*The Korean Journal of Applied Statistics* (2020) DOI: https://doi.org/10.5351/KJAS.2020.33.4.395
**33**(4), 395–407

# Index of union and other accuracy measures

Chong Sun Hong[a,1] · So Yeon Choi[a] · Dong Hui Lim[a]

[a]Department of Statistics, Sungkyunkwan University

(Received May 6, 2020; Revised June 1, 2020; Accepted June 13, 2020)

**Abstract**

Most classification accuracy measures for optimal threshold are divided into two types: one is expressed with cumulative distribution functions and probability density functions, the other is based on ROC curve and AUC. Unal (2017) proposed the index of union (IU) as an accuracy measure that considers two types to get them. In this study, ten kinds of accuracy measures (including IU) are divided into six categories, and the advantages of the IU are studied by comparing the measures belonging to each category. The optimal thresholds of these measures are obtained by setting various normal mixture distributions; subsequently, the first and second type of errors as well as the error sums corresponding to each threshold are calculated. The properties and characteristics of the IU statistic are explored by comparing the discriminative power of other accuracy measures based on error values.The values of the first type error and error sum of IU statistic converge to those of the best accuracy measures of the second category as the mean difference between the two distributions increases. Therefore, IU could be an accuracy measure to evaluate the discriminant power of a model.

Keywords: accuracy, discriminant, sensitivity, specificity, threshold

## 1. Introduction

두 분포함수의 혼합분포로부터 판별력을 극대화하는 분류점(threshold, cut-off point)을 추정하는 방법은 신용평가와 의학통계분야 등에서 많이 연구되고 있다. 본 연구에서는 신용평가에서 차주(borrower)의 신용가치를 기준으로 대출상환능력에 따라 부도(default; $d$)와 정상(non-default; $n$) 상태를 판별하는 문제를 고려한다.

신용을 측정하는 확률변수 $X$는 연속형 변수이며, 부도와 정상 상태에 대해 전체 모수공간을 $\Theta = \{\theta_d, \theta_n\}$로 간주한다. 차주의 부도와 정상 상태에서 조건부 누적분포함수를 각각 $F_d(x) = P(X \leq x|\theta_d)$와 $F_n(x) = P(X \leq x|\theta_n)$이며, 누적분포함수 $F(x)$는 $F_d(x)$와 $F_n(x)$의 선형결합인 $F(x) = \gamma F_d(x) + (1-\gamma)F_n(x)$, 여기서 $\gamma$는 전체부도율(total probability of default)로 설정한다. $F_d(x)$와 $F_n(x)$는 확률 분류자(probability classifier)이며, 일반적으로 모든 $x$에 대하여 $F_d(x) \geq F_n(x)$을 가정한다 (Metz와 Kronman, 1980; Hsieh와 Turnbull, 1996; Provost와 Fawcett, 2001; Engelmann 등, 2003; Pepe, 2003; Fawcett, 2006).

Receiver operating characteristic (ROC) 곡선은 임의의 분류점에서 구한 비율 중, 부도를 부도로 정확하게 예측한 비율인 Sensitivity (민감도, true positive rate, hit rate, recall; $F_d(x)$)와 실제 정상

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

을 부도로 잘못 예측하는 비율인 1 − Specificity (1 − 특이도, false positive rate, false alarm rate; $F_n(x)$)를 각각 세로축과 가로축으로 갖는 좌표 위에 표현된다. 임의의 분류점 $x$에 대한 ROC 곡선의 좌표는 $(1 - \text{Specificity}, \text{Sensitivity}) = (u, \text{ROC}(u))$, 여기서 ROC 함수는 $\text{ROC}(u) = F_d(F_n^{-1}(u))$, $u \in [0, 1]$이다 (Egan, 1975; Spackman, 1989; Zweig와 Campbell, 1993; Bradley, 1997; Fawcett와 Provost, 1997; Pepe, 2000; Tasche, 2006; Hong과 Lee, 2018). 판별 모형의 성능을 평가하는 대표적인 측도 중 하나로 ROC 곡선 아래의 면적을 측정하는 area under the ROC curve (AUC)는 ROC 함수의 적분식으로 표현된다 (Hanley와 McNeil, 1982; Centor, 1991; Bradley, 1997; Krzanowski와 Hand, 2009; Engelmann 등, 2003; Vuk와 Curk, 2006).

최적분류점(optimal threshold)을 선택하는 대부분의 정확도 측도들은 크게 두 가지를 기반으로 정의하는데 우선 차주의 부도와 정상 상태의 누적분포함수 $F_d(x)$, $F_n(x)$와 이에 대응하는 확률밀도함수 $f_d(x)$, $f_n(x)$를 이용하여 Table 2.1과 같이 정의하는 방법과, ROC 곡선과 AUC를 이용하여 표현하는 방법으로 구분된다. 그러나 Unal (2017)은 두 방법을 혼합하여 누적분포함수 $F_d(x)$와 $F_n(x)$ 그리고 AUC를 모두 고려하는 새로운 정확도 측도 index of union (IU)를 식 (1.1)과 같이 제안하였다.

$$\begin{aligned} \text{IU} &= \min\{|\text{Sensitivity} - \text{AUC}| + |\text{Specificity} - \text{AUC}|\} \\ &= \min\{|F_d(x) - \text{AUC}| + |1 - F_n(x) - \text{AUC}|\}. \end{aligned} \tag{1.1}$$

Unal (2017)은 부도와 정상 상태를 정규분포와 감마분포에서 다양한 크기의 표본을 추출을 하여 IU 측도의 최적분류점을 구하였다. 그리고 세 종류의 정확도 측도인 Youden 지수($J$) (Youden, 1950), $(0, 1)$까지 최단거리기준(closest-to-$(0, 1)$ criterion; $C$) (Perkins와 Schisterman, 2006), 정확도면적(accuracy area; AA) (Brasil, 2010)에 대응하는 최적분류점을 구하고, IU 측도와 비교하기 위하여 상대 편차의 평균과 평균제곱오차의 평균을 살펴보았다.

본 연구는 Unal (2017)이 고려한 정확도 측도 이외에 Cantor 등 (1999), Greiner와 Gardner (2000), Freeman과 Moisen (2008), Liu 등 (2009), 그리고 Hong 등 (2011) 등의 다양한 정확도 측도들에 대하여 연구한다. 특히 Yoo와 Hong (2011)은 Unal (2017)의 연구에서 비교한 세 종류의 정확도 측도 외에 여섯 종류인 최대수직거리(maximum vertical distance; MVD) (Krzanowski와 Hand, 2009), 수정된 $(0, 1)$까지 최단거리기준(the amended closest-to-$(0, 1)$ criterion; AC) (Perkins과 Schisterman, 2006), 민감도와 특이도의 합(sum of sensitivity and specificity; SSS) (Connell과 Koepsell, 1985), 대칭점(symmetric point; SP) (Moses 등, 1993; Pepe, 2003), 전체정확도(total accuracy; TA) (Lambert와 Lipkovich, 2008), 진실율(true rate; TR) (Hong 등, 2010)을 추가하여 아홉 종류의 정확도 측도들을 IU 측도와 비교한다.

정확도 측도들을 비교하기 위하여 다양한 모평균과 모분산 그리고 전체부도율로 정의된 정규혼합분포(normal mixture)에서 각각의 측도들에 대응하는 최적분류점을 직접 구한다. 각 분류점에 대응하는 제1종 오류와 제2종 오류 그리고 제1종과 2종 오류의 합을 구하고, 오류들의 크기를 통해 정확도 측도들의 판별력을 비교한다. 이 결과를 바탕으로 새로운 정확도 측도인 IU 통계량의 성격을 파악하고, 기존의 다른 측도들에 비해 어떤 특징이 있는지 살펴본다.

본 연구의 2절에서는 IU 통계량을 포함한 열 종류의 정확도 측도들을 여섯 범주로 나누어 각 범주에 속한 정확도 측도의 성격과 조건을 설명한다. 3절에서는 다양한 정규혼합분포를 설정하여 부도와 정상분포의 분산의 크기에 대하여 세 가지 경우로 구분한다. 각각의 경우에 속하는 다양한 부도와 정상분포에서 여섯 범주들의 최적분류점을 구하고, 이에 대응하는 제1종과 제2종 오류들 그리고 오류합을 계산한다. 그리고 IU 통계량을 기준으로 여섯 범주에 대한 오류들의 크기를 비교하면서 IU 통계량의 장점을 탐색한다. 마지막으로 4절에서는 IU 통계량의 활용방법에 대해 논의하면서 결론을 유도한다.

**Table 2.1.** Conditions on optimal thresholds

| Category | Accuracy measure | Condition |
|---|---|---|
| 1 | closest-to-$(0,1)$ criterion | $f_d(x)/f_n(x) = F_n(x)/(1-F_d(x))$ |
| 2 | Youden index, True rate, Accuracy area, Maximum vertical distance, Sum of sensitivity and specificity | $f_d(x) = f_n(x)$ |
| 3 | Symmetric point | $F_n(x) = 1 - F_d(x)$ |
| 4 | Total accuracy | $f_d(x)/f_n(x) = (1-\gamma)/\gamma$ |
| 5 | Accuracy area | $f_d(x)/f_n(x) = F_d(x)/(1-F_n(x))$ |
| 6 | Index of Union | $\min\{\lvert F_d(x) - \mathrm{AUC}\rvert + \lvert 1 - F_n(x) - \mathrm{AUC}\rvert\}$ |

## 2. Accuracy measures

Hong 등 (2011)과 Yoo과 Hong (2011)은 최적분류점을 설정하는 아홉 종류의 정확도 측도들을 차주의 부도와 정상 상태의 누적분포함수 $F_d(x)$와 $F_n(x)$ 그리고 확률밀도함수 $f_d(x)$와 $f_n(x)$에 대한 조건으로 구분한 다섯 범주로 나누어 여러 가지 상황에서 비교 분석하였다. Unal (2017)의 IU 통계량을 포함한 열 종류의 분류정확도 측도들을 여섯 범주로 나누고, 각 범주에 대한 정확도 측도 조건과 이를 만족하는 측도의 종류들을 Table 2.1과 같이 정리하였다.

- **제1범주**에 속하는 정확도 측도는 '$(0,1)$까지 최단거리기준($C$)'이며, 오분류율 제곱의 합에 기반을 둔다.

$$C = \min\left\{(1-F_d(x))^2 + F_n(x)^2\right\} \quad \text{또는} \quad \min\sqrt{(1-F_d(x))^2 + F_n(x)^2}.$$

- **제2범주**에 속하는 정확도 측도는 'Youden 지수($J$)', '민감도와 특이도의 합(SSS)', '수정된$(0,1)$까지 최단거리기준(AC)', '진실율(TR)', '최대 수직거리(MVD)' 등이며, 두 누적분포의 차이를 기준으로 설정한다.

$$\begin{aligned} J &= \max\{F_d(x) - F_n(x)\}, \\ \mathrm{SSS} &= \max\{F_d(x) + (1 - F_n(x))\}, \\ \mathrm{AC} &= \min\{1 - F_d(x) + F_n(x)\}, \\ \mathrm{TR} &= \max\left\{\frac{1}{2}[F_d(x) + (1 - F_n(x))]\right\}, \\ \mathrm{MVD} &= \max|\mathrm{ROC}(u) - u|, \end{aligned}$$

  여기서 $\mathrm{ROC}(u) = F_d(F_n^{-1}(u))$. 제2범주를 대표하는 측도로 가장 많이 사용되는 $J$를 설정한다.

- **제3범주**에 속하는 정확도 측도는 '대칭점(SP)'이고 SP는 다음의 조건을 만족하는 $u$이다.

$$\mathrm{ROC}(u) = 1 - u.$$

  이 조건은 $F_d(x) = 1 - F_n(x)$ 또는 $F_d(x) + F_n(x) = 1$로 표현된다.

- **제4범주**에 속하는 정확도 측도는 '전체정확도 (TA)'이다. TA는 민감도와 특이도의 가중평균으로 다음과 같이 정의된다.

$$\mathrm{TA} = \max\left\{\frac{\mathrm{TP} + \mathrm{TN}}{p+n}\right\} = \max\{\gamma F_d(x) + (1-\gamma)(1 - F_n(x))\},$$

  여기서 전체부도율 $\gamma$의 추정량은 $p/(p+n)$이다.

- **제5범주**에 속하는 정확도 측도는 '정확도면적 (AA)'이며 AA는 민감도와 특이도의 곱을 최대로 하는 기준으로 다음과 같이 정의한다.

$$\mathrm{AA} = \max\{F_d(x)(1 - F_n(x))\}.$$

- **제6범주**에 속하는 정확도 측도는 IU 통계량으로 식 (1.1)과 같이 정의되며, 민감도와 특이도가 AUC와 유사한 값을 가지면서 동시에 최대가 되는 분류점을 최적분류점으로 간주한다.

본 연구에서 범주화된 정확도 측도들은 모두 누적분포함수인 $F_d(x)$와 $F_n(x)$로 정의되었으며, 조건들은 누적분포함수뿐만 아니라 확률밀도함수인 $f_d(x)$와 $f_n(x)$로 표현된다. 본 연구에서는 제1부터 5범주 까지의 정확도 측도들과 마지막 제6범주에서 정리한 IU 통계량을 오류의 관점에서 비교하기 위해, 다양한 정규혼합분포를 설정하여 각각의 최적분류점을 구하고 이에 대응하는 제1종 오류($\alpha$), 제2종 오류($\beta$) 그리고 오류합($\alpha + \beta$)을 비교한다.

## 3. Comparison with other accuracy measures

ROC 곡선의 가정인 부도와 정상 상태의 분포함수들이 모든 $x$에 대하여 $F_d(x) \geq F_n(x)$를 만족하면서, 부도 상태의 분포함수 $F_d(x)$를 $\mu_d = 0$이고, $\sigma_d^2 = 1$인 표준정규분포 $\Phi(x; 0, 1)$로, 정상 상태의 분포함수 $F_n(x)$는 $\mu_n > 0$이고, 다양한 $\sigma_n^2$의 값을 갖는 정규분포 $\Phi(x; \mu_n, \sigma_n^2)$로 설정한다. 신용을 측정하는 확률변수 $X$의 누적분포함수 $F(x)$를 다음과 같이 표현한다.

$$F(x) = \gamma\Phi\left(\frac{x - \mu_d}{\sigma_d}\right) + (1 - \gamma)\Phi\left(\frac{x - \mu_n}{\sigma_n}\right).$$

전체부도율 $\gamma \in (0, 1)$에 대하여 각 범주에서의 최적분류점을 구하고, 각각의 최적분류점에 대한 제1종 오류 ($\alpha$), 제2종 오류 ($\beta$) 그리고 오류합 ($\alpha + \beta$)을 구하여 제6범주인 IU 통계량을 중심으로 다른 범주의 통계량과 비교하여 IU 통계량의 특징을 비교한다.

본 절에서는 부도과 정상분포의 분산이 같은 경우 ($\sigma_d^2 = \sigma_n^2$)와 부도분포의 분산이 정상분포의 분산보다 큰 경우 ($\sigma_d^2 > \sigma_n^2$) 그리고 부도분포의 분산이 정상분포의 분산보다 작은 경우 ($\sigma_d^2 < \sigma_n^2$)로 나누어 탐색한다. 2절에서 논의한 여섯 범주의 정확도 측도들의 최적분류점과 이에 대응하는 오류들의 크기를 구한 결과를 Table 3.1부터 Table 3.3에 정리하고, Figure 3.1부터 Figure 3.6에 표현하였다. 이때 전체부도율 $\gamma$에 영향을 받는 제4범주(TA)는 $\gamma = 0.3551$ 이상의 값을 가질 때 최적분류점이 존재하기 때문에 (Yoo와 Hong, 2011, pp.349–351) $\gamma = 0.4$로 설정하고 각각의 $\alpha$, $\beta$, $\alpha + \beta$를 구하였다.

### 3.1. Case with equal variance of the default and non-default

3.1절에서는 정상분포의 분산을 1로 설정하여 부도와 정상분포의 분산이 동일한 경우 ($\sigma_d^2 = \sigma_n^2 = 1$)를 살펴본다. 본 연구에서 $F_d(x) \geq F_n(x)$를 가정하였고 부도분포 $F_d(x)$를 표준정규분포로 고정하고, 정상분포 $F_n(x)$는 다양한 평균을 갖는 분포로 다음과 같이 설정한다.

$$F_d(x) = \Phi(x : 0, 1), F_n(x) = \Phi(x : \mu_n, 1).$$

부도와 정상분포의 분산이 동일한 경우에 여섯 범주에 대한 최적분류점과 이에 대응하는 오류들 $\alpha$, $\beta$, $\alpha + \beta$을 구하고, Table 3.1에는 정상분포의 평균 $\mu_n = \{2, 3, 4\}$들에 대하여 정리하였다.

**Table 3.1.** Thresholds and their errors with $\sigma_d^2 = \sigma_n^2 (= 1)$and $\gamma = 0.4$

| Category | $\mu_n$ | Threshold | $\alpha$ | $\beta$ | $\alpha + \beta$ |
|---|---|---|---|---|---|
| 1 | 2 | 1.0000 | 0.1587 | 0.1587 | 0.3174 |
| 2 | | 1.0000 | 0.1587 | 0.1587 | 0.3174 |
| 3 | | 1.0000 | 0.1587 | 0.1587 | 0.3174 |
| 4 | | 0.7973 | 0.2126 | 0.1145 | 0.1145 |
| 5 | | 1.0000 | 0.1587 | 0.1587 | 0.3174 |
| 6 (IU) | | 1.0000 | 0.1587 | 0.1587 | 0.3174 |
| 1 | 3 | 1.5000 | 0.0668 | 0.0668 | 0.1336 |
| 2 | | 1.5000 | 0.0668 | 0.0668 | 0.1336 |
| 3 | | 1.5000 | 0.0668 | 0.0668 | 0.1336 |
| 4 | | 1.3648 | 0.0862 | 0.0510 | 0.1372 |
| 5 | | 1.5000 | 0.0668 | 0.0668 | 0.1336 |
| 6 (IU) | | 1.5000 | 0.0668 | 0.0668 | 0.1336 |
| 1 | 4 | 2.0000 | 0.0228 | 0.0228 | 0.0456 |
| 2 | | 2.0000 | 0.0228 | 0.0228 | 0.0456 |
| 3 | | 2.0000 | 0.0228 | 0.0228 | 0.0456 |
| 4 | | 1.8986 | 0.0288 | 0.0178 | 0.0466 |
| 5 | | 2.0000 | 0.0228 | 0.0228 | 0.0456 |
| 6 (IU) | | 2.0000 | 0.0228 | 0.0228 | 0.0456 |

제4범주 (TA)를 제외한 모든 측도들 (제1범주 ($C$), 2범주 ($J$), 3범주 (SP), 5범주 (AA), 6범주 (IU))은 정상과 부도분포의 평균들의 중간 지점인 $\mu_n/2$에서 최적분류점을 가지며, 이 분류점에 대응하는 $\alpha$와 $\beta$는 동일한 값을 가지므 오류합은 $\alpha$또는 $\beta$의 두 배로 나타난다. 또한 $\mu_n$이 증가할수록 두 분포간의 거리가 멀어져 모든 오류들이 감소하는 것을 탐색할 수 있다. 전체부도율이 $\gamma = 0.5$가 되면, 제4범주 (TA)가 제2범주 ($J$, TR)의 조건과 같아지므로 제4범주 (TA)는 곧 제2범주 ($J$)와 같은 최적분류점과 오류값을 갖게 되어 모든 측도들의 최적분류점과 오류들이 동일함을 발견하였다.

$$\alpha_{\mathrm{IU}} = \alpha_C = \alpha_J = \alpha_{\mathrm{SP}} = \alpha_{\mathrm{AA}} \leq \alpha_{\mathrm{TA}},$$
$$\beta_{\mathrm{IU}} = \beta_C = \beta_J = \beta_{\mathrm{SP}} = \beta_{\mathrm{AA}} \geq \beta_{\mathrm{TA}},$$
$$(\alpha+\beta)_{\mathrm{IU}} = (\alpha+\beta)_C = (\alpha+\beta)_J = (\alpha+\beta)_{\mathrm{SP}} = (\alpha+\beta)_{\mathrm{AA}} \leq (\alpha+\beta)_{\mathrm{TA}}.$$

따라서 부도 및 정상분포의 분산이 같은 경우 ($\sigma_d^2 = \sigma_n^2$), 제6범주 (IU)는 제1범주 ($C$), 2범주 ($J$), 3범주 (SP), 5범주 (AA)와 $\alpha$, $\beta$ 그리고 $\alpha + \beta$가 동일하다. 그리고 제6범주 (IU)는 제4범주 (TA)보다 작은 제1종 오류와 오류합을 가지며, 제4범주 (TA)의 제2종 오류보다 크다. 그러므로 IU 통계량은 두 분포의 분산이 같은 경우에는 제4범주 (TA)를 제외한 모든 통계량들과 같은 제1종과 2종오류, 오류합을 가지며 특히 제1종 오류와 오류합은 최소값을 갖는다.

### 3.2. Case when the variance of the default is greater than the non-default

본 절에서는 정상분포의 분산 $\sigma_n^2$을 $0.5^2$과 $0.75^2$으로 설정하고, 부도 분포의 분산이 정상분포보다 큰 일반적인 경우 ($\sigma_d^2 > \sigma_n^2$)를 고려한다. 3.1절에서와 유사하게, 부도분포는 표준정규분포로 고정하고, 정상분포는 다양한 평균을 갖는 분포로 다음과 같다.

$$F_d(x) = \Phi(x : 0, 1), F_n(x) = \Phi(x : \mu_n, \sigma_n^2).$$

**Table 3.2.** Thresholds and their errors with $\sigma_d^2 > \sigma_n^2$and $\gamma = 0.4$

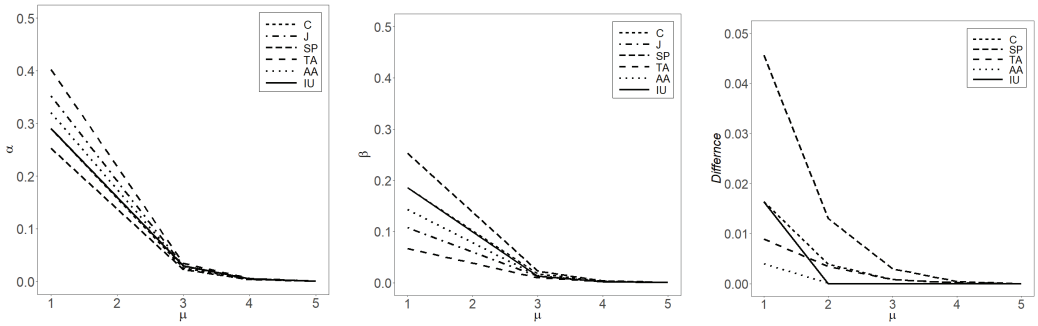
| Category | $\mu_n$ | $\sigma_n^2 = 0.5^2$ | | | | $\sigma_n^2 = 0.75^2$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Threshold | $\alpha$ | $\beta$ | $\alpha+\beta$ | Threshold | $\alpha$ | $\beta$ | $\alpha+\beta$ |
| 1 | 2 | 1.2611 | 0.1036 | 0.0697 | 0.1733 | 1.0987 | 0.1360 | 0.1147 | 0.2507 |
| 2 | | 1.1694 | 0.1211 | 0.0483 | 0.1694 | 1.0361 | 0.1501 | 0.0994 | 0.2495 |
| 3 | | 1.3333 | 0.0912 | 0.0912 | 0.1824 | 1.1429 | 0.1265 | 0.1265 | 0.2530 |
| 4 | | 1.0818 | 0.1397 | 0.0331 | 0.1728 | 0.8920 | 0.1862 | 0.0698 | 0.2560 |
| 5 | | 1.1864 | 0.1177 | 0.0518 | 0.1695 | 1.0545 | 0.1458 | 0.1037 | 0.2495 |
| 6 (IU) | | 1.1694 | 0.1211 | 0.0483 | 0.1694 | 1.0361 | 0.1501 | 0.0994 | 0.2495 |
| 1 | 3 | 1.9477 | 0.0257 | 0.0177 | 0.0434 | 1.6821 | 0.0463 | 0.0394 | 0.0857 |
| 2 | | 1.8857 | 0.0297 | 0.0129 | 0.0426 | 1.6408 | 0.0504 | 0.0350 | 0.0854 |
| 3 | | 2.0000 | 0.0228 | 0.0228 | 0.0456 | 1.7143 | 0.0432 | 0.0432 | 0.0864 |
| 4 | | 1.8242 | 0.0341 | 0.0093 | 0.0434 | 1.5442 | 0.0613 | 0.0261 | 0.0874 |
| 5 | | 1.8905 | 0.0293 | 0.0132 | 0.0425 | 1.6466 | 0.0498 | 0.0356 | 0.0854 |
| 6 (IU) | | 1.8857 | 0.0297 | 0.0129 | 0.0426 | 1.6408 | 0.0504 | 0.0350 | 0.0854 |
| 1 | 4 | 2.6259 | 0.0043 | 0.0030 | 0.0073 | 2.2606 | 0.0119 | 0.0102 | 0.0221 |
| 2 | | 2.5788 | 0.0050 | 0.0022 | 0.0072 | 2.2307 | 0.0129 | 0.0092 | 0.0221 |
| 3 | | 2.6667 | 0.0038 | 0.0038 | 0.0076 | 2.2857 | 0.0111 | 0.0111 | 0.0222 |
| 4 | | 2.5322 | 0.0057 | 0.0017 | 0.0074 | 2.1550 | 0.0156 | 0.0069 | 0.0225 |
| 5 | | 2.5840 | 0.0049 | 0.0023 | 0.0072 | 2.2332 | 0.0128 | 0.0092 | 0.0220 |
| 6 (IU) | | 2.5788 | 0.0050 | 0.0022 | 0.0072 | 2.2307 | 0.0129 | 0.0092 | 0.0221 |

Table 3.2에 정상분포의 평균 $\mu_n = \{2, 3, 4\}$에서의 여섯 범주에 대한 최적분류점과 그에 따른 오류들 $\alpha$, $\beta$, $\alpha+\beta$을 정리했으며 Figures 3.1과 3.2에는 $\mu_n = \{1, 2, 3, 4, 5\}$에서의 모든 오류들을 표현하였다.

Table 3.2를 살펴보면, 제6범주 (IU)의 최적분류점은 정상분포의 평균 $\mu_n$과 분산 $\sigma_n^2$의 크기와 관계없이 제4범주 (TA)의 최적분류점보다는 크고, 제5범주 (AA)의 최적분류점보다는 작으며 제2범주 ($J$)의 최적분류점과는 동일한 최적분류점을 갖는다 ($x_{\mathrm{TA}} < x_{\mathrm{IU}} = x_J < x_{\mathrm{AA}} < x_c < x_{\mathrm{SP}}$). 따라서 제6범주 (IU)의 최적분류점에 대응하는 $\alpha$는 최적분류점의 크기 비교와 반대로 제5범주 (AA)의 $\alpha$보다는 크고, 제4범주 (TA)의 $\alpha$보다는 작으며 제2범주 ($J$)의 $\alpha$와는 동일한 값을 갖는다. 제6범주 (IU)와 제2범주 ($J$)를 제외한 다른 범주들의 $\alpha$ 또한 정상분포의 평균 $\mu_n$이 커질수록 제6범주 (IU)와 제2범주 ($J$)의 $\alpha$값에 근접하므로 여섯 범주의 제1종 오류 ($\alpha$)에 대한 크기 순위 $\alpha_{\mathrm{TA}} \geq \alpha_{\mathrm{IU}} = \alpha_J \geq \alpha_{\mathrm{AA}} \geq \alpha_C \geq \alpha_{\mathrm{SP}}$이다.

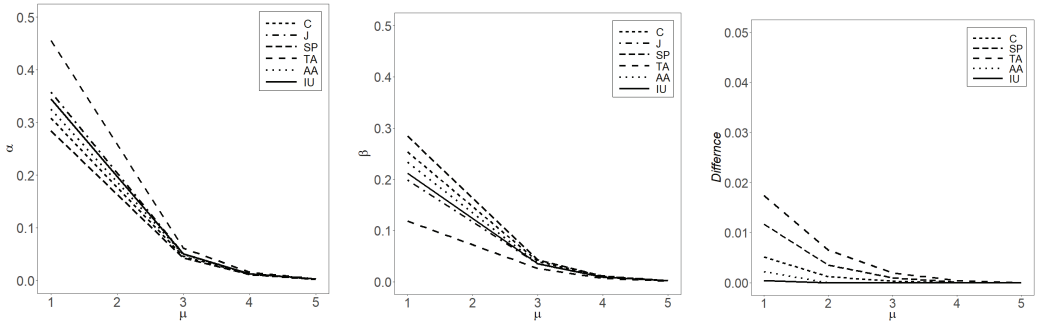
다음에는 Table 3.2에서 제2종 오류 ($\beta$)에 대하여 살펴보자. 제6범주 (IU)의 $\beta$는 최적분류점의 크기 비교와 마찬가지로 제4범주 (TA)의 $\beta$보다는 크고, 제5범주 (AA)의 $\beta$보다는 작으며 제2범주 ($J$)의 $\beta$와는 동일한 값을 갖는다. 정상분포의 평균 $\mu_n$이 커질수록 제6범주 (IU)와 제2범주 ($J$)를 제외한 다른 범주들의 $\beta$ 또한 제6범주 (IU)와 제2범주 ($J$)의 $\beta$값에 가까운 값을 가지며 여섯 범주의 제2종 오류 ($\beta$)에 대한 크기 순위는 $\beta_{\mathrm{TA}} \leq \beta_{\mathrm{IU}} = \beta_J \leq \beta_{\mathrm{AA}} \leq \beta_C \leq \beta_{\mathrm{SP}}$이다.

마지막으로 Table 3.2에서 오류합 ($\alpha+\beta$)에 대하여 살펴보면 다음과 같다. 제6범주 (IU)의 오류합은 제2범주 ($J$)를 제외한 모든 범주의 오류합보다는 작고, 제2범주 ($J$)와는 동일한 오류합을 가져 제2범주 ($J$)와 함께 오류합에 대한 최소 범주임을 알 수 있다. 정상분포의 평균 $\mu_n$이 커질수록 제6범주 (IU)와 제2범주 ($J$)를 제외한 다른 범주들의 오류합 또한 제6범주 (IU)와 제2범주 ($J$)의 값에 매우 가까운 값을 가지며 여섯 범주의 오류합 ($\alpha+\beta$)에 대한 크기 순위는 $(\alpha+\beta)_{\mathrm{IU}} = (\alpha+\beta)_J \approx (\alpha+\beta)_{\mathrm{AA}} \leq (\alpha+\beta)_C \leq (\alpha+\beta)_{\mathrm{SP}} \leq (\alpha+\beta)_{\mathrm{TA}}$이다.




(a) (b) (c)

**Figure 3.1.** $\alpha, \beta, \alpha+\beta$ with $\sigma_d^2 > \sigma_n^2 = 0.5^2$.




(a) (b) (c)

**Figure 3.2.** $\alpha, \beta, \alpha+\beta$ with $\sigma_d^2 > \sigma_n^2 = 0.75^2$.

Figures 3.1과 3.2는 정상분포의 분산 $\sigma_n^2$이 각각 $0.5^2$과 $0.75^2$인 경우에 대하여 여섯 범주에 대한 $\alpha$, $\beta$ 의 크기를 정상분포의 평균 $\mu_n$의 변화에 따라 표현하였으며, 오류합 $\alpha+\beta$에 대해서는 제2범주 $(J)$와 다른 범주들 간의 차이를 구현하였다. 정상분포의 평균의 범위를 Table 3.2보다 확대하여 $\mu_n = \{1, 2, 3, 4, 5\}$에 대하여 살펴보았다. Figures 3.1과 3.2의 왼쪽과 가운데 그림들을 통해 정상분포의 분산 $\sigma_n^2$이 $0.5^2$인 경우보다 $\sigma_n^2$이 $0.75^2$인 경우에서 모든 정상분포의 평균 $\mu_n$의 모든 구간에 대한 제6범주 (IU)의 $\alpha$, $\beta$가 제2범주 $(J)$의 $\alpha$, $\beta$에 더 가까운 것을 확인할 수 있다. 또한 Figures 3.1과 3.2의 오른쪽 그림들을 통해 정상분포의 분산 $\sigma_n^2$이 $0.5^2$인 경우에는 정상분포의 평균 $\mu_n$이 2일 때 제2범주 $(J)$와 제6범주 (IU)의 차이가 0에 수렴하는 반면, $\sigma_n^2$이 $0.75^2$인 경우에는 대부분의 정상분포의 평균 $\mu_n$ 구간에서 두 범주의 $\alpha+\beta$ 차이가 0으로 수렴하는 것을 볼 수 있다. 따라서 정상분포의 분산 $\sigma_n^2$이 커질수록 모든 정상분포의 평균 $\mu_n$ 구간에서 제6범주 (IU)의 $\alpha$, $\beta$가 제2범주 $(J)$의 $\alpha$, $\beta$에 더 근접한 값을 가지며, 두 범주의 $\alpha+\beta$ 차이가 빠르게 감소하는 것을 탐색할 수 있다. 이러한 현상을 통해 부도분포의 분산이 정상분포보다 큰 경우 $(\sigma_d^2 > \sigma_n^2)$, 정상분포의 분산 $\sigma_n^2$이 커질수록 제6범주 (IU)에 대한 오류들의 값이 제2범주 $(J)$에 대한 오류들의 값에 더 빨리 수렴하는 것을 파악할 수 있다.

$$\alpha_{\mathrm{TA}} \geq \alpha_{\mathrm{IU}} \approx \alpha_J \geq \alpha_{\mathrm{AA}} = \alpha_C \geq \alpha_{\mathrm{SP}},$$
$$\beta_{\mathrm{TA}} \leq \beta_{\mathrm{IU}} \approx \beta_J \leq \beta_{\mathrm{AA}} \leq \beta_C \leq \beta_{\mathrm{SP}},$$
$$(\alpha+\beta)_{\mathrm{IU}} \approx (\alpha+\beta)_J \approx (\alpha+\beta)_{\mathrm{AA}} \leq (\alpha+\beta)_C \leq (\alpha+\beta)_{\mathrm{SP}} \leq (\alpha+\beta)_{\mathrm{TA}}.$$

**Table 3.3.** Thresholds and their errors with $\sigma_d^2 < \sigma_n^2$and $\gamma = 0.4$

| Category | $\mu_n$ | $\sigma_n^2 = 1.25^2$ | | | | $\sigma_n^2 = 1.5^2$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Threshold | $\alpha$ | $\beta$ | $\alpha+\beta$ | Threshold | $\alpha$ | $\beta$ | $\alpha+\beta$ |
| 1 | 2 | 0.9421 | 0.1731 | 0.1987 | 0.3718 | 0.9105 | 0.1813 | 0.2338 | 0.4151 |
| 2 | | 1.0253 | 0.1526 | 0.2178 | 0.3704 | 1.0847 | 0.1390 | 0.2709 | 0.4099 |
| 3 | | 0.8889 | 0.1870 | 0.1870 | 0.3740 | 0.8000 | 0.2119 | 0.2119 | 0.4238 |
| 4 | | 0.7726 | 0.2199 | 0.1631 | 0.3830 | 0.7990 | 0.3121 | 0.2117 | 0.4238 |
| 5 | | 0.9904 | 0.1610 | 0.2096 | 0.3706 | 1.0048 | 0.1575 | 0.2535 | 0.4110 |
| 6 (IU) | | 1.0253 | 0.1526 | 0.2178 | 0.3704 | 1.0847 | 0.1390 | 0.2709 | 0.4099 |
| 1 | 3 | 1.3730 | 0.0849 | 0.0965 | 0.1814 | 1.2834 | 0.0997 | 0.1262 | 0.2259 |
| 2 | | 1.4236 | 0.0773 | 0.1036 | 0.1809 | 1.3965 | 0.0813 | 0.1425 | 0.2238 |
| 3 | | 1.3333 | 0.0912 | 0.0912 | 0.1824 | 1.2000 | 0.1151 | 0.1151 | 0.2302 |
| 4 | | 1.2559 | 0.1046 | 0.0815 | 0.1861 | 1.1995 | 0.1152 | 0.1150 | 0.2302 |
| 5 | | 1.4148 | 0.0786 | 0.1024 | 0.1810 | 1.3689 | 0.0855 | 0.1384 | 0.2239 |
| 6 (IU) | | 1.4236 | 0.0773 | 0.1036 | 0.1809 | 1.3961 | 0.0813 | 0.1425 | 0.2238 |
| 1 | 4 | 1.8089 | 0.0352 | 0.0398 | 0.0750 | 1.6664 | 0.0478 | 0.0599 | 0.1077 |
| 2 | | 1.8445 | 0.0326 | 0.0423 | 0.0749 | 1.7475 | 0.0403 | 0.0666 | 0.1069 |
| 3 | | 1.7778 | 0.0377 | 0.0377 | 0.0754 | 1.6000 | 0.0548 | 0.0548 | 0.1096 |
| 4 | | 1.7210 | 0.0426 | 0.0341 | 0.0767 | 1.5999 | 0.0548 | 0.0548 | 0.1096 |
| 5 | | 1.8445 | 0.0326 | 0.0423 | 0.0749 | 1.7403 | 0.0409 | 0.0660 | 0.1069 |
| 6 (IU) | | 1.8445 | 0.0326 | 0.0423 | 0.0749 | 1.7475 | 0.0403 | 0.0666 | 0.1069 |

그러므로 부도분포의 분산이 정상분포보다 큰 경우 ($\sigma_d^2 > \sigma_n^2$)에는 제6범주 (IU)의 제1종 오류 ($\alpha$)와 제2종 오류 ($\beta$) 그리고 오류합 ($\alpha+\beta$)의 크기가 최소는 아니지만, 정상분포의 평균 $\mu_n$이 증가할수록 제6범주 (IU)의 제1종 오류가 제2범주 ($J$)의 제1종 오류에 수렴하는 것을 발견하였으며, 제6범주 (IU)의 오류합 또한 제2범주 ($J$)의 오류합에 수렴하여 최소 오류합을 갖는 정확도 측도는 제6범주의 IU 통계량임을 확인하였다. 또한 정상분포의 분산 $\sigma_n^2$이 커질수록 제6범주 (IU)에 대한 오류들의 값이 제2범주 ($J$)에 대한 오류들의 값에 더 빠르게 수렴하는 것을 발견하였다.

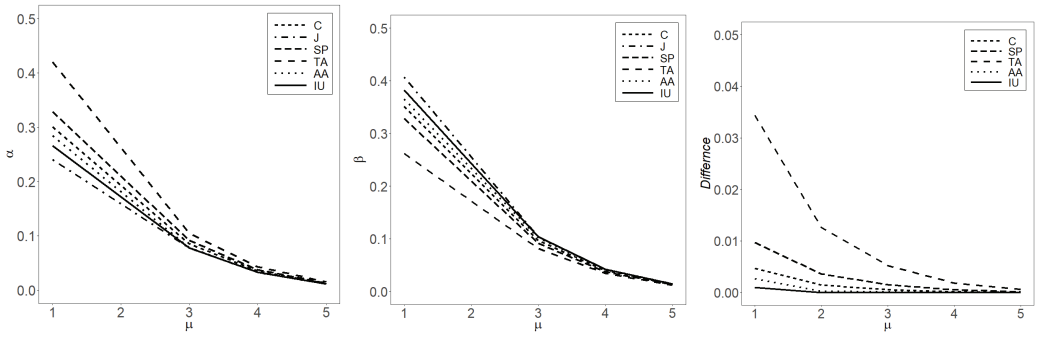
### 3.3. Case when the variance of the non-default is greater than the default

정상분포의 분산 $\sigma_n^2$을 $1.25^2$과 $1.5^2$인 정규분포와 일반적인 현상은 아니지만, 부도분포의 분산이 정상분포보다 작은 ($\sigma_d^2 < \sigma_n^2$) 정규분포를 다음과 같이 설정한다.

$$F_d(x) = \Phi(x : 0, 1), F_n(x) = \Phi(x : \mu_n, \sigma_n^2),$$

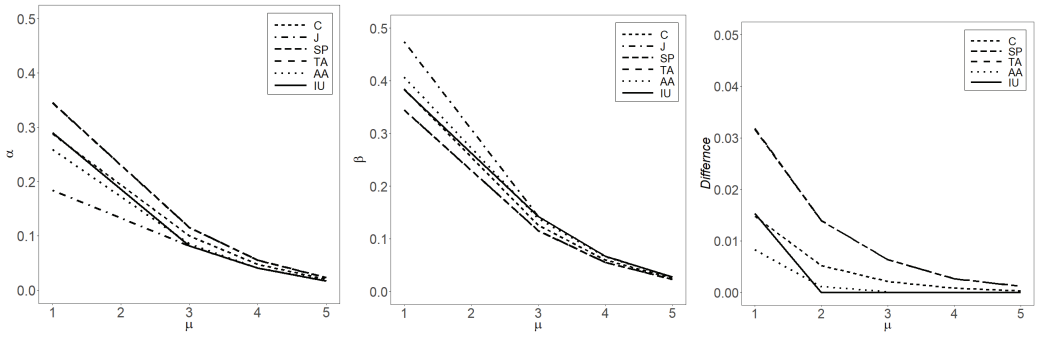
정상분포의 다양한 평균일 때의 여섯 범주에 대한 최적분류점과 이에 대응하는 오류들 $\alpha$, $\beta$, $\alpha+\beta$을 구하고, Table 3.3에는 $\mu_n = \{2, 3, 4\}$만을 정리하였으며 Figures 3.3과 3.4에서는 $\mu_n = \{1, 2, 3, 4, 5\}$에서의 모든 오류들을 표현하였다.

Table 3.3에서 제6범주 (IU)의 최적분류점은 정상분포의 평균 $\mu_n$과 분산 $\sigma_n^2$의 크기와 관계없이 제2범주 ($J$)를 제외한 다른 범주들의 최적분류점보다는 크고, 제2범주 ($J$)와는 동일한 최적분류점을 갖는다 ($x_{\mathrm{IU}} = x_J > x_{\mathrm{AA}} > x_C > x_{\mathrm{SP}} > x_{\mathrm{TA}}$). 따라서 제6범주 (IU)의 최적분류점에 대응하는 $\alpha$는 최적분류점의 크기 비교와 반대로 제2범주 ($J$)를 제외한 다른 범주들의 $\alpha$보다는 작고, 제2범주 ($J$)의 $\alpha$와는 동일한 $\alpha$값을 갖는다. 제6범주 (IU)와 제2범주 ($J$)를 제외한 다른 범주들의 $\alpha$ 또한 정상분포의 평균 $\mu_n$이 커질수록 제6범주 (IU)와 제2범주 ($J$)의 값에 가까운 $\alpha$값을 가지며 여섯 범주의 제1종 오류 ($\alpha$)에 대한 크기 순위는 ($\alpha_{\mathrm{IU}} = \alpha_J \leq \alpha_{\mathrm{AA}} \leq \alpha_C \leq \alpha_{\mathrm{SP}} \leq \alpha_{\mathrm{TA}}$)가 된다.




(a) (b) (c)

**Figure 3.3.** $\alpha, \beta, \alpha+\beta$ with $\sigma_d^2 < \sigma_n^2 = 1.25^2$.




(a) (b) (c)

**Figure 3.4.** $\alpha, \beta, \alpha+\beta$ with $\sigma_d^2 < \sigma_n^2 = 1.5^2$.

다음으로 Table 3.3에서 제2종 오류 ($\beta$)에 대하여 살펴보면, 제6범주 (IU)의 $\beta$는 최적분류점의 크기 비교와 마찬가지로 제2범주 ($J$)를 제외한 다른 범주들의 최적분류점보다는 크고, 제2범주 ($J$)의 $\beta$와는 동일한 값을 갖는다. 제6범주 (IU)와 제2범주 ($J$)를 제외한 다른 범주들의 $\beta$도 정상분포의 평균 $\mu_n$이 커질수록 제6범주 (IU)와 제2범주 ($J$)의 $\beta$값에 근접하여 여섯 범주의 제2종 오류 ($\beta$)에 대한 크기 순위는 ($\beta_{\mathrm{IU}} = \beta_J \geq \beta_{\mathrm{AA}} \geq \beta_C \geq \beta_{\mathrm{SP}} \geq \beta_{\mathrm{TA}}$)이다.

Table 3.3에서 오류합 ($\alpha+\beta$)에 대하여 살펴보면 다음과 같다. 제6범주 (IU)의 오류합은 $\alpha$의 크기 비교와 마찬가지로 제2범주 ($J$)를 제외한 다른 범주들의 오류합보다는 작고, 제2범주 ($J$)의 오류합과는 동일한 값을 갖는다. 제2범주 ($J$)를 제외한 다른 범주들의 오류합도 정상분포의 평균 $\mu_n$이 커질수록 제6범주 (IU)와 제6범주 (IU)와 제2범주 ($J$)의 오류합에 가까운 값을 가지므로 여섯 범주의 오류합 ($\alpha+\beta$)에 대한 크기 순위는 ($(\alpha+\beta)_{\mathrm{IU}} = (\alpha+\beta)_J \leq (\alpha+\beta)_{\mathrm{AA}} \leq (\alpha+\beta)_C \leq (\alpha+\beta)_{\mathrm{SP}} \leq (\alpha+\beta)_{\mathrm{TA}}$)가 된다.

Figures 3.3과 3.4는 Table 3.3의 결과를 여섯 범주에 대한 $\alpha$, $\beta$의 크기를 정상분포의 평균 $\mu_n$이 1부터 5까지의 변화에 따라 표현하였으며, 오류합 ($\alpha+\beta$)에 대해서는 제2범주와 다른 범주들 간의 차이를 구현하였다. Figures 3.3과 3.4의 왼쪽과 가운데 그림들을 통해 정상분포의 분산 $\sigma_n^2$이 $1.5^2$인 경우보다 $\sigma_n^2$이 $1.25^2$인 경우에서 모든 정상분포의 평균 $\mu_n$ 구간에 대한 제6범주 (IU)의 $\alpha$, $\beta$가 제2범주 ($J$)의 $\alpha$, $\beta$에 더 가까운 것을 확인할 수 있다. 또한 Figures 3.3과 3.4의 오른쪽 그림들을 통해 정상분포의 분

산 $\sigma_n^2$이 $1.5^2$인 경우에는 정상분포의 평균 $\mu_n$이 2일 때 제2범주 ($J$)와 제6범주 (IU)의 오류합 차이가 0에 수렴하는 반면, $\sigma_n^2$이 $1.25^2$인 경우에는 정상분포의 평균 $\mu_n$의 전체 구간에서 두 범주의 오류합 차이가 0으로 수렴하는 것을 볼 수 있다. 따라서 정상분포의 분산 $\sigma_n^2$이 작아질수록 모든 정상분포의 평균 $\mu_n$ 구간에서 제6범주 (IU)의 $\alpha$, $\beta$가 제2범주 ($J$)의 $\alpha$, $\beta$에 더 근접한 값을 가지며, 두 범주의 오류합 차이가 빠르게 감소하는 것을 탐색할 수 있다. 이러한 현상을 통해 부도분포의 분산이 정상분포보다 작은 경우 ($\sigma_d^2 < \sigma_n^2$) 정상분포의 분산 $\sigma_n^2$이 작아질수록 제6범주 (IU)에 대한 오류들의 값이 제2범주 ($J$)에 대한 오류들의 값에 더 빨리 수렴하는 것을 알 수 있다.

$$\alpha_{\mathrm{IU}} \approx \alpha_J \le \alpha_{\mathrm{AA}} \le \alpha_C \le \alpha_{\mathrm{SP}} \le \alpha_{\mathrm{TA}},$$
$$\beta_{\mathrm{IU}} \approx \beta_J \ge \beta_{\mathrm{AA}} \ge \beta_C \ge \beta_{\mathrm{SP}} \ge \beta_{\mathrm{TA}},$$
$$(\alpha+\beta)_{\mathrm{IU}} \approx (\alpha+\beta)_J \le (\alpha+\beta)_{\mathrm{AA}} \le (\alpha+\beta)_C \le (\alpha+\beta)_{\mathrm{SP}} \le (\alpha+\beta)_{\mathrm{TA}}.$$

그러므로 부도분포의 분산이 정상분포보다 작은 경우 ($\sigma_d^2 < \sigma_n^2$)에는 제6범주 (IU)의 제1종 오류 ($\alpha$)와 제2종 오류 ($\beta$) 그리고 오류합 ($\alpha+\beta$)의 크기가 최소는 아니지만, 정상분포의 평균 $\mu_n$이 증가할수록 제6범주 (IU)의 제1종오류와 오류합이 제2범주 ($J$)의 제1종 오류와 오류합에 수렴하므로 제1종 오류와 오류합에 대한 최소 범주가 되는 것을 확인하였다. 또한 정상분포의 분산 $\sigma_n^2$이 작아질수록 제6범주 (IU)에 대한 오류들의 값이 제2범주 ($J$)에 대한 오류들의 값에 더 빠르게 수렴하는 것을 발견하였다.

전체부도율이 $\gamma = 0.4$일 때 3절의 결과를 제6범주 IU 통계량을 중심으로 살펴보면, 두 분포의 분산이 같은 경우 ($\sigma_d^2 = \sigma_n^2 = 1$)에는 제6범주 (IU)가 제4범주 (TA)를 제외한 나머지 범주 (제1범주 (C), 2범주 ($J$), 3범주 (SP), 5범주 (AA))와 함께 제1종 오류와 오류합에 대하여 최소값을 갖는 범주가 된다.

3.2절과 3.3절의 결과를 바탕으로 두 분포의 분산이 다른 경우 ($\sigma_d^2 > \sigma_n^2$와 $\sigma_d^2 < \sigma_n^2$), 제6범주 (IU)의 제1종 오류와 오류합은 정상분포의 평균 $\mu_n$이 증가할수록 제2범주 ($J$)의 제1종 오류와 오류합에 근접하는 것을 발견하였다. 이 경우 제6범주 (IU)는 2범주 ($J$)와 함께 오류합에 대하여 최소값을 갖는 범주가 되며, 특히 부도분포의 분산이 정상분포보다 작은 경우 ($\sigma_d^2 < \sigma_n^2$)에는 제1종 오류도 2범주 ($J$)와 함께 가장 작은 값을 갖는 것을 확인하였다.

또한 각 경우에서 정상분포의 분산 $\sigma_n^2$ 크기에 따른 제6범주 (IU)에 대한 오류들을 살펴보면, 부도분포의 분산이 정상분포보다 큰 경우 ($\sigma_d^2 > \sigma_n^2$)에는 정상분포의 분산 $\sigma_n^2$이 커질수록 제6범주 (IU)에 대한 오류들의 값이 제2범주 ($J$)에 대한 오류들의 값에 더 빠르게 수렴하고, 부도분포의 분산이 정상분포보다 작은 경우 ($\sigma_d^2 < \sigma_n^2$)에는 이 $\sigma_n^2$작아질수록 제6범주 (IU)에 대한 오류들의 값이 제2범주 ($J$)에 대한 오류들의 값에 더 빠르게 수렴하여 제2범주 ($J$)와 함께 가장 작은 오류합을 갖는 것을 발견하였다.

## 4. Conclusion

Unal (2017)은 새롭게 제안한 정확도 측도인 Index of Union (IU)의 성능을 세 종류의 정확도 측도들 (Youden 지수, $(0,1)$까지 최단거리기준, 정확도면적)의 성능과 비교하기 위해 부도와 정상분포를 정규분포로 설정하고, 표본추출한 자료에 대하여 최적분류점들을 각각 구하였다. 각각의 최적분류점에 대응하는 상대 편차의 평균과 평균제곱오차의 평균 크기 비교를 통해 상대 편차와 평균제곱오차의 관점에서 IU 통계량이 $(0,1)$까지 최단거리기준과 더불어 가장 좋은 측도임을 보였다. 나아가 AUC에 가장 가까운 최대 민감도(sensitivity)와 특이도(specificity)를 찾음으로써 오분류율을 최소화시킨다는 측면에서 IU가 $(0,1)$까지 최단거리기준보다 의미 있는 측도임을 강조하였다.

본 연구에서는 Hong 등 (2011)이 정리한 아홉 종류의 정확도 측도들과 IU 통계량을 모두 여섯 범주로 구분하여 비교 연구하였다. 다양한 정규혼합분포를 설정하여 IU 통계량이 포함된 여섯 범주들의 최적

분류점을 각각 구하고, 이에 대응하는 제1종과 2종 오류 그리고 두 종류의 오류합을 비교하였다. 또한 다양한 분포 상황에서 제6범주인 IU 통계량에 대한 세 종류의 오류들이 제2범주의 대표 측도인 Youden 지수에 대한 세 종류의 오류들과 동일한 값으로 수렴하기까지의 속도를 비교하여 IU 통계량의 장점을 탐색하였다.

본 연구는 오류의 관점에서 부도와 정상분포에 대한 분산 크기에 따라 IU 통계량의 특징을 정리하였다. 먼저 두 분포의 분산이 같은 경우에 IU 통계량은 정상분포의 평균과 관계없이 전체부도율에 영향을 받는 제4범주를 제외한 모든 측도들과 동일한 제1종과 2종오류, 오류합을 가지며, 특히 제1종 오류와 오류합에 대해 항상 최소값을 갖는다. 다음으로 부도분포의 분산이 정상분포보다 큰 경우에는 IU 통계량의 제1종 오류와 오류합은 정상분포의 평균이 증가할수록 Youden 지수의 제1종 오류와 오류합에 수렴하여 제2범주와 같이 최소의 오류합을 가지며, IU 통계량의 모든 오류들은 정상분포의 분산이 커질수록 Youden 지수의 오류들과 동일한 값으로 수렴하는 속도가 더 빨라지는 것을 발견하였다. 마지막으로 부도분포의 분산이 정상분포보다 작은 경우 IU 통계량의 제1종 오류와 오류합은 정상분포의 평균이 증가할수록 Youden 지수의 제1종 오류와 오류합에 수렴하여 제2범주와 같이 최소의 오류합을 가지며, IU 통계량의 모든 오류들은 정상분포의 분산이 작아질수록 Youden 지수의 오류들과 동일한 값으로 더 빠르게 수렴하는 것을 확인하였다.

본 연구에서는 오류의 관점에서 IU 통계량이 Youden 지수와 더불어 대부분의 경우에서 제1종의 오류와 오류합이 최소임을 보였다. 그러므로 IU 통계량이 다른 정확도 측도들보다 효율적인 특성을 갖는다고 판단할 수 있으며, 다양한 분포 상황에서 IU 통계량을 포함한 여섯 범주의 최적분류점과 이에 대응하는 세 종류의 오류의 크기 순서를 바탕으로 분석할 자료의 상황에 적합한 최적의 분류정확도를 사용하여 제1종과 제2종 오류와 오류합을 최소화함으로써 모형의 판별력 향상을 기대할 수 있다.

3차원 ROC 곡면(surface)과 volume under the ROC surface (VUS) 그리고 4차원 이상으로 확장한 ROC 다면체(manifold)와 hypervolume under the ROC manifold (HUM)를 이용하여, 이차원에서의 IU를 $k$차원으로 확장한 IU 통계량 ($IU^k$)은 다음과 같이 정의할 수 있다.

$$\mathrm{IU}^k = \sum_{i=1}^{k} |(F_i(x_{o,i+1}) - F_i(x_{o,i})) - \mathrm{HUM}|, \tag{4.1}$$

여기서 $x_{o,i}$는 두 순차적인 분포함수 $F_{i-1}(x)$와 $F_i(x)$의 분류점이며, $x_{o,1} = -\infty$이고, $x_{o,k+1} = \infty$이다. 또한 3차원 IU 통계량은 $\mathrm{IU}^3 = |F_1(x_{o,2}) - \mathrm{VUS}| + |(F_2(x_{o,3}) - F_2(x_{o,2})) - \mathrm{VUS}| + |(1 - F_3(x_{o,3})) - \mathrm{VUS}|$으로 표현할 수 있다. 본 연구에서의 이차원 IU 통계량을 $k$차원으로 확장한 $\mathrm{IU}^k$에 대하여 다른 정확도 측도들과의 비교 연구를 향후 과제로 남겨둔다.

## References


Brasil, P. (2010). Diagnostic test accuracy evaluation for medical professionals, Package DiagnosisMed in R.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms, *Pattern Recognitions*, **30**, 1145–1159.

Cantor, S. B., Sun, C. C., Tortolero-Luna, G., Richards-Kortum, R., and Follen, M. (1999). A comparison of C/B ratios from studies using receiver operating characteristic curve analysis, *Journal of Clinical Epidemiology*, **52**, 885–892.

Centor, R. N. (1991). Signal detectability: The use of ROC curves and their analyses, *Medical Decision Making*, **11**, 102–106.

Connell, F. A. and Koepsell, T. D. (1985). Measures of gain in certainty from a diagnostic test, *American Journal of Epidemiology*, **121**, 744–753.

Egan, J. P. (1975). Signal detection theory and ROC analysis, Academic Press, New York.

Engelmann, B., Hayden, E., and Tasche, D. (2003). Testing rating accuracy, *Risk*, **16**, 82–86.

Fawcett, T. (2006). An introduction to ROC analysis, *Pattern Recognition Letters*, **27**, 861–874.

Fawcett, T. and Provost, F. (1997). Adaptive fraud detection, *Data Mining and Knowledge Discovery*, **1**, 291–316.

Freeman, E. A. and Moisen, G. G. (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa, *Ecological Modelling*, **217**, 48–58.

Greiner, M. M. and Gardner, I. A. (2000). Epidemiologic issues in the validation of veterinary diagnostic tests, *Preventive Veterinary Medicine*, **45**, 3–22.

Hanley, J. A. and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology*, **143**, 29–36.

Hong, C. S. and Lee S. J. (2018). TROC curve and accuracy measures, *Journal of the Korean Data & Information Science Society*, **29**, 861–872.

Hong, C. S., Joo, J. S., and Choi, J. S. (2010). Optimal thresholds from mixture distributions, *The Korean Journal of Applied Statistics*, **23(1)**, 13–28.

Hong, C. S., Lin, M. H., Hong, S. W., and Kim, G. C. (2011). Classification accuracy measures with minimum error rate for normal mixture, *Journal of the Korean Data & Information Science Society*, **22**, 619–630.

Hsieh, F. and Turnbull, B. W. (1996). Nonparametric and semiparametric estimation of the receiver operating characteristic curve, *The Annals of Statistics*, **24**, 25–40.

Krzanowski, W. J. and Hand, D. J. (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**.

Liu, C., White, M., and Newell1, G. (2009). Measuring the accuracy of species distribution models: a review. In *Proceedings 18th World IMACs/MODSIM Congress*, **4241**, 4247.

Metz, C. E. and Kronman H. B. (1980). Statistical significance tests for binormal ROC curves, *Journal of Mathematical Psychology*, **22**, 218–243.

Moses, L. E., 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, M. S. (2000). Receiver operating characteristic methodology, *Journal of the American Statistical Association*, **95**, 308–311.

Pepe, M. S. (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*, Oxford University Press, Oxford.

Perkins, N. J. and Schisterman, E. F. (2006). The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve, *American Journal of Epidemiology*, **163**, 670–675.

Provost, F. and Fawcett, T. (2001). Robust classification for imprecise environments, *Machine Learning*, **42**, 203–231.

Spackman, K. A. (1989). Signal detection theory: valuable tools for evaluating inductive learning, *The Analytics of Risk Model Validation*, San Mateo, 160–163.

Tasche, D. (2006). Validation of internal rating systems and PD estimates, *The Analytics of Risk Model Validation*, 169–196.

Unal, I. (2017). Defining an optimal cut-point value in ROC analysis: an alternative approach, *Computational & Mathematical Methods in Medicine*, **2017**, 1–14.

Vuk, M. and Curk, T. (2006). ROC curve, lift chart and calibration plot, *Metodološki Zvezki*, **3**, 89–108.

Yoo, H. S. and Hong, C. S. (2011). Optimal criterion of classification accuracy measures for normal mixture, *Communications for Statistical Applications and Methods*, **18**, 343–355.

Youden, W. J. (1950). Index for rating diagnostic test, *Cancer*, **3**, 32–35.

Zweig, M. and Campbell, G. (1993). Receiver-operating characteristics (ROC) plots: A fundamental evaluation tool in clinical medicine, *Clinical Chemistry*, **39**, 561–577.

# Index of Union와 다른 정확도 측도들

홍종선[a,1] · 최소연[a] · 임동휘[a]

[a]성균관대학교 통계학과

(2020년 5월 6일 접수, 2020년 6월 1일 수정, 2020년 6월 13일 채택)

**요 약**

최적분류점에 대한 대부분의 정확도 측도들은 두 종류의 누적분포함수와 확률밀도함수를 기반으로 정의하거나 또는 ROC 곡선과 AUC를 기반으로 정의하는 방법으로 구분하는데, Unal (2017)은 두 가지 방법을 혼합하여 누적분포함수와 AUC를 모두 고려하는 정확도 측도 Index of Union (IU) 통계량을 제안하였다. 본 연구에서는 IU 통계량을 포함한 열 개의 정확도 측도들을 여섯 종류의 범주로 구분하여 각 범주에 속하는 측도들을 비교하면서 IU의 장점을 연구한다. 다양한 정규혼합분포를 설정하여 각각의 측도들에 대응하는 최적분류점들을 구하고 각 분류점에 대응하는 제1종과 제2종 오류 그리고 두 종류의 오류합을 구해서 오류들의 크기를 비교하면서 분류정확도 측도들의 판별력을 비교하면서 IU의 성격과 특징을 탐색한다. 두 종류 분포들의 평균 차이가 증가할수록 IU 통계량의 제1종 오류와 오류합의 크기가 최고의 분류정확도를 갖는 제2범주의 정확도 측도의 오류에 수렴하는 것을 발견하였다. 그러므로 IU는 모형의 판별력을 평가하는 정확도 측도로 활용할 수 있다.

주요용어: 민감도, 분류점, 정확도, 판별, 특이도

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