УДК 519.246.8

АНАЛИЗ СЛУЧАЙНОСТИ И СТЕПЕНИ ГЕТЕРОГЕННОСТИ ВРЕМЕННОГО РЯДА

Кондрашова Н.В.
Международный научно-учебный центр информационных технологий и систем НАН та МОН Украины
nkondrashova@ukr.net

У роботі проведений статистичний аналіз нестаціонарного ряду світової ціни на нафту. На основі його емпіричної функції розподілу і методів варіаційної статистики визначена ступень гетерогенності та отримані точні довірчі інтервали для невідомої функції ймовірності зміни змінної ряду.
Ключові слова: аналіз випадковості, гетерогенний часовий ряд, емпірична функція розподілу, довірчий інтервал, непараметричні критерії

The statistical analysis of nonstationary series of world oil prices is carried out. The degree of heterogeneity is determined and the exact confidence intervals for the unknown theoretical frequency of change in the time series are obtained on the basis of its empirical distribution function and the methods of variation statistics.
Keywords: an analysis of randomness, heterogeneous time series, empirical function distribution, confidence intervals, nonparametric criteria

В работе проделан статистический анализ нестационарного ряда мировой цены на нефть. На основе его эмпирической функции распределения и методов вариационной статистики определена степень гетерогенности и получены точные доверительные интервалы для неизвестной функции вероятности изменения переменной ряда.
Ключевые слова: анализ случайности, гетерогенный временной ряд, эмпирическая функция распределения, доверительный интервал, непараметрические критерии

Прогнозирование временного ряда мировой цены на нефть с использованием некоторых популярных методов в работе [1] не было достаточно точным. Анализ случайности для определения степени предсказуемости процесса случайной динамической природы и анализ его эмпирической функции распределения является предметом рассмотрения данной работы.

Исходный ряд является гетерогенным, т.к. уровень, скорость роста, разброс колебаний и проч. все время изменяются, в качестве ухода от нестационарности используется интеграция – последовательное получение конечных разностей [2]. На рис. 1 представлены исходный ряд мировых цен на нефть и ряд первых разностей. Ряд первых разностей имеет меньший разброс значений, чем исходный ряд.

Постановка проблемы

Имеется выборка значений первых разностей $X=(x_1, x_2, ..., x_k, \dots, x_N)$ ряда последовательных ежемесячных данных (наблюдений) мировой цены на нефть марки Brent, где N – объем выборки.

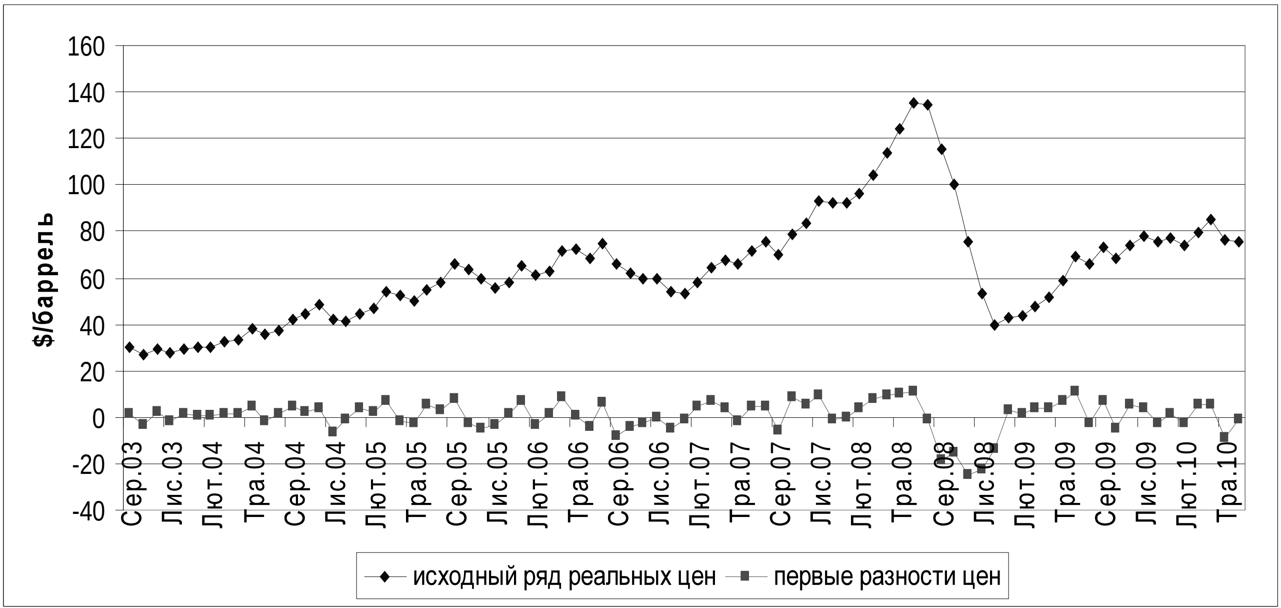

Рис.1 Ряд мировых цен на нефть марки Brent и соответствующий ему ряд первых разностей

Прежде чем приступать к изучению связи между последовательными значениями цен, была предпринята попытка выяснить, не является ли исходный ряд абсолютно случайным, «белым шумом», в котором отсутствует зависимость между значениями ряда, относящимися к различным моментам, и направление движения которого в будущем равновероятно, т.е. непредсказуемо. Для этого ряд испытан с помощью критериев случайности. В качестве таких критериев использовались: *критерий поворотных точек*, *критерий серий*, *критерий распределения длины фазы*, *критерий, основанный на знаках разностей*, *и критерии, основанные на ранговой корреляции* [3].

Анализ случайности

Критерий поворотных точек

Для ряда подсчитывается сумма локальных экстремумов (поворотных точек) L. Если выполняется условие [4]:

$$L > L_{kp}, \quad L_{kp} = \left[\frac{2(N-2)}{3} - 1{,}96\sqrt{\frac{16N-29}{90}} \right],$$

то ряд x_i является случайным с вероятностью α=0.95. Здесь N –число наблюдений, а квадратные скобки означают, что от результата берется целая часть. Для ряда первых разностей, представленного на рис.1, наблюдаемое число локальных экстремумов L=51 больше критического для уровня значимости β=0.05 значения $L_{kp} = 45$ (51>45), поэтому ряд по данному критерию является случайным.

Критерий серий

Находится медиана m вариационного ряда (срединное значение при нечетном N или среднее арифметическое двух срединных значений при четном N). Сравнивается каждое значение исходной последовательности x_i с m и ставится «+1», если значение $x_i > m$ и «-1», если значение $x_i < m$. В случае равенства $x_i = m$ ставится ноль. В результате получается последовательность из +1 и -1, общее число которых меньше N. Последовательность, подряд идущих +1 или -1, называется серией. Протяженность самой длинной серии $K_{\max}$, а ν – число серий. Выборка признается случайной, если выполняются следующие условия (при 5% уровне значимости) [5]:

$$K_{\max} < \left[3{,}3(\log N + 1)\right]$$

$$\nu > \left[\tfrac{1}{2}(N+1) - 1{,}96\sqrt{N-1}\right] \tag{1}$$

[·] – функция определения целой части числа. Если хотя бы одно из неравенств (1) нарушается, то гипотеза о случайном характере ряда отвергается. Для ряда первых разностей m=1,625. И хотя первое неравенство выполняется $K_{\max} = 6 < 10$, но не выполняется второе неравенство, т.к. $\nu = 21 < 23$, поэтому гипотеза о случайном характере ряда отклоняется.

Коэффициенты ранговой корреляции

Проверяются два непараметрических критерия. Коэффициент Спирмана – аналог коэффициента парной корреляции между следующими рядами. Исследуемому ряду ставится в соответствие ранжированный по возрастанию (убыванию) ряд, а второй ряд – исследуемый хронологический порядок наблюдения. Если ряд всегда возрастает, то ранжированный ряд совпадает (совпадает в обратном порядке) с порядком наблюдения и – коэффициент корреляции укажет на их тесную связь. Если порядки независимы, то коэффициент укажет на чисто случайный процесс (гипотеза H_0). Частный случай коэффициента Спирмана при отсутствии повторения (связанных) рангов имеет вид [6]:

$$r_s = 1 - \frac{6}{N(N^2-1)} \sum_{i=1}^{N} (\mu_i - i)^2 ,$$

где μ_i – ранг присвоенный i-му наблюдению. При наличии связанных рангов формула имеет вид:

$$\tilde{r}_s = \frac{r_s + T}{\sqrt{1-T}}, \; T = \frac{3}{N(N^2-1)} \sum_j k_j (k_j - 1)$$

где k_j - количество повторений в j-й подпоследовательности связанных рангов.

Для чисто случайных процессов r_s имеет нулевое математическое ожидание и дисперсию, равную $\frac{1}{\sqrt{N-1}}$. В больших выборках r_s имеет нормальное распределение с указанными параметрами. Для малых выборок предпочтительнее пользоваться статистикой $r_s\sqrt{\frac{N-2}{1-r_s{}^2}}$, которая приближенно имеет распределение Стьюдента с (N-2) степенями свободы. Если искомая расчетная величина по модулю меньше двусторонней критической границы распределения Стьюдента, то нулевая гипотеза о случайности процесса принимается и утверждается, что тенденция отсутствует. Коэффициент Спирмана равен r_s = 0,067393. Для N=80 и уровня значимости β=0.05 критическое значение r_s^{β}=0.22 [7]. Гипотеза H_0 принимается, корреляция между рядами не достигает уровня статистической значимости.

Ещё одним вариантом парных ранговых коэффициентов корреляции является коэффициент Кендалла. В этом методе одна переменная представляется в виде монотонной последовательности в порядке возрастания величин; другой переменной присваивается соответствующее ранговое место, аналогично, как при вычислении коэффициента Спирмана. Количество инверсий (нарушений монотонности по сравнению с первым рядом) используется в формуле для корреляционных коэффициентов. Формула для его расчета [6]:

$$\tau = \frac{2}{N(N-1)} \sum_{i,j=1}^{N} sign\left[(\mu_{i1} - \mu_{j1})(\mu_{i1} - \mu_{j1})\right],\ \tau \in [-1;1],$$

где μ_{ij} – ранг присвоенный i-му наблюдению j-й переменной (j=2). При наличии связанных рангов в расчетах используется скорректированный коэффициент корреляции Кендалла:

$$\tau = \frac{1}{\sqrt{\left[\frac{N(N-1)}{2} - T_1\right]\left[\frac{N(N-1)}{2} - T_2\right]}} \sum_{i,j=1}^{N} sign\left[(\mu_{i1} - \mu_{j1})(\mu_{i1} - \mu_{j1})\right]$$

$$T_j = \frac{1}{2}\sum_i k_{ji}(k_{ji} - 1), \quad j = 1,2,$$

где k_{ji} - количество повторений в i-й подпоследовательности связанных (повторяющихся) рангов j-й последовательности, [·] функция определения целой части числа. Применение коэффициента Кендалла является предпочтительным, если в исходных данных встречаются выбросы. Коэффициенту τ=0.052 соответствует уровень значимости β=0.498>>0.05, что не позволяет отклонить гипотезу о случайности ряда.

Критерий распределения длины фазы

Чтобы обнаружить фазу в ряде первых разностей длины d, необходимо обнаружить цепочку длины d+3, имеющею вид

$$x_i < x_{i+1} > x_{i+2} > \ldots > x_{i+d+1} < x_{i+d+2}$$

(фаза убывания) или такую же, в которой все неравенства заменены на противоположные (фаза возрастания). Для (чисто) случайного временного ряда из N членов среднее количество фаз длины d_i равно [4]:

$$y_i = \mathrm{M}\left[n(d_i)\right] = \frac{2(N-d_i-2)(d_i^2+3d_i+1)}{(d_i+3)!}, \qquad (2)$$

математическое ожидание числа фаз во всем случайном ряду приближенно равно $\bar{y} = \mathrm{M}\left[\sum_{i=1}^{N-3} n(d_i)\right]$=(2$N$-7)/3. Наблюдаемое число фаз заданной длины y_i, ожидаемое их число y_i, рассчитанное по формуле (2), а также суммарное значение равное 49 и среднее суммарное $\bar{y}$=51 представлены во второй и третьей строках таблицы 1.

Таблица 1

Сравнение количества фаз

Распре-деление фаз	длина фазы d_i	1	2	3	4	6	
	Число фаз наблюдаемых y_i	33	11	3	1	1	49
	Число фаз вычисленных y_i	12,83	6,33	2,08	0,499	0,015	51
Разность (y_i- y)		20,17	4,67	0,92	0,501	0,99	-2
Ранг μ_{i1}		6	5	2	1	3	-4

Поскольку распределение числа длин фаз не является нормальным и даже не сходится к нему при увеличении числа N, то критерий Стьюдента применять нельзя. Насколько сходятся ряды (наблюдаемый y и вычисленный y), представленные в таблице 1 можно убедиться визуально, и статистически с помощью критериев Кендалла и Спирмана.

Если рассчитать корреляционный коэффициент Кендалла (τ=0.949), то видно, что в данном случае он значительно ниже, чем корреляционный коэффициент Спирмана (r_s=0.975), и соответствует более высокому уровню значимости (β=0.05), чем у коэффициента Спирмана (β=0.01) (см. таблицу 2).

Гипотеза о случайном характере ряда принимается по обоим критериям, поскольку ряд y соответствует абсолютно случайному ряду.

Таблица 2

Correlations

			VAR00001	VAR00002
Kendall's tau_b	VAR00001	Correlation Coefficient	1,000	,949*
		Sig. (1-tailed)	.	,011
		N	5	5
	VAR00002	Correlation Coefficient	,949*	1,000
		Sig. (1-tailed)	,011	.
		N	5	5
Spearman's rho	VAR00001	Correlation Coefficient	1,000	,975**
		Sig. (1-tailed)	.	,002
		N	5	5
	VAR00002	Correlation Coefficient	,975**	1,000
		Sig. (1-tailed)	,002	.
		N	5	5

*. Correlation is significant at the 0.05 level (1-tailed).

**. Correlation is significant at the 0.01 level (1-tailed).

Критерий знаков и Т-критерий Вилкоксона

Для сравнения двух зависимых выборок, элементы которых представлены в неметрических шкалах, или вследствие других причин, определяющих невозможность использования соответствующих параметрических критериев, применяются критерий знаков и Т-критерий Вилкоксона.

Критерий знаков применяется для проверки гипотезы об однородности генеральных совокупностей по попарно связанным выборкам. Критерий знаков основан на подсчете количества раз, когда элементы одной выборки превышают парные элементы другой выборки ($y_i > [y_i]$), и наоборот ($y_i < [y_i]$, $[\cdot]$ – функция округления числа с избытком). Нулевая гипотеза о независимости выборок предполагает равенство этих случаев. В случае данных Таблицы 1 все значения одной выборки превосходят значения другой выборки, а значит – гипотеза о чистой случайности ряда первых разностей отвергается, т. к. у производного наблюдаемого ряда y имеется значительное расхождение на уровне β =0,001 с рядом y, распределение фаз у которого случайно.

Более чувствительным, чем критерий знаков, при сравнении двух связанных выборок является Т-критерий Вилкоксона. Его целесообразно использовать в случаях, когда неприменим t-критерий Стьюдента. Условием его применения является существование непрерывной функции распределения.

Он основан на вычислении разности μ_i значений в каждой i-й паре элементов связанных выборок. Затем эти разности ранжируются от меньших к большим по модулю безотносительно знака. Если несколько разностей имеют одинаковое значение, им приписываются средние арифметические значения рангов. Нулевые значения не учитываются. Затем подсчитываются суммы рангов положительных (T^+) и отрицательных – разностей (T^-). Меньшая из сумм принимается в качестве эмпирического значения и сравнивается с критическим значением. Нулевая гипотеза отклоняется, если эмпирическое

Проверим распределение фаз данным критерием. Исследуем являются ли оба ряда независимыми (нулевая гипотеза H_0 – да, являются). Полученные данные и результаты их обработки представлены в 4-ой и 5-ой строчках Таблицы 1. Сумма положительных рангов T^{+}=17. Сумма отрицательных разностей T^{-} =4. Следовательно, за эмпирическое значение критерия Вилкоксона принимается T^{-}=4. Критическое значение находится из таблицы в работе [8] при уровне значимости β =0,05 и n равном количеству ненулевых разностей (для данного примера n=6). Критическое значение критерия $T_{кр}$<2. В этом тесте эмпирическое значение должно быть меньше критического для отклонения нулевой гипотезы. Следовательно, принимается нулевая гипотеза, и делается содержательный вывод о статистической независимости эмпирического ряда и случайного. Значит, эмпирический ряд не является чисто случайным. Результаты исследований по различным критериям случайности приведены в таблице 3.

Таблица 3

Результат проверки временного ряда по критериям случайности

поворотных точек	серий	коэффициент Спирмана	коэффициент Кендалла	распределение длины фаз	знаков	Вилкоксона
+	–	+	+	+	–	–

В табл. 3 знак «+» означает, что гипотеза о случайности ряда H_0 принимается; знак «–» означает, что гипотеза H_0 отклоняется.

Интересно, что анализ одних и тех же попарно связанных выборок y и y из таблицы 1 с помощью двух групп непараметрических критериев привел к прямо противоположным выводам. К первой группе относятся критерии Кендалла и Спирмана, а ко второй – критерий знаков и критерий Вилкоксона.

В целом проведенный анализ ряда первых разностей среднемесячных цен нефти Brent показал, что, несмотря на сильную колеблемость данных, они не являются абсолютно случайными.

В работе [9] утверждается, что условия практического использования теории вероятности трактуются в зависимости от приложения к экспериментам и эксперименты, в которых нет никакой стойкости исходов испытаний, являются предметом изучения теории вероятности. К сказанному можно добавить, что предметом изучения теории вероятности являются также эксперименты с некоторой степенью стойкости исходов.

Весь диапазон значений, которые могут принимать ряды (на рис.1) разбиваем на одинаковое число интервалов. Анализ графиков распределения частот исходного ряда и ряда первых разностей, выявляет полимодальность этих распределений (см. рис.2 и рис.3).

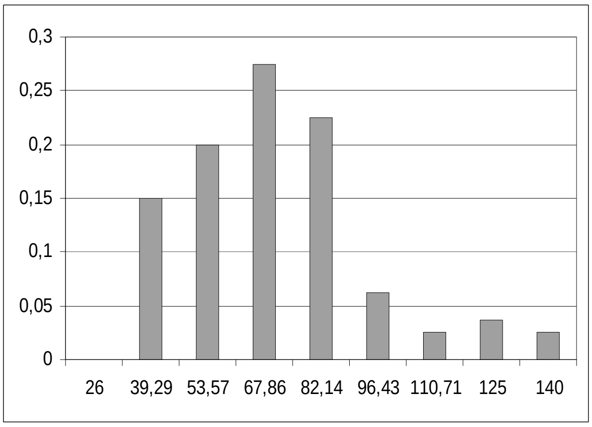

Рис. 2 Гистограмма распределения частот исходного ряда

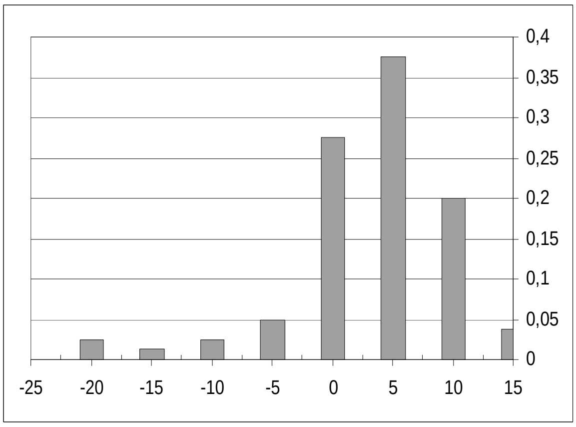

Рис.3 Гистограмма распределения частот первых разностей

Решим задачу о выявлении двух частей суммарной выборки, которые принадлежат разным генеральным совокупностям (задачу о точке изменения распределения). Такая задача возникает тогда, когда необходимо определить точку изменения тренда регрессии (точку перехода регрессии).

Далее будем обрабатывать временные ряды методом идентификации смешанных выборок, основываясь при этом на анализе графиков центральной части кривых Кетле [10] (рис.4) и спейсинг анализе [11] (рис.5).

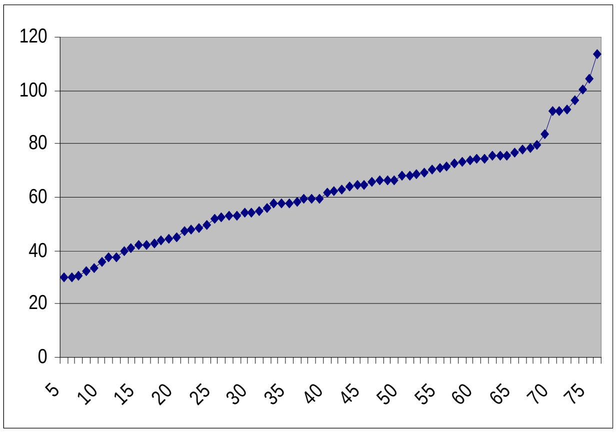

Рис. 4 Кривая Кетле среднемесячных цен на нефть Brent

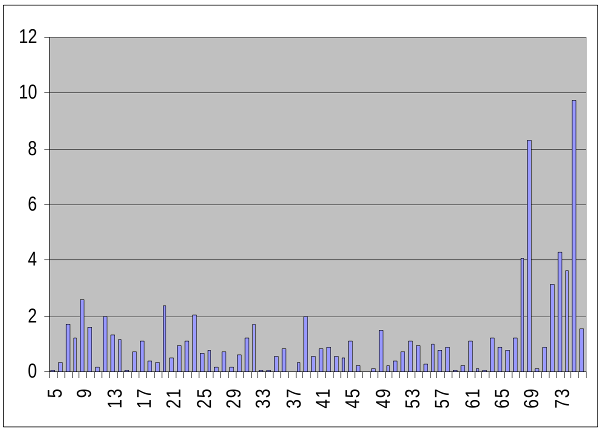

Рис.5 Спейсинги исходного ряда

Из графика кривой Кетле на рис.4 видно, что выборка среднемесячных цен принадлежит гетерогенной генеральной совокупности (ГС), является смесью трех ГС $X = \mathbf{x}_1 \cup \mathbf{x}_2 \cup \mathbf{x}_3$, т.к. вариационная кривая имеет не одну точку перегиба. Первая точка перегиба соответствует значению 79.7 $/баррель, максимальный спейсинг совокупности $\mathbf{x}_2$ приблизительно равен 8.5 и превышает максимальный спейсинг совокупности $\mathbf{x}_1$ 2.3 в 3.5 раза. Вторая точка перегиба соответствует значению 93.2 $/баррель и максимальный спейсинг совокупности $\mathbf{x}_3$ равен 9.76, что превышает спейсинг совокупности $\mathbf{x}_1$ в 4.25 раза. Откуда следует, что выборочное среднее на первом отрезке при больших N будет статистически отличаться от выборочного среднего на другом и, особенно, третьем отрезке кривой Кетле со статистической значимостью расхождения по критерию $3s_1$. Где s_1 – стандартное отклонение порядковой статистики $x_{(1)} \le x_{(2)} \le \ldots \le x_{(i_1)}$, принадлежащей первому модальному классу $\mathbf{x}_1$. Оценки математического ожидания и дисперсии имеют вид:

$$m_1 \approx \bar{x}^{(1)} = \frac{1}{i_1}\sum_{k=1}^{i_1} x_{(k)}\,,\ \sigma_1^2 \approx s_1^2 = \frac{1}{i_1 - 1}\sum_{k=1}^{i_1}(x_{(k)} - \bar{x}^{(1)})^2$$

Однако для выборки объема N=80 $\bar{x}^{(1)}$=57.14, s_1=13.9, $\bar{x}^{(2)}$=81.75, s_2=2.9, $\bar{x}^{(3)}$=98.9, s_3=8.3 значимого расхождения (β=0.05) между этими классами нет, поскольку все доверительные интервалы на уровне $3s_1$ пересекаются (57.14±41.74), (81.75±8.67) (98.9±24.1). Степень гетерогенности $\lambda(\mathbf{x}_1, \mathbf{x}_2 \cup \mathbf{x}_3)$=0.13 есть отношение объемов двух подвыборок (меньшей $n_{2\cup 3} = 9$ к большей $n_1 = 63$). Выборка среднемесячных цен на нефть является слабо гетерогенной, т.к. $\lambda(\mathbf{x}_1, \mathbf{x}_2 \cup \mathbf{x}_3)$<0.3. Функция вероятности первого класса является непрерывной, имеет симметричный закон и в соответствии с гипотезой Хилла [12] и правилом 3σ [13] максимальное значение приращения цены нефти марки Brent не должно превысить 98.88 $/баррель в месяц, а минимальное не будет меньше, чем 15.4 $/баррель в месяц с уровнем значимости $2\beta \le 0.05$. Относительно смеси $\mathbf{x}_2 \cup \mathbf{x}_3$ нельзя утверждать ничего определенного, т.к. число точек $n_{2\cup 3} = 9 < 30$ меньше минимального их количества необходимого для обеспечения гарантированного уровня значимости $\beta \le 0.05$ доверительного интервала 3σ. Рассмотрим график вариационного ряда первых разностей и соответствующий спейсинг (рис.6.)

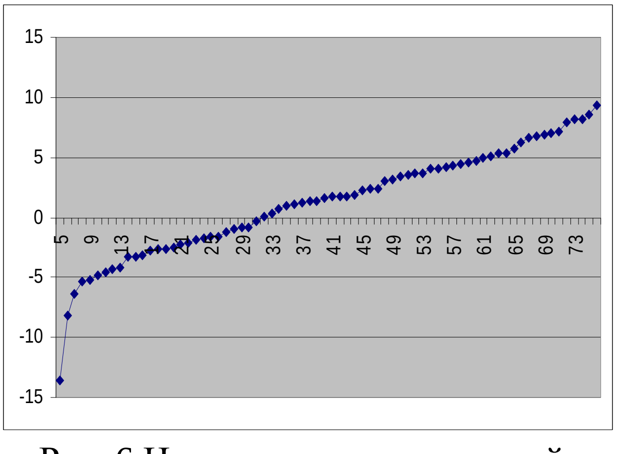

Рис. 6 Центральная часть кривой Кетле первой разности

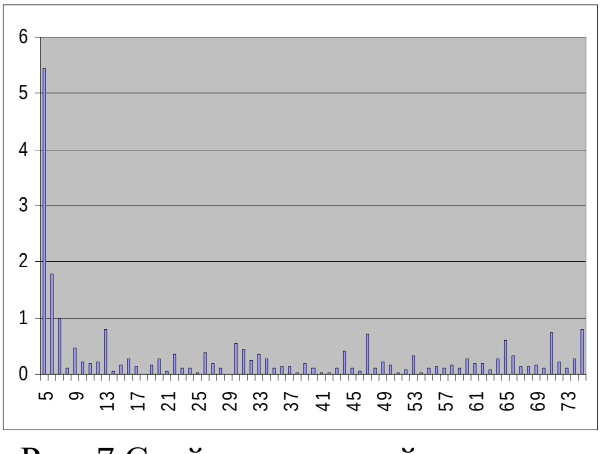

Рис. 7 Спейсинги первой разности

Из графика рис.6 видно, что выборка первых разностей является смесью двух подвыборок $X = \mathbf{x}_1 \cup \mathbf{x}_2$. Точка перелома соответствует значению -6.34, максимальный спейсинг совокупности $\mathbf{x}_1$ приблизительно равен 5.45 и превышает максимальный спейсинг совокупности $\mathbf{x}_2$ равный 0.79 приблизительно в 7 раз. Оценки математических ожиданий и стандартных отклонений порядковых статистик, принадлежащих первому и второму модальным классам соответственно равны: $\bar{x}^{(1)}$=-9.36, s_1=3.78, $\bar{x}^{(2)}$=1.7, s_2=3.9. Значимого (β=0.05) расхождения между этими классами также нет, поскольку доверительные интервалы на уровне $3s_1$ пересекаются (-9.36±11.32), (1.7±11.69). Степень гетерогенности выборки $\lambda(\mathbf{x}_1, \mathbf{x}_2)$=0.043 ($n_1 = 3$, $n_2 = 69$, см. рис. 7). Выборка первых разностей является слабо гетерогенной, т.к. $\lambda(\mathbf{x}_1, \mathbf{x}_2)<0.3$. Выборка первых разностей является более однородной, чем исходная выборка, т.к. $\lambda(\mathbf{x}_1, \mathbf{x}_2) < \lambda(\mathbf{x}_1, \mathbf{x}_2 \cup \mathbf{x}_3)$, т.е. почти гомогенной. Если про распределение первого малочисленного класса отрицательных значений ничего определенного сказать нельзя, т.к. слишком мал объем его выборки ($n_1 = 3$), то функция вероятности второго класса является непрерывной, имеет симметричный закон. Максимальное значение приращения цены нефти марки Brent не должно превысить 13.4 $/баррель в месяц, а минимальное не будет меньше, чем -9.99 $/баррель в месяц с уровнем значимости $2\beta \le 0.05$.

Выводы. Исследованные гетерогенные временные ряды не являются стопроцентно абсолютно случайными. Об этом свидетельствуют некоторые критерии случайности. Особо обнадеживает то, что именно непараметрические критерии, для применения которых не является необходимым знание функции вероятности, не подтверждают абсолютную случайность ряда. Во временном ряде первых разностей среднемесячных значений проявляется некоторая инерционность движения, и это позволяет надеяться на применимость статистических методов для прогнозирования курсов. В то же время многие исследованные характеристики ряда не очень сильно отличаются от тех, которые теоретически выведены для совершенно случайного ряда. Вследствие этого нельзя ожидать от статистических прогнозов большой степени осуществимости.

Для определения точки изменения тренда регрессии с помощью методов вариационного анализа для каждого ряда определены точки перелома для выявления двух и более частей суммарной выборки. Определены степени гетерогенности суммарных выборок. Поскольку выборка первых разностей является почти гомогенной, то переход на иной тренд маловероятен. Эта выборка имеет непрерывную симметричную функцию вероятности, а значит для нее определены предельные границы изменения положительных и отрицательных приращений с гарантированным уровнем значимости. Анализ степени гетерогенности выборки первых разностей помог установить граничные оценки изменения исходного ряда с высокой вероятностью (95%) не выхода за их пределы. Но поскольку гистограмма резко обрывается, подойдя к граничной точке в области положительных значений, следует ожидать в будущем, больших, чем $x^{+}=15$ $ за баррель в месяц положительных приростов мировой цены нефти, а в периоды будущих экономических кризисов – серий из больших, чем x^{+} приращений.

Литература

1. Кондрашова Н. В., Павлов Я.В. Сравнительный анализ методов сглаживания и МГУА для прогнозирования временных рядов// Індуктивне моделювання складних систем. Збірник наукових праць – Київ: МННЦ ІТС, 2009.– С. 84-101.

2. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. (Вып. 1, 2.). – М.: Мир, 1972, – 405с. – 197 с.

3. Кендалл М. Временные ряды. - М.: Финансы и статистика, 1981.

4. Дронов B.C. Многомерный статистический анализ: Учебное пособие. Барнаул. Из-во Алт. гос. ун-та. 2003.– 213 с.

5. http://xxl.alterxp.com/unprot/RADOST.DOC

6.Давнис В. В., Тинякова В. И. Прогнозные модели экспертных предпочтений. Воронеж, Изд-во Воронеж, гос. ун-та, 2005. – 248 с.

7. http://www.psychol-ok.ru/statistics/spearman/

8. Cooper, D.R. and Shindler, P.S. Business Research Methods. Irwin/McGraw-Hill, 1995. – P. 686.

9 . Тутубалин В.Н. Теория вероятностей и случайных процессов.– М.: Изд-во МГУ, 1992.

10. Ван дер Варден Б. Л. Математическая статистика. - М.: И-Л. 1960. – 435 с.

11. Руke R., Spacings. J. Roy. Statist. Soc, 1965.– ser. B., 27.– P. 395—436.

12. Hill B.M. Posteriori distribution of percentiles: Bayes' theorem for sampling from a population// Journal of the American Statistical Association.-1968.- Vol. 63, No 322. – P.677-691.

13.Матвейчук С.А., Петунин Ю.И. Обобщение схемы Бернулли, возникающее в вариационной статистике I,II // Укр. Мат. журнал. – 1991.– т.42.– № 4.– С.518-528.– т.48.– № 6.– С.779-785.