Summer School on

Stereology and Geometric Tomography

20–25 May, 2000

Sandbjerg Manor

Denmark

Supported by

- MaPhySto — Centre for Mathematical Physics and Stochastics, University of Aarhus
- StocLab — Laboratory for Computational Stochastics, University of Aarhus
- International Society for Stereology

Introduction

In the days 20–25 May, 2000 a Summer School on Stereology and Geometric Tomography was held in the beautiful surroundings of Sandbjerg Manor in the Southern parts of Denmark. The aim of the summer school was to give an overview of modern stereology and its relation to geometric tomography, including both the mathematical and statistical theory, and the practical applications.

In this small booklet we have gathered the program, abstracts, notes, and various other information relating to the summer school. It is our hope that this collection can be of use both to the participants of the school as well as to other researchers/students working in the field(s).

I take pleasure in thanking the other lecturers for their dedicated work and the participants for their positive attitude which were instrumental in creating the agreeable atmosphere that was evident during the summer school. Last but certainly not least I want to thank Søren Have Hansen and Oddbjørg Wethelund for their excellent organisation of the summer school which we all benefited from.

Eva B. Vedel Jensen

Contents

School Program (Final)

Legend:

(L) : Lectures
(TE) : Theoretical Exercises
(SL) : Special Invited Lectures
(PL) : Participants' Lectures
(PE) : Practical Exercises

Saturday 20 May 2000

15.00—16.00	Coffee/tea
16.00—16.30	**Welcome:** *Presentation of teachers and participants.*
16.30—17.10	**Eva B. Vedel Jensen:** *Introduction to stereology.*
	Richard Gardner: *Introduction to geometric tomography.*
17.15—18.00	**Adrian Baddeley:** *Analogy between stereology and survey sampling (L).*
18.00	Dinner
20.30—	Get-together

Sunday 21 May 2000

09.00—09.45	**Richard Gardner:** *Star bodies, chord functions, and section functions (L).*
10.00—10.45	**Eva B. Vedel Jensen:** *Stereological estimation of number (L).*

10.45—11.15	Coffee/tea
11.15—12.00	**Niels Væver Hartvig:** *Stereological estimation of number (TE).*
12.00—14.00	Lunch
14.00—14.45	**Richard Gill:** *Quantum tomography as a statistical inverse problem (SL).*
14.45—15.15	Coffee/tea
15.15—16.00	**Aljoša Volčič:** *Determination of a convex body by sections I (SL).*
16.10—16.50	**Markus Kiderlen:** *Endomorphisms of convex bodies (PL).*
17.00—17.30	**Boris Rubin:** *Arithmetrical properties of generalized Minkowski-Funk transforms and small denominators on the sphere (PL).*
18.00	Dinner
20.00—21.00	**Adrian Baddeley:** *The fractionator (PE).*

Monday 22 May 2000

09.00—09.45	**Richard Gardner:** *The spherical Radon transform and Funk's theorem (L).*
10.00—10.45	**Hans Jørgen G. Gundersen:** *Counting and sampling in 3D (demonstration at the microscope).*

10.45—11.15	Coffee/tea
11.15—12.00	**Adrian Baddeley:** *Stereological sampling designs, volume estimation (L).*
12.00—14.00	Lunch
14.00—14.45	**Aljoša Volčič:** *Determination of a convex body by sections II (SL).*
14.45—15.15	Coffee/tea
15.15—16.00	**Martin Bøgsted Hansen:** *Statistical aspects of inverse problems (SL).*
16.10—16.50	**Niels Holm Olsen:** *3D-reconstruction, light microscopy and optics of embryo (PL).*
17.00—17.30	**Andrew Olenko:** *Closeness of random fields in different multidimensional metrics (PL).*
19.00—	Summer school dinner

Tuesday 23 May 2000

09.00—09.45	**Richard Gardner:** *Determination by chord and section functions (L).*
10.00—10.45	**Eva B. Vedel Jensen:** *Length and surface area estimation under isotropy (L).*
10.45—11.15	Coffee/tea
11.15—12.00	**Kiên Kiêu:** *Length and surface area estimation under isotropy (TE).*

12.00—13.00	LUNCH
13.00—17.30	EXCURSION
18.00	DINNER

Wednesday 24 May 2000

09.00—09.45	**Richard Gardner:** *Affine inequalities and volume estimates (L).*
10.00—10.45	**Adrian Baddeley:** *Vertical sections (L).*
10.45—11.15	COFFEE/TEA
11.15—12.00	**Kiên Kiêu:** *Variance of planar area estimators based on systematic sampling (L).*
12.00—14.00	LUNCH
14.00—14.45	**Boris Rubin:** *Continuous wavelet transforms in geometric tomography (SL).*
14.45—15.15	COFFEE/TEA
15.15—15.45	**Eugene Spodarev:** *One isoperimetrical problem for stationary flat processes (PL).*
15.45—16.15	**Martin Bøgsted Hansen:** *Nonparametric estimation of the chord length distribution (PL).*

16.30—17.00	**Marta Garcia-Finana:** *Fractional trend of the variance under systematic sampling on R (PL).*
17.00—17.30	**Stephan Böhm:** *On Laslett's test for Boolean model (PL).*
18.00	DINNER
20.00—21.00	**Hans Jørgen G. Gundersen:** *How to estimate the volume and surface area of a banana (PE).*

Thursday 25 May 2000

09.00—09.45	**Eva B. Vedel Jensen:** *Local stereology (L).*
10.00—10.45	**Kiên Kiêu:** *Variance of planar area estimators based on systematic sampling (TE).*
10.45—11.15	COFFEE/TEA
11.15—12.00	**Hans Jørgen G. Gundersen:** *Connectivity (L).*

Abstracts of Lectures

Adrian Baddeley

University of Western Australia

Analogy between stereology and survey sampling

ABSTRACT: The statistical principles which underlie modern stereological methods are closely analogous to the classical theory of survey sampling. A slice through a three-dimensional object is analogous to a survey sample of a population. The validity of estimators is guaranteed by the randomness of the sampling design, which is under our control, rather than by assumptions about the population or the observations. This lecture will develop the analogy, and show that many of the key concepts of survey sampling are also important in stereology. In particular, sampling bias is frequently present in stereological experiments, and the Horvitz-Thompson device is frequently used to correct for sampling bias.

Stereological sampling designs, volume estimation

ABSTRACT: This lecture discusses some of the main types of stereological sampling design with emphasis on the estimation of absolute volume and volume fraction.

Stereological methods for estimating geometrical quantities are based on identities in integral geometry. For each integral formula, there are several possible stochastic interpretations which will yield an unbiased estimator of the same quantity; these correspond to different stereological sampling designs. It is also easy to construct erroneous designs which superficially appear correct but which yield invalid estimators.

In biological applications of stereology, it is typical for the experimental protocol to consist of several consecutive stages. At each stage the available material is subdivided or sectioned, and a random sample of these pieces is taken. This "nested" design will be discussed and the use of ratio estimators outlined.

Vertical sections

ABSTRACT: In some stereological experiments, the section plane is constrained to be normal to a fixed "horizontal" plane, or equivalently parallel to a fixed "vertical" axis. Examples include cross-sections of metal fracture surfaces (taken normal to the macroscopic plane of fracture); sections of a large biological organ cut perpendicular to the laboratory bench; and transverse sections of skin. This sampling design does not satisfy the usual requirement of isotropy, and indeed has a palpable sampling bias. Nevertheless, it is possible to obtain an unbiased estimator of surface area from vertical sections. We derive this estimator and sketch some practical implementations. The underlying principle is one of the prototypes for 'local' stereology.

Richard Gardner

Western Washington University

Star bodies, chord functions, and section functions

Abstract: We begin with a brief review of the principal methods of computer tomography (namely, linear algebra and the Fourier transform), and a reminder of the main theme of the lectures, mentioned in the introductory talk: the mysterious duality between results and methods concerning projections of convex bodies and results and methods concerning concurrent sections of star bodies. This duality, first noticed by E. Lutwak, is not at all explained by the well-known polar duality; however, the latter already involves the radial function ρ_L of a star body L , the function giving the signed distance from the origin to the boundary of L. The i-chord functions $\rho_{i,L}$ are generalizations of the radial function that can be defined for any real number i. They are particularly useful when i is an integer strictly between 0 and n, the dimension of the space (but other values are also relevant, as in the various forms of the notorious equichordal problem). For these values of i, the i-chord function is closely related (via the polar coordinate formula for volume) to the ith section function of a star body, the function giving the i-dimensional volumes of its intersections with i-dimensional subspaces. When $i = 1$, the ith section function coincides with the 1-chord function, also known as the point X-ray at the origin (or, in the computer tomography literature, the fan-beam X-ray at the origin). The $(n-1)$th section function is simply called the section function.

This and the other lectures are based on parts of the speaker's book *Geometric Tomography*, Cambridge University Press, New York, 1995, henceforth referred to as [G]. The contents of Lecture 1 can be found in Sections 0.7, 0.8, 6.1, 6.3, and 7.2 of [G].

The spherical Radon transform and Funk's theorem

Abstract: The spherical Radon transform Rf of a Borel function f on the unit sphere is another function defined on the unit sphere whose value at a point u is the integral of f over the great sphere orthogonal to u. The section function of a star body containing the origin is, up to a constant, just $R(\rho_L^{n-1})$. Using this and the injectivity of R on even functions (a property that seems to require spherical harmonics for its proof), it can be proved that under mild restrictions, two star bodies have equal ith section functions if and only if they have equal i-chord functions. It follows that origin-symmetric star bodies are determined by their ith section function (a general form of Funk's theorem), and therefore that an origin-symmetric star body of constant i-section must be a ball with center at the origin. On the other hand, there are for each i non-spherical convex bodies of constant i-section. Analogous results from the classical Brunn-Minkowski theory are Aleksandrov's

projection theorem, which states that origin-symmetric convex bodies are determined by their ith projection function, and the existence of non-spherical convex bodies of constant i-brightness (that is, constant ith projection function). Certain symmetric bodies are also discussed: the i-chordal symmetrals, and some classical counterparts, the central symmetral and the Blaschke body.

Lecture 2 is based on material in Sections 6.3, 7.2, and C.2 of [G].

Determination by chord and section functions

Abstract: The first part of this lecture continues the theme of determination of star bodies by chord and section functions. We ask when a star body is determined by its i- and j-chord functions, where i and j are different real numbers. In general one can only say that a star body is locally determined, up to reflection in the origin, by such data (even when its i-chord functions for all real i are known). However, additional conditions, for example that the body contains the origin in its interior and has real analytic boundary, or that the body does not contain the origin and is connected, suffice for determination up to reflection in the origin. One can also conclude that a star body of constant i-section and constant j-section, $i \neq j$, must be a ball with center at the origin. Once again, there are classical analogs of these results. It is known that a convex body of constant i-brightness and constant j-brightness must be a ball if its boundary has positive Gaussian curvature everywhere; without the extra smoothness assumption, however, it is even unknown whether a convex body in three dimensions of constant width and constant brightness must be a ball.

The second part of the lecture begins to consider how the volume of a star body might be estimated from its section function. Using Hölder's inequality, one can obtain E. Lutwak's dual isoperimetric inequality. This, together with the dual of Cauchy's surface area formula, quickly leads to a lower bound for the volume in terms of the average of the section function of the body, in which equality holds precisely for balls with center at the origin.

Lecture 3 is taken from Sections 6.2, 7.2, and B.4 of [G].

Affine inequalities and volume estimates

Abstract: In this lecture we continue to study the question of how to estimate volume from section functions. It turns out that the estimate obtained in Lecture 3 can be dramatically improved: the average of the section function can be replaced by the nth mean of the section function. Moreover, this improvement is best possible in that the nth mean cannot be replaced by the pth mean for $p > n$, and it is affine invariant, equality holding precisely for ellipsoids with center at the origin. The result can be reformulated using the notion of the intersection body IL of a star body L containing the origin. The radial function of IL equals the section function of L. Then the affine-invariant estimate can be written as an upper bound for the volume of IL in terms of the volume

of L and is known as Busemann's intersection inequality. The classical analog is the Petty projection inequality, an affine-invariant inequality stronger than the isoperimetric inequality, giving an upper bound for the volume of the polar projection body of a convex body K in terms of the volume of K. The proof of Busemann's intersection inequality, which uses Steiner symmetrization and the Blaschke-Petkantschin formula, is sketched.

Finally, some remarks are made about estimating upper bounds for volume from section functions. This question quickly leads to one of the most important open problems in the area, the so-called slicing problem concerning central sections of an origin-symmetric convex body. Several formulations are briefly discussed.

Lecture 4 comes from Section 9.4 and Note 9.6 in [G].

Hans Jørgen G. Gundersen

Stereological Research Laboratory, University of Aarhus

Connectivity

Abstract: The Euler number and the connectivity of an arbitrary object is defined, and it is illustrated why the connectivity of an n-dimensional object cannot be estimated in an $(n-1)$-dimensional section. The disector-principle for 3D counting of the Euler-events is illustrated in cancellous bone. The correct handling for unbiased counting of events at artificial edges is outlined. A nomogram for predicting the precision of an estimate is provided.

Eva B. Vedel Jensen

University of Aarhus

Stereological estimation of number

Abstract: In this lecture, we discuss how to estimate the number of objects in a finite population. For a population of planar objects, estimates of population number based on 2D disector sampling, Gundersen's tiling rule, the associated point rule and plus sampling are derived. All designs are special cases of a generalized type of cluster sampling and the estimators are derived by the Horwitz-Thompson procedure. We also discuss disector sampling in 3D as well as 3D counting. Finally, we will shortly mention fractionator sampling.

Length and surface area estimation under isotropy

Abstract: The stereological method of estimating the length of a planar curve from counting the number of intersection between the curve and a line grid goes back to Buffon (1777). He found the probability that a randomly dropped needle crosses a grid line. We will start by rederiving his result and explain how it can be used for estimating the length of a planar curve, from intersection counts with a uniform and isotropic line grid. We will thereafter discuss estimation of length and surface area in 3D, using isotropic spatial line or plane grids. Finally, we will briefly mention the use of vertical sampling planes in 3D.

Local steoreology

Abstract: Local stereology is a collection of stereological designs based on sections through reference points. In this lecture we will give an introduction to local stereology in $\mathbb{E}^n$ and relate this field to geometric tomography. We will begin by a simple example, viz. local estimation of planar area. Next, we will present the local stereological volume estimators in $\mathbb{E}^n$ and show that for a large class of bodies, they are proportional to the section functions from geometric tomography. Finally, we will give a review of the local stereological volume estimators in $\mathbb{E}^n$.

Kiên Kiêu (with Marianne Mora)

Institut National de la Recherche Agronomique

Variance of planar area estimators based on systematic sampling

Abstract: Systematic sampling is widely used in practical stereology. Examples of systematic sampling probes are serial sections, line and point grids. Assessing the precision of such designs is not a trivial task because of the statistical spatial dependency of the data. First methods for assessing the precision of systematic geometric sampling are due to Kendall (1948, 1953) and Matheron (1965, 1971).

We present their approach on a particular case: the estimation of planar area based on sampling by parallel lines. First, the estimation variance is expressed in terms of the Fourier transform of the indicator function associated to the investigated body. Using classical tools from analysis (Gauss-Green formula, method of the stationary phase), an asymptotic approximation of the Fourier transform is derived. This yields an asymptotic approximation of the estimation variance involving some simple geometric features of the body boundary.

Abstracts of Special Invited Lectures

Richard Gill

University of Utrecht

Quantum tomography as a statistical inverse problem

ABSTRACT: The idea that one could reconstruct the state of a quantum system by taking appropriate measurements of many copies of the system has been a matter of speculation in physics for many many years. It was actually done in the laboratory for the first time in 1993 by Raymer, in the field of quantum optics. The non-stochastic version of the problem turns out to be equivalent to the problem of recovering a function on $\mathbb{R}^2$ from all its one-dimensional projections, hence the name 'quantum tomography'. Only very recently have physicists started to take account of the randomness of the outcomes coming from what they call 'finite statistics': the fact that only a finite number of copies of the system can be measured. Randomness is intrinsic to any measurement of a quantum system, and the theory actually specifies the probability density of the measurements.

The problem can now be formulated as an inverse statistical problem: the data has a density which is a given functional of an unknown infinite-dimensional parameter, and one wants to estimate various functionals of the parameter. So far only ad hoc statistical approaches have been tried. The challenge is to see if modern statistical theory (curve estimation, sieved maximum likelihood estimation, ...?) can provide practically useful estimators with good statistical properties.

Interestingly, recent work by the physicists in this area suggests also new approaches for classical tomography (inversion of the Radon transform).

Introductory material on quantum statistics and quantum tomography can be found on the speakers webpage `http://www.math.uu.nl/people/gill`

Martin Bøgsted Hansen

MaPhySto & Aalborg University

Nonparametric estimation of the chord length distribution

Abstract: The distribution of the length of a typical chord of a stationary random set is an interesting feature of the set's whole distribution. A nonparametric (maximum likelihood) estimator of the chord length distribution is given and its properties will be studied. The estimator will be compared on simulated as well as real data.

Statistical aspects of inverse problems

Abstract: The purpose of the talk is to give an introduction to inverse problems, which arise in many scientific areas. Informally, one can state a direct problem as calculating the effect of some given causes, whereas the inverse problem is to derive the causes given some effects.

An example is scattering of waves (e.g. ultrasound imaging). The direct problem is to calculate the scattered waves given the scattering medium. On the opposite the inverse problem is to find the scattering media given the wave source and the scattered waves. As noise and uncertainty about the specified model are inevitable in many experiments the talk will focus on statistical aspects of inverse problems.

Boris Rubin

The Hebrew University of Jerusalem

Continuous wavelet transforms in geometric tomography

Abstract: This is a survey lecture devoted to application of continuous wavelet transforms to explicit inversion of various Radon-type transforms in Geometric Tomography. We consider totally geodesic Radon transforms on real spaces of constant curvature, the Minkowski-Funk transforms on the n-sphere, exponential k-plane transforms, horocycle transforms in the hyperbolic space and related problems. The key idea is to include Radon transforms under consideration into suitable analytic families of fractional integrals which give rise to the relevant continuous wavelet transforms. Some open problems leading to harmonic analysis and number theory are indicated.

Aljoša Volčič

University of Trieste

Determination of a convex body by sections

1. Position of the problem

There are several papers in the literature which study the determination of convex bodies from the measures of their intersections with lines and, in higher dimension, with affine subspaces. These kind of questions where initiated by Hammer's problem [H] which motivated many papers since then. For an exhaustive account on the subject see [G], Chapters 2, 5, 6 and 7.

The existing results can be divided in two classes. One group of papers gives conditions on the position and number of points p_1, p_2, ...,p_k which assure that a convex body $K \in \mathbb{R}^m$ is uniquely detemined by the measures of the intersections of K with all the affine subspaces of a given dimension i, $1 \leq i \leq m-1$, through p_h, $1 \leq h \leq k$ (see [H], [F]). We may also allow that some of the points are at infinity.

In other papers just one point p is fixed, and the measures of intersections correspond to affine subspaces of (at least) two different dimensions (see [GSV] and [GV]).

In these lectures we will not address this second type of questions.

We will be concerned in fact with a more general situation, since the problem will be put in the frame of the so-called i-chord functions, first introduced in [G1]. This concept is not so appealing from the geometric point of view, but turns out to be an indispensable tecnical mean.

2. Definitions and notations

A convex body in $\mathbb{R}^n$ is a compact convex set with non empty interior. We shall denote by int$K$ the interior of K and by ∂K its boundary. If p_1, p_2 are two points, we shall denote by p_1p_2 the line joining them. If p_1, p_2, p_3 are three non collinear points in $\mathbb{R}^m$, with $m \geq 3$, we will denote by $p_1p_2p_3$ the 2-dimensional plane containing them.

If K is a convex body, its *radial function* ρ_K is defined, for all $u \in S^{m-1}$ such that the line through the origin parallel to u intersects K, by

$$\rho_K(u) = \max\{c : cu \in K\}\,.$$

Let us denote by $\mathcal{G}(m,i)$ the set of all i-dimensional subspaces of $\mathbb{R}^m$. By λ_i we denote the i-dimensional Lebesgue measure.

If K is an m-dimensional convex body and i is a fixed integer, with $1 \leq i \leq m-1$, then the i-section function of K at the origin is the function which assigns to every $G \in \mathcal{G}(n,i)$ the number $\lambda_i(G \cap K)$.

In particular, if $i = 1$, the i-section function is also called the X-ray function of K at p.

We will consider also the X-ray function of a convex body K corresponding to a direction, i.e. to a point at infinity. If u is a vector belonging to the unit sphere S^{m-1} of $\mathbb{R}^m$, then the X-ray function of K in direction u is the function which assigns, to every line L parallel to u, the number $\lambda_1(E \cap L)$.

For $i \in \mathbb{R}$, let us define the i-chord function of a convex body K in $\mathbb{R}^m$. If $i \leq 0$, we have to assume that the origin does not belong to the boundary of K. The *i-chord function $\rho_{i,K}$ of K at the origin* is the function defined for all $u \in S^{m-1}$ as follows. If the line through the origin, parallel to u, does not intersect K, we define $\rho_{i,K}(u) = 0$. Otherwise, if $i \neq 0$, we let

$$\rho_{i,K}(u) = \begin{cases} \rho_K^i(u) + \rho_K^i(-u) & \text{if } K \text{ contains the origin} \\ \big| |\rho_K(u)|^i - |\rho_K(-u)|^i \big| & \text{if } K \text{ does not contain the origin.} \end{cases}$$

If $i = 0$, we let

$$\rho_{i,K}(u) = \begin{cases} \rho_K(u)\rho_K(-u) & \text{if } K \text{ contains the origin} \\ \exp\big|\log|\rho_K(u)/\rho_K(-u)|\big| & \text{if } K \text{ does not contain the origin.} \end{cases}$$

The *i-chord function of K at $p \in \mathbb{R}^m$* is simply the i-chord function at the origin of the set $K + p = \{x \in \mathbb{R}^m : x + p \in K\}$.

Note that for $i = 1$, the i-chord function and the i-section function are just the same.

Let us state now the following useful statement, which goes back to P. Funk, [Fu] and gives a link between i-chord functions and i- section functions for $1 \leq i \leq m-1$. A more general result can be found in [GV].

Proposition 2.1. *Suppose that K_1 and K_2 are convex bodies in $\mathbb{R}^m$ and let $1 \leq i \leq m-1$. Then*

$$\rho_{i,K_1}(u) = \rho_{i,K_2}(u) \ \text{ for all } u \in S^{m-1}$$

if and only if

$$\lambda_i(K_1 \cap G) = \lambda_i(K_2 \cap G) \ \text{ for all } G \in \mathcal{G}(m,i)\,.$$

The link between i-chord and i-section functions in a given direction u is given by the following statement. Let us denote by $u^\perp$ the plane through the origin orthogonal to u and by $\mathcal{G}_u(m,i)$ the set of all the i-dimensional affine subspaces orthogonal to $u^\perp$.

Proposition 2.2. *Suppose that K_1 and K_2 are m-dimensional convex bodies, and let $1 < i \leq m-1$ and let u be a direction. Then*

$$\lambda_i(K_1 \cap G) = \lambda_i(K_2 \cap G)\,, \tag{1}$$

for each $G \in \mathcal{G}_u(m,i)$ if and only if

$$\lambda_1(K_1 \cap L) = \lambda_1(K_2 \cap L)\,, \tag{2}$$

for every line in direction u.

This proposition confirms that there is no natural definition of i-chord functions of a convex body *in a direction* u for $i \neq 1$.

Definition. Let $i \in \mathbb{R}$ and consider the subspace $G = \{(x_1, x_2, \dots, x_m) : x_m = 0\}$. For a bounded measurable set E in $\mathbb{R}^m$ let us define the measure

$$\nu_i(E) = \int_E |x_m|^{i-m} dx_1 \dots dx_m \ .$$

The following proposition is one of the important tools which are used for proving the uniqueness results listed in Section 3.

Proposition 2.3. *Suppose that K_1 and K_2 are two convex bodies in $\mathbb{R}^m$ not intersecting the plane $\{x = (x_1, x_2, \dots, x_m) : x_m = 0\}$. Suppose moreover that K_1 and K_2 have, for some $i \in \mathbb{R}$, the same i-chord function at $p = (t_1, t_2, \dots, t_{m-1}, 0)$. Let A and A' be two components of $\operatorname{int}(K_1 \triangle K_2)$ such that a line through p intersects A if and only if it intersects A'. Then*

$$\nu_i(A) = \nu_i(A') \, .$$

3. The main results

We will list here the main theorems concerning the determination of convex bodies from i-chord functions at two or more points.

Theorem 3.1 ([GM], [G]). *Suppose K and H are convex bodies in $\mathbb{R}^2$ and suppose P is a finite set of points on a line l not intersecting K, not projectively equivalent to a subset of directions of the edges of a regular polygon and such that K and H have the same i-chord function at the points in P.*

Then $K = H$.

Theorem 3.2 ([F], [G$_1$], [V], [G]). *Suppose K and H are convex bodies in $\mathbb{R}^m$ and suppose p_1 and p_2 are points in $\mathbb{R}^m$ such that K and H have the same i-chord function at p_j, $j = 1, 2$. with $i > 0$. Suppose moreover that the line through p_1 and p_2 intersects K.*

Then $K = H$ if

a) the line $p_1 p_2$ meets $\operatorname{int} K$ and p_1, p_2 do not belong to $\operatorname{int} K$, and K and H meet the same component of $p_1 p_2 \setminus \{p_1, p_2\}$,

b) p_1, p_2 belong to $\operatorname{int} K$ or

c) the line through p_1 and p_2 supports K.

Theorem 3.3. *Suppose K and H are convex bodies in $\mathbb{R}^m$ and suppose p_1 and p_2 are points in $\mathbb{R}^m$ not belonging to K such that K and H have the same i-chord function at p_1, and the same j-chord function at p_2, with $i, j \in \mathbb{R}$.*

Then $K = H$ *if* $\operatorname{int} K$ *and* $\operatorname{int} H$ *meet the segment* $[p_1, p_2]$.

Remark. There is no restriction on i and j.

Theorem 3.4 ([V]) . *Suppose* K *and* H *are convex bodies in* $\mathbb{R}^2$ *and suppose* p_1, p_2, p_3 *and* p_4 *are points in generic position in* $\mathbb{R}^2$ *such that* K *and* H *have the same X-rays at* p_j, $j = 1, 2, 3, 4$.

Then $K = H$.

Theorem 3.5. *Suppose* K *and* H *are convex bodies in* $\mathbb{R}^m$ *and suppose* p_1, p_2 *and* p_3 *are non collinear points in* $\mathbb{R}^m$ *such that* K *and* H *have the same* i*-chord function at* p_j, $j = 1, 2, 3$, *with* $i > 1$.

Then $K = H$ *if* $\operatorname{int} K$ *meets the plane* $p_1p_2p_3$ *and* $p_j \notin K$ *for* $j = 1, 2, 3$.

Remark. If $p_j \in K$ for $j = 1, 2, 3$, then $K = H$ follows from Theorem 3.1, for all $i > 0$.

References

[F$_1$] K. J. Falconer, *X-ray problems for point sources*, Proc. London Math. Soc. **46** (1983), 241–262.

[Fu] P. Funk, *Über Flächen mit lauter geschlossenen geodätischen Linien*, Math. Ann. **74** (1913), 278–300.

[G$_1$] R. J. Gardner, *Chord functions of convex bodies*, J. London Math. Soc. (2) **36** (1987), 314–326.

[G] R. J. Gardner, *Geometric Tomography*, Cambridge Univ. Press, 1995.

[GM] R. J. Gardner and P. McMullen, *On Hammer's X-ray problem*, J. London Math. Soc. (2) **21** (1980), 171–175.

[GSV] R. J. Gardner, A. Soranzo and A. Volčič, *Determination of Star and Convex Bodies by Section Functions*, Discrete Comput. Goem. **21** (1999), 69–85.

[GV] R. J. Gardner and A. Volčič, *Tomography of Convex and Star Bodies*, Advances in Math. **108** (1994), 367–399.

[H] P. C. Hammer, *Problem 2; In: Proc. Symp. Pure Math., vol. VII: Convexity*, Amer. Math. Soc., 1963, pp. 498–499.

[V] A. Volčič, *A three-point solution to Hammer's X-ray problem*, J. London Math. Soc. (2) **34** (1986), 340–359.

Abstracts of Participants' Lectures

Stephan Böhm

University of Ulm

On Laslett's Test for Boolean Model

ABSTRACT: In this talk plain images with black and white areas are considered, which are, for example, extracted from data of groundwater examinations. The black areas correspond to those geographical locations, where some groundwater index is above a certain threshold. These areas are modeled as random closed sets. In order to prove the hypothesis if the observed image could be the realization of a stationary Boolean Model with compact and convex grains, Laslett's Test is applied. Furthermore, a significance test for the area fraction is considered, which is based on asymptotic normality of the estimated area fraction and on consistency of the estimated covariance function. If one can assume that the underlying random closed set is a stationary Boolean Model, the assumtions of this significance test are satisfied reasonably well; see Baddeley (1980) and Mase (1982). Finally, results of data analysis and implementation techniques will be presented.

References

Baddeley A.J. (1980): A limit theorem for statistics of spatial data. *Adv. Appl. Probab.* **12**, 447-461.

Mase S. (1982:) Asymptotic properties of stereological estimators of volume fraction for stationary random sets. *J. Appl. Probab.* **19**, 111-126.

Marta García-Fiñana (with Luis M. Cruz-Orive)
University of Cantabria

Fractional trend of the variance in systematic sampling on $\mathbb{R}$

Abstract: Unidimensional systematic sampling is widely used in design stereology. For this reason it is important to predict the variance of an estimator under this kind of sampling. The corresponding theory has undergone significant advances in the last years but it still has unsolved important points.

The problem is to estimate the integral Q of a non-random measurement function $f : \mathbb{R} \rightarrow \mathbb{R}^+$ of bounded support by systematic sampling of period T on $\mathbb{R}$. The current theory connects the variance of the estimator $\widehat{Q}$ of Q with the smoothness properties of f. The singularities of the first non-continuous derivative of f, $f^{(m)}$, $(m = 0, 1, ...)$, determine the main contribution to the variance, called the *extension term*, which is of $\mathrm{O}\left(T^{2m+2}\right)$ when T is small. If the singularities of $f^{(m)}$ are not finite, however, then the current theory cannot be applied, as we illustrate with some examples. To handle the problem we present an extended version of the Euler-MacLaurin summation formula which allows a more general representation of the variance. We show that the extension term decreases with fractional powers of the number of observations, that is:

$$\mathrm{Var}_E(\widehat{Q}) = O\left(T^{2q+2}\right), \quad q \geq 0,$$

and we express $\mathrm{Var}_E(\widehat{Q})$, and the remaining variance components, in terms of fractional q-derivatives of f. Our results lead to more general variance estimators. Finally, we present a class of measurement functions which will still require a more general theory.

Marcus Kiderlen
University of Karlsruhe

Endomorphisms of convex bodies

Abstract: The talk 'Endomorphisms of Convex Bodies' describes properties of certain mappings, the 'endomorphisms', from the set $\mathcal{K}$ of convex bodies in $\mathbb{R}^d$ into itself. These mappings are compatible with the most basic geometrical structures on $\mathcal{K}$. More precisely they are assumed to be continuous, Minkowski-additive and to commute with all proper rotations. An example is the function that maps a convex body to the invariant mean (in the sense of Minkowski-addition) of orthogonal projections on k-planes $(0 < k < d)$.

We state a representation theorem for endomorphisms using the generalized spherical Radon transform. Injectivity properties are discussed. We will see that the only convex bodies that are 'fixed points' of (nontrivial monotonous) endomorphisms are balls.

Andrew Olenko

Kyiv University

Closeness of random fields in different multidimensional metrics

Abstract: In many cases mathematical models for spatial phenomenon or images are obtained as particular instances of random fields. Their correlation or spectral functions often characterize models of this type reasonably well. There are a lot of numerical methods for estimating the values of correlation or spectral functions. Since we usually have a finite number of observations, it is clear that these methods build estimates only for finite area. In this presentation there will be given some estimates of the closeness in different metrics of the spectral and correlation functions of random fields.

Niels Holm Olsen

University of Copenhagen

3D-reconstruction, light microscopy and optics of embryo

Abstract: The evaluation of the fertility of living embryo is of practical importance in the daily routines at fertility clinics as well as an ongoing biological research topic. Living embryo may be studied through a light microscope. By focusing the microscope at different optical sections, the three-dimensional structure of the embryo may be studied. The reconstruction of the three-dimensional structure of living embryo from optical sectional images, must be based on a model of the image formation. The image formation is a result of the optical chararacteristics of the embryo and the microscope optics used.

Boris Rubin

The Hebrew University of Jerusalem

Arithmetrical properties of generalized Minkowski-Funk transforms and small denominators on the sphere

Abstract: The Cauchy problem for the Euler-Poisson-Darboux equation on the sphere and various transforms of integral geometry (including those of Minkowski and Funk) give rise to a family of fractional integrals associated with a spherical cap of fixed radius θ. These fractional integrals are called the generalized Minkowski-Funk transforms.

Investigation of injectivity, invertibility and boundedness of these transforms in Sobolev spaces leads to small denominators for spherical harmonic expansions and to some delicate problems related to arithmetical properties of zeros of the associated Legendre functions.

Problems of such a type can be stated also for Jacobi (or Gegenbauer) polynomials. For example, how does the number of common zeros of such polynomials, say, $P_j(t)$, $t = cos\theta$, depend on θ? In other words, how many j's satisfy the equation $P_j(t) = 0$ for t fixed? The results are different depending on whether θ is a rational or irrational multiple of π. In simplest cases the usual techniques of diophantine approximations are applicable.

Eugene Spodarev

Friedrich-Schiller-Universität Jena

One isoperimetrical problem for stationary flat processes

Abstract: Consider any stationary process Φ of k-flats in $\mathbb{R}^d$ with intensity λ and the directional distribution $\theta(\cdot)$. Intersections of any two k-planes of Φ induce the new stationary $2k-d$ flat process whose intensity is called *the intersection density of* Φ. A series of researchers (R. Davidson (1974), J. Janson and O. Kallenberg (1981), J. Mecke and C. Thomas (1984, 1988), J. Keutel (1992)) dealt with the following isoperimetrical problem concerning Φ: one has to find such an extremal directional distribution $\theta(\cdot)$ that would maximize the above intersection density of Φ provided that λ is fixed. In the case of hyperplanes ($k=d-1$) the solution is unique and coincides with Haar measure on the sphere. In other particular cases the whole class of extremal measures $\theta(\cdot)$ was described but nevertheless there are still some open questions there (e. g. when d is not divisible by $d-k$) .

The main result of our research yields the necessary conditions of extrema for arbitrary dimensions d and k which are expressed in terms of the rose of intersections of Φ; we get also retrieval formulae for the directional distribution of any stationary process of k-flats in $\mathbb{R}^d$ from its rose of intersections for some important particular cases. Then the characteristical properties of a rose of intersections are obtained.

The proofs involve the ideas of variational calculus on the cone of measures and harmonic analysis on Grassman manifolds.

References

[1] I. Molchanov, S. Zuev "Variational analysis of functionals of a Poisson process" *Rapport de Recherche N 3302 INRIA, 1997*

[1] J. Mecke, C. Thomas "On an extreme value problem for flat processes" *Commun. Statist.-Stochastic models 2(2) (1986), 273-280*

[3] E. Spodarev "On the rose of intersections for stationary flat processes" *Jenaer Schriften zur Mathematik und Informatik, April 2000*

[4] E. Spodarev "Isoperimetrical problems and roses of intersections for stationary flat processes" *to appear*

Theoretical Exercises

Length and surface area estimation under isotropy

Exercise 1. Let Z be a line-segment in the plane $\mathbb{R}^2$ of length l. Let $L_{\theta,u}$ be the line with unit normal $(\cos\theta, \sin\theta)$, $\theta \in [0, \pi)$, and signed distance $u \in \mathbb{R}$ to the origin, i.e.

$$L_{\theta,u} = \{(x, y) : x\cos\theta + y\sin\theta = u\}.$$

A uniform and isotropic line G_1 is then given by

$$G_1 = \{L_{\Theta,U+j\Delta} : j = \ldots, -1, 0, 1, \ldots\},$$

where Θ is uniform random in $[0, \pi)$ and U is independent of Θ and uniform in $[0, \Delta)$.

1. Show under the assumption $l < \Delta$ that, conditionally on Θ,

$$P(G_1 \text{ hits } Z|\Theta = \theta) = \frac{L(\pi_\theta Z)}{\Delta} = \frac{|\cos(\theta - \theta_0)| \cdot l}{\Delta},$$

where π_θ is the orthogonal projection onto the line

$$\{u(\cos\theta, \sin\theta) : u \in \mathbb{R}\}$$

and $\theta_0 \in [0, \pi)$ is the angle that the line-segment Z makes with the $x-$axis.
Hint. Note that

$$Z \cap L_{\Theta,U+j\Delta} \neq \emptyset \Leftrightarrow (U + j\Delta)(\cos\theta, \sin\theta) \in \pi_\theta Z.$$

2. Show that the unconditional hitting probability becomes

$$P(G_1 \text{ hits } Z) = \frac{2l}{\pi\Delta}.$$

Hint. Without loss of generality we can assume that $\theta_0 = 0$.

Exercise 2. Let Z be a line-segment in $\mathbb{R}^3$ of length l. Let $L_{\omega,u}$ be the plane in $\mathbb{R}^3$ with unit normal $\omega \in S^2_+$ and signed distance $u \in \mathbb{R}$ to the origin. A uniform and isotropic plane grid G_2 in $\mathbb{R}^3$ is then given by

$$G_2 = \{L_{\Omega,U+j\Delta} : j = \ldots, -1, 0, 1, \ldots\},$$

where $\Omega \in S^2_+$ is an isotropic direction and U is independent of Ω and uniform in $[0, \Delta)$.

Using the same type of reasoning as in Exercise 1, it can be shown that

$$P(G_2 \text{ hits } Z) = \frac{L(Z)}{2\Delta}.$$

The important new step, compared to the planar case described in Exercise 1, is to show that the random variable $L(\pi_\Omega Z)/L(Z)$ is uniform in $[0, 1]$. Here, π_Ω is the orthogonal projection onto the line spanned by Ω.

Show that $L(\pi_\Omega Z)/L(Z)$ is uniform in $[0, 1]$.
Hint. Without loss of generality, it can be assumed that Z is parallel to the $z-$axis. In that case

$$L(\pi_\Omega Z)/L(Z) = \cos\Theta,$$

where Θ is the angle between Ω and the $z-$axis. For an isotropic Ω, Θ has probability density

$$p(\theta) = \sin\theta, \theta \in [0, \frac{\pi}{2}).$$

Exercise 3. Let X be a bounded flat surface in $\mathbb{R}^3$. Let P be the horizontal plane through the origin. For any point $z \in P$, let L_z be the vertical line through z. A uniform random grid of vertical lines is then given by

$$G_1 = \left\{ L_{U+\Delta j} : j \in Z^2 \right\},$$

where $\Delta > 0$ and U is a uniform random point in $[0, \Delta) \times [0, \Delta) \times \{0\}$.

Below, it is assumed that X is contained in a ball with radius smaller than $\Delta/2$.

1. Let $\theta \in [0, \pi/2]$ be the angle between the normal to X and the vertical axis. Let $S(X)$ be the surface area of X. Compute the area $S(\pi X)$ of the projection πX of X onto the horizontal plane P.
   Hint. The projection area may be written as

   $$\int_E l\left(\pi(X \cap (E^\perp + e))\right)\, de,$$

   where E is a line parallel to X and P.

2. Compute the probability that X is hit by G_1.

3. Consider the case where an isotropic random rotation r is applied to G_1. Compute the probability that X is hit by rG_1.

Solution to Exercise 1.

1. Since $l < \Delta$, the line-segment Z can be hit by at most one line in the grid G_1. If we let $1\{\cdot\}$ denote the indicator function, we therefore have

$$\begin{aligned}
&P(G_1 \text{ hits } Z|\Theta = \theta) \\
&= \sum_{j=-\infty}^{\infty} P(L_{\Theta,U+j\Delta} \text{ hits } Z|\Theta = \theta) \\
&= \sum_{j=-\infty}^{\infty} P(L_{\theta,U+j\Delta} \text{ hits } Z) \\
&= \sum_{j=-\infty}^{\infty} \int_0^{\Delta} 1\{L_{\theta,u+j\Delta} \text{ hits } Z\} \frac{du}{\Delta} \\
&= \sum_{j=-\infty}^{\infty} \int_0^{\Delta} 1\{(u+j\Delta)(\cos\theta, \sin\theta) \in \pi_\theta Z\} \frac{du}{\Delta} \\
&= \int_{-\infty}^{\infty} 1\{u(\cos\theta, \sin\theta) \in \pi_\theta Z\} \frac{du}{\Delta} \\
&= \frac{L(\pi_\theta Z)}{\Delta}.
\end{aligned}$$

It is easy to show that $L(\pi_\theta Z) = |\cos(\theta - \theta_0)| \cdot l$, where $l = L(Z)$ is the length of the line-segment.

2. Let $\theta_0 = 0$. Then,

$$\begin{aligned}
P(G_1 \text{ hits } Z) &= \int_0^{\pi} P(G_1 \text{ hits } Z|\Theta = \theta) \frac{d\theta}{\pi} \\
&= \int_0^{\pi} \frac{|\cos\theta| \cdot l}{\Delta} \frac{d\theta}{\pi} \\
&= \frac{2l}{\pi\Delta}.
\end{aligned}$$

Solution to Exercise 2. The only thing that has to be proved is that $\cos\Theta$ is uniform in $[0, 1]$, when Θ has probability density $p(\theta) = \sin\theta, \theta \in [0, \frac{\pi}{2})$. We find for $u \in [0, 1]$

$$\begin{aligned}
P(\cos\Theta \leq u) &= P(\Theta \geq \arccos u) \\
&= \int_{\arccos u}^{1} \sin\theta d\theta \\
&= \int_0^u dv \\
&= u.
\end{aligned}$$

Solution to Exercise 3.

1. The segment $X \cap (E^{\perp}+e)$ makes an angle equal to θ with the horizontal plane P. Therefore,

$$l\left(X \cap (E^{\perp}+e)\right) = l\left(X \cap (E^{\perp}+e)\right)\cos\theta.$$

The projection area can be now written as

$$\cos\theta \int_E l\left(X \cap (E^{\perp}+e)\right)\ de = \cos\theta S(X).$$

Hence, we get the result

$$S(\pi X) = \cos\theta S(X).$$

2. G_1 hits X if and only if it hits πX. The probability that G_1 hits πX is equal to the ratio $S(\pi X)/\Delta^2$. Therefore

$$P\left(G_1 \text{ hits } X\right) = \frac{\cos\theta S(X)}{\Delta^2}.$$

3. Let Θ be the random angle between the lines of rG_1 and the normal to X. Since r is isotropic random, Θ's distribution has density: $\sin\theta$, $\theta \in [0, \pi/2)$.

Given Θ, the conditional probability that rG_1 hits X is given by

$$P\left(rG_1 \text{ hits } X|\Theta\right) = \frac{S(X)}{\Delta^2}\cos\Theta.$$

The unconditional hitting probability is obtained by integration:

$$\begin{aligned} P\left(G_1 \text{ hits } X\right) &= \frac{S(X)}{\Delta^2}\int_0^{\pi/2}\cos\theta\sin\theta\ d\theta = \frac{S(X)}{2\Delta^2}\int_0^{\pi/2}\sin 2\theta\ d\theta \\ &= \frac{S(X)}{4\Delta^2}\int_0^{\pi/2} 2\sin 2\theta\ d\theta = \frac{S(X)}{2\Delta^2}. \end{aligned}$$

Stereological estimation of number

Exercise 1. This exercise concerns the systematic 2D-disector sampling design, applied to the population of objects shown below.

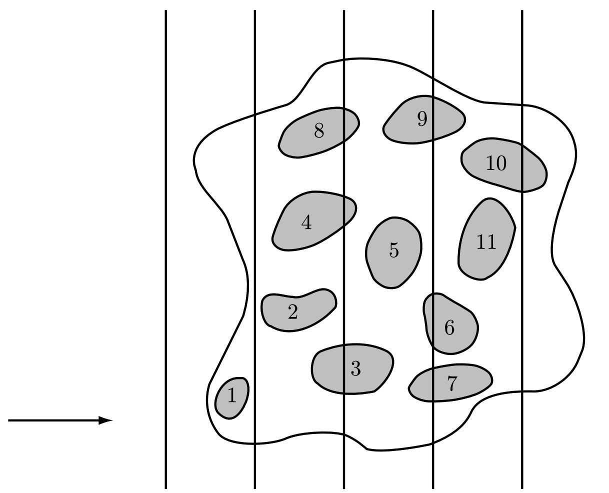
1. The cluster $\mathcal{P}_j$ consists of those objects first seen in the j'th strip (vertical linear band). The strips are ordered from left to right. Find $\mathcal{P}_j, j = 1, \ldots, 5$.

2. Systematic sampling of every second cluster is used, with a random start. Find the resulting two different samples of objects.

3. Find the distribution of $\hat{N}$.

4. What is the distribution of $\hat{N}$ if the ordering is reversed?

5. How can the two estimators of N be combined?



Exercise 2. Let $\mathcal{P} = \{1, \ldots, N\}$ be a finite population and let $\mathcal{S}$ be a random sample from the population. Suppose that all sampling probabilities are positive, i.e.

$$p_i = P(i \in \mathcal{S}) > 0, i \in \mathcal{P}.$$

Let

$$\hat{N} = \sum_{i \in \mathcal{S}} p_i^{-1}$$

be the Horvitz-Thompson estimator of N.

1. Show that $\hat{N}$ is unbiased for N, i.e..

$$E\hat{N} = N.$$

2. Show that the variance of $\hat{N}$ is given by

$$\mathrm{Var}\hat{N} = \sum_{i\in\mathcal{P},j\in\mathcal{P}} \frac{p_{i,j}}{p_i p_j} - N^2 = \sum_{i\in\mathcal{P}} \frac{1-p_i}{p_i} + \sum_{i\in\mathcal{P},j\in\mathcal{P},i\neq j} \frac{p_{i,j} - p_i p_j}{p_i p_j},$$

where $p_{i,j} = P(i \in \mathcal{S}, j \in \mathcal{S})$.

Solution to Exercise 1. With the ordering, indicated on the illustration, let $\mathcal{P}_j$ be the set of objects first seen in the j'th strip. Then,

$$\mathcal{P}_1 = \{1\}, \mathcal{P}_2 = \{2,3,4,8\}, \mathcal{P}_3 = \{5,6,7,9\}, \mathcal{P}_4 = \{10,11\}, \mathcal{P}_5 = \emptyset$$

The two possible samples become

$$\mathcal{S}_1 = \mathcal{P}_1 \cup \mathcal{P}_3 \cup \mathcal{P}_5 = \{1,5,6,7,9\} \text{ and } \mathcal{S}_2 = \mathcal{P}_2 \cup \mathcal{P}_4 = \{2,3,4,8,10,11\},$$

which have each probability 1/2 of being sampled. Since

$$\hat{N} = 2 \cdot \#S,$$

we get

$$P(\hat{N} = 10) = P(\hat{N} = 12) = 1/2.$$

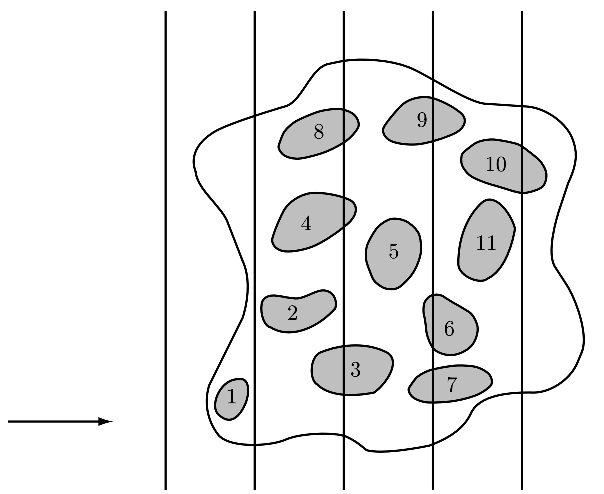
If the ordering is reversed the clusters become

$$\mathcal{P}_1 = \{10\}, \mathcal{P}_2 = \{6,7,9,11\}, \mathcal{P}_3 = \{3,4,5,8\}, \mathcal{P}_4 = \{2\}, \mathcal{P}_5 = \{1\}$$

Here, we also find

$$P(\hat{N} = 10) = \hat{P}(\hat{N} = 12) = 1/2.$$

A combined estimator gives the answer 11, irrespectively of the sample chosen.



Solution to Exercise 2.

1. If we let $1\{\cdot\}$ denote indicator function, we get

$$E\hat{N} = E \sum_{i \in \mathcal{P}} 1\{i \in \mathcal{S}\} p_i^{-1} = \sum_{i \in \mathcal{P}} P(i \in \mathcal{S}) p_i^{-1} = N.$$

2. For the first equality sign, it suffices to show that

$$E(\hat{N}^2) = \sum_{i\in\mathcal{P},j\in\mathcal{P}} \frac{p_{i,j}}{p_i p_j}.$$

We find

$$\begin{aligned} E(\hat{N}^2) &= E(\sum_{i\in\mathcal{P},j\in\mathcal{P}} 1\{i \in \mathcal{S}\}1\{j \in \mathcal{S}\}p_i^{-1}p_j^{-1}) \\ &= \sum_{i\in\mathcal{P},j\in\mathcal{P}} p_{i,j}p_i^{-1}p_j^{-1}. \end{aligned}$$

For the second equality sign, we use that $p_{i,i} = p_i$.

STEREOLOGICAL SAMPLING DESIGNS

Adrian Baddeley [1]
University of Western Australia
adrian@maths.uwa.edu.au

Notes from Sandbjerg lecture 2
1 June 2000

Introduction

Stereology [53, 54] originated as the problem of studying a three-dimensional physical object from random two-dimensional plane sections or projections, and in particular of determining geometrical parameters such as volume, surface area, length and total curvature. A typical result of classical stereology states that (under suitable conditions) the fraction of volume occupied by holes in Emmenthaler cheese can be statistically estimated from a random thin slice of cheese, by measuring the fraction of area occupied by holes.

There is a very strong analogy between stereology and the classical methods of survey sampling [4, 52] — the statistical estimation of properties of a population from observations made on random samples of the population. In order to estimate the porosity of Emmenthaler cheese by this method we do not need to know the spatial position of the slice of cheese, and indeed the position of the section plane must have been random in a specific sense. Hence this is a matter of statistical sampling inference, rather than computer tomographic reconstruction. Many applied scientists are surprised to learn that the volume of a three-dimensional object can be measured from plane sections without having to perform a three-dimensional reconstruction of its shape.

Modern stereology can be regarded as *sampling theory for spatial populations.* It embraces a wide class of 'geometrical random sampling' operations, such as clipping a two-dimensional image inside a window, taking one-dimensional linear probes, or sampling a spatial pattern at the points of a rectangular grid. The modern theory has links with stochastic geometry, spatial statistics, image analysis and empirical processes [1, 8, 12, 25, 50].

This lecture is a very brief introduction to stereological sampling designs. For more detail see [1, 8, 12, 25, 50, 53, 54, 55].

1 Basic concepts

1.1 Setting

A generic stereological sampling scheme is sketched in Figure 1. The experimental material is a given, fixed set $X \subset \mathbb{R}^d$ (the *'specimen'*) containing an unknown subset $Y \subset X$ (the *'feature of interest'*). We generate a random *probe* T intersecting X and we are able to observe the intersections $X \cap T$, $Y \cap T$ of the probe with X and Y.

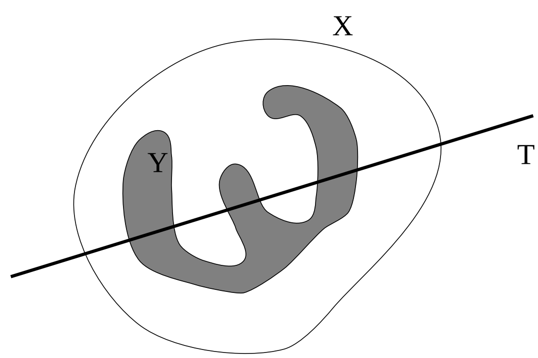


Figure 1: Basic scheme: specimen X and feature Y intersected by probe T

For example X may be a three-dimensional solid object and Y may be a solid component (or hole) within X, or a curved surface, a space curve, etc. The probe T could be a randomly-positioned two-dimensional section plane, a stack of parallel section planes, a sampling window of fixed size, a one-dimensional line probe, etc.

Our objective is to statistically estimate geometrical properties of Y (such as its volume or surface area) from the observed information in $X \cap T$, which includes $Y \cap T$. This is analogous to a survey sampling problem [4, 52]. We may regard the specimen X as a 'population' and the probe T as yielding a 'sample' $X \cap T$ from which we draw inferences.

The very close relation between stereology and survey sampling theory [4] was first emphasised by Miles and Davy [13, 14, 42, 43, 44, 45] and developed by Cruz Orive, Jensen, Gundersen and others [7, 11, 26, 28].

The goal of stereological estimation is typically one of the following quantities.

absolute size: the 'size' $\phi(Y)$ of the feature of interest Y, where ϕ is some appropriate geometrical measure of size. Examples are the total volume $V(Y)$ of neocortex Y in a human brain X available for sampling at autopsy, and the surface area $S(Y)$ of the gas exchange surface Y in a gazelle lung X.

relative size: the ratio $\phi(Y)/\psi(X)$ of 'sizes' of the feature of interest Y to the *reference space* X, where ϕ, ψ are possibly different measures of size. Examples are the volume fraction $V_V = V(Y)/V(X)$ of neocortex Y in brain X, and the surface area per unit volume $S_V = S(Y)/S(X)$ of the gas exchange surface Y in a gazelle lung X.

particle average size: if Y consists of discretely identifiable objects or 'particles' $Y_1, Y_2, \ldots$ then the parameter of interest may be the average of the individual particle values $\phi(Y_i)$ of a measure of 'size' ϕ, such as the particle volumes $v_i = V(Y_i)$, surface areas $s_i = S(Y_i)$, etc. and their higher moments if possible. Examples are the mean volume $\overline{v}_N = \sum_i V(Y_i)/N$ of individual air bubbles in neoprene rubber, and the volume-weighted mean volume $\overline{v}_V = \sum_i V(Y_i)^2 / \sum_i V(Y_i)$ of biological cell nuclei in cancer diagnostic samples. [The total number N of particles in a finite population of particles $Y_1, \ldots, Y_N$ is regarded as an 'absolute' quantity.]

Under the analogy with survey sampling, measures of absolute size $\phi(Y)$ are analogous to population totals, while ratios $\phi(Y)/\psi(X)$ are analogous to population means, where the 'population' is the specimen X. Particle averages are also analogous to population means, but with respect to the population of discrete particles.

In applications it is very important to distinguish carefully these three types of parameters, as their practical interpretations are quite different. Changes in the absolute size (volume, surface area) of a biological organ are relatively easy to interpret since the size is usually a measure of the functional capacity of the organ. An increase in the surface area of the lung's gas exchange surface implies an increased flow rate of oxygen into the blood. However an increase in the relative size (volume fraction, surface area fraction) of a feature Y relative to a reference space X may occur either because Y has increased in size, X has decreased in size, or both have increased in size with Y increasing proportionally more, or both have *decreased* in size with Y decreasing proportionally less. Some spectacular errors in biological science have been caused by misinterpretations of ratios.

1.2 Integral geometry

As we noted in the Introduction, many researchers are surprised to learn that the volume of a three-dimensional object can be measured from plane sections without having to perform a three-dimensional reconstruction of its shape. This is possible because of the volume of a three-dimensional object is the integral of the areas of its plane slices:

$$\int_{-\infty}^{\infty} A(Y \cap T_h)\, dh = V(Y) \tag{1}$$

where T_h is the plane $\{(x, y, z) : x = h\}$. This is a straightforward consequence of Fubini's theorem.

[On a historical note, (1) was one of the main motivating examples for the development of the integral calculus. An important half-way step in this development was the discovery of *Cavalieri's principle*, namely, that two solid objects which have equal plane sections (on all planes T_h, say) have equal volumes. Hence we could say it is Cavalieri's principle which enables stereology to work.]

Certain other geometrical parameters, such as the surface area of a curved surface in $\mathbb{R}^3$, can also be determined from plane sections or line probes. This relies on a class of identities analogous to (1) which have the general form

$$\int \alpha(Y \cap T)\, \mathrm{d}T = c\, \beta(Y) \tag{2}$$

where α, β are geometrical quantities and $c = c_{\alpha,\beta}$ is a constant. The integral is over all possible positions of the probe T, and $\mathrm{d}T$ is the appropriate 'uniform integration' over positions of T. Such results hold under minimal regularity conditions without regard to the 'shape' of the object Y, and hence these estimation techniques have very wide application. However there is a limited set of quantities β for which such results exist. The study of such representations is integral geometry [48, 55].

1.3 Sampling interpretations

Each integral-geometric formula of the form (2) has many stochastic interpretations, i.e. it can be applied in several different ways to set up a random sampling experiment and derive an estimator of the desired geometrical quantity $\beta(Y)$.

Consider the identity (1) relating the volume of a solid Y to the areas of its plane sections $A(Y \cap T_h)$ where again T_h denotes the plane with x-coordinate equal to h. Three possibilities are the following.

simple random sample: equation (1) implies that we can estimate the volume $V(Y)$ from a *single* randomly-chosen plane section. Let $[a, b]$ be the interval obtained by projecting Y onto the x-axis. Generate a random plane T by generating a random variable H uniformly distributed on the interval $[a, b]$ and taking $T = T_H$. Slice the object Y through the plane T, evaluate the section area $A(Y \cap T)$, and estimate $V(Y)$ by

$$\widehat{V} = (b-a)\, A(Y \cap T).$$

Then we have

$$\begin{aligned} \mathbb{E}[\widehat{V}\,] &= \int_a^b (b-a)\, A(Y \cap T_h)\, \frac{1}{b-a}\, \mathrm{d}h \\ &= V(Y) \end{aligned}$$

so that $\widehat{V}$ is an unbiased estimator of $V(Y)$.

systematic random sample: we can estimate the volume $V(Y)$ by measuring the areas of intersection of Y with a sequence of equally spaced, parallel, section planes:

$$\widehat{V} = \Delta \sum_m A(Y \cap T_{H+m\Delta})$$

where $\Delta > 0$ is a fixed spacing, the sum is over all integers m, and H is uniformly distributed over $[0, \Delta]$. This estimator can be viewed simply as a finite sum approximation to the integral of the function $f(x) = A(Y \cap T_x)$ over the real line, based on a sequence of equally-spaced sample points $(h,\ h + \Delta,\ h + 2\Delta,\ \ldots)$. However, note that the position of the sample is randomised, by assigning a value to h which is uniformly distributed over an interval of length Δ. Under these conditions the sample points form a systematic random sample of points on the real line. This method is a direct analogue of systematic sampling in finite populations. The estimator is unbiased because

$$\begin{aligned} \mathbb{E}[\widehat{V}\,] &= \int_0^\Delta \Delta \sum_m A(Y \cap T_{h+m\Delta})\, \frac{1}{\Delta}\, \mathrm{d}h \\ &= \sum_m \int_0^\Delta A(Y \cap T_{h+m\Delta})\, \mathrm{d}h \\ &= \sum_m \int_{m\Delta}^{(m+1)\Delta} A(Y \cap T_h)\, \mathrm{d}h \end{aligned}$$

$$
\begin{aligned}
&= \int A(Y \cap T_h)\,\mathrm{d}h \\
&= V(Y).
\end{aligned}
$$

In applied stereology this method is often called "the estimation of volume by Cavalieri's principle."

estimation of a ratio: as noted above, we may wish to estimate the volume fraction $V_V = V(Y)/V(X)$ of a solid Y inside another solid X. However there is an important caveat here. Recall that the expectation of a ratio of two random variables is generally not equal to the ratio of their expectations:

$$\mathbb{E}\left[\frac{A}{B}\right] \neq \frac{\mathbb{E}A}{\mathbb{E}B}$$

Hence if A, B are unbiased estimators of $V(Y), V(X)$ respectively, then the ratio A/B is typically a *biased* estimator of $V(Y)/V(X)$. This is analogous to the problem of *variable sample size* in survey sampling: the sample mean is of a uniform random sample is *not* an unbiased estimator of the population mean if the sample size is random.

In the stereological context, this problem was first pointed out by Miles and Davy [14, 44]. Suppose we take a single random plane section $T = T_H$ where H is uniformly distributed on an interval $[a, b]$ which contains the projections of X and Y onto the x-axis. Then the ratio of section areas $A(Y \cap T)/A(X \cap T)$ is a *biased* estimator of $V(Y)/V(X)$, because

$$\mathbb{E}\left[\frac{A(Y \cap T)}{A(X \cap T)}\right] = \frac{1}{b-a}\int \frac{A(Y \cap T_h)}{A(X \cap T_h)}\,\mathrm{d}h \neq \frac{\int_a^b A(Y \cap T_h)\,\mathrm{d}h}{\int_a^b A(X \cap T_h)\,\mathrm{d}h} = \frac{V(Y)}{V(X)}.$$

In order that $A(Y \cap T)/A(X \cap T)$ be an unbiased estimator of $V(Y)/V(X)$ we need to change the probability distribution of the section plane T. Let T have probability density proportional to the section area $A(X \cap T)$. Thus H has probability density

$$p(h) = \frac{A(X \cap T_h)}{\int A(X \cap T_z)\,\mathrm{d}z} = \frac{A(X \cap T_h)}{V(X)}$$

the *area-weighted* density. Then

$$\mathbb{E}\left[\frac{A(Y \cap T)}{A(X \cap T)}\right] = \int \frac{A(Y \cap T_h)}{A(X \cap T_h)}\frac{A(X \cap T_h)}{V(X)}\,\mathrm{d}h$$

$$= \int \frac{A(Y \cap T_h)}{V(X)} \, \mathrm{d}h$$
$$= \frac{V(Y)}{V(X)}.$$

This technique is analogous to the survey sampling method of *probability proportional to size (pps)* sampling. Miles and Davy [14, 44] developed the theory of stereological estimation using weighted random probes and showed how these could be implemented. This approach was adopted in Weibel's classic textbook [53, 54].

stationary random set: a quite different formulation of stereological methods is possible if the three-dimensional material has effectively infinite extent, and can be assumed to be "homogeneous" in a statistical sense. This might be appropriate for the study of samples of metal taken from a steel mill, or samples of volcanic rock.

Assume that Z is a random closed set in $\mathbb{R}^3$ (see [51]) which is *stationary* in the sense that its probability distribution is invariant under translations of $\mathbb{R}^3$. The *volume density* V_V of Z can be defined as the *expected* volume fraction of Z within any fixed reference region X of finite positive volume:

$$V_V = \frac{\mathbb{E}[V(Z \cap X)]}{V(X)}. \tag{3}$$

Stationarity implies that this ratio does not depend on X.

Now let S be a fixed, bounded subset of a two-dimensional plane in $\mathbb{R}^3$, having finite positive area $A(S)$. The area fraction $A_A = A(Z \cap S)/A(S)$ is an unbiased estimator of V_V, by the following argument. Without loss of generality assume S is a subset of the (y, z) coordinate plane and construct the prism $X = [0, 1] \times S$ with unit height on the base T. Applying (1) to the numerator of (3) we have

$$\begin{aligned} V_V &= \frac{1}{V(X)} \mathbb{E}\left[\int_0^1 A(Z \cap S_h) \, \mathrm{d}h\right] \\ &= \frac{1}{A(S)} \int_0^1 \mathbb{E}[A(Z \cap S_h)] \, \mathrm{d}h \end{aligned}$$

where $S_h = T_h \cap S = \{h\} \times S$ is the plane section of X at height h. Now S_h is the translation of S by a distance h parallel to S; by stationarity, $A(Z \cap S)$ has the same distribution as $A(Z \cap S_h)$, so $\mathbb{E}[A(Z \cap S_h)] =$

$\mathbb{E}[A(Z \cap S)]$. We obtain $V_V = \mathbb{E}[A(Z \cap S)]/A(S)$, so the estimator is unbiased.

This argument is similar to the original derivation by Delesse [15]. This approach to stereology is called *model-based* because it imposes model assumptions such as stationarity. It is the approach adopted in [51, chapter 10].

2 Sampling inference

Miles [42] first emphasised the importance of the correct specification of stereological sampling experiments. In this section we give details of how such specifications are made.

2.1 Formulations

Miles [42] distinguished three kinds of inference in stereology:

'restricted case': the specimen X and feature Y are non-random, bounded sets which are the sole object of interest (e.g. X is a whole organ or tumour);

'extended case': the specimen X available for examination is but a portion sampled from a much larger object W (e.g. a rock sample from a large rock outcrop);

'random case': the internal structure of the sample is a realization of a spatial random process. That is, the specimen X is a fixed set, but the feature Y inside X is generated as $Y = X \cap Z$ where Z is a random closed set (e.g. a sample from a continuous roll of steel formed under given conditions).

The restricted case corresponds to finite population survey inference, the extended case roughly to infinite population inference, and the random case to superpopulation inference.

These three sampling contexts are different with regard to their sampling requirements and the inferences which may be drawn from the sample, as Miles [42] explains. In the restricted case, we do not need to assume anything about the geometrical configuration of the feature Y and specimen X, but the probe T must adhere strictly to a random sampling protocol. The estimates of parameters $\phi(Y)$, $\phi(Y)/\psi(X)$ etc. obtained under these conditions are estimates of properties of the contents of the specimen X itself. This is "randomised design based survey sampling".

In the extended and random cases, assumptions of 'statistical homogeneity' are made about the contents of the specimen, and the position of the probe T is often irrelevant. The parameter estimates are not interpretable in terms of the specimen X; rather, they are estimates of averages over larger (super-)populations. Sampling variability (due to the random placement of T) is ignored, or conflated with other sources of variability.

Examples of the restricted case are the estimation of the total number and total volume of tumour cells in a tumour excised (in its entirety) from a patient; and estimation of the total length of glomerular tubules in the kidney of a laboratory rat. In both examples we need to estimate these quantities for the specific individual patient or animal. It would not be prudent to assume these biological structures exhibit any kind of spatial homogeneity.

Examples of the random case are the estimation of the volume fraction V_V of tungsten carbide in a composite abrasive material, and estimation of the surface area per unit volume S_V of metal grain surfaces in a certain type of steel. In both examples the parameter of interest is the *average* composition of the material. This implies that we average over any large-scale spatial variation in the characteristics of the material.

2.2 Definition and interpretation of parameters

These three inferential contexts also affect the definition and interpretation of stereological quantities.

Absolute geometrical quantities such as volume $V(Y)$ and surface area $S(Y)$ are well-defined only in the 'restricted' case. Under the analogy with survey sampling, absolute quantities correspond to population totals.

Ratios such as the volume fraction $V_V = V(Y)/V(X)$ occupied by Y within X, and the surface area per unit volume $S_V = S(Y)/V(X)$, are also well-defined only in the 'restricted' case. They correspond to population means in finite population sampling. However, in the 'extended' and 'random' cases, we can define analogous quantities called *densities*, which correspond to averages over an infinite population or superpopulation, respectively. Examples in the 'random' case are the mean volume fraction $V_V = \mathbb{E}[V(Y \cap X)]/V(X)$ and mean surface area fraction $S_V = \mathbb{E}[S(Y \cap X)]/V(X)$ of the random set Y. We need to assume that the underlying random set process Z is stationary so that these quantities do not depend on X and can genuinely be interpreted as average densities.

Particle average quantities are defined when Y consists of discretely identifiable objects or 'particles' $Y_1, Y_2, \ldots$. The averages of individual particle quantities $\phi(Y_i)$ over the particle population may be defined in all three settings. In the 'random' case we need the methods of stochastic geometry [51] to describe a stationary random process of particles (compact sets) in order to take an expectation over particles.

2.3 Containing space and reference space

Baddeley and Gundersen [3] make the further distinction between a *containing space* and a *reference space*. A *containing space* serves simply as a 'container' X which we sample at random in order to obtain random samples of the feature of interest Y. This is often the case when we are estimating an absolute geometrical quantity. The containing space is required to completely contain Y, but its size and extent need not be known exactly. For example, in order to estimate the total volume of a tumour, it suffices to remove enough tissue from the patient to ensure that the entire tumour has been removed, and then to sample (correctly!) from this excised tissue.

In contrast, a *reference space* is a well-defined object X of known or measurable size $\psi(X)$. The size $\phi(Y)$ of the feature of interest Y is expressed relative to X using the ratio $\phi(Y)/\psi(X)$. The reference space X need not contain the feature Y, although it often does. The boundary of the reference space must be clearly defined and its size $\psi(X)$ must be known exactly, in order to form the ratio $\phi(Y)/\psi(X)$. For example the volume fraction $V_V = V(Y)/V(X)$ of neocortex Y in brain X is well-defined (and comparable between different publications) only if the spatial extent of the brain is defined unambiguously (for example we must clearly define the boundary between brain and spinal cord).

2.4 Overview of sampling designs

The early development of stereology focused on the estimation of ratios and densities such as V_V and S_V in the 'extended' and 'random' cases. Some very elegant and simple estimators exist in this setting, and they are the standard methods in materials science.

In biological applications, the 'restricted' case is the natural setting, because biological structures have finite extent and cannot be assumed to be spatially homogeneous. The development of a valid statistical basis for stereology in this context was hindered by the fact that many of the simplest stereological sampling operations (such as taking a single plane section through an organ) have variable sample size. Fixed sample size is obtainable using a sampling window or 'quadrat' [7, 11, 13, 43, 45] but only in the extended or random cases. Variable sample size hampers the estimation of population means such as V_V, S_V because the sample mean is not an unbiased estimator of the population mean if the sample size is random. Miles and Davy [14, 44] were the first to point this out, and statistical research in the 1980's gave considerable attention to this problem [13, 26]. Miles and Davy [14, 44] also suggested the use of *size-weighted sampling designs* which have probability proportional to sample size (analogous to pps sampling in finite population survey sampling) and which do yield unbiased estimators of ratios such as V_V, S_V. This approach to stereological sampling is presented in Weibel's classic book [53, 54].

Systematic random samples are also possible, in the form of a stack of equally-spaced parallel section planes, a regular grid of test lines, etc. The great importance of systematic sampling for estimating population totals $V(Y), S(Y)$ was not realised until somewhat later [10, 7, 21, 26]. The modern theory of stereological estimation depends very much on the use of systematic samples.

In classical survey sampling, it is possible to use sampling schemes which have nonuniform sampling probabilities and variable sample size, provided that when we form an estimator, the contribution from each unit in the sample is weighted by the reciprocal of its sampling probability. This is the well-known device of Horvitz and Thompson [24]. A very similar device can be used in stereology to derive unbiased estimators of geometrical parameters from nonuniform sampling designs, including plane sections through a fixed point, line probes through a fixed point, and plane sections constrained to be perpendicular to a given reference plane. This is the field of *local stereology* [28].

3 Integralgeometric identities

3.1 Overview

As discussed in section 1.2, our ability to estimate geometrical parameters by stereological methods rests on the *section formulae* of integral geometry [48, 55]. These have the general form

$$\int \alpha(Y \cap T)\, \mathrm{d}T = c\, \beta(Y) \tag{4}$$

where α, β are geometrical quantities and $c = c_{\alpha,\beta}$ is a constant. The integral is over all possible positions of the probe T, and $\mathrm{d}T$ is the appropriate 'uniform integration' over positions of T. Such results hold under minimal regularity conditions without regard to the 'shape' of the object Y, and hence these estimation techniques have very wide application. However there is a limited set of quantities β for which such results exist. Following is a thumbnail sketch of some of the most well-known examples.

The *mean content formulae* enable us to determine the k-dimensional content of a k-dimensional subset Y in $\mathbb{R}^d$ (where $0 < k \leq d$) by integrating over all m-dimensional section planes T, where $d - k \leq m < d$. The intersection $Y \cap T$ has dimension $m + k - d$ generically. (In the example above we had $d = 3$, $m = 2$ and $k = 3$ so that $m + k - d = 2$.) If μ_n denotes n-dimensional content (Hausdorff measure), the formulae assert that

$$\int_{m\text{-planes}} \mu_{m+k-d}(Y \cap T)\, \mathrm{d}T = c_{k,m,d}\, \mu_k(Y)$$

where the integral is over all m-dimensional planes T in $\mathbb{R}^d$, and $\mathrm{d}T$ denotes 'uniform integration' over all T in an appropriate sense. Again $c_{k,m,d}$ is a geometrical constant.

In three dimensions, the mean content formulae state that:

- the volume of a solid can be determined from the areas of its plane sections ($m = 2, k = 3$);
- the surface area of a curved surface in $\mathbb{R}^3$ can be determined from the lengths of its intersection curves with section planes ($m = 2, k = 2$);
- the length of a curve in space can be determined from the number of intersection points it makes with section planes ($m = 2, k = 1$);
- the volume of a solid can be determined from the lengths of its intersections with straight line probes ($m = 1, k = 3$);
- the surface area of a curved surface in $\mathbb{R}^3$ can be determined from the number of times it meets a straight line probe ($m = 1, k = 2$).

In the rest of this section we spell out the details of these mean content formulae.

3.2 Plane sections in $\mathbb{R}^3$

The position of a two-dimensional plane in $\mathbb{R}^3$ is determined by its unit normal vector u and its distance s from the origin,

$$T_{u,s} = \left\{ x \in \mathbb{R}^3 \,:\, x \cdot u = s \right\}$$

where $\cdot$ denotes inner product of vectors. It is convenient to take the unit normal u to lie on the hemisphere $S^2_+ = \{(x, y, z) : x^2 + y^2 + z^2 = 1, z \geq 0\}$

and to allow signed distances $s \in \mathbb{R}$. It turns out [48] that the appropriate uniform measure for integrating over all planes T is $\mathrm{d}T = \mathrm{d}u \ \mathrm{d}s$ where $\mathrm{d}u$ is the usual uniform (area) measure on the hemisphere, with total value 2π. This measure is invariant under translations and rotations of $\mathbb{R}^3$.

Integrating (1) uniformly over all orientations u we obtain

$$\int_{\text{planes}} A(Y \cap T)\,\mathrm{d}T = 2\pi V(Y) \tag{5}$$

where the integral is now over all planes in $\mathbb{R}^3$. This states that the volume of a set Y (satisfying minimal conditions for measurability) can be determined by integrating over all planes T in $\mathbb{R}^3$ the area of intersection between Y and T. This obviously has direct application to stereology.

Other results are more surprising. Let S be a two-dimensional curved surface in $\mathbb{R}^3$ of finite surface area $A(S)$, satisfying certain regularity and rectifiability conditions. Then

$$\int_{\text{planes}} L(S \cap T)\,\mathrm{d}T = \frac{\pi^2}{2} A(S) \tag{6}$$

where $L(S \cap T)$ is the length of the curve $S \cap T$ of intersection between the surface S and the plane T. Applications of this formula allow the area of a curved surface to be statistically estimated from the lengths of plane section curves.

Let C be a (one-dimensional) space curve in $\mathbb{R}^3$ of finite length $L(C)$, satisfying certain regularity and rectifiability conditions. Then

$$\int_{\text{planes}} n(C \cap T)\,\mathrm{d}T = \pi L(C) \tag{7}$$

where $n(C \cap T)$ is the number of points of intersection between the curve C and the plane T. Applications of this formula allow the length of a curved filament to be statistically estimated from the number of crossings it makes with plane sections.

3.3 Line probes in $\mathbb{R}^3$

Instead of two-dimensional plane sections, we may consider random probes of other types. A one-dimensional straight line probe can be used to estimate the area of a curved surface or the volume of a solid.

To determine the position of a one-dimensional infinite straight line L in $\mathbb{R}^3$ we specify its direction vector u (a unit vector parallel to L) and its vector displacement x from the origin. We may take x to be perpendicular to u so that it lies in the two-dimensional plane $u^\perp$ perpendicular to u. The appropriate uniform measure for integration over all lines in $\mathbb{R}^3$ turns out to be $\mathrm{d}L = \mathrm{d}x\ \mathrm{d}u$ where $\mathrm{d}u$ is uniform integration over the sphere as before, and $\mathrm{d}x$ is uniform integration over the plane $u^\perp$. [Note the set of all pairs (u, x) with $x \in u^\perp$ is not a Cartesian product, so this statement only makes sense when we integrate over $x \in u^\perp$ with u fixed in the innermost integral.]

The volume $V(Y)$ of a measurable subset $Y \subset \mathbb{R}^3$ can be recovered from

$$\int_{\text{lines}} \ell(Y \cap L)\,\mathrm{d}L = 2\pi V(Y) \tag{8}$$

where $\ell(Y \cap L)$ is the length of the intersection $Y \cap L$ between the feature Y and the line probe L. This is a straightforward consequence of Fubini's theorem. Applications of this formula allow volumes to be statistically estimated from intersections with systems of test lines.

Let S be a two-dimensional curved surface in $\mathbb{R}^3$ of finite surface area $A(S)$, satisfying certain regularity and rectifiability conditions. Then

$$\int_{\text{lines}} n(S \cap L)\,\mathrm{d}L = \pi A(S) \tag{9}$$

where $n(S \cap L)$ is the number of intersection points between the surface S and the line probe L. Applications of this formula allow the area of a curved surface to be statistically estimated from its intersections with systems of test lines.

4 Stochastic interpretations

As explained in section 1.3 there are several possible random sampling interpretations of an integral geometric identity (4). The interpretation depends on the desired type of sampling inference, following Miles' trichotomy described in section 2.1.

Four 'template' sampling schemes were sketched in section 1.3 for the special case of estimating volume from plane section area:

- uniform sampling
- systematic sampling
- size-weighted sampling (for estimating a ratio)
- arbitrary sample of a stationary random set.

In the general context, these four templates operate as follows.

4.1 Uniform sampling

Here X and Y are fixed, bounded subsets of $\mathbb{R}^d$ with $Y \subset X$ and the extent of X is known. Thus X is both a containing space and a reference space (see section 2.3). We generate a random probe T intersecting X with the *uniform* probability distribution which has probability element

$$\mathrm{d}P = \frac{1}{\mu(X)}\,\mathrm{d}T \tag{10}$$

with $\mathrm{d}T$ denoting uniform integration in the sense of the appropriate integral geometric identity (4), and where

$$\mu(T) = \int \mathbf{1}\left\{T \cap X \neq \emptyset\right\}\,\mathrm{d}T$$

is the measure of all positions of the probe T in which T intersects X. Then

$$\mathbb{E}\;\alpha(Y \cap T) = \frac{\beta(Y)}{\mu(X)} \tag{11}$$

so that $\mu(X)\alpha(Y \cap T)$ is an unbiased estimator of $\beta(Y)$.

For example, if T is a two-dimensional plane in $\mathbb{R}^3$, then T is uniformly distributed in the sense of (10) if the (direction, distance) parameters (u, s) described in Section 3.2 are jointly uniformly distributed over the set

$$\left\{(u, s) \in S^2_+ \times \mathbb{R}\;:\;T_{u,s} \cap X \neq \emptyset\right\}.$$

This is called an **isotropic, uniformly random (IUR)** plane.

Note carefully that an IUR plane does *not* have a marginally uniformly distributed orientation u.

4.2 Systematic sampling

Here X, Y are fixed, bounded subsets of $\mathbb{R}^d$ with $Y \subset X$. We only need X to serve as a containing space (so that the exact dimensions of X do not need to be known).

Systematic sampling is easier to implement than IUR sampling in almost all applications, and simpler to justify, since the variability of the sample size does not present difficulties. However, a general definition of systematic sampling in an abstract setting is mathematically involved. See [48] for some general integral-geometric identities for systematic arrays of geometrical objects. A systematic array can be defined as the image of a single geometrical object (such as a line or plane) under a finitely-generated group of Euclidean rigid transformations. The resulting integral geometric identities have the same general form as (4), except that T is a systematic array of planes, lines, etc., and the total measure $\int \mathrm{d}T = \kappa$ is a known, finite geometrical constant.

The prime example is that of serial section planes in $\mathbb{R}^3$. Defining $T_{u,s} = \{x \in \mathbb{R}^3 \,:\, x \cdot u = s\}$ as before, a serial section stack of constant spacing $\Delta > 0$ is of the form

$$S_{u,s} = \{T_{u,s+m\Delta} \,:\, m \in \mathbb{Z}\}$$

for $u \in S^2_+$ and $s \in [0, \Delta]$. The total measure of all serial section stacks is $\kappa = \int \mathrm{d}s\,\mathrm{d}u = 2\pi\Delta$. An isotropic, uniformly random (IUR) serial section stack is generated by taking u and s to be independent and uniformly distributed over S^2_+ and $[0, \Delta]$ respectively. The identities (5)–(7) yield respectively

$$\begin{aligned} \mathbb{E}\left[\sum_m A(Y \cap T_m)\right] &= \frac{1}{\Delta} V(Y) \\ \mathbb{E}\left[\sum_m L(S \cap T_m)\right] &= \frac{\pi}{4\Delta} A(S) \\ \mathbb{E}\left[\sum_m n(C \cap T_m)\right] &= \frac{1}{2\Delta} L(C) \end{aligned}$$

in the same notation, where $T_m = T_{u,s+m\Delta}$. Thus an IUR serial section stack enables us to estimate volume, surface area and length knowing only the spacing Δ between the sections.

Other commonly-used examples include a rectangular grid of points in $\mathbb{R}^2$ (a "test point grid"), a sequence of equally spaced parallel lines in $\mathbb{R}^2$ (a "test line grid") and a systematic array of parallel lines in $\mathbb{R}^3$ (a "fakir's bed").

4.3 Weighted sampling

The setting is the same as in section 4.1, except that we exploit *two* identities of the form (4) for pairs of quantities (α, β) and (α', β') in order to estimate the ratio $\beta'(Y)/\beta(X)$. We generate a random probe T intersecting X with the *α-weighted* probability distribution which has probability element

$$\mathrm{d}P = \frac{\alpha(X \cap T)}{c\beta(X)}\,\mathrm{d}T \tag{12}$$

where $c = c_{\alpha,\beta}$ is the geometrical constant appearing in (4). Then

$$\mathbb{E}\left[\frac{\alpha'(Y \cap T)}{\alpha(X \cap T)}\right] = b\frac{\beta'(Y)}{\beta(X)} \tag{13}$$

where $b = c_{\alpha',\beta'}/c_{\alpha,\beta}$ is another geometrical constant. Thus,

$$\frac{1}{b}\ \frac{\alpha'(Y \cap T)}{\alpha(X \cap T)}$$

is an unbiased estimator of the ratio $\beta'(Y)/\beta(X)$.

As explained in section 1.3, this is analogous to the survey sampling method of *probability proportional to size (pps)* sampling. Miles and Davy [14, 44] developed the theory of stereological estimation using weighted random probes and showed how these could be implemented. This approach was adopted in Weibel's classic textbook [53, 54].

The connection with ratio estimation was developed and it was later realised that a more practical alternative is to invoke (13) using the ratio-of-sums estimator

$$\frac{\sum_{i=1}^{m} \alpha(Y \cap T_i)}{\sum_{i=1}^{m} \alpha'(X \cap T_i)}$$

based on a sufficient number of independent uniform samples $T_1, \ldots, T_m$.

4.4 Arbitrary sample of homogeneous material

In the 'random' and 'extended' cases we adopt a completely different approach to sampling. The feature of interest Y is taken to be a random set in $\mathbb{R}^d$ which is assumed to be statistically homogeneous. The probe T may as well be a fixed k-dimensional subset of $\mathbb{R}^d$ with known, bounded extent. Then we obtain interpretations of (4) in the form

$$\frac{\mathbb{E}\alpha(Y \cap T)}{\mu_k(T)} = c\frac{\mathbb{E}\beta(Y \cap X)}{V(X)} \tag{14}$$

where X is an arbitrary solid in $\mathbb{R}^3$, and on both sides $\mathbb{E}$ denotes expectation with respect to the distribution of Y. Stationarity implies that the right hand side of (14) does not depend on X and equals a desired parameter of the distribution of Y (a density or 'superpopulation average'). Again c is a geometrical constant and $\mu_k(T)$ denotes the k-dimensional content (Hausdorff measure) of T.

Differences and links between design-based and model-based inference are discussed in [1, 25, 49, 50]. Biological applications typically require a design-based approach, e.g. because biological organs are highly organised; materials science applications often lend themselves to a model-based approach.

5 Strategies for designing stereological experiments

`This section of the notes is fragmentary at the moment.`

In practical applications of stereology, the sampling protocol must be designed with a view to ensuring the validity of the estimators and to minimising the variance contribution from sampling variation.

This section describes some of the main strategies which can be used to develop good sampling protocols. We refer only to the 'restricted' case where X, Y are fixed bounded subsets of $\mathbb{R}^3$ and $Y \subset X$.

5.1 Stratification

Stratification of a population is a standard technique for reducing variance in survey sampling. In stereology this method can be applied when the quantity of interest is an absolute geometrical quantity $\phi(Y)$ which is additive in the sense that $\phi(Y_1 \cup Y_2) = \phi(Y_1) + \phi(Y_2)$ when Y_1, Y_2 are disjoint.

We may stratify the specimen X by physically dividing X into disjoint pieces $X_1, \ldots, X_K$ (with consequent unseen division of Y into pieces $Y_i = Y \cap X$). Then we treat each piece X_i as a separate specimen and estimate the desired *absolute* geometrical property $\phi(Y_i)$ by sampling X_i. Finally we form an estimate of $\phi(Y) = \sum_{i=1}^{K} \phi(Y_i)$ by summing the estimates of each $\phi(Y_i)$.

Advantages of stratification include the ability to sample with different intensities or sample sizes in each subpopulation X_i, and the fact that it allows us to randomise over the section orientation (the common orientation of the section planes through X_i is random and different in different pieces X_i).

5.2 Ratio estimation

In survey sampling theory, "ratio estimation" means the estimation of a population parameter θ by a two-stage process in which we first estimate another parameter η that is easier to estimate, and then estimate the ratio θ/η of the two parameters by different means.

Ratio estimation has many applications in biological stereology because of the widely differing scales of organisation of biological structures. Suppose we wish to estimate the total volume $V(Y)$ of the cortex Y of a kidney X. Write

$$V(Y) = \frac{V(Y)}{V(X)} \times V(X).$$

We can estimate $V(X)$ macroscopically, for example by measuring the volume of fluid displaced by the kidney, or by using "Cavalieri's principle" at low magnification. Then taking plane sections of the kidney and using higher magnification we can estimate $V_V = \frac{V(Y)}{V(X)}$ by test point counting.

Ratio estimation can be applied repeatedly at multiple scales. Suppose we wish to determine the total volume of the tubules Z in the cortex Y of a kidney X. We can write

$$V(Z) = \frac{V(Z)}{V(Y)} \times \frac{V(Y)}{V(X)} \times V(X);$$

estimating $V(X)$ and $\frac{V(Y)}{V(X)}$ as before, we can estimate the tubule/cortex fraction $V_V = \frac{V(Z)}{V(Y)}$ at higher magnification, with test point counting.

5.3 Nesting

The typical stereological experiment is a nested design involving several levels of subsampling [11, 18]. For example, in many biological applications, organs are taken from each of several animals; several tissue blocks are sampled from each organ; several thin sections are cut from each block; and several sampling windows are photographed on each section.

5.4 Granularity and dimension of probes

Additionally one can choose from a variety of 'test probes' of different geometries and dimensions. For example the volume of a three-dimensional object can be estimated using random 2-dimensional plane sections, random 1-dimensional linear probes, or random test points. There is a lively discussion over their relative merits, particularly because low-dimensional probes can be performed manually and quickly, whereas higher-dimensional probes require computer image processing [19, 20, 40].

6 Information about variances

`This is very rough at the moment`

It is difficult to make general statements about the variance or efficiency of stereological estimators, and this is still a matter of controversy [13, 16, 17, 27]. Some progress has been made on the variance of systematic sampling [9, 22, 30, 37, 41]. `See Kiên Kiêu's lectures for up-to-date information and references.`

Cruz Orive [6] developed a theory of best linear unbiased estimation of ratios in stereology, which however appears to be successful only for fixed sample size designs [27].

6.1 Variance components in nested designs

The typical stereological experiment is a nested design involving several levels of subsampling [11, 18].

Standard techniques of nested ANOVA can be applied to estimate variances and variance contributions empirically. In biological experiments, the results of ANOVA frequently indicate that the between-animal variance is substantially larger than other sources of variation. The optimal allocation of sampling effort is then to sample a large number of animals and spend relatively little effort on measuring the data in each sampling window ("do more less well" [11, 18]).

6.2 Efficiency of systematic sampling

`See Kiên's talk.`

Point count estimators using a random test grid have been much studied [7, 23, 26, 31, 32, 45], [47, chap. 3], [35, 36, chap. 3] and recent work [9, 22, 30, 37, 41] has shown that they are very accurate in most situations.

6.3 Rao-Blackwell

A typical issue [18, 19, 20, 40] concerns the relative efficiency of the quadrat and point-counting estimators of area fraction in the plane.

Intuitively one expects a point-counting estimator to have higher variance than a corresponding quadrat sample estimator, because it is based on a subsample. Davy and Miles [14, sec. 6] proved that, for the standard IUR and WUR sampling probes, estimators based on lower-dimensional probes have higher variances. However, Jensen & Gundersen [26, sec. 6] constructed an example in which point counting is *more* efficient than quadrat sampling. Ohser [46] and Baddeley & Cruz [2] noted that the length density of a random line process can sometimes be estimated more efficiently by counting the number of intersections with a test line grid than by measuring lengths inside a sampling quadrat.

The variance result of Miles and Davy is in fact an instance of the Rao-Blackwell theorem, in a stereological version which we have proved in [2] (see also [33, 34]). Similar results were stated by Lantuéjoul [33] as a consequence of "Cartier's formula" $\mathbb{E}[U|V] = V$ whenever this holds for real random variables U, V. Jensen and Gundersen recently showed [29, eq. (5.1)] that the 'nucleator' and 'rotator' estimates of mean particle volume are related by the Rao-Blackwell process.

The comparison of variances is related to the 'change of support' problem in geostatistics [5, §5.2, pp. 284–289], [33, 38, 39].

References

[1] A.J. Baddeley. Stereology. In *Spatial Statistics and Digital Image Analysis*, chapter 10, pages 181–216. National Research Council USA, Washington DC, 1991.

[2] A.J. Baddeley and L.M. Cruz-Orive. The Rao-Blackwell theorem in stereology and some counterexamples. *Advances in Applied Probability*, 27:2–19, 1995.

[3] A.J. Baddeley and H.J.G. Gundersen. *Stereology: general principles and biological applications.* Publisher not determined, 2001.

[4] W. G. Cochran. *Sampling Techniques.* John Wiley and Sons, 3rd edition, 1977.

[5] N.A.C. Cressie. *Statistics for spatial data.* John Wiley and Sons, New York, 1991.

[6] L.M. Cruz-Orive. Best linear unbiased estimators for stereology. *Biometrics*, 36:595–605, 1980.

[7] L.M. Cruz-Orive. The use of quadrats and test systems in stereology, including magnification corrections. *Journal of Microscopy*, 125:89–102, 1982.

[8] L.M. Cruz-Orive. Stereology: recent solutions to old problems and a glimpse into the future. *Acta Stereologica*, 6/III:3–18, 1987.

[9] L.M. Cruz-Orive. On the precision of systematic sampling: a review of Matheron's transitive methods. *Journal of Microscopy*, 153:315–333, 1989.

[10] L.M. Cruz-Orive and A.O. Myking. Estimation of volume ratios by systematic sections. *Journal of Microscopy*, 122:143–157, 1981.

[11] L.M. Cruz-Orive and E.R. Weibel. Sampling designs for stereology. *Journal of Microscopy*, 122:235–257, 1981.

[12] L.M. Cruz-Orive and E.R. Weibel. Recent stereological methods for cell biology: a brief survey. *American Journal of Physiology*, 258:L148–L156, 1990. Lung Cell. Mol. Physiol. 2.

[13] P.J. Davy. *Stereology: a statistical viewpoint.* Ph.D. thesis, Australian National University, 1978.

[14] P.J. Davy and R.E. Miles. Sampling theory for opaque spatial specimens. *Journal of the Royal Statistical Society, series B*, 39:56–65, 1977.

[15] A. Delesse. Procédé mécanique pour déterminer la composition des roches. *Comptes Rendues de l'Académie des Sciences (Paris)*, 25:544, 1847.

[16] A. Dhani. *Problèmes d'éstimation en stéréologie.* Mémoire, Licencié en sciences mathématiques, Faculté de Science, Facultés Universitaires de Notre Dame de la Paix, Namur, Belgium, 1979.

[17] A.S. Downie. *Efficiency of statistics in stereology.* Ph.D. thesis, Imperial College London, 1991.

[18] H.J. Gundersen and R. Østerby. Optimizing sampling efficiency of stereological studies in biology: or 'Do more less well!'. *Journal of Microscopy*, 121:65–74, 1981.

[19] H.J.G. Gundersen, M. Boysen, and A. Reith. Comparison of semiautomatic digitizer-tablet and simple point counting performance in morphometry. *Virchows Archiv B (Cell Pathology)*, 37:317–325, 1981.

[20] H.J.G. Gundersen, M. Boysen, and A. Reith. Digitizer-tablet or point counting in biomorphometry? In *Proceedings of the 3rd European Symposium on Stereology*, volume 3 (suppl.), pages 205–210, 1981.

[21] H.J.G. Gundersen et al. Some new, simple and efficient stereological methods and their use in pathological research and diagnosis. *Acta Pathologica Microbiologica et Immunologica Scandinavica*, 96:379–394, 1988.

[22] H.J.G. Gundersen and E.B. Jensen. The efficiency of systematic sampling in stereology and its prediction. *Journal of Microscopy*, 147:229–263, 1987.

[23] Peter Hall. *An introduction to the theory of coverage processes.* John Wiley and Sons, New York, 1988.

[24] D.G. Horvitz and D.J. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47:663–685, 1952.

[25] E.B. Jensen, A.J. Baddeley, H.J.G. Gundersen, and R. Sundberg. Recent trends in stereology. *International Statistical Review*, 53:99–108, 1985.

[26] E.B. Jensen and H.J.G. Gundersen. Stereological ratio estimation based on counts from integral test systems. *Journal of Microscopy*, 125:51–66, 1982.

[27] E.B. Jensen and R. Sundberg. Statistical models for stereological inference about spatial structures; on the applicability of best linear unbiased estimators in stereology. *Biometrics*, 42:735–751, 1986.

[28] E.B.V. Jensen. *Local Stereology*. Singapore. World Scientific Publishing, 1997.

[29] E.B.V. Jensen and H.J.G. Gundersen. The rotator. *Journal of Microscopy*, 170:35–44, 1993.

[30] A.M. Kellerer. Exact formulae for the precision of systematic sampling. *Journal of Microscopy*, 153:285–300, 1989.

[31] D.G. Kendall. On the number of lattice points inside a random oval. *Quarterly Journal of Mathematics (Oxford)*, 19:1–26, 1948.

[32] D.G. Kendall and R.A. Rankin. On the number of points of a given lattice in a random hypersphere. *Quarterly Journal of Mathematics (Oxford)*, 4:178–189, 1953.

[33] Ch. Lantuéjoul. Some stereological and statistical consequences derived from Cartier's formula. *Journal of Microscopy*, 151:265–276, 1988.

[34] E.L. Lehmann. *Theory of point estimation*. John Wiley and Sons, New York, 1983.

[35] B. Matérn. Spatial variation. *Meddelanden från Statens Skogsforskningsinstitut*, 49(5):1–114, 1960.

[36] B. Matérn. *Spatial Variation*. Number 36 in Lecture Notes in Statistics. Springer Verlag, New York, 1986.

[37] B. Matérn. Precision of area estimation: a numerical study. *Journal of Microscopy*, 153:269–284, 1989.

[38] G. Matheron. *Les variables régionalisées et leur estimation.* Masson, Paris, 1965.

[39] G. Matheron. The intrinsic random functions and their applications. *Advances in Applied Probability*, 5:439–468, 1973.

[40] O. Mathieu, L.M. Cruz-Orive, H. Hoppeler, and E. R. Weibel. Measuring error and sampling variation in stereology: comparison of the efficiency of various methods for planar image analysis. *Journal of Microscopy*, 121:75–88, 1981.

[41] T. Mattfeldt. The accuracy of one-dimensional systematic sampling. *Journal of Microscopy*, 153:301–313, 1989.

[42] R.E. Miles. The importance of proper model specification in stereology. In R E Miles and J Serra, editors, *Geometrical Probability and Biological Structures: Buffon's 200th Anniversary*, Lecture Notes in Biomathematics, No 23, pages 115–136, Berlin-Heidelberg-New York, 1978. Springer Verlag.

[43] R.E. Miles. The sampling, by quadrats, of planar aggregates. *Journal of Microscopy*, 113:257–267, 1978.

[44] R.E. Miles and P.J. Davy. Precise and general conditions for the validity of a comprehensive set of stereological fundamental formulae. *Journal of Microscopy*, 107:211–226, 1976.

[45] R.E. Miles and P.J. Davy. On the choice of quadrats in stereology. *Journal of Microscopy*, 110:27–44, 1977.

[46] J. Ohser. *Grundlagen und praktische Möglichkeiten der Charakterisierung struktureller Inhomogenitäten von Werkstoffen.* Dr. sc. techn. thesis, Bergakademie Freiberg, Freiberg (Sachsen), Germany, 1990.

[47] B.D. Ripley. *Spatial statistics.* John Wiley and Sons, New York, 1981.

[48] L. A. Santaló. *Integral Geometry and Geometric Probability.* Encyclopedia of Mathematics and Its Applications, vol. 1. Addison-Wesley, 1976.

[49] C.E. Särndal. Design-based and model-based inference in survey sampling (with discussion). *Scandinavian Journal of Statistics*, 5:27–52, 1978.

[50] D. Stoyan. Stereology and stochastic geometry. *International Statistical Review*, 58:227–242, 1990.

[51] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. John Wiley and Sons, Chichester, second edition, 1995.

[52] S.K. Thompson. *Sampling*. John Wiley and Sons, 1992.

[53] E.R. Weibel. *Stereological Methods, 1. Practical Methods for Biological Morphometry*. Academic Press, London, 1979.

[54] E.R. Weibel. *Stereological Methods, 2. Theoretical Foundations.* Academic Press, London, 1980.

[55] W. Weil. Stereology: a survey for geometers. In P M Gruber and J M Wills, editors, *Convexity and its applications*, pages 360–412. Birkhauser, Basel, Boston, Stuttgart, 1983.

Eva B. Vedel Jensen: Supplementary notes on Local Stereology and its relation to Geometric Tomography

A simple example — local estimation of planar area

Let K be a compact subset of the Euclidean plane $\mathbb{E}^2$. Let us suppose that we want to estimate its area $A(K)$ by local stereological methods.

For this purpose, let l_θ be the random line through o, the origin, that makes a uniform angle $\theta \in [0, \pi)$ with a fixed axis. Let us find the probability that the line hits an infinitesimal element dx of K. We will use the well-known transformation from polar coordinates to Cartesian coordinates

$$(r, \theta) \in \mathbb{R} \times [0, \pi) \to (r \cos\theta, r \sin\theta).$$

We have

$$\lambda_2(dx) = |r| dr d\theta,$$

where $\lambda_2(dx)$ is the area (2-dimensional Lebesgue measure) of dx. Therefore, the probability that the line hits dx is

$$\frac{d\theta}{\pi} = \frac{\lambda_2(dx)}{\pi |r| dr}.$$

Having determined the sampling probabilities the Horvitz-Thompson procedure can be applied. The estimator of $A(K)$ becomes a sum over those infinitesimal elements of K which are hit by l_θ and is given by

$$\begin{aligned} &\sum \lambda_2(dx) / \frac{d\theta}{\pi} \\ &= \sum \lambda_2(dx) \pi |r| dr / \lambda_2(dx) \\ &= \pi \sum |r| dr. \end{aligned}$$

The corresponding continuous version will be used as an estimator of $A(K)$,

$$\widehat{A}(K) = \pi \int_{K \cap l_\theta} d(x, o) \lambda_1(dx).$$

Here, we integrate along the intersection between K and the random line l_θ, $d(x, o)$ is the distance from x to the origin o and λ_1 is the 1-dimensional Lebesgure measure.

In what follows, we will discuss local stereological volume estimators in $\mathbb{E}^n$. In Section 2, we extend the definitons of chord functions and section functions, known from geometric tomography. In Section 3, we derive the local stereological volume estimators and show that for a general class of bodies in $\mathbb{E}^n$ the estimators are proportional to section functions. In Section 4, we give a brief summary of local stereological volume estimators in $\mathbb{E}^3$.

Extensions of chord and section functions

Let K be a star-shaped (at o) body. The body K has thus the property that its intersection with any line through o consists of one line segment.

Let $\rho_K(u), u \in S^{n-1}$, be the restriction of the radial function of K to S^{n-1}. If we let l_u be the line through o with direction $u \in S^{n-1}$, the radial function $\rho_K(u)$ is given by

$$\rho_K(u) = \begin{cases} \max\{c : cu \in K\} & \text{if } l_u \cap K \neq \emptyset \\ 0 & \text{if } l_u \cap K = \emptyset. \end{cases}$$

Note that if l_u hits K, then $\rho_K(u)$ is the maximal signed distance in the direction u from o to the boundary of K.

The $i-$chord function $\rho_{i,K}$ of K at o is defined for $u \in S^{n-1}$ in terms of the radial function. If $l_u \cap K = \emptyset$, we let $\rho_{i,K}(u) = 0$. Otherwise, for $i \neq 0$

$$\rho_{i,K}(u) = \begin{cases} \rho_K(u)^i + \rho_K(-u)^i & \text{if } o \in K \\ ||\rho_K(u)|^i - |\rho_K(-u)|^i| & \text{if } o \notin K. \end{cases} \tag{1}$$

Below we will only consider this function for $i \in \{1, \dots, n\}$.

It turns out to be interesting to extend the i-chord function to not-necessarily star-shaped K. We extend the definition to bodies K in $\mathbb{E}^n$ which have the property that $K \cap l$ consists of a finite number of line segments for any line $l \in \mathcal{G}(n, 1)$. The set of bodies satisfying this property will be called the star ring and denoted by $\mathcal{S}(st)$.

($\mathcal{G}(n, 1)$ is the notation used for lines in $\mathbb{E}^n$ through o. More generally, $\mathcal{G}(n, k)$ is the notation used for $k-$dimensional linear subspaces of $\mathbb{E}^n$. A $k-$dimensional linear subspace is called a $k-$subspace below.)

For $K \in \mathcal{S}(st)$, let $\rho_{i,K}(u) = 0$ if $l_u \cap K = \emptyset$. Otherwise, $l_u \cap K$ consists of a non-empty finite union of line segments. Let e_u be the set of end points of the line segments with positive length. For $x \in e_u \setminus \{o\}$, let $\alpha(x)$ be the number of elements of e_u which are on the same side of o as x, but at a longer distance from o than x. If $o \in e_u$, we let $\alpha(o) = 0$. Then we define

$$\rho_{i,K}(u) = \sum_{x \in e_u} (-1)^{\alpha(x)} d(x, o)^i. \tag{2}$$

This is an extension of the i-chord function, defined for star-shaped bodies. To see this, let K be star-shaped and let $u \in S^{n-1}$ be chosen such that l_u hits K. Let us concentrate on the case where $l_u \cap K$ is a line segment of positive length.

If $o \in K$, then $e_u = \{x_+, x_-\}$ consists of two elements, both with α-value 0. The definition (2) yields

$$\rho_{i,K}(u) = d(x_+, o)^i + d(x_-, o)^i$$

which coincides with (1), upper case. If $o \notin K$, then $e_u = \{x_+, x_-\}$ consists of two elements, on the same side of o. Let $d(x_+, 0) > d(x_-, 0)$. Then, $\alpha(x_+) = 0, \alpha(x_-) = 1$ and (2) yields

$$\rho_{i,K}(u) = d(x_+, o)^i - d(x_-, o)^i$$

which coincides with (1), lower case.

It is also natural to extend the definition of section functions to $\mathcal{S}(st)$. Since the section functions are defined, using i-chord functions, the generalization is immediate. For $K \in \mathcal{S}(st)$ and $S \in \mathcal{G}(n,k), k = 1, \ldots, n-1$, we define the section function by

$$\widetilde{V}_{i,k}(K \cap S) = \frac{1}{2k} \int_{S^{n-1} \cap S} \rho_{i,K}(u) \lambda_{k-1}(du), \tag{3}$$

where $\rho_{i,K}$ is defined in (2).

Local stereological volume estimators in $\mathbb{E}^n$

Let K be a body in $\mathbb{E}^n$. The object is to estimate its volume ($n-$dimensional Lebesgue measure) $V(K)$, based on information in an isotropic $k-$subspace $S \in \mathcal{G}(n,k)$. Such a random subspace has as distribution the unique rotation invariant probability measure on $\mathcal{G}(n,k)$.

The developments will be centered around a version of the Blaschke-Petkantschin formula, cf. Jensen (1998, Proposition 4.5). For any non-negative Borel function g on $\mathbb{E}^n$, we have

$$\int_{\mathbb{E}^n} g(x) \lambda_n(dx) = \frac{\omega_n}{\omega_k} \int_{\mathcal{G}(n,k)} \int_S g(x) d(x,o)^{n-k} \lambda_k(dx) dS, \tag{4}$$

where $\lambda_n(dx)$ is the element of $n-$dimensional Lebesgue measure, $\omega_n = 2(\pi)^{\frac{1}{2}n} \Gamma(\frac{1}{2}n)^{-1}$ is the surface area of the unit sphere S^{n-1} in $\mathbb{E}^n$ and dS is the element of the unique rotation invariant probability measure on $\mathcal{G}(n,k)$.

In order to derive a local volume estimator, consider a volume element dx of K. Using (4), we find

$$\lambda_n(dx) = \frac{\omega_n}{\omega_k} d(x,o)^{n-k} \lambda_k(dx) dS. \tag{5}$$

Therefore, using the Horvitz-Thompson procedure on an infinitesimal level, we are led to consider the following estimator

$$\begin{aligned} & \sum \lambda_n(dx)/dS \\ &= \sum \lambda_n(dx) \frac{\omega_n}{\omega_k} d(x,o)^{n-k} \lambda_k(dx)/\lambda_n(dx) \\ &= \frac{\omega_n}{\omega_k} \sum d(x,o)^{n-k} \lambda_k(dx), \end{aligned}$$

where the sum is over those infinitesimal volume elements of K hit by the isotropic $k-$subspace S. We use the continuous form of this estimator

$$\widehat{V}_{n,k}(K \cap S) = \frac{\omega_n}{\omega_k} \int_{K \cap S} d(x, o)^{n-k} \lambda_k(dx). \tag{6}$$

For $k = 1$, (6) reduces to

$$\widehat{V}_{n,1}(K \cap l) = \frac{\pi^{n/2}}{\Gamma(n/2)} \int_{K \cap l} d(x, o)^{n-1} \lambda_1(dx), \tag{7}$$

where $l \in \mathcal{G}(n, 1)$. It is not difficult to see that for $K \in \mathcal{S}(st)$ and $u \in S^{n-1}$, cf. (2),

$$\widehat{V}_{n,1}(K \cap l_u) = \frac{\pi^{n/2}}{n\Gamma(n/2)} \rho_{n,K}(u). \tag{8}$$

For $K \in \mathcal{S}(st)$, the local stereological estimator of volume, based on information along a line, is thus proportional to the $n-$chord function.

The estimators based on subspaces of different dimensions are related by a conditional mean-value operation. For $k_1 \leq k_2$,

$$\widehat{V}_{n,k_2}(K \cap S) = \int_{\mathcal{G}(k_2,k_1)} \widehat{V}_{n,k_1}(K \cap T) dT, \tag{9}$$

where $\mathcal{G}(k_2, k_1)$ is the set of k_1-subspaces, contained in $S \in \mathcal{G}(n, k_2)$. The result (9) can be obtained by combining (4) and (6). Note that (9) implies that for $k_1 \leq k_2$

$$\mathrm{Var}\widehat{V}_{n,k_2}(K \cap S) \leq \mathrm{Var}\widehat{V}_{n,k_1}(K \cap T).$$

If we use (9) for $k_1 = 1$ and $k_2 = k$, we get

$$\widehat{V}_{n,k}(K \cap S) = \int_{\mathcal{G}(k,1)} \widehat{V}_{n,1}(K \cap l) dl.$$

Using that the rotation invariant probability measure on $\mathcal{G}(k, 1)$ can be constructed by lifting the normalized Hausdorff measure on $S^{k-1} = S^{n-1} \cap S$, using the mapping $u \to l_u$, we get from (9) that

$$\widehat{V}_{n,k}(K \cap S) = \int_{S^{n-1} \cap S} \widehat{V}_{n,1}(K \cap l_u) \frac{\lambda_{k-1}(du)}{\omega_k}. \tag{10}$$

It follows from (3), (8) and (10) that for $K \in \mathcal{S}(st)$

$$\widehat{V}_{n,k}(K \cap S) = \frac{\pi^{n/2}}{n\Gamma(n/2)} \frac{2k}{\omega_k} \widetilde{V}_{n,k}(K \cap S),$$

$S \in \mathcal{G}(n, k)$. The local stereological volume estmator is thus proportional to the section function of geometric tomography.

It is also possible to construct a local volume estimator, based on an isotropic $k-$subspace S, containing a *fixed* $r-$subspace T, say, where $r < k$. This estimator takes the form

$$\widehat{V}_{n,k(r)}(K \cap S) = \frac{\omega_{n-r}}{\omega_{k-r}} \int_{K \cap S} d(x,T)^{n-k} \lambda_k(dx).$$

Note that with this notation, $\widehat{V}_{n,k(0)} = \widehat{V}_{n,k}$. Using a decomposition of Lebesgue measure, it is not difficult to see that

$$\widehat{V}_{n,k(r)}(K \cap S) = \int_T \widehat{V}_{n-r,k-r}((K - x) \cap S \cap T^\perp) \lambda_r(dx). \tag{11}$$

For $K \in \mathcal{S}(st)$, we therefore have

$$\begin{aligned}&\widehat{V}_{n,k(r)}(K \cap S)\\ &= \frac{\pi^{(n-r)/2}}{(n-r)\Gamma((n-r)/2)} \frac{2(k-r)}{\omega_{k-r}} \int_T \widetilde{V}_{n-r,k-r}((K - x) \cap S \cap T^\perp) \lambda_r(dx).\end{aligned}$$

In this formula, section functions $\widetilde{V}_{i,k}$ with $i < n$ appear.

Note also that (10) implies that $\widehat{V}_{n,n-1}$ is proportional to the spherical Radon transform of $\widehat{V}_{n,1}$. Thus, it follows from (10) that for $u \in S^{n-1}$,

$$\begin{aligned}\widehat{V}_{n,n-1}(K \cap u^\perp) &= \frac{1}{\omega_{n-1}} \int_{S^{n-1} \cap u^\perp} \widehat{V}_{n,1}(K \cap l_v) \lambda_{n-2}(dv)\\ &= \frac{1}{\omega_{n-1}} Rf(u),\end{aligned}$$

where $f(u) = \widehat{V}_{n,1}(K \cap l_u)$.

Local stereological volume estimators in $\mathbb{E}^3$

In $\mathbb{E}^3$, we have 3 different local stereological volume estimators, viz. $\widehat{V}_{3,1}$, $\widehat{V}_{3,2}$ and $\widehat{V}_{3,2(1)}$.

The estimator $\widehat{V}_{3,1}$ is based on information along an isotropic line l through o and is given by

$$\begin{aligned}\widehat{V}_{3,1}(K \cap l) &= 2\pi \int_{K \cap l} d(x,o)^2 \lambda_1(dx)\\ &= \frac{2\pi}{3} \rho_{3,K}(u),\end{aligned}$$

where the last equality holds if $K \in \mathcal{S}(st)$. Often, measurements along two perpendicular directions are combined. In that case, the estimator is called the nucleator.

The estimator $\widehat{V}_{3,2}$ is based on information in an isotropic plane S through o. From (6), we find

$$\widehat{V}_{3,2}(K \cap S) = 2 \int_{K \cap S} d(x,o) \lambda_2(dx).$$

The planar integral can be discretized using a line grid in the plane S. The discretized version is called the isotropic rotator in the stereological literature.

The estimator $\widehat{V}_{3,2(1)}$ is based on an isotropic plane S, containing a fixed line l. It is given by

$$\widehat{V}_{3,2(1)}(K \cap S) = \pi \int_{K \cap S} d(x, l) \lambda_2(dx).$$

Using (11), we get

$$\widehat{V}_{3,2(1)}(K \cap S) = \int_l \widehat{V}_{2,1}((K - x) \cap S \cap l^{\perp}) \lambda_1(dx).$$

If $K \in \mathcal{S}(st), \widehat{V}_{2,1}$ can be expressed in terms of the 2-chord function of K. A discretized version of $\widehat{V}_{3,2(1)}$ is called the vertical rotator.

Variance of planar area estimators based on systematic sampling

Kiên Kiêu & Marianne Mora

Contents

1 Introduction

Systematic sampling is widely used in practical stereology. Examples of systematic sampling probes are serial sections, line and point grids. Assessing the precision of such designs is not a trivial task because of the statistical spatial dependency of the data. First methods for assessing the precision of systematic geometric sampling are due to Kendall [3, 4] and Matheron [5, 6].

We present their approach on a particular case: the estimation of planar area based on sampling by parallel lines. First, the estimation variance is expressed in terms of the Fourier transform of the indicator function associated with the investigated body. Using classical tools from analysis (Gauss-Green formula, method of the stationary phase), an asymptotic approximation of the Fourier transform is derived. This yields an asymptotic approximation of the estimation variance involving some simple geometric features of the body boundary.

During the exercise session, we will apply the general approach in order to derive variance approximations for other stereological estimators. Also, we will see how to use the variance formulae in practice for assessing and improving stereological designs.

2 Sampling and estimation

Let X be a random bounded convex body X in the Euclidean plane $\mathbb{R}^2$.

We assume that the following regularity conditions hold:

1. The mean area and the mean boundary length of X are finite.
2. The boundary ∂X of X is almost surely (a.s.) C^4.
3. The radius of curvature $R(x)$ of ∂X at $x \in \partial X$ is uniformly bounded on ∂X:

$$\sup_{x \in \partial X} R < R_{\max} \text{ a.s.}$$

The parameter to be estimated is the mean area A of X. We consider the case where X is sampled by parallel lines with fixed orientation (e.g. vertical) and uniform random location, see Figure 1. Available measurements are intercept lengths. Let $L(x_1)$, $x_1 \in \mathbb{R}$, be the intercept length for the vertical line with abscissa x_1. The observed intercept lengths can be written as

$$L\left((U+k)T\right),$$

where $k \in \mathbb{Z}$ and $T > 0$ is the line spacing. The area estimator is

$$\widehat{A} = T \sum_k L\left((U+k)T\right). \tag{1}$$

This estimator is conditionally unbiased given X.

The variance of the estimator $\widehat{A}$ can be decomposed as follows

$$\operatorname{Var} \widehat{A} = \operatorname{Var} A + \operatorname{E} \operatorname{Var}[\widehat{A}|X]. \tag{2}$$

The term $\operatorname{E}\operatorname{Var}[\widehat{A}|X]$ is called below the *sampling variance*. Hence the estimation variance is the sum of the area variance and of the sampling variance. The sampling variance depends both on T and on the distribution of X. Also, note that the estimation variance cannot be less than the area variance.

The estimation of the mean area $\operatorname{E} A$ is usually based on a sample $X_1, X_2 \dots X_n$ of X as shown in Figure 1. If $\widehat{A}_1, \widehat{A}_2, \dots, \widehat{A}_n$ are the estimated areas, the mean area is estimated by the average estimated area:

$$\bar{A} = \frac{1}{n} \sum_{i=1}^{n} \widehat{A}_i. \tag{3}$$

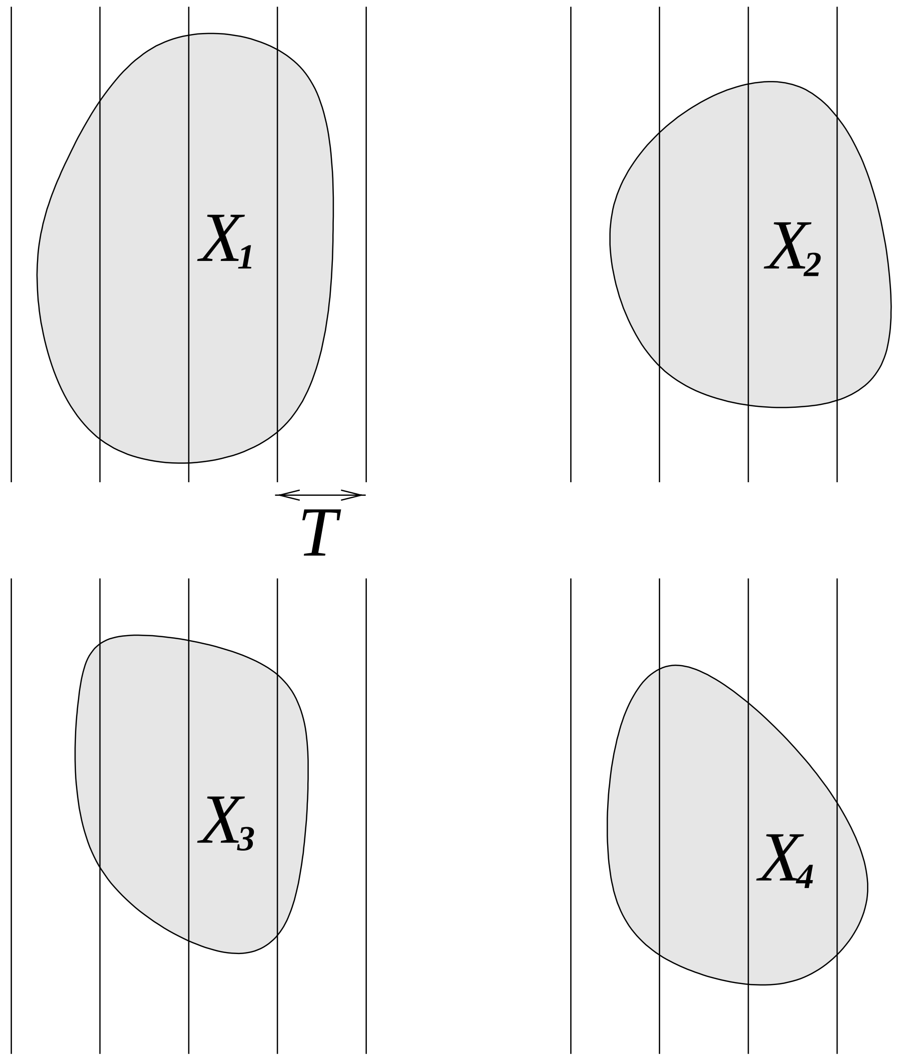


Figure 1: A sample of 4 planar bodies. Each body is sampled by parallel lines. The area of each body is estimated by the sampling distance times the total intercept length. The mean body area is estimated by the average estimated area.

This estimator is unbiased. The variance of the estimator $\bar{A}$ is given by

$$\operatorname{Var} \bar{A} = \frac{1}{n} \operatorname{Var} \widehat{A} = \frac{1}{n} \left(\operatorname{Var} A + \operatorname{E} \operatorname{Var}[\widehat{A} | X] \right). \tag{4}$$

It decreases to 0 as the sample size n increases to infinity.

Below, we focus on the sampling variance $\operatorname{E} \operatorname{Var}[\widehat{A}|X]$.

3 Estimation variance and Fourier transform

The estimator (1) may be considered as a periodic function:

$$u \in \mathbb{R} \mapsto T \sum_k L\left((u+k)T\right). \tag{5}$$

Note that the function (5) is integrable over $[0, 1[$ (its integral is equal to A). Let C_j be the Fourier coefficients of (5):

$$C_j = T \sum_k \int_0^1 L\left((u+k)T\right) \exp\left(-2\pi i u j\right) \, du.$$

Using the change of variable $x_1 = (u+k)T$, the integral is written as

$$T^{-1} \int_{kT}^{(k+1)T} L\left(x_1\right) \exp\left(-2\pi i x_1 / T j\right) \, dx_1.$$

Summing over all $k \in \mathbb{Z}$, we obtain the Fourier transform of L:

$$C_j = \mathcal{F}L\left(\frac{j}{T}\right).$$

Since X is bounded, the estimator $\widehat{A}$ is bounded and squared integrable. From Parseval equality, it follows

$$\int_0^1 \left(\widehat{A}\left(u\right) - A\right)^2 \, du = \sum_j{}' \left|C_j\right|^2 = \sum_j{}' \left| \mathcal{F}L\left(\frac{j}{T}\right) \right|^2,$$

where the dash indicates summation over $j \neq 0$.

The intercept length $L(x_1)$ can be obtained by integrating the indicator function I of X along the vertical line with abscissa x_1:

$$L(x_1) = \int_{\mathbb{R}} I(x_1, x_2) \, dx_2.$$

It follows that

$$\mathcal{F}L\left(y_1\right) = \int_{\mathbb{R}} \int_{\mathbb{R}} I(x_1, x_2) \exp\left(-2\pi i x_1 y_1\right) \, dx_1 dx_2.$$

The right-hand side is just the Fourier transform of I along the horizontal axis:

$$\mathcal{F}L\left(y_1\right) = \mathcal{F}I\left(y_1, 0\right).$$

Hence, we get

$$\operatorname{Var}[\widehat{A}|X] = \sum_k{}' \left|\mathcal{F}I\left(\frac{k}{T}, 0\right)\right|^2, \tag{6}$$

and

$$\operatorname{E}\operatorname{Var}[\widehat{A}|X] = \sum_k{}' \operatorname{E}\left|\mathcal{F}I\left(\frac{k}{T}, 0\right)\right|^2. \tag{7}$$

4 Asymptotic approximation of the Fourier transform

In this section, we derive asymptotic approximations of $\mathcal{F}I$.

Let φ be a continuous integrable function on $\mathbb{R}$ such that

$$y_1 \in \mathbb{R} \mapsto y_1\varphi(y_1)$$

is integrable. Then the derivative of $\mathcal{F}\varphi$ is given by

$$\left(\mathcal{F}\varphi\right)'\left(x_1\right) = -2\pi i \int_{\mathbb{R}} y_1 \exp\left(-2\pi i x_1 y_1\right)\varphi(y_1)\, dy_1. \tag{8}$$

Let us consider the integral

$$2\pi i \int_{\mathbb{R}} y_1 \mathcal{F}I(y_1, 0)\varphi(y_1)\, dy_1.$$

Expanding the Fourier transform of the indicator function, we get

$$2\pi i \int_X \int_{\mathbb{R}} y_1 \exp\left(-2\pi i x_1 y_1\right)\varphi(y_1)\, dx dy_1.$$

Identifying the derivative of the Fourier transform of φ, we obtain the identity

$$2\pi i \int_{\mathbb{R}} y_1 \mathcal{F}I(y_1, 0)\varphi(y_1)\, dy_1 = -\int_X \left(\mathcal{F}\varphi\right)'\left(x_1\right)\, dx.$$

Using the Gauss-Green formula (see Section A.1 of the Appendix), the right-hand side of the above equality can be written as an integral over the boundary ∂X of X:

$$\int_X \left(\mathcal{F}\varphi\right)'\left(x_1\right)\, dx = -\int_{\partial X} \left(\mathcal{F}\varphi\right)\left(x_1\right) n_1(x)\, dx.$$

Hence, we get

$$2\pi i \int_{\mathbb{R}} y_1 \mathcal{F}I\left(y_1, 0\right)\varphi(y_1)\, dy_1 = \int_{\mathbb{R}} \int_{\partial X} \exp\left(-2\pi i x_1 y_1\right) n_1\left(x\right)\varphi(y_1)\, dx dy_1.$$

The pointwise identity follows from the continuity of $\mathcal{F}I$:

$$2\pi i y_1 \mathcal{F}I\,(y_1, 0) = \int_{\partial X} \exp\left(-2\pi i x_1 y_1\right) n_1\,(x)\;\,dx. \tag{9}$$

Now, consider a parametric representation of ∂X:

$$\theta \in [0, 2\pi[\;\mapsto x\,(\theta) = (x_1(\theta), x_2(\theta)) \in \partial X.$$

The right-hand side of (9) can be written as

$$\int_0^{2\pi} \exp\left(-2\pi i x_1(\theta) y_1\right) n_1\,(x(\theta))\; Jx(\theta)\;\,d\theta.$$

When y_1 tends to ∞, the asymptotic behavior of the integral above is determined by the local behavior of $x_1(\theta)$ and $n_1\,(x(\theta))\; Jx(\theta)$ in the neighbourhood of critical points for the function:

$$\theta \mapsto x_1(\theta).$$

Observe that $x_1'(\theta) = 0$ if the tangent to X at $x(\theta)$ is vertical (i.e. parallel to the sampling lines).

An explicit approximation is obtained using the method of the stationary phase (see Sections A.2 and B of the Appendix). For the squared modulus of $\mathcal{F}I$, we get

$$\left|\mathcal{F}I\,(y_1, 0)\right|^2 = \frac{1}{4\pi^2}|y_1|^{-3}\sum R(x) + |y_1|^{-3} Z\,(y_1) + \mathrm{O}\left(|y_1|^{-7/2}\right), \tag{10}$$

where the sum is taken over the two points x_- and x_+ on the boundary ∂X where the tangent is vertical, $R(x)$ is the radius of curvature at x and Z is the oscillating function defined by

$$Z\,(y_1) = -\frac{1}{2\pi^2}\sqrt{R(x_-)R(x_+)}\,\sin\left(2\pi H |y_1|\right),$$

H being the horizontal breadth (distance between the two vertical support lines) of X.

Next, consider the case where X is isotropic random. According to a well-known result from differential geometry (see e.g. Santaló's book [7, page 3]), we have for any convex set K with a C^2 boundary

$$\int_0^{\pi} \sum R(x)\;d\theta = B,$$

where the sum is taken over the two points on the boundary of K where the angle between the tangent and a given axis is equal to θ and B is the boundary length of K. For an isotropic random X, it follows

$$\mathrm{E}\sum R(x) = \frac{1}{\pi}\,\mathrm{E}\,B.$$

Hence, we get

$$\mathrm{E}\left|\mathcal{F}I\,(y_1, 0)\right|^2 \simeq \frac{1}{4\pi^3}|y_1|^{-3}\,\mathrm{E}\,B + |y_1|^{-3}\,\mathrm{E}\,Z\,(y_1)\,.$$

Under some additional regularity conditions on X, it can be shown that the oscillating term can be neglected. We have

$$|\mathrm{E}\, Z\,(y_1)| \leq \frac{1}{2\pi^2} R_{\max}\, |\mathrm{E}\sin\left(2\pi H y_1\right)|\,.$$

If the distribution of H has a density h with respect to the Lebesgue measure on $\mathbb{R}$, the mean in the right-hand side is equal to the imaginary part of the Fourier transform of h. It tends to 0 as y_1 tends to ∞. Hence, we get

$$\mathrm{E}\,|\mathcal{F}I\,(y_1,0)|^2 \simeq \frac{1}{4\pi^3}|y_1|^{-3}\,\mathrm{E}\,B. \tag{11}$$

The speed of convergence of $\mathrm{E}\,Z$ to 0 depends on the distribution of H. For common statistical distributions (Chi square, Gamma, Beta...), we have

$$|\mathrm{E}\, Z(y_1)| = \mathrm{O}\left(|y_1|^{-\epsilon}\right),$$

where $\epsilon > 0$.

5 Approximation and estimation of the variance

In this section, we consider only the case where X is isotropic random. It is also assumed that the breadth H is distributed such that the oscillating term $\mathrm{E}\,Z$ is a $\mathrm{O}\left(y_1^{-\epsilon}\right)$ with $\epsilon > 0$.

Combining formulae (6) and (11), we get the following asymptotic approximation (T small):

$$\mathrm{E}\,\mathrm{Var}[\widehat{A}|X] \simeq \frac{T^3}{4\pi^3}\,\mathrm{E}\,B \sum_k{}' \,|k|^{-3} = \frac{\zeta(3)}{2\pi^3} T^3\,\mathrm{E}\,B, \quad \zeta(3) = 1.202. \tag{12}$$

Hence, the estimation variance can be determined from the sampling distance T and the mean boundary length $\mathrm{E}\,B$. When X is sampled by parallel lines, its boundary length can be estimated by

$$\frac{\pi}{2}TI, \tag{13}$$

where I is the number of intersection points between the boundary ∂X and the sampling lines. When X is convex, I is a.s. twice the number of lines hitting X.

6 Exercises

Exercise 1 (Length estimation for a finite union of intervals in $\mathbb{R}$) *Let* $X = \bigcup_{j=1}^{n} [a_j, b_j]$ *be a (deterministic) finite union of* n *intervals in* $\mathbb{R}$*. Let* L *denote the total length of* X*.*

The sampling probe is the lattice of points with uniform random location $\{T\,(U+k) :\ k \in \mathbb{Z}\}$, $T > 0$*, where* U *is a uniform random point in* $[0, 1[$.

Let I denote the indicator function of X. Consider the length estimator

$$\widehat{L} = T \sum_{k \in \mathbb{Z}} I\left(T\left(U+k\right)\right),$$

based on the count of points of the lattice which intersect X.

1. *Show that*
$$\operatorname{Var} \widehat{L} = \sum_{k}{}' \left| \mathcal{F}I\left(\frac{k}{T}\right) \right|.$$

2. *Compute the Fourier transform $\mathcal{F}I$.*

3. *Let $\mathcal{E}$ be the set of unordered pairs of distinct endpoints of X. Show that*
$$\left|\mathcal{F}I\left(y\right)\right|^2 = \frac{1}{2\pi^2\left|y\right|^2}\left(n + \sum_{\{s,t\}\in\mathcal{E}} \chi\left(s,t\right)\cos\left(2\pi y\left(s-t\right)\right)\right)$$
where $\chi\left(s,t\right) = 1$ if s and t are both left or right endpoints and $\chi\left(s,t\right) = -1$ otherwise.

4. *Calculate* $\operatorname{Var} \hat{L}$. *Note: use the identity*
$$\frac{1}{\pi^2}\sum_{k=1}^{\infty}\frac{\cos(2\pi kx)}{k^2} = x^2 - x + \frac{1}{6}, \quad x \in [0,1],$$
and $\zeta(2) = \pi^2/6$.

In the following exercises, area estimation in $\mathbb{R}^2$ is considered. The body X under study is supposed to be isotropic random and satisfies all the necessary regularity conditions. The parameter to be estimated is the mean area $\operatorname{E} A$ of X.

Exercise 2 (Area estimation in $\mathbb{R}^2$ by point sampling) *The estimation is performed on a single observation X. The sampling probe is the square lattice of points with uniform random location $\left\{T\left(U+k\right);\, k \in \mathbb{Z}^2\right\}$, $T > 0$, where U is a uniform random point in $\left[0,1\right[^2$ (see Figure 2).*

Consider the area estimator

$$\widehat{A} = T^2 \sum_{k \in \mathbb{Z}^2} I\left(T\left(U+k\right)\right).$$

Using the same arguments as in Section 3 and Exercise 1, we can check that

$$\operatorname{E}\operatorname{Var}[\widehat{A}|X] = \sum_{k\in\mathbb{Z}^2}{}' \operatorname{E}\left|\mathcal{F}I\left(\frac{k}{T}\right)\right|^2.$$

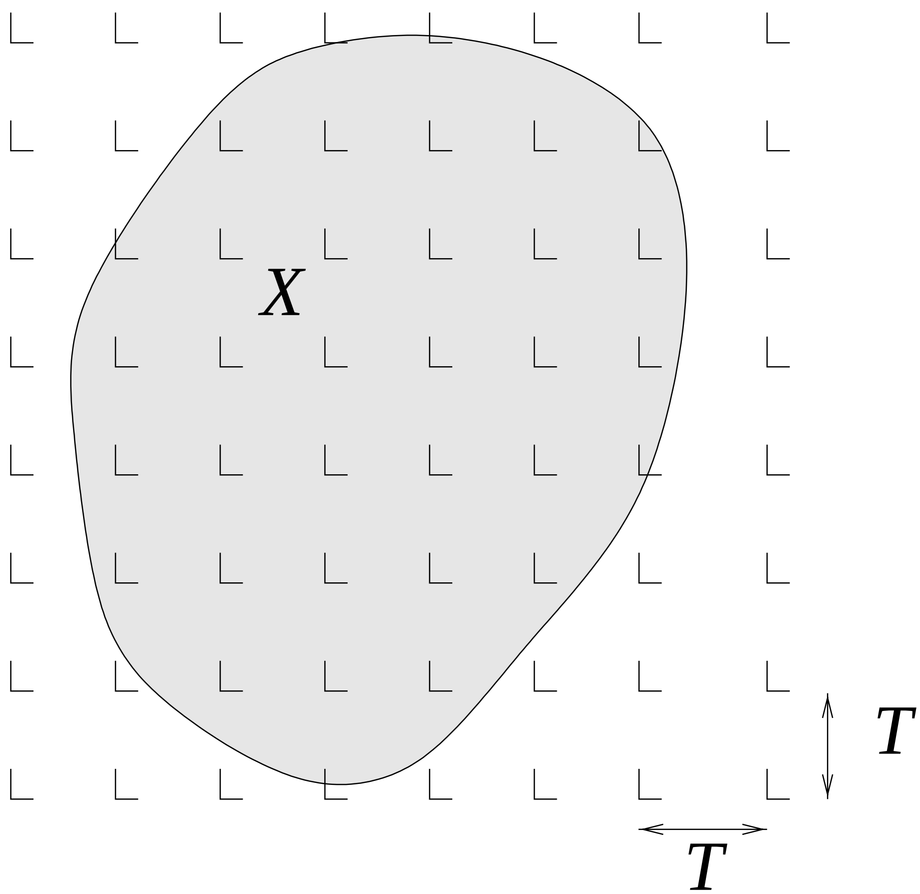


Figure 2: The planar body X is sampled by a square lattice of points. Its area is estimated by the area of the fundamental tile of the lattice times the total number of lattice points hitting X.

Write an asymptotic approximation formula for $\mathrm{E}\,\mathrm{Var}[\widehat{A}|X]$ *using the approximation of* $\mathcal{F}I$ *given by Formula (11) of Section 4. Note: you can use the two-dimensional Epstein zeta function*

$$\mathcal{Z}\,(3) = \sum_{k\in\mathbb{Z}^2}{}' \|k\|^{-3} \simeq 9.\,033\,6.$$

Exercise 3 (Area estimation in $\mathbb{R}^2$ by strip sampling) *The estimation is performed on a single observation* X.

For $s \in \mathbb{R}$, *let* $S\,(s)$ *be the vertical strip of width* $w > 0$ *defined by* $S\,(s) = [s, s + w] \times \mathbb{R}$.

The sampling probe is the series of strips

$$\{S\,(T\,(U + k))\,;\, k \in \mathbb{Z}\}\,,\, T > w,$$

where U *is a uniform random point in* $[0, 1[$ *(see Figure 3). The distance between neighbour strips is* $T - w$ *and the total sampling fraction is* $f = w/T$.

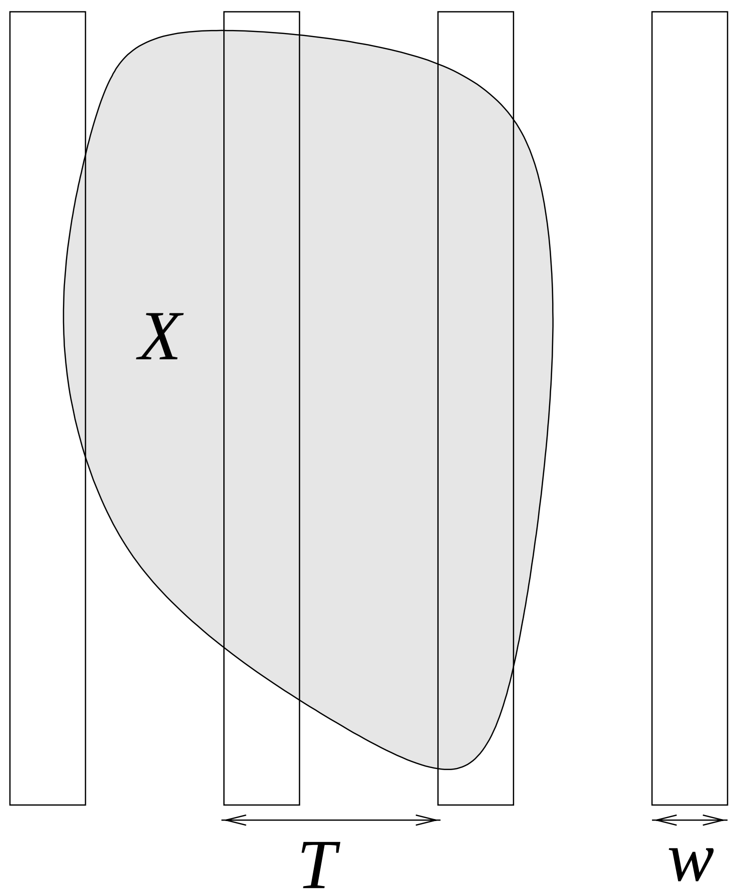


Figure 3: The planar body X is sampled by vertical strips. Its area is estimated by the total area of X contained in the strips divided by the sampling fraction.

1. *For $s \in \mathbb{R}$, let $A(s)$ be the area of X measured in the strip $S(s)$. Show that*

$$A(s) = w\, L * p(s)$$

where $L(x_1)$ is the intercept length for the vertical line with abscissa x_1 and $p(s) = \frac{1}{w} I_{[-w,0]}(s)$.

2. *Consider the area estimator*

$$\widehat{A} = T \sum_{k \in \mathbb{Z}} \frac{A(T(U+k))}{w}.$$

   (a) *Show that*

$$\mathrm{E}\,\mathrm{Var}[\widehat{A}|X] = \sum_k{}' \, \mathrm{E} \left| \mathcal{F}I\left(0, \frac{k}{T}\right)\right|^2 \frac{(1-\cos(2\pi f k))}{2\pi^2 k^2 f^2}.$$

   (b) *Derive an asymptotic approximation formula for $\mathrm{E}\,\mathrm{Var}[\widehat{A}|X]$. Note: you may use the polylogarithm function Li_5:*

$$\mathrm{Li}_5(\exp(2\pi i x)) = \sum_{k=1}^{\infty} \frac{\exp(2\pi i k x)}{k^5}, \quad \mathrm{Li}_5(0) = \zeta(5).$$

   The polylogarithm function can be calculated by standard mathematical softwares such as Maple and Mathematica.

Exercise 4 (Practical exercise) *Let $X_1, X_2, \ldots, X_6$ be a sample of a random planar body X in $\mathbb{R}^2$. Each X_i has been sampled by parallel lines separated by a distance $T =$* 1 cm.

The intercept lengths are given in Table 1.

1. *Calculate for each X_i the estimated area $\widehat{A}_i$ and the estimated boundary length $\widehat{B}_i$.*

2. *Calculate the estimations of $\mathrm{E}\,A$, $\mathrm{E}\,B$, $\mathrm{Var}\,\widehat{A}$, $\mathrm{E}\,\mathrm{Var}[\widehat{A}|X]$ and $\mathrm{Var}\,A$:*

$$\begin{aligned}
\mathrm{E}\,A &: \bar{A} = \frac{1}{6}\sum_{i=1}^{6} \widehat{A}_i \\
\mathrm{E}\,B &: \bar{B} = \frac{1}{6}\sum_{i=1}^{6} \widehat{B}_i \\
\mathrm{Var}\,\widehat{A} &: S_e^2 = \frac{1}{5}\sum_{i=1}^{6} \left(\widehat{A}_i - \bar{A}\right)^2 \\
\mathrm{E}\,\mathrm{Var}[\widehat{A}|X] &: S_s^2 = \frac{\zeta(3)}{2\pi^3} T^3 \bar{B} \\
\mathrm{Var}\,A &: S_e^2 - S_s^2
\end{aligned}$$

intercept lengths measured on					
X_1	X_2	X_3	X_4	X_5	X_6
2.3	1.5	2.3	2.4	1.5	2.9
3.9	3.8	2.7	4.4	2.6	3.9
4.7	4.9	2.9	5.4	3.1	4.3
5.0	5.0	3.1	5.7	3.4	4.4
4.6	4.7	3.1	5.2	3.5	3.9
3.7	3.8	3.1	3.7	3.4	2.7
	0.7	2.8		3.1	
		2.2		2.5	
		1.1		1.0	

Table 1: Lengths are given in cm.

3. *Calculate the coefficient of error of the mean area estimator $\bar{A}$:*

$$\mathrm{CE}\,\bar{A} = \sqrt{\frac{\mathrm{Var}\,\bar{A}}{\mathrm{E}^2\,\bar{A}}} = \sqrt{\frac{\mathrm{Var}\,A}{6\,\mathrm{E}^2\,A}}.$$

4. *In order to get $\mathrm{CE}\,\bar{A} = 0.03$, we may change the sample size (less X_i's) or the value of the sampling distance T.*
   (a) *Compute the required number of X_i's for fixed T.*
   (b) *Compute the required value of T for fixed sample size.*

References

[1] Federer, H. (1969) *Geometric Measure Theory*. Springer-Verlag.

[2] Hörmander (1990) *The Analysis of Linear Partial Differential Operators I*. Second edition. Springer-Verlag.

[3] Kendall, D.G. (1948) On the number of points of a given lattice inside a random oval. *Quart. J. Math. Oxford*, **19**, 1–26.

[4] Kendall, D.G. & Rankin, R.A. (1953) On the number of points of a given lattice in a random hypersphere. *Quart. J. Math. Oxford(2)*, **4**, 178-189.

[5] Matheron, G. (1965) *Les variables régionalisées et leur estimation*. Masson.

[6] Matheron, G. (1971) The theory of regionalized variables and its applications. *Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau*, fascicule **5bis**. École Nationale Supérieure des Mines de Paris.

[7] Santaló, L. (1976) *Integral Geometry and Geometric Probability*. Addison-Wesley.

A Toolbox

A.1 Gauss-Green formula

The Gauss-Green formula is a classical result from analysis. A general version can be found in Federer's textbook [1].

Theorem 1 *Let C be a closed C^2 curve in the Euclidean plane $\mathbb{R}^2$. Below, X denotes the domain bounded by C. Let ψ be a differentiable vector field on $\mathbb{R}^2$. Then we have*

$$\int_X \operatorname{div} \psi(x)\, dx = -\int_C \psi(x) \cdot n(x)\, dx, \tag{14}$$

where $n(x)$ is the outer normal to X at $x \in C$.

A.2 Method of the stationary phase

The following theorem is a consequence of much more general results. In particular, one can find in Hörmander's textbook [2] further results for higher-dimensional spaces and other regularity conditions.

Theorem 2 *Let f and u be real-valued periodic (period $= 2\pi$) functions on $\mathbb{R}$. We assume that $f \in C^4$ and $u \in C^3$. The set of non-degenerate critical points for f*

$$K_0 = \{x_0 \in [0, 2\pi[: f'(x_0) = 0, f''(x_0) \neq 0\}$$

is supposed to be finite. Then, we have

$$\int_0^{2\pi} u(x) \exp\left(iyf(x)\right)\, dx = y^{-1/2} \sum_{x_0 \in K_0} \exp\left(iyf(x_0)\right) \left(\frac{f''(x_0)}{2\pi i}\right)^{-1/2} u(x_0) + \mathrm{O}\left(y^{-1}\right). \tag{15}$$

B An application of the method of the stationary phase

Let X be a bounded convex body in the Euclidean plane $\mathbb{R}^2$. The boundary C of X is assumed to be C^4. Furthermore, the radius of curvature is supposed to be bounded on C.

If $r : \mathbb{R} \to \mathbb{R}_+$ is the radius function of X, the closed curve C can be parametrized by

$$\theta \in \mathbb{R} \mapsto x(\theta) = (r(\theta)\cos\theta, r(\theta)\sin\theta) \in C.$$

Note that the periodic function x is C^4.

Integration along the curve C is defined by

$$\int_C \varphi(x)\ dx = \int_0^{2\pi} \varphi(x(\theta))Jx(\theta)\ d\theta,$$

where the Jacobian $Jx = \|x'\|$.

Tangent vectors t to C at x are such that

$$t \propto x'.$$

Normal vectors n to C at x are such that

$$n \perp x'.$$

In particular, the unit outer normal to X at x is

$$n(x) = (x')^*/Jx,$$

where $(x')^*$ is the dual of x':

$$(x')^* = (x_2', -x_1')\,.$$

The radius of curvature on C is given by the formula

$$R = \frac{(Jx)^2}{x'' \cdot n}.$$

In this appendix, we use the method of the stationary phase (Theorem 2, Section A.2) in order to derive an asymptotic approximation of

$$\int_C \exp\left(-2\pi i x_1 y_1\right) n_1(x)\ dx$$

for large positive y_1's. The integral can be written as

$$\int_0^{2\pi} \exp\left(-2\pi i x_1(\theta) y_1\right) n_1\left(x(\theta)\right) Jx(\theta)\ d\theta.$$

We apply the method of the stationary phase (Theorem 2 of Section A.2) with

$$\begin{aligned} f(x) &\equiv x_1(\theta) \\ y &\equiv -2\pi y_1 \\ u(x) &\equiv n_1\left(x(\theta)\right) Jx(\theta). \end{aligned}$$

Note that θ is a critical point for x_1 if the tangent to C at $x(\theta)$ is vertical, i.e. if $n(x(\theta)) = (\pm 1, 0)$. Let $\theta_\pm$ be the two values such that $n(x(\theta_\pm)) = (\pm 1, 0)$ and let $x_- = x(\theta_-)$ and $x_+ = x(\theta_+)$.

It is easy to check that

$$\begin{aligned} x_1''(\theta_\pm) &= \pm\frac{Jx(\theta_\pm)^2}{R(x_\pm)} \\ n_1\,(x_\pm)\,Jx(\theta_\pm) &= \pm Jx(\theta_\pm). \end{aligned}$$

Hence, Formula (15) yields

$$\int_C \exp\left(-2\pi i x_1 y_1\right) n_1(x)\; dx = \frac{\exp\left(-2\pi i y_1 x_1(\theta_+) - i\pi/4\right)\sqrt{R(x_+)}}{\sqrt{y_1}} - \frac{\exp\left(-2\pi i y_1 x_1(\theta_-) - i3\pi/4\right)\sqrt{R(x_-)}}{\sqrt{y_1}} + \mathrm{O}\left(y_1^{-1}\right).$$

In particular, the squared modulus of the integral is given by

$$\frac{R(x_-) + R(x_+) - 2\sqrt{R(x_-)R(x_+)}\,\sin\left(2\pi y_1\left(x_1(\theta_+) - x_1(\theta_-)\right)\right)}{y_1}.$$

C Solutions

Solution 1 (Length estimation for a finite union of intervals in $\mathbb{R}$)

$$\operatorname{Var}\widehat{L} = \frac{nT^2}{6} + T^2 \sum_{\{s,t\}\in\mathcal{E}} \chi(s,t)\left(r\left(\frac{s-t}{T}\right)^2 - r\left(\frac{s-t}{T}\right)^2 + \frac{1}{6}\right),$$

where $r(x) = x - [x]$ is the fractional part of x.

Solution 2 (Area estimation in $\mathbb{R}^2$ by point sampling)

$$\operatorname{E}\operatorname{Var}[\widehat{A}|X] \simeq \frac{\mathcal{Z}(3)}{4\pi^3} T^3 \operatorname{E} B.$$

Solution 3 (Area estimation in $\mathbb{R}^2$ by strip sampling)

$$\operatorname{E}\operatorname{Var}[\widehat{A}|X] \simeq \frac{\zeta(5) - \Re\operatorname{Li}_5\left(\exp(2\pi i f)\right)}{4\pi^5}\frac{T^3}{f^2}\operatorname{E} B,$$

where $\Re$ indicates real part.

Solution 4 (Practical exercise)

1.
   - $\bar{A} = 24.1\ \mathrm{cm}^2$,
   - $\bar{B} = 22.5\ \mathrm{cm}$,
   - $S_e^2 = 2.40\ \mathrm{cm}^4$,
   - $S_s^2 = 0.44\ \mathrm{cm}^4$,
   - $S_e^2 - S_s^2 = 1.97\ \mathrm{cm}^4$.
2. *Estimate of* $\operatorname{CE}\bar{A} = 0.026\,2$.
3. (a) $\operatorname{CE}\bar{A} \leq 0.03$ *if the sample size is greater than or equal to* 5.

   (b) $\operatorname{CE}\bar{A} \leq 0.03$ *if* $T \leq 1.39\ \mathrm{cm}$.

List of additional hand-outs/notes

- *Geometric tomography* by Richard Gardner, Notices Amer. Math. Soc. **42** (1995), 422–429. [`http://www.ams.org/notices/199504/199504-toc-ps.html`]

- Copies of slides for R. Gardner's talks.

- *Some new, simple and efficient stereological methods and their use in pathological research and diagnosis* by Hans Jørgen G. Gundersen, Acta Pathologica, Microbiologica et Immunologica Scandinavica **96** (1988), 379-394.

- *The new stereological tools — disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis* by Hans Jørgen G. Gundersen, Acta Pathologica, Microbiologica et Immunologica Scandinavica **96** (1988), 857-881.

- Chapter 1 and List of References from *Local Stereology* by Eva B. Vedel Jensen, Advanced Series on Statistical Science & Applied Probability **5**, World Scientific, 1998.

- *The Computer Assisted Stereological Toolbox*, brochure about the C.A.S.T. grid system, Olympus Denmark. [`stereology@olympus.dk`].

List of participants

Adrian Baddeley
Department of Mathematics
University of Western Australia
Nedlands WA 6907
Australia
Email: adrian@maths.uwa.edu.au

Stephan Böhm
Dept. of Stochastics
University of Ulm
Helmholtzstrasse 18
D-89081 Ulm, Germany
Email: sboehm@mathematik.uni-ulm.de

Joanne Chia
Dept. Mathematics and Statistics
University of Western Australia
Nedlands WA 6907
Australia
Email: chia@maths.uwa.edu.au

Karl-Anton Dorph-Petersen
Stereological Research Laboratory
Building 182
University of Aarhus
DK-8000 Aarhus C, Denmark
Email: stereokadp@svfcd.au.dk

Marta Garcia-Finana
Dept. of Mathematics, Statistics and Computation
Faculty of Sciences, University of Cantabria
Avda Los Castros, s/n
E-39005 Santander, Spain
Email: martaf@matesco.unican.es

Richard Gardner
Department of Mathematics
Western Washington University
Bellingham
WA 98225-9063, U.S.A.
Email: gardner@baker.math.wwu.edu

Iris Geisler
German Cancer Research Center
Dept. of Biostatistics (R0700)
Im Neuenheimer Feld 280
D-69120 Heidelberg, Germany
Email: i.geisler@dkfz.de

Richard Gill
University of Utrecht
Mathematical Institute
Budapestlaan 6
NL-3585 CD Utrecht, The Netherlands
Email: Richard.Gill@math.ruu.nl

Hans Jørgen G. Gundersen
Stereological Research Laboratory
University of Aarhus
DK-8000 Aarhus C
Denmark
Email: stereo@svfcd.au.dk

Ute Hahn
Dept. of Mathematics and Statistics
University of Western Australia
Nedlands WA 6907
Australia
Email: uhahn@maths.uwa.edu.au

Martin Bøgsted Hansen
Dept. of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7E
DK-9220 Aalborg Ø, Denmark
Email: mbh@math.auc.dk

Niels Væver Hartvig
Dept. of Mathematical Sciences
University of Aarhus
DK-8000 Aarhus C
Denmark
Email: vaever@imf.au.dk

Eva B. Vedel Jensen
Dept. of Mathematical Sciences
Ny Munkegade
University of Aarhus
DK-8000 Aarhus C, Denmark
Email: eva@imf.au.dk

Patrick Juhl
NeuroSearch
Pederstrupvej 93
DK-2750 Ballerup
Denmark
Email: pmj@neurosearch.dk

Salme Karkkainen
Dept. of Statistics
University of Jyväskylä
P.O. Box 35 (MaD)
SF-40351 Jyväskylä, Finland
Email: samk@louhi.stat.jyu.fi

Marcus Kiderlen
University of Karlsruhe
Postfach 6980
D-76128 Karlsruhe
Germany
Email: kiderlen@math.uni-karlsruhe.de

Kiên Kiêu
Unité de Biométrie
Institut National de la Recherche Agronomique
Domaine de Vilvert
F-78352 Jouy-en-Josas, France
Email: kien.kieu@jouy.inra.fr

Simone Klenk
Dept. of Stochastics
University of Ulm
Helmholtzstrasse 18
D-89081 Ulm, Germany
Email: goetz@mathematik.uni-ulm.de

Tomas Mrkvicka
VSK 17.listopadu
Patkova 3, pokol B413
CZ-18000 Praha
Czech Republic
Email: mrkvicka@karlin.mff.cuni.cz

Alberto Nettel
Statistical Research Group
University of Calgary
1206 Tuscarora Manor NW
T3L 2J9 Calgary, Alberta, Canada
Email: nettel@math.ucalgary.ca

Andrew Olenko
Statistics Dept., Mathematics Faculty
Kyiv University
Volodimirska 64
252033 Kyiv, Ukraine
Email: olenk@ukma.kiev.ua

Niels Holm Olsen
3D-Lab, School of Dentistry
University of Copenhagen
Nørre Allé 20
DK-2200 Copenhagen N, Denmark
Email: qwer@lab3d.odont.ku.dk

Jonas Kildegaard Pedersen
NeuroSearch
Pederstrupvej 93
DK-2750 Ballerup
Denmark
Email: jkp@neurosearch.dk

Pietro Polesello
Institut de Mathématiques
Université Paris VI Pierre et Marie Curie
175 rue de Chevaleret
F-75013 Paris, France
Email: pietro@math.jussieu.fr

Michaela Prokesova
Dept. of Probability and Statistics
Charles University
Sokolovska 83
CZ-18675 Praha 8, Czeck Republic
Email: misa@kma14.karlin.mff.cuni.cz

Boris Rubin
Dept. of Mathematics
Hebrew University
91904 Jerusalem
Israel
Email: boris@math.huji.ac.il

Hanspeter Schmidli
Dept. of Mathematical Sciences
Aarhus University
DK-8000 Aarhus C
Denmark
Email: schmidli@imf.au.dk

Antje Schwella
Institut für Stochastik
Friedrich-Schiller-Universität Jena
Ernst-Abbee-Platz 1-4
D-07743 Jena, Germany
Email: antje@mipool.uni-jena.de

Eugene Spodarev
Institut für Stochastik
Friedrich-Schiller-Universität Jena
Ernst-Abbee-Platz 1-4
D-07743 Jena, Germany
Email: seu@minet.uni-jena.de

Jens Jakob Thune
Neurological Research Laboratory
Bartholin Institute, Kommunehospitalet
Øster Farimagsgade 5
DK-1399 Copenhagen K, Denmark
Email: jjthune@hotmail.com

Uffe Thygesen
DFU – Danish Institute for Fisheries Research
Charlottenlund Slot
Jægersborg Allé 1
DK-2920 Charlottenlund, Denmark
Email: uht@dfu.min.dk

Aljoša Volčič
Dipartimento di Scienze Matematiche
Universitá degli Studi di Trieste
I-34100 Trieste
Italy
Email: volcic@univ.trieste.it

Dvoralai Wulfsohn
Institut for Maskinteknik
Aalborg Universitet
Pontoppidanstræde 101
DK-9220 Aalborg Ø, Denmark
Email: dw@ime.auc.dk