ПРЕДЕЛЬНОЕ РАСПРЕДЕЛЕНИЕ ОЦЕНОК ПЕРВЫХ ДВУХ МОМЕНТОВ СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА С ПУАССОНОВСКИМИ НЕРЕГУЛЯРНЫМИ НАБЛЮДЕНИЯМИ

Т. И. Воротницкая

Белорусский государственный университет
Минск, Беларусь
E-mail: iliukevich@bsu.by

Исследуется предельное распределение оценок математического ожидания и ковариационной функции стационарного случайного процесса с пуассоновскими пропусками наблюдений.

Ключевые слова: m-зависимый случайный процесс, математическое ожидание, ковариационная функция, амплитудная модуляция, кумулянтный подход.

Рассмотрим стационарный в широком смысле случайный процесс $X(t)$, $t \in Z$ с нулевым математическим ожиданием m^X, ковариационной функцией $R^X(\tau)$, $\tau \in Z$, спектральной плотностью $f^X(\lambda)$, $\lambda \in \Pi = [-\pi, \pi]$, семиинвариантной спектральной плотностью четвертого порядка $f_4^X(\lambda_1, \lambda_2, \lambda_3)$, $\lambda_i \in \Pi$, $i = \overline{1,3}$, смешанным моментом четвертого порядка $m_4^X(t_1, t_2, t_3)$, $t_i \in \mathrm{Z}$, $i = \overline{1,3}$ и смешанным семиинвариантом четвертого порядка $c_4^X(t_1, t_2, t_3)$, $t_i \in \mathrm{Z}$, $i = \overline{1,3}$.

Пусть в результате некоторого эксперимента получено T последовательных наблюдений через равные промежутки времени

$$Y(0), Y(1), \ldots, Y(T-1) \tag{1}$$

за процессом $Y(t)$, $t \in \mathrm{Z}$, который связан с процессом $X(t)$, $t \in Z$ следующим соотношением

$$Y(t) = X(t)d(t), \tag{2}$$

где $d(t)$, $t \in Z$ - последовательность независимых случайных величин, распределенных по закону Пуассона с параметром $\alpha > 0$. Предположим, что $d(t)$, $t \in Z$ не зависит от процесса $X(t)$, $t \in Z$.

В качестве оценки математического ожидания стационарного случайного процесса, построенной по наблюдениям (1) за процессом $Y(t)$, рассмотрим статистику вида

$$\hat{m}^X = \frac{1}{T\alpha} \sum_{t=0}^{T-1} Y(t), \tag{3}$$

для которой имеет место следующий результат.

Теорема 1 [6]. Для оценки математического ожидания (3) справедливы следующие утверждения:

1. Если ковариационная функция $R^X(\tau)$ такова, что

$$\sum_{\tau=0}^{\infty} R^X(\tau) < \infty,$$

то

$$\lim_{T\to\infty} TD\hat{m}^X = 2\sum_{\tau=0}^{\infty} R^X(\tau) + \frac{R^X(0)(1-\alpha)}{\alpha} + \frac{(m^X)^2}{\alpha}.$$

2. Если спектральная плотность $f^X(\lambda)$ непрерывна в точке $\lambda = 0$ и ограничена на Π, то

$$\lim_{T\to\infty} TD\hat{m}^X = 2\pi f^X(0) + \frac{1}{\alpha}\int_{\Pi} f^X(y)dy + \frac{(m^X)^2}{\alpha}.$$

Из теоремы 1 следует, что оценка (3) является состоятельной в среднеквадратическом смысле оценкой для математического ожидания процесса $X(t)$.

Найдем предельное распределение статистики $\hat{m}^X$ в случае, когда $X(t)$ является m-зависимым стационарным случайным процессом.

Определение [5]. Случайный процесс $X(t)$ называется m-зависимым, если для любого целого n, любых $t_j \in Z$, $j = \overline{1,n}$, $(0 < t_1 < t_2 < \ldots < t_n)$ $\{X(t_1),\ldots X(t_n)\}$ распределены независимо от $\{\ldots, X(t_1 - m - 1)\}$ и $\{X(t_n + m + 1), \ldots\}$ для некоторого m, m=1, 2,….

Теорема 2. Пусть $X(t)$, $t \in Z$ - m-зависимый случайный процесс с нулевым математическим ожиданием, тогда статистика

$$\sqrt{T}\,\hat{m}^X = \frac{1}{\sqrt{T}\alpha}\sum_{t=0}^{T-1} Y(t) \qquad (4)$$

имеет предельное нормальное распределение с нулевым математическим ожиданием и дисперсией, равной

$$\frac{\alpha^2 + \alpha}{\alpha^2} R^X(0) + 2R^X(1) + \ldots + 2R^X(m).$$

Доказательство. Из формулировки теоремы следует, что математическое ожидание оценки (4) равно нулю. Найдем дисперсию оценки, учитывая независимость $X(t)$ и $d(t)$, $t \in Z$. Имеем

$$D(\sqrt{T}\,\hat{m}^X) = \frac{1}{T\alpha^2}\sum_{t=0}^{T-1}\sum_{s=0}^{T-1} MX(t)X(s)Md(t)d(s) =$$

$$= \frac{1}{T\alpha^2}\sum_{t=0}^{T-1} MX^2(0)\left(\alpha + \alpha^2\right) + \frac{1}{T\alpha^2}\sum_{\substack{t,s=0\\ t\neq s}}^{T-1} MX(t)X(s)\alpha^2.$$

Так как процесс $X(t)$, $t \in Z$ является m-зависимым и стационарным, перепишем правую часть выражения дисперсии оценки (4) в виде

$$\frac{1}{T}\left\{\frac{T\left(\alpha + \alpha^2\right)}{\alpha^2} R^X(0) + 2(T-1)R^X(1) + 2(T-2)R^X(2) + \ldots + 2(T-m)R^X(m)\right\}.$$

Следовательно

$$\lim_{T\to\infty} D(\sqrt{T}\,\hat{m}^X) = \frac{(\alpha+\alpha^2)}{\alpha^2} R^X(0) + 2R^X(1) + 2R^X(2) + \ldots + 2R^X(m) = W .$$

В дальнейшем доказательство теоремы проведем, используя подход, предложенный при доказательстве теоремы Хоэффдинга и Робинса [3]. Пусть $T = hk + r$, где $k > 2m$, $0 \le r < k$. Представим статистику (4) в следующем виде

$$\sqrt{T}\,\hat{m}^X = S' + S'',$$

где

$$S' = \frac{1}{\sqrt{T}} \sum_{i=1}^{h} U_i,$$

$$S'' = \frac{1}{\sqrt{T}} \sum_{i=1}^{h} V_i,$$

$$U_i = \sum_{j=0}^{k-m-1} \frac{Y((i-1)k+j)}{\alpha}, i = \overline{1,h},$$

$$V_i = \begin{cases} \sum\limits_{j=0}^{m-1} \dfrac{Y(ik-m+j)}{\alpha}, i = \overline{1,h-1}, \\ \sum\limits_{j=0}^{m+r-1} \dfrac{Y(hk-m+j)}{\alpha}, i = h. \end{cases}$$

Статистика S' есть сумма независимых и одинаково распределенных слагаемых. Далее имеем

$$MU_i = \sum_{j=0}^{k-m-1} \frac{MX[(i-1)k+j]Md[(i-1)k+j]}{\alpha} = (k-m)m^X .$$

Учитывая стационарность процесса $X(t)$, получим

$$DU_i = \frac{1}{\alpha^2} \sum_{j,s=0}^{k-m-1} MX[(i-1)k+j]X[(i-1)k+s]Md[(i-1)k+j]d[(i-1)k+j] =$$

$$= k\left\{(1-\frac{m}{k})\frac{(\alpha^2+\alpha)}{\alpha^2} R^X(0) + 2\left(1-\frac{m+1}{k}\right)R^X(1) + \ldots + 2\left(1-\frac{2m}{k}\right)R^X(m)\right\} = kW_k .$$

Применив центральную предельную теорему для одинаково распределенных случайных величин, получим, что статистика $\sqrt{\frac{T}{h}}S'$ имеет предельное нормальное распределение с математическим ожиданием $\sqrt{h}(k-m)m^X$ и дисперсией kW_k при $h \to \infty$.

Следовательно, $\sqrt{\frac{T}{hk}}S'$ имеет предельное нормальное распределение с математическим ожиданием $\sqrt{h}k\left(1-\frac{m}{k}\right)m^X$ и дисперсией W_k. Так как $W_k \xrightarrow[k\to\infty]{} W$, а

$\sqrt{T}\left(1-\frac{m}{k}\right)m^X \xrightarrow[k\to\infty]{} \sqrt{T}m^X$, то, применив теорему Бернштейна [3], получим требуемый результат.

Для окончательного доказательства теоремы покажем, что статистика $S^{''}$ сходится по вероятности к нулю. Вычислим дисперсию $V_i, i=\overline{1,h}$

$$DV_i=\begin{cases}\frac{1}{\alpha^2}\sum\limits_{t,j=0}^{m-1}MX(ik-m+t)X(ik-m+j)+\left(\frac{m}{\alpha}-\frac{m}{\alpha^2}\right)MX^2(0), i=\overline{1,h-1}.\\ \frac{1}{\alpha^2}\sum\limits_{t,j=0}^{m+r-1}MX(hk-m+t)X(hk-m+j)+\left(\frac{m+r}{\alpha}-\frac{m+r}{\alpha^2}\right)MX^2(0), i=h.\end{cases}$$

Далее, используя неравенство Коши-Буняковского, получим

$$DV_i\le\begin{cases}\frac{m^2}{\alpha^2}MX^2(0)+\left(\frac{m}{\alpha}-\frac{m}{\alpha^2}\right)MX^2(0), i=\overline{1,h-1},\\ \frac{(m+r)^2}{\alpha^2}MX^2(0)+\left(\frac{m+r}{\alpha}-\frac{m+r}{\alpha^2}\right)MX^2(0), i=h.\end{cases}$$

Таким образом,

$$DS^{''}\le\frac{1}{T}R^X(0)\left\{\frac{m^2}{\alpha^2}+\left(\frac{m}{\alpha}+\frac{m+r}{\alpha}\right)\left(1-\frac{1}{\alpha}\right)+\frac{(m+r)^2}{\alpha^2}\right\},$$

следовательно, $DS^{''}\xrightarrow[T\to\infty]{P}0$. Теорема доказана.

В качестве оценки ковариационной функции процесса $X(t)$, $t\in Z$ построенной по наблюдениям (1) рассмотрим статистику вида [2]

$$\hat{R}^X(\tau)=\frac{1}{(T-\tau)C_\tau^d}\sum_{t=0}^{T-\tau-1}Y(t+\tau)Y(t), \tau=\overline{0,T-1}, \tag{5}$$

$$\hat{R}^X(-\tau)=\hat{R}^X(\tau), \hat{R}^X(\tau)=0, |\tau|\ge T,$$

где

$$C_\tau^d=\begin{cases}\alpha+\alpha^2, \tau=0,\\ \alpha^2, \tau\ne 0,\end{cases} \tag{6}$$

α - параметр пуассоновского распределения.

Дисперсия оценки (5) удовлетворяет следующему соотношению

$$\begin{aligned}(T-\tau)D\hat{R}^X(\tau)&=2\pi\int_\Pi\Phi_{T-\tau}(z)\left[\iint_{\Pi^2}f^X(x_1,z-x_1,x_3)e^{i\tau(x_1+x_3)}dx_1dx_3+\right.\\ &\left.+\int_\Pi f^X(x_1)f^X(z-x_1)dx_1+\int_\Pi f^X(x_1)f^X(z-x_1)e^{i\tau(z-2x_1)}dx_1\right]dz+\\ &+K\left[\iiint_{\Pi^3}f^X(x_1,x_2,x_3)e^{i\tau(x_1+x_3)}dx_1dx_2dx_3+2\left(\int_\Pi f^X(x)e^{i\tau x}dx\right)^2+\left(\int_\Pi f^X(x)dx\right)^2\right],\end{aligned} \tag{7}$$

где

$$K = \frac{(\alpha+\alpha^2)^2}{\alpha^4} - 1, \text{ при } \tau \neq 0,$$

$$K = \frac{\alpha^3 + 6\alpha^2 + 7\alpha + 1}{1+\alpha} - 1, \text{ при } \tau = 0.$$

Для нахождения предельного распределения оценки ковариационной функции нам потребуется вычислить кумулянт n-го порядка

$$cum\left\{\hat{R}^X(\tau_1),\ldots,\hat{R}^X(\tau_n)\right\}, \tau_i = \overline{0,T-1}, i = \overline{1,n}$$

и исследовать его асимптотическое поведение. Введем некоторые вспомогательные определения и утверждения.

Рассмотрим таблицу:

$$\begin{array}{cc} (1,1) & (1,2) \\ (2,1) & (2,2) \\ \ldots & \\ (n,1) & (n,2) \end{array} \tag{8}$$

и обозначим P={(1,1), (1,2), (2,1), (2,2), …, (n,1), (n,2)}.

Теорема 3 [1]. Рассмотрим набор случайных величин $x_{ij}, j = 1,\ldots,J_i, i = \overline{1,I}$. Введем I случайных величин $y_i = x_{i1}x_{i2}, i = \overline{1,I}$. Тогда совместный кумулянт $cum(y_1,\ldots,y_I)$ задается формулой

$$cum\{y_1,\ldots,y_I\} = \sum_{\bigcup_{k=1}^{M} P_k = P} cum\{x_{ij};(i,j)\in P_1\}\times\ldots\times cum\{x_{ij};(i,j)\in P_M\},$$

где суммирование ведется по всем неразложимым разбиениям множества P таблицы (8), а $cum\{x_{ij},(i,j)\in P_k\}$ означает, что берется смешанный семиинвариант $cum\{x_{i_1 j_1},\ldots,x_{i_{p_k} j_{p_k}}\}$, где $\{(i_1,j_1),\ldots,(i_{p_k},j_{p_k})\} = P_k, k = \overline{1,M}$.

Следующая лемма понадобится при доказательстве теорем о предельном распределении исследуемых статистик.

Лемма 1 [1]. Пусть $Y^{(T)}, T = 1,2,\ldots,$ будет такой последовательностью случайных векторов с r комплексными переменными, что все кумулянты величины $\left\{Y_1^{(T)},\overline{Y}_1^{(T)},\ldots,Y_r^{(T)},\overline{Y}_r^{(T)}\right\}$ существуют и стремятся к кумулянтам величины $\left\{Y_1,\overline{Y}_1,\ldots,Y_r,\overline{Y}_r\right\}$, которая определяется своими моментами. Тогда $Y^{(T)}$ сходится по распределению к величине, имеющей компоненты $Y_1,\ldots,Y_r$.

Теорема 4 [4]. Если все семиинварианты процесса $X_n^r(t) = \left\{X_a^{(n)}(t), a = \overline{1,r}\right\}$, $t\in Z, n = 1,2,\ldots,$ сходятся к семиинвариантам процесса $X^r(t) = \left\{X_a(t), a = \overline{1,r}\right\}$ при $n\to\infty$, а конечномерные распределения процесса $X^r(t)$ однозначно определяются своими моментами, то конечномерные распределения процесса $X_n^r(t)$ при $n\to\infty$ сходятся к соответствующим конечномерным распределениям процесса $X^r(t)$, $t\in Z$.

Следствие [4]. Если все семиинварианты действительного случайного процесса $X_n^r(t) = \left\{X_a^{(n)}(t), a = \overline{1,r}\right\}, t \in Z, n = 1,2,...,$ выше второго порядка стремятся к нулю при $n \to \infty$, а

$$m_a^{(n)}(t) = MX_a^{(n)}(t) \xrightarrow[n\to\infty]{} m_a(t),$$

$$R_{ab}^{(n)}(t_1,t_2) = M\left[X_a^{(n)}(t_1) - m_a(t_1)\right]\left[X_b^{(n)}(t_2) - m_b(t_2)\right] \xrightarrow[n\to\infty]{} R_{ab}(t_1,t_2),$$

$a,b = \overline{1,r}$, $t_1,t_2 \in Z$, то конечномерные распределения случайного процесса сходятся к конечномерным распределениям гауссовского процесса с вектором математических ожиданий $m(t) = \{m_a(t), a = \overline{1,r}\}, t \in Z$, и ковариационной матрицей $V = \{R_{ab}(t_1,t_2), a,b = \overline{1,r}\}$, $V = \{R_{ab}(t_1,t_2), a,b = \overline{1,r}\}, t_1,t_2 \in Z$. В этом случае говорят, что случайный процесс $X_n^r(t), n = 1,2,..,t \in Z$, имеет асимптотическое нормальное распределение.

Теорема 5. При выполнении условия

$$\sum_{t_1=-\infty}^{\infty} \dots \sum_{t_{n-1}=-\infty}^{\infty} \left|c_n^X(t_1,...,t_{n-1})\right| < \infty, n = 2,3,... \tag{9}$$

оценка, задаваемая соотношением (4), имеет асимптотическое нормальное распределение с математическим ожиданием равным $R^X(\tau)$, $\tau \in Z$, и предельной дисперсией, удовлетворяющей соотношению

$$\lim_{T\to\infty}(T-\tau)D\hat{R}^X(\tau) = 2\pi\left[g_1(0) + g_2(0) + g_3(0)\right] + KG(\tau), \tag{10}$$

где

$$g_1(z) = \iint_{\Pi^2} f^X(x_1, z - x_1, x_3)e^{i\tau(x_1+x_3)}dx_1dx_3,$$

$$g_2(z) = \int_{\Pi} f^X(x_1)f^X(z - x_1)dx_1,$$

$$g_3(z) = \int_{\Pi} f^X(x_1)f^X(z - x_1)e^{i\tau(z-2x_1)}dx_1,$$

$$G(\tau) = \iiint_{\Pi^3} f^X(x_1,x_2,x_3)e^{i\tau(x_1+x_3)}dx_1dx_2dx_3 + 2\left(\int_{\Pi} f^X(x)e^{i\tau x}dx\right)^2 + \left(\int_{\Pi} f^X(x)dx\right)^2.$$

Доказательство. По свойствам смешанных семиинвариантов, приведенных в работе Бриллинджера Д. [1], имеем

$$cum\left\{\hat{R}^X(\tau_1),...,\hat{R}^X(\tau_n)\right\} = \frac{1}{C_{\tau_1}^d(T-\tau_1)} \cdots \frac{1}{C_{\tau_n}^d(T-\tau_n)} \times$$

$$\times \sum_{t_1=0}^{T-\tau_1-1} \dots \sum_{t_n=0}^{T-\tau_n-1} cum\{Y(t_1+\tau_1)Y(t_1),...,Y(t_n+\tau_n)Y(t_n)\}.$$

Введем обозначения $Z_i = \prod_{j=1}^{2} Y\left(t_i + \left[\frac{j}{2}\right]\tau_i\right), i = \overline{1,n}$, где $\left[\frac{j}{2}\right]$ означает целую часть числа $\frac{j}{2}$.

На основании теоремы 3 правую часть предыдущего равенства перепишем в виде:

$$\frac{1}{C^{d}_{\tau_1}(T-\tau_1)}\cdots\frac{1}{C^{d}_{\tau_n}(T-\tau_n)}\times$$

$$\sum_{\bigcup_{k=1}^{M}P_k=P}\ \sum_{t_1=0}^{T-\tau_1-1}\cdots\sum_{t_n=0}^{T-\tau_n-1} cum\left\{Y\left(t_i+\left[\frac{j}{2}\right]\tau_i\right);(i,j)\in P_1\right\}\times\ldots\times cum\left\{Y\left(t_i+\left[\frac{j}{2}\right]\tau_i\right);(i,j)\in P_M\right\},$$

где суммирование ведется по всем неразложимым разбиениям множества P таблицы (8), а $cum\left\{Y\left(t_i+\left[\frac{j}{2}\right]\tau_i\right);(i,j)\in P_M\right\}$ означает смешанный семиинвариант наблюдений случайного процесса $Y\left(t_i+\left[\frac{j}{2}\right]\tau_i\right)$ с индексами $(i,j)\in P_k, k=\overline{1,M}$.

Учитывая независимость процессов $X(t)$ и $d(t)$, $t\in Z$, соотношения между смешанными моментами и смешанными семиинвариантами [4] и условие теоремы (9), получим, что

$$cum\left\{\hat{R}^{X}(\tau_1),\ldots,\hat{R}^{X}(\tau_n)\right\}\xrightarrow[T\to\infty]{}0$$

равномерно по τ_1, τ_2, ..., τ_n, $n\geq 2$, $\tau_j=\overline{0,T-1}$.

Очевидно, что оценка является несмещенной, а из соотношения (7) и свойств ядра Фейера следует (10). Окончательное доказательство теоремы следует из теоремы 4.

ЛИТЕРАТУРА

1. *Бриллинджер, Д.* Временные ряды. Обработка данных и теория / Д. Бриллинжер. М.: Мир, 1980.

2. *Воротницкая, Т.И.* О предельном распределении оценки ковариационной функции стационарных случайных процессов с пуассоновскими нерегулярными наблюдениями / Т. И. Воротницкая // Материалы республиканской научно-практической конференции «Информационные технологии управления в экономике -2007», БрГУ им. А.С. Пушкина, БрГУ, 2007. стр.28.

3. *Драймз, Ф.* Распределенные лаги: Проблемы выбора и оценивания модели /Ф. Драймз. М.: Финансы и статистика, 1982.

4. *Труш, Н. Н.* Асимптотические методы статистического анализа временных рядов / Н. Н. Труш. Мн.: БГУ, 1999. 218 с.

5. *Sen, P. K.* Asymptotic Normality of Sample Quantiles for m-Dependent Processes / P. K. Sen // Annals of Mathematical Statistics, 1968. Vol. 39. No. 5. pp. 1724-1730.

6. *Varatnitskaya, T. I.* The investigation of estimates of characteristics of random process with non-regular observations / T. I. Varatnitskaya // Proceedings of the Eighth International Conference Computer Data Analysis and Modeling, Minsk, September 12-15, 2007, pp. 233-236.