УДК 519.24

Т.В. Цеховая, Н.Н. Труш

ПРЕДЕЛЬНОЕ РАСПРЕДЕЛЕНИЕ ОЦЕНОК СЕМИВАРИОГРАММЫ СТАЦИОНАРНЫХ СЛУЧАЙНЫХ ПОЛЕЙ

В данной статье построена оценка семивариограммы стационарного в узком смысле m - зависимого двумерного случайного поля с дискретным временем. Получены выражения для математического ожидания и дисперсии исследуемой статистики. Найдено предельное распределение изучаемой оценки семивариограммы.

Ключевые слова: *случайное поле, m - зависимое случайное поле, семивариограмма, бивариограмма, оценка, предельное распределение.*

В настоящее время для исследования задач геологии, гидрологии, метеорологии, экологии и других областей знаний наряду с использованием классических методов теории случайных процессов, статистического анализа временных рядов, возникла необходимость разработки и применения новых методов.

Вариограммный анализ представляет собой сравнительно новый раздел статистического анализа временных рядов, применяемый для исследования таких задач. Вариограмма является одной из основных характеристик случайных процессов и полей во временной области.

За последние десятилетия значительно увеличился интерес к исследованию статистических свойств оценок вариограммы. Это можно объяснить тем, что оценки несут полезную и необходимую информацию о внутренней структуре процессов и полей.

Асимптотические распределения оценок вариограммы стационарных в широком смысле случайных процессов с дискретным и непрерывным временем находились, например, в [1, 2]. Статистические свойства оценки вариограммы гауссовского случайного процесса исследовались в статье [3]. Данная работа посвящена нахождению предельного распределения оценок семивариограммы стационарного в узком смысле случайного поля.

Введем определения [4], которые понадобятся нам в дальнейшем.

Определение. Вариограммой k – мерного случайного поля $X(s)$, $s \in \mathbf{Z}^k$, $\mathbf{Z} = \{0, \pm 1, \pm 2, \ldots\}$, $k \in \mathbf{N}$, $\mathbf{N}$ – множество натуральных чисел, называется функция

$$2\gamma(s) = D\{X(t+s) - X(t)\},$$

$t, s \in \mathbf{Z}^k$, функция $\gamma(s)$ – семивариограмма рассматриваемого поля.

Определение. Бивариограммой k – мерного случайного поля $X(s)$, $s \in \mathbf{Z}^k$, называется функция вида

$$\beta(s_1, s_2, s_3) = 0{,}25\, M\{(X(s_1+t) - X(t))^2\,(X(s_3+t) - X(s_2+t))^2\},$$

$s_1, s_2, s_3, t \in \mathbf{Z}^k$.

Определение. Случайное поле $X(s)$, $s \in \mathbf{N}^k$, называется m – зависимым, $m, k \in \mathbf{N}$, если $X(s_1)$, $s_1 \in C$, и $X(s_2)$, $s_2 \in D$, независимы для любых подмножеств C и D множества $\mathbf{N}^k$, для которых выполняется условие

$$d(C, D) > m,$$

где

$$d(C, D) = \inf\{\| s_1 - s_2 \|\},\ s_1 \in C,\ s_2 \in D,$$

$$\| s \| = \max\{| s^{(i)} |,\ i = 1, \ldots, k\},\ s = (s^{(1)}, s^{(2)}, \ldots, s^{(k)}) \in \mathbf{N}^k.$$

Заметим, что в случае $k = 1$ вышеуказанные определения соответствуют одномерному случайному процессу.

Рассмотрим стационарное в широком смысле m - зависимое случайное поле $X(t)$, $t \in \mathbf{Z}^2$, с нулевым математическим ожиданием и ковариационной функцией $R_X(t)$, $t \in \mathbf{Z}^2$, $m \in \mathbf{N}$.

Пусть имеется выборка $X(t^{(1)}, t^{(2)})$ объема n^2, $n \in \mathbf{N}$, $n >> m$, двумерного поля $X(t)$, $t \in \mathbf{Z}^2$, в точках

$$Q = \{(t^{(1)},\ t^{(2)}),\ \ t^{(j)} = \overline{1,n},\ \ j = 1,2\}.$$

Обозначим $X_n = \sum\limits_{\substack{t^{(j)}=1\\ j=1,2}}^{n} X(t^{(1)},\ t^{(2)})$. Докажем вспомогательный результат.

Лемма 1. Дисперсия статистики X_n удовлетворяет следующему соотношению:

$$D\{X_n\} = n^2 \sum_{\substack{t^{(j)}=-m\\ j=1,2}}^{m} R_X(t^{(1)},\ t^{(2)}) \prod_{k=1}^{2} \Big(1 - \frac{| t^{(k)} |}{n}\Big), \tag{1}$$

где $R_X(t^{(1)}, t^{(2)})$, $t^{(1)}, t^{(2)} \in \mathbf{Z}$, – ковариационная функция случайного поля $X(t)$, $t \in \mathbf{Z}^2$.

Доказательство. Используя определение дисперсии, стационарность рассматриваемого поля, запишем

$$D\{X_n\} = \sum_{\substack{t^{(j)}=1\\ j=1,2}}^{n} \ \sum_{\substack{s^{(j)}=1\\ j=1,2}}^{n} \operatorname{cov}\{X(t^{(1)},\ t^{(2)}), X(s^{(1)},\ s^{(2)})\} =$$

.

$$= \sum_{\substack{t^{(j)}=1\\ j=1,2}}^{n} \sum_{\substack{s^{(j)}=1\\ j=1,2}}^{n} R_X(t^{(1)} - s^{(1)},\ t^{(2)} - s^{(2)}) .$$

Сделаем замену переменных суммирования: $t - s = t$, $s = s$, где $t = (t^{(1)}, t^{(2)})$, $s = (s^{(1)}, s^{(2)})$. Тогда получим

$$D\{X_n\} = \sum_{\substack{t^{(j)}=1-n\\ j=1,2}}^{0} \sum_{\substack{s^{(j)}=1-t^{(j)}\\ j=1,2}}^{n} R_X(t^{(1)},\ t^{(2)}) + \sum_{\substack{t^{(j)}=1\\ j=1,2}}^{n-1} \sum_{\substack{s^{(j)}=1\\ j=1,2}}^{n-t^{(j)}} R_X(t^{(1)},\ t^{(2)}) =$$

$$= \sum_{\substack{t^{(j)}=1-n\\ j=1,2}}^{0} R_X(t^{(1)},\ t^{(2)}) \prod_{k=1}^{2} (n + t^{(k)}) + \sum_{\substack{t^{(j)}=1\\ j=1,2}}^{n-1} R_X(t^{(1)},\ t^{(2)}) \prod_{k=1}^{2} (n - t^{(k)}) =$$

$$= \sum_{\substack{t^{(j)}=1-n\\ j=1,2}}^{n-1} R_X(t^{(1)},\ t^{(2)}) \prod_{k=1}^{2} (n - |t^{(k)}|).$$

В силу того, что случайное поле $X(t)$, $t \in \mathbf{Z}^2$, является m - зависимым, $m << n$, имеем

$$D\{X_n\} = \sum_{\substack{t^{(j)}=-m\\ j=1,2}}^{m} R_X(t^{(1)},\ t^{(2)}) \prod_{k=1}^{2} (n - |t^{(k)}|) ,$$

что и требовалось доказать.

Рассмотрим m - зависимое стационарное в узком смысле случайное поле $Y(t)$, $t \in \mathbf{Z}^2$, с нулевым математическим ожиданием, конечным вторым моментом, неизвестными семивариограммой $\gamma(t)$, $t \in \mathbf{Z}^2$, и бивариограммой $\beta(t_1, t_2, t_3)$, $t_1, t_2, t_3 \in \mathbf{Z}^2$.

Для фиксированных $n \in \mathbf{N}$ и вектора $h = (h^{(1)}, h^{(2)})$, имеющего направление одной из координатных осей, где $h^{(1)}, h^{(2)} = \overline{0, n-1}$, обозначим

$$T^h(t) = \frac{(Y(t+h) - Y(t))^2}{2N(h)} , \tag{2}$$

где $t = (t^{(1)},\ t^{(2)})$, $t^{(j)} = \overline{1, n - h^{(j)}}$, $j = 1, 2$, $N(h) = (n - h^{(1)})(n - h^{(2)})$.

Теорема 1. Случайное поле $T^h(t)$, $t = (t^{(1)},\ t^{(2)})$, $t^{(j)} = \overline{1, n - h^{(j)}}$, $j = 1, 2$, $h^{(1)}, h^{(2)} = \overline{0, n-1}$, $n \in \mathbf{N}$, вида (2) является $(m + |h|)$ - зависимым, где $|h|$ – длина вектора $h = (h^{(1)}, h^{(2)})$.

Доказательство. Из определения m – зависимого k – мерного случайного поля вытекает, что $T^h(t_1)$ и $T^h(t_2)$ независимы, если

$$d(t_1, t_2) > m ,$$
$$d(t_1 + h, t_2) > m ,$$
$$d(t_1, t_2 + h) > m ,$$
$$d(t_1 + h, t_2 + h) > m .$$

Объединяя полученные условия, имеем $d(t_1, t_2) > m + |h|$, $m \in \mathbf{N}$. Теорема доказана.

Теорема 2. Поле $T^h(t)$, $t = (t^{(1)},\ t^{(2)})$, $t^{(j)} = \overline{1, n - h^{(j)}}$, $j = 1, 2$, $h^{(1)}, h^{(2)} = \overline{0, n-1}$, $n \in \mathbf{N}$, является стационарным в узком смысле.

Доказательство. В силу определения l - мерной функции распределения случайного поля $T^h(t)$, учитывая равенство (2), запишем

$$F(x_1, ...,\ x_l; t_1 + \tau, ...,\ t_l + \tau) = P\{T^h(t_1 + \tau) < x_1, ..., T^h(t_l + \tau) < x_l\} =$$
$$= P\{\frac{(Y(t_1 + h + \tau) - Y(t_1 + \tau))^2}{2N(h)} < x_1, ..., \frac{(Y(t_l + h + \tau) - Y(t_l + \tau))^2}{2N(h)} < x_l\},$$

$x_i \in \mathbf{R}$, $t_i \in \mathbf{Z}^2$, $i = \overline{1, l}$, $\tau \in \mathbf{Z}^2$.

Поскольку $Y(t)$, $t \in \mathbf{Z}^2$, стационарное в узком смысле случайное поле, то

$$F(x_1, ...,\ x_l;\ t_1 + \tau, ...,\ t_l + \tau) =$$
$$= P\{\frac{(Y(t_1 + h) - Y(t_1))^2}{2N(h)} < x_1, ...,\ \frac{(Y(t_l + h) - Y(t_l))^2}{2N(h)} < x_l\} =$$
$$= P\{T^h(t_1) < x_1, ...,\ T^h(t_l) < x_l\} = F(x_1, ...,\ x_l;\ t_1, ...,\ t_l).$$

Таким образом, получаем требуемое.

Заметим, что поле $T^h(t)$, $t=(t^{(1)},\ t^{(2)})$, $t^{(j)}=\overline{1,n-h^{(j)}}$, j = 1, 2, $h^{(1)},h^{(2)}=\overline{0,n-1}$, $n\in$ **N**, является также стационарным в широком смысле. В силу определения вариограммы случайного поля $Y(t)$, $t\in \mathbf{Z}^2$,

$$M\{T^h(t)\}=\frac{M\{(Y(t+h)-Y(t))^2\}}{2N(h)}=\frac{\gamma(h)}{N(h)}.$$

Покажем, что для поля $T^h(t)$ существует момент второго порядка. Из определения бивариограммы, учитывая явный вид (2) поля $T^h(t)$, имеем

$$\beta(h,0,h)=0{,}25\,M\{Y(t+h)-Y(t)\}^4=N^2(h)\,M\{T^h(t)\}^2.$$

Пусть для фиксированных $n\in$ **N** и $h=(h^{(1)},h^{(2)})$, где $h^{(1)},h^{(2)}=\overline{0,n-1}$, выполняется условие $\dfrac{\beta(h,0,h)}{N^2(h)}<\infty$, тогда $M\{T^h(t)\}<\infty$.

Отсюда и из утверждения теоремы 2 вытекает, что $T^h(t)$ является стационарным в широком смысле случайным полем.

В качестве оценки семивариограммы двумерного случайного поля $Y(t)$, $t\in \mathbf{Z}^2$, рассмотрим статистику

$$\hat{\gamma}(h)=\sum_{i=1}^{N(h)}T^h(t_i)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}(Y(t_i+h)-Y(t_i))^2, \tag{3}$$

построенную по выборке $Y(t_i)=Y(t_i^{(1)},\ t_i^{(2)})$, $t_i^{(j)}=\overline{1,n}$, $j=1,2$, $i=\overline{1,n^2}$, $n\in$ **N**, случайного поля $Y(t)$, $t\in \mathbf{Z}^2$, причем t_i, $t_i+h\in Q$, h имеет направление одной из координатных осей.

Нетрудно видеть, что $M\{\hat{\gamma}(h)\}=\gamma(h)$. Найдем выражение для момента второго порядка статистики (3).

Лемма 2. Дисперсия оценки семивариограммы $\hat{\gamma}(h)$ вида (3) удовлетворяет соотношению

$$D\{\hat{\gamma}(h)\}=N(h)\sum_{\substack{t^{(k)}=-(m+|h|)\\ k=1,2}}^{m+|h|}R(t^{(1)},\ t^{(2)})\prod_{l=1}^{2}\Big(1-\frac{|t^{(l)}|}{n-h^{(l)}}\Big), \tag{4}$$

где $R(t^{(1)},\ t^{(2)})$, $t^{(1)},\ t^{(2)}\in$ **Z**, – ковариационная функция поля $T^h(t)$, t = $(t^{(1)},\ t^{(2)})$, $t^{(j)}=\overline{1,n-h^{(j)}}$, j = 1, 2, $h^{(1)},h^{(2)}=\overline{0,n-1}$, $n\in$ **N**.

Доказательство. Используя определение дисперсии, стационарность рассматриваемого поля, запишем

$$D\{\hat{\gamma}(h)\}=\sum_{i,\,j=1}^{N(h)}\operatorname{cov}\{T^h(t_i),T^h(t_j)\}=\sum_{i,\,j=1}^{N(h)}R(t_i-t_j).$$

где $t_i=(t_i^{(1)},\ t_i^{(2)})$, $R(t)$, $t\in \mathbf{Z}^2$, – ковариационная функция поля $T^h(t)$ вида (2).

Поскольку $t^{(j)}=\overline{1,n-h^{(j)}}$, j = 1, 2, то по аналогии с доказательством леммы 1 дисперсию оценки (3) семивариограммы можно представить в виде

$$D\{\hat{\gamma}(h)\}=\sum_{\substack{t_i^{(k)}=1\\ k=1,2}}^{n-h^{(k)}}\ \sum_{\substack{t_j^{(k)}=1\\ k=1,2}}^{n-h^{(k)}}R(t_i-t_j).$$

Сделаем замену переменных суммирования: $t_i-t_j=t$, $t_j=s$, где $t=(t^{(1)},t^{(2)})$, $s=(s^{(1)},s^{(2)})$.

$$D\{\hat{\gamma}(h)\}=\sum_{\substack{t^{(k)}=1-n+h^{(k)}\\ k=1,2}}^{0}\ \sum_{\substack{s^{(k)}=1-t^{(k)}\\ k=1,2}}^{n-h^{(k)}}R(t)+\sum_{\substack{t^{(k)}=1\\ k=1,2}}^{n-h^{(k)}-1}\ \sum_{\substack{s^{(k)}=1\\ k=1,2}}^{n-h^{(k)}-t^{(k)}}R(t)=$$

$$=\sum_{\substack{t^{(k)}=1-n+h^{(k)}\\ k=1,2}}^{0}R(t)\prod_{l=1}^{2}(n-h^{(l)}+t^{(l)})\ +\sum_{\substack{t^{(k)}=1\\ k=1,2}}^{n-h^{(k)}-1}R(t)\prod_{l=1}^{2}(n-h^{(l)}-t^{(l)})=$$

$$=\sum_{\substack{t^{(k)}=1-n+h^{(k)}\\ k=1,2}}^{n-h^{(k)}-1}R(t)\prod_{l=1}^{2}(n-h^{(l)}-|t^{(l)}|).$$

Учитывая утверждение теоремы 1, имеем

$$D\{\hat{\gamma}(h)\}=N(h)\sum_{\substack{t^{(k)}=-(m+|h|)\\ k=1,2}}^{m+|h|}R(t)\prod_{l=1}^{2}\Big(1-\frac{|t^{(l)}|}{n-h^{(l)}}\Big),$$

что и требовалось доказать.

Замечание. Если имеет место соотношение $\sum_{\substack{t^{(j)}=-\infty \\ j=1,2}}^{+\infty} |R(t)| < \infty$, $t = (t^{(1)}, t^{(2)})$, то легко показать, что $\lim_{N(h)\to\infty} D\{\hat{\gamma}(h)\} = 0$.

В силу соотношения $M\{\hat{\gamma}(h)\} = \gamma(h)$ и вышеуказанного замечания получаем, что $\hat{\gamma}(h)$ является состоятельной в среднеквадратическом смысле оценкой для семивариограммы $\gamma(h)$, $h \in \mathbf{Z}^2$, случайного поля $Y(t)$, $t \in \mathbf{Z}^2$.

Найдем предельное распределение статистики (3). Обозначим через $\xrightarrow{D}$ сходимость по распределению.

Теорема 3. Оценка семивариограммы $\hat{\gamma}(h)$ вида (3) удовлетворяет предельному соотношению

$$\frac{\hat{\gamma}(h) - \gamma(h)}{\sqrt{D\{\hat{\gamma}(h)\}}} \xrightarrow{D} N(0,1) \quad \text{при} \quad N(h) \to \infty,$$

где $\gamma(h)$, $h \in \mathbf{Z}^2$, – семивариограмма случайного поля $Y(t)$, $t \in \mathbf{Z}^2$, а дисперсия $D\{\hat{\gamma}(h)\}$ задана равенством (4).

Доказательство непосредственным образом вытекает из утверждения теоремы работы [4, стр.191].

ЛИТЕРАТУРА

1. *Цеховая Т. В*. Предельное распределение оценки вариограммы стационарного случайного процесса // Вестник БГУ. Сер. 1, Физ. Мат. Мех. 2002. №1. С. 104 – 105.
2. *Tsekhavaya. T. V., Troush N. N.* Asymptotic distribution of the variogram estimator of stationary stochastic process with continuous time // Transaction of XXIV International Seminar on Stability Problems for Stochastic Models, Jurmala, Latvia, September 10 – 17, 2004 / Transport and Telecommunication Institute. Jurmala. 2004. P. 292 – 297.
3. *Цеховая Т. В.* Асимптотическое распределение оценки вариограммы // Вестник БрГУ им. А.С. Пушкина. Серия естественных наук. 2008. №2(31). С. 32 – 37.
4. *Davis B.M., Borgman L.E.* A Note on the Asymptotic Distribution of the Sample Variogram // Jour. Inter. Assoc. Mathematical Geology. 1982. Vol. 14. № 2. P. 189 – 193.