 doi:10.20944/preprints201608.0017.v4

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.:+ 09 276 221 21 21 - 2585

Abstract: The systematic sampling is used as a method to get the quantitative results from the tissues and the radiological images. Systematic sampling on real line ($\mathbb{R}$) is a very attractive method within which the biomedical imaging is consulted by the practitioners. For the systematic sampling on $\mathbb{R}$, the measurement function (*MF*) is occurred by slicing the three-dimensional object equidistant systematically. The currently used covariogram model in variance approximation proposed by [28,29] is tested for the different measurement functions in a class to see the performance on the variance estimation of systematically sampled $\mathbb{R}$. This study is an extension of [17], and an exact calculation method is proposed to calculate the constant $\lambda(q, N)$ of 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

The 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, length. In the systematic sampling principle, geometrical objects are sampled with probes, such as lines, 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 systematically sampling on geometrical objects, we will have the estimated values for each replication of sampling of these objects. The estimated values obtained from each replication of this sampling on geometrical objects produce a fluctuation [14,17,18,24,32]. So, this fluctuation is modelled by Fourier transformation $\mathcal{F}_1$, which was considered by [28,29]. Since we have the estimated values for parameter Q, the variance estimation is needed. For the systematic sampling on $\mathbb{R}$, also called as Cavalieri sampling, previous studies shed some light on the variance estimation for the systematic sampling on real line, [9,16,19,20,25–27,30]. Main source in these studies is of inspiration in 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 performance of his covariogram model, because the sampling directions (axises) which are axial, coronal or sagittal produce the different measurement functions, as implied by [17]. There is no study addressing the performance evaluation of the covariogram model for the different measurement functions. In this context, the papers [2–5,14,15,18,20,21,23,24,31,36–39] used the 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 $\hat{Q}$. [17] proposes two approaches for the coefficient of confidence interval. The first one is based on the statistical theory. In this study, the second proposition of [17] is examined. Main motivation is to find a calculation method for the constant $\lambda(q, N)$ which is used to construct 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 at Section 3. A section 4 is considered for the discussions on 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 variance approximation formula are introduced briefly.

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

2.1.1. Estimation for Volume Q

Suppose that we have three-dimensional geometrical object. This object is a fixed, bounded, nonvoid with 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 as the 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,

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

T is a constant distance among the slices obtained from three-dimensional object. $\mathbf{u}$ is an 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 sampling, 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 [6,9,12,16].

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 [9] for these properties. The variance of Cavalieri sampling changes with a fractional power of T [16,17]. The fact that there is a fractional power for T is not pointed out by [26,27]. In this sense, [16] is an extension of [26]. For this reason, we want to test the performance of covariogram model in Eq. (11) and the variance extension term in Eq. (19) for the different measurement functions.

For the practical purpose, 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 [9] and the properties of measurement functions proposed by [16].

Firstly, we will introduce the empirical true variance for $\hat{Q}$. Secondly, the true variance for $\hat{Q}$ is defined in Eq. (4). It includes the covariogram function. After the covariogram function is defined, the measurement functions will be introduced.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

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 [6,8,36].

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

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

which is in the context of coefficient of error square.

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

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

In the simulation section, Eq. (3) is used to test the performance of variance approximation formula in Eq. (19). The true variance is calculated by a formula in Eq. (4). A calculation of the true variance by means of Eq. (4) is intractable (see [9] for more details). The true variance comprises of three components which are the variance extension term, Zitterbewegung and the higher-order terms, respectively. These terms were gotten in [16] (see section 6 in [16] for detailed expressions).

$$Var(\hat{Q}) = T\sum_{k=-\infty}^{\infty} g(kT) - \int_{-\infty}^{\infty} g(h)dh, \tag{4}$$

where

$$g(h) = \int_{\mathbb{R}} f(x)\cdot f(x+h)dx, \;\; h \in \mathbb{R} \tag{5}$$

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

$$G(t) = \mathcal{F}_1 g(h) = \int_{-\infty}^{\infty} g(h)exp(-2\pi ith)dh. \tag{6}$$

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

Suppose that we have measurement functions in Eqs. (7)-(10). These functions are used for the simulation. They can represent the biological objects. They are given with following forms and the function f has positive values for each values of variable x, namely $f : \mathbb{R} \to \mathbb{R}^+$.

$$f(x) = (1-x^2)^q, \;\; x \in [-1,1], \;\; q \in [0,1], \tag{7}$$

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

where q is a parameter of smoothness constant.

$$f(x) = exp(-sin(-x^3)), \;\; x \in [-38\pi/100, 53\pi/100], \tag{9}$$

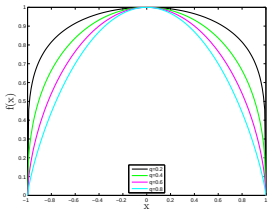
and

$$f(x) = (5/112)(-54x^4 - 25x^3 + 48x^2 + 25x + 6), \;\; x \in [-1,1]. \tag{10}$$

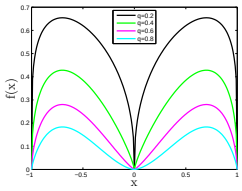
They represent the area measured on the each slices of three-dimensional biological objects. The measurement functions in Eqs. (7) and (10) are used by [16,22], respectively. Each function f and its

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

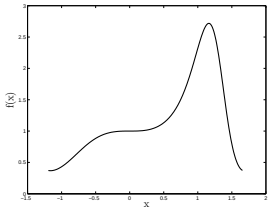
covariogram function can be an example for the representation of q even though the function f does not have the parameter q. Two approximately same functions can be proposed, $f_i \approx f_j$, where $i \neq j$, $i, j = \mathbb{N}$, which have the different mathematical expressions, however they can be a neighborhood to each other. For example, the measurement functions in Eqs. (9) - (10) can be more modified version of MF in Eq. (8). MF in Eq. (8) can be a modified version of MF in Eq. (7).



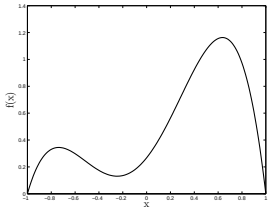
(a) MF of Eq. (7)



(b) MF of Eq. (8)



(a) MF of Eq. (9)



(b) MF of Eq. (10)

Since the covariogram functions of measurement functions in Eqs. (7)-(10) can not be calculable to get the integral values, a model for the covariogram functions should be proposed. In this sense, the covariogram functions can be modelled by a polynomial with the fractional power,

$$g(h) = b_0 + b_j|h|^j + b_2h^2,\ j = 2q + 1,\ q \in [0, 1]. \tag{11}$$

Eq. (11) is defined to be a covariogram model [6,9,16,28,29].

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

By using Eq. (4) and the properties of $G(t)$, we finally get

$$Var(\hat{Q}) \quad = \quad T^{-1}\int_0^T \hat{Q}^2 du - Q^2 = \sum_{k=-\infty}^{\infty} G(k/T) - G(0) = 2\sum_{k=1}^{\infty} G(k/T) \tag{12}$$

(Detailed expressions are given in [9,16,28,29]).

The Fourier transformations of parts h^0 and h^2 in the covariogram model in Eq. (11) are zero. Thus, the variance extension term of systematic sampling on $\mathbb{R}$ is obtained by using the formula given in Eq. (9.1) and Eq. (9.2) in [9] as follow,

$$\mathcal{F}_1|h|^j \quad = \quad b(j,1)\varrho^{-(j+1)},\ \ (j > -1,\ \text{non-even}) \tag{13}$$

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

By using Eq. (11), Eq. (12), and Eq. (13), Eq. (14) is obtained.

$$\begin{aligned} Var_E(\hat{Q}) &= 2b_j b(j,1)T^{j+1}\zeta(j+1) \\ Var_E(\hat{Q}) &= \alpha(q)(3g(0) - 4g(T) + g(2T)) \\ Var_E(\hat{Q}) &= \alpha(q)(3C_0 - 4C_1 + C_2)T^2 \end{aligned} \tag{14}$$

where

$$\alpha(q) = \frac{2\pi^{-(2q+3/2)}\Gamma(q+1)\zeta(2q+2)}{(2^{2q+1}-4)\Gamma(-\frac{2q+1}{2})}, \quad q \in \mathbb{R}, \tag{15}$$

$$\hat{g}(kT) = TC_k, \tag{16}$$

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

Eq. (15) is gotten according to Fourier transformation basically given in [9]. However, Eq. (17) is obtained according to the generalized version of refined Euler-MacLaurin, proposed by [16].

$$\alpha(q) = \frac{\Gamma(2q+2)\zeta(2q+2)cos(q\pi)}{(2\pi)^{2q+2}(1-2^{2q-1})}, \quad q \in [0,1], \tag{17}$$

$$Var_E(\hat{Q}) = \frac{2\pi^{-(2q+3/2)}\Gamma(q+1)\zeta(2q+2)}{(2^{2q+1}-4)\Gamma(-\frac{2q+1}{2})}(3C_0 - 4C_1 + C_2)T^2. \tag{18}$$

Note that Eq. (15) and Eq. (17) give the same results, because the connection between the generalized version of refined Euler-MacLaurin summation formula and the Matheron's transitive theory is expressed by [16].

We calculate the estimated coefficient of error for the systematic sampling on $\mathbb{R}$ by means of Eq. (19) given below,

$$\begin{aligned} \widehat{var}_E(\hat{Q})/Q^2 &= \alpha(q)(3C_0 - 4C_1 + C_2)T^2/Q^2 \\ \hat{ce}(\hat{Q}) &= \sqrt{\alpha(q)(3C_0 - 4C_1 + C_2)}(\sum_{i=1}^{n} f_i)^{-1} \end{aligned} \tag{19}$$

where

$$C_k = \sum_{i=1}^{n-k} f_i \cdot f_{i+k}, \; k = 0, 1, ..., n-1. \tag{20}$$

This is defined to be a coefficient of error of Matheron's covariogram model. Eq. (5) and Eq. (20) have in fact inheritance. $n \geq 2k+1$ observations are required due to Eq. (20) [9–11,15,16,28,29].

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. [2–5,14,15,18,20,21,23,24,31,36–39] used currently the covariogram model for the different measurement functions. We aim to show the performance of the covariogram model in Eq. (11) for the different measurement functions. In other words, the capability of 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 Eq. (14), the estimation formula of parameter q is given into following subsection.

2.1.4. Estimation Formula of Parameter q

The estimation formula of smoothness constant q is proposed by [16]. In this subsection, we will get it via the variance extension term in Eq. (14) in the framework of Fourier transformation. The formula we got for q is the same with the formula in [16].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

The covariogram model g can be declared by integer k values. If $h = iT$ is near zero, Eq. (6) has more information. By using Eqs. (11) and (14), respectively,

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

$$\begin{aligned} Var_E(\hat{Q}) &= 2b_{2q+1}b(2q+1,1)T^{2q+2}\zeta(2q+2) \\ Var_E(\hat{Q}) &= 2\pi^{-(2q+3/2)}\frac{\Gamma(q+1)}{\Gamma(-\frac{2q+1}{2})}\zeta(2q+2) \\ &\quad \frac{T\cdot 3g(0) - 4g(kT) + g(2kT)}{k^{2q+1}(2^{2q+1}-4)} \end{aligned} \tag{21}$$

can be obtained. By using Eq. (18) and Eq. (21),

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

is obtained and then

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

where, $k = 2, 3, ...$
can be obtained. By using Eq. (16),

$$\hat{q} = \frac{1}{2log(k)}\cdot log\left(\frac{3C_0 - 4C_k + C_{2k}}{3C_0 - 4C_1 + C_2}\right) - \frac{1}{2},\ k = 2, 3, ... \tag{22}$$

The estimator $\hat{q}$ is gotten when the MF is obtained by planimetry. The planimetry is also called as an 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 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 [16] and [17] 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

$$|\hat{Q} - Q| \leq T^{q+1}P^*_{q+1}\sum_{i=1}^{N}|Sf^{(q)}(a_i)| \tag{23}$$

where

$$P^*_{q+1} = max_{\{\Delta,\beta\}}\{|\ \frac{-2}{(2\pi)^{q+1}}\sum_{j=1}^{\infty}\frac{1}{j^{q+1}}cos(2\pi\Delta j - \frac{\pi}{2}(q+1) + \beta)\ |\}, \tag{24}$$

$\Delta \in [0, 1)$ and $\beta \in [0, 2\pi]$. $\{a_i, i = 1, 2, ..., N\}$ is the set of real line at which function $f^{(q)}$ exhibits discontinuities. N is the number of set. The difference $\hat{Q} - Q$ and $(Sf^{(q)}(a_i))^2$ are given by Eq. (6) and Eq. (11) in [17]. If N is greater than 2, the information in $\hat{Q} - Q$ decreases (For detailed

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

expressions, see [17]). The preliminaries were given. Now, we will use the tools in [16,17]. After some straightforward calculations in the following steps, the formula of $\lambda(q,N)$ will be obtained.

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

$$\left(\sum_{i=1}^{N}|Sf^{(q)}(a_i)|\right)^2 \leq N\sum_{i=1}^{N}\left(Sf^{(q)}(a_i)\right)^2. \tag{25}$$

When Eq. (23) and Eq. (25) are used,

$$\begin{aligned}\frac{|\hat{Q}-Q|}{T^{q+1}P^*_{q+1}} &\leq \sum_{i=1}^{N}|Sf^{(q)}(a_i)| \\ |\hat{Q}-Q| &\leq T^{q+1}P^*_{q+1}\sqrt{N}\sqrt{\sum_{i=1}^{N}\left(Sf^{(q)}(a_i)\right)^2}\end{aligned} \tag{26}$$

is obtained.

The following formula having the parameter q is proposed by [16],

$$Var_E(\hat{Q}) = T^{2q+2}\frac{P_{2q+2,T}(0)}{cos(\pi q)}\sum_{i=1}^{N}\left(Sf^{(q)}(a_i)\right)^2. \tag{27}$$

When $\sum_{i=1}^{N}\left(Sf^{(q)}(a_i)\right)^2$ in Eq. (27) is written in Eq. (26), Eq. (28) is obtained.

$$\begin{aligned}|\hat{Q}-Q| &\leq P^*_{q+1}\sqrt{N}\sqrt{\frac{cos(\pi q)}{P_{2q+2,T}(0)}}\sqrt{Var_E(\hat{Q})}, \\ |\hat{Q}-Q| &\leq \lambda(q,N)\sqrt{Var_E(\hat{Q})},\end{aligned} \tag{28}$$

where

$$\lambda(q,N) = P^*_{q+1}\sqrt{N}\sqrt{\frac{cos(\pi q)}{P_{2q+2,T}(0)}},\ q\in[0,1]. \tag{29}$$

$\lambda(q,N)$ is a function of q and N. In the following steps, we will give a definition for the function $P_{k,T}$ in [1]. The function P^*_{q+1} in Eq. (24) is required to apply Theorem 2.1 for the exact calculation of $\lambda(q,N)$.

From Eq. (28), for true parameter Q a bounded interval (or 100% confidence interval) is given as [17,18]

$$\left(\hat{Q}-\lambda(q,N)\sqrt{Var_E(\hat{Q})}, \hat{Q}+\lambda(q,N)\sqrt{Var_E(\hat{Q})}\right). \tag{30}$$

We need the following periodic function with period T to get the values of constant $\lambda(q,N)$ [1,16],

$$P_{k,T}(x) = P_k\left(\frac{x}{T}-\left[\frac{x}{T}\right]\right) = \frac{-2}{(2\pi)^k}\sum_{j=1}^{\infty}\frac{cos(2\pi j(x/T)-(1/2)\pi k)}{j^k}, \tag{31}$$

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

From Eq. (31), $P_{2q+2,T}(0)$ is obtained,

$$\begin{aligned} P_{2q+2,T}(0) &= \frac{-2}{(2\pi)^{2q+2}} \sum_{j=1}^{\infty} \frac{cos(2\pi j(0/T) - (1/2)\pi(2q+2))}{j^{2q+2}} \\ P_{2q+2,T}(0) &= \frac{2}{(2\pi)^{2q+2}} \zeta(2q+2) cos(\pi q). \end{aligned} \tag{32}$$

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

Theorem 2.1. *1. Supposing that we have a right-open side interval $[a,b)$, the maximum value of the interval is*

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

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

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

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

$$\lim_{h\to 0}(a+h) = \lim_{h\to\infty}(a+1/h),$$
$$\lim_{h\to 0}(b-h) = \lim_{h\to\infty}(b-1/h).$$

In order to calculate P^*_{q+1} in Eq. (24) exactly, Δ is replaced with $1-h$, because we must take the maximum value of Δ. When $h \to 0$, we get the result for the function P^*_{q+1} given in the following form,

$$P^*_{q+1} = \lim_{h\to 0} \mid \frac{-2}{(2\pi)^{q+1}} \sum_{j=1}^{\infty} \left[\frac{1}{j^{q+1}} cos(2\pi(1-h)j - \frac{\pi}{2}(q+1) + 2\pi) \right] \mid . \tag{33}$$

Now we can explain how to get the values of constant $\lambda(q,N)$ below. In 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 Eq. (29) are replaced with 2 and Eq. (32), respectively, we get an equation,

$$\lambda(q,N) = P^*_{q+1}\sqrt{2}\sqrt{\frac{(2\pi)^{2q+2}}{2\zeta(2q+2)}}. \tag{34}$$

The codes written in Mathematica 7 or higher versions 8 and 9 do not compute the value of $\lambda(0,2)$ and give infinity. So, for the calculation of $q=0$, we use Eq. (35) in order to give the result.

By means of Eq. (34), we get $\lambda(0,2)$:

$$\begin{aligned} P^*_{q+1} &= \lim_{h\to 0} \mid \frac{-2}{(2\pi)} \sum_{j=1}^{\infty} \frac{1}{j} cos(2\pi(1-h)j + \frac{3\pi}{2}) \mid = 1/2 \\ \lambda(0,2) &= \frac{1}{2} \cdot \sqrt{\frac{(2\pi)^2}{\zeta(2)}} = 2.44949. \end{aligned} \tag{35}$$

By using Eq. (34), for $q = 0.1$,

$$\begin{aligned} \lambda(0.1,2) &= \lim_{h\to 0} \mid \frac{-2}{(2\pi)^{1.1}} \sum_{j=1}^{\infty} \frac{1}{j^{1.1}} cos(2\pi(1-h)j - \frac{\pi}{2}(1.1) + 2\pi) \mid \cdot \sqrt{\frac{(2\pi)^{2.2}}{\zeta(2.2)}} \\ &= 2.71243 \end{aligned} \tag{36}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

is gotten. For other q values, similar procedure is proceeded. By using the formula given in Eq. (29), the values at Table 1 are obtained for $N = 2$, $N = 3$ and $N = 4$ [13].

It is attested that $\lim_{h\to\infty}(1 - 1/h)$ and $\lim_{h\to 0}(1 - h)$ 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 data set,

$$\hat{Q} - \lambda(q, N) \cdot \hat{ce}(\hat{Q}) \cdot \hat{Q} \leq Q \leq \hat{Q} + \lambda(q, N) \cdot \hat{ce}(\hat{Q}) \cdot \hat{Q} \tag{37}$$

are used, where $\lambda(q, N)$ is the coefficient of confidence interval. $\hat{ce}$ is an estimated coefficient of error [17,18].

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 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 sequence in codes prepared by [17] is increased, the computational values of $\lambda(q, N)$ convergence 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 sequence 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 $N = 2$

$S_1 : \lambda(q,N)$	$S_2 : \lambda(q,N)$	$S_3 : \lambda(q,N)$	$S_4 : \lambda(q,N)$
2.402871	2.418465	-	-
2.671264	-	-	-
2.89612	-	-	2.926749
3.07818	3.105091	3.111937	3.114245
3.216635	-	-	3.258353
3.323898	-	-	3.365038
3.397839	3.408876	3.412577	3.414001
3.418973	-	-	3.426855
3.383912	-	3.38736	3.387638
3.297052	3.29831	3.298648	3.298765
3.161317	-	-	-

3. Results

This section will give the simulation results for the estimation of parameter q in the variance extension term in Eq. (19). Together with the estimated $\hat{q}$ with $k = 2$ for Eqs. (7)-(10), 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 the Matheron's covariogram model and the empirical true variance estimations are used. Real data examples are given to test the performance of 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

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 success for each replication of sampling are counted and divided by the number of replication. Thus, the values of coverage probability of 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 2, because $h = iT$ must be near zero in order to increase the information in Eq. (6).

In the simulation performed, the number of replication is 3000 and the numbers of systematic sampling, n, on $\mathbb{R}$ are from 5 to 20 for the measurement functions in Eqs. (7)-(10). The number of systematic sampling is taken to be 20 at most for the practical purpose.

The following lines represent the producing 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 into [9], it is obvious that the changing of MF affects the variance extension term given in Eq. (14). In this sense, the covariogram model in Eq. (11) can be adopted for the some measurement functions. The simulation results in Tables $5-16$ show that the covariogram model can be used for the measurement functions given in Eqs. (7), (8) and (10). Note that the integral of function in Eq. (8) is computed by means of the numerical integration in MATLAB 2013a, but the integrals of functions in Eq. (7), (9) and (10) are computed by means of the 'int' function which is a function for the exact calculation of integral in MATLAB 2013a. The covariogram model in Eq. (11) is proposed intuitively by [28,29]. We want to use it to check the performance of variance estimation in Eq. (19). By using the covariogram model in Eq. (11), the estimation of q can be said as an open problem for the measurement functions in Eqs. (7)- (10) except to Eq. (7) with $q = 0.4$. It is observed from the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

simulation results that the performance of q formula depends on the covariogram model in Eq. (11) and the variance extension term in Eq. (19). For irregular MF, such as Eqs. (9)-(10), 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 equidistant systematically sampling version of f changes. When we have regular patterns, such as Eqs. (7)-(8), the estimations of q are stable according to the sample size.

Suppose that we have a flexible balloon filled by gas. This three-dimensional balloon can have different shapes after modification done by hands. Thus, each measurement function can be in a class for the functions with parameter q. As it is seen, Eqs. (9) and (10) 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 Eqs. (11) and (19).

The estimated values for q of the measurement functions in Eq. (7) with $q = 0.8$ and also Eq. (8) with $q = 0.4$ and $q = 0.8$ would be around the true parameter values. It is seen that the covariogram model in Eq. (11) can not 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 Eqs. (7), (8), and (10) contain the same numbers of zero digits at the fraction part at most times under Eqs. (11) and (19). It should be noted that even though MF in Eq. (10) does not have the parameter q, the values of $CEMC$ and $CEET$ contain the same number of zero in fraction part when the sample sizes are from 6 to 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 Eq. (10) are more approximate than that of Eqs. (7), (8) generally, which shows an 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 modelled 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 Eq. (9) is not as good as that of Eq. (10) because of Eqs. (11) and (19). The covariogram function of Eq. (9) can not be modelled by the covariogram model. The precision of systematic sampling of MF in Eq. (9) should be improved. This shows us that the proposed measurement functions are also used while making a research in the applied science when Eq. (19) is used. As a result, for Eqs. (7), (8) and (10), Eqs. (11) and (19) give the values around $CEET$, however for Eq. (9) the values of $CEMC$ and $CEET$ do not contain the same numbers of zero digits at 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 Eq. (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 the Matheron's covariogram model are 100%; but the coverage probabilities of confidence intervals based on the coefficient of error for the empirical true variance can not 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 covariogram model with parameter q is that the bandwidth of confidence can be more accurate for a class of measurement functions, such as Eqs. (7), (8) and (10).

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

$CEMC$ can be approximate to $CEET$ for Eqs. (7), (8) and (10), and so the coverage probabilities of the confidence intervals based on $CEMC$ and $CEET$, which are called as $CPCEMC$ 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 confidence interval whether or not it includes the true value, we should focus on $CPCEMC$ and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

$CPCEET$ must give similar results together. In this point, the trustiness of confidence interval is acceptable, because $CPCEET$ is at least 95% approximately. In all performed simulation, N in $\lambda(q,N)$ is taken to be 2, which shows us that the confidence interval keeps itself on 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 3, the bandwidths of confidence interval in $\lambda(q,3)$ case are larger than that of $\lambda(q,2)$. In such a case, we would have a non useful information for the confidence band of $\hat{Q}$. The 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 Eq. (9), since $CPCEET$ is 100%, it seems the constant $\lambda(q,N)$ produces the confidence intervals, which includes the true value of 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 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 that of $CEMC$. When we look at Table 16, $CPCEMC$ and $CPCEET$ have same values except that the number of sampling n is equal to 5. The more number of sampling means the more information gained from the MF. This is one of the reasons why the proposed MF must be used. For MF in Eq. (10), 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 a right value for the constant $\lambda(q,N)$. Finally, proposing a class of measurement functions for the researches helps us to control the trustworthiness of the researches in the applied science if the experts in area can propose them correctly.

As a final comment for the simulation section, the covariogram model in Eq. (11) and the variance extension term in Eq. (19) can be used. However, the proposed measurement functions, such as Eqs. (7) - (10), should be systematically sampled so that one can get the more precise decision on the application for the biomedical imaging. As an another solution, Eq. (19) may be reformulated in the framework of fractional Fourier transformation proposed by [7,34].

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 5 mm slice thickness. The true values of volume for each brain were obtained by using Archimedean principle repeated in 6 times. The arithematic mean of 6 results for each brain was used as a true value of volume of a brain. After MRI scanned the sheep brains, the area of each digital images for slices of brains was computed by pixel counting. While doing the pixel counting, the edges of each digital images were detected by experts in the anatomy area. After that, they were estimated by the slices in coronal plane. The results are given in Tables 3-4. Tables 3-4 show the true values for the volume (Q), the estimated volume ($\hat{Q}$), the estimated smoothness constant ($\hat{q}$), the estimated coefficient of error ($\hat{ce}$), 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 at the Figure 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

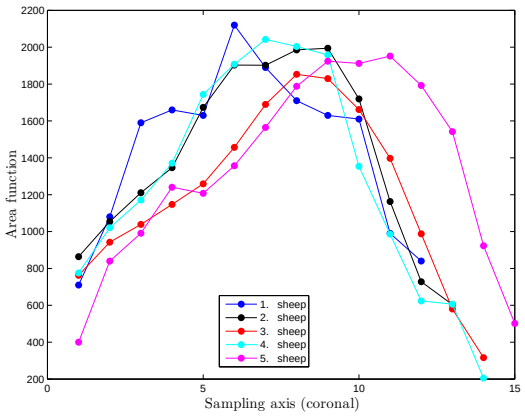


Figure 3. Area functions for each brain.

Table 3. 5 Sheep brain volumes from automatic pixel counting (mm^2) after determining border of slices with 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 Eq. (22)

Brains	Q	$\hat{Q}$	$\hat{q}$	$\hat{ce}$	$\hat{Q}_{lower}$	$\hat{Q}_{upper}$	Num. Slc.
1	85000	87300	.20	.0234	81289.05	93310.95	12
2	94000	90738	.34	.0106	87731.40	93744.60	13
3	83000	84608	.50	.0066	82722.64	86493.36	14
4	88000	88846	.42	.0087	86318.13	91372.87	14
5	100000	99675	.67	.0039	98351.66	100998.34	15

When N in Eq. (29) is replaced with 3, $\lambda(0.3, 3) = 3.82689$ and $\lambda(0.34, 3) = 3.90405$ are found. The confidence intervals of brain 2 for $q = 0.3$ and $q = 0.34$ are (87055.68,94420.32) and (86981.44, 94494.56), respectively. These confidence intervals include the true value of volume for the brain 2.

Table 4. 5 Sheep brain volumes from automatic pixel counting (mm^2) after determining border of slices with 5 mm thickness and their confidence intervals for $N = 2$: $\hat{q}$ values with $k = 2$ in Eq. (22)

Brains	Q	$\hat{Q}$	$\hat{q}$	$\hat{ce}$	$\hat{Q}_{lower}$	$\hat{Q}_{upper}$	Num. Slc.
1	85000	87300	-.06	.0234	82646.98	91953.02	12
2	94000	90738	.18	.0131	87249.31	94226.69	13
3	83000	84608	.41	.0075	82464.84	86751.16	14
4	88000	88846	.28	.0105	85934.10	91756.90	14
5	100000	99675	.63	.0041	98283.46	101066.54	15

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 brains. 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 2. 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 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 to Eq. (7) 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 $\lambda(q, N)$ are bigger than 2 is not useful due to loss on the information, however we do not have a solution for the real data case. In real data case, we have to come across trying the different values of k and N

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

and also the proposed MF must be used to increase the trustworthiness of researches in the applied science. As a result, using two propositions is required.

4. Conclusions and Discussions

The studies [9–12,16,19–21,25–27,30] focused on the variance estimation for the systematic sampling. As implied by [11,15,16], the estimation of q is important to avoid the biasedness of the variance estimator in systematic sampling on $\mathbb{R}$. Unbiasedness of variance extension term estimator leads to have the accurate lower and upper bounds of confidence interval for the systematic sampling on $\mathbb{R}$. In fact, the true variance in Eq. (4) is the combination of three components, so we always have biased variance estimation, however for the practical purpose, we are only interested in the variance extension term. It is observed from the simulation results that $CEMC$ and $CEET$ have the values which contain the same numbers of zero digits at the fraction part for the measurement functions in Eqs. (7), (8), (10) at most times under Eqs. (11) and (19). However, the same performance of the formula in Eq. (19) is not observed for the variance approximation of systematic sampling of MF in Eq. (9) even if the number of sample size increases. So, the formulae in Eqs. (11) and Eq. (19) are not appropriate to use. It should be noted that due to the definition of systematic sampling given in Eq. (1), a fluctuation can be observed, so evaluating the performance of $CEMC$ and $CEET$ according to the behaviour of the sample size can not be sensible. However, when the sample size increases, it is logical to expect the information coming from MF will be increased. In this context, Table $5-16$ show that when the sample size increases, the values of $CEMC$ and $CEET$ are decreased. As implied by p. 319 into [9], if a measurement function has irregularities, such as Eqs. (8), (9), and (10), 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 Eqs. (8), (9), and (10), Table 10-16 support this implication into p. 319 of [9]. The $CEET$ values of MF in Eq. (8) 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 $CEMC$ which are approximate to $CEET$.

A three-dimensional object is represented by a flexible balloon filled by gas. The shape of this balloon can be modified by means of hands. The volume of balloon can change after each modification, however it is not important whether or not there is a changing of the volume of balloon. The modified versions of the balloon are represented by the measurement functions given in Eqs. (7) - (10). The measurement functions in Eqs. (9)-(10) do not have the parameter q, however they can be a neighborhood of MF having parameter q. Since we have different form of the balloon, we will have a lot of forms of the measurement functions representing the three-dimensional object exactly. In real data case, knowing the exact form of three-dimensional object is impossible. The covariogram functions coming from the MF without q will be 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 can not be known exactly due to fact that we can have MF without q which is a neighborhood of MF with q. In the 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 a request. The estimation of q is an open problem even if we know the exact form of the measurement functions except to Eq. (7) with $q = 0.4$. In other perspective of our discussion, the covariogram model can not be a very good approximation for the covariogram functions of measurement functions. However,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

the values of $CEMC$ and $CEET$ contain the same numbers of zero digits at the fraction part at most times for the measurement functions in Eqs. (7), (8), (10) under Eqs. (11) and (19).

The values of $CPCEMC$ and $CPCEET$ are other criteria to approve the performance of exact values of the constant $\lambda(q, N)$, as observed from the simulation results. A numerical computation for the constant $\lambda(q, N)$ of confidence interval was done by [17]. The more precise values of the constant $\lambda(q, N)$ mean the more precise 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 [17] 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 Eqs. (7), (8) and (10).

It is observed that the synthetic data can approve the real data for non-vivo brains if the real data has an exactly similar form with the synthetic data. Since the real data can not 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 estimation of q in real data. Eq. (22) can produce the negative estimated $\hat{q}$ value for the area function of 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 can need constant $\lambda(q, N)$ values with different N. In this context, the edge of digital images has 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 [33] can be used.

The proposed measurement functions should be systematically sampled while conducting a research on the biomedical imaging to increase the information in the decision rule for now. To be able to increase the performance of 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 [35,40] will be light for the computational integration of other MF types that we will use. For construction of MF via the digital images which have coronal, axial or sagittal directions, the new edge detection methods proposed by [33] may be applied to get more precise area values of slices. We will prepare a free statistical software R package with a macro of Mathematica for all methodologies given here and in future.

Appendix A

Table 5. $(1-x^2)^q$: $q = 0.4$

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 6. $(1-x^2)^q$: $q = 0.4$

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2016 doi:10.20944/preprints201608.0017.v4

Table 7. $(1-x^2)^q$: $q = 0.8$

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 8. $(1-x^2)^q$: $q = 0.8$

n	$CEMC$	$CEET$	$CPCEMC\%$	$CPCEET\%$
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 9. $((1-cos(x))(1-x^2))^q$: $q = 0.4$

n	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
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 10. $((1-cos(x))(1-x^2))^q$: $q = 0.4$

n	$CEMC$	$CEET$	$CPCEMC\%$	$CPCEET\%$
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 11. $((1-cos(x))(1-x^2))^q$: $q = 0.8$

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 12. $((1-cos(x))(1-x^2))^q$: $q = 0.8$

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 13. $exp(-sin(-x^3))$: $q = 0.95$

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 14. $exp(-sin(-x^3))$: $q = 0.95$

n	$CEMC$	$CEET$	$CPCEMC\%$	$CPCEET\%$
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 15. $(5/112)(-54x^4 - 25x^3 + 48x^2 + 25x + 6) : q = 0.9$

n	$\hat{q}$	$\widehat{Var}(\hat{q})$	$\widehat{MSE}(\hat{q})$
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 16. $(5/112)(-54x^4 - 25x^3 + 48x^2 + 25x + 6) : q = 0.9$

n	$CEMC$	$CEET$	$CPCEMC\%$	$CPCEET\%$
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

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

References

1. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions*, Dover, New York, **1970**, 1058p.
2. 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.
3. 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. *Journal of Craniofacial Surgery*, **2009**, 20, 921-938.
4. 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 Research*, **2010**, 1310, 29 -36.
5. 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. *Aesthetic plastic sur*. **2004**, 28, 275 - 280.
6. Baddeley, A.J.; Vedel Jensen, E.B. *Stereology for Statisticians*, Monographs on Statistics and Applied Probability, Chapman & Hall/CRC, USA. **2005**, 380p.
7. Baleanu D.; Wu G.C.; Duan J.S. Some Analytical Techniques in Fractional Calculus: Realities and Challenges, Discontinuity and Complexity in Nonlinear Physical Systems, Volume 6 of the series Nonlinear Systems and Complexity. **2013**, 35-62.
8. Bellhouse, D.R.; Krishnaiah, P.R.; Rao, C.R. (Editors), *Sampling: Handbook of Statistics*, Vol. 6, North-Holland, Amsterdam. **1988**, 125-145.
9. Cruz-Orive, L.M. On the Precision of Systematic Sampling:A Review of Matheron's Transitive Methods. *Journal of Microscopy*. **1989**, 153, 315-333.
10. Cruz-Orive, L.M. Systematic Sampling in Stereology. Bull. Intern. Statis. Inst., Proceedinngs 49th Session, Firenze. **1993**, 55, 451-468.
11. Cruz-Orive, L.M. A General Variance Predictor for Cavalieri Slices. *Journal of Microscopy*. **2006**, 222, 158-165.

12. Cruz-Orive, L.M. Variance predictors for isotropic geometric sampling, with applications in forestry. *Statistical Methods & Applications*. **2012**, 22, 3-31.
13. Çankaya, M.N. Stereological Estimation and Inference with Applications. Unpublished Master Thesis, Council of Higher Education. **2010**, 168p.
14. 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. *British Institute of Radiology*. **2010**, 83, 307-317.
15. 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.
16. García-Fiñana, M.; Cruz-Orive, L.M. Improved Variance Prediction for Systematic Sampling on **R**. *Taylor and Francis Group, Statistics*. **2004**, 38, 243-272.
17. García-Fiñana, M. Confidence intervals in Cavalieri Sampling. *Journal of Microscopy*. **2006**, 222, 146-157.
18. García-Fiñana, M.; Keller, S.S.; Roberts, N. Confidence Intervals for the Volume of Brain Structures in Cavalieri Sampling with Local Errors. *Journal of Neuroscience Methods*. **2009**, 179, 71-77.
19. Gual Arnau, X.; Cruz-Orive, L.M. Variance prediction under systematic sampling with geometric probes. Adv. Appl. Prob. **1998**, 30, 889-903.
20. Gundersen, H.J.G.; Jensen, E.B. The Efficiency of Systematic Sampling in Stereology and its Prediction. *Journal of Microscopy*. **1987**, 147, 229-263.
21. Gundersen, H.J.G.; Jensen, E.B.V.; Kiêu, K.; Nielsen, J. The Efficiency of Systematic Sampling in Stereology - Reconsidered. *Journal of Microscopy*. **1999**, 193, 199-211.
22. Hall, P.; Ziegel J. Distributions estimators and confidence intervals for stereological volumes, Biometrika. **2011**, 98, 417-431.
23. Howard, M.A.; Roberts N.; García-Fiñana, M.; Cowell, P.E. Volume Estimation of Prefrontal Cortical Subfields using MRI and Stereology. *Brain Research Protocols*. **2003**, 10, 125-138.
24. 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. *The British Journal of Radioology*. **1999**, 72, 236–245.
25. Kellerer, A.M. Exact formulae for the precision of systematic sampling. J. Microsc. **1989**, 153, 285–300.
26. Kiêu, K. Three Lectures on Systematic Geometric Sampling., Memoirs 13. Department of Theoretical Statistics, University of Aarush. **1997**, 100p.
27. Kiêu K.; Souchet, S.; Istas, J. Precision of Systematic Sampling and Transitive Methods, *Journal of Statistical Planning and Inference*. **1999**, 77, 263-279.
28. Matheron, G. Variables régionalisées et leur estimation [Les], Masson et CIE, éditeurs. **1965**.
29. Matheron, G. The Theory of Regionalized Variables and Its Applications. Les Cahires du Centre de Morphologie Mathématique de Fontainebleau, No. 5. Ecole Nationale Supérieure des Mines de Paris. **1971**, F. Fontainebleau.
30. Mattfeldt, T. The accuracy of one-dimensional systematic sampling. J. Microsc. **1989**, 153, 301–313.
31. Maudsley, R.; García-Fiñana, M. Sampling Intensity with Fixed Precision When Estimating Volume of Human Brain Compartments. *Image Anal Stereol*. **2008**, 27, 143-149.
32. McNulty V.; Cruz-Orive L.M.; Roberts N.; Holmes C.J.; Gual-Arnau, X. Estimation of Brain Compartment Volume from MR Cavalieri Slices. *Journal of Computer Assisted Tomography*. **2000**, 24, 466-477.
33. Peters J.F. *Computational Proximity Excursions in the Topology of Digital Images*, Intelligent Systems Reference Library Vol. 102, Springer, Switzerland. **2016**, 433p.
34. Poularikas A.D. *Transforms and Applications Handbook*, Third Edition. CRC Press: Taylor and Fracis Group. **2010**.
35. 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.
36. 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, *Journal of Microscopy*. **1993**, 171, 239-253.
37. Roberts N.; Pubdephat M.J.; McNulty V. The Benefit of Stereology for Quantitative Radiology, *The British Journal of Radiology*. **2000**, 73, 679-697.

526 38. Şahin, B.; Emirzeoğlu, M.; Uzun, A.; İncesu, L.; Bek, Y.; Bilgiç, S.; Kaplan, S. Unbiased Estimation of the
527 Liver Volume by the Cavalieri Principle using Magnetic Resonance Images. *European Journal of Radiology*.
528 **2003**, 47, 164-170.
529 39. Şahin, B.; Ergür, H. Assessment of the Optimum Section Thickness for the Estimation of Liver Volume
530 Using Magnetic Resonance Images: A Stereological Gold Standard Study. European Journal of Radiology.
531 **2005**, 57, 96-101.
532 40. Zeng Y. Approximate solutions of three integral equations by the new Adomian decomposition method. *J.*
533 *Comput. Complex. Appl.* **2016**, 2, 38-43.