KU LEUVEN

MASTER THESIS

Estimation of ROC curvesin the presence of measurement errors

Author:
Youhee KIL

Supervisor:

February 16, 2021

Declaration of Authorship

I, Youhee KIL, declare that this thesis titled, "Estimation of ROC curvesin the presence of measurement errors" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."

Dave Barry

KU LEUVEN

Abstract

Faculty Name
Department or School Name

Estimation of ROC curvesin the presence of measurement errors

by Youhee KIL

work in progress

This document lays out methodologies and application for estimation of Receiver Operating Characteristic (ROC) curves in the presence of measurement errors.

Acknowledgements

work in progress

Contents

List of Figures

List of Tables

List of Abbreviations

ROC	**R**eceiver **O**perating **C**haracteristic
PTP	**P**robability of a **T**rue **P**ositive
PFP	**P**robability of a **F**alse **P**ositive
AUC	**A**rea **U**nder the **C**urve

List of Symbols

a	distance	m
P	power	W ($\mathrm{J\,s^{-1}}$)
ω	angular frequency	rad

For/Dedicated to/To my…

Chapter 1

Introduction

1.1 Preliminary Considerations

The problem of estimation receiver operating characteristic (ROC) curves in the presence of measurement errors is concerned. A ROC curve enunciates the probability of a true positive (PTP) as a function of the probability of a false positive (PFP) for all possible values of the cut-off between cases and controls. The area under the curve (AUC), θ, can measure globally how well the separator variable distinguishes between cases and control. Therefore, AUC, θ, is widely used as a summary measure of diagnostic accuracy. We propose a smooth non-parametric ROC curve derived from Bernstein type polynomial estimates to obtain the ROC curve and AUC. The features of the Bernstein type polynomials can take the noted drawbacks of non-parametric ROC curve in hands. The aim of the paper is the estimation of the ROC curve and the AUC when predictors are measured with error. The classical measurement error is one of the most commonly used measurement error where the observed variables are measured with an additive error. The classical measurement error model states that $W_{ij} = X_i + U_{ij}$ where W_{ij} is an unbiased measure of X_i, and U_{ij} is mean-zero error which could be homoscedastic or heteroscedatic. For example, in context of disease study, this model can be interpreted as the observed dose (W_{ij}) equals the true dose (X_i) plus classical measurement error (U_{ij}). Carroll et al., 2006 indicated the effects of measurement error in covariates causes biases. There are many statistical methods aimed to correct for biases of estimation caused by measurement error. As pointed out by Coffin and Sukhatme, 1997, if the separator is measured with error, then the usual non-parametric estimate is also biased as well. As Carroll et al., 2006 states not taking account the measurement error will lead to serious consequences such as bias in the estimators of ROC curves and AUC in non-parametric cases and even in parametric cases. Hence, the measurement error has to be considered in order to derive well grounded inference by some bias-correct methods. The exact error distribution is frequently required for those bias-correct methods, however, as Bertrand, Van Keilegom, and Legrand, 2019 mentioned it is almost impossible to carry out the exact error distribution (variance of measurement error) when neither validation nor auxiliary data are available due to complexity.

1.2 Literature Reviews

There are many literatures dealing of the effect of random measurment error on ROC curves and AUC.

- Faraggi, 2000 studied confidence interval for the the effect of random measurement error on ROC curves and AUC with assumption of parametric normal model. Two different cases, with the effect of ignoring the measurement error on the cofidence interval for the area and Illustrations are performed

1.3 Objectives

1.4 Organization

Therefore, in this paper, we proposed one of the nonparametric methods for the estimation of the ROC curve and AUC with the measurement error variance by a Bernstein type polynomial.

The paper will be arranged as follows. In **Chapter 2**, we described basic knowledge about measurement error and ROC curves. In **Chapter 3**, the Bernstein likelihood with mixture of beta distribution approach is presented as a way of obtaining the maximum Bernstein likelihood estimates of ROC curve in presence of measurement error. Moreover, a method of choosing the ideal Bernstein polynomial model degree (m) based on the results of simulation is presented. The results of the estimation of the density AUC found on the estimated ROC curve with measurement errors are reported in **Chapter 4**. In **Chapter 5**, a comparison of the efficiency of the proposed maximum Bernstein likelihood ROC estimators with other nonparametric ROC estimators with SIMEX algorithm is performed. The proposed methods on estimation of ROC curve in presence of measurement error are applied to a real data set in **Chapter 6**.

Chapters

- Chapter 1: Introduction to the thesis topic
- Chapter 2: Background information and theory
- Chapter 3: (Laboratory) experimental setup
- Chapter 4: Details of experiment 1
- Chapter 5: Details of experiment 2
- Chapter 6: Discussion of the experimental results
- Chapter 7: Conclusion and future directions

This chapter layout is specialised for the experimental sciences, your discipline may be different.

Figures – this folder contains all figures for the thesis. These are the final images that will go into the thesis document.

Chapter 2

Basic Methods

2.1 Measurement Error

As we have discussed, there is increasing awareness that measurement error must take into account for accurate statistical analysis especially for epidemiological data. The measurement error can occur for many reasons, in this paper, classical measurement error will be focused. Historically, there are two major defining characteristics of the taxonomy of measurement error. The true (unobserved) values for covariate are denoted as X. The error subject to values are denoted as U, and its mean-zero which could be homoscedastic or heteroscedastic. Therefore, the classical measurement error model is :

$$W = X + U \tag{2.1}$$

where W is the observed (mis-measured) covariate, besides, X and U are assumed to be independent to each other. We further assume that the classical measurement error model with a Gaussian error of unknown variance τ^2, so that $U \sim N(0, \tau^2)$.

2.2 Receiver Operating Characteristics (ROC) curve

The ROC is needed in order to provide an assessment of the classifier over the whole range of t values of the threshold in a particular classification rule rather at just a single chosen one. The ROC must lie within the border of [(0,0), (0,1)] and [(0,1), (1,1)]. Therefore, in practice, the ROC curve will be a continuous curve lying in the upper triangle of the graph such as between two extremes points of the graph (Krzanowski and Hand, 2009).

Let test variable X denotes the continuous scores, F_1 and F_0 be distribution functions of X for the diseased group (1) and non-diseased group (0) (Faraggi, 2000). For the cut point (threshold value) t, Sensitivity/true positive rate denoted as $SE(t) = P(X > t|group = diseased(1)) = 1 - F_1(t)$ while specificity(s) is denoted as $SP(t) = P(X < t|group = nondiseased(0)) = F_0(t)$, and false positive rate is denoted as $1 - SP(t)$. The ROC curve is a plot of the sensitivity against 1-specificity $(1 - s)$, in other words, of the true positive fraction (TPF) against false positive fraction (FPF) using a threshold t. The equation of the ROC curve is

$$\begin{aligned} ROC(\cdot) &= (FPF(t), TPF(t)) \qquad t \in (-\inf, \inf) \\ ROC(s) &= 1 - F_1[F_0^{-1}(1 - s)] \qquad (0 \le s \le 1) \end{aligned} \tag{2.2}$$

The most common approach to estimate the ROC Curve is empirical approach where the data modeled samples from the relevant population without assumptions

(Krzanowski and Hand, 2009). Then the ROC curve is estimated based on the data sampled from the true positive rates ($SE(t) = 1 - F_1(t)$), specificity($SP(t) = F_0(t)$), however, the result empirical ROC estimator is jagged in appearance than underlying continuous form and not accurate enough for the smaller sample sizes. In this paper, we introduce one of nonparametric approaches to estimate the ROC curve in presence of measurement errors with Bernstein type polynomials. This approach is built on the work of Schwarz and Van Bellegem, 2010, Bertrand, Van Keilegom, and Legrand, 2019, and Guan, 2016. We firstly review several results on Bernstein distribution function estimators and measurement error, then present the precise definition of the maximum Bernstein likelihood density estimation of the measurement error of the ROC curve. Asymptotic properties and weak convergence of the Bernstein polynomial estimator was studied by the following studies.Babu, Canty, and Chaubey, 2002 implemented Bernstein polynomial estimation to the univariate distribution and density function, later Leblanc, 2010 showed the Bernstein estimator outperforms the classical empirical distribution regarding the asymptotic variance by proving point-wise asymptotic normality. Guan, 2016 proposed maximum Bernstein likelihood density estimation method. This Bernstein polynomial method relying on only one regularization parameter which is advantageous, it was possible because it is the same as using a mixture of Beta density function with known parameters. Besides, it approximates any continuous function with compact support. Guan, 2016 mentioned a flexible approximate parametric estimation is favoured over nonparametric estimation due to a lower rate of convergence to the true density function. The semi-nonparametric or mixture function density methods are not be able to used for here because of the compact support property of the ROC curve. A likelihood-based approach for estimating error variance was derived by Schwarz and Van Bellegem, 2010 with only weak assumption. The work of Guan, 2016 and Schwarz and Van Bellegem, 2010 carried out the study of Bertrand, Van Keilegom, and Legrand, 2019 about flexible parametric approach to classical measurement error variance estimation when Auxiliary data is not available.

In this paper, we consider the classical measurement error model with a Gaussian error of unknown variance τ^2, $U \sim N(0, \tau^2)$, from equation 2.1 where X is only assumed to be continuous and to have compact support. The objective is to obtain an estimate of τ which will then be available for use in a bias-correction method. Therefore, the key idea of this is to build the probability density function of W with only weak assumptions on the distribution of X. Therefore, we can assume that

$$X = \alpha S + \beta \tag{2.3}$$

where α and β are two unknown constants with $\alpha > 0$ for identifiability reasons, and S is a continuous random variable taking values in $[0, 1]$. Let f_X, f_S, f_U and f_W be the density functions of X, S, U, and W, respectively, under the addictive measurement error model(2.1), besides, value of X are in $[\beta, \alpha + \beta]$ according to the equation of (2.3). The classical additive measurement error model(2.1) with normally distributed error is $\frac{1}{\tau}\phi\left(\frac{w-x}{\tau}\right)$ where $\phi(\cdot)$ is a normal distribution with zero mean and unit variance. Carroll et al., 2006 noted that the density function f_W is the convolution of f_X and f_U.

$$\begin{aligned}
f_W(w) &= \frac{1}{\tau}\int_{x=\beta}^{\alpha+\beta} f_X(x)\phi\left(\frac{w-x}{\tau}\right)dx \\
&= \frac{1}{\tau}\int_{x=\beta}^{\alpha+\beta} f_S(u)\phi\left(\frac{w-x}{\tau}\right)du \qquad \textit{where } u = \frac{x-\beta}{\alpha}, du = \frac{1}{\alpha}dx \\
&= \frac{1}{\alpha\tau}\int_{x=\beta}^{\alpha+\beta} f_S\left(\frac{x-\beta}{\alpha}\right)\phi\left(\frac{w-x}{\tau}\right)dx
\end{aligned} \tag{2.4}$$

where the probability density function of S is denoted as f_s which is specified with minimal assumption.

Chapter 3

Methodology

3.1 Bernstein Polynomials

Non-parametric approach is getting more attention, in fact, it is very generically applicable and undoubtedly a useful tool for estimation of ROC curve. The first non-parametric ROC curves was introduced by Zou, Hall, and Shapiro, 1997 based on kernel density methods. Kernel estimation of density function is given by

$$\hat{f}_n(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \tag{3.1}$$

where K(·) is the kernel function in population and h is the bandwidth. The random sample drawn from f is denoted as $X_1, X_2, ..., X_n$ in here. Unfortunately, according to Peng and Zhou, 2004, the resulting estimators from these standard kernel estimation method have some drawbacks in our case where the support of the density function f to be estimated is compact support.

Bernstein polynomials is known as a very smooth estimator with acceptable behavior at the boundaries (Leblanc, 2010).

As Bernstein (1912) defined, the Bernstein polynomial of order m for a given function u is

$$B_{m,u}(x) = \sum_{k=0}^{m} u\left(\frac{k}{m}\right)\binom{m}{k} x^k (1-x)^{(m-k)} \tag{3.2}$$

if u is continuous in the closed interval [0,1] the Bernstein polynomial $B_{m,u}(x)$ converges uniformly to $u(x)$ (Bernstein, 1912)

When the underlying density f is known to be continuous and compactly supported, following Babu, Canty, and Chaubey, 2002, the Bernstein estimator of f of order $m > 0$ is

$$\hat{f}_{m,n}(x) = m\sum_{k=0}^{m-1}\left[F_n\left(\frac{k+1}{m}\right) - F_n\left(\frac{k}{m}\right)\right] P_{k,m-1}(x) \tag{3.3}$$

where $P_{k,m}(x) = \binom{m}{k} x^k (1-x)^{(m-k)}$ are binomial probabilities, as it is defined by Babu, Canty, and Chaubey, 2002. As Vitale, 1975 stated the form of the Bernstein polynomial estimate of $f(x)$ is one of a linear combination of beta densities with random coefficients based on the observation. Note that $\hat{f}_{m,n}(x)$ is the derivative of $\hat{F}_{m,n}(x)$ with respect to x. Hence, the polynomial of degree m with coefficients depending on the data , $\hat{F}_{m,n}(x)$, is defined as

$$\hat{F}_{m,n}(x) = \sum_{k=0}^{m} F_n\left(\frac{k}{m}\right) P_{k,m}(x) \tag{3.4}$$

where F_n denotes the empirical distribution function obtained from a random sample of size n. Hence, the Bernstein polynomial of order m of F is denoted as $B_m(x)$:

$$B_m(x) = \sum_{k=0}^{m} F\left(\frac{k}{m}\right) P_{k,m}(x) \tag{3.5}$$

It is clear to inspect $E[\hat{F}_{m,n}(x)] = B_m(x)$ for all $x \in [0, 1]$ and all $n \geq 1$. For all k, $F_{m,n}$ is non-decreasing and non-negative first derivative in x emanated from

$$F_{m,n}(0) = 0 = F(0) = B_m(0) \ \ and \ \ F_{m,n}(1) = 1 = F(1) = B_m(1) \tag{3.6}$$

the asymptotic properties of $\hat{F}_{m,n}$, $\hat{f}_{m,n}$, and other more details are examined in a paper by Babu, Canty, and Chaubey, 2002.

In view of discussed properties of Bernstein polynomial, the Bernstein polynomial of order $m > 0$ for the f_s can be defined as

$$B_m(s) = \sum_{k=0}^{m} f\left(\frac{k}{m}\right) P_{k,m}(s) \tag{3.7}$$

where $P_{k,m}(x) = \binom{m}{k} x^k (1-x)^{(m-k)}$, for $k = 0, ..., m$. The theorem presented by Lorentz, 2013, any continuous f is defined on [0,1],

$$\lim_{m \to +\infty} B_m(s) = f(s) \tag{3.8}$$

As Guan, 2016 mentioned one of advantages of the Bernstein polynomial method is approximating any continuous function with compact support. These features of the Bernstein polynomial method allowed us to apply maximum Bernstein likelihood density estimation approach to the estimation of τ and the approximation of f_s.

The approximation of f_s was introduced by Bertrand, Van Keilegom, and Legrand, 2019 and used by Guan, 2016 which is integrated with a mixture of the beta distribution. The equation of the approximating f_s is

$$\begin{aligned}
\tilde{f}_{s,m}(s; \bar{\theta}_m) &= \sum_{k=0}^{m} f_s\left(\frac{k}{m}\right) p_{k,m}(s) \\
&= \sum_{k=0}^{m} f_s\left(\frac{k}{m}\right) \binom{m}{k} s^k (1-s)^{(m-k)} \\
&= \sum_{k=0}^{m} \theta_{k,m} \binom{m}{k} s^k (1-s)^{(m-k)} \\
&= \sum_{k=0}^{m} f_s\left(\frac{k}{m}\right) \binom{m}{k} \frac{k!(m-k)!}{(m+1)!} s^k (1-s)^{(m-k)} \\
&= \frac{1}{m+1} \sum_{k=0}^{m} f_s\left(\frac{k}{m}\right) s^k (1-s)^{(m-k)}
\end{aligned} \tag{3.9}$$

where $\bar{\theta} = (\theta_{0,m}, ..., \theta_{m,m})$ and $s \in [0, 1]$. The approximation of f_s is done by a mixture of $Beta(k+1, m-k+1)$ densities with known parameters. The probability density function of X , $f_x(\cdot;\alpha,\beta)$, from the equation (2.3) can be approximated by implementing the context of density estimation of equation (3.9).

$$\begin{aligned} X &= \alpha S + \beta \\ S &= \frac{X-\beta}{\alpha} \end{aligned} \tag{3.10}$$

The probability density function of X can be approximated with a $beta_{\alpha,\beta}(\cdot)$ which is probability density function of a Beta-distributed random variables based on the parameter α, β where $x \in [\beta, \alpha+\beta]$.

$$\begin{aligned} \tilde{f}_{x,m}(x;\alpha,\beta,\bar{\theta}_m) &= \frac{1}{(\alpha+\beta)-\beta}\tilde{f}_{s,m}\left(\frac{X-\beta}{\alpha};\bar{\theta}_m\right) \\ &= \frac{1}{\alpha}\sum_{k=0}^{m}\theta_{k,m}\binom{m}{k}\left(\frac{X-\beta}{\alpha}\right)^k\left(1-\left(\frac{X-\beta}{\alpha}\right)\right)^{(m-k)} \\ &= \frac{1}{\alpha}\sum_{k=0}^{m}\theta_{k,m}beta_{(k+1,m-k+1)}\left(\frac{X-\beta}{\alpha}\right) \end{aligned} \tag{3.11}$$

The density of the classical measurement error model (2.1) can be approximated and applied on the equation (2.4) based on the result of the approximated probability density function of X denoted as $\tilde{f}_{x,m}(x;\alpha,\beta,\bar{\theta}_m)$.

$$\begin{aligned} \tilde{f}_{w,m}(w;\tau,\alpha,\beta,\bar{\theta}_m) &= \frac{1}{\tau}\int f_{x,m}(x;\alpha,\beta,\bar{\theta}_m)\phi\left(\frac{w-x}{\tau}\right)dx \\ &= \frac{1}{\tau}\int\left\{\frac{1}{\alpha}\sum_{k=0}^{m}\theta_{k,m}beta_{(k+1,m-k+1)}\left(\frac{x-\beta}{\alpha}\right)\right\}\phi\left(\frac{w-x}{\tau}\right)dx \\ &= \frac{1}{\alpha\tau}\sum_{k=0}^{m}\theta_{k,m}\int beta_{(k+1,m-k+1)}\left(\frac{x-\beta}{\alpha}\right)\phi\left(\frac{w-x}{\tau}\right)dx \end{aligned} \tag{3.12}$$

The theorem 3.8 by Lorentz, 2013 is also applied to $\tilde{f}_{w,m}(w;\tau,\alpha,\beta,\bar{\theta}_m)$ since f_s is continuous. Hence, we apply the Bernstein log-likelihood fo the set of unknown parameters $\tau,\alpha,\beta,\bar{\theta}_m$ in following sections.

3.1.1 The Bernstein likelihood

The Bernstein likelihood was defined by Guan, 2016, the maximiser the set of estimated parameters of log-likelihood is called MBLE. In this case, the Bernstein log-likelihood function of the set of parameters $(\tau,\alpha,\beta,\bar{\theta})$ given the observed data is

$$\begin{aligned} l_n(\tau,\alpha,\beta,\bar{\theta}_m) &= \sum_{i=1}^{n} log f_{w,m}(w;\tau,\alpha,\beta,\bar{\theta}_m) \\ &= \sum_{i=1}^{n} log\left[\frac{1}{\alpha\tau}\sum_{k=0}^{m}\theta_{k,m}\int beta_{(k+1,m-k+1)}\left(\frac{x-\beta}{\alpha}\right)\phi\left(\frac{W_i-x}{\tau}\right)dx\right]. \end{aligned} \tag{3.13}$$

where a collection of sample of W are independent and identically distributed, $W_1, W_2, ..., W_n \sim W$. Therefore, according to the Guan, 2016, the maximizer $\hat{\tau}, \hat{\alpha}, \hat{\beta}, \hat{\theta}$ of $l(\tau, \alpha, \beta, \bar{\theta_m})$ is called the MBLE of $\tau, \alpha, \beta, \bar{\theta}_m$ and the MBLEs $\hat{f}_w(w) = f_w(w; \hat{\tau}_m, \hat{\alpha}_m, \hat{\beta}_m, \hat{\theta}_m)$ and $\hat{F}_w(w) = F_w(w; \hat{\tau}_m, \hat{\alpha}_m, \hat{\beta}_m, \hat{\theta}_m)$ of $f(w)$ and $F(w)$ are called "the Bernstein probability density function" and "the Bernstein cumulative distribution function" respectively.

Estimation of ROC curve by using Maximum Bernstein likelihood density estimation

In this section we present the Maximum Bernstein likelihood density estimation procedure for ROC curve in presence of measurement error given in equation 2.2 above. As we supposed that diagnostic test results $x_1^1, x_2^1, ...x_m^1$ and $x_1^0, x_2^0, ..., x_n^0$ denoted as X_1 and X_0 are from the diseased and non-diseased population having cumulative distribution function F_1 and F_0 respectively.

The proposed refining nonparametric approach is to provide a smooth ROC curve in following presence of measurement error using mixture of flexible approximate parametric estimation with Bernstein type polynomials. Hence, the proposed nonparametric estimation method for the ROC curve involves replacing F_1 and F_0 by their distribution functions $\hat{F}_{1m}(s_1)$ and $\hat{F}_{0n}(s_0)$, respectively, by estimating the $\bar{\theta}$ parameters via maximum Bernstein likelihood due to the presence of the measurement error in the estimator of F.

As Lorentz, 2013 discussed in the equation (3.8) that f can be approximated with the same rate as $B_{mf}^k(s)$ by a Bernstein type polynomial of the form $f_s(x; \theta_m) = \sum_{k=0}^{m} \theta_{k,m} P_{k,m}(s)$, where $\bar{\theta}_m = (\theta_{0,m}, \theta_{1,m}, ..., \theta_{m,m})^T$, and its positive, which is called "a polynomial with positive coefficients in the literature of polynomial approximation" in the paper of Guan, 2016. Therefore, in this context, density f_{s1} and f_{s0} can be approximately modeled and parameterised by $f_{s_1,m}(s_1; \bar{\theta}_{1m})$ and $f_{s0,n}(s_0; \bar{\theta}_{0n})$ respectively as a mixture of the beta distribution and estimate the $\bar{\theta}_m$ parameters via the maximum likelihood method.

As we chose Bernstein polynomial method and estimate the probability density function (PDF) $f_1(x) = F_1'(x)$ by the Bernstein estimator of f of order $m > 0$ and F_E denotes the empirical distribution function.

$$\hat{f}_{1m}(x) = \frac{d}{dx}\hat{F}_{1m} = m \sum_{k=0}^{m-1} \left[F_{1E}\left(\frac{k+1}{m}\right) - F_{1E}\left(\frac{k}{m}\right) \right] P_{k,m-1}(x) \tag{3.14}$$

where $P_{k,m}(x) = \binom{m}{k} x^k (1-x)^{(m-k)}$ are binomial probabilities, and F_E denotes as the empirical distribution function obtained. The integration of the estimator of $\hat{f}_1$ is

$$\hat{F}_{1m,l}(x) = \int \hat{f}_{1m}(x)dx \tag{3.15}$$

and $\hat{F}_{0n}(x)$ is obtained similarly as $\hat{F}_{1m}(x)$ with n of the Bernstein polynomial degree.

$$\hat{f}_{0n}(x) = \frac{d}{dx}\hat{F}_{0n} = n \sum_{k=0}^{n-1} \theta_{k,n} \left[F_{0E}\left(\frac{k+1}{n}\right) - F_{0E}\left(\frac{k}{n}\right) \right] P_{k,n-1}(x) \tag{3.16}$$

F_{0E} denotes as the empirical distribution of F_0 obtained. The integration of the estimator of $\hat{f}_1$ is

$$\hat{F_0}_n(x) = \int \hat{f_0}_n(x)dx \tag{3.17}$$

$$\begin{aligned}\hat{f}_1(x;\bar{\theta}_{1m}) &= \sum_{k=0}^{m} \theta_{k,m} P_{k,m}(x) \\ &= P_{0,m} + \bar{\theta}_m \bar{P}_m(x)\end{aligned} \tag{3.18}$$

where $\bar{\theta}_m = (\theta_{1,m}, ..., \theta_{m,m})^T$, $\bar{P}_m = (P_{1,m} - P_{0,m}, ..., P_{m,m} - P_{0,m})^T$, and $P_{0,m} = \frac{1}{m+1}$. Clearly, $F_1(x) = \int f_1(x)dx$ is corresponding continuous cumulative density function. The F_1 also can be obtained approximately as

$$\hat{F}_1(x;\bar{\theta}_{1m}) = \sum_{k=0}^{m} \theta_{k,m} \int_0^x P_{k,m}(u)du \tag{3.19}$$

similarly, $\tilde{f}_0(x)$ and $\tilde{F}_0(x)$ can obtained similarly with a polynomial with n positive coefficients $\bar{\theta}_{0n}$. After obtaining the estimated $\hat{F}_1$ and $\hat{F}_0$ respectively, $\hat{F_0}^{-1}$ can be acquired by doing inverse of $\hat{F}_0$. Eventually, the ROC curve can be estimated by plugging in the estimated $\hat{F}_1, \hat{F_0}^{-1}$ into the equation (3).

Each probability density function of X_1 and X_0 can be approximated with a $beta_{\alpha,\beta}(\cdot)$ which is probability density function of a beta-distributed random variables based on the parameters α_1, β_1 and α_0, β_0 respectively from the assumption equation (2.3). Hence, the equations of the the approximated probability density function of X_1 and X_0 denoted as $\tilde{f}_{1x,m}(x_1; \alpha_1, \beta_1, \bar{\theta}_{1m})$ and $\tilde{f}_{0x,n}(x_0; \alpha_0, \beta_0, \bar{\theta}_{0n})$ respectively are

$$\begin{aligned} f_{1x,m}(x;\alpha_1,\beta_1,\bar{\theta}_{1m}) &= \frac{1}{(\alpha_1+\beta_1)-\beta_1} \tilde{f}_{s_1,m}\left(\frac{X_1-\beta_1}{\alpha_1}; \bar{\theta}_{1m}\right) \\ &= \frac{1}{\alpha_1} \sum_{k=0}^{m} \theta_{k,m} beta_{(k+1,m-k+1)} \left(\frac{X_1-\beta_1}{\alpha_1}\right) \end{aligned} \tag{3.20}$$

As Lorentz, 2013 stated since both of each f_{s1} and f_{s2} are continuous, the theorem 3.8 applied to the approximated density of $f_{w_1,m}(w_1,\tau_1,\alpha_1,\beta_1,\bar{\theta}_{1m})$

$$\begin{aligned} \tilde{f}_{w_1,m}(w_1;\tau_1,\alpha_1,\beta_1,\bar{\theta}_{1m}) &= \frac{1}{\tau_1}\int f_{x_{(1)},m}(x_{(1)};\alpha_1,\beta_1,\bar{\theta}_{1m}\phi\left(\frac{w_1-x_1}{\tau_1}\right)dx_1 \\ &= \frac{1}{\tau_1}\int \left\{\frac{1}{\alpha_1}\sum_{k=0}^{m}\theta_{k,m} beta_{(k+1,m-k+1)}\left(\frac{x_1-\beta_1}{\alpha_1}\right)\right\}\phi\left(\frac{w_1-x_1}{\tau_1}\right)dx_1 \\ &= \frac{1}{\alpha_1\tau_1}\sum_{k=0}^{m}\theta_{k,m}\int beta_{(k+1,m-k+1)}\left(\frac{x_1-\beta_1}{\alpha_1}\right)\phi\left(\frac{w_1-x_1}{\tau_1}\right)dx_1 \end{aligned} \tag{3.21}$$

which results similarly to $f_{w_0,n}(w_0,\tau_0,\alpha_0,\beta_0,\bar{\theta}_{0n})$. Finally, the Bernstein log-likelihood function as defined as Guan, 2016 of the set of parameters, $(\tau_1,\alpha_1,\beta_1,\bar{\theta}_{1m})$ given the observed data is

$$l_p(\tau_1,\alpha_1,\beta_1,\bar{\theta}_{1m}) = \sum_{i=1}^{n} log f_{w_1,m}(w_1;\tau_1,\alpha_1,\beta_1,\bar{\theta}_{1m})$$
$$= \sum_{i=1}^{n} log \left[\frac{1}{\alpha_1\tau_1} \sum_{k=0}^{m} \theta_{k,m} \int beta_{(k+1,m-k+1)} \left(\frac{x_1-\beta_1}{\alpha_1} \right) \phi \left(\frac{W_{1i}-x_1}{\tau_1} \right) dx_1 \right]. \tag{3.22}$$

where a collection of sample of W_1 are independent and identically distributed, W_{11}, W_{12}
$, W_{13}, ..., W_{1p} \sim W_1$. Besides, it applies to the Bernstein log-likelihood function of the set of parameters, $(\tau_0,\alpha_0,\beta_0,\bar{\theta}_{0m})$ denoted as $l_q(\tau_0,\alpha_0,\beta_0,\bar{\theta}_{0n})$ similarly where W_0 is a collection of sample with i.i.d , $W_{01}, W_{02}, W_{03}, ..., W_{0q} \sim W_0$.

The degree of the Bernstein polynomial, m,n , for each distribution is crucial to determine the optimal value because it determines model of the Bernstein polynomial. But, as we mentioned as an advantage of the maximum Bernstein likelihood density estimation is only one regularization parameter for each density to choose. Such as the choice of optimal bandwidth and kernel function is difficult, even the selecting tuning parameters of the Bayesian approach is more complicated. In this case, the Bernstein polynomial model is only determined by the each positive integer m,n.

The choice of the each optimal value, m,n, is mostly based on the result of the simulation. Guan, 2016 used EM algorithm for finding the maximum likelihood estimates $(\hat{\tau}_m,\hat{\alpha}_m,\hat{\beta}_m,\hat{\theta}_m)$ of $(\tau_1,\alpha_1,\beta_1,\bar{\theta}_{1m})$ and $(\hat{\tau}_n,\hat{\alpha}_n,\hat{\beta}_n,\hat{\theta}_n)$ of $(\tau_0,\alpha_0,\beta_0,\bar{\theta}_{0n})$ by using information criterion AIC/BIC which was created to find an appropriate penalty term.

Appendix A

Frequently Asked Questions

A.1 How do I change the colors of links?

The color of links can be changed to your liking using:
`\hypersetup{urlcolor=red}`, or
`\hypersetup{citecolor=green}`, or
`\hypersetup{allcolor=blue}`.
If you want to completely hide the links, you can use:
`\hypersetup{allcolors=.}`, or even better:
`\hypersetup{hidelinks}`.
If you want to have obvious links in the PDF but not the printed text, use:
`\hypersetup{colorlinks=false}`.

Bibliography

Arnold, A. S. et al. (Mar. 1998). “A Simple Extended-Cavity Diode Laser”. In: *Review of Scientific Instruments* 69.3, pp. 1236–1239. URL: http://link.aip.org/link/?RSI/69/1236/1.

Babu, G Jogesh, Angelo J Canty, and Yogendra P Chaubey (2002). “Application of Bernstein polynomials for smooth estimation of a distribution and density function”. In: *Journal of Statistical Planning and Inference* 105.2, pp. 377–392.

Bernstein, Serge (1912). “Demo istration du th'eoreme de Weierstrass fondee sur le calcul des probabilites”. In: *i* 13.1, pp. 1–2.

Bertrand, Aurélie, Ingrid Van Keilegom, and Catherine Legrand (2019). “Flexible parametric approach to classical measurement error variance estimation without auxiliary data”. In: *Biometrics* 75.1, pp. 297–307.

Carroll, Raymond J et al. (2006). *Measurement error in nonlinear models: a modern perspective*. CRC press.

Coffin, Marie and Shashikala Sukhatme (1997). “Receiver operating characteristic studies and measurement errors”. In: *Biometrics*, pp. 823–837.

Faraggi, David (2000). “The effect of random measurement error on receiver operating characteristic (ROC) curves”. In: *Statistics in Medicine* 19.1, pp. 61–70.

Guan, Zhong (2016). “Efficient and robust density estimation using Bernstein type polynomials”. In: *Journal of Nonparametric Statistics* 28.2, pp. 250–271.

Hawthorn, C. J., K. P. Weber, and R. E. Scholten (Dec. 2001). “Littrow Configuration Tunable External Cavity Diode Laser with Fixed Direction Output Beam”. In: *Review of Scientific Instruments* 72.12, pp. 4477–4479. URL: http://link.aip.org/link/?RSI/72/4477/1.

Krzanowski, Wojtek J and David J Hand (2009). *ROC curves for continuous data*. Crc Press.

Leblanc, Alexandre (2010). “A bias-reduced approach to density estimation using Bernstein polynomials”. In: *Journal of Nonparametric Statistics* 22.4, pp. 459–475.

Lorentz, George G (2013). *Bernstein polynomials*. American Mathematical Soc.

Peng, Liang and Xiao-Hua Zhou (2004). “Local linear smoothing of receiver operating characteristic (ROC) curves”. In: *Journal of Statistical Planning and Inference* 118.1-2, pp. 129–143.

Schwarz, Maik and Sébastien Van Bellegem (2010). “Consistent density deconvolution under partially known error distribution”. In: *Statistics & probability letters* 80.3-4, pp. 236–241.

Vitale, Richard A (1975). “A Bernstein polynomial approach to density function estimation”. In: *Statistical inference and related topics*. Elsevier, pp. 87–99.

Wieman, Carl E. and Leo Hollberg (Jan. 1991). “Using Diode Lasers for Atomic Physics”. In: *Review of Scientific Instruments* 62.1, pp. 1–20. URL: http://link.aip.org/link/?RSI/62/1/1.

Zou, Kelly H, WJ Hall, and David E Shapiro (1997). “Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests”. In: *Statistics in medicine* 16.19, pp. 2143–2156.