



# Density and Generating Function Estimations of Random Sum

Sy-Mien Chen, Yu-Sheng Hsu

## ► To cite this version:

Sy-Mien Chen, Yu-Sheng Hsu. Density and Generating Function Estimations of Random Sum. Annales de l'ISUP, 2001, XXXXV (2-3), pp.29-45. <hal-03649090>

**HAL Id: hal-03649090**

**<https://hal.science/hal-03649090v1>**

Submitted on 22 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

*Pub. Inst. Stat. Univ. Paris*  
 XXXXV, fasc. 2-3, 2001, 29 à 45

## Density and Generating Function Estimations of Random Sum

**Sy-Mien Chen**

Department of Mathematics  
 Fu Jen University, Taipei, Taiwan 242, ROC  
 Tel: 886-3-4200197. Fax: 886-2-29044509  
 e-mail: math1013@mails.fju.edu.tw

and

**Yu-Sheng Hsu**

Department of Mathematics  
 National Central University, Chungli, Taiwan 320, ROC

*keywords and phrases:* kernel estimator, density function, generating  
 function, random sum, central limit theorem.

running head: Estimations of Random Sum

*AMS Classification:* Primary 62.

### Abstract

Given  $n$  realizations of a random sum  $S_N$ , we investigate the estimations of its probability density function  $f_{S_N}$  and generating function  $G$  by kernel and empirical methods respectively. In each case, we discuss two estimators and make comparisons based on their central limit theorems.

### 1. Introduction.

Statistical inferences for stochastic process data grow rapidly recently. Developments in this field can be found in Taniguchi and Kakizawa [14] for time series, Prakasa Rao [10] for diffusion type processes, Prakasa Rao [11] for



semimartingales, Bosq [3] for nonparametric inferences, Andersen et al. [1] for counting processes, Guttorp [7] for branching processes, and Shorack and Wellner [12] for empirical processes, among others.

In this paper, we will study some nonparametric inferences for random sum data. Let  $X_1, X_2, \dots, X_n$  denote a sequence of i.i.d. random variables with common generating function  $f$ . Define random sum

$$S_N = X_1 + X_2 + \dots + X_N,$$

where  $N$  is an integer-valued random variable with common generating function  $g$ . We assume that  $N$  and  $X_i, i = 1, \dots, n$ , are independent. Important processes like random walk, compound Poisson process, and branching process are all special cases of random sum.

Let  $f_{S_N}$  and  $G$  denote the density and generating functions of  $S_N$  respectively. For simplicity, we assume  $N$  is a positive integer-valued random variable in density estimation case. Then

$$f_{S_N}(c) = \sum_{k=1}^{\infty} g_k(c) p_k,$$

where  $g_k$  is the density function of  $X_1 + X_2 + \dots + X_k$ ,  $p_k = P(N = k)$ ,  $k$  is a positive integer, and  $G(c) = g(f(c))$  (See Feller [5] page 287).

Both density function and generating function are important in computing some interesting probabilities, see chapter 11 of Feller [5] for an interesting introduction. In practice, density and generating function are rarely known and must be inferred from samples. Given  $n$  i.i.d. observations

$$(X_1^{(i)}, X_2^{(i)}, \dots, X_{N_i}^{(i)}), i = 1, 2, \dots, n,$$

the purpose of this paper is to study the estimations of  $f_{S_N}$  and  $G$ . The rest of this paper is organized as follows. Section 2 is focused on the kernel estimation of  $f_{S_N}$ . Section 3 is devoted to the empirical estimation of  $G$ . Section 4 provides some conclusions.

## 2. Density Estimation.

In this section, we investigate the density estimation of  $f_{S_N}$  based on kernel method that is currently the most popular techniques applied in nonparametric density estimation. Comprehensive reviews can be found from influential books like Tapia and Thompson [15], Prakasa Rao [9], Silverman [13] and Pagan and Ullah [8], among others.

Let

$$T_i = X_1^{(i)} + X_2^{(i)} + \dots + X_{N_i}^{(i)}, i = 1, 2, \dots, n,$$

$$J_k = \{i | N_i = k, i = 1, 2, \dots, n\},$$

and  $M_{nk} = \sum_{i=1}^n I_{\{k\}}(N_i)$ , where  $k$  is a positive integer and  $I_{\{k\}}$  is the indicator function of  $\{k\}$ . Then  $f_{S_N}$  can be estimated by

$$\hat{f}_{DS_N}(c) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{c - T_i}{h}\right),$$

directly or by

$$\hat{f}_{IS_N}(c) = \sum_{k=1}^{\infty} \left[ \sum_{i \in J_k} \frac{1}{M_{nk}h} K_k\left(\frac{c - T_i}{h}\right) \right] \frac{M_{nk}}{n} = \sum_{k=1}^{\infty} \left[ \sum_{i \in J_k} \frac{1}{nh} K_k\left(\frac{c - T_i}{h}\right) \right],$$



indirectly, where  $K$  and  $K_k$  are kernel functions and  $h$  is the bandwidth. Clearly, if  $K = K_k$  then  $\hat{f}_{DS_N}(c) = \hat{f}_{IS_N}(c)$ . It is then interesting to investigate the effect of different kernel functions.

In order to make comparison, we will find central limit theorems for  $\hat{f}_{DS_N}(c)$  and  $\hat{f}_{IS_N}(c)$ . Assume (i)  $K$  is nonnegative and  $\int K(s)ds = 1$ ,  $\int sK(s)ds < \infty$ , (ii)  $nh \rightarrow \infty$ ,  $nh^3 \rightarrow 0$ , (iii)  $f_{S_N}$  has bounded continuous second order derivatives, then

$$\sqrt{nh}(\hat{f}_{DS_N}(c) - f_{S_N}(c)) \xrightarrow{d} N(0, f_{S_N}(c) \int K^2(t)dt),$$

by standard kernel estimation arguments. The following theorem is needed to prove the central limit theorem of  $\hat{f}_{IS_N}(c)$ .

**Theorem 1:** Let  $Y_1, Y_2, \dots, Y_n$  denote a sequence of i.i.d. random variables with common density function  $v$  such that  $v'$  is bounded. Assume  $M_n$  is a positive integer-valued random variable satisfies  $\frac{M_n}{n} \xrightarrow{p} d$  as  $n \rightarrow \infty$ . Then

$$\sqrt{ndh} \left\{ \frac{1}{M_n h} \sum_{i=1}^{M_n} K\left(\frac{c - Y_i}{h}\right) - v(c) \right\} \xrightarrow{d} N(0, v(c) \int K^2(t)dt).$$

**Proof:** First note that

$$\begin{aligned} & E \frac{1}{h} K\left(\frac{c - Y_1}{h}\right) \\ &= \int \frac{1}{h} K\left(\frac{c - y}{h}\right) v(y) dy \\ &= \int K(s) v(c - hs) ds \\ &= \int K(s) [v(c) + O(hs)] ds \\ &= v(c) + O(h), \end{aligned}$$

by the boundedness of  $v'$ , and analogously,

$$\begin{aligned}
 & E \frac{1}{h} K^2 \left( \frac{c - Y_1}{h} \right) \\
 &= \int \frac{1}{h} K^2 \left( \frac{c - y}{h} \right) v(y) dy \\
 &= \int K^2(s) v(c - hs) ds \\
 &= \int K^2(s) [v(c) + O(hs)] ds \\
 &= v(c) \int K^2(s) ds + O(h).
 \end{aligned}$$

Therefore,

$$Var \frac{1}{h} K \left( \frac{c - Y_1}{h} \right) = \frac{v(c)}{h} \int K^2(s) ds + O(1).$$

Consequently, by Anscombe theorem ( see Anscombe [2] or Ghosh et al. [6] page 45 )

$$\sqrt{ndh} \left\{ \frac{1}{M_n h} \sum_{i=1}^{M_n} K \left( \frac{c - Y_i}{h} \right) - E \frac{1}{h} K \left( \frac{c - Y_1}{h} \right) \right\} \xrightarrow{d} N(0, v(c) \int K^2(t) dt).$$

It remains to show that

$$\sqrt{ndh} \left\{ E \frac{1}{h} K \left( \frac{c - Y_1}{h} \right) - v(c) \right\} \rightarrow 0.$$

Since  $nh^3 \rightarrow 0$ , the proof is completed.  $\square$

We are now ready to prove the central limit theorem for  $\hat{f}_{IS_N}(c)$ .

**Theorem 2:** Under the assumptions (i), (ii) and (iii) with  $K_k$  replaces  $K$ ,

$$\sqrt{nh} (\hat{f}_{IS_N}(c) - f_{S_N}(c)) \xrightarrow{d} N(0, \sum_{k=1}^{\infty} g_k(c) p_k \int K_k^2(t) dt).$$

**Proof:** Denote

$$\hat{g}_k(c) = \sum_{i \in J_k} \frac{1}{M_{nk} h} K_k \left( \frac{c - T_i}{h} \right),$$



and

$$\hat{p}_k = \frac{M_{nk}}{n}.$$

Then

$$\frac{M_{nk}}{n} \xrightarrow{p} p_k$$

and theorem 1 together imply that

$$\sqrt{nhp_k}(\hat{g}_k(c) - g_k(c)) \xrightarrow{d} N(0, g_k(c) \int K_k^2(t) dt).$$

Clearly

$$\sqrt{n}(\hat{p}_k - p_k) \xrightarrow{d} N(0, p_k(1 - p_k)),$$

therefore

$$\begin{aligned} & \sqrt{nh} \left\{ \sum_{i \in J_k} \frac{1}{nh} K_k\left(\frac{c - T_i}{h}\right) - g_k(c)p_k \right\} \\ &= \sqrt{nh} \{ [\hat{g}_k(c) - g_k(c)]\hat{p}_k + g_k(c)(\hat{p}_k - p_k) \} \\ &\xrightarrow{d} N(0, g_k(c)p_k \int K_k^2(t) dt), \end{aligned}$$

and hence

$$\begin{aligned} & \sqrt{nh}(\hat{f}_2(c) - f_{S_N}(c)) \\ &= \sqrt{nh} \sum_{k=1}^{\infty} \left\{ \sum_{i \in J_k} \frac{1}{nh} K_k\left(\frac{c - T_i}{h}\right) - g_k(c)p_k \right\} \\ &\xrightarrow{d} N(0, \sum_{k=1}^{\infty} g_k(c)p_k \int K_k^2(t) dt), \end{aligned}$$

since

$$\sum_{i \in J_k} \frac{1}{nh} K_k\left(\frac{c - T_i}{h}\right) - g_k(c)p_k, k = 1, \dots, n,$$

are independent.  $\square$

It is then easy to see that  $\hat{f}_{DS_N}$  is asymptotically better(worse) than  $\hat{f}_{IS_N}$  if

$$f_{S_N}(c) \int K^2(t) dt < (>) \sum_{k=1}^{\infty} g_k(c) p_k \int K_k^2(t) dt.$$

If  $K_k = K$ , then  $\hat{f}_{DS_N}$  is asymptotically equivalent to  $\hat{f}_{IS_N}$  as mentioned above. Under mild assumptions on  $K$ , optimal  $K$  minimizing  $\int K^2(t) dt$  can be obtained as Epanechnikov [4] kernel. If we restrict  $K$  to be defined on compact interval  $[a, b]$ , then the optimal kernel  $K^*$  is given by

$$K^*(t) = \frac{1}{b-a} I_{[a,b]}(t),$$

since

$$\int_a^b K^2(t) dt \int_a^b 1 dt \geq \left[ \int_a^b K(t) dt \right]^2 = 1,$$

or equivalently,

$$\int_a^b K^2(t) dt \geq \frac{1}{b-a} = \int_a^b (K^*)^2(t) dt,$$

by Cauchy-Schwartz inequality. In practice, the optimal kernel function can be applied and hence  $\hat{f}_{DS_N}$  and  $\hat{f}_{IS_N}$  are asymptotically equivalent.

### 3. Generating Function Estimation.

In this section, we will discuss the estimation of generating function based on empirical distribution. The generating function  $Ec^{T_1} = G(c) = g(f(c))$  can be estimated by

$$\hat{G}_1(c) = \frac{1}{n} \sum_{i=1}^n c^{T_1},$$

directly or by

$$\hat{G}_2(c) = \hat{g}(\hat{f}(c)),$$



indirectly where

$$\hat{g}(c) = \frac{1}{n} \sum_{i=1}^n c^{N_i},$$

and

$$\hat{f}(c) = \frac{1}{\sum_{i=1}^n N_i} \sum_{i=1}^n \sum_{j=1}^{N_i} c^{X_j^{(i)}}.$$

Clearly,

$$\sqrt{n}[\hat{G}_1(c) - G(c)] \xrightarrow{d} N(0, A(f, g, c)),$$

where  $A(f, g, c) = G(c^2) - G^2(c)$ . The following theorem gives the asymptotic normality of  $\hat{G}_2(c)$ .

**Theorem 3:** Given  $n$  i.i.d. observations

$$(X_1^{(i)}, X_2^{(i)}, \dots, X_{N_i}^{(i)}), i = 1, 2, \dots, n,$$

$$\sqrt{n}[\hat{G}_2(c) - G(c)] \xrightarrow{d} N(0, B(f, g, c)),$$

where

$$B(f, g, c) = g([f(c)]^2) - [g(f(c))]^2 + \frac{[g'(f(c))]^2 \{f(c^2) - [f(c)]^2\}}{EN_1}.$$

**Proof:** It is easily seen that

$$\begin{aligned} & \sqrt{n}\{\hat{G}_2(c) - G(c)\} \\ &= \sqrt{n}\left\{\frac{1}{n} \sum_{i=1}^n [\hat{f}(c)]^{N_i} - G(c)\right\} \\ &= \sqrt{n}\left\{\frac{1}{n} \sum_{i=1}^n [f(c)]^{N_i} - G(c)\right\} \\ & \quad + \sqrt{n}[\hat{f}(c) - f(c)] \frac{1}{n} \sum_{i=1}^n N_i [f(c)]^{N_i-1} \\ & \quad + \sqrt{n}[\hat{f}(c) - f(c)]^2 \frac{1}{2n} \sum_{i=1}^n N_i(N_i - 1) [f^*(c)]^{N_i-2} \end{aligned}$$

$$\begin{aligned}
&= \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n [f(c)]^{N_i} - G(c) + [\hat{f}(c) - f(c)] g'(f(c)) \right\} \\
&\quad + \sqrt{n} [\hat{f}(c) - f(c)] \left[ \frac{1}{n} \sum_{i=1}^n N_i [f(c)]^{N_i-1} - g'(f(c)) \right] \\
&\quad + \sqrt{n} [\hat{f}(c) - f(c)]^2 \frac{1}{2n} \sum_{i=1}^n N_i (N_i - 1) [f^*(c)]^{N_i-2} \\
&= I + II + III,
\end{aligned}$$

where  $f^*(c)$  lies between  $f(c)$  and  $\hat{f}(c)$ .

Since

$$\frac{\sum_{i=1}^n N_i}{n} \xrightarrow{p} EN_1,$$

by Anscombe theorem and law of large numbers, we have

$$\sqrt{n} [\hat{f}(c) - f(c)] \xrightarrow{d} N(0, \frac{f(c^2) - [f(c)]^2}{EN_1}),$$

and

$$\frac{1}{n} \sum_{i=1}^n N_i [f(c)]^{N_i-1} \xrightarrow{p} g'(f(c)),$$

and hence

$$II + III \xrightarrow{p} 0.$$

It suffices to find the asymptotic normality of  $I$ . Since

$$\begin{aligned}
&I \\
&= \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n [f(c)]^{N_i} - G(c) + [\hat{f}(c) - f(c)] g'(f(c)) \right\} \\
&= \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n \{ [f(c)]^{N_i} - G(c) \} \right\} \\
&\quad + \frac{\sqrt{n} g'(f(c))}{\frac{1}{n} \sum_{i=1}^n [N_i - EN_1] + EN_1} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} \{ c^{X_j^{(i)}} - f(c) \},
\end{aligned}$$



the asymptotic normality of  $I$  comes from the fact that  $I$  is a function of

$$\left(\frac{1}{n} \sum_{i=1}^n \{[f(c)]^{N_i} - G(c)\}, \frac{1}{n} \sum_{i=1}^n \{N_i - EN_1\}, \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} \{c^{X_j^{(i)}} - f(c)\}\right),$$

that is jointly asymptotically normal ( since linear combinations of the elements of above vector is a sum of i.i.d. random variables and hence is asymptotically normal ).

The calculation of limiting variance is straightforward if

$$\frac{1}{n} \sum_{i=1}^n \{[f(c)]^{N_i} - G(c)\} = \hat{g}(f(c)) - G(c),$$

and

$$\hat{f}(c) - f(c),$$

are uncorrelated. It is true by the unbiasedness of  $\hat{f}(c)$  and  $\hat{g}(f(c))$  and the following observations

$$\begin{aligned} & E\hat{f}(c)\hat{g}(f(c)) \\ &= E[\hat{g}(f(c)) \frac{1}{\sum_{i=1}^n N_i} E(\sum_{i=1}^n \sum_{j=1}^{N_i} c^{X_j^{(i)}} | N_1, N_2, \dots, N_n)] \\ &= E[\hat{g}(f(c))f(c)] \\ &= g(f(c))f(c). \end{aligned}$$

□

The following examples illustrate the different large sample performances of  $\hat{G}_1$  and  $\hat{G}_2$ .

**Example 1:** In general, if  $A(f, g, c) < B(f, g, c)$ , then  $\hat{G}_1$  is asymptotically better than  $\hat{G}_2$ . □

For convenience, in the following examples, let  $X$  and  $T$  denote two random variables having the same distributions as  $X_j^{(i)}$  and  $T_i$  respectively.

**Example 2:** Let  $X$  and  $N$  denote Bernoulli and Poisson random variables with parameters  $p$  and  $t$  respectively. Let  $q = 1 - p$ , then  $f(c) = pc + q$ ,  $g(c) = e^{-t+tc}$ ,  $G(c) = g(f(c)) = e^{-t+t(pc+q)} = e^{-tp+tpc}$ . In this example,  $A(f, g, c)$  and  $B(f, g, c)$  depend on  $p, t$  and  $c$ , let us denote them by  $A(p, t, c)$  and  $B(p, t, c)$ . Consider the special case when  $p = \frac{1}{2}$  and  $c = \frac{1}{2}$ ,

$$\begin{aligned} A\left(\frac{1}{2}, t, \frac{1}{2}\right) &= e^{-\frac{t}{2} + \frac{t}{2} \cdot \frac{1}{2}} - (e^{-\frac{t}{2} + \frac{t}{2} \cdot \frac{1}{2}})^2 \\ &= e^{-\frac{3}{8}t} - (e^{-\frac{t}{4}})^2, \end{aligned}$$

and

$$\begin{aligned} B\left(\frac{1}{2}, t, \frac{1}{2}\right) &= g\left(f^2\left(\frac{1}{2}\right)\right) - [g(f(\frac{1}{2}))]^2 + (g'[f(\frac{1}{2})])^2 [f(\frac{1}{4}) - f^2(\frac{1}{2})] \frac{1}{t} \\ &= g\left(\frac{9}{16}\right) - g^2\left(\frac{3}{4}\right) + (te^{-t+\frac{3}{4}t})^2 \left(\frac{5}{8} - \left(\frac{3}{4}\right)^2\right) \frac{1}{t} \\ &= e^{-t+\frac{9}{16}t} - (e^{-t+\frac{3}{4}t})^2 + (te^{-t/4})^2 \frac{1}{16t} \\ &= e^{-\frac{7}{16}t} - (e^{-\frac{t}{4}})^2 + te^{-\frac{t}{2}} \frac{1}{16}. \end{aligned}$$

We will prove that  $A(\frac{1}{2}, t, \frac{1}{2}) > B(\frac{1}{2}, t, \frac{1}{2})$ , or equivalently,  $e^{\frac{t}{16}} + \frac{t}{16} < e^{\frac{2}{16}t}$ , for all  $t > 0$ . Let  $H(y) = y^2 - y - \ln y$ . It is sufficient to prove that  $H'(y) > 0$  for  $y > 1$ , since then  $H$  is increasing on  $(1, \infty)$  and



$$H(y) > H(1) = 0$$

for all  $y > 1$ . In particular,  $H(e^{\frac{1}{16}}) > 0$  as desired. Clearly,

$$\begin{aligned} H'(y) &= 2y - 1 - \frac{1}{y} \\ &= \frac{1}{y}(2y + 1)(y - 1) \\ &> 0, \end{aligned}$$

for  $y > 1$ . Therefore,  $\hat{G}_2$  is asymptotically better than  $\hat{G}_1$  in this special case.

□

The generating function is particularly useful in moment computation. The following example makes a comparison in this subject.

**Example 3:** Since

$$\hat{G}'_2(1) = \hat{g}'(\hat{f}(1))\hat{f}'(1) = \left[\frac{1}{n} \sum_{i=1}^n N_i\right] \left[\frac{1}{\sum_{i=1}^n N_i} \sum_{i=1}^n \sum_{j=1}^{N_i} X_j^{(i)}\right] = \frac{1}{n} \sum_{i=1}^n T_i = \hat{G}'_1(1),$$

the two estimators are equal in estimating  $G'(1) = ET$ .

We proceed to discuss the estimation of  $G''(1) = ET(T-1)$ . In order to make a comparison, we will find the central limit theorems of  $\hat{G}'_1(1)$  and  $\hat{G}'_2(1)$ . Clearly

$$\sqrt{n}[\hat{G}'_1(1) - G''(1)] \xrightarrow{d} N(0, \text{Var}T(T-1)).$$

To find the asymptotic normality of  $\hat{G}'_2(1)$ , first note that

$$\hat{G}''_2(1) = \hat{g}'(\hat{f}(1))[\hat{f}''(1)] + \hat{g}''(\hat{f}(1))[\hat{f}'(1)]^2$$

where

$$\hat{g}'(\hat{f}(1)) = \frac{1}{n} \sum_{i=1}^n N_i,$$

$$\hat{g}''(\hat{f}(1)) = \frac{1}{n} \sum_{i=1}^n N_i(N_i - 1),$$

$$\hat{f}'(1) = \frac{1}{\sum_{i=1}^n N_i} \sum_{i=1}^n \sum_{j=1}^{N_i} X_j^{(i)},$$

and

$$\hat{f}''(1) = \frac{1}{\sum_{i=1}^n N_i} \sum_{i=1}^n \sum_{j=1}^{N_i} X_j^{(i)}(X_j^{(i)} - 1).$$

Let  $D(w, x, y, z) = (x - w)[\frac{y}{w}]^2 + z - y$ , then

$$\hat{G}_2'''(1) = D\left(\frac{1}{n} \sum_{i=1}^n N_i, \frac{1}{n} \sum_{i=1}^n N_i^2, \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} X_j^{(i)}, \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} [X_j^{(i)}]^2\right),$$

and

$$G_2'''(1) = D(EN, EN^2, EXEN, EX^2EN).$$

( Since  $ET = EXEN$  and  $ET^2 = ENVarX + EN^2(EX)^2$ . )

It is easy to see that

$$\begin{aligned} & \sqrt{n} \left\{ \left( \frac{1}{n} \sum_{i=1}^n N_i, \frac{1}{n} \sum_{i=1}^n N_i^2, \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} X_j^{(i)}, \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} [X_j^{(i)}]^2 \right) \right. \\ & \quad \left. - (EN, EN^2, EXEN, EX^2EN) \right\} \end{aligned}$$

is asymptotically normally distributed with covariance matrix

$$\Sigma = \begin{pmatrix} VarN & Cov(N, N^2) & 0 & 0 \\ Cov(N, N^2) & VarN^2 & 0 & 0 \\ 0 & 0 & VarXEN & Cov(X, X^2)EN \\ 0 & 0 & Cov(X, X^2)EN & VarX^2EN \end{pmatrix}$$



( Here, we use the fact that the first two elements and the last two element of

$$\left(\frac{1}{n} \sum_{i=1}^n N_i, \frac{1}{n} \sum_{i=1}^n N_i^2, \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} X_j^{(i)}, \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{N_i} [X_j^{(i)}]^2\right) \\ - (EN, EN^2, EXEN, EX^2EN)$$

are uncorrelated. )

Consequently,

$$\sqrt{n}[\hat{G}_2''(1) - G''(1)] \xrightarrow{d} N(0, \sigma^*),$$

where  $\sigma^* = v \Sigma v'$ , and

$$v = \left( \frac{\partial D}{\partial w}, \frac{\partial D}{\partial x}, \frac{\partial D}{\partial y}, \frac{\partial D}{\partial z} \right) = \left( \left(1 - \frac{2x}{w}\right) \frac{y^2}{w^2}, \frac{y^2}{w^2}, (x-w) \frac{2y}{w^2} - 1, 1 \right),$$

evaluated at  $(w, x, y, z) = (EN, EN^2, EXEN, EX^2EN)$ . Thus,

$$v = ((EX)^2 \left[1 - \frac{2EN^2}{EN}\right], (EX)^2, 2\left[\frac{EXEN^2}{EN} - EX\right] - 1, 1).$$

The above results shows that if  $VarT(T-1) < \sigma^*$ , then  $\hat{G}_1''(1)$  is asymptotically better than  $\hat{G}_2''(1)$ , , the converse holds if  $VarT(T-1) > \sigma^*$ . Consider the special case when the distribution of  $N$  is Poisson with parameter  $t$  and  $P(X=1) = \frac{1}{2} = P(X=-1)$ . Then  $EX=0$ ,  $v=(0,0,-1,1)$  and

$$\Sigma = \begin{pmatrix} t & 2t^2+t & 0 & 0 \\ 2t^2+t & 4t^3+6t^2+t & 0 & 0 \\ 0 & 0 & t & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore  $\sigma^* = t$ . It is straightforward to find

$$ET^4 = 3t^2 + t, ET^3 = 0, ET^2 = t, ET = 0,$$

therefore,  $VarT(T-1) = 2t^2 + 2t$ . Since  $2t^2 + 2t > t$ ,  $\hat{G}_2''(1)$  is asymptotically better than  $\hat{G}_1''(1)$  in this special case.  $\square$

#### 4. Conclusions.

In this paper, we discuss the estimations of density and generating functions for a random sum  $S_N$ . In both cases, we present two estimators, find their central limit theorems and make comparisons. The major findings are : (i) two kernel estimators of density function  $f_{S_N}$  are asymptotically equivalent if the same kernel function and bandwidth are applied, (ii) two empirical process based estimators of generating function dominate each other in different distributions.

## References

- [1] Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993). "Statistical Models Based on Counting Processes," Springer-Verlag.
- [2] Anscombe, F. J. (1952). "Large sample theory of sequential estimation," *Proc. Cambridge Philos. Soc.*, 48, 600-607.
- [3] Bosq, D. (1998). "Nonparametric Statistics for Stochastic Processes," 2nd ed, Springer.



- [4] Epanechnikov, V. A. (1968). "Nonparametric estimation of a multidimensional probability density," *Theor. Probab. Appl.*, 14, 153-158.
- [5] Feller, W. (1968). "An Introduction to Probability Theory and Its Applications," Vol 1, 3rd edition. Academic Press.
- [6] Ghosh, M., Mukhopadhyay, N. and Sen, P. K. (1997). "Sequential Estimation," Wiley.
- [7] Guttorp, P. (1991). "Statistical Inference for Branching Processes," Wiley.
- [8] Pagan, A. and Ullah, A. (1999). "Nonparametric Econometrics," Cambridge University Press.
- [9] Prakasa Rao, B. L. S. (1983). "Nonparametric Functional Estimation," Academic Press.
- [10] Prakasa Rao, B. L. S. (1999a). "Statistical Inference for Diffusion Type Process," Arnold.
- [11] Prakasa Rao, B. L. S. (1999b). "Semimartingales and their Statistical Inference," Chapman and Hall.
- [12] Shorack and Wellner (1986). "Empirical Processes with Applications to Statistics," Wiley.
- [13] Silverman, B. W. (1986). "Density Estimation for Statistics and Data Analysis," Chapman and Hall.

- [14] Taniguchi, M. and , Kakizawa, Y. (2000). "Asymptotic Theory of Statistical Inference for Time Series," Springer.
- [15] 'Tapia, R. A. and Thompson, J. R. (1977). "Nonparametric Probability Density Estimation." Johns Hopkins University Press.

*Department of Statistics, Alborz University, Alborz, Iran*

**Abstract.** We propose a ridge estimator as an alternative estimator in the presence of collinearity among the elements of unobservable values in functional measurement error models. Modifications of the estimator are also considered and properties derived which are analogous to those known in ordinary linear models. Our simulation results show that the modified ridge estimate will have shorter length, mean square errors and standard errors than other estimates of the coefficients of the model.

**Keywords:** biased estimators, collinearity, functional relationship, parametric bootstrap, ridge parameter, variance reduction.

## 1. Introduction

Collinearity among elements and the effects of its presence on different aspects of linear models is well known (see, for instance, Belsley, 1991). When there is collinearity among the predictor variables in ordinary linear model, alternative estimation procedures have been proposed which are generally biased. Among these procedures ridge estimators have received particular attention. Hoerl and Kennard (1970, a, b) have introduced the ridge estimator and a large number of papers on the subject have been written since then. Recently, the ridge estimation techniques have been extended to generalized linear models. One can refer to Schaefer et al (1984) and Nyquist (1991) for different versions of the extended ridge techniques in generalized linear models. Gamboa and Sarda (1992) have considered collinearity in measurement error models and defined collinearity in terms of a relationship between elements of unobservable values. Apart from this work, little attention has been given to the effects of collinearity on different aspects of the estimates of parameters of functional