Communications for Statistical Applications and Methods
2015, Vol. 22, No. 3, 223–232

DOI: http://dx.doi.org/10.5351/CSAM.2015.22.3.223
Print ISSN 2287-7843 / Online ISSN 2383-4757

VUS and HUM Represented with Mann-Whitney Statistic

Chong Sun Hong[1,a], Min Ho Cho[a]

[a]Department of Statistics, Sungkyunkwan University, Korea

Abstract

The area under the ROC curve (AUC), the volume under the ROC surface (VUS) and the hypervolume under the ROC manifold (HUM) are defined and interpreted with probability that measures the discriminant power of classification models. AUC, VUS and HUM are expressed with the summation and integration notations for discrete and continuous random variables, respectively. AUC for discrete two random samples is represented as the nonparametric Mann-Whitney statistic. In this work, we define conditional Mann-Whitney statistics to compare more than two discrete random samples as well as propose that VUS and HUM are represented as functions of the conditional Mann-Whitney statistics. Three and four discrete random samples with some tie values are generated. Values of VUS and HUM are obtained using the proposed statistic. The values of VUS and HUM are identical with those obtained by definition; therefore, both VUS and HUM could be represented with conditional Mann-Whitney statistics proposed in this paper.

Keywords: AUC, classification, HUM, manifold, nonparametric, ROC, surface, VUS

1. Introduction

The receiver operating characteristic (ROC) curve is a technique to visualize and organize the classification model (or classifiers) based on performance. ROC curve has been widely used in signal detection theory to depict the tradeoff between Sensitivity and 1-Specificity of classifiers. ROC analysis has been extended to visualize and analyze the behavior of diagnostic systems (Egan, 1975; Swets, 1988; Swets *et al.*, 2000). The medical decision-making community has extensive literature on the use of ROC graphs for diagnostic testing. Research regarding the property of ROC curve and information about application of ROC analysis can be found in many papers, such as Provost and Engelmann *et al.* (2003), Fawcett (2003), Hong (2009), Hong and Choi (2009), Hong *et al.* (2010), Provost and Fawcett (2001), Sobehart and Keenan (2001) and Zou *et al.* (2007). The area under the ROC curve (AUC) is used as an objective statistic to measure the power of discriminant through ROC curve (Bradley, 1997; Hanley and McNeil, 1982).

Most circumstances in real world consist of multiple categories rather than only two that require methods to measure the power of discriminant of multiple category classification model are required. For example, the triple categories (Nondefault, Warning, Default) are more realistic rather than two categories (Nondefault, Default) in order to classify the credit assessment models. Whereas both ROC curve and AUC are for two dimensions, ROC surface and the volume under the ROC surface (VUS) are extended to three dimensions (Dreiseitl *et al.*, 2000; Fawcett, 2003; Heckerling, 2001; Hong *et al.*, 2013; Mossmann, 1999; Nakas and Yiannoutsos, 2004; Nakas *et al.*, 2010; Patel and Markey, 2005; Scurfield, 1996; Wandishin and Mullen, 2009, Zou *et al.*, 2007). By extending the ROC surface

[1] Corresponding author: Department of Statistics, Sungkyunkwan University, Seoul 110-745, Korea.
E-mail: cshong@skku.edu

Published 31 May 2015 / journal homepage: http://csam.or.kr

and VUS for three dimensions, Li and Fine (2008) and Hong and Jung (2014) defined and suggested both the ROC manifold and the hypervolume under the ROC manifold (HUM) statistic to discriminate more than three categories.

The properties of discrete form of both VUS and HUM are discussed and explained. We will also express VUS and HUM as functions of Mann-Whitney statistics in this paper because AUC can be represented with the well-known nonparametric Mann-Whitney statistic. First, the conditional Mann-Whitney statistic is suggested and defined. Then VUS and HUM for discrete random variables might be represented with conditional Mann-Whitney statistics. Hence values of the VUS and HUM could be obtained using the conditional Mann-Whitney statistics proposed in this paper.

In Section 2 of this paper, we explain the definitions of AUC, VUS and HUM. Especially, these are explored for discrete variables and the HUM is concerned only for four categories of the classification model. Section 3 suggests and defines the conditional Mann-Whitney statistics for multiple random samples. Then VUS and HUM are proposed to represent with conditional Mann-Whitney statistics. Some illustrative examples consisting of three and four random variables with some tie values are generated. Then values of VUS and HUM are obtained and compared with values calculated by definitions in Section 4. Section 5 provides the conclusion and future works.

2. ROC Manifold and HUM

Suppose k random variables $X_1, X_2, \ldots, X_k$ and their cumulative distribution functions, $F_1(\cdot), F_2(\cdot), \ldots, F_k(\cdot)$ satisfying $F_1(x) \geq F_2(x) \geq \cdots \geq F_k(x)$ for all x. Nakas and Yiannoutsos (2004) and many others defined AUC, VUS and HUM for ROC curve, surface and manifold, respectively, such as

$$\begin{aligned} \text{AUC} &= P(X_1 \leq X_2), \\ \text{VUS} &= P(X_1 \leq X_2 \leq X_3), \\ \text{HUM}^k &= P(X_1 \leq X_2 \leq \cdots \leq X_k) \quad \text{for } k \geq 4. \end{aligned}$$

The AUC, VUS and HUM could be expressed by the following integral and probability notations for continuous and discrete random variables, respectively (Hong and Jung, 2014). When random variables are continuous, AUC, VUS and HUM^4 are represented as

$$\begin{aligned} \text{AUC} &= \int_0^1 \text{ROC}(u_1)du_1 = \int_0^1 F_1\left(F_2^{-1}(u_1)\right)du_1, \\ \text{VUS} &= \int_0^1 \int_0^{F_1\left(F_3^{-1}(1-u_3)\right)} \text{ROC}_s(u_1, u_3)du_1du_3 = \int_0^1 F_1\left(F_2^{-1}(1-u)\right) - F_3\left(F_2^{-1}(u)\right)du, \\ \text{HUM}^4 &= \int_0^1 \int_0^{F_2\left(F_3^{-1}(u_3)\right)} F_1\left(F_2^{-1}(u_2)\right)\left[1 - F_4\left(F_3^{-1}(u_3)\right)\right]du_2du_3. \end{aligned}$$

When random variables are discrete, values of AUC, VUS and HUM^4 are expressed such as

$$\text{AUC} = P(X_1 < X_2) + \frac{1}{2}P(X_1 = X_2), \tag{2.1}$$

$$\text{VUS} = P(X_1 < X_2 < X_3) + \frac{1}{2}P(X_1 = X_2 < X_3) + \frac{1}{2}P(X_1 < X_2 = X_3) + \frac{1}{2^2}P(X_1 = X_2 = X_3), \tag{2.2}$$

$$\begin{aligned}\text{HUM}^4 &= P(X_1 < X_2 < X_3 < X_4) + \frac{1}{2}P(X_1 = X_2 < X_3 < X_4) + \frac{1}{2}P(X_1 < X_2 = X_3 < X_4)\\ &+ \frac{1}{2}P(X_1 < X_2 < X_3 = X_4) + \frac{1}{2^2}P(X_1 = X_2 < X_3 = X_4) + \frac{1}{2^2}P(X_1 = X_2 = X_3 < X_4)\\ &+ \frac{1}{2^2}P(X_1 < X_2 = X_3 = X_4) + \frac{1}{2^3}P(X_1 = X_2 = X_3 = X_4). \qquad (2.3)\end{aligned}$$

AUC in (2.1) can be represented with Mann-Whitney statistic such as

$$\text{AUC} = \frac{1}{n_1 n_2}\left[U_{X_1<X_2} + \frac{1}{2}U_{X_1=X_2}\right],$$

where $U_{X_1<X_2} \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} I(X_{1i} < X_{2j})$, $U_{X_1=X_2} \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} I(X_{1i} = X_{2j})$ (Rosset, 2004). In this work, we would like to express VUS and HUM in (2.2) and (2.3), respectively, as functions of Mann-Whitney statistics.

3. VUS and HUM Represented with Conditional Mann-Whitney Statistics

3.1. VUS with conditional Mann-Whitney statistics

Suppose three random samples $\{X_{1i}\}, \{X_{2j}\}, \{X_{3k}\}$ of sizes n_1, n_2, n_3, respectively. And with their cumulative distribution functions, $F_1(\cdot)$, $F_2(\cdot)$, $F_3(\cdot)$, assume that $F_1(x) - F_2(x) \geq 0$ and $F_2(x) - F_3(x) \geq 0$ for all x.

By using the property of the conditional probability, the VUS under the assumption $F_1(x) \geq F_2(x) \geq F_3(x)$ can be defined as follows.

$$\begin{aligned}\text{VUS} &= P(X_1 < X_2 < X_3) + \frac{1}{2}P(X_1 = X_2 < X_3) + \frac{1}{2}P(X_1 < X_2 = X_3) + \frac{1}{2^2}P(X_1 = X_2 = X_3)\\ &= P(X_2 < X_3|X_1 < X_2)P(X_1 < X_2) + \frac{1}{2}P(X_2 < X_3|X_1 = X_2)P(X_1 = X_2)\\ &+ \frac{1}{2}P(X_2 = X_3|X_1 < X_2)P(X_1 < X_2) + \frac{1}{2^2}P(X_2 = X_3|X_1 = X_2)P(X_1 = X_2). \qquad (3.1)\end{aligned}$$

Consider a paired sample $\{(X_{1i}, X_{2j})\}$ satisfying $X_{1i} < X_{2j}$ for each pair. The $\{X_{2j}\}$ in $\{(X_{1i}, X_{2j})|X_{1i} < X_{2j}\}$ is denoted as $\{X_{2j}; X_{1i} < X_{2j}\}$. In order to compare $\{X_{3k}\}$ with $\{X_{2j}; X_{1i} < X_{2j}\}$, we will define the conditional Mann-Whitney statistic.

Definition 1. *Definition of the conditional Mann-Whitney statistic*

$$U_{X_2<X_3|X_1<X_2} = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j} < X_{3k}|X_{1i} < X_{2j}\right).$$

The conditional probability $P(X_2 < X_3|X_1 < X_2)$ in (3.1) could be defined with the conditional Mann-Whitney statistic which is calculated from two random samples $\{X_{3k}\}$ and $\{X_{2j}; X_{1i} < X_{2j}\}$. Now we propose alternative statistic to obtain the VUS for ROC surface.

Theorem 1. *The VUS could be obtained by using the conditional Mann-Whitney statistic, such as*

$$VUS_{MW} = \frac{1}{n_1 n_2 n_3}\left[U_{X_2<X_3|X_1<X_2} + \frac{1}{2}U_{X_2=X_3|X_1<X_2} + \frac{1}{2}U_{X_2<X_3|X_1=X_2} + \frac{1}{2^2}U_{X_2=X_3|X_1=X_2}\right],$$

where

$$U_{X_2<X_3|X_1<X_2} \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}<X_{3k}|X_{1i}<X_{2j}\right), \quad U_{X_2=X_3|X_1<X_2} \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}=X_{3k}|X_{1i}<X_{2j}\right),$$

$$U_{X_2<X_3|X_1=X_2} \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}<X_{3k}|X_{1i}=X_{2j}\right), \quad U_{X_2=X_3|X_1=X_2} \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}=X_{3k}|X_{1i}=X_{2j}\right).$$

Proof: Since

$$\frac{1}{n_1n_2n_3}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i}<X_{2j}<X_{3k}\right) = \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}<X_{3k}|X_{1i}<X_{2j}\right)}{n_3\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} I\left(X_{1i}<X_{2j}\right)} \times \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} I\left(X_{1i}<X_{2j}\right)}{n_1n_2}$$
$$= \frac{1}{n_1n_2n_3}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}<X_{3k}|X_{1i}<X_{2j}\right),$$

the VUS is obtained that

$$\frac{1}{n_1n_2n_3}\left[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i}<X_{2j}<X_{3k}\right) + \frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i}=X_{2j}<X_{3k}\right)\right.$$
$$\left. +\frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i}<X_{2j}=X_{3k}\right) + \frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i}=X_{2j}=X_{3k}\right)\right]$$
$$= \frac{1}{n_1n_2n_3}\left[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}<X_{3k}|X_{1i}<X_{2j}\right) + \frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}=X_{3k}|X_{1i}<X_{2j}\right)\right.$$
$$\left. +\frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}<X_{3k}|X_{1i}=X_{2j}\right) + \frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{2j}=X_{3k}|X_{1i}=X_{2j}\right)\right]$$
$$= \frac{1}{n_1n_2n_3}\left[U_{X_2<X_3|X_1<X_2} + \frac{1}{2}U_{X_2=X_3|X_1<X_2} + \frac{1}{2}U_{X_2<X_3|X_1=X_2} + \frac{1}{2^2}U_{X_2=X_3|X_1=X_2}\right].$$

□

3.2. HUM^4 with conditional Mann-Whitney statistics

Now we suppose four random samples $\{X_{1i}\}, \{X_{2j}\}, \{X_{3k}\}, \{X_{4l}\}$ of sizes n_1, n_2, n_3, n_4, respectively. For cumulative distribution functions, $F_1(\cdot), F_2(\cdot), F_3(\cdot), F_4(\cdot)$, assume that $F_1(x) \geq F_2(x) \geq F_3(x) \geq F_4(x)$ for all x.

By using the property of the conditional probability, the HUM^4 under the assumption $F_1(x) \geq F_2(x) \geq F_3(x) \geq F_4(x)$ can be represented as follows.

$$\text{HUM}^4 = P(X_1<X_2<X_3<X_4) + \frac{1}{2}P(X_1=X_2<X_3<X_4) + \frac{1}{2}P(X_1<X_2=X_3<X_4)$$
$$+\frac{1}{2}P(X_1<X_2<X_3=X_4) + \frac{1}{2^2}P(X_1=X_2<X_3=X_4) + \frac{1}{2^2}P(X_1=X_2=X_3<X_4)$$
$$+\frac{1}{2^2}P(X_1<X_2=X_3=X_4) + \frac{1}{2^3}P(X_1=X_2=X_3=X_4)$$

$$
\begin{aligned}
&= P(X_3 < X_4 | X_1 < X_2 < X_3) P(X_1 < X_2 < X_3) + \frac{1}{2} P(X_3 < X_4 | X_1 = X_2 < X_3) P(X_1 = X_2 < X_3) \\
&+ \frac{1}{2} P(X_3 < X_4 | X_1 < X_2 = X_3) P(X_1 < X_2 = X_3) + \frac{1}{2} P(X_3 = X_4 | X_1 < X_2 < X_3) P(X_1 < X_2 < X_3) \\
&+ \frac{1}{2^2} P(X_3 = X_4 | X_1 = X_2 < X_3) P(X_1 = X_2 < X_3) + \frac{1}{2^2} P(X_3 < X_4 | X_1 = X_2 = X_3) P(X_1 = X_2 = X_3) \\
&+ \frac{1}{2^2} P(X_3 = X_4 | X_1 < X_2 = X_3) P(X_1 < X_2 = X_3) + \frac{1}{2^3} P(X_3 = X_4 | X_1 = X_2 = X_3) P(X_1 = X_2 = X_3). \quad (3.2)
\end{aligned}
$$

The probability $P(X_1 < X_2 < X_3)$, $P(X_1 = X_2 < X_3)$, $P(X_1 < X_2 = X_3)$ and $P(X_1 = X_2 = X_3)$ in (3.2) can be generalized by using the conditional probabilities in (3.1). Therefore, the conditional probability $P(X_3 < X_4 | X_1 < X_2 < X_3)$ could be represented with the conditional Mann-Whitney statistic which is calculated from two random samples $\{X_{4l}\}$ and $\{X_{3k}; X_{1i} < X_{2j} < X_{3k}\}$. Then the conditional Mann-Whitney statistic, $U_{X_3<X_4|X_1<X_2<X_3}$ could be defined as $\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I(X_{3k} < X_{4l} | X_{1i} < X_{2j} < X_{3k})$. Other conditional probabilities are also represented with conditional Mann-Whitney statistics. Therefore, HUM for ROC manifold for four random variables could be represented with the following conditional Mann-Whitney statistics.

Theorem 2. *The HUM^4 could be obtained by using the conditional Mann-Whitney statistic such as*

$$
\begin{aligned}
HUM^4_{MW} = \frac{1}{n_1 n_2 n_3 n_4} \Bigg[& U_{X_3<X_4|X_1<X_2<X_3} + \frac{1}{2} U_{X_3=X_4|X_1<X_2<X_3} + \frac{1}{2} U_{X_3<X_4|X_1=X_2<X_3} + \frac{1}{2^2} U_{X_3=X_4|X_1=X_2<X_3} \\
& + \frac{1}{2} U_{X_3<X_4|X_1<X_2=X_3} + \frac{1}{2^2} U_{X_3=X_4|X_1<X_2=X_3} + \frac{1}{2^2} U_{X_3<X_4|X_1=X_2=X_3} + \frac{1}{2^3} U_{X_3=X_4|X_1=X_2=X_3} \Bigg], \quad (3.3)
\end{aligned}
$$

where

$$
\begin{aligned}
U_{X_3<X_4|X_1<X_2<X_3} &\equiv \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I\left(X_{3k} < X_{4l} | X_{1i} < X_{2j} < X_{3k}\right), \\
U_{X_3=X_4|X_1<X_2<X_3} &\equiv \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I\left(X_{3k} = X_{4l} | X_{1i} < X_{2j} < X_{3k}\right), \\
U_{X_3<X_4|X_1=X_2<X_3} &\equiv \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I\left(X_{3k} < X_{4l} | X_{1i} < X_{2j} = X_{3k}\right), \\
U_{X_3=X_4|X_1=X_2<X_3} &\equiv \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I\left(X_{3k} = X_{4l} | X_{1i} = X_{2j} < X_{3k}\right), \\
U_{X_3<X_4|X_1<X_2=X_3} &\equiv \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I\left(X_{3k} = X_{4l} | X_{1i} = X_{2j} < X_{3k}\right), \\
U_{X_3=X_4|X_1<X_2=X_3} &\equiv \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I\left(X_{3k} = X_{4l} | X_{1i} < X_{2j} = X_{3k}\right), \\
U_{X_3<X_4|X_1=X_2=X_3} &\equiv \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \sum_{l=1}^{n_4} I\left(X_{3k} < X_{4l} | X_{1i} = X_{2j} = X_{3k}\right),
\end{aligned}
$$

$$U_{X_3=X_4|X_1=X_2=X_3} \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{3k} = X_{4l}|X_{1i} = X_{2j} = X_{3k}\right).$$

Proof: The HUM[4] can be defined by using indicator functions.

$$\frac{1}{n_1n_2n_3n_4}\left[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{1i} < X_{2j} < X_{3k} < X_{4l}\right) + \frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{1i} = X_{2j} < X_{3k} < X_{4l}\right)\right.$$
$$+\frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i} < X_{2j} = X_{3k} < X_{4l}\right) + \frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i} < X_{2j} < X_{3k} = X_{4l}\right)$$
$$+\frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1,i} = X_{2,j} < X_{3,k} = X_{4,l}\right) + \frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i} = X_{2j} = X_{3k} < X_{4l}\right)$$
$$\left.+\frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i} < X_{2j} = X_{3k} = X_{4l}\right) + \frac{1}{2^3}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i} = X_{2j} = X_{3k} = X_{4l}\right)\right].$$

Since

$$\frac{1}{n_1n_2n_3n_4}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{1i} < X_{2j} < X_{3k} < X_{4l}\right)$$
$$= \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{3k} < X_{4l}|X_{1i} < X_{2j} < X_{3k}\right)}{n_4\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i} < X_{2j} < X_{3k}\right)} \times \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{1i} < X_{2j} < X_{3k}\right)}{n_1n_2n_3}$$
$$= \frac{1}{n_1n_2n_3n_4}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{3k} < X_{4l}|X_{1i} < X_{2j} < X_{3k}\right),$$

the HUM[4] could be obtained that

$$\frac{1}{n_1n_2n_3n_4}\left[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{3k} < X_{4l}|X_{1i} < X_{2j} < X_{3k}\right) + \frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{l=1}^{n_4} I\left(X_{3k} = X_{4l}|X_{1i} < X_{2j} < X_{3k}\right)\right.$$
$$+\frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{3k} < X_{4l}|X_{1i} = X_{2j} < X_{3k}\right) + \frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{3k} = X_{4l}|X_{1i} = X_{2j} < X_{3k}\right)$$
$$+\frac{1}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{3k} < X_{4l}|X_{1i} < X_{2j} = X_{3k}\right) + \frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{3k} = X_{4l}|X_{1i} < X_{2j} = X_{3k}\right)$$
$$\left.+\frac{1}{2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{3k} < X_{4l}|X_{1i} = X_{2j} = X_{3k}\right) + \frac{1}{2^3}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I\left(X_{3k} = X_{4l}|X_{1i} = X_{2j} = X_{3k}\right)\right].$$

□

4. Some Illustrative Examples

4.1. Example of VUS

Table 1 shows three samples $\{X_{1i}\}, \{X_{2j}\}, \{X_{3k}\}$ of sizes $n_1 = 5$, $n_2 = 6$, $n_3 = 5$, respectively. The ROC surface could consist of three positive rates such as $(F_1(x), F_2(y) - F_2(x), 1 - F_3(y))$, for all x, y $(x < y)$

Table 1: Three random samples

X_1	11	17		23	39	44							$n_1 = 5$
X_2		17	22		39		48	57		72			$n_2 = 6$
X_3					39			57	63		89	94	$n_3 = 5$

Table 2: Subsamples and the conditional Mann-Whitney statistics

$\{(X_1, X_2)\|X_1 < X_2\}$	$X_3(> X_2)$	$\{(X_1, X_2)\|X_1 = X_2\}$	$X_3(> X_2)$	$\{(X_1, X_2)\|X_1 < X_2\}$	$X_3(= X_2)$	$\{(X_1, X_2)\|X_1 = X_2\}$	$X_3(= X_2)$
(11, 17)		(17, 17)		(11, 17)		(17, 17)	
(11, 22)				(11, 22)			
(17, 22)				(17, 22)			
(11, 39)	39	(39, 39)	39	(11, 39)	39	(39, 39)	39
(17, 39)				(17, 39)			
(23, 39)				(23, 39)			
(11, 48)				(11, 48)			
(17, 48)				(17, 48)			
(23, 48)				(23, 48)			
(39, 48)				(39, 48)			
(44, 48)				(44, 48)			
(11, 57)	57		57	(11, 57)	57		57
(17, 57)				(17, 57)			
(23, 57)				(23, 57)			
(39, 57)				(39, 57)			
(44, 57)				(44, 57)			
	63		63		63		63
(11, 72)				(11, 72)			
(17, 72)				(17, 72)			
(23, 72)				(23, 72)			
(39, 72)				(39, 72)			
(44, 72)				(44, 72)			
	89		89		89		89
	94		94		94		94
$U_{X_2<X_3\|X_1<X_2} = 72$		$U_{X_2<X_3\|X_1=X_2} = 9$		$U_{X_2=X_3\|X_1<X_2} = 8$		$U_{X_2=X_3\|X_1=X_2} = 1$	

on unit dice (Hong *et al.*, 2013).

There are two subsamples $\{(X_1, X_2)|X_1 < X_2\}$ and $\{(X_1, X_2)|X_1 = X_2\}$ collected in Table 2. Then $\{X_3\}$ is compared with $\{X_2\}$ in these subsamples $\{X_2; X_1 < X_2\}$ and $\{X_2; X_1 = X_2\}$, so that conditional Mann-Whitney statistics are obtained in Table 2.

Then with conditional Mann-Whitney statistics in Table 2, the VUS is obtained that

$$\begin{aligned} \text{VUS}_{MW} &= \frac{1}{n_1 n_2 n_3}\left[U_{X_2<X_3|X_1<X_2} + \frac{1}{2}U_{X_2<X_3|X_1=X_2} + \frac{1}{2}U_{X_2=X_3|X_1<X_2} + \frac{1}{2^2}U_{X_2=X_3|X_1=X_2}\right] \\ &= \frac{1}{5\times 6\times 5}\left(72 + \frac{9}{2} + \frac{8}{2} + \frac{1}{4}\right) = 0.5383. \end{aligned}$$

4.2. Example of HUM[4]

There are four illustrative random samples $\{X_{1i}\}$, $\{X_{2j}\}$, $\{X_{3k}\}$, $\{X_{4l}\}$ of sizes $n_1 = 4$, $n_2 = 4$, $n_3 = 6$, $n_4 = 5$, respectively, in Table 3.

Four subsamples $\{(X_1, X_2, X_3)|X_1 < X_2 < X_3\}$, $\{(X_1, X_2, X_3)|X_1 = X_2 < X_3\}$, $\{(X_1, X_2, X_3)|X_1 < X_2 = X_3\}$ and $\{(X_1, X_2, X_3)|X_1 = X_2 = X_3\}$ are collected in Table 4. Then $\{X_4\}$ is compared of $\{X_3\}$ in these four subsamples, so that conditional Mann-Whitney statistics are obtained in Table 4.

Table 3: Four random samples

X_1	11	17		23		45											$n_1 = 4$
X_2			22			45		61			77						$n_2 = 4$
X_3					29	45	54			72		83		90			$n_3 = 6$
X_4						45			69				88		95	100	$n_4 = 5$

Table 4: Subsamples and the conditional Mann-Whitney statistics

$\{(X_1, X_2, X_3) \mid X_1 < X_2 < X_3\}$	X_4	$\{(X_1, X_2, X_3) \mid X_1 = X_2 < X_3\}$	X_4	$\{(X_1, X_2, X_3) \mid X_1 < X_2 = X_3\}$	X_4	$\{(X_1, X_2, X_3) \mid X_1 = X_2 = X_3\}$	X_4
(11, 22, 29)							
(17, 22, 29)							
(11, 22, 45)	45		45	(11, 45, 45)			
(17, 22, 45)				(17, 45, 45)	45	(45, 45, 45)	45
(11, 22, 54)		(45, 45, 54)		(23, 45, 45)			
(17, 22, 54)							
(11, 45, 54)							
(17, 45, 54)							
(23, 45, 54)							
	69		69		69		69
(11, 22, 72)		(45, 45, 72)					
(17, 22, 72)							
(11, 45, 72)							
(17, 45, 72)							
(23, 45, 72)							
(11, 61, 72)							
(17, 61, 72)							
(23, 61, 72)							
(45, 61, 72)							
(11, 22, 83)		(45, 45, 83)					
(17, 22, 83)							
(11, 45, 83)							
(17, 45, 83)							
(23, 45, 83)							
(11, 61, 83)							
(17, 61, 83)							
(23, 61, 83)							
(45, 61, 83)							
(11, 77, 83)							
(17, 77, 83)							
(23, 77, 83)							
(45, 77, 83)							
	88		88		88		88
(11, 22, 90)		(45, 45, 90)					
(17, 22, 90)							
(11, 45, 90)							
(17, 45, 90)							
(23, 45, 90)							
(11, 61, 90)							
(17, 61, 90)							
(23, 61, 90)							
(45, 61, 90)							
(11, 77, 90)							
(17, 77, 90)							
(23, 77, 90)							
(45, 77, 90)							
	95		95		95		95
	100		100		100		100
$U_{X_3<X_4 \mid X_1<X_2<X_3} = 130$		$U_{X_3<X_4 \mid X_1=X_2<X_3} = 12$		$U_{X_3<X_4 \mid X_1<X_2=X_3} = 12$		$U_{X_3<X_4 \mid X_1=X_2=X_3} = 4$	
$U_{X_3=X_4 \mid X_1<X_2<X_3} = 2$		$U_{X_3=X_4 \mid X_1=X_2<X_3} = 0$		$U_{X_3=X_4 \mid X_1<X_2=X_3} = 3$		$U_{X_3=X_4 \mid X_1=X_2=X_3} = 1$	

Then, the HUM^4 is obtained that

$$\begin{aligned}\text{HUM}^4_{MW} &= \frac{1}{n_1n_2n_3n_4}\Big[U_{X_3<X_4|X_1<X_2<X_3} + \frac{1}{2}U_{X_3=X_4|X_1<X_2<X_3} + \frac{1}{2}U_{X_3<X_4|X_1=X_2<X_3} + \frac{1}{2^2}U_{X_3=X_4|X_1=X_2<X_3} \\ &\quad + \frac{1}{2}U_{X_3<X_4|X_1<X_2=X_3} + \frac{1}{2^2}U_{X_3=X_4|X_1<X_2=X_3} + \frac{1}{2^2}U_{X_3<X_4|X_1=X_2=X_3} + \frac{1}{2^3}U_{X_3=X_4|X_1=X_2=X_3}\Big] \\ &= \frac{1}{4\times4\times6\times5}\left[130 + \frac{2}{2} + \frac{12}{2} + \frac{0}{4} + \frac{12}{2} + \frac{3}{4} + \frac{4}{4} + \frac{1}{8}\right] = 0.3018.\end{aligned}$$

5. Conclusion

The Mann-Whitney test statistic is used to compare two location parameters of discrete random samples. The conditional Mann-Whitney statistics are proposed in order to compare more than equal to three random samples. Whereas the Mann-Whitney statistic, $U_{X_1<X_2}$, may be defined as $\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} I(X_{1i} < X_{2j})$ from two random samples $\{X_{1i}\}$ and $\{X_{2j}\}$ of sizes n_1 and n_2, respectively, the conditional Mann-Whitney statistic, $U_{X_2<X_3|X_1<X_2}$, could be defined as $\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} I(X_{2j} < X_{3k}|X_{1i} < X_{2j})$ from two random samples $\{X_{3k}\}$ and $\{X_{2j};\ X_{1i} < X_{2j}\}$, where the subsample $\{X_{2j};\ X_{1i} < X_{2j}\}$ is collected satisfying the state $\{X_{1i} < X_{2j}\}$ from two random samples $\{X_{1i}\}$ and $\{X_{2j}\}$.

It is known that the Mann-Whitney statistic, $U_{X_1<X_2}$, can be used to define the conditional probability $P(X_1 < X_2)$. Moreover it is found that the conditional Mann-Whitney statistic, $U_{X_2<X_3|X_1<X_2}$, is used to derive the probability $P(X_2 < X_3|X_1 < X_2)$.

With the similar argument that AUC has a linear relation with the Mann-Whitney statistics, $U_{X_1<X_2}$ and $U_{X_1=X_2}$, it might be concluded that VUS could be represented with conditional Mann-Whitney statistics, $U_{X_2<X_3|X_1<X_2}$, $U_{X_2=X_3|X_1<X_2}$, $U_{X_2<X_3|X_1=X_2}$ and $U_{X_2=X_3|X_1=X_2}$. In addition, the HUM with more than three random variables is proposed to define with the conditional Mann-Whitney statistics in this work.

The Mann-Whitney statistic for two random samples is also known to have a linear relation with Wilcoxon rank sum statistic; therefore, it might be derived that the conditional Mann-Whitney statistic for more than two random samples has a relationship with some modified Wilcoxon rank sum statistic in a further study.

References

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.

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.

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*, 82–86.

Fawcett, T. (2003). ROC graphs: notes and practical considerations for data mining researchers, Available from: http://www.hpl.hp.com/techreports/2003/HPL-2003-4.pdf.

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.

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

Hong, C. S. (2009). Optimal threshold from ROC and CAP curves, *Communications in Statistics -*

Simulation and Computation, **38**, 2060–2072.
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. and Jung, D. G. (2014). Standard criterion of hypervolume under the ROC manifold, *Journal of the Korean Data & Information Science Society*, **25**, 473–483.
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.
Li, J. and Fine, J. P. (2008). ROC analysis with multiple classes and multiple tests: Methodology and its application in microarray studies, *Biostatistics*, **9**, 566–576.
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, *Proceedings of SPIE*, **5749**, 581–589.
Provost, F. and Fawcett, T. (2001). Robust classification for imprecise environments, *Machine Learning*, **42**, 203–231.
Rosset, S. (2004). Model Selection via the AUC, In *Proceedings of the 21st International Conference on Machine Learning (ICML'04)*, Banff, Canada, 89.
Scureld, B. K. (1996). Multiple-event forced-choice tasks in the theory of signal detectability, *Journal of Mathematical Psychology*, **40**, 253–269.
Sobehart, J. and Keenan, S. (2001). Measuring default accurately, *Risk - Credit Risk Special Report*, **14**, 31–33.
Swets, J. A. (1988). Measuring the accuracy of diagnostic systems, *Science*, **240**, 1285–1293.
Swets, J. A., Dawes, R. M. and Monahan, J. (2000). Better decisions through science, *Scientific American*, **283**, 82–87.
Wandishin, M. S. and Mullen, S. J. (2009). Multiclass ROC analysis, *Weather and Forecasting*, **24**, 530–547.
Zou, K. H., O'Malley, A. J. and Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models, *Circulation*, **115**, 654–657.

Received December 16, 2014; Revised February 4, 2015; Accepted February 5, 2015