

Null Distribution of Volume Under Ordered Three-Class ROC Surface (VUS) with Continuous Measurements

Shun Liu, Xu Sun *Member, IEEE*, Weichao Xu*, *Member, IEEE*, Yun Zhang, and Jisheng Dai

Abstract—Receiver operating characteristic (ROC) analysis has become an indispensable tool in medical care, with a major application to characterizing the performance of binary diagnostic tests in clinical practice. In many circumstances, however, the diagnostic test has three outcomes, that is, the abnormalities are two-sided. To deal with this scenario, this paper develops a recursive algorithm for computing the exact null distribution of the volume under the ordered three-class receiver operating characteristic surface (VUS) for samples following continuous distributions. Based on the asymptotic normality, an approximately normal distribution with exact mean and variance is also proposed, which is hoped to be useful for large-sample scenarios. Moreover, an efficient rank-based formula, in linearithmic time, is established for nonparametric estimation of VUS. Monte Carlo simulations verify the usefulness of the theoretical and algorithmic findings in this work.

Index Terms—Null distribution, Three-class ROC surface, Volume under the ROC surface (VUS).

I. INTRODUCTION

Originated from the detection theory in early 1950s [1]–[4], receiver operating characteristic (ROC) analysis has become an indispensable tool in medical care, with a major application to characterizing the performance of binary diagnostic tests in clinical practice [5]. Essentially, ROC analysis is a supervised methodology requiring the prior knowledge of the sample membership (positive vs. negative). Given such knowledge, an ROC curve, which is a plot of false-positive rate against true-positive rate, can be defined according to various decision threshold settings [6]–[8]. The area under the curve (AUC) can then be computed, either analytically or empirically, as a figure of merit to summarize the overall performance of the associated binary diagnostic test [5], [9], [10].

In many circumstances, however, the diagnostic test has three outcomes, i.e., the abnormalities are two-sided. For instance, the heart signal is, according to the heart rhythm, classified as Bradycardia (slower rhythm), normal, and Tachycardia (faster rhythm); and the blood pressure, based on the

readings of hemomanometer, falls into three categories, as Hypotension (lower pressure), normal, and Hypertension (higher pressure). For such cases with three ordered alternatives, ROC curve and AUC have been extended in parallel to ROC surface and volume under the surface (VUS) in the literature [11].

Despite the extensive studies concerning VUS in the ordered three-class case [12]–[17], an important problem, i.e., the exact null distribution of VUS, remains unaddressed to the best of our knowledge. The distributional information under the null assumption is mandatory when performing a hypothesis test, where false positive probability must be accurately controlled. Motivated by this unsatisfactory situation, in this work we focus on the null distribution of VUS. The contribution of this letter is the development of a recursive algorithm capable of computing the exact null distribution of VUS.

The rest of this paper is structured as follows. Section II describes the basic concepts on the volume under three-class ROC surface (VUS), as well the unbiased estimate for the variance of VUS. Sections III and IV are devoted to establishing the exact and approximate null distributions, respectively. Section V presents numerical results concerning all the theoretical and algorithmic findings achieved. Finally in Section VI we draw our conclusion on this work.

II. VOLUME UNDER ROC SURFACE

A. Probabilistic interpretation

Let $\{X_i\}_{i=1}^m$, $\{Y_i\}_{i=1}^n$ and $\{Z_k\}_{k=1}^l$ be independent and identically distributed (i.i.d) samples drawn from three universes with continuous cumulative distribution functions (cdf). As elaborated in the literature [11], [12], [14], the probability

$$\theta = \Pr(X < Y < Z) \quad (1)$$

can be interpreted as a volume under an ROC surface (VUS) confined within a unit cube. This expression means that θ is the probability that the three random variables X , Y and Z are in ascending order. Then it follows that $\theta = 1$ if, from left to right, X , Y and Z are completely separable, and $\theta = 1/6$ if X , Y and Z are all overlapped together (the null case).

B. Sample version of θ

From (1), a nonparametric estimator of the VUS can be constructed, as [12], [14]

$$\hat{\theta} \triangleq \frac{1}{mnl} \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^l \mathcal{I}(X_i < Y_j < Z_k) \quad (2)$$

This work was jointly supported in part by National Natural Science Foundation of China (Projects 61771148, U1501251 and 61571211), and in part by Guangzhou Science and Technology Plan (Project 201607010290).

S. Liu, W. Xu, and Y. Zhang are with the Department of Automatic Control, School of Automation, Guangdong University of Technology, Guangzhou, 510006, P. R. China (e-mail: sliu5222@foxmail.com; wxu@gdut.edu.cn; yz@gdut.edu.cn).

X. Sun is with the Baidu, Inc. (e-mail: pamisun@gmail.com).

J. Dai is with the School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, 212013, China (e-mail: jsdai@ujs.edu.cn).

*Corresponding Author.

Tel:+86-20-39322552 Fax:+86-20-39322563.

where the indicator function $\mathcal{I}(\cdot)$ equals unity (zero) if the relationship inside the bracket is true (false). It is easy to verify that $\hat{\theta}$ is an unbiased estimator of θ , namely $E(\hat{\theta}) = \theta$.

C. Fast computation of $\hat{\theta}$

It is obvious that the time complexity of $\hat{\theta}$ in (2) is of cubic order $\mathcal{O}(mnl)$, which is inefficient when the sample sizes are large. However, by a similar technique as in our previous work [9], [10], we can transform (2) into an equivalent form which possesses a linearithmic time complexity.

We first rewrite (2) as

$$\hat{\theta} = \frac{1}{mnl} \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^l \mathcal{H}(Y_j - X_i) \mathcal{H}(Z_k - Y_j) \quad (3)$$

where $\mathcal{H}(t) = 1$ for $t > 0$ and $\mathcal{H}(t) = 0$ for $t \leq 0$. Now we proceed to develop an efficient algorithm based on the relationship between ranks and $\mathcal{H}(\cdot)$. For convenience, write $W_1 \triangleq \{X_i\}_{i=1}^m \cup \{Y_j\}_{j=1}^n$ and $W_2 \triangleq \{Y_j\}_{j=1}^n \cup \{Z_k\}_{k=1}^l$. Let $\mathbb{S}_Y(j)$ ($= j$) denote the rank of Y_j with respect to the Y -array, if Y_j is the j th smallest in the rearranged Y -array which is sorted in ascending order [18]–[20]. Similarly, we denote $\mathbb{T}_Y(j)$ and $\mathbb{R}_Y(j)$ as the ranks of Y_j in the combined W_1 -array and W_2 -array, respectively. Then, from our previous work [9], [10], we have

$$\sum_{i=1}^m \mathcal{H}(Y_j - X_i) = \mathbb{T}_Y(j) - \mathbb{S}_Y(j), \quad (4)$$

$$\begin{aligned} \sum_{k=1}^l \mathcal{H}(Z_k - Y_j) &= \sum_{k=1}^l [1 - \mathcal{H}(Y_j - Z_k)] \\ &= l - [\mathbb{R}_Y(j) - \mathbb{S}_Y(j)]. \end{aligned} \quad (5)$$

A substitution of (4) and (5) into (3) leads readily to

$$\hat{\theta} = \frac{1}{mnl} \sum_{j=1}^n [\mathbb{T}_Y(j) - \mathbb{S}_Y(j)] [l - \mathbb{R}_Y(j) + \mathbb{S}_Y(j)]. \quad (6)$$

The most time-consuming operations in (6) are ranking Y_j 's in Y -array, W_1 -array and W_2 -array, respectively, which can be accomplished in linearithmic time. This means that the formula (6) is much faster than the original version (2), which is in cubic time $\mathcal{O}(mnl)$.

Remark 1. Note that Waegeman et al [14] has developed another linearithmic algorithm for $\hat{\theta}$ based on dynamic programming. See Fig. 1 for the comparison.

III. EXACT NULL DISTRIBUTION OF $\hat{\theta}$

Theorem 1. Under the assumption $F_X = F_Y = F_Z$, the probability mass function of $\hat{\theta}$ defined in (2) is

$$Pr\left(\hat{\theta} = \frac{v}{mnl}\right) = \frac{m!n!l!}{(m+n+l)!} \sum_{u=0}^{mn} S_{m,n,l}(u, v) \quad (7)$$

where $S_{m,n,l}(u, v)$ satisfies the following recursive formula

$$\begin{aligned} S_{m,n,l}(u, v) &= \\ S_{m-1,n,l}(u, v) &+ S_{m,n-1,l}(u-m, v) + S_{m,n,l-1}(u, v-u), \end{aligned} \quad (8)$$

with $u = 0, 1, \dots, mn$ and $v = 0, 1, \dots, mnl$.

Proof. Let $\Omega_{m,n,l}$ denote all strings containing m X 's, n Y 's and l Z 's. Then $\Omega_{m,n,l}$ can be constructed from three disjoint and mutually exclusive sets of substrings $\Omega_{m-1,n,l}$, $\Omega_{m,n-1,l}$ and $\Omega_{m,n,l-1}$, as

$$\Omega_{m,n,l} = \begin{cases} \Omega_{m-1,n,l} + X \\ \Omega_{m,n-1,l} + Y \\ \Omega_{m,n,l-1} + Z \end{cases}. \quad (9)$$

Let $S_{m,n,l}(u, v)$ be the number of strings $\in \Omega_{m,n,l}$ containing u $X < Y$ pairs and v $X < Y < Z$ triplets. Then the following three scenarios concerning the values of u and v happen:

S1) For strings $\in \Omega_{m-1,n,l}$, appending an X at the rear changes neither u nor v , that is,

$$S_{m,n,l}(u, v) = S_{m-1,n,l}(u, v); \quad (10)$$

S2) For strings $\in \Omega_{m,n-1,l}$, appending a Y at the rear does not change v , but increases u by m , that is,

$$\begin{aligned} S_{m,n,l}(u+m, v) &= S_{m,n-1,l}(u, v) \\ \Rightarrow S_{m,n,l}(u, v) &= S_{m,n-1,l}(u-m, v); \end{aligned} \quad (11)$$

S3) For strings $\in \Omega_{m,n,l-1}$, appending a Z at the rear does not change u , but increases v by u , that is,

$$\begin{aligned} S_{m,n,l}(u, v+u) &= S_{m,n,l-1}(u, v) \\ \Rightarrow S_{m,n,l}(u, v) &= S_{m,n,l-1}(u, v-u). \end{aligned} \quad (12)$$

Combining (9)–(12) gives the formula (8).

Noticing that the total number of strings $\in \Omega_{m,n,l}$ is $(m+n+l)!$, and the number of strings $\in \Omega_{m,n,l}$ producing $S_{m,n,l}(u, v)$ is $m!n!l!$, we finally obtain (7) by summing out the auxiliary term u . \square

Remark 2. The null distribution of $\hat{\theta}$ determined by (7) depends only on the sample sizes m , n and l , other than the functional form F of the parent distribution. In other words, $\hat{\theta}$ is distribution-free under the null case.

Algorithm 1: Calculation of $S_{m,n,l}(u, v)$

Data: m, n, l, u, v

Result: $S_{m,n,l}(u, v)$

```

1 begin
2   if  $m = n = l = u = v = 0$  then
3     | return 1
4   end
5   if  $m < 0$  or  $n < 0$  or  $l < 0$  then
6     | return 0
7   end
8   if  $u \notin [0, mn]$  or  $v \notin [0, mnl]$  then
9     | return 0
10  end
11  return  $S_{m-1,n,l}(u, v) + S_{m,n-1,l}(u-m, v) +$ 
       $S_{m,n,l-1}(u, v-u)$ 
12 end
```

Based on Theorem 1, a recursive algorithm for computing $S_{m,n,l}(u, v)$ follows readily, as shown in Algorithm 1 above,

with some necessary stopping rules contained in Lines 1 to 10. The pmf of $\hat{\theta}$ can then be obtained upon substitution of $S_{m,n,l}(u,v)$ based on Algorithm 1 in (7).

Remark 3. *Despite its simplicity for implementation, Algorithm 1 has a time complexity of exponential order $\mathcal{O}(3^{m+n+l}m^2n^2l)$, which makes it soon become computationally intractable, even for sample sizes as small as 10. Computational cost can be reduced if the intermediate results are restored on the disk for later reuse. After such procedure, the time-complexity can be reduced to a polynomial order of $\mathcal{O}(m^3n^3l^2)$.*¹

IV. APPROXIMATE NULL DISTRIBUTION OF $\hat{\theta}$

Thus far we have, on a theoretical and algorithmic level, solved the problem of computing the exact null distribution of $\hat{\theta}$ based on Theorem 1 and Algorithm 1, respectively. However, as remarked before, the recursive algorithm is computationally very inefficient, even with the improvement just mentioned. For large samples, we have to resort to other methods. Fortunately, it follows that, in the null case, i.e., $F_X = F_Y = F_Z$, $\hat{\theta}$ converges in distribution to a normal distribution with $\mathbb{E}(\hat{\theta}) = 1/6$ and variance $\mathbb{V}(\hat{\theta})$ depending only on the sample sizes [21]. Then, for large samples, we only need to focus on the variance of $\hat{\theta}$, which is [21]

$$\sigma_{\text{null}}^2 = \frac{1}{180} \frac{1}{mnl} (4 + 5m + 5l + 2n + 4mn + 4nl + ml). \quad (13)$$

Proof. Let X' , Y' and Z' be i.i.d. copies of X , Y and Z , respectively. Since in the null case, X , Y , Z , X' , Y' and Z' are i.i.d., we have, based on the formulas in [12],

$$\begin{aligned} \mathbb{V}(\hat{\theta}) &= \frac{1}{mnl} [\theta(1-\theta) + (l-1)(q_{12} - \theta^2) \\ &\quad + (n-1)(q_{13} - \theta^2) + (m-1)(q_{23} - \theta^2) \\ &\quad + (n-1)(l-1)(q_1 - \theta^2) \\ &\quad + (m-1)(l-1)(q_2 - \theta^2) \\ &\quad + (m-1)(n-1)(q_3 - \theta^2)] \end{aligned} \quad (14)$$

where

$$\theta = \Pr(X < Y < Z) = \frac{1}{3!} = \frac{1}{6} \quad (15)$$

$$q_{12} = \Pr(X < Y < Z \cap X < Y < Z') = \frac{2}{4!} = \frac{1}{12} \quad (16)$$

$$q_{13} = \Pr(X < Y < Z \cap X < Y' < Z) = \frac{2}{4!} = \frac{1}{12} \quad (17)$$

$$q_{23} = \Pr(X < Y < Z \cap X' < Y < Z) = \frac{2}{4!} = \frac{1}{12} \quad (18)$$

$$q_1 = \Pr(X < Y < Z \cap X < Y' < Z') = \frac{6}{5!} = \frac{1}{20} \quad (19)$$

$$q_2 = \Pr(X < Y < Z \cap X' < Y < Z') = \frac{4}{5!} = \frac{1}{30} \quad (20)$$

$$q_3 = \Pr(X < Y < Z \cap X' < Y' < Z) = \frac{6}{5!} = \frac{1}{20} \quad (21)$$

¹The authors have computed the exact pmfs for sample sizes up to $m = n = l = 20$, and listed into a table the decision thresholds corresponding to various p -values. However, due to its large size (90 pages), the table is not allowed to be provided in the supplemental material. This table is available to interested readers upon request.

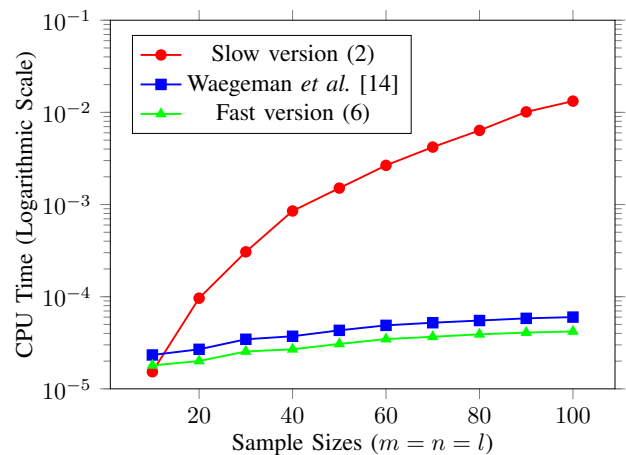


Fig. 1. Comparative results of CPU time between the algorithms of slow version in (2), Waegeman *et al.* in [14] and rank-based version in (6). A log scale is used for better visual effect.

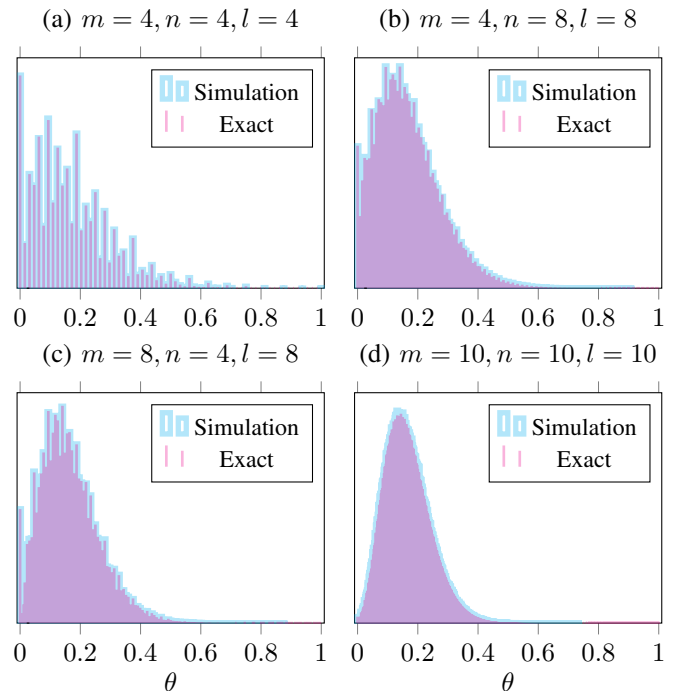


Fig. 2. Exact null distributions calculated based on Theorem 1 and Algorithm 1. For purpose of verification, the exact results are superimposed on empirical histograms from Monte Carlo simulations. Perfect agreements are observed between the exact and simulation results.

which leads readily to (13) upon substitution of (15)–(21) into (14) with some tidying up. \square

From the result (14) just established, we can assume the null distribution follows approximately the normal distribution

$$\hat{\theta} \sim \mathcal{N}\left(\frac{1}{6}, \sigma_{\text{null}}^2\right) \quad (22)$$

when the sample sizes m, n, l are large.

V. NUMERIC RESULTS

This section presents numeric results concerning all the theoretical and algorithmic achievements established in the

previous sections. Results with respect to Algorithm 1 are obtained by C++ language; while others are obtained in Matlab environment on a personal computer. Since the null distribution of $\hat{\theta}$ is *distribution free*, the distributions of X , Y and Z are all set to be standard normal $\mathcal{N}(0, 1)$ for convenience in this work. All samples are generated by functions in Matlab Statistics Toolbox™. For reason of accuracy, the number of Monte Carlo trials are all set to be 10^4 .

A. Fast computation of $\hat{\theta}$

Fig. 1 compares the computational loads between the two sample versions of $\hat{\theta}$ in (2) and (6), and the dynamic programming based algorithm in [14]. For simplicity, the sample sizes of the three classes are set to be identical. It is seen that when the sample size is small, the CPU time of all three versions is comparable. However, with an increase of the sample sizes, the computational time of (2) soars up rapidly, suggesting its inferiority in terms of computational load. Moreover, our rank-based algorithm runs a little bit faster than that in [14]. However, since the difference is so tiny, we do not consider the advantage of our method significant.

B. Exact null distribution for small samples

The exact null distributions of $\hat{\theta}$ along with empirical histograms (background) with respect to different sample sizes are plotted in Fig. 2. The foreground parts of all four subplots are obtained based on our recursive Algorithm 1. It is obvious that, when the sample sizes are small, the exact null distribution is far from normal. Specifically, the null distribution is significantly skewed towards the right. Therefore, the normal approximation cannot be used in these scenarios, especially for samples sizes ≤ 10 .

C. Goodness of normal approximation

Fig. 3 compares the null distribution based on Monte Carlo simulation with the normal approximate version (22) when the sample sizes are relatively large ($m = n = l = 20$, and 50, respectively). It is seen that, for sample sizes ≥ 20 , the null distribution can be well approximated by the normal distribution established in (22), although deviation from normal distribution is still noticeable in Fig. 3 (a).

D. A real example

Before concluding this section, we demonstrate the usefulness of the recursive algorithm developed in this work based on a textbook example [22], which concerns the relationship between sleep deprivation and aggressive behavior. Table I lists the data collected from three groups (each of size 3) on the relationship between the hours of sleep deprivation and the aggression scores assigned by experts [22]. As shown in Table I, there are ties in the data (three 6's), which violates the requirement of continuity of data's parent distribution. To solve this problem, i.e., to break the ties, we added a zero-mean Gaussian noise with a tiny variance (10^{-4} in this work), which is a tie-breaking technique suggested in [23]. After this preprocessing, the distribution of our data becomes continuous.

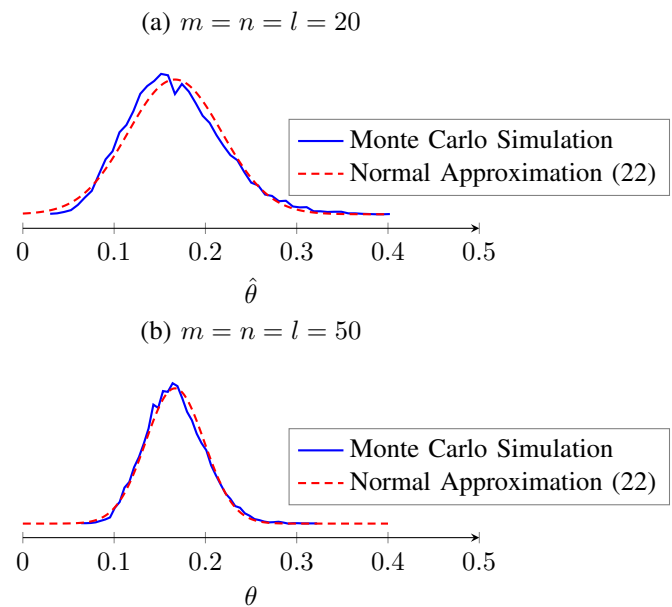


Fig. 3. Illustration of the convergence of $\hat{\theta}$ to normal distribution when the sample sizes are relatively large ($m = n = l = 20, 50$ in this case). Good agreement between simulation and normal approximation can be observed.

TABLE I
AGGRESSION SCORES WITH RESPECT TO TIME OF SLEEP DEPRIVATION

0 hour	24 hour	48 hour
$X_1 = 0$	$Y_1 = 3$	$Z_1 = 6$
$X_2 = 4$	$Y_2 = 6$	$Z_2 = 8$
$X_3 = 2$	$Y_3 = 6$	$Z_3 = 10$

Then, by (2) along with Theorem 1 and Algorithm 1, we obtain $\hat{\theta} = 0.7778$ and the associated p -value being 0.0048 (under the right-tailed test), which is less than 0.05. Therefore, from the point of hypothesis test, the null hypothesis that sleep deprivation affects no aggressive behavior must be rejected, at the significance level of 0.05.

VI. CONCLUSION

This paper proposes a recursive algorithm for computing the exact null distribution of VUS with respect to ordered three-class ROC analysis with continuous measurements. An approximate null distribution with exact mean and variance was also presented, which can be employed when the sample sizes are large. Moreover, an efficient rank-based formula for nonparametric estimation of VUS was established which has a linearithmic time complexity. Simulation studies suggest that: 1) for sample sizes below 10, the null distribution is far from normal. This means that one has to resort to the recursive algorithm to obtain the exact null distribution, in order to accurately control the false positive probability in hypothesis test, 2) it would be safe to use the normal approximation when the sample sizes are over 20, and 3) for sample sizes within 10 to 20, other methodology, such as resampling based one, might be more suitable due to the excessive computational load of the recursive Algorithm 1 (but see the footnote in Remark 3).

REFERENCES

- [1] W. Peterson, T. Birdsall, and W. Fox, "The theory of signal detectability," *IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 171–212, 1954.
- [2] D. Van Meter and D. Middleton, "Modern statistical approaches to reception in communication theory," *IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 119–145, 1954.
- [3] A. Wald, "Statistical decision functions," *The Annals of Mathematical Statistics*, pp. 165–205, 1949.
- [4] J. A. Hanley, "Receiver operating characteristic (ROC) methodology: the state of the art," *Critical Reviews in Diagnostic Imaging*, vol. 29, no. 3, pp. 307–335, 1989.
- [5] D. Mossman, "Three-way ROCs," *Medical Decision Making*, vol. 19, no. 1, pp. 78–89, 1999.
- [6] S. Gopakumar, T. Tran, T. D. Nguyen, D. Phung, and S. Venkatesh, "Stabilizing high-dimensional prediction models using feature graphs," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 1044–1052, May 2015.
- [7] L. Clifton, D. A. Clifton, M. A. F. Pimentel, P. J. Watkinson, and L. Tarassenko, "Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 3, pp. 722–730, May 2014.
- [8] H. Han, L. Li, F. Han, B. Song, W. Moore, and Z. Liang, "Fast and adaptive detection of pulmonary nodules in thoracic ct images using a hierarchical vector quantization scheme," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 648–659, March 2015.
- [9] W. Xu, J. Dai, Y. Hung, and Q. Wang, "Estimating the area under a receiver operating characteristic (roc) curve: Parametric and nonparametric ways," *Signal Processing*, vol. 93, no. 11, pp. 3111–3123, 2013.
- [10] X. Sun and W. Xu, "Fast implementation of delong's algorithm for comparing the areas under correlated receiver operating characteristic curves," *IEEE Signal Processing Letters*, vol. 21, no. 11, pp. 1389–1393, Nov 2014.
- [11] B. K. Scurfield, "Multiple-event forced-choice tasks in the theory of signal detectability," *Journal of Mathematical Psychology*, vol. 40, no. 3, pp. 253 – 269, 1996.
- [12] C. T. Nakas and C. T. Yiannoutsos, "Ordered multiple-class roc analysis with continuous measurements," *Statistics in Medicine*, vol. 23, no. 22, pp. 3437–3449, 2004.
- [13] J. Li and X.-H. Zhou, "Nonparametric and semiparametric estimation of the three way receiver operating characteristic surface," *Journal of Statistical Planning and Inference*, vol. 139, no. 12, pp. 4133 – 4142, 2009.
- [14] W. Waegeman, B. D. Baets, and L. Boullart, "On the scalability of ordered multi-class ROC analysis," *Computational Statistics & Data Analysis*, vol. 52, no. 7, pp. 3371 – 3388, 2008.
- [15] L. Kang, C. Xiong, and L. Tian, "Estimating confidence intervals for the difference in diagnostic accuracy with three ordinal diagnostic categories without a gold standard," *Computational Statistics & Data Analysis*, vol. 68, pp. 326 – 338, 2013.
- [16] L. Kang and L. Tian, "Estimation of the volume under the ROC surface with three ordinal diagnostic categories," *Computational Statistics & Data Analysis*, vol. 62, pp. 39 – 51, 2013.
- [17] L. Tian, C. Xiong, C.-Y. Lai, and A. Vexler, "Exact confidence interval estimation for the difference in diagnostic accuracy with three ordinal diagnostic groups," *Journal of Statistical Planning and Inference*, vol. 141, no. 1, pp. 549 – 558, 2011.
- [18] W. Xu, C. Chang, Y. S. Hung, S. K. Kwan, and P. C. W. Fung, "Order statistics correlation coefficient as a novel association measurement with applications to biosignal analysis," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5552–5563, dec. 2007.
- [19] W. Xu, C. Chang, Y. S. Hung, and P. C. W. Fung, "Asymptotic properties of order statistics correlation coefficient in the normal cases," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2239–2248, Jun. 2008.
- [20] W. Xu, Y. S. Hung, M. Niranjana, and M. Shen, "Asymptotic mean and variance of Gini correlation for bivariate normal samples," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 522–534, Feb. 2010.
- [21] W. Xu, P. Chen, Y. S. Hung, and S. Y. Kung, "A novel three-class ROC method for eQTL analysis," in *2010 International Conference on Machine Learning and Cybernetics*, vol. 6, July 2010, pp. 3056–3061.
- [22] R. S. Witte and W. J. S., *Statistics*, 9th ed. New York: Wiley & Sons, 2010.
- [23] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference*, 3rd ed. New York: M. Dekker, 1992.