

Odds curve and optimal threshold

Chong Sun Hong^{1,a}, Tae Gyu Oh^a, Se Hyeon Oh^a

^aDepartment of Statistics, Sungkyunkwan University

Abstract

Various accuracy measures that can be explained on the odds curve are discussed, and an alternative accuracy measure, the maximum square, is proposed based on the characteristics of the odds curve. Thresholds corresponding to these accuracy measures are obtained by considering various probability distribution functions and an illustrative example. Their characteristics are discussed while comparing many kinds of statistics measuring thresholds. Therefore, we can conclude that optimal thresholds could be explored from the odds curve, similar to the ROC curve, and that the maximum square measure can be used as a good accuracy measure that can improve the performance of the binary classification model.

Keywords: accuracy, binary model, confusion matrix, threshold

1. 서론

의학통계와 신용평가 분야 등에서 이진 분류모형(binary classification model)의 성능(performance)을 탐색하는 유용한 시각적인 방법인 receiver operating characteristic (ROC) 곡선은 true positive rate, sensitivity (TPR)와 false positive rate, 1-specificity (FPR)의 비율 변화를 구현한다 (Altman과 Bland, 1994; Green과 Swets, 1966; Bamber, 1975; Egan과 Egan, 1975; Metz, 1978; Hanley와 McNeil, 1982; Swets, 1988; Centor, 1991; Zweig와 Campbell, 1993; Vuk와 Curk, 2006; Tasche, 2008). Pontius와 Si (2014)는 비율 대신에 빈도수를 이용하여 평행 사변형에서 구현하는 total operating characteristic (TOC) 곡선을 제안하였으며, Hong과 Lee (2018)는 세로와 가로축을 혼동행렬의 빈도수인 true positive (TP)와 false positive (FP)로 구현하여 직사각형에서의 곡선으로 표현하는 total receiver operating characteristic (TROC) 곡선을 제안하였다. ROC 곡선은 비율로 TOC와 TROC 곡선은 빈도수로 구현하지만, Hong 등 (2021)은 두 종류의 오즈(odds)인 FPR/TPR 과 $(1-TPR)/(1-FPR)$ 을 단위 면적인 정사각형에서의 곡선으로 구현하는 오즈 곡선(odds curve)이라고 제안하였다. 오즈 곡선을 분류 모형에 적용하고 활용하기 위하여 ‘오즈 곡선의 아래 면적(area under the odds curve)’과 ‘오즈 비 제곱(squared odds ratio)’ 측도를 개발하여, Joseph (2005)이 제안한 ROC 곡선에서의 13단계의 판단 기준과 유사하게 오즈 곡선의 판단 기준을 제안하였다.

최적분류점(optimal threshold, cut-off point)을 설정하는 정확도 측도(accuracy measures)들 중에서 ‘(0, 1)까지 최단 거리 기준(closest to (0,1) criterion; Perkins와 Schisterman, 2006)’, ‘Youden지수 (Youden, 1950)’, ‘최대 수직 거리(maximum vertical distance; Krzanowski와 Hand, 2009)’, ‘수정된(0, 1)까지 최단거리기준(amended closest-to-(0,1) criterion; Perkins와 Schisterman, 2006)’, ‘민감도와 특이도의 합(sum of sensitivity and specificity; Connell과 Koepsell, 1985)’, ‘진실율(true rate; Hong, 2009; Hong 등, 2010)’, ‘대칭점(symmetric point;

¹ Corresponding author: Department of Statistics, Sungkyunkwan University, 25-2, Sungkyunkwan-Ro, Jongno-Gu, Seoul 03063, Korea. E-mail: cshong@skku.edu

Moses 등, 1993; Pepe, 2003)', '전체 정확도(total accuracy; Lambert와 Lipkovich, 2008)', '정확도 면적(accuracy area; Brasil, 2010)' 등은 ROC 곡선으로 설명되며 구할 수 있다 (Hong과 Choi, 2009; Cho와 Hong, 2015). 본 연구에서는 오즈 곡선의 활용성을 높이기 위하여 ROC 곡선과 유사하게 오즈 곡선으로 설명되고, 오즈 곡선으로부터 최적분류점을 찾는 방법을 연구한다. 오즈 곡선에서 설명이 가능한 정확도 측도들의 존재를 탐색하여 오즈 곡선으로부터 최적분류점을 선정하는 정확도 측도들을 정리하고 오즈 곡선으로 설명한다. 그리고 오즈 곡선의 특성을 이용하여 대안적인 최적분류점을 선정하는 방법을 제안한다.

우선 2절에서는 잘 알려진 다양한 정확도 측도들이 오즈 곡선에서 어떻게 표현이 되는지를 살펴보고 오즈 곡선으로부터 발견할 수 있는 두 종류의 정확도 측도를 설명한다. 그리고 오즈 곡선의 성질을 바탕으로 오즈 곡선으로부터 최적분류점을 선정할 수 있는 대안적인 정확도 측도를 제안한다. 3절에서는 이진 분류모형의 다양한 확률분포함수를 고려하여 2절에서 논의한 두 종류의 정확도 측도와 제안한 정확도 측도로부터 최적분류점을 구하고, 분류점의 정확성과 정밀도 등을 측정하는 다양한 통계량값들을 비교하면서 성질과 특징을 토론한다. 4절에서는 실증예제를 표현한 오즈 곡선으로부터 정확도 측도에 대응하는 최적분류점들을 구하여 오분류율과 다양한 분류점 측정 통계량들을 비교하면서 설명한다. 마지막 5절에서는 ROC, TOC 그리고 TROC 곡선들에서 설명할 수 있듯이 오즈 곡선으로부터도 잘 알려진 정확도 측도들뿐만 아니라 본 연구에서 제안한 정확도 측도를 만족하는 최적분류점들을 쉽게 발견하고 설명할 수 있으며, 새로운 정확도 측도는 이미 알려진 측도들과 더불어 이진 분류모형의 성능을 향상시킬 수 있는 정확도 측도로 활용할 수 있는 결론을 유도한다.

2. 오즈 곡선과 정확도 측도

의학통계와 신용평가 분야 등에서 환자를 정상(non-disease/negative)상태와 질병(disease/positive)상태로 구분한 분류모형의 이항 결과(binary result)를 2×2 의 분할표인 혼동행렬(confusion matrix)로 나타낼 수 있다. 임의의 분류점에 대하여 true positive (TP)와 true negative (TN)은 각각 정확하게 분류된 질병집단과 정상집단의 수를 false positive (FP)와 false negative (FN)은 각각 질병을 정상으로 예측한 집단의 수와 정상을 질병으로 예측한 집단의 수를 의미한다. 표본의 크기 $p=TP+FN$ 인 질병 상태를 나타내는 확률변수 X 의 누적분포함수 $F_d(x)$ 와 크기 $q=FP+TN$ 인 정상 상태를 나타내는 분포함수 $F_n(x)$ 로 설정하면서 모든 x 에 대하여 $F_d(x) \geq F_n(x)$ 를 가정한다.

혼동행렬의 각 칸의 비율인 TPR과 FPR, false negative rate (FNR)와 true negative rate (TNR)는 다음과 같이 정의한다.

$$\begin{aligned} \text{TPR} &= \frac{\text{TP}}{p} = \hat{F}_d(x), & \text{FPR} &= \frac{\text{FP}}{q} = \hat{F}_n(x), \\ \text{TNR} &= \frac{\text{TN}}{q} = 1 - \hat{F}_n(x), & \text{FNR} &= \frac{\text{FN}}{p} = 1 - \hat{F}_d(x). \end{aligned}$$

Hong 등 (2021)이 제안한 오즈 곡선은 가로와 세로축의 좌표를 $(\text{FPR}/\text{TPR}, (1 - \text{TPR})/(1 - \text{FPR}))$ 또는 $(F_n(x)/F_d(x), (1 - F_d(x))/(1 - F_n(x)))$ 으로 설정하여 단위면적인 정사각형에서의 오즈의 변화를 곡선으로 구현한다. 우선, 오즈 곡선으로부터 설명할 수 있는 정확도 측도를 살펴보자.

2.1. 오즈 곡선과 최적분류점

최적분류점을 설정하는 아홉 종류의 정확도 측도인 '(0, 1)까지 최단 거리 기준', 'Youden지수', '최대 수직 거리', '수정된(0, 1)까지 최단 거리 기준', '민감도와 특이도의 합', '진실율', '대칭점', '전체 정확도', '정확도 면적'에 대하여, Hong 등 (2011)과 Yoo와 Hong (2011)은 누적분포함수 $F_d(x)$ 와 $F_n(x)$, 확률밀도함수 $f_d(x)$ 와

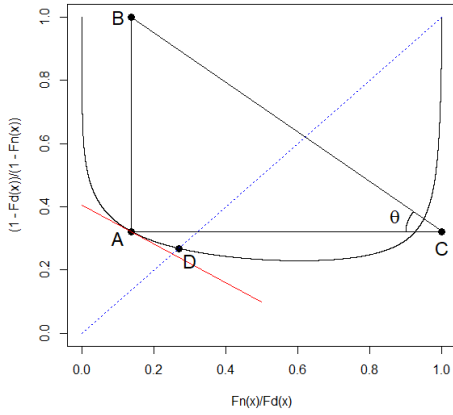


Figure 1: Threshold on odds curve.

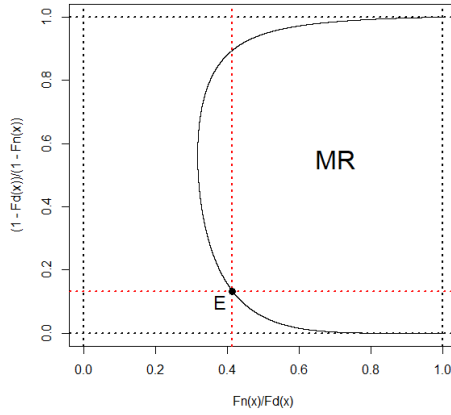


Figure 2: MR on odds curve.

$f_n(x)$ 에 대한 조건을 다섯 범주로 구분하였는데, 오즈 곡선에서 표현이 가능한 정확도 측도는 다섯 범주 중에서 두 번째에 속하는 Youden지수 (J), 최대수직거리(MVD), 수정된(0, 1)까지 최단거리기준(AC), 민감도와 특이도의 합(SSS), 진실율(TR)과 세 번째 범주에 속하는 대칭점(SP) 측도이다. 두 번째 범주에 속하는 다섯 종류의 정확도 측도 J, SSS, AC, TR, MVD의 조건은 모두 동일하게 확률밀도함수들로 표현되므로 진실율 (TR)을 대표 정확도 측도라고 설정한다. 다섯 범주 중에서 오즈 곡선으로 설명이 가능한 두 번째와 세 번째 범주에 속하는 정확도 측도들에 대하여 살펴본다.

Lemma 1. 오즈 곡선의 임의의 점 $A = (F_n(u)/F_d(u), (1 - F_d(u))/(1 - F_n(u)))$ 의 기울기는 다음과 같다.

$$\frac{[f_n(u) - f_d(u) + f_d(u)F_n(u) - f_n(u)F_d(u)]F_d^2(u)}{[f_n(u)F_d(u) - f_d(u)F_n(u)](1 - F_n(u))^2}. \tag{2.1}$$

Proof: 오즈 곡선의 점 A를 지나는 접선의 기울기는,

$$\Delta \frac{1 - F_d(u)}{1 - F_n(u)} / \Delta \frac{F_n(u)}{F_d(u)} = \frac{[-f_d(u)(1 - F_n(u)) + f_n(u)(1 - F_d(u))]F_d^2(u)}{[f_n(u)F_d(u) - f_d(u)F_n(u)](1 - F_n(u))^2}$$

이며, 이를 정리하면 식 (2.1)과 같다. □

Lemma 2. 오즈 곡선의 임의의 점 $A = (F_n(u)/F_d(u), (1 - F_d(u))/(1 - F_n(u)))$ 에서 사각형의 윗변과 오른쪽 변까지의 점을 각각 $B = (F_n(u)/F_d(u), 1)$ 와 $C = (1, (1 - F_d(u))/(1 - F_n(u)))$ 로 설정하고, Figure 1에서의 세 개의 점 A, B, C로 구성된 직각 삼각형에서 각도 θ 에 대한 $\tan^2 \theta$ 는 정분류율의 비율(오즈)의 제곱과 같다.

$$\tan^2 \theta = \left[\frac{TPR}{TNR} \right]^2. \tag{2.2}$$

Proof: Figure 1에서 직각 삼각형에서 각도 θ 에 대한 $\tan^2 \theta$ 는 다음과 같다.

$$\tan^2 \theta = \left[\frac{(F_d(u) - F_n(u))/(1 - F_n(u))}{(F_d(u) - F_n(u))/F_d(u)} \right]^2 = \left[\frac{F_d(u)}{1 - F_n(u)} \right]^2 = \left[\frac{TPR}{TNR} \right]^2.$$

□

Theorem 1. 오즈 곡선의 임의의 점을 지나는 접선의 기울기가 Lemma 2에서 정의한 정분류율의 오즈인 TPR/TNR의 제곱의 음수와 일치하면, 정확도 측도 중에서 진실율(TR)를 만족하는 분류점을 얻는다(J, SSS, AC, MVD 측도 포함).

Proof: Lemma 1에서 얻은 오즈 곡선위의 점의 기울기 식 (2.1)과 Lemma 2에서의 식 (2.2)를 이용하여 다음과 같은 등식으로 설정하면,

$$\frac{[f_n(u) - f_d(u) + f_d(u)F_n(u) - f_n(u)F_d(u)] F_d^2(u)}{[f_n(u)F_d(u) - f_d(u)F_n(u)](1 - F_n(u))^2} = - \left[\frac{F_d(u)}{1 - F_n(u)} \right]^2$$

$$f_n(u) - f_d(u) + f_d(u)F_n(u) - f_n(u)F_d(u) = f_d(u)F_n(u) - f_n(u)F_d(u)$$

이며, $f_d(u)F_n(u) - f_n(u)F_d(u) \neq 0$ 이므로 $f_d(u) = f_n(u)$ 의 관계를 갖는다. 이것은 정확도 측도인 TR의 조건과 일치하므로 이와 같은 정확도 측도들의 분류점이다. □

기울기가 +1인 접선과 만나는 ROC 곡선의 점으로부터 TR을 만족하는 최적분류점을 발견하듯이, Figure 2에서 직각 삼각형의 점 B와 C를 연결하는 직선의 기울기는 $-\sqrt{\tan^2 \theta} = -(TPR/TNR)$ 이므로 오즈 곡선의 점 A에 대응하는 접선의 기울기와 $-\tan^2 \theta$ 가 일치하면, 오즈 곡선에서도 TR을 만족하는 최적분류점을 발견할 수 있다. Figure 2의 각도 θ 가 45도 ($\pi/4$)보다 작으면, 최적분류점 A를 지나는 접선의 기울기는 더욱 작게 나타나고, 직각 삼각형의 각도 θ 가 45도 ($\pi/4$)보다 크면, 점 A를 지나는 접선의 기울기는 크게 표현되어 최적분류점을 시각적으로 설명하고 발견할 수 있다.

Theorem 2. 기울기가 1이고 원점을 지나는 직선과 오즈 곡선의 교차점으로부터 정확도 측도 중의 하나인 대칭점(SP)을 만족하는 최적분류점을 얻는다.

Proof: 기울기가 1이고 원점을 지나는 직선과 오즈 곡선의 교차점은 Figure 1에서 점 D이며 다음을 만족한다.

$$\frac{F_n(u)}{F_d(u)} = \frac{1 - F_d(u)}{1 - F_n(u)}$$

이 식으로부터 $F_n(u)(1 - F_n(u)) = F_d(u)(1 - F_d(u))$ 의 관계를 유도할 수 있으며, ROC 함수의 일반적인 조건 $F_d(u) \geq F_n(u)$ 에서는 $F_n(u) = 1 - F_d(u)$ 을 만족한다. 따라서 이 조건으로부터는 SP를 만족하는 최적분류점이다. □

그러므로 정확도 측도 중에서 진실율(TR: J, SSS, AC, MVD 측도 포함)과 대칭점(SP) 측도를 만족하는 최적분류점들은 오즈 곡선으로부터 설명할 수 있다.

2.2. 오즈 곡선과 대안적인 최적분류점

오즈 곡선으로부터 설정할 수 있는 대안적인 정확도 측도인 최대사각형(maximum rectangle, MR)을 다음과 같이 제안한다.

Definition 1. Figure 2에서 오즈 곡선의 한 점 $E = (F_n(u)/F_d(u), (1 - F_d(u))/(1 - F_n(u)))$ 과 점 (1, 1)로 형성되는 사각형 면적을 최대로 하는 측도를 최대사각형(MR)이라 하며, 이 점에 대응하는 u 를 최적분류점으로

정의한다.

$$\begin{aligned} MR &= \max \left\{ \left(1 - \frac{F_n(u)}{F_d(u)} \right) \times \left(1 - \frac{1 - F_d(u)}{1 - F_n(u)} \right) \right\}, \\ &= \max \left\{ \frac{(F_d(u) - F_n(u))^2}{F_d(u)(1 - F_n(u))} \right\}. \end{aligned} \quad (2.3)$$

Figure 1과 2의 오즈 곡선 형태가 다른 이유는 Hong 등 (2021)에서 설명하였듯이 오즈 곡선을 구성하는 수평과 수직축의 오즈는 두 분류분포의 분산에 민감하기 때문이며, 3절에서 다양한 오즈 곡선에 대하여 토론한다.

대안적인 MR 측도는 Hong 등 (2011)과 Yoo와 Hong (2011)이 구분한 다섯 범주의 조건에는 포함되지 않지만, 오즈 곡선으로부터 쉽게 설정할 수 있는 최적분류점이 된다. 두 표본의 분포함수를 비교하는 비모수 검정법 중의 하나인 Anderson-Darling 검정 통계량은 다음과 같다 (Pettitt, 1976).

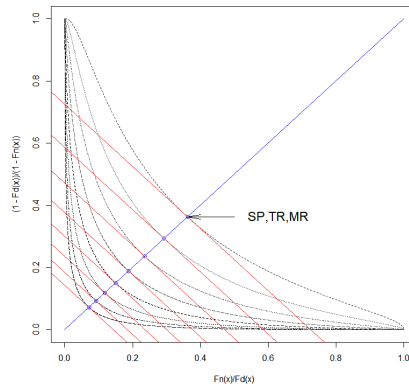
$$A^2 = \frac{pq}{(p+q)} \int_{-\infty}^{\infty} \frac{(F_d(u) - F_n(u))^2}{F(u)(1 - F(u))} dF(u), \quad (2.4)$$

여기서 $F(u) = (pF_d(u) + qF_n(u))/(p+q)$ 으로 전체 표본크기 $p+q$ 에 대한 누적분포함수이다. 그러나 MR 측도의 식 (2.3)은 식 (2.4)의 적분 안의 식과 유사하며 $F(u)$ 가 $F_d(u)$ 와 $F_n(u)$ 의 선형결합이 아니라 $F_d(u)$ 와 $F_n(u)$ 로 각각 대체한 식이라는 것을 파악할 수 있다. 따라서 본 연구에서 제안한 MR 측도는 오즈 곡선으로부터 설명할 수 있는 타당성 있는 측도로 생각할 수 있으며, 다른 정확도 측도들과의 성능은 3절에서 비교 한다.

3. 오즈 곡선으로부터 얻는 최적분류점들

이진 분류모형의 다양한 확률분포함수를 고려하여 두 종류의 정확도 측도인 TR과 SP와 본 연구에서 제안한 MR 측도에 대응한 최적분류점들을 구하고 비교한다. ROC 곡선에서의 판단 기준을 설정할 때에 Joseph (2005)은 $F_d(x)$ 를 표준정규분포 $N(\mu_d = 0, \sigma_d^2 = 1)$ 그리고 $F_n(x)$ 를 평균과 분산이 각각 μ_n, σ_n^2 인 정규분포로 가정하고, 두 분포의 표준화된 평균 차이 (standardized mean difference, $SMD = (\mu_n - \mu_d) / \sqrt{(\sigma_d^2 + \sigma_n^2)}$)를 기반으로 하였다. 본 연구에서도 유사하게 설정하고, 두 분포의 분산이 동일한 경우와 건강 상태의 분산이 서로 다른 두 종류의 경우인 $\sigma_n^2 = 1, 0.5^2, 1.5^2$ 인 세 종류의 경우 각각에 대하여 0보다 큰 값을 갖는 μ_n 의 변화에 대응하는 다양한 오즈 곡선으로부터 분류점을 구한다. 분류점을 구하고 표본수가 $p = 200, q = 800$ 인 경우와 $p = 800, q = 200$ 인 경우를 고려하여, 정확도(accuracy), 민감도(sensitivity, recall), 경보도(alarm rate), negative predictive value (NPV), 정밀도(precision, positive predictive value), 특이도(specificity, 1-false alarm rate) 통계량 값을 비교하고자 한다. 위의 여섯 가지 통계량의 대한 정의는 다음과 같다 (Bradley, 1997; Hong 등, 2010).

$$\begin{aligned} \text{정확도} &= \frac{TP + TN}{TP + TN + FN + FP}, \\ \text{민감도} &= \frac{TP}{TP + FN}, \\ \text{경보도} &= \frac{TP + FP}{TP + TN + FN + FP}, \\ \text{NPV} &= \frac{TN}{TN + FN}, \\ \text{정밀도} &= \frac{TP}{TP + FP}, \\ \text{특이도} &= \frac{TN}{FP + TN}. \end{aligned}$$

Figure 3: Thresholds on Odds curve when $\sigma_d^2 = \sigma_n^2$.Table 1: Thresholds when $\sigma_d^2 = \sigma_n^2$

SMD	μ_n	Measures	Thresholds	Coordinates	α	β	$\alpha + \beta$
0.8839	1.25	TR, SP, MR	0.625	(0.3624, 0.3624)	0.2660	0.2660	0.5320
1.0607	1.50	TR, SP, MR	0.750	(0.2930, 0.2930)	0.2266	0.2266	0.4532
1.2374	1.75	TR, SP, MR	0.875	(0.2358, 0.2358)	0.1908	0.1908	0.3816
1.4142	2.00	TR, SP, MR	1.000	(0.1886, 0.1886)	0.1587	0.1587	0.3174
1.5910	2.25	TR, SP, MR	1.125	(0.1498, 0.1498)	0.1303	0.1303	0.2606
1.7678	2.50	TR, SP, MR	1.250	(0.1181, 0.1181)	0.1056	0.1056	0.2112
1.9445	2.75	TR, SP, MR	1.375	(0.0924, 0.0924)	0.0846	0.0846	0.1692
2.1213	3.00	TR, SP, MR	1.500	(0.0716, 0.0716)	0.0668	0.0668	0.1336

3.1. 경우 1: 두 분포의 분산이 동일한 경우($\sigma_d^2 = \sigma_n^2$)

질병과 건강 상태의 분포의 분산이 동일한 경우 1에서의($\sigma_d^2 = \sigma_n^2 = 1.0$) 오즈 곡선들을 Joseph (2005)의 SMD와 동일하게 설정하고 특히 판별 기준의 13단계 중에서 판별 가능한 여섯 단계부터 살펴보기 위하여 건강 상태의 분포의 평균 μ_n 을 1.25부터 3.0까지에 대응시켜 Figure 3에 구현하였다.

경우 1의 오즈 곡선들은 Hong 등 (2021)에서 상세히 설명하였는데 Figure 3에서의 오즈 곡선들의 제일 위의 곡선은 건강 상태의 분포의 평균 $\mu_n = 1.25$ 이고 제일 아래 곡선은 $\mu_n = 3.0$ 이다, Figure 3에서 원점과 점 (1, 1)을 지나는 대각선과 오즈 곡선이 교차하는 점은 SP를 만족하는 분류점이다. 그리고 오즈 곡선의 접선의 기울기가 Lemma 2에서 정의한 정분류율 TPR/TNR의 제곱의 음수와 일치하는 점은 TR을 만족하는 분류점으로 SP와 MR을 만족하는 분류점과 일치하는 것을 탐색할 수 있다. 정확도 측도 TR, SP, MR을 만족하는 분류점과 이에 대응하는 오즈곡선의 좌표를 Table 1에 구하였다. Table 1을 살펴보면, TR, SP, MR의 최적분류점은 모두 동일하며 두 모평균의 가운데 위치한다($X_{TR} = X_{SP} = X_{MR}$).

각각의 최적분류점들의 오류를 살펴보면, 제1종과 제2종의 오류를 각각 $FNR=\alpha$ 와 $FPR=\beta$ 그리고 두 오류합($FNR+FPR$)을 구하여 Table 1에 정리하였다. Table 1을 통하여 TR, SP, MR의 측도를 만족하는 최적분류점은 질병과 건강분포의 평균들의 중간지점이므로 최적분류점에 대응하는 FNR과 FPR는 동일한 값을 가지므로 $FNR+FPR$ 은 FNR 또는 FPR의 두배로 나타난다. 또한 μ_n 이 증가할수록 두 분포간의 거리가 멀어져 모든 오류들이 감소하는 것을 탐색할 수 있다. 세 종류의 정확도 측도들의 오류와 합의 크기는 다음과 같은 관계를

Table 2: Various measures

$\sigma_d^2 = \sigma_n^2$	$p = 200, q = 800$	$p = 800, q = 200$
	TR, SP, MR	TR, SP, MR
Accuracy	0.8413	0.8413
Sensitivity	0.8413	0.8413
Alarm rate	0.2952	0.7048
NPV	0.9550	0.5700
Precision	0.5700	0.9550
Specificity	0.8413	0.8413

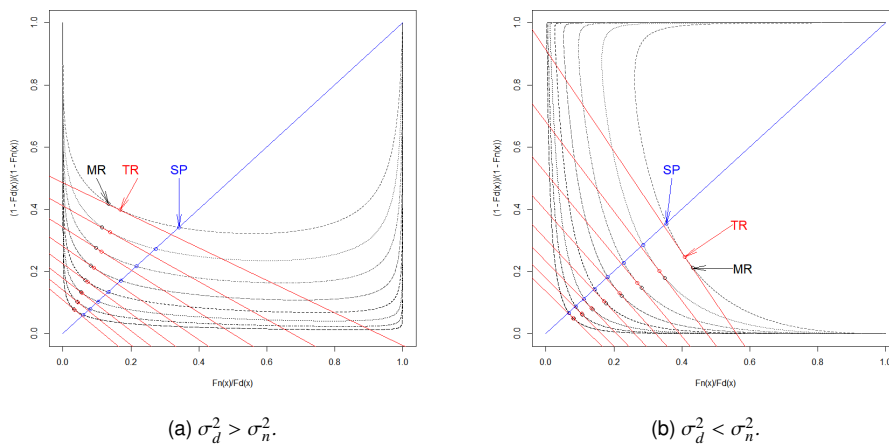


Figure 4: Thresholds on Odds curve when $\sigma_d^2 \neq \sigma_n^2$.

갖는다.

$$\alpha_{TR} = \alpha_{SP} = \alpha_{MR}, \quad \beta_{TR} = \beta_{SP} = \beta_{MR}, \quad (\alpha + \beta)_{TR} = (\alpha + \beta)_{SP} = (\alpha + \beta)_{MR}.$$

그러므로 두 분포의 분산이 동일할 때에 오류의 관점에서 MR 측도는 기존에 가장 많이 사용되는 TR, SP 측도처럼 보편적으로 사용할 수 있다.

판별 기준의 13단계 중 예를 들어 아홉번째 단계인 $SMD=1.4142$ 일 때, 질병과 건강 상태의 표본수를 각각 $p = 200, q = 800$ 와 $p = 800, q = 200$ 으로 하는 두 가지 경우에서 TR, SP, MR에 대응하는 측도들을 Table 2에 정리하였다. 우선, 표본수에 상관없이 정확도, 민감도, 특이도는 동일하며, 질병 상태의 표본수가 커지면 경보도와 정밀도는 커지고, NPV는 작아진다. 이는 SP 측도의 특징인데, 표본수가 대칭으로 바뀌면 NPV와 정밀도가 서로 바뀌며, 경보도의 합은 1이 되도록 바뀌고 정확도, 민감도, 특이도는 동일하다는 것을 파악할 수 있다. 또한, 질병과 건강 상태의 분포의 분산이 동일하고 질병 상태의 표본수가 작은 $p < q$ 인 경우에는 NPV 값이 크며, 질병 상태의 표본수가 커지는 $p > q$ 일 때는 경보도와 정밀도 값이 크다.

3.2. 경우 2: 건강 상태의 분산이 작은 경우($\sigma_d^2 > \sigma_n^2 = 0.5^2$)

건강 상태의 분포의 분산이 작은 경우 2의($\sigma_d^2 > \sigma_n^2 = 0.5^2$) 오즈 곡선들을 경우 1의 SMD와 일치하기 위하여 SMD기준 여섯 단계부터인 건강 상태의 분포의 평균 μ_n 을 0.9882부터 2.3727까지에 대응시켜 Figure 4(a)에 구현하였다.

Table 3: Thresholds when $\sigma_d^2 > \sigma_n^2$

SMD	μ_n	Measures	Thresholds	Coordinates	α	β	$\alpha + \beta$
0.8839	0.9882	TR	0.3710	(0.1683, 0.3986)	0.3553	0.1085	0.4638
		SP	0.6588	(0.3423, 0.3423)	0.2550	0.2550	0.5100
		MR	0.2953	(0.1346, 0.4186)	0.3839	0.0829	0.4668
1.0607	1.1859	TR	0.5385	(0.1386, 0.3271)	0.2951	0.0977	0.3928
		SP	0.7906	(0.2732, 0.2732)	0.2146	0.2146	0.4292
		MR	0.4810	(0.1158, 0.3424)	0.3153	0.0793	0.3946
1.2374	1.3835	TR	0.6989	(0.1128, 0.2650)	0.2423	0.0855	0.3278
		SP	0.9223	(0.2168, 0.2168)	0.1782	0.1782	0.3564
		MR	0.6559	(0.0979, 0.2760)	0.2559	0.0728	0.3287
1.4142	1.5811	TR	0.8539	(0.0908, 0.2120)	0.1966	0.0729	0.2695
		SP	1.0541	(0.1709, 0.1709)	0.1459	0.1459	0.2918
		MR	0.8223	(0.0813, 0.2196)	0.2055	0.0646	0.2701
1.5910	1.7788	TR	1.0048	(0.0722, 0.1677)	0.1575	0.0608	0.2183
		SP	1.1859	(0.1336, 0.1336)	0.1178	0.1178	0.2356
		MR	0.9819	(0.0663, 0.1727)	0.1631	0.0555	0.2186
1.7678	1.9765	TR	1.1527	(0.0568, 0.1310)	0.1245	0.0497	0.1742
		SP	1.3177	(0.1035, 0.1035)	0.0938	0.0938	0.1876
		MR	1.1361	(0.0532, 0.1342)	0.1280	0.0464	0.1744
1.9445	2.1740	TR	1.2978	(0.0441, 0.1012)	0.0972	0.0399	0.1371
		SP	1.4493	(0.0795, 0.0795)	0.0736	0.0736	0.1472
		MR	1.2860	(0.0420, 0.1031)	0.0992	0.0379	0.1371
2.1213	2.3717	TR	1.4412	(0.0339, 0.0772)	0.0748	0.0314	0.1062
		SP	1.5811	(0.0604, 0.0604)	0.0569	0.0569	0.1138
		MR	1.4328	(0.0327, 0.0783)	0.0760	0.0302	0.1062

질병과 건강 상태의 분포의 분산이 동일하지 않은 경우의 오즈 곡선들도 Hong 등 (2021)에서 상세히 설명하였는데 Figure 4(a)에서의 오즈 곡선들은 건강 상태의 분포의 분산이 작은 경우이고, 제일 위의 곡선은 건강 상태의 분포의 평균 $\mu_n = 0.9882$ 이고 제일 아래 곡선은 $\mu_n = 2.3727$ 이다. Figure 4(a)에서 대각선과 오즈 곡선이 교차하는 점은 SP를 만족하는 분류점이고, 오즈 곡선의 접선의 기울기가 정분류율 TPR/TNR의 제곱의 음수와 일치하는 점은 TR을 만족하는 분류점으로 SP를 만족하는 분류점보다 왼쪽에 위치한다. 그리고 본 연구에서 제안한 MR 측도를 만족하는 분류점은 TR을 만족하는 분류점보다 더 왼쪽에 위치하는 것을 탐색한다. 따라서 경우 2에서의 세 종류의 분류점은 SP 분류점이 제일 오른쪽에 그리고 MR 분류점이 제일 왼쪽에 위치하며 가운데에 TR 분류점이 위치한다. 그리고 건강 상태의 분포의 평균 μ_n 이 증가할수록 세 분류점이 가까이 모이는데 특히 MR 분류점이 TR 분류점에 수렴하는 것을 파악할 수 있다.

경우 2의 TR, SP, MR을 만족하는 분류점과 이에 대응하는 오즈곡선의 좌표를 Table 3에 구하였다. Table 3을 통하여 SMD기준 여섯 단계부터 SP의 분류점이 가장 크며, MR이 가장 작으며 SMD가 커질수록 MR은 TR에 수렴함을 확인할 수 있다 ($X_{MR} \approx X_{TR} < X_{SP}$).

TR, SP, MR의 측도를 만족하는 최적분류점에 대응하는 제1종 오류는 SP측도가 가장 작으며 TR, MR측도 순으로 작다. 제2종 오류는 MR 측도가 가장 작으며 TR, SP측도 순으로 작다. 오류합은 TR측도가 가장 작으며 MR, SP측도 순으로 작다. 세 종류의 정확도 측도들의 오류와 합의 크기는 다음과 같은 관계를 갖는다.

$$\alpha_{SP} < \alpha_{TR} < \alpha_{MR}, \quad \beta_{MR} < \beta_{TR} < \beta_{SP}, \quad (\alpha + \beta)_{TR} < (\alpha + \beta)_{MR} < (\alpha + \beta)_{SP}.$$

Table 4: Various measures

$\sigma_d^2 > \sigma_n^2$	$p = 200, q = 800$			$p = 800, q = 200$		
	TR	SP	MR	TR	SP	MR
Accuracy	0.9023	0.8541	0.9073	0.8282	0.8541	0.8227
Sensitivity	0.8034	0.8541	0.7945	0.8034	0.8541	0.7945
Alarm rate	0.2190	0.2875	0.2105	0.6573	0.7125	0.6485
NPV	0.9497	0.9590	0.9479	0.5411	0.5940	0.5322
Precision	0.7336	0.5940	0.7549	0.9778	0.9590	0.9801
Specificity	0.9271	0.8541	0.9335	0.9271	0.8541	0.9355

그러므로 경우 2에서 MR 측도는 제2종 오류의 관점에서 TR, SP측도들보다 좋으며 오류합 관점에서도 TR 보다는 높지만 매우 가깝기 때문에 TR, SP 측도만큼 좋은 측도라고 할 수 있다.

경우 1에서와 동일하게 질병 상태 분포의 분산이 건강 상태 분포의 분산보다 큰 경우 2에서도 판별 기준의 13단계 중 아홉번째 단계인 $SMD = 1.4142$ 일 때를 고려하면서 질병과 건강 상태의 표본수가 각각 $p = 200, q = 800$ ($p < q$)이고 $p = 800, q = 200$ ($p > q$)일 때를 Table 4에 정리하였다. MR 측도가 가장 큰 정확도, 정밀도, 특이도를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. 그리고 SP, TR, MR 순으로 좋은 측도는 민감도, 경보도, NPV로 민감도, 경보도, NPV에서 MR 측도가 가장 작은 값을 갖는다.

다음에는 질병과 건강 상태의 표본수가 각각 $p = 800, q = 200$ ($p > q$)일 때는 MR 측도가 가장 큰 정밀도, 특이도를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 측도는 정확도, 민감도, 경보도, NPV이며 MR 측도가 가장 작은 값을 갖는다.

경우 2의 $p < q$ 와 $p > q$ 인 상태에서는 MR 측도가 가장 큰 값과 작은 값을 갖는 통계량이 정확도를 제외 하면 동일하다. 즉, 정밀도와 특이도에서는 MR 측도가 가장 큰 값을 가지며, 민감도, 경보도, NPV에서는 MR 측도가 가장 작은 값을 갖는다. 그리고 정확도 통계량에서는 MR 측도와 TR 측도의 통계량의 값 차이가 매우 적다는 것을 파악할 수 있다. 따라서 질병 상태 분포의 분산이 건강 상태 분포의 분산보다 큰 경우 2에서는 MR 측도가 가장 큰 값과 작은 값을 갖는 통계량이 유사하다는 것을 탐색하였다.

경우 1에서 언급한 SP의 특징으로, 표본수가 대칭으로 바뀔 때, NPV와 정밀도가 서로 바뀌며, 경보도의 합은 1이 되도록 바뀌어 지고, 정확도, 민감도, 특이도는 동일한데 이 또한 경우 2에서도 같은 특징을 가짐을 알 수 있다.

$\sigma_d^2 > \sigma_n^2$ 인 경우에 최적분류점의 분류 성과를 측정하는 여섯 가지 통계량 중에서 MR을 만족하는 최적분류점이 가장 큰 값을 갖는 통계량은 정밀도, 특이도와 정확도($p < q$ 경우)이고, 가장 작은 값을 갖는 통계량은 민감도, 경보도, NPV와 정확도($p > q$ 경우)이다.

3.3. 경우 3: 건강 상태의 분산이 큰 경우($\sigma_d^2 < \sigma_n^2 = 1.5^2$)

건강 상태의 분포의 분산이 큰 경우 3($\sigma_d^2 < \sigma_n^2 = 1.5^2$)에서의 오즈 곡선들을 경우 1과 2의 SMD와 일치하기 위하여 SMD기준 여섯 단계부터인 건강 상태의 분포의 평균 μ_n 을 1.5934부터 3.8243까지에 대응시켜 Figure 4(b)에 구현하였다. 경우 3에서 오즈 곡선들의 제일 오른쪽 곡선은 건강 상태의 분포의 평균 $\mu_n = 1.5934$ 고 제일 왼쪽 곡선은 $\mu_n = 3.8243$ 이다. 대각선과 오즈 곡선이 교차하는 점은 SP를 만족하는 분류점이고, 오즈 곡선의 접선의 기울기가 정분류율 TPR/TNR의 제곱의 음수와 일치하는 점은 TR을 만족하는 분류점으로 SP를 만족하는 분류점보다 오른쪽에 위치한다. 본 연구에서 제안한 MR 측도를 만족하는 분류점은 TR을 만족하는 분류점보다 더 오른쪽에 위치하는 것을 탐색한다. 따라서 경우 3에서의 세 종류의 분류점은, 경우 2와 반대로, SP 분류점이 제일 왼쪽에 그리고 MR 분류점이 제일 오른쪽에 위치하며 가운데에 TR 분류점이 위치한다.

Table 5: Thresholds when $\sigma_d^2 < \sigma_n^2$

SMD	μ_n	Measures	Thresholds	Coordinates	α	β	$\alpha + \beta$
0.8839	1.5934	TR	0.9871	(0.4092, 0.2463)	0.1618	0.3430	0.5048
		SP	0.6374	(0.3549, 0.3549)	0.2619	0.2619	0.5238
		MR	1.1167	(0.4324, 0.2114)	0.1321	0.3753	0.5074
1.0607	1.9121	TR	1.0635	(0.3338, 0.2013)	0.1438	0.2858	0.4296
		SP	0.7648	(0.2857, 0.2857)	0.2222	0.2222	0.4444
		MR	1.1563	(0.3506, 0.1787)	0.1238	0.3072	0.4310
1.2374	2.2308	TR	1.1523	(0.2697, 0.1631)	0.1246	0.2361	0.3607
		SP	0.8923	(0.2287, 0.2287)	0.1861	0.1861	0.3722
		MR	1.2185	(0.2812, 0.1487)	0.1115	0.2499	0.3614
1.4142	2.5495	TR	1.2497	(0.2159, 0.1310)	0.1057	0.1931	0.2988
		SP	1.0198	(0.1819, 0.1819)	0.1539	0.1539	0.3078
		MR	1.2968	(0.2236, 0.1220)	0.0974	0.2018	0.2992
1.5910	2.8682	TR	1.3532	(0.1713, 0.1043)	0.0880	0.1562	0.2442
		SP	1.1473	(0.1437, 0.1437)	0.1256	0.1256	0.2512
		MR	1.3867	(0.1762, 0.0987)	0.0828	0.1617	0.2445
1.7678	3.1868	TR	1.4610	(0.1347, 0.0823)	0.0720	0.1250	0.1970
		SP	1.2747	(0.1126, 0.1126)	0.1012	0.1012	0.2024
		MR	1.4848	(0.1377, 0.0789)	0.0688	0.1283	0.1971
1.9445	3.5056	TR	1.5723	(0.1048, 0.0643)	0.0579	0.0987	0.1566
		SP	1.4022	(0.0875, 0.0875)	0.0804	0.0804	0.1608
		MR	1.5891	(0.1067, 0.0623)	0.0560	0.1007	0.1567
2.1213	3.8243	TR	1.6861	(0.0807, 0.0497)	0.0459	0.0770	0.1229
		SP	1.5297	(0.0673, 0.0673)	0.0630	0.0630	0.1260
		MR	1.6979	(0.0818, 0.0486)	0.0448	0.0782	0.1230

경우 3에서는 경우 2와 유사하게 건강 상태의 분포의 평균 μ_n 이 증가할수록 세 분류점이 가까이 모이고 특히 MR 분류점이 TR 분류점에 수렴하는 것을 파악할 수 있다.

경우 3에서 정확도 측도 TR, SP, MR을 만족하는 분류점과 이에 대응하는 오즈곡선의 좌표를 Table 5에 구하고, 이를 바탕으로 SMD기준 여섯 단계부터 SP가 가장 작으며, MR이 가장 크고, SMD가 커질수록 MR은 TR에 수렴하며 ($X_{SP} < X_{TR} \approx X_{MR}$), 건강 상태의 분산이 작은 경우의 반대 현상으로 나타난다. Table 5를 통하여 각각의 최적분류점에 대응하는 제1종 오류는 MR측도가 가장 작으며 TR, SP측도 순으로 작다. 제2종 오류는 SP측도가 가장 작으며 TR, MR측도 순으로 작다. 오류합은 TR측도가 가장 작으며 MR, SP측도 순으로 작다. 세 종류의 정확도 측도들의 오류와 합 크기는 다음과 같은 관계를 갖는다.

$$\alpha_{MR} < \alpha_{TR} < \alpha_{SP}, \quad \beta_{SP} < \beta_{TR} < \beta_{MR}, \quad (\alpha + \beta)_{TR} < (\alpha + \beta)_{MR} < (\alpha + \beta)_{SP}.$$

그러므로 MR 측도는 제1종 오류의 관점에서 TR, SP측도들보다 좋으며 오류합 관점에서도 TR보다는 높지만 매우 가깝기 때문에 TR, SP 측도만큼 좋은 측도라고 할 수 있다.

경우 1과 2에서와 동일하게 질병 상태 분포의 분산이 건강 상태 분포의 분산보다 작은 경우 3에서도 판별 기준의 13단계 중 아홉번째 단계인 SMD = 1.4142일 때를 고려하면서 질병과 건강 상태의 표본수가 각각 $p = 200$, $q = 800$ ($p < q$)인 경우에 TR, SP, MR에 대응하는 측도들을 살펴보면, MR 측도가 가장 큰 민감도, 경보도, NPV를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 측도는 정확도, 정밀도, 특이도이며 MR 측도가 가장 작은 값을 갖는다.

Table 6: Various measures

$\sigma_d^2 < \sigma_n^2$	$p = 200, q = 800$			$p = 800, q = 200$		
	TR	SP	MR	TR	SP	MR
Accuracy	0.8243	0.8461	0.8190	0.8768	0.8461	0.8818
Sensitivity	0.8944	0.8461	0.9027	0.8944	0.8461	0.9027
Alarm rate	0.3334	0.2923	0.3420	0.7541	0.7077	0.7625
NPV	0.9683	0.9565	0.9704	0.6563	0.5788	0.6722
Precision	0.5365	0.5788	0.5278	0.9488	0.9565	0.9471
Specificity	0.8068	0.8461	0.7981	0.8068	0.8461	0.7981

이번에는 질병과 건강 상태의 표본수가 각각 $p = 200, q = 800$ ($p < q$)이고 $p = 800, q = 200$ ($p > q$) 일 때를 Table 6에 정리하였다. $p = 800, q = 200$ ($p > q$)인 경우를 살펴보면, MR 측도가 가장 큰 정확도, 민감도, 경보도, NPV를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 통계량은 정밀도, 특이도이며 MR 측도가 가장 작은 값을 갖는다. 경우 3에서 MR 측도가 가장 큰 값과 작은 값을 갖는 통계량들이 경우 2와 반대인 현상을 보인다. 즉, $p < q$ 인 경우에 경우 2에서 MR 측도가 가장 작은 값을 갖는 민감도, 경보도, NPV 통계량이 경우 3에서는 MR 측도가 가장 큰 값을 가지며, 가장 큰 값을 갖는 정확도, 정밀도, 특이도 통계량이 경우 3에서는 MR 측도가 가장 작은 값을 갖는다. 그리고 $p > q$ 인 경우에 경우 2에서 MR 측도가 가장 큰 값을 갖는 정밀도, 특이도 통계량이 경우 3에서는 MR 측도가 가장 작은 값을 가지며, 경우 2에서 MR 측도가 가장 작은 값을 갖는 정확도, 민감도, 경보도, NPV 통계량이 경우 3에서는 MR 측도가 가장 큰 값을 갖는다. 경우 3에서도 경우 2와 동일하게 정확도 통계량 값만이 차이가 나는데 MR 측도와 TR 측도의 정확도 통계량의 값 차이가 매우 적다는 것을 파악할 수 있다. 따라서 질병 상태 분포의 분산이 건강 상태 분포의 분산보다 작은 경우 3에서도 경우 2와 마찬가지로 MR 측도가 가장 큰 값과 작은 값을 갖는 통계량이 유사하다는 것을 발견하였다. 경우 1과 경우 2에서 언급한 SP의 특징으로, 표본수가 대칭으로 바뀔 때, NPV와 정밀도가 서로 바뀌며, 경보도의 합은 1이 되도록 바뀌어 지고, 정확도, 민감도, 특이도는 동일한데 이 또한 경우 3에서도 같은 특징을 갖는다.

경우 3($\sigma_d^2 < \sigma_n^2$)에 최적분류점의 분류 성과를 측정하는 여섯 가지 통계량 중에서 MR을 만족하는 최적분류점은 경우 2($\sigma_d^2 > \sigma_n^2$)인 경우와 반대되는 현상으로 가장 작은 값을 갖는 통계량은 정밀도, 특이도, 정확도 ($p < q$ 경우)이며, 가장 큰 값을 갖는 통계량은 민감도, 경보도, NPV, 정확도 ($p > q$ 경우)이다. 따라서 최적분류점의 분류 성과를 측정하는 여섯 가지 통계량 중에서 MR을 만족하는 최적분류점이 가장 큰 값을 갖을 때는 $\sigma_d^2 > \sigma_n^2$ 인 경우 2에서 정밀도, 특이도, 정확도 ($p < q$)이며, $\sigma_d^2 < \sigma_n^2$ 인 경우 3에서는 민감도, 경보도, NPV, 정확도 ($p > q$)이다. 그러므로 MR을 만족하는 최적분류점이 가장 많이 사용하는 TR을 만족하는 최적분류점보다 $\sigma_d^2 > \sigma_n^2$ 인 경우에 정밀도, 특이도, 정확도 ($p < q$) 측면에서 분류 성과가 우수하고, $\sigma_d^2 < \sigma_n^2$ 인 경우에 민감도, 경보도, NPV, 정확도 ($p > q$) 측면에서 분류 성과가 우수한 최적분류점이라고 판단할 수 있다.

4. 실증예제

2018년도에 국내 K 은행의 신용평가모형을 위한 RS 등급(risk score rating) 자료를 실증예제로 채택하여 오즈 곡선으로 구현하여 본 논문에서 토론한 정확도 측도들에 대한 분류점을 구하고 이에 대응하는 오분류율을 비교한다. 표본크기는 65,455로 3단계로 분류하며, 높을수록 파산확률이 높다고 간주하는 신용을 20개의 RS 등급으로 분류했다. Hong과 Jang (2020)은 표본크기가 작은 세 번째 단계를 제외하고 첫 번째와 두 번째 단계를 부도(default)와 정상(non-default) 상태로 간주하여 Table 7에 정리하였다.

부도와 정상 상태의 분포함수를 각각 정규분포 $N(7.89, 2.74^2)$ 과 $N(12.4, 2.99^2)$ 으로 추정하고 각각의 경

Table 7: RS rating data of K bank

RS grade	1	2	3	4	5	6	7	8	9	10	
Default	745	1424	1837	2514	38060	5893	7676	8985	9455	5676	
Non-default	0	0	0	55	56	175	189	279	518	421	
RS grade	11	12	13	14	15	16	17	18	19	20	Total
Default	3759	2854	1546	1078	5	16	12	0	0	0	57335
Non-default	875	761	631	1102	555	965	347	0	0	0	6929

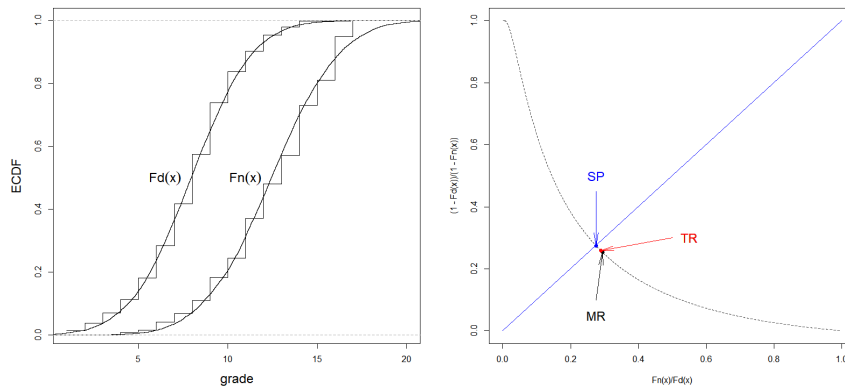


Figure 5: Two cdf's and an odds curve.

험 누적분포함수들과 함께 Figure 5의 왼쪽에 구현하였다. 그리고 추정된 누적분포함수들에 대응하는 오즈 곡선과 세 종류의 정확도 측도로부터 얻은 최적분류점들을 Figure 5의 오른쪽에 나타내었다.

부도 상태의 분포를 표준정규분포로 설정하면, 정상 상태의 분포는 $N(4.59, 1.09^2)$ 으로 표준화시킬 수 있다. 표준화된 정상 상태의 표준편차 $\sigma_n = 1.09$ 은 $\sigma_d = 1.0$ 보다 크기 때문에 3.3절의 $\sigma_d^2 < \sigma_n^2$ 경우인 Table 3에서와 유사하게 가장 좋고 많이 사용되는 TR을 만족하는 최적분류점이 SP보다 MR을 만족하는 최적분류점에 더 근접하며 $X_{SP} = 10.0466 < X_{TR} = 10.2047 \approx X_{MR} = 10.2625$ 임을 탐색할 수 있다. 세 종류의 최적분류점과 이에 대응하는 혼동행렬을 Table 8에 정리하고, 오즈 곡선에 정확도 측도에 대응하는 최적분류점을 Figure 5의 오른쪽에 구현하였다. 그러나 $\sigma_n = 1.09$ 은 $\sigma_d = 1.0$ 과 차이가 거의 없기 때문에, 3.1절의 결과와 유사하게 세 최적분류점들이 근사하게 모여있다고 판단할 수 있다.

세 종류의 정확도 측도들의 오류와 합의 크기는 다음과 같은 관계를 갖는다. 이는 3.3절의 건강상태의 분산이 작을 때인 경우 3과 같은 특징을 가진다.

$$\alpha_{MR} < \alpha_{TR} < \alpha_{SP}, \quad \beta_{SP} < \beta_{TR} < \beta_{MR}, \quad (\alpha + \beta)_{TR} < (\alpha + \beta)_{MR} < (\alpha + \beta)_{SP}.$$

그러므로 본 연구에서 제안한 MR 측도가 실증예제에서도 성립함을 알 수 있고, 오류의 관점에서도 가장 많이 사용되는 TR, SP측도만큼 충분히 보편적이고 효율적으로 사용할 수 있다.

부도와 정상 상태의 표본수는 각각 $p = 57,335$, $q = 6,929$ ($p > q$)인 경우에, TR, SP, MR에 대응하는 측도들은 Table 9에 나타내었다. MR 측도가 가장 큰 정확도, 민감도, 경보도, NPV를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 통계량은 정밀도, 특이도이며 MR 측도가 가장 작은 값을 갖는다. 또한, MR 측도와 TR 측도에서 정확도 통계량의 값의 차이가 매우 작다. 이 현상은 3절의 부도 상태 분포의 분산이 정상 상태 분포의 분산보다 작은 경우 3에서 표본수가 $p > q$ 인 경우인 Table 6의 결과와

Table 8: Thresholds and confusion matrices

			SP		TR		MR	
Threshold			10.0466		10.2047		10.2625	
Confusion matrix	TP	FP	44973	1494	45921	1604	46250	1644
	FN	TN	12362	5435	11415	5325	11085	5285

Table 9: Various measures

	SP	TR	MR
Accuracy	0.7844	0.7974	0.8019
Sensitivity	0.7844	0.8009	0.8067
Alarm rate	0.7231	0.7395	0.7453
NPV	0.3054	0.3181	0.3228
Precision	0.9678	0.9663	0.9657
Specificity	0.7844	0.7685	0.7627

비교하면, 실증예제와 동일하게 해석할 수 있다. 따라서 3절에서 언급했던 부도 상태의 분포의 분산과 정상 상태의 분포의 분산크기로 나는 세 가지 경우와 표본수에 따라 나는 두 가지 경우가 실증예제에서도 마찬가지로 잘 적용됨을 확인할 수 있다.

MR을 만족하는 최적분류점에서 민감도, 경보도, NPV, 정확도 통계량이 가장 큰 값을 가지며, 정밀도, 특이도 통계량이 가장 작은 값을 갖는다. 따라서, TR을 만족하는 최적분류점보다 MR을 만족하는 최적분류점이 민감도, 경보도, NPV와 정확도 측면에서 우수한 분류 성과를 나타낸다.

5. 결론

본 연구에서는 정확도 측도 중에서 진실율(TR: J, SSS, AC, MVD 측도 포함)과 대칭점(SP) 측도를 만족하는 분류점들은 오즈 곡선으로부터 설명할 수 있음을 발견하였다. 그리고 오즈 곡선으로부터 구할 수 있는 최대사각형(MR)이라는 정확도 측도를 제안하였다.

우선, 질병과 건강 상태의 분포의 분산이 동일한 경우에는 TR, SP, MR의 최적분류점은 모두 동일하며 두 모평균의 가운데 위치한다. 표본수에 상관없이 정확도, 민감도, 특이도는 동일하며, 질병 상태의 표본수가 커지면 경보도와 정밀도는 커지고, NPV는 작아진다. 또한, 질병 상태의 표본수가 작은 $p < q$ 인 경우에는 NPV 값이 크며, 경보도와 정밀도는 작아진다.

건강 상태의 분산이 질병 상태의 분산보다 작은 경우에는 SP를 만족하는 최적분류점이 가장 크며, MR을 만족하는 최적분류점이 가장 작으며 SMD가 커질수록 MR을 만족하는 최적분류점은 TR을 만족하는 최적분류점에 수렴한다($X_{MR} \approx X_{TR} < X_{SP}$). 질병 상태의 표본수가 작은 경우($p < q$)에서는 MR 측도가 가장 큰 정확도, 정밀도, 특이도를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 측도는 민감도, 경보도, NPV로 민감도, 경보도, NPV에서 MR 측도가 가장 작은 값을 갖는다. 다음에는 질병 상태의 표본수가 큰 경우($p > q$)일 때는 MR 측도가 가장 큰 정밀도, 특이도를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 측도는 정확도, 민감도, 경보도, NPV이며 MR 측도가 가장 작은 값을 갖는다. 그리고 $p < q$ 와 $p > q$ 인 상태에서는 MR 측도가 가장 큰 값과 작은 값을 갖는 통계량이 정확도를 제외하면 동일하다. 즉 정밀도와 특이도에서는 MR 측도가 가장 큰 값을 가지며, 민감도, 경보도, NPV에서는 MR 측도가 가장 작은 값을 갖는다. 그리고 정확도 통계량에서는 MR 측도와 TR 측도의 통계량의 값 차이가 매우 적다는 것을 파악할 수 있다. 따라서 질병 상태 분포의 분산이 건강 상태 분포의 분산보다 큰

경우에서는 MR 측도가 가장 큰 값과 작은 값을 갖는 통계량이 유사하다는 것을 탐색하였다.

건강 상태의 분산이 질병 상태의 분산보다 큰 경우에는 SP를 만족하는 최적분류점이 가장 작으며, MR을 만족하는 최적분류점이 가장 크고, SMD가 커질수록 MR을 만족하는 최적분류점은 TR을 만족하는 최적분류점에 수렴하며 ($X_{SP} < X_{TR} \approx X_{MR}$), 건강 상태의 분산이 작은 경우의 반대 현상으로 나타난다. 질병 상태의 표본수가 작은 경우에 ($p < q$) TR, SP, MR에 대응하는 측도들을 살펴보면, MR 측도가 가장 큰 민감도, 경보도, NPV를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 측도는 정확도, 정밀도, 특이도이며 여기에서는 MR 측도가 가장 작은 값을 갖는다. 질병 상태의 표본수가 큰 경우 ($p > q$)에서는 MR 측도가 가장 큰 정확도, 민감도, 경보도, NPV를 갖고 있으며 다음으로 TR, SP 순으로 좋은 측도이다. SP, TR, MR 순으로 좋은 통계량은 정밀도, 특이도이며 MR 측도가 가장 작은 값을 갖는데, 특히, MR 측도가 가장 큰 값과 작은 값을 갖는 통계량들이 경우 2와 반대인 현상을 보인다. 그리고 질병 상태 분포의 분산이 건강 상태 분포의 분산보다 작거나 큰 경우에서 본 연구에서 제안한 MR 측도가 가장 큰 값과 작은 값을 갖는 통계량이 유사하다는 것을 탐색하였다.

본 연구에서는 ROC, TOC와 TROC 곡선들과 유사하게 진실율(TR: J, SSS, AC, MVD 측도 포함)과 대칭점(SP)인 정확도 측도들을 만족하는 분류점들을 오즈 곡선으로부터 설명할 수 있음을 발견하였고, 최대사각형(MR)이라는 정확도 측도도 오즈 곡선으로부터 구할 수 있다. 제1종과 제2종 오류 및 오류합의 관점에서 MR 측도를 만족하는 최적분류점이 TR, SP 측도를 만족하는 최적분류점 만큼 효율적인 특성을 갖는다고 판단할 수 있으며, MR을 만족하는 최적분류점이 가장 많이 사용하는 TR을 만족하는 최적분류점보다 건강 상태의 분산이 질병 상태보다 작은($\sigma_d^2 > \sigma_n^2$) 경우에 정밀도, 특이도, 정확도($p < q$ 경우) 측면에서 분류 성과가 우수하고, 건강 상태의 분산이 질병 상태보다 큰($\sigma_d^2 < \sigma_n^2$) 경우에 민감도, 경보도, NPV, 정확도($p > q$ 경우) 측면에서 분류 성과가 우수한 최적분류점임을 발견하였다.

실증예제와 같은 일반적인 상황($\sigma_d^2 < \sigma_n^2, p > q$)에서는 MR을 만족하는 최적분류점이 TR을 만족하는 최적분류점보다 민감도, 경보도, NPV, 정확도 측면에서 우수한 분류 성과를 나타내므로, 오즈 곡선과 곡선으로부터 구할 수 있는 최대사각형 측도를 만족하는 최적분류점을 이진 분류모형 분석에 유용하게 활용하여 모형의 판별력 향상을 기대할 수 있다.

References

- Altman DG and Bland JM (1994). Diagnostic tests. 1: Sensitivity and specificity, *BMJ: British Medical Journal*, **308**, 1552.
- Bamber D (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph, *Journal of Mathematical Psychology*, **12**, 387—415.
- Brasil P (2010). *Diagnostic Test Accuracy Evaluation for Medical Professionals*, Package DiagnosisMed in R.
- Bradley AP (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms, *Pattern Recognitions*, **30**, 1145—1159.
- Centor RN (1991). Signal detectability: The use of ROC curves and their analyses, *Medical Decision Making*, **11**, 102—106.
- Cho MH and Hong CS (2015). Two optimal threshold criteria for ROC analysis, *Journal of the Korean Data & Information Science Society*, **26**, 255—260.
- Connell FA and Koepsell TD (1985). Measures of gain in certainty from a diagnostic test, *American Journal of Epidemiology*, **121**, 744—753.
- Egan JP and Egan JP (1975). *Signal Detection Theory and ROC-Analysis*, Academic press, New York.
- Green DM and Swets JA (1966). *Signal detection theory and psychophysics (vol. 1)*, New York.

- Hanley JA and McNeil BJ (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology*, **143**, 29—36.
- Hong CS (2009). Optimal threshold from ROC and CAP curves, *Communications in Statistics-Simulation and Computation*, **38**, 2060—2072.
- Hong CS and Choi JS (2009). Optimal threshold from ROC and CAP curves, *The Korean Journal of Applied Statistics*, **22**, 911—921.
- Hong CS and Jang DH (2020). Validation ratings for the length of the ROC curve, *Journal of the Korean Data & Information Science Society*, **31**, 851—863.
- Hong CS, Joo JS, and Choi JS (2010). Optimal thresholds from mixture distributions, *The Korean Journal of Applied Statistics*, **23**, 13—28.
- Hong CS and Lee SJ (2018). TROC curve and accuracy measures, *Journal of the Korean Data & Information Science Society*, **29**, 861—872.
- Hong CS, Lin MH, Hong SW, and Kim GC (2011). Classification accuracy measures with minimum error rate for normal mixture, *Journal of the Korean Data & Information Science Society*, **22**, 619—630.
- Hong CS, Oh SH, and Oh TG (2021). Odds curve for two classification distributions, *The Korean Journal of Applied Statistics*, **34**, 225—238.
- Joseph MP (2005). A PD validation framework for Basel II internal ratings-based systems, *Credit Scoring and Credit Control IV*
- Krzanowski WJ and Hand DJ (2009). *ROC Curves for Continuous Data*. Chapman & Hall/CRC, Boca Raton.
- Lambert J and Lipkovich I (2008). A macro for getting more out of your ROC curve, *SAS Global Forum*, **231**.
- Metz CE (1978). Basic principles of ROC analysis, *Seminars in Nuclear Medicine*, **8**, 283—298.
- Moses LE, Shapiro D, and Littenberg B (1993). Combining independent studies of a diagnostic test into a summary ROC curve: Data-analytic approaches and some additional considerations, *Statistics in Medicine*, **12**, 1293—1316.
- Pepe MS (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*, Oxford University Press, Oxford.
- Perkins NJ and Schisterman EF (2006). The inconsistency of “optimal” cutpoints obtained using two criteria based on the receiver operating characteristic curve, *American Journal of Epidemiology*, **163**, 670—675.
- Pettitt AN (1976). A two-sample Anderson-Darling rank statistic, *Biometrika*, **63**, 161—168.
- Pontius Jr, RG and Si K (2014). The total operating characteristic to measure diagnostic ability for multiple thresholds, *International Journal of Geographical Information Science*, **28**, 570—583.
- Swets JA (1988). Measuring the accuracy of diagnostic systems, *Science*, **240**, 1285—1293.
- Tasche D (2008). Validation of internal rating systems and PD estimates, *The Analytics of Risk Model Validation*, 169—196.
- Vuk M and Curk T (2006). ROC curve, lift chart and calibration plot, *Metodološki Zvezki*, **3**, 89—108.
- Yoo HS and Hong CS (2011). Optimal criterion of classification accuracy measures for normal mixture, *Communications for Statistical Applications and Methods*, **18**, 343—355.
- Youden WJ (1950). Index for rating diagnostic test, *Cancer*, **3**, 32—35.
- Zweig MH and Campbell G (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine, *Clinical chemistry*, **39**, 561—577.

오즈 곡선과 최적분류점

홍종선^{1,a}, 오탈규^a, 오세현^a

^a성균관대학교 통계학과

요 약

오즈 곡선으로 설명이 가능한 정확도 측도들을 살펴보고, 오즈 곡선의 성질을 바탕으로 대안적인 최대 사각형 정확도 측도를 제안한다. 다양한 확률분포함수와 실증예제를 고려하여 정확도 측도들에 대응하는 분류점을 구하고, 분류점을 측정하는 통계량들을 비교하면서 특징을 토론한다. 그러므로 ROC 곡선 등과 유사하게 오즈 곡선으로부터도 최적분류점들을 발견하고 설명할 수 있으며, 최대사각형 측도는 이진 분류모형의 성능을 향상시킬 수 있는 정확도 측도로 활용할 수 있다.

주요용어: 분류모형, 정확도, 최적분류점, 혼동행렬

¹교신저자: (03063) 서울특별시 중로구 성균관로 25-2, 성균관대학교 통계학과. E-mail: cshong@skku.edu