

## 선택편향이 존재할 때, 수정 층화우도비를 이용한 최적절사점의 결정

김후남<sup>1)</sup> 박용규<sup>2)</sup>

### 요 약

진단검사에서 민감도와 특이도가 선택편향에 영향을 받을 때, 진단검사의 각 층에서의 민감도와 1-특이도의 비로 얻어지는 층화우도비도 편이가 존재하게 된다. 따라서 편이가 있는 층화우도비로 찾아진 최적절사점도 잘못된 값이 된다. 본 연구에서는 Begg와 Greenes (1983) 에 의해 제안된 수정 민감도와 특이도를 층화우도비에 적용하여, 선택편향이 수정되는 최적절사점을 찾아보았다. 그리고 선택편향이 최적절사점에 미치는 영향을 선택편향 수정인자를 통해 설명하였다.

주요용어: 선택편향, 층화우도비, 최적절사점.

### 1. 서론

검증 절차 (verification procedure)란 진단검사의 결과가 옳은지를 확인하기 위해, 질병 상태에 대한 참값 (gold standard)을 얻는 과정을 말한다. 그러나 진단검사 값이나 그 외 증상의 유무 등을 고려하여, 진단검사를 받는 모든 사람들이 아닌, 일부만 검증 절차를 수행하게 되는 경우가 많으며, 이로 인해 발생하는 오류를 선택편향 (verification bias)이라고 한다 (Begg와 Greenes, 1983; Gray 등, 1984; Greenes와 Begg, 1985; Diamond, 1986). 이러한 선택편향은 진단검사의 유용성을 판단하는 지표인 민감도 (sensitivity)와 특이도 (specificity)의 추정에 편이를 나타나게 한다. 예를 들어, 진단검사 값이 높을수록 검증 절차를 수행할 가능성이 높다면, 최종 분석자료에는 진단검사의 절사점 (cutoff point) 이상의 값을 갖는 사람들이 많아지므로, 민감도는 참값보다 높게 추정되고, 특이도는 참값보다 낮게 추정된다.

연속형 진단검사 결과를 값의 크기에 따라 여러 개의 층으로 나눈 후, 각 층에서의 민감도와 1-특이도의 비로 표현되는 층화우도비 (stratum-specific likelihood ratios, SSLR)는, 진단검사의 최적절사점 결정에 사용된다 (Peirce와 Cornell, 1993; Lee, 2005). 만약 선택편향이 존재한다면, 각 층에서 구한 민감도와 특이도에 편이가 발생하여 층화우도비에 영향을 미치고, 이를 이용하여 얻은 최적절사점 또한 잘못된 값이 된다.

1) (137-701) 서울시 서초구 반포동 505, 가톨릭대학교 대학원 의학통계학과, 석사과정  
E-mail: khnsky@dreamwiz.com

2) (교신저자)(137-701) 서울시 서초구 반포동 505, 가톨릭대학교 대학원 의학통계학과, 교수  
E-mail: ygpark@catholic.ac.kr

본 연구에서는 선택편향 수정인자를 통해 편의발생 요인과 그것이 수정 총화우도비에 미치는 영향을 설명하고, 수정 총화우도비를 이용하여 최적절사점을 찾는 방법을 제시한다. 2절에서는 선택편향이 민감도와 특이도에 미치는 영향을 구체적인 예를 통해 보이고, 3절에서는 선택편향을 수정한 총화우도비와 이를 이용하여 최적절사점을 찾는 방법과, 수정 총화우도비와 수정하지 않은 총화우도비간의 관계를 살펴본다. 4절에서는 선택편향의 정도에 변화를 준 여러 상황에서 총화우도비와 수정 총화우도비로 구한 최적절사점들간의 관계를 모의실험을 통해 살펴보고, 실제 예제자료에 수정 총화우도비를 적용하여 선택편향이 수정된 최적절사점을 찾아본다. 마지막으로 5절에서는 연구결과를 요약, 정리한다.

## 2. 기호, 정의 및 가정

### 2.1. 선택편향

표 2.1은 선택편향이 민감도와 특이도에 미치는 영향을 예시자료를 통해 나타낸 것이다. (a)는 환자 1000명과 정상인 1000명으로 구성된 전체 자료의 진단검사 결과 (민감도 = 특이도 = 90%)를 나타낸다. (b)는 진단검사의 양성반응자가 음성반응자보다, (c)는 진단검사의 음성반응자가 양성반응자보다 검증받을 비율이 높은 예로, 두 경우 모두 전체 자료 중 절반만 검증받는다고 가정하였다.

표 2.1: 선택편향의 예시자료

		(a) 전체 자료				
		진단검사 결과			민감도	
		양성	음성	계	/특이도	
질병상태	유	900	100	1000	0.90 (900/1000)	
	무	100	900	1000	0.90 (900/1000)	
	계	1000	1000	2000		

		(b) 검증 자료 i			(c) 검증 자료 ii								
		진단검사 결과			민감도		진단검사 결과		민감도				
		양성	음성	계	/특이도		양성	음성	계	/특이도			
질병상태	유	720	20	740	0.97 (720/740)		유	180	80	260	0.69 (180/260)		
	무	80	180	260	0.69 (180/260)		상태 무	20	720	740	0.97 (720/740)		
	계	800	200	1000			계	200	800	1000			
검증비율		0.8	0.2	0.5			검증비율		0.2	0.8	0.5		

또한, 실제 검증받기 전까지 대상자 각각의 질병상태는 알려지지 않은 정보이므로, 전체 대상자 중에서 검증될 사람의 선택은 질병상태와 무관하다는 가정하에, 검증된 자료에서 진단검사의 양성반응자 중 환자의 비율과 음성반응자 중 환자의 비율은, 전체 자료에서의 비율과 각각 동일하게 구성하였다.

(b)와 같이 검증비율이 음성보다 양성반응자에게서 높은 경우, 검증 자료에서의 민감도는 97%로 전체 자료의 민감도보다 높게 추정되고, 특이도는 69%로 낮게 추정된다. 반대로, (c)와 같이 양성보다 음성반응자에게서 검증비율이 높은 경우, 검증 자료에서의 민감도는 69%로 전체 자료의 민감도보다 낮게 추정되고 특이도는 97%로 높게 추정된다. 따라서 선택편향이 존재할 때, 검증 자료에서 추정되는 민감도와 특이도는 편의가 존재한다.

### 2.2. 자료의 형태와 기호의 정의

표 2.2는 진단검사의 결과, 질병의 유무와 검증여부에 따라 분류된 자료를 나타낸다. 여기서  $V$ 는 검증여부를 나타내며,  $V+$ 는 검증 절차를 받은 경우를,  $V-$ 는 받지 않은 경우를 말한다. 참의 질병상태를  $D$ 로 정의하며,  $D+$ 와  $D-$ 는 각각 질병이 있음과 없음을 나타낸다.

$R$ 은 진단검사 결과를 나타내는 연속형 변수이며, 값이 높을수록 질병이 있을 가능성이

표 2.2: 선택편향이 존재할 때 질병의 유무 진단을 위한 자료

절사점	진단검사	검증된 사람( $V+$ )		검증된	검증되지 않은	계
	결과	질병 유	질병 무	사람 전체	사람 전체	
	$R$	( $D+$ )	( $D-$ )	( $V+$ )	( $V-$ )	
첫 번째 절사점 →						
	$R_1$	$x_{111}$	$x_{121}$	$x_{1.1}$	$x_{1.2}$	$x_{1..}$
두 번째 절사점 →						
	$R_2$	$x_{211}$	$x_{221}$	$x_{2.1}$	$x_{2.2}$	$x_{2..}$
⋮						
	⋮	⋮	⋮	⋮	⋮	⋮
$i$ 번째 절사점 →						
	$R_i$	$x_{i11}$	$x_{i21}$	$x_{i.1}$	$x_{i.2}$	$x_{i..}$
⋮						
	⋮	⋮	⋮	⋮	⋮	⋮
$I$ 번째 절사점 →						
	$R_I$	$x_{I11}$	$x_{I21}$	$x_{I.1}$	$x_{I.2}$	$x_{I..}$
$I+1$ 번째 절사점 →						
	계	$x_{.11}$	$x_{.21}$	$x_{.1}$	$x_{.2}$	$x_{...}$

높다고 가정한다.  $R$ 의 최소값을 첫 번째 절사점, 최대값을 마지막 절사점으로 하고, 중간에 임의의  $I-1$ 개 값을 절사점으로 하여 증화하였을 때,  $R_i$ 는  $i$ 번째 층에 해당하는 진단검사 값을 나타낸다.

관측치  $x_{ijk}$ 는  $i$ 번째 층의 진단검사 결과에서, 질병상태가  $j$ (1:  $D+$ , 2:  $D-$ )이고, 검증 수행여부가  $k$ (1:  $V+$ , 2:  $V-$ )에 해당하는 사람수를 나타내고, .는 해당 층자에 대한 합을 표시한다. 이후 연구에서 다루게 될 확률들은 다음과 같다.

- $P(R_i)$  진단검사 결과가  $i$ 번째 층에 속할 확률,  
 $P(D)$  질병상태에 대한 확률.  $P(D+)$ : 환자의 확률,  $P(D-)$ : 정상인의 확률,  
 $P(D|R_i)$   $i$ 번째 층의 진단검사 결과가 주어졌을 때, 해당 질병상태에 대한 조건부확률.  $P(D+|R_i)$ :  $i$ 번째 층에서 환자의 확률,  $P(D-|R_i)$ :  $i$ 번째 층에서 정상인의 확률,  
 $P(V+|R_i)$   $i$ 번째 층의 진단검사 결과를 가진 사람이 검증 절차를 받을 조건부확률,  
 $P(R_i|D)$  질병상태의 참값 정보가 주어졌을 때,  $i$ 번째 층의 진단검사 결과에 대한 조건부확률.  $P(R_i|D+)$ : 환자 중 진단검사 값이  $R_i$ 에 속할 확률,  $P(R_i|D-)$ : 정상인 중 진단검사 값이  $R_i$ 에 속할 확률. 만약 진단검사 결과가  $R_i$ 와 같거나 클 때 양성( $R+$ )으로 분류하고,  $R_{i-1}$ 과 같거나 작을 때 음성( $R-$ )으로 분류한다면,  $P(R+|D+)$ : 질병이 있을 때 양성으로 옳게 진단할 확률 (민감도),  $P(R-|D-)$ : 질병이 없을 때 음성으로 옳게 진단할 확률 (특이도).

위에서 정의된 확률들을 검증된 자료에서는  $P(R_i|V+)$ ,  $P(D|V+)$ ,  $P(D|R_i, V+)$ ,  $P(R_i|D, V+)$ 로 각각 표현한다.

### 2.3. 가정

선택편향을 수정하기 위해서 질병상태  $D$ 와 검증 수행여부  $V$ 가 서로 독립 또는 조건부 독립임을 가정한다. 먼저,  $D$ 와  $V$ 의 독립 가정은, 검증되기 전까지는 질병상태에 대해 알 수 없으므로, 질병상태는 대상자의 검증 수행여부 결정에 영향을 미치지 않고, 단지 진단검사의 결과  $R$ 만 검증 수행여부 결정에 영향을 미치는 것을 말한다.  $D$ 와  $V$ 의 조건부 독립 가정은, 알려지지 않은 질병상태는 검증 수행여부의 결정에 직접적으로 영향을 주지 않고, 질병상태와 서로 연관되어 있는 진단검사 값을 통해서만 검증 수행여부 결정에 영향을 미치는 것을 말한다.

$D$ 와  $V$ 의 독립 또는 조건부 독립 가정이 만족된다면, 아래의 식 (2.1)이 성립된다.

$$P(V|R) = P(V|R, D). \quad (2.1)$$

또한, 진단검사 결과가 주어졌을 때, 질병상태의 분포가 전체 자료에서와 검증 자료에서 동일하게 되므로, 아래의 식 (2.2)도 성립된다.

$$P(D|R) = P(D|R, V). \tag{2.2}$$

### 3. 수정 총화우도비

$i$ 번째 층의 우도비  $SSLR_i$ 는 식 (3.1)과 같이, 환자 중 진단검사 값이  $R_i$ 에 속할 확률과 정상인 중 진단검사 값이  $R_i$ 에 속할 확률의 비, 즉  $i$ 번째 층에서의 민감도와 1-특이도의 비로 정의된다.

$$SSLR_i = \frac{P(R_i|D+)}{P(R_i|D-)}. \tag{3.1}$$

#### 3.1. 총화우도비

검증 자료의 진단검사 결과  $i$ 번째 층에서 선택편향을 수정하지 않은 민감도와 1-특이도,  $P(R_i|D, V+)$ 는 베이지 정리에 의해 식 (3.2)와 같다.

$$P(R_i|D, V+) = \frac{P(R_i|V+)P(D|R_i, V+)}{P(D|V+)}. \tag{3.2}$$

여기서  $P(D|V+) = \sum_{t=1}^I P(R_t|V+)P(D|R_t, V+)$ 이다. 따라서 식 (3.1)에 식 (3.2)의  $P(R_i|D+, V+)$ 와  $P(R_i|D-, V+)$ 를 대입하여 얻은 선택편향을 수정하지 않은  $SSLR_i$ 는 식 (3.3)과 같다.

$$\begin{aligned} SSLR_i &= \frac{P(R_i|D+, V+)}{P(R_i|D-, V+)} = \frac{P(D+|R_i, V+) \cdot P(D-|V+)}{P(D-|R_i, V+) \cdot P(D+|V+)} \\ &= \frac{P(D+|R_i, V+) \left\{ \sum_{t=1}^I P(R_t|V+)P(D-|R_t, V+) \right\}}{P(D-|R_i, V+) \left\{ \sum_{t=1}^I P(R_t|V+)P(D+|R_t, V+) \right\}}. \end{aligned} \tag{3.3}$$

이때,  $P(R_i|D_j, V+)$ 의 최대우도추정량은  $\hat{P}(R_i|D_j, V+) = x_{ij1}/x_{.j1}$ 이므로,  $SSLR_i$ 의 최대우도추정량은 식 (3.4)와 같다.

$$\hat{SSLR}_i = \frac{\frac{x_{i11}}{x_{.11}}}{\frac{x_{i21}}{x_{.21}}}. \tag{3.4}$$

$\text{Var}(\ln SSLR_i)$ 의 최대우도추정량은 식 (3.5)와 같으며 (Simel 등, 1991),

$$\hat{\text{Var}}(\ln \hat{SSLR}_i) = \frac{1}{x_{i11}} - \frac{1}{x_{.11}} + \frac{1}{x_{i21}} - \frac{1}{x_{.21}}, \tag{3.5}$$

$SSLR_i$ 의 95% 신뢰구간은 다음과 같다.

$$\left( e^{\ln \hat{SSLR}_i - 1.96 \sqrt{\hat{\text{Var}}(\ln \hat{SSLR}_i)}}, e^{\ln \hat{SSLR}_i + 1.96 \sqrt{\hat{\text{Var}}(\ln \hat{SSLR}_i)}} \right).$$

### 3.2. 수정 증화우도비

진단검사 결과의  $i$ 번째 층에서 선택편향을 수정한 민감도와 1-특이도,  $P_a(R_i|D)$ 는 베이즈 정리와 식 (2.2)에 의해 다음과 같다.

$$P_a(R_i|D) = \frac{P(R_i)P(D|R_i)}{P(D)} = \frac{P(R_i)P(D|R_i, V+)}{P(D)}. \quad (3.6)$$

여기서  $P(D) = \sum_{t=1}^I P(R_t)P(D|R_t, V+)$ 이다. 따라서 식 (3.1)에 식 (3.6)의  $P_a(R_i|D+)$ 와  $P_a(R_i|D-)$ 를 대입하여 얻은 선택편향을 수정한  $SSLR_{a,i}$ 는 식 (3.7)과 같다.

$$\begin{aligned} SSLR_{a,i} &= \frac{P_a(R_i|D+)}{P_a(R_i|D-)} = \frac{P(D+|R_i, V+) \cdot P(D-)}{P(D-|R_i, V+) \cdot P(D+)} \\ &= \frac{P(D+|R_i, V+) \left\{ \sum_{t=1}^I P(R_t)P(D-|R_t, V+) \right\}}{P(D-|R_i, V+) \left\{ \sum_{t=1}^I P(R_t)P(D+|R_t, V+) \right\}}. \end{aligned} \quad (3.7)$$

이때,  $P(R_i)$ 와  $P(D_j|R_i, V+)$ 의 최대우도추정량  $\hat{P}(R_i) = x_{i..}/x_{...}$ 와  $\hat{P}(D_j|R_i, V+) = x_{ij1}/x_{i.1}$  ( $j = 1, 2$ )를 대입한  $SSLR_{a,i}$ 의 최대우도추정량은 식 (3.8)과 같다.

$$\hat{SSLR}_{a,i} = \frac{x_{i11} \left( \sum_{t=1}^I \frac{x_{t..} x_{t21}}{x_{...} x_{t.1}} \right)}{x_{i21} \left( \sum_{t=1}^I \frac{x_{t..} x_{t11}}{x_{...} x_{t.1}} \right)}. \quad (3.8)$$

$\text{Var}(\ln SSLR_{a,i})$ 의 최대우도추정량과 유도과정은 부록에 제시하였으며,  $SSLR_{a,i}$ 의 95% 신뢰구간은 다음과 같다.

$$\left( e^{\ln \hat{SSLR}_{a,i} - 1.96 \sqrt{\hat{\text{Var}}(\ln \hat{SSLR}_{a,i})}}, e^{\ln \hat{SSLR}_{a,i} + 1.96 \sqrt{\hat{\text{Var}}(\ln \hat{SSLR}_{a,i})}} \right).$$

### 3.3. 수정 증화우도비와 증화우도비의 관계

$SSLR_{a,i}$ 와  $SSLR_i$ 의 관계를

$$SSLR_{a,i} = z \times SSLR_i$$

와 같이 표현하면, 수정인자  $z$ 는 식 (3.3)과 (3.7)로부터  $P(R_i)$ 와  $P(R_i|V+)$ 에 관한 아래의 식으로 표현된다.

$$\begin{aligned} z &= \frac{P(D+|V+) \cdot P(D-)}{P(D-|V+) \cdot P(D+)} \\ &= \frac{\sum_{t=1}^I P(R_t|D-, V+) \frac{P(R_t)}{P(R_t|V+)}}{\sum_{t=1}^I P(R_t|D+, V+) \frac{P(R_t)}{P(R_t|V+)}}. \end{aligned}$$

여기서,  $P(R_i|V+) \neq 0$ 이고  $P(R_i) \neq 0$ 이다. 따라서 검증절차가 랜덤하게 이루어져, 전체 자료에서의  $R_i$ 분포와 검증 자료에서의  $R_i$ 의 분포가 동일한,  $P(R_i)/P(R_i|V+) = 1(z = 1)$ 인 경우에만  $SSLR_{a,i}$ 와  $SSLR_i$ 는 같게 된다.

진단검사 결과  $R$ 이 하나의 절사점에 의해  $R+$ 와  $R-$ 의 두 층으로 나누어질 때,  $P(R)$ 과  $P(R|V+)$ 의 차이에 따른  $z$ 의 변화는 아래 그림 3.1과 같다. 이때,  $P(R+|D+, V+) = 0.8$ ,  $P(R+|D-, V+) = 0.2$ 로 고정시켰다. 검증 비율이 음성보다 양성반응자에게서 높은 경우, 검증 자료에서의  $R+$ 비율이 전체 자료보다 높게 되고, 이때  $z$ 는 1보다 커진다 ( $SSLR_{a,i} > SSLR_i$ ). 반대로 검증 비율이 음성반응자에게서 높은 경우, 검증 자료에서의  $R+$ 비율이 전체 자료보다 낮게 되고  $z$ 는 1보다 작아진다 ( $SSLR_{a,i} < SSLR_i$ ). 표 3.1에는 그림 3.1에서 수정인자  $z$ 가 1을 기준으로 변하는 경우를 제시하였다.

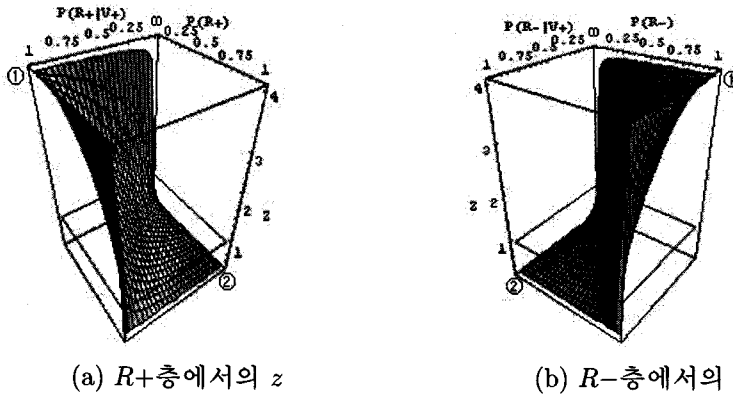


그림 3.1:  $P(R)$ 과  $P(R|V+)$ 의 차이에 따른 수정인자  $z$ 의 변화

표 3.1:  $P(R)$ 과  $P(R|V+)$ 의 차이에 따른 수정인자  $z$ 의 변화

	R+ 층	R- 층
① $z$ 가 1보다 커짐 ( $SSLR_{a,i} > SSLR_i$ )	i. $P(R+) < P(R+ V+)$ 인 경우 ii. $P(R+ V+) \simeq 1$ 인 경우 iii. $P(R+) \simeq 0$ 인 경우	i. $P(R-) > P(R- V-)$ 인 경우 ii. $P(R- V+) \simeq 0$ 인 경우 iii. $P(R-) \simeq 1$ 인 경우
② $z$ 가 1보다 작아짐 ( $SSLR_{a,i} < SSLR_i$ )	i. $P(R+) > P(R+ V+)$ 인 경우 ii. $P(R+ V+) \simeq 0$ 인 경우 iii. $P(R+) \simeq 1$ 인 경우	i. $P(R-) < P(R- V-)$ 인 경우 ii. $P(R- V+) \simeq 1$ 인 경우 iii. $P(R-) \simeq 0$ 인 경우

### 3.4. 수정 증화우도비를 이용한 최적절사점 결정

$SSLR_i$ 를 이용한 최적절사점을 찾는 방법은 Peirce와 Cornell (1993), Lee (2005)에 의해 제시되었으며,  $SSLR_{a,i}$ 도 동일한 방법으로 최적절사점을 찾을 수 있다.  $SSLR_{a,i}$ 는 진단검사 결과  $R$ 의  $i$ 번째 층에서의 ROC 곡선 기울기를 의미하며, 다음과 같은 방법에 의해 최적절사점을 얻을 수 있다.

- i.  $SSLR_{a,i}$ 가  $SSLR_{a,i+1}$ 의 95% 신뢰구간에 포함되거나,  $SSLR_{a,i+1}$ 이  $SSLR_{a,i}$ 의 95% 신뢰구간에 포함되면, 두 층을 하나로 병합한다.
- ii. 병합된 층과 인접한 층에 대해 더 이상 병합되지 않는 층들로 구성될 때까지  $i$ 의 과정을 반복한다.
- iii.  $i$ 과  $ii$ 로부터 얻어지는 더 이상 병합되지 않는 최적의 층에서  $SSLR_{a,i}$ 가 1미만에서 1이상으로 변하는 층의 진단검사 값의 시작점이 최적절사점이 된다.

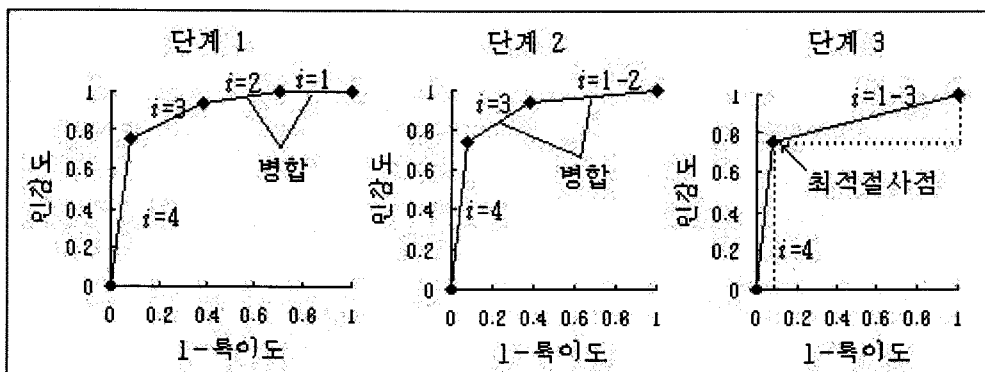


그림 3.2:  $SSLR_{a,i}$ 를 이용한 최적절사점 결정 과정

그림 3.2는  $SSLR_{a,i}$ 를 이용하여 최적절사점을 찾는 예이다. 4개의 층으로 증화된 진단 검사 결과  $R_i$ 에 대해 첫 번째 층  $SSLR_{a,1}$ 과 두 번째 층  $SSLR_{a,2}$ 가 서로의 신뢰구간에 포함되어, 두 층을 하나로 병합한다 (단계 1). 단계 1의 첫 번째와 두 번째 층의 병합으로 얻어진  $SSLR_{a,1-2}$ 와 인접한 세 번째 층의  $SSLR_{a,3}$ 이 또다시 서로의 신뢰구간에 포함되므로, 이를 하나로 병합한다 (단계 2). 단계 1, 2로부터 최종적으로 얻어진 최적의 층에서  $SSLR_{a,1-3}$ 은 1보다 작고,  $SSLR_{a,4}$ 는 1보다 크므로, 네 번째 층  $R_4$ 의 시작점이 최적절사점이 된다 (단계 3).

## 4. 모의실험과 예제

### 4.1. 모의실험

질병이 있는 집단과 없는 집단의 진단검사 결과  $R$ 은 각각  $N(35, 15^2)$ 와  $N(15, 15^2)$ 에서 각각 500개씩 난수를 생성하였다. 따라서 양성(+)과 음성(-)으로 분류되는 최적절사점은 25이



며, 민감도와 특이도 각각 74.75%,  $P(R+) = P(D+) = 0.5$ 이다. 선택편향을 위해 진단검사 결과의 두 층  $R-:(-\infty, 25)$ ,  $R+:[25, \infty)$ 에서 각각 다른 검증 비율로 랜덤하게 질병상태를 결측시켰으며, 이는 식 (2.1)의 가정을 만족한다. 또한  $P(V+) = P(R+) = 0.5$ 이므로,  $P(V+|R) = P(R|V+)$ 가 된다.

표 4.1의 상황 5는  $R-$ 와  $R+$ 에서 검증비율이 동일하여 선택편향이 존재하지 않는 경우이며, 상황 1과 상황 9는 각각  $R+$ 와  $R-$ 에서 검증비율이 높아 선택편향이 가장 심한 경우이다. 상황 1~9에서 진단검사 결과의 각 층  $R_i$ 의 간격을 1로 하여  $SSLR_{a,i}$ 와  $SSLR_i$ 로 최적절사점을 찾는 과정을 SAS를 통해 각각 10000번씩 반복하였으며, 이후 제시될 최적절사점과 민감도, 특이도는 그 값들의 평균값이다.

그림 4.1은 상황 1~9에서  $SSLR_{a,i}$ 와  $SSLR_i$ 로 얻어진 최적절사점을 나타낸다. 선택편향이 없는 상황 5에서  $SSLR_{a,i}$ 와  $SSLR_i$ 의 최적절사점은 전체 자료의 최적절사점 25와 유사하게 찾아졌다. 그러나  $R+$ 에서 검증비율이 높아질수록 (상황 4→상황 1)  $SSLR_i$ 로 얻어지는 최적절사점은 25보다 높아지고,  $R-$ 에서 검증비율이 높아질수록 (상황 6→상황 9) 25보다 낮아진다. 반면, 선택편향을 수정한  $SSLR_{a,i}$ 는 25에 가까운 최적절사점을 찾는다.

그림 4.2는 그림 4.1의  $SSLR_{a,i}$ 와  $SSLR_i$ 로 찾은 최적절사점을 전체 자료에 적용하여 얻은 민감도와 특이도이다.  $R+$ 에서 검증비율이 높은 경우 (상황 1~상황 4),  $SSLR_i$ 에 의한 최적절사점이 전체 자료의 최적절사점 25보다 높게 나타나므로, 민감도는 낮아지고 특이도는 높아진다. 반대로  $R-$ 에서 검증비율이 높은 경우 (상황 6~상황 9),  $SSLR_i$ 의 최적절사점이 25보다 낮게 나타나므로, 민감도는 높아지고 특이도는 낮아진다. 한편,  $SSLR_{a,i}$ 의 최적절사점으로 구한 민감도와 특이도는 전체 자료에서 얻은 값에 근접한다.

### 4.2. 예제자료 분석

새 간암 진단키트인 헤파чек (hepacheck) (Yoon 등, 2004) 자료에 선택편향을 수정한  $SSLR_{a,i}$ 를 적용하여 최적절사점을 찾아보았다. 이 자료는 정상인 188명, 간경변 환자 125명, 간암 환자 207명을 대상으로 한 것으로, 본 예제에서는 간경변 환자를 정상인으로

표 4.1: 각 상황별  $P(R)$ 과  $P(R|V+)$ 의 비율,  $P(V+) = P(R+) = 0.5, P(V+|R) = P(R|V+)$

	상황1		상황2		상황3		상황4		상황5	
	$R-$	$R+$	$R-$	$R+$	$R-$	$R+$	$R-$	$R+$	$R-$	$R+$
$P(V+ R)$	0.1	0.9	0.2	0.8	0.3	0.7	0.4	0.6	0.5	0.5
$\frac{P(R)}{P(R V+)}$	5.00	0.56	2.50	0.63	1.67	0.71	1.25	0.83	1.00	1.00
	상황6		상황7		상황8		상황9			
	$R-$	$R+$	$R-$	$R+$	$R-$	$R+$	$R-$	$R+$		
$P(V+ R)$	0.6	0.4	0.7	0.3	0.8	0.2	0.9	0.1		
$\frac{P(R)}{P(R V+)}$	0.83	1.25	0.71	1.67	0.63	2.50	0.56	5.00		

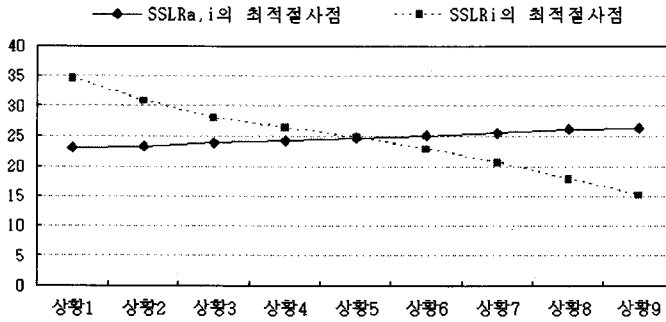


그림 4.1:  $SSLR_{a,i}$ 와  $SSLR_i$ 로 얻어진 최적결사점 비교

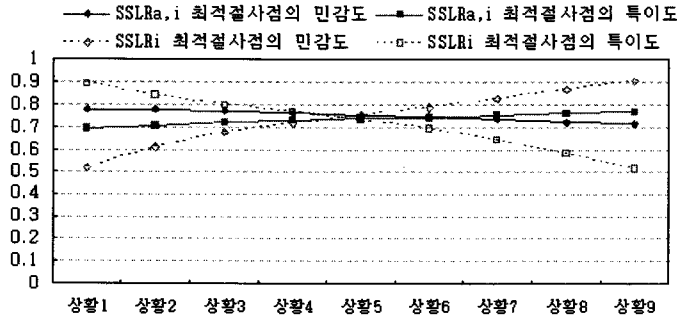


그림 4.2: 전체 자료에서 그림 4.1의  $SSLR_{a,i}$ 와  $SSLR_i$ 로 찾아진 최적결사점에 의한 민감도와 특이도

간주하여 정상인 313명, 간암 환자 207명의 자료로 이용하였다.

표 4.2는 전체 자료를 대상으로 진단검사 결과의 각 층  $R_i$ 의 간격을 2로 한 후,  $SSLR_i$ 로 병합과정을 거쳐 얻은 결과를 나타낸다. 이때  $SSLR_i$ 가 1미만에서 1이상으로 변하는 값, 10이 최적결사점이 되며, 이를 기준으로 얻어지는 전체 자료의 민감도는 98%, 특이도는 61%이다.

표 4.3은 선택편향이 존재하는 검증 자료에서  $SSLR_{a,i}$ 를 이용하여 최적결사점을 결정하는 과정을 보여준다. 이때,  $R_-:(-\infty, 10)$ 층에는  $P(V+|R_-) = 0.2$ 의 검증비율을,  $R_+:[10, \infty)$ 층에는  $P(V+|R_+) = 0.8$ 의 검증비율을 부여하였으며,  $P(R_+) = 0.62$ ,  $P(R_-) = 0.38$ 이고,  $P(R+|V+) = 0.87$ ,  $P(R-|V+) = 0.13$ 이다.

초기 층화된 A-1에서  $R_1$ 의  $SSLR_{a,1}$ 이 0이므로,  $R_1$ 과  $R_2$ 를 하나로 병합하고, 다음으로  $R_3$ 과  $R_4$ 의  $SSLR_{a,i}$ 가 서로의 신뢰구간에 포함되므로, 이 두 층을 하나로 병합한다. 또한,  $R_{12}$ 의  $SSLR_{a,12}$ 가  $\infty$ 이므로,  $R_{12}$ 와  $R_{13}$ 을 하나로 병합한다.

표 4.2: 전체 자료에서  $SSLR_i$ 로 구한 최적절사점

	최적의 층			계	민감도 /특이도
	2 ~	10 ~	22 ~		
간암 환자	5	117	85	207	0.98
정상인	191	106	16	313	0.61
효과책 전체	196	223	101	520	
$\hat{SSLR}_i$	0.04	1.67	8.03		
95% 신뢰구간 하한	0.02	1.37	4.89		
95% 신뢰구간 상한	0.09	2.03	13.20		

A-2는 이 병합된 층에서  $SSLR_{a,i}$ 를 구한 결과이다. 또한, A-2에서 전 단계에서 병합으로 얻어진 층  $R_{3-4}$ 와 인접한 층  $R_5$ 의  $SSLR_{a,i}$ 가 서로의 신뢰구간에 포함되므로, 이를 하나로 병합할 수 있다. 이러한 과정을 반복하여 더 이상 병합되지 않는 최적의 층 A-3이 얻어지고, 이때  $SSLR_{a,i}$  값이 1을 기준으로 변하는 10이 최적절사점이 되며, 이는 전체 자료에서 얻어진 최적절사점 10과 동일하다.

같은 검증 자료에  $SSLR_i$ 를 적용했을 때에는 표 4.4의 결과를 얻었다. 이때 최적절사점은 22로, 전체 자료에서 얻어진 최적절사점 10보다 큰 값이 결정되었다. 또한, 전체 자료에

표 4.3: 검증 자료에서  $SSLR_{a,i}$ 로 이용한 최적절사점 과정

A-1 층	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	$R_6$	$R_7$	$R_8$
	2 ~	8 ~	10 ~	12 ~	14 ~	16 ~	18 ~	20 ~
간암 환자	0	2	14	16	16	16	18	13
정상인	29	8	12	13	18	15	13	7
효과책 전체	143	53	29	37	48	42	37	30
$\hat{SSLR}_{a,i}$	0	0.35	1.64	1.73	1.25	1.50	1.94	2.61
95% 신뢰구간 하한	.	0.10	0.78	0.88	0.69	0.79	1.00	1.14
95% 신뢰구간 상한	.	1.25	3.42	3.38	2.27	2.82	3.79	5.96

층	$R_9$	$R_{10}$	$R_{11}$	$R_{12}$	$R_{13}$	$R_{14}$	$R_{15}$	$R_{16}$
	22 ~	24 ~	26 ~	28 ~	30 ~	32 ~	34 ~	36 ~
간암 환자	19	10	7	6	5	4	6	16
정상인	4	4	3	0	2	1	1	0
효과책 전체	27	15	11	6	7	6	8	21
$\hat{SSLR}_{a,i}$	6.67	3.51	3.28	$\infty$	3.51	5.62	8.43	$\infty$
95% 신뢰구간 하한	2.35	1.13	0.88	$\infty$	0.69	0.67	1.05	$\infty$
95% 신뢰구간 상한	18.94	10.92	12.25	$\infty$	17.99	47.10	67.38	$\infty$

A-2	층	$R_{1-2}$ 2 ~	$R_{3-4}$ 10 ~	$R_5$ 14 ~	$R_6$ 16 ~	$R_7$ 18 ~	$R_8$ 20 ~
	간암 환자	2	30	16	16	18	13
	정상인	37	25	18	15	13	7
	해파책 전체	196	66	48	42	37	30
	$\hat{SSLR}_{a,i}$	0.08	1.69	1.25	1.51	1.95	2.62
	95% 신뢰구간 하한	0.02	1.04	0.69	0.80	1.00	1.15
	95% 신뢰구간 상한	0.28	2.75	2.28	2.84	3.80	5.99

	층	$R_9$ 22 ~	$R_{10}$ 24 ~	$R_{11}$ 26 ~	$R_{12-13}$ 28 ~	$R_{14}$ 32 ~	$R_{15}$ 34 ~	$R_{16}$ 36 ~
	간암 환자	19	10	7	11	4	6	16
	정상인	4	4	3	2	1	1	0
	해파책 전체	27	15	11	13	6	8	21
	$\hat{SSLR}_{a,i}$	6.70	3.53	3.29	7.76	5.64	8.47	$\infty$
	95% 신뢰구간 하한	2.36	1.13	0.88	1.73	0.67	1.06	$\infty$
	95% 신뢰구간 상한	19.03	10.97	12.30	34.79	47.32	67.69	$\infty$

A-3	층	$R_{1-2}$ 2 ~	$R_{3-8}$ 10 ~	$R_{9-16}$ 22 ~
	간암 환자	2	93	73
	정상인	37	78	15
	해파책 전체	196	223	101
	$\hat{SSLR}_{a,i}$	0.08	1.69	6.90
	95% 신뢰구간 하한	0.02	1.33	4.09
	95% 신뢰구간 상한	0.28	2.15	11.64

표 4.4: 검증 자료에서  $SSLR_i$ 로 구한 최적절사점

층	2 ~	10 ~	22 ~
간암 환자	2	93	73
정상인	37	78	15
해파책 전체	39	171	88
$\hat{SSLR}_i$	0.04	0.92	3.77
95% 신뢰구간 하한	0.01	0.76	2.29
95% 신뢰구간 상한	0.15	1.12	6.20

서  $SSLR_i$ 로 찾아진 최적절사점 22에 의한 민감도는 41%로 낮고, 특이도는 95%로 높게 얻어졌다 (표 4.5).

표 4.5: 전체 자료에서  $SSLR_i$ 의 최적절사점 22에 의한 민감도와 특이도

	효과척 (R)			민감도 / 특이도
	~ 21.9	22.0 ~	계	
간암 환자	122	85	207	0.41
정상인	297	16	313	0.95
계	419	101	520	

### 5. 결론

선택편향이 존재할 때, 총화우도비로 얻어지는 최적절사점은 편의가 발생된다. 본 연구에서는 총화우도비로 최적절사점을 결정하는 방법을 수정 총화우도비에 적용하여 보았다. 그 결과, 선택편향이 심할수록 수정하지 않은 총화우도비로 찾아지는 최적절사점은 참값에서 멀어지며, 수정 총화우도비로 얻어지는 최적절사점이 수정하지 않은 총화우도비에 비해 참값에 가깝게 찾아짐을 밝혔다. 따라서 수정된 총화우도비로 찾아지는 최적절사점은 수정하지 않은 총화우도비보다 정확한 진단검사의 유용성 (민감도와 특이도)을 제공하게 된다.

본 연구의 문제점으로는 수정 총화우도비로 최적절사점을 결정하는 과정에서 각 층의 충분하지 않은 자료수가 병합 과정에 영향을 미쳐, 잘못된 최적절사점이 찾아질 수 있으므로, 각 층의 대상자가 충분히 확보될 수 있도록 층을 나뉘어야 한다.

추후 연구과제로 적절한 층의 수와 너비를 구하는 방법과, 진단검사 결과 외에 선택편향에 영향을 주는 또 다른 공변량이 있는 경우의 수정 총화우도비를 구하고, 이를 통해 최적절사점을 찾는 방법을 고려할 수 있을 것이다.

### 부록

$\ln \hat{SSLR}_{a,i}$ 의 분산은 아래 식 (A.1)과 같이  $\ln \hat{P}_a(R_i|D)$ 의 분산을 통해 얻을 수 있다.

$$\begin{aligned} \text{Var}(\ln \hat{SSLR}_{a,i}) &= \text{Var}[\ln \hat{P}_a(R_i|D+) - \ln \hat{P}_a(R_i|D-)] \\ &= \text{Var}[\ln \hat{P}_a(R_i|D+)] + \text{Var}[\ln \hat{P}_a(R_i|D-)]. \end{aligned} \quad (A.1)$$

이 때, 선택편향이 수정된  $P_a(R_i|D)$ 는  $R_i$ 가 서로 배반이므로 다음과 같이 표현될 수 있다.

$$P_a(R_i|D) = \frac{P(R_i)P(D|R_i, V+)}{P(R_i)P(D|R_i, V+) + P(\bar{R}_i)P(D|\bar{R}_i, V+)}. \quad (A.2)$$

여기서,  $\bar{R}_i$ 는  $R_i$ 의 여집합이다.

$\hat{P}(R_i)$ ,  $\hat{P}(D|R_i, V+)$ ,  $\hat{P}(D|\bar{R}_i, V+)$ 가 서로 독립이라는 점을 바탕으로, 식 (A.2)의 추정량에 대한 점근분산은 식 (A.3)과 같이 얻을 수 있다 (Begg과 Greenes, 1983).

$$\begin{aligned} \ln \frac{P_a(R_i|D)}{1 - P_a(R_i|D)} &= \ln \frac{P(R_i)}{1 - P(R_i)} + \ln P(D|R_i, V+) - \ln P(D|\bar{R}_i, V+) \\ \Rightarrow \text{Var} \left\{ \ln \frac{\hat{P}_a(R_i|D)}{1 - \hat{P}_a(R_i|D)} \right\} &= \text{Var} \left\{ \ln \frac{\hat{P}(R_i)}{1 - \hat{P}(R_i)} + \ln \hat{P}(D|R_i, V+) - \ln \hat{P}(D|\bar{R}_i, V+) \right\} \\ \text{좌변; Var} \left\{ \ln \frac{\hat{P}_a(R_i|D)}{1 - \hat{P}_a(R_i|D)} \right\} &= \left[ \frac{1}{P_a(R_i|D)\{1 - P_a(R_i|D)\}} \right]^2 \text{Var} \{ \hat{P}_a(R_i|D) \} \\ \text{우변; Var} \left\{ \ln \frac{\hat{P}(R_i)}{1 - \hat{P}(R_i)} + \ln \hat{P}(D|R_i, V+) - \ln \hat{P}(D|\bar{R}_i, V+) \right\} \\ &= \left[ \frac{1}{P(R_i)\{1 - P(R_i)\}} \right]^2 \text{Var} \{ \hat{P}(R_i) \} + \left\{ \frac{1}{P(D|R_i, V+)} \right\}^2 \text{Var} \{ \hat{P}(D|R_i, V+) \} \\ &\quad + \left\{ \frac{1}{P(D|\bar{R}_i, V+)} \right\}^2 \text{Var} \{ \hat{P}(D|\bar{R}_i, V+) \} \\ \text{Var} \{ \hat{P}_a(R_i|D) \} &= P_a(R_i|D)^2 \{1 - P_a(R_i|D)\}^2 \times \\ &\quad \left[ \frac{1}{P(R_i)^2 \{1 - P(R_i)\}^2} \text{Var} \{ \hat{P}(R_i) \} + \frac{1}{P(D|R_i, V+)^2} \text{Var} \{ \hat{P}(D|R_i, V+) \} \right. \\ &\quad \left. + \frac{1}{P(D|\bar{R}_i, V+)^2} \text{Var} \{ \hat{P}(D|\bar{R}_i, V+) \} \right]. \end{aligned} \tag{A.3}$$

한편, 델타 방법에 의해 식 (A.1)의  $\text{Var} \{ \ln \hat{P}_a(R_i|D) \} = \{1/P_a(R_i|D)\}^2 \text{Var} \{ \hat{P}_a(R_i|D) \}$ 가 되므로, 결국  $\text{Var} \{ \ln \hat{P}_a(R_i|D) \}$ 는 식 (A.4)와 같이 유도된다.

$$\begin{aligned} \text{Var} \{ \ln \hat{P}_a(R_i|D) \} &= \{1 - P_a(R_i|D)\}^2 \times \\ &\quad \left[ \frac{1}{P(R_i)^2 \{1 - P(R_i)\}^2} \text{Var} \{ \hat{P}(R_i) \} + \frac{1}{P(D|R_i, V+)^2} \text{Var} \{ \hat{P}(D|R_i, V+) \} \right. \\ &\quad \left. + \frac{1}{P(D|\bar{R}_i, V+)^2} \text{Var} \{ \hat{P}(D|\bar{R}_i, V+) \} \right]. \end{aligned} \tag{A.4}$$

이때,  $P(R_i)$ ,  $P(D_j|R_i, V+)$ 와  $P(D_j|\bar{R}_i, V+)$ 의 최대우도추정량  $\hat{P}(R_i) = x_{i..}/x_{...}$ ,  $\hat{P}(D_j|R_i, V+) = x_{ij1}/x_{i.1}$ 과  $\hat{P}(D_j|\bar{R}_i, V+) = (x_{.j1} - x_{ij1})/(x_{.1} - x_{i.1})$ 을 대입하여 얻은  $\text{Var} \{ \ln \hat{P}_a(R_i|D+) \}$ 와  $\text{Var} \{ \ln \hat{P}_a(R_i|D-) \}$ 의 최대우도추정량은 다음과 같다.

$$\begin{aligned} \hat{\text{Var}}\{\ln\hat{P}_a(R_i|D+)\} &= \left(1 - \frac{\frac{x_{i..} x_{i11}}{x_{...} x_{i.1}}}{\sum_{t=1}^I \frac{x_{t..} x_{t11}}{x_{...} x_{t.1}}}\right)^2 \\ &\times \left\{ \frac{x_{...}}{x_{i..}(x_{...} - x_{i..})} + \frac{x_{i.1} - x_{i11}}{x_{i.1}x_{i11}} + \frac{(x_{..1} - x_{i.1}) - (x_{.11} - x_{i11})}{(x_{..1} - x_{i.1})(x_{.11} - x_{i11})} \right\}. \end{aligned} \tag{A.5}$$

$$\begin{aligned} \hat{\text{Var}}\{\ln\hat{P}_a(R_i|D-)\} &= \left(1 - \frac{\frac{x_{i..} x_{i21}}{x_{...} x_{i.1}}}{\sum_{t=1}^I \frac{x_{t..} x_{t21}}{x_{...} x_{t.1}}}\right)^2 \\ &\times \left\{ \frac{x_{...}}{x_{i..}(x_{...} - x_{i..})} + \frac{x_{i.1} - x_{i21}}{x_{i.1}x_{i21}} + \frac{(x_{..1} - x_{i.1}) - (x_{.21} - x_{i21})}{(x_{..1} - x_{i.1})(x_{.21} - x_{i21})} \right\}. \end{aligned} \tag{A.6}$$

따라서  $\hat{\text{Var}}(\ln\hat{SSLR}_{a,i}) = (A.5) + (A.6)$  이 된다.

### 참고문헌

Begg, C. B. and Greenes, R. A. (1983). Assessment of diagnostic tests when disease verification is subject to selection bias, *Biometrics*, **39**, 207-215.

Diamond, G. A. (1986). Reverend Bayes' silent majority: An alternative factor affecting sensitivity and specificity of exercise electrocardiography, *The American Journal of Cardiology*, **57**, 1175-1180.

Gray, R., Begg, C. B. and Greenes, R. A. (1984). Construction of receiver operating characteristic curves when disease verification is subject to selection bias, *Medical Decision Making*, **4**, 151-164.

Greenes, R.A. and Begg, C. B. (1985). Assessment of diagnostic technologies: Methodology for unbiased estimation from samples of selectively verified patients, *Investigative Radiology*, **20**, 751-756.

Lee, J. H. (2005). Estimation of optimal cut-off point using VUS(Volume Under the surface) and SSLR(Stratum-Specific Likelihood Ratio), Thesis, Catholic University, Seoul.

Peirce, J. C. and Cornell, R. G. (1993). Integrating stratum-specific likelihood ratios with the analysis of ROC curves, *Medical Decision Making*, **13**, 141-151.

Simel, D. L., Samsa, G. P. and Matchar, D. B. (1991). Likelihood ratios with confidence: sample size estimation for diagnostic test studies, *Journal of Clinical Epidemiology*, **44**, 763-770.

Yoon, S. K., Lim, N. K., Ha, S.-A., Park, Y. G., Choi, J. Y., Chung, K. W., Sun, H. S., Choi, M. J., Chung, J., Wands, J. R. and Kim, J. W. (2004). The human cervical cancer oncogene protein is a biomarker for human hepatocellular carcinoma, *Cancer Research*, **64**, 5434-5441.

[ 2007년 4월 접수, 2007년 5월 채택 ]

## Determination of the Optimal Cutoff Point using Adjusted Stratum-Specific Likelihood Ratios when Disease Verification is subject to Verification Bias

Hu-Nam Kim<sup>1)</sup> Yong-Gyu Park<sup>2)</sup>

### ABSTRACT

Stratum-specific likelihood ratio, which is ratio of the sensitivity to 1-the specificity in each stratum of the test, could be biased if the sensitivity and specificity of diagnostic test are affected by verification bias. Therefore, the optimal cutoff point determined by biased stratum-specific likelihood ratios is incorrect. In this study, we derived adjusted stratum-specific likelihood ratios using the adjusted sensitivity and specificity, and obtained the adjusted optimal cutoff point. The influence of the verification bias on the optimal cutoff point was described through the relation between adjusted and unadjusted stratum-specific likelihood ratios.

*Keywords:* Verification bias, stratum-specific likelihood ratios, optimal cutoff point.

---

1) Graduate Student, Department of Biostatistics, The Catholic University of Korea, Seoul 137-701, Korea

E-mail: khnsky@dreamwiz.com

2) (Corresponding author) Professor, Department of Biostatistics, The Catholic University of Korea, Seoul 137-701, Korea

E-mail: ygpark@catholic.ac.kr