작은 값을

보여주었으며 세 가지  $c$ -index 값들은 큰 차이가 없었다.

## 6. 맺음말

본 논문에서 생존 분석 모형에 기반한 마커의 예측력을 평가하기 위해 널리 사용되는 시간 가변 ROC 곡선과 AUC에 대한 연구를 정리하고 실제 자료에 적용하였다. 특히 시간을 통해 관측되는 사건 발생 시간의 특성을 반영하기 위해 다양한 방법론이 비교되었다. 예를 들어 중도 절단 여부를 어떻게 처리하는가에 따라 사례군과 대조군의 처리가 달리 정의되었으며 가중치의 적용을 통해 표본 편이를 보정하고자 하였다. 본 연구에서 제시된 다양한 방법론은 R 패키지로 구축되어 있으며 각 자료의 특성에 따라 적절한 방법론이 적용될 수 있을 것이다. 본 연구의 모의 분석 결과에 의하면 중도 절단이 마커값과 관련되어 있는 경우에는 마커값을 고려한 중도 절단 생존 함수의 추정량이 가중치로 사용되는 것이 가장 바람직할 것이다. 이와 비슷한 결과와 적용 과정은 Blanche 등 (2013)에 의해 시행되었다. 본 연구에서는 그들과 비슷한 주제를 다루면서 좀 더 다양한 방법론을 적용하고자 하였으며 경쟁위험모형에 대한 추정치도 고려하였다. 본 연구의 확장 분야로는 경시적 마커값에 대한 연구이다. 관련 연구는 Zheng과 Heagerty (2004, 2007)와 Rizopoulos (2011)에 의해 각각 스무딩 기법의 확장과 결합 모형(joint model)의 확장이 제안되었다. 치매 자료 또한 연구 시작 시점부터 반복 측정된 인지 테스트 점수들을 이용하여 누적된 정보에 근거한 시간 가변 마커를 이용하게 된다면 더 높은 예측력을 가질 수 있을 것이다.

## References

- Akritis MG (1994). Nearest neighbor estimation of a bivariate distribution under random censoring, *Annals of Statistics*, **22**, 1299—1327.
- Blanche P, Dartigues JF, and Jacqmin-Gadda H (2013). Review and comparison of ROC curve estimators for a time-dependent outcome with marker-dependent censoring, *Biometrical Journal*, **55**, 687—704.
- Blanche P, Dartigues JF, and Jacqmin-Gadda H (2013b). Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks, *Statistics in Medicine*, **32**, 5381—5397.
- Chambless LE and Diao G (2006). Estimation of time-dependent area under the ROC curve for long-term risk prediction, *Statistics in Medicine*, **25**, 3474—3486.
- Fine JP and Gray RJ (1999). A proportional hazards model for the subdistribution of a competing risk, *Journal of the American Statistical Association*, **94**, 496—509.
- Foucher Y, Giral M, Soullillou JP, and Daures JP (2010). Time-dependent ROC analysis for a three-class prognostic with application to kidney transplantation, *Statistics in Medicine*, **29**, 3079—3087.
- Gerds TA, Kattan MW, Schumacher M, and Yu C (2013). Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring, *Statistics in Medicine*, **32**, 2173—2184.
- Harrell F, Kerry L, and Mark D (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, *Statistics in Medicine*, **15**, 361—387.
- Heagerty PJ, Lumley T, and Pepe MS (2000). Time-dependent ROC curves for censored survival data and a diagnostic marker, *Biometrics*, **56**, 337—344.
- Heagerty PJ and Zheng Y (2005). Survival model predictive accuracy and ROC curves, *Biometrics*, **61**, 92—105.
- Hung H and Chiang CT (2010). Estimation methods for time-dependent AUC models with survival data, *The Canadian Journal of Statistics*, **38**, 8—26.

- Letenneur L, Commenges D, Dartigues JF, and Barberger-Gateau P (1994). Incidence of dementia and alzheimer's disease in elderly community residents of southwestern France, *International Journal of Epidemiology*, **23**, 1256–1261.
- Rizopoulos D (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data, *Biometrics*, **67**, 819–829.
- Uno H, Cai TX, Tian L, and Wei LJ (2007). Evaluating prediction rules for t-year survivors with censored regression models, *Journal of American Statistical Association*, **102**, 527–537.
- Zheng Y and Heagerty PJ (2004). Semiparametric estimation of time-dependent ROC curves for longitudinal marker data, *Biostatistics*, **5**, 615–632.
- Zheng Y and Heagerty PJ (2007). Prospective accuracy for longitudinal markers, *Biometrics*, **63**, 332–341.

*Received September 1, 2021; Revised October 13, 2021; Accepted October 19, 2021*

# 생존 분석 자료에서 적용되는 시간 가변 ROC 분석에 대한 리뷰

김양진<sup>1,a</sup>

<sup>a</sup>숙명여자 대학교, 통계학과

---

## 요 약

Receiver operating characteristic (ROC) 곡선은 이항 반응 자료에 대한 마커의 분류 예측력을 측정하기 위해 널리 적용되어왔으며 최근에는 생존 분석에서도 매우 중요한 역할을 하고 있다. 여러 가지 유형의 중도 절단과 원인 불명 등 다양한 종류의 결측 자료를 포함한 생존 자료 분석에서 마커의 사건 발생 여부에 대한 예측력을 판단하기 위해 기존의 통계량을 확장하였다. 생존 분석 자료는 각 시점에서의 사건 발생 여부로 이해할 수 있으며, 따라서 시점마다 ROC 곡선과 AUC를 구할 수 있다. 본 논문에서는 우중도 절단과 경쟁 위험 모형하에서 사용되는 다양한 방법론과 관련 R 패키지를 소개하고 각 방법의 특성을 설명하고 비교하였으며 이를 검토하기 위해 간단한 모의실험을 시행하였다. 또한, 프랑스에서 수집된 치매 자료의 마커 분석을 시행하였다.

주요용어: AUC, 역확률 중도 절단 가중, 예측, 생존 분석, 시간 가변 ROC

---

본 연구는 연구재단(NRF-2020R1A2C1A01100755)의 지원을 받았습니다.

<sup>1</sup>(04310) 서울 특별시 용산구 청파로 47길, 숙명여자대학교, 통계학과 E-mail: yjin@sookmyung.ac.kr