Georgia State University
Digital Archive @ GSU

Mathematics Theses | Department of Mathematics and Statistics

11-28-2007

Empirical Likelihood Based Confidence Intervals for the Difference between Two Sensitivities of Continuous-scale Diagnostic Tests at a Fixed Level of Specificity

Suqin Yao

Follow this and additional works at: http://digitalarchive.gsu.edu/math_theses

Recommended Citation
Yao, Suqin, "Empirical Likelihood Based Confidence Intervals for the Difference between Two Sensitivities of Continuous-scale Diagnostic Tests at a Fixed Level of Specificity" (2007). *Mathematics Theses.* Paper 36.

This Thesis is brought to you for free and open access by the Department of Mathematics and Statistics at Digital Archive @ GSU. It has been accepted for inclusion in Mathematics Theses by an authorized administrator of Digital Archive @ GSU. For more information, please contact digitalarchive@gsu.edu.

Empirical Likelihood Based Confidence Intervals for the Difference between Two Sensitivities of Continuous-scale Diagnostic Tests at a Fixed Level of Specificity

by

SUQIN YAO

Under the Direction of Gengsheng Qin

ABSTRACT

Diagnostic testing is essential to distinguish non-diseased individuals from diseased individuals. The sensitivity and specificity are two important indices for the diagnostic accuracy of continuous-scale diagnostic tests. If we want to compare the effectiveness of two tests, it is of interest to construct a confidence interval for the difference of the two sensitivities at a fixed level of specificity. In this thesis, we propose two empirical likelihood based confidence intervals (HBELI and HBELII) for the difference of two sensitivities at a predetermined specificity level. Simulation studies show that when correlation between the two test results exists, HBELI and HBELII intervals perform better than the existing bootstrap based BCa, BTI and BTII intervals due to shorter interval lengths. However, when there is no correlation, BCa, BTI and BTII intervals outperform HBELI and HBELII intervals due to better coverage probability in most simulation settings.

INDEX WORDS: Empirical likelihood, diagnostic test, sensitivity, specificity

Empirical Likelihood Based Confidence Intervals for the Difference between Two Sensitivities of Continuous-scale Diagnostic Test at a Fixed Level of Specificity

by

SUQIN YAO

A Thesis Submitted in Partial Fulfillment of the requirements for the Degree of

Master of Science
in the College of Arts and Sciences
Georgia State University

2007

Copyright by
Suqin Yao
2007

Empirical Likelihood Based Confidence Intervals for the Difference between Two Sensitivities of Continuous-scale Diagnostic Test at a Fixed Level of Specificity

by

SUQIN YAO

Major Professor: Dr. Gengsheng (Jeff) Qin
Committee: Dr. Xu Zhang
Dr. Yixin Fang

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Science
Georgia State University
December 2007

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Gengsheng Qin, for his sharp knowledge, great leadership and patience throughout this process. Words can not express how truly grateful I am for all of your guidance. I would also like to thank the committee members for accepting my invitation and your assistance. I thank them very much for their comments and suggestions on the improvement of my thesis.

I would like to thank Brains & Behavior Program and acknowledge the funding for my graduate study and research in Georgia State University.

Finally, I thank my husband, my parents and all other family members, for your continuous support and endless encouragement throughout my graduate study and thank my little son for his special love to mommy.

Table of Contents

List of Tables vi

Chapter 1 Introduction 1

Chapter 2 Existing methods 5

2.1 Normal-approximation-based interval 5

2.2 Bootstrap based intervals 7

2.2.1 Paired uncorrelated samples 8

2.2.2 Paired dependent samples 10

2.2.3 New bootstrap intervals for $D(p_0)$ 10

Chapter 3 Hybrid empirical likelihood based intervals for the difference between two sensitivities 12

Chapter 4 Simulation 17

Chapter 5 Dermatoscope example 20

Chapter 6 Discussion 22

References 23

Appendix I: Simulation tables 25

Appendix II S-plus code for Simulation 34

1. Normal distribution 34

2. Exponential distribution – no correlation 40

3. Exponential distribution – correlation 45

4. Dermatoscope example 47

List of Tables

Table a- 1 Level of 95 per cent confidence interval for D(p)=0. Bivariate normal distribution with $\rho = 0$ 25

Table a- 2 Level of 95 per cent confidence interval for D(p)=0. Bivariate normal distribution with $\rho = 0.5$ 27

Table a- 3 Level of 95 per cent confidence interval for D(p)=0. Bivariate exponential distribution with $\rho = 0$ 29

Table a- 4 Level of 95 per cent confidence interval for D(p)=0. Bivariate exponential distribution with $\rho \neq 0$ (Using 0.02 to generate diseased random sample) 31

Table a- 5 95 per cent confidence interval for the difference of sensitivities between the two clinical assessments with and without the use of dermatoscopy 33

Chapter 1 Introduction

Diagnostic tests play a key role in modern medicine by screening a specific population for evidence of disease. The interpretation to the result of a diagnostic test depends on the discriminatory accuracy of the test to distinguish diseased patients from non-diseased subjects (Shapiro, 1999). Sensitivity and specificity are two measurements to describe the discriminatory accuracy of a test, which are defined as the probability of the test correctly identifying the diseased and non-diseased subjects respectively.

A diagnostic test is named continuous, dichotomous, or ordinal test depending on whether the test generate a continuous result (e.g. blood pressure), a dichotomous outcome (e.g. positive or negative), or an ordinal conclusion (e.g. confidence rating for presence of disease-definitely, probably, possibly, probably not, definitely not) (Shapiro, 1999). The main focus in this thesis is on continuous-scale diagnostic tests.

In continuous-scale diagnostic tests , it is common to define a threshold or a cut-off point γ and classify the subject as diseased if the test result Y is greater than or equal to γ and non-diseased if the test result X is less than γ. Thus, sensitivity and specificity are defined for each cut-off point γ as:

$$\begin{aligned} R &= P(Y \geq \gamma) = 1 - G(\gamma), \\ Sp &= P(X < \gamma) = F(\gamma), \end{aligned} \tag{1-1}$$

respectively, where G and F are the distribution functions of Y and X respectively. Let $X_1, X_2, ..., X_m$ be the test results of a random sample of non-diseased subjects, and $Y_1, Y_2, ..., Y_n$ be the test results of a random sample of diseased subjects. As we can see, when γ decreases,

sensitivity increases but specificity decreases; as γ increases, specificity increases at the expense of sensitivity. Therefore, there is a compromise between sensitivity and specificity when cut-off point changes, which is accounted for assessing discriminatory accuracy. From equation (1-1), the relationship between sensitivity and specificity can be set up without knowing the exact value of cut-off point γ. Let the specificity of a test be p ($0 \le p \le 1$), the corresponding sensitivity of the test is

$$R(p) = 1 - G(F^{-1}(p)) \ , \qquad (1\text{-}2)$$

where F^{-1} is the inverse function of F.

Using equation (1-2), we can estimate the sensitivity of a test at a fixed level of specificity based on test results from the diseased and non-diseased subjects. It is also of interest to construct confidence intervals for the sensitivity $R(p)$. However, if we have two (or more) continuous-scale diagnostic tests to the same set of subjects, some of whom are non-diseases, some diseased, we may be more interested in knowing which test is better, especially when only a particular value of specificity is relevant (e.g. 70%, 80%, 90%). There are studies in literature for comparing the accuracy of two or more diagnostic tests, including comparing ROC curves and comparing summary accuracy indices (such as AUC, partial AUC, sensitivity and specificity). Some studies used 'unpaired' design, in which each diagnostic test is applied to a different group of subjects. The other studies utilized paired design, in which the diagnostic tests are applied to the same subjects (Shapiro, 1999). We focus on the comparison of sensitivities of two tests at a common specificity in this thesis.

Greenhouse and Mantel (1950) provided normal-theory that a diagnostic test has at least a specified sensitivity (e.g. ≥ 0.9) with specificity higher than a specified value (e.g. ≥ 0.95). Based on the result of Greenhouse and Mantel, Linnet (1987) proposed both parametric and non-

parametric methods for constructing confidence intervals for the sensitivity of a test at a fixed value of specificity, accounting for the random variation associated with the estimated cut-off point. Wieand et al. (1989) studied asymptotic behaviors of these non-parametric procedures and generalized them to a comparison of two weighted average of sensitivities. Their theory can be used to construct a normal approximation based confidence interval (WGJ interval) for the difference between two sensitivities. Qin et al. (2006) proposed three new bootstrap based intervals (BCa, BTI, BTII) that have better coverage accuracy than the WGJ interval.

Empirical likelihood (EL) (Owen, 1990, 2001) is a popular non-parametric method traditionally used for providing confidence intervals for means. 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; empirical likelihood based confidence regions are range preserving and transformation respecting; the regularity conditions for empirical likelihood based methods are weak and natural. However, the empirical likelihood method has not been used widely in the study of accuracy of diagnostic tests. Qin (2007) proposed empirical likelihood based confidence intervals for the sensitivity of a single test at a fixed level of specificity. In this thesis, we are going to expand Qin's finding (2007) in one continuous-scale test to construct EL-based confidence intervals for the difference between the sensitivities of two continuous-scale tests at a fixed level of specificity.

The thesis is organized as follow. In Chapter 2, we review some existing methods for the interval estimation of the difference between two sensitivities. In Chapter 3, we propose new hybrid empirical likelihood and bootstrap confidence intervals for the difference between two sensitivities at a pre-determined specificity, by using the asymptotic scaled chi-square distribution of the empirical likelihood ratio statistic. In Chapter 4, simulation studies are

conducted to compare the relative performance of the proposed empirical likelihood based intervals with the existing bootstrap intervals (BCa, BTI, and BTII). In Chapter 5, the new empirical likelihood based confidence intervals for the difference between two sensitivities are applied to a real example. A discussion is given in Chapter 6, and simulation tables and S-plus code are provided in the Appendix I and II.

Chapter 2 Existing methods

For two continuous-scale diagnostic tests, it is of interest to compare their sensitivities at a predetermined level of specificity. In this chapter, we give a review of the existing normal-approximation based interval proposed by Wieand (1989) and three bootstrap based intervals recently proposed by Qin et al. (2006) for the difference between two sensitivities at a fixed level of specificity.

2.1 Normal-approximation-based interval

Greenhouse and Mantel (1950) and Linnet (1987) proposed non-parametric procedures for the comparison of two sensitivities at a fixed level of specificity. Wieand et al. (1989) studied asymptotic behaviors of these non-parametric procedures and generalized them to a comparison of two weighted average of sensitivities.

Let T_1 and T_2 be two diagnostic tests that yield continuous measurements. It is assumed that both tests are performed on the same m controls (non-diseased) and n cases (diseased). $(X_{1i}, X_{2i}), i = 1,2,...,m$ are i.i.d. bivariate outcomes from the population with a joint distribution $F(x_1, x_2)$ that represents the non-diseased group, $(Y_{1j}, Y_{2j}), j = 1,2,...,n$ are i.i.d. bivariate outcomes from the population with a joint distribution $G(y_1, y_2)$ that represents the diseased group. The marginal distribution functions of X_k and Y_k are denoted by $F_k(x_i)$ and $G_k(y_j)$ respectively, $k = 1,2$. For a given cut-off point γ, the sensitivity and specificity of the test $T_k, k = 1,2$ are defined by

$$R_k = P(Y_k \geq \gamma) = 1 - G_k(\gamma), Sp_k = P(X_k < \gamma) = F_k(\gamma) \ , \qquad (2\text{-}1)$$

respectively. Thus, the sensitivity of test T_k at a fixed value of specificity p, is

$$R_k(p) = 1 - G_k(F_k^{-1}(p)),$$

where $F_k^{-1}(p) = \inf\{t : F_k(t) \geq p\}, k = 1,2$. The parameter of interest is the difference between two sensitivities at the same fixed value of specificity p_0,

$$D(p_0) = R_1(p_0) - R_2(p_0) \ . \tag{2-2}$$

Let $\hat{G}_k$ be the empirical distribution of G_k, based on the sample $X_{k1},...,X_{km}$, and $\hat{F}_k^{-1}(p)$ be the empirical estimate for the p-th quantile of $F_k, k = 1,2$, based on the sample $Y_{k1},...,Y_{kn}$. The non-parametric estimator for $D(p_0)$ proposed by Linnet (1987) and Wieand et al. (1989) is given as follows:

$$\hat{D}(p_0) = \hat{R}_1(p_0) - \hat{R}_2(p_0) \ , \tag{2-3}$$

where $\hat{R}_k(p_0) = 1 - \hat{G}_k(\hat{F}_k^{-1}(p_0))$.

Let N=m+n. Wieand et al. (1989) showed that

$$N^{1/2}(\hat{D}(p_0) - D(p_0)) \xrightarrow{d} N(0, \sigma^2), \tag{2-4}$$

where

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_{12},$$

$$\sigma_k^2 = (1-\lambda)^{-1} R_k(p_0)(1 - R_k(p_0)) + \lambda^{-1}(1-p_0)p_0 \frac{g_k^2(F_i^{-1}(p_0))}{f_k^2(F_i^{-1}(p_0))} \quad (k = 1,2),$$

$$\sigma_{12} = (1-\lambda)^{-1}\{G(F_1^{-1}(p_0), F_2^{-1}(p_0)) - G_1(F_1^{-1}(p_0))G_2(F_2^{-1}(p_0))] +$$

$$\lambda^{-1}[F(F_1^{-1}(p_0), F_2^{-1}(p_0)) - p_0^2] \frac{g_1(F_1^{-1}(p_0))g_2(F_2^{-1}(p_0))}{f_1(F_1^{-1}(p_0))f_2(F_2^{-1}(p_0))},$$

$$\lambda = m/(m+n),$$

where f_k and g_k are the density functions of F_k and G_k respectively.

If a good estimate for σ^2 is available, the normal approximation equation (2-4) can be used to construct a confidence interval for the difference between two sensitivities at the same fixed level of specificity. However, the estimation of σ^2 requires the estimation of density functions f_k and g_k, the estimation of bivariate distribution functions $F(x_1, x_2)$ and $G(y_1, y_2)$, and the estimation of quantiles $F_k^{-1}(p)$. Therefore, the performance of the normal approximation based interval is very sensitive to the choice of the smoothing parameters in density and distribution estimations. Selection of satisfactory smoothing parameters in this context is problematic.

2.2 Bootstrap based intervals

Qin et al. (2006) proposed three intervals called BCa, BTI and BTII intervals for the difference between sensitivities of two diagnostic tests at a fixed value of specificity by using bootstrap method. The major advantage of these intervals over the normal approximation based interval is that no density and distribution estimation is needed. And the new intervals are computationally easy to implement in practice.

The difference between two sensitivities at the same fixed value of specificity p_0 is the difference between two proportions:

$$D(p_0) = R_1(p_0) - R_2(p_0) = P(Y_{1K} \geq F_1^{-1}(p_0)) - P(Y_{2K} \geq F_2^{-1}(p_0)).$$

If F_k were known, an obvious estimator of $D(p_0)$ would be the difference between the observed sensitivities at p_0-th quantiles $F_1^{-1}(p_0)$ and $F_2^{-1}(p_0)$, which would be defined as

$$\tilde{D}(p_0) = \frac{1}{n}\sum_{j=1}^{n} I_{[Y_{1j} \geq F_1^{-1}(p_0)]} - \frac{1}{n}\sum_{j=1}^{n} I_{[Y_{2j} \geq F_2^{-1}(p_0)]}, \qquad (2\text{-}5)$$

where I_A is the indicator function of *A*. We can also regard $\tilde{D}(p_0)$ as the difference between two sample proportions of binomial distributions with proportions $R_k(p_0)$, $k = 1,2$. However, F_k's are unknown, by replacing $F_k^{-1}(p_0)$ by $\hat{F}_k^{-1}(p_0)$ in equation (2-5), we acquire an estimator $\hat{D}(p_0)$ for $D(p_0)$.

$$\hat{D}(p_0) = \frac{1}{n}\sum_{j=1}^{n} I_{[Y_{1j} \geq \hat{F}_1^{-1}(p_0)]} - \frac{1}{n}\sum_{j=1}^{n} I_{[Y_{2j} \geq \hat{F}_2^{-1}(p_0)]} \qquad (2\text{-}6)$$

Because the indicator variables $I_{[Y_{i1} \geq \hat{F}_1^{-1}(p_0)]}, I_{[Y_{i2} \geq \hat{F}_1^{-1}(p_0)]}, \ldots, I_{[Y_{in} \geq \hat{F}_1^{-1}(p_0)]}$ are not independent, $\hat{D}(p_0)$ is no longer the difference between two simple binomial proportions. Depending on whether there is a correlation between the test results from two diagnostic tests, Qin et al. (2006) proposed the following different procedures for the confidence intervals of $D(p_0)$ by combining bootstrap method with the technique provided by Agresti and Caffo (2000).

2.2.1 Paired uncorrelated samples

If the test results from two diagnostic tests are conditionally uncorrelated, $\hat{D}(p_0)$ can be considered as the difference between two independent sample proportions. Qin et al. (2006) proposed the following estimator for $D(p_0)$ instead of $\hat{D}(p_0)$:

$$\hat{D}(p_0) = \hat{R}_1(p_0) - \hat{R}_2(p_0), \qquad (2\text{-}7)$$

where

$$\hat{R}_k(p_0) = \frac{\sum_{j=1}^{n} I_{[Y_{kj} \geq \hat{F}_i^{-1}(p_0)]} + Z_{1-\alpha/2}^2 / 2}{n + Z_{1-\alpha/2}^2}, k = 1,2 \qquad (2\text{-}8)$$

$Z^2_{1-\alpha/2}$ is the $(1-\alpha/2)$-th quantile of standard normal distribution. The procedure for computing the bootstrap variance is as follows:

1. For each $k=1,2$, draw a resample of size n, $Y^*_{kj}(j=1,...,n)$ with replacement from the diseased patient sample $Y_{kj}(j=1,...,n)$, and a separate resample of size m, $X^*_{ki}(i=1,...,m)$ with replacement from the non-diseased patient sample $X_{ki}(i=1,...,m)$.
2. Calculate the bootstrap versions of $\hat{R}_k(p_0)$ ($k=1,2$) and $\hat{D}(p_0)$,

$$\hat{R}^*_k(p_0)=\frac{\sum_{j=1}^{n} I_{[Y^*_{kj}\geq \hat{F}^{*-1}_k(p_0)]}+Z^2_{1-\alpha/2}/2}{n+Z^2_{1-\alpha/2}}, k=1,2$$

$$\hat{D}^*(p_0)=\hat{R}^*_1(p_0)-\hat{R}^*_2(p_0),$$

where $\hat{F}^{*-1}_k(p_0)$ is the p_0-th sample quantile based on the bootstrap resample $X^*_{ki}s$.

3. Repeat the first two steps B times to obtain the set of bootstrap replications:

$\{\hat{R}^*_{kb}(p_0):b=1,2,...,B\}$, and $\{\hat{D}^*_b(p_0):b=1,2,...,B\}$, $k=1,2$.

Then, the bootstrap estimate V^* for the variance of $\hat{D}(p_0)$ is defined as follows:

$$V^*=V^*_1+V^*_2,$$

where

$$V^*=\frac{1}{B-1}\sum_{b=1}^{B}(\hat{R}^*_{kb}(p_0)-\bar{R}^*_k(p_0))^2, k=1,2$$

$$\bar{R}^*_k(p_0)=\frac{1}{B}\sum_{b=1}^{B}\hat{R}^*_{kb}(p_0), k=1,2$$

The above procedure can also be used to the case of two independent samples with different sample size.

2.2.2 Paired dependent samples

When two diagnostic tests are applied to the same patients, the test results from two diagnostic tests are most likely correlated. Qin et al. (2006) proposed to use the following estimates for the sensitivities:

$$\hat{R}_k(p_0) = \frac{\sum_{j=1}^{n} I_{[Y_{kj} \geq \hat{F}_k^{-1}(p_0)]} + 1}{n+2}, k = 1,2.$$

The bootstrap estimate V^* for the variance of $\hat{D}(p_0)$ is defined as follows:

$$V^* = V_1^* + V_2^* - 2V_{12}^*,$$

where V_k^* ($k = 1,2$) are defined as before, and

$$V_{12}^* = \frac{1}{B-1}\sum_{b=1}^{B}(\hat{R}_{1b}^*(p_0) - \bar{R}_1^*(p_0))(\hat{R}_{2b}^*(p_0) - \bar{R}_2^*(p_0))$$

2.2.3 New bootstrap intervals for $D(p_0)$

Qin et al. (2006) proposed three new intervals for $D(p_0)$. The first two $(1-\alpha)100$ per cent confidence intervals for $D(p_0)$ are bootstrap intervals based on the bootstrap variance estimate V^*. They are defined as follows:

(i) The first one, called BTI interval, is

$$(\hat{D}(p_0) - z_{1-\alpha/2}\sqrt{V^*}, \hat{D}(p_0) + z_{1-\alpha/2}\sqrt{V^*})$$

where $\hat{D}(p_0)$ is defined by equation (2-7)

(ii) The second one, called BTII interval, is

$$(\bar{D}^*(p_0) - z_{1-\alpha/2}\sqrt{V^*}, \bar{D}^*(p_0) + z_{1-\alpha/2}\sqrt{V^*})$$

Where $\bar{D}^*(p_0) = \frac{1}{B}\sum_{b=1}^{B} \hat{D}_b^*(p_0)$

The above two intervals require variance estimation of $\hat{D}(p_0)$. The third interval for $D(p_0)$ proposed by Qin et al. (2006) is a BCa-type bootstrap interval in which the direct variance estimation is not needed:

$$(\hat{D}^*_{(B\hat{\alpha}/2)}(p_0), \hat{D}^*_{(B(1-\hat{\alpha}/2))}(p_0)),$$

where

$$\hat{\alpha} = \Phi(w + \frac{w + z_\alpha}{1 - \alpha(w + z_\alpha)})$$

$$w = \Phi^{-1}(\frac{1}{B}\sum_{b=1}^{B} I_{[\hat{D}_b^*(p_0) \le \hat{D}(p_0)]})$$

$$\alpha = \frac{1}{6}\frac{\sum_{k=1}^{n} l_k^3}{(\sum_{k=1}^{n} l_k^2)^{3/2}}$$

$$l_k = (I_{[Y_{1K} \ge \hat{F}_1^{-1}(p_0)]} - I_{[Y_{2K} \ge \hat{F}_2^{-1}(p_0)]}) - (\hat{R}_1(p_0) - \hat{R}_2(p_0)),$$

and Φ is the standard normal distribution function, and $\hat{D}^*_{(b)}(p_0)$ is the *b*-th ordered value among $\{\hat{D}^*_b(p_0), b = 1, 2, ..., B\}$.

Through simulation study, Qin et al. (2006) showed that BTI and BTII intervals perform better than the normal approximation based interval for independent samples, and BCa interval performs better than the normal approximation based interval for paired dependent samples. In addition, BTI and BTII intervals are computationally simpler than the normal approximation based interval. Therefore, we only use BCa, BTI and BTII intervals as a comparison in this thesis.

Chapter 3 Hybrid empirical likelihood based intervals for the difference between two sensitivities

We recently developed an empirical likelihood based method to construct the confidence interval for the difference between two sensitivities from two diagnostic tests at a fixed level of specificity. An introduction of this method is given in this chapter.

Pepe (2003) defined a placement value for a given test value *Y* from a diseased subject as

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

This value is the proportion of the non-diseased population with a test value greater than *Y*. It marks the placement of *Y* within non-diseased distribution.

It is evident that

$$E(I(U \le 1-p)) = P(F(Y) \ge p) = P(Y \ge F^{-1}(p)) = R(p).$$

For two diagnostic tests T_1 and T_2 that yield continuous measurements, we have

$$U_k = 1 - F_k(y_k), k = 1,2;$$

$$E[I(U_k \le 1-p)] = P(Y_k \ge F_k^{-1}(p)) = 1 - G_k(F_k^{-1}(p)) = R_k(p).$$

Therefore,

$$D(p) = R_1(p) - R_2(p) = E[I(U_1 \le 1-p)] - E[I(U_2 \le 1-p)]$$

Based on this relationship between $D(p)$ and the placement value U_k's, an empirical likelihood procedure is derived for the difference between two sensitivities. Let

$P_k = (p_{k1}, p_{k2}, ..., p_{kn}), k = 1,2$ be two probability vectors, i.e,. $\sum_{j=1}^{n} p_{kj} = 1$ and $p_{kj} \geq 0$ for all *j*. The profile EL for *D(p)* can be defined as

$$L(D(p)) = \sup\{\prod_{k=1,2}\prod_{j=1}^{n} p_{kj} : \sum_{j=1}^{n} p_{kj} = 1, \sum_{j=1}^{n} p_{kj} W_{kj}(p) = 0, \sum_{j=1}^{n} p_{1j} V_{1j} - \sum_{j=1}^{n} p_{2j} V_{2j} = D(p), k = 1,2\}, \tag{3-1}$$

where

$$W_{kj}(p) = I(U_{kj} \leq 1-p) - R_k(p) \equiv V_{kj}(p) - R_k(p),$$

$$V_{kj}(p) = I(U_{kj} \leq 1-p), k = 1,2.$$

The placement values, U_{kj}'s $(k = 1,2)$, depend on the unknown distribution functions F_k's $(k = 1,2)$ of the non-diseased populations. Therefore, by replacing F_k by its empirical distribution $\hat{F}_k$, we get an adjusted empirical likelihood for *D(p)*:

$$\hat{L}(D(p)) = \sup\{\prod_{k=1,2}\prod_{j=1}^{n} p_{kj} : \sum_{j=1}^{n} p_{kj} = 1, \sum_{j=1}^{n} p_{kj} \hat{W}_{kj}(p) = 0, \sum_{j=1}^{n} p_{1j} \hat{V}_{1j} - \sum_{j=1}^{n} p_{2j} \hat{V}_{2j} = D(p), k = 1,2\}$$

where

$$\hat{W}_{kj}(p) = I(\hat{U}_{kj} \leq 1-p) - R_k(p) \equiv \hat{V}_{kj}(p) - R_k(p),$$

$$\hat{V}_{kj}(p) = I(\hat{U}_{kj} \leq 1-p), k = 1,2.$$

By using the Lagrange multiplier method, we get the corresponding log-EL ratio statistic:

$$l(D(p)) = 2(\sum_{j=1}^{n} \log(1 + 2t\hat{w}_{1j}(p)) + \sum_{j=1}^{n} \log(1 - 2t\hat{w}_{2j}(p))), \tag{3-2}$$

where $t, R_1(p), R_2(p)$ are determined by

$$\begin{cases} \frac{1}{n}\sum_{j=1}^{n}\frac{\hat{V}_{1j}-R_1(p)}{1+2t(\hat{V}_{1j}-R_1(p))}=0 \\ \frac{1}{n}\sum_{j=1}^{n}\frac{\hat{V}_{2j}-R_2(p)}{1-2t(\hat{V}_{2j}-R_2(p))}=0 \\ \frac{1}{n}\sum_{j=1}^{n}\frac{\hat{V}_{1j}}{1+2t(\hat{V}_{1j}-R_1(p))}-\frac{1}{n}\sum_{j=1}^{n}\frac{\hat{V}_{2j}}{1-2t(\hat{V}_{2j}-R_2(p))}=D(p) \end{cases}$$

Qin (2007) established the following theorem for the asymptotic distribution of the log-EL likelihood ratio statistic.

Theorem 3.1. *If* $D_0(p)$ *is the true value of* $D(p)=R_1(p)-R_2(p)$ *at a fixed level p of specificity, then the limiting distribution of* $l(D(p))$*, defined by equation (3-2), is a scale chi-square distribution with one degree of freedom. That is,*

$$r(p)l(D_0(p))\xrightarrow{d}\chi_1^2,$$

where the scale constant $r(p)$ *is*

$$r(p)=\frac{R_1(p)(1-R_1(p))+R_2(p)(1-R_2(p))}{(1-\lambda)\sigma^2}.$$

The scale constant $r(p)$ in Theorem 3.1 is still unknown. In order to construct confidence intervals for $D(p)$, we propose to use bootstrap method to estimate $r(p)$. The procedure is as follows:

Step 1: Draw resample of size m, $X_{ki}^*'s$, with replacement from the non-diseased sample $X_{ki}'s$ and a separate resample of size n, $Y_{kj}^*'s$, with replacement from the diseased sample $Y_{kj}'s$.

Step 2: Calculate the bootstrap versions $\hat{R}_k^*(p)$ of $R_k(p), k=1,2$.

$$\hat{R}_k^*(p) = \frac{\sum_{i-1}^{nz} I[Y_{ki}^* \geq \hat{F}_k^{*-1}(p) + Z_{1-\alpha/2}^2/2}{n + Z_{1-\alpha/2}^2} \quad k=1,2 .$$

Setp 3: Repeat Steps 1-2 B ($B \geq 150$) times, we get $\{\hat{R}_{1b}^*(p), \hat{R}_{2b}^*(p) : b=1...B\}$ and

$$V_k^* = \frac{1}{B-1}\sum_{b=1}^{B}(\hat{R}_{kb}^*(p) - \bar{R}_k^*(p))^2 \ , k=1,2 ,$$

$$V_{12}^* = \frac{1}{B-1}\sum_{b=1}^{B}((\hat{R}_{1b}^*(p) - \bar{R}_1^*(p)(\hat{R}_{2b}^*(p) - \bar{R}_2^*(p)) ,$$

$$V^* = V_1^* + V_2^* - 2V_{12}^* ,$$

where $\bar{R}_k^*(p_0) = \frac{1}{B}\sum_{b=1}^{B}\hat{R}_{kb}^*(p_0)$, *k=1,2.*

Hence, the scale constant $r(p)$ can be consistently estimated by

$$r_1^*(p) = \frac{\bar{R}_1^*(p)(1-\bar{R}_1^*(p)) + \bar{R}_2^*(p)(1-\bar{R}_2^*(p))}{n * V^*} ,$$

or

$$r_2^*(p) = \frac{\hat{R}_1(p)(1-\hat{R}_1(p)) + \hat{R}_2(p)(1-\hat{R}_2(p))}{n * V^*} .$$

By using these estimates for $r(p)$, we propose two hybrid bootstrap and empirical likelihood based confidence intervals for $D(p)$.

The first one, called HBELI interval, is defined by

$$\{D(p) : r_1^*(p) l(D(p)) \leq \chi_1^2(1-\alpha)\}, \tag{3-5}$$

where $\chi_1^2(1-\alpha)$ is the $(1-\alpha)$-th quantile of χ_1^2.

The second one, called HBELII interval, is defined by

$$\left\{D(p): r_2^*(p) l(D(p)) \le \chi_1^2(1-\alpha)\right\}. \tag{3-6}$$

Chapter 4 Simulation

In this chapter, we conduct two simulation studies using bivariate normal distribution and exponential distribution to evaluate coverage accuracy and interval length of the newly proposed intervals for $D(p)$, the difference of the two sensitivities, when the specificity p is taken to be 0.70, 0.80 or 0.90 in finite-sample sizes. In both studies, we generated 1000 random samples of size n from $G(y_1, y_2)$ for test responses of diseased patients, and another set of independent random samples of size m from $F(x_1, x_2)$ for test responses of non-diseased patients. In this thesis, we didn't use the normal approximation based interval as a comparison because Qin et al. (2006) have already shown that BTI and BTII intervals perform better than the normal approximation based interval for independent samples, and BCa performs better than the normal approximation based interval for paired dependent samples, and these three intervals are computationally much simpler than the normal approximation based interval.

In the first study, $G(y_1, y_2)$ is chosen to be a bivariate normal distribution having mean $E(Y_1) = \mu_1$, $E(Y_2) = \mu_2$ and with a common standard deviation 2 and correlation ρ; $F(x_1, x_2)$ is chosen to be a bivariate normal distribution having means $E(X_1) = 0$, $E(X_2) = 0$ and with a common standard deviation 1 and correlation ρ. ρ is chosen as 0 and 0.5 respectively. Thus,

$$R_k(p) = 1 - \Phi\{(\Phi^{-1}(p) - \mu_k)/2\} \text{ for } k=1,2.$$

For $D(p)$=0, we choose $\mu_1 = \mu_2$ such that the sensitivity $R_k(p)$ of the test $T_k(k = 1,2)$ varies over the points 0.95, 0.90, 0.80, 0.70, 0.60, 0.50, 0.40, 0.30, 0.20, 0.10, respectively .

In the second study, the distributions $G(y_1, y_2)$, $F(x_1, x_2)$ are chosen to be different bivariate exponential distributions that have exponential distributions as their marginal

distributions. Depending on the possible correlation between the test results from two diagnostic tests, we use two different procedures to generate the random samples of test response.

First we choose the correlation as zero (ρ=0), and then we generate two independent samples, $X_{11}, X_{12}, ..., X_{1m}$ and $X_{21}, X_{22}, ..., X_{2m}$, from standard exponential distribution; and two independent samples, $Y_{11}, Y_{12}, ..., Y_{1n}$ and $Y_{21}, Y_{22}, ..., Y_{2n}$ from exponential distributions with rates λ_1, λ_2 , respectively. Therefore,

$$R_k(p) = \exp(\lambda_k \log(1-p)], \quad \text{for } k = 1,2.$$

Similar to the first simulation study, we choose $\lambda_k, l_k (k = 1,2)$ such that $D(p)$= 0 as the sensitivity $R_i(p)$ of the test $T_i (i = 1,2)$ varies over the points 0.95, 0.90, 0.80, 0.70, 0.60, 0.50, 0.40, 0.30, 0.20, 0.10 respectively.

Secondly, we choose a positive correlation for the bivariate exponential distribution (ρ>0). We first generate random sample, $U_{k1}, U_{k2}, ..., U_{km}$, from an exponential distribution with rate 0.5, for $k = 1,2,3;$ and random samples, $V_{k1}, V_{k2}, ..., V_{kn}$, from an exponential distribution with rate l_i, for $i = 1,2.$; and a random sample, $V_{31}, V_{32}, ..., V_{3n}$ from an exponential distributions with rate 0.02. Then the simulated test responses for a non-diseased patients are $X_{ki} = \min(U_{ki}, U_{3i}), k = 1,2, i = 1,2,...,m,$ which are random samples from two standard exponential distributions with correlation ρ; and those for diseased patients are $Y_{kj} = \min(V_{kj}, V_{3j}), k = 1,2, j = 1,2,...,n,$ which are random samples from two exponential distributions with correlation ρ and rates $l_1 + 0{,}02, l_2 + 0{,}02$,respectively. Under this setting,

$$R_k(p) = \exp[(l_k + 0.02)(\lambda_k \log(1-p)], \text{ for } k = 1,2.$$

We choose $\lambda_k, l_k (k=1,2)$ such that $D(p)$=0 as the sensitivity $R_k(p)$ of the test $T_k(k=1,2)$ varies over the points 0.95, 0.90, 0.80, 0.70 respectively.

In the bootstrap step, we draw B=150 bootstrap re-samples from the original samples. We construct 95% confidence intervals for *D(p).* The results of the simulation study are shown in Table I to Table VI in Appendix I. From these tables, the following observations are made.

(1) When the correlation ρ=0 and *D(p)*=0, the BCa, BTI and BTII intervals have better coverage probability, but HBELI and HBELII intervals have shorter interval length.

(2) When the correlation ρ>0 and *D(p)*=0, the five intervals have similar coverage probability, but HBELI and HBELII intervals have shorter interval length.

(3) When the correlation ρ is positive, bigger sample sizes ($m,n \geq 150$) are needed to get better coverage accuracy for all the intervals.

In summary, when correlation exists, the hybrid empirical likelihood and bootstrap based intervals HBELI and HBELII perform better than the bootstrap intervals due to the shorter interval length. When there is no correlation, the bootstrap based intervals BCa, BTI, BTII perform better than the HBELI and HBELII intervals due to better coverage probability.

Chapter 5 Dermatoscope example

Melanoma is a malignant tumor of melanocytes which are found predominantly in skin but also in the bowel and the eye. It is one of the rarer types of skin cancer but causes the majority of skin cancer related deaths. Around 160,000 new cases of melanoma are diagnosed worldwide each year, and it is more frequent in males and Caucasians, especially in Caucasian populations living in sunny climates than other groups. According to the WHO Report about 48,000 melanoma related deaths occur worldwide per annum. Despite many years of intensive laboratory and clinical research, the sole effective cure is surgical resection of the primary tumor before it achieves a thickness greater than 1mm (Wikipedia 2007). Therefore, early diagnose of Melanoma is critical to increase the change to cure the disease.

Dermatoscopy is a hand-held instrument with a dermatoscope, a magnifier with a light and a liquid medium between the instrument and the skin, thus illuminating the skin without reflected light. Dermatoscopy is a noninvasive diagnostic technique for the early diagnosis of melanoma and the evaluation of other pigmented and non-pigmented lesions on the skin that are not as well seen with the unaided eye. Stolz et al. (1994) studied the accuracy of clinical evaluations with or without the aid of Dermatoscopy in detecting malignant Melanoma (MM) by using the ABCD rule (Asymmetry, irregular border, different colors, and Diameter larger than 6mm). In this study, two tests were used for detecting MM on the same subjects. The first test is the clinical assessment without the aid of dermatoscopy, and the second test is the clinical assessment with the aid of dermatoscopy. The data set we used here includes 21 patients with MM and 51 patients with benign melanocytic lesions. The goal is to find out whether the use of dermatoscopy can improve for detecting MM. We estimate the difference between two sensitivities of the two tests and construct confidence intervals for the difference by using BCa,

BTI, BTII, HBELI and HBELII methods. The 95% confidence intervals for the difference between two sensitivities when the specificity is fixed at 0.9 or 0.95 respectively are shown in Appendix I Table V.

All the confidence intervals from above five methods contain zero. In summary, we conclude that there is no significant advantage in adopting the clinical assessment with the aid of dermatoscopy in detecting MM. The same conclusion has been obtained in Qin et al. (2006).

Chapter 6 Discussion

When a new method for continuous-scale tests is developed, comparing its effectiveness with existing methods is necessary. Using the confidence intervals for the difference between two sensitivities of two tests is straightforward. In many cases, only a particular value of specificity is relevant (e.g., 70%, 80%, 90%). Therefore, it is of interest to construct a confidence interval for the sensitivity of the test at a fixed level of specificity.

Qin et al. (2006) proposed three bootstrap-based intervals (BCa, BTI and BTII) for the difference between two sensitivities and showed that these intervals outperform the normal-approximation-based interval. In this thesis, we have proposed another two hybrid empirical likelihood and bootstrap confidence intervals (HBELI and HBELII) for the difference between two sensitivities. Simulation studies show that when correlation exists, HBELI and HBELII intervals perform better than the existing bootstrap based intervals (BCa, BTI and BTII) due to shorter interval length. However, when there is no correlation, BCa, BTI and BTII intervals outperform HBELI and HBELII intervals due to better coverage probability in most simulation settings.

References

Agresti A, Caffo BA (2000). Simple and effective confidence intervals for proportions and difference of proportions result from adding two successes and two failures. American Statistician, **54**:280-288

Greenhouse SW, Mantel N (1950). The evaluation of diagnostic tests. Biometrics, **6**:399-412

Linnet K (1987). Comparison of quantitative diagnostic tests: type I error, power, and sample size. Statistics in Medicine, **6**: 147-158

Owen A (1990). Empirical likelihood ratio confidence regions. Annals of Statistics, **18**: 90-120

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

Pepe MS (2003). The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford University Press, Cary, NC

Platt RW, Hanley JA and Yang H (2000). Bootstrap confidence intervals for the sensitivity of a quantitative diagnostic test. Statistics in Medicine, **19**: 313-322.

Qin GS (2007). Empirical likelihood based confidence intervals for the sensitivity of a continuous-scale diagnostic test at a fixed level of specificity. Manuscript.

Qin GS, Hsu YS, and Zhou XH (2006). New confidence intervals for the difference between two sensitivities at a fixed level of specificity. Statistics in Medicine, **25**: 3487-3502

Shapiro DE (1999). The interpretation of diagnostic tests. Statistical Methods in Medical Research, **8**: 113-134

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

Wieand S, Gail MH, James BR, James KR (1989). A family of non-parametric statistics for comparing diagnostic markers with paired and unpaired data. Biometrics, **76**:585-592

Wikipedia (2007). Retrieved November 9,2007, from <http://en.wikipedia.org/wiki/Melanoma>

Zhou XH, and Qin GS (2005). Improved confidence intervals for the sensitivity at a fixed level of specificity of a continuous-scale diagnostic test. Statistical in Medicine, **24**: 465-477.

Appendix I: Simulation tables

Table a- 1 Level of 95 per cent confidence interval for D(p)=0. Bivariate normal distribution with $\rho = 0$

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(20,20)	BCa	0.8200	0.3864	0.9312	0.4175	0.9331	0.4300
	BTI	0.9030	0.4534	0.9557	0.4797	0.9584	0.4966
	BTII	0.9080	0.4534	0.9645	0.4797	0.9701	0.4966
	HBELI	0.8877	0.2591	0.8714	0.2421	0.8729	0.1902
	HBELII	0.8896	0.2591	0.8729	0.2421	0.8745	0.1903
(50,50)	BCa	0.8905	0.2892	0.9388	0.3101	0.9418	0.3222
	BTI	0.9330	0.3191	0.9555	0.3382	0.9588	0.3535
	BTII	0.9380	0.3191	0.9645	0.3382	0.9682	0.3535
	HBELI	0.9191	0.2810	0.8714	0.2421	0.9178	0.1940
	HBELII	0.9203	0.2810	0.8729	0.2421	0.9189	0.1941
(80,80)	BCa	0.9105	0.2434	0.8886	0.2452	0.8944	0.2507
	BTI	0.9430	0.2590	0.9280	0.2630	0.9420	0.2747
	BTII	0.9460	0.2590	0.9350	0.2630	0.9440	0.2747
	HBELI	0.9341	0.2490	0.9225	0.2205	0.9323	0.1795
	HBELII	0.9343	0.2490	0.9222	0.2205	0.9332	0.1795
(150,150)	BCa	0.9505	0.1942	0.9433	0.1969	0.9390	0.2054
	BTI	0.9544	0.2003	0.9522	0.2037	0.9496	0.2136
	BTII	0.9361	0.2003	0.9573	0.2037	0.9576	0.2136
	HBELI	0.9413	0.1912	0.9312	0.1741	0.9364	0.1464
	HBELII	0.9413	0.1912	0.9308	0.1741	0.9366	0.1464

Table a- 1 Level of 95 per cent confidence interval for D(p)=0. Bivariate normal distribution with $\rho = 0$ (continued)

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(50,30)	BCa	0.8635	0.3318	0.9360	0.3308	0.9418	0.3511
	BTI	0.9110	0.3753	0.9561	0.3614	0.9603	0.3842
	BTII	0.9220	0.3753	0.9667	0.3614	0.9687	0.3942
	HBELI	0.9067	0.2866	0.8956	0.2590	0.8896	0.2035
	HBELII	0.9061	0.2866	0.8957	0.2590	0.8899	0.2035
(100,80)	BCa	0.9377	0.2496	0.9371	0.2530	0.9425	0.2627
	BTI	0.9530	0.2636	0.9519	0.2675	0.9540	0.2790
	BTII	0.9594	0.2636	0.957	0.2675	0.9621	0.2790
	HBELI	0.9329	0.2451	0.9273	0.2201	0.9366	0.1813
	HBELII	0.9330	0.2451	0.928	0.2201	0.9365	0.1813

Table a- 2 Level of 95 per cent confidence interval for D(p)=0. Bivariate normal distribution with $\rho = 0.5$

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(20,20)	BCa	0.9667	0.4437	0.9623	0.4493	0.9646	0.4649
	BTI	0.9180	0.3705	0.9193	0.3760	0.9212	0.3917
	BTII	0.9407	0.3705	0.9448	0.3760	0.9470	0.3917
	HBELI	0.9263	0.2848	0.9099	0.2554	0.9100	0.1909
	HBELII	0.9271	0.2848	0.9094	0.2554	0.9113	0.1910
(50,50)	BCa	0.9687	0.3183	0.9687	0.3225	0.9688	0.3347
	BTI	0.9111	0.2479	0.9155	0.2525	0.9241	0.2647
	BTII	0.9294	0.2479	0.9336	0.2525	0.9332	0.2647
	HBELI	0.9538	0.2910	0.9425	0.2516	0.9413	0.1945
	HBELII	0.9534	0.2910	0.9425	0.2516	0.9414	0.1946
(80,80)	BCa	0.9719	0.2598	0.9752	0.2647	0.9696	0.2743
	BTI	0.9114	0.1985	0.9122	0.2016	0.9115	0.2115
	BTII	0.9264	0.1985	0.9247	0.2016	0.9315	0.2115
	HBELI	0.9563	0.2556	0.9477	0.2266	0.9532	0.1789
	HBELII	0.9564	0.2556	0.9481	0.2266	0.9535	0.1789
(150,150)	BCa	0.9733	0.1958	0.9738	0.1991	0.9714	0.2066
	BTI	0.9141	0.1460	0.9175	0.1485	0.9112	0.1558
	BTII	0.9247	0.1460	0.9267	0.1485	0.9254	0.1558
	HBELI	0.9714	0.1931	0.9563	0.1783	0.9534	0.1470
	HBELII	0.9716	0.1931	0.9561	0.1783	0.9541	0.1470

Table a- 2 Level of 95 per cent confidence interval for D(p)=0. Bivariate normal distribution with $\rho = 0.5$ (continued)

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(50,30)	BCa	0.9565	0.3589	0.9696	0.3435	0.9708	0.3598
	BTI	0.9114	0.2984	0.9142	0.2700	0.9182	0.2864
	BTII	0.9273	0.2984	0.9378	0.2700	0.9414	0.2864
	HBELI	0.9399	0.3052	0.9358	0.2695	0.9233	0.2017
	HBELII	0.9393	0.3052	0.9345	0.2695	0.9235	0.2017
(100,80)	BCa	0.9701	0.2547	0.9701	0.2583	0.9716	0.2672
	BTI	0.9080	0.1938	0.9136	0.1970	0.9141	0.2056
	BTII	0.9191	0.1938	0.9230	0.1970	0.9302	0.2056
	HBELI	0.9609	0.2515	0.9540	0.2237	0.9366	0.1813
	HBELII	0.9606	0.2515	0.9551	0.2267	0.9365	0.1813

Table a- 3 Level of 95 per cent confidence interval for D(p)=0. Bivariate exponential distribution with $\rho = 0$

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(20,20)	BCa	0.9323	0.4553	0.9311	0.4484	0.9277	0.4430
	BTI	0.9593	0.5367	0.9644	0.5311	0.9591	0.5303
	BTII	0.9695	0.5367	0.9730	0.5311	0.9712	0.5303
	HBELI	0.8425	0.2392	0.8714	0.2421	0.8978	0.0420
	HBELII	0.8443	0.2391	0.8729	0.2421	0.9041	0.0420
(50,50)	BCa	0.9390	0.3450	0.9370	0.3393	0.9358	0.3434
	BTI	0.9632	0.3808	0.9576	0.3766	0.9580	0.3810
	BTII	0.9703	0.3808	0.9668	0.3766	0.9688	0.3810
	HBELI	0.8897	0.2660	0.8963	0.1508	0.9178	0.1940
	HBELII	0.8912	0.2660	0.897	0.1509	0.9178	0.1941
(80,80)	BCa	0.9393	0.2859	0.9360	0.2829	0.9435	0.2864
	BTI	0.9550	0.3089	0.9550	0.3062	0.9605	0.3119
	BTII	0.9613	0.3089	0.9617	0.3062	0.9688	0.3119
	HBELI	0.9208	0.2449	0.9183	0.1474	0.9023	0.0100
	HBELII	0.9204	0.2449	0.9176	0.1474	0.9028	0.0100
(150,150)	BCa	0.9409	0.2168	0.9400	0.2154	0.9407	0.2170
	BTI	0.9568	0.2299	0.9568	0.2287	0.9560	0.2322
	BTII	0.9633	0.2299	0.9621	0.2287	0.9647	0.2322
	HBELI	0.9387	0.1949	0.9299	0.1296	0.9192	0.0500
	HBELII	0.9389	0.1949	0.9302	0.1296	0.9200	0.0500

Table a- 3 Level of 95 per cent confidence interval for D(p)=0. Bivariate exponential distribution with $\rho = 0$ (continued)

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(50,30)	BCa	0.9283	0.3767	0.9325	0.3695	0.8828	0.4303
	BTI	0.9581	0.4348	0.9555	0.4121	0.9337	0.4910
	BTII	0.9662	0.4348	0.9671	0.4121	0.9443	0.4910
	HBELI	0.8694	0.2690	0.863	0.1523	0.8840	0.0225
	HBELII	0.8689	0.2690	0.8635	0.1524	0.8849	0.0225
(100,80)	BCa	0.9419	0.2782	0.9400	0.2741	0.9389	0.2772
	BTI	0.9566	0.2990	0.9549	0.2958	0.9602	0.3006
	BTII	0.9648	0.2990	0.9611	0.2958	0.9674	0.3006
	HBELI	0.9173	0.2383	0.9182	0.1450	0.8985	0.0015
	HBELII	0.9177	0.2383	0.9183	0.1450	0.8987	0.0014

Table a- 4 Level of 95 per cent confidence interval for D(p)=0. Bivariate exponential distribution with $\rho \neq 0$ (Using 0.02 to generate diseased random sample)

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(20,20)	BCa	0.9517	0.3842	0.9573	0.3748	0.9700	0.3713
	BTI	0.9161	0.3312	0.9135	0.3266	0.9354	0.3248
	BTII	0.9437	0.3312	0.9386	0.3266	0.9503	0.3248
	HBELI	0.8618	0.0937	0.9005	0.0440	0.9358	0.0081
	HBELII	0.8648	0.0937	0.9005	0.0440	0.9364	0.0081
(50,50)	BCa	0.9565	0.2735	0.9670	0.2717	0.9664	0.2883
	BTI	0.9047	0.2178	0.9202	0.2158	0.9153	0.2301
	BTII	0.9194	0.2178	0.9320	0.2158	0.9323	0.2301
	HBELI	0.9085	0.1566	0.9044	0.1039	0.9403	0.0500
	HBELII	0.9089	0.1566	0.9039	0.1039	0.9408	0.0580
(80,80)	BCa	0.9619	0.2238	0.9668	0.2232	0.9702	0.2376
	BTI	0.8913	0.1743	0.9159	0.1734	0.9125	0.1843
	BTII	0.9071	0.1743	0.9234	0.1734	0.9267	0.1843
	HBELI	0.9319	0.1606	0.9103	0.1102	0.9493	0.0290
	HBELII	0.9321	0.1606	0.9088	0.1102	0.9488	0.0290
(150,150)	BCa	0.9659	0.1708	0.9669	0.1704	0.9737	0.1712
	BTI	0.8990	0.1288	0.9108	0.1284	0.9250	0.1274
	BTII	0.9058	0.1288	0.9195	0.1284	0.9358	0.1274
	HBELI	0.9565	0.1301	0.9469	0.1026	0.9215	0.0480
	HBELII	0.9565	0.1302	0.9469	0.1026	0.9220	0.0480

Table a- 4 Level of 95 per cent confidence interval for D(p)=0. Bivariate exponential distribution with $\rho \neq 0$ (Using 0.02 to generate diseased random sample) (continued)

Sample size	Method	Specificity=0.7		Specificity=0.8		Specificity=0.9	
		Coverage probability	Length	Coverage probability	Length	Coverage probability	Length
(50,30)	BCa	0.9479	0.3067	0.9511	0.3044	0.9612	0.3068
	BTI	0.9068	0.2622	0.9094	0.2621	0.9300	0.2628
	BTII	0.9190	0.2622	0.9232	0.2621	0.9431	0.2628
	HBELI	0.8808	0.1395	0.8883	0.0903	0.9253	0.0679
	HBELII	0.8780	0.1395	0.8870	0.0903	0.9253	0.0680
(100,80)	BCa	0.9732	0.2182	0.9658	0.2182	0.9593	0.2194
	BTI	0.9200	0.1702	0.9107	0.1691	0.9009	0.1700
	BTII	0.9275	0.1702	0.9210	0.1691	0.9102	0.1700
	HBELI	0.9260	0.1580	0.9110	0.1180	0.9108	0.0700
	HBELII	0.9258	0.1581	0.9115	0.1181	0.9095	0.0700

Table a- 5 95 per cent confidence interval for the difference of sensitivities between the two clinical assessments with and without the use of dermatoscopy

Specificity	Bca	BTI	BTII	HBELI	HBELII
0.90	(-0.261,0.261)	(-0.220,0.394)	(-0.302,0.312)	(-0.183,0.183)	(-0.183,0.183)
0.95	(-0.609,0.479)	(-0.346,0.346)	(-0.336,0.357)	(-0.052,0.052)	(-0.052,0.052)

Appendix II S-plus code for Simulation

1. Normal distribution

```
##############################################################################
#        Functions
#Get sensitivity from abnorm and norm samples at fixed specificity p
# m: number of bootstrap
#################################################################################
sensb<-function(abnorm, norm, p, m)
{
        result <- rep(NA, m)
        if(m > 1) {
                for(i in 1:m) {
                        t <- sample(abnorm, length(abnorm), replace = T)
                        u <- sample(norm, length(norm), replace = T)
                        if(max(t) < min(u)) {
                                result[i] <- 0
                        }
                        else {
                                #result[i] <- (sum(t > quantile(u, p))+ k^2/2)/(length(t)+k^2)
        result[i] <- sum(t > quantile(u, p))/length(t)
                        }
                }
        }
        else result[1] <- sum(abnorm > quantile(norm, p))/length(abnorm)
        return(result)
}

 solveNonlinear<-function(f,y0,x)
 {
   g<-function(x,y0,f) sum((f(x)-y0)^2)
   g$y0<-y0
   g$f<-f
   nlmin(g,x,max.fcal=100,max.iter=100)
 }

#############################################################
#                                                           #
#            Main Program                                   #
#############################################################

mm<-1000                # number of repetition
m<-80           # sample sizes of non-diseased samples
n<-80

tt<-0.7    # Specificity level        tt
#tt<-0.8
#tt<-0.7

rho=0

alpha<-0.05
```

```
Cvalue<-qchisq(1-alpha,1)  #chi-sq(1-alpha,1)

# sensitivity 1 -sensitivity 2 =0
ss1<-c(0.95,0.90,0.80,0.70,0.60,0.50,0.40,0.30,0.20,0.10)  #R1(t)  sensitivity 1
ss2<-c(0.95,0.90,0.80,0.70,0.60,0.50,0.40,0.30,0.20,0.10)  #R2(t)  sensitivity 2

inrange1<-0
logltzero<-0  #record the number which less than zero in LDp

Dp.low<-0
Dp.up<-0
nlow<-0   #if the function is converage, nlow+1
nup<-0

Dp2.low<-0
Dp2.up<-0
nlow2<-0   #idicator whether the function is converage, nlow2+1
nup2<-0

inrange2<-0

for (i in 1:length(ss1))
{
        cov1<-cov2<-0                                   # coverage

        for (j in 1:mm)
        {
         mud1<-qnorm(tt,0,1)-2*qnorm(1-ss1[i],0,1)            # mean of the first
                                 diseased population
         mud2<-qnorm(tt,0,1)-2*qnorm(1-ss2[i],0,1)            # mean of the second
                                  diseased population
        Rtt1<-1-pnorm(qnorm(tt,0,1),mud1,2)        #the first true sensitivity
        Rtt2<-1-pnorm(qnorm(tt,0,1),mud2,2)
        Rtt<-Rtt1-Rtt2              #the difference of two true sensitivities


#############  step 1 ############################
# Generate diseased and non-diseased distribution
##################################################
# generate two samples from the nondiseased populations:

        xx<-rmvnorm(m, mean=c(0,0), cov=matrix(c(1,rho,rho,1),2))

        x10<-xx[,1]    # the sample from the first nondiseased population
        x20<-xx[,2]    # the sample from the second nondiseased population

# generate two samples from the deseased populations:
        yy<-rmvnorm(n, mean=c(mud1, mud2), ov=matrix(c(4,rho*2*2,rho*2*2,4),2))

# the sample from 2-dimensinal multinomial distribution with mean=c(mud1, mud2),sd=2, and correlation=0.5
        y11<-yy[,1]          # the sample from the first diseased population
        y21<-yy[,2]          # the sample from the second diseased population

# Two estimated sensitivities at specificity (tt):

    sens1<-sum((yy[,1] >=quantile(xx[,1],tt)))/n    # estimated sensitivity
```

```
                                        from the first sample
            sens2<-sum((yy[,2] >=quantile(xx[,2],tt)))/n    # estimated sensitivity
                                        from the second sample


 #############  step 2 ##############################
 #Bootstrap
 # Generate diseased and non-diseased distribution
#######################################################
B=150

 #get sensitivity from bootstrap samples
 Rb1<-sensb(y11, x10,tt,B)
 Rb2<-sensb(y21, x20,tt,B)

 vb1<-sum((Rb1-mean(Rb1))^2)/(B-1)                          # V_1^*(t)
 vb2<-sum((Rb2-mean(Rb2))^2)/(B-1)

 vb12<-0
 if (rho!=0)
            vb12<-sum( (Rb1-mean(Rb1))*(Rb2-mean(Rb2)) )/(B-1)

 vb<-vb1+vb2-2*vb12                                          # Bootstrap variance estimate

 #############  step 3 ##############################
 R1<-0
 R2<-0
 if(vb!=0)
 {
            R1<-(mean(Rb1)*(1-mean(Rb1))+mean(Rb2)*(1-mean(Rb2)))/(n*vb)  #estimate
                                        for the scale constant
            R2<-( sens1*(1-sens1) + sens2*(1-sens2) )/(n*vb)

 ############Calculate L(D(p))#########
 f1<-2
 f2<-2

  u11hat<-rep(100,n)
  u22hat<-rep(100,n)

 for(ii in 1:n)                             # hat Uk=1-F(Yk)
 {
  u11hat[ii]<-1-mean(x10<=y11[ii])
  u22hat[ii]<-1-mean(x20<=y21[ii])
 }

 v11hat<-(u11hat<=tt)*1    # indicator function of U:I(U_j<=p)
 v22hat<-(u22hat<=tt)*1

 ########solove R_1(p), R_2(p) and lambda ###########

 g<-function(x,v1h=v11hat, v2h=v22hat)
    {
                        y_numeric(3)

                        y[1]_mean( (v11hat-x[1])/(1-2*x[3]*(v11hat-x[1])) )
```

```
                    y[2]_mean( (v22hat-x[2])/(1+2*x[3]*(v22hat-x[2])) )
                    y[3]_mean( v22hat/(1+2*x[3]*(v22hat-x[2]))) - mean( v11hat/(1-
                2*x[3]*(v11hat-x[1])))
          y
          }

 sol<-solveNonlinear(g,c(0,0,Rtt),c(Rtt1,Rtt2,0))
                                        # c(Rtt1,Rtt2,0) are initial values of
                    c(R_1(p), R_2(p), lambda)

 newr1<-sol$x[1]
 newr2<-sol$x[2]
 lambda<-sol$x[3]

 w11hat<-v11hat-newr1
 w22hat<-v22hat-newr2

###### test the number when (1-2*lambda*w11hat or 1+2*lambda*w22hat <0
flag<-0
for(ii in 1:n)   {
 if ((1-2*lambda*w11hat[ii])<0 || (1+2*lambda*w22hat[ii])<0)
          flag<-1
}
if(flag=1) logltzero<-logltzero+1

 LDp<-2*( sum(log(abs(1-2*lambda*w11hat)))+sum(log(abs(1+2*lambda*w22hat))))

###using abs here
# LDp<-2*( sum(log(1-2*lambda*w11hat))+sum(log(1+2*lambda*w22hat)))

inrange1<-inrange1 + (R1*LDp<qchisq(1-alpha,1))*1
inrange2<-inrange2 + (R2*LDp<qchisq(1-alpha,1))*1

########solove R_1(p), R_2(p) ,lambda, D(p) to find confidence interval of D(p)  04/24/2007###########

 f<-function(x,v1h=v11hat, v2h=v22hat c=Cvalue)
   {
                    y_numeric(4)

                    y[1]_mean( (v11hat-x[1])/(1-2*x[3]*(v11hat-x[1])) )
                    y[2]_mean( (v22hat-x[2])/(1+2*x[3]*(v22hat-x[2])) )
                    y[3]_mean( v22hat/(1+2*x[3]*(v22hat-x[2]))) - mean(v11hat/(1-2*x[3]*(v11hat-x[1])))-x[4]
     y[4]_R1*2*( sum(log(abs(1-2*x[3]*(v11hat-x[1]))))+sum(log(abs(1+2*x[3]*(v22hat-x[2])))))-Cvalue

                    y
          }

 solf1<-solveNonlinear(f,c(0,0,0,0),c((Rtt1+0.1),(Rtt2-0.1),0,0.2)) #initial values
 if(solf1$converged = T)
{
          nlow<-nlow+1
  Dp.low<-Dp.low+solf1$x[4]
 }
 solf2<-solveNonlinear(f,c(0,0,0,0),c((Rtt1-0.1),(Rtt2+0.1),0,-0.2))
 if(solf2$converged = T)
{
```

```
        nup<-nup+1
  Dp.up<-Dp.up+solf2$x[4]
 }


########solove R_1(p), R_2(p) ,lambda, D(p) to find confidence interval of D(p) by using R2
04/24/2007###########

 f<-function(x,v1h=v11hat, v2h=v22hat c=Cvalue)
   {
                y_numeric(4)

                y[1]_mean( (v11hat-x[1])/(1-2*x[3]*(v11hat-x[1])) )
                y[2]_mean( (v22hat-x[2])/(1+2*x[3]*(v22hat-x[2])) )
                y[3]_mean( v22hat/(1+2*x[3]*(v22hat-x[2]))) - mean(v11hat/(1-2*x[3]*(v11hat-x[1])))-x[4]
     y[4]_R2*2*( sum(log(abs(1-2*x[3]*(v11hat-x[1]))))+sum(log(abs(1+2*x[3]*(v22hat-x[2])))))-Cvalue

                y
        }

 solf3<-solveNonlinear(f,c(0,0,0,0),c((Rtt1+0.1),(Rtt2-0.1),0,0.2)) #initial values
 if(solf3$converged = T)
{
        nlow2<-nlow2+1
  Dp2.low<-Dp.low+solf1$x[4]
 }
 solf4<-solveNonlinear(f,c(0,0,0,0),c((Rtt1-0.1),(Rtt2+0.1),0,-0.2))
 if(solf4$converged = T)
{
        nup2<-nup2+1
  Dp2.up<-Dp.up+solf2$x[4]
 }


} #end of if(vb!=0)

} #end of loop for (j in 1:mm)

 } #end of loop for (i in 1:length(ss1))

newcov1<-inrange1/(10*mm)
newcov2<-inrange2/(10*mm)
if(nlow & nup)
{
        Dplow<-min(Dp.low/nlow,Dp.up/nup)
        Dpup<-max(Dp.low/nlow,Dp.up/nup)
}

Dplength<-max(Dpup,Dplow)-min(Dpup,Dplow)

if(nlow2 & nup2)
{
        Dp2low<-min(Dp2.low/nlow2, Dp2.up/nup2)
        Dp2up<-max(Dp2.low/nlow2, Dp.up/nup2)
}
```

```
Dplength2<-max(Dp2up,Dp2low)-min(Dp2up, Dp2low)


#Result Output
sink("D:\\Suqin\\normalresult1.txt",append = T)

cat("#######################################\n");
cat(" specifciity=",tt, "\n")
cat(" rho=",rho, "\n")
cat(" Non-disease sample m=", m, " disease sample n=", n, "iteration mm=", mm, "\n")
cat(" Number of log <0 ",logltzero,"\n\n");

cat(" Coverage1=", newcov1,"\n");
cat(" Dp  Lower bound 1=", Dplow, " Up bound 1=", Dpup, "\n")
cat(" Coverage length 1 =", Dplength,"\n")
cat(" Number of converge nlow1=", nlow, "  nup1=", nup,"\n\n")

cat(" Coverage2=", newcov2,"\n");
cat(" Dp  Lower bound 2=", Dp2low, " Up bound 2=", Dp2up, "\n")
cat(" Coverage length 2 =", Dplength2,"\n")
cat(" Number of converge nlow2=", nlow2, "  nup2=", nup2,"\n")

cat("#######################################\n");

sink();
```

2. Exponential distribution – no correlation

```
#######################################################
#                                                     #
#                Main Program                         #
#######################################################

mm<-1000      # number of repetition
m<-150        # sample sizes of non-diseased samples
n<-80        # sample sizes of diseased samples

tt<-0.9     # Specificity level  tt
#tt<-0.8
#tt<-0.7

rho=0


alpha<-0.05
Cvalue<-qchisq(1-alpha,1)  #chi-sq(1-alpha,1)

# sensitivity 1 -sensitivity 2 =0
ss1<-c(0.95,0.90,0.80,0.70,0.60,0.50,0.40,0.30,0.20,0.10)  #R1(t)  sensitivity 1
ss2<-c(0.95,0.90,0.80,0.70,0.60,0.50,0.40,0.30,0.20,0.10)  #R2(t)  sensitivity 2


inrange1<-0
logltzero<-0  #record the number which less than zero in LDp

Dp.low<-0
Dp.up<-0
nlow<-0   #if the function is converage, nlow+1
nup<-0

Dp2.low<-0
Dp2.up<-0
nlow2<-0   #idicator whether the function is converage, nlow2+1
nup2<-0

inrange2<-0

for (i in 1:length(ss1))
{
   cov1<-cov2<-0                 # coverage

    l1<-log(ss1[i])/log(1-tt)            # rate of the Exp(l1) (first diseased
   population) (rate=1/expectation)
    l2<-log(ss2[i])/log(1-tt)            # rate of the Exp(l2) (second diseased
   population)

   Rtt1<-exp(l1*log(1-tt))
   Rtt2<-exp(l2*log(1-tt))
   Rtt<-Rtt1-Rtt2                              # the difference of two true
   sensitivities

   for (j in 1:mm)
   {
   #  mud1<-qnorm(tt,0,1)-2*qnorm(1-ss1[i],0,1)   # mean of the first diseased
   population
   #  mud2<-qnorm(tt,0,1)-2*qnorm(1-ss2[i],0,1)   # mean of the second diseased
   population
```

```
   #   Rtt1<-1-pnorm(qnorm(tt,0,1),mud1,2)          #the first true sensitivity
   #   Rtt2<-1-pnorm(qnorm(tt,0,1),mud2,2)
   #   Rtt<-Rtt1-Rtt2                                # the difference of two true
   sensitivities


#############  step 1 ############################
# Generate diseased and non-diseased distribution
##################################################

#Exponential distribution
# two dependent samples from the nondiseased populations:

 x10<-rexp(m,1)         # Exp(1):the sample from the first nondiseased population
 x20<-rexp(m,1)         # Exp(1):the sample from the second nondiseased population

# two dependent samples from the deseased populations:

 y11<-rexp(n,l1)        # Exp(l1):the sample from the first diseased population
 y21<-rexp(n,l2)        # Exp(l2):the sample from the second diseased population

   sens1<-sum((y11 >=quantile(x10,tt)))/n    # estimated sensitivity from the first
   sample
   sens2<-sum((y21 >=quantile(x20,tt)))/n    # estimated sensitivity from the second
   sample



 #############  step 2 ############################
 #Bootstrap
 # Generate diseased and non-diseased distribution
##################################################
B=150

 #get sensitivity from bootstrap samples
 Rb1<-sensb(y11, x10,tt,B)
 Rb2<-sensb(y21, x20,tt,B)

 vb1<-sum((Rb1-mean(Rb1))^2)/(B-1)                              # V_1^*(t)
 vb2<-sum((Rb2-mean(Rb2))^2)/(B-1)

 vb12<-0
 if (rho!=0)
    vb12<-sum( (Rb1-mean(Rb1))*(Rb2-mean(Rb2)) )/(B-1)

 vb<-vb1+vb2-2*vb12              # Bootstrap variance estimate

 #############  step 3 ############################
 R1<-0
 R2<-0
 if(vb!=0)
 {
   R1<-(mean(Rb1)*(1-mean(Rb1))+mean(Rb2)*(1-mean(Rb2)))/(n*vb)  #estimate for the
   scale constant
   R2<-( sens1*(1-sens1) + sens2*(1-sens2) )/(n*vb)

 ############Calculate L(D(p))#########
 f1<-2
 f2<-2

  u11hat<-rep(100,n)
  u22hat<-rep(100,n)
```

```
 for(ii in 1:n)                # hat Uk=1-F(Yk)
 {
  u11hat[ii]<-1-mean(x10<=y11[ii])
  u22hat[ii]<-1-mean(x20<=y21[ii])
 }

 v11hat<-(u11hat<=tt)*1    # indicator function of U:I(U_j<=p)
 v22hat<-(u22hat<=tt)*1

 ########solve R_1(p), R_2(p) and lambda ###########

 g<-function(x,v1h=v11hat, v2h=v22hat)
    {
       y_numeric(3)

       y[1]_mean( (v11hat-x[1])/(1-2*x[3]*(v11hat-x[1])) )
       y[2]_mean( (v22hat-x[2])/(1+2*x[3]*(v22hat-x[2])) )
       y[3]_mean( v22hat/(1+2*x[3]*(v22hat-x[2]))) - mean( v11hat/(1-2*x[3]*(v11hat-
   x[1])))

       y
    }

 sol<-solveNonlinear(g,c(0,0,Rtt),c(Rtt1,Rtt2,0))
                # c(Rtt1,Rtt2,0) are initial values of c(R_1(p), R_2(p), lambda)

 newr1<-sol$x[1]
 newr2<-sol$x[2]
 lambda<-sol$x[3]

 w11hat<-v11hat-newr1
 w22hat<-v22hat-newr2

###### test the number when (1-2*lambda*w11hat or 1+2*lambda*w22hat <0
flag<-0
for(ii in 1:n)   {
 if ((1-2*lambda*w11hat[ii])<0 || (1+2*lambda*w22hat[ii])<0)
    flag<-1
}
if(flag=1) logltzero<-logltzero+1

 LDp<-2*( sum(log(abs(1-2*lambda*w11hat)))+sum(log(abs(1+2*lambda*w22hat))) )
    ###using abs here
# LDp<-2*( sum(log(1-2*lambda*w11hat))+sum(log(1+2*lambda*w22hat)))

inrange1<-inrange1 + (R1*LDp<qchisq(1-alpha,1))*1
inrange2<-inrange2 + (R2*LDp<qchisq(1-alpha,1))*1

########solve R_1(p), R_2(p) ,lambda, D(p) to find confidence interval of D(p)
    04/24/2007###########

 f<-function(x,v1h=v11hat, v2h=v22hat c=Cvalue)
    {
       y_numeric(4)

       y[1]_mean( (v11hat-x[1])/(1-2*x[3]*(v11hat-x[1])) )
       y[2]_mean( (v22hat-x[2])/(1+2*x[3]*(v22hat-x[2])) )
       y[3]_mean( v22hat/(1+2*x[3]*(v22hat-x[2]))) - mean(v11hat/(1-2*x[3]*(v11hat-
   x[1])))-x[4]
        y[4]_R1*2*( sum(log(abs(1-2*x[3]*(v11hat-x[1]))))+sum(log(abs(1+2*x[3]*(v22hat-
   x[2])))))-Cvalue
```

```
      y
   }

 solf1<-solveNonlinear(f,c(0,0,0,0),c((Rtt1+0.1),(Rtt2-0.1),0,0.2)) #initial values
 if(solf1$converged = T)
{
   nlow<-nlow+1
   Dp.low<-Dp.low+solf1$x[4]
 }
 solf2<-solveNonlinear(f,c(0,0,0,0),c((Rtt1-0.1),(Rtt2+0.1),0,-0.2))
 if(solf2$converged = T)
{
   nup<-nup+1
   Dp.up<-Dp.up+solf2$x[4]
 }


########solove R_1(p), R_2(p) ,lambda, D(p) to find confidence interval of D(p) by
   using R2  04/24/2007###########

 f<-function(x,v1h=v11hat, v2h=v22hat c=Cvalue)
    {
      y_numeric(4)

      y[1]_mean( (v11hat-x[1])/(1-2*x[3]*(v11hat-x[1])) )
      y[2]_mean( (v22hat-x[2])/(1+2*x[3]*(v22hat-x[2])) )
      y[3]_mean( v22hat/(1+2*x[3]*(v22hat-x[2]))) - mean(v11hat/(1-2*x[3]*(v11hat-
   x[1])))-x[4]
       y[4]_R2*2*( sum(log(abs(1-2*x[3]*(v11hat-x[1]))))+sum(log(abs(1+2*x[3]*(v22hat-
   x[2])))))-Cvalue

      y
   }

 solf3<-solveNonlinear(f,c(0,0,0,0),c((Rtt1+0.1),(Rtt2-0.1),0,0.2)) #initial values
 if(solf3$converged = T)
{
   nlow2<-nlow2+1
   Dp2.low<-Dp.low+solf1$x[4]
 }
 solf4<-solveNonlinear(f,c(0,0,0,0),c((Rtt1-0.1),(Rtt2+0.1),0,-0.2))
 if(solf4$converged = T)
{
   nup2<-nup2+1
   Dp2.up<-Dp.up+solf2$x[4]
 }


} #end of if(vb!=0)



   } #end of loop for (j in 1:mm)


 } #end of loop for (i in 1:length(ss1))


newcov1<-inrange1/(10*mm)
newcov2<-inrange2/(10*mm)
if(nlow & nup)
{
   Dplow<-min(Dp.low/nlow,Dp.up/nup)
```

```
    Dpup<-max(Dp.low/nlow,Dp.up/nup)
}

#Dplow<-Dp.low/(10*mm)
#Dpup<-Dp.up/(10*mm)
Dplength<-max(Dpup,Dplow)-min(Dpup,Dplow)

if(nlow2 & nup2)
{
    Dp2low<-min(Dp2.low/nlow2, Dp2.up/nup2)
    Dp2up<-max(Dp2.low/nlow2, Dp.up/nup2)
}

Dplength2<-max(Dp2up,Dp2low)-min(Dp2up, Dp2low)


#Result Output
sink("D:\\Suqin\\Expind_rhoeq0.txt",append = T)

cat("#######Exponential Distribution ##############################\n");
cat(" specificity=",tt, "\n")
cat(" rho=",rho, "\n")
cat(" Non-disease sample m=", m, " disease sample n=", n, "iteration mm=", mm, "\n")
cat(" Number of log <0 ",logltzero,"\n\n");

cat(" Coverage1=", newcov1,"\n");
cat(" Dp  Lower bound 1=", Dplow, " Up bound 1=", Dpup, "\n")
cat(" Coverage length 1 =", Dplength,"\n")
cat(" Number of converge nlow1=", nlow, "  nup1=", nup,"\n\n")

cat(" Coverage2=", newcov2,"\n");
cat(" Dp  Lower bound 2=", Dp2low, " Up bound 2=", Dp2up, "\n")
cat(" Coverage length 2 =", Dplength2,"\n")
cat(" Number of converge nlow2=", nlow2, "  nup2=", nup2,"\n")

cat("#####################################\n");

sink();
```

3. Exponential distribution – correlation

```
…( Specificity and sample size setting are the same as exponential distribution
   without correlation)…

  expcov<-0.02

 l1<-log(ss1[i])/log(1-tt)-expcov           # rate of the Exp(l1) (first diseased
   population) (rate=1/expectation)
 l2<-log(ss2[i])/log(1-tt)-expcov             # rate of the Exp(l2) (second diseased
   population)


 Rtt1<-exp((l1+expcov)*log(1-tt))
 Rtt2<-exp((l2+expcov)*log(1-tt))
 Rtt<-Rtt1-Rtt2                             # the difference of two true sensitivities

 mnb1<-mnb2<-mnb3<-0
 LUb1<-LUb3<-0

for (j in 1:mm)
{

explambda<-0.5

 u1<-rexp(m,explambda)
 u2<-rexp(m,explambda)
 u3<-rexp(m,explambda)


# two dependent samples from the nondiseased populations:
 x10<-x20<-0
 for (k in 1:m)
{
 x10[k]<-min(u1[k],u3[k])         # Exp(1):the sample from the first nondiseased
   population
 x20[k]<-min(u2[k],u3[k])         # Exp(1):the sample from the second nondiseased
   population
}

# two dependent samples from the deseased populations:

 v1<-rexp(n,l1)
 v2<-rexp(n,l2)
 v3<-rexp(n,expcov)

 y11<-y21<-0
 for (k in 1:n)
{
 y11[k]<-min(v1[k],v3[k])             # Exp(l1+0.01):the sample from the first
   diseased population
 y21[k]<-min(v2[k],v3[k])             # Exp(l2+0.01):the sample from the second
   diseased population
}

# Two estimated sensitivities at specificity (tt):

X1.hat<-sum((y11 >=quantile(x10,tt)))
sens1<-X1.hat/n          # estimated sensitivity from the first sample

X2.hat<-sum((y21 >=quantile(x20,tt)))
sens2<-X2.hat/n          # estimated sensitivity from the second sample
```

```
…(Same as exponential distribution without correlation)…
}

#Result Output

sink("D:\\Suqin\\Exponential_Rhogt0_result.txt",append = T)

cat("#######Exponential Distribution ##############################\n");
cat(" specificity=",tt, "\n")
cat("Exponential lambda=",explambda," expcov=",expcov,"\n")
cat(" rho>0","inlambda=",inlambda, "\n")
cat(" Non-disease sample m=", m, " disease sample n=", n, "iteration mm=", mm, "\n")
cat(" Number of log <0 ",logltzero,"\n\n");

cat(" Coverage1=", newcov1,"\n");
cat(" Dp  Lower bound 1=", Dplow, " Up bound 1=", Dpup, "\n")
cat(" Coverage length 1 =", Dplength,"\n")
cat(" Number of converge nlow1=", nlow, "  nup1=", nup,"\n\n")

cat(" Coverage2=", newcov2,"\n");
cat(" Dp  Lower bound 2=", Dp2low, " Up bound 2=", Dp2up, "\n")
cat(" Coverage length 2 =", Dplength2,"\n")
cat(" Number of converge nlow2=", nlow2, "  nup2=", nup2,"\n")

cat("#####################################\n");

sink();
}
```

4. Dermatoscope example

```
mm<-1000          # number of repetition

tt<-0.9
#tt<-0.95

realdata<-read.table("D:\\Suqin\\exam4.SSC",header=F,skip=3)
##get non-disease nx and disease number ny;
nx<-0;
for(i in 1:72) if(realdata[i,4]==0) nx<-nx+1;
ny<-0;
for(i in 1:72) if(realdata[i,4]==1) ny<-ny+1;
m<-nx;   #assign to non-disease group
n<-ny;   #assign to disease group

xx<-matrix(0,nx,2)
yy<-matrix(0,ny,2)

xi<-1;
yi<-1;
for (r1 in 1:72)
{
     if (realdata[r1,4]==0) { xx[xi,1]<-realdata[r1,2];xx[xi,2]<-realdata[r1,3]; xi<-xi+1}
     else {     yy[yi,1]<-realdata[r1,2]; yy[yi,2]<-realdata[r1,3];yi<-yi+1;}
}

x10<-xx[,1]           # the sample from the first nondiseased population
x20<-xx[,2]           # the sample from the second nondiseased population

y11<-yy[,1]           # the sample from the first diseased population
y21<-yy[,2]
```

{The following part is the same as normal distribution}