Brazilian Journal of Probability and Statistics (2002), 16, pp. 47–58.

SUPERPOPULATIONS MODELS AND STRATIFICATION IN DIAGNOSTIC ABILITY STUDIES

Carlos Bouza

Facultad de Matematica y Computacion, Universidad de La Habana, San Lazaro y L., Habana, CP 10 400, Cuba. Email: bouza@matcom.uh.cu

Summary

The ability in diagnosis of a separator variable is studied using a superpopulation model for stratified sampling. The estimated errors of the predictors and the estimators are similar. The model errors are computed easily. They averaged smaller computing time in the analyzed examples.

1 Introduction

A common problem in medical research is to detect a certain disease. An expensive method may be used for preparing the diagnosis but the physician may be interested in evaluating the performance of an alternative one. A comparison between them is needed. This problem has been analyzed by different authors. Recently Bloch (1997) used a kappa index and Sukhatme and Beam (1994) derived non parametric based estimators of the diagnosis ability of a variable. In medical research commonly the aims of the physician is to evaluate the convenience of using a cheaper but less accurate procedure for diagnosing. For example he/she wants to evaluate if X-rays may be used instead of Computer Aided Axial Tomography for detecting the existence of a certain kind of tumors.

The ability of the auxiliary variable S (separator variable) in the characterization of the diagnosis is the aim of many basic medical research problems. The performance of S is evaluated by its ability (effectiveness) for detecting patients with the disease. To measure S is easier and/or cheaper than to use the other method. Without losing in generality we will consider that a large value of S suggests that the patient is classified as *positive* when $S > A$. He/she is considered as a *case*. Small values of S are expected in the patients without the disease. They are denominated negative and are denominated *controls*. Various methods have been proposed for comparing the performance of two diagnostic tests by using the measure of sensitivity

$$P_1 = Prob\,(\text{the test is positive}|\text{the patient has the disease})$$

The sensitivity of S evaluates its performance in the classification of cases. We denote by U the population of cases. The specificity is measured by

$$P_2 = Prob\,(\text{the test is negative}|\text{the patient does not have the disease})$$

The behavior in the classification of controls is evaluated by the sensitivity and specificity. The population of controls will be denoted by V. The trace of P_1 versus $Q_2 = 1 - P_2$, for S permits to evaluate the deviation of the test . The area under the resulting curve measures the diagnosis ability (DA). This has been called *diagnosability* by Pereira and Pericchi (1990). They developed its estimation using the Bayesian approach . This area summarizes the capability of the diagnostic test.

The random variable S , when the sample s_U is selected from U is modeled by $Y_h = S(h|h \in s_U)$, and by $X_t = S(t|(t \in s_V)$ plays the same role denoting by s_V a sample selected randomly from V. The probability $P(Y_h - X_t > 0)$ plays a key role in the study of the DA of S. Sukhatme and Beam (1994) employed stratified sampling for estimating it. Mann-Whitney statistics were derived for estimating and testing.

Three possible situations may arise in the applications:

Case 1 - U and V are divided into k strata

Case 2 - U is stratified but an unstratified sample is selected from V.

Case 3 - An unstratified sample is selected from U and V is stratified.

We will consider that the physicians have an aprioristic information that permits to characterize the conditional distribution of S. It is used for deriving a superpopulation model that links a Bernoulli random variable with the probability $Prob(Y > X) = \theta$. In our paper this model approach is used, within the inferential frame developed by Pothoff, Woodbury and Manton (1992). The results obtained provide a model-based procedure which is an alternative to the usual design-based developed by Sukhatme and Beam (1994). As Pothoff, Woodbury and Manton (1992) obtained approximations for computing confidence intervals when the weighted mean of parameters we used the corresponding theory for deriving them in the studied problem.

The behavior of the estimators proposed by Sukhatme and Beam (1994) is studied . Section 2 is devoted to this problem. The superpopulation rationale is used for developing the study. The derived model variances of the predictors are of easy computation in comparison with the counterparts obtained by Sukhatme and Beam (1994). In Section 3 different tests of hypothesis are considered. We propose to use ranks as strata weights. The design based estimators proposed by Sukhatme and Beam (1994) are compared with the model based predictors derived in this paper. A comparison is developed by using data from a study of the incidence of bronchial asthma. The DA of different concomitant variables is evaluated. The results of simulation experiments suggest that the design based tests may be considered as equivalent to their model based counterparts in terms of accuracy. The analysis of the averaged computing time sustained that the model based procedures are considerably faster.

2 Estimation of the area

The population of cases U has N persons and V has M. Usually homogeneous groups exist and the researcher aims to observe a representation of each of them in the sample. Then, if the existing information permits to determine non overlapping set of patients Stratified Sampling may be used. In addition this design ensures a gain in the accuracy of the estimators. We will consider that the populations are divided into k strata when possible. The cumulative distribution function (cdf) of S in the stratum U_h is $G_h(s)$. For V_j this role is played by $F_j(s)$. Take N_h and M_j as the number of patients that belong to U_h and V_j respectively, the weights $u_h = N_h/N$ and $\mathrm{v}_j = M_j/M$. The corresponding cdf's are given by:

$$G(z) = \sum_{h=1}^{k} u_h G_h(z)$$

and

$$F(z) = \sum_{j=1}^{k} v_j F_j(z)$$

Independent samples of size n_h and m_j are selected from the strata U_h and V_j. S_{pt} denotes the value S of in the $t - t_h$ patient from stratum p. Then the estimators of the within stratum cdf's are given by:

$$g_h(z) = \sum_{t=1}^{n_h} \frac{z_{ht}(x)}{n_h}$$

and

$$f_j(z) = \sum_{t=1}^{m_j} \frac{z_{jt}(x)}{m_j}$$

where

$$z_{pt}(r) = \begin{cases} 1 \text{ if } \ S_{pt} < z \\ 0 \text{ otherwise} \end{cases}$$

and

$$S_{pt} = \begin{cases} X_{pt} \text{ if } r = x \\ Y_{pt} \text{ if } r = y \end{cases}$$

Using the classic weighted mean estimator we derive easily the cdf's of the populations. They are

$$g(x) = \sum_{h=1}^{k} u_h g_h(x)$$

and

$$f(x) = \sum_{j=1}^{k} v_j f_j(x)$$

Sukhatme and Beam (1994) deduced, using the sampling design approach, that

$$\theta = Prob\,(Y - Z > 0) = \sum_{h=1}^{k}\sum_{j=1}^{k} u_h v_j \theta_{hj} = \sum_{h=1}^{k}\sum_{j=1}^{k} w_{hj}\theta_{hj}$$

where

$$\theta_{hj} = Prob(Y_{hi} - X_{jt} > 0) = \int F_j(x)\partial G_h(y)$$

θ is estimated unbiasedly by:

$$\widehat{\theta} = \sum_{h=1}^{k}\sum_{j=1}^{k} w_{hj} \sum_{i=1}^{n_h}\sum_{t=1}^{m_j} \frac{z_{hj}(it)}{n_h m_j} = \sum_{h=1}^{k}\sum_{j=1}^{k} u_h v_j \theta_{hj} = \sum_{h=1}^{k}\sum_{j=1}^{k} w_{hj}\widehat{\theta}_{hj} \qquad (2.1)$$

where

$$z_{hj}\,(it) = \begin{cases} 1 \text{ if } Y_{hi} - X_{jt} > 0 \\ 0 \text{ otherwise} \end{cases}$$

has a Bernoulli distribution with parameter θ_{hj} which is estimated unbiasedly by the corresponding Mann-Whitney statistic $\widehat{\theta}_{hj}$. Therefore the unbiasedness of (2.1), as an estimator of the area under the curve P_1 vs. Q_2 follows easily, see Sukhatme and Beam (1994).

The knowledge of the researcher may permit to characterize the expected behavior of X and Y by means of a superpopulation model . We will assume that the a priori distribution that describes each superpopulation belongs to a class such that:

$$E\,(z_{hj}(it)) = \theta_{hj}$$

and

$$Cov\left(z_{hj}(it), z_{h'j'}(i't')\right) = \begin{cases} \sigma^2_{hj} \text{ if } i = i',\ t = t',\ j = j' \text{ and } h = h' \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases}$$

Note that $\widehat{\theta}$ is model unbiased because under the superpopulation model each $\widehat{\theta}_{kj}$ is unbiased. Therefore its structure suggests that an adequate predictor is:

$$\widetilde{\theta} = \sum_{h=1}^{k}\sum_{j=1}^{k} w*_{hj}\frac{\widehat{\theta}_{hj}}{n*} \tag{2.2}$$

where

$$n* = \frac{(NM)^2}{\sum_{h=1}^{k}\sum_{j=1}^{k}(N_h M_j)^2}$$

and

$$w*_{hj} = n * w_{hj}$$

$n*$ is called "equivalent sample size".

Pothoff et. al. (1992) derived that if the parameter of interest belongs to the class

$$C = \left\{\mu|\mu = \frac{\sum_{r=1}^{R}\lambda_r\mu_r}{\sum_{r=1}^{R}\lambda_r}\right\}$$

then the use of predictors with the structure

$$\widehat{\mu} = \frac{\sum_{r=1}^{R}\lambda_r\widehat{\mu}_r}{\sum_{r=1}^{R}\lambda_r}$$

permits to develop T-Student inference whenever each $\widehat{\mu}_r$ is unbiased. As

$$w = \sum_{h=1}^{k}\sum_{j=1}^{k} w_{hj} = 1$$

the conditions established by Pothoff et. al. *1992(are satisfied in our case. The model variance of (2.2) is

$$\nu^2 = \sum_{h=1}^{k}\sum_{j=1}^{k} w^2_{hj}\frac{\sigma^2_{hj}}{n_h m_j}$$

An estimator of it is

$$\widehat{\nu}^2_1 = \sum_{h=1}^{k}\sum_{j=1}^{k} w^2_{hj}\sum_{i=1}^{n_h}\sum_{t=1}^{m_j}\frac{\left(z_{hj}(it) - \widetilde{\theta}\right)^2}{n*}$$

Under the model $E\left(z_{hj}(it)\right) = \theta_{hj}^2 + \sigma_{hj}^2$ and $E\left(\widetilde{\theta}\right) = \theta^2 + \nu^2$. Hence

$$E\left(\widehat{\nu}_1^2\right) = \sum_{h=1}^{k}\sum_{j=1}^{k} w_{hj}\left(\theta_{hj}^2 - \theta^2\right) + \sum_{h=1}^{k}\sum_{j=1}^{k} \frac{\left(n_h m_j^3 - n * w_{hj}\right) w_{hj}\sigma_{hj}^2}{n_h m_j n*}$$

The model unbiasedness of

$$\widehat{\sigma}_{hj}^2 = \sum_{i=1}^{n_h}\sum_{t=1}^{m_j} \frac{\left(z_{hj}(it) - \widehat{\theta}_{hj}\right)^2}{(n_h m_j - 1)}$$

with respect to σ_{hj}^2 suggests that we may use

$$\widehat{\nu}_2^2 = \sum_{h=1}^{k}\sum_{j=1}^{k} \frac{w_{hj}^2 \widehat{\sigma}_{hj}^2}{n_h m_j}$$

which unbiasedness is easily derived.

Sukhatme and Beam (1994) used a design based approach for obtaining that the variance of (2.2) is given by:

$$V(\widetilde{\theta}) = \sum_{q=1}^{4} V_q$$

where

$$\begin{aligned}
V_1 &= \sum_{h=1}^{k}\sum_{j=1}^{k} w_{hj} \frac{\left(P_{hj1} - P_{hj1}^2\right) + (n_h - 1)\left(P_{hj2} - P_{hj2}^2\right)}{n_h m_j} \\
V_2 &= \sum_{h=1}^{k}\sum_{j=1}^{k} w_{hj} \frac{(m_j - 1)\left(P_{hj3} - P_{hj3}^2\right)}{n_h m_j} \\
V_3 &= \sum_{h=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k} w_{hj} w_{hr} \frac{\left(P_{hjr1} - P_{hj1} P_{hr1}\right)}{m_h} \\
V_4 &= \sum_{h=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{k} w_{hj} w_{rj} \frac{\left(P_{hjr2} - P_{hj2} P_{rj2}\right)}{n_h m_j}
\end{aligned}$$

The involved parameters are defined as

$$P_{hj1} = Prob(Y_{jt} - X_{ih} > 0) = \int F_h(y) \partial G_j(y) \qquad (2.3)$$

$$\begin{aligned}
P_{hj2} &= Prob(Y_{jt} - X_{hi} > 0 \text{ and } Y_{jt'} - X_{hi'} > 0 | t \neq t', i \neq i') \\
&= \int (1 - G_j(y))^2 \partial F_h(y) \qquad (2.4) \\
P_{hj3} &= Prob(Y_{jt} - X_{hi} > 0 \text{ and } Y_{jt} - X_{hi'} > 0 | i \neq i') \\
&= \int F_h(y)^2 \partial G_j(y) \qquad (2.5) \\
P_{hjr1} &= Prob(Y_{jt} - X_{hi} > 0 \text{ and } Y_{rt'} - X_{hi'} > 0 | h \neq j) \\
&= \int (1 - G_j(y))^2 \partial F_h(y) \qquad (2.6) \\
P_{hjr2} &= Prob(Y_{jt} - X_{hi} > 0 \text{ and } Y_{jt} - X_{h'r} > 0 | i \neq r) \\
&= \int F_h(y) F_{h'}(y) \partial G_j(y) \qquad (2.7)
\end{aligned}$$

Following their proposal a naive estimation of $V(\widetilde{\theta})$ may be computed by placing the corresponding sample proportions instead of the unknown probabilities. Denoting it by $v_{(p)}\left(\widetilde{\theta}\right)$ Sukhatme and Beam (1994) argued that:

$$Z_{(p)} = \frac{\widetilde{\theta} - \theta}{\sqrt{V(\widetilde{\theta})}} \sqrt{M + N} \qquad (2.8)$$

is asymptotically normally distributed with zero mean and variance one. Therefore

$$T_{(p)} = \frac{\widetilde{\theta} - \theta}{\sqrt{\widehat{v}_{(;p)}(\widetilde{\theta})}} \sqrt{M + N} \qquad (2.9)$$

has a T-Student distribution that is approximated by the standard normal distribution when the sample size is sufficiently large

We will assume that the hypothesis of the normality of the errors of the superpopulation model $\theta = z_{ht}(it) + \epsilon_{ht}(it)$ holds. Then

$$Z_{(1)} = \frac{\widetilde{\theta} - \theta}{v} \qquad (2.10)$$

follows asymptotically the same distribution as $Z_{(p)}$ and $T_{(q)} = (\widetilde{\theta} - \theta)/\widehat{v}_q$, $q = 1, 2$ has a T-Student distribution with η, the integer closer to $n*$, degrees of freedom.

We will analyze the case in which only U is stratified. The probability $\theta_j = Prob\,(Y_{jh} - X > 0)$ is estimated unbiasedly by:

$$\widetilde{\theta} = \sum_{h=1}^{k} \sum_{t=1}^{m_j} \frac{z_{hj}(t)}{n m_j}$$

where

$$z_{hj}(t) = \begin{cases} 1 \text{ if } Y_{jt} - X > 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{2.11}$$

Hence, if $w_j = M_j/M$

$$\widetilde{\theta} = \sum_{h=1}^{k} w_j \widetilde{\theta}_j$$

is an unbiased predictor of it. When the classical approach is used the design variance of $\widetilde{\theta}_U$, see Sukhatme and Beam (1994), is:

$$\begin{aligned} V_U^2 = V\left(\widetilde{\theta}_U\right) &= \sum_{j=1}^{k} w_j^2 \frac{P_{j1} - P_{j1}^2 + (m_j - 1)\left(P_{j2} - P_{j1}^2\right) + (n-1)(P_{j3} - P_{j1})}{n m_j} \\ &+ \sum_{j=1}^{k} \sum_{j' \neq j}^{k} w_j w_{j'} \frac{P_{jj'}}{n} \end{aligned}$$

The involved probabilities are defined in terms of their distributions. Their explicit formulae are:

$$\begin{aligned} P_{j1} &= P\left(Y_{jt} - X > 0\right) = \int F(y) \partial G_j(y) \\ P_{j2} &= P\left(Y_{jt} - X > 0 \text{ and } Y_{jt'} - X > 0 | t \neq t'\right) = \int \left(1 - G_j(y)\right) \partial F(y) \\ P_{j3} &= P\left(Y_{jt} - X' > 0 \text{ and } Y_{jt} - X'' > 0 | t \neq t'\right) = \int F(y)^2 \partial G_j(y) \\ P_{jj'} &= P\left(Y_{jt} - X > 0 \text{ and } Y_{j't} - X > 0 | j \neq j'\right) = \\ &= \int \left(1 - G_j(y)\right) \left(1 - G_{j'}(y(\right) \partial F(y) \end{aligned}$$

See Suhatme and Beam (1994) for a detailed discussion. They argued that the test statistic $Z_U = (\widetilde{\theta}_{(U)} - \theta)/V_{(U)}$ has a standard normal distribution when n is sufficiently large. If $v_{(U)p}$ is the estimator of the variance derived by replacing the unknown probabilities P_{jq} by the sample proportions p_{jq}, $q = 1, 2, 3$. The corresponding T-Student test statistic is $t_{(U)} = (\widetilde{\theta}_{(U)} - \theta)/v_{(U)p}$. In this case the equivalent sample size in this model is given by:

$$n*_{(U)} = \frac{M^2}{\sum_{j=1}^{k} M_j^2}$$

Then the weight of U_j is $w_{(U)j} = n_{(U)}u_j$. The use of the same criteria suggests to use the predictor

$$\widetilde{\theta}_{(U)m} = \sum_{j=1}^{k} \frac{w_{(U)j}\widetilde{\theta}_j}{n*_{(U)}}$$

and to estimate the variance using:

$$v^2(\widetilde{\theta}_{(U)m}) = \sum_{j=1}^{k} \frac{w^2_{(U)j}(\widetilde{\theta}_j - \theta)^2}{n *_{(U)} -1}$$

and the bias is

$$B\left(v^2(\widetilde{\theta}_{(U)m})\right) = \sum_{h=1}^{k} \frac{\left(w_{(U)j} - w_{(U)j}\right)\sigma_j^2 + w_{(U)j}\left(\theta_j - \theta\right)^2}{n *_{(U)} -1}$$

Note that when $\theta_j = \theta$ for any stratum the estimator is model unbiased because $\sigma_j^2 = \theta(1-\theta)$ for any j. Note that the predictors belong to the class analyzed by Pothoff et. al [1992] therefore the distribution of the statistic $\mathrm{t}_{(U)m} = (\widetilde{\theta}_{(U)} - \theta)/v_{(U)m}$ can be approximated a T-Student with $n(u) * -1$ degrees of freedom approximately.

3 Comparisons of the tests

In the stratified case the test statistics depend on the calculus of the probabilities (2.3)-(2.7) or of their estimations. If the hypothesis to be tested is $H_0 : \theta = \theta_0$ we have that $\sigma_0^2 = \theta_0(1-\theta_0)$ and

$$Z_0 = \frac{\widetilde{\theta} - \theta}{\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\frac{w_{ij}}{n_i m_j}\right)^{\frac{1}{2}}\sigma_0}$$

is the test statistic.

If proportional allocation is used for fixing the stratum sample sizes then $\nu_0^2 = v_{AP}^2 = \frac{\sigma_0^2}{nm}$

The interest of the researcher is to test if S has a significative Diagnosis Ability (DA). If this hypothesis is accepted the cheap medical test may be used massively for classifying the patients at a first stage.

The null hypothesis $H_0 : \theta_{ij} = 0,5$ for any strata establishes if the classification can be regarded as a random result.. We will suppose that the strata with larger values of S have higher ranks of the test variable.

Then the researcher may assign ranks for weighting the importance of the strata, instead of s the sizes. We denote by r_h the rank of stratum U_ha and by R_j the rank of V_j. As the proposed model imposes the constraint that the weights must belong to the interval [0,1] and their sum is equal to one, we will use the rank based weights $u{*}_h = 2r_h/(k(k+1))$ and $v{*}_j = 2R_j/(k(k+1))$. If the DA is not significantly different from $\frac{1}{2}$ a proportional allocation is used and $v^2_{AP} = nm/4$. When the allocation criteria is $n_h = n/k$ and $m_j = m/k$ and the strata are weighted by $u{*}_h$ and $v{*}_j$ we have that $v^2_0 = v^2_{AE}$. Note that our proposal permits to evaluate the expected variability of the predictors if the developed functions of the ranks are used for weighting.

A simulation experiment was performed for comparing the behavior of the tests, based on the frequentist approach of Sukhatme and Beam (1994) and on the model based predictors proposed in our paper.. The data from 2.500 allergic patients, ascribed to anti allergic treatment, were used for establishing *controls* and *cases* of acute bronchial asthma. $N = M = 1.250$ and the overall sample sizes were $n = m = 125$. The probabilities (2.3)-(2.7) were calculated for four possible separator variables:

1. Biliar juice's PH. (BJPH).
2. Evaluation of Biliar Juice (EBJ).
3. Dyspepsia (DP).
4. Main Digestive Intolerance (MDI).

The laboratory analysis of biliar juice was evaluated and a score between 0 and 100 was assigned to each patient. They established a value to DP using qualitative information obtained from an interview. A similar procedure produced the necessary information for evaluating MDI. The analysis of the data yielded that the BJPH was normally distributed. EBJ and MDI had heavy tailed distribution functions, that were approximated by a Double Exponential Distribution. A contamination of two normal cdf's fitted the data collected on DP. The theoretical variances were calculated using this distributions.

Five strata were defined for BJPH and three for EBJ. Four levels of dyspepsia fixed the strata for DP and eight types of disturbances those for MDI. 1.000 samples were randomly selected and evaluated. The errors and their estimates were calculated and the computing time was measured. The behavior of the procedures was studied by analyzing the percentage of samples in which the true hypothesis was rejected. Table 1 gives the results obtained when Allocation Proportional to Size was used. The effect of accepting the normal approximation seems to be adequate. The differences between the signification level and the observed percent of rejects can be considered small because the sample sizes were only moderately large. In general the results of the tests based on the frequentist criteria are similar to those obtained using the proposed model approach. The distribution of $t_{(p)}$ and $t_{(U)}$) are approximated by the standard normal cdf but the T-Student tests, of the model based procedures, used the equivalent degrees of freedom for fixing the reject region. A slightly better behavior of the model T-tests was obtained. The computed variances for the estimators were between 0,7 and 1,9 minutes and their estimates averaged 0,56. The model counterparts moved within the interval [0,06 0,09] minutes. Hence the use of the tests based on the proposed predictors are considerably less costly.

Table 1

Percent of rejections of the true hypothesis $H_0 : \theta = \theta_0$ for $\alpha = 0,05$.
Allocation Proportional to Size

Separator		Stratification				Unstratification			
Variable	$z_{(p)}$	$t_{(p)}$	$z_{(1)}$	$t_{(1)}$	$t_{(2)}$	$z_{(U)}$	$t_{(U)}$	$z_{(U(m}$	$t_{(U)m}$
BJPH	8,1	6,7	8,7	7,8	7,3	8,8	7,9	8,5	6,7
RBJ	9,5	9,7	8,9	8,0	7,6	7,9	8,1	7,7	5,3
DP	7,2	8,4	8,9	6,9	6,4	7,8	7,4	7,6	6,7
MDI	5,6	5,9	6,6	5,3	6,1	9,6	8,3	7,5	6,3

The ranks of the strata were assigned by the decision makers taking into account the value of the computed within mean of each separator variable. The results are given in Table 2. They exhibit a pattern which is very similar to the results obtained for Allocation Proportional to Size. Note that the T-Student tests have a behavior closer to the fixed α than the normal ones. The intervals obtained for the computing time remain valid because the kernel of the computation of these tests is the calculation of the variances.

Table 2

Percent of rejections of the true hypothesis $H_0 : \theta = \theta_0$ for $\alpha = 0,05$.
Allocation Proportional to Ranks

Separator		Stratification				Unstratification			
Variable	$z_{(p)}$	$t_{(p)}$	$z_{(1)}$	$t_{(1)}$	$t_{(2)}$	$z_{(U)}$	$t_{(U)}$	$z_{(U(m}$	$t_{(U)m}$
BJPH	6,1	6,4	5,7	4,8	5,3	8,8	7,4	8,5	6,7
RBJ	9,8	7,0	6,9	5,1	5,2	7,9	6,1	7,7	5,7
DP	6,8	5,4	8,9	5,9	5,4	7,8	7,4	7,6	5,4
MDI	6,8	6,8	6,4	5,2	4,0	8,1	7,3	6,8	5,9

Acknowledgements

This paper was developed while the author was visiting the Departamento de Matemàticas Puras y Aplicadas of Universidad Simòn Bolìvar.The last revision took place during a sejour to Universidad A Coruña. The author thanks to the referees for their helpful comments which permitted to improve its presentation.

(Received January, 2001. Revised November, 2001.)

References

Bloch, D. A. (1997). Comparing two diagnostic tests against the same "gold standard" in the same sample. *Biometrics*, **53**, 73–85.

Bouza, C. (1995). Linear rank tests derived from superpopulation models. *Biometrical Journal*, **37**, 37–85.

Pereira, C. A. B. and Pericchi. L. R. (1990). Analysis of diagnosability. *Applied Statistics*, **39**, 189–204.

Pothoff, R. A. Woodbury, M. A. and Manton, K. G. (1992). Equivalent sample sizes and equivalent degrees of freedom:refinements for inference using survey weights under superpopulation models. *J. Amer. Stat. Ass.*, **87**, 383–396.

Sukhatme, S. and Beam, C.A. (1994). Stratification in nonparametric ROC studies. *Biometrics*, **50**, 147–163.