UW Biostatistics Working Paper Series

12-7-2005

Empirical Likelihood Inference for the Area Under the ROC Curve

Gengsheng Qin
Georgia State University, gqin@gsu.edu

Xiao-Hua Zhou
University of Washington, azhou@u.washington.edu

Suggested Citation
Qin, Gengsheng and Zhou, Xiao-Hua, "Empirical Likelihood Inference for the Area Under the ROC Curve" (December 2005). *UW Biostatistics Working Paper Series*. Working Paper 268.
http://biostats.bepress.com/uwbiostat/paper268

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.
Copyright © 2011 by the authors

1. Introduction

The accuracy of a diagnostic test can be measured by its sensitivity and specificity, which are defined as the probabilities of correctly identifying the diseased and non-diseased subjects, respectively. When the response of a test is continuous, we need to choose a cut-off point for positivity of disease in order to compute sensitivity and specificity of the test. Let Y and X be the results of a continuous-scale test for a diseased and a non-diseased subject, respectively. For a given cut-off point c, the sensitivity and specificity of the test are defined as

$$Se = P(Y \geq c) = 1 - G(c), \qquad Sp = P(X \leq c) = F(c),$$

respectively, where F and G are the distribution functions of X and Y, respectively. Therefore, for a fixed value of specificity at p, the corresponding sensitivity of the test is $R(p) = 1 - G(F^{-1}(p))$, where $F^{-1}(p)$ is the inverse function of $F(p)$. The ROC curve, $R(p)$, is a plot of sensitivity against 1-specificity as the cut-off point c runs through the whole range of possible test values. In practice, the optimal cut-off point is not known because the distribution function F of the non-diseased population is unknown. The area under the curve (AUC), defined as $\delta = \int_o^1 R(p)dp$, is the most commonly used summary measure of diagnostic accuracy for a continuous-scale diagnostic test.

Bamber (1975) showed that the AUC $\delta = P(Y \geq X)$. This can be interpreted as the probability that in a randomly selected pair of diseased and non-diseased subjects, the test value of the diseased subject is higher than or equal to that of the non-diseased subject. In a more general context, Wolfe and Hogg (1971) recommended the use of this index as a general measure for the differences between two distributions.

One important problem for the inference on the AUC is how to construct a confidence interval for δ. Let $X_1, \ldots, X_m$ be test results of a random sample of non-diseased subjects and $Y_1, \ldots, Y_n$ be test results of a random sample of diseased subjects. Traditionally, the

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

well known Mann-Whitney (hereafter MW) two-sample rank statistic, defined by

$$\widehat{\delta} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} I(Y_j \geq X_i),$$

is used as a non-parametric estimator of the AUC. Based on the asymptotic normality of MW statistic, we can construct a confidence interval (hereafter MW interval) for AUC. Although the MW estimator of the AUC is known to be unbiased, and the coverage probability of the MW interval is asymptotically correct, the normal approximation based MW interval suffers from low coverage accuracy for high values of the AUC (e.g., 0.80 to 0.95, which are of most interest in diagnostic tests) when sample sizes for diseased and non-diseased subjects are small and unequal. Therefore, it is desirable to find an reliable alternative approach for constructing a confidence interval of the AUC.

Empirical likelihood (hereafter EL) (Owen, 1990, 2001) is a popular non-parametric method traditionally used for providing confidence intervals for the mean. The EL method has many advantages over other non-parametric methods. For example, it has better small sample performance than approaches based on normal approximation, and it studentizes internally, thereby eliminating the need for a pivot; see Hall and LaScala (1990), and Owen (2001) for a more discussion.

In this paper, we develop an empirical likelihood (EL) approach for the inference on the AUC. The main difficulty of developing the EL method for the AUC is that the standard EL method can not be directly applied to the AUC setting when F and G are unknown. We must use an adjusted EL method (cf. section 2). These kinds of adjustment are simple and work well for many applications (Qin and Jing, 2001, 2003; Wang, Oliver and Härdle, 2004; Wang and Rao, 2002). In Section 2 we develop the EL theory for the AUC and propose an EL based confidence interval for the AUC. In Section 3 we extend the EL inference on the AUC developed in Section 2 to stratified samples. In Section 4, we conduct simulation studies to compare the relative performance of the proposed EL based interval with the existing normal approximation based intervals and bootstrap intervals

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

for the AUC. In Section 5, we apply our EL interval for the AUC to a real example. The proof of the EL theory is deferred until the appendix.

2. Empirical Likelihood Method for AUC

For test value Y from a diseased subject, Pepe and Cai (2002) defined the placement value as

$$U = 1 - F(Y).$$

This value can be interpreted as the proportion of non-diseased subjects with their test values larger than Y. It essentially marks the placement of Y within the non-diseased distribution. It is easily seen that

$$E(1-U) = E(F(Y)) = P(Y \geq X) = \delta.$$

Noticing this relationship between the AUC (δ) and the placement value U, we can derive an empirical likelihood procedure for the AUC. Let $\mathbf{p} = (p_1, p_2, \cdots, p_n)$ be a probability vector, i.e., $\sum_{j=1}^{n} p_j = 1$ and $p_j \geq 0$ for all j. The profile empirical likelihood for AUC, evaluated at the true value δ_0 of δ, can be defined as

$$\tilde{L}(\delta_0) = \sup\left\{ \prod_{j=1}^{n} p_j : \sum_{j=1}^{n} p_j = 1, \quad \sum_{j=1}^{n} p_j(1 - U_j - \delta_0) = 0 \right\},$$

where $U_j = 1 - F(Y_j), \quad j = 1, 2, \cdots, n$. Since placement values U_j's depend on the unknown distribution function F of the non-diseased population, replacing F by its empirical distribution $\widehat{F}$, we get an estimated empirical likelihood for AUC:

$$L(\delta_0) = \sup\left\{ \prod_{j=1}^{n} p_j : \sum_{j=1}^{n} p_j = 1, \quad \sum_{j=1}^{n} p_j(1 - \widehat{U}_j - \delta_0) = 0 \right\},$$

where $\widehat{U}_j = 1 - \widehat{F}(Y_j), \quad j = 1, 2, \cdots, n$. Then, by the Lagrange multiplier, we can easily get

$$p_j = \frac{1}{n}\left\{1 + \lambda(1 - \widehat{U}_j - \delta_0)\right\}^{-1}, \qquad j = 1, \cdots, n,$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

where λ is the solution of

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1-\widehat{U}_j-\delta_0}{1+\lambda(1-\widehat{U}_j-\delta_0)}=0. \tag{1}$$

Note that $\prod_{j=1}^{n} p_j$, subject to $\sum_{j=1}^{n} p_j = 1$, attains its maximum n^{-n} at $p_j = n^{-1}$. So we define the empirical likelihood ratio at δ_0 as

$$R(\delta_0)=\prod_{j=1}^{n}(np_j)=\prod_{j=1}^{n}\{1+\lambda(1-\widehat{U}_j-\delta_0)\}^{-1}.$$

Hence the corresponding empirical log-likelihood ratio is

$$l(\delta_0)=-2\log R(\delta_0)=2\sum_{j=1}^{n}\log\{1+\lambda(1-\widehat{U}_j-\delta_0)\}, \tag{2}$$

where λ is the solution to the equation (1).

Since $\widehat{U}_j$'s are not independent, we cannot apply the standard empirical likelihood theory to derive the asymptotic distribution of our empirical likelihood ratio $l(\delta_0)$. However, in Appendix A, we show that $l(\delta_0)$ still follows a scaled χ^2 distribution. We summarize the asymptotic distribution of the empirical log-likelihood ratio for the AUC in the following theorem.

The following theorem establishes the asymptotic distribution of the empirical log-likelihood ratio for AUC.

Theorem 2.1. *If δ_0 is the true value of AUC, then the limiting distribution of $l(\delta_0)$, defined by (2), is a scaled chi-square distribution with one degree of freedom. That is,*

$$r(\delta_0)l(\delta_0)\xrightarrow{\mathcal{L}}\chi_1^2, \tag{3}$$

where the scale constant $r(\delta_0)$ is

$$r(\delta_0)=\frac{m}{m+n}\frac{\sum_{j=1}^{n}(1-\widehat{U}_j-\delta_0)^2}{nS^2},$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

with

$$
\begin{aligned}
S^2 &= \frac{mS_{01}^2 + nS_{10}^2}{m+n}, \\
S_{10}^2 &= \frac{1}{(m-1)n^2}\left[\sum_{i=1}^{m}(R_i - i)^2 - m\left(\bar{R} - \frac{m+1}{2}\right)^2\right], \\
S_{01}^2 &= \frac{1}{(n-1)m^2}\left[\sum_{j=1}^{n}(S_j - j)^2 - n\left(\bar{S} - \frac{n+1}{2}\right)^2\right], \\
\bar{R} &= \frac{1}{m}\sum_{i=1}^{m} R_i, \qquad\qquad \bar{S} = \frac{1}{n}\sum_{j=1}^{n} S_j.
\end{aligned}
$$

Here R_i is the rank of $X_{(i)}$ (the i-th ordered value among X_i's) in the combined sample of X_i's and Y_j's, and S_j is the rank of $Y_{(j)}$ (the j-th ordered value among Y_j's) in the combined sample of X_i's and Y_j's.

Empirical likelihood confidence intervals for the AUC can be constructed as follows. Let

$$R_\alpha(\delta) = \{\delta : r(\widehat{\delta})l(\delta) \leq \chi_1^2(1-\alpha)\}, \tag{4}$$

where $\chi_1^2(1-\alpha)$ is the $(1-\alpha)$-*th* quantile of the chi-square distribution χ_1^2. Then from Theorem 2.1, $R_\alpha(\delta)$ gives an approximate confidence interval for the AUC with asymptotically correct coverage probability $1-\alpha$, i.e.,

$$P(\delta_0 \in R_\alpha(\delta)) = 1 - \alpha + o(1).$$

3. EL Intervals for AUC with Stratified Samples

In this section we extend the EL inference on the AUC, developed in Section 2, to stratified samples. Suppose we have L institutions participating in an ROC study of a continuous-scale diagnostic test. We use l to index the lth institution. Let X_l and Y_l

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

be the results of a continuous-scale test for a non-diseased and a diseased subject in the lth institution, F_l and G_l are the distribution functions of X_l and Y_l, respectively. Let $X_{l1}, \ldots, X_{lm_l}$ be test results of a random sample of non-diseased patients and $Y_{l1}, \ldots, Y_{ln_l}$ be test results of a random sample of diseased subjects in the lth institution.

In this paper, we don't assume that F_l's and G_l's are homogeneous across institutions. Instead, we allow F_l's and G_l's to be different by institutions and only assume the accuracy of the test is the same across institutions. That is, we assume that $\delta_1 = \delta_2 = \cdots \delta_L = \delta$, where δ_l denotes the AUC for the lth institution. This is the same assumption made in Tsimikas, Bosch, Coull and Barmi (2002) when the response of a diagnostic test is categorical (ordinal rating).

Let $\mathbf{p}_l = (p_{l1}, p_{l2}, \cdots, p_{ln_l})$ be a probability vector for $l = 1, 2, \cdots, L$. Similar to those in Section 2, we can define the profile EL for the common diagnostic accuracy as measured by the AUC, δ, as follows:

$$L(\delta) = \sup \left\{ \prod_{l=1}^{L} \prod_{j=1}^{n_l} p_{lj} : \sum_{j=1}^{n_l} p_{lj} = 1, \quad \sum_{j=1}^{n_l} p_{lj}(1 - \widehat{U}_{lj} - \delta) = 0, \quad l = 1, 2, \cdots, L \right\},$$

where $\widehat{U}_{lj} = 1 - \widehat{F}_l(Y_{lj}), \quad l = 1, 2, \cdots, L, \quad j = 1, 2, \cdots, n_l$, and $\widehat{F}_l$ is the empirical distribution of F_l.

The corresponding empirical log-likelihood ratio is

$$l(\delta) = 2 \sum_{l=1}^{L} \sum_{j=1}^{n_l} \log\{1 + \lambda_l(1 - \widehat{U}_{lj} - \delta)\}, \tag{5}$$

where λ_l is the solution to the following equation:

$$\frac{1}{n_l} \sum_{j=1}^{n_l} \frac{1 - \widehat{U}_{lj} - \delta}{1 + \lambda_l(1 - \widehat{U}_{lj} - \delta)} = 0, \quad l = 1, 2, \cdots, L. \tag{6}$$

Once again, we cannot use the standard empirical likelihood theory to derive the asymptotic distribution of $l(\delta_0)$ because $\widehat{U}_{lj}$'s are not independent. In Appendix, we show that

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

the asymptotic distribution of $l(\delta_0)$ is still a weighted sum of independent χ^2 distributions. We summarize the asymptotic result in the following theorem.

Theorem 3.1. *If δ_0 is the true value of the common AUC, then the limiting distribution of $l(\delta_0)$ is a weighted sum of independent chi-square distributions with one degree of freedom. That is,*

$$l(\delta_0) \xrightarrow{\mathcal{L}} w_1\chi^2_{1,1} + \cdots + w_L\chi^2_{L,1}, \tag{7}$$

where the weights $w_l = \lim_{m_l,n_l\to\infty} \widehat{w}_l(\delta_0)$, $1 \le i \le L$, with

$$\begin{aligned}
\widehat{w}_l(\delta_0) &= \frac{m_l+n_l}{m_l}\frac{n_l S_l^2}{\sum_{j=1}^{n_l}(1-\widehat{U}_{lj}-\delta_0)^2}, \\
S_l^2 &= \frac{m_l S_{01}^2(l) + n_l S_{10}^2(l)}{m_l+n_l}, \\
S_{10}^2(l) &= \frac{1}{(m_l-1)n_l^2}\left[\sum_{i=1}^{m_l}(R_i(l)-i)^2 - m_l\left(\bar{R}_l - \frac{m_l+1}{2}\right)^2\right], \\
S_{01}^2(l) &= \frac{1}{(n_l-1)m_l^2}\left[\sum_{j=1}^{n_l}(S_j(l)-j)^2 - n_l\left(\bar{S}_l - \frac{n_l+1}{2}\right)^2\right], \\
\bar{R}_l &= \frac{1}{m_l}\sum_{i=1}^{m_l} R_i(l), \qquad\qquad \bar{S}_l = \frac{1}{n_l}\sum_{j=1}^{n_l} S_j(l).
\end{aligned}$$

Here $R_i(l)$ is the rank of $X_{l(i)}$ (the i-th ordered value among $X_{l1},\ldots,X_{lm_l}$) in the l-th combined sample of X_{li}'s and Y_{lj}'s, and $S_j(l)$ is the rank of $Y_{l(j)}$ (the j-th ordered value among $Y_{l1},\ldots,Y_{ln_l}$) in the l-th combined sample of X_{li}'s and Y_{lj}'s.

We can then construct the EL confidence interval for the common AUC as follows. Let

$$R_\alpha(\delta) = \{\delta : l(\delta) \le c_{1-\alpha}\}, \tag{8}$$

where $c_{1-\alpha}$ is the $(1-\alpha)$-*th* quantile of the weighted chi-square distribution $\widehat{w}_1(\widehat{\delta})\chi^2_{1,1}+\cdots+\widehat{w}_L(\widehat{\delta})\chi^2_{L,1}$. In practice, a simple Monte Carlo simulation is needed to calculate the critical

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

value $c_{1-\alpha}$ in (8). This can be done by first generating a large number of realizations of $\widehat{w}_1(\widehat{\delta})\chi^2_{1,1} + \cdots + \widehat{w}_L(\widehat{\delta})\chi^2_{L,1}$ and then taking $c_{1-\alpha}$ to be the $(1-\alpha)$-*th* sample quantile. Then, from the earlier discussion, $R_\alpha(\delta)$ gives an approximate confidence interval for the AUC with asymptotically correct coverage probability $1-\alpha$.

4. Simulation Studies for EL Intervals

In this section, we report the three simulation studies for evaluating coverage accuracy and interval length of the newly proposed EL interval for the AUC when the AUC is taken to be 0.8 (moderate accuracy), 0.9 and 0.95 (high accuracy) in finite-sample sizes. For simplicity, we take $L = 1$ in these simulation studies. In each study, we generated 10,000 random samples of size m from F for test responses of non-diseased subjects, and another 10,000 independent random samples of size n from G for test responses of diseased subjects.

Since the AUC is restricted to [0,1], Pepe (2003) has argued that an asymmetric confidence interval within (0,1) should be preferred. Using a logistic transformation, the lower and upper limits of $(1-\alpha)\%$ confidence interval for logit $AUC = log(AUC/(1-AUC))$ are

$$\text{LL} = \log\frac{\widehat{\delta}}{1-\widehat{\delta}} - z_{1-\alpha/2}\frac{\sqrt{var(\widehat{\delta})}}{\widehat{\delta}(1-\widehat{\delta})}, \qquad \text{UL} = \log\frac{\widehat{\delta}}{1-\widehat{\delta}} + z_{1-\alpha/2}\frac{\sqrt{var(\widehat{\delta})}}{\widehat{\delta}(1-\widehat{\delta})},$$

respectively. Therefore, the $(1-\alpha)\%$ confidence interval for the AUC is

$$\left[\frac{exp(LL)}{1+exp(LL)}, \frac{exp(UL)}{1+exp(UL)}\right].$$

Pepe (2003, p107) recommended the use of the logit transformation based confidence intervals (Hereafter LT interval) for the AUC. Because bootstrap confidence intervals usually are good competitors to the EL based intervals; hence for comparison, we include

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

the LT interval, the normal approximation based interval (MW interval), the standard percentile bootstrap (PB) interval, and percentile-t bootstrap (PTB) interval for the AUC in these simulation studies.

In the first simulation study, F is chosen to be the standard normal distribution (with mean $\mu_0 = 0$ and standard deviation $\sigma_0 = 1$), and G is chosen to be a normal distribution $\mathrm{N}(\mu, \sigma^2)$ having mean $\mu = \sqrt{5}\Phi^{-1}(\delta)$ (Φ is the cdf of the standard normal distribution) and standard deviation $\sigma = 2$. In the second study, F is chosen to be the standard log-normal distribution, and G a log-normal distribution Log-$\mathrm{N}(\mu, \sigma^2)$ with $\mu = \sqrt{401}\Phi^{-1}(\delta)$ and $\sigma = 20$. Note that in these simulation settings, δ represents the true AUC and is related to the mean and standard deviation by the following relation:

$$\delta = \Phi\left(\frac{\mu - \mu_0}{\sqrt{\sigma^2 + \sigma_0^2}}\right). \tag{9}$$

Thus, δ can be estimated by substituting the maximum likelihood estimates of the mean and standard deviation into (9). Obuchowski and McClish (1997, formula (6)) gave the variance estimate for this estimator of AUC. Therefore another normal approximation based confidence interval for the AUC (hereafter ML interval) can be constructed by applying the asymptotic normality of the ML based estimate for the AUC.

In the third study, F is chosen to be the standard exponential distribution (with rate $\upsilon = 1$), and G a exponential distribution with rate $\theta = 1/\delta - 1$. The AUC is

$$\delta = \frac{\upsilon}{\upsilon + \theta} \tag{10}$$

and can be estimated by replacing υ and θ by its maximum likelihood estimates. This maximum likelihood (ML) based estimate for the AUC is

$$\widetilde{\delta} = \frac{\bar{Y}}{\bar{X} + \bar{Y}}.$$

Using the delta method, we can find the asymptotic variance of $\widetilde{\delta}$ and hence obtain the

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

normal approximation based ML interval for AUC as follows:

$$\left[\widetilde{\delta} \mp z_{1-\alpha/2}\sqrt{\frac{1}{m}+\frac{1}{n}}\frac{\bar{X}\bar{Y}}{(\bar{X}+\bar{Y})^2}\right],$$

where $z_{1-\alpha/2}$ is $(1-\alpha/2)$-th quantile of standard normal distribution.

For each of the parameter and distribution settings mentioned above, two sided 90% and 95% EL intervals, MW intervals, ML intervals and LT intervals for the AUC are constructed for five combinations of sample sizes of (m, n), followed by the calculation of coverage probabilities, average interval lengths. The computation results are presented in Tables 1-6.

Place Tables 1-6 here

Tables 1-6 indicate that EL intervals always have better coverage accuracy than MW intervals. Particularly, for high values of the AUC, EL intervals greatly outperform MW intervals. All intervals have similar interval lengths.

The ML interval for the AUC has good coverage accuracy when the underlying parametric assumptions are true. Our EL intervals have similar coverage accuracy to the ML intervals for parametric models considered here. While the validity of ML intervals is relied on the parametric distributional assumptions, the EL interval is purely a nonparametric approach and makes no assumptions about the distributional forms for diseased and non-diseased populations. On additional problem with the parametric ML interval is that it does not have the invariant property, which requires the ROC curve of a test be invariant to any monotone increasing transformation of test results, a fundamental property of an ROC curve.

The LT interval for the AUC also has good coverage accuracy but has the longest interval length. One serious problem with the LT interval is that it can break down when

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

the observed AUC $\widehat{\delta}$ is close to one (see Tables 3-4 when true AUC=0.95. "NA" means "not available"). Particularly, if the estimated AUC $\widehat{\delta} = 1$, no LT interval for the AUC can be constructed. For those situations when we can compute the LT intervals, our EL intervals have similar coverage accuracy as the LT intervals.

Comparing with the standard bootstrap intervals for the AUC, EL intervals have better coverage accuracy than the PB intervals. The PTB intervals over-cover the AUC and have bigger coverage errors than the EL intervals in most settings considered here. Furthermore, the PTB is the computationally the most extensive method among all the methods considered here.

In summary, we suggest the use of EL intervals for the AUC when the underlying distributions for diseased and non-diseased populations are unknown. The ML interval for the AUC should be used when the underlying parametric models are believed to be true.

5. A Real Application: Dermoscope Example

We illustrate the application of the proposed EL interval for AUC in the study of the accuracy of dermatoscopy in distinguishing patients with malignant melanoma (MM) from those without MM.

The most deadly kind of skin disease is malignant melanoma (MM), and early detection of MM combined with excision of MM is the only way to cure patients with MM. Stolz et al. (1994) studied the accuracy of clinical evaluation with the aid of dermatoscopy in detecting malignant melanoma by using the ABCD (Asymmetry, irregular Border, different Colors, and Diameter larger than 6mm) rule. The dermatoscopy is a hand-held instrument for skin surface microscopy at 10 times magnification (Stolz et al., 1989). Dermatologists want to know what the diagnostic accuracy of dermatoscopy is in detecting

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

patients with MM. The study sample consists of 21 patients with MM and 51 patients with benign melanocytic lesions, and the gold standard used in the study is biopsy. Because the distributions of the measurement from dermatoscopy for the two group of patients are unknown and the sample sizes are unequal, based our simulation studies, we would use the EL interval for the AUC as the range of the diagnostic accuracy of dermatoscopy. The estimated AUC of dermatoscopy based on MW estimator is 0.900. The 95% EL interval for the AUC is [0.812, 0.955]. This EL interval suggests that the dermatoscopy has moderate to high level of diagnostic accuracy in detecting patients with MM.

6. Discussion

In this paper we have proposed an empirical likelihood (EL) based confidence interval for the area under the ROC curve of a continuous-scale test. We have shown that the proposed method has some nice theoretical property, and our simulation results have indicated the proposed EL based interval outperforms the existing methods, particularly when the AUC is close to 1.0.

Our proposed method is an additional contribution for constructing an confidence interval for the AUC of a highly accurate diagnostic test. When the response of a diagnostic test is ordinal, Tsimika et al (2002) have proposed a profile-likelihood method for constructing such a interval. However, their method cannot be applied when the response of the test is continuous.

6. Appendix: Proof of Theorems

We need the following lemma for the proofs of Theorems 2.1 and 3.1.

Lemma 1. *Under the same conditions as in Theorem 2.1, we have*

(i). $\frac{1}{n}\sum_{j=1}^{n}(1-\widehat{U}_j-\delta_0)^2 \xrightarrow{p} \sigma_0^2$, where $\sigma_0^2 = E\left[F^2(Y)\right] - \delta_0^2$.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

(ii). $\left(\frac{mn}{m+n}\right)^{1/2} \frac{\widehat{\delta_0}-\delta_0}{S} \xrightarrow{\mathcal{L}} N(0,1)$, where $\widehat{\delta_0} = \frac{1}{n}\sum_{j=1}^{n}(1-\widehat{U}_j)$.

Proof of Lemma 1.

Lemma 1(i) follows from the uniform consistency of empirical distribution $\widehat{F}$ and the fact that

$$\frac{1}{n}\sum_{j=1}^{n}(1-U_j-\delta_0)^2 \xrightarrow{p} E\left[F(Y)-\delta_0\right]^2 = \sigma_0^2$$

Lemma 1(ii) is from Sen (1967).

Proof of Theorem 2.1:

Using Lemma 1 and the similar arguments used in Owen (1990), we can prove that $|\lambda| = O_p(n^{-1/2})$. Noting that $\max_{1\le j\le n}|1-\widehat{U}_j-\delta_0| = O(1)$, a.s., by Taylor's expansion to (2), we have

$$\begin{aligned} l(\delta_0) &= \sum_{j=1}^{n}\log\{1+\lambda(1-\widehat{U}_j-\delta_0)\} \\ &= \sum_{j=1}^{n}\left[\lambda(1-\widehat{U}_j-\delta_0)-\frac{1}{2}\left(\lambda(1-\widehat{U}_j-\delta_0)\right)^2\right]+r_n \end{aligned}$$

with

$$|r_n| \le C\sum_{j=1}^{n}|\lambda(1-\widehat{U}_j-\delta_0)|^3 \le C|\lambda|^3 n = O\left(n^{-1/2}\right).$$

From (1), it follows that

$$\lambda = \frac{\sum_{j=1}^{n}(1-\widehat{U}_j-\delta_0)}{\sum_{j=1}^{n}(1-\widehat{U}_j-\delta_0)^2} + O_p\left(n^{-1/2}\right),$$

$$\sum_{j=1}^{n}\lambda(1-\widehat{U}_j-\delta_0) = \sum_{j=1}^{n}\left[\lambda(1-\widehat{U}_j-\delta_0)\right]^2 + o_p(1).$$

Therefore,

$$rl(\delta_0) = r\sum_{j=1}^{n}\lambda(1-\widehat{U}_j-\delta_0)+o_p(1)$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

$$
\begin{aligned}
&= r \cdot \frac{\left[\frac{1}{\sqrt{n}} \sum_{j=1}^{n} (1 - \widehat{U}_j - \delta_0)\right]^2}{\frac{1}{n} \sum_{j=1}^{n} (1 - \widehat{U}_j - \delta_0)^2} + o_p(1) \\
&= r \cdot \frac{\left[\sqrt{n}(\widehat{\delta}_0 - \delta_0)\right]^2}{\frac{1}{n} \sum_{j=1}^{n} (1 - \widehat{U}_j - \delta_0)^2} + o_p(1) \\
&= \left[\left(\frac{mn}{m+n}\right)^{1/2} \frac{\widehat{\delta}_0 - \delta_0}{S}\right] + o_p(1) \\
&\xrightarrow{\mathcal{L}} \chi_1^2.
\end{aligned}
$$

The proof of Theorem 2.1 is thus completed.

Proof of Theorem 3.1.

Similar to the proof of Theorem 2.1, we have

$$
\lambda_l = \frac{\sum_{j=1}^{n_l} (1 - \widehat{U}_{lj} - \delta_0)}{\sum_{j=1}^{n_l} (1 - \widehat{U}_{lj} - \delta_0)^2} + O_p\left(n_l^{-1/2}\right),
$$

$$
\sum_{j=1}^{n_l} \lambda_l (1 - \widehat{U}_{lj} - \delta_0) = \sum_{j=1}^{n_l} \left[\lambda_l (1 - \widehat{U}_{lj} - \delta_0)\right]^2 + o_p(1),
$$

and

$$
\begin{aligned}
l(\delta_0) &= \sum_{l=1}^{L} \sum_{j=1}^{n_l} \lambda_l (1 - \widehat{U}_{lj} - \delta_0) + o_p(1) \\
&= \sum_{l=1}^{L} \widehat{w}_l \cdot \frac{1}{\widehat{w}_l} \sum_{j=1}^{n_l} \lambda_l (1 - \widehat{U}_{lj} - \delta_0) + o_p(1) \\
&= \sum_{l=1}^{L} \widehat{w}_l \cdot \left[\left(\frac{m_l n_l}{m_l + n_l}\right)^{1/2} \frac{\widehat{\delta_{l0}} - \delta_0}{S_l}\right] + o_p(1) \\
&\xrightarrow{\mathcal{L}} w_1 \chi_{1,1}^2 + \cdots + w_L \chi_{L,1}^2,
\end{aligned}
$$

where $\widehat{\delta_{l0}} = \frac{1}{n_l} \sum_{j=1}^{n_l} (1 - \widehat{U}_{lj})$. The proof of Theorem 3.1 is completed.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

References

Bamber DC (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic curve graph. *J Math Psychology*, **12**, 387-415.

Obuchowski NA, and McClish DK (1997). Sample size determination for diagnostic accuracy studies involving binomal ROC curve indices. *Statistics in Medicine*, **16**, 1529-1542.

Owen A (1990). Empirical likelihood ratio confidence regions. *Ann Statist*, **18**, 90-120.

Owen A (2001). Empirical likelihood. *Chapman & Hall/CRC*. New York.

Pepe MS and Cai T (2004). The analysis of placement values for evaluating discriminatory measures. *Biometrics*, **60**, 528-535.

Pepe MS (2003). *Statistical Evaluation of Diagnostic Tests and Biomakers*. Oxford University Press.

Qin, GS and Jing, BY (2001). Censored partial linear models and empirical likelihood. *J Multi Anal*, **78**, 37-61.

Qin, GS and Jing, BY (2001). Empirical likelihood for censored linear regression. *Scand J Statistics*, **28**, 661-673.

Qin, GS and Tsao, M (2003). Empirical likelihood inference for median regression models of censored survival data. *J Multi Anal*, **85**, 416-430.

Sen PK (1967). A note on asymptotically distribution-free confidence bounds for P(X ¡ Y) based on two independent samples. *Sankhyã, Series A*, **29**, 95-102

Stolz W, Bilek P, Landthaler M, Merkle T, and Braun-Falco O (1989). Skin surface microscopy. *Lancet* **2**, 864-865.

Stolz W, Riemann A, Cognetta AB, Pillet L, Abmayr W, Holzel D, Bilek P, Nachbar F, and Landthaler M, Braun-Falco O (1994). ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. *European Journal of Dermatology* **4**, 521-527.

Tsimikas JV, Bosch RJ, Coull BA and Barmi HE (2002). Profile-likelihood inference for

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

highly accurate diagnostic tests. *Biometrics*, **58**, 946-956.

Wang, QH, Linton, O and Härdle, W (2004). Semiparametric regression analysis with missing response at random. *J Amer Statist Assoc*, **99**, 334-345.

Wang, QH, Rao, JNK (2002). Empirical likelihood-based inference under imputation for missing response data. *Ann Statist*, **30**, 896-924.

Wang, QH, Rao, JNK (2002). Empirical likelihood-based inference in linear errors-in-covariables models with validation data. *Biometrika*, **89**, 345-358.

Wolfe DA, Hogg RV (1971). On constructing statistics and reporting data. *American Statistician*, **25**, 27-30.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

Table 1. Normal distribution: coverage probabilities and average lengths of 90% confidence intervals for AUC. The average lengths are the numbers in the rows below the corresponding coverage probabilities.

AUC	(m,n)	EL	MW	ML	LT	PB	PTB
0.80	(50,50)	0.8938	0.8911	0.8858	0.9101	0.8930	0.9240
		0.1499	0.1519	0.1433	0.1519	0.1502	0.1626
	(80,80)	0.8996	0.8975	0.8859	0.9016	0.8860	0.9200
		0.1188	0.1197	0.1139	0.1200	0.1185	0.1259
	(100,100)	0.8988	0.8988	0.8956	0.9010	0.9000	0.9190
		0.1064	0.1071	0.1019	0.1073	0.1071	0.1126
	(50,80)	0.8917	0.8886	0.8955	0.9036	0.8830	0.9100
		0.1246	0.1257	0.1197	0.1262	0.1252	0.1327
	(70,100)	0.8909	0.8842	0.8943	0.9018	0.8860	0.9090
		0.1102	0.1109	0.1056	0.1111	0.1107	0.1165
0.90	(50,50)	0.8780	0.8713	0.8794	0.8937	0.8790	0.9350
		0.1070	0.1069	0.1007	0.1101	0.1058	0.1252
	(80,80)	0.8941	0.8801	0.8849	0.9006	0.8740	0.9230
		0.0849	0.0846	0.0801	0.0860	0.0841	0.0930
	(100,100)	0.8987	0.8853	0.8917	0.9006	0.8710	0.9100
		0.0759	0.0757	0.0716	0.0767	0.0752	0.0817
	(50,80)	0.8991	0.8869	0.8891	0.8976	0.8920	0.9230
		0.0886	0.0882	0.0836	0.0902	0.0883	0.0973
	(70,100)	0.8968	0.8857	0.8971	0.8970	0.8750	0.9140
		0.0783	0.0781	0.0740	0.0794	0.0782	0.0847
0.95	(50,50)	0.8528	0.8407	0.8689	0.8804	0.8500	0.9080
		0.0723	0.0711	0.0659	0.0765	0.0706	0.1013
	(80,80)	0.8762	0.8595	0.8826	0.8924	0.8660	0.9250
		0.0582	0.0571	0.0523	0.0600	0.0569	0.0704
	(100,100)	0.8852	0.8711	0.8863	0.8928	0.8630	0.9240
		0.0522	0.0513	0.0469	0.0531	0.0512	0.0603
	(50,80)	0.8845	0.8696	0.8803	0.8946	0.8730	0.9250
		0.0604	0.0593	0.0546	0.0626	0.0595	0.0727
	(70,100)	0.8892	0.8752	0.8870	0.8956	0.8820	0.9270
		0.0537	0.0527	0.0483	0.0546	0.0522	0.0612

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Table 2. Normal distribution: coverage probabilities and average lengths of 95% confidence intervals for AUC. The average lengths are the numbers in the rows below the corresponding coverage probabilities.

AUC	(m,n)	EL	MW	ML	LT	PB	PTB
0.80	(50,50)	0.9407	0.9379	0.9351	0.9538	0.9300	0.9690
		0.1783	0.1808	0.1712	0.1808	0.1746	0.1971
	(80,80)	0.9431	0.9418	0.9427	0.9530	0.9450	0.9690
		0.1419	0.1431	0.1357	0.1430	0.1379	0.1508
	(100,100)	0.9489	0.9454	0.9394	0.9514	0.9380	0.9590
		0.1269	0.1278	0.1215	0.1279	0.1246	0.1346
	(50,80)	0.9468	0.9412	0.9393	0.9513	0.9310	0.9610
		0.1487	0.1501	0.1428	0.1501	0.1458	0.1594
	(70,100)	0.9472	0.9440	0.9459	0.9513	0.9290	0.9570
		0.1310	0.1320	0.1260	0.1326	0.1287	0.1391
0.90	(50,50)	0.9352	0.9204	0.9257	0.9468	0.9150	0.9700
		0.1281	0.1271	0.1199	0.1326	0.1228	0.1591
	(80,80)	0.9411	0.9264	0.9347	0.9471	0.9310	0.9670
		0.1015	0.1008	0.0955	0.1035	0.0976	0.1136
	(100,100)	0.9468	0.9356	0.9356	0.9486	0.9220	0.9540
		0.0909	0.0904	0.0853	0.0922	0.0877	0.0997
	(50,80)	0.9458	0.9281	0.9349	0.9547	0.9340	0.9630
		0.1061	0.1053	0.0995	0.1084	0.1025	0.1188
	(70,100)	0.9434	0.9297	0.9384	0.9498	0.9290	0.9580
		0.0936	0.0931	0.0882	0.0951	0.0908	0.1025
0.95	(50,50)	0.8964	0.8818	0.9166	0.9289	0.8840	0.9490
		0.0874	0.0850	0.0788	0.0930	0.0814	0.1360
	(80,80)	0.9252	0.9064	0.9264	0.9414	0.9180	0.9610
		0.0700	0.0680	0.0624	0.0727	0.0658	0.0905
	(100,100)	0.9340	0.9142	0.9298	0.9366	0.9180	0.9660
		0.0628	0.0612	0.0559	0.0643	0.0594	0.0758
	(50,80)	0.9269	0.9060	0.9258	0.9462	0.9150	0.9720
		0.0728	0.0707	0.0649	0.0756	0.0690	0.0936
	(70,100)	0.9351	0.9138	0.9314	0.9478	0.9230	0.9660
		0.0643	0.0625	0.0575	0.0663	0.0605	0.0773

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

Table 3. Log-normal distribution: coverage probabilities and average lengths of 90% confidence intervals for AUC. The average lengths are the numbers in the rows below the corresponding coverage probabilities.

AUC	(m,n)	EL	MW	ML	LT	PB	PTB
0.80	(50,50)	0.8941	0.8789	0.8800	0.9053	0.8840	0.917
		0.1771	0.1805	0.1492	0.1812	0.1518	0.1642
	(80,80)	0.8969	0.8931	0.8873	0.9061	0.8970	0.9250
		0.1413	0.1430	0.1186	0.1433	0.1191	0.1262
	(100,100)	0.8996	0.8955	0.8909	0.9046	0.8810	0.9080
		0.1268	0.1280	0.1064	0.1281	0.1071	0.1128
	(50,80)	0.8962	0.8926	0.8924	0.9097	0.8930	0.9220
		0.1415	0.1433	0.1188	0.1434	0.1247	0.1322
	(70,100)	0.8990	0.8961	0.8921	0.9016	0.9180	0.9330
		0.1269	0.1281	0.1064	0.1283	0.1110	0.1168
0.90	(50,50)	0.8776	0.8632	0.8806	NA	0.8720	0.9360
		0.1315	0.1329	0.1078	NA	0.1057	0.1242
	(80,80)	0.8915	0.8834	0.8839	0.9052	0.8960	0.9230
		0.1056	0.1063	0.0859	0.1092	0.0853	0.0944
	(100,100)	0.8958	0.8872	0.8924	0.9165	0.8890	0.9190
		0.0946	0.0951	0.0771	0.0973	0.0765	0.0829
	(50,80)	0.8864	0.8800	0.8902	0.9129	0.8800	0.9280
		0.1054	0.1061	0.0861	0.1093	0.0872	0.0961
	(70,100)	0.8971	0.8838	0.8905	0.9077	0.8780	0.9080
		0.0944	0.0950	0.0771	0.0973	0.0781	0.0847
0.95	(50,50)	0.8699	0.8629	0.8661	NA	0.8570	0.9230
		0.0928	0.0921	0.0719	NA	0.0712	0.1022
	(80,80)	0.8726	0.8697	0.8792	NA	0.8630	0.9130
		0.0754	0.0751	0.0575	NA	0.0573	0.0703
	(100,100)	0.8936	0.8695	0.8813	NA	0.8680	0.9240
		0.0680	0.0678	0.0515	NA	0.0509	0.0601
	(50,80)	0.8678	0.8676	0.8817	NA	0.8610	0.9280
		0.0751	0.0748	0.0577	NA	0.0588	0.0722
	(70,100)	0.8980	0.8609	0.8857	NA	0.8670	0.9150
		0.0681	0.0678	0.0517	NA	0.0527	0.0620

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Table 4. Log-normal distribution: coverage probabilities and average lengths of 95% confidence intervals for AUC. The average lengths are the numbers in the rows below the corresponding coverage probabilities.

AUC	(m,n)	EL	MW	ML	LT	PB	PTB
0.80	(50,50)	0.9354	0.9312	0.9313	0.9564	0.9410	0.9650
		0.2102	0.2151	0.1778	0.2162	0.1764	0.1994
	(80,80)	0.9437	0.9423	0.9442	0.9537	0.9380	0.9660
		0.1680	0.1704	0.1415	0.1709	0.1381	0.1508
	(100,100)	0.9442	0.9392	0.9417	0.9558	0.9380	0.9630
		0.1507	0.1524	0.1266	0.1529	0.1245	0.1347
	(50,80)	0.9421	0.9361	0.9430	0.9545	0.9380	0.9610
		0.1678	0.1703	0.1416	0.1711	0.1449	0.1588
	(70,100)	0.9516	0.9456	0.9413	0.9561	0.9530	0.9680
		0.1508	0.1525	0.1266	0.1530	0.1291	0.1392
0.90	(50,50)	0.9344	0.8975	0.9250	NA	0.9130	0.9690
		0.1562	0.1578	0.1286	NA	0.1226	0.1575
	(80,80)	0.9427	0.9253	0.9306	0.9561	0.9370	0.9650
		0.1254	0.1263	0.1023	0.1317	0.0991	0.1151
	(100,100)	0.9456	0.9346	0.9362	0.9632	0.9340	0.9600
		0.1129	0.1135	0.0918	0.1171	0.0888	0.1007
	(50,80)	0.9415	0.9229	0.9347	0.9581	0.9270	0.9730
		0.1254	0.1263	0.1026	0.1318	0.1013	0.1177
	(70,100)	0.9431	0.9303	0.9355	0.9558	0.9260	0.9510
		0.1128	0.1134	0.0918	0.1171	0.0905	0.1024
0.95	(50,50)	0.8975	0.8941	0.9101	NA	0.9010	0.9520
		0.1109	0.1094	0.0864	NA	0.0824	0.1387
	(80,80)	0.9344	0.8945	0.9215	NA	0.9090	0.9510
		0.0905	0.0897	0.0685	NA	0.0661	0.0905
	(100,100)	0.9490	0.8989	0.9312	NA	0.9060	0.9610
		0.0813	0.0806	0.0614	NA	0.0589	0.0756
	(50,80)	0.9504	0.8969	0.9205	NA	0.9030	0.9650
		0.0902	0.0894	0.0687	NA	0.0681	0.0929
	(70,100)	0.9474	0.8946	0.9279	NA	0.9080	0.9660
		0.0810	0.0804	0.0613	NA	0.0611	0.0780

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268

Table 5. Exponential distribution: coverage probabilities and average lengths of 90% confidence intervals for AUC. The average lengths are the numbers in the rows below the corresponding coverage probabilities.

AUC	(m,n)	EL	MW	ML	LT	PB	PTB
0.80	(50,50)	0.8927	0.8905	0.8947	0.9039	0.8840	0.9040
		0.1444	0.1462	0.1054	0.1466	0.1454	0.1561
	(80,80)	0.8940	0.8915	0.8984	0.9008	0.8860	0.9150
		0.1143	0.1152	0.0832	0.1155	0.1146	0.1208
	(100,100)	0.8957	0.8940	0.8935	0.9049	0.8900	0.9110
		0.1025	0.1032	0.0743	0.1032	0.1029	0.1081
	(50,80)	0.8961	0.8911	0.8951	0.9038	0.8950	0.9200
		0.1238	0.1249	0.0946	0.1250	0.1240	0.1310
	(70,100)	0.8965	0.8942	0.8924	0.9060	0.9100	0.9240
		0.1085	0.1092	0.0818	0.1094	0.1092	0.1146
0.90	(50,50)	0.8886	0.8791	0.8921	0.9051	0.8800	0.9310
		0.1048	0.1047	0.0597	0.1076	0.1035	0.1200
	(80,80)	0.8941	0.8849	0.8987	0.9011	0.8820	0.9110
		0.0830	0.0828	0.0471	0.0841	0.0827	0.0908
	(100,100)	0.8958	0.8872	0.9030	0.8939	0.8870	0.9220
		0.0741	0.0740	0.0421	0.0751	0.0739	0.0800
	(50,80)	0.8953	0.8889	0.9032	0.8986	0.8800	0.9030
		0.0873	0.0871	0.0535	0.0886	0.0873	0.0954
	(70,100)	0.8981	0.8875	0.8968	0.9059	0.8950	0.9280
		0.0769	0.0767	0.0464	0.0777	0.0765	0.0826
0.95	(50,50)	0.8613	0.8479	0.9010	NA	0.8660	0.9150
		0.0729	0.0718	0.0315	NA	0.0726	0.1043
	(80,80)	0.8793	0.8627	0.8945	0.8898	0.8780	0.9240
		0.0585	0.0576	0.0249	0.0603	0.0577	0.0714
	(100,100)	0.8902	0.8746	0.8994	0.8958	0.8680	0.9290
		0.0525	0.0518	0.0222	0.0537	0.0507	0.0602
	(50,80)	0.8801	0.8617	0.8981	0.8994	0.8780	0.9410
		0.0603	0.0593	0.0284	0.0625	0.0591	0.0725
	(70,100)	0.8867	0.8692	0.8912	0.8957	0.8890	0.9300
		0.0537	0.0529	0.0245	0.0552	0.0523	0.0611

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Table 6. Exponential distribution: coverage probabilities and average lengths of 95% confidence intervals for AUC. The average lengths are the numbers in the rows below the corresponding coverage probabilities.

AUC	(m,n)	EL	MW	ML	LT	PB	PTB
0.80	(50,50)	0.9446	0.9394	0.9433	0.9551	0.9270	0.9570
		0.1725	0.1746	0.1256	0.1748	0.1692	0.1887
	(80,80)	0.9459	0.9432	0.9422	0.9547	0.9280	0.9540
		0.1366	0.1376	0.0992	0.1378	0.1336	0.1450
	(100,100)	0.9507	0.9456	0.9500	0.9478	0.9380	0.9540
		0.1223	0.1230	0.0886	0.1232	0.1196	0.1286
	(50,80)	0.9469	0.9444	0.9442	0.9550	0.9370	0.9620
		0.1478	0.1489	0.1128	0.1489	0.1439	0.1568
	(70,100)	0.9477	0.9410	0.9481	0.9541	0.9450	0.9660
		0.1292	0.1301	0.0978	0.1303	0.1269	0.1370
0.90	(50,50)	0.9321	0.9200	0.9428	0.9482	0.9240	0.9740
		0.1254	0.1247	0.0710	0.1290	0.1198	0.1501
	(80,80)	0.9423	0.9291	0.9432	0.9514	0.9240	0.9600
		0.0990	0.0984	0.0560	0.1010	0.0961	0.1106
	(100,100)	0.9499	0.9380	0.9493	0.9485	0.9340	0.9610
		0.0887	0.0882	0.0502	0.0898	0.0857	0.0964
	(50,80)	0.9467	0.9312	0.9454	0.9535	0.9270	0.9610
		0.1044	0.1036	0.0639	0.1064	0.1012	0.1163
	(70,100)	0.9467	0.9364	0.9459	0.9520	0.9310	0.9670
		0.0919	0.0913	0.0552	0.0933	0.0886	0.0996
0.95	(50,50)	0.8977	0.8817	0.9421	NA	0.9000	0.9460
		0.0881	0.0859	0.0377	NA	0.0838	0.1427
	(80,80)	0.9296	0.9090	0.9471	0.9398	0.9160	0.9660
		0.0705	0.0688	0.0297	0.0732	0.0670	0.0930
	(100,100)	0.9412	0.9174	0.9441	0.9473	0.9110	0.9660
		0.0630	0.0616	0.0265	0.0650	0.0589	0.0758
	(50,80)	0.9329	0.9080	0.9424	0.9446	0.9220	0.9750
		0.0726	0.0707	0.0338	0.0759	0.0683	0.0935
	(70,100)	0.9356	0.9117	0.9437	0.9435	0.9340	0.9710
		0.0643	0.0629	0.0292	0.0668	0.0607	0.0768

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/uwbiostat/paper268