

Article

# Propositions for Confidence Interval in Systematic Sampling on Real Line

Mehmet Niyazi Çankaya

Department of Statistics, Faculty of Arts and Science, University of Uşak,  
Ankara-İzmir Karayolu 8.Km. 1.Eylül Kampüsü, UŞAK 64200, Turkey; mehmet.cankaya@usak.edu.tr;  
Tel.: +90-276-221-21-21 (ext. 2585)

Academic Editor: Raúl Alcaraz Martínez

Received: 14 June 2016; Accepted: 21 September 2016; Published: 28 September 2016

**Abstract:** Systematic sampling is used as a method to get the quantitative results from tissues and radiological images. Systematic sampling on a real line ( $\mathbb{R}$ ) is a very attractive method within which biomedical imaging is consulted by practitioners. For the systematic sampling on  $\mathbb{R}$ , the measurement function (*MF*) occurs by slicing the three-dimensional object equidistant systematically. The currently-used covariogram model in variance approximation is tested for the different measurement functions in a class to see the performance on the variance estimation of systematically-sampled  $\mathbb{R}$ . An exact calculation method is proposed to calculate the constant  $\lambda(q, N)$  of the confidence interval in the systematic sampling. The exact value of constant  $\lambda(q, N)$  is examined for the different measurement functions, as well. As a result, it is observed from the simulation that the proposed *MF* should be used to check the performances of the variance approximation and the constant  $\lambda(q, N)$ . Synthetic data can support the results of real data.

**Keywords:** biomedical imaging; covariogram; design-based stereology; estimation of volume; systematic sampling

## 1. Introduction

Systematic sampling, often used in the area of biomedical imaging, is a design-based approach for estimating a parameter  $Q$  of the geometrical quantities, such as volume, area, surface area and length. In the systematic sampling principle, geometrical objects are sampled with probes, such as lines, a regular grid or designed patterns. The probes superimposed on the geometrical objects are tools for us to get the quantitative values of geometrical objects. If we have an equidistant systematic sampling on geometrical objects, we will have the estimated values for each replication of the sampling of these objects. The estimated values obtained from each replication of this sampling on geometrical objects produce a fluctuation [1–5]. Therefore, this fluctuation is modeled by Fourier transformation  $\mathcal{F}_1$ , which was considered by [6,7]. Since we have the estimated values for parameter  $Q$ , the variance estimation is needed. For the systematic sampling on  $\mathbb{R}$ , also called Cavalieri sampling, previous studies shed some light on the variance estimation for the systematic sampling on a real line [8–15]. The main source in these studies is the inspiration of Matheron's theory. Matheron proposed his covariogram model  $g(h)$  intuitively. In this study, we aim to focus on testing the performance of this covariogram model for the different measurement functions in the regular and irregular forms. It is important to test the capability of fitting the performance of his covariogram model, because the sampling directions (axes), which are axial, coronal or sagittal, produce the different measurement functions, as implied by [2]. There is no study addressing the performance evaluation of the covariogram model for the different measurement functions. In this context, the works in Ref. [1–11,13,14,16–31] used Matheron's covariogram model for the different measurement functions.

After having the estimated value for the parameter  $Q$ , said as  $\hat{Q}$ , we should focus on the confidence interval for  $\hat{Q}$ . The work in Ref. [2] proposes two approaches for the coefficient of the confidence interval. The first one is based on statistical theory. In this study, the second proposition of [2] is examined. The main motivation is to find a calculation method for the constant  $\lambda(q, N)$ , which is used to construct the confidence interval for  $\hat{Q}$ . In this sense, we will give the exact calculation method for  $\lambda(q, N)$ , so we have more accurate information for the confidence interval. We also test the exact values of  $\lambda(q, N)$  for the different measurement functions.

The paper is organized as follows. Section 2 introduces the materials in the systematic sampling on  $\mathbb{R}$  and describes the exact calculation method for constant  $\lambda(q, N)$ , as well. A simulation study and real data examples are given in Section 3. Section 4 considers the discussions of the results.

## 2. Materials and Methods

The estimation for volume  $Q$ , the empirical true variance, the true variance, the variance approximation and the confidence interval for systematic sampling on  $\mathbb{R}$  with the variance approximation formula are introduced briefly.

### 2.1. Estimations for Volume $Q$ and the Variance of $\hat{Q}$

#### 2.1.1. Estimation for Volume $Q$

Suppose that we have a three-dimensional geometrical object. This object is a fixed, bounded, non-void with a piecewise smooth boundary of finite surface area of  $Q$ , except fractals. The volume value of the object is parameter  $Q$ . Let parameter  $Q$  be estimated. To get the estimation value for the parameter  $Q$ , Cavalieri planes called measurement function  $f$  are used. The function  $f$  has to represent the shape of the geometrical object. The mathematical expression for the volume estimation with Cavalieri planes is given in the following form,

$$\hat{Q}(\mathbf{u}, T) = T \sum_{j=0}^{n-1} f((\mathbf{u} + j)T). \quad (1)$$

$T$  is a constant distance among the slices obtained from the three-dimensional object.  $\mathbf{u}$  is a uniform random variable in the interval  $[0, T)$ .  $\hat{Q}$  depends on the random variable  $\mathbf{u}$  and the slice thickness or the number of systematic samplings,  $n$ .

The problem is about predicting  $Var(\hat{Q}) = \mathbb{E}(\hat{Q} - Q)^2 = \mathbb{E}(\hat{Q}^2) - Q^2$ . Since the uniform distribution defined at  $[0, T)$  is used,  $\mathbb{E}(\hat{Q}^2) = 1/T \int_0^T \hat{Q}^2 du$ . In following subsections, we will introduce some important formulae for the variance estimation and its counterparts [6–15,18–22].

#### 2.1.2. Variance of $\hat{Q}$

The behavior of the variance of the Cavalieri estimator is strongly connected to analytical properties of the measurement function  $f$ . An aspect coming from Matheron's transitive theory is given in [8] for these properties. The variance of Cavalieri sampling changes with a fractional power of  $T$  [2,9]. The fact that there is a fractional power for  $T$  is not pointed out by [13,14]. In this sense, [9] is an extension of [13]. For this reason, we want to test the performance of the covariogram model in Equation (12) and the variance extension term in Equation (20) for the different measurement functions.

For practical purposes, since the true  $MF$  is not known, we do not get the true variance of systematic sampling on  $\mathbb{R}$ . For this reason, the variance approximation known as a variance extension ( $Var_E(\hat{Q})$ ) term is given. It is in the two forms coming from Fourier transformation basically given in [8] and the properties of measurement functions proposed by [9].

Firstly, we will introduce the empirical true variance for  $\hat{Q}$ . Secondly, the true variance for  $\hat{Q}$  is defined in Equation (5). It includes the covariogram function. After the covariogram function is defined, the measurement functions will be introduced.

The empirical true variance of systematic sampling on  $\mathbb{R}$  is calculated by the formula given below,

$$Var(\hat{Q}) = \frac{1}{m} \sum_{r=1}^m (\hat{Q}_r - Q)^2, \tag{2}$$

where  $m$  is the number of replication for the systematic sampling on  $\mathbb{R}$ .  $\hat{Q}_r$  is an estimated value at a uniformly random starting point  $u$  for the area.  $Q$  is the true value for the area under the MF [16,18,32].

The empirical true variance of systematic sampling on  $\mathbb{R}$  is also:

$$CE^2(\hat{Q}) = Var(\hat{Q}/Q), \tag{3}$$

which is in the context of the coefficient of the error square.

The coefficient of error for the empirical true variance is calculated by:

$$CE(\hat{Q}) = \sqrt{Var(\hat{Q})/Q^2}. \tag{4}$$

In the simulation section, Equation (4) is used to test the performance of the variance approximation formula in Equation (20). The true variance is calculated by a formula in Equation (5). A calculation of the true variance by means of Equation (5) is intractable (see Ref. [8] for more details). The true variance is comprised of three components, which are the variance extension term, Zitterbewegung and the higher-order terms, respectively. These terms were obtained in Ref. [9] (see Section 6 in Ref. [9] for detailed expressions).

$$Var(\hat{Q}) = T \sum_{k=-\infty}^{\infty} g(kT) - \int_{-\infty}^{\infty} g(h)dh, \tag{5}$$

where

$$g(h) = \int_{\mathbb{R}} f(x) \cdot f(x+h)dx, \quad h \in \mathbb{R} \tag{6}$$

is a covariogram function of the measurement function  $f$ . It is proposed by Matheron's transitive theory and is known to be the convolution of  $f$  with its reflection,  $G(t) = \mathcal{F}_1 g = (\mathcal{F}_1 f)(\overline{\mathcal{F}_1 f})$ ;  $\mathcal{F}_1$  expresses the Fourier transform defined as:

$$G(t) = \mathcal{F}_1 g(h) = \int_{-\infty}^{\infty} g(h)exp(-2\pi i t h)dh. \tag{7}$$

Note that the subscript is for the dimension. In this case, we are interested in the one-dimensional systematic sampling.

Suppose that we have measurement functions in Equations (8)–(11). These functions are used for the simulation. They can represent biological objects. They are given with the following forms, and the function  $f$  has positive values for each value of variable  $x$ , namely  $f : \mathbb{R} \rightarrow \mathbb{R}^+$ .

$$f(x) = (1 - x^2)^q, \quad x \in [-1, 1], \quad q \in [0, 1], \tag{8}$$

$$f(x) = ((1 - \cos(x))(1 - x^2))^q, \quad x \in [-1, 1], \quad q \in [0, 1], \tag{9}$$

where  $q$  is a parameter of the smoothness constant.

$$f(x) = exp(-sin(-x^3)), \quad x \in [-38\pi/100, 53\pi/100], \tag{10}$$

and

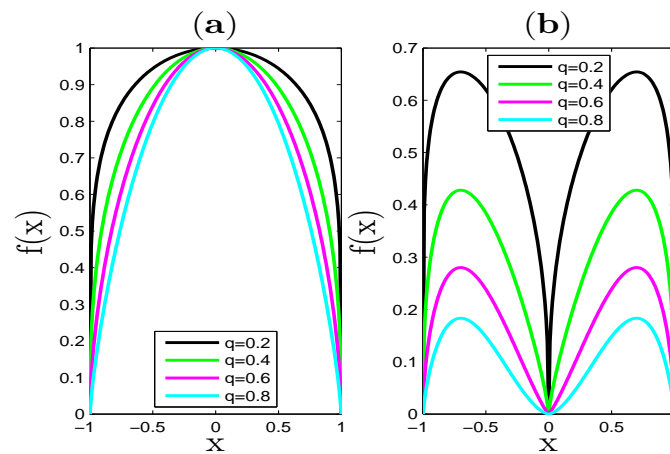
$$f(x) = (5/112)(-54x^4 - 25x^3 + 48x^2 + 25x + 6), \quad x \in [-1, 1]. \tag{11}$$

They represent the area measured on the each slice of three-dimensional biological objects. The measurement functions in Equations (8) and (11) are used by [9,33], respectively. Each function  $f$  and its covariogram function can be an example for the representation of  $q$ , even though the function  $f$  does not have the parameter  $q$ . The two approximately same functions can be proposed,  $f_i \approx f_j$ , where  $i \neq j, i, j = \mathbb{N}$ , which have different mathematical expressions; however, they can be a neighborhood of each other. For example, the measurement functions in Equations (10) and (11) can be a more modified version of  $MF$  in Equation (9).  $MF$  in Equation (9) can be a modified version of  $MF$  in Equation (8).

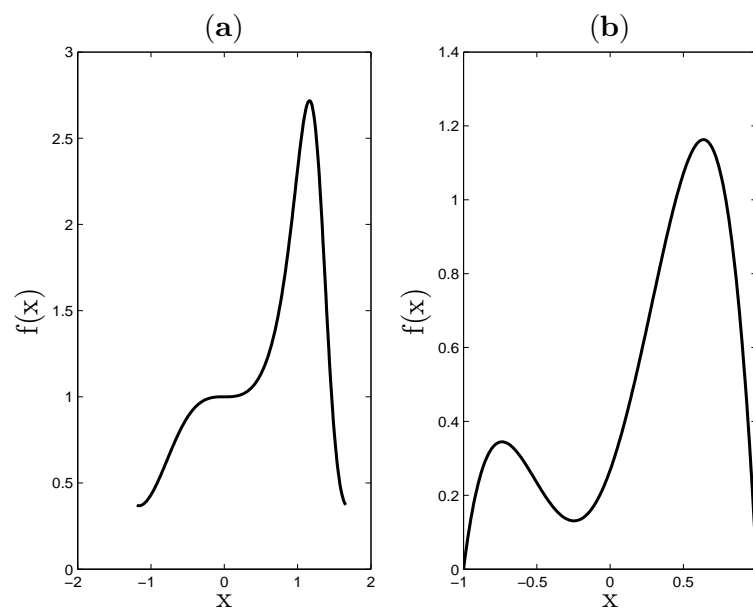
Since the covariogram functions of measurement functions in Equations (8)–(11) (displayed by Figures 1 and 2) cannot be calculable to get the integral values, a model for the covariogram functions should be proposed. In this sense, the covariogram functions can be modeled by a polynomial with the fractional power,

$$g(h) = b_0 + b_1|h|^j + b_2h^2, \quad j = 2q + 1, \quad q \in [0, 1]. \tag{12}$$

Equation (12) is defined to be a covariogram model [6–9,18].



**Figure 1.** Measurement functions with parameter  $q$ . (a) Measurement function ( $MF$ ) of Equation (8); (b)  $MF$  of Equation (9).



**Figure 2.** Measurement functions without parameter  $q$ . (a)  $MF$  of Equation (10); (b)  $MF$  of Equation (11).

### 2.1.3. Approximate Variance of $\hat{Q}$

By using Equation (5) and the properties of  $G(t)$ , we finally get:

$$Var(\hat{Q}) = T^{-1} \int_0^T \hat{Q}^2 du - Q^2 = \sum_{k=-\infty}^{\infty} G(k/T) - G(0) = 2 \sum_{k=1}^{\infty} G(k/T) \tag{13}$$

(detailed expressions are given in Ref. [6–9]).

The Fourier transformations of parts  $h^0$  and  $h^2$  in the covariogram model in Equation (12) are zero. Thus, the variance extension term of systematic sampling on  $\mathbb{R}$  is obtained by using the formula given in Equations (9.1) and (9.2) in [8] as follows,

$$\mathcal{F}_1|h|^j = b(j,1)e^{-(j+1)}, \quad (j > -1, \text{ non-even}) \tag{14}$$

which is the Fourier transformation of  $|h|^j$  for the one-dimensional systematic sampling.  $b(j,1) = \pi^{-j-1/2} \frac{\Gamma(\frac{j+1}{2})}{\Gamma(-\frac{j}{2})}$ .

By using Equations (12)–(14), Equation (15) is obtained.

$$\begin{aligned} Var_E(\hat{Q}) &= 2b_j b(j,1) T^{j+1} \zeta(j+1) \\ Var_E(\hat{Q}) &= \alpha(q)(3g(0) - 4g(T) + g(2T)) \\ Var_E(\hat{Q}) &= \alpha(q)(3C_0 - 4C_1 + C_2) T^2 \end{aligned} \tag{15}$$

where

$$\alpha(q) = \frac{2\pi^{-(2q+3/2)} \Gamma(q+1) \zeta(2q+2)}{(2^{2q+1} - 4) \Gamma(-\frac{2q+1}{2})}, \quad q \in \mathbb{R}, \tag{16}$$

$$\hat{g}(kT) = TC_k, \tag{17}$$

$\Gamma, \zeta$  are Gamma and Riemann’s Zeta functions in [34], respectively.  $\hat{g}(kT)$  is unbiased point estimators of  $g(kT)$ ,  $k = 0, 1, 2, \dots$

Equation (16) is obtained according to Fourier transformation basically given in [8]. However, Equation (18) is obtained according to the generalized version of the refined Euler–MacLaurin, proposed by [9].

$$\alpha(q) = \frac{\Gamma(2q+2) \zeta(2q+2) \cos(q\pi)}{(2\pi)^{2q+2} (1 - 2^{2q-1})}, \quad q \in [0, 1], \tag{18}$$

$$Var_E(\hat{Q}) = \frac{2\pi^{-(2q+3/2)} \Gamma(q+1) \zeta(2q+2)}{(2^{2q+1} - 4) \Gamma(-\frac{2q+1}{2})} (3C_0 - 4C_1 + C_2) T^2. \tag{19}$$

Note that Equations (16) and (18) give the same results, because the connection between the generalized version of the refined Euler–MacLaurin summation formula and Matheron’s transitive theory is expressed by [9].

We calculate the estimated coefficient of error for the systematic sampling on  $\mathbb{R}$  by means of Equation (20) given below,

$$\begin{aligned} \widehat{var}_E(\hat{Q})/Q^2 &= \alpha(q)(3C_0 - 4C_1 + C_2) T^2 / Q^2 \\ \hat{ce}(\hat{Q}) &= \sqrt{\alpha(q)(3C_0 - 4C_1 + C_2)} (\sum_{i=1}^n f_i)^{-1} \end{aligned} \tag{20}$$

where

$$C_k = \sum_{i=1}^{n-k} f_i \cdot f_{i+k}, \quad k = 0, 1, \dots, n - 1. \tag{21}$$

This is defined to be a coefficient of error of Matheron’s covariogram model. Equations (6) and (21) have in fact inheritance.  $n \geq 2k + 1$  observations are required due to Equation (21) [6–9,17,20,21].

The estimation values of parameter  $q$  from  $(1 - x^2)^q$ ,  $((1 - \cos(x))(1 - x^2))^q$ ,  $\exp(-\sin(-x^3))$  and  $(5/112)(-54x^4 - 25x^3 + 48x^2 + 25x + 6)$  measurement functions will be obtained at the simulation section. The works in Ref. [1–11,13,14,16–31] used the covariogram model for the different measurement functions. We aim to show the performance of the covariogram model in Equation (12) for the different measurement functions. In other words, the capability of the covariogram model for fitting performance on the covariogram functions is tested, so the information gained by this model is displayed by the simulation results. By using this model and the variance extension term in Equation (15), the estimation formula of parameter  $q$  is given in the following subsection.

### 2.1.4. Estimation Formula of Parameter $q$

The estimation formula of smoothness constant  $q$  is proposed by [9]. In this subsection, we will get it via the variance extension term in Equation (15) in the framework of the Fourier transformation. The formula we got for  $q$  is the same as the formula in [9].

The covariogram model  $g$  can be declared by integer  $k$  values. If  $h = iT$  is near zero, Equation (7) has more information. By using Equations (12) and (15), respectively,

$$g(iT) = b_0 + b_{2q+1}(|iT|)^{2q+1} + b_2(iT)^2, \quad i = 0, k, 2k,$$

$$\begin{aligned} Var_E(\hat{Q}) &= 2b_{2q+1}b(2q + 1, 1)T^{2q+2}\zeta(2q + 2) \\ Var_E(\hat{Q}) &= 2\pi^{-(2q+3/2)} \frac{\Gamma(q+1)}{\Gamma(-\frac{2q+1}{2})} \zeta(2q + 2) \\ &\quad \frac{T \cdot 3g(0) - 4g(kT) + g(2kT)}{k^{2q+1}(2^{2q+1} - 4)} \end{aligned} \tag{22}$$

can be obtained. By using Equations (19) and (22),

$$\frac{1}{k^{2q+1}} [3g(0) - 4g(kT) + g(2kT)] = [3g(0) - 4g(T) + g(2T)] \tag{23}$$

is obtained and then:

$$q = \frac{1}{2\log(k)} \cdot \log\left(\frac{3g(0) - 4g(kT) + g(2kT)}{3g(0) - 4g(T) + g(2T)}\right) - \frac{1}{2}, \tag{24}$$

where,  $k = 2, 3, \dots$

By using Equation (17),

$$\hat{q} = \frac{1}{2\log(k)} \cdot \log\left(\frac{3C_0 - 4C_k + C_{2k}}{3C_0 - 4C_1 + C_2}\right) - \frac{1}{2}, \quad k = 2, 3, \dots \tag{25}$$

The estimator  $\hat{q}$  is obtained when the MF is obtained by planimetry. The planimetry is also called automatic pixel counting given in the section for real data.

## 2.2. Confidence Interval in Systematic Sampling on $\mathbb{R}$

We will give an exact calculation of  $\lambda(q, N)$  in a generalized version of the refined Euler–MacLaurin summation formula with a fractional power of measurement functions. A brief introduction and some tools will be given in the next subsection to get the formula of  $\lambda(q, N)$  for the confidence interval of  $\hat{Q}$ .

### 2.2.1. Confidence Interval of $\hat{Q}$ : $\lambda(q, N)$

Some definitions expressed in Ref. [2,9] will be given. A theorem will be proposed for a tool in the constant  $\lambda(q, N)$ .

A bounded interval for the difference  $(\hat{Q} - Q)$ , which is defined as a generalized version of the refined Euler–MacLaurin summation formula, is:

$$|\hat{Q} - Q| \leq T^{q+1} P_{q+1}^* \sum_{i=1}^N |Sf^{(q)}(a_i)| \tag{26}$$

where

$$P_{q+1}^* = \max_{\{\Delta, \beta\}} \left\{ \left| \frac{-2}{(2\pi)^{q+1}} \sum_{j=1}^{\infty} \frac{1}{j^{q+1}} \cos(2\pi\Delta j - \frac{\pi}{2}(q+1) + \beta) \right| \right\}, \tag{27}$$

$\Delta \in [0, 1)$  and  $\beta \in [0, 2\pi]$ .  $\{a_i, i = 1, 2, \dots, N\}$  is the set of a real line at which function  $f^{(q)}$  exhibits discontinuities.  $N$  is the number of sets. The difference  $\hat{Q} - Q$  and  $(Sf^{(q)}(a_i))^2$  are given by Equations (6) and (11) in Ref. [2]. If  $N$  is greater than 2, the information in  $\hat{Q} - Q$  decreases (for detailed expressions, see Ref. [2]). The preliminaries were given. Now, we will use the tools in Ref. [2,9]. After some straightforward calculations in the following steps, the formula of  $\lambda(q, N)$  will be obtained.

By means of the Cauchy–Schwarz inequality  $(\sum_{i=1}^N |y_i|)^2 \leq N \sum_{i=1}^N y_i^2$ , where  $y_i = Sf^{(q)}(a_i)$ ,

$$\left( \sum_{i=1}^N |Sf^{(q)}(a_i)| \right)^2 \leq N \sum_{i=1}^N (Sf^{(q)}(a_i))^2. \tag{28}$$

When Equations (26) and (28) are used,

$$\begin{aligned} \frac{|\hat{Q} - Q|}{T^{q+1} P_{q+1}^*} &\leq \sum_{i=1}^N |Sf^{(q)}(a_i)| \\ |\hat{Q} - Q| &\leq T^{q+1} P_{q+1}^* \sqrt{N} \sqrt{\sum_{i=1}^N (Sf^{(q)}(a_i))^2} \end{aligned} \tag{29}$$

is obtained.

The following formula having the parameter  $q$  is proposed by [9],

$$Var_E(\hat{Q}) = T^{2q+2} \frac{P_{2q+2, T}(0)}{\cos(\pi q)} \sum_{i=1}^N (Sf^{(q)}(a_i))^2. \tag{30}$$

When  $\sum_{i=1}^N (Sf^{(q)}(a_i))^2$  in Equation (30) is written in Equation (29), Equation (31) is obtained.

$$|\hat{Q} - Q| \leq P_{q+1}^* \sqrt{N} \sqrt{\frac{\cos(\pi q)}{P_{2q+2, T}(0)}} \sqrt{Var_E(\hat{Q})}, \tag{31}$$

$$|\hat{Q} - Q| \leq \lambda(q, N) \sqrt{Var_E(\hat{Q})},$$

$$\lambda(q, N) = P_{q+1}^* \sqrt{N} \sqrt{\frac{\cos(\pi q)}{P_{2q+2, T}(0)}}, \quad q \in [0, 1]. \tag{32}$$

$\lambda(q, N)$  is a function of  $q$  and  $N$ . In the following steps, we will give a definition for the function  $P_{k, T}$  in [34]. The function  $P_{q+1}^*$  in Equation (27) is required to apply Theorem 1 for the exact calculation of  $\lambda(q, N)$ .

From Equation (31), for true parameter  $Q$ , a bounded interval (or 100% confidence interval) is given as [2,3]:

$$\left( \hat{Q} - \lambda(q, N) \sqrt{Var_E(\hat{Q})}, \hat{Q} + \lambda(q, N) \sqrt{Var_E(\hat{Q})} \right). \tag{33}$$

We need the following periodic function with period  $T$  to get the values of constant  $\lambda(q, N)$  [9,34],

$$P_{k, T}(x) = P_k \left( \frac{x}{T} - \left[ \frac{x}{T} \right] \right) = \frac{-2}{(2\pi)^k} \sum_{j=1}^{\infty} \frac{\cos(2\pi j(x/T) - (1/2)\pi k)}{j^k}, \tag{34}$$

where  $x \in \mathbb{R}$ ,  $k = 2, 3, \dots$

From Equation (34),  $P_{2q+2,T}(0)$  is obtained,

$$\begin{aligned}
 P_{2q+2,T}(0) &= \frac{-2}{(2\pi)^{2q+2}} \sum_{j=1}^{\infty} \frac{\cos(2\pi j(0/T) - (1/2)\pi(2q+2))}{j^{2q+2}} \\
 P_{2q+2,T}(0) &= \frac{2}{(2\pi)^{2q+2}} \zeta(2q+2) \cos(\pi q).
 \end{aligned}
 \tag{35}$$

We will give the following theorem to get the values  $P_{q+1}^*$  exactly.

**Theorem 1.**

1. Supposing that we have a right-open side interval  $[a, b)$ , the maximum value of the interval is:

$$\lim_{h \rightarrow 0} (b - h) = \lim_{h \rightarrow \infty} (b - 1/h).$$

2. Supposing that we have a left-open side interval  $(a, b]$ , the maximum value of the interval is:

$$\lim_{h \rightarrow 0} (a + h) = \lim_{h \rightarrow \infty} (a + 1/h).$$

3. Supposing that we have a left-open side and right-open side interval  $(a, b)$ , the maximum value of the interval is:

$$\begin{aligned}
 \lim_{h \rightarrow 0} (a + h) &= \lim_{h \rightarrow \infty} (a + 1/h), \\
 \lim_{h \rightarrow 0} (b - h) &= \lim_{h \rightarrow \infty} (b - 1/h).
 \end{aligned}$$

In order to calculate  $P_{q+1}^*$  in Equation (27) exactly,  $\Delta$  is replaced with  $1 - h$ , because we must take the maximum value of  $\Delta$ . When  $h \rightarrow 0$ , we get the result for the function  $P_{q+1}^*$  given in the following form,

$$P_{q+1}^* = \lim_{h \rightarrow 0} \left| \frac{-2}{(2\pi)^{q+1}} \sum_{j=1}^{\infty} \left[ \frac{1}{j^{q+1}} \cos(2\pi(1-h)j - \frac{\pi}{2}(q+1) + 2\pi) \right] \right|. \tag{36}$$

Now, we can explain how to get the values of constant  $\lambda(q, N)$  below. In the following lines, we will give the limit values done by Mathematica 7, 8 or 9.

When  $N$  and  $P_{2q+2,T}(0)$  given in Equation (32) are replaced with 2 and Equation (35), respectively, we get an equation,

$$\lambda(q, N) = P_{q+1}^* \sqrt{2} \sqrt{\frac{(2\pi)^{2q+2}}{2\zeta(2q+2)}}. \tag{37}$$

The codes written in Mathematica 7 or higher versions 8 and 9 do not compute the value of  $\lambda(0, 2)$  and give infinity. Therefore, for the calculation of  $q = 0$ , we use Equation (38) in order to give the result.

By means of Equation (37), we get  $\lambda(0, 2)$ :

$$\begin{aligned}
 P_{q+1}^* &= \lim_{h \rightarrow 0} \left| \frac{-2}{(2\pi)} \sum_{j=1}^{\infty} \frac{1}{j} \cos(2\pi(1-h)j + \frac{3\pi}{2}) \right| = 1/2 \\
 \lambda(0, 2) &= \frac{1}{2} \cdot \sqrt{\frac{(2\pi)^2}{\zeta(2)}} = 2.44949.
 \end{aligned}
 \tag{38}$$

By using Equation (37), for  $q = 0.1$ ,

$$\begin{aligned}
 \lambda(0.1, 2) &= \lim_{h \rightarrow 0} \left| \frac{-2}{(2\pi)^{1.1}} \sum_{j=1}^{\infty} \frac{1}{j^{1.1}} \cos(2\pi(1-h)j - \frac{\pi}{2}(1.1) + 2\pi) \right| \cdot \sqrt{\frac{(2\pi)^{2.2}}{\zeta(2.2)}} \\
 &= 2.71243
 \end{aligned}
 \tag{39}$$

is obtained. For other  $q$  values, a similar procedure is performed. By using the formula given in Equation (32), the values in Table 1 are obtained for  $N = 2, N = 3$  and  $N = 4$  [35].



It is attested that  $\lim_{h \rightarrow \infty} (1 - 1/h)$  and  $\lim_{h \rightarrow 0} (1 - h)$  give the same values for  $\lambda(q, N)$ , which is used to define the right-open side interval.

Moreover, in order to construct the confidence interval for the dataset,

$$\hat{Q} - \lambda(q, N) \cdot \hat{ce}(\hat{Q}) \cdot \hat{Q} \leq Q \leq \hat{Q} + \lambda(q, N) \cdot \hat{ce}(\hat{Q}) \cdot \hat{Q} \tag{40}$$

are used, where  $\lambda(q, N)$  is the coefficient of the confidence interval.  $\hat{ce}$  is an estimated coefficient of error [2,3].

### 2.2.2. Exact and Computational Values of $\lambda(q, N)$

We got the values of  $\lambda(q, N)$  for different  $q$  and  $N$  values by using the exact calculation. This calculation was done by the Mathematica 7 or the higher Versions 8 and 9. One can get the values of  $\lambda(q, N)$  for the different  $q$  values with  $N = 2, N = 3$  and  $N = 4$ . When the number of sequences in codes prepared by [2] is increased, the computational values of  $\lambda(q, N)$  converge to the exact values of  $\lambda(q, N)$  (see Tables 1 and 2). Because getting the computational values of  $\lambda(q, N)$  for  $N = 2$  would be unlikely to be useful, we give up computing some of them.  $S_i$ : the number of sequences increased by a user,  $i : 1, 2, 3, 4$ .

**Table 1.**  $q$  and its exact  $\lambda(q, N)$  values for different  $N$  values.

$q$	$\lambda(q, 2)$	$\lambda(q, 3)$	$\lambda(q, 4)$
0	2.44949	-	-
0.1	2.71243	3.32203	3.83595
0.2	2.93821	3.59855	4.15525
0.3	3.12464	3.82689	4.41891
0.4	3.26925	4.004	4.62342
0.5	3.36968	4.12699	4.76544
0.6	3.42394	4.19345	4.84218
0.7	3.43064	4.20165	4.85165
0.8	3.38906	4.15073	4.79285
0.9	3.29929	4.04079	4.6659
1	3.16228	3.87298	4.47214

**Table 2.**  $q$  and its computational  $\lambda(q, N)$  values for  $N = 2$ .

$q$	$S_1 : \lambda(q, N)$	$S_2 : \lambda(q, N)$	$S_3 : \lambda(q, N)$	$S_4 : \lambda(q, N)$
0	2.402871	2.418465	-	-
0.1	2.671264	-	-	-
0.2	2.89612	-	-	2.926749
0.3	3.07818	3.105091	3.111937	3.114245
0.4	3.216635	-	-	3.258353
0.5	3.323898	-	-	3.365038
0.6	3.397839	3.408876	3.412577	3.414001
0.7	3.418973	-	-	3.426855
0.8	3.383912	-	3.38736	3.387638
0.9	3.297052	3.29831	3.298648	3.298765
1	3.161317	-	-	-

## 3. Results

This section will give the simulation results for the estimation of parameter  $q$  in the variance extension term in Equation (20). Together with the estimated  $\hat{q}$  with  $k = 2$  for Equations (8)–(11), it is planned to see whether or not the confidence interval includes the true value of volume. While constructing the confidence interval, the approximate variance based on Matheron’s covariogram

model and the empirical true variance estimations are used. Real data examples are given to test the performance of the confidence interval.

### 3.1. Simulation

#### 3.1.1. Plan and Output of Simulation

*CEMC*, *CEET*, *CPCEMC* and *CPCEET* are the abbreviations for the coefficient of error of Matheron's covariogram model, the coefficient of error for the empirical true variance, the coverage probability of confidence interval based on the coefficient of error for the Matheron's covariogram model and the coverage probability of confidence interval based on the coefficient of error for the empirical true variance, respectively. The last two abbreviations are for the coverage probabilities of confidence intervals based on *CEMC* and *CEET*. If the true value  $Q$  is in the confidence interval, the confidence interval is successful. The numbers of successes for each replication of sampling are counted and divided by the number of replications. Thus, the values of the coverage probability of the confidence interval for a parameter  $Q$  are computed.  $\widehat{Var}(\hat{q})$  and  $\widehat{MSE}(\hat{q})$  are the simulated variance and the simulated mean squared error of the estimator  $\hat{q}$ , respectively. The  $k$  in  $q$  formula was taken to be two, because  $h = iT$  must be near zero in order to increase the information in Equation (7).

In the simulation performed, the number of replication is 3000, and the numbers of systematic samplings,  $n$ , on  $\mathbb{R}$  are from 5–20 for the measurement functions in Equations (8)–(11). The number of systematic samplings is taken to be 20 at most for practical purposes.

The following lines represent the produced  $x$  values of *MF*, the values of *MF*, the estimated values of parameter  $Q$  and the coefficient of error for the empirical true variance in three blocks given below.

```
FOR i FROM 1 TO the number of replication
  FOR j FROM 0 TO the number of systematic sampling minus one
    T(j)=Length of interval of domain of function / the number of systematic sampling
    FOR k FROM 1 TO the number of systematic sampling
      x(k,j,i)=Left end point of the domain of the measurement function+(u(j,i)+j)T(j)
      y(k,j,i)=Measurement function(x(k,j,i))
    END FOR
  END FOR
END FOR

FOR i FROM 1 TO the number of replication
  FOR j FROM 1 TO the number of systematic sampling
    Estimated Q in replicated sampling=(Summation of y in replicated sampling)T(j)
  END FOR
END FOR

FOR i FROM 1 TO the number of systematic sampling
  Summation of empirical variance in each number of systematic sampling is assigned to be zero
  FOR j FROM 1 TO the number of replication
    Empirical variance in each replication=(Estimated Q in each replication-True value)^2;
    Summation of empirical variance in each replication
  END FOR
  Mean of empirical variance=Summation of empirical variance in each replication/The number of replication
  The coefficient of error for the empirical true variance=The root of mean of empirical variance/True value
END FOR
```

#### 3.1.2. Estimations for Parameter $q$ and Variance Approximation of $\hat{Q}$

As implied by p. 319 in [8], it is obvious that the changing of *MF* affects the variance extension term given in Equation (15). In this sense, the covariogram model in Equation (12) can be adopted for

some measurement functions. The simulation results in Tables A1–A12 show that the covariogram model can be used for the measurement functions given in Equations (8), (9) and (11). Note that the integral of the function in Equation (9) is computed by means of the numerical integration in MATLAB 2013a, but the integrals of the functions in Equation (8), (10) and (11) are computed by means of the 'int' function, which is a function for the exact calculation of the integral in MATLAB 2013a. The covariogram model in Equation (12) is proposed intuitively by [6,7]. We want to use it to check the performance of variance estimation in Equation (20). By using the covariogram model in Equation (12), the estimation of  $q$  can be said to be an open problem for the measurement functions in Equations (8)–(11), except Equation (8) with  $q = 0.4$ . It is observed from the simulation results that the performance of the  $q$  formula depends on the covariogram model in Equation (12) and the variance extension term in Equation (20). For irregular  $MF$ , such as Equations (10) and (11), the estimations of  $q$  are unstable, namely, they change according to the sample size. It should be noted that for each sample size, the information on the equidistant systematic sampling version of  $f$  changes. When we have regular patterns, such as Equations (8) and (9), the estimations of  $q$  are stable according to the sample size.

Suppose that we have a flexible balloon filled with gas. This three-dimensional balloon can have different shapes after modification done by hand. Thus, each measurement function can be in a class for the functions with parameter  $q$ . As is seen, Equations (10) and (11) do not have the parameter  $q$ , but we want to estimate it to get the values of variance estimation precisely for them. The true parameter values of them are accepted to be  $q = 0.95$ ,  $q = 0.9$ , respectively, because it seems that the estimated values of parameter  $q$  are around these values under Equations (12) and (20).

The estimated values for  $q$  of the measurement functions in Equation (8) with  $q = 0.8$  and also Equation (9) with  $q = 0.4$  and  $q = 0.8$  would be around the true parameter values. It is seen that the covariogram model in Equation (12) cannot be a very good representative for these  $q$  values of two measurement functions. However, when it is considered on the performance of variance estimation, the values of  $CEMC$  and  $CEET$  for the systematic sampling of Equations (8), (9) and (11) contain the same numbers of zero digits in the fraction part the most times under Equations (12) and (20). It should be noted that even though  $MF$  in Equation (11) does not have the parameter  $q$ , the values of  $CEMC$  and  $CEET$  contain the same number of zeros in the fraction part when the sample sizes are from 6–20. When  $n = 5$ , there is not enough information due to the irregular form of  $MF$ . It is seen from the simulation results that the values of  $CEMC$  for  $MF$  in Equation (11) are more approximate than that of Equations (8) and (9) generally, which shows evidence that we can propose a function without a parameter  $q$ , which is a neighborhood of functions with parameter  $q$ . In other words, we have a family for the covariogram functions. This family can be modeled by the covariogram model. This is one of the reasons why the proposed  $MF$  must be used. The efficiency for the systematic sampling of  $MF$  in Equation (10) is not as good as that of Equation (11) because of Equations (12) and (20). The covariogram function of Equation (10) cannot be modeled by the covariogram model. The precision of the systematic sampling of  $MF$  in Equation (10) should be improved. This shows us that the proposed measurement functions are also used while making research into the applied science when Equation (20) is used. As a result, for Equations (8), (9) and (11), Equations (12) and (20) give the values around  $CEET$ ; however, for Equation (10), the values of  $CEMC$  and  $CEET$  do not contain the same number of zero digits in the fraction part even if the sample size increases. The fluctuation of  $CEET$  is an expected result, because the idea of estimation for parameter  $Q$  via the systematic sampling in Equation (1) is used.

### 3.1.3. Confidence Interval of $\hat{Q}$ : Empirical True and Approximation for $Var(\hat{Q})$

The theoretical coverage probability of the confidence interval for a parameter  $Q$  was 100%. In this sense, the coverage probabilities of confidence intervals based on the coefficient of error for Matheron's covariogram model are 100%; but the coverage probabilities of the confidence intervals based on the coefficient of error for the empirical true variance cannot be 100%. It is observed that

the coverage probabilities of confidence intervals based on the coefficient of error for the empirical true variance are 95% approximately. The benefit of using the covariogram model with parameter  $q$  is that the bandwidth of confidence can be more accurate for a class of measurement functions, such as Equations (8), (9) and (11).

### 3.1.4. Confidence Interval of $\hat{Q}$ : $\lambda(q, N)$

*CEMC* can be approximated to *CEET* for Equations (8), (9) and (11), and so, the coverage probabilities of the confidence intervals based on *CEMC* and *CEET*, which are called *CPCEMC* and *CPCEET*, respectively, can be considered to have similar values, which shows that our investigation on the performance of  $\lambda(q, N)$  values is reasonable. In other words, when we look at the performance of the confidence interval for whether or not it includes the true value, we should focus on *CPCEMC* and *CPCEET*, which must give similar results together. To this point, the trustiness of the confidence interval is acceptable, because *CPCEET* is at least 95% approximately. In all performed simulations,  $N$  in  $\lambda(q, N)$  is taken to be two, which shows us that the confidence interval keeps itself in the optimal sense. This optimality is supported by *CPCEMC* and *CPCEET*, because they have similar results. In other words, suppose that if  $N$  is taken to be three, the bandwidths of the confidence interval in the  $\lambda(q, 3)$  case are larger than that of  $\lambda(q, 2)$ . In such a case, we would have non-useful information for the confidence band of  $\hat{Q}$ . Another point we should focus on is the values in Tables 1 and 2, because the exact values of constant  $\lambda(q, N)$  provide useful information for the confidence band. For *MF* in Equation (10), since *CPCEET* is 100%, it seems the constant  $\lambda(q, N)$  produces the confidence intervals, which includes the true value of the area. However, the accurateness of the constant  $\lambda(q, N)$  with the empirical true variance and the approximate variance estimations should be examined together. For the same values of  $\lambda(q, N)$ , the bandwidths of the confidence intervals with *CEET* are smaller than that of *CEMC*, which shows us that the bandwidths of confidence intervals constructed with *CEET* are more accurate than those of *CEMC*. When we look at Table A12, *CPCEMC* and *CPCEET* have the same values, except that the number of samplings  $n$  is equal to five. The more number of samplings means the more information gained from the *MF*. This is one of the reasons why the proposed *MF* must be used. For *MF* in Equation (11), it is important to take a value for  $q$ , such that we can not only estimate the variance of systematic sampling more precisely, but also choose the right value for the constant  $\lambda(q, N)$ . Finally, proposing a class of measurement functions for research helps us to control the trustworthiness of the research in the applied science if the experts in the area can propose them correctly.

A final comment for the simulation section is that the covariogram model in Equation (12) and the variance extension term in Equation (20) can be used. However, the proposed measurement functions, such as Equations (8)–(11), should be systematically sampled, so that one can get a more precise decision on the application for biomedical imaging. As another solution, Equation (20) may be reformulated in the framework of the fractional Fourier transformation proposed by [36,37].

## 3.2. Real Data

Five different sheep brains, which were 12–18 months old, were removed from their skull via craniotomy in the laboratory for anatomy. These brains were immersed in formalin (5%) for 10 days. Brains were scanned with standard  $T2$ -weighted 0.5 Tesla magnetic resonance imaging (MRI) in the coronal plane with a 5-mm slice thickness. The true values of volume for each brain were obtained by using the Archimedean principle repeated six times. The arithmetic mean of six results for each brain was used as a true value of the volume of a brain. After MRI scanned the sheep brains, the area of each digital image for slices of brains was computed by pixel counting. While doing the pixel counting, the edges of each digital image were detected by experts in the anatomy area. After that, they were estimated by the slices in the coronal plane. The results are given in Tables 3 and 4. Tables 3 and 4 show the true values for the volume ( $Q$ ), the estimated volume ( $\hat{Q}$ ), the estimated smoothness constant ( $\hat{q}$ ), the estimated coefficient of error ( $\hat{c}_e$ ), the lower bound of the estimated volume ( $\hat{Q}_{lower}$ ), the upper

bound of the estimated volume ( $\hat{Q}_{upper}$ ) and the number of slices (Num. Slc.). The area values of each slice obtained from the coronal axis are depicted in the Figure 3.

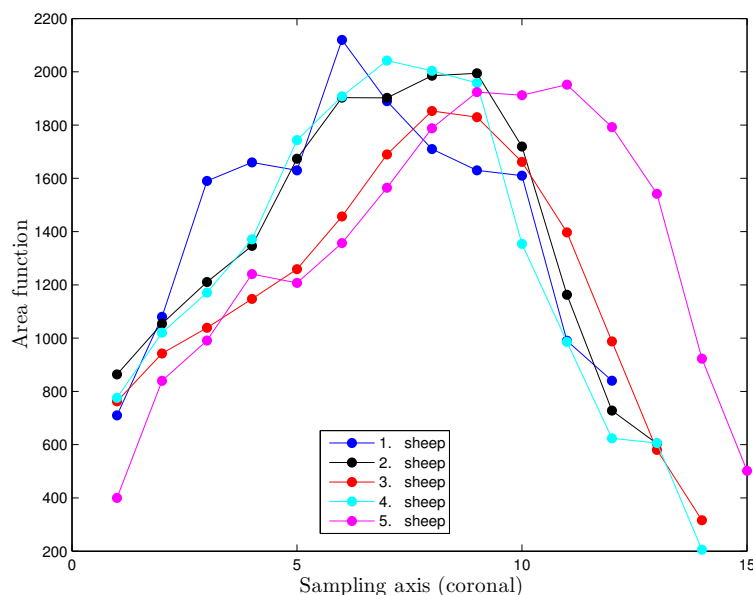
When  $N$  in Equation (32) is replaced with three,  $\lambda(0.3, 3) = 3.82689$  and  $\lambda(0.34, 3) = 3.90405$  are found. The confidence intervals of Brain 2 for  $q = 0.3$  and  $q = 0.34$  are (87,055.68, 94,420.32) and (86,981.44, 94,494.56), respectively. These confidence intervals include the true value of volume for Brain 2.

**Table 3.** Five sheep brain volumes from automatic pixel counting ( $\text{mm}^2$ ) after determining the border of slices with a 5-mm thickness and their confidence intervals for  $N = 2$ : arithmetic mean of  $\hat{q}$  values with  $k = 2, 3, 4, 5, 6, 7$  in Equation (25). Num. Slc., number of slices.

Brains	$Q$	$\hat{Q}$	$\hat{q}$	$\hat{c}e$	$\hat{Q}_{lower}$	$\hat{Q}_{upper}$	Num. Slc.
1	85,000	87,300	0.20	0.0234	81,289.05	93,310.95	12
2	94,000	90,738	0.34	0.0106	87,731.40	93,744.60	13
3	83,000	84,608	0.50	0.0066	82,722.64	86,493.36	14
4	88,000	88,846	0.42	0.0087	86,318.13	91,372.87	14
5	100,000	99,675	0.67	0.0039	98,351.66	100,998.34	15

**Table 4.** Five sheep brain volumes from automatic pixel counting ( $\text{mm}^2$ ) after determining the border of slices with a 5-mm thickness and their confidence intervals for  $N = 2$ :  $\hat{q}$  values with  $k = 2$  in Equation (25).

Brains	$Q$	$\hat{Q}$	$\hat{q}$	$\hat{c}e$	$\hat{Q}_{lower}$	$\hat{Q}_{upper}$	Num. Slc.
1	85,000	87,300	-0.06	0.0234	82,646.98	91,953.02	12
2	94,000	90,738	0.18	0.0131	87,249.31	94,226.69	13
3	83,000	84,608	0.41	0.0075	82,464.84	86,751.16	14
4	88,000	88,846	0.28	0.0105	85,934.10	91,756.90	14
5	100,000	99,675	0.63	0.0041	98,283.46	101,066.54	15



**Figure 3.** Area functions for each brain.

In Table 4, when the estimated  $\hat{q}$  with  $k = 2$  is taken, the confidence interval includes the true values of volume for each non-vivo brain. It is seen that estimating accurately the parameter  $q$  affects the variance estimation and the confidence interval, as well. For this reason,  $\lambda(q, N)$  values for

$N = 2, 3, 4$  in Table 1 are computed. The simulation results show that  $N$  should be two. The more precise measurement of the slice area is needed for the real data case, because measuring the area of slices more precisely produces the shape of  $MF$  more accurately. Having a more accurate shape of  $MF$  means that the estimated values of parameter  $q$  can be determined more accurately. However, it should be noted that the estimation of parameter  $q$  except Equation (8) with  $q = 0.4$  is an open problem, as well, even if the exact form of  $MF$  is known. Taking the values of  $k$  in  $q$  formula and  $N$  in the constant  $\lambda(q, N)$  as bigger than two is not useful due to the loss of information; however, we do not have a solution for the real data case. In the real data case, we have to try different values of  $k$  and  $N$ , and also, the proposed  $MF$  must be used to increase the trustworthiness of research in the applied science. As a result, using two propositions is required.

#### 4. Conclusions and Discussions

The works in Ref. [6–15,19–22] focused on the variance estimation for systematic sampling. As implied by [9,17,21], the estimation of  $q$  is important to avoid the biasedness of the variance estimator in systematic sampling on  $\mathbb{R}$ . The unbiasedness of the variance extension term estimator leads to having accurate lower and upper bounds of the confidence interval for the systematic sampling on  $\mathbb{R}$ . In fact, the true variance in Equation (5) is the combination of three components, so we always have biased variance estimation; however, for practical purposes, we are only interested in the variance extension term. It is observed from the simulation results that *CEMC* and *CEET* have values that contain the same number of zero digits in the fraction part for the measurement functions in Equations (8), (9) and (11) at most times under Equations (12) and (20). However, the same performance of the formula in Equation (20) is not observed for the variance approximation of the systematic sampling of  $MF$  in Equation (10) even if the number of sample sizes increases. Therefore, the formulae in Equation (12) and Equation (20) are not appropriate to use. It should be noted that due to the definition of systematic sampling given in Equation (1), a fluctuation can be observed, so evaluating the performance of *CEMC* and *CEET* according to the behavior of the sample size cannot be sensible. However, when the sample size increases, it is logical to expect the information coming from  $MF$  to be increased. In this context, Tables A1–A12 show that when the sample size increases, the values of *CEMC* and *CEET* are decreased. As implied by p. 319 in [8], if a measurement function has irregularities, such as Equations (9)–(11), the values of *CEMC* are high when the sample size is small, because the information is not enough to catch the approximate shape of the three-dimensional object. For the measurement functions in Equations (9)–(11), Tables A6, A8, A10 and A12 support this implication on p. 319 of [8]. The *CEET* values of  $MF$  in Equation (9) with  $q = 0.4$  can approve the fluctuation when the sample sizes are 5–20. For other measurement functions, the fluctuation on *CEET* is not observed; in fact, there is a fluctuation; however, the effect of sample size covers the fluctuation. Otherwise, proposing a Fourier transformation to model the estimated values of parameter  $Q$  would not give the results for *CEMC* that are approximate to *CEET*.

A three-dimensional object is represented by a flexible balloon filled with gas. The shape of this balloon can be modified by means of hands. The volume of the balloon can change after each modification; however, it is not important whether or not there is a change of the volume of the balloon. The modified versions of the balloon are represented by the measurement functions given in Equations (8)–(11). The measurement functions in Equations (10) and (11) do not have the parameter  $q$ ; however, they can be a neighborhood of  $MF$  having parameter  $q$ . Since we have different forms of the balloon, we will have many forms of the measurement functions representing the three-dimensional object exactly. In the real data case, knowing the exact form of the three-dimensional object is impossible. The covariogram functions coming from the  $MF$  without  $q$  will be the neighborhood of the covariogram functions coming from  $MF$  with  $q$ . This is why the covariogram model has the parameter  $q$ . The true  $MF$  with  $q$  cannot be known exactly due to fact that we can have  $MF$  without  $q$ , which is a neighborhood of  $MF$  with  $q$ . In other words, we can have approximately the same functions, which are different mathematical expressions.

A method showing how to calculate constant  $\lambda(q, N)$  is proposed. It is expected that this method can be used as a new tool in mathematics when it is needed. A package program written in MATLAB 2013a gives the lower and upper levels of confidence and the quantitative values of stereology when the data obtained from a single replication is typed. The program can be supplied on request. The estimation of  $q$  is an open problem, even if we know the exact form of the measurement functions, except Equation (8) with  $q = 0.4$ . In the other perspective of our discussion, the covariogram model cannot be a very good approximation for the covariogram functions of measurement functions. However, the values of *CEMC* and *CEET* contain the same numbers of zero digits in the fraction part at most times for the measurement functions in Equations (8), (9) and (11) under Equations (12) and (20).

The values of *CPCEMC* and *CPCEET* are the other criteria to approve the performance of the exact values of the constant  $\lambda(q, N)$ , as observed from the simulation results. A numerical computation for the constant  $\lambda(q, N)$  of the confidence interval was done by [2]. The more precise the values of the constant  $\lambda(q, N)$  means the more precise the confidence interval. It is obvious that the proposition of exact calculation has to be preferred, because the computation of the constant  $\lambda(q, N)$  proposed by [2] is not as good as the results displayed by Tables 1 and 2, as seen by the conformity between *CPCEMC* and *CPCEET* for exact values of  $\lambda(q, N)$ , especially for the measurement functions in Equations (8), (9) and (11).

It is observed that the synthetic data can approve the real data for non-vivo brains if the real data have an exactly similar form as the synthetic data. Since the real data cannot be represented by the exactly similar measurement functions, we should prefer to use different  $N$  values. For this reason, we chose the different  $k$  values for the estimation of  $q$  in real data. Equation (25) can produce the negative estimated  $\hat{q}$  value for the area function of the real data given in Figure 3. The constant  $\lambda(q, N)$  values for different  $N$  values must be given, because the real data shows that we need constant  $\lambda(q, N)$  values with different  $N$ . In this context, the edges of digital images have to be found more precisely. Especially, when the edges of images are more irregular, the precision of getting area values for each image has to be decreased significantly. To do this, new edge detection methods proposed by [38] can be used.

The proposed measurement functions should be systematically sampled while conducting research on biomedical imaging to increase the information in the decision rule for now. To be able to increase the performance of the variance approximation, the fractional Fourier transformation may be applied to get a new variance estimation of the equidistant systematically sampled on  $\mathbb{R}$ . Other types of *MF* may also be used. In this case, we need the computational integral techniques. The studies done by [39,40] will be enlightening for the computational integration of other *MF* types that we will use. For the construction of *MF* via digital images, which have coronal, axial or sagittal directions, the new edge detection methods proposed by [38] may be applied to get more precise area values of slices. We will prepare a free statistical software *R* package with a macro for Mathematica for all methodologies given here and in the future.

**Acknowledgments:** I am very grateful to Luis M. Cruz-Orive who sincerely invited me, to Marta García-Fiñana for sending the codes in *R* and to Niyazi Acer for kindly providing the data. I am also so indebted to the Council of Higher Education for giving the funding for my research. I am indebted to James F. Peters from the University of Manitoba, Winnipeg, Canada, for critical reading. I would like to thank sincerely the four referees. Without their help, this paper would not have been improved. I also wish to thank my great parents. In the memory of my great brother.

**Author Contributions:** Mehmet Niyazi Çankaya finished the manuscript. The author has read and approved the final manuscript.

**Conflicts of Interest:** The author declares no conflict of interest.

## Appendix

Table A1.  $(1 - x^2)^q$ :  $q = 0.4$ .

$n$	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
5	0.469211	0.000755	0.005545
6	0.449412	0.001009	0.003450
7	0.438921	0.000994	0.002509
8	0.430876	0.000992	0.001945
9	0.426702	0.001029	0.001742
10	0.422139	0.001089	0.001580
11	0.419202	0.001029	0.001398
12	0.417282	0.001049	0.001348
13	0.414414	0.001213	0.001421
14	0.413815	0.001090	0.001281
15	0.411180	0.001210	0.001335
16	0.411575	0.001060	0.001194
17	0.409010	0.001232	0.001314
18	0.409631	0.001119	0.001211
19	0.409595	0.001086	0.001178
20	0.408593	0.001079	0.001153

Table A2.  $(1 - x^2)^q$ :  $q = 0.4$ .

$n$	CEMC	CEET	CPCEMC%	CPCEET%
5	0.031387	0.027150	100.0	98.8
6	0.024935	0.021804	100.0	99.1
7	0.020275	0.017246	100.0	99.1
8	0.016988	0.014385	100.0	99.3
9	0.014433	0.011948	100.0	98.7
10	0.012535	0.010502	100.0	99.1
11	0.011000	0.009178	100.0	99.5
12	0.009751	0.008085	100.0	99.0
13	0.008768	0.007458	100.0	99.0
14	0.007891	0.006564	100.0	99.1
15	0.007211	0.006121	100.0	99.3
16	0.006562	0.005408	100.0	99.3
17	0.006069	0.005181	100.0	99.2
18	0.005580	0.004672	100.0	99.5
19	0.005167	0.004271	100.0	99.4
20	0.004820	0.003961	100.0	99.3

Table A3.  $(1 - x^2)^q$ :  $q = 0.8$ .

$n$	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
5	0.798607	0.014187	0.014189
6	0.821274	0.015347	0.015800
7	0.827312	0.016704	0.017450
8	0.838035	0.016471	0.017918
9	0.835330	0.017350	0.018599
10	0.840301	0.017530	0.019154
11	0.840649	0.017641	0.019294
12	0.839555	0.018128	0.019693
13	0.842700	0.017997	0.019821
14	0.837016	0.018080	0.019450
15	0.838411	0.018319	0.019794
16	0.840756	0.018341	0.020002
17	0.837451	0.018479	0.019881
18	0.831963	0.019301	0.020323
19	0.839645	0.019642	0.021214
20	0.835757	0.019321	0.020599



**Table A4.**  $(1 - x^2)^q$ :  $q = 0.8$ .

<i>n</i>	<i>CEMC</i>	<i>CEET</i>	<i>CPCEMC%</i>	<i>CPCEET%</i>
5	0.018256	0.022835	100.0	100.0
6	0.012532	0.016332	100.0	100.0
7	0.009405	0.012425	100.0	100.0
8	0.007190	0.009670	100.0	100.0
9	0.005913	0.007842	100.0	100.0
10	0.004853	0.006570	100.0	100.0
11	0.004053	0.005437	100.0	100.0
12	0.003491	0.004684	100.0	100.0
13	0.003021	0.004037	100.0	100.0
14	0.002674	0.003554	100.0	100.0
15	0.002380	0.003191	100.0	100.0
16	0.002116	0.002814	100.0	100.0
17	0.001885	0.002528	100.0	100.0
18	0.001702	0.002280	100.0	100.0
19	0.001549	0.002043	100.0	100.0
20	0.001429	0.001903	100.0	100.0

**Table A5.**  $((1 - \cos(x))(1 - x^2))^q$ :  $q = 0.4$ .

<i>n</i>	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
5	-0.059737	0.009231	0.220590
6	0.257864	0.001882	0.022085
7	0.412975	0.000590	0.000758
8	0.483350	0.000779	0.007726
9	0.501458	0.000922	0.011216
10	0.507921	0.001314	0.012961
11	0.507945	0.000837	0.012489
12	0.507897	0.001435	0.013076
13	0.503777	0.000699	0.011469
14	0.501835	0.001445	0.011816
15	0.497033	0.000737	0.010153
16	0.494786	0.001527	0.010511
17	0.489366	0.000710	0.008696
18	0.488828	0.001432	0.009323
19	0.483590	0.000714	0.007702
20	0.481851	0.001531	0.008230

**Table A6.**  $((1 - \cos(x))(1 - x^2))^q$ :  $q = 0.4$ .

<i>n</i>	<i>CEMC</i>	<i>CEET</i>	<i>CPCEMC%</i>	<i>CPCEET%</i>
5	0.136986	0.035193	100.0	94.6
6	0.075842	0.056090	100.0	97.4
7	0.048192	0.022847	100.0	98.0
8	0.035892	0.036584	100.0	99.1
9	0.029172	0.017402	100.0	98.3
10	0.024801	0.026910	100.0	99.5
11	0.021386	0.013797	100.0	98.6
12	0.018852	0.020350	100.0	99.4
13	0.016763	0.010922	100.0	98.9
14	0.015121	0.016248	100.0	99.7
15	0.013697	0.009184	100.0	98.5
16	0.012545	0.013334	100.0	99.4
17	0.011526	0.007842	100.0	98.9
18	0.010623	0.011027	100.0	99.7
19	0.009865	0.006687	100.0	98.5
20	0.009211	0.009652	100.0	99.6

**Table A7.**  $((1 - \cos(x))(1 - x^2))^q$ :  $q = 0.8$ .

$n$	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
5	-0.127902	0.025653	0.886655
6	0.259307	0.009587	0.301935
7	0.471353	0.008843	0.116851
8	0.591335	0.008365	0.051906
9	0.667866	0.009158	0.026618
10	0.712836	0.010045	0.017643
11	0.743469	0.010628	0.013824
12	0.769775	0.011272	0.012186
13	0.782195	0.012278	0.012595
14	0.794971	0.013111	0.013137
15	0.802343	0.013406	0.013412
16	0.810423	0.013491	0.013599
17	0.814130	0.013994	0.014194
18	0.822037	0.014498	0.014984
19	0.822426	0.014351	0.014854
20	0.822911	0.015298	0.015823

**Table A8.**  $((1 - \cos(x))(1 - x^2))^q$ :  $q = 0.8$ .

$n$	CEMC	CEET	CPCEMC%	CPCEET%
5	0.179574	0.070757	100.0	95.8
6	0.098853	0.058929	100.0	98.7
7	0.056666	0.040635	100.0	100.0
8	0.037906	0.035454	100.0	100.0
9	0.027210	0.025762	100.0	100.0
10	0.021142	0.023491	100.0	100.0
11	0.016973	0.018652	100.0	100.0
12	0.014014	0.016925	100.0	100.0
13	0.011669	0.013700	100.0	100.0
14	0.010083	0.012682	100.0	100.0
15	0.008732	0.010624	100.0	100.0
16	0.007799	0.010070	100.0	100.0
17	0.006862	0.008591	100.0	100.0
18	0.006093	0.008067	100.0	100.0
19	0.005530	0.007070	100.0	100.0
20	0.005030	0.006753	100.0	100.0

**Table A9.**  $\exp(-\sin(-x^3))$ :  $q = 0.95$ .

$n$	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
5	-0.088096	0.050826	1.128469
6	0.041988	0.031897	0.856384
7	0.189917	0.018652	0.596379
8	0.320166	0.008843	0.405534
9	0.449291	0.007235	0.257945
10	0.549659	0.004972	0.165245
11	0.639576	0.002916	0.099279
12	0.714919	0.001908	0.057171
13	0.776112	0.001817	0.032055
14	0.827798	0.001763	0.016697
15	0.868178	0.001724	0.008419
16	0.899500	0.001604	0.004154
17	0.928089	0.001326	0.001806
18	0.942607	0.001128	0.001183
19	0.959425	0.001025	0.001114
20	0.967014	0.000817	0.001106

**Table A10.**  $\exp(-\sin(-x^3))$ :  $q = 0.95$ .

<i>n</i>	<i>CEMC</i>	<i>CEET</i>	<i>CPCEMC%</i>	<i>CPCEET%</i>
5	0.123899	0.042779	100.0	100.0
6	0.096852	0.020019	100.0	100.0
7	0.068757	0.008892	100.0	100.0
8	0.046235	0.003329	100.0	100.0
9	0.031395	0.001138	100.0	100.0
10	0.022537	0.000614	100.0	100.0
11	0.016419	0.000538	100.0	100.0
12	0.012289	0.000462	100.0	100.0
13	0.009477	0.000366	100.0	100.0
14	0.007494	0.000282	100.0	100.0
15	0.006081	0.000220	100.0	100.0
16	0.005047	0.000184	100.0	100.0
17	0.004236	0.000158	100.0	100.0
18	0.003677	0.000132	100.0	100.0
19	0.003202	0.000117	100.0	100.0
20	0.002850	0.000103	100.0	100.0

**Table A11.**  $(5/112)(-54x^4 - 25x^3 + 48x^2 + 25x + 6)$  :  $q = 0.9$ .

<i>n</i>	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
5	0.119701	0.021089	0.629956
6	0.349198	0.018577	0.321959
7	0.506408	0.017147	0.172062
8	0.607155	0.017009	0.102768
9	0.677396	0.017249	0.066802
10	0.720912	0.016863	0.048935
11	0.765776	0.016458	0.034474
12	0.786878	0.015944	0.028741
13	0.815932	0.015711	0.022778
14	0.838310	0.015351	0.019157
15	0.856894	0.015623	0.017481
16	0.865002	0.016006	0.017231
17	0.883260	0.015086	0.015367
18	0.887106	0.015674	0.015840
19	0.896326	0.015430	0.015444
20	0.904643	0.014824	0.014845

**Table A12.**  $(5/112)(-54x^4 - 25x^3 + 48x^2 + 25x + 6)$  :  $q = 0.9$ .

<i>n</i>	<i>CEMC</i>	<i>CEET</i>	<i>CPCEMC%</i>	<i>CPCEET%</i>
5	0.131979	0.061047	100.0	97.7
6	0.076664	0.044234	100.0	100.0
7	0.047065	0.031891	100.0	100.0
8	0.032165	0.024913	100.0	100.0
9	0.023411	0.019979	100.0	100.0
10	0.018146	0.015962	100.0	100.0
11	0.014053	0.013222	100.0	100.0
12	0.011617	0.011093	100.0	100.0
13	0.009464	0.009315	100.0	100.0
14	0.007890	0.007985	100.0	100.0
15	0.006690	0.006992	100.0	100.0
16	0.005872	0.006274	100.0	100.0
17	0.005015	0.005386	100.0	100.0
18	0.004499	0.004952	100.0	100.0
19	0.003983	0.004410	100.0	100.0
20	0.003544	0.003891	100.0	100.0

## References

1. Eriksen, N.; Rostrup, E.; Andersen, K.; Lauritzen, M.J.; Larsen, V.A.; Dreier, J.P.; Strong, A.J.; Hartings, J.A.; Fabricius, M.; Pakkenberg, B. Application of stereological estimates in patients with severe head injuries using CT and MR scanning images. *Br. Inst. Radiol.* **2010**, *83*, 307–317.
2. García-Fiñana, M. Confidence intervals in Cavalieri Sampling. *J. Microsc.* **2006**, *222*, 146–157.
3. García-Fiñana, M.; Keller, S.S.; Roberts, N. Confidence Intervals for the Volume of Brain Structures in Cavalieri Sampling with Local Errors. *J. Neurosci. Methods* **2009**, *179*, 71–77.
4. Hussain, Z.; Roberts, N.; Whitehouse, G.H.; García-Fiñana, M.; Percy, D. Estimation of Breast Volume and its Variation During the Menstrual Cycle using MRI and Stereology. *Br. J. Radiol.* **1999**, *72*, 236–245.
5. McNulty, V.; Cruz-Orive, L.M.; Roberts, N.; Holmes, C.J.; Gual-Arnau, X. Estimation of Brain Compartment Volume from MR Cavalieri Slices. *J. Comput. Assist. Tomogr.* **2000**, *24*, 466–477.
6. Matheron, G. *Variables Régionalisées et Leur Estimation [Les]*; Masson: Paris, France, 1965.
7. Matheron, G. The Theory of Regionalized Variables and Its Applications. In *Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau, No. 5*; Ecole Nationale Supérieure des Mines de Paris: Paris, France, 1971.
8. Cruz-Orive, L.M. On the Precision of Systematic Sampling: A Review of Matheron's Transitive Methods. *J. Microsc.* **1989**, *153*, 315–333.
9. García-Fiñana, M.; Cruz-Orive, L.M. Improved Variance Prediction for Systematic Sampling on R. *Statistics* **2004**, *38*, 243–272.
10. Gual Arnau, X.; Cruz-Orive, L.M. Variance prediction under systematic sampling with geometric probes. *Adv. Appl. Probab.* **1998**, *30*, 889–903.
11. Gundersen, H.J.G.; Jensen, E.B. The Efficiency of Systematic Sampling in Stereology and its Prediction. *J. Microsc.* **1987**, *147*, 229–263.
12. Kellerer, A.M. Exact formulae for the precision of systematic sampling. *J. Microsc.* **1989**, *153*, 285–300.
13. Kiêu, K. *Three Lectures on Systematic Geometric Sampling*; Memoirs 13; Department of Theoretical Statistics, University of Aarhus: Aarhus, Denmark, 1997; p. 100.
14. Kiêu, K.; Souchet, S.; Istas, J. Precision of Systematic Sampling and Transitive Methods. *J. Stat. Plan. Inference* **1999**, *77*, 263–279.
15. Mattfeldt, T. The accuracy of one-dimensional systematic sampling. *J. Microsc.* **1989**, *153*, 301–313.
16. Roberts, N.; Cruz-Orive, L.M.; Reid, M.K.; Brodie, D.A.; Bourne, M.; Edwards, H.T. Unbiased Estimation of Human Body Composition by the Cavalieri Method Using Magnetic Resonance Imaging. *J. Microsc.* **1993**, *171*, 239–253.
17. García-Fiñana, M.; Cruz-Orive, L.M.; Mackay Clare, E.; Pakkenberg, B.; Roberts, N. Comparison of MR Imaging Against Physical Sectioning to Estimate the Volume of Human Cerebral Compartments. *NeuroImage* **2003**, *18*, 505–516.
18. Baddeley, A.J.; Vedel Jensen, E.B. *Stereology for Statisticians*; Monographs on Statistics and Applied Probability; Chapman & Hall/CRC: London, UK, 2005; p. 380.
19. Cruz-Orive, L.M. Variance predictors for isotropic geometric sampling, with applications in forestry. *Stat. Methods Appl.* **2012**, *22*, 3–31.
20. Cruz-Orive, L.M. Systematic Sampling in Stereology. In Proceedings of the 49th Session of the International Statistical Institute, Florence, Italy, 25 August–3 September 1993; Volume 55, pp. 451–468.
21. Cruz-Orive, L.M. A General Variance Predictor for Cavalieri Slices. *J. Microsc.* **2006**, *222*, 158–165.
22. Gundersen, H.J.G.; Jensen, E.B.V.; Kiêu, K.; Nielsen, J. The Efficiency of Systematic Sampling in Stereology—Reconsidered. *J. Microsc.* **1999**, *193*, 199–211.
23. Acer, N.; Şahin, B.; Usanmaz, M.; Tatoğlu, H.; Irmak, Z. Comparison of Point Counting and Planimetry Methods for the Assessment of Cerebellar Volume in Human Using Magnetic Resonance Imaging: A Stereological Study. *Surg. Radiol. Anat.* **2008**, *30*, 335–339.
24. Acer, N.; Şahin, B.; Emirzeoğlu, M.; Uzun, A.; İncesu, L.; Bek, Y.; Bilgiç, S.; Kaplan, S. Stereological Estimation of the Orbital Volume: A Criterion Standard Study. *J. Craniofacial Surg.* **2009**, *20*, 921–938.
25. Acer, N.; Çankaya, M.N.; İşçi, Ö.; Baş, O.; Çamurdanoğlu, M.; Turgut, M. Estimation of Cerebral Surface Area using Vertical Sectioning and Magnetic Resonance Imaging: A Stereological Study. *Brain Res.* **2010**, *1310*, 29–36.

26. Akbaş, H.; Şahin, B.; Eroğlu, L.; Odacı, E.; Bilgiç, S.; Kaplan, S.; Uzun, A.; Ergür, H.; Bek, Y. Estimation of Breast Prosthesis Volume by the Cavalieri Principle Using Magnetic Resonance Images. *Aesthet. Plast. Surg.* **2004**, *28*, 275–280.
27. Howard, M.A.; Roberts, N.; García-Fiñana, M.; Cowell, P.E. Volume Estimation of Prefrontal Cortical Subfields using MRI and Stereology. *Brain Res. Protoc.* **2003**, *10*, 125–138.
28. Maudsley, R.; García-Fiñana, M. Sampling Intensity with Fixed Precision When Estimating Volume of Human Brain Compartments. *Image Anal. Stereol.* **2008**, *27*, 143–149.
29. Roberts, N.; Pubdephat, M.J.; McNulty, V. The Benefit of Stereology for Quantitative Radiology. *Br. J. Radiol.* **2000**, *73*, 679–697.
30. Şahin, B.; Emirzeoğlu, M.; Uzun, A.; İncesu, L.; Bek, Y.; Bilgiç, S.; Kaplan, S. Unbiased Estimation of the Liver Volume by the Cavalieri Principle Using Magnetic Resonance Images. *Eur. J. Radiol.* **2003**, *47*, 164–170.
31. Şahin, B.; Ergür, H. Assessment of the Optimum Section Thickness for the Estimation of Liver Volume Using Magnetic Resonance Images: A Stereological Gold Standard Study. *Eur. J. Radiol.* **2005**, *57*, 96–101.
32. Bellhouse, D.R.; Krishnaiah, P.R.; Rao, C.R. (Eds.) *Sampling: Handbook of Statistics*; Elsevier Science: Amsterdam, The Netherlands, 1988; Volume 6, pp. 125–145.
33. Hall, P.; Ziegel, J. Distributions estimators and confidence intervals for stereological volumes. *Biometrika* **2011**, *98*, 417–431.
34. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions*; U.S. Government Printing Office: Washington, DC, USA, 1970; p. 1058.
35. Çankaya, M.N. Stereological Estimation and Inference with Applications. Master's Thesis, Council of Higher Education, Ankara, Turkey, 2010; p. 168.
36. Baleanu, D.; Wu, G.C.; Duan, J.S. Some Analytical Techniques in Fractional Calculus: Realities and Challenges. In *Discontinuity and Complexity in Nonlinear Physical Systems*; The Series Nonlinear Systems and Complexity; Springer: Heidelberg, Germany, 2013; Volume 6, pp. 35–62.
37. Poularikas, A.D. *Transforms and Applications Handbook*, 3rd ed.; CRC Press, Taylor and Francis Group: Boca Raton, FL, USA, 2010.
38. Peters, J.F. *Computational Proximity Excursions in the Topology of Digital Images*; Intelligent Systems Reference Library; Springer: Basel, Switzerland, 2016; Volume 102, p. 433.
39. Qin, Y.M.; Shi, Y.G. A new approximate method to conjugacies between a family of unimodal interval maps. *J. Comput. Complex. Appl.* **2016**, *2*, 163–169.
40. Zengn, Y. Approximate solutions of three integral equations by the new Adomian decomposition method. *J. Comput. Complex. Appl.* **2016**, *2*, 38–43.



© 2016 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).