Bulletin of the Iranian Mathematical Society Vol. 28 *No. 2 (2002), pp 37-49.*

SCREENING FOR A MULTIVARIATE MIXTURE NORMAL DISTRIBUTION

HOSSEIN MAHJUB AND TREVOR F. COX

Abstract. The screening problem has been studied by many authors mostly focused on individual multivariate normal model with screening for normal distribution when all or part of the parameters are known, or the performance variable is dichotomous. In this paper a screening method is presented when the screening variable is a mixture of two multivariate normal distributions, meanwhile the performance variable is dichotomous. The method is used for the case when the parameters are known or estimated from separate samples. To reduce the dimensionality of the problem and therefore the scale of the computation, a Fisher's linear discrimination function is applied to find coefficients of a standard linear combination of the variables used in the proposed method. A comparison of methods is made for Conn's syndrome date. The results of the study are equivalent to the predictive screening approach.

1. Introduction

Sometimes a variable cannot be measured directly, or measurement of it may be expensive, take a long time, or even be destructive. Instruments of screening are cheap and/or quick to use. However, they are not typically perfect and are usually used as a first stage. In screening the probability of the correct classification of cases within the certain category can be controlled in advance. Variables which are considered in screening are performance and screening variables. The variable which is not easy to measure, is called the performance variable. The

MSC(2000): Primary 62H30; Secondary 62P10,
Keywords: Mixture distribution, Multivariate normal, Screening
Received:15 March 2001 , Revised: 27 May 2002

other variable which is correlated with the performance variable and where measurement is simpler, is called the screening variable. Sometimes, the performance variable is dichotomous. For example, consider a population where some people have a specific disease and others not. Because of high expense, lack of enough specialists, time taken and so on, selection of patients by clinical tests, is not possible. For these reasons, screening methods are recommended. The variable which is considered as the screening variable, should have some specifications, such as high correlation with the disease, and ease of use compared to clinical tests. Screening will not detect all the patients with the disease, but will be a sort indicator of those who do and those who do not. After the selection of people who are suspected of having the disease, clinical tests can then be carried out for confirmation. In this situation, the performance variable, T say, is dichotomous; $T = 0$; $T = 1$, and the screening variable may be continuous or discrete. The aim is to find a limit, L, to achieve $\delta > \gamma$, where,

$$\gamma = P(T = 1) \qquad \text{and} \qquad \delta = P(T = 1|X > L) \ .$$

The proportion of cases which will be selected in the screening test, is

$$\beta = P(X > L) \, .$$

In the rejected group, the proportion of cases that would have been recorded as a success is

$$\varepsilon = P(T = 1|X \leq L).$$

In Medical statistics, δ is known as the positive predictive value, and $1 - \varepsilon$ is the negative predictive value. One field of application in the medical situation is, where the problem of diagnosis of the form of disease from which a patient suffers is often of paramount importance. For example, for the undiagnosed patient it is important to assess which form of disease is appropriate. It is clearly important that to be fairly sure that patients for whom , for example, surgery is decided, surgery operation is essential, say type D. A criteria is to formulate a decision rule which ensure that the probability that a patient for whom is decided on surgery, is in fact of the type D, takes some prespecified value δ. So, one of the most important part of the screening procedure is to achieved a fixed positive predictive value of δ such as 0.90, 0.95, 0.99 and so on.

The screening problem has been investigated for a long time. Owen et al. [10] proposed a jointly normal model where all parameters of

the distribution are known. Owen and Boddie [9] extended the approach with unknown means and unknown variances. Owen and Su [11] considered a solution for screening where the performance and screening variable are jointly normally distributed, when all parameters are unknown. Thomas and Owen [13] used the trivariate normal distribution in order to have two screening variables when all parameters are known. They presented the conditional probability of success on one-sided specification of variables. Also, they proposed a selection criterion based on a minimum cut-off for a linear combination of the two screening variables. Li and Owen [7] extended the two-sided screening procedure under the assumption of bivariate normality. The procedures were based on known and unknown parameters. Haas et al.[4] proposed a screening method with several screening variables, where the performance variable and the screening variables all have a joint multivariate normal distribution, and the specification limits on the performance variables are two-sided model which emphasised the dichotomy of a performance variable, within the diagnostic paradigm, that is based on the predictive probability function of the performance variable given the screening variable, and sampling paradigm, which is based on the conditional distribution function of the screening variable on the performance variable. Dunsmore and Boys [2] considered the screening procedure when the screening variable has a multivariate normal distribution and the performance variable is dichotomous, by using a predictive distribution. Tang and Tang [12] reviewed the literature in the area of screening. De'Moraes and Dunsmore [1] developed the "local approach" to the screening problem derived by Dunsmore and Boys [3] to the polychotomous multivariate case, and derived the approach, where each component of the explanatory variables is discrete or continuous. Lee and Jang [6] used optimum target values for a production process with three-class screening with normally distributed with an unknown mean and a known variance. Optimum mean value and screening limits for production processes with multiclass screening is suggested by Hong et al.[5].

In the present paper a different case where the screening variable is a multivariate mixture normal distribution and the performance variable is dichotomous is considered. The model that is used, however, is an extension of a normal model with single screening variable to multivariate case.

2. Screening for the mixture normal distribution when the parameters are known

In this section a screening method where the screening variable is mixture of two multivariate normal distributions, and the performance variable is dichotomous is presented. Consider

$$f(\mathbf{x};\boldsymbol{\mu}_1,\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_0,p) = pg_1(\mathbf{x};\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1) + qg_0(\mathbf{x};\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0),$$

where $g_i(\mathbf{x};\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i)$, $(i = 0,1)$ are two probability density functions of multivariate normal random variables with parameters $(\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i)$, $(i = 0,1)$ and $q = 1-p$, $0 \le p \le 1$.
The function $f(\mathbf{x})$ is the density function of a mixture of two multivariate normal distributions with mixing proportions p and q.
Suppose that the screening variable, $\mathbf{X}$, has a multivariate mixture distribution and the performance variable is dichotomous, $(T = 0$ or $T = 1)$ and the screening variable in each category is multivariate normal with parameters $(\boldsymbol{\mu}_j,\boldsymbol{\Sigma}_j)$, $(j = 0,1)$. Suppose that $P(T = 1) = \gamma$ and $P(T = 0) = 1-\gamma$. The aim is to determine a specification region, $C_{\mathbf{x}}$, such that

$$\delta = P(T = 1|\mathbf{x} \in C_{\mathbf{x}}) \ ,$$

with $\delta > \gamma$. The form of $C_{\mathbf{x}}$ is restricted to a linear combination of the variables, and so

$$C_{\mathbf{x}} = \left\{\mathbf{x} : \mathbf{a}'\mathbf{x} \ge \omega\right\} \ ,$$

where $\mathbf{a}$ is a constant vector and ω is a constant. A standardised linear combination of variables is applied, so that, $\mathbf{a}'\mathbf{a} = 1$. There is not a unique solution for the problem. So, another restriction is made by choosing the values of $\mathbf{a}$ and ω which satisfy δ and which minimise the error probability

$$\varepsilon = P(T = 1|\mathbf{x} \notin C_{\mathbf{x}}) \ .$$

Now

$$\delta = \frac{P(\mathbf{a}'\mathbf{X} \ge \omega|T = 1)P(T = 1)}{P(\mathbf{a}'\mathbf{X} \ge \omega|T = 1)P(T = 1) + P(\mathbf{a}'\mathbf{X} \ge \omega|T = 0)P(T = 0)} \ .$$

When the parameters are known, if $(\mathbf{X}|T = j)$, $(j = 0,1)$ is a multivariate normal distribution $N_p(\boldsymbol{\mu}_j,\boldsymbol{\Sigma}_j)$, $(j = 0,1)$, then $(\mathbf{a}'\mathbf{X}|T = j)$, $(j = 0,1)$ has a univariate normal distribution of $N_1(\mathbf{a}'\boldsymbol{\mu}_j, \mathbf{a}'\boldsymbol{\Sigma}_j\mathbf{a})$, $(j = 0,1)$. So,

$$\delta = \frac{\gamma\left[1 - P\left(\frac{\mathbf{a}'\mathbf{X} - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}} < \frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)\right]}{\gamma\left[1-P\left(\frac{\mathbf{a}'\mathbf{X} - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}} < \frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)\right]+(1-\gamma)\left[1-P\left(\frac{\mathbf{a}'\mathbf{X} - \mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_0\mathbf{a}}} < \frac{\omega - \mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_0\mathbf{a}}}\right)\right]}.$$

Therefore,

$$\delta = \frac{\gamma\left[1 - \Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)\right]}{\gamma\left[1 - \Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)\right] + (1-\gamma)\left[1 - \Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_0\mathbf{a}}}\right)\right]},$$

where $\Phi(z)$ is the standard normal distribution function. Since, the above expression does not have a unique solution for $C_{\mathbf{x}} = \left\{\mathbf{x} : \mathbf{a}'\mathbf{x} \geq \omega\right\}$, the aim is to find a region such that ε is minimised, where,

$$\varepsilon = \frac{P(\mathbf{a}'\mathbf{X} < \omega | T = 1)P(T = 1)}{P(\mathbf{a}'\mathbf{X} < \omega | T = 1)P(T = 1) + P(\mathbf{a}'\mathbf{X} < \omega | T = 0)P(T = 0)}.$$

Therefore,

$$\varepsilon = \frac{\gamma\Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)}{\gamma\Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right) + (1-\gamma)\Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_0\mathbf{a}}}\right)}.$$

Some different methods were applied to solve the problem analytically. Since it was not possible, a FORTRAN-77 program was written to obtain $\mathbf{a}$ and ω numerically, so that δ is achieved and ε minimised. The required time for computation is high, especially when the number of screening variables is large. So to reduce the dimensionality of the problem and therefore the scale of the computation, Fisher's linear discrimination function can be applied to obtain a particular $\mathbf{a}$. Fisher's linear discrimination function is

$$\mathbf{L}'\mathbf{x} \qquad \text{where} \qquad \mathbf{L} = \boldsymbol{\Sigma}^{-1}(\mu_{\mathbf{1}} - \mu_{\mathbf{0}}),$$

where $\boldsymbol{\Sigma}$ is variance-covariance matrix pooling $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$. So, the specification region becomes $C_{\mathbf{x}} = \left\{\mathbf{x} : (\mu_{\mathbf{1}} - \mu_{\mathbf{0}})'\Sigma^{-1}\mathbf{x} > \omega\right\}$. The only unknown value to be found is ω which can be obtained from

$$\delta = \frac{\gamma\left[1 - \Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)\right]}{\gamma\left[1 - \Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)\right] + (1-\gamma)\left[1 - \Phi\left(\frac{\omega - \mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_0\mathbf{a}}}\right)\right]}.$$

Note that the linear discrimination function is obtained by using a pooled variance- covariance matrix for the two populations, but in the expression for δ, respective variance-covariance matrices for the two populations are used. Of course, when the two variance-covariance matrices for the two populations are equal ($\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}$), δ and ε are given by

$$\delta = \frac{\gamma\left[1-\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}}}\right)\right]}{\gamma\left[1-\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}}}\right)\right]+(1-\gamma)\left[1-\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}}}\right)\right]}$$

and

$$\varepsilon = \frac{\gamma\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}}}\right)}{\gamma\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}}}\right)+(1-\gamma)\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}}}\right)} .$$

In general, the proportion of cases retained in the screening process is

$$\beta = \gamma\left[1-\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_1}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_1\mathbf{a}}}\right)\right]+(1-\gamma)\left[1-\Phi\left(\frac{\omega-\mathbf{a}'\boldsymbol{\mu}_0}{\sqrt{\mathbf{a}'\boldsymbol{\Sigma}_0\mathbf{a}}}\right)\right] .$$

An extensive numerical study was shown, the advantage of using Fisher's linear discriminant for finding an $\mathbf{a}$ for the proposed method is that computation time is much reduced, but the disadvantage is that ε will not be necessarily a minimum.
By using principal component analysis, it might be possible to reduce the dimensionality of the screening variables. Principal components are formed using $\boldsymbol{\Sigma}$ and then the important ones used in screening. Transforming the data to principal components will still leave the appropriate distribution a mixture multivariate normal.

3. Screening for the mixture normal distribution when the parameters are unknown

If the parameters of the mixture distribution are unknown, assume two separate samples from the multivariate normal distributions with sizes N_j, $(j = 0, 1)$ are available. The estimation of the parameters of the normal distributions are

$$\hat{\boldsymbol{\mu}}_1 = \overline{\mathbf{x}}_1 \quad , \quad \hat{\boldsymbol{\Sigma}}_1 = \mathbf{s}_1 \quad , \quad \hat{\boldsymbol{\mu}}_0 = \overline{\mathbf{x}}_0 \quad , \quad \hat{\boldsymbol{\Sigma}}_0 = \mathbf{s}_0 .$$

Also, the mixture proportion, γ, is estimated by a sample of size N, from the mixture population of $\mathbf{X}$. So, $\hat{\gamma} = \frac{R}{N}$, where R, is the number of cases with $(T = 1)$. Now

$$(\mathbf{X} - \overline{\mathbf{X}}_j)(\mathbf{X} - \overline{\mathbf{X}}_j)' = (N_j - 1)\mathbf{S}_j \sim W_p(\mathbf{\Sigma}_j, N_j - 1)$$

where $W_p(\mathbf{\Sigma}_j, N_j - 1)$ is a Wishart distribution with parameters $(\mathbf{\Sigma}_j, N_j - 1)$, $(j = 0, 1)$, see (Mardia et al., Chap 4) [7]. Also,

$$(N_j - 1)\frac{\mathbf{a}'\mathbf{S_j}\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_j\mathbf{a}} \sim \chi^2_{(N_j-1)}\,.$$

With some simplification

$$\sqrt{\frac{N_j}{N_j + 1}}\frac{\mathbf{a}'(\mathbf{X}_j - \overline{\mathbf{X}}_j)}{\sqrt{\mathbf{a}'\mathbf{S}_j\mathbf{a}}} \sim t_{(N_j-1)}\ ,$$

where, $t_{(N_j-1)}$ is a univariate Student's t distribution with $N_j - 1$, $(j = 0, 1)$ degrees of freedom.
In the screening procedure when the parameters are unknown,

$$\delta = \frac{\hat{\gamma}\left[1 - P\left(\frac{\mathbf{a}'\mathbf{X} - \mathbf{a}'\overline{\mathbf{X}}_1}{\sqrt{\mathbf{a}'\mathbf{S}_1\mathbf{a}}} < \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)\right]}{\hat{\gamma}\left[1-P\left(\frac{\mathbf{a}'\mathbf{X} - \mathbf{a}'\overline{\mathbf{X}}_1}{\sqrt{\mathbf{a}'\mathbf{S}_1\mathbf{a}}} < \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)\right]+(1-\hat{\gamma})\left[1-P\left(\frac{\mathbf{a}'\mathbf{X} - \mathbf{a}'\overline{\mathbf{X}}_0}{\sqrt{\mathbf{a}'\mathbf{S}_0\mathbf{a}}} < \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_0}{\sqrt{\mathbf{a}'\mathbf{s}_0\mathbf{a}}}\right)\right]},$$

and hence

$$\delta = \frac{\hat{\gamma}\left[1 - F_1\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)\right]}{\hat{\gamma}\left[1 - F_1\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)\right] + (1-\hat{\gamma})\left[1 - F_0\left(\sqrt{\frac{N_0+1}{N_0}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_0}{\sqrt{\mathbf{a}'\mathbf{s}_0\mathbf{a}}}\right)\right]},$$

where $F_j = F_{t,N_j-1}(t)$, $(j = 0, 1)$ is the probability distribution function of the univariate Student's t distribution with $N_j - 1$, $(j = 0, 1)$ degrees of freedom. Since, the above expression does not have a unique solution for $C_{\mathbf{x}} = \left\{\mathbf{x} : \mathbf{a}'\mathbf{x} \geq \omega\right\}$, the aim is to find a region $C_{\mathbf{x}}$ such that $\hat{\varepsilon}$ is minimised, where,

$$\hat{\varepsilon} = \frac{P(\mathbf{a}'\mathbf{X} < \omega | T = 1)P(T = 1)}{P(\mathbf{a}'\mathbf{X} < \omega | T = 1)P(T = 1) + P(\mathbf{a}'\mathbf{X} < \omega | T = 0)P(T = 0)}\ ,$$

i.e.,

$$\hat{\varepsilon} = \frac{\hat{\gamma} F_1\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)}{\hat{\gamma} F_1\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right) + (1-\hat{\gamma}) F_0\left(\sqrt{\frac{N_0+1}{N_0}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_0}{\sqrt{\mathbf{a}'\mathbf{s}_0\mathbf{a}}}\right)}.$$

Then $\mathbf{a}$ and ω are found numerically to obtain the specified δ and where $\hat{\varepsilon}$ is minimised.
Again, the required computation time is high, especially when the number of screening variables is large. So, to reduce the dimensionality of the problem and therefore the scale of the computation, Fisher's linear discrimination function can be applied in the proposed method to obtain $\mathbf{a}$. Fisher's linear discrimination function when the parameters are unknown is

$$\mathbf{L}'\mathbf{x} \qquad \text{where} \qquad \mathbf{L} = \mathbf{S}_{pooled}^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_0)\ .$$

So, the specification region becomes $C_{\mathbf{x}} = \left\{\mathbf{x} : (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_0)\mathbf{S}_{pooled}^{-1}\mathbf{x} > \omega\right\}$, where $\mathbf{S}_{pooled}$ is the estimation of the pooled variance-covariance matrix of the two populations. In fact

$$\mathbf{S}_{pooled} = \frac{(N_1-1)\mathbf{S}_1 + (N_0-1)\mathbf{S}_0}{N_1+N_0-2}\ .$$

Then the only unknown value remaining to be found is ω, which can be obtained from

$$\delta = \frac{\hat{\gamma}\left[1 - F_1\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)\right]}{\hat{\gamma}\left[1 - F_1\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)\right] + (1-\hat{\gamma})\left[1 - F_0\left(\sqrt{\frac{N_0+1}{N_0}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_0}{\sqrt{\mathbf{a}'\mathbf{s}_0\mathbf{a}}}\right)\right]}.$$

Note again that the linear discrimination function is obtained by using a pooled variance-covariance matrix, but in the expression for δ, unequal variance matrices for the two populations are used. When the two variance-covariance matrices for the two populations are equal, δ and $\hat{\varepsilon}$ are given by

$$\delta = \frac{\hat{\gamma}\left[1 - F\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_{pooled}\mathbf{a}}}\right)\right]}{\hat{\gamma}\left[1 - F\left(\sqrt{\frac{N_1+1}{N_1}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_{pooled}\mathbf{a}}}\right)\right] + (1-\hat{\gamma})\left[1 - F\left(\sqrt{\frac{N_0+1}{N_0}}\frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_0}{\sqrt{\mathbf{a}'\mathbf{s}_{pooled}\mathbf{a}}}\right)\right]}$$

and

$$\hat{\varepsilon} = \frac{\hat{\gamma} F\left(\sqrt{\frac{N_1+1}{N_1}} \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_{pooled}\mathbf{a}}}\right)}{\hat{\gamma} F\left(\sqrt{\frac{N_1+1}{N_1}} \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_{pooled}\mathbf{a}}}\right) + (1-\hat{\gamma}) F\left(\sqrt{\frac{N_0+1}{N_0}} \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_0}{\sqrt{\mathbf{a}'\mathbf{s}_{pooled}\mathbf{a}}}\right)}.$$

where, $F = F_{t,N_1+N_2-2}(t)$ is the probability distribution function of the univariate Student's t distribution with $(N_1 + N_2 - 2)$ degree of freedom.
In general, the estimation of the probability of retaining a case in the screening process is

$$\hat{\beta} = \hat{\gamma}\left[1 - F_1\left(\sqrt{\frac{N_1+1}{N_1}} \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_1}{\sqrt{\mathbf{a}'\mathbf{s}_1\mathbf{a}}}\right)\right] + (1-\hat{\gamma})\left[1 - F_0\left(\sqrt{\frac{N_0+1}{N_0}} \frac{\omega - \mathbf{a}'\overline{\mathbf{x}}_0}{\sqrt{\mathbf{a}'\mathbf{s}_0\mathbf{a}}}\right)\right].$$

Example 3.1. The example used by Dunsmore and Boys [2] for predictive screening is now considered and comparison is made. The data relate to Conn's syndrome which is a rare form of hypertension. Two form of syndrome exist, namely :
A : benign tumour in the adrenal cortex, (adenoma),
B : more diffuse condition of the adrenal glands, (bilateral hyperplasia).
The treatment for A is surgical operation to remove the adrenal glad. For B drug therapy is the recognised treatment, and surgery is inadvisable.
Let $T = 1$ for A, and $T = 0$ for B. The aim is to find a region using the feature vector $\mathbf{x}$ to attempt to screen out the B cases and to retain the A cases with a fixed positive predictive value of $\delta = 0.95$, where $\mathbf{x}$ is the three variables of concentrations (meq/l) in blood plasma: sodium (Na), potassium (K) and carbon dioxide (CO_2).
The data are given in Table 1, and log (concentrations) for the basic variables $\mathbf{x}$ is used to remove much of the skewness apparent in the data.

Table 1
Conn's Syndrome Data log**(concentration, meq/l) in blood plasma**

Type	Patient	Na	K	CO_2
		x_1	x_2	x_3
A	1	4.9459	0.8329	3.4112
	2	4.9628	1.1314	3.2995
	3	4.9416	1.0986	3.2958
	4	4.9836	1.0296	3.4965
	5	4.9323	1.2809	3.1822
	6	4.9677	1.1314	3.3322
	7	4.9222	0.9163	3.3878
	8	4.9488	0.9163	3.4012
	9	4.9684	0.8755	3.4720
	10	4.9740	1.0647	3.3844
	11	4.9381	0.8329	3.2581
	12	4.9698	0.7885	3.5175
	13	4.9767	0.9933	3.4965
	14	4.9431	1.1314	3.37.7
	15	4.9747	1.0647	3.3105
	16	4.9345	1.1314	3.4468
	17	4.9754	0.6419	3.5116
	18	4.9816	1.3083	3.3105
	19	4.9698	0.7885	3.4965
	20	4.9663	0.9933	3.3142
B	21	4.9438	1.4586	3.1527
	22	4.9488	1.1632	3.2189
	23	4.9502	1.2809	3.2504
	24	4.9558	1.0986	3.0910
	25	4.9663	1.4351	3.3250
	26	4.9395	1.2238	3.3322
	27	4.9495	1.2809	3.2189
	28	4.9488	1.3350	3.2581
	29	4.9452	1.1939	3.2958
	30	4.9416	1.2809	3.2581
	31	4.9416	1.4816	3.2426

The abstracted information from the data is

$$\overline{\mathbf{x}}_1 = \begin{bmatrix} 4.96 \\ 1.00 \\ 3.38 \end{bmatrix} \quad , \quad \overline{\mathbf{x}}_0 = \begin{bmatrix} 4.95 \\ 1.29 \\ 3.24 \end{bmatrix} .$$

with the estimated variance-covariance matrices

$$\mathbf{s_1} = \begin{bmatrix} 0.035 & -0.028 & 0.073 \\ -0.028 & 2.974 & -1.026 \\ 0.073 & -1.026 & 0.920 \end{bmatrix} \times 10^{-2},$$

$$\mathbf{s_0} = \begin{bmatrix} +0.006 & 0.000 & 0.003 \\ 0.000 & +1.546 & 0.186 \\ 0.003 & 0.186 & +0.503 \end{bmatrix} \times 10^{-2},$$

$$\mathbf{s}_{pooled} = \begin{bmatrix} 0.025 & -0.018 & 0.046 \\ -0.018 & 2.481 & -0.608 \\ 0.046 & -0.608 & 0.776 \end{bmatrix} \times 10^{-2}.$$

As was mentioned it is possible to solve the screening problem by different methods in the proposed method. Table 2 shows the results of screening procedure by different methods, and also compares the obtained results with the results of predictive screening used by Dunsmore and Boys [2].

Table 2
Comparison between two different methods for optimal specification region of Conn's Syndrome Data

Method	Different approaches	a_1	a_2	a_3	ω	$\hat{\varepsilon}$	$\hat{\beta}$	$\hat{\gamma}$
Proposed method	multivariate	0.91	−0.24	0.34	5.34	0.25	0.56	0.65
	multivariate $\Sigma_1 = \Sigma_0$	0.77	−0.44	0.46	4.90	0.39	0.46	0.65
	linear discriminant	0.76	−0.43	0.48	4.87	0.25	0.56	0.65
	linear discriminant $\Sigma_1 = \Sigma_0$	0.76	−0.43	0.48	4.94	0.39	0.46	0.65
Predictive Screening	multivariate	0.91	−0.24	0.34	5.34	0.25	0.56	0.65
	multivariate $\Sigma_1 = \Sigma_0$	0.77	−0.43	0.47	4.95	0.39	0.46	0.65
	linear discriminant	0.76	−0.43	0.48	4.87	0.26	0.56	0.65
	linear discriminant $\Sigma_1 = \Sigma_0$	0.76	−0.43	0.48	4.94	0.39	0.46	0.65

The results show that overall there is no difference between the proposed method and the predictive screening approach of Dunsmore and Boys [2]. However, using predictive probability does make the problem more complicated. Also, it can be seen that by using Fisher's linear discrimination analysis in the proposed method to determine **a** gives a value of $\hat{\varepsilon}$ very close to that obtained by the direct method in the proposed model. Although a value of $\delta = 0.95$ is achieved, $\hat{\varepsilon}$ is rather large at 0.25. Assuming equality of variance in the two populations gives a value of $\hat{\varepsilon}$ much greater than that obtained without the assumption.

References

[1] A.R. De'Moraes and I.R. Dunsmore, Predictive comparisons in ordinal models. *Commun. Statist. - Theory Meth.*, **24** (1995) 2145-2164.

[2] I.R. Dunsmore and R.J. Boys, Predictive screening method in binary response models. In *Probability and Bayesian Statistics* (ed. R. Viert),(1987) pp. 151-8. New York: Plenum.

[3] I.R. Dunsmore and R.J. Boys, Global *v* local screening. In *Baysian Statistics 3* (eds. J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith), pp. 593-9. Oxford: Oxford University Press(1988).

[4] R.W. Haas, R.F. Gunst and D.B. Owen, Screening procedures using quadratic forms. *Commun. Statist. - Theory Meth.*, **14** (1985) 1393-1404.

[5] S.H. Hong, E.A. Elsayed and M.K. Lee, Optimum mean value and screening limits for production processes with multi-class screening. *International Journal of Production Research*, **37** (1999) 155-163.

[6] M.K. Lee and J.S. Jang, The optimum target values for a production process with three-class screening. *International Journal of Production Economics*, **49** (1997) 91-99.

[7] L. Li and D.B. Owen, Two-sided screening Procedures in the bivariate case. *Technomertics*, **21** (1979) 79-85.

[8] K.V. Mardia, J.T. Kent and J.M. Bibby, *Multivariate Analysis*, London: Academic Press Limited(1979).

[9] D.B. Owen and J.W. Boddie, A screening method for increasing acceptable product with some parameters unknown. *Technometrics*, **18** (1976) 195-199.

[10] D.B. Owen, D. Mcintire and E. Seymour, Tables using one or two screening variables to increase accptable product under one-sided specifications. *Journal of Quality Technology*, **7** (1975) 127-138.

[11] D.B. Owen and Y.H. Su, Screening based on normal variables. *Technometrics*, **19** (1977) 65-68.

[12] K. Tang and J. Tang, Design of screening procedures - a review. *Journal of Quality Technology*, **26** (1994) 209-226.

[13] J.G. Thomas and D.B. Owen, Improving the use of educational tests as selection tools. *Journal of Educational Statistics*, **2** (1977) 55-77.

Department of Biostatistics
Hamadan Med. Sci. University
P.O.Box 689, Hamadan, Iran
e-mail:H_Mahjub@yahoo.com

Department of Statistics
University of Newcastle
Newcastle, NE1 7RU, UK
e-mail:Trevor.Cox@ncl.ac.uk