Journal of the Korean Data & Information Science Society
2019, **30**(4), 895–905

http://dx.doi.org/10.7465/jkdi.2019.30.4.895
한국데이터정보과학회지

Partial VUS and optimal thresholds

Chong Sun Hong[1] · Min Sub Jung[2] · Hye Soo Shin[3]

[123]Department of Statistics, Sungkyunkwan University

Received 10 May 2019, revised 3 June 2019, accepted 3 June 2019

Abstract

Based on many literature for the partial VUS (volume under the ROC surface), a partial VUS is proposed alternatively in this work. This partial VUS is represented as the probability and formulated with the integral equations using two kinds of the ROC surface functions. We derive some relationships among the first type of the ROC surface function, its derivatives and the second partial derivative of the partial VUS. It is found that the optimal thresholds for the ROC surface are derived by using derivatives of the ROC surface function or the third partial derivative of the partial VUS. And the second type ROC surface function is discussed and explained with the results obtained from the first type of the ROC surface function. Various normal distribution functions are illustrated to find two ordered optimal threshold using the partial VUS.

Keywords: Accuracy, classifier, discrimination, sensitivity, specificity.

1. Introduction

The ROC (Receiver operating characteristic) curve is a useful method to visualize and evaluate classification models or classifiers based on its performance. It has been used for a long time in the signal detection theory which indicates a trade-off between the sensitivity and 1-specificity of the classifier. Many researches on the characteristics of ROC curves and the application of ROC analysis in decision-making and medical diagnosis systems can be found in many literature including Zweig and Campbell (1993), Berry and Linoff (1999), Provost and Fawcett (2001), Zhou (2002), Fawcett (2003, 2006), Pepe (2003, 2006), Tasche (2006), Vuk and Curk (2006), Hong and Choi (2009), Hong *et al*. (2010) etc.

A well-known statistic for determining the discriminant power for the ROC curve is the AUC (area under the ROC curve) and a specicfic area in the AUC as the 'partial AUC' (Hanley and McNeil, 1982; McClish, 1989; Thompson and Zucchini, 1989; Centor, 1991; Jiang *et al*., 1996; Bradley, 1997; Swets *et al*., 2000; Joseph, 2005; Krzanowski and Hand, 2009). Hong and Cho (2019) suggested another partial AUCs which represent the special parts of the total area under the ROC and CAP curves, and discussed to find optimal

[1] Corresponding author: Professor, Department of Statistics, Sungkyunkwan University, Seoul 03063, Korea. E-mail: cshong@skku.edu
[2] Graduate student, Department of Statistics, Sungkyunkwan University, Seoul 03063, Korea.
[3] Graduate student, Department of Statistics, Sungkyunkwan University, Seoul 03063, Korea.

thresholds using two well known accuracy measures such as TA (total accuracy) and TR (true rate) based on the first and second differential equations of the partial AUC functions.

Whereas the partial AUC has some restriction only on the TPR (true positive rate) which is the horizontal axis of the ROC curve, Yang *et al.* (2019) proposed the two-way partial area under the ROC curve which has some conditions on both the horizontal and vertical axises of the ROC curve, TPR and FPR (false positive rate). Moreover, they explained this with integration equations.

In real world, there are many cases that are classified into more than two categories, so that a method to measure the discrimination power of the multi-category classification model is required. Therefore, the ROC surface method which represents a model that classifies the three categories was developed. The volume under the ROC surface (VUS) statistic is also a well known measure to evaluate the discriminant power for the ROC surface. See Mossman (1999), Dreiseitl *et al.* (2000), Heckerling (2001), Fawcett (2003), Nakas and Yiannoutsos (2004), Nakas *et al.* (2010), Patel and Markey (2005), Hong *et al.* (2013) for more detail.

In this work, we will concentrate on the ROC surface, VUS and its related fields. For the ROC surface, Xiong *et al.* (2006) developed some methods to define and derive the partial VUS. Hong *et al.* (2019) explained the two-way partial area under the ROC curve more exactly and clearly than Yang *et al.* (2019) with probability concepts. And in order to calculate the two-way partial AUC, its probability was represented with integration equations which are able to explain the relation between the partial and two-way partial AUC. Furthermore, using the probability concepts of the partial VUS proposed by Xiong *et al.* (2006), various definitions of the partial VUS are expressed with integration equations.

In this work, the definitions of the partial AUC for the ROC curve and the partial VUS for the ROC surface are derived consistently using probability concepts, and propose an alternative partial VUS by extending results of Hong and Cho (2019). And the partial VUS defined in this work is formulated with easy integration equations. Furthermore, with some partial differential equations of this partial VUS, some methods of finding two ordered optimal threshold are explored by using some accuracy measures with various distribution functions.

In Section 2, the definitions of the partial VUS for the ROC surface proposed by Xiong *et al.* (2006) and Hong *et al.* (2019) are compared and explained consistently with probability concepts. And a partial VUS is proposed alternatively in Section 3. This partial VUS is described as the probability. Moreover, based on two kinds of the ROC surface functions, the partial VUS is expressed with the double integral equations and a single integral equation as well. In Section 4, it is found that the first type of the ROC surface function has a relation with the second partial derivative of the partial VUS, and optimal thresholds for the ROC surface can be derived by using derivatives of the ROC surface function or the third partial derivative of the partial VUS. Moreover, the first derivatives of the second type ROC surface function are discussed and explained with the results obtained from the first type of the ROC surface function. Various normal distribution functions are illustrated to find two ordered optimal thresholds in Section 5. Finally, some conclusions are mentioned in Section 6.

2. VUS for the ROC surface

Let $X_1,\ X_2,\ X_3$ be three score random variables whose corresponding cumulative distribution functions are $F_1(\cdot),\ F_2(\cdot),\ F_3(\cdot)$, respectively. For any $x \in (-\infty, \infty)$, assume that

$F_1(x) \geq F_2(x) \geq F_3(x)$. Consider a discrimination task involving three events (classes), denoted $\{E_1,\ E_2,\ E_3\}$ assigned by the testing procedure, and an observer attempts to discriminate among a set of 3 events, denoted $\{D_1,\ D_2,\ D_3\}$.

The probability, $P(D_i|E_j)$ for $i,\ j = 1, 2, 3$, means that the observer will make a particular decision, D_i, given that a particular event, E_j, occurred. The six probabilities, $P(D_i|E_j),\ i \neq j$, are the false classification rates, and the three probabilities, $P(D_k|E_k)$, $k = 1,\ 2,\ 3$, are the true classification rates: $P(D_1|E_1) = F_1(x_{01}),\ P(D_2|E_2) = F_2(x_{02}) - F_2(x_{01}),\ P(D_3|E_3) = 1 - F_3(x_{02})$ for two ordered decision markers (thresholds) $x_{01},\ x_{02}\ (x_{01} < x_{02})$.

The surface generated by three true classification rates is called the ROC surface with the two decision markers over the domain of X. This is constructed by plotting the points $(F_1(x_1),\ F_2(x_2) - F_2(x_1),\ 1 - F_3(x_2))$ in three dimensional space. There are two kinds of well-known functional forms of the ROC surface such as

$$ROC_s(p_1, p_3) = F_2(F_3^{-1}(1 - p_3)) - F_2(F_1^{-1}(p_1)),\quad 0 \leq p_1,\ p_3 \leq 1, \tag{2.1}$$

$$ROC_s(p) = F_1(F_2^{-1}(p))[1 - F_3(F_2^{-1}(p))],\quad 0 \leq p \leq 1, \tag{2.2}$$

where $p_1 = F_1(x_1),\ p_3 = 1 - F_3(x_2)$ and $p = F_2(x_2)$.

The VUS is defined based on two kinds of the ROC surface functions (1.1) and (1.2) as the followings: (see Nakas and Yiannoustsos, 2004; Hong *et al.* 2013 for more details)

$$\begin{aligned} VUS &= P(X_1 \leq X_2 \leq X_3) \\ &= \int_0^1 \int_0^{F_1(F_3^{-1}(1-p_3))} [F_2(F_3^{-1}(1 - p_3)) - F_2(F_1^{-1}(p_1))]dp_1 dp_3 \qquad (2.3) \\ &= \int_0^1 F_1(F_2^{-1}(p))[1 - F_3(F_2^{-1}(p))]dp. \qquad (2.4) \end{aligned}$$

The integration expressions of VUS in (2.3) and (2.4) are equivalent to the equations (1) and (2) of Xiong *et al.* (2006), respectively. The partial AUC for the ROC curve was discussed and defined in many literature. Furthermore, the definitions of the partial VUS for the ROC surface are also found in many works. Among them, Xiong *et al.* (2006) proposed the partial VUS which belongs to the interval between the minimum desired specificity and the minimum desired sensitivity. And Hong *et al.* (2019) defined the partial VUS alternatively in Definition 3.1 such as

$$P(x_{01} \leq X_1 \leq X_2 \leq X_3 \leq x_{02}\ \cup\ x_{03} \leq X_1 \leq X_2 \leq X_3 \leq x_{04}\ \cup\ x_{01} \leq X_2 \leq x_{04})$$

for four specific points x_{01} to x_{04} $(x_{01} < x_{02} < x_{03} < x_{04})$. In this work, we will focus the probability and its integration expression of the partial VUS.

3. The partial VUS

By extending the partial VUS of Xiong *et al.* (2006) and Hong *et al.* (2019), we propose alternative partial VUS. By using the probability concept, the partial VUS, $pVUS(u_1, u_2)$, is defined as the following.

Definition 3.1 For two specific points x_{01}, x_{02} $(x_{01} < x_{02})$ and $u_1 = F_1(x_{01}),\ u_2 = 1 - F_3(x_{02})$, the partial VUS is set with the probability:

$$pVUS(u_1, u_2) = P(X_1 \leq X_2 \leq X_3\ \cap\ X_1 \leq x_{01}\ \cap\ x_{02} \leq X_3).$$

The relationship among three random variables X_1, X_2 and X_3, and their corresponding density functions $f_1(\cdot), f_2(\cdot)$ and $f_3(\cdot)$ is represented in the left plot of Figure 3.1. Moreover, the partial VUS appears as a shaded volume shown in the right plot of Figure 3.1.

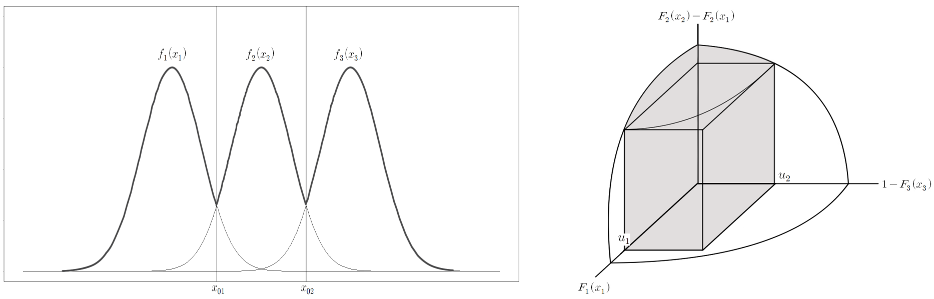

Figure 3.1 The partial VUS

This partial VUS in Definition 3.1 is proposed with integration expressions like in Theorem 3.1 and 3.2.

Theorem 3.1 The partial VUS in Definition 3.1 is expressed with the double integral equations using the first type of the ROC surface function. For specific points $x_{01},\ x_{02}$ $(x_{01} < x_{02})$ and $u_1 = F_1(x_{01}),\ u_2 = 1 - F_3(x_{02})$,

$$pVUS(u_1, u_2) = \int_0^{u_2} \int_0^{u_1} [F_2(F_3^{-1}(1-p_3)) - F_2(F_1^{-1}(p_1))] dp_1 dp_3,$$

where $p_1 = F_1(x_1)$ and $p_3 = 1 - F_3(x_2)$.

Proof:

$$\begin{aligned} pVUS(u_1, u_2) &= \int_{x_{02}}^{\infty} \int_{-\infty}^{x_{01}} [F_2(x_3) - F_2(x_1)] f_1(x_1) dx_1 f_3(x_3) dx_1 \\ &= -\int_{u_2}^{0} \int_0^{u_1} [F_2(F_3^{-1}(1-p_3)) - F_2(F_1^{-1}(p_1))] dp_1 dp_3 \\ &= \int_0^{u_2} \int_0^{u_1} [F_2(F_3^{-1}(1-p_3)) - F_2(F_1^{-1}(p_1))] dp_1 dp_3. \end{aligned}$$

□

The partial VUS in Theorem 3.1 can be easily obtained by exchanging p_1 and p_3 such as

$$pVUS(u_1, u_2) = \int_0^{u_1} \int_0^{u_2} [F_2(F_3^{-1}(1-p_3)) - F_2(F_1^{-1}(p_1))] dp_3 dp_1.$$

Theorem 3.2 The partial VUS in Definition 3.1 is expressed with the single integral equations using the second type of the ROC surface function. For specific points $x_{01},\ x_{02}$ $(x_{01} < x_{02})$ and $u_1 = F_1(x_{01}),\ u_2 = 1 - F_3(x_{02})$,

$$pVUS(u_1, u_2) = \int_{u_1}^{u_2} [F_1(F_2^{-1}(p)) - F_1(F_2^{-1}(u_1))][F_3(F_2^{-1}(u_2)) - F_3(F_2^{-1}(p))] dp,$$

where $p = F_2(x_2)$.

Proof:

$$\begin{aligned} pVUS(u_1, u_2) &= \int_{x_{01}}^{x_{02}} [F_1(x_2) - F_1(x_{01})][F_3(x_{02}) - F_3(x_2)]dF_2(x_2) \\ &= \int_{u_1}^{u_2} [F_1(F_2^{-1}(p)) - F_1(F_2^{-1}(u_1))][F_3(F_2^{-1}(u_2)) - F_3(F_2^{-1}(p))]dp. \end{aligned}$$

□

4. Relationship between the partial VUS and optimal thresholds

The partial VUS discussed in Section 3 is obtained using two kinds of the ROC surface functions. In this section, we derive some useful information for optimal thresholds using the ROC surface functions based on the partial VUS.

First, the relationship between the partial VUS defined in Theorem 3.1 and the first type ROC surface function in (2.1) is explored. Since the second partial derivative of the partial VUS in theorem 3.1 with respect to u_1 and u_2 is the first type of the ROC surface function of u_1 and u_2, two optimal thresholds could be obtained by using two following facts: One is two of the first partial derivative of the first type ROC surface function with respect to u_1 or u_2, respectively. The other is two of the third partial derivative of the partial VUS with respect to twice of u_1 and one of u_2 or one of u_1 and twice of u_2.

Theorem 4.1 The second partial derivative of the partial VUS in Theorem 3.1 turns to be the first type ROC surface function in (2.1). That is, for $u_1 = F_1(x_{01}),\ u_2 = 1 - F_3(x_{02})$,

$$\frac{\partial^2}{\partial u_1 \partial u_2} pVUS(u_1, u_2) = F_2(F_3^{-1}(1 - u_2)) - F_2(F_1^{-1}(u_1)).$$

Proof: One can find that result of the second partial derivative of the partial VUS with respect to u_1 first and next u_2 is the same as that of the second partial derivative of the partial VUS with respect to u_2 first and next u_1.

$$\begin{aligned} \frac{\partial^2}{\partial u_1 \partial u_2} pVUS(u_1, u_2) &= \frac{\partial}{\partial u_1}\left[-\int_{u_1}^{F_1(F_3^{-1}(1-u_2))} [F_2(F_3^{-1}(1 - u_2)) - F_2(F_1^{-1}(p_1))]dp_1 \right] \\ &= F_2(F_3^{-1}(1 - u_2)) - F_2(F_1^{-1}(u_1)). \end{aligned}$$

□

From Theorem 4.1, it is found that the second partial derivative of the partial VUS is the first type of the ROC surface functions. Next, in order to find optimal thresholds using the ROC surface, we take derivatives of the first type of the ROC surface functions in Theorem 4.2.

Theorem 4.2 The first partial derivatives of the first type ROC surface function with respect to u_1 and u_2, respectively, are the ratios of two the probability density functions.

$$\begin{aligned} \frac{\partial}{\partial u_1} ROC_s(u_1, u_2) &= -\frac{f_2(F_1^{-1}(u_1))}{f_1(F_1^{-1}(u_1))}, \\ \frac{\partial}{\partial u_2} ROC_s(u_1, u_2) &= -\frac{f_2(F_3^{-1}(1 - u_2))}{f_3(F_3^{-1}(1 - u_2))}. \end{aligned}$$

Proof: The first partial derivatives of the first type ROC surface function with respect to u_1 is

$$\begin{aligned}
\frac{\partial}{\partial u_1} ROC_s(u_1,u_2) &= \frac{d}{du_1}\left(-F_2(F_1^{-1}(u_1))\right) \\
&= \frac{dF_1^{-1}(u_1)}{du_1}\frac{d}{dF_1^{-1}(u_1)}\left(-F_2(F_1^{-1}(u_1))\right) \\
&= \left(\frac{d}{dx_{01}}F_1(x_{01})\right)^{-1}\left(-\frac{d}{dx_{01}}F_2(x_{01})\right) \\
&= -\frac{f_2(x_{01})}{f_1(x_{01})} = -\frac{f_2(F_1^{-1}(u_1))}{f_1(F_1^{-1}(u_1))}.
\end{aligned}$$

And the first partial derivatives of the first type of the ROC surface function with respect to u_2 turns to be

$$\begin{aligned}
\frac{\partial}{\partial u_2} ROC_s(u_1,u_2) &= \frac{d}{du_2}\left(F_2(F_3^{-1}(1-u_2))\right) \\
&= \left(-\frac{d}{dx_{02}}F_3(x_{02})\right)^{-1}\frac{d}{dF_3^{-1}(1-u_2)}\left(F_2(F_3^{-1}(1-u_2))\right) \\
&= -\frac{f_2(x_{02})}{f_3(x_{02})} = -\frac{f_2(F_3^{-1}(1-u_2))}{f_3(F_3^{-1}(1-u_2))}.
\end{aligned}$$

□

We may conclude that these results in Theorem 4.2 are similar with that of Hong and Cho (2019) which is the ratio of two the probability density functions for the ROC curve. Since the second partial derivative of the partial VUS is the first type ROC surface function, we take the third partial derivative of the partial VUS like in Lemma 4.1.

Lemma 4.1 Two kinds of the third partial derivative of the partial VUS in Theorem 3.1 are also represented the corresponding ratios:

$$\frac{\partial^3}{\partial u_1^2 \partial u_2} pVUS(u_1,u_2) = -\frac{f_2(x_{01})}{f_1(x_{01})}, \quad \frac{\partial^3}{\partial u_1 \partial u_2^2} pVUS(u_1,u_2) = -\frac{f_2(x_{02})}{f_3(x_{02})}.$$

Therefore, one might conclude that two optimal thresholds can be found since there are two axises for the first type of the ROC surface functions. By using these derivatives in Theorem 4.2 and Lemma 4.1, optimal thresholds corresponding to accuracy measures can be derived. For example, the optimal threshold for the well known accuracy measure, TR (true rate), is obtained when these ratios in Theorem 4.2 and Lemma 4.1 are set to be -1. Some examples to find optimal thresholds for some accuracy measures are explored with various kinds of distribution functions in Section 5.

Next, the relation between the optimal thresholds and the second type ROC surface function is explored. The first derivatives of the second type ROC surface function in (2.2) is discussed in Theorem 4.3.

Theorem 4.3 For any x_0, the first derivatives of the second type ROC surface function with respect to $u = F_2(x)$ is derived such as

$$\frac{d}{du} ROC_s(x)|_{u=u_0} = \frac{f_1(x_0)}{f_2(x_0)} [1 - F_3(x_0)] - \frac{f_3(x_0)}{f_2(x_0)} F_3(x_0).$$

Proof:

$$\begin{aligned} \frac{d}{du} ROC_s(u)|_{u=u_0} &= \frac{dF_2^{-1}(u)}{du} \frac{dF_1(F_2^{-1}(u))[1 - F_3(F_2^{-1}(u))]}{dF_2^{-1}(u)}|_{u=u_0} \\ &= \left(\frac{dF_2(x))}{dx}|_{x=x_0} \right)^{-1} [f_1(x_0)(1 - F_3(x_0)) - F_1(x_0) f_3(x_0)] \\ &= \frac{f_1(x_0)}{f_2(x_0)} [1 - F_3(x_0)] \ - \ \frac{f_3(x_0)}{f_2(x_0)} F_1(x_0). \end{aligned}$$

□

Based on Theorem 4.3, one can derive approximation results:

$$\frac{d}{du_0} ROC_s(x) \approx \frac{f_1(x_0)}{f_2(x_0)} \quad \text{if } F_3(x_0) \to 0, \tag{4.1}$$

$$\frac{d}{du_0} ROC_s(x) \approx -\frac{f_3(x_0)}{f_2(x_0)} \quad \text{if } F_1(x_0) \to 1. \tag{4.2}$$

From the above approximation results based on the first derivatives of the second type ROC surface function, we cannot get the same results obtained from the partial differential equation of the first type ROC surface functions relationships in Theorem 4.2 and Lemma 4.1. However, analogous relations such as the results in Theorem 4.2 and Lemma 4.1 could be derived.

Based on the result in (4.1), $F_3(x_0) \to 0$ means that the probability density function of X_3 locates far away from the right side of the density function of X_2, so that the one optimal threshold could be determined only the ratio of probability density functions of X_1 and X_2. And $F_1(x_0) \to 1$ in (4.2) can be interpreted as the probability density function of X_1 locates far away from the left side of density function of X_2. Hence only one optimal threshold could be obtained the ratio of probability density functions of X_2 and X_3. Note that the first type ROC surface function is the function of two variables, and the second type of the ROC surface functions is the function of only one variable.

Yoo and Hong (2011) discussed that when the ratio of two the probability density functions has a certain value, the corresponding optimal threshold could be obtained. Also, Hong and Cho (2019) found that the second derivative of the partial AUC which is the first derivative of the ROC function turns to be the ratio of two the probability density functions. For example, when this ratio is set to be one, the optimal threshold can be obtained by the accuracy criterion such as TR (true rate), J (Youden Index), SSS (sum of sensitivity and specificity), MVD (maximum vertical distance). The optimal threshold for TR will be mentioned in next section.

5. Some simulation results with normal distributions

5.1. Optimal thresholds for normal distributions

Let three probability density functions be assumed to be normal distributions such as: $f_1(x) = \phi(x|\mu_1, \sigma_1^2)$, $f_2(x) = \phi(x|\mu_2, \sigma_2^2)$, $f_3(x) = \phi(x|\mu_3, \sigma_3^2)$, respectively. And suppose that $\mu_1 < \mu_2 < \mu_3$. Based on Theorem 4.2, we knew that the first partial derivative of the first type of the ROC surface functions is the ratio of two probability density functions. For the TR (true rate) which is the well known accuracy measure, the ratio of two probability density functions should be one to find the optimal threshold. Hence, in order to find the first ordered optimal threshold, set the ratio to be one.

$$\frac{f_2(x)}{f_1(x)} = \frac{\sigma_1}{\sigma_2}\, exp\left(\frac{-(x-\mu_2)^2}{2\sigma_2^2} + \frac{-(x-\mu_1)^2}{2\sigma_1^2}\right) = 1.$$

The optimal threshold is

$$x_{01} = \begin{cases} (\mu_1+\mu_2)/2, & when\ \sigma_1 = \sigma_2, \\ \left(\frac{(\mu_2\sigma_1^2-\mu_1\sigma_2^2)^2}{(\sigma_1^2-\sigma_2^2)^2} - \frac{\mu_2^2\sigma_1^2-\mu_1^2\sigma_2^2}{\sigma_1^2-\sigma_2^2} + \frac{2\sigma_1^2\sigma_2^2}{\sigma_1^2-\sigma_2^2}\ln(\frac{\sigma_1}{\sigma_2})\right)^{0.5} + \left(\frac{\mu_2\sigma_1^2-\mu_1\sigma_2^2}{\sigma_1^2-\sigma_2^2}\right), & when\ \sigma_1 \le \sigma_2, \\ -\left(\frac{(\mu_2\sigma_1^2-\mu_1\sigma_2^2)^2}{(\sigma_1^2-\sigma_2^2)^2} - \frac{\mu_2^2\sigma_1^2-\mu_1^2\sigma_2^2}{\sigma_1^2-\sigma_2^2} + \frac{2\sigma_1^2\sigma_2^2}{\sigma_1^2-\sigma_2^2}\ln(\frac{\sigma_1}{\sigma_2})\right)^{0.5} + \left(\frac{\mu_2\sigma_1^2-\mu_1\sigma_2^2}{\sigma_1^2-\sigma_2^2}\right), & when\ \sigma_1 > \sigma_2, \end{cases}$$

where the terms in the square root should be positive.

Then the second ordered optimal threshold can be obtained to set the ratio $f_1(x)/f_3(x)$ to be one. The process of obtaining the second ordered optimal threshold is similar to that of the first optimal threshold and is omitted. In order to get simulation results for these works, let the second probability density function be a standard normal distribution $\mu_2 = 0,\ \sigma_2^2 = 1$, and set $\mu_3 = -\mu_1 = 2.0,\ 3.0,\ 4.0$ and $\sigma_1 = \sigma_3 = 1.0,\ 1.5,\ 2.0$ for other two normal distributions.

5.2. Optimal thresholds and errors according to means and variances

When the mean $\mu_3 = -\mu_1$ is changed from 2.0 to 4.0 at intervals of 1.0 with a fixed value 1 of the standard deviation, the optimal threshold corresponding to the accuracy measure, TR, is obtained for these distributions. And the corresponding type 1 error (α), type 2 error (β) and sum of errors ($\alpha + \beta$) are obtained. These results are summarized in Table 5.1, and shown in Figure 5.1.

Table 5.1 Optimal thresholds and errors according to means

	$\mu_3 = -\mu_1$		
	2	3	4
x_0	1	1.5	2
α	0.159	0.067	0.023
β	0.159	0.067	0.023
$\alpha+\beta$	0.317	0.134	0.046

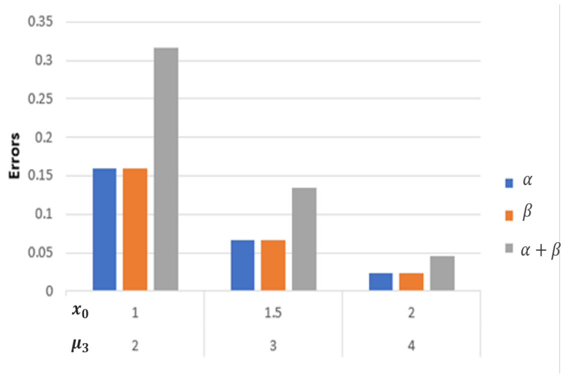

Figure 5.1 Optimal thresholds and errors according to means

From Table 5.1 and Figure 5.1, it is found that the optimal thresholds corresponding to TR are $\mu_1/2 = -\mu_3/2$ with equal variance of these distributions. When μ_1 or μ_3 are getting far away to zero, all values of α, β, $\alpha + \beta$ are decreasing exponentially. Note that $\alpha = \beta$.

When the standard deviations σ_1 and σ_3 are changed from 1.0 to 2.0 at intervals of 0.5 with fixed $\mu_3 = -\mu_1 = 2$, the optimal threshold corresponding to TR, α, β and $\alpha + \beta$ are obtained. Then these results are summarized in Table 5.2 and expressed in Figure 5.2.

Table 5.2 Optimal thresholds and errors according to variances

	σ_1			σ_3		
	1	1.5	2	1	1.5	2
x_0	-1	-1.087	-1.238	1	1.087	1.238
α	0.159	0.139	0.108	0.159	0.271	0.352
β	0.159	0.271	0.352	0.159	0.139	0.108
$\alpha+\beta$	0.317	0.41	0.459	0.317	0.41	0.459

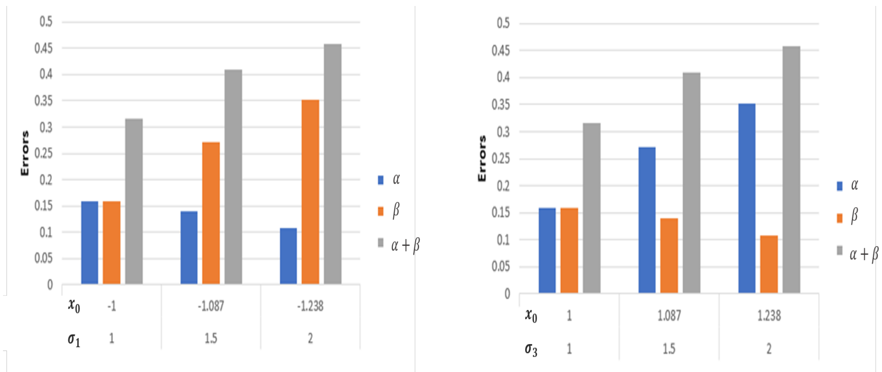

Figure 5.2 Optimal thresholds and errors according to variances

Based on Table 5.2 and Figure 5.2, as σ_1 gets larger, the optimal threshold moves farther away from zero and has a negative value. But when σ_3 increases, the optimal threshold locates farther away from zero and has a positive value. Moreover, it is found that as σ_1 and

σ_3 are getting larger, $\alpha + \beta$ is increasing. In particular, when σ_1 is increasing, α decrease slowly and β is increasing faster than α. On the contrary, β decrease slowly and α gets larger than β when σ_3 increases. Therefore, since increasing errors get larger than decreasing errors, sum of errors is increasing.

6. Conclusion

The ROC curve which is a useful method to visualize and evaluate classification models or classifiers has been used for a long time in the signal detection theory, and it has been widely applied in decision-making and medical diagnosis systems. An well-known statistic for determining the discriminant power using the ROC curve is the AUC.

A specific area in the AUC is called as the partial AUC (pAUC). There are several kind of definitions of the partial AUC. Hong and Cho (2019) suggested another partial AUC for the ROC and CAP curves, and discussed relationships between the first and second differential equations of the partial AUC and optimal thresholds corresponding to some accuracy measures. Yang *et al*. (2019) proposed the two-way partial area under the ROC curve (tpAUC) which have some conditions on both the horizontal and vertical axises of the ROC curve. Moreover, the tpAUC was expressed as integral equations. Hong *et al*. (2019) explained the tpAUC with probability concepts and represented the tpAUC in terms of integration equations which are able to explain the relation between pAUC and tpAUC.

An extending method to measure the discrimination power of the three category classification model is the ROC surface method with three categories. The volume under the ROC surface (VUS) statistic is also a well known measure to evaluate the discriminant power using the ROC surface. For the ROC surface, Xiong *et al*. (2006) suggested the partial VUS (pVUS). Using the probability concepts of the partial VUS, Hong *et al*.. (2019) discussed and expressed various definitions of the partial VUS with integration equations.

In this work, an alternative partial VUS for the ROC surface is represented as the probability by extending results of Hong and Cho (2019) and formulated with double and single integral equations. Furthermore, it is found that the first type of the ROC surface function has a relation with the second partial derivative of the partial VUS. And two ordered optimal thresholds using some accuracy measures can be derived by using derivatives of the ROC surface function or the third partial derivative of the partial VUS. Various normal distribution functions are explored for finding two optimal thresholds and two kind of errors corresponding to their optimal thresholds. Therefore, one might conclude that since using the partial VUS formulated with integral equations, these methods can be widely used and applied in real data in order to measure the discrimination power of the three category classification model.

References

Berry, M. J. A. and Linoff, G. (1999). *Data mining techniques: For marketing, sales, and customer support*, 3rd Ed., John Wiley Sons, New York.

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

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

Dreiseitl, S., Ohno-Machado, L. and Binder, M. (2000). Comparing three-class diagnostic tests by three-way ROC analysis. *Medical Decision Making*, **20**, 323-331.

Fawcett, T. (2003). ROC Graphs: Notes and practical considerations for data mining researchers. *Technical report hpl-2003-4, HP Laboratories, Palo Alto, CA, USA*.

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

Hanley, A. and McNeil, B. (1982). The meaning and use of the area under a receiver operating characteristics curve. *Diagnostic Radiology*, **143**, 29-36.

Heckerling, P. S. (2001). Parametric three-way receiver operating characteristic surface analysis using Mathematica. *Medical Decision Making*, **21**, 409-417.

Hong, C. S. and Cho, H. S. (2019). Partial AUC and optimal thresholds. *The Korean Journal of Applied Statistics*, **32**, 187-198.

Hong, C. S. and Choi, J. S. (2009). Optimal Threshold from ROC and CAP Curves. *The Korean Journal of Applied Statistics*, **22**, 911-921.

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

Hong, C. S. Jung, E. S. and Jung, D. G. (2013). Standard criterion of VUS for ROC surface. *The Korean Journal of Applied Statistics*, **26**, 977-985.

Hong, C. S., Jung, M. S. and Shin, H. S. (2019). Three-way partial VUS. *Journal of The Korean Data & Information Science Society*, **30**, 445-454.

Jiang, Y., Metz, C. and Nishikawa, R. (1996). A receiver operating characteristic partial area index for highly sensitive diagnostic tests. *Radiology*, **201**, 745-750.

Joseph, M. P. (2005). A PD validation framework for Basel II internal ratings-based systems. *Credit Scoring and Credit Control IV*.

Krzanowski, W. J. and Hand, D. J. (2009). *ROC curves for continuous data*, 1st Ed., CRC Press, New York.

McClish, D. (1989). Analyzing a portion of the ROC curve. *Medical Decision Making*, **9**, 190-195.

Mossman, D. (1999). Three-way ROCs. *Medical Decision Making*, **19**, 78-89.

Nakas, C. T., Alonzo, T. A. and Yiannoutsos, C. T. (2010). Accuracy and cut-off point selection in three-class classification problems using a generalization of the Youden index. *Statistics in Medicine*, **29**, 2946-2955.

Nakas, C. T. and Yiannoutsos, C. T. (2004). Ordered multiple-class ROC analysis with continuous measurements. *Statistics in Medicine*, **23**, 3437-3449.

Patel, A. C. and Markey, M. K. (2005). Comparison of three-class classification performance metrics: a case study in breast cancer CAD. *International Society for Optical Engineering*, **5749**, 581-589.

Pepe, M. S. (2003). *The statistical evaluation of medical test for classification and prediction*, 17th Ed., Oxford University Press, Oxford.

Pepe, M. S., Cai, T. and Longton, G. (2006). Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics*, **62**, 221-229.

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

Swets, J. A., Dawes, R. M. and Monahan, J. (2000). Better decisions through science. *Scientific American*, **283**, 82-87.

Tasche, D. (2006). Validation of internal rating systems and PD estimates. *arXiv.org, eprint arXiv: physics /0606071*.

Thompson, M. L. and Zucchini, W. (1989). On the statistical analysis of ROC curves. *Statistics in Medicine*, **8**, 1277-1290.

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

Xiong, C., Van Belle, G., Miller, J. P. and Morris, J. C. (2006). Measuring and estimating diagnostic accuracy when there are three ordinal diagnostic groups. *Statistics in Medicine*, **25**, 1251-1273.

Yang, H., Lu, K., Lyu, X. and Hu, F. (2019). Two-way partial AUC and its properties. *Statistical Methods in Medical Research*, **28**, 184-195.

Yoo, H. S. and Hong, C. S. (2011). Optimal criterion of classification accuracy measures for normal mixture. *The Korean Journal of Applied Statistics*, **18**, 343-355.

Zhou, K. H. (2002). Receiver operating characteristic literature research. *On-line bibliography available from: http://www.spl.harvard.edu/pages/ppl/zou/roc.html*.

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