Recent Advances in Probability
Indian Statistical Institute, Kolkata
December 11–15, 2007

Differentiation of sets in measure

E.V.Khmaladze
Victoria University of Wellington

Based on the paper of the same title in *J. Mathemat. Analysis Appl.*, 334 (2007),1055-1072.

Other references - at the final pages.

The problem.

Set-valued mapping:

$$\varepsilon \to C(\varepsilon), \quad \varepsilon \in [0,1],$$

such that each $C(\varepsilon) \subset \mathbb{R}^d$ is bounded and Borel.

If $C(\varepsilon) \to C(0)$ in Hausdorf metric, then

$$A(\varepsilon) = C(\varepsilon) \triangle C(0)$$

"shrinks" towards something on the boundary $\partial C(0)$.

We want to differentiate $A(\varepsilon)$ and $C(\varepsilon)$ w.r.t. ε.

Given a measure $\mathbb{P}$ in $\mathbb{R}^d$, as a side result, we will give meaning to the equality

$$\frac{d}{d\varepsilon}\mathbb{P}(A(\varepsilon))|_{\varepsilon=0} = \mathbb{Q}(\frac{d}{d\varepsilon}A(\varepsilon)|_{\varepsilon=0}),$$

where the measure $\mathbb{Q}$ depends only on $\mathbb{P}$ and on the initial set $C(0)$ but not on the mapping $C(\varepsilon)$.

Local Poisson processes.

Poisson point processes N_n in $\mathbb{R}^d$, $n = 1, 2, \ldots$, with

$$EN_n(A) = n\mathbb{P}(A),$$

where $\mathbb{P}$ is some given measure.

Let $\mathcal{V}_\varepsilon(\partial C)$ be a narrow strip around the boundary ∂C:

$$\mathcal{V}_\varepsilon(\partial C) = \{z \in \mathbb{R}^d : \|z - \partial C\| \leq \varepsilon\},$$

and consider restriction of N_n on $\mathcal{V}_\varepsilon(\partial C)$,

$$N_{n\varepsilon} = \{N_n(A), A \subseteq \mathcal{V}_\varepsilon(\partial C)\}.$$

Call $N_{n\varepsilon}$, when $n \to \infty$ and $\varepsilon \to 0$ simultaneously, *local Poisson process in the neighborhood of* ∂C.

What is the limit of this process and where does this limit "live"?

$$\{N_n(A), A \subseteq \mathcal{V}_\varepsilon(\partial C)\} \xrightarrow{d} \quad ?$$

For given choice of "shrinking" sets $A(\varepsilon) \subseteq \mathcal{V}_\varepsilon(\partial C)$, suppose $n\mathbb{P}(A(\varepsilon)) \to l < \infty$. Therefore

$$N_n(A(\varepsilon)) \xrightarrow{d} N_0, \quad \text{where } N_0 \sim Poi(l)$$

But what is this N_0 "physically"?

If we have a family of "shrinking" sets, where can we put the family of these limiting Poisson random variables?

Not on the boundary ∂C as we will see.

And why should a statistician be interested in these local point processes?

see on the blackboard

Simple example: in $\mathbb{R}^2$, suppose

$$\{a \leq x \leq b, y = 0\} \subset \partial C.$$

Suppose $\mathbb{P}$ has a density $p(x, y)$ continuous in y at $y = 0$. For g positive and bounded, consider the sets

$$A(\varepsilon) = \{a \leq x \leq b, 0 \leq y < \varepsilon g(x)\},$$

For $n \sim \varepsilon^{-1}$

$$\begin{aligned} n\mathbb{P}\{A(\varepsilon)\} &\sim \varepsilon^{-1} \int_a^b \int_0^{\varepsilon g(x)} p(x, y) dx dy \\ &\sim \int_a^b g(x) p(x, 0) dx = l, \end{aligned}$$

However, for

$$A^-(\varepsilon) = \{a \leq x \leq b, -\varepsilon g(x) \leq y < 0\},$$

$$n\mathbb{P}\{A^-(\varepsilon)\} \to l \ \textit{ as well.}$$

But since $A(\varepsilon)$ and $A^-(\varepsilon)$ are disjoint, the $N_n(A(\varepsilon))$ and $N_n(A^-(\varepsilon))$ are independent and converge to independent Poisson r.v.

Moreover, consider disjoint sets of the form

$$A^k(\varepsilon) = \{a \leq x \leq b, (k-1)\varepsilon g(x) \leq y < k\varepsilon g(x)\}$$

for different integer values of k. Then $N_n(A^k(\varepsilon))$ are independent and converge to independent Poisson r.v. with the same intensity.

It will not be good to label all these limiting Poisson r.v. by the same label – either the bit of the boundary, $[a, b]$, or with their intensity.

Actually, with these, or any, "shrinking" sets we could associate a sequence of generalized function

$$\int_{\mathbb{R}^d} \phi(x, y)[\varepsilon^{-1}\mathbb{I}_{A^k(\varepsilon)}(x, y)]dxdy \to \int_a^b \phi(x, 0)g(x, 0)dx$$

with the *same* limit for all k.

So, the above means that the generalized functions provide more "crude" description of how these sets "shrink" than is needed for local processes.

What we suggest is the following: there are sets

$$B^k = \{a \leq x \leq b, (k-1) \leq s < kg(x)\},$$

completely specified by these $A^k(\varepsilon)$, and a measure

$$\mathbb{Q}(dx, ds) = p(x, 0)dxds,$$

quite independent of the choice of $A^k(\varepsilon)$, such that

$$\varepsilon^{-1}\mathbb{P}\{A^k(\varepsilon)\} \sim \frac{d}{d\varepsilon}\mathbb{P}\{A^k(\varepsilon)\} = \mathbb{Q}(B^k).$$

These B^k we interpret as the derivatives of $A^k(\varepsilon)$ at $\varepsilon = 0$. Then, as first prototype of our limit theorem, we have

$$N_n(A^k(\varepsilon)) \xrightarrow{d} N_0 N_0 N_0(\frac{d}{d\varepsilon}A^k(\varepsilon)|_{\varepsilon=0})$$

so that *the limiting process lives on the derivative sets.*

So far the derivative sets belonged to the same $\mathbb{R}^2$ as the sets $A^k(\varepsilon)$. But in general it would not be really possible to stay in $\mathbb{R}^2$

bb

The normal cylinder and local magnification map.

Assume that $C(0) = C$ is convex. We could take finite unions of convex sets or sets of positive reach.

We hope (see D.Hug, G.Last and W.Weil [11]), that $C(0)$ just bounded can also be treated.

About $C(\varepsilon)$ we do not, basically, assume anything except it is differentiable :-).

For $z \in \mathbb{R}^d$ denote $P_{\partial C}(z)$ its nearest point from ∂C

The *skeleton* of ∂C is the set $S_{\partial C}$ defined as

$$S_{\partial C} = \{z \in \mathbb{R}^d : P_{\partial C}(z) \text{ is not unique}\}.$$

We know that $\mu_d(S_{\partial C}) = 0$.

The *generalized normal bundle* of C is defined as

$Nor(C) = \{(x,u) : x \in \partial C,\ u-\text{outer normal at } x\}$

We use it to construct the cylinder

$$\Sigma = \mathbb{R} \times Nor(C)$$

It will be often easier to consider the cyllinder

$$\Gamma = \mathbb{R} \times \partial C$$

and to project sets of Σ onto sets from Γ by letting $(t,x,u) \mapsto (t,x)$.

But we have to work on Σ.

Any point $z \in \mathbb{R}^d \setminus S_{\partial C}$ can be written as

$$z = x + d(z)u,$$

where $x = P_{\partial C}(z)$ and $u \in N(x)$ and $d(z)$ the signed distance function

$$d(z) = \begin{cases} \|z - P_{\partial C}(z)\|, & \text{if } \ z \in \mathbb{R}^d \setminus C \\ -\|z - P_{\partial C}(z)\|, & \text{if } \ z \in C. \end{cases}$$

Hence, the *local magnification map*,

$$\tau_\varepsilon(z) = (\,\frac{d(z)}{\varepsilon}, P_{\partial C}(z), u),$$

is defined a.s. on $\mathbb{R}^d$.

Important tool will be the *local Steiner formula*:

$$\int_{\mathbb{R}^d} f(z)\mu_d(dz) =$$

$$\sum_{j=1}^{d} \binom{d-1}{j-1} \int_{Nor(C)} \int_{-r(x)}^{\infty} f(x+tu)t^{j-1} dt \theta_{d-j}(d(x,u)).$$

Here $\theta_{d-1}(B), \ldots, \theta_0(B)$ are finite measures on $Nor(C)$ called *support measures* of C [17]. In particular, $\theta_{d-1}(\cdot)$ is Hausdorff measure on ∂C.

Definition of differentiability.

Associate with a set-valued mapping $C(\varepsilon)$ its “increments”

$$A^{+}(\varepsilon) = C(\varepsilon) \setminus C, \quad A^{-}(\varepsilon) = C \setminus C(\varepsilon)$$

and

$$A(\varepsilon) = A^{+}(\varepsilon) \cup A^{-}(\varepsilon) = C(\varepsilon) \triangle C.$$

Denote

$$\tau_\varepsilon(A(\varepsilon)) \in \Sigma$$

the image of $A(\varepsilon)$. Denote M the measure on Σ defined as

$$M(ds, d(x,u)) = ds \times \theta_{d-1}(d(x,u)).$$

Definition. Call the $A(\varepsilon), 0 \leq \varepsilon \leq 1$, *differentiable* at ∂C at $\varepsilon = 0$ if, for $\varepsilon \to 0$,

$$\exists B \in \Sigma : \quad M(\tau_\varepsilon(A(\varepsilon)) \triangle B) \to 0.$$

Call the set B the *derivative* of $A(\varepsilon)$ at ∂C.

Define the *derivative* of $C(\varepsilon)$ at C is as the derivative of $C(\varepsilon)\triangle C$ at ∂C:

$$\frac{d}{d\varepsilon}C(\varepsilon)|_{\varepsilon=0} = \frac{d}{d\varepsilon}C(\varepsilon)\triangle C|_{\varepsilon=0} = B.$$

Some set-theoretic properties of the differentiation.

Lemma 1 *(i) If* $A_1(\varepsilon), A_2(\varepsilon)$ *– differentiable (at* ∂C*), then*

$$A_1(\varepsilon) \cup A_2(\varepsilon),\ \ A_1(\varepsilon) \setminus A_2(\varepsilon),\ \ \ A_1(\varepsilon) \cap A_2(\varepsilon)$$

all differentiable at ∂C

(ii) If $C_1(\varepsilon)$ *– differentiable at* C *and* $A_2(\varepsilon)$ *– differentiable at* ∂C*, then both*

$$C_1(\varepsilon) \cup A_2(\varepsilon),\ \ \ C_1(\varepsilon) \setminus A_2(\varepsilon)$$

differentiable at C

(iii) Let $f(\varepsilon) \nearrow$ *be differentiable at* 0 *and* $f(0) = 0$*. If* $C(\varepsilon)$ *is differentiable, then* $C(f(\varepsilon))$ *is also differentiable.*

Suppose $\mathbb{P}$ is absolutely continuous and finite on bounded sets. Suppose the density $p(z)$ can be approximated in the $\mathcal{V}_\varepsilon$ by a function depending on $P_{\partial C}(z)$ only.

$$\frac{1}{\varepsilon}\int_{\mathcal{V}_\varepsilon} |p(z) - \bar{p}(P_{\partial C}(z))|\mu_d(dz) \to 0 \qquad (1)$$

Now define a measure $\mathbb{Q}$ on Σ as follows:

$$\mathbb{Q}(ds, d(x,u)) = \bar{p}(x)M(ds, d(x,u)). \qquad (2)$$

Let θ^c_{d-j} denote the part of θ_{d-j} absolutely continuous with respect to θ_{d-1}.

Theorem 2 *Suppose* $\mathbb{P}$ *satisfies condition* (1) *and* $\bar{p}$ *is integrable w.r.t.* $\theta^c_{d-j}, j = 1, \ldots, d$. *If* $A(\varepsilon)$ *is differentiable at* ∂C *then*

$$\frac{d}{d\varepsilon}\mathbb{P}(A(\varepsilon))|_{\varepsilon=0} = \mathbb{Q}(\frac{d}{d\varepsilon}A(\varepsilon)|_{\varepsilon=0}).$$

Corollary 3 *Under conditions of the theorem*

$$\frac{d}{d\varepsilon}\mathbb{P}(C(\varepsilon))|_{\varepsilon=0} = \mathbb{Q}(\frac{d}{d\varepsilon}A^+(\varepsilon)|_{\varepsilon=0}) - \mathbb{Q}(\frac{d}{d\varepsilon}A^-(\varepsilon)|_{\varepsilon=0}).$$

Reduction to many 1-dim. differentiations.

Let $(x,u) \in Nor(C)$. Consider a *section* of a set A by the line $z = x + tu$ $(t \in \mathbb{R})$:

$$A_{(x,u)} = \{z \in A : P_{\partial C}(z) = x, z - x \in \mathbb{R}u\}.$$

Similarly, the set

$$\tau_\varepsilon(A_{(x,u)}) = \tau_\varepsilon(A)_{(x,u)} = B_{(x,u)}$$

is the section of $B \in \Sigma$ by the line $\mathbb{R} \times (x,u)$.

Definition 3. Call the section $A_{(x,u)}(\varepsilon)$ of $A(\varepsilon)$ *differentiable* at $(x,u) \in Nor(C)$ if (essentially :-)) $\exists B_{(x,u)} \in \mathbb{R} \times (x,u)$:

$$\int_{-r(x)/\varepsilon}^{\infty} \mathbb{I}_{B_{(x,u)}(\varepsilon) \triangle B_{(x,u)}}(s)ds \to 0.$$

Here $r(x)$ is the local reach of C at x.

Theorem 4 *Suppose ... something (see the paper). Then* $A(\varepsilon)$ *is differentiable (at* ∂C *at* $\varepsilon = 0$*) if sections* $A_{(x,u)}(\varepsilon)$ *are differentiable at* θ_{d-1} *-a.a.* $(x, u) \in \partial C$*. The derivative of* $A(\varepsilon)$ *is the set*

$$B = \cup_{(x,u)\in Nor(F)} B_{(x,u)}$$

where $B_{(x,u)}$ *is the derivative of* $A_{(x,u)}(\varepsilon)$*.*

Example. Let Q_ε be some positive definite matrix,
...

Example. Let again $C = C(0)$ be the unit ball, but this time assume that with ε increasing new "flanks" can branch away from it ...

Further properties. Some discussion and examples.

$A(\varepsilon)$ as subgraphs.

Let $h_\varepsilon, \varepsilon \in [0,1]$, be a family of the functions on $Nor(C)$, and let h_ε^+ and h_ε^- be positive and negative parts of h_ε. Consider the sets in $\mathbb{R}^d$

$$\begin{aligned} A^+(h_\varepsilon) &= \{z \in \mathbb{R}^d \setminus C : 0 < d(z) \le h_\varepsilon^+(x,u)\} \\ A^-(h_\varepsilon) &= \{z \in C \setminus S_{\partial C} : -h_\varepsilon^-(x,u) < d(z) \le 0\} \\ A(h_\varepsilon) &= A^+(h_\varepsilon) \cup A^+(h_\varepsilon) \qquad (3) \end{aligned}$$

where, as always, $x = P_{\partial C}(z)$ and u is the outer normal at x.

For a function g on $Nor(F)$, call the subsets of Σ defined as

$$\begin{aligned} g_{sub}^+ &= \{(t,x,u) : 0 < t \le g^+(x,u)\}, \\ g_{sub}^- &= \{(t,x,u) : -g^-(x,u) < t \le 0\} \\ g_{sub} &= g_{sub}^+ \cup g_{sub}^- \end{aligned}$$

the subgraphs of g^+, g^- and g respectively.

Consider the norm of h in the space $\mathfrak{L}_j(\theta_{d-j})$:

$$\begin{aligned}\|h\|_j =&[\int_{Nor(C)}(h^+(x,u))^j\theta_{d-j}(d(x,u))]^{1/j}\\ &+[\int_{Nor(C)}(h^-(x,u))^j\theta^c_{d-j}(d(x,u))]^{1/j}\end{aligned}$$

We say that h_ε is $\mathfrak{L}_1$-differentiable if

$$\exists g\in\mathfrak{L}_1(\theta_{d-1}):\quad \|\varepsilon^{-1}h_\varepsilon - g\|_1\to 0$$

.

Theorem 5 *If $\|h_\varepsilon\|_j = o(\varepsilon^{1/j}), j=2,\ldots,d,$ then $A(h_\varepsilon)$ is differentiable if and only if h_ε is $\mathfrak{L}_1$-differentiable. In this case*

$$\frac{d}{d\varepsilon}A(h_\varepsilon)|_{\varepsilon=0} = g_{sub}.$$

Shifts.

Let $C(\varepsilon) = C + \varepsilon A$, A - a convex body. This mapping is called the *affine mapping*. We can show that $C(\varepsilon)$ is differentiable at C with the derivative

$$B^+ = s_A^+(\cdot)_{sub}, \quad B^- = s_A^-(\cdot)_{sub}$$

where s_A is the support function of A.

More generally, for "smooth" shifts of differentiable mappings we have:

Theorem 6 *Suppose $C(\varepsilon)$ is "regularly differentiable" (see the paper :-)) at C with derivative set B and suppose*

$$\varepsilon^{-1} a(\varepsilon) \longrightarrow a' \in \mathbb{R}^d.$$

Then the mapping $C(\varepsilon) + a(\varepsilon)$ is differentiable at C and the derivative is $\bar{B}$ with

$$\bar{B}^+ = (B^+ + \langle a', \cdot \rangle)^+ \cup \langle a', \cdot \rangle_{sub}^+ \setminus (B^- + \langle a', \cdot \rangle)^+$$

and

$$\bar{B}^- = ((B^+ + \langle a', \cdot \rangle)^- \setminus \langle a', \cdot \rangle_{sub}^-) \cup (B^- + \langle a', \cdot \rangle)^-.$$

Sets defined through inequalities. Quasi-affine mappings.

Example. Consider a polytope

$$C = \{x \in \mathbb{R}^d : \langle c_i, x\rangle \le b_i, i = 1, \ldots, m\}$$

defined through the minimal set of linear inequalities. Consider its perturbation

$$C(\varepsilon) = \{x \in \mathbb{R}^d : \langle c_i(\varepsilon), x\rangle \le b_i(\varepsilon), i = 1, \ldots, m\}$$

where we only assume that

$$c_i(\varepsilon) \sim c_i + \varepsilon c_i',\ b_i(\varepsilon) \sim b_i + \varepsilon b_i',\ \ \varepsilon \to 0.$$

Although each $C(\varepsilon)$ is convex, the graph of it,

$$\{(C(\varepsilon), \varepsilon), 0 \le \varepsilon \le 1\},$$

does not have to be and typically is not convex. Hence it is not what is called *quasi-affine*. Often $C(\varepsilon)$ can not be approximated by a quasi-affine mapping with accuracy $o(\varepsilon)$ and therefore is not differentiable in the sense suggested in the literature.

However,

Theorem 7 *Let* $g_i(x) = b_i' - \langle c_i', x\rangle$. *Then*

$$\frac{d}{d\varepsilon} A(\varepsilon)|_{\varepsilon=0} = \cup_{i=1}^m g_{i,sub}.$$

In general, let a quasi-affine mapping be defined as

$$C(\varepsilon) = \cap_{\|\psi\|=1}\{x \in \mathbb{R}^d : \langle x, \psi \rangle \leq s(\psi) + \varepsilon c(\psi)\}$$

where $s(\cdot)$ is support function of the set C and $c(\cdot)$ is some positively homogeneous function.

Theorem 8 *Quasi-affine mapping is differentiable in the sense of Definition 2.*

Corollary. *Set-valued mapping $C(\varepsilon), \varepsilon[0,1]$, differentiable in the sense that there exists an quasi-affine mapping $C'(\varepsilon), \varepsilon[0,1]$, which approximates $C(\varepsilon)$ in Hausdorf metric with the rate $o(\varepsilon)$, is differentiable in the sense of Definition 2.*

Derivatives as measures.

on bb

Derivatives as tangent cones. (Connections with contingent derivatives of J.-P. Aubin and Clarke's derivative.)

on bb

References

[1] Z. Artstein. A calculus of set-valued maps and set-valued evolution equations, *Set-Valued Anal.*, **3**(1995), 213-261

[2] Z. Artstein. Invariant measures of set-valued maps, *J. Mathemat. Analysis Appl.* **252** (2000), 696-709.

[3] J.-P. Aubin, H. Frankowska. *Set-valued analysis*, Birkhäuser, 1990

[4] F. Bernardin. Multivalued Stochastic Differential Equation: Convergence of Numerical Scheme, *Set-Valued Analysis*, **11** (2003), 393-415.

[5] J.M. Borwein, Q.J. Zhu. A survey of subdifferential calculus with applications. *Nonlinear Analysis*,**38** (1999), 687-773.

[6] P. Deheuvels, D.M. Mason. Functional laws of the iterated logarithm for local empirical pro-

cesses indexed by sets, *Ann. Probab.*, **22** (1994), 1619–1661.

[7] P. Deheuvels, D.M. Mason. Nonstandard local empirical processes indexed by sets, *J. Statist. Plann. Inference*, **45** (1995), 91–112.

[8] J.H.J. Einmahl. Poisson and Gaussian approximation of weighted local empirical processes, *Stochastic Process. Appl.*, **70** (1997), 31–58.

[9] H. Federer. Curvature measures, *Trans. Amer. Math. Soc.*, **93** (1959), 418–491.

[10] S. Gautier. Affine and eclipsing multifunctions, *Numer. Funct. Anal. Optim.*, **11**(1990), 679-699

[11] D. Hug, G. Last, W. Weil, A local Steiner-type formula for general closed sets and applications, *Math. Z.* **246** (2004), 237–272.

[12] E. Khmaladze. Goodness of fit tests for "chimeric" alternatives, *Statist. Neerl.*, **52** (1998), 90-111.

[13] Byoung Kyun Kim, Jai Heui Kim. Stochastic Integrals of Set-valued Processes and Fuzzy Processes, *J. Mathemat. Analysis Appl.*, **236** (1999), 480-502.

[14] C. Lemaréchal and J. Zowe. The eclipsing concept to approximate a multi-valued mapping, *Optimization*, **22** (1991), 3-37

[15] G.Ch. Pflug. *Optimization of Stochastic Models*, Kluwer Acad.Publ., Dordrecht, 1996

[16] R.T. Rockafellar and R.J.-B. Wets. *Variational Analysis*, Springer, 1998

[17] R. Schneider, *Convex Bodies: the Brunn-Minkowski Theory*, Encyclopedia of Mathematics and its Applications, **44**, Cambridge University Press, Cambridge, 1993.

[18] D. Silin. On set-valued differentiation and integration. *Set-Valued Analysis*, **5** (1997),107-146.

[19] H. Weisshaupt. A Measure-Valued Approach to Convex Set-Valued Dynamics, *Set-Valued Analysis*, **9**, (2001), 337- 373