The Third International Conference "Problems of Cybernetics and Informatics"
September 6-8, 2010, Baku, Azerbaijan. The plenary paper
www.pci2010.science.az/4/09.pdf

ON TWO SIMPLE STOCHASTIC MODELS

Zurab Kvatadze[1], Tinatin Mumladze[1], and Tengiz Shervashidze[2,3,1]

[1]Georgian Technical University, Tbilisi, Georgia
[2]A. Razmadze Mathematical Institute, Tbilisi, Georgia
[3]I. Vekua Institute of Applied Mathematics of I. Javakhishvili Tbilisi State University, Tbilisi, Georgia, *sher@rmi.acnet.ge*

1. A Simple Stochastic Model of Hail Clouds Emerging. Under a simple stochastic model of hail clouds emerging over a circular domain *C* with the center *O*=(0,0) and radius *R* which identifies a cloud with its plane projection rectangle Δ_ω of fixed sizes $2l\times 2h$, $h<1$, centered at ω and is based on the uniformity of the location of ω in the outer parallel set of the basic rectangle Δ_0, the Minkowski sum $\Delta_0 \oplus C$ [1,2,3], and the isotropy of rectangle orientation on the one hand and identity of probabilities and stochastic independence of covering of *O* by random rectangles Δ_ω on the other hand, and using the normal approximation for the binomial probability distribution of the random number of such coverings the confidence interval is constructed which gives the bounds for the unknown number *n* of hail clouds over the domain by the number ξ of hail clouds observed over the center.

Encouraged by the advanced studies in stochastic modeling completed by R. Chitashvili and E. Khmaladze at I.Vekua Institute of Applied Mathematics, the third author who was involved into stochastic modeling of hail clouds emerging by G. Sulakvelidze performed the above mentioned research at the same institute in early 1970ies. The paper with a description of that research [4] available online is out of print only recently.We decided to present its shortened version to the international audience at PCI2010 in Baku, Azerbaijan.

Under our assumption that each of *n* clouds observable at random and independently from others covers the point *O* with the same probability *p* the random variable ξ has a binomial distribution with the parameters *n* and *p* .Maximum likelihood estimator $\hat{n}$ for *n* under the observed value ξ when *p* is known, equals to

$$\hat{n} = \left[\frac{\xi}{p}\right], \tag{1}$$

where [*x*] is the integral part of a real number *x*.

Assume that an unknown *n* is large enough. Due to the De Moivre–Laplace theorem we obtan, that $|\xi - np| < \sqrt{npq}\, t_\alpha$ with probability α , where $q = 1-p$, $\Phi(t) = (2\pi)^{-1/2}\int_0^t e^{-u^2/2}du$ and t_α is chosen such that $\Phi(t_\alpha) = \alpha/2$ .Solving this inequality with respect to *n*, we have the following asymptotic confidence interval for *n*

$$(\xi/p - a(\xi,p,\alpha), \xi/p + b(\xi,p,\alpha))$$

with the confidence probability α , i.e.,

$$P(\xi/p - a(\xi,p,\alpha) < n < \xi/p + b(\xi,p,\alpha)) \approx \alpha, \tag{2}$$

where

$$a(\xi,p,\alpha) = \frac{\sqrt{t_\alpha^2 q(t_\alpha^2 q + 4\xi)} - t_\alpha^2 q}{2p}, \qquad b(\xi,p,\alpha) = a(\xi,p,\alpha) + \frac{t_\alpha^2}{p}. \tag{3}$$

Now we assign a meaningful value to the probability *p* using the notion of geometric probability. Assume that the cloud is observable from the circle *C* if the above-mentioned rectangle intersects with the circle. Under registration of the cloud over the point *O* let us mean the hitting of the point *O* into the rectangle. Thus we have to calculate the probability that the rectangle $2l \times 2h \ (l > h)$, randomly chosen from those rectangles which intersect with the circle *C* of radius *R*, will cover the center of the circle.

The position of the rectangle on the plane is characterized by that of its center and angle between the fixed line, passing through the point *O*, and the rectangle basis. By the symmetry, we can fix this angle .

For any $u = (s,t) \in R^2$ denote $\Delta_u = [s-l, s+l] \times [t-h, t+h]$ the rectangle of fixed sizes with the center at *u*, Δ_0 being the basic rectangle $[-l,l] \times [-h,h]$. Evidently, $u + \Delta_0 = \Delta_u$ and the inclusions $u \in \Delta_v$ and $v \in \Delta_u$ are equivalent for any two $u, v \in R^2$.

Let us now construct the set Ω of positions of the rectangle center $\omega = (x, y)$ when the rectangle Δ_ω intersects with the circle *C* of radius *R*. Place the origin of the Cartesian coordinate system at *O* and assume that the *Ox*-axis is a straight line for the angle counting out. For the sake of simplicity, we assume that the angle between the rectangle basis and the *Ox*-axis is equal to zero.

From the definition of the Minkowski sum $A \oplus B = \{a + b \mid a \in A, b \in B\}$ of two sets *A* and *B* in Euclidean space (see, e.g., [1], [2], [3]) it is easy to derive the representation

$$\Omega = C \oplus \Delta_0$$

for the set

$$\Omega = \{\omega \mid \Delta_\omega \cap C \neq \varnothing\}$$

as Minkowski sum of the basic rectangle Δ_0 and the given circle *C* called an outer parallel set of Δ_0 [2] (for the proof see [4]).

If instead of Δ_0 a general convex set *K* is meant and K_R denotes its outer parallel set on the distance *R*, then according to [2, Ch. I] we have the following formulas for the perimeter L_R and area (Lebesgue measure) of K_R:

$$L_R = L + 2\pi R, \quad F_R = F + LR + \pi R^2 , \tag{4}$$

where *L* is the perimeter of *K* and *F* is its area..

Thus if we assume that all the positions of ω are uniformly distributed on Ω for the probability that a random rectangle Δ_ω covers the point *O* we obtain from (4)

$$p_{l,h} = \frac{F}{F_R} = \frac{4lh}{4lh + 4(l+h)R + \pi R^2} . \tag{5}$$

(The notation $p_{l,h}$ emphasizes that the probability is calculated for rectangles of fixed sizes.) If extra randomness is introduced assuming that *l* and *h* are random variables with a known joint distribution, then the unknown probability would be equal to mathematical expectation $p = E(p_{l,h})$.

Note that if we can indicate a priori the numbers $\boldsymbol{l}_0$, $\boldsymbol{l}$, $\boldsymbol{h}_0$, $\boldsymbol{h}$ such that

$$\boldsymbol{l}_0 < l < \boldsymbol{l}, \quad \boldsymbol{h}_0 < h < \boldsymbol{h}, \quad \boldsymbol{l}_0 \gg R, \quad \boldsymbol{h}_0 \gg R, \tag{6}$$

then

$$p \approx 4\frac{E(lh)}{\pi R^2} . \tag{7}$$

But if *l* and *h* are not correlated, then

$$p \approx 4\frac{E(l)E(h)}{\pi R^2} . \tag{8}$$

The expectations *E*(*lh*), *E*(*l*), *E*(*h*) may be unknown but on the basis of suitable sampling data they can be approximated reliably by the empirical means $\overline{lh}$, $\overline{l}$ and $\overline{h}$.

Thus with a high reliability

$$p \approx 4\frac{\overline{lh}}{\pi R^2} \qquad (7')$$

and in the case of uncorrelated *l* and *h*, when

$$p \approx 4\frac{\overline{l}\ \overline{h}}{\pi R^2} \qquad (8')$$

(8') can be obtained by the choice from the very beginning of a rectangle of sizes $2\overline{l} \otimes 2\overline{h}$ by passing from

$$p \approx 4\frac{\overline{l}\ \overline{h}}{4\overline{l}\ \overline{h} + 4(\overline{l} + \overline{h})R + \pi R^2} \qquad (5')$$

to (8') under the condition (6).

The set of formulas (1)–(3), (5), (5') and (8') allow us to estimate the unknown number *n*.

Remark 1. According to our best knowledge no proper application of the proposed technique was done. A problem to collect data to test model quality by comparison the values of areas damaged by hail and its model values posed by G. Sulakvelidze and for which later much efforts by J. Mdinaradze were spent without any success, is still open for collaboration.

Remark 2. Note that if in a role of basic set Δ_0 one takes the circle of radius *l* or ellipse with half-axes *l* and *h* $(h < l)$ one obtains some meaningful extensions of our model which may have an interest for, say, biological, ecological and even meteorological modelling. For circle our ratio equals to $p_l = \frac{F}{F_R} = \left(\frac{l}{R+l}\right)^2$. For ellipse we have $p_{l,h} = \frac{F}{F_R} = \frac{\pi lh}{\pi lh + 4lE(e)R + \pi R^2}$, where *E*(*e*) stands for the complete elliptic integral of the second kind and $e = \frac{\sqrt{l^2 - h^2}}{l}$ for the eccentricity of ellipse.

2. A simple stochastic trade model and related quantile optimality. Let now a random variable ξ having a distribution function $F(u)$, $u \ge 0$, which is absolutely continuous with the density function $f(u)$, $u \ge 0$ w.r.t. Lebesgue measure describe the random demand in kg of certain continuously varied goods. If a businessman offers to certain market a quantity *x* of these goods which is less or equal to random demand ξ then he gains USD *A* per kg, i.e. USD *Ax* in whole, if vice versa, he gains USD $A\xi$ and loses USD $B(x-\xi)$. If $G(x,\xi)$ stands for gain function and $I(\xi > x)$ and $I(\xi \le x)$ denote indicators of corresponding events, we have

$$G(x,\xi) = I(\xi > x)Ax + I(\xi \le x)\left[A\xi - B(x-\xi)\right]$$

and under the notation $\eta_+ = \eta I(\eta \ge 0)$ for the nonnegative part of a random variable η we obtain, that

$$G(x,\xi) = Ax - (A+B)(x-\xi)_+ .$$

Introducing notation for the expectation $M(x) = EG(x,\xi)$ we have

$$M(x) = Ax - (A+B)E(x-\xi)_+$$

and performing simple calculations

$$E(x-\xi)_+ = \int_0^x (x-u)dF(u) = xF(x) - \int_0^x udF(u) = xF(x) - uF(u)\Big|_0^x + \int_0^x F(u)du = \int_0^x F(u)du$$

we obtain

$$M(x) = Ax - (A+B)\int_0^x F(u)du .$$

Now let us find $x^* = \arg\max_x M(x)$. Taking derivative of $M(x)$, we obtain

$$M'(x) = A - (A+B)F(x) = 0,$$

that is

$$x = F^{-1}\left(\frac{A}{A+B}\right).$$

Further, for a.e. $x \in \text{supp}(F)$ we have $M''(x) = -(A+B)F'(x) > 0$,and

$$x^* = \arg\max_x M(x) = F^{-1}\left(\frac{A}{A+B}\right) .$$

As for variance

$$D(x) = DG(x,\xi) = (A+B)^2 D(x-\xi)_+ = (A+B)^2\left[E(x-\xi)_+^2 - \left(E(x-\xi)_+\right)^2\right],$$

we have

$$D(x) = (A+B)^2\left[2\int_0^x (x-u)F(u)du - \left(\int_0^x F(u)du\right)^2\right], D'(x) = (A+B)^2 2(1-F(x))\int_0^x F(u)du$$

and

$$D''(x) = 2(A+B)^2\left[F(x)(1-F(x)) - f(x)\int_0^x F(u)du\right].$$

This leads to

$$D(x) = 2(A+B)^2\int_0^x\left[(1-F(x))\int_0^t F(u)du\right]dt$$

and special properties of $x = 0$ and finite $x = b = F^{-1}(1)$ of having least and greatest variances.

For the known $A/(A+B)$ -quantile, we have a solution of the initial optimality problem and when it is unknown we obtain the following interesting statistical problem

How to choose $x_1, x_2, \ldots, x_m$ such that by mn independent observations $\xi_{i1}, \ldots \xi_{in}$, $i = 1, \ldots m$, to construct a good procedure based on the vector

$$\left(G(x_i, \xi_{i1}) + \cdots + G(x_i, \xi_{in})\right)/n, \quad i = 1, \ldots m,$$

to estimate

$$\arg\max_{x>0} M(x) = F^{-1}\left(A/(A+B)\right)?$$

References

1. Matheron G. Random Sets and Integral Geometry, Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley & Sons, 1975.
2. Santalo L.A. Integral Geometry and Geometric Probability, with a foreword by Mark Kac. Encyclopedia of Mathematics and its Applications, Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1, 1976; Russian translation: Moscow, Nauka, 1983.
3. Schneider R. Convex Bodies: the Brunn-Minkowski Theory, Encyclopedia of Mathematics and its Applications Cambridge University Press, Cambridge, 44, 1993.
4. Shervashidze T. On a simple stochastic model of hail clouds emerging over a circular domain. Sem. I. Vekua Inst. Appl. Math. Rep. 32–33 (2006—2007), 15–19. http://www.viam.science.tsu/report/32-33/shervashidze.pdf