ERROR CONVERGENCE RATES FOR ESTIMATES OF
MULTIDIMENSIONAL INTEGRALS OF RANDOM FUNCTIONS

BY

ALBERTO TUBILLA

TECHNICAL REPORT NO. 72
AUGUST 23, 1975

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT MPS74-21416

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ERROR CONVERGENCE RATES FOR ESTIMATES OF
MULTIDIMENSIONAL INTEGRALS OF RANDOM FUNCTIONS

by

ALBERTO TUBILLA

TECHNICAL REPORT NO. 72
AUGUST 23, 1975

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT MPS74-21416

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ACKNOWLEDGMENTS

I would like to express my sincere thanks to my advisor, Professor Paul Switzer, for his helpful guidance and encouragement. I would also like to thank Professors Tom Cover and Ingram Olkin, who read this thesis and made many helpful suggestions. My thanks go also to Laurie Yadon for the fine typing job she did.

Part of this research was supported by the National Science Foundation under Grant GP-30711X-2.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>PRELIMINARIES</td>
<td></td>
</tr>
<tr>
<td>1.1 Statement of the Problem, Summary and Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Stationary and Isotropic Stochastic Processes</td>
<td>8</td>
</tr>
<tr>
<td>1.3 The Euler-MacLaurin Summation Formula and its Extensions</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>MEAN SQUARE ERRORS</td>
<td></td>
</tr>
<tr>
<td>2.1 Preliminaries</td>
<td>21</td>
</tr>
<tr>
<td>2.2 Mean Square Error for Random and Stratified Sampling</td>
<td>21</td>
</tr>
<tr>
<td>2.3 M.S.E. for Systematic and Midpoint Sampling in One Dimension</td>
<td>24</td>
</tr>
<tr>
<td>2.4 M.S.E. for Systematic and Midpoint Sampling in d-Dimensions</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>RATES OF CONVERGENCE FOR SMOOTH COVARIANCE FUNCTIONS</td>
<td></td>
</tr>
<tr>
<td>3.1 Preliminaries</td>
<td>32</td>
</tr>
<tr>
<td>3.2 Random and Stratified Sampling</td>
<td>33</td>
</tr>
<tr>
<td>3.3 Systematic and Midpoint Sampling</td>
<td>36</td>
</tr>
<tr>
<td>3.4 The Use of Mean Square Derivatives in Obtaining Rates of Convergence</td>
<td>41</td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>RATES OF CONVERGENCE FOR A CLASS OF ISOTROPIC COVARIANCE FUNCTIONS</td>
<td></td>
</tr>
<tr>
<td>4.1 Preliminaries</td>
<td>45</td>
</tr>
<tr>
<td>4.2 Stratified and Random Sampling</td>
<td>46</td>
</tr>
<tr>
<td>4.3 A Geometrical Interpretation of $h'(0)$ for an Isotropic Set Process</td>
<td>47</td>
</tr>
<tr>
<td>4.4 Systematic and Midpoint Sampling</td>
<td>52</td>
</tr>
<tr>
<td>4.5 Some Numerical Results on Systematic and Midpoint Sampling</td>
<td>72</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>77</td>
</tr>
</tbody>
</table>
Chapter 1

Preliminaries

1.1 Statement of the Problem, Summary and Background

Scientists and technicians in several fields are often interested in the properties of a numerical variable \( z(x) \) distributed in space. Of particular importance in many applications is the average value of \( z(x) \) over a certain region \( H \). For example, one may wish to evaluate the average grade of ore in a specified ore body or the proportion of cloud cover in satellite imagery.

Often the function \( z(x) \) will vary in a complex unknown way, and its average value will have to be estimated from samples of \( z(x) \) taken at a particular set of locations.

Naturally we are concerned with the precision of the estimation procedure. The experimenter can usually decide how to sample the function, and his choice should lead to a low estimation error. An assessment of the precision of a sample scheme could bring great savings in many applications where the process of sampling is a costly or time-consuming operation.

It is clear that not much can be said about precision if we impose no further conditions on the function being sampled. A natural restriction would be to consider only those functions that have a certain mathematical property. A greater flexibility is obtained, however, if we assume that \( z(x) \) belongs to a certain space of functions \( \Omega \) where a measure \( P \) has been defined.
We are thus led to consider the numerical variable \( z(x) \) as the realization of a stochastic process and to express the accuracy of a sample scheme as an average accuracy over the class \( \Omega \). The pair \((\Omega, P)\) may then be interpreted as a prior distribution for the unknown function \( z(x) \). The average measure of the accuracy of a sampling scheme becomes then a Bayes risk with respect to this prior.

To be precise, consider a real stochastic process \( z(x) \) defined in the \( d \)-dimensional Euclidean space \( \mathbb{R}^d \). Let \( H \) be the \( d \)-dimensional unit hypercube. We will consider the estimation of the random variable given by the stochastic integral

\[
I(H) = \int_H z(x) \, dx
\]

by means of the statistic

\[
\bar{z} = \frac{1}{N} \sum_{i=1}^{N} z(x_i)
\]

where \( x_1, x_2, \ldots, x_N \) represent \( N \) selected sample points in \( H \) where the process is observed.

The set of rules by which the sample points, \( x_1 \), are chosen will be called a sample design. The sample design may or may not involve a random experiment, but if it does, it will be assumed that this experiment is independent of \( z(x) \).

The performance of a sample design may be judged on the basis of its mean square error (M.S.E.),

\[
\sigma^2 = \mathbb{E}_P \mathbb{E}_R (\bar{z} - I(H))^2,
\]
where $P$ and $R$ stand, respectively, for expectations taken over the underlying process and the randomness, if any, of the sample design. Note that in (1.1.3) both $\bar{Z}$ and $I(H)$ are random variables.

The main object of this thesis is to study the asymptotic behavior of $\sigma^2$, as $N \to \infty$, for four particular sample designs and under two sets of assumptions on the covariance structure of $Z(x)$.

The particular designs we will consider are the following (we assume that $n = N^{1/d}$ is an integer):

**Simple random sampling.** Each sample point is chosen independently of the others and with a uniform distribution over $H$.

**Stratified sampling.** The $i$-th sample point is chosen independently of the others, uniformly on $H_i$, where $H_i$, $i = 1, 2, \ldots, N$, stand for the $N$ disjoint hypercubes of side $\frac{1}{n}$ into which $H$ may be subdivided.

**Systematic sampling.** The sample point $x_1$ is chosen uniformly on $H_1$. For $i = 2, 3, \ldots, N$ translate $H_1$ so as to make it coincide with $H_i$. The sample point $x_i$ will then be the image of $x_1$ under the translation.

**Midpoint sampling.** The sample points are chosen to be the centers of the hypercubes $H_i$.

We now outline the organization of this thesis. In Sections 1.2 and 1.3 we review some preliminary material. Exact formulas for the mean square errors are found in Chapter 2. These formulas are then used, in the remaining chapters, to study the asymptotic behavior of $\sigma^2$ for smooth covariance functions (Chapter 3) and for a certain class of isotropic covariance functions (Chapter 4).
The remainder of this section contains a summary of some of the results in later chapters and a discussion of related literature. Some of the results in Chapters 3 and 4 are displayed in Table 1.1.1. Each entry in this table represents the order of magnitude of $\sigma^2$ for the particular situation there described. The displayed term appears in the expansion for $\sigma^2$ multiplied by a constant which under special circumstances might vanish. We have, in the corresponding theorems, discussed conditions under which this happens.

<table>
<thead>
<tr>
<th></th>
<th>Chapter 3</th>
<th>Chapter 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>General d *</td>
<td>$\frac{1}{N}$</td>
<td>$\frac{1}{N}$</td>
</tr>
<tr>
<td>d = 1 **</td>
<td>$\frac{1}{N}$</td>
<td>$\frac{1}{N}$</td>
</tr>
<tr>
<td>d = 2 ***</td>
<td>$\frac{1}{N}$</td>
<td>$\frac{1}{N}$</td>
</tr>
<tr>
<td>Random Sampling</td>
<td>$N^{-1}$</td>
<td>$N^{-1}$</td>
</tr>
<tr>
<td>Stratified Sampling</td>
<td>$N^{-1+2/d}$</td>
<td>$N^{-2}$</td>
</tr>
<tr>
<td>Systematic Sampling</td>
<td>$N^{-2/d}$</td>
<td>$N^{-2}$</td>
</tr>
<tr>
<td>Midpoint Sampling</td>
<td>$N^{-4/d}$</td>
<td>$N^{-2}$</td>
</tr>
</tbody>
</table>

Table 1.1.1
Rates of Convergence of the Mean Square Errors

* Theorems 4 and 5
** Theorem 6 and Remark 1 of Section 3.3
*** Theorems 6 and 7

We shall make some remarks with respect to Table 1.1.1. Note how, in the first column, the rates of convergence of the last three designs decrease as the dimension increases. Random sampling has a rate of
convergence which is independent of the dimension and practically inde-
pendent of any assumption whatsoever on the process. The further
assumption of isotropy in the first column would not, in general,
increase the rate of convergence of \( \sigma^2 \). This can easily be seen by
checking that the constants do not vanish for the isotropic smooth
covariance function \( \exp(-x^2 - y^2) \) using the results of Theorems 4 and 5.

Chapter 4 deals with a certain class of isotropic covariance func-
tions defined in terms of smoothness conditions. In \( \mathbb{R}_1 \) any covariance
function is isotropic and the results of column two apply to any covari-
ance function that has the differentiability properties required in
that chapter.

In the case of the two-dimensional isotropic covariance functions
of Chapter 4, Table 1.1.1 shows that stratified and midpoint sampling
have the same rate of convergence. It can easily be seen, however,
that the constants that multiply the term \( N^{-3/2} \) are different in each
case, the smallest one being that of midpoint sampling. Here systematic
sampling has a slower rate of convergence than either stratified or
midpoint sampling. This result is a little surprising and it is
undoubtedly due to the border effects in \( H \). We remark, finally, that
the results of column three in Table 1.1.1 apply to many instances of
isotropic zero-one set processes.

We finish this section with a discussion of related results in
some of the existing literature. An early paper that considers methods
of sampling in the plane is Quenuville (1949). He develops formulas
for the mean square errors incorporating the idea of a sampled population
which is itself a sample from an infinite population. This is the same approach we follow in this thesis by considering \( z(x) \) to be a stochastic process. [The idea had been considered also by Cochran (1946) in connection with the one-dimensional case.] Although not in the context of rates of convergence, Quenouille makes some comparisons between the sample designs. He concludes that for a variety of cases systematic sampling is more accurate than random sampling.

Zubrzycki (1958) gives formulas for the mean square errors of random, stratified and systematic sampling in the context of two-dimensional isotropic covariance functions. He, however, does not exploit the fact that these formulas represent the difference between an integral and its approximation by means of a finite sum. For the exponential correlation function, \( k \exp(-\alpha \sqrt{x^2+y^2}) \), he shows that, for fixed domains, systematic sampling will be better than stratified sampling provided \( \alpha \) is big enough. How big \( \alpha \) should be is determined by the shape and size of the domains.

In the general topic of random functions on Euclidean spaces as models for natural phenomena there are two important references: Matérn (1960) and Matheron (1965). They will often be cited in this thesis.

The work of Matérn contains a very complete discussion of the structure of covariance functions in \( \mathbb{R}^d \) plus a variety of concrete examples of stochastic processes that have been considered useful in the description of natural phenomena. In Chapter 5 of his work he compares the performance of different sampling designs under the assumption that the sampling is done in the infinite plane but assuming
that the average number of sample points per unit area is fixed. He then proceeds to calculate numerically the "variance per sample point" for the various sampling designs using particular covariance functions. In this thesis we study the asymptotic behavior of \( \sigma^2 \) by letting the sampling density increase on a fixed bounded region. Matérn's approach is exactly the opposite and neglects a very important border effect produced by the geometrical shape of the region where the sampling takes place. This border effect shows clearly in the striking differences between systematic and midpoint sampling in Table 1.1.1. These two designs are completely equivalent in Matérn's formulation due to the stationarity assumption.

The work of Matheron is divided into two main parts. In the first part he considers \( z(x) \) as a given fixed function and in the second part he introduces the notion of a stochastic process. The results in these two parts are formally very similar since the analysis for fixed functions is based mainly on the "empirical covariance function." He develops formulas for the numerical approximation of the mean square errors using the same limiting procedure as Matérn. In the case of systematic sampling he obtains, by the use of spectral theory, numerical approximations valid for isotropic covariance functions. In contrast, we have not used spectral methods at all and our results are based instead on multidimensional Euler-MacLaurin expansions which greatly simplify the study of convergence rates when sampling on the unit cube.

Finally we point out that the question of optimal sampling designs has not been discussed at all in this thesis. Optimal designs require,
for their application, detailed prior information about the structure of the stochastic process being sampled. Some references in this field are Matheron (1965) and Sacks and Ylvisaker (1971).

1.2 **Stationary and Isotropic Stochastic Processes**

In this section we review some concepts in the theory of stochastic processes that will be needed in the sequel. Details, proofs and further references can be found in Cramér and Leadbetter (1967), Matérn (1960) and Yaglom (1962).

Let \( z(\mathbf{x}) \) be a stochastic process defined in the \( d \)-dimensional Euclidean space \( \mathbb{R}^d \) such that

\[
(1.2.1) \quad \mathbb{E} z^2(\mathbf{x}) < \infty.
\]

The mean function of \( z(\mathbf{x}) \) is

\[
(1.2.2) \quad \mathbb{E} z(\mathbf{x}) = \mu(\mathbf{x});
\]

and the covariance function is defined as

\[
(1.2.3) \quad \mathbb{E} \left( z(\mathbf{x}) - \mu(\mathbf{x}) \right) \left( z(\mathbf{y}) - \mu(\mathbf{y}) \right) = c(\mathbf{x}, \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d.
\]

Condition (1.2.1) implies that both (1.2.2) and (1.2.3) are finite.

Any function \( \mu(\mathbf{x}) \) is the mean function of a stochastic process \( z(\mathbf{x}) \); for covariance functions the following is true: \( c(\mathbf{x}, \mathbf{y}) \) is the covariance function of a process \( z(\mathbf{x}) \) if and only if \( c(\mathbf{x}, \mathbf{y}) \) is non-negative definite.
If the random variable $z(x)$ takes only the values 0 or 1 we shall say that $z(x)$ is a set process. We will then call

$$A = \{x | z(x) = 1\}$$

the random set. In this context the mean function has the interpretation

$$M(x) = Pr(x \in A)$$

and the covariance function

$$c(x, y) = Pr(x \in A, y \in A) - Pr(x \in A) Pr(y \in A).$$

A process $z(x)$ satisfying (1.2.1) is said to be stationary in the wide sense if both the mean and covariance functions are invariant under translations. This implies that $m(x)$ is a constant and that $c(x, y)$ depends on $x$ and $y$ only through the difference vector $x - y$. We will then simply write $c(x - y)$ instead of $c(x, y)$.

A further simplification is obtained if we require invariance of the mean and covariance functions under both translations and rotations. If this is the case we shall say that $z(x)$ is isotropic. The covariance function will then have the form

$$c(x - y) = h(r),$$

where $r = |x - y|$ denotes the Euclidean distance between $x$ and $y$. 

9
We shall, from now on, deal exclusively with stationary processes, although some of the following concepts can be defined in the general setting of a process satisfying (1.2.1).

The process \( z(x) \) is said to be continuous in quadratic mean (abbreviated q.m.) if at every point \( x_0 \),

\[
E \left[ z(x) - z(x_0) \right]^2 \to 0,
\]

as \( |x - x_0| \to 0 \). The process \( z(x) \) is continuous in q.m. if and only if \( c(x-y) \) is continuous at the origin, which in turn implies that \( c \) is continuous everywhere.

Let \( S \) be a bounded region in \( \mathbb{R}^d \) and let \( g(x) \) be a non-random function continuous on \( S \). If \( c \) is continuous, the random variable

\[
(1.2.4) \quad I(S) = \int_S z(x) \, g(x) \, dx
\]

is defined as the limit, in the mean square sense, of the well known approximating sums formally associated with the integral. Suppose that, with probability 1, the realizations of the process are integrable functions over \( S \) in the ordinary Riemann sense. It can then be shown, under certain regularity conditions, that for almost all sample paths the value of \( I(S) \) will be equal to the value of the ordinary Riemann integral obtained by integrating, over \( S \), the product of the sample function and \( g(x) \).

Let \( S_1 \) and \( S_2 \) be two measurable sets in \( \mathbb{R}^d \) and assume that the two stochastic integrals \( I(S_1) \) and \( I(S_2) \) are defined; then
(1.2.5) \[ \text{EI}(S_1) = p u(S_1), \]

where \( p = m(x) = \mathbb{E}z(x) \) and \( u(S_1) \) is the Lebesque measure of the set \( S_1 \). Furthermore,

(1.2.6) \[ \text{cov} \left( I(S_1), I(S_2) \right) = \int_{S_1} \int_{S_2} c(x-y) dx \, dy, \]

Let \( R_1 \) and \( R_2 \) be two finite subsets of \( R_d \) containing, respectively, \( N_1 \) and \( N_2 \) points. Let \( \bar{z}(R_1) \) and \( \bar{z}(R_2) \) stand for the corresponding average of the process over each one; then

(1.2.7) \[ \text{cov} \left( \bar{z}(R_1), \bar{z}(R_2) \right) = \frac{1}{N_1 N_2} \sum_{x \in R_1} \sum_{y \in R_2} c(x-y), \]

and

(1.2.8) \[ \text{cov} \left( \bar{z}(R_1), I(S_1) \right) = \frac{1}{N_1} \sum_{x \in R_1} \int_{S_1} c(x-y) dy. \]

Assume now that \( z(x) \) is a one-dimensional stationary process. If

\[ \lim_{h \to 0} \mathbb{E} \left[ \frac{z(x+h)-z(x)}{h} - z'(x) \right]^2 = 0, \]

we shall say that \( z'(x) \) is the quadratic mean derivative of \( z(x) \) at the point \( x \). If the second derivative of the covariance function exists at the origin it will then exist everywhere else and the process \( z(x) \) will have a q.m. derivative \( z'(x) \) at all points. Then \( z'(x) \)
will define a new stationary process whose covariance function will be

\[ -\frac{d^2c(x)}{dx^2} \]

with mean

\[ Ez'(x) = 0. \]

Again, if \( z(x) \) has an ordinary sample function derivative (w.p.l) this will be equal to the q.m. derivative with probability one.

Higher derivatives in q.m. can be defined in a completely similar way and they are analogously related to the higher derivatives of the covariance function \( c(x) \). The theory can also be extended to the partial q.m. derivatives of a process defined in \( \mathbb{R}^d \) by considering the one-dimensional processes obtained by keeping fixed all the variables but one.

Derivatives and integrals in q.m. possess the ordinary properties of calculus; we quote for instance the formulas

\[
(1.2.9) \quad z(x) - z(x_0) = \int_{x_0}^{x} z'(x)dx ,
\]

and

\[
(1.2.10) \quad \int_{a}^{b} g(x) z'(x)dx = g(x) \left[ z(x) \right]_{a}^{b} - \int_{a}^{b} g'(x) z(x)dx ,
\]

where we have assumed that the deterministic function \( g(x) \) has a continuous derivative in \( [a,b] \).
We briefly mention the following fundamental result in the theory of stationary stochastic processes. If the covariance function of \( z(\mathbf{x}) \), \( \mathbf{x} \in \mathbb{R}_d \), is continuous at the origin there exists a \( d \)-dimensional random vector whose characteristic function is the covariance function of \( z(\mathbf{x}) \). The c.d.f. of this random vector is called the spectral distribution of the process. If the corresponding joint density exists it will be called the spectral density.

A great variety of concrete examples of stationary stochastic processes can be found in the references given above, in particular Matérn (1960, p. 10-50). For our purposes it will be enough to illustrate some of the ideas by means of the following example:

**Example 1.2.1.** An isotropic stochastic set process defined in \( \mathbb{R}_2 \).

Identify each line in \( \mathbb{R}_2 \) with the polar coordinates \((\rho, \theta)\) of the point where the perpendicular from the origin meets it. Assume that in the \((\rho, \theta)\) plane a two-dimensional Poisson point process has been defined (see Matérn (1960, p. 26)). Each realization of this Poisson process produces a collection of lines in \( \mathbb{R}_2 \) that subdivide the plane in an infinite number of disjoint convex cells. Independently of the others each cell is included in the random set with probability \( p \), i.e., \( z(\mathbf{x}) = 1 \) if \( \mathbf{x} \) is in one of the selected cells and \( z(\mathbf{x}) = 0 \) otherwise.

The set process so obtained is isotropic and exhibits a Markovian property. Along any line, the restricted real process \( z(\lambda) = z(\mathbf{x}_0 + \lambda \mathbf{x}_1) \) is a two-state continuous time Markov chain (Switzer (1965)).
The covariance function of $z(x)$ is of the form

\begin{equation}
(1.2.11) \quad h(r) = k \exp(-ar), \quad r > 0,
\end{equation}

for some positive constants $k$ and $a$. The spectral density of $z(x)$ is given by

\begin{equation}
(1.2.12) \quad \frac{ka}{2\pi} (a^2 + x^2 + y^2)^{-3/2}.
\end{equation}

1.3 The Euler-MacLaurin Estimation Formula and Its Extensions

The purpose of this section is to present, in an organized way, certain facts about the Euler-MacLaurin summation formula that will be essential in our study of the rates of convergence of the mean square errors. A nice treatment of the usual one-dimensional version can be found in Knopp (1947) and the extension to $d$-dimensions has been discussed by Lyness (1964). The fact that a similar formula holds for stochastic integrals and derivatives is a simple extension and will be discussed briefly.

Let $f(x)$ be a function defined on the interval $[0,1]$ such that the $(2k+1)$-derivative $f^{(2k+1)}(x)$ exists and is continuous. For any positive integer $N$ let $T_N(f)$ denote the trapezoidal sum associated with $f$, i.e.,

\begin{equation}
T_N(f) = \frac{1}{N} \left[ \frac{f(0)}{2} + f\left(\frac{1}{N}\right) + \cdots + f\left(\frac{N-1}{N}\right) + \frac{f(1)}{2} \right].
\end{equation}
Then

\[ (1.3.1) \quad T_N(f) - \int_0^1 f(x) \, dx = \sum_{i=1}^{k} \frac{B_{2i}}{(2i)!} \cdot \frac{1}{N^{2i}} \left( f^{(2i-1)}(1) - f^{(2i-1)}(0) \right) + R_k, \]

(Euler-Maclaurin summation formula), where the constants \( B_{2i} \), \( i = 1, 2, \ldots \) are the so called Bernoulli numbers

\[ B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_8 = -\frac{1}{30}, \ldots, \quad B_{2i} = 0 \quad \text{for} \ i \ \text{odd} \]

and the remainder \( R_k \) is given by

\[ R_k = \frac{1}{N^{2k+1}} \int_0^1 P_{2k+1}(Nx) f^{(2k+1)}(x) \, dx. \]

The functions \( P_k(x), \ k = 1, 2, 3, \ldots, \) appearing in the expression for the remainder are periodic functions of period one and have the following Fourier Series representations:

\[ P_{2k}(x) = (-1)^{k-1} \sum_{i=1}^{\infty} \frac{2 \cos 2i \pi x}{(2i \pi)^{2k}}, \]

and

\[ P_{2k+1}(x) = (-1)^{k-1} \sum_{i=1}^{\infty} \frac{2 \sin 2i \pi x}{(2i \pi)^{2k+1}}, \quad k = 0, 1, 2, \ldots. \]

For \( k = 2, 3, \ldots \) they are continuously differentiable with the property

\[ \frac{dP_{k+1}(x)}{dx} = P_k(x); \]
furthermore

\[ P_{2k}(0) = P_{2k}(1) = \frac{B_{2k}}{(2k)!} , \]

\[ P_{2k+1}(0) = P_{2k+1}(1) = 0 . \]

Restricted to the interval \([0,1]\) \(P_k(x)\), \(k = 1,2,\ldots\) is a polynomial of degree \(k\) (the Bernoulli polynomials); we list the first few:

\[ P_1(x) = x - \frac{1}{2} , \]

\[ P_2(x) = \frac{x^2}{2} - \frac{x}{2} + \frac{1}{12} , \]

\[ P_3(x) = \frac{x^3}{6} - \frac{x^2}{4} + \frac{x}{12} . \]

**Proof of (1.3.1).** The proof is based solely on the properties of the functions \(P_k\) and the use of repeated integration by parts. For \(i = 0,1,2,\ldots,N-1\) let \(a_i = i/n\); then

\[ \int_{a_i}^{a_{i+1}} f(x)dx = \frac{1}{2N} \left( f\left(\frac{i+1}{N}\right) + f\left(\frac{i}{N}\right) \right) - \frac{1}{N} \int_{a_i}^{a_{i+1}} f'(x) P_1(Nx)dx , \]

\[ \int_{a_i}^{a_{i+1}} f'(x) P_1(Nx)dx = \frac{B_2}{2!} \left( f\left(\frac{i+1}{N}\right) - f\left(\frac{i}{N}\right) \right) - \frac{1}{N} \int_{a_i}^{a_{i+1}} f''(x) P_2(Nx)dx , \]

and

\[ \int_{a_i}^{a_{i+1}} f''(x) P_2(Nx)dx = -\frac{1}{N} \int_{a_i}^{a_{i+1}} f'''(x) P_3(Nx)dx . \]
If we now combine these three equalities into a single one and add the result for \( i = 0, 1, 2, \ldots, N-1 \), we get (1.3.1) for the particular case \( k = 1 \). (Note that we had to split the interval of integration because \( \mathcal{P}_1'(nx) \) does not exist at the points \( \frac{i}{N} \).) We can continue this procedure as many times as \( f \) is differentiable; we will then obtain the general form of (1.3.1).

It is interesting to observe that if we let \( k \) tend to infinity in (1.3.1) the series on the righthand side will in general diverge. This is due to the fact that the Bernoulli numbers increase very rapidly, in fact,

\[
|B_{2k}| = \frac{2(2k)!}{(2\pi)^{2k}} \theta, \quad 1 < \theta < 2.
\]

However, if \( N \) is large, the first few terms of the series will approximate very accurately the lefthand side of (1.3.1). Examination of the remainder \( R_k \) shows that it is, in general, smaller in absolute than the first term neglected. If \( N \) is large the terms of the series will be very small at first, increasing only later to a high value.

Let \( M_N(f) \) stand for the midpoint sum associated with \( f \), i.e.

\[
M_N(f) = \frac{1}{N} \left[ f\left(\frac{1}{2N}\right) + f\left(\frac{1}{2N} + \frac{1}{N}\right) + \cdots + f\left(1 - \frac{1}{2N}\right) \right].
\]

The relationship

\[
M_N(f) = 2T_{2N}(f) - T_N(f)
\]
is easily checked and can be used to obtain the following expansion:

\[(1.3.2) \quad M_N(f) = \int_0^1 f(x)dx = \sum_{i=1}^k \frac{B_{2i}}{(2i)!} \cdot \frac{2^{-2i}2^i}{2^{2i}} \cdot \frac{1}{N^{2i}} \left[ f^{(2i+1)}(1) - f^{(2i+1)}(0) \right] + R'_k, \]

where

\[R'_k = \frac{1}{N^{2k+1}} \left[ \int_0^1 \left( \frac{P_{2k+1}(2Nt)}{2^{2k}} - P_{2k+1}(Nt) \right) f^{(2k+1)}(t)dt \right].\]

For a function of d-variables \( f \), defined on the unit hypercube \( H \), we will define what is meant by \( T_N(f) \) and \( M_N(f) \). Let \( n \) be a positive integer with \( N = n^d \); then

\[M_N(f) = \frac{1}{N} \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_d=1}^n f \left( \frac{2i_1-1}{2n}, \frac{2i_2-1}{2n}, \ldots, \frac{2i_d-1}{2n} \right),\]

and

\[T_N(f) = \frac{1}{N} \sum_{i_1=0}^n \sum_{i_2=0}^n \cdots \sum_{i_d=0}^n \frac{1}{2} w(i_1, i_2, \ldots, i_d) f \left( \frac{i_1}{n}, \frac{i_2}{n}, \ldots, \frac{i_d}{n} \right),\]

where

\[w(i_1, i_2, \ldots, i_d) = \text{number of } i's \text{ that are equal to zero or to } n.\]

In the d-dimensional case we have the following generalizations of (1.3.1) and (1.3.2): let \( f(x_1, x_2, \ldots, x_d) \) be defined on \( H \) where its partial derivatives up to order \( 2k+1 \) exist and are continuous; then
(1.3.3) \[ T_N(f) = \int_H f \, d\lambda \]
\[ = \frac{k}{n} \sum_{i=1}^{\frac{1}{2k}} t_{2S_1}^{2S_2} \cdots t_{2S_d}^{2S_2} \int_H \frac{2i f}{2S_1^{2S_2} \cdots 2S_d} \, dx + D_k, \]

and

(1.3.4) \[ M_N(f) = \int_H f \, d\lambda \]
\[ = \frac{k}{n} \sum_{i=1}^{\frac{1}{2k}} m_{2S_1}^{2S_2} \cdots m_{2S_d}^{2S_2} \int_H \frac{2i f}{2S_1^{2S_2} \cdots 2S_d} \, dx + D'_k, \]

where the inner sums are taken over all sequences \( S = (S_1, S_2, \ldots, S_d) \) such that \( \sum_{j=1}^{d} S_j = i \). The \( t \)'s and \( m \)'s are defined as follows:

\[ t_{2S_j} = \frac{B_{2S_j}}{(2S_j)!}, \quad t_0 = 1, \]

\[ m_{2S_j} = \frac{B_{2S_j}}{(2S_j)!} \frac{2-2S_j}{2S_j}, \quad m_0 = 1. \]

Lyness (1964) does not give expressions for the remainders \( D_k \) and \( D'_k \).

It may be shown, however, that they have the form

\[ \frac{1}{n^{2k+1}} \int_H \sum_{S} \frac{2^{k+1} f}{S_1 S_2 \cdots S_d} Q_S(n\lambda) \, dx, \]
where $\sum s_i = 2k+1$ and the functions $Q_S(x)$, which are different for $D_k$ and $D'_k$, are all bounded in $R_d$. In fact, in Chapter 4 we will have to find the exact form of $D_k$ for the particular case $k = 1$. Here it will be enough to observe that if the $2k+1$ partial derivatives of $f$ are integrable then $D_k$ and $D'_k$ are

$$0 \left( \frac{1}{n^{2k+1}} \right).$$

We have seen how the proof of the Euler-Maclaurin summation formula relies solely on the formula for integration by parts of ordinary calculus. This formula is also true for the mean square differential and integral calculus defined in Section 1.2 (see (1.2.10)). It should then be clear that the results of this section, in particular (1.3.3) and (1.3.4), hold true if the non-random function $f$ is replaced by a stationary stochastic process $z(x)$, having a smooth covariance function. Derivatives and integrals will then be considered as derivatives and integrals in quadratic mean.
Chapter 2

Mean Square Errors

2.1 Preliminaries

The calculation of the mean square errors will be made under the assumption that the stochastic process \( z(\mathbf{x}) \), \( \mathbf{x} \in \mathbb{R}^d \), is stationary in the wide sense. We assume that the sampling is always done in the \( d \)-dimensional hypercube \( H \).

In the sequel \( \sigma^2_{SR} \), \( \sigma^2_{ST} \), \( \sigma^2_{SY} \) and \( \sigma^2_{MR} \) will denote, respectively, the mean square error of simple random, stratified, systematic and midpoint sampling. The theorems in this chapter give exact expressions for the values of these quantities in terms of \( c(\mathbf{x}) \), \( \mathbf{x} \in \mathbb{R}^d \), the covariance function of the process. With the possible exception of midpoint sampling these expressions are well known (see, for instance, Zubrzycki (1958)). However, the proofs given here are fairly simple and the results have been reinterpreted to facilitate the study of convergence rates in later chapters.

2.2 Mean Square Error for Random and Stratified Sampling

Theorem 1.

\[
(2.2.1) \quad \sigma^2_{SR} = \frac{1}{N} \left( c(Q) - \int_H \int_H c(\mathbf{x}-\mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} \right),
\]

\[
(2.2.2) \quad \sigma^2_{ST} = \frac{1}{N} \left( c(Q) - N^2 \int_{H_1} \int_{H_1} c(\mathbf{x}-\mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} \right),
\]
where \( H_1 = \{ x \mid 0 \leq x_1 \leq \frac{1}{n}, \ i = 1, 2, \ldots, d \} \) and \( N = n^d \) is the total number of sample points.

**Proof.** We start by proving (2.2.1). Given the process the random variables \( z(x_i), \ i = 1, 2, \ldots, N \), are independent and identically distributed with common mean

\[
I(H) = \int_H z(x) \, dx .
\]

Let the symbol \( E_p \) stand for conditional expectation given the process. Then

\[
c^2_{SR} = E(z-I(H))^2 = E E_p(z-I(H))^2 = \frac{1}{N} E E_p(z^2(x_1)-I^2(H)) = \frac{1}{N} (\text{Var } z(x_1) - \text{Var } I(H)) ,
\]

where the last equality follows from the fact that, from (1.2.5),

\[
E z(x_1) = E I(H) = p .
\]

Note that although \( x_1 \) is a random point we can still write

\[
\text{Var } z(x_1) = c(Q) ,
\]

because of the stationarity assumption. The result now follows using Formula (1.2.6) for the variance of \( I(H) \). To prove (2.2.2) recall that for stratified sampling the unit hypercube \( H \) has been subdivided
into \( N \) small hypercubes \( H_i \), \( i = 1, 2, \ldots, N \). We can then write

\[
I(H) = \sum_{i=1}^{N} I(H_i) .
\]

Conditionally on the process the random variables \( z(x_i) \), \( i = 1, 2, \ldots, N \), are independent with

\[
Ez(x_i) = NI(H_i) , \quad i = 1, 2, \ldots, N ,
\]

because \( x_i \) is uniformly distributed over \( H_i \). We have, using these facts, that

\[
\sigma_{ST}^2 = E_E[ (\bar{z} - I(H))^2 ] = \frac{1}{N^2} \sum_{i=1}^{N} E_E[z(x_i) - N z(H_i))^2
\]

\[
= \frac{1}{N^2} \sum_{i=1}^{N} E_E[z^2(x_i) - N^2 I^2(H_i)] = \frac{1}{N^2} \sum_{i=1}^{N} [Var z(x_i) - N^2 Var I(H_i)] ,
\]

since, from (1.2.5),

\[
ENI(H_i) = Npu(H_i) = p .
\]

Again, by the stationarity assumption,

\[
Var z(x_i) = Var z(x_0) = c(0) ,
\]

and

\[
Var I(H_i) = Var I(H_0) = \int_{H_i} \int_{H_i} c(x-y) \, dx \, dy .
\]

The result follows.
2.3 M.S.E. for Systematic and Midpoint Sampling in One Dimension

In the case of systematic and midpoint sampling, the calculations become more involved. Initially, we will treat the one-dimensional case only. We have

**Theorem 2.** If $d = 1$, i.e., $H$ is the interval $[0,1]$, then

\[
\sigma_{SY}^2 = \frac{N}{N N} \sum_{i=-N}^{i=N} \left( 1 - \frac{|i|}{N} \right) c \left( \frac{i}{N} \right) - 2 \int_0^1 (1-x) c(x) \, dx ,
\]

and

\[
\sigma_{MR}^2 = \frac{N}{N N} \sum_{i=-N}^{i=N} \left( 1 - \frac{|i|}{N} \right) c \left( \frac{i}{N} \right) + 2 \int_0^1 (1-x) c(x) \, dx \]

\[ - \frac{2}{N} \sum_{i=1}^{N} \int_0^1 c(x_i-x) \, dx ,
\]

where

\[
x_i = \frac{1}{2N} + \frac{i}{N} , \quad i = 1, 2, \ldots, N .
\]

**Proof.** We have, in general, that

\[
E\bar{Z} = E\text{I}(H) = p ,
\]

and consequently

\[
c^2 = \text{Var} \bar{Z} + \text{Var} \text{I}(H) - 2 \text{Cov}(\bar{Z}, \text{I}(H)) .
\]

It was already pointed out that for systematic sampling

\[
E \frac{\bar{Z}}{p} = \text{I}(H) ,
\]
which implies that
\[
\text{Cov}(\bar{Z}, I(H)) = E\bar{Z}I(H) - p^2 = E\frac{\bar{Z}I(H)}{p} - p^2 = EI(H) - p^2 = \text{Var} I(H).
\]

We can then write
\[
c_{SY}^2 = \text{Var} \bar{Z} - \text{Var} I(H),
\]
and
\[
c_{MR}^2 = \text{Var} \bar{Z} + \text{Var} I(H) - 2 \text{Cov}(\bar{Z}, I(H)).
\]

Note that in \( R \), the sample points of both systematic and midpoint lie on a uniform grid whose points are \( 1/N \) units apart. For any such fixed grid
\[
\text{Var} \bar{Z} = \frac{1}{N} \sum_{i=-N}^{N} (1 - \frac{|i|}{N}) c(\frac{i}{N})
\]
as can be seen by using Formula (1.2.7) and collecting likewise terms.

Formula (1.2.6) gives
\[
\text{Var} I = \int_{0}^{1} \int_{0}^{1} c(x-y) \, dx \, dy.
\]

Make the transformation
\[
u = x - y, \quad v = x,
\]
so that the integral takes the form
\[ \text{Var I} = \int_{-1}^{0} \int_{0}^{1+u} c(u) \, dv \, du + \int_{0}^{1} \int_{u}^{1} c(u) \, dv \, du \]

\[ = \int_{-1}^{0} (1+u) \, c(u) \, du + \int_{0}^{1} (1-u) \, c(u) \, du \]

\[ = 2 \int_{0}^{1} (1-u) \, c(u) \, du , \]

where we have used the fact that \( c(u) = c(-u) \).

Finally, from (1.2.8), we have that

\[ \text{cov}(\bar{z}, I(H)) = \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{1} c(x_i - x) \, dx . \]

This concludes the proof of Theorem 2.

It is possible to rewrite Formulas (2.3.1) and (2.3.2) in a more suggestive form. The limiting properties of \( \sigma^2 \) for these two designs will then become apparent.

Define

\[ \phi(x) = 2(1-x) \, c(x) , \]

\[ P(y) = \int_{0}^{1} c(y - x) \, dx . \]

Note that

\[ \int_{0}^{1} P(y) \, dy = 2 \int_{0}^{1} (1-x) \, c(x) \, dx , \]

26
\[ \phi(1) = 0 \]

so that

\[
\sigma_{SY}^2 = \left[ \frac{1}{N} \left( \frac{\phi(0)}{2} + \phi\left(\frac{1}{N}\right) + \cdots + \phi\left(\frac{N-1}{N}\right) + \frac{\phi(1)}{2} \right) - \int_0^1 \phi(x) \, dx \right],
\]

and

\[
\sigma_{MR}^2 = \left[ \frac{1}{N} \left( \frac{\phi(0)}{2} + \phi\left(\frac{1}{N}\right) + \cdots + \frac{\phi(1)}{2} \right) - \int_0^1 \phi(x) \, dx \right]
- 2 \left[ \frac{1}{N} \sum_{i=1}^N P(x_i) - \int_0^1 P(x) \, dx \right].
\]

Each term in brackets, in the above expressions, represents the difference between an integral and its approximation by a finite sum. They include familiar quadrature methods, i.e., the trapezoidal and the midpoint rule. Later we shall use those representations to study the behavior of \( \sigma^2 \).

2.4 M.S.E. for Systematic and Midpoint Sampling in d Dimensions

We conclude this chapter with the generalization of Formulas (2.3.1) and (2.3.2) to the d-dimensional case.

It will be convenient to introduce some extra notation. In the sequel, the covariance between two points will be frequently written as a function of the coordinates of their difference vector. It should then be clear what is meant by \( c(x_1, x_2, \ldots, x_d) \).

Given a covariance function \( c \), consider the "associated functions"
\[ \phi(x) = (1-x_1)(1-x_2) \cdots (1-x_d) \prod c(tx_1, tx_2, \ldots, tx_d), \]

and

\[ P(x) = \int_H c(x_1 - y_1, x_2 - y_2, \ldots, x_d - y_d) dy_1 \cdots dy_d, \]

where the sum in the first one is taken over all \(2^d\) possible combinations of plus and minus signs and \( x = (x_1, x_2, \ldots, x_d) \in H \).

Let \( N = n^d \), where \( n \) is an integer. Introduce the following operators:

\[ T_N(\phi) = \frac{1}{N} \sum_{i_1=0}^{n} \sum_{i_2=0}^{n} \cdots \sum_{i_d=0}^{n} 2^{-w(i_1, i_2, \ldots, i_d)} \phi \left( \frac{i_1}{n}, \frac{i_2}{n}, \ldots, \frac{i_d}{n} \right), \]

where \( w(i_1, i_2, \ldots, i_d) \) = number of \( i \)'s that are zero or \( n \), and

\[ M_N(P) = \frac{1}{N} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_d=1}^{n} P \left( \frac{1}{2n} + \frac{i_1-1}{n}, \frac{1}{2n} + \frac{i_2-1}{n}, \ldots, \frac{1}{2n} + \frac{i_d-1}{n} \right). \]

These two operators are, respectively, the generalizations to \( d \)-dimensions of the trapezoidal and midpoint rules.

We are now ready to write, in a compact manner, the following theorem:

**Theorem 3.** For the general \( d \)-dimensional case, and using the above notation, we have

\[ (2.4.1) \quad \phi^2 = T_N(\phi) - \int_H \phi(x) \, dx, \]
and

\[(2.4.2) \quad \sigma_{MR}^2 = T_N(\phi) - \int_H \phi(x) \, dx - 2 \left[ M_N(\mathbb{P}) - \int_H \mathbb{P}(x) \, dx \right].\]

**Proof.** The proof of Theorem 2 \((d = 1)\) shows that there are only two equalities left to prove, namely

\[\text{Var } \bar{Z} = T_N(\phi),\]

and

\[\text{Var } I(H) = \int_H \phi(x) \, dx.\]

If we use formula (1.2.7) and collect terms we get

\[\text{Var } \bar{Z} = \frac{1}{N} \sum_{i_1 = -n}^{n} \ldots \sum_{i_d = -n}^{n} \left(1 - \frac{|i_1|}{n}\right) \ldots \left(1 - \frac{|i_d|}{n}\right) c\left(\frac{i_1}{n}, \frac{i_2}{n}, \ldots, \frac{i_d}{n}\right).\]

We are now tempted to write

\[\text{Var } \bar{Z} = \frac{1}{N} \sum_{i_1 = 0}^{n} \ldots \sum_{i_d = 0}^{n} (1 - \frac{i_1}{n}) \ldots (1 - \frac{i_d}{n}) \sum_{\pm} \sum_{\pm} \ldots \sum_{\pm} c\left(\pm \frac{i_1}{n}, \pm \frac{i_2}{n}, \ldots, \pm \frac{i_d}{n}\right),\]

but care must be taken, because if we do so, the terms that have \(k\) indices vanishing will be counted \(2^k\) times. Note that all terms vanish when one index is \(n\). The second formula will then be correct if we use the weights

\[-w(i_1, i_2, \ldots, i_d).\]
Looking at the definitions of $\phi(x)$ and $T_N(\phi)$ we realize that we have proved the first equality. The second equality,

$$\text{Var } I(H) = \int_H \phi(x) \, dx,$$

will be proved in a somewhat different fashion. Recall that

$$\text{Var } I(H) = \int_H \int_H c(x-y) \, dx \, dy.$$

We can also write

$$\text{Var } I(H) = \mathbb{E} c(x_1-y_1, x_2-y_2, \ldots, x_d-y_d),$$

where $x_i$ and $y_i$, $i = 1, \ldots, d$ are independent random variables, uniformly distributed on $(0,1)$.

The random variables

$$U_i = x_i - y_i, \quad i = 1, 2, \ldots, d,$$

have a common density function

$$1 - |t|, \quad -1 \leq t \leq 1.$$

Using these random variables we can write
\[ \text{Var } I(H) \]
\[ = E \ c(u_1, u_2, \ldots, u_d) \]
\[ = \int_{-1}^{1} \int_{-1}^{1} \ldots \int_{-1}^{1} c(u_1, u_2, \ldots, u_d) (1-u_1)(1-u_2)\ldots(1-u_d) \, du_1 \ldots du_d \]
\[ = \int_{0}^{1} \ldots \int_{0}^{1} \sum \sum \ldots \sum c(tu_1, tu_2, \ldots, tu_d) (1-u_1)\ldots(1-u_d) \, du_1 \ldots du_d \]
\[ = \int_{H} \phi(x) \, dx \]

which concludes the proof of Theorem 3.

The relationship

\[ (2.4.3) \quad \int_{H} \int_{H} c(x-y) \, dx \, dy \]
\[ = \int_{H} \sum \sum \ldots \sum c(tx_1, tx_2, \ldots, tx_d) (1-x_1)\ldots(1-x_d) \, dx \]

valid for any integrable function \( c \), will be used again in later chapters.
Chapter 3

Rates of Convergence for Smooth Covariance Functions

3.1 Preliminaries

We shall now use the formulas developed in the preceding chapter to study the asymptotic behavior of \( \sigma^2 \) for the different designs. These formulas express \( \sigma^2 \) as the difference between two quantities involving the covariance function of the process: an integral and a finite sum. For some of the designs it will be the case that, as \( N \to \infty \), the finite sum will converge to the value of the integral. How rapidly this convergence takes place is a question we can investigate by methods commonly used in the field of Numerical Analysis.

In this chapter we will assume that the covariance function is smooth. In general we shall say that a function \( c(x) \), \( x \in \mathbb{R}^d \), is smooth at \( x \) if, for all \( k \), all the partial derivatives

\[
\frac{\partial^k c(x)}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_d^{k_d}}, \quad k_1 + k_2 + \cdots + k_d = k,
\]

exist at \( x \). A function will be called smooth over a region if it is smooth at every point in the region. If we omit mentioning a particular region, it will be understood that the function is smooth everywhere in \( \mathbb{R}^d \).

Consider now the special case where \( c(x) \) is the covariance function of a stationary stochastic process. The results of Section 1.2 show that \( c(x) \) is smooth if and only if \( c(x) \) is smooth at the
origin. This will also imply the existence, everywhere in \( \mathbb{R}_d \), of the partial quadratic mean derivatives of the process \( z(\mathbf{x}) \) as defined in Section 1.2.

This regularity imposed on the covariance function will allow us to study the asymptotic behavior of \( \sigma^2 \) in terms of Taylor series and Euler-Maclaurin expansions. The idea of using the ordinary one-dimensional version of the Euler-Maclaurin summation formula in connection with \( \sigma^2_{SY} \) is not new and has been discussed by Williams (1956) and Mathéron (1965). In this chapter we show how its extensions can be applied to study both \( \sigma^2_{SY} \) and \( \sigma^2_{MR} \) in the general \( \mathbb{R} \)-dimensional case.

### 3.2 Random and Stratified Sampling

**Theorem 4.** If the covariance function of the stationary process \( z(\mathbf{x}) \), \( \mathbf{x} \in \mathbb{R}_d \), is smooth, then

\[
(3.2.1) \quad \lim_{N \to \infty} N \sigma^2_{SR} = c(\mathcal{O}) - \int_H \int_H c(\mathbf{x} - \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y},
\]

and

\[
(3.2.2) \quad \lim_{N \to \infty} \frac{1}{d} \frac{\sigma^2_{ST}}{N} = - \frac{1}{6} \sum_{i=1}^d \left. \frac{\partial^2 c}{\partial x_i^2} \right|_{\mathbf{x} = 0}.
\]

The quantities on the righthand sides of (3.2.1) and (3.2.2) vanish if and only if

\[
c(\mathbf{x}) = c(\mathcal{O}), \quad \forall \mathbf{x} \in H.
\]
Proof. (3.2.1) is a trivial consequence of (2.2.1) in Theorem 1. To prove (3.2.2), expand \( c(x) \), \( x \in \mathcal{H}_1 \), in a Taylor series around the origin. We get

\[
c(x_1, x_2, \ldots, x_d) = c(0) + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} x_i x_j \frac{\partial^2 c}{\partial x_i \partial x_j}(x) \bigg|_{x=0} + O\left(\frac{1}{n^{3/2}}\right),
\]

where we have used the fact that the first partials vanish at the origin since \( c(x) = c(-x) \). We then have that

\[
\sigma_{ST}^2 = \frac{1}{N} \left( c(0) - N^{-2} \int_{\mathcal{H}_1} \int_{\mathcal{H}_1} c(u-v) \, du \, dv \right)
\]

\[
= N \left( \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \frac{\partial^2 c}{\partial x_i \partial x_j}(0) \bigg|_{x=0} \int_{\mathcal{H}_1} \int_{\mathcal{H}_1} (u_i - v_i)(u_j - v_j) \, du \, dv \right) + o\left(\frac{1}{N^{1+2/d}}\right)
\]

\[
= \frac{1}{N^{1+2/d}} \cdot \frac{1}{6} \left( -\sum_{i=1}^{d} \frac{\partial^2 c}{\partial x_i^2}(0) \bigg|_{x=0} \right) + o\left(\frac{1}{N^{1+2/d}}\right).
\]

This proves (3.2.2).

Suppose now that the r.h.s. of (3.2.1) vanishes. Note that under the assumptions \( c(x) \), \( x \in \mathcal{H} \), is a continuous function. A straightforward application of the Cauchy-Schwarz inequality shows that

\[
|c(x)| \leq c(0), \quad \forall x.
\]

Under these conditions

\[
c(0) - \int_{\mathcal{H}} \int_{\mathcal{H}} c(x-y) \, dx \, dy
\]
can only vanish if
\[ c(x) = c(0), \quad \forall x \in \mathbb{R}. \]

Let us suppose finally that the r.h.s. of (3.2.2) vanishes. To keep notation simple, we will assume that \( d = 2 \). Note that the function
\[ -\frac{\partial^2 c}{\partial x^2}(x,0), \quad x \in \mathbb{R}, \]
is the covariance function of the process \( \frac{\partial c}{\partial x}(x,0) \). Consequently the vanishing of the r.h.s. of (3.2.2) implies that
\[ \frac{\partial^2 c}{\partial x^2}(0,0) = 0. \]
The Cauchy-Schwarz inequality gives
\[ \left| -\frac{\partial^2 c}{\partial x^2}(x,0) \right| \leq -\frac{\partial^2 c}{\partial x^2}(0,0) = 0. \]
which shows that the second partial vanishes at all points \( (x,0) \).
We must then have
\[ \frac{\partial c}{\partial x}(x,0) = \frac{\partial c}{\partial x}(0,0) = 0, \]
which in turn implies that
\[ c(x,0) = c(0,0), \quad \forall x \in \mathbb{R}. \]
Similarly
\[ c(0,y) = c(0,0), \quad \forall y \in \mathbb{R}_\perp. \]

We can then write
\[
[c(0,0) - c(x,y)]^2 = \left[ \mathbb{E} \left( z(x,y) - z(x,0) \right) \left( z(x,0) - z(0,0) \right) \right]^2
\]
\[ \leq \mathbb{E} \left( z(x,y) - z(x,0) \right)^2 \mathbb{E} \left( z(x,0) - z(0,0) \right)^2 \]
\[ = 4 \left( c(0,0) - c(0,y) \right) \left( c(0,0) - c(x,0) \right) \]
\[ = 0, \]
so that
\[ c(0,0) = c(x,y), \quad \forall \ x, \ y \in \mathbb{R}_\perp. \]

This concludes the proof of Theorem 4.

3.3 Systematic and Midpoint Sampling

Theorem 5. If the covariance function of \( z(x), \ x \in \mathbb{R}_d \), is smooth then for any positive integer \( k \) we have

\[ \sigma_{SY}^2 = \sum_{i=1}^{k} \frac{a_i}{n^{2i}} + O \left( \frac{1}{n^{2k+1}} \right), \tag{3.3.5} \]

and

\[ \sigma_{MR}^2 = \sum_{i=1}^{k} \frac{a_i - 2b_i}{n^{2i}} + O \left( \frac{1}{n^{2k+1}} \right), \tag{3.3.6} \]
where \( a_i, b_i, i = 1, 2, \ldots, k \), are constants independent of \( n \), and \( N = n^d \) is the total number of sample points. Furthermore,

\[(3.3.7) \quad a_1 - 2b_1 = 0.\]

In particular we will have that

\[
\lim_{N \to \infty} N^{2/d} \sigma_{SY}^2 = a_1,
\]

and

\[
\lim_{N \to \infty} N^{4/d} \sigma_{MR}^2 = a_2 - 2b_2.
\]

**Proof.** Recall formulas (2.4.1) and (2.4.2),

\[
\sigma_{SY}^2 = T_N(\phi) - \int_H \phi(x) \, dx,
\]

\[
\sigma_{MR}^2 = \sigma_{SY}^2 - 2[M_N(p) - \int_H p \, dx],
\]

where

\[
\phi(x) = (1-x_1) \cdots (1-x_d) \sum \cdots \sum c(x_1, x_2, \ldots, x_d),
\]

and

\[
P(x) = \int_H c(x-y) \, dy.
\]
Both functions \( \phi \) and \( P \) are smooth and we can now apply the d-dimensional Euler-MacLaurin expansions discussed in Section 1.3. From Formulas (1.3.3) and (1.3.4) we obtain (3.3.5) and (3.3.6). All we have to show now is (3.3.7). From the definitions in Section 1.3, we have that

\[
a_1 = \frac{1}{12} \int_{\mathcal{H}} \sum_{i=1}^{d} \frac{\partial^2 \phi}{\partial x_i^2} \, dx,
\]

and

\[
b_1 = -\frac{1}{12\cdot 2} \int_{\mathcal{H}} \sum_{i=1}^{d} \frac{\partial^2 P}{\partial x_i^2} \, dx,
\]

from which it follows that

\[
(3.3.8) \quad a_1 - 2b_1 = \frac{1}{12} \int_{\mathcal{H}} \left( \sum_{i=1}^{d} \frac{\partial^2 \phi}{\partial x_i^2} + \frac{\partial^2 P}{\partial x_i^2} \right) \, dx.
\]

Using the definitions of \( P \) and \( \phi \) we have that

\[
(3.3.9) \quad \frac{\partial^2 \phi}{\partial x_1^2} = (1-x_2) \ldots (1-x_d)
\]

\[
\times \left[ (1-x_1) \sum_{\pm} \frac{\partial^2 \phi}{\partial x_1^2} (\pm x_1, x_2, \ldots, x_d) - 2 \sum_{\pm} \frac{\partial \phi}{\partial x_1} (\pm x_1, \ldots, \pm x_d) \right],
\]

and, differentiating under the integral sign,

\[
(3.3.10) \quad \int_{\mathcal{H}} \frac{\partial P}{\partial x_1^2} \, dx = \int_{\mathcal{H}} \left[ \frac{\partial^2 c(x-y)}{\partial x_1^2} \right] \, dx \, dy.
\]
Recall Formula (2.4.3), i.e., for any integrable function \( f \)

\[
\int_{H} \int_{H} f(x-y)dx \, dy = \int_{H} (1-x_1)(1-x_2)\ldots(1-x_d) \sum f(x_1, x_2, \ldots, x_d)dx .
\]

If we apply Formula (2.4.3) to \( f(x) = \frac{\partial^2}{\partial x_1^2} c(x) \) and combine (3.3.9) and (3.3.10), we get

\[
(3.3.11) \quad \int_{H} \left[ \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right] dx = 2 \left[ \int_{H} \int_{H} \frac{\partial^2}{\partial x_1^2} c(x-y)dx \, dy - \int_{H} (1-x_2)\ldots(1-x_d) \sum \frac{\partial^2}{\partial x_1^2} (\pm x_1, \pm x_2, \ldots, \pm x_d)dx \right] .
\]

Let \( H^1 \) be the \( d-1 \) dimensional hypercube. Also let \( x^1 = (x_2, x_3, \ldots, x_d) \) and \( y^1 = (y_2, y_3, \ldots, y_d) \). Then

\[
\int_{H} \int_{H} \frac{\partial^2}{\partial x_1^2} c(x-y)dx \, dy = \int_{H^1} \int_{H^1} \int_{H^1} \frac{\partial^2}{\partial x_1^2} c(x-y^1)dx_1 \, dy_1 \, dx_1 \, dy_1 .
\]

Applying Formula (2.4.3) to the \((d-1)\times(d-1)\)-fold inner integral we obtain

\[
\int_{H} \int_{H} \frac{\partial^2}{\partial x_1^2} c(x-y)dx \, dy = \int_{0}^{1} \int_{0}^{1} (1-x_2)\ldots(1-x_d) \sum \frac{\partial^2}{\partial x_1^2} c(x_1-y_1, x_2, \ldots, x_d)dx_1 \, dy_1 \, dx_1 \, dy_1 = \int_{H} (1-x_2)\ldots(1-x_d) \sum \frac{\partial}{\partial x_1} c(\pm x_1, \pm x_2, \ldots, \pm x_d)dx ,
\]

39
where the last equality is obtained by integrating out \( y_1 \). This shows, from (3.3.11), that

\[
\int_H \left( \frac{\partial^2 \Phi}{\partial x_1^2} + \frac{\partial^2 P}{\partial x_1^2} \right) \, dx = 0 .
\]

By an analogous argument,

\[
\int_H \left( \frac{\partial^2 \Phi}{\partial x_i^2} + \frac{\partial^2 P}{\partial x_i^2} \right) \, dx = 0 , \quad i = 2, 3, \ldots, d .
\]

Expression (3.3.8) now becomes

\[
a_1 - 2b_1 = 0 ,
\]

which is (3.3.7). Theorem 5 has been proved.

**Remark 1.** The conditions under which Theorem 5 is true may be relaxed somewhat. Suppose that the covariance function \( c(x_1, x_2, \ldots, x_d) \) is such that it is smooth everywhere except maybe at the points on the surfaces

\[
x_i = 0 , \quad i = 1, 2, \ldots, d ,
\]

where the right- and left-hand partials with respect to \( x_i \) exist but might not be equal to each other. For such a covariance function, \( \Phi(x) \) and \( P(x) \) will still be smooth functions on \( H \) if, at the boundaries, we only require the existence of the appropriate one-sided derivatives. This is, however, all that is needed to apply the
Euler-Maclaurin expansions used in the proof of Theorem 5, since the applicability of such expansions does not depend on how the functions are defined outside of \( H \). We conclude then that (3.3.5) and (3.3.6) are still valid in this case. We remark, however, that expression (3.3.7) is no longer necessarily true.

An important example where Remark 1 is relevant is the so-called two-state stationary Markov chain in continuous time. This process, defined on the real positive axis, has been used to describe the behavior of a device that is either on or off corresponding to, say, \( z(t) = 1 \) or \( z(t) = 0 \). In this context

\[
\int_{0}^{1} z(t) \, dt
\]

represents the total usage of the device in the time interval \([0,1]\).

The covariance function of \( z(t) \) has the form

\[
c(x) = k e^{-\lambda |x|}, \quad |x| < \infty,
\]

for some positive constants \( k \) and \( \lambda \). Clearly, it is not smooth at the origin, but by Remark 1 the expansions (3.3.5) and (3.3.6) of Theorem 5 still hold true.

3.4 The Use of Mean Square Derivatives in Obtaining Rates of Convergence

This section detracts somewhat from the main stream of this thesis. Its purpose is to present an alternate way of obtaining the expansion for \( c_{MR}^2 \) for smooth covariance functions and to show how this expansion
may be used to obtain estimators whose M.S.E. converges to zero faster than any given power of \( N \).

For simplicity we will restrict ourselves to the one-dimensional case. In Section 1.3 we have seen that the Euler-MacLaurin summation formula holds true in terms of stochastic integrals and derivatives, i.e.

\[
M_N(z) - \int_0^1 z(x) \, dx = \sum_{i=1}^{k} \frac{B_{2i}}{(2i)!} \cdot \frac{2 - 2^i}{2^{2i}} \cdot \frac{1}{N^{2i}} \left[ z(2i-1)(1) - z(2i-1)(0) \right] + R_k,
\]

where \( z(t) \) is a stationary process whose \( 2k + 1 \) q.m. derivative exists.

We can now calculate \( \sigma^2_{MR} \) by squaring both sides of (3.4.1) and taking expectations. Using the fact that the covariance function of \( z^{(k)}(x) \) is

\[
(-1)^k \frac{d^{2k}c(x)}{dx}, \quad |x| < \infty
\]

we can easily check that we obtain the same result as before, namely (3.3.6). Expansion (3.4.1) has, however, other uses. It shows clearly that we can improve our estimators of

\[
\int_0^1 z(x) \, dx
\]
by the use of what numerical analysts call "extrapolation to the limit". We illustrate this process in its simplest form.

Suppose we have the two midpoint estimators $M_N(z)$ and $M_{2N}(z)$ based, respectively, on $N$ and $2N$ sample points. According to (3.4.1) we have

$$M_N(z) - \int_0^1 z \, dx = \sum_{i=1}^{\infty} \frac{W_i}{N^{2i}},$$

$$M_{2N}(z) - \int_0^1 z \, dx = \sum_{i=1}^{\infty} \frac{W_i}{2^{2i} N^{2i}},$$

where the $W_i$'s depend on the stochastic derivatives but are the same in both expansions.

If we form the linear combination

$$z^* = \frac{1}{3} M_{2N}(z) - \frac{1}{3} M_N(z),$$

we will then have

$$(3.4.2) \quad z^* - \int_0^1 z \, dx = \sum_{i=2}^{\infty} \frac{W_i}{N^{2i}},$$

i.e., $z^*$ has a M.S.E. of order $1/N^8$ since the first term in the r.h.s. of (3.4.2) is of the order $1/N^4$. The midpoint estimator, based on the same number of sample points as $z^*$, is $M_{3N}(z)$. It has a mean square error of order $1/(3N)^4$ and if $N$ is large, $z^*$ will be a big improvement over it.
This process can be continued; for any \( k \) we can form the linear combination

\[
z_k^* = \sum_{i=1}^{k} a_i M_{2^{i-1}N}(z),
\]

selecting the coefficients \( a_i \) such that they satisfy

\[
\sum a_i = 1,
\]

and

\[
\sum_{i=1}^{k} \frac{a_i}{2^{(i-1)2j}} = 0, \quad j = 1, 2, \ldots, k-1.
\]

We will then have that

\[
z_k^* - \int_{0}^{1} z(x) \, dx = O\left( \frac{1}{N^{2k}} \right),
\]

which shows that if the covariance function is smooth we can construct estimators \( z_k^* \), for any \( k \), such that their M.S.E. is

\[
O\left( \frac{1}{N^{4k}} \right).
\]
Chapter 4

Rates of Convergence for a Class of Isotropic Covariance Functions

4.1 Preliminaries

In this chapter we shall study the asymptotic behavior of $\sigma^2$ under the assumption that the covariance function of the process is of the form

$$c(x) = h(r), \quad r = \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2},$$

where we assume that

$$\frac{d^3h}{dr^3}(r)$$

exists for $r > 0$ (at $r = 0$ derivatives mean righthand derivatives).

For reasons discussed in Section 4.3 we are mainly interested in the case

$$h'(0) < 0.$$

(It is clear that, necessarily, $h'(0) \leq 0$ since $|c(x)| \leq c(0)$, $\forall x \in \mathbb{R}_d$).

An example of this type of covariance function has been discussed in Section 1.2, namely

$$c(x) = k e^{-ar}.$$
4.2 Stratified and Random Sampling

Theorem 6. If the covariance function of $z(x)$, $x \in \mathbb{R}^d$, has the form

$$c(x_1, x_2, \ldots, x_d) = h(r), \quad r = \sqrt{x_1^2 + \cdots + x_d^2},$$

where $h$ has a continuous second derivative, then

$$\lim_{N \to \infty} N \sigma_{SR}^2 = h(0) - \int_H \int_H h(|x-y|) \, dx \, dy,$$

and

$$\lim_{N \to \infty} N^{1+1/d} \sigma_{ST}^2 = -h'(0) \int_H \int_H |x-y| \, dx \, dy,$$

where $N$ is the total number of sample points and $|x|$ denotes the Euclidean norm of the vector $x$.

Proof. Equality (4.2.1) is included here just for completeness; it is clear that under any reasonable assumption the rate of convergence of $\sigma_{SR}^2$ is always the same, namely $1/N$.

To prove (4.2.2), recall the formula

$$\sigma_{SR}^2 = \frac{1}{N} \left( c(Q) - N^2 \int_{H_1} \int_{H_1} c(x-y) \, dx \, dy \right),$$

where

$$H_1 = \left\{ (x_1, x_2, \ldots, x_d) \mid 0 \leq x_i \leq \frac{1}{n}, \ i = 1, 2, \ldots, d, \ n^d = N \right\}.$$
Using the Taylor expansion
\[ h(r) = h(0) + rh'(0) + \frac{r^2}{2} h''(0), \quad 0 \leq \theta \leq r, \]
we obtain
\[ \sigma_{SR}^2 = N \left( -h'(0) \int_{H} \int_{H} |x-y| \, dx \, dy + \int_{H} \int_{H} |x-y|^2 \, h''(0) \, dx \, dy \right). \]

Rescaling the integrals by setting \( x_i' = nx_i \) and \( y_i' = ny_i \), and using the fact that the second derivative of \( h \) is bounded in every finite interval, we get
\[ \sigma_{SR}^2 = -h'(0) \frac{N}{n^{2d+1}} \int_{H} \int_{H} |x-y| \, dx \, dy + o \left( \frac{N}{n^{2d+2}} \right), \]
and the proof of Theorem 6 is completed.

4.3 A Geometrical Interpretation of \( h'(0) \) for an Isotropic Set Process

In our study of rates of convergence for isotropic processes the quantity \( h'(0) \) plays an important role. For an isotropic set process this quantity has a rather tangible geometrical meaning. Matérn (1960), based on geometrical considerations, has conjectured that if \( \delta B \) is the boundary of the random set and \( S(\delta B) \) the surface content of \( \delta B \) per unit volume, then

\[ (4.3.1) \quad ES(\delta B) = -2h'(0)\sqrt{\pi} \Gamma \left( \frac{d+1}{2} \right) / \Gamma \left( \frac{d}{2} \right), \]
where $\Gamma$ is the ordinary gamma function and $d$ is the dimension of the Euclidean space where the process is defined. Matheron (1967) has given a proof of (4.3.1) for the special case of random sets whose boundaries have a tangent hyperplane at every point.

In this section we will point out the connection that exists between the two-dimensional version of (4.3.1) and some formulas in integral geometry. We will then use these results to get a set of conditions under which (4.3.1) is true.

The connection with integral geometry will be established by considering an oriented segment $\ell$ whose position and orientation in the plane is random. Randomness is defined in terms of the kinematic measure $K$ of integral geometry (Santaló (1953, p. 21)). Let $1/n$ be the length of the segment and let $a$ and $b$ denote its endpoints. Define

$$M = \text{number of intersections of } \ell \text{ and } \delta B$$

where $\delta B$ denotes that part of the boundary of the random set which is contained in the unit square $H$. Let the underlying isotropic set process $z(x)$ be defined by the pair $(\Omega,\mathcal{P})$ where $\mathcal{P}$ is a probability measure on the space $\Omega$. By Fubini's theorem

$$\int_{\mathcal{A} \times \Omega} z(a)(1-z(b))d\mathcal{P} \times dK = \int_{\Omega} \int_{\mathcal{A}} z(a)(1-z(b))dKd\mathcal{P}$$

$$= \int_{\mathcal{A}} \int_{\Omega} z(a)(1-z(b))d\mathcal{P}dK,$$

where $\mathcal{P} \times K$ represents the product measure of $\mathcal{P}$ and $K$ and $\mathcal{A}$ is the set of all those positions of the segment where it intersects the unit square.

48
Because of the isotropy of the zero-one set process \( z(x) \), we have that

\[
(4.3.2) \quad \int_A \int_\Omega z(a)(1-z(b)) \, dP \, dK = \int_A \int_\Omega (z^2(a) - z(a)z(b)) \, dP \, dK
\]

\[
= \int_A (h(0) - h(\frac{1}{n})) \, dK
\]

\[
= [h(0) - h(\frac{1}{n})] K(A) ,
\]

where \( h \) is the covariance function of the process. On the other hand,

\[
(4.3.3) \quad \int_\Omega \int_A z(a)(1-z(b)) \, dK \, dP = \frac{1}{2} \int_\Omega \int_A [z(a)(1-z(b)) + z(b)(1-z(a))] \, dK \, dP
\]

\[
= \frac{1}{2} \int_\Omega K(M=\text{o dd}) \, dP ,
\]

because the integrand of the second integral has the value one if the segment intersects \( \delta B \) at an odd number of points and is zero otherwise.

The kinematic measure of the set \( A \) is given by (Santaló (1953, p. 26))

\[
K(A) = 2(\pi + \frac{h}{n}) .
\]

Equating (4.3.2) and (4.3.3) we now get

\[
(4.3.4) \quad h(0) - h(\frac{1}{n}) = \frac{1}{4(\pi + \frac{1}{n})} \int_\Omega K(M=\text{o dd}) \, dP .
\]

We are now going to show that if, with probability one, \( \delta B \) is rectifiable and
\[(4.3.5) \quad \lim_{n \to \infty} n K(\{M=1\}) = \lim_{n \to \infty} \int M \, dK, \]

then \((4.3.1)\) is true.

It is shown in integral geometry (Santaló (1953, p. 31)) that if \(\delta B\) and \(\ell\) are any two rectifiable curves, then

\[(4.3.6) \quad \int M \, dK = 4S(\delta B)S(\ell), \]

where \(S(\delta B)\) and \(S(\ell)\) are the lengths of \(\delta B\) and \(\ell\). Formula \((4.3.6)\) is known as Poincaré's Formula. Let the symbol \(E\) stand for expectations taken over the measure \(P\). Assume first that \(E S(\delta B) < \infty\). We have, from \((4.3.6)\), that

\[n K(\{M=\text{odd}\}) \leq n \int M \, dK = 4S(\delta B).\]

Then, from \((4.3.4)\) and Lebesgue's dominated convergence theorem,

\[
\lim_{n \to \infty} n(h(0) - h(\frac{1}{n})) = \frac{1}{4\pi} \int_\Omega \lim_{n \to \infty} n K(\{M=\text{odd}\}) \, dP.
\]

Clearly, under \((4.3.5)\),

\[
\lim_{n \to \infty} n K(\{M=\text{odd}\}) = 4S(\delta B).
\]

We finally obtain

\[h'(0) = \frac{E S(\delta B)}{\pi},\]

which is the two-dimensional version of \((4.3.1)\). Assume now that \(E S(\delta B) = \infty\); then, by Fatou's lemma,
\[
\lim_{n \to \infty} n(h(0) - h(\frac{1}{n})) \geq \frac{1}{16\pi} \mathbb{E} S(\delta B),
\]

which implies that \( h'(0) = \infty \). We see then that equation (4.3.1) is again true, provided that \( \infty \) is allowed as a value for either side of the equation.

As an application we will now check that condition (4.3.5) is satisfied in the case of set processes whose boundaries are, with probability one, the union of a finite number of polygonal lines with a finite number of sides. An example of this type of process has been discussed in Section 1.2. For \( n \) big enough, the segment can only intersect \( \delta B \) in either one or two points. The kinematic measure of the positions where a segment of length \( 1/n \) intersects both sides of an angle \( \omega \) is (Santaló (1953, p. 28))

\[
\frac{1}{2n^2} (1 + (\pi - \omega) \cot \omega) = 0 \left( \frac{1}{n^2} \right).
\]

We then have

\[
\lim_{n \to \infty} n \int M \, dK = \lim_{n \to \infty} \left[ n K(\{M=1\}) + 2n K(\{M=2\}) \right]
\]

\[
= \lim_{n \to \infty} \left[ n K(\{M=1\}) \right],
\]

which is condition (4.3.5).
4.4 Systematic and Midpoint Sampling

The main purpose of this section is to state and prove Theorem 7, which gives the rates of convergence of $\sigma_{SY}^2$ and $\sigma_{MR}^2$ for the class of isotropic covariance functions

$$c(x,y) = h\left(\sqrt{x^2 + y^2}\right),$$

where $h$ has a continuous third derivative. The theorem is stated for the particular case $d = 2$, i.e., when the sampling is done in the unit square. Extensions to general $d$ are not straightforward and have not been attempted.

The proof of Theorem 7 is complicated by the fact that the function $c(x,y)$ is, in general, highly irregular at the origin. We have seen in Section 4.3 that for a certain class of isotropic set processes,

$$h'(0) = -\frac{\pi}{2} E\delta B$$

where $E\delta B$ is the mean boundary length of the random set per unit area. In particular $h'(0)$ will be different from zero for any non-trivial process in this class and the higher order partials of $c$ will not even be bounded in the neighborhood of the origin. The second partial,

$$\frac{\partial^2 c}{\partial x^2} = h'(r) \left[ \frac{1}{r} - \frac{x^2}{r^3} \right] + h''(r) \frac{x^2}{r^2},$$

already exhibits this behavior since at the origin it is $0 \left(\frac{1}{r}\right)$.  

52
Having to deal with functions of this type, it will be convenient to have the following terminology:

Let $S$ be a region in $\mathbb{R}^2$ and let $u(x,y)$ be defined on $S$. The function $u$ is said to have properties $A_k$, $A'_k$ or $B_k$ over $S$ if

$a)$ $u$ is continuous in $S$,

$b)$ the partial derivatives of $u$ up to the $k$-th order exist and are continuous in $S$.

$A'_k$: same as $A_k$ except that $b)$ is known to be satisfied only in the region

$$S \cap \{(0,0)\}^c,$$

where $c$ denotes the set complement operation.

$B_k$: $u$ has property $A'_k$ over $S$.

$b)$ $u$ and its partial derivatives up to order $k$ are absolutely integrable over $S$.

We shall mainly use this notation for $k = 1, 2$ or 3.

As an example and for future reference we point out here that the function $r$ over $H$

- does not have property $A_k$ for any $k$,
- has property $A'_k$ for $k = 1, 2, \ldots$,
- has property $B_k$ only for $k = 1$ and 2.
Its first few partial derivatives are:

\[
\frac{\partial r}{\partial x} = \frac{x}{r},
\]

\[
\frac{\partial^2 r}{\partial x^2} = \frac{1}{r} - \frac{x^2}{r^3},
\]

\[
\frac{\partial^3 r}{\partial x^3} = -3 \left( \frac{x}{r^3} - \frac{x^3}{r^5} \right),
\]

\[
\frac{\partial^3 r}{\partial x \partial y^2} = -\frac{x}{r^3} + \frac{3xy^2}{r^5}.
\]

Our main result in this section is:

**Theorem 7.** Suppose we sample an isotropic stochastic process on the unit square by means of an \( n \times n = N \) grid. Assume further that the covariance function of the process is of the form

\[
c(x,y) = h(r), \quad r = \sqrt{x^2 + y^2},
\]

with \( h \) having a continuous third derivative. Then

\[
(4.4.1) \quad \sigma_{SY}^2 = \frac{a}{n^2} + \frac{4a \gamma}{n^3} + o \left( \frac{1}{n^3} \right),
\]

and

\[
(4.4.2) \quad \sigma_{MR}^2 = \frac{4a \gamma}{n^3} + o \left( \frac{1}{n^3} \right),
\]

where
\[ \alpha = h'(0) , \]
\[ a = \frac{1}{3} \int_H \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \left( (1-x)(1-y)c(x,y) \right) \, dx \, dy , \]

and \( \gamma \) is a constant whose value does not depend on the particular shape of \( c \).

**Proof.** The proof is long and complicated and will be broken up in a sequence of lemmas. We had previously obtained the general expressions (Formulas (2.4.1) and (2.4.2)):

\[ \sigma_{SY}^2 = T_N(\phi) - \int_H \phi \, dx , \]

and

\[ \sigma_{MR}^2 = \sigma_{SY}^2 - 2 \left[ M_N(P) - \int_H P \, dx \right] , \]

where

\[ (4.4.3) \quad \phi(x,y) = (1-x)(1-y)[c(x,y) + c(-x,y) + c(x,-y) + c(-x,-y)] , \]

and

\[ (4.4.4) \quad P(x,y) = \int_H c(x-u,y-v)du \, dv . \]

We shall first prove (4.4.1). The strategy of the proof is as follows:
Using Lemma 1 we show that we can write

\[(4.4.5) \quad \phi(x,y) = 4ar + f(x,y)\]

where \(f\) is a function having property \(B_3\) over \(H\). We will then have that

\[(4.4.6) \quad \sigma^2_{\text{SY}} = 4a \left[ T_N(r) - \int_H r \, dx \right] + \left[ T_N(f) - \int_H f \, dx \right].\]

Using Lemma 2 we show next that \(\sigma^2_{\text{SY}}\) has a limited Euler-MacLaurin expansion, i.e.

\[\sigma^2_{\text{SY}} = \frac{a}{n^2} + \frac{b}{n^3}.\]

We then finally show, using Lemmas 3 and 4, that the sequence \(b_n\) converges to a constant \(\gamma\) whose value does not depend on the particular shape of \(c\).

\textbf{Lemma 1.} If \(u(x), \ x > 0\), has a continuous third derivative for \(x \in [0,1]\), the function

\[v(r) = u\left(\sqrt{2x + y^2}\right) - u'(0)\sqrt{2x + y^2}\]

has property \(B_3\) over the unit square \(H\).

\textbf{Proof.} We may use the chain rule to calculate the partial derivatives of \(v(r)\) with respect to \(x\) and \(y\). Since \(r\) is \(B_2\) but not \(B_3\) over \(H\), the only thing we have to prove is that the third
partials of \( v \) are integrable. We will prove this if we show that

\[
v'(r) \frac{\partial^3 r}{S_1^2 S_2^2}, \quad S_1 + S_2 = 3,
\]

is integrable on \( H \).

We have, for \( x \in [0,1] \),

\[
\left| \frac{v'(x)}{x} \right| = \left| \frac{u'(x) - u'(0)}{x} \right| = u''(0) \leq A,
\]

where \( 0 \in [0,1] \) and \( A \) is a bound for the second derivative of \( u \) on \( [0,1] \). This shows that

\[
v'(x) = O(x),
\]

or

\[
v'(r) = O(r) \quad \text{on} \quad H.
\]

Consequently a term of the form

\[
v'(r) \frac{\partial^3 r}{S_1^2 S_2^2}, \quad S_1 + S_2 = 3,
\]

is \( O \left( \frac{1}{r} \right) \), because the third partials of \( r \) are \( O \left( \frac{1}{r^2} \right) \). Since the function \( \frac{1}{r} \) is integrable over \( H \), the proof of Lemma 1 is finished.

The covariance function can now be written as

\[
h(r) = ar + g(r),
\]

57
with $g$ having property $B_3$ over $H$. This immediately proves (4.4.5) and (4.4.6). We now need

**Lemma 2.** If $\mu(x,y)$ has property $A_3$ over

$$H_{ij} = \{(x,y) \mid \frac{i}{n} \leq x \leq \frac{i+1}{n}, \frac{j}{n} \leq y \leq \frac{j+1}{n}\}$$

for $i$ and $j$ integers, then

$$T_{H_{ij}}^2(\mu) - \int_{H_{ij}} \mu \, dx = \frac{1}{n^2} \int_{H_{ij}} \left(\frac{\partial^2 \mu}{\partial x^2} + \frac{\partial^2 \mu}{\partial y^2}\right) \, dx \, dy + \frac{1}{n^3} \int_{H_{ij}} \mu_n(x) \, dx,$$

where

$$T_{H_{ij}}^2(\mu) = \frac{1}{n^2} \times \text{average of } \mu \text{ on the 4 vertices of } H_{ij}$$

and

$$\mu_n(x) = \left[P_3(ny) \frac{\partial^3 \mu}{\partial y^3} + P_3(nx) \frac{\partial^3 \mu}{\partial x^3} + P_1(nx)[P_2(0) - P_2(ny)] \frac{\partial^3 \mu}{\partial y^2 \partial x}\right],$$

the functions $P_k(x)$, $k = 1, 2, 3$, being the first three Bernoulli functions defined in Section 1.3.

For $i = j = 0$ expansion (4.4.7) is also valid in either one of the following two cases:

a) $\mu$ has property $B_3$ over $H$,  

58
b) \( \mu \) has properties \( A' \) and \( B_2 \) over \( H \) and its third partials are bounded in absolute value by the function

\[
\frac{A}{r^2}
\]

for some constant \( A \).

**Proof.** Using the properties of the periodic functions \( P_k \), see Section 1.4, and the formula for differentiating a product, we get

\[
\frac{\partial^2 P_1(nx)P_1(ny)}{\partial x \partial y} = n^2 \mu + nP_1(ny) \frac{\partial \mu}{\partial y} + nP_1(nx) \frac{\partial \mu}{\partial x} + P_1(nx)P_1(ny) \frac{\partial^2 \mu}{\partial x \partial y},
\]

\[
\frac{\partial}{\partial y} \left( P_2(ny) \frac{\partial \mu}{\partial y} \right) = nP_1(ny) \frac{\partial \mu}{\partial y} + P_2(ny) \frac{\partial^2 \mu}{\partial y^2},
\]

\[
\frac{\partial}{\partial y} \left( P_3(ny) \frac{\partial^2 \mu}{\partial y^2} \right) = nP_2(ny) \frac{\partial^2 \mu}{\partial y^2} + P_3(ny) \frac{\partial^3 \mu}{\partial y^3},
\]

\[
P_1(nx) \frac{\partial}{\partial y} \left( P_2(ny) \frac{\partial^2 \mu}{\partial x \partial y} \right) = nP_1(nx)P_1(ny) \frac{\partial^2 \mu}{\partial x \partial y} + P_1(nx)P_2(ny) \frac{\partial^3 \mu}{\partial y \partial x}.
\]

Combining these equalities we obtain

\[
(4.4.9) \quad \frac{1}{n^2} \frac{\partial^2 P_1(nx)P_1(ny)\mu}{\partial x \partial y} - \mu = \frac{1}{n^2} \left[ \frac{\partial}{\partial y} \left( P_2(ny) \frac{\partial \mu}{\partial y} \right) + \frac{\partial}{\partial x} \left( P_2(nx) \frac{\partial \mu}{\partial x} \right) \right]
\]

\[
+ \frac{1}{n^3} \left[ P_3(ny) \frac{\partial^3 \mu}{\partial y^3} + P_3(nx) \frac{\partial^3 \mu}{\partial x^3} \right]
\]

\[
- \frac{1}{n^3} \left[ \frac{\partial}{\partial y} \left( P_3(ny) \frac{\partial^2 \mu}{\partial y^2} \right) + \frac{\partial}{\partial x} \left( P_3(nx) \frac{\partial^2 \mu}{\partial x^2} \right) \right]
\]

\[
+ \frac{1}{n^3} \left[ P_1(nx) \frac{\partial}{\partial y} \left( P_2(ny) \frac{\partial^2 \mu}{\partial x \partial y} \right) - P_1(nx)P_2(ny) \frac{\partial^3 \mu}{\partial y \partial x} \right].
\]
Integrate \((4.4.9)\) over \(H_{ij}\). Some of the resulting integrals may be simplified if we recall that for every integer \(i\)

\[
\lim_{x \to i} P_1(x) = -\lim_{x \to i} P_1(x) = \frac{1}{2}
\]

and

\[P_2(i) = \frac{B_2}{2},\]

\[P_3(i) = 0 .\]

We can then show

\[
\int_{H_{ij}} \frac{\partial}{\partial y} \left( P_2(ny) \frac{\partial u}{\partial y} \right) dy = \int_{\frac{i}{n}}^{i+1} \left[ P_2(j+1) \frac{\partial u}{\partial y} (x, \frac{j+1}{n}) - P_2(j) \frac{\partial u}{\partial y} (x, \frac{j}{n}) \right] dx
\]

\[= \frac{B_2}{2} \int_{H_{ij}} \frac{\partial^2 u}{\partial y^2} dx dy ,\]

\[
\int_{H_{ij}} \frac{\partial}{\partial y} \left( P_3(ny) \frac{\partial^2 u}{\partial y^2} \right) dx dy = 0 ,
\]

and

\[
\int_{H_{ij}} \frac{\partial}{\partial y} \left( P_2(ny) \frac{\partial^2 u}{\partial x \partial y} \right) dx dy = \frac{B_2}{2} \int_{H_{ij}} \frac{\partial^3 u}{\partial x^2 \partial y} dx dy .
\]

These equalities and the similar ones for the variable \(x\) lead directly to \((4.4.7)\) and the first part of Lemma 2 is proved. Let us now prove the second part. Note that under either a) or b) the function
\[ \mu(x + \frac{\varepsilon}{n}, y + \frac{\varepsilon}{n}) \]

has property \( A_3 \) over \( H_{00} \) for any small positive \( \varepsilon \). If we use (4.4.7) and make a change of variables, we obtain

\[
T_{H_\varepsilon}^\mu (\mu) - \int_{H_\varepsilon} \mu(x) dx = \frac{1}{n^2} \frac{B_2}{2} \int_{H_\varepsilon} \left( \frac{\partial^2 \mu}{\partial x^2} + \frac{\partial^2 \mu}{\partial y^2} \right) dx \, dy + \frac{1}{n^3} \int_{H_\varepsilon} \mu_n^{(\varepsilon)}(x) dx,
\]

where

\[
H_\varepsilon = \left\{ (x,y) \mid \frac{\varepsilon}{n} < x < \frac{1+\varepsilon}{n}, \frac{\varepsilon}{n} < y < \frac{1+\varepsilon}{n} \right\}
\]

and

\[
\mu_n^{(\varepsilon)}(x) = P_3(nx-\varepsilon) \frac{\partial^3 \mu}{\partial y^3} + P_3(nx-\varepsilon) \frac{\partial^3 \mu}{\partial x^3} + P_1(nx-\varepsilon)[P_2(0) - P_2(ny-\varepsilon)] \frac{\partial^3 \mu}{\partial y^2 \partial x}.
\]

Let \( \varepsilon \to 0 \). Under either a) or b) \( \mu \) is continuous, and together with its second partials it is absolutely integrable over \( H \). This shows that

\[
T_{H_\varepsilon}^\mu (\mu) + T_{H_{00}}^\mu (\mu), \quad \int_{H_\varepsilon} \mu \, dx \to \int_{H_{00}} \mu \, dx,
\]

\[
\int_{H_\varepsilon} \left( \frac{\partial^2 \mu}{\partial x^2} + \frac{\partial^2 \mu}{\partial y^2} \right) dx \to \int_{H_{00}} \left( \frac{\partial^2 \mu}{\partial x^2} + \frac{\partial^2 \mu}{\partial y^2} \right) dx,
\]

as \( \varepsilon \to 0 \). We will finish the proof of the second part of Lemma 2 if
we can show that

\[(4.4.10) \quad \int_{H_\varepsilon} \nu_n^{(c)} dx + \int_{H_{00}} \nu_n dx \quad \text{as} \quad \varepsilon \to 0.\]

Recall that the Bernoulli functions \( P_k(x) \) are periodic with period one and that for \( x \in [0,1] \)

\[P_1(x) = x - \frac{1}{2},\]
\[P_2(x) = \frac{x^2}{2} - x + \frac{1}{12},\]
\[P_3(x) = \frac{x^3}{6} - \frac{x^2}{4} + \frac{x}{12}.\]

We then have, for some constant \( M \), that

\[(4.4.11) \quad |P_3(x)| \leq M|x| \leq Mr,\]
\[(4.4.12) \quad |P_2(0) - P_2(x)| \leq M|x| \leq Mr,\]

for every \( x \). Let \( I_{H_\varepsilon} \) be the characteristic function of the set \( H_\varepsilon \). Inequalities \( (4.4.11) \) and \( (4.4.12) \) give

\[(4.4.13) \quad |I_{H_\varepsilon} \nu_n^{(c)}(x)| \leq Mr \left( \frac{3^3 u}{3^3 y^3} + \frac{3^3 u}{3^3 y^3} + \frac{3^3 u}{3^3 y^3} \right), \quad x \in H.\]

Under either a) or b) the righthand side of \( (4.4.13) \) is an
integrable function. Expression (4.4.10) is now obtained by the use of Lebesque's Dominated Convergence Theorem and Lemma 2 is fully proved.

Let us return to the proof of Theorem 7. Recall expression (4.4.6)

$$
\sigma^2_{SY} = 2a \left[ T_N(r) - \int_H r \, dx \right] + \left[ T_N(f) - \int_H f \, dx \right].
$$

Note that we can substitute $u$ by $f$ or $r$ in (4.4.7) and (4.4.8) for $i = 0, 1, 2, \ldots, n-1$ and $j = 0, 1, 2, \ldots, n-1$, since in every case the assumptions of Lemma 2 are satisfied. If we add these expressions we get

$$(4.4.14) \quad \sigma^2_{SY} = \frac{a}{n^2} + \frac{b}{n^3}.$$ 

where

$$a = \int_H \left( \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} \right) \, dx \, dy,$$

$$\phi = 4ar + f$$

and

$$(4.4.15) \quad b_n = \int_H r_n(x) \, dx + \int_H f_n(x) \, dx.$$ 

Here $r_n(x)$ and $f_n(x)$ are the remainders associated with $r$ and $f$ as given by (4.4.8).
To show that $b_n$ converges we need the following two Lemmas:

**Lemma 3.** Let $P(x,y)$ and $u(x,y)$ be two functions such that:

i) $P(x+m, y+n) = P(x,y) \forall m, n \text{ integers, i.e., } P$ is doubly periodic with both periods equal to 1,

ii) \[ \int_H P(x) dx = 0 \text{ where } H \text{ is the unit square,} \]

iii) \[ \int_H \frac{P(x)}{r^2} dx \text{ exists, } r = \sqrt{x^2 + y^2}, \]

iv) $u(r \cos \theta, r \sin \theta) = s(\theta)$ for any $r > 0$ with $s(\theta)$ having a bounded derivative in $0 \leq \theta \leq \frac{\pi}{2}$.

Define

\[ (4.4.16) \quad b_n^* = \int_0^1 \int_0^1 \frac{P(nx, ny)u(x,y)}{r^2} dx \, dy = \int_0^1 \int_0^1 \frac{P(x,y)}{r^2} u(x,y) dx \, dy. \]

Then

\[ \lim_{n \to \infty} b_n^* \]

exists.

**Proof.** We will need the following elementary fact (mean value theorem for multiple integrals): Assume that $f$ and $g$ are integrable functions in a bounded set $S \subset \mathbb{R}^d$, suppose also that $f$ is continuous in $S$ and that $g(x) \geq 0 \forall x \in S$; then

\[ \int_S fg \, dx = f(x_0) \int_S g \, dx \]

64
where \( x_0 \in S \). We have from (4.4.16) that if \( m > n \)

\[
|b^*_m - b^*_n| \leq A_n + B_n,
\]

where

\[
A_n = \left| \int_0^\infty \int_0^\infty \frac{P(x,y)}{r^2} \mu \, dx \right|,
\]

and

\[
B_n = \left| \int_n^\infty \int_0^\infty \frac{P(x,y)}{r^2} \mu \, dx \right|.
\]

In order to prove that \( b^*_n \) converges, it is clearly enough to show that

\[
(4.4.17) \quad \lim_{n \to \infty} A_n = 0,
\]

and

\[
(4.4.18) \quad \lim_{n \to \infty} B_n = 0.
\]

We shall only prove (4.4.17) since the proof of (4.4.18) is entirely similar.

In

\[
H_{ij} = \left\{ (x,y) \mid i \leq x \leq i+1, \ j \leq y \leq j+1 \right\},
\]

let

\[
P_{ij} = \left\{ (x,y) \mid P(x,y) \geq 0 \right\},
\]
and
\[ N_{ij} = \left\{ (x,y) \mid P(x,y) < 0 \right\}. \]

Note that because of properties (i) and (ii)

\[ \int_{P_{ij}} P \, dx = \int_{P_{kl}} P \, dx = -\int_{N_{kj}} P \, dx = -\int_{N_{ij}} P \, dx \]

for all non-negative integers \( i, j, k \) and \( \ell \). By the Mean Value Theorem and using (4.4.19)

\[ A_n \leq \int_{H_{00}} P \, dx \sum_{i=0}^{\infty} \sum_{j=n}^{\infty} \frac{\left| \mu(x_{ij}) \right|}{r^2(x_{ij})} - \frac{\left| \mu(y_{ij}) \right|}{r^2(y_{ij})}, \]

where \( x_{ij} \in P_{ij} \) and \( y_{ij} \in N_{ij} \). Because of iv)

\[ |\mu(x_{ij}) - \mu(y_{ij})| = |s(\theta) - s(\theta')| \leq K_1 |\theta - \theta'|, \]

where \( K_1 \) is a bound for the derivative of \( s \). Now, for any two points \( (x,y) \) and \( (x',y') \) in \( H_{ij} \) with
\[ x = r \cos \theta, \quad y = r \sin \theta, \]

and
\[ x' = r' \cos \theta', \quad y' = r' \sin \theta', \]

we have
\[ (4.4.22) \quad 1 \geq |r' \cos \theta' - r \cos \theta| \]
\[ \geq \left| \cos \theta |r'-r| - |r||\cos \theta - \cos \theta'| \right| . \]

Also
\[ |r'-r| \leq \sqrt{2} \]
and \((4.4.22)\) together imply
\[ |\cos \theta - \cos \theta'| \leq \frac{K_2}{|r|} . \]

Therefore
\[ (4.4.23) \quad |\theta - \theta'| \leq \frac{K_2}{|r|} , \]

since necessarily \(0 \leq \theta , \theta' \leq \frac{\pi}{4} \). Using \((4.4.21)\) and \((4.4.23)\) we get
\[ \left| \frac{\mu(x_{i,j})}{r^2(x_{i,j})} - \frac{\mu(y_{i,j})}{r^2(y_{i,j})} \right| \leq \frac{|r^2(x_{i,j})-r^2(y_{i,j})| |\mu(x_{i,j})| + |\mu(x_{i,j})-\mu(y_{i,j})||r^2(y_{i,j})|}{r^2(x_{i,j}) r^2(y_{i,j})} \]
\[ \leq \frac{|r(y_{i,j})|K_4}{|r^2(x_{i,j})||r^2(y_{i,j})|} = \frac{K_4}{r^3(x_{i,j})} \]

where we have assumed that \(r(x_{i,j}) \leq r(y_{i,j})\). This shows, from \((4.7.20)\), that
\[ 0 \leq A_n \leq K_2 \int_0^\infty \int_n^\infty \frac{dx}{x^3} . \]

Since the function \(1/r^3\) is integrable in any closed region that does not contain the origin, we have proved that
\[
\lim_{n \to \infty} A_n = 0 ,
\]
and the proof of Lemma 3 has been completed.

**Lemma 4.** Let \( P(x,y) \) and \( u(x,y) \) be two functions such that:

(i) \( P(x+m,y+n) = P(x,y) \) \( \forall \) \( m, n \) integers,

(ii) \( \int_H P(x,y)dx \, dy = 0 \) with \( P(x,y) \) bounded on the unit square \( H \),

(iii) \( u(x,y) \) is a continuous function everywhere in \( H \), except possibly at the origin. It is also Riemann integrable on \( H \).

Then

\[
\lim_{n \to \infty} \int_0^1 \int_0^1 P(nx,ny) \, u(x,y) dx \, dy = 0 .
\]

**Proof.** Let \( a \) be a constant such that \( P(x,y)+a \) is non-negative in \( H \). Assume first that \( h \) is continuous everywhere on \( H \). By the mean value theorem we have

\[
\int_H P(nx,ny) \, u(x,y) dx \, dy = \sum_{i,j} \int_{H_{ij}} \left[ P(nx,ny)+a \right] \, u dx - a \int_H u dx \\
= \sum_{i,j} u(x_{ij}) \frac{a}{n^2} - a \int_H u dx,
\]

where
\[ x_{ij} \in H_{ij} = \left\{ (x,y) \mid \frac{i}{n} < x < \frac{i+1}{n}, \frac{j}{n} < y < \frac{j+1}{n} \right\}. \]

Note that we have used the fact that
\[ \int_{H_{ij}} P(nx, ny) dx \, dy = 0. \]

The result is now established, since by hypothesis
\[ \lim_{n \to \infty} \frac{1}{n^2} \sum_{i,j} \mu(x_{ij}) = \int_{H} \mu \, dx. \]

Now, in the general case it is clear that for any \( 0 < \varepsilon < 1 \)
\[ \lim_{n \to \infty} \int_{\varepsilon}^{1-\varepsilon} \int_{\varepsilon}^{1-\varepsilon} P(nx, ny) \mu(x, y) dx \, dy = 0, \]
but \( \mu \) is integrable on \( H \) so
\[ \left| \int_{H} \mu \, dx - \int_{\varepsilon}^{1-\varepsilon} \int_{\varepsilon}^{1-\varepsilon} \mu \, dx \right| < K \varepsilon \]
for some constant \( K \). This concludes the proof of Lemma 4.

We now show that the sequence
\[ b_n = \int_{H} r_n(x) dx + \int_{H} f_n(x) dx, \]
as given by (4.4.15), converges. Recall that
\[ r_n(x) = \left[ P_3(ny) \frac{\partial^3 r}{\partial y^3} + P_3(nx) \frac{\partial^3 r}{\partial x^3} + P_1(nx) \left( P_2(0) - P_2(ny) \right) \frac{\partial^3 r}{\partial y^2 \partial x} \right]. \]

If we define

\[ P_1(x,y) = P_3(y), \quad P_2(x,y) = P_3(x), \quad P_3(x,y) = P_1(x) \left( P_2(0) - P_2(y) \right). \]

and

\[ \mu_1(x,y) = r^2 \frac{\partial^3 r}{\partial x^3}, \quad \mu_2(x,y) = r^2 \frac{\partial^3 r}{\partial y^3}, \quad \mu_3(x,y) = r^2 \frac{\partial^3 r}{\partial x \partial y^2}. \]

we can use Lemma 3 to show that

\[ \int_{H} \frac{P_i(nx,ny) \mu_i(x,y)}{r^2} \, dx \, dy \]

converges for \( i = 1,2 \) and 3 and we can then write

\[ \gamma = \lim_{n \to \infty} \int_{H} r_n(x) \, dx = \lim_{n \to \infty} \frac{3}{2} \sum_{i=1}^{3} \int_{H} \frac{P_i(nx,ny) \mu_i}{r^2} \, dx \, dy. \]

The function \( f = \phi(x,y) - 4ar \) has property \( B_3 \) over \( H \), this immediately implies, by Lemma 4, that

\[ \lim_{n \to \infty} \int_{H} f_n(x) \, dx = 0. \]
We have now completely established the first part of Theorem 7, namely

\[ \sigma^2_{SY} = \frac{\phi}{n^2} + \frac{4\phi\gamma}{n^3} + o\left(\frac{1}{n^3}\right). \]

To conclude the proof recall

\[ \sigma^2_{MR} = \sigma^2_{SY} - 2 \left[ M_N(P) - \int_H P \, dx \right], \]

where \( P(x,y) \) is given by (4.4.4). Using the relationship

\[ M_N(P) = 2T_{2N}(P) - T_N(P) \]

we obtain, by Lemma 2, that

\[ M_N(P) - \int H P \, dx = -\frac{B^2}{4n^2} \int H \left( \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} \right) \, dx + \frac{1}{n^3} \int H \left[ P_{2N}(x) + P_N(x) \right] \, dx. \]

By Lemma 4

\[ \lim_{n \to \infty} \int H P_N(x) \, dx = 0. \]

Furthermore, we proved in Theorem 5 that

\[ \int H \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} \, dx = - \int H \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} \, dx. \]

We now get
\[ a_{MR}^2 = \frac{a}{n^2} + \frac{4aY}{n^3} + o\left(\frac{1}{n^3}\right) - \frac{a}{n^2} + o\left(\frac{1}{n^3}\right) = \frac{4aY}{n^3} + o\left(\frac{1}{n^3}\right). \]

Theorem 7 is now fully proved.

4.5 Some Numerical Results on Systematic and Midpoint Sampling

A FORTRAN program was written to illustrate the content of Theorem 7. The particular choice

\[ c(x,y) = h(r) = r \]

was used to calculate, for different values of \( N \), the quantities

\[(4.5.1) \quad R_n = T_n(\phi) - \int_H \phi \, dx,
\]

and

\[(4.5.2) \quad S_n = R_n - 2 \left[ M_n(P) - \int_H P \, dx \right],
\]

where

\[ n^2 = N, \]

\[ \phi(x,y) = (1-x)(1-y) \sqrt{x^2+y^2} \]

and

\[ P(x,y) = \int_0^1 \int_0^1 \sqrt{(x-\mu)^2+(y-\nu)^2} \, d\mu \, d\nu. \]
Note that $r$ is not a covariance function and we can not interpret the quantities $R_n$ and $S_n$ as mean square errors. However, the expansions

$$(4.5.2) \quad R_n = \frac{a}{n^2} + \frac{b\gamma}{n^3} + o\left(\frac{1}{n^3}\right),$$

and

$$(4.5.3) \quad S_n = \frac{b\gamma}{n^3} + o\left(\frac{1}{n^3}\right)$$

are still valid since the proof of Theorem 7 did not use the fact that $c(x,y)$ is a covariance function. The choice of $r$ does allow us to calculate exactly, within the precision of the computer arithmetic, the quantities $R_n$ and $S_n$. In fact, straightforward but lengthy calculations give

$$\int_H \phi \, dx = \int_H P \, dx = \frac{1}{15} (2 + \sqrt{2}) + \frac{1}{3} \ln(1 + \sqrt{2}) = .5214 \ldots$$

and

$$P(x,y) = f(x,y) + f(1-x,y) + f(x,1-y) + f(1-x,1-y)$$

where

$$f(x,y) = \frac{1}{6} \left[ 2xy \sqrt{x^2 + y^2} + x^3 \ln \left( \frac{\sqrt{x^2 + y^2} + x}{x} \right) + y^3 \ln \left( \frac{\sqrt{x^2 + y^2} + y}{y} \right) \right].$$

73
All the arithmetic was done in double-precision. The results, for increasing values of \( n = \sqrt{N} \), are displayed in Table 4.5.1.

<table>
<thead>
<tr>
<th>( n = \sqrt{N} )</th>
<th>( R_n )</th>
<th>( S_n )</th>
<th>( \frac{n^3}{4} \cdot S_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-0.9463 \times 10^{-1}</td>
<td>-0.3004 \times 10^{-1}</td>
<td>-0.06009</td>
</tr>
<tr>
<td>4</td>
<td>-0.1940 \times 10^{-1}</td>
<td>-0.3700 \times 10^{-2}</td>
<td>-0.05920</td>
</tr>
<tr>
<td>8</td>
<td>-0.4345 \times 10^{-2}</td>
<td>-0.4570 \times 10^{-3}</td>
<td>-0.05849</td>
</tr>
<tr>
<td>16</td>
<td>-0.1025 \times 10^{-2}</td>
<td>-0.5664 \times 10^{-4}</td>
<td>-0.05799</td>
</tr>
<tr>
<td>32</td>
<td>-0.2490 \times 10^{-3}</td>
<td>-0.7041 \times 10^{-5}</td>
<td>-0.05768</td>
</tr>
<tr>
<td>64</td>
<td>-0.6137 \times 10^{-4}</td>
<td>-0.8771 \times 10^{-6}</td>
<td>-0.05747</td>
</tr>
<tr>
<td>128</td>
<td>-0.1523 \times 10^{-4}</td>
<td>-0.1094 \times 10^{-6}</td>
<td>-0.05736</td>
</tr>
</tbody>
</table>

**Table 4.5.1**

Calculation of \( R_n \) and \( S_n \) -- Formulas (4.5.1) and (4.5.2)

Note that from (4.5.2) and (4.5.3) we should have, approximately, that

\[ R_{2n} \approx \frac{1}{4} R_n \]

and

\[ S_{2n} \approx \frac{1}{8} S_n \]

These results are confirmed in Table 4.5.1.
The last column in Table 4.5.1 should be converging to the constant

\[ \gamma = A + B, \]

where

\[ A = \lim_{n \to \infty} -6 \int_{H} P_{3}(nx) \left( \frac{x}{r^3} - \frac{x}{r^5} \right) \, dx \, dy \]

and

\[ B = \lim_{n \to \infty} \int_{H} P_{1}(nx) \left[ P_{2}(0) - P_{2}(ny) \right] \left[ -\frac{x}{r^3} + \frac{3xy^2}{r^5} \right] \, dx \, dy. \]

We have been unable to obtain a closed expression for \( B \). The quantity \( A \) can be calculated as follows:

On integrating with respect to \( y \),

\[ A = \lim_{n \to \infty} 2 \int_{0}^{1} \frac{P_{3}(nx)}{x} g(x) \, dx, \]

where

\[ g(x) = -\frac{3}{\sqrt{1+x^2}} + 3 \left[ \frac{1}{\sqrt{1+x^2}} - \frac{1}{3} \frac{1}{(1+x^2)^{3/2}} \right]. \]

We can write

\[ g(x) = -1 + x h(x) \]

where \( h(x) \) is an integrable function on \([0,1]\); then by Lemma 4 and using the expansion

75
\[ P_3(nx) = \sum_{i=1}^{\infty} \frac{2 \sin 2\pi ix}{(2\pi i)^3} \]

we finally get

\[ A = -2 \sum_{i=1}^{\infty} \int_{0}^{\infty} \frac{2 \sin 2\pi ix}{(2\pi i)^3} \, dx = -\frac{1}{4\pi^2} \sum_{i=1}^{\infty} \frac{1}{i^3} = -0.03044 \ldots \]
REFERENCES


