



Published in final edited form as:

IEEE Trans Med Imaging. 2008 February ; 27(2): 215–227. doi:10.1109/TMI.2007.905822.

Performance Analysis of 3-class Classifiers: Properties of a 3D ROC Surface and the Normalized Volume Under the Surface for the Ideal Observer

Berkman Sahiner^{*}[Member, IEEE], Heang-Ping Chan, and Lubomir Hadjiiski[Member, IEEE]
Department of Radiology, the University of Michigan, 1500 E. Medical Center Drive, Ann Arbor, MI USA

Abstract

Classification of a given observation to one of three classes is an important task in many decision processes or pattern recognition applications. A general analysis of the performance of three-class classifiers results in a complex six-dimensional (6D) receiver operating characteristic (ROC) space, for which no simple analytical tool exists at present. We investigate the performance of an ideal observer under a specific set of assumptions that reduces the 6D ROC space to 3D by constraining the utilities of some of the decisions in the classification task. These assumptions lead to a 3D ROC space in which the true-positive fraction (TPF) can be expressed in terms of the two types of false-positive fractions (FPFs). We demonstrate that the TPF is uniquely determined by, and therefore is a function of, the two FPFs. The domain of this function is shown to be related to the decision boundaries in the likelihood ratio plane. Based on these properties of the 3D ROC space, we can define a summary measure, referred to as the normalized volume under the surface (NVUS), that is analogous to the area under the ROC curve (AUC) for a two-class classifier. We further investigate the properties of the 3D ROC surface and the NVUS for the ideal observer under the condition that the three class distributions are multivariate normal with equal covariance matrices. The probability density functions (pdfs) of the decision variables are shown to follow a bivariate log-normal distribution. By considering these pdfs, we express the TPF in terms of the FPFs, and integrate the TPF over its domain numerically to obtain the NVUS. In addition, we performed a Monte Carlo simulation study, in which the 3D ROC surface was generated by empirical “optimal” classification of case samples in the multi-dimensional feature space following the assumed distributions, to obtain an independent estimate of NVUS. The NVUS value obtained by using the analytical pdfs was found to be in good agreement with that obtained from the Monte Carlo simulation study. We also found that, under all conditions studied, the NVUS increased when the difficulty of the classification task was reduced by changing the parameters of the class distributions, thereby exhibiting the properties of a performance metric in analogous to AUC. Our results indicate that, under the conditions that lead to our 3D ROC analysis, the performance of a 3-class classifier may be analyzed by considering the ROC surface, and its accuracy characterized by the NVUS.

Index Terms

3-class classification; ROC analysis; performance index; ideal observer

^{*}(phone: 734-674-7429; fax: 734-615-5513; berki@umich.edu).

I. INTRODUCTION

Receiver operating characteristic (ROC) analysis has been a very important tool for the assessment of classification problems in which the ground truth is a binary reference standard (for example, presence or absence of disease), i.e., two-class classification problems. Many problems in pattern recognition involve multi-class classification, where the task can be defined as assigning a given observation vector to one of several classes. Some practical applications in the medical field in recent literature include: (1) diagnosis of oral cancer as high-grade squamous cell carcinoma, low-grade squamous cell carcinoma, leukoplakia, and normal squamous tissue [1]; (2) classification of the voxels in knee MRI volumes as tibial cartilage, femoral cartilage or the background [2]; (3) classification of ECG signals as normal, ventricular premature and supraventricular premature beats [3]; and (4) classification of pigmented skin lesions as benign, dysplastic nevi, or cutaneous melanoma [4]. In all of the applications above, the task of the designed classifier is to solve the multi-class discrimination task, as opposed to solving a number of binary decision tasks. In three of them, the percentage correct classification was used as the performance metric. From ROC analysis, it is known that the percentage correct classification is an incomplete, and sometimes misleading measure. This motivates us to investigate the performance limits for an ideal observer model for three-class classification, and whether a summary measure, analogous to the area under the ROC curve (AUC) in two-class problems can be defined for the three-class task.

The interest on three-class (or, more generally, multi-class) classification has prompted many researchers to investigate how ROC analysis can be generalized to multi-class problems [5–13]. Some of these investigations [7,9–13] studied the three-class ROC problem when the decision variables are related to two likelihood ratios generated by the ideal observer, (see (3) below), while others studied the problem for one or two decision variables that may not necessarily represent the optimal decision variables [5,6,8]. Several of these studies involved assumptions that resulted in simplified ROC surfaces [5–8,12,13], while others investigated the ideal-observer performance in a general 6D ROC space [9–11]. We are interested in the application of three-class classifiers to computer-aided diagnosis (CAD), where many problems involve three possible outcomes. Examples include the classification of a region of interest (ROI) as malignant, normal, or benign in the problem of mammographic lesion detection/characterization, and the classification of a detected mass ROI as a spiculated malignant, non-spiculated malignant, or benign mass in the problem of mammographic lesion characterization.

In this study, we investigate the three-class ROC problem based on the decision variables used by the ideal observer. We reduce the 6D ROC space generated by these decision variables to 3D by constraining the utilities of some of the decisions in the classification task. This approach was previously taken by our research group [7,12], as well as recently by others [13]. The difference between our approach and that of He et al. is that we chose constraints such that the 3D ROC analysis can be performed in terms of the trade-off between the true-positive fraction (TPF) and two false-positive fractions (FPFs), and therefore is analogous to the two-class ROC curve. The specific set of assumptions imposed by He et al. [13] results in a three-class ROC surface in which the axes are the correct decision fractions for the three classes. The area under the ROC curve plays an important role in the evaluation of two-class classifiers. In analogy, we propose a normalized volume under the ROC surface (NVUS) as a figure of merit (FOM) for 3-class classification under our utility constraints. We investigate the properties of the 3D ROC surface and the properties of the NVUS, and present simulation results under the equal covariance Gaussian assumption for the three classes.

II. METHODS

A. 3D ROC surface

In a two-class medical diagnosis task, the ROC curve is defined as the locus of TPF and FPF decision pairs, which represents the tradeoff between the sensitivity and specificity. For a three-class problem, a similar analysis of the tradeoff among different correct and incorrect decisions results in a 5-dimensional ROC hypersurface in a 6-dimensional probability space. A complete characterization of the ROC hypersurface even for simple class distributions is an important but difficult task because of the high dimensionality of the probability space. An avenue that might lead to a practical analysis method, as well as shed light on the general 3-class classification task, is to make simplifying assumptions that focus on a certain useful subclass while still preserving the nature of the problem. A specific problem of interest in the computer-aided diagnosis (CAD) community is that of classifying a case (or an ROI) as malignant, benign, and normal. In this study, we make simplifying assumptions that may result in a practical ROC analysis for the type of problems that satisfy the assumptions, explore the properties of the resulting ROC hypersurface, and define an FOM based on the ROC hypersurface that may be useful for summarizing the classification performance.

In a three-class classification problem, the task is to assign a given D -dimensional observation vector \vec{w} , to one of three classes. In this study, motivated by the classification of cases as malignant, benign, and normal, these classes are denoted by π_M , π_B , and π_N , respectively. Since there are three possibilities of true class membership, denoted by t_M , t_B , and t_N , and three possibilities of class assignment for the observation, denoted by d_M , d_B , and d_N , there are a total of nine distinct true membership-assignment pairs. When \vec{w} is a random vector, the class assignment \mathbf{d} is a random variable, that can take three distinct values, $\mathbf{d} = d_M$, $\mathbf{d} = d_B$, or $\mathbf{d} = d_N$ (random variables are denoted by boldface letters). The probability of $\mathbf{d} = d_i$ given t_j i.e., the probability of deciding that the observation belongs to class π_j when the true class membership is π_j , is denoted by $P_i \equiv P(\mathbf{d} = d_i | t_j)$, where $i \in \{M, B, N\}$ and $j \in \{M, B, N\}$. Since the sum of these probabilities add to unity for each true class ($P_{Mj} + P_{Bj} + P_{Nj} = 1$ for $j \in \{M, B, N\}$), the probability space in a three-class problem is six-dimensional.

In a Bayes' test, one assigns a utility to each of these nine true membership-assignment pairs. Each decision made for a random vector \vec{w} is therefore associated with a utility. The Bayes' test is designed so that, on the average, the utility is as large as possible, i.e., the expected value of the utility is maximized. Let U_{ij} ($i \in \{M, B, N\}$, $j \in \{M, B, N\}$), denote the utility of d_i given t_j i.e., the utility of deciding that the observation belongs to class π_j when the true class membership is π_j . If the random vector \vec{w} belongs to class π_j , the expected utility conditioned on its class membership is

$$E\{U|\pi_j\} = \sum_{i \in \{M, B, N\}} P_{ij} U_{ij}, \quad (1)$$

The expected utility of the decision rule is then given by averaging over all classes π_j

$$E\{U\} = \sum_{j \in \{M, B, N\}} P(\pi_j) \sum_{i \in \{M, B, N\}} P_{ij} U_{ij}, \quad (2)$$

where $P(\pi_j)$ is the *a-priori* probability of class π_j .

Let $p_{\vec{w}}(\vec{w})$ denote the probability density function (pdf) of the random vector \vec{w} , and $p_{\vec{w}|\pi_j}(\vec{w}|\pi_j)$ denote the conditional pdf of \vec{w} given class π_j . Defining

$$\Lambda_{BM}(\vec{w}) = \frac{P_{\vec{w}|\pi_B}(\vec{w}|\pi_B)}{P_{\vec{w}|\pi_M}(\vec{w}|\pi_M)} \quad \text{and} \quad \Lambda_{NM}(\vec{w}) = \frac{P_{\vec{w}|\pi_N}(\vec{w}|\pi_N)}{P_{\vec{w}|\pi_M}(\vec{w}|\pi_M)}, \quad (3)$$

one can show that the application of the Bayes' test leads to the following decision rules [10,14]:

Assign $\mathbf{d} = d_M$ if and only if

$$P(\pi_M)(U_{MM} - U_{BM}) + P(\pi_B)(U_{MB} - U_{BB})\Lambda_{BM}(\vec{w}) + P(\pi_N)(U_{MN} - U_{BN})\Lambda_{NM}(\vec{w}) \geq 0, \quad \text{and} \quad (4)$$

$$P(\pi_M)(U_{MM} - U_{NM}) + P(\pi_B)(U_{MB} - U_{NB})\Lambda_{BM}(\vec{w}) + P(\pi_N)(U_{MN} - U_{NN})\Lambda_{NM}(\vec{w}) \geq 0 \quad (5)$$

Assign $\mathbf{d} = d_B$ if and only if (4) is not satisfied, and

$$P(\pi_M)(U_{BM} - U_{NM}) + P(\pi_B)(U_{BB} - U_{NB})\Lambda_{BM}(\vec{w}) + P(\pi_N)(U_{BN} - U_{NN})\Lambda_{NM}(\vec{w}) \geq 0 \quad (6)$$

Assign $\mathbf{d} = d_N$ if and only if after performing the tests in (4)–(6), neither $\mathbf{d} = d_M$ nor $\mathbf{d} = d_B$ is assigned.

Equations (4)–(6) define the decision boundary lines in the likelihood-ratio plane, defined by $\Lambda_{BM}(\vec{w})$ and $\Lambda_{NM}(\vec{w})$ as the abscissa and ordinate, respectively. These three lines intersect at one point [14], and partition the decision plane (the likelihood-ratio plane) into three distinct decision regions. The relationship among these decision boundary lines has recently been investigated by Edwards et al. [10].

In a 2-class problem, the ROC curve for the ideal observer can be considered as the relationship between the sensitivity and (1-specificity) as the threshold of the likelihood ratio varies, i.e., as the likelihood ratio line is partitioned into two regions. The ROC hypersurface for a 3-class problem can similarly be considered as the relationship among various decision probabilities as the partitioning of the decision plane, defined by (4)–(6), varies. An independent set of three equations can be characterized by three slopes and three intercepts. However, since the decision boundary lines are not independent, but intersect at one point, one degree of freedom is lost, resulting in a five dimensional ROC hypersurface in a six-dimensional probability space.

In this study, we reduce the dimensionality of the ROC hypersurface by imposing constraints on (4)–(6). For the purpose of normalization, we assume that $0 \leq U_{ij} \leq 1$. The utility of a correct decision is assumed to be at a maximum of 1, thus $U_{MM} = 1$, $U_{BB} = 1$, and $U_{NN} = 1$. If a malignant case is misdiagnosed as normal or benign, the utilities will be at a minimum of 0, thus $U_{BM} = 0$, and $U_{NM} = 0$. If a normal case is called benign or vice versa, it may not be very harmful or costly, so that the utilities are still high with $U_{BN} = 1$ and $U_{NB} = 1$. The utilities of misdiagnosing benign or normal to be malignant, U_{MB} and U_{MN} , depend on the further course of action, and will be variable between 0 and 1. Under these conditions, we note that (4) and (5) coincide, and can be written as:

$$P(\pi_M) + P(\pi_B)(U_{MB} - 1)\Lambda_{BM}(\vec{w}) + P(\pi_N)(U_{MN} - 1)\Lambda_{NM}(\vec{w}) \geq 0, \quad (7)$$

while all coefficients in (6) equal zero and it becomes indeterminate. For the assumed values of the utilities, the classification problem reduces to deciding whether the observation belongs to π_M or not, i.e., assign \vec{w} to class M if and only if (7) holds. However, since U_{MB}

and U_{MN} are variable, and are not necessarily equal, there is still a distinction between the benign and normal classes, and the problem is still a three-class classification task.

With these assumptions, the expected utility in (2) becomes a function of only three of the P_{ij} 's, namely, P_{MM} (probability of correctly deciding that the observation is malignant) P_{MB} (probability of falsely deciding that the observation is malignant when the observation belongs to the benign class), and P_{MN} (probability of falsely deciding that the observation is malignant when the observation belongs to the normal class). One can thus analyze how the TPF (P_{MM}) varies as a function of the two FPFs (P_{MB} and P_{MN}). To differentiate the resulting surface in 3D space from both the ROC curve for a 2-class problem (2D ROC curve) and from the general ROC hypersurface for a 3-class problem, we refer to it as the "3D ROC surface". Note that, similar to a 2D ROC curve, our 3D ROC surface describes the variation of the TPF as a function of the FPFs, and results from the simplification of the general ROC hypersurface into 3D because of the assumptions made for the utilities.

We note that the assumptions for the utilities described above are only a special case of the choices for the utilities that will degenerate the three decision lines into a single line. Specifically, if

$$U_{BM}=U_{NM}=\nu_1, U_{BB}=U_{NB}=\nu_2, \text{ and } U_{BN}=U_{NN}=\nu_3 \quad (8)$$

then (4) and (5) coincide, and (6) becomes undefined. These assumptions are summarized in Table 1, and imply that (i) the utility of classifying a malignant tumor as benign is the same as that of classifying it as normal, (ii) the utility of classifying a benign lesion as normal is the same as that of classifying it correctly, and (iii) the utility of classifying a normal region as a benign lesion is the same as that of classifying it correctly. In this case, the expected utility in (2) can be written as a function of P_{MM} (sensitivity), P_{MB} and P_{MN} (two kinds of FPFs):

$$E\{U\}=(U_{MM}-U_{BM})P(\pi_M)P_{MM}+(U_{MB}-U_{BB})P(\pi_B)P_{MB}+(U_{MN}-U_{BN})P(\pi_N)P_{MN}+U_{BM}P(\pi_M)+U_{BB}P(\pi_B)+U_{BN}P(\pi_N)$$

B. The likelihood-ratio plane

Figure 1 describes the decision line provided by (7) in the likelihood ratio plane, where the abscissa and ordinate are Λ_{BM} and Λ_{NM} , respectively. The equation for the decision line can be expressed as

$$\Lambda_{NM}=m\Lambda_{BM}+n, \quad (10)$$

where m and n are the slope and intercept of the line, respectively

$$m=-\frac{P(\pi_B)}{P(\pi_N)}\frac{1-U_{MB}}{1-U_{MN}}, \quad (11)$$

and

$$n=\frac{P(\pi_M)}{P(\pi_N)}\frac{1}{1-U_{MN}}. \quad (12)$$

Since all utilities are by definition less than or equal to unity, we have $-\infty < m < 0$, and $0 < n < \infty$ as the utilities and the *a-priori* class probabilities vary.

From Fig. 1, we observe that for a given \vec{w} we assign $\mathbf{d} = d_M$ if the two likelihood ratios are such that the point $(\Lambda_{BM}(\vec{w}), \Lambda_{NM}(\vec{w}))$ is within the triangle S , and assign $\mathbf{d} = d_M$ otherwise. Since the likelihood ratios are functions of the random vector \vec{w} , $\Lambda_{BM}(\vec{w})$ and $\Lambda_{NM}(\vec{w})$ are random variables, and have a joint pdf under π_M , π_B , and π_N . The TPF and FPFs in which we are interested can be expressed as

$$P_{MM} = \iint_S p_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM} \quad (13)$$

$$P_{MB} = \iint_S p_{\Lambda_{BM}, \Lambda_{NM} | \pi_B}(\Lambda_{BM}, \Lambda_{NM} | \pi_B) d\Lambda_{BM} d\Lambda_{NM} \quad (14)$$

$$P_{MN} = \iint_S p_{\Lambda_{BM}, \Lambda_{NM} | \pi_N}(\Lambda_{BM}, \Lambda_{NM} | \pi_N) d\Lambda_{BM} d\Lambda_{NM} \quad (15)$$

In the ROC plots that we construct, we chose to express the TPF as a function of the two FPFs. To gain insight into the interdependence among these three quantities of interest, we first study how the two FPFs are related. Consider a fixed value of $m = m^*$. By a change of variables $\Lambda_{BM} = u$ and $\Lambda_{NM} = v + m^* u$, (14) can be written as

$$P_{MB} = \int_{v=0}^n dv \int_{u=0}^{-v/m^*} p_{\Lambda_{BM}, \Lambda_{NM} | \pi_B}(u, v + m^* u) du = \int_{v=0}^n f_{B, m^*}(v) dv \quad (16)$$

where

$$f_{B, m^*}(v) = \int_{u=0}^{-v/m^*} p_{\Lambda_{BM}, \Lambda_{NM} | \pi_B}(u, v + m^* u) du. \quad (17)$$

Although the new expression for P_{MB} seems more complicated, it has a meaningful interpretation. $f_{B, m^*}(v)$ can be interpreted as the integral of $p_{\Lambda_{BM}, \Lambda_{NM} | \pi_B}(\Lambda_{BM}, \Lambda_{NM} | \pi_B)$, i.e., the joint pdf of the likelihood ratios under class B, along a line that is parallel to the $\Lambda_{NM} = m^* \Lambda_{BM}$ line, and that intersects the Λ_{NM} axis at $\Lambda_{NM} = v$ (Fig. 2a). P_{MB} is then obtained by integrating $f_{B, m^*}(v)$ up to n , similar to the integration one might perform in 2-class ROC analysis to obtain the false-negative fraction. Similarly, we also express P_{MN} as (Fig. 2b)

$$P_{MN} = \int_{v=0}^n dv \int_{u=0}^{-v/m^*} p_{\Lambda_{BM}, \Lambda_{NM} | \pi_N}(u, v + m^* u) du = \int_{v=0}^n f_{N, m^*}(v) dv \quad (18)$$

where

$$f_{N, m^*}(v) = \int_{u=0}^{-v/m^*} p_{\Lambda_{BM}, \Lambda_{NM} | \pi_N}(u, v + m^* u) du. \quad (19)$$

With this interpretation, it can be seen that as n varies while $m = m^*$ is kept constant, the relationship between P_{MB} and P_{MN} is similar to an ROC curve (Fig. 3) in the FPF plane, referred to as C_{m^*} below. In particular, we observe that $n=0$ implies $P_{MB} = P_{MN} = 0$, and as $n \rightarrow \infty$, $P_{MB} = P_{MN} = 1$.

C. Properties of the ROC surface

In this subsection, we explore two properties of the ROC surface of the ideal observer obtained under the constraints given by (8). Specifically, we show that for a given FPF pair (P_{MB}, P_{MN}) , multiple values for sensitivity (P_{MM}) cannot exist. In other words, P_{MM} is a function of P_{MB} and P_{MN} . We also determine the domain of this function. These properties are used in Section III to numerically evaluate the 3D ROC surface.

Property 1: Let $P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1)$ and $m_1 > m_0$. Then $P_{MN}(m_0, n_0) > P_{MN}(m_1, n_1)$.

This property is proven in Appendix 1. The relative locations of the curves C_{m_0} and C_{m_1} as a result of Property 1 are shown in Fig. 4. Let us define $C_0(P_{MB})$ as the loci of P_{MN} as P_{MB} is varied for $m = 0$, and $C_{-\infty}(P_{MB})$ as the loci of P_{MN} as P_{MB} is varied for $m = -\infty$. Since $-\infty < m < 0$, it can be concluded that the loci of P_{MN} in the FPF plane is bounded from above by the curve $C_{-\infty}(P_{MB})$ and below by the curve $C_0(P_{MB})$.

Property 2: If the pairs (m_0, n_0) and (m_1, n_1) are such that $P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1)$ and $P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1)$, then $P_{MM}(m_0, n_0) = P_{MM}(m_1, n_1)$.

Proof: The proof of this claim is similar to that of Property 1, and is provided in Appendix 2. Property 2 asserts that P_{MM} can be expressed as a function of P_{MB} and P_{MN} , although it may be difficult to express the functional form $\varphi(P_{MB}, P_{MN}) \equiv P_{MM}$ explicitly for given distributions of the malignant, normal, and benign classes.

Property 3: When P_{MB} is held constant, P_{MM} is a non-decreasing function of P_{MN} . Similarly, when P_{MN} is held constant, P_{MM} is a non-decreasing function of P_{MB} .

Proof: Please refer to Appendix 3.

D. Normalized volume under the surface

For a two-class classification task, the AUC is a commonly used FOM to measure the classification accuracy. AUC can be interpreted as the sensitivity averaged over all possible values of specificities. As described above, in a three-class classification task with the utilities assumed to be those in (8), the sensitivity can be expressed as a function of the two FPFs. It is then natural to define an FOM for a three-class classification task as a function of the average sensitivity over all possible values FPF pairs. The height of the surface $P_{MM} = \varphi(P_{MB}, P_{MN})$ is the sensitivity at a given value of the FPF pair (P_{MB}, P_{MN}) . As described above, in general, the FPF pair (P_{MB}, P_{MN}) does not take all possible values in the square $[0, 1] \times [0, 1]$. The total volume V_{total} under the surface P_{MM} therefore does not provide the average sensitivity over all possible values of FPF pairs. The domain (P_{MB}, P_{MN}) is bounded by the curves $C_{-\infty}$ and C_0 in the FPF plane as shown in Fig. 4, and the average sensitivity over all possible values of FPF pairs is the volume under the surface P_{MM} , normalized by the area A_{FPF} bounded by the curves $C_{-\infty}$ and C_0 . The normalized volume under the surface (NVUS) is thus defined as

$$NVUS = \frac{\int_{P_{MB}=0}^1 dP_{MB} \int_{P_{MN}=C_0(P_{MB})}^{C_{-\infty}(P_{MB})} \varphi(P_{MB}, P_{MN}) dP_{MN}}{\int_{P_{MB}=0}^1 dP_{MB} \int_{P_{MN}=C_0(P_{MB})}^{C_{-\infty}(P_{MB})} dP_{MN}} \equiv \frac{V_{total}}{A_{FPF}} \quad (20)$$

The need for normalization may also be deduced by noting that A_{FPF} is related to the separation between classes B and N . If the distributions of these two classes are similar, then the FPF values P_{MB} and P_{MN} will also be similar, regardless of how the decision line is selected. Therefore, both A_{FPF} and the spread of the surface $P_{MM} = \varphi(P_{MB}, P_{MN})$ will be small. On the other hand, relatively dissimilar distributions of the classes B and N will result in a larger value of A_{FPF} and a larger spread of the surface, which will lead to a larger volume under the surface, provided that the separation between classes M and B and between classes M and N remain constant. However, as seen from (9), the separation between classes B and N has no impact on the expected utility, and therefore the effect of the A_{FPF} on V_{total} needs to be removed by normalization.

Gaussian distributions—In the development of parametric ROC analysis for a two-class problem, a frequently made assumption is that the classifier scores (i.e., the decision variable) for the two classes have a binormal distribution. One can then either fit a maximum likelihood curve to the classifier scores [15] or to the likelihood ratio [16]. Although we do not attempt to fit surfaces to our 3D ROC plots in this work, we believe that it is important to examine the shape of the ROC surface and to demonstrate the feasibility of our approach to using NVUS as a performance metric for the important sub-category of Gaussian distributed data. We thus assume that the observation \vec{w} is normally distributed under each class, π_M , π_B , and π_N . To facilitate the analysis, we also assume a simple case in which the covariance matrices of the Gaussian distribution under the three classes are equal, and the mean vectors are different. Without loss of generality, one can then diagonalize the covariance matrices, and translate the distributions under the three classes so that the mean for class π_A is 0 (Appendix 4). Then, under the equal covariance assumption, one can write

$$\begin{aligned} \vec{w} &\sim N(0, I) \text{ under } \pi_M \\ \vec{w} &\sim N(\vec{\mu}_B, I) \text{ under } \pi_B \\ \vec{w} &\sim N(\vec{\mu}_N, I) \text{ under } \pi_N \end{aligned} \quad (21)$$

where $\vec{\mu}_B$ and $\vec{\mu}_N$ are the mean vectors under π_B and π_N , respectively, and I is the identity matrix. The likelihood ratios are then

$$\Lambda_{BM}(\vec{w}) = \frac{P_{\vec{w}|\pi_B}(\vec{w}|\pi_B)}{P_{\vec{w}|\pi_M}(\vec{w}|\pi_M)} = \exp \left\{ -\frac{1}{2} \left[(\vec{w} - \vec{\mu}_B)^T (\vec{w} - \vec{\mu}_B) - \vec{w}^T \vec{w} \right] \right\} = \exp \left\{ \vec{\mu}_B^T \vec{w} - \vec{\mu}_B^T \vec{\mu}_B \right\}$$

$$\Lambda_{NM}(\vec{w}) = \frac{P_{\vec{w}|\pi_N}(\vec{w}|\pi_N)}{P_{\vec{w}|\pi_M}(\vec{w}|\pi_M)} = \exp \left\{ -\frac{1}{2} \left[(\vec{w} - \vec{\mu}_N)^T (\vec{w} - \vec{\mu}_N) - \vec{w}^T \vec{w} \right] \right\} = \exp \left\{ \vec{\mu}_N^T \vec{w} - \vec{\mu}_N^T \vec{\mu}_N \right\}$$

Defining

$$\vec{z} = \log(\vec{\Lambda}) = \begin{pmatrix} \log(\Lambda_{BM}) \\ \log(\Lambda_{NM}) \end{pmatrix}, \quad (24)$$

it can be seen that \vec{z} is a linear function of \vec{w} , with

$$\vec{z} = P\vec{w} + K, \quad (25)$$

where

$$P = \begin{pmatrix} \vec{\mu}_B^T \\ \vec{\mu}_N^T \end{pmatrix}, \quad (26)$$

and

$$K = -\frac{1}{2} \begin{pmatrix} \vec{\mu}_B^T \vec{\mu}_B \\ \vec{\mu}_N^T \vec{\mu}_N \end{pmatrix}. \quad (27)$$

Therefore, \vec{z} is Gaussian under all three classes with covariance matrix

$$\Sigma_z = PP^T = \begin{pmatrix} \vec{\mu}_B^T \vec{\mu}_B & \vec{\mu}_B^T \vec{\mu}_N \\ \vec{\mu}_B^T \vec{\mu}_N & \vec{\mu}_N^T \vec{\mu}_N \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad (28)$$

where

$$\sigma_1 = \sqrt{\vec{\mu}_B^T \vec{\mu}_B}, \sigma_2 = \sqrt{\vec{\mu}_N^T \vec{\mu}_N}, \text{ and } r = \frac{\vec{\mu}_B^T \vec{\mu}_N}{\sqrt{\vec{\mu}_B^T \vec{\mu}_B} \sqrt{\vec{\mu}_N^T \vec{\mu}_N}} \quad (29)$$

Note that σ_1 and σ_2 are the magnitudes of the vectors $\vec{\mu}_B$ and $\vec{\mu}_N$, while r is the cosine of the angle between them. The mean of \vec{z} under classes M, B, and N are calculated as

$$\begin{aligned} \bar{z}_M &= \begin{pmatrix} \bar{z}_M(1) \\ \bar{z}_M(2) \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \end{pmatrix}, \\ \bar{z}_B &= \begin{pmatrix} \bar{z}_B(1) \\ \bar{z}_B(2) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sigma_1^2 \\ 2r\sigma_1\sigma_2 - \sigma_2^2 \end{pmatrix}, \text{ and} \\ \bar{z}_N &= \begin{pmatrix} \bar{z}_N(1) \\ \bar{z}_N(2) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2r\sigma_1\sigma_2 - \sigma_1^2 \\ \sigma_2^2 \end{pmatrix}, \end{aligned} \quad (30)$$

where $\bar{z}_j(i)$, $i = 1, 2$, denotes the i^{th} component of the mean vector of \vec{z} under class j , $j = M, B, N$.

Since $\vec{\Lambda} = \exp(\vec{z})$, $\vec{\Lambda}$ follows a bivariate log-normal distribution. For given values of m and n (defined by (11) and (12)) the sensitivity and false-positive fractions can then be calculated using (13)–(15). For example,

$$P_{MM} = \int_{\Lambda_{BM}=0}^{-n/m} \int_{\Lambda_{NM}=0}^{m\Lambda_{BM}+n} P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM} \quad (31)$$

where

$$P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \frac{1}{\Lambda_{BM}} \frac{1}{\Lambda_{NM}} \exp \left\{ \frac{-1}{2(1-r^2)} \times \left[\frac{(\log \Lambda_{BM} - \bar{z}_M(1))^2}{\sigma_1^2} - \frac{2r(\log \Lambda_{BM} - \bar{z}_M(1))(\log \Lambda_{NM} - \bar{z}_M(2))}{\sigma_1\sigma_2} + \frac{(\log \Lambda_{NM} - \bar{z}_M(2))^2}{\sigma_2^2} \right] \right\}$$

is the joint pdf for the bivariate log-normal random variables Λ_{BM} and Λ_{NM} . The equations for P_{MB} and P_{MN} are similar, with \bar{z}_B and \bar{z}_N replacing \bar{z}_M in (32), respectively:

$$P_{MB} = \int_{\Lambda_{BM}=0}^{-n/m} \int_{\Lambda_{NM}=0}^{m\Lambda_{BM}+n} d\Lambda_{BM} d\Lambda_{NM} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \frac{1}{\Lambda_{BM}} \frac{1}{\Lambda_{NM}} \exp \left\{ \frac{-1}{2(1-r^2)} \times \left[\frac{(\log \Lambda_{BM} - \bar{z}_B(1))^2}{\sigma_1^2} - \frac{2r(\log \Lambda_{BM} - \bar{z}_B(1))(\log \Lambda_{NM} - \bar{z}_B(2))}{\sigma_1\sigma_2} + \frac{(\log \Lambda_{NM} - \bar{z}_B(2))^2}{\sigma_2^2} \right] \right\}$$

$$P_{MN} = \int_{\Lambda_{BM}=0}^{-n/m} \int_{\Lambda_{NM}=0}^{m\Lambda_{BM}+n} d\Lambda_{BM} d\Lambda_{NM} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \frac{1}{\Lambda_{BM}} \frac{1}{\Lambda_{NM}} \exp \left\{ \frac{-1}{2(1-r^2)} \times \left[\frac{(\log \Lambda_{BM} - \bar{z}_N(1))^2}{\sigma_1^2} - \frac{2r(\log \Lambda_{BM} - \bar{z}_N(1))(\log \Lambda_{NM} - \bar{z}_N(2))}{\sigma_1\sigma_2} + \frac{(\log \Lambda_{NM} - \bar{z}_N(2))^2}{\sigma_2^2} \right] \right\}$$

It is seen that for given m and n , the TPF and the FPFs can be completely described in terms of three parameters, σ_1 , σ_2 , and r . Since the 3D ROC surface is obtained by varying m and n over their allowable range ($-\infty < m < 0$, and $0 < n < \infty$), the 3D ROC surface is a function of only these three parameters.

Under the equal covariance Gaussian assumption, the upper and lower bounds for P_{MN} on the FPF plane ($C_{-\infty}(P_{MB})$ and $C_0(P_{MB})$, respectively) can be determined analytically (Appendix 5):

$$C_{-\infty}(P_{MB}) = \Phi \left[\Phi^{-1}(P_{MB}) + \frac{\bar{z}_B(1) - \bar{z}_N(1)}{\sigma_1} \right] \quad (35)$$

and

$$C_0(P_{MB}) = \Phi \left[\Phi^{-1}(P_{MB}) + \frac{\bar{z}_B(2) - \bar{z}_N(2)}{\sigma_2} \right] \quad (36)$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \quad (37)$$

III. SIMULATION RESULTS

We have performed simulation studies under the equal covariance Gaussian assumption in order to study some properties of the ROC surface and compare the NVUS for different

values of the parameters σ_1 , σ_2 , and r . We also implemented two different methods to construct the ROC surface, and compared the NVUS values obtained by these two methods.

We first investigated the characteristics of the C_m curves, i.e., the relationship between P_{MB} and P_{MN} for a given m . For this purpose, we numerically evaluated (33) and (34) for given n and m values to obtain P_{MB} and P_{MN} ; and plotted the relationship between P_{MB} and P_{MN} as n is varied and $\theta = \tan^{-1}(m)$ is fixed. The resulting plots for different values of the parameters σ_1 , σ_2 , and r are shown in Fig. 5 for $\theta = -90^\circ, -67.5^\circ, -45^\circ, -22.5^\circ$, and 0° . The curves for $\theta = -90^\circ$ and 0° (defined as $C_{-\infty}(P_{MB})$ and $C_0(P_{MB})$ above) are important because they define the domain of $P_{MM} = \varphi(P_{MB}, P_{MN})$. It is seen that this domain can take a number of different convex and non-convex shapes, and the area of the domain increases as r decreases. In the limiting case where μ_B and μ_N are equal ($r=1$ and $\sigma_1 = \sigma_2$), there is no distinction between the classes π_B and π_N , and the domain of the ROC surface reduces to the diagonal line that passes through (0,0) and (1,1) in the (P_{MB}, P_{MN}) plane and the ROC surface reduces to the 2-class ROC curve.

Fig. 6 shows examples of the ROC surfaces obtained for two sets of values for the ROC parameters. In Fig. 6a, $r=0.5$ and $\sigma_1^2=\sigma_2^2=3$ so that the resulting ROC surface is symmetric with respect to the $P_{MB} = P_{MN}$ line. In Fig. 6b, $r=0.5$, $\sigma_1^2=1$ and $\sigma_2^2=3$, so that the surface is not symmetric, although the domain is still convex. Figures 6c and 6d show the same surfaces as grayscale images, where the brightness of the pixel is proportional to the value of P_{MM} . From these examples, it is observed that P_{MM} is a non-decreasing function of P_{MB} and P_{MN} , which is a result of Property 3.

To evaluate the NVUS for different values of the parameters σ_1 , σ_2 , and r , we implemented two methods. In Method 1, which uses (31) to (34) for the pdf of \mathbf{A} directly, we numerically integrated the volume over the domain of $P_{MM} = \varphi(P_{MB}, P_{MN})$. For this purpose, we used recursive adaptive Simpson quadrature [17], in the form that it was implemented in MATLAB for double integration. Adaptive Simpson quadrature attempts to minimize the number of function calculations by using small subdivisions of the interval only where required for the desired tolerance [18]. To perform the integration, one needs to numerically express P_{MM} as a function of P_{MB} and P_{MN} . For a given pair of P_{MB} and P_{MN} , we first numerically solved the non-linear equations (33) and (34) to determine which values of m and n result in the given FPFs. We then substituted these values into (31) and performed numerical integration to obtain P_{MM} . The desired tolerance in the recursive adaptive Simpson quadrature was set as 10^{-5} for both the integration to find P_{MM} and to find the volume under the surface. The area of the domain of $\varphi(P_{MB}, P_{MN})$ was likewise obtained by numerically integrating the area between the curves given by (35) and (36).

In Method 2, we used a Monte-Carlo method to compute the NVUS. For given values of the parameters σ_1 , σ_2 , and r , $N/3$ observations were randomly generated from each of the three classes π_M , π_B , and π_N according to the pdfs given by (21). We then computed the likelihood ratios Λ_{BM} and Λ_{NM} using (22) and (23) for each of the N observations. For a given value of the pair (m,n) , the three quantities of interest (P_{MM} , P_{MB} and P_{MN}) can then be determined by how many $(\Lambda_{BM}, \Lambda_{NM})$ pairs from each class fall into the shaded region in Fig. 1. By systematically varying m between 0 and $-\infty$, and n between 0 and ∞ , one can generate pairs (P_{MB}, P_{MN}) that cover the region between $C_{-\infty}(P_{MB})$ and $C_0(P_{MB})$, and the corresponding values of the sensitivity P_{MM} . To find the volume under the surface and the area of the domain of the 3D ROC surface, we used Delaunay triangulation, as described in Appendix 6. Since our purpose was to compare with Method 1 rather than to evaluate the small-sample dependence of the Monte-Carlo method, we used a large number ($N=300,000$) for the number of observations, which was experimentally chosen as a trade-off between computational efficiency and the accuracy of the estimated volume. One hundred fifty

different values were used for both m and n to generate (P_{MB}, P_{MN}) pairs that cover the domain of the 3D ROC surface.

Fig. 7 plots the NVUS results for different values of σ_1 when $\sigma_2^2=1$. Three plots are shown for $r=-0.5$, $r=0$, and $r=0.5$, for both Methods 1 and 2. Figure 8 is the same as Fig. 7, except that $\sigma_2^2=6$. Similar trends were observed for other values of σ_2^2 between 1 and 9. Methods 1 and 2 result in very similar NVUS values. For fixed r and σ_2 , NVUS increases with σ_1 . Also, comparing Figs. 7 and 8, it is seen that for fixed r and σ_1 , NVUS increases with σ_2 . For fixed values of σ_1 and σ_2 , the dependence of NVUS values on r is shown in Fig. 9.

IV. DISCUSSION

From the results presented in Figs. 7–9, it can be observed that the trends for the NVUS are consistent with the characteristics of a proper FOM, namely, as the difficulty of the classification problem decreases, the value of the FOM for the optimal classifier increases. First, NVUS increases with σ_1 for fixed r and σ_2 . From the definitions of these parameters, if σ_1 increases while the other parameters are fixed, the difference in the magnitude between the mean vectors of the malignant and benign classes increases. If the malignant and benign classes are farther apart, then the classification problem is easier and the value of the FOM should increase. Similarly, if σ_2 increases while the other parameters are fixed, the value of the FOM should increase, which is the trend observed by comparing Figs. 7 and 8. Finally, if r increases while σ_1 and σ_2 are fixed, the angle between the vectors μ_B and μ_N decreases, and the normal and benign classes become more similar to each other. Therefore, if all the other parameters are fixed while r increases, the portion of the malignant class that overlaps with the union of the normal and benign class distributions decreases. For large r , if the classifier works well for separating the normal and malignant classes, then it will have a similar performance in separating the benign and malignant classes as well. In contrast, if the angle between the vectors μ_B and μ_N is large (small r), then the benign and normal classes are quite different, hence a classifier that works well to differentiate the normal and malignant classes may not work well to differentiate the benign and malignant classes. In summary, a larger r increases the separation between the malignant class and the other two classes, while it decreases the separation between the benign and normal classes. Due to our choice of the utilities, the separation between the benign and normal classes has no effect on the expected utility (9), and the value of the NVUS is determined by the separation between the malignant and the other two classes. We therefore expect that the value of the FOM would usually increase with r , as demonstrated by the trend for the NVUS in Fig. 9.

With the utilities given as in (8), the decision boundary between the benign and normal classes is indeterminate. The expected utility in (9) does not contain terms related to the misclassification probabilities P_{BN} and P_{NB} , therefore, the decision between classes B and N does not affect the task performance. However, the expected utility, and thus the task performance, is uniquely determined by P_{MM} , P_{MB} and P_{MN} . Furthermore, since U_{MB} and U_{MN} are not necessarily equal, there is still a distinction between the benign and normal classes, and the problem remains a three-class classification task.

As in 2D ROC analysis, comparison of 3D ROC surfaces and the summary measures for two classification tasks requires caution. Similar to the condition that two 2D ROC curves may intersect, the 3D ROC surfaces for the ideal observer for two classification tasks may also intersect. Under this condition, one does not know whether the sensitivity for the task with the larger NVUS value is higher than that with the smaller NVUS value at all FPF values. An additional need for caution for the comparison of the NVUS values of two 3D ROC surfaces results from the possibility that the FPF domains for the two tasks may be different, as shown in Fig. 5 for six different tasks. Consider two tasks T_1 and T_2 , and

assume that the domain of the 3D ROC surface for T_1 is larger than and encloses the entire domain for T_2 . It is conceivable that the sensitivity of the ideal observer may be higher for T_1 than T_2 at all FPF values in the domain of T_2 , and despite this, the NVUS value for T_1 can be lower if the sensitivity for T_1 at FPF values outside the domain of T_2 is very low.

In our previous [7,12] and current studies, our assumptions for the utilities were that $U_{MM} = U_{BB} = U_{NN} = U_{BN} = U_{NB} = 1$, $U_{BM} = U_{NM} = 0$, and that U_{MB} and U_{MN} are variable in $[0,1]$. These assumptions simplify the formulations of the slope and intercept of the decision line, (Fig. 1 and (11) and (12)) and facilitate the subsequent analyses. However, the conditions on the utilities can be more general, namely $U_{BM} = U_{NM}$, $U_{BB} = U_{NB}$, and $U_{BN} = U_{NN}$ (8), and the three decision lines of the ideal observer will still degenerate into a single line that separates the malignant class from the other two classes. With the additional assumption that the utility of an incorrect decision is not larger than that of a correct decision (i.e., $U_{BM} \leq U_{MM}$, $U_{MB} \leq U_{BB}$, and $U_{MN} \leq U_{NN}$), the slope and intercept for this line will satisfy $-\infty < m < 0$ and $0 < n < \infty$, respectively. Since the properties of the 3D ROC surface found in this study depend only on the situation of a single decision line separating the malignant class from the other two classes in the likelihood ratio plane, the conclusions of our study is applicable to the more general condition given by (8).

V. CONCLUSION

In this study, we investigated the 3-class classification problem under some simplifying, but realistic, assumptions for the utilities that result in a single decision line in the likelihood ratio plane separating the malignant decisions from “non-malignant” decisions. Under these assumptions (8), we showed that the sensitivity of the ideal observer can be expressed as a function of two false-positive fractions, thereby one can define a 3D ROC surface above the FPF plane, and that the domain of the ROC surface in the FPF plane depends on the class distributions. We extended the definition of the area under the ROC curve for a two-class problem to a normalized volume under the ROC surface for a three-class problem, where the normalization is required because of the dependence of the domain of the 3D ROC surface on the classification task.

We studied the properties of the 3D ROC surface and the NVUS for the ideal observer under the condition that the three-class distributions in the feature space are multivariate normal with equal covariance matrices. Under this assumption for class distributions, it can be shown that the probability density functions of the likelihood ratios follow a bivariate log-normal distribution, and that they depend only on three parameters. We investigated two methods for calculating the NVUS, and found that the results were in good agreement. We also found that the NVUS satisfies some of the intuitive properties of a proper FOM for characterizing the classification performance. Although the analysis in this study was performed under simplifying assumptions, we expect that it will provide guidance for the challenging problem of designing performance metric and analysis for 3-class classification tasks.

Acknowledgments

This work was supported in part by USPHS grant CA95153 and U. S. Army Medical Research and Materiel Command grant DAMD17-02-1-0214.

APPENDIX 1

We first show that

$$P_{MB} = \iint_S \Lambda_{BM} P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM} \quad (\text{A.1})$$

and

$$P_{MN} = \iint_S \Lambda_{NM} P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM}. \quad (\text{A.2})$$

Let $\Lambda_{BM} = f_1(\vec{w})$, $\Lambda_{NM} = f_2(\vec{w})$, and let $\Omega(m, n)$ denote the set of vectors \vec{w} that satisfy $f_2(\vec{w}) - mf_1(\vec{w}) - n < 0$. Noting that $P_{\vec{w} | \pi_B}(\vec{w} | \pi_B) = \Lambda_{BM}(\vec{w}) P_{\vec{w} | \pi_M}$, the false positive probability for the benign class can be written as the D-dimensional integral

$$P_{MB} = \int_{\Omega(m,n)} P_{\vec{w} | \pi_B}(\vec{w} | \pi_B) d\vec{w} = \int_{\Omega(m,n)} \Lambda_{BM}(\vec{w}) P_{\vec{w} | \pi_M}(\vec{w} | \pi_M) d\vec{w} \quad (\text{A.3})$$

The integrand in (A.3) can be integrated by first finding the set of all vectors \vec{w} in the D-dimensional observation space such that

$$\Lambda_{BM} < f_1(\vec{w}) \leq \Lambda_{BM} + \Delta\Lambda_{BM}, \text{ and } \Lambda_{NM} < f_2(\vec{w}) \leq \Lambda_{NM} + \Delta\Lambda_{NM}. \quad (\text{A.4})$$

The sum of all $P_{\vec{w} | \pi_M}(\vec{w} | \pi_M)$ such that (A.4) is satisfied is (by definition) $P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) \Delta\Lambda_{BM} \Delta\Lambda_{NM}$, while $\vec{w} \in \Omega(m, n)$ is equivalent to $\Lambda_{NM} - m\Lambda_{BM} - n < 0$. Integrating over the entire D-dimensional observation space, we find that

$$\begin{aligned} \int_{\Omega(m,n)} \Lambda_{BM}(\vec{w}) P_{\vec{w} | \pi_M}(\vec{w} | \pi_M) d\vec{w} &= \iint_{\Lambda_{NM} - m\Lambda_{BM} - n < 0} \Lambda_{BM} P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \\ &\Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM} = \iint_S \Lambda_{BM} P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \\ &\Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM} \end{aligned} \quad (\text{A.5})$$

Eq. (A.2) can be obtained in a similar manner.

We now show if $P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1)$ and $m_1 > m_0$, then $P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1)$.

For notational simplicity, let us define $x = \Lambda_{BM}$, $y = \Lambda_{NM}$, and $f(x, y) = P_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M)$. First consider Case (A), in which the two lines $y = m_0x + n_0$ and $y = m_1x + n_1$ (i.e., the lines $\Lambda_{NM} = m_0\Lambda_{BM} + n_0$ and $\Lambda_{NM} = m_1\Lambda_{BM} + n_1$) do not intersect in the positive quadrant (Fig. A1). Using (A.1), and noting that $f(x, y)$ is non-negative, $P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1)$ implies $f(x, y) = 0$ in the region D shown in Fig. A1. Using (A.2), this implies $P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1)$.

We now consider Case (B), in which the two lines $y = m_0x + n_0$ and $y = m_1x + n_1$ intersect at a point (x^*, y^*) in the positive quadrant (Fig. A2). Using (A.1), it is seen that

$$P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1) \Rightarrow \iint_{S_0} xf(x, y) dx dy = \iint_{S_1} xf(x, y) dx dy, \quad (\text{A.6})$$

where the two regions S_0 and S_1 are the two shaded triangles in Fig. A2. For any x in S_0 , we have $x^* \leq x$, and for any x in S_1 , we have $x \leq x^*$. Therefore,

$$x^* \iint_{S_0} f(x, y) dx dy \geq \iint_{S_0} xf(x, y) dx dy = \iint_{S_1} xf(x, y) dx dy \geq x^* \iint_{S_1} f(x, y) dx dy, \quad (\text{A.7})$$

which indicates

$$P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1) \Rightarrow \iint_{S_0} f(x, y) dx dy \geq \iint_{S_1} f(x, y) dx dy. \quad (\text{A.8})$$

Similarly, for any y in S_0 , we have $y \leq y^*$, and for any y in S_1 , we have $y^* \leq y$. Therefore,

$$\iint_{S_0} yf(x, y) dx dy \geq y^* \iint_{S_0} f(x, y) dx dy \geq y^* \iint_{S_1} f(x, y) dx dy \geq \iint_{S_1} yf(x, y) dx dy, \quad (\text{A.9})$$

where the middle inequality is due to (A.8). We therefore conclude that

$$\iint_{S_0} yf(x, y) dx dy \geq \iint_{S_1} yf(x, y) dx dy, \quad (\text{A.10})$$

which, along with (A.2), implies that $P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1)$. We also note that $P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1)$ if and only if $f(x, y) = 0$ in $S_0 \cup S_1 \setminus (x^*, y^*)$.

APPENDIX 2

In this appendix, we show that a given (P_{MB}, P_{MN}) pair uniquely determines P_{MM} , i.e., we cannot have more than one sensitivity value for a given FPF pair. Let (m_0, n_0) and (m_1, n_1) be such that $P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1)$ and $P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1)$. Proceeding in a similar fashion to the proof of Property 1, we see that if the lines $y = m_0x + n_0$ and $y = m_1x + n_1$ do not intersect in the positive quadrant, or if $m_1 = m_0$, then $P_{MM}(m_0, n_0) = P_{MM}(m_1, n_1)$. If these two lines intersect at one point in the positive quadrant, then, without loss of generality, let us assume $m_1 > m_0$ so that we can refer to Fig. A2. We have already shown (A.8) that if $m_1 > m_0$, then

$$P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1) \Rightarrow \iint_{S_0} f(x, y) dx dy \geq \iint_{S_1} f(x, y) dx dy \Rightarrow P_{MM}(m_0, n_0) \geq P_{MM}(m_1, n_1) \quad (\text{A.11})$$

Referring to Fig. A2, for any y in S_0 , we have $y \leq y^*$, and for any y in S_1 , we have $y^* \leq y$. Therefore,

$$P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1) \Rightarrow \iint_{S_0} yf(x, y) dx dy = \iint_{S_1} yf(x, y) dx dy \Rightarrow \iint_{S_0} y^* f(x, y) dx dy \leq \iint_{S_1} y^* f(x, y) dx dy \Rightarrow \iint_{S_0} f(x, y) dx dy \leq \iint_{S_1} f(x, y) dx dy$$

from which we conclude

$$P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1) \Rightarrow P_{MM}(m_0, n_0) \leq P_{MM}(m_1, n_1) \quad (\text{A.13})$$

Equations (A.11) and (A.13) imply that $P_{MM}(m_0, n_0) = P_{MM}(m_1, n_1)$.

APPENDIX 3

We will only prove the first part, i.e., if $P_{MB}(m_0, n_0) = P_{MB}(m_1, n_1)$ and $P_{MN}(m_0, n_0) > P_{MN}(m_1, n_1)$, then $P_{MM}(m_0, n_0) > P_{MM}(m_1, n_1)$. The proof of the second part follows similar steps. The assumptions about P_{MB} and P_{MN} imply that $m_1 > m_0$. This is because if m_0 were larger than or equal to m_1 , then, by Property 1, we would have $P_{MN}(m_1, n_1) > P_{MN}(m_0, n_0)$. We also note that the lines $\Lambda_{NM} = m_0\Lambda_{BM} + n_0$ and $\Lambda_{NM} = m_1\Lambda_{BM} + n_1$ intersect in the positive quadrant, because otherwise we would have $P_{MN}(m_0, n_0) = P_{MN}(m_1, n_1)$. Under these conditions, we have already shown that (A.8)

$$\iint_{S_0} p_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM} \geq \iint_{S_1} p_{\Lambda_{BM}, \Lambda_{NM} | \pi_M}(\Lambda_{BM}, \Lambda_{NM} | \pi_M) d\Lambda_{BM} d\Lambda_{NM},$$

which implies that $P_{MM}(m_0, n_0) > P_{MM}(m_1, n_1)$.

APPENDIX 4

In this appendix, we show that, for the equal-covariance Gaussian distributions discussed in Section 2.E., diagonalization and translation do not affect the 3D ROC surface. This is equivalent to proving that a straight line in the original likelihood ratio plane remains a straight line after diagonalization. Let \vec{w}' denote a random vector to be classified before the diagonalization, and let

$$\begin{aligned} \vec{w}' &\sim N(\vec{\mu}'_M, \Sigma') \text{ under } \pi_M \\ \vec{w}' &\sim N(\vec{\mu}'_B, \Sigma') \text{ under } \pi_B \\ \vec{w}' &\sim N(\vec{\mu}'_N, \Sigma') \text{ under } \pi_N \end{aligned} \quad (\text{A.14})$$

The likelihood ratio between π_B and π_M before the diagonalization and translation is given by

$$\begin{aligned}
\Lambda'_{\mathbf{BM}}(\vec{\mathbf{w}}') &= \frac{P_{\vec{\mathbf{w}}'|\pi_B}(\vec{\mathbf{w}}'|\pi_B)}{P_{\vec{\mathbf{w}}'|\pi_M}(\vec{\mathbf{w}}'|\pi_M)} \\
&= \exp \left\{ -\frac{1}{2} \left[(\vec{\mathbf{w}}' - \vec{\mu}'_B)^T \Sigma'^{-1} (\vec{\mathbf{w}}' - \vec{\mu}'_B) - (\vec{\mathbf{w}}' - \vec{\mu}'_M)^T \Sigma'^{-1} (\vec{\mathbf{w}}' - \vec{\mu}'_M) \right] \right\} \\
&= \exp \left\{ -\frac{1}{2} \left(\vec{\mu}'_B{}^T \Sigma'^{-1} \vec{\mu}'_B - \vec{\mu}'_M{}^T \Sigma'^{-1} \vec{\mu}'_M \right) \right\} \exp \left\{ (\vec{\mu}'_B - \vec{\mu}'_M)^T \Sigma'^{-1} \vec{\mathbf{w}}' \right\}
\end{aligned} \tag{A.15}$$

We will show that the likelihood ratio between π_B and π_M after the diagonalization and translation, $\Lambda_{\mathbf{BM}}(\vec{\mathbf{w}})$, is proportional to $\Lambda'_{\mathbf{BM}}(\vec{\mathbf{w}}')$. The transformation for diagonalization and translation between $\vec{\mathbf{w}}$ and $\vec{\mathbf{w}}'$ is

$$\vec{\mathbf{w}} = \mathbf{L}^{-1/2} \mathbf{V}^T (\vec{\mathbf{w}}' - \vec{\mu}'_M), \tag{A.16}$$

where \mathbf{V} is the matrix composed of the eigenvectors of Σ' , and \mathbf{L} is the diagonal matrix constructed from the corresponding eigenvalues. The covariance matrix after the diagonalization is the identity matrix, and the class means are

$$\vec{\mu}'_M = 0, \vec{\mu}'_B = \mathbf{L}^{-1/2} \mathbf{V}^T (\vec{\mu}'_B - \vec{\mu}'_M), \quad \text{and} \quad \vec{\mu}'_N = \mathbf{L}^{-1/2} \mathbf{V}^T (\vec{\mu}'_N - \vec{\mu}'_M) \tag{A.17}$$

The likelihood ratio between π_B and π_M after the diagonalization and translation is

$$\begin{aligned}
\Lambda_{\mathbf{BM}}(\vec{\mathbf{w}}) &= \frac{P_{\vec{\mathbf{w}}|\pi_B}(\vec{\mathbf{w}}|\pi_B)}{P_{\vec{\mathbf{w}}|\pi_M}(\vec{\mathbf{w}}|\pi_M)} \\
&= \exp \left\{ -\frac{1}{2} \left[(\vec{\mathbf{w}} - \vec{\mu}'_B)^T (\vec{\mathbf{w}} - \vec{\mu}'_B) - \vec{\mathbf{w}}^T \vec{\mathbf{w}} \right] \right\} \\
&= \exp \left\{ \vec{\mu}'_B{}^T \vec{\mathbf{w}} - \frac{1}{2} \vec{\mu}'_B{}^T \vec{\mu}'_B \right\} \\
&= \exp \left\{ -\frac{1}{2} \vec{\mu}'_B{}^T \vec{\mu}'_B \right\} \exp \left\{ \vec{\mu}'_B{}^T \vec{\mathbf{w}} \right\} \\
&= \exp \left\{ -\frac{1}{2} \vec{\mu}'_B{}^T \vec{\mu}'_B \right\} \exp \left\{ \vec{\mu}'_B{}^T \mathbf{L}^{-1/2} \mathbf{V}^T (\vec{\mathbf{w}}' - \vec{\mu}'_M) \right\} \\
&= \exp \left\{ -\frac{1}{2} \vec{\mu}'_B{}^T \vec{\mu}'_B - \vec{\mu}'_B{}^T \mathbf{L}^{-1/2} \mathbf{V}^T \vec{\mu}'_M \right\} \cdot \exp \left\{ (\vec{\mu}'_B - \vec{\mu}'_M)^T \mathbf{V} \mathbf{L}^{-1/2} \cdot \mathbf{L}^{-1/2} \mathbf{V}^T \vec{\mathbf{w}}' \right\} \\
&= \exp \left\{ -\frac{1}{2} \vec{\mu}'_B{}^T \vec{\mu}'_B - \vec{\mu}'_B{}^T \mathbf{L}^{-1/2} \mathbf{V}^T \vec{\mu}'_M \right\} \cdot \exp \left\{ (\vec{\mu}'_B - \vec{\mu}'_M)^T \Sigma'^{-1} \vec{\mathbf{w}}' \right\}
\end{aligned} \tag{A.18}$$

Comparing to (A.15), we find that $\Lambda_{\mathbf{BM}}(\vec{\mathbf{w}}) = \mathbf{K}_B \Lambda'_{\mathbf{BM}}(\vec{\mathbf{w}}')$, where \mathbf{K}_B is a constant of proportionality. Similarly, it can be shown that $\Lambda_{\mathbf{NM}}(\vec{\mathbf{w}}) = \mathbf{K}_N \Lambda'_{\mathbf{NM}}(\vec{\mathbf{w}}')$, and therefore a straight line in the original likelihood ratio plane remains a straight line after diagonalization.

APPENDIX 5

We will demonstrate the derivation of (36) here, i.e., the relationship between P_{MB} and P_{MN} for $m=0$. The integrals can be evaluated in a similar manner to find the relationship between P_{MB} and P_{MN} for $m = -\infty$ (35). First, for $m=0$, and with a change of variables

$$\begin{aligned} u &= \log \Lambda_{BM} - \bar{z}_B(1) \\ v &= \log \Lambda_{NM} - \bar{z}_B(2), \end{aligned} \quad (\text{A.19})$$

Eq. (33) can be written as

$$P_{MB} = \int_{v=-\infty}^{\log n - \bar{z}_B(2)} \int_{-\infty}^{\infty} \frac{dudv}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} G(u, v), \quad (\text{A.20})$$

where

$$\begin{aligned} G(u, v) &= \exp \left\{ \frac{-1}{2(1-r^2)} \left[\frac{u^2}{\sigma_1^2} - \frac{2ruv}{\sigma_1\sigma_2} + \frac{v^2}{\sigma_2^2} \right] \right\} \\ &= \exp \left\{ \frac{-1}{2(1-r^2)} \left[\frac{u}{\sigma_1} - \frac{rv}{\sigma_2} \right]^2 - \frac{v^2}{2\sigma_2^2} \right\}. \end{aligned} \quad (\text{A.21})$$

Therefore,

$$\begin{aligned} P_{MB} &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int_{v=-\infty}^{\log n - \bar{z}_B(2)} dv \exp(-v^2/2\sigma_2^2) \cdot \int_{-\infty}^{\infty} \exp \left\{ \frac{-1}{2(1-r^2)} \left[\frac{u}{\sigma_1} - \frac{rv}{\sigma_2} \right]^2 \right\} du \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} \int_{v=-\infty}^{\log n - \bar{z}_B(2)} dv \exp(-v^2/2\sigma_2^2) \\ &= \Phi \left[\frac{\log n - \bar{z}_B(2)}{\sigma_2} \right] \end{aligned} \quad (\text{A.22})$$

Similarly,

$$P_{MN} = \Phi \left[\frac{\log n - \bar{z}_N(2)}{\sigma_2} \right] \quad (\text{A.23})$$

Solving for $\log n$ in terms of P_{MB} in (A.22) and substituting into (A.23), we obtain (36).

APPENDIX 6

Delaunay triangulation is closely related to the Voronoi diagram for a set of Q points in the plane. Given Q points P_q , $q = 1, \dots, Q$, one can partition the plane into Q partitions S_q , $q = 1, \dots, Q$, where S_q^* is the region of the plane that is closer to the point P_q^* than to any of the other $Q-1$ points. S_q^* is called the Voronoi domain of the point q^* . If one draws a line between any two points whose Voronoi domains are adjacent to each other, a set of triangles is obtained, known as the Delaunay triangulation. The Delaunay triangulation has the property that the circumcircle of no triangle contains any points of the triangulation. Figure A3 shows a Delaunay triangulation for $Q=9$ points. The Voronoi cell S_7 for the point P_7 is shown as the shaded polygon. When the domain of the 3D ROC is convex, the area of the domain was approximated by the sum of all such triangles. When the domain was not convex, we excluded from the computation of the area any triangle whose interior contained any point outside the region bounded by $C_{-\infty}(P_{MB})$ and $C_0(P_{MB})$, obtained empirically using the Monte Carlo simulation. The volume under the surface was approximated by

summing up the volume of each prism whose base was given by a Delaunay triangle used in the area estimation.

REFERENCES

1. Majumder SK, Gupta A, Gupta S, Ghosh N, Gupta PK. Multi-class classification algorithm for optical diagnosis of oral cancer. *Journal of Photochemistry and Photobiology B-Biology*. 2006; vol. 85:109–117.
2. Folkesson J, Dam E, Olsen OF, Pettersen P, Christiansen C. Automatic segmentation of the articular cartilage in knee MRI using a hierarchical multi-class classification scheme. *Lecture Notes in Computer Science*. 2005; vol. 3749:327–334.
3. Silipo R, Vergassola R, Zong W, Berthold MR. Knowledge-based and data-driven models in arrhythmia fuzzy classification. *Methods of Information in Medicine*. 2001; vol. 40:397–402. [PubMed: 11776738]
4. Dreiseitl S, Ohno-Machado L, Binder M. Comparing three-class diagnostic tests by three-way ROC analysis. *Medical Decision Making*. 2000; vol. 20:323–331. [PubMed: 10929855]
5. Scurfield BK. Generalization of the theory of signal detectability to n-even t m-dimensional forced-choice tasks. *Journal of Mathematical Psychology*. 1998; vol. 42:5–31. [PubMed: 9606159]
6. Mossman D. Three-way ROCs. *Medical Decision Making*. 1999; vol. 19:78–89. [PubMed: 9917023]
7. Chan HP, Sahiner B, Hadjiiski LM, Petrick N, Zhou C. Design of three-class classifiers in computer-aided diagnosis: Monte Carlo simulation study. *Proceedings of the SPIE - Medical Imaging*. 2003; vol. 5032:567–578.
8. Nakas CT, Yiannoutsos CT. Ordered multiple-class ROC analysis with continuous measurements. *Statistics in Medicine*. 2004; vol. 23:3437–3449. [PubMed: 15505886]
9. Edwards DC, Metz CE, Kupinski MA. Ideal observers and optimal ROC hypersurfaces in N-class classification. *IEEE Transactions on Medical Imaging*. 2004; vol. 23:891–895. [PubMed: 15250641]
10. Edwards DC, Metz CE. Restrictions on the three-class ideal observer's decision boundary lines. *IEEE Transactions on Medical Imaging*. 2005; vol. 24:1566–1573. [PubMed: 16350917]
11. Edwards DC, Metz CE, Nishikawa RM. The hypervolume under the ROC hypersurface of "near-guessing" and "near-perfect" observers in N-class classification tasks. *IEEE Transactions on Medical Imaging*. 2005; vol. 24:293–299. [PubMed: 15754980]
12. Sahiner B, Chan HP, Hadjiiski LM. Performance analysis of 3-class classifiers: Properties of the 3D ROC surface and the normalized volume under the surface. *Proceedings of the SPIE - Medical Imaging*. 2006; vol. 6146:61460C1–61460C7.
13. He X, Metz CE, Tsui BMW, Links JM, Frey EC. Three-class ROC analysis -- A decision theoretic approach under the ideal observer framework. *IEEE Transactions on Medical Imaging*. 2006; vol. 25:571–581. [PubMed: 16689261]
14. Van Trees, HL. *Detection, estimation, and modulation theory*. New York: John Wiley and Sons; 1968.
15. Dorfman DD, Alf E Jr. Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals-rating method data. *Journal of Mathematical Psychology*. 1969; vol. 6:487–496.
16. Metz CE, Pan X. "Proper" binormal ROC curves: Theory and maximum-likelihood estimation. *Journal of Mathematical Psychology*. 1999; vol. 43:1–33. [PubMed: 10069933]
17. Lyness JN. Notes on the adaptive Simpson quadrature routine. *Journal of the Association for Computing Machinery*. 1969; vol. 16:483–495.
18. McKeeman WM. Algorithm 145: Adaptive numerical integration by Simpson's rule. *Communications of the Association for Computing Machinery*. 1962; vol. 5:604.

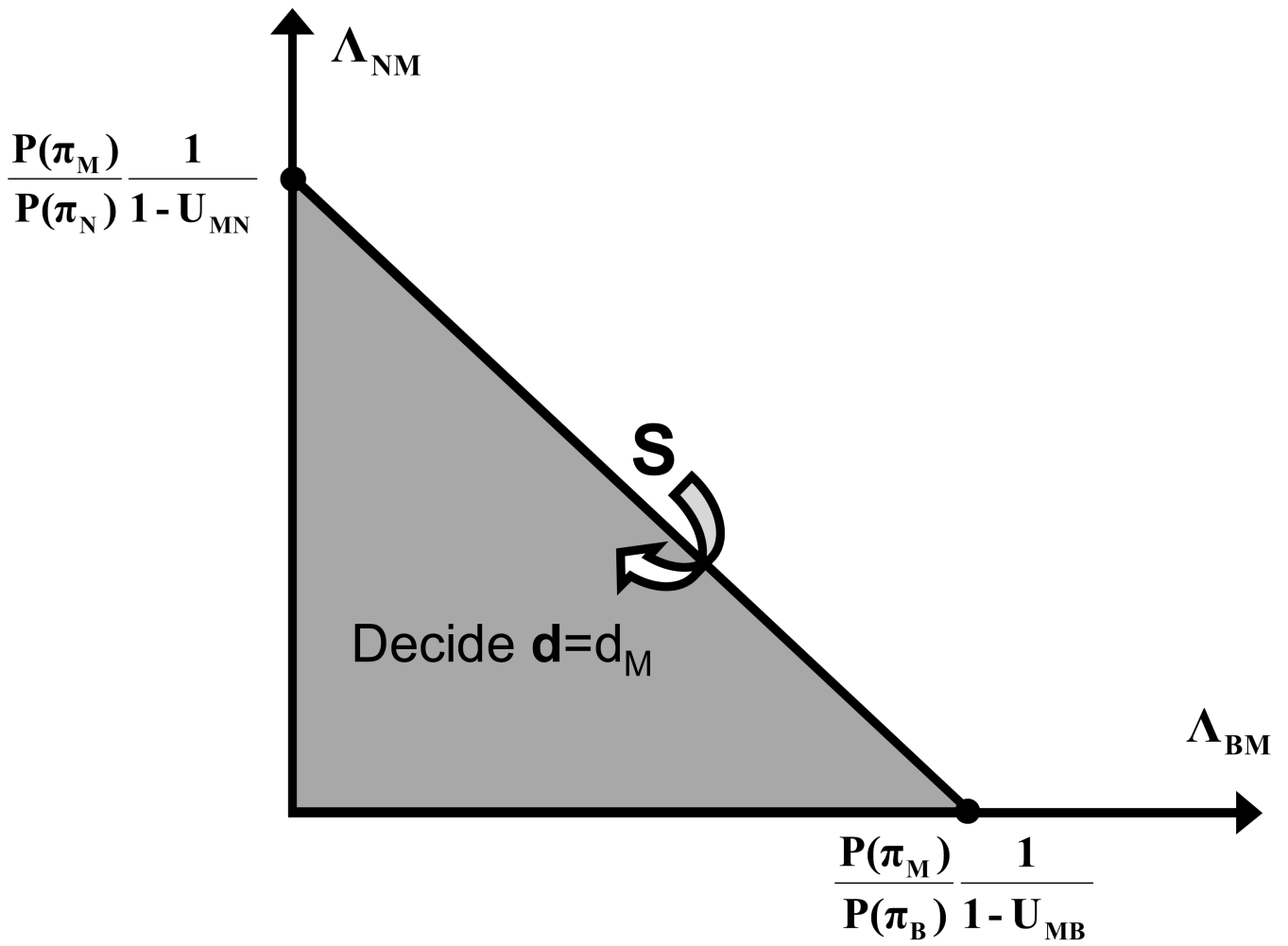


Fig. 1.
The likelihood ratio (LR) plane and the decision line.

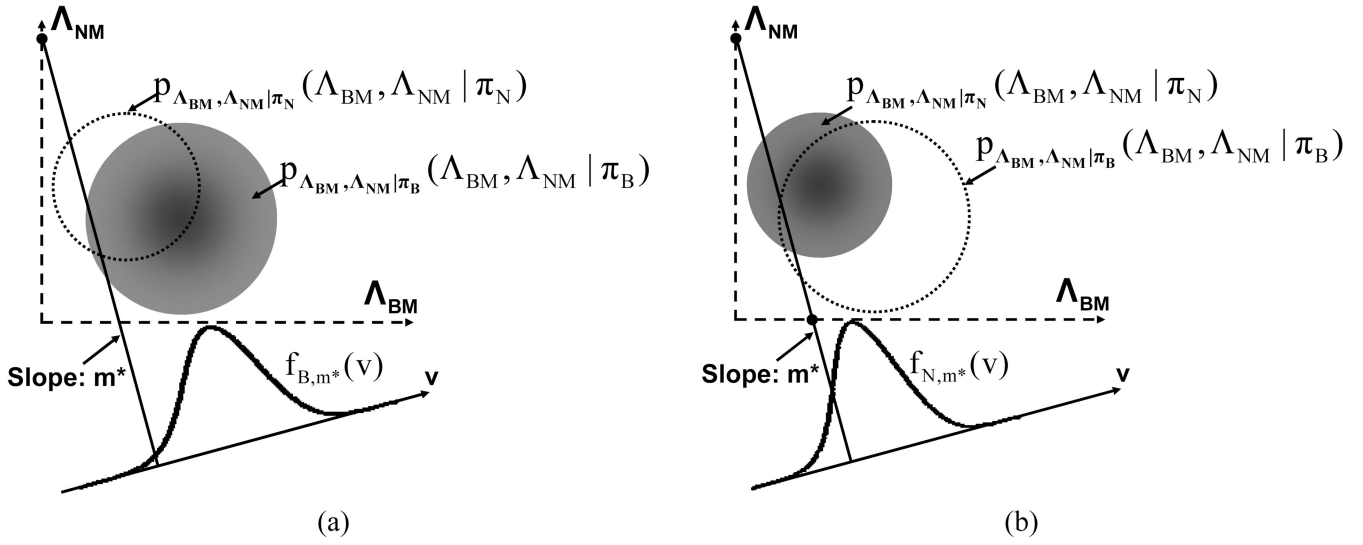


Fig. 2. (a) The integral of the joint pdf of the likelihood ratios under class B, along lines with slope m^* . The resulting integral is a 1-D function, and is denoted as $f_{B,m^*}(v)$. (b) The integral of the joint pdf of the likelihood ratios under class N, along lines with slope m^* . The resulting integral is a 1-D function, and is denoted as $f_{N,m^*}(v)$. Note that $f_{B,m^*}(v)$ and $f_{N,m^*}(v)$ are not Gaussian; in particular, they are zero for $v < 0$ and $u < 0$, respectively.

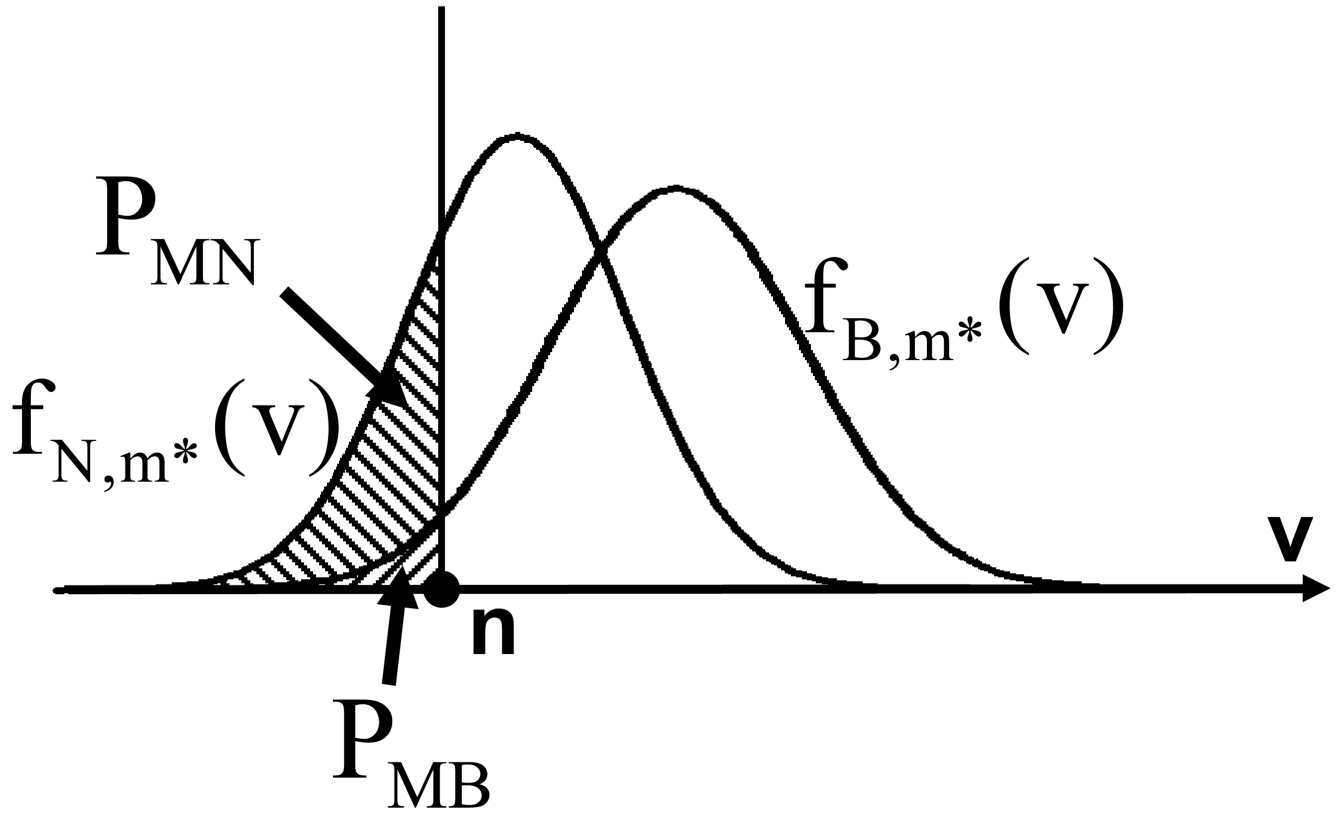


Fig. 3.

The interpretation of P_{MN} and P_{MB} in terms of the areas under $f_{B,m^*}(v)$ and $f_{N,m^*}(v)$ below a threshold n . This interpretation is similar to that for true-negative and false-negative fractions in a 2-class classification problem.

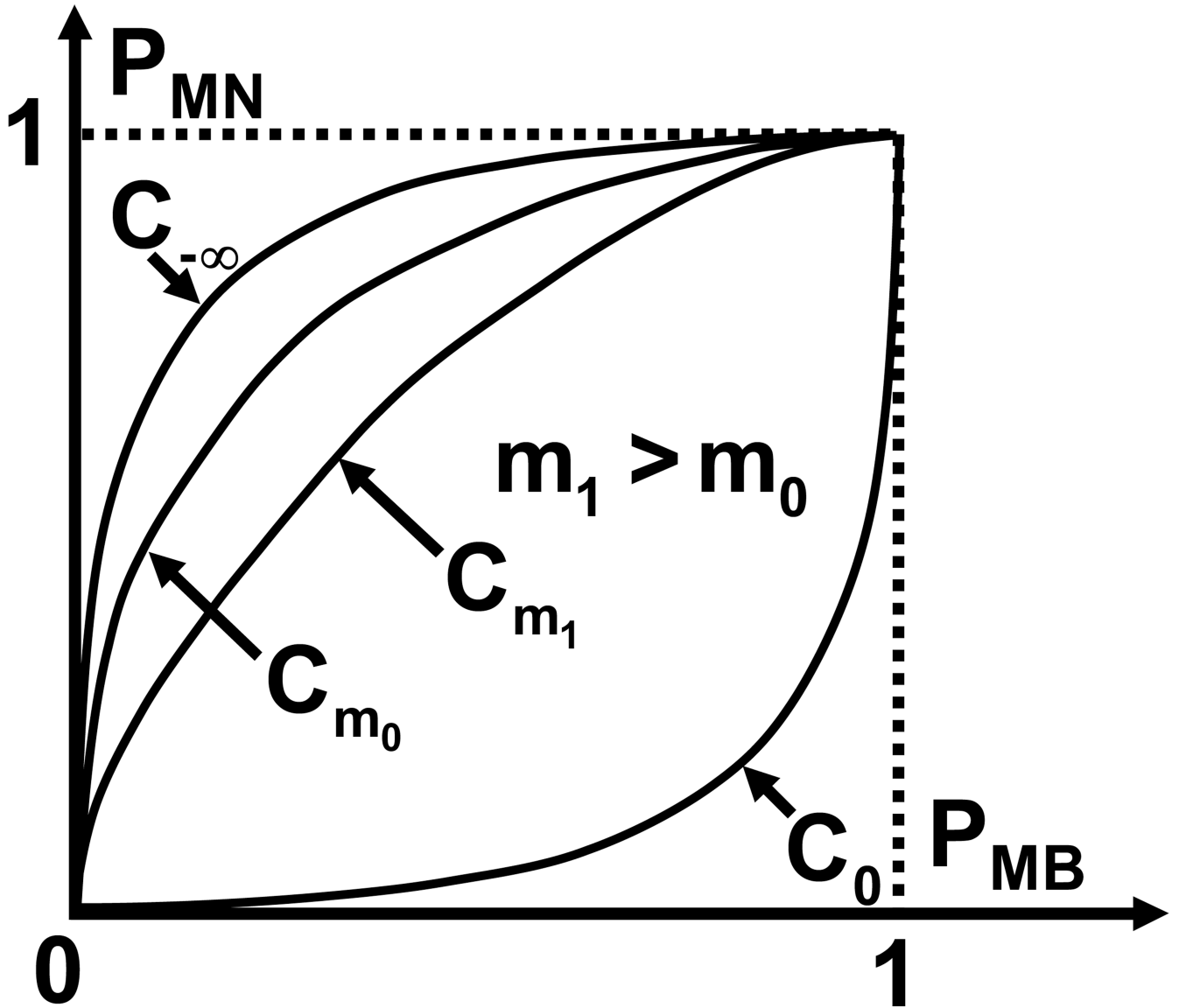


Fig. 4. The curve C_{m_i} plots the relationship between P_{MB} and P_{MN} in the FPF plane for $m = m_i$. Four such curves are plotted: $C_{-\infty}$, C_{m_0} , C_{m_1} , and C_0 . The domain of the ROC surface is defined by the curves $C_{-\infty}$ and C_0 . The figure shows that when $m_1 > m_0$, the curve C_{m_1} is below the curve C_{m_0} .

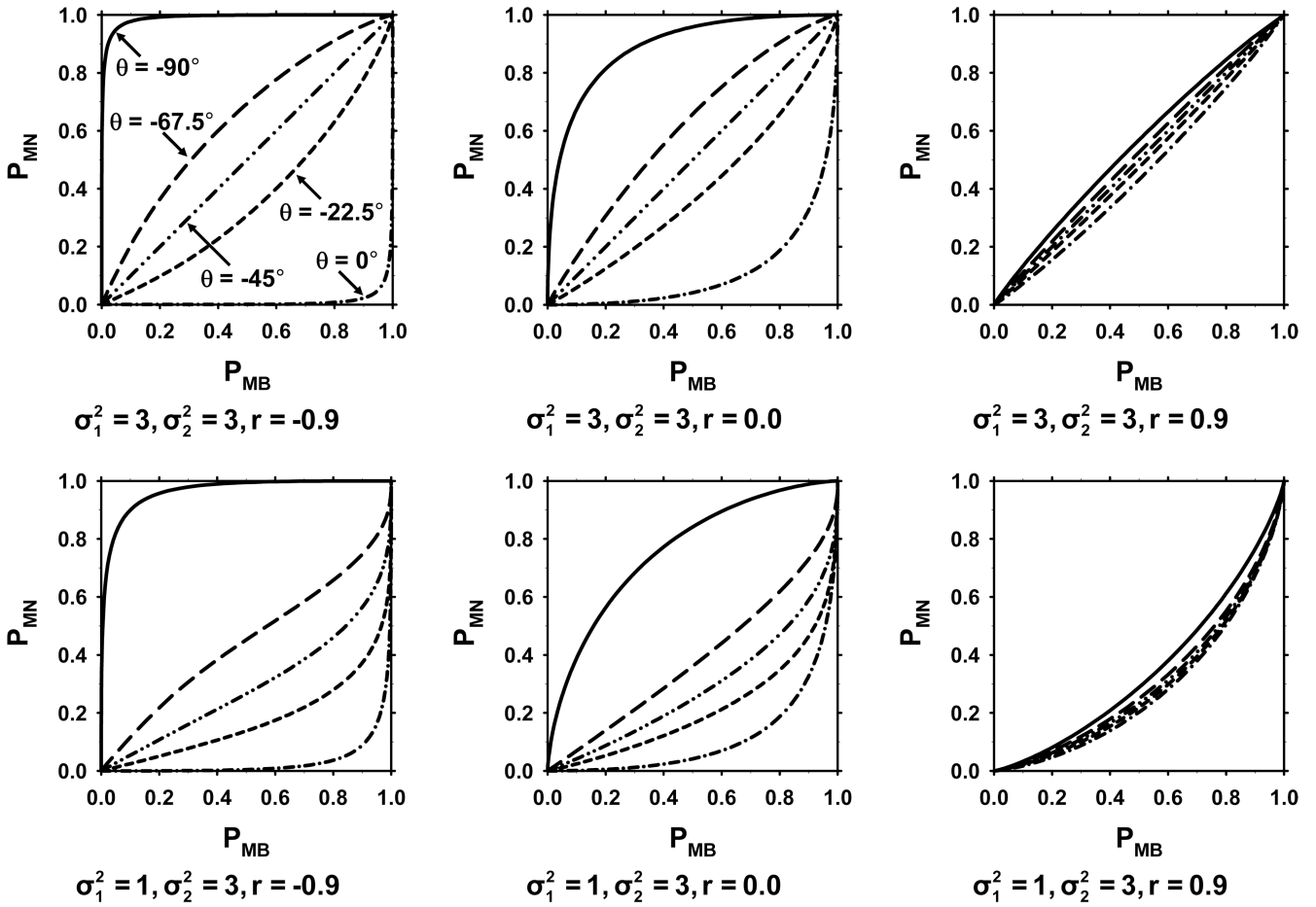


Fig. 5.

The relationship between P_{MN} and P_{MB} for different values of the slope of the decision line (m) in the likelihood ratio plane. Each graph shows five values for $\theta = \tan^{-1}(m)$: $\theta = -90^\circ, -67.5^\circ, -45^\circ, -22.5^\circ$, and 0° . In the top row of graphs, $\sigma_1^2 = \sigma_2^2 = 3$, and in the bottom row of graphs, $\sigma_1^2 = 1$ and $\sigma_2^2 = 3$. The value of r is $-0.9, 0$, and 0.9 for the graphs on the left, middle, and right, respectively.

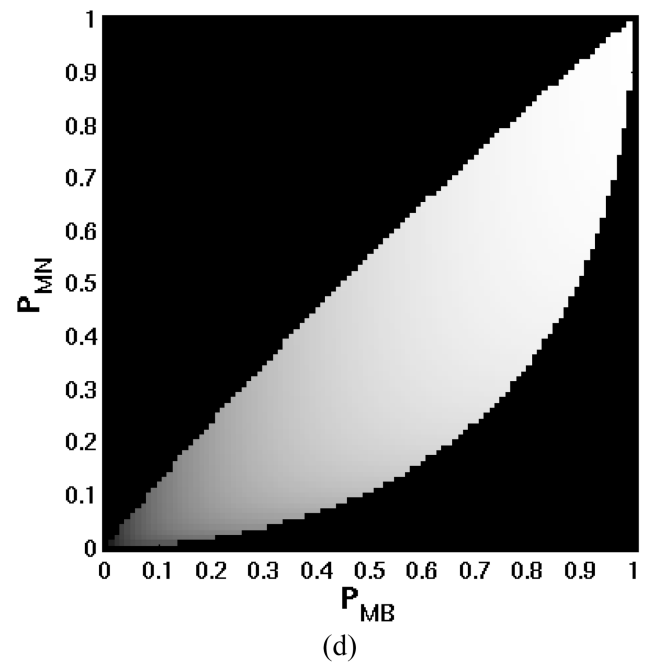
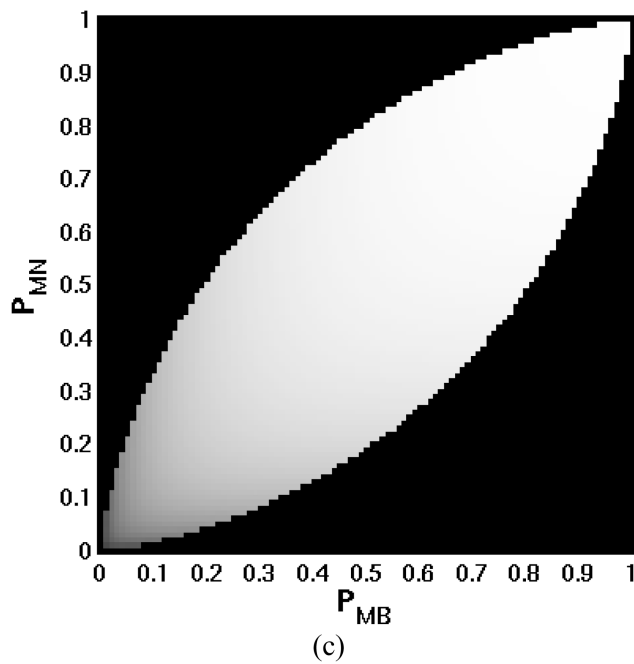
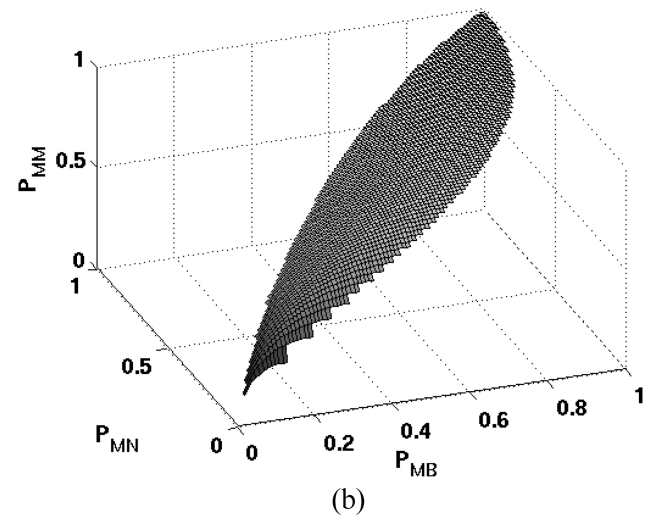
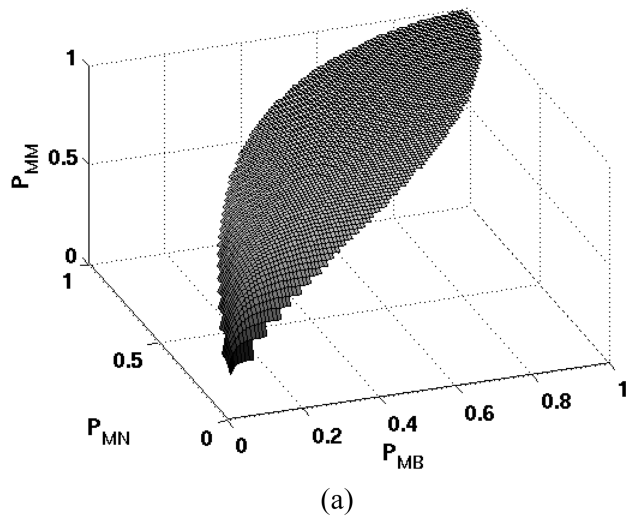


Fig. 6. Examples of 3D ROC surfaces: (a) and (b) are surface representations for Gaussian distributions with $r=0.5$, $\sigma_1^2=\sigma_2^2=3$, and $r=0.5$, $\sigma_1^2=1$, $\sigma_2^2=3$, respectively. (c) and (d) are grayscale representations of (a) and (b), respectively.

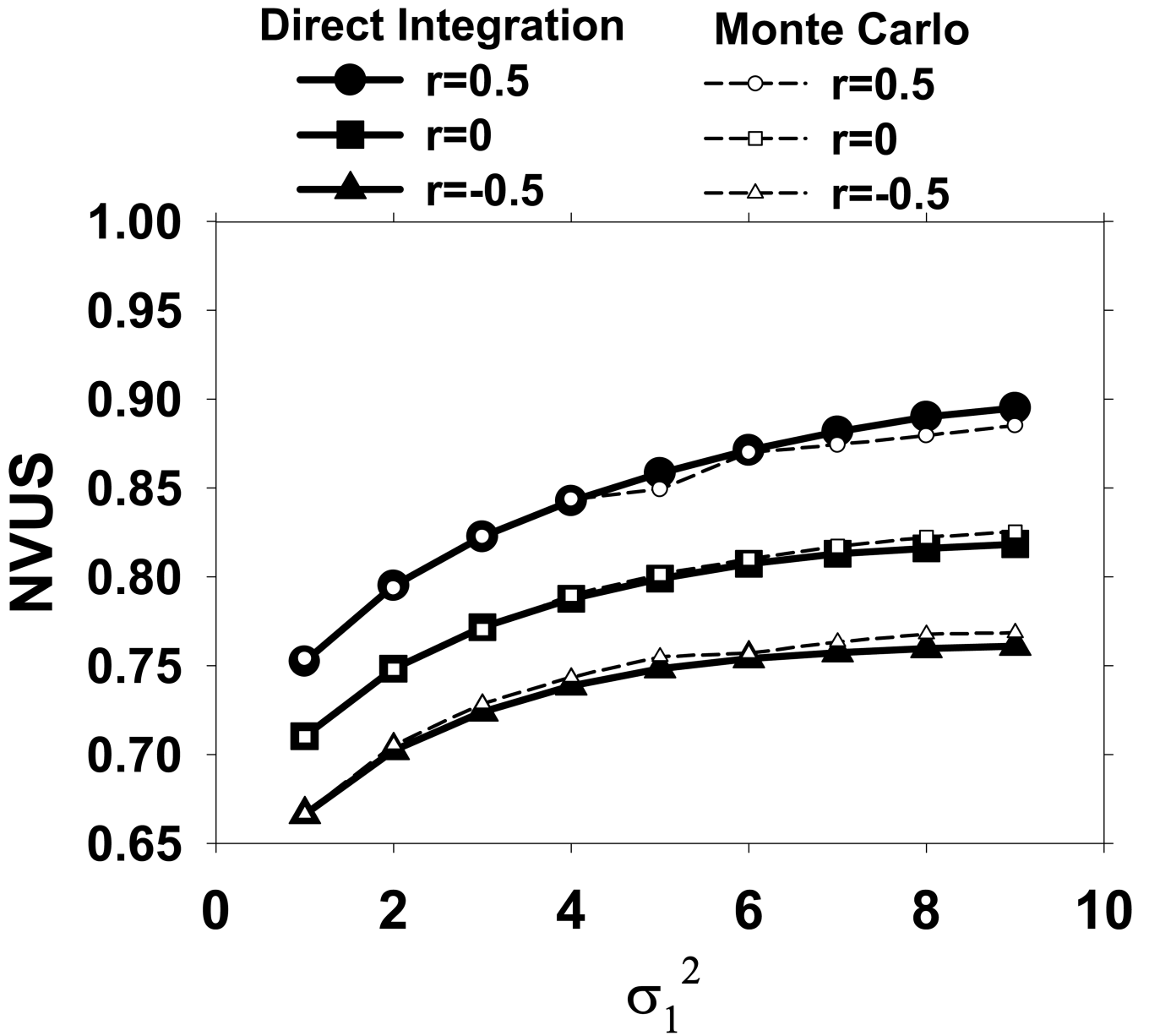


Fig. 7. Comparison of the NVUS values obtained by the direct integration and Monte-Carlo methods for $\sigma_2^2=1$. The NVUS values are plotted as a function of σ_1^2 for three values of r ($r=-0.5, 0, \text{ and } 0.5$).

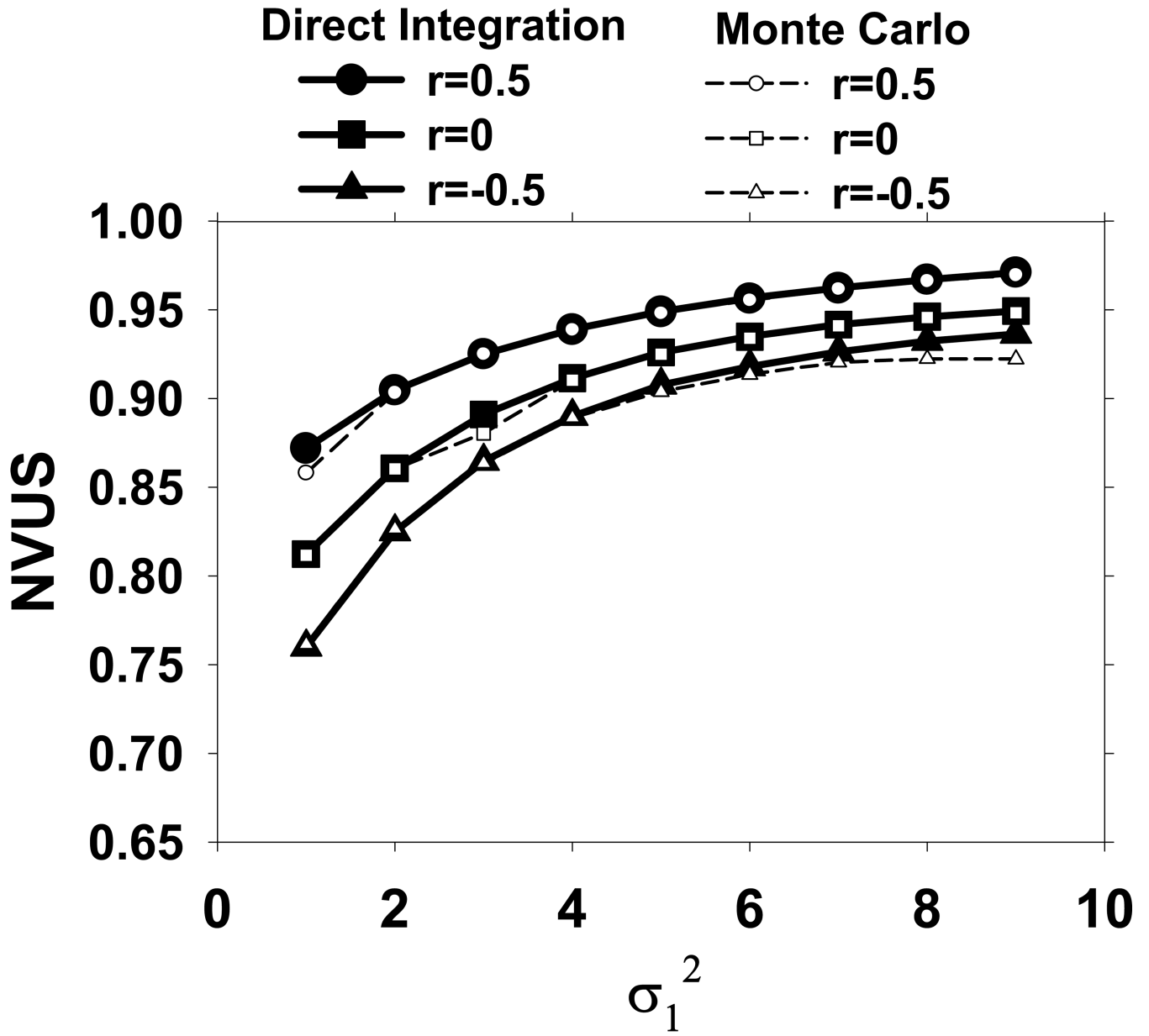


Fig. 8. Comparison of the NVUS values obtained by the direct integration and Monte-Carlo methods for $\sigma_2^2=6$. The NVUS values are plotted as a function of σ_1^2 for three values of r ($r=-0.5, 0, \text{ and } 0.5$).

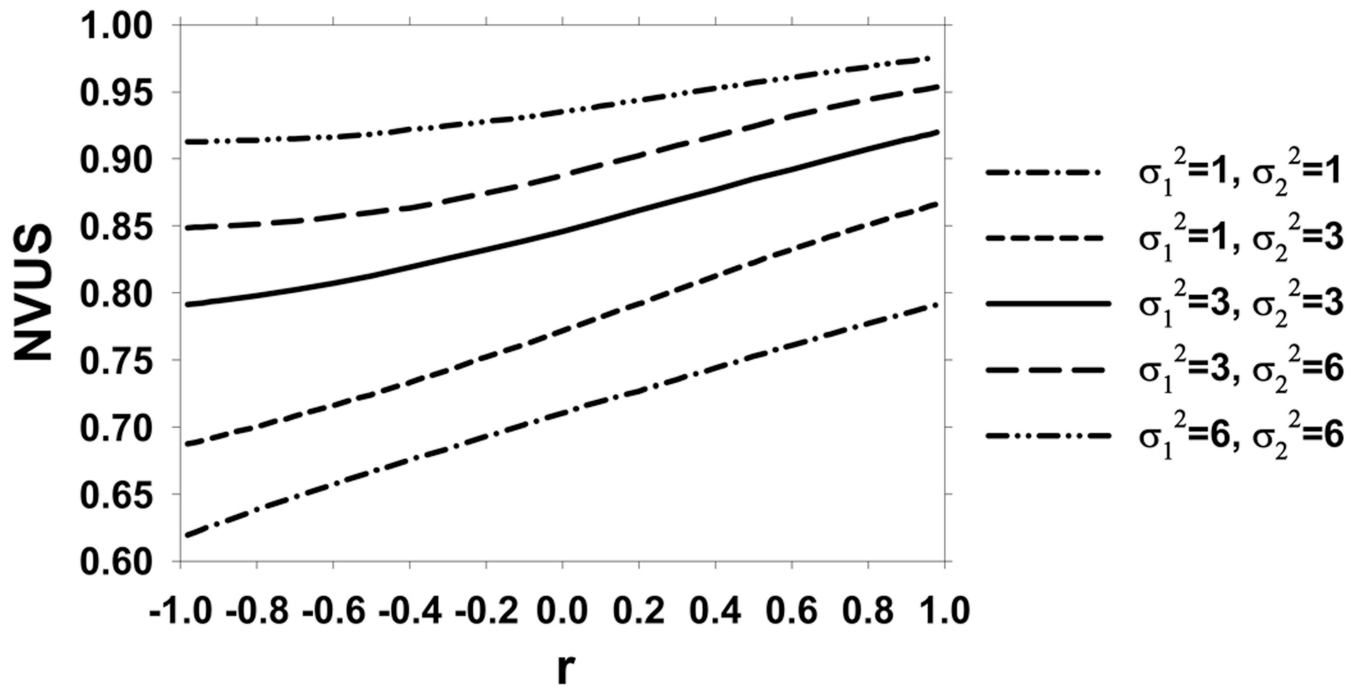


Fig. 9.

NVUS as a function of r for different combinations of values for σ_1^2 and σ_2^2 .

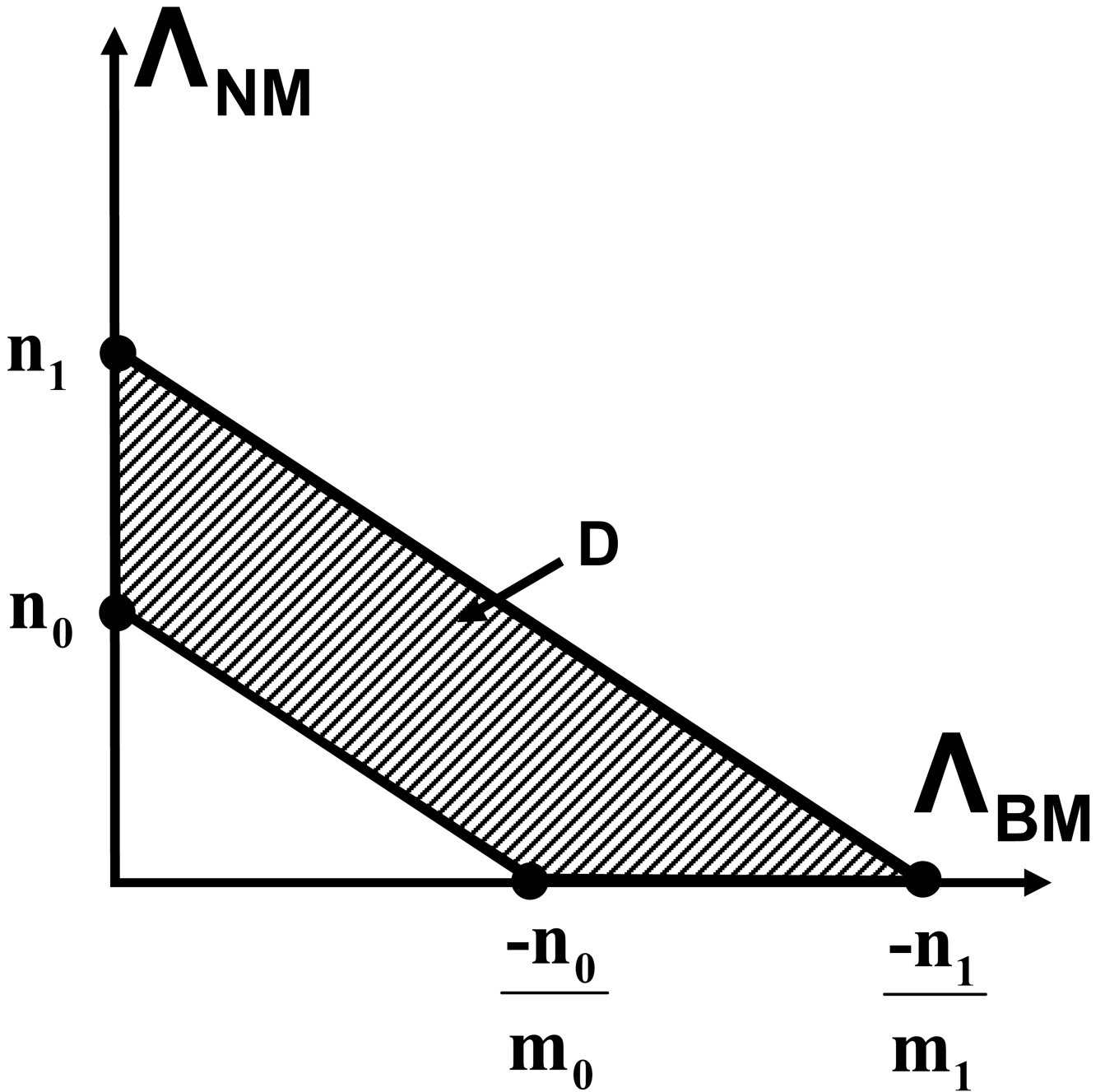


Fig. A1.

Case (A) for Property 1: The lines $\Lambda_{NM} = m_0 \Lambda_{BM} + n_0$ and $\Lambda_{NM} = m_1 \Lambda_{BM} + n_1$ do not intersect in the positive quadrant. The difference between $P_{MB}(m_0, n_0)$ and $P_{MB}(m_1, n_1)$ is the integral of the joint pdf of the likelihood ratios under class B in the shaded region D.

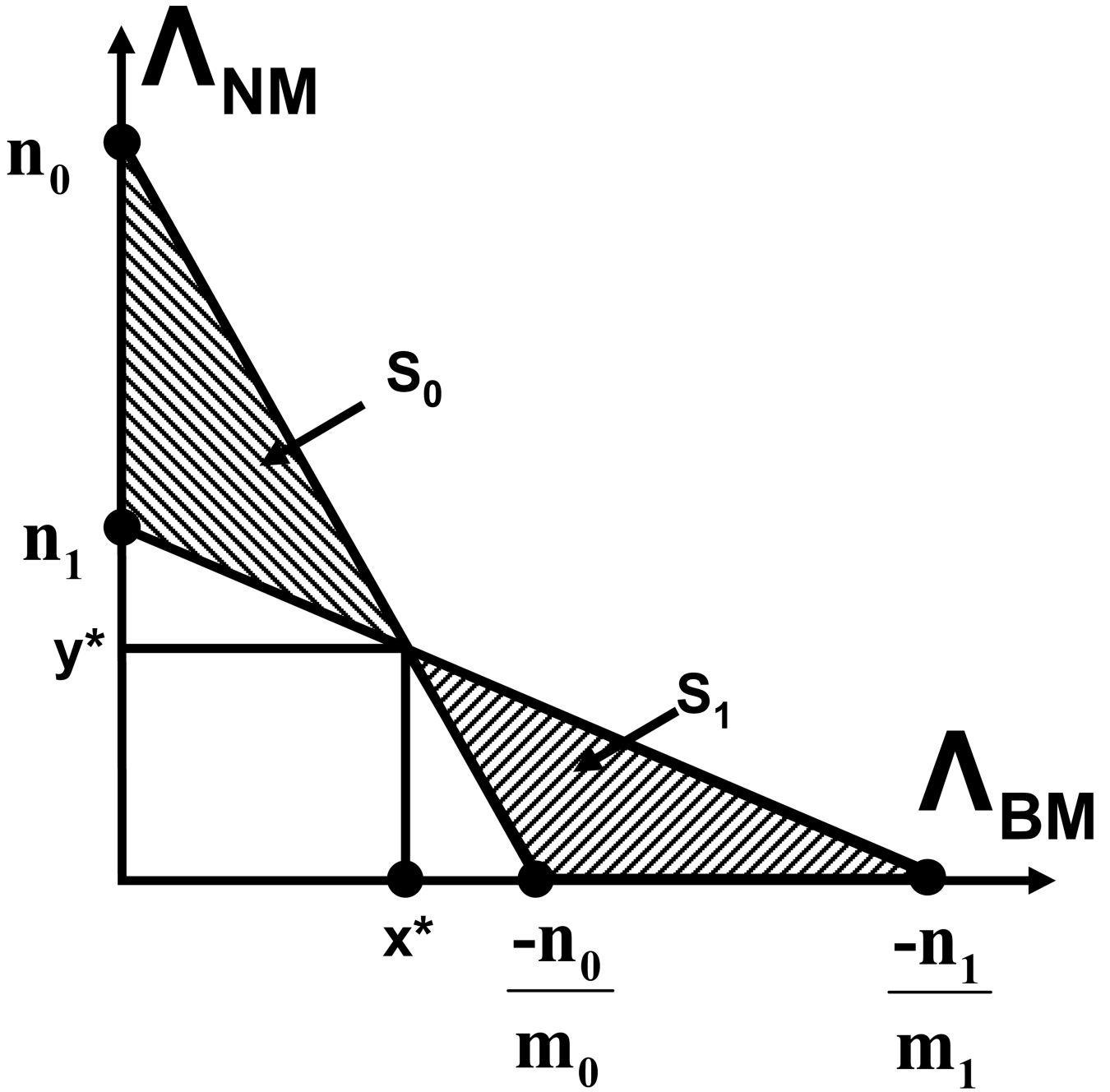


Fig. A2.
 Case (B) for Property 1: The lines $\Lambda_{NM} = m_0 \Lambda_{BM} + n_0$ and $\Lambda_{NM} = m_1 \Lambda_{BM} + n_1$ intersect at a point (x^*, y^*) in the positive quadrant. The difference between $P_{MB}(m_0, n_0)$ and $P_{MB}(m_1, n_1)$ is the difference of the integral of the joint pdf of the likelihood ratios under class B in the shaded regions S_0 and S_1 .

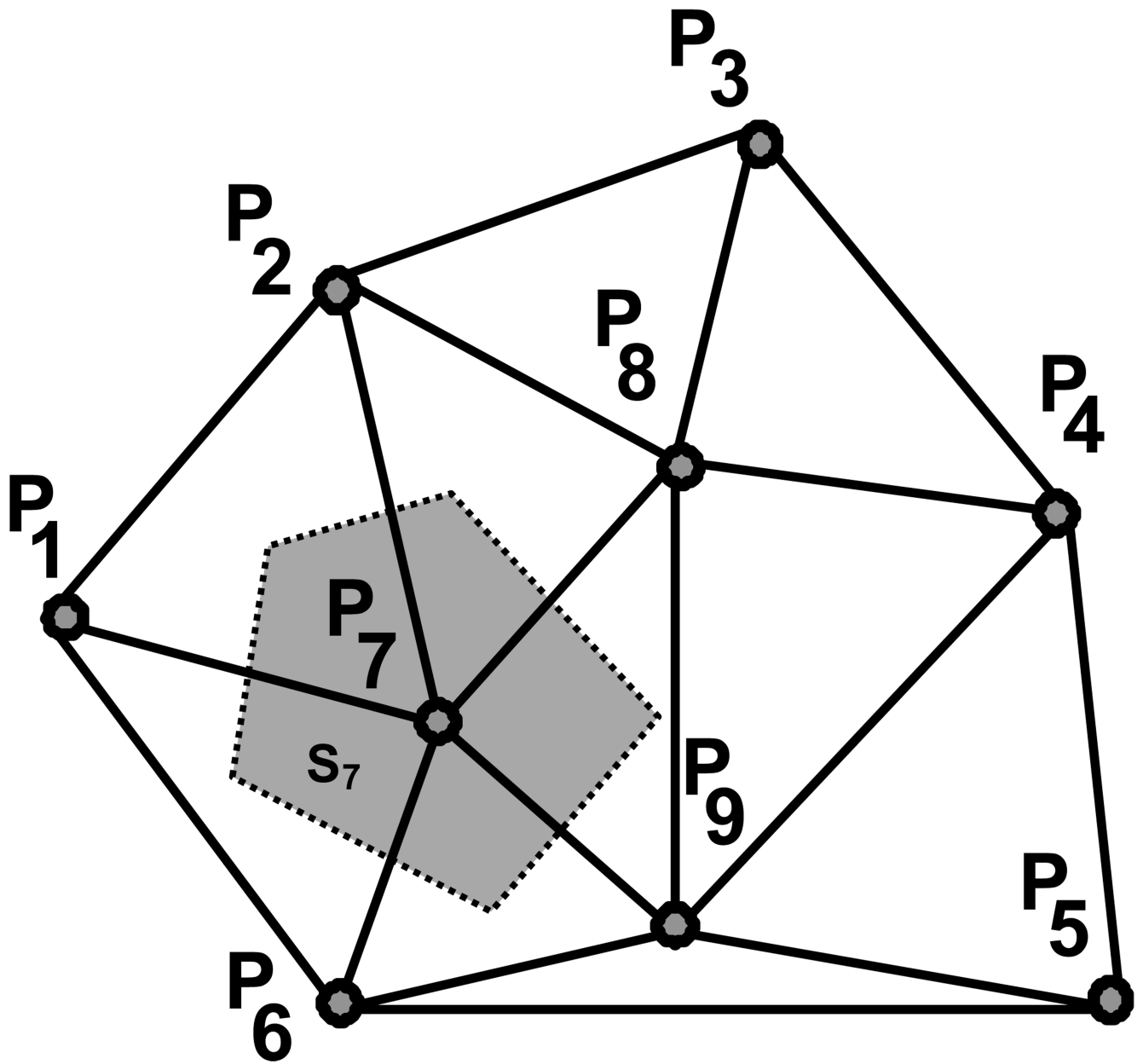


Fig. A3. Delaunay triangulation for $Q=9$ points in the plane. The Voronoi cell S_7 for the point P_7 is shown as the shaded polygon.

Table I

The conditions for the utilities that will degenerate the three decision lines into a single line. For a problem such as classification of a case as malignant, benign and normal, the utilities may be chosen as $v_0=v_2=v_3=1$ and $v_1=0$.

True class membership	M	B	N
Assigned class membership			
M	v_0	U_{MB}	U_{MN}
B	v_1	v_2	v_3
N	v_1	v_2	v_3