TWMS J. App. and Eng. Math. V.11, N.3, 2021, pp. 717-723

CONVERGENCY RATE OF MEAN INTEGRATED SQUARE ERROR IN DENSITY ESTIMATION BASED ON HERMITE POLYNOMIALS

E. ERÇELIK[1], §

Abstract. The estimate of a probability density using a sequence $X_1, X_2, ..., X_N$ of independent and identically distributed (i.i.d) random variables with common unknown density function is studied by means of Hermite functions. We obtained convergency rate of the mean integrated square error (MISE) of density estimators by using delta sequences which are based on the hermite functions when the support of the density function is infinite. Then we obtained convergency rate of the mean square error (MSE) and MISE of estimator for the densities having compact support. The contribution of this work is improving the results of former publications about rate of convergence of estimators based on Hermite series.

Keywords: Delta Sequences; Rates of Convergence; Hermite Functions; Orthogonal Polynomials.

AMS Subject Classification: 62G07, 41A25, 33C45.

1. Introduction

Let $X_1, X_2, ..., X_N$ be a sequence of independent and identical distributed (i.i.d) random variables with a common unknown density function $f(x)$. A lot of methods have been used to estimate the density function such as kernel method studied by Rosenblatt (1956) and Parzen (1962); interpolation method studied by Wahba (1971) and Wahba (1975); delta sequence method studied by Bleuez and Bosq (1976), Bosq (1977), Walter and Blum (1979), Susarla and Walter (1981), Letellier (1997), Nadar (2011) and Nadar and Erçelik (2017); the orthogonal series method which was first introduced by Chencov (1962).

The Hermite series estimator is a special type of orthogonal series estimator which are considered by Schwartz (1967), Walter (1977), Greblicki and Pawlak (1984) and Liebscher (1990). For the first time Hermite series estimators are used by Schwartz (1967). They studied consistency and MISE and MSE rate of convergence of the estimators based on Hermite series. Then, Walter (1977) used Hermite series estimators and obtained rates of convergence of MISE which improved the results of Schwartz (1967) by using better bounds on the Hermite functions. They also obtained the estimates for the characteristic function and their rates of convergence owing to the fact that the Hermite functions are

[1] Istanbul Technical University - Department of Mathematical Engineering - Istanbul- Turkey.
e-mail:ercelikel@itu.edu.tr; ORCID: https://orcid.org/0000-0002-2008-8033.
§ Manuscript received: July 30, 2019; accepted: December 21, 2020.

eigenfunctions of the Fourier transform. Later, Greblicki and Pawlak (1984) investigated the estimate of a density and its derivatives using the orthogonal Hermite series. They obtained better results than former studies about convergence rate of MISE as well as the uniform convergence in the MSE. Also, when the density is not continuous, the behavior of Hermite series estimators are studied and assertions about MSE, MISE and asymptotic normality are given in the paper of Liebscher (1990).

Hermite polynomials are encountered in a wide variety of applications in the mathematical statistics, mathematical physics, and mechanics. They appear in the problems of the integration of Laplace equation and Helmholtz equation in parabolic coordinates (see Lebedev, 1965). Also, eigensolution of the quantum harmonic oscillator contains Hermite polynomials. Moreover, Hermite functions are useful in applied work involving nonparametric techniques. (see Jones and Silverman, 1989). They are much faster to calculate than kernel density estimators for large N. The calculations are only based on the recurrence relations for the Hermite functions.

In this paper, we used delta sequence density estimation method based on Hermite polynomials for an unknown density function f. We obtain the convergency rate of the MISE of a density estimator based on Hermite polynomials for density function with infinite support and having rth derivative. Moreover, the convergency rate of the MISE and MSE of density estimator are studied for the densities which have compact support. The contribution of this work is improving the results of former studies about the rate of convergence of estimators based on Hermite series by using certain delta sequence method.

In section 2, standard properties and some preliminary results of the Hermite functions are given. Then, the rate of convergency of MISE is obtained for the densities having infinite support. In section 3, the convergency rate of MSE and MISE are derived for densities having compact support. Finally, Section 4 is devoted to conclusions.

2. Density Estimation Using Delta Sequence Based on Hermite Polynomials

The Hermite orthonormal system over the real line R given by

$$h_k(x) = \left(2^k k! \pi^{\frac{1}{2}}\right)^{-\frac{1}{2}} H_k(x) e^{-\frac{x^2}{2}}, \quad k = 0, 1, 2, ...$$

where

$$H_k(x) = (-1)^k e^{x^2} \left(\frac{d^k}{dx^k}\right) e^{-x^2}$$

is the kth Hermite polynomial.

It is known that the normalized Hermite functions $\{h_k\}$ are the complete orthonormal system in $L_2(-\infty, \infty)$ and they satisfy the recurrence formulas

$$xh_k = \left(\frac{k}{2}\right)^{\frac{1}{2}} h_{k-1} + \left(\frac{k+1}{2}\right)^{\frac{1}{2}} h_{k+1}, \quad k = 1, 2, ...$$

and

$$\frac{d}{dx} h_k = \left(\frac{k}{2}\right)^{\frac{1}{2}} h_{k-1} - \left(\frac{k+1}{2}\right)^{\frac{1}{2}} h_{k+1}, \quad k = 1, 2, ...$$

and satisfy the following inequalities (Szegö, 1939);

$$|h_k(x)| \leq \frac{c_1}{(k+1)^{1/12}}, \quad x \in (-\infty, \infty), \quad k = 0, 1, 2... \tag{1}$$

and

$$|h_k(x)| \leq \frac{c_2}{(k+1)^{1/4}}, \quad x \in (-a, a), \quad k = 0, 1, 2... \tag{2}$$

where a is any nonnegative integer and the constants c_1 and c_2 are independent of x and k.

Let $X_1, X_2, ..., X_N$ be a sequence of i.i.d. random variables with unknown density function $f(x)$. Then, an unknown density function f can be written by means of the Hermite series

$$f(x) = \sum_{k=0}^{\infty} a_k h_k(x)$$

with the Hermite coefficients defined by

$$a_k = \int f(x) h_k(x) dx.$$

Throughout this chapter, we shall assume that $f(x)$ is square integrable and we use c or c_i, $i = 1, 2, ..., m$ for any positive constant, independent of f. Now, by using Hermite polynomials delta sequence density estimator can be defined as

$$\widehat{f}_{N,n}(x) = \frac{1}{N} \sum_{i=1}^{N} \delta_n(x, X_i) \tag{3}$$

where

$$\delta_n(x, X_i) = \sum_{k=0}^{n} \frac{H_k(x) e^{-\frac{x^2}{2}}}{\left(2^k k! \pi^{\frac{1}{2}}\right)^{\frac{1}{2}}} \frac{H_k(X_i) e^{-\frac{x^2}{2}}}{\left(2^k k! \pi^{\frac{1}{2}}\right)^{\frac{1}{2}}} = \sum_{k=0}^{n} h_k(x) h_k(X_i).$$

Since orthogonal series density estimate can take negative values, then it is proposed that density estimate at x is the max $\left[0, \widehat{f}_{N,n}(x)\right]$. For the discussion of the MISE and MSE, following two lemmas, proved by Schwartz (1967), are necessary.

Lemma 2.1. *Assume that the function* $\left(x - \frac{d}{dx}\right)^r f \in L_2(-\infty, \infty)$ *for some integer* $r > 0$. *Then the coefficients* a_k, $k = 1, 2, ...$ *satisfy the bound*

$$|a_k| \leq \frac{c_3}{(2k)^{\frac{r}{2}}}$$

where c_3 *is the* L_2 *norm of* $\left(x - \frac{d}{dx}\right)^r f$.

Lemma 2.2. *Let f(x) be continuous, of bounded variation,* L_1 *and* L_2 *in* $(-\infty, \infty)$. *Then, the series in*

$$f(x) = \sum_{k=0}^{\infty} a_k h_k(x)$$

converges uniformly in any interval interior to $(-\infty, \infty)$.

2.1. Convergency rate of MISE of estimators for densities having infinite support.

Theorem 2.1. *Let* $\left(x - \frac{d}{dx}\right)^j f \in L_2(-\infty, \infty)$ *for* $j = 1, 2, ..., r$; *then the MISE rate of (3) satisfies*

$$MISE(\widehat{f}_{N,n}(x)) = O\left(N^{-\frac{12r}{6r+23}}\right).$$

Proof. First, lets investigate the integrated variance term

$$N \int_{-\infty}^{\infty} var(\widehat{f}_{N,n}(x)) dx = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \delta_n^2(x, t) f(t) dt - \left(\int_{-\infty}^{\infty} \delta_n(x, t) f(t) dt \right)^2 \right\} dx$$

it is sufficient to investigate the rate of grows of $\int\int \delta_n^2(x,t)f(t)dtdx$.

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta_n^2(x,t)f(t)dtdx &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sum_{k=0}^{n} h_k(x)h_k(t)\right)^2 f(t)dtdx \\
&\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sum_{k=0}^{n} h_k^2(x)\sum_{k=0}^{n} h_k^2(t)\right) f(t)dtdx \\
&= \int_{-\infty}^{\infty}\sum_{k=0}^{n} h_k^2(x)dx \int_{-\infty}^{\infty}\sum_{k=0}^{n} h_k^2(t)f(t)dt \\
&\leq \int_{-\infty}^{\infty}\sum_{k=0}^{n} h_k^2(x)dx \sum_{k=0}^{n}\frac{1}{(k+1)^{\frac{1}{12}}}\int_{-\infty}^{\infty} h_k(t)f(t)dt
\end{aligned}$$

From Lemma 1, it can be deduced that

$$|a_k| \leq \frac{c}{(2k+2)^{r/2}}, \quad k = 0, 1, 2, ...$$

Since

$$\int_{-\infty}^{\infty} e^{-x^2}H_k^2(x)dx = \sqrt{\pi}2^n n!$$

and using above bound, an order for the integrated variance term can be obtained as

$$\int_{-\infty}^{\infty} var\left(\widehat{f}_{N,n}(x)\right)dx = O\left(\frac{n^{\frac{23-6r}{12}}}{N}\right) \tag{4}$$

Now, the squared bias term can be written as

$$\begin{aligned}
Bias^2\left(\widehat{f}_{N,n}(x)\right) &= \left\{\int_{-\infty}^{\infty} \delta_n(x,t)f(t)dt - f(x)\right\}^2 \\
&= \left\{\int_{-\infty}^{\infty}\sum_{k=0}^{n} h_k(x)h_k(t)f(t)dt - \sum_{k=0}^{\infty} a_k h_k(x)\right\}^2 \\
&= \left\{\sum_{k=n+1}^{\infty} a_k h_k(x)\right\}^2
\end{aligned}$$

since Hermite functions are orthonormal in L_2, then the integrated squared bias is written as follows

$$\int_{-\infty}^{\infty} Bias^2\left(\widehat{f}_{N,n}(x)\right) = \sum_{k=n+1}^{\infty} a_k^2$$

In order to find a bound for integrated squared bias, the inequality $a_k^2 \leq k^{-r}b_{k+r}^2$ is used which is derived by Walter (1977) under the condition $(x - d/dx)^r f(x)$. Where b_k is the kth coefficients of the expansion of $\left(x - \frac{d}{dx}\right)^r f$ in the Hermite series and the series $\sum b_k^2$ converges. So, by virtue of the result of Walter (1977),

$$\int_{-\infty}^{\infty} Bias^2\left(\widehat{f}_{N,n}(x)\right) = \sum_{k=n+1}^{\infty} a_k^2 = O(n^{-r}). \tag{5}$$

Since MISE can be expressed as the sum of integrated squared bias and integrated variance, by combining (4) and (5)

$$MISE(\widehat{f}_{N,n}(x)) = O\left(N^{\frac{-12r}{6r+23}}\right).$$

□

Remark 2.1. *Hermite series method was used by Schwartz (1967) and Walter (1977) and the convergency rate of MISE of the density estimator was obtained as* $O\left(N^{-\frac{(r-1)}{r}}\right)$ *in paper of Schwartz (1967) and for the estimate of the pth derivative and assuming that the density has* r *derivatives where* $0 \leq p < r$*, the MISE rate obtained as* $O\left(N^{-\frac{6(r-p)-5}{6r}}\right)$ *by Walter (1977). Note that, for comparison reasons take* $p = 0$ *for the proposed estimator and also for the estimator used by Schwartz (1967). The hypothesis are the same but faster result are obtained using delta sequence method.*

Remark 2.2. *For* $r > 2$*, the rate of convergency of MISE is better than* $O\left(N^{-\frac{2(r-p)}{2r+1}}\right)$ *which was obtained by Greblicki and Pawlak (1984).*

2.2. **Convergency Rate of MSE and MISE of Estimators for Densities Having Compact Support.** In this subsection, the rate of convergence of MSE and MISE of an estimator of densities having compact support and based on Hermite functions are investigated. Since the density function has compact support then the hypothesis weaken a little.

Theorem 2.2. *Let* f *have compact support and suppose* $\left(\frac{d}{dx}\right)^j f \in L_2$ *for* $j = 1, 2, ..., r$*; then the MSE of the estimate (3) satisfies*

$$MSE(\widehat{f}_{N,n}(x)) = O(N^{-\frac{2r+1}{2r+2}})$$

Proof. First, lets obtain bound for delta sequence by using (2)

$$|\delta_n(x,t)| = \left|\sum_{k=0}^{n} h_k(x)h_k(t)\right| \leq \sum_{k=0}^{n} |h_k(x)|\,|h_k(t)| \leq \sum_{k=0}^{n} \frac{1}{(1+k)^{1/2}} \tag{6}$$

If a convenient integral is used as the upper bound for the (6), then

$$|\delta_n(x,t)| = O\left(n^{1/2}\right) \tag{7}$$

For the variance term, since $\int \delta_m(x,t)f(t)dt \leq \|f\|_\infty$

$$Var(\widehat{f}_{N,n}(x)) \leq \frac{1}{N}\int \delta_n^2(x,t)f(t)dt \leq \frac{c_4}{N}\sup|\delta_n(x,t)| \leq c_5\frac{n^{1/2}}{N}.$$

The bias term can be written as

$$\begin{aligned} bias^2(\widehat{f}_{N,n}(x)) &= \left(\int \delta_n(x,t)f(t)dt - f(x)\right)^2 \\ &= \left(\int \sum_{k=0}^{n} h_k(x)h_k(t)f(t)dt - \sum_{k=0}^{\infty} a_k h_k(x)\right)^2 = \left(\sum_{k=n+1}^{\infty} a_k h_k(x)\right)^2 \\ &\leq \left(\sum_{k=n+1}^{\infty} b_{k+r}\frac{1}{(2k)^{\frac{r}{2}}}\frac{1}{(k+1)^{\frac{1}{4}}}\right)^2 \leq \left(\sum_{k=n+1}^{\infty} b_{k+r}k^{\frac{-2r-1}{4}}\right)^2 \\ &\leq \left((n+1)^{-\frac{2r+1}{4}} \sum_{k=n+1+r}^{\infty} b_k\right)^2 \end{aligned}$$

Note that, since f has compact support and $D^r f \in L^2$, $x^p D^s f \in L^2$ for all integers $p \geq 0$ and $0 \leq s \leq r$. So, it follows that $(x-D)^r f \in L^2$. Then, the bounds for $\left|a_k^2\right|$ obtained by Walter (1977) can be used to obtain an order for the squared bias term

$$bias^2(\widehat{f}_{N,n}(x)) = O(n^{-\frac{2r+1}{2}}).$$

So, MSE is

$$MSE(\widehat{f}_{N,n}(x)) = O\left(\frac{n^{\frac{1}{2}}}{N} + \frac{1}{n^{\frac{2r+1}{2}}}\right).$$

If $n = N^{\frac{1}{r+1}}$ is chosen, then the MSE rate of estimator based on Hermite functions is obtained as below

$$MSE(\widehat{f}_{N,n}(x)) = O(N^{-\frac{2r+1}{2r+2}})$$

□

Theorem 2.3. *Let f have compact support and suppose $\left(\frac{d}{dx}\right)^j f \in L_2$ for $j = 1, 2, ..., r$; then the MISE of the estimate (3) satisfies*

$$MISE(\widehat{f}_{N,n}(x)) = O(N^{-\frac{2r+1}{2r+2}}).$$

Proof. The proof is similar to earlier one. Notice that the bounds of delta sequence in (7) is used for the integrated variance term. □

Remark 2.3. *The rates obtained in this study are better than those reported by Walter (1977) for the densities having compact support. Moreover, the rates of convergence obtained in this work is also better than those suggested by Letellier (1997) who used delta sequence method to obtain rate of convergence of estimator based on Jacobi polynomials. It is reported that, the convergency rate of MISE and MSE as $O\left(N^{-\frac{1}{3}}\right)$ by Letellier (1997). They are considerably slower than the rates obtained in this study.*

3. Conclusions

In this study, for densities which have r derivatives, we obtained convergency rate of the MISE of estimators of densities with infinite support by using delta sequences which are based on the hermite functions. Then, we obtained convergency rate of the MSE and MISE of estimator for the densities having compact support. By using delta sequence method, we improve the results of Schwartz (1967) and Walter (1977). Also, we extend the conclusions of Schwartz (1967) to the case $r = 1$. Moreover, for $r > 2$, the rate of MISE for densities having infinite support obtained in this study is better than the rate reported by Greblicki and Pawlak (1984). However, in the paper of Letellier (1997), the negativity problem of orthogonal series estimators based on Jacobi polynomials was solved by using certain summability methods. In this work, the negativity problem of Hermite series estimators could not be solved, so it is assumed that density estimate at x is the max $\left[0, \widehat{f}_{N,n}(x)\right]$ to avoid the negative values of the estimator. So, the problem of obtaining a nonnegative orthogonal series estimator based on Hermite functions may be a challenging work for further study.

References

[1] Bleuez, J., Bosq, D., (1976), Conditions necessaires et suffisantes de convergence pour une classe d'estima teurs de la densite', C.R. Acad. Sci. Paris Ser. A, 282: 63-66.

[2] Bosq, D. (1977), Study of a class of density estimators. In: Barra, J.R. et al., eds., Recent Developments in Statistics, Amsterdam: North Holland.

[3] Chencov N.N. (1962), Estimation of an unknown density using observations, Doklady of academy of sciences, 147: 45-48 (in Russian).

[4] Greblicki W., Pawlak M., (1984), Hermite Series Estimates of a Probability Density and Its Derivatives, Journal of Multivariate Analysis, 15, 174-182.

[5] Jones, M.C., Silverman B. W., (1989), An orthogonal series density estimation approach to reconstructing positron emission tomography images, Journal of Applied Statistics, Vol.16, 177-191.

[6] Lebedev, N. N., (1965), Special functions and their applications, Prentice Hall.

[7] Letellier J. A., (1997), Rates of convergence for an estimator of a density function based on Jacobi polynomials, Communication in Statistics- Theory and Methods, 26:1, 197-220.

[8] Liebscher E. (1990), Hermite series estimators for probability densities, Metrika, 37: 321-343.

[9] Nadar M., (2011), Local convergence rate of mean squared error in density estimation, Communication in Statistics-Theory and Methods, Vol 40, pp. 176-185.

[10] Nadar, M., Erçelik E., (2017), Local convergency rate of MSE in density estimation using the second order modulus of smoothness, Communications in Statistics - Theory and Methods, 46:7, 3164-3173.

[11] Parzen, E., (1962), On the estimation of probability density and mode, Ann. Math. Statist., 33: 1065-1076.

[12] Rosenblatt, P., (1956), Remarks on some non-parametric estimates of a density functions, Ann. Math. Statist., 27: 832-837.

[13] Schwartz S. C., (1967), Estimation of probability density by an orthogonal series, Ann. Math.Statist., 27:932-937.

[14] Susarla, V., Walter, G., (1981), Estimation of a multivariate density function using delta sequences, Ann. Statist., 9: 347-355.

[15] Szegö G., (1939), Orthogonal Polynomials, American Mathematical Society.

[16] Wahba, G., (1971), A polynomial algorithm for density estimation, Ann. Math. Statist., 42: 1870-1886.

[17] Wahba, G., (1975), Optimal convergence properties of variable knot, kernel, and orthogonal series methods for density estimation, Ann. Statist., 3: 30-48.

[18] Walter G., Blum.J., (1979), Probability density estimation using delta sequences, Annals of Statistics, Vol.7, pp. 328-340.

[19] Walter G., (1977), Properties of Hermite Series Estimation of Probability Density, Annals of Statistics, Vol.5., No.6, pp.1258- 1264.

Dr. Elif Erçelik received her Ph.D degree in 2019 from Istanbul Technical University, Department of Mathematical Engineering, Istanbul, Turkey. Her area of interest focuses mainly on nonparametric density estimation problem.