Journal of the Chinese
Statistical Association
Vol. 61, (2023) 75–107

A Novel Approach of the Classification by the Bernstein Monte-Carlo Smoothed Quantiles

Tsung-Lin Cheng[†] **and Tai-Chieh Yang**

Graduate Institute of Statistics and Information Science, National Changhua University of Education

ABSTRACT

The kernel density approaches depend heavily on choosing the best bandwidth corresponding to their related score functions. The procedures of classification using kernel density are popular but very complicated and inconvenient to implement. In this paper, we propose an interesting smoothing method of the estimated functionals of the given sample, which we call it "Bernstein Monte-Carlo Smoother" (BMCS), to serve as a classifier of the recruited junior high school students according to their monthly exam scores, IQ test scores, and Aptitude test scores. Instead of the kernel-based non-parametric discrimination procedures proposed by Lin *et al.* (2004), among others, we consider a much simpler method using the smoothed quantiles of the sample. In our study, we found that when the difference among the underlying groups is small, the un-smoothed curves cannot be distinguished clearly. It is fortunate that, after smoothing, the smoothed quantile plots are much smoother, which enable us to distinguish them much easier.

Based on a sample from a junior high school, the boys perform a slight better in IQ tests and the Aptitude tests while the girls perform a slight better in the monthly exams. It suggests that if we would like to classify the students despite the gender difference, we may adopt the weighted scores or the average scores suggested in this study.

Key words and phrases: quantile estimation, Bernstein approximation, classification.

JEL classification: C14, C15.

[†]Corresponding to: Tsung-Lin Cheng
E-mail: tlcheng@cc.ncue.edu.tw

1. Introduction and Motivation

As mentioned in Lin *et al.*, "the kernel-based nonparametric estimators (including LSCV: Least Squares Cross Validation, LKCV: Likelihood Cross Validation, and GKNN: Generalized kth-Nearest Neighbor) as desirable competitors to Fisher's linear discriminant rule for handling problems of educational placemen", kernel-based estimators are very popular in the nonparametric setting. They noted that "the LSCV and LKCV procedures failed to accomplish a complete separation of remedial group (RG) and advanced group (AG) students from the normal group (NG), resulting in some students being placed erroneously in the opposite direction". As mentioned by Wand and Jones and Kakizawa, since the kernel-based procedure is prone to obtain larger biased estimates for the kernel density estimation evaluated on the boundary, we plan to deal with the classification problem directly by a modified Bernstein approximation of the quantile functions.

For $n \geq 1$, define $b(k; n, x) := \binom{n}{k} x^k (1-x)^{n-k}$, $x \in [0,1]$, $k = 0, 1, \ldots, n$, to be the probability mass function of a binomial distribution related to the number of heads in tossing a coin n times, each with probability x of success, for $x \in [0,1]$, and $0 \leq k \leq n$.

Suppose that $G : [0,1] \to \Re$ is a continuous function. Define the Bernstein polynomial, of order m, with respect to the function G by

$$B_m^G(x) = \sum_{k=0}^{m} G(\frac{k}{m}) b(k; m, x), \ x \in [0,1]. \tag{1.1}$$

Note that the Bernstein polynomial is a very smooth function of x. Unfortunately, in general, the function G is unknown. For a random sample $X_1, X_2, \ldots, X_N$ with the distribution function F and associated density function f supported on a unit interval, Babu *et al.* used the empirical distribution function $F_N(x)$ instead of the distribution function $F(x)$, namely

$$B_m^{F_N}(x) = \sum_{j=0}^{m} F_N(\frac{j}{m}) b(j; m, x), \ x \in [0,1], \ m = 1, 2, \ldots. \tag{1.2}$$

Kakizawa considered some types of Bernstein-based generalized (boundary) kernel density estimators and investigates their bias and variance properties. He also compared his estimator with those proposed by Chen in terms of the Asymptotic Mean Integrated Square Errors (AMISE).

Leblanc has shown that the Bernstein estimator defined by (**1.2**) outperforms the empirical distribution function in terms of asymptotic mean square errors (AMSE) and mean integrated square errors (MISE).

Clearly, the Bernstein estimator (**1.2**) is unbiased for the corresponding Bernstein polynomial $B_m^F(x) = \sum_{k=0}^{m} F(\frac{k}{m})b(k; m, x)$ but not unbiased for the original distribution function F. From the definition of the Bernstein approximation, we find that the computation is a little complicated and therefore it requires a simpler method for its implementation.

Lin *et al.* adopted the kernel density estimation to divide students into three groups; the bottom 5% group, the top 5% group, and the remaining middle 80% group, according to their GPA scores. In particular, students with GPAs between the 90% and 95% quantiles and between the 5% and 10% quantiles—were excluded from their study. This is due to that there're no agreements could be reached on how to identify students falling in these zones (Glass). All students in every junior high school that participated in this study were going to be arranged in the three reference groups in compliance with the remedial, advanced, and regular curriculum programs, respectively. From Lin *et al.*'s results, we can see that these three approaches depend heavily on choosing the best bandwidth corresponding to their related score functions. Such procedures of classification might be complicated and inconvenient to implement. Remember that our classification problem involves much with the boundary which kernel smoothing is not able to deal with it unbiasedly. We are not aiming to propose a better estimation of the density function other than the standard kernel density estimation but to conduct the classification directly by our proposed Bernstein Monte-Carlo quantile estimation.

In this paper, we will propose a much simpler approach which is based on the Bernstein Monte-Carlo smoothed quantiles of the scores of the tests conducted for the students in the recruited junior high school. The smoothed quantile can serve as a good classifier. In fact, when we compare the IQ scores among different students, we usually adopt the percentiles to justify if they are in the “gifted children” group, the “normal group”, etc. However, the estimated empirical quantile function is usually not smooth and it sometimes affects the accuracy of the classification. In this case, we may use the Bernstein approximation to remedy the problem.

This paper is organized as below: in Section 2, we propose the so-called Bernstein Monte Carlo Smoother and investigate its viability; in Section 3, we recruit a sample of junior-high students from a junior high school in Changhua County, Taiwan, and apply our proposed Bernstein Monte Carlo approach to analyze the given data.

2. A New Bernstein Monte-Carlo Approximation of the Quantile Functions

In this section, we aim to propose to predict the group to which a student with given scores belongs. In practice, we may use many tests, e.g. IQ test, questionnaires, GPA, etc. among others, to classify the recruited students into the designed levels of learning ability.

Suppose that $G : [0, 1] \to \mathfrak{R}$ is a continuous function. Recall from the m-th order Bernstein polynomial of G given by,

$$B_m^G(x) = \sum_{j=0}^{m} G(\frac{j}{m})b(j; m, x), \ x \in [0, 1].$$

For any $\delta > 0$, define the modulus of a function f

$$\omega_f(\delta) = \sup_{|x-y|\leq\delta} |f(x) - f(y)|, \tag{2.1}$$

if the supremum does exist.

If $\omega_G(\delta) = O(\delta^H)$, as $\delta = m^{-\alpha} \to 0$ and we set $-\alpha H = -2 + 4\alpha$, then $\alpha = \frac{2}{4+H}$. We then have for all $x \in [0, 1]$,

$$\left|B_m^G(p) - G(x)\right| = O\left(m^{\frac{-2H}{4+H}}\right), \text{ as } n \to \infty. \tag{2.2}$$

In fact, in a similar way to but more tedious fashion than that of Theorem 2.2.2 in Chow and Teicher (1997), one might have an interesting generalization with a better convergent rate as below,

$$\sum_{|k-mp|>m\varepsilon} b(k;m,x) = O\left(\frac{1}{m^k\varepsilon^{2k}}\right), \text{ for } k = 1, 2, \dots. \qquad \textbf{(2.3)}$$

If we set $-\alpha H = -k+2k\alpha$, then $\alpha = \frac{k}{2k+H} \to \frac{1}{2}$, as $k \to \infty$. From this observation, we conclude that, following such an approach the best rate of convergence is $\omega_g\left(m^{-1/2}\right)$ which coincides with the result by Popoviciu (1935).

In general, the function G is unknown and remains to be estimated. Suppose that $\hat{G}$ is a consistent estimator of the function G on the points $\{\frac{k}{m},\ k = 0, 1, 2, \dots, m\}$. Through the $m+1$ points, $\hat{G}\left(\frac{k}{m}\right),\ k = 0, 1, 2, \dots, m$, we can define the Bernstein Monte-Carlo smoother as below:

Step1. For any given $x \in [0, 1]$, generate a large sample $T_1, T_2, \dots, T_n$ from a binomial distribution with parameter m, x, i.e. the probability mass function of each T_i is $f(k) = \begin{pmatrix} m \\ k \end{pmatrix} x^k(1-x)^{m-k},\ k = 0, 1, \dots, m$. Note that n can be chosen very large, say $n = 10^4$ or larger.

Step2. Set

$$\tilde{G}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \hat{G}\left(\frac{T_i}{m}\right) \qquad \textbf{(2.4)}$$

be the sample mean of the observations $\hat{G}\left(\frac{T_i}{m}\right),\ i = 1, 1, \dots, n$.

We call the smoother $\tilde{G}_n(x)$ defined by (**2.4**) the "Bernstein Monte-Carlo Smoother of the function G". Note that when m is not large, the estimated curve $\hat{G}$ of the functional G might not be smooth. The Bernstein Monte-Carlo smoother can fix this problem.

Why does this smoother work? Observe that if we treat the consistent estimated $\hat{G}\left(\frac{T}{m}\right)$ to be a (deterministic) function of the random variable T from $Binomial(m, x)$, then the expectation $E\hat{G}\left(\frac{T}{m}\right) = \sum_{k=0}^{n} \hat{G}\left(\frac{k}{m}\right) b(k;m,x),\ x \in [0, 1]$, which is exactly the Bernstein estimator defined by (**1.1**). Since $\{T_j,\ j = 1, 2, \dots, n\}$ are random variables from $Binomial(m, x)$, under some mild conditions, the Law of Large Numbers ensures the convergence of the sample mean (**2.1**) to the expectation $E\hat{G}\left(\frac{T}{m}\right)$ stochastically. In

Table 1: Mean and Standard-Deviation of the five Scores for Boys and Girls.

Scores	Monthly test		IQ test		Aptitude test		Weighted		Average	
Gender	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Mean	65.1	68.2	70.6	69.0	41.9	40.2	60.7	61.4	59.2	59.1
Standard Deviation	20.6	18.9	10.2	9.1	27.8	26.9	17.7	16.6	17.4	16.4

the following section, we will consider the function $\hat{G} = \hat{Q}$ to be the estimated empirical quantile function.

3. Real Data Analysis

The data is collected by our second author who has been teaching at the Chang-Tai Junior High School, Changhua County, for more than fifteen years. This data set contains three monthly examinations of five courses, the IQ scores, and the Aptitude Test scores of the 9$^{\text{th}}$ grade students. This data is collected and analyzed by deleting all personal information except for the gender. We summarize the data in Table 1.

If $\hat{G}$ is a consistent estimator of G, then as the sample size is large, it will approximate the true values of the function G. If $X = (X_1, X_2, \ldots, X_d)$ is a vector of test scores under consideration, then we may calculate the weighted sum $\tilde{X} = \sum_{i=1}^{d} \alpha_i X_i$ to conduct the logistic regression for the weighted variable $\tilde{X}$ and then to determine whether a student is gifted or not. While logistic regression is a popular tool of classification, there are only two possible values, 0 and 1, for the response in logistic regression. If we have to classify the students into more than two groups, it requires three possible values for the response. To remedy this drawback, for continuous variables we can use the empirical quantile function directly to classify the students into more than two groups by our proposed Bernstein smoother. For example, by the classification according to the quantiles at $x = 0.05$ and $x = 0.95$ as the cut points, we may use the following Bernstein Monte Carlo type smoother to estimate the quantiles and then to predict the group to which a student belongs.

Step1. Collect a large sample of the monthly exam scores of the five main courses taught in the junior high schools, the IQ scores (standardized to between 0

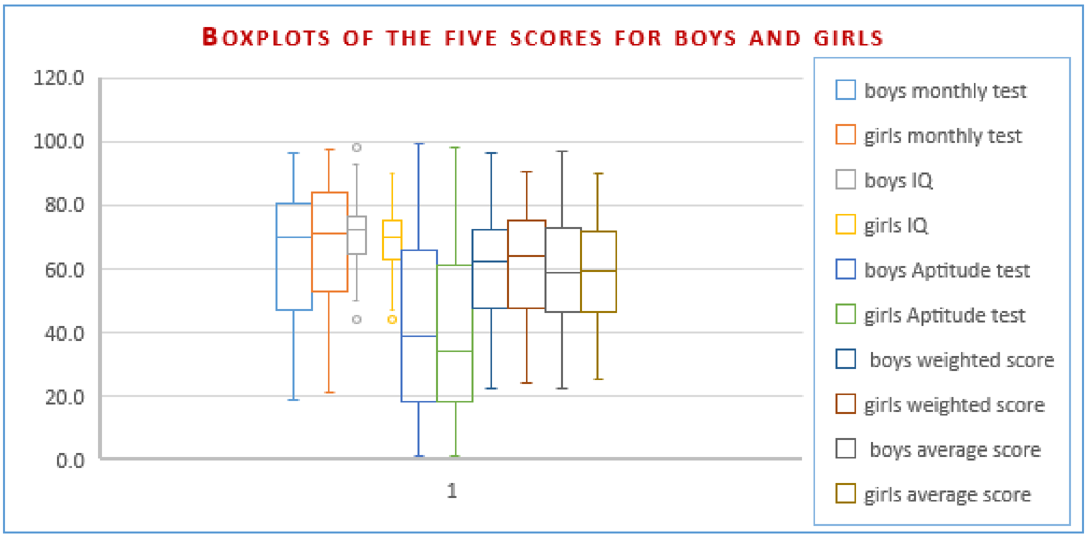


Figure 1: Boxplots of the five Scores for Boys and Girls.

~100), and the aptitude test scores (standardized to between 0 ~100) from the students in a junior high school of Changhua County.

Step2. Calculate the weighted scores or the average scores of each students and calculate the empirical quantiles from 0.05 to 0.95 by step size of 0.05, say $Q_N(0.05), Q_N(0.10), Q_N(0.15), \ldots, Q_N(0.95)$.

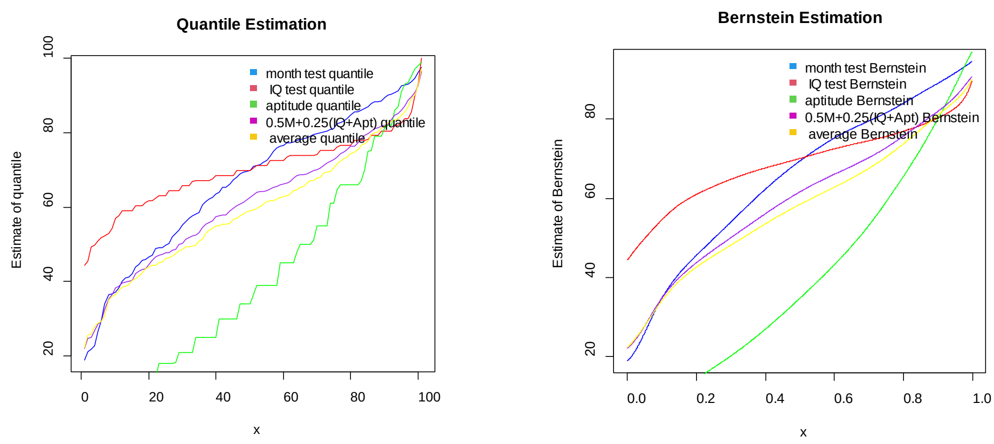
Step3. Based on the estimated $Q_N(0.05), Q_N(0.10), Q_N(0.15), \ldots, Q_N(0.95)$, we use the Bernstein Monte Carlo smoother to calculate the predicted $\tilde{Q}(x)$ for each $x \in [0, 1]$.

In addition to the monthly test, IQ test (scaled to 0 ~100 points) and Aptitude test scores, we consider two types of mixed scores:

Type1. Weighted score= 0.5 average monthly test score + 0.25(IQ test score + aptitude test score).

Type2. Average score= (average monthly test score + IQ test score+ aptitude test score)/3.

In this study, we recruit 99 boys and 95 girls as participants. In Figure 1, we sketch the boxplots of the five scores for boys and girls respectively. The basic descriptive statistics of the five scores are illustrated in Table 1.



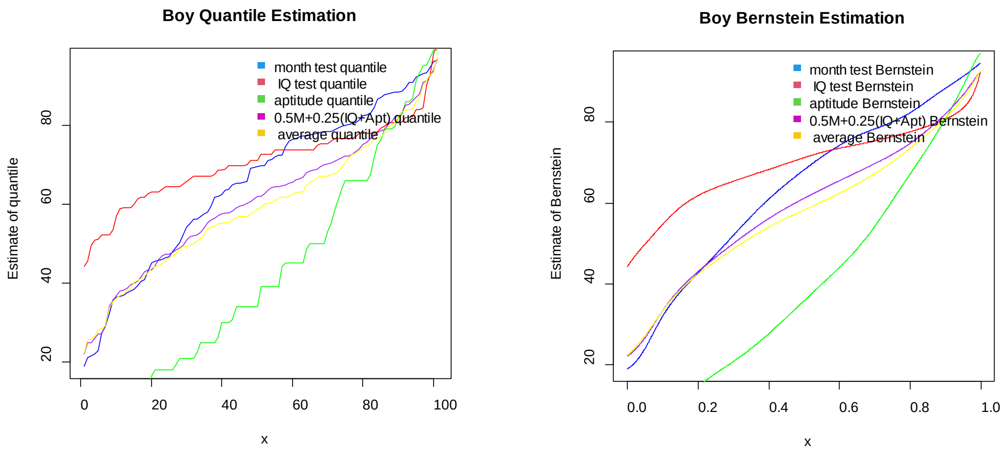
(a) (b)

Figure 2: (a)Quantiles of the five scores for all students. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the five scores for all students.

The following discussion is to illustrate the difference between the empirical quantile plots (Figure 2(a)-9(a), Table 2a-9a) and our proposed Bernstein Monte-Carlo Smoothed (BMCS) quantile plots (Figure 2(b)-9(b), Table 2b-9b). We will consider the quantiles (resp. the BMCS quantiles) of the five scores including the monthly test scores, the IQ test score, the Aptitude test scores, the weighted scores, and the average scores. In particular, analyzing the quantiles of the scores of the girls and the boys is interesting, which enables us to have a wider scope of the learning abilities of the slight difference of girls and boys in the junior high school.

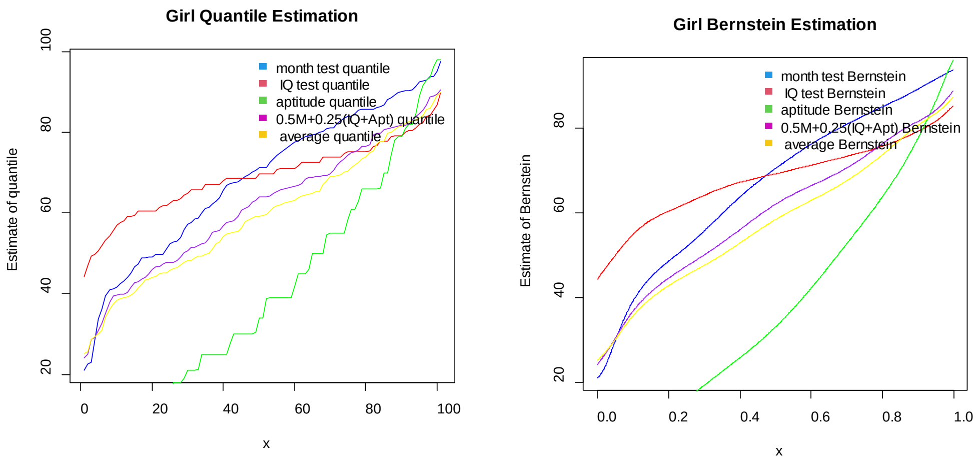
First, in Figure 2(a) we can have a quick look at the corresponding quantiles of the five scores of the pooled students. From the quantile plots in Figure 2(a), we can roughly observe that both of the IQ scores and the monthly test scores are easily distinguishable from the Aptitude test scores. This fact suggests that if we are only concerned with the intellectual aspect, then we might only consider the IQ test scores and the monthly test scores. On the other hand, the quantiles of the weighted scores and the average scores are similar and might achieve the same conclusion. Table 2a is the empirical quantiles of the five scores for the pooled sample.

In order to know if there are differences between the quantiles of the boys and those of the girls, we sketch the quantile plot for boys in Figure 3(a) and the quantile



(a) (b)

Figure 3: (a)Quantiles of the five scores for boys. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the five scores for boys.

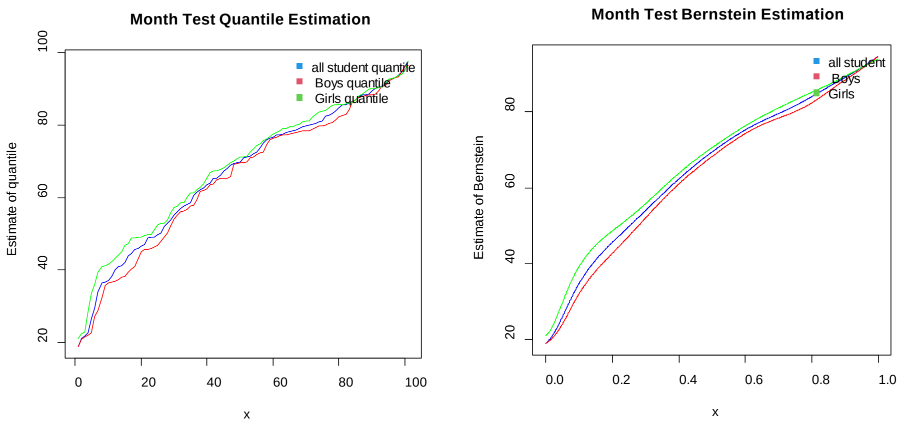


(a) (b)

Figure 4: (a)Quantiles of the five scores for girls. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the five scores for girls.

plot for girls in Figure 4(a). From Figure 3(a) and Figure 4(a), we can observe that there is a slight but not clear difference between boys and girls through the empirical quantile plots.

In order to know if there are differences between the quantiles of the boys and



(a) (b)

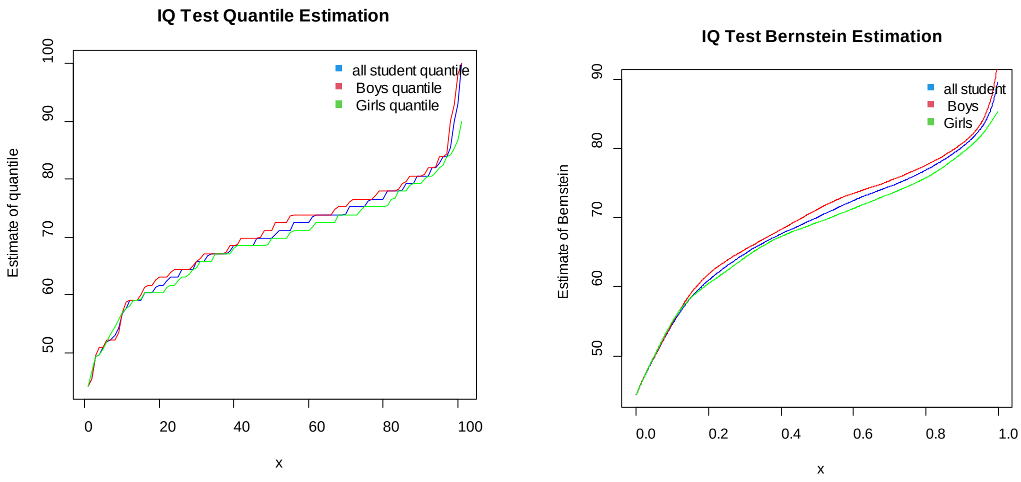
Figure 5: (a)Quantiles of the weighted monthly tests scores of the pooled students, boys and girls. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the weighted monthly tests scores of the pooled students, boys and girls.

those of the girls, we sketch the quantile plot for boys in Figure 3(a) and the quantile plot for girls in Figure 4(a). From Figure 3(a) and Figure 4(a), we can observe that there is a slight but not clear difference between boys and girls through the empirical quantile plots.

In Figure 5(a)-7(a), we compare the empirical quantiles of the weighted monthly tests, the IQ tests, and the Aptitude tests scores of the boys with those of the girls. In Figure 5(a), we can clearly distinguish the quantiles of the weighted monthly tests scores of the boys from those of the girls; however, in Figure 6(a) and 7(a), it is not easy to distinguish the IQ scores quantiles and the Aptitude scores quantiles of the boys from those of the girls. It is due to the roughness of the empirical quantile plots.

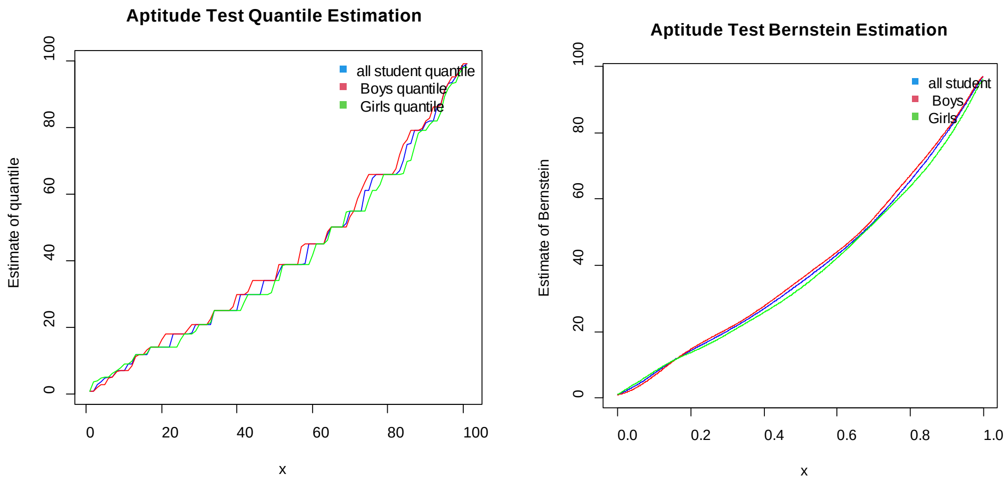
From Figure 8(a)-9(a), we can see the quantile plots of the weighted scores and the average scores of the pooled students, the boys, and the girls are twisted. In addition, we refer the readers to Tables 3a-9a which are the empirical quantiles used to plot Figure 3(a)-9(a).

From Figures 2(b) to 9(b), we illustrate the smoothed quantile plots by the proposed Bernstein Monte-Carlo Smoothing. In Figure 2(b), we may easily see the differences of the quantile plots of the five scores for the pooled students. From the smoothed quantile plots, the quantile curves are much smoother and can tell more stories.



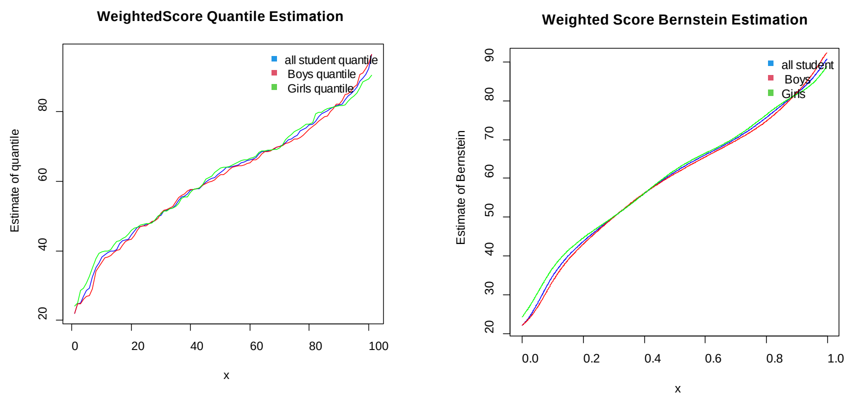
(a) (b)

Figure 6: (a)Quantiles of the IQ tests scores of the pooled students, boys and girls. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the IQ tests scores of the pooled students, boys and girls.



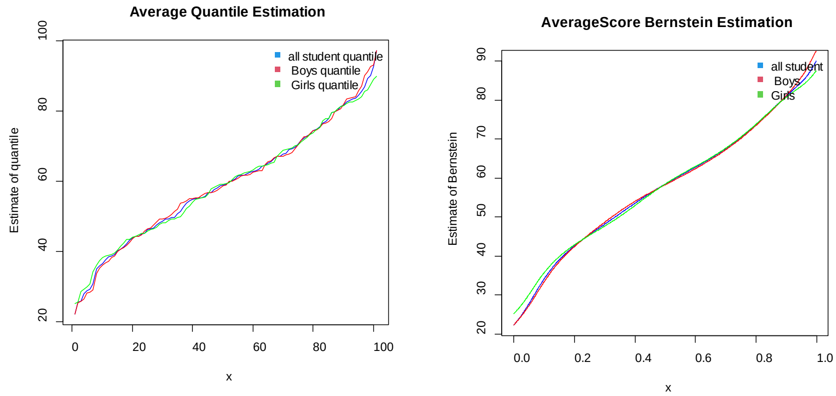
(a) (b)

Figure 7: (a)Quantiles of the aptitude test scores of the pooled students, boys and girls. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the aptitude test scores of the pooled students, boys and girls.



(a) (b)

Figure 8: (a)Quantiles of the weighted scores of the pooled students, boys and girls. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the weighted scores of the pooled students, boys and girls.



(a) (b)

Figure 9: (a)Quantiles of the average scores of the pooled students, boys and girls. (b)Bernstein Monte-Carlo Smoothing of the quantiles of the average scores of the pooled students, boys and girls.

Figure 3(b) is the smoothed quantile plot of the five scores for the boys and Figure 4(b) is the smoothed one for the girls. In order to better compare the scores of the boys with those of the girls, we sketch the smoothed quantile plots of the scores of the pooled students, the boys and the girls in the same figure. It is interesting that, from Figure 5(b), the smoothed quantile plots of the weighted monthly tests scores for the pooled students, boys and girls, show that the quantiles of the girls are a little higher than that of the boys. Note that the interesting phenomenon cannot be clearly illustrated by the original empirical quantile plots. It is an apparent benefit of our proposed approach. At the same quantile level, the junior high school girls have better performance than the boys on the standing of the weighted monthly exam scores. This fact agrees with the real educational environments.

From Figure 6(b), on the contrast, the smoothed quantile plots show that, at the same IQ score quantile level, the boys have a slight better score than the girls. This fact can't be observed by the original empirical quantile plots. Figure 7(b) has the same story for the Aptitude test scores as that depicted in Figure 6(b). At the same level of quantile level, the boys got a slight higher score.

From Figure 5(b)-7(b), we can see that the boys perform a slight better in IQ tests and the Aptitude tests and the girl perform a slight better in the monthly exams. If we will classify the students despite the gender difference, we may adopt the weighted scores or the average scores suggested in this study. As noted in Figure 8b and 9b, we observe that even if we use the smoothed quantiles of the weighted scores as well as the average scores, the difference between the boys and the girls is not clear.

4. Concluding Remarks

Empirical quantile can itself serve as a good classifier of the levels of the IQ scores, the Aptitude Test scores, and the monthly examination scores, among the students. While empirical quantile plots can help us distinguish the performance of different groups in a sample, it can't help us tell the difference among the groups when the difference is not significant enough. The proposed Bernstein Monte-Carlo smoother can simplify the computation of the original Bernstein approximation just as the Monte Carlo method can help us approximate the integration of a complicated or unknown function in a closed interval. From the smoothed quantile plots, we see that the boys outperform a slight better in IQ tests and the Aptitude tests and the girl outperform a slight better in the monthly exams. If we would like to classify the students despite the gender difference, we may adopt the weighted scores or the average scores suggested in this study.

Table 2a: Quantiles of the five scores for all students.

Quantile	MonthTest	IQtest	AptitudeTest	WeightedScore	Average
1.0	19.0	44.3	1.0	22.1	22.3
2.5	21.1	45.5	1.0	24.9	25.6
5.0	26.5	51.0	4.7	28.5	28.8
7.5	34.1	52.3	5.0	32.4	30.8
10.0	37.2	56.9	7.0	38.2	36.6
12.5	40.0	59.1	9.0	39.6	38.5
15.0	42.2	59.1	12.0	40.4	40.2
17.5	44.6	60.4	14.0	42.9	41.4
20.0	46.7	61.7	14.0	44.6	44
22.5	48.9	62.4	14.0	46.7	44.4
25.0	49.8	63.1	18.0	47.7	46.2
27.5	52.1	64.4	18.0	48.3	46.8
30.0	55.2	65.8	21.0	50.3	48.7
32.5	57.0	65.8	21.0	51.7	49.3
35.0	58.6	67.1	25.0	53.3	50.8
37.5	61.4	67.1	25.0	55.4	52.8
40.0	63.4	68.5	25.0	57.4	54.9
42.5	65.2	68.5	30.0	57.8	55.3
45.0	67.4	68.5	30.0	59.8	56.8
47.5	69	69.8	34.0	60.9	57.7
50.0	69.9	69.8	34.0	62.6	59.1
52.5	71.2	71.1	39.0	64.0	59.8
55.0	72.5	71.1	39.0	64.5	61.5
57.5	75.0	72.5	39.0	65.3	61.8
60.0	76.6	72.5	45.0	66.3	62.9
62.5	77.3	73.8	45.0	66.8	63.5
65.0	78.3	73.8	50.0	68.7	65.3
67.5	78.7	73.8	50.0	69.0	66.4
70.0	79.6	74.0	55.0	70.0	67.7
72.5	80.2	75.2	55.0	71.2	69.1
75.0	81.1	75.2	61.0	72.8	70.5
77.5	82.7	76.5	66.0	74.4	72.4
80.0	84.7	76.5	66.0	76.3	74.4
82.5	85.7	77.9	66.0	77.3	75.5
85.0	87.0	78.1	75.0	79.8	77.8
87.5	88.2	79.2	79.0	81.0	79.9
90.0	89.3	80.5	81.3	82.1	81.7
92.5	90.4	80.5	82.0	83.5	83.1
95.0	92.5	82.7	91.0	86.1	84.6
97.5	93.1	83.9	93.6	88.8	86.9

Table 2b: Bernstein Monte-Carlo Smoothing of the quantiles of the five scores for all students.

BQuantile	MonthTest	IQtest	AptitudeTest	WeightedScore	Average
2.0	19.1	44.4	1.0	22.2	22.4
2.5	21.5	47.1	2.2	24.3	24.7
5.0	25.7	49.7	3.7	27.5	27.6
7.5	30.4	52.1	5.4	31.2	30.8
10.0	34.6	54.4	7.4	34.6	34.0
12.5	38.0	56.5	9.3	37.5	36.7
15.0	40.8	58.2	11.1	39.8	38.9
17.5	43.3	59.7	12.7	41.8	40.8
20.0	45.5	60.9	14.2	43.6	42.5
22.5	47.6	62.0	15.7	45.3	44.0
25.0	49.6	62.9	17.2	46.9	45.4
27.5	51.7	63.9	18.6	48.4	46.8
30.0	53.9	64.7	20.1	50.0	48.1
32.5	56.0	65.5	21.7	51.5	49.4
35.0	58.1	66.3	23.3	53.0	50.8
37.5	60.2	67.0	25.1	54.5	52.1
40.0	62.2	67.6	26.9	56.0	53.5
42.5	64.1	68.2	28.8	57.5	54.8
45.0	65.9	68.8	30.7	58.9	56.0
47.5	67.7	69.4	32.6	60.2	57.2
50.0	69.3	70.0	34.7	61.5	58.4
52.5	70.9	70.6	36.7	62.8	59.5
55.0	72.4	71.3	38.8	63.9	60.6
57.5	73.8	71.8	40.9	65.0	61.6
60.0	75.1	72.4	43.2	66.0	62.7
62.5	76.3	73.0	45.4	67.0	63.8
65.0	77.5	73.5	47.8	68.0	65.0
67.5	78.5	74.0	50.3	69.1	66.2
70.0	79.5	74.5	53.0	70.2	67.5
72.5	80.4	75.0	55.9	71.4	68.8
75.0	81.5	75.6	58.9	72.7	70.3
77.5	82.7	76.2	62.1	74.1	71.9
80.0	83.9	76.9	65.4	75.6	73.6
82.5	85.2	77.6	68.7	77.2	75.4
85.0	86.4	78.3	72.1	78.7	77.2
87.5	87.7	79.2	75.9	80.3	79.1
90.0	88.9	80.2	79.8	81.8	80.9
92.5	90.2	81.3	83.8	83.5	82.7
95.0	91.6	82.7	88.2	85.5	84.5
97.5	93.0	84.9	92.6	88.0	86.7

Table 3a: Quantiles of the five scores for boys.

Quantile	MonthTest	IQtest	AptitudeTest	WeightedScore	Average
1.0	19.0	44.3	1.0	22.1	22.3
2.5	21.1	45.6	1.0	24.8	25.5
5.0	22.8	51.0	3.0	27.0	28.2
7.5	28.9	52.3	5.0	29.1	29.0
10.0	36.5	57.0	7.0	36.8	36.2
12.5	37.0	59.1	8.6	38.3	37.2
15.0	38.4	60.0	12.0	40.1	40.2
17.5	40.4	61.7	14.0	41.6	41.2
20.0	45.0	63.1	16.5	43.3	43.8
22.5	45.8	63.1	18.0	46.0	44.4
25.0	46.8	64.4	18.0	47.3	46.5
27.5	49.1	64.4	19.4	48.5	47.5
30.0	54.1	65.8	21.0	50.8	49.3
32.5	56.1	67.1	21.0	51.9	49.9
35.0	57.6	67.1	25.0	54.0	52.2
37.5	59.4	67.1	25.0	56.0	54.1
40.0	62.5	68.5	30.0	57.6	55.1
42.5	63.7	69.8	30.0	57.8	55.4
45.0	65.3	69.8	34.0	59.4	56.8
47.5	65.8	69.9	34.0	60.1	56.9
50.0	69.6	71.1	34.1	62.0	58.8
52.5	69.8	72.5	39.0	62.6	60.0
55.0	71.9	73.7	39.0	64.3	60.9
57.5	72.5	73.8	44.3	64.6	61.7
60.0	76.3	73.8	45.0	65.5	62.6
62.5	77.0	73.8	45.0	66.3	63.1
65.0	77.4	73.8	50.0	68.5	65.6
67.5	77.9	74.8	50.0	68.7	66.7
70.0	78.4	75.2	53.1	70.1	67.2
72.5	78.7	76.5	58.5	70.9	67.7
75.0	79.9	76.5	66.0	72.1	70.3
77.5	80.3	76.5	66.0	72.5	72.6
80.0	82.3	77.9	66.0	75.0	74.4
82.5	83.0	77.9	67.5	76.5	75.3
85.0	87.0	79.2	76.3	78.5	77.0
87.5	88.0	80.5	79.0	80.0	79.8
90.0	88.3	80.5	82.0	82.3	82.1
92.5	89.4	81.9	86.0	85.0	83.7
95.0	92.1	83.9	91.2	87.1	85.8
97.5	93.1	84.4	95.0	90.8	90.1

Table 3b: Bernstein Monte-Carlo Smoothing of the quantiles of the five scores for boys.

BQuantile	MonthTest	IQtest	AptitudeTest	WeightedScore	Average
2.0	19.1	44.4	1.0	22.2	22.4
2.5	20.8	47.1	1.8	24.0	24.5
5.0	23.9	49.7	3.0	26.6	27.1
7.5	27.9	52.2	4.7	29.8	30.1
10.0	31.8	54.6	6.7	33.2	33.2
12.5	35.2	56.9	8.8	36.3	36.0
15.0	37.9	58.8	10.8	38.8	38.4
17.5	40.3	60.5	12.8	41.0	40.5
20.0	42.6	61.7	14.7	42.9	42.3
22.5	44.9	62.8	16.3	44.8	44.0
25.0	47.2	63.7	17.8	46.5	45.6
27.5	49.5	64.6	19.3	48.2	47.1
30.0	51.9	65.3	20.8	49.9	48.6
32.5	54.3	66.1	22.3	51.5	50.0
35.0	56.6	66.8	24.0	53.1	51.4
37.5	58.8	67.5	25.7	54.7	52.7
40.0	60.9	68.2	27.6	56.1	54.0
42.5	62.8	69.0	29.6	57.4	55.1
45.0	64.7	69.7	31.6	58.7	56.2
47.5	66.4	70.4	33.7	59.9	57.2
50.0	68.1	71.1	35.7	61.1	58.2
52.5	69.8	71.8	37.8	62.2	59.2
55.0	71.3	72.5	39.8	63.3	60.2
57.5	72.8	73.0	41.9	64.4	61.3
60.0	74.2	73.5	44.0	65.4	62.4
62.5	75.4	73.9	46.2	66.5	63.5
65.0	76.5	74.4	48.5	67.5	64.6
67.5	77.4	74.8	51.2	68.6	65.9
70.0	78.3	75.3	54.1	69.6	67.2
72.5	79.1	75.8	57.3	70.8	68.6
75.0	80.0	76.4	60.5	71.9	70.1
77.5	81.0	76.9	63.8	73.2	71.7
80.0	82.2	77.6	67.0	74.6	73.4
82.5	83.6	78.2	70.2	76.2	75.2
85.0	85.2	78.9	73.6	78.0	77.1
87.5	86.8	79.8	77.0	79.9	79.0
90.0	88.3	80.7	80.6	81.9	81.1
92.5	89.7	81.8	84.5	84.1	83.2
95.0	91.2	83.4	88.7	86.6	85.8
97.5	92.8	86.2	93.2	89.5	88.9

Table 4a: Quantiles of the five scores for girls.

Quantile	MonthTest	IQtest	AptitudeTest	WeightedScore	Average
1.0	21.1	44.3	1.0	24.2	25.2
2.5	22.6	46.8	3.8	25.0	25.7
5.0	33.7	50.7	5.0	30.8	30.0
7.5	39.4	53.2	6.3	35.1	34.2
10.0	41.7	57.0	9.0	39.7	38.4
12.5	43.2	58.2	10.0	39.9	39.0
15.0	46.8	59.3	12.0	42.6	40.4
17.5	48.9	60.4	14.0	43.5	42.3
20.0	49.0	60.4	14.0	45.9	44.1
22.5	49.8	61.4	14.0	46.8	44.9
25.0	52.4	62.5	16.2	47.8	45.8
27.5	53.0	63.1	18.0	48.1	46.4
30.0	57.3	64.8	21.0	50.6	48.2
32.5	58.5	65.8	21.0	51.5	48.8
35.0	61.2	67.0	25.0	52.7	49.7
37.5	62.2	67.1	25.0	55.1	51.0
40.0	65.4	68.0	25.0	56.7	54.1
42.5	67.4	68.5	27.7	57.9	55 .0
45.0	68.3	68.5	30.0	60.6	56.4
47.5	69.5	68.5	30.0	61.5	58.3
50.0	71.2	69.8	34.0	63.9	59.1
52.5	71.3	69.8	38.7	64.1	59.6
55.0	74.3	70.8	39.0	65.2	61.7
57.5	75.7	71.1	39.0	65.9	62.4
60.0	77.5	71.1	41.8	66.6	63.2
62.5	78.5	72.5	45.0	67.3	64.3
65.0	79.4	72.5	50.0	68.8	64.8
67.5	80.0	72.6	50.0	69.0	65.4
70.0	81.1	73.8	55.0	69.6	68.9
72.5	82.1	73.8	55.0	72.0	69.2
75.0	83.9	75.2	58.4	74.3	70.5
77.5	84.6	75.2	61.0	75.1	71.9
80.0	85.6	75.2	66.0	76.4	73.8
82.5	85.8	76.5	66.0	79.4	75.4
85.0	86.9	77.9	69.8	80.6	77.9
87.5	88.5	79.0	74.2	81.1	79.9
90.0	90.1	79.2	79.0	81.7	81.5
92.5	90.4	80.5	82.0	82.1	82.6
95.0	92.5	81.9	89.1	84.5	83.4
97.5	93.0	83.9	93.0	86.9	85.6

Table 4b: Bernstein Monte-Carlo Smoothing of the quantiles of the five scores for girls.

BQuantile	MonthTest	IQtest	AptitudeTest	WeightedScore	Average
2.0	21.2	44.4	1.1	24.3	25.3
2.5	24.1	47.1	2.7	26.9	27.3
5.0	29.5	49.8	4.4	30.2	30.0
7.5	34.7	52.5	6.1	33.6	32.9
10.0	39.0	54.8	7.9	36.7	35.6
12.5	42.2	56.8	9.6	39.2	37.9
15.0	44.7	58.3	11.1	41.2	39.7
17.5	46.8	59.4	12.5	43.0	41.4
20.0	48.6	60.4	13.7	44.6	42.8
22.5	50.2	61.3	14.9	46.0	44.1
25.0	51.9	62.3	16.3	47.4	45.3
27.5	53.7	63.2	17.7	48.7	46.4
30.0	55.7	64.2	19.3	50.1	47.6
32.5	57.7	65.0	20.9	51.5	48.8
35.0	59.7	65.9	22.5	52.9	50.1
37.5	61.7	66.6	24.1	54.4	51.4
40.0	63.7	67.2	25.8	56.0	52.8
42.5	65.5	67.8	27.4	57.6	54.3
45.0	67.2	68.3	29.2	59.2	55.7
47.5	68.9	68.8	31.0	60.6	57.1
50.0	70.4	69.3	33.0	62.0	58.4
52.5	72.0	69.7	35.1	63.3	59.6
55.0	73.4	70.3	37.3	64.4	60.8
57.5	74.8	70.8	39.7	65.5	61.9
60.0	76.2	71.3	42.1	66.4	62.9
62.5	77.5	71.8	44.6	67.4	64.0
65.0	78.7	72.3	47.3	68.4	65.1
67.5	79.8	72.9	49.9	69.4	66.3
70.0	80.9	73.4	52.6	70.6	67.6
72.5	82.0	73.9	55.3	71.9	69.0
75.0	83.0	74.5	58.0	73.3	70.4
77.5	84.1	75.1	60.8	74.8	72.0
80.0	85.1	75.7	63.7	76.3	73.7
82.5	86.0	76.5	66.7	77.8	75.5
85.0	87.0	77.3	70.0	79.2	77.2
87.5	88.1	78.3	73.6	80.5	79.0
90.0	89.3	79.3	77.5	81.6	80.6
92.5	90.4	80.4	81.7	82.8	82.1
95.0	91.6	81.6	86.2	84.2	83.5
97.5	92.8	83.3	91.3	86.3	85.2

Table 5a: Quantiles of the weighted monthly tests scores of the pooled students, boys and girls.

Quantile	AllStudents	Boys	Girls
1.0	22.4	22.4	25.3
2.5	24.8	24.6	27.4
5.0	27.8	27.3	30.1
7.5	31.0	30.3	33.0
10.0	34.1	33.3	35.7
12.5	36.8	36.1	37.9
15.0	39.0	38.5	39.8
17.5	40.9	40.6	41.4
20.0	42.6	42.4	42.9
22.5	44.1	44.1	44.1
25.0	45.5	45.6	45.3
27.5	46.8	47.2	46.5
30.0	48.1	48.6	47.6
32.5	49.5	50.1	48.8
35.0	50.9	51.5	50.1
37.5	52.2	52.8	51.5
40.0	53.5	54.0	52.9
42.5	54.8	55.1	54.4
45.0	56.1	56.2	55.8
47.5	57.3	57.2	57.1
50.0	58.4	58.2	58.4
52.5	59.5	59.3	59.7
55.0	60.6	60.3	60.8
57.5	61.7	61.3	61.9
60.0	62.8	62.4	63.0
62.5	63.9	63.5	64.1
65.0	65.0	64.7	65.2
67.5	66.2	65.9	66.4
70.0	67.5	67.2	67.7
72.5	68.9	68.6	69.0
75.0	70.4	70.2	70.5
77.5	72.0	71.8	72.1
80.0	73.7	73.5	73.8
82.5	75.5	75.3	75.5
85.0	77.3	77.1	77.3
87.5	79.2	79.1	79.1
90.0	81.0	81.1	80.7
92.5	82.7	83.3	82.1
95.0	84.6	85.9	83.6
97.5	86.8	89.0	85.3

Table 5b: Bernstein Monte-Carlo Smoothing of the quantiles of the weighted monthly tests scores of the pooled students, boys and girls.

BernsteinQuantile	AllStudents	Boys	Girls
2.0	19.1	19.1	21.1
2.5	21.5	20.8	24.1
5.0	25.7	23.9	29.5
7.5	30.4	27.9	34.8
10.0	34.6	31.8	39.0
12.5	38.0	35.1	42.2
15.0	40.8	37.9	44.7
17.5	43.3	40.3	46.8
20.0	45.5	42.6	48.5
22.5	47.6	44.9	50.2
25.0	49.6	47.2	51.9
27.5	51.8	49.5	53.8
30.0	53.9	51.9	55.7
32.5	56.0	54.3	57.7
35.0	58.1	56.5	59.7
37.5	60.2	58.8	61.7
40.0	62.2	60.8	63.6
42.5	64.1	62.8	65.5
45.0	65.9	64.7	67.2
47.5	67.7	66.4	68.9
50.0	69.3	68.1	70.4
52.5	70.9	69.8	71.9
55.0	72.4	71.3	73.4
57.5	73.8	72.8	74.9
60.0	75.1	74.2	76.2
62.5	76.4	75.4	77.5
65.0	77.5	76.5	78.7
67.5	78.5	77.4	79.8
70.0	79.5	78.3	80.9
72.5	80.4	79.1	82.0
75.0	81.5	80.0	83.0
77.5	82.7	81.0	84.1
80.0	83.9	82.2	85.1
82.5	85.2	83.7	86.0
85.0	86.5	85.2	87.1
87.5	87.7	86.8	88.1
90.0	88.9	88.3	89.3
92.5	90.2	89.7	90.4
95.0	91.6	91.2	91.6
97.5	93.0	92.8	92.8

Table 6a: Quantiles of the IQ tests scores of the pooled students, boys and girls.

Quantile	AllStudents	Boys	Girls
1.0	22.4	22.4	25.3
2.5	24.8	24.6	27.4
5.0	27.8	27.3	30.1
7.5	31.0	30.3	33.0
10.0	34.1	33.3	35.7
12.5	36.8	36.1	37.9
15.0	39.0	38.5	39.8
17.5	40.9	40.6	41.4
20.0	42.6	42.4	42.9
22.5	44.1	44.1	44.1
25.0	45.5	45.6	45.3
27.5	46.8	47.2	46.5
30.0	48.1	48.6	47.6
32.5	49.5	50.1	48.8
35.0	50.9	51.5	50.1
37.5	52.2	52.8	51.5
40.0	53.5	54.0	52.9
42.5	54.8	55.1	54.4
45.0	56.1	56.2	55.8
47.5	57.3	57.2	57.1
50.0	58.4	58.2	58.4
52.5	59.5	59.3	59.7
55.0	60.6	60.3	60.8
57.5	61.7	61.3	61.9
60.0	62.8	62.4	63.0
62.5	63.9	63.5	64.1
65.0	65.0	64.7	65.2
67.5	66.2	65.9	66.4
70.0	67.5	67.2	67.7
72.5	68.9	68.6	69.0
75.0	70.4	70.2	70.5
77.5	72.0	71.8	72.1
80.0	73.7	73.5	73.8
82.5	75.5	75.3	75.5
85.0	77.3	77.1	77.3
87.5	79.2	79.1	79.1
90.0	81.0	81.1	80.7
92.5	82.7	83.3	82.1
95.0	84.6	85.9	83.6
97.5	86.8	89.0	85.3

Table 6b: Bernstein Monte-Carlo Smoothing of the quantiles of the IQ test scores of the pooled students, boys and girls.

BernsteinQuantile	AllStudents	Boys	Girls
2.0	44.4	44.4	44.4
2.5	47.1	47.1	47.1
5.0	49.7	49.7	49.8
7.5	52.1	52.2	52.5
10.0	54.4	54.6	54.8
12.5	56.5	56.9	56.8
15.0	58.3	58.9	58.3
17.5	59.7	60.5	59.4
20.0	60.9	61.8	60.4
22.5	62.0	62.8	61.3
25.0	62.9	63.7	62.2
27.5	63.9	64.6	63.2
30.0	64.7	65.3	64.1
32.5	65.5	66.1	65.1
35.0	66.3	66.8	65.9
37.5	67.0	67.5	66.6
40.0	67.6	68.2	67.2
42.5	68.2	69.0	67.8
45.0	68.8	69.7	68.3
47.5	69.4	70.4	68.8
50.0	70.0	71.1	69.3
52.5	70.6	71.8	69.7
55.0	71.2	72.5	70.2
57.5	71.8	73.0	70.8
60.0	72.4	73.5	71.3
62.5	73.0	73.9	71.8
65.0	73.5	74.4	72.3
67.5	74.0	74.8	72.9
70.0	74.5	75.3	73.4
72.5	75.0	75.8	73.9
75.0	75.6	76.4	74.5
77.5	76.2	77.0	75.1
80.0	76.8	77.6	75.7
82.5	77.6	78.2	76.5
85.0	78.3	78.9	77.3
87.5	79.2	79.8	78.3
90.0	80.2	80.7	79.3
92.5	81.3	81.8	80.4
95.0	82.7	83.4	81.6
97.5	84.9	86.1	83.3

Table 7a: Quantiles of the aptitude test scores of the pooled students, boys and girls.

Quantile	AllStudents	Boys	Girls
1.0	22.4	22.4	25.3
2.5	24.8	24.6	27.4
5.0	27.8	27.3	30.1
7.5	31.0	30.3	33.0
10.0	34.1	33.3	35.7
12.5	36.8	36.1	37.9
15.0	39.0	38.5	39.8
17.5	40.9	40.6	41.4
20.0	42.6	42.4	42.9
22.5	44.1	44.1	44.1
25.0	45.5	45.6	45.3
27.5	46.8	47.2	46.5
30.0	48.1	48.6	47.6
32.5	49.5	50.1	48.8
35.0	50.9	51.5	50.1
37.5	52.2	52.8	51.5
40.0	53.5	54.0	52.9
42.5	54.8	55.1	54.4
45.0	56.1	56.2	55.8
47.5	57.3	57.2	57.1
50.0	58.4	58.2	58.4
52.5	59.5	59.3	59.7
55.0	60.6	60.3	60.8
57.5	61.7	61.3	61.9
60.0	62.8	62.4	63
62.5	63.9	63.5	64.1
65.0	65.0	64.7	65.2
67.5	66.2	65.9	66.4
70.0	67.5	67.2	67.7
72.5	68.9	68.6	69.0
75.0	70.4	70.2	70.5
77.5	72.0	71.8	72.1
80.0	73.7	73.5	73.8
82.5	75.5	75.3	75.5
85.0	77.3	77.1	77.3
87.5	79.2	79.1	79.1
90.0	81.0	81.1	80.7
92.5	82.7	83.3	82.1
95.0	84.6	85.9	83.6
97.5	86.8	89.0	85.3

Table 7b: Bernstein Monte-Carlo Smoothing of the quantiles of the aptitude test scores of the pooled students, boys and girls.

BernsteinQuantile	AllStudents	Boys	Girls
2.0	1.0	1.0	1.1
2.5	2.3	1.7	2.7
5.0	3.7	3.0	4.4
7.5	5.4	4.7	6.1
10.0	7.3	6.7	7.9
12.5	9.3	8.7	9.6
15.0	11.1	10.8	11.1
17.5	12.7	12.8	12.5
20.0	14.2	14.7	13.7
22.5	15.7	16.3	14.9
25.0	17.2	17.8	16.2
27.5	18.6	19.3	17.7
30.0	20.1	20.8	19.3
32.5	21.7	22.3	20.9
35.0	23.3	23.9	22.5
37.5	25.1	25.7	24.1
40.0	26.9	27.6	25.8
42.5	28.7	29.6	27.4
45.0	30.7	31.6	29.1
47.5	32.6	33.6	31.0
50.0	34.6	35.7	33.0
52.5	36.7	37.8	35.1
55.0	38.8	39.9	37.3
57.5	40.9	41.9	39.7
60.0	43.1	44.0	42.1
62.5	45.4	46.2	44.6
65.0	47.8	48.6	47.3
67.5	50.3	51.1	49.9
70.0	53.0	54.1	52.6
72.5	55.8	57.2	55.3
75.0	58.9	60.5	58.0
77.5	62.1	63.8	60.8
80.0	65.4	67.0	63.7
82.5	68.7	70.2	66.7
85.0	72.3	73.5	70.0
87.5	76.0	77.0	73.6
90.0	79.8	80.6	77.5
92.5	83.9	84.5	81.7
95.0	88.2	88.7	86.3
97.5	92.7	93.3	91.4

Table 8a: Quantiles of the weighted scores of the pooled students, boys and girls.

Quantile	AllStudents	Boys	Girls
1.0	22.4	22.4	25.3
2.5	24.8	24.6	27.4
5.0	27.8	27.3	30.1
7.5	31.0	30.3	33.0
10.0	34.1	33.3	35.7
12.5	36.8	36.1	37.9
15.0	39.0	38.5	39.8
17.5	40.9	40.6	41.4
20.0	42.6	42.4	42.9
22.5	44.1	44.1	44.1
25.0	45.5	45.6	45.3
27.5	46.8	47.2	46.5
30.0	48.1	48.6	47.6
32.5	49.5	50.1	48.8
35.0	50.9	51.5	50.1
37.5	52.2	52.8	51.5
40.0	53.5	54.0	52.9
42.5	54.8	55.1	54.4
45.0	56.1	56.2	55.8
47.5	57.3	57.2	57.1
50.0	58.4	58.2	58.4
52.5	59.5	59.3	59.7
55.0	60.6	60.3	60.8
57.5	61.7	61.3	61.9
60.0	62.8	62.4	63.0
62.5	63.9	63.5	64.1
65.0	65.0	64.7	65.2
67.5	66.2	65.9	66.4
70.0	67.5	67.2	67.7
72.5	68.9	68.6	69.0
75.0	70.4	70.2	70.5
77.5	72.0	71.8	72.1
80.0	73.7	73.5	73.8
82.5	75.5	75.3	75.5
85.0	77.3	77.1	77.3
87.5	79.2	79.1	79.1
90.0	81.0	81.1	80.7
92.5	82.7	83.3	82.1
95.0	84.6	85.9	83.6
97.5	86.8	89.0	85.3

Table 8b: Bernstein Monte-Carlo Smoothing of the quantiles of the weighted scores of the pooled students, boys and girls.

BernsteinQuantile	AllStudents	Boys	Girls
2.0	22.2	22.2	24.3
2.5	24.3	24.0	26.9
5.0	27.5	26.6	30.2
7.5	31.2	29.9	33.6
10.0	34.6	33.2	36.7
12.5	37.5	36.3	39.2
15.0	39.8	38.8	41.2
17.5	41.8	41.0	43.0
20.0	43.6	42.9	44.6
22.5	45.3	44.8	46.0
25.0	46.9	46.5	47.4
27.5	48.4	48.2	48.7
30.0	50.0	49.9	50.1
32.5	51.5	51.5	51.5
35.0	53.0	53.1	52.9
37.5	54.5	54.7	54.4
40.0	56.0	56.1	56.0
42.5	57.5	57.5	57.6
45.0	58.9	58.7	59.2
47.5	60.3	60.0	60.6
50.0	61.5	61.1	62.0
52.5	62.8	62.2	63.3
55.0	63.9	63.3	64.4
57.5	65.0	64.4	65.4
60.0	66.0	65.5	66.4
62.5	67.0	66.5	67.4
65.0	68.0	67.5	68.4
67.5	69.1	68.6	69.4
70.0	70.2	69.7	70.6
72.5	71.4	70.8	71.9
75.0	72.7	71.9	73.3
77.5	74.1	73.2	74.8
80.0	75.6	74.7	76.3
82.5	77.1	76.2	77.8
85.0	78.7	78.0	79.2
87.5	80.3	79.9	80.5
90.0	81.9	81.9	81.6
92.5	83.5	84.1	82.8
95.0	85.5	86.6	84.2
97.5	88.0	89.5	86.3

Table 9a: Quantiles of the average scores of the pooled students, boys and girls.

Quantile	AllStudents	Boys	Girls
1.0	22.4	22.4	25.3
2.5	24.8	24.6	27.4
5.0	27.8	27.3	30.1
7.5	31.0	30.3	33.0
10.0	34.1	33.3	35.7
12.5	36.8	36.1	37.9
15.0	39.0	38.5	39.8
17.5	40.9	40.6	41.4
20.0	42.6	42.4	42.9
22.5	44.1	44.1	44.1
25.0	45.5	45.6	45.3
27.5	46.8	47.2	46.5
30.0	48.1	48.6	47.6
32.5	49.5	50.1	48.8
35.0	50.9	51.5	50.1
37.5	52.2	52.8	51.5
40.0	53.5	54.0	52.9
42.5	54.8	55.1	54.4
45.0	56.1	56.2	55.8
47.5	57.3	57.2	57.1
50.0	58.4	58.2	58.4
52.5	59.5	59.3	59.7
55.0	60.6	60.3	60.8
57.5	61.7	61.3	61.9
60.0	62.8	62.4	63.0
62.5	63.9	63.5	64.1
65.0	65.0	64.7	65.2
67.5	66.2	65.9	66.4
70.0	67.5	67.2	67.7
72.5	68.9	68.6	69.0
75.0	70.4	70.2	70.5
77.5	72.0	71.8	72.1
80.0	73.7	73.5	73.8
82.5	75.5	75.3	75.5
85.0	77.3	77.1	77.3
87.5	79.2	79.1	79.1
90.0	81.0	81.1	80.7
92.5	82.7	83.3	82.1
95.0	84.6	85.9	83.6
97.5	86.8	89.0	85.3

Table 9b: Bernstein Monte-Carlo Smoothing of the quantiles of the average scores of the pooled students, boys and girls.

BernsteinQuantile	AllStudents	Boys	Girls
2.0	22.4	22.4	25.3
2.5	24.7	24.5	27.3
5.0	27.6	27.1	30.0
7.5	30.9	30.1	32.9
10.0	34.0	33.2	35.6
12.5	36.7	36.0	37.8
15.0	38.9	38.4	39.7
17.5	40.8	40.5	41.4
20.0	42.5	42.3	42.8
22.5	44.0	44.0	44.1
25.0	45.4	45.6	45.3
27.5	46.8	47.1	46.4
30.0	48.1	48.6	47.6
32.5	49.4	50.0	48.8
35.0	50.8	51.4	50.1
37.5	52.1	52.7	51.5
40.0	53.5	54.0	52.8
42.5	54.8	55.1	54.3
45.0	56.0	56.2	55.7
47.5	57.2	57.2	57.1
50.0	58.4	58.2	58.4
52.5	59.5	59.2	59.6
55.0	60.6	60.2	60.8
57.5	61.6	61.3	61.9
60.0	62.7	62.3	62.9
62.5	63.8	63.5	64.0
65.0	65.0	64.6	65.1
67.5	66.2	65.9	66.3
70.0	67.5	67.2	67.6
72.5	68.8	68.6	69.0
75.0	70.3	70.1	70.5
77.5	71.9	71.7	72.0
80.0	73.6	73.4	73.7
82.5	75.4	75.2	75.4
85.0	77.2	77.0	77.3
87.5	79.1	79.0	79.0
90.0	80.9	81.1	80.6
92.5	82.7	83.2	82.1
95.0	84.5	85.8	83.5
97.5	86.7	88.9	85.2

References

[1] Babu, G. J., Canty, A. J., and Chaubey, Y. P. (2002). Application of Bernstein polynomials for smooth estimation of a distribution and density function. *Journal of Statistical Planning and Inference*, 105(2), pages 377-392.

[2] Bahadur, R. R. (1966). A note on quantiles in large samples. *The Annals of Mathematical Statistics*, 37(3), pages 577-580.

[3] Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71(2), pages 353-360.

[4] Braess, D., and Sauer, T. (2004). Bernstein polynomials and learning theory. *Journal of Approximation Theory*, 128(2), pages 187-206.

[5] Chen, S. X. (1996). Empirical likelihood confidence intervals for nonparametric density estimation. *Biometrika*, 83(2), pages 329-341.

[6] Chen, S. X. (1999). Beta kernel estimators for density functions. *Computational Statistics & Data Analysis*, 31(2), pages 131-145.

[7] Chow, Y. S., and Teicher, H. (1997). Probability Theory: Independent, Interchangeability, Martingales, 3rd Edition. Springer.

[8] Cryer, J. D., and Chan, K. S. (2008). Time Series Analysis: With Applications in R. Springer Science & Business Media.

[9] Duin, R. P. W. (1976). On the Choice of Smoothing Parameters for Parzen Estimators of Probability Density Functions. *IEEE Transactions on Computers*, 25(11), pages 1175-1179.

[10] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2), pages 179-188.

[11] Glass, G. V. (1978). Standards and criteria. *Journal of educational measurement*, 15(4), pages 237-261.

[12] Habbema, J. D. F., Hermans, J., and Van Der Broek, K. (1974). A stepwise discriminant analysis program using density estimation. In *1974 Compstat*, pages 100-110.

[13] Hand, D. J. (1981). Discrimination and Classification (Wiley Series in Probability and Statistics - Applied Probability and Statistics Section). John Wiley & Sons.

[14] Hand, D. J. (1982). Kernel Discriminant Analysis. John Wiley & Sons.

[15] Härdle, W. K. (1991). Smoothing techniques: with implementation in S. Springer Science & Business Media.

[16] Kakizawa, Y. (2004). Bernstein polynomial probability density estimation. *Journal of Nonparametric Statistics*, 16(5), pages 709-729.

[17] Khuri, A. I. (2003). Advanced calculus with applications in statistics. John Wiley & Sons.

[18] Leblanc, A. (2012). On the boundary properties of Bernstein polynomial estimators of density and distribution functions. *Journal of Statistical Planning and Inference*, 142(10), pages 2762-2778.

[19] Lin, M. H., Huang, S. Y., and Chang, Y. C. (2004). Kernel-based discriminant techniques for educational placement. *Journal of Educational and Behavioral Statistics*, 29(2), pages 219-240.

[20] Moore, D. S., and Yackel, J. W. (1977). Consistency properties of nearest neighbor density function estimators. *The Annals of Statistics*, 5(1), pages 143-154.

[21] Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3), pages 1065-1076.

[22] Popoviciu, T. (1935). Sur l'approximation des fonctions convexes d'ordre supérier. *Mathematica (Cluj)*, 10, pages 49-54.

[23] Rao, C. R. (1973). Linear Statistical Inference and its Applications: Second Edition. John Wiley & Sons.

[24] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27(3), pages 832-837.

[25] Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, 9, pages 65-78.

[26] Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis. Chapman & Hall.

[27] Stone, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *The Annals of Statistics*, 10(4), pages 1040-1053.

[28] Stone, C. J. (1984). An asymptotically optimal window selection rule for kernel density estimates. *The Annals of Statistics*, 12(4), pages 1285-1297.

[29] Wand, M. P., and Jones, M. C. (1995). Kernel smoothing. Chapman & Hall.

[30] Whittle, P. (1958). On the smoothing of probability density functions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), pages 334-343.

[Received July 2022; accepted November 2022.]

一種由伯恩斯坦蒙地卡羅平滑分位數來進行分類的新方法

鄭宗琳† **楊岱潔**

國立彰化師範大學統計資訊研究所

摘　要

核密度估計法非常倚賴選擇最佳帶寬以獲得較好的估計，雖然用核密度估計來做分類的方式非常普遍，但是這個方法也比較複雜，實際應用上也不是太方便。本文提出一個新的分類方法，我們稱之為伯恩斯坦蒙地卡羅平滑法。我們使用了針對彰化縣某國中的月考成績、IQ 分數、以及性向測驗分數所估計出來的分位數來做分類的依據，並使用我們所提出的新方法來測試。由於我們的方法不但可以用來內插尚未估計的分位數，在視覺上它比原來的原始分位數更能區分不同組別的差異。例如，不同性別在成績表現上或 IQ 分數的表現上，有著些微的差距，雖然原始分位數的圖形不易分辨出，但是平滑之後的分位數圖形就比較容易分辨他們的差異。在使用的三種原始分數的分析中，我們發現若想要忽略性別所造成的影響，可以使用加權分數或平均分數來計算原始分位數，然後再使用我們的平滑法來產生平滑分位數並據此來做分類。

關鍵詞：伯恩斯坦蒙地卡羅平滑法、分位數近似、分類。

JEL classification: C14, C15.

†通訊作者: 鄭宗琳
E-mail: tlcheng@cc.ncue.edu.tw