Discussiones Mathematicae
Probability and Statistics 36 (2016) 115–135
doi:10.7151/dmps.1186

GENERALIZED LEHMANN FAMILY FOR META ANALYSIS BASED UPON SUMMARY RECEIVER OPERATING CHARACTERISTIC CURVES

Savita Jain, Kanchan Jain

and

Suresh K. Sharma

Department of Statistics
Panjab University
Chandigarh – 160014, India

e-mail: bansal.savita29@gmail.com
jaink14@gmail.com
ssharma643@yahoo.co.in

Abstract

In this paper, we consider the modelling of Summary Receiver Operating Characteristic (SROC) curve used in meta-analysis of diagnostic studies. This is done through Generalized Lehmann model which relates the log-sensitivity and the log false positive rate across various studies. The nuisance parameters (study specific false positive rates) are eliminated through the use of profile likelihood. The estimation for the parameters of the Generalized Lehmann family has been carried out using adjusted profile likelihood. This model is extended further to accommodate unobserved heterogeneity by allowing the constant of proportionality to vary across studies.

Keywords: meta analysis, sensitivity, false positive rate, generalized Lehmann family, profile likelihood, adjusted profile likelihood, SROC.

2010 Mathematics Subject Classification: 62P10.

1. Introduction

Meta-analysis is a statistical methodology that combines or integrates the results of several independent clinical trials that are considered to be combinable by an

analyst. However, different studies may not be comparable as they use different cut-off values for the continuous and ordered categorical diagnostic tests. To cope up with this problem, the Summary Receiver Operating Characteristic (ROC) Curve is considered which summarizes the performance of a diagnostic test. This curve indicates the relationship between the true positive rate (TPR) and the false positive rate (FPR) of the test at various thresholds used to differentiate between diseased cases from non-diseased cases. More recently, the Summary Receiver Operating Characteristic (SROC) curve and the area under the curve (AUC) have been proposed to assess diagnostic accuracy in the context of meta-analysis (Ref. Krzanowski and Hand (2009)).

Meta analysis has many applications in the fields of medicine, education, social sciences, business, environmental and agriculture sciences. For example, in the field of medicine, different study results related to risk of a particular disease from various regions, may be combined. In education, meta analysis is useful for combining studies about coaching effectiveness to improve Scholastic Aptitude Test (SAT) scores. In social sciences, one may combine several studies of gender differences in separate categories of quantitative, verbal and visual spatial ability. For more applications of meta-analysis, one may refer to Hedges and Olkin (2014), Cooper and Hedges (1994), Hasselblad and Hedges (1995), Irwig *et al.* (1995), Swets (2014), Sutton *et al.* (2000), Egger *et al.* (2001), Schulze *et al.* (2003) and Krzanowski and Hand (2009).

If $D = 0(1)$ means absence (presence) of a disease and $T = 0(1)$ if the test result is negative(positive), then the specificity and sensitivity are defined as

$$(1 - u) = P(T = 0|D = 0)$$

and

$$p = P(T = 1|D = 1).$$

Here the false positive rate is given by u.

Suppose there are k studies under consideration. Let x_i be the number of false positives out of n_i healthy individuals and y_i be the number of true positives out of m_i diseased individuals for $i = 1, \ldots, k$. For i^{th} study, the estimators of u_i and p_i are given by

$$\widehat{u_i} = \frac{x_i}{n_i} \text{ and } \widehat{p_i} = \frac{y_i}{m_i}.$$

Summary Receiver Operating Characteristic (SROC) curves are used to cope with different cut off values and comparatibility problem for sensitivity and specificity in case of independent samples and different diagnostic studies. SROC curve has been recommended to represent the performance of a diagnostic test and is preferred to Youden index (Ref. Youden (1950)) or the diagnostic odds ratio (Ref. Glas *et al.* (2003)). It is intended to represent the relationship between

TPR and FPR across studies, recognizing the fact that they may have used different thresholds.

For the continuous test T with potential value t, SROC curves are plotted by using the pair $(u(t), p(t))$ where

$$u(t) = P(T \geq t|D = 0)$$

and

$$p(t) = P(T \geq t|D = 1).$$

For k possible unknown cut-off values $t_1, \ldots, t_k$, the pairs $(u(t_i), p(t_i))$ can be estimated by

$$(\widehat{u_i}, \widehat{p_i}) = \left(\frac{x_i}{n_i}, \frac{y_i}{m_i}\right) \text{ for } i = 1, \ldots, k.$$

where we write $u_i = u(t_i)$ and $p_i = p(t_i)$.

Holling *et al.* (2012) proposed Lehmann model for analysis of SROC curves. They studied the elimination of the nuisance parameters through profile likelihood which led to a proper Gaussian likelihood after adjustment. We generalise this model by introducing an additional parameter and explore estimation of parameters through adjusted profile likelihood (APL).

In Section 2, we propose the Generalized Lehmann model for the analysis of SROC curves. In this model, the nuisance parameter FPR or (1-specificity) is eliminated by means of profile likelihood (PL) which is discussed in Section 3. In Section 4, the adjusted profile likelihood is derived. Section 5 discusses the problem of heterogeneity. Simulations have been carried out for estimation of unknown parameters in Section 6. A real life data set is analysed in Section 7. This section also consists of forest, crosshair and ROCellipse plots for checking heterogeneity in the data.

2. The Generalized Lehmann model

We propose a Generalized Lehmann family which relates the sensitivity and false positive rate as

$$p(t) = [u(t)]^{\theta^\alpha}, \qquad \theta > 0 \quad \text{and} \quad \alpha > 0. \tag{2.1}$$

If $u(t) \in [0, 1]$, then $p(t) \in [0, 1]$ for $\theta, \alpha > 0$. The introduction of factor α leads to more flexibility to Lehmann model since higher order relationships can also be established. The parameters θ and α represent the diagnostic accuracy of the model. For (2.1), the ratio of $logp(t)$ and $logu(t)$ is θ^α, which is constant, a property satisfied by proportional hazard model (PHM) (Ref. Breslow (1975)).

The ROC curves for this model for different values of θ and α are shown in Figure 1.

From Figure 1, one can easily see that

- for fixed value of θ and varying α, the diagnostic accuracy is quite good.
- the diagnostic accuracy is not good when α is fixed but θ varies. This is evident as the ROC curves are below the diagonal line.

It may be noted that the two diagnostic tests represented by different values of θ and α can be easily compared. The area under the curve (AUC) for this model is given by

$$AUC = \int_0^1 u^{\theta^\alpha}\, du = \frac{1}{1+\theta^\alpha} \quad \text{for } \theta, \alpha > 0.$$

3. Profile likelihood

In the following discussion, we consider the profile likelihood (PL) which helps in estimation in the presence of a nuisance parameter. This method reduces the infinite dimensional estimation problem to a finite dimensional one. It is a widely used method to eliminate a nuisance parameter and also possesses invariance property.

For i^{th} study, let X_i and Y_i be Binomial random variables with parameters (n_i, u_i) and (m_i, p_i) respectively. For sake of simplicity, we suppress the index i and the product-binomial likelihood can be written as

$$\binom{n}{x} u^x (1-u)^{n-x} \binom{m}{y} p^y (1-p)^{m-y},$$

where $x = 0, 1, \ldots, n$ and $y = 0, 1, \ldots, m$, $0 < p < 1$ and $0 < u < 1$.

Using Delta Method (Ref. Casella and Berger (2002)), we have that for large n and m, $\log X$ and $\log Y$ follow Normal distribution with

$$E[\log X] = \log nu;$$
$$V[\log X] = nu(1-u)\left[\frac{d}{d(nu)}\log(nu)\right]^2 = \frac{1-u}{nu}$$

and

$$E[\log Y] = \log mp;$$
$$V[\log Y] = mp(1-p)\left[\frac{d}{d(mp)}\log(mp)\right]^2 = \frac{1-p}{mp}.$$

Using estimators of u as $\frac{x}{n}$ and p as $\frac{y}{m}$, we can write for large n and m

$$\widehat{V}[\log X] = \frac{1}{x} - \frac{1}{n} \quad \text{and} \quad \widehat{V}[\log Y] = \frac{1}{y} - \frac{1}{m}.$$

Hence writing $\widehat{V}[\log X]$ as t^2 and $\widehat{V}[\log Y]$ as s^2, the joint distribution of $\log X_i$ and $\log Y_i$ can be written as

$$\frac{1}{\sqrt{2\pi t^2}} \exp\left[\frac{-1}{2t^2}[logx - log(nu)]^2\right] \frac{1}{\sqrt{2\pi s^2}} \exp\left[\frac{-1}{2s^2}[logy - log(mp)]^2\right]. \tag{3.1}$$

If x_i and y_i are assumed to be positive, the estimated variances for the log-proportions, that is, $\log\left(\frac{x_i}{n_i}\right)$ and $\log\left(\frac{y_i}{m_i}\right)$ are given by

$$\widehat{V}ar(\log \widehat{u}_i) = t_i^2 = \frac{1}{x_i} - \frac{1}{n_i} \text{ and } \widehat{V}ar(\log \widehat{p}_i) = s_i^2 = \frac{1}{y_i} - \frac{1}{m_i}, \quad i = 1, 2, \ldots, k.$$

Let $w_i = log(x_i) - log(n_i)$ and $z_i = log(y_i) - log(m_i)$ be the log-true positive and the log-false positive rates. Since in diagnostic studies, the no. of observations in each study is not small, hence the normal approximation is valid. The index i will be suppressed for convenience of notation in the sequel.

Using (3.1), the relevant part of the log-likelihood is

$$\frac{-1}{2t^2}(logx - logn - logu)^2 - \frac{1}{2s^2}(logy - logm - logp)^2. \tag{3.2}$$

If $u' = logu$, then (3.2) can be written as

$$l(\theta, \alpha, u^{'}) = \frac{-1}{2t^2}(w - u^{'})^2 - \frac{1}{2s^2}(z - \theta^{\alpha}u^{'})^2. \tag{3.3}$$

For fixed θ and α, differentiating $l(\theta, \alpha, u^{'})$ w.r.t. $u^{'}$ and equating to zero leads to

$$\frac{1}{t^2}(w - u^{'}) + \frac{1}{s^2}(z - \theta^{\alpha}u^{'})(\theta^{\alpha}) = 0$$

$$\Rightarrow \widehat{u}^{'} = \frac{zt^2\theta^{\alpha} + ws^2}{t^2\theta^{2\alpha} + s^2}.$$

Substituing $\widehat{u}^{'}$ in (3.3), the profile log-likelihood becomes

$$l(\theta, \alpha) = -\frac{1}{2t^2}\left[w - \left(\frac{zt^2\theta^{\alpha} + ws^2}{t^2\theta^{2\alpha} + s^2}\right)\right]^2 - \frac{1}{2s^2}\left[z - \theta^{\alpha}\left(\frac{zt^2\theta^{\alpha} + ws^2}{t^2\theta^{2\alpha} + s^2}\right)\right]^2.$$

Simplifying, we get

$$l(\theta,\alpha)=\frac{-(z-\theta^{\alpha}w)^2}{2(t^2\theta^{2\alpha}+s^2)}.$$

In the next subsection, we explore properties of Profile log-likelihood.

3.1. Properties of profile log-likelihood

1) Invariance Property:

Generalized Lehmann model can be written as $\log p=\theta^{\alpha}\log u$ or $\log u=\frac{1}{\theta^{\alpha}}\log p$. In the first form, log-sensitivity can be regressed on the log-false positive rate, whereas in the latter, the log-false positive rate is regressed on the log-sensitivity. Both these regression problems are known to have different solutions. Now the profile maximum likelihood is invariant to the choice of nuisance parameter. For example, if u or p is chosen to be nuisance parameter, then

$$l(\theta,\alpha,\widehat{u})=l(\theta,\alpha,\widehat{p}).$$

Since the labelling of axis is arbitrary in the ROC diagram, hence, irrespective of the choice of model for analysis, the profile log-likelihood is suitable for the inference. This happens because the choice of nuisance parameter (sensitivity or false positive rate) will ultimately not affect the inference about the parameter of interest which gives the invariance property of log-likelihood.

2) Approximation to Gaussian log-likelihood:

Now we show that $l(\theta,\alpha)$ is almost a Gaussian log-likelihood. Since $p=u^{\theta^{\alpha}}\Rightarrow \log p=\theta^{\alpha}\log u\Rightarrow\log\frac{y}{m}=\theta^{\alpha}\log\frac{x}{n}$. Thus $z=\theta^{\alpha}\log\frac{x}{n}$ and $E(z)=\theta^{\alpha}[E(\log x)-E(\log n)]=\theta^{\alpha}[\log(nu)-\log(n)]=\theta^{\alpha}log(u)=\theta^{\alpha}w$.

We also have $V(z)=V\left(\log\left(\frac{y}{m}\right)\right)=V(\log y)=s^2$ and
$V(\theta^{\alpha}w)=\theta^{2\alpha}V\left(\log\left(\frac{x}{n}\right)\right)=\theta^{2\alpha}V(\log x)=\theta^{2\alpha}t^2$ so that
$V(z-\theta^{\alpha}w)=s^2+\theta^{2\alpha}t^2=\sigma^2(\theta,\alpha)(say)$.

This implies that $(z-\theta^{\alpha}w)\sim N[0,\sigma^2(\theta,\alpha)]$ which shows $l(\theta,\alpha)$ is almost a Gaussian log-likelihood and

$$l(\theta,\alpha)=\frac{-(z-\theta^{\alpha}w)^2}{2(s^2+\theta^{2\alpha}t^2)}=\frac{-(z-\theta^{\alpha}w)^2}{2(\sigma^2(\theta,\alpha))}. \tag{3.4}$$

If $L(\theta,\alpha)$ denotes the proper log-likelihood, then

$$L(\theta,\alpha)=\frac{-\log\sigma^2(\theta,\alpha)}{2}-\frac{(z-\theta^{\alpha}w)^2}{2(\sigma^2(\theta,\alpha))}.$$

The profile log-likelihood $l(\theta,\alpha)$ differs from $L(\theta,\alpha)$ only by $-\frac{\log\sigma^2(\theta,\alpha)}{2}$ and this shows that $l(\theta,\alpha)$ is not a proper log-likelihood.

In particular, first and second-order properties don't hold necessarily and the curvature of the profile likelihood does not provide estimate of the variance. Since the profile likelihood considers the estimated nuisance parameter as a true parameter value, it can underestimate the variance of the parameter of interest (Ref. Barndorff-Nielsen (1983), Cox and Reid (1987), Lee *et al.* (2006)). In addition, the profile log-likelihood $l(\theta,\alpha)$ fails in case of heterogeneity if further variance components are incorporated. Because of these shortcomings, we use the concept of adjusted profile likelihood.

4. The adjusted profile likelihood

For fixed θ and α, the observed Fisher information $\widehat{I}(u)$ evaluated at $\widehat{u}$ is the modified or adjusted profile likelihood and is given by

$$\begin{aligned}\widehat{I}(\widehat{u}) &= \frac{-\partial^2 l(\theta,\alpha,u^{'})}{\partial u^{'2}} = \frac{\partial^2}{\partial u^{'2}}\left[\frac{1}{2t^2}(w-u^{'})^2+\frac{1}{2s^2}(z-\theta^{\alpha}u^{'})^2\right]\\ &= \frac{\theta^{2\alpha}t^2+s^2}{s^2t^2} = \frac{\sigma^2(\theta,\alpha)}{s^2t^2}.\end{aligned}$$

$$\begin{aligned}L(\theta,\alpha) &= \frac{-1}{2}\log(\widehat{I}(\widehat{u}))+l(\theta,\alpha)+constant\\ &= \frac{-\log\sigma^2(\theta,\alpha)}{2}+\frac{log(s^2t^2)}{2}+l(\theta,\alpha)+constant\end{aligned} \tag{4.1}$$

where constant is independent of θ and α. $L(\theta,\alpha)$ becomes a proper log-likelihood if the constant term is chosen to be $\left(-\frac{\log(s^2t^2)}{2}\right)$ which is independent of θ and α.

For k studies, using (3.4) and (4.1), the full-sample adjusted profile log-likelihood can be written as $L(\theta,\alpha) = -\sum_{i=1}^{k}\frac{\log\sigma_i^2(\theta,\alpha)}{2}-\frac{1}{2}\sum_{i=1}^{k}\frac{(z_i-\theta^{\alpha}w_i)^2}{\sigma_i^2(\theta,\alpha)}$ where $\sigma_i^2(\theta,\alpha)=\theta^{2\alpha}t_i^2+s_i^2$. This form implies that $Z_i\sim N(\theta^{\alpha}w_i,\sigma_i^2(\theta,\alpha))$. For formulation of the model without w_i, we write $\frac{Z_i}{w_i}=\theta^{\alpha}+\epsilon_i$, where ϵ_i are the error terms following $N\left[0,\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}\right]$. Since $E\left(\frac{Z_i}{w_i}\right)=\theta^{\alpha}$ and $V\left(\frac{Z_i}{w_i}\right)=\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}$, hence the associated log-likelihood takes the form

$$L(\theta,\alpha)=\frac{-1}{2}\left[\sum_{i=1}^{k}\log\left(\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}\right)+\sum_{i=1}^{k}\frac{(z_i/w_i-\theta^{\alpha})^2}{\sigma_i^2(\theta,\alpha)/w_i^2}\right].$$

This gives

$$\frac{\partial L(\theta,\alpha)}{\partial\theta} = -\sum_i \frac{(2\alpha t_i^2\theta^{2\alpha-1})}{2(\sigma_i^2(\theta,\alpha))} + \sum_i \frac{\alpha(z_i-\theta^\alpha w_i)w_i\theta^{\alpha-1}}{\sigma_i^2(\theta,\alpha)} + \frac{1}{2}\sum_i \frac{(z_i-\theta^\alpha w_i)^2(2\alpha t_i^2\theta^{2\alpha-1})}{(\sigma_i^2(\theta,\alpha))^2}.$$

Putting $v_i = \frac{1}{\sigma_i^2(\theta,\alpha)}$,

$$\begin{aligned}\frac{\partial L(\theta,\alpha)}{\partial\theta} &= \alpha\theta^{\alpha-1}\left[-\sum_{i=1}^k t_i^2\theta^\alpha v_i + \sum_{i=1}^k w_i v_i(z_i-\theta^\alpha w_i) + \sum_{i=1}^k (z_i-\theta^\alpha w_i)^2 t_i^2\theta^\alpha v_i^2\right] \\ &= \alpha\theta^{\alpha-1}\left[\theta^\alpha\sum_{i=1}^k v_i t_i^2\left[(z_i-\theta^\alpha w_i)^2 v_i - 1\right] + \sum_{i=1}^k w_i v_i(z_i-\theta^\alpha w_i)\right]. \qquad (4.2)\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{\partial L(\theta,\alpha)}{\partial\alpha} &= -\sum_{i=1}^k \frac{(t_i^2\theta^{2\alpha}\log\theta)}{2(\sigma_i^2(\theta,\alpha))} - \sum_{i=1}^k \frac{(z_i-\theta^\alpha w_i)(-w_i\theta^\alpha\log\theta)}{\sigma_i^2(\theta,\alpha)} \\ &\quad + \sum_{i=1}^k \frac{(z_i-\theta^\alpha w_i)^2(t_i^2\theta^{2\alpha}\log\theta)}{2(\sigma_i^2(\theta,\alpha))^2} \\ &= \frac{\theta^\alpha\log\theta}{2}\left[-\sum_{i=1}^k v_i t_i^2\theta^\alpha + 2\sum_{i=1}^k w_i v_i(z_i-\theta^\alpha w_i)\right. \\ &\quad \left. + \sum_{i=1}^k t_i^2 v_i(z_i-\theta^\alpha w_i)^2\theta^\alpha v_i\right] \\ &= \frac{\theta^\alpha\log\theta}{2}\left[\theta^\alpha\sum_{i=1}^k v_i t_i^2[(z_i-\theta^\alpha w_i)^2 v_i - 1] + 2\sum_{i=1}^k w_i v_i(z_i-\theta^\alpha w_i)\right]. \qquad (4.3)\end{aligned}$$

Using $E[z_i-\theta^\alpha w_i] = 0$ and $E[z_i-\theta^\alpha w_i]^2 = \sigma_i^2(\theta,\alpha)$, we get $E\left[\frac{\partial L(\theta,\alpha)}{\partial\theta}\right] = 0$ and $E\left[\frac{\partial L(\theta,\alpha)}{\partial\alpha}\right] = 0$. Hence the expected values of the scores of the adjusted profile log-likelihood satisfy the conventional first-order property. This also gives that $\frac{z_i-\theta^\alpha w_i}{\sigma_i(\theta,\alpha)} = \frac{z_i/w_i-\theta^\alpha}{\sigma_i(\theta,\alpha)/w_i}$ is approximately a standard normal variate and $\chi^2_{k-2} = \sum_{i=1}^k \frac{(\widehat{\theta}_i^\alpha-\widehat{\theta}^\alpha)^2}{\sigma_i^2(\widehat{\theta},\widehat{\alpha})/w_i^2}$ has an approximate χ^2 distribution with $(k-2)$ df. This chi-square statistic is used for testing goodness of fit of the model.

For finding maximum likelihood estimates (MLEs) of θ and α, we equate (4.2) and (4.3) to zero and solve them. Since this results into two non linear

equations, hence numerical approximation is used for estimating θ and α. For this, we use OPTIM function in R software. This will provide the maximum likelihood estimates using the adjusted profile likelihood called the adjusted profile maximum likelihood estimates (APMLEs).

5. HETEROGENEITY

Different studies have different false positive rates but identical proportionality parameters θ and α. However, θ and α may also vary from study to study. For the diagnostic accuracy of test in case of heterogenity, $\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}$ is replaced by $\left(\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}+\tau^2\right)$ in $L(\theta,\alpha)$, where τ^2 is the appropriate random effect variance component parameter. This is accomplished by extending the fixed effect model using a further random effect δ_i which helps us in writing

$$\frac{Z_i}{w_i}=\theta^\alpha+\delta_i+\epsilon_i\sim N\left(\theta^\alpha,\left(\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}+\tau^2\right)\right),$$

where δ_i is the random effect independent of ϵ_i with $E(\delta_i)=0$, $Var(\delta_i)=\tau^2$ and ϵ_i are error terms following $N\Big(0,\left(\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}+\tau^2\right)\Big)$.

The full-sample adjusted profile log-likelihood with random effect is written as

$$L(\theta,\alpha,\tau^2)=-\sum_i\frac{\log(\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}+\tau^2)}{2}-\sum_i\frac{(\widehat{\theta}_i^\alpha-\theta^\alpha)^2}{2(\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}+\tau^2)}. \tag{5.1}$$

The score vector $\mathbf{U}=\left(\frac{\partial L}{\partial\theta},\frac{\partial L}{\partial\alpha},\frac{\partial L}{\partial\tau^2}\right)$, where using $r_i=\frac{1}{\left(\frac{\sigma_i^2(\theta,\alpha)}{w_i^2}+\tau^2\right)}$,

$$\begin{aligned}\frac{\partial L}{\partial\theta}&=\sum_i(\widehat{\theta}_i^\alpha-\theta^\alpha)r_iw_i(\alpha\theta^{\alpha-1})+\sum_i(\widehat{\theta}_i^\alpha-\theta^\alpha)^2\frac{(\alpha t_i^2\theta^{2\alpha-1}r_i^2)}{w_i^2}\\&\quad-\sum_i\frac{\alpha t_i^2\theta^{2\alpha-1}r_i}{w_i^2};\end{aligned} \tag{5.2}$$

$$\begin{aligned}\frac{\partial L}{\partial\alpha}&=-\sum_i\frac{r_it_i^2\theta^{2\alpha}\log\theta}{2w_i^2}+\sum_i\frac{(\widehat{\theta}_i^\alpha-\theta^\alpha)^2r_i^2t_i^2\theta^{2\alpha}\log\theta}{2w_i^2}\\&\quad+\sum_ir_iw_i(\widehat{\theta}_i^\alpha-\theta^\alpha)(\theta^\alpha\log\theta);\end{aligned} \tag{5.3}$$

and

$$\frac{\partial L}{\partial \tau^2} = -\sum_i \frac{1}{2(\frac{\sigma_i^2(\theta^\alpha)}{w_i^2} + \tau^2)} + \sum_i \frac{(\widehat{\theta}_i^\alpha - \theta^\alpha)^2}{2(\frac{\sigma_i^2(\theta^\alpha)}{w_i^2} + \tau^2)^2}$$

$$= \sum_i \frac{(\widehat{\theta}_i^\alpha - \theta^\alpha)^2 r_i^2}{2} - \sum_i \frac{r_i}{2}. \tag{5.4}$$

The score vector $\mathbf{U} = (\frac{\partial L}{\partial \theta}, \frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \tau^2})$ satisfies the first-order property $E(\mathbf{U}) = \mathbf{0}$ because for $\widehat{\theta}_i^\alpha \sim N(\theta^\alpha, \frac{1}{r_i})$,

$$E\left(\frac{\partial L}{\partial \theta}\right) = \sum_i \alpha(E(\widehat{\theta}_i^\alpha - \theta^\alpha)) r_i w_i \theta^{\alpha-1} + \sum_i (E(\widehat{\theta}_i^\alpha - \theta^\alpha)^2) \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i^2)}{w_i^2}$$

$$- \sum_i \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i)}{w_i^2}$$

$$= \sum_i \left(\alpha \times 0 \times r_i \times w_i \times \theta^{\alpha-1}\right) + \sum_i \frac{1}{r_i} \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i^2)}{w_i^2} - \sum_i \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i)}{w_i^2} = 0\,; \tag{5.5}$$

$$E\left(\frac{\partial L}{\partial \alpha}\right) = -\sum_i \frac{r_i t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2} + \sum_i (E(\widehat{\theta}_i^\alpha - \theta^\alpha)^2) \frac{r_i^2 t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2}$$

$$+ \sum_i r_i w_i (E(\widehat{\theta}_i^\alpha - \theta^\alpha)) \theta^\alpha \log\theta$$

$$= -\sum_i \frac{r_i t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2} + \sum_i \frac{r_i t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2} + \sum_i (r_i \times w_i \times 0 \times \theta^\alpha \times \log\theta) = 0\,; \tag{5.6}$$

$$E\left(\frac{\partial L}{\partial \tau^2}\right) = \sum_i \frac{(E(\widehat{\theta}_i^\alpha - \theta^\alpha)^2) r_i^2}{2} - \sum_i \frac{r_i}{2} = 0. \tag{5.7}$$

The adjusted χ^2 for testing goodness of fit in case of heterogeneity is

$$\chi^2_{het} = \sum_{i=1}^{k} \frac{(\widehat{\theta}_i^\alpha - \widehat{\theta}^\alpha)^2}{(\sigma_i^2(\widehat{\theta}, \widehat{\alpha})/w_i^2 + \widehat{\tau}^2)},$$

which follows chi-square distribution with $(k-3)$ degrees of freedom.

In the next section, we carry out simulations to find estimates of θ, α and also the heterogeneity parameter τ^2 by using numerical approximation and the non-linear equations given in Sections 4 and 5.

and

$$\frac{\partial L}{\partial \tau^2} = -\sum_i \frac{1}{2(\frac{\sigma_i^2(\theta^\alpha)}{w_i^2} + \tau^2)} + \sum_i \frac{(\widehat{\theta}_i^\alpha - \theta^\alpha)^2}{2(\frac{\sigma_i^2(\theta^\alpha)}{w_i^2} + \tau^2)^2}$$

$$= \sum_i \frac{(\widehat{\theta}_i^\alpha - \theta^\alpha)^2 r_i^2}{2} - \sum_i \frac{r_i}{2}. \tag{5.8}$$

The score vector $\mathbf{U} = (\frac{\partial L}{\partial \theta}, \frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \tau^2})$ satisfies the first-order property $E(\mathbf{U}) = \mathbf{0}$ because for $\widehat{\theta}_i^\alpha \sim N(\theta^\alpha, \frac{1}{r_i})$,

$$E\left(\frac{\partial L}{\partial \theta}\right) = \sum_i \alpha(E(\widehat{\theta}_i^\alpha - \theta^\alpha)) r_i \theta^{\alpha-1} + \sum_i (E(\widehat{\theta}_i^\alpha - \theta^\alpha)^2) \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i^2)}{w_i^2}$$

$$- \sum_i \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i^2)}{w_i^2}$$

$$= \sum_i \left(\alpha \times 0 \times r_i \times \theta^{\alpha-1}\right) + \sum_i \frac{1}{r_i} \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i^2)}{w_i^2} - \sum_i \frac{(\alpha t_i^2 \theta^{2\alpha-1} r_i)}{w_i^2} = 0\,; \tag{5.9}$$

$$E\left(\frac{\partial L}{\partial \alpha}\right) = -\sum_i \frac{r_i t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2} + \sum_i (E(\widehat{\theta}_i^\alpha - \theta^\alpha)^2) \frac{r_i^2 t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2}$$

$$+ \sum_i r_i (E(\widehat{\theta}_i^\alpha - \theta^\alpha)) \theta^\alpha \log\theta$$

$$= -\sum_i \frac{r_i t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2} + \sum_i \frac{r_i t_i^2 \theta^{2\alpha} \log\theta}{2w_i^2} + \sum_i (r_i \times 0 \times \theta^\alpha \times \log\theta) = 0\,; \tag{5.10}$$

$$E\left(\frac{\partial L}{\partial \tau^2}\right) = \sum_i \frac{(E(\widehat{\theta}_i^\alpha - \theta^\alpha)^2) r_i^2}{2} - \sum_i \frac{r_i}{2} = 0. \tag{5.11}$$

The adjusted χ^2 for testing goodness of fit in case of heterogeneity is

$$\chi^2_{het} = \sum_{i=1}^{k} \frac{(\widehat{\theta}_i^\alpha - \widehat{\theta}^\alpha)^2}{(\sigma_i^2(\widehat{\theta}, \widehat{\alpha})/w_i^2 + \widehat{\tau}^2)},$$

which follows chi-square distribution with $(k-3)$ degrees of freedom.

In the next section, we carry out simulations to find estimates of θ, α and also the heterogeneity parameter τ^2 by using numerical approximation and the non-linear equations given in Sections 4 and 5.

6. Simulations

A simulation study is carried out to find the adjusted profile maximum likelihood estimates (APMLE) of parameters for different no. of studies by taking $k = 5, 10, 15, 20$ and 50. n and m are generated from Poisson distributions with means 25 and 50, respectively. False positive rates $u_i^{'}s$ are sampled from a Uniform distribution on $(0, 1)$. Sensitivities $p_i^{'}s$ are calculated according to the Generalised Lehmann model given by (2). Finally, $x_i^{'}s$ are sampled from a Binomial distribution with parameters n_i and u_i whereas $y_i^{'}s$ are sampled from a Binomial distribution with parameters m_i and p_i for $i = 1, 2, \ldots, k$. The no. of repetitions is taken to be 10000. The used numerical method is L-BFGS-B in optim function in R.

6.1. Estimation in case of no heterogeneity

Tables 1–3 give the estimates and corresponding Root Mean Square Errors (RMSEs) of θ and α for k studies, We consider the following cases

1. Both θ and α are unknown (Table 1);
2. θ is unknown and α is known (Table 2);
3. α is unknown and θ is known (Table 3).

From the Tables 1–3, it can be concluded that as the number of studies increases, the RMSE of the estimates gets lower.

6.2. Estimation in case of heterogeneity

It is assumed that there is heterogeneity present in the data set and the corresponding parameter is τ^2. Tables 4–6 provide estimates of θ, α and τ^2 under different setups.

In Table 4, we give maximum likelihood estimates of θ, α and τ^2 and the corresponding Root Mean Square Errors (RMSEs) by using profile likelihood and adjusted profile likelihood for purpose of comparison. The initial values are taken to be $\theta = 0.1$, $\alpha = 0.4$ and $\tau^2 = 0.01$. PMLE gives the maximum likelihood estimate (MLE) using Profile likelihood function and APMLE is the MLE using adjusted profile likelihood function.

It is seen from Table 4 that the RMSEs are lower when APL function is used. This supports the advantange of using APL in place of PL function for finding MLEs. This is true for other setups as well and hence the Tables that follow, show only MLEs obtained through APL function.

Table 5 gives the ML estimates and corresponding RMSEs when one of the three parameters is known. Table 6 gives the ML estimates and corresponding RMSEs when two of the three parameters are unknown.

On the basis of values of RMSEs of the estimates given in Tables 5 and 6, it can be concluded that as more studies are included in the analysis, the estimates perform better.

7. Data analysis

7.1. Example 1

We consider the data set used by Doust *et al.* (2004) for the diagnostic accuracy of Brain Natriuretic Peptides (BNP) for Heart Failure. It is found that $\widehat{\theta} = 0.3153677$ and $\widehat{\alpha} = 2.1678286$. This corresponds to an AUC of 0.92, which indicates good diagnostic accuracy. Holling *et al.* (2012) found the AUC of Lehmann model as 0.83. So, our model performs better in case of diagnostic accuracy of BNP than the Lehmann model considered by Holling *et al.* (2012).

7.2. Example 2

Alcohol Use Disorder Identification Test (AUDIT) is a recommended instrument for screening all forms of unhealthy alcohol use (risky drinking, alcohol abuse, alcohol dependence). The full AUDIT consists of 10 items and has been extensively investigated in several settings and countries (Ref. Reinert and Allen (2002)). We consider a meta-analysis study in measuring the Diagnostic Accuracy of the AUDIT data provided by Kriston *et al.* (2008) and given in Table 7. The meta-analysis on AUDIT and alcohol disorders provides adjusted profile maximum likelihood estimates (APMLEs) as $\widehat{\theta} = 0.3952426$ and $\widehat{\alpha} = 1.7096843$. The estimate of area under the ROC, that is, AUROC is 0.83 which indicates good diagnostic accuracy. It is observed that value of χ^2 statistic is 25.26 (df = 12) which provides strong evidence of heterogeneity in the data set. In case of heterogeneity, $\widehat{\theta} = 0.0496030$, $\widehat{\alpha} = 0.5278279$ and $\widehat{\tau}^2 = 0.0010000$ for the given data. For the AUDIT data, we also provide Forest plots, Crosshair and ROCellipse plots.

Forest plots used for diagnostic test accuracy report the estimated sensitivity and specificity along with their confidence intervals. The plots are also known as coupled forest plots as they contain two graphical sections: one depicting sensitivity and the other specificity. The Forest plots for estimated sensitivity and specificity are given in Figure 2 and Figure 3.

The middle part shows the confidence interval for estimated sensitivities and specificities for different studies.

Two high level plots viz Crosshair plot (Ref. Phillips *et al.* (2010)) and ROCellipse plot can be obtained using **mada** function in R. Crosshair plot displays the individual studies in ROC space with paired confidence intervals representing sensitivity and specificity and allow for the results of meta-analysis to be overlaid on the plot. It can be used for comparison of the heterogeneity of two variables in the same plot. ROCellipse plot gives individual confidence regions (in the form of ellipses) for estimate from each of the studies. These regions show the uncertainty of the pair of sensitivity and false positive rate.

Figures 4 and 5 display the Crosshair and ROCellipse plots respectively for the AUDIT data.

Acknowledgments

The first author is grateful to Department of Science and Technology, Govt. of India for providing financial assistance for carrying out this work.

References

[1] O.E. Barndorff-Nielsen, *On a formula for the distribution of the maximum likelihood estimator*, Biometrika **70** (1983) 343–65. doi:10.1093/biomet/70.2.343

[2] N.E. Breslow, *Analysis of survival data under the proportional hazards model*, International Statistical Review / Revue Internationale de Statistique **43** (1975) 45–57.

[3] G. Casella and R.L. Berger, Statistical Inference, Cengage Learning, 2nd ed. 2002.

[4] H. Cooper and L. Hedges, The Handbook of Research Synthesis (New York, Russell Sage Foundation, 1994).

[5] D.R. Cox and N. Reid, *Parameter orthogonality and approximatic conditional inference*, Journal of the Royal Statistical Society, Ser. B (Methodological) **49** (1987) 1–39.

[6] J.A. Doust, P.P. Glasziou, E. Pietrzak and A.J. Dobson, *A system reviews of diagnostic accuracy of natriyretic peptides for heart failure*, Archive of Internal Medicine **164** (2004) 1978–84.

[7] M. Egger, G.D. Smith and D.G. Altman, Systematic Reviews in Health Care: Meta-Analysis in Context (London, BMJ Publishing Group, 2001). doi:10.1002/9780470693926

[8] A.S. Glas, J.G. Lijmer, M.H. Prins, G.J. Bonsel and P.M.M. Bossuyt, *The diagnostic odds ratio: a single indicator of test performance*, Journal of Clinical Epidemiology **56** (2003) 1129–35. doi:10.1016/S0895-4356(03)00177-X

[9] V. Hasselblad and L.V. Hedges, *Meta-analysis of screening and diagnostic tests*, Psychological Bulletin **117** (1995) 167–78. doi:10.1037/0033-2909.117.1.167

[10] L.V. Hedges and I. Olkin, Statistical Methods for Meta-Analysis (New York, Academic Press, 2014).

[11] H. Holling, W. Bohning and D. Bohning, *Meta-analysis of diagnostic studies based upon SROC-curves: a mixed model approach using the Lehmann family*, Statistical Modelling **12** (2012) 347–375. doi:10.1177/1471082X1201200403

[12] L. Irwig, P. Macaskill, P. Glasziou and M. Fahey, *Meta-analytic methods for diagnostic test accuracy*, Journal of Clinical Epidemiology **48** (1995) 119–30. doi:10.1016/0895-4356(94)00099-C

[13] L. Kriston, L. Holzel, A. Weiser, M.B. Berner and M. Harter, *Meta-analysis: are 3 questions enough to detect unhealthy alcohol use*?, Annals Internal Medicine **149** (2008) 879–88. doi:10.7326/0003-4819-149-12-200812160-00007

[14] W.J. Krzanowski and D.J. Hand, ROC Curves for Continuous Data, Boca Raton (FL) (CRC Press Taylor and Francis Group, Chapman and Hall, 2009).

[15] Y. Lee, J.A. Nelder and Y. Pawitan, Generalized linear models with random effects unified analysis via H-likelihood, Boca Raton (FL): (Chapman and Hall/CRC, Taylor and Francis Group, 2006),

[16] B. Phillips, L. Stewart and A. Sutton, *Cross hairs plots for diagnostic meta-analysis*, Research Synthesis Methods **1** (2010) 308–315. doi:10.1002/jrsm.26

[17] D.F. Reinert and J.P. Allen, *The alcohol use disorders identification test (AUDIT): a review or recent research*, Alcohol: Clinical and Experimental Research **26** (2002) 272–79.

[18] R. Schulze, H. Holling and D. Bohning, Meta-Analysis. New Developments and Applications in Medical and Social Sciences, Gottingen: Hogrefe, Huber, 1st ed. 2003.

[19] A.J. Sutton, K.R. Abrams, D.R. Jones, T.A. Sheldon and F. Song, Methods for Meta-Analysis in Medical Research (New York, Wiley, 2000).

[20] J.A. Swets, Signal Detection Theory and ROC Analysis in Psychology and Diagnostics Collected Papers (New York, Psychology Press, Taylor and Francis Group, 2014).

[21] D. Youden, *Index for rating diagnostic tests*, Cancer **3** (1950) 32–5. doi:10.1002/1097-0142(1950)3:1

Received 28 July 2016

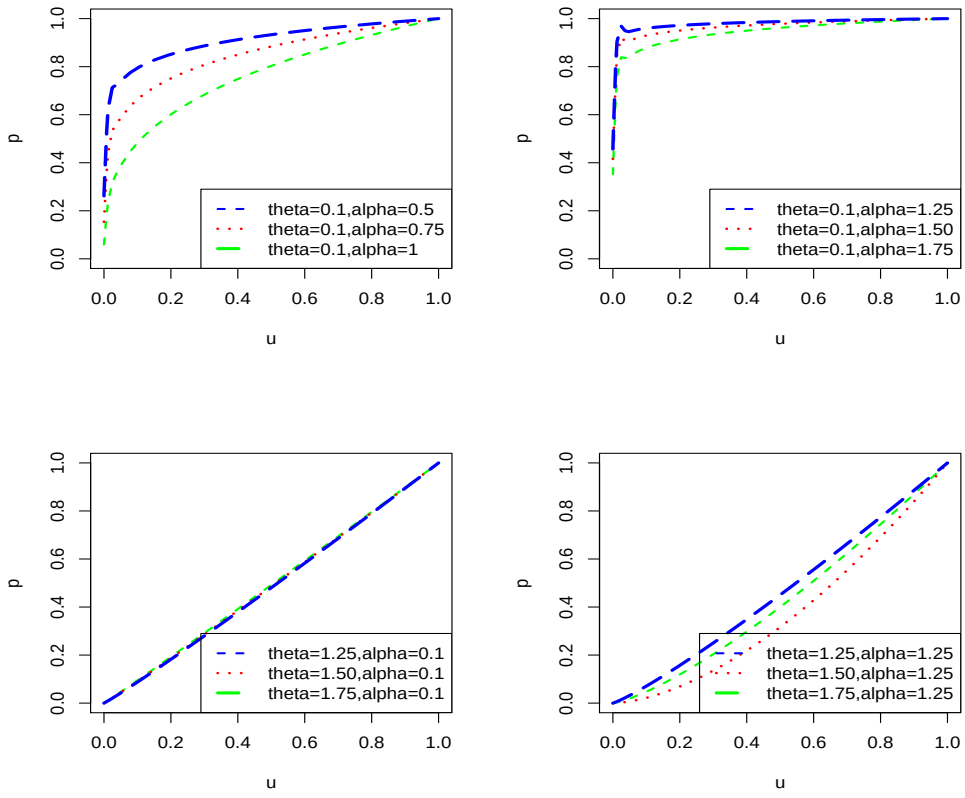

Figure 1. ROC curves for Generalized Lehmann family.

In Table 1, the initial values of θ and α are assumed to be 0.1 and 0.4 respectively.

Table 1. Estimates of θ and α and corresponding RMSEs.

k	$\widehat{\theta}$	$\widehat{\alpha}$	RMSE($\widehat{\theta}$)	RMSE($\widehat{\alpha}$)
5	0.11079	0.45122	0.04165	0.09994
10	0.10624	0.44575	0.02880	0.08176
15	0.10489	0.44196	0.02244	0.07020
20	0.10384	0.44105	0.01991	0.06503
50	0.10134	0.43648	0.00179	0.04614

In Table 2, $\alpha = 0.4$ is assumed to be known and initial value of θ is taken to be 0.1.

Table 2. Estimates of θ and corresponding RMSEs for known α.

k	$\widehat{\theta}$	RMSE($\widehat{\theta}$)
5	0.08899	0.04820
10	0.08571	0.03325
15	0.08462	0.02849
20	0.08392	0.02601
50	0.08263	0.02143

In the following Table, $\theta = 0.1$ is known and initial value of $\alpha = 0.4$.

Table 3. Estimates of α and corresponding RMSEs for known θ.

k	$\widehat{\alpha}$	RMSE ($\widehat{\alpha}$)
5	0.43803	0.09483
10	0.43458	0.06990
15	0.43529	0.06084
20	0.43378	0.05441
50	0.43468	0.04414

Table 4. Estimates of θ, α and τ^2 and corresponding RMSE.

k	5	10	15	20	50
APMLE ($\widehat{\theta}$)	0.10784	0.11063	0.11533	0.11423	0.10927
RMSE ($\widehat{\theta}$)	0.06278	0.05684	0.06326	0.04880	0.01783
PMLE ($\widehat{\theta}$)	0.56306	0.78348	0.50063	0.62242	0.94809
RMSE ($\widehat{\theta}$)	0.65128	0.74343	0.63812	0.66355	0.85227
APMLE ($\widehat{\alpha}$)	0.43230	0.42192	0.42144	0.42133	0.40932
RMSE ($\widehat{\alpha}$)	0.14490	0.09247	0.08084	0.07231	0.02063
PMLE ($\widehat{\alpha}$)	77.4904	84.5658	210.3141	173.0123	242.4411
RMSE ($\widehat{\alpha}$)	121.3141	108.5433	297.0212	270.6695	278.8208
APMLE ($\widehat{\tau}^2$)	0.00678	0.00858	0.00958	0.00982	0.00809
RMSE ($\widehat{\tau}^2$)	0.01675	0.01475	0.01463	0.01173	0.00466
PMLE ($\widehat{\tau}^2$)	432254.0	414663.6	153350.1	125864.0	446019.0
RMSE ($\widehat{\tau}^2$)	782146.6	652924.7	202888.1	234503.0	873458.4

Table 5. ML Estimates and RMSEs in case of one unknown parameter.

	θ, α known and τ^2 unknown Initial value of $\tau^2 = 0.01$		θ, τ^2 known and α unknown Initial value of $\alpha = 0.4$		α, τ^2 known and θ unknown Initial value of $\theta = 0.1$	
k	$\widehat{\tau}^2$	RMSE ($\widehat{\tau}^2$)	$\widehat{\alpha}$	RMSE ($\widehat{\alpha}$)	$\widehat{\theta}$	RMSE ($\widehat{\theta}$)
5	0.01578	0.03839	0.43362	0.11224	0.10489	0.08869
10	0.01349	0.02369	0.42183	0.08367	0.10645	0.08049
15	0.01275	0.01860	0.41718	0.07038	0.10737	0.07876
20	0.01188	0.01562	0.41425	0.06277	0.10660	0.06266
50	0.01083	0.00929	0.40919	0.04054	0.10355	0.02498

Table 6. ML Estimates and RMSEs in case of two unknown parameters.

k		5	10	15	20	50
θ known,	$\widehat{\alpha}$	0.43265	0.41957	0.41247	0.40934	0.40330
α and τ^2 unknown	RMSE ($\widehat{\alpha}$)	0.11537	0.07476	0.06041	0.05172	0.02851
(Initial value of $\alpha = 0.1$	$\widehat{\tau}^2$	0.00740	0.01003	0.01063	0.01028	0.01047
and $\tau^2 = 0.01$)	RMSE ($\widehat{\tau}^2$)	0.01771	0.01797	0.01698	0.01418	0.00916
α known,	$\widehat{\theta}$	0.09931	0.10102	0.10259	0.10400	0.10474
θ and τ^2 unknown	RMSE ($\widehat{\theta}$)	0.06213	0.04755	0.04102	0.03473	0.02087
(Initial value of $\theta = 0.4$	τ^2	0.00766	0.00954	0.00999	0.01023	0.01052
and $\tau^2 = 0.01$)	RMSE ($\widehat{\tau}^2$)	0.01965	0.01755	0.01509	0.01382	0.00922
τ^2 known,	$\widehat{\theta}$	0.14409	0.14942	0.13505	0.12969	0.11436
θ and α unknown	RMSE ($\widehat{\theta}$)	0.36472	0.34314	0.31863	0.29260	0.03565
(Initial value of $\theta = 0.4$	$\widehat{\alpha}$	0.45689	0.45001	0.43853	0.43007	0.42170
and $\alpha = 0.1$)	RMSE ($\widehat{\alpha}$)	0.17965	0.16151	0.11802	0.09132	0.05433

Table 7. Data for Meta-analysis of diagnostic accuracy of the AUDIT.

Alcohol disorder			No disorder		Total
Study i	y_i(TP)	$m_i - y_i$(FN)	$n_i - x_i$(TN)	x_i(FP)	$n_i + m_i$
1	48	7	738	101	894
2	138	39	1506	309	1992
3	24	5	173	31	233
4	37	2	227	127	393
5	137	12	936	234	1319
6	73	13	127	30	243
7	53	14	508	27	602
8	571	180	5707	496	6954
9	54	10	172	19	255
10	148	44	2687	672	3551
11	143	18	334	130	625
12	47	13	464	76	600
13	34	1	65	12	112
14	154	49	261	92	555

In the above table, m_i and n_i denote the number of persons with and without alcohol disorder for i^{th} study. x_i and y_i denote the number of false positives and true positives.

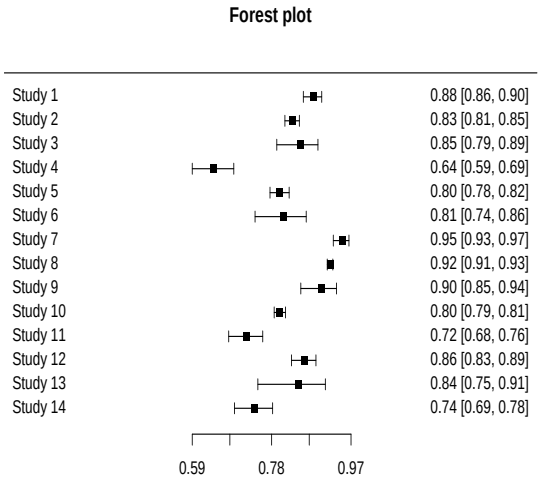

Figure 2. Forest plot for estimated sensitivity for AUDIT data.

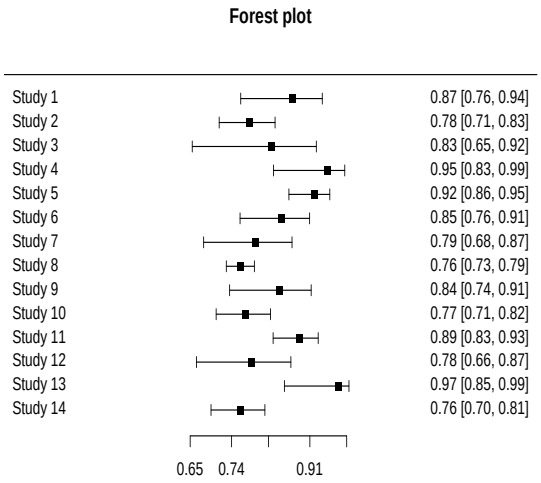

Figure 3. Forest plot for estimated specificity for AUDIT data.

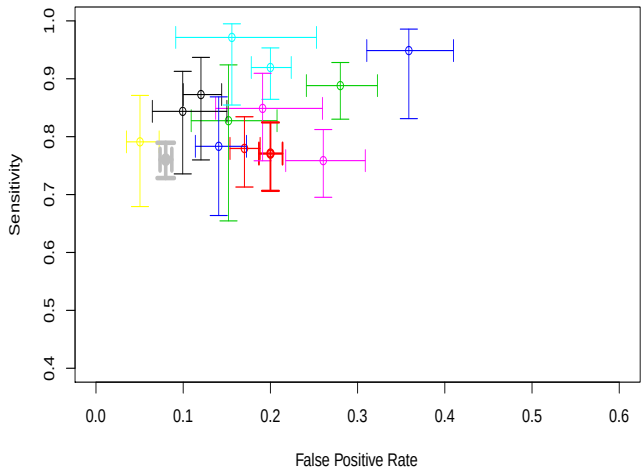

Figure 4. Crosshair Plot.

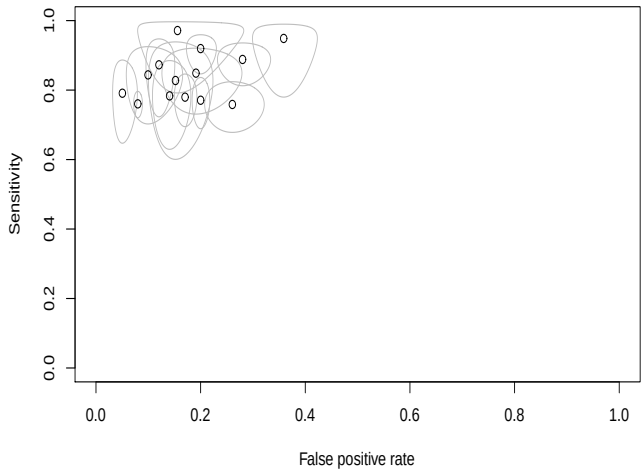

Figure 5. ROCellipse Plot.