

ФІЗИКО-МАТЕМАТИЧНІ НАУКИ

УДК 519.2

СТАТИСТИЧЕСКОЕ РАСПОЗНАВАНИЕ ЧИСЛОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ МЕТОДОМ АНАЛИЗА ЗНАКОВЫХ ПРИЗНАКОВ ПЕРЕМЕННЫХ РАЗНОСТЕЙ В ПОЛЕ АНАЛИЗА СИСТЕМАМИ С ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ

Крутик А.В.

Скворская государственная районная администрация

В этой работе автор предлагает составлять статистические портреты событий, линейная корреляционная обработка числовых характеристик которых не дает существенных результатов, разработанным им методом корреляционно-регрессионного анализа. Как и в методе Фехнера, в предлагаемом методе анализируются знаковые признаки отклонений от средних значений (математических ожиданий) величин но не только исследуемых последовательностей но и их переменных разностей в предложенном автором поле анализа. Новый метод значительно расширяет количество анализируемых факторов что улучшает качественные характеристики конечного продукта исследования – статистического портрета события. Метод дает широкие возможности для выявления статистических зависимостей между событиями при проведении научных исследований в любых областях нашей жизни и деятельности. Что автор и показывает на примере исследования цен на золото в долларах США в годы, что предшествовали финансовым кризисам и в итоге составлением статистического портрета финансового кризиса. Конечной целью рассмотренного метода, безусловно, есть желание улучшения качества нашей жизни и деятельности в различных их областях.

Ключевые слова: статистическое распознавание, корреляционно-регрессионный анализ, знаковые признаки, переменные разности, искусственный интеллект, цена золота, финансовые кризисы.

Постановка проблемы. Основными математическими методами для распознавания были и остаются формулы вычисления коэффициентов корреляции между двумя выборками чисел. Одна из этих последовательностей может являться исследуемой а другая эталонная. После проводится сравнение коэффициентов с пороговым значением, что дает возможность сделать простой вывод – принадлежит ли исследуемая последовательность чисел или не принадлежит статистически к эталонной последовательности (исследуемому явлению) с заданными параметрами качества распознавания, такими как вероятность правильного распознавания (P_n) и вероятность ложного распознавания (P_l). Но когда они не соответствуют одни одному, то есть когда коэффициенты корреляции имеют низкие значения и не взирая на взаимную подстройку числовых трендов по частоте их следования и по задержкам по оси их следования, все равно коэффициенты корреляции не увеличиваются – как тогда решить вопрос об их, последовательностей, «родстве»? Особенно когда мы точно знаем что эти последовательности характеризуют один и тот же объект исследования и имеют одну и ту же природу и условия возникновения. Ответ очевиден – необходимы другие подходы и математические методы корреляционно-регрессионного анализа с включением новых факторов в уравнение регрессии для решения этой задачи – задачи распознавания.

Анализ последних исследований и публикаций. Тема нахождения зависимостей между выборками чисел в итоге статистического распознавания событий корреляционно-регрессионными методами, хорошо изучена и представлена в литературе и электронном ресурсе [1]. Для статистического анализа двух выборок чисел широко используются на практике основные методы вычисления:

- Линейного коэффициента корреляции (или коэффициент корреляции Пирсона);
- Непараметрического коэффициента ранговой корреляции Кендалла;
- Коэффициента ранговой корреляции Спирмена;
- Коэффициента корреляции знаков Фехнера;
- Коэффициента множественной ранговой корреляции (конкордации).

Предлагаемый метод в отличие от существующих позволяет найти статистическую зависимость одновременно не только между двумя выборками чисел, а между неограниченным количеством трендов описывающих исследуемое явление, что дает возможность составить уравнение регрессии с включением множества факторов статистического распознавания характерных для всех известных выборок чисел характеризующих событие. Это уравнение в работе автор называет «статистическим портретом» исследуемого события.

Выделения нерешенных ранее частей общей проблемы. Одной из основных конечных целей

разработки систем с искусственным интеллектом (и.и.) в любых областях где-бы их не применяли, есть принятие решения на основании оценок параметров характеризующих объект управления. Сегодня эти исходные данные системы с и.и. в основном получают в уже в обработанном и пригодном для применения виде. В недалеком будущем эти системы будут самостоятельно получать информацию ее анализировать и конечно распознавать то есть самостоятельно принимать решение принадлежит или не принадлежит цифровая последовательность к конкретному событию, объекту управления. Поэтому составление математического портрета события есть частью общей проблемы для достижения конечной цели – принятия качественного решения.

Формирования целей статьи. Можно ли решить задачу распознавания события другими еще не изученными корреляционно-регрессионными математическими методами, иными словами можно ли создать статистический портрет события с высокими качественными статистическими характеристиками новыми методами цифровой обработки? На мой взгляд такие методы существуют и чем их больше будем исследовать и применять тем более качественней будет конечный продукт исследования – статистический портрет события.

Постановка задания. И так, задача представленной вашему вниманию работы – создать новый метод корреляционно-регрессионной обработки цифровой информации для числовых последовательностей. Ее целью есть статистическое распознавание различных событий и составление их статистических портретов с заданными качественными характеристиками.

Изложение основного материала.

1. Статистическое распознавание событий методом анализа распределений знаковых признаков переменных разностей в поле анализа.

1.1 Приведение числовых последовательностей к удобному виду для последующей их обработки.

Зачастую, исследуемые выборки чисел имеют разное количество членов, разные среднеквадратические отклонения и математические ожидания и значения их могут сниматься в разные моменты времени с разной частотой и в различных внешних условиях, что в итоге может привести при их обработке к значительным погрешностям, неточностям и несогласованиям. Поэтому для исключения в последующей обработке этих ошибок предлагается исследуемые последовательности:

1. Масштабировать – определить начало последовательности, установить одинаковый размер (количество чисел), выбрать одинаковый частотный диапазон из общего спектра, то есть частоту дискретизации, упорядочить числа по другим признакам.

2. Нормировать – привести все исследуемые ряды к единому диапазону по величине (СКО=1) и к единому среднему значению (МАТОЖ=0)

1.2 Метод переменных разностей. Матрица-оператор переменных разностей.

Статистический анализ временных рядов методом переменных разностей хорошо изучен

и описан в литературе [2, с. 76-94]. Благодаря некоторым свойствам этого метода он широко применяется при решении задач математического прогнозирования. А на заре зарождения вычислительной техники метод был положен в основу при создании первой программированной механической вычислительной аналитической машины Чарльза Бэббиджа.

В матричном виде переменную разность (R) (для 8 степени разности и ряда с 17 чисел) можно представить как

$$R = Or \cdot X, \quad \text{где} \quad (1)$$

Or – матрица-оператор переменной разности

$$Or = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -3 & 3 & -1 \\ 0 & 0 & 0 & 0 & 1 & -4 & 6 & -4 & 1 \\ 0 & 0 & 0 & 1 & -5 & 10 & -10 & 5 & -1 \\ 0 & 0 & 1 & -6 & 15 & -20 & 15 & -6 & 1 \\ 0 & 1 & -7 & 21 & -35 & 35 & -21 & 7 & -1 \\ 1 & -8 & 28 & -56 & 70 & -56 & 28 & -8 & 1 \end{pmatrix} \quad (2)$$

Интересное свойство оператора – любой член каждого столбца есть разность между двумя числами находящимися на один ряд выше, числом столбца справа и числом этого столбца, к примеру, $-56 = -21 - 35$ (см.обведенные числа матрицы)

X – матрица чисел исследуемого ряда:

$$X = \begin{pmatrix} X_9 & X_{10} & X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{16} & X_{17} \\ X_8 & X_9 & X_{10} & X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{16} \\ X_7 & X_8 & X_9 & X_{10} & X_{11} & X_{12} & X_{13} & X_{14} & X_{15} \\ X_6 & X_7 & X_8 & X_9 & X_{10} & X_{11} & X_{12} & X_{13} & X_{14} \\ X_5 & X_6 & X_7 & X_8 & X_9 & X_{10} & X_{11} & X_{12} & X_{13} \\ X_4 & X_5 & X_6 & X_7 & X_8 & X_9 & X_{10} & X_{11} & X_{12} \\ X_3 & X_4 & X_5 & X_6 & X_7 & X_8 & X_9 & X_{10} & X_{11} \\ X_2 & X_3 & X_4 & X_5 & X_6 & X_7 & X_8 & X_9 & X_{10} \\ X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 & X_8 & X_9 \end{pmatrix} \quad (3)$$

Свойство переменных разностей, на которое я ссылаюсь в методе распознавания.

Каждая очередная по степени переменная разность увеличивает среднеквадратическое отклонение чисел в количество раз, K которое равно корню от суммы квадратов множителей соответствующей строки оператора разности (для последовательностей имеющих случайный характер).

Например:

$$j = 1 \quad K = \sqrt{1 \cdot 1 + 1 \cdot 1} = 1.414;$$

$$j = 2 \quad K = \sqrt{1 \cdot 1 + (-2) \cdot (-2) + 1 \cdot 1} = 2.449;$$

$$j = 3 \quad K = \sqrt{1 \cdot 1 + (-3) \cdot (-3) + 3 \cdot 3 + 1 \cdot 1} = 4.472;$$

$$j = 4 \quad K = \sqrt{1 \cdot 1 + (-4) \cdot (-4) + 6 \cdot 6 + (-4) \cdot (-4) + 1 \cdot 1} = 8.366$$

В обычном, принятом виде уравнение дисперсии разности для нормированных величин имеет следующий вид [2, с. 80].

$$\sigma^2 = \frac{(2j)!}{(j!)^2}, \quad (4)$$

где j – степень разности

1.3 Знаковые признаки (Z) переменных разностей. Поле анализа знаковых признаков переменных разностей.

Каждое число в матрице R можно расписать как:

$$R_{i,j} = Z_{i,j} \cdot |R_{i,j}| \quad (5)$$

Где $Z_{i,j}$ – знаковый признак: i – порядковый номер; j – степень разности, он может принимать значения: «+1», «-1» (или «0»)

$$Z_{i,j} = \text{ОКРУГЛ}\left(\frac{R_{i,j}}{|R_{i,j}| + \Delta}; 0\right) \quad (6)$$

(Δ – очень маленькое постоянное число необходимо, чтобы не допустить деление на 0 при составлении компьютерных программ)

Далее для удобства анализа распределения указанных выше признаков Z предлагается расписать все признаки Z матрицы R всех исследуемых выборок чисел построчно один под другим в соответствии с выбранной очередностью их анализа. Полученную таким образом таблицу назовем «**полем анализа знаков переменных разностей исследуемых рядов**». Далее в поле анализа найдем повторяющиеся определенным образом знаковые признаки « Z ». При этом важно чтобы повторение носило системный характер для всех исследуемых последовательностей. Под системным характером понимается метод анализа повторений знаковых признаков или их групп в поле анализа – по диагонали, по столбцам, по строкам, буквой «Г» или другим обнаруженным способом. Далее обнаруженные повторения знаковых признаков объединим в общее уравнение регрессии, или другими словами составим статистический портрет исследуемого события.

1.4 Расчет весовых коэффициентов «Kv». Формула приблизительного восстановления исходной последовательности чисел исходя из знаковых признаков ее переменных разностей.

В формуле распознавания (уравнения №№: 12-18), каждый знаковый признак должен иметь свой весовой коэффициент Kv так как не известна значимость каждого значения Z входящих в формулу с точки зрения их влияния на степень «искривления» восстановленной последовательности в случае изменения этого признака на противоположный. Для установления этого влияния предлагается вычислить формулу приблизительного восстановления исходной последовательности исходя из Z признаков.

1.4.1 Формула приблизительного восстановления исходной последовательности чисел исходя из знаковых признаков ее переменных разностей.

Вывод формулы, для лучшего восприятия, рассмотрим на примере расчета переменных разностей, матрицы R – 4 степени разности, для X_5 (пятого элемента исследуемой последовательности чисел)

1. Запишем переменную разность 4 степени в матричном виде

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 1 & -3 & 3 & -1 \\ 1 & -4 & 6 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_5 & X_6 & X_7 & X_8 & X_9 \\ X_4 & X_5 & X_6 & X_7 & X_8 \\ X_3 & X_4 & X_5 & X_6 & X_7 \\ X_2 & X_3 & X_4 & X_5 & X_6 \\ X_1 & X_2 & X_3 & X_4 & X_5 \end{pmatrix} \quad (7)$$

2. Выпишем все разностные уравнения в которых есть X_5 (см. табл. 1)

Таблица 1

№ п/п	Степень разности	уравнения
1	0	$R_{5,0} = X_5$
2	1	$R_{4,1} = X_5 - X_4$
3		$R_{3,1} = X_6 - X_5$
4	2	$R_{5,2} = X_3 - 2X_4 + X_5$
5		$R_{4,2} = X_4 - 2X_5 + X_6$
6		$R_{3,2} = X_5 - 2X_6 + X_7$
7	3	$R_{2,3} = X_2 - 3X_3 + 3X_4 - X_5$
8		$R_{3,3} = X_3 - 3X_4 + 3X_5 - X_6$
9		$R_{4,3} = X_4 - 3X_5 + 3X_6 - X_7$
10		$R_{5,3} = X_5 - 3X_6 + 3X_7 - X_8$
11	4	$R_{1,4} = X_1 - 4X_2 + 6X_3 - 4X_4 + X_5$
12		$R_{2,4} = X_2 - 4X_3 + 6X_4 - 4X_5 + X_6$
13		$R_{3,4} = X_3 - 4X_4 + 6X_5 - 4X_6 + X_7$
14		$R_{4,4} = X_4 - 4X_5 + 6X_6 - 4X_7 + X_8$
15		$R_{5,4} = X_5 - 4X_6 + 6X_7 - 4X_8 + X_9$

3. Полученные уравнения изменим, переписав их относительно X_5 , получим: (см. табл. 2)

Таблица 2

№ п/п	Степень разности	уравнения
1	0	$X_5 = R_{5,0} = X_5$
2	1	$X_5 = R_{4,1} + X_4$
3		$X_5 = -R_{3,1} + X_6$
4	2	$X_5 = R_{3,2} - X_3 + 2X_4$
5		$2X_5 = -R_{4,2} + X_4 - X_6$
6		$X_5 = R_{5,2} + 2X_6 - X_7$
7	3	$X_5 = -R_{2,3} + X_2 - 3X_3 + 3X_4$
8		$3X_5 = R_{3,3} - X_3 + 3X_4 + X_6$
9		$3X_5 = -R_{4,3} + X_4 + 3X_6 - X_7$
10		$X_5 = R_{5,3} + 3X_6 - 3X_7 + X_8$
11	4	$X_5 = R_{1,4} - X_1 + 4X_2 - 6X_3 + 4X_4$
12		$4X_5 = -R_{2,4} + X_2 - 4X_3 + 6X_4 + X_6$
13		$6X_5 = R_{3,4} - X_3 + 4X_4 + 4X_6 - X_7$
14		$4X_5 = -R_{4,4} + X_4 + 6X_6 - 4X_7 + X_8$
15		$X_5 = R_{5,4} + 4X_6 - 6X_7 + 4X_8 - X_9$

4. Далее полученную систему линейных уравнений решим относительно X_5 суммировав их левые и правые части, в результате получим:

$$X_5 = 0.32(R_{4,1} - R_{5,1} + R_{3,2} - R_{4,2} + R_{5,2} - R_{2,3} + R_{3,3} - R_{4,3} + R_{5,3} + R_{1,4} - R_{2,4} + R_{3,4} - R_{4,4} + R_{5,4}) - 0.03X_1 + 0.194X_2 - 0.52X_3 + 0.839X_4 + 0.032X_5 + 0.839X_6 - 0.52X_7 + 0.194X_8 - 0.03X_9 \quad (8)$$

5. Формула приблизительного восстановления исходной последовательности чисел исходя из знаковых признаков ее переменных разностей.

Проведем замену переменных R и X в формуле 8 с учетом их знаковых признаков Z.

Для этого переменные в уравнении заменим на их математические ожидания. При этом припустим, что закон распределения плотности вероятностей Гауссовского типа и таков, что математические ожидания положительных и отрицательных значений разностей тождественно равны их среднеквадратическим отклонениям умноженным на знаковый признак Z.

Как уже отмечалось выше, для нормированных значений «X» (см. свойство раздел 1.2) СКО(R) зависит от степени разности и равно для:

$$R_{j=1} - 1.414; R_{j=2} - 2.449; R_{j=3} - 4.472; R_{j=4} - 8.366$$

Заменив переменные в уравнении (8) указанным выше образом, решим его относительно X_5 получим:

$$X_5 = 0.046Z_{4,1} - 0.046Z_{5,1} + 0.079Z_{3,2} - 0.08Z_{4,2} + 0.79Z_{5,2} - 0.14Z_{2,3} + 0.144Z_{3,3} - 0.14Z_{4,3} + 0.144Z_{5,3} + 0.27Z_{1,4} - 0.27Z_{2,4} + 0.27Z_{3,4} - 0.27Z_{4,4} + 0.27Z_{5,4} - 0.03Z_{1,0} + 0.194Z_{2,0} - 0.52Z_{3,0} + 0.839Z_{4,0} + 0.032Z_{5,0} + 0.839Z_{6,0} - 0.52Z_{7,0} + 0.194Z_{8,0} - 0.03Z_{9,0} \quad (9)$$

Предлагается ряд улучшений коэффициента линейной корреляции между восстановленной последовательностью по формуле 9 и исходной последовательностью. Таких как: исключение при выводе формулы уравнений №№ 11, 15 из табл. 2; приведение правых частей уравнений из указанной выше таблицы к СКО=1. В результате, указанных выше преобразований формулы 9, получим:

$$X_5 = 0.13Z_{4,1} - 0.13Z_{5,1} + 0.12Z_{3,2} - 0.14Z_{4,2} + 0.12Z_{5,2} - 0.12Z_{2,3} + 0.13Z_{3,3} - 0.13Z_{4,3} + 0.12Z_{5,3} - 0.12Z_{2,4} + 0.13Z_{3,4} - 0.12Z_{4,4} + 0.04Z_{2,0} - 0.23Z_{3,0} + 0.609Z_{4,0} + 0.162Z_{5,0} + 0.609Z_{6,0} - 0.23Z_{7,0} + 0.04Z_{8,0} \quad (10)$$

(Для получения значения X_6 необходимо к порядковым номерам слагаемых "i" в формуле 10 добавить 1, так же и для получения других X_i . Способ получения формулы восстановления для более высоких порядков переменных разностей аналогичный рассмотренному способу, потому в данной работе не рассматривается)

В таблице 3 записаны значения коэффициента линейной корреляции Пирсона r между исходной последовательностью чисел заданных формулой: $X(i) = \text{COS}((6,28/N)*i)$ (где N – количество измерений на период, i – номер числа в последовательности) и восстановленной последовательностью чисел рассчитанными в соответствии с формулой 9 – (столбик 3 табл.) и формулой 10 (столбик 4 табл.).

Таблица 3

№ п/п	N (количество измерений на период)	Коэффициент «r» Без улучшений	Коэффициент «r» с улучшениями
1	0,6	0,945	0,987
2	1,2	0,927	0,979
3	1,8	-0,98	-0,93
4	2,4	-0,91	-0,78
5	3	0,93	0,992
6	3,6	0,877	0,927
7	4,2	0,914	0,959
8	4,8	0,895	0,962
9	5,4	0,871	0,956
10	6	0,93	0,981
11	6,6	0,884	0,953
12	7,2	0,941	0,973
13	7,8	0,958	0,975
14	8,4	0,91	0,969

Как видно из анализа значений коэффициента r при количестве измерений меньше 3 на период можно получить противоположную ожидаемому результату последовательность (см. отрицательные значения r рядки №№ 3, 4 табл.).

Формула 10 дает возможность решить 3, на мой взгляд, важные задачи:

1. Определить весовые коэффициенты в уравнении регрессии (распознавания).
2. Рассчитать примерную последовательность чисел исходя из значений знаковых признаков переменных разностей.
3. Вывчислить коэффициент линейной корреляции между восстановленной исходя из знаковых признаков последовательностью и исходной тем самым проконтролировать качество восстановления.

1.4.2 Весовые коэффициенты «Kv» в формулах расчета распознавательных признаков Pr.

Весовые коэффициенты Kv находящихся при каждом члене $Z_{i,j}$ уравнений регрессии Pr (см. уравнения №№ 12-18) могут быть представлены как значения модулей коэффициентов для каждой степени разности формулы восстановления 10, или упрощенно, как их математические ожидания:

$$Kv_{i,j} = \text{MATOЖ} | Kz_{i,j} | \quad (11)$$

Для формулы восстановления с максимальным порядком разности 4 Kv (11) равны: для 0-й степени разности – 0,27; для 1-й – 0,13; для 2-й – 0,127; для 3-й – 0,125; для 4-й – 0,123.

Таким образом, при расчете признака распознавания Pr, будем учитывать вес каждого члена уравнения $Z_{i,j}$, как $Kv_{i,j}$ (11)

2. Составление статистического портрета на примере исследования цен на золото в долларах США в годы, что предшествовали современным мировым финансовым кризисам.

2.1 Анализ среднегодовых цен за 1 тройскую унцию золота в долларах США за период с 1966 по 2017 год [3] (см. табл. 4, рис. 1).

За исследуемый период времени (1966-2017 год), как известно, было 4 мировых финансовых кризиса, которые начинались в США, имели разные первоначальные причины

Таблица 4

год	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980
цена	38	41	43	48	40	60	75	110	150	180	160	170	200	390	600
год	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
цена	510	395	398	350	325	380	480	450	398	395	385	365	387	399	400
год	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
цена	395	355	305	301	298	296	341	385	410	494	602	745	898	1000	1230
год	2011	2012	2013	2014	2015	2016	2017								
цена	1610	1700	1450	1250	1166	1250	1240								

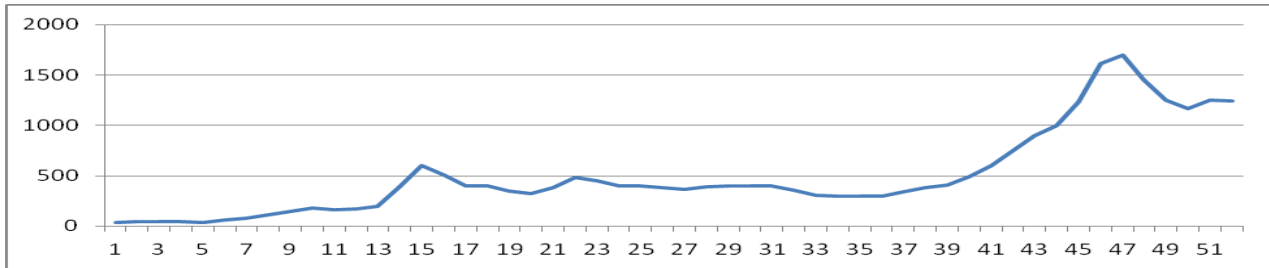


Рис. 1

и разные последствия для экономики США и всего мира [4]. Это финансовые кризисы: 1973, 1987, 200-2003, 2008-20012 годов.

Для проведения исследования влияния цен золота в годы, что предшествовали этим кризисным годам, на начало кризиса выберем максимальную степень разности $R_{max}=4$, количество анализируемых лет предшествующих

кризису – 7 (значение периода средней частоты изменения цен, см. рис. 1).

То есть, для составления математического портрета выберем 4, так сказать «эталонные», на наш взгляд характеризующие кризис, последовательности цен на золото за 7 лет, что предшествовали началам кризисов. Пронормировав их, в результате получим:

F1:	(1966-1972)	-0,83	-0,61	-0,46	-0,1	-0,69	0,791	1,899
F2:	(1980-1986)	1,564	-0,16	-0,11	-0,83	-1,2	-0,38	1,115
F3:	(1993-1999)	1,003	1,024	0,921	0,095	-0,94	-1,02	-1,08
F4:	(2001-2007)	-1,03	-0,82	-0,7	-0,29	0,235	0,928	1,669

Вычислим между каждыми двумя последовательностями коэффициент корреляции Пирсона r , получим: $r_{1,2} = 0.252$; $r_{1,3} = -0.56$; $r_{1,4} = 0.751$; $r_{2,3} = 0.256$; $r_{2,4} = 0.01$; $r_{3,4} = -0.92$.

Из анализа r видно, что невозможно опреде-

лить «похожести» выборок обычным линейным корреляционным методом.

2.2 Вычислим каждой из последовательностей F переменную разность и заменим ее члены знаковыми признаками. В результате получим:

Для F1:							Для F2:						
-0,83	-0,61	-0,46	-0,1	-0,69	0,791	1,899	1,564	-0,16	-0,11	-0,83	-1,2	-0,38	1,115
0,222	0,148	0,369	-0,59	1,477	1,108	-1,72	0,045	-0,72	-0,37	0,822	1,495		
-0,07	0,222	-0,96	2,068	-0,37		1,764	-0,76	0,344	1,196	0,673			
0,295	-1,18	3,028	-2,44			-2,53	1,106	0,852	-0,52				
-1,48	4,21	-5,47				3,633	-0,25	-1,38					
-1	-1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1	1
1	1	1	-1	1	1		-1	1	-1	-1	1	1	
-1	1	-1	1	-1			1	-1	1	1			
1	-1	1	-1				-1	1	1	-1			
-1	1	-1					1	-1	-1				
Для F3:							Для F4:						
1,003	1,024	0,921	0,095	-0,94	-1,02	-1,08	-1,03	-0,82	-0,7	-0,29	0,235	0,928	1,669
0,021	-0,1	-0,83	-1,03	-0,08	-0,06		0,213	0,121	0,407	0,524	0,693	0,742	
-0,12	-0,72	-0,21	0,95	0,021			-0,09	0,286	0,116	0,17	0,048		
-0,6	0,516	1,156	-0,93				0,378	-0,17	0,053	-0,12			
1,115	0,64	-2,09					-0,55	0,223	-0,17				
1	1	1	1	-1	-1	-1	-1	-1	-1	-1	1	1	1
1	-1	-1	-1	-1	-1		1	1	1	1	1	1	
-1	-1	-1	1	1			-1	1	1	1	1		
-1	1	1	-1				1	-1	1	-1			
1	1	-1					-1	1	-1				

Таблица 5

	Kv=0,27						Kv=0,13						Kv=0,127				Kv=0,125			Kv=0,123					
	j = 0						j = 1						j = 2				j = 3			j = 4					
73	-1	-1	-1	-1	-1	1	1	1	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1		
87	1	1	-1	-1	-1	-1	-1	-1	1	-1	-1	1	-1	1	-1	1	1	1	-1	1	1	-1	1	-1	
00	1	1	1	1	-1	-1	-1	1	1	-1	-1	-1	-1	-1	-1	-1	1	1	-1	1	1	-1	1	1	
08	-1	-1	-1	-1	1	1	1	1	1	1	1	1	-1	-1	1	1	1	-1	1	-1	1	1	-1	1	
№ Pr	1	1	2	2	3	1	1	4	4	2	3	1	2	5	6	3	2	7	3	3	3	7	3	3	6
I	1	2	3	4	5	6	7	1	2	3	4	5	6	1	2	3	4	5	1	2	3	4	1	2	3

2.3 Поле анализа знаковых признаков Z исследуемых рядов.

Расписав знаковые признаки переменных разностей исследуемых рядов построчно в соответствии с очередностью их следования по времени в таблицу 5, «-1» выделив оранжевым цветом, получим (см. табл. 5).

2.4 Найдем в поле анализа таб. 5 повторяющиеся определенным образом знаковые признаки « Z » с учетом их весовых коэффициентов « Kv »

Не трудно заметить размещения особым образом в поле анализа знаковых признаков Z повторяющихся по столбцам. Проанализировав, найдем 7 таких признаков (см. ряд «№ Pr» табл. 7). Расписав их с учетом коэффициентов Kv , получим 7 формул:

$$\text{Pr 1} = 0.27(-Z_{1,0} - Z_{2,0} + Z_{6,0} + Z_{7,0}) + 0.13Z_{5,1}; \quad (12)$$

$$\text{Pr 2} = 0.27(-Z_{3,0} - Z_{4,0}) + 0.13(Z_{3,1} + Z_{6,1}) + 0.127Z_{4,2}; \quad (13)$$

$$\text{Pr 3} = 0.27Z_{3,0} + 0.13Z_{4,1} + 0.127Z_{3,2} + 0.125(-Z_{1,3} + Z_{2,3} - Z_{3,3}) + 0.123(Z_{1,4} - Z_{2,4}); \quad (14)$$

$$\text{Pr 4} = 0.13(Z_{1,1} + Z_{2,1}); \quad (15)$$

$$\text{Pr 5} = -0.127Z_{1,2}; \quad (16)$$

$$\text{Pr 6} = -0.127Z_{2,2} + 0.123Z_{3,4}; \quad (17)$$

$$\text{Pr 7} = 0.127Z_{5,2} + 0.125Z_{4,3}; \quad (18)$$

Признаки Pr 1 – Pr 7, кроме признака Pr 5 могут быть как положительные так и отрицательные по знаку при этом чем больше они по величине (модулю) тем исследуемая последовательность больше соответствует портрету события. Поэтому, при расчете общего признака распознавания (статистического портрета), эти признаки необходимо учитывать по величине их модулей. С учетом вышесказанного, определим формулу общего признака распознавания, математического портрета, для рассматриваемого исследования:

$$\text{Pr}(\text{общ}) = |\text{Pr 1}| + |\text{Pr 2}| + |\text{Pr 3}| + |\text{Pr 4}| + \text{Pr 5} + |\text{Pr 6}| + |\text{Pr 7}|; \quad (19)$$

Для лет, что предшествовали исследуемым мировым финансовым кризисам $\text{Pr}(\text{общ}) = 4.174$.

2.5 Оценка качества рассчитанного математического портрета. Определение порогового значения найденного общего признака распознавания исследуемого события. Выводы.

Для определения порогового значения, как уже отмечалось выше, оценим качественные характеристики портрета распознавания заданного формулой общего признака распознавания $\text{Pr}(\text{общ})$ (19). Для этого проведем расчет $\text{Pr}(\text{общ})$ по формуле 19 для каждого года из тренда цен на золото в долларах США за период с 1973 по 2017 год, в результате получим: (см. табл. 6).

Вероятности правильного распознавания события, P_{np} и вероятности ложного распознавания $P_{лр}$ (ложных тревог), оценим исходя из существующих данных. Анализ результатов исследования значений общего признака статистического распознавания финансовых кризисов $\text{Pr}(\text{общ})$ за исключением 1973, 1987, 2000, 2008 лет, показывает, что за период с 1973 по 2017 год максимального значения $\text{Pr}(\text{общ})=4,174$ не было. Что свидетельствует об не плохом качестве рассчитанного статистического портрета за исследуемый период:

$$P_{лр} \leq 1 / (2017 + 1 - 1973) = 0,022;$$

$$P_{np} \geq 1 - 1 / (2017 + 1 - 1973) = 0,978.$$

Из проведенного исследования также видно, что максимальное значение $\text{Pr}(\text{общ.})$, за исключением кризисных годов, было в 1976 году и равнялось 3,66. Его и будем считать пороговым значением:

$$\text{Pr}(\text{общ})_{\text{пор}} = 3,66$$

Таким образом критерий распознавания исследуемого события, статистический портрет, запишем как неравенство:

$$\text{Pr}(\text{кризиса}) > 3.66 \quad 20$$

Выводы:

1. Рассмотренный пример составления статистического портрета финансового кризиса, методом анализа знаковых признаков переменных

Таблица 6

Год	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982
$\text{Pr}(\text{общ.})$	4,174	3,406	3,404	3,66	2,668	2,116	1,86	3,374	3,15	3,4
Год	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
$\text{Pr}(\text{общ.})$	1,618	2,35	1,348	1,798	4,174	0,774	1,818	1,364	1,27	1,348
Год	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
$\text{Pr}(\text{общ.})$	1,834	3,164	2,118	2,04	1,574	1,074	2,61	4,174	2,418	2,08
Год	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
$\text{Pr}(\text{общ.})$	2,624	2,37	3,16	2,37	2,164	4,174	2,87	3,92	2,674	2,87
Год	2013	2014	2015	2016	2017					
$\text{Pr}(\text{общ.})$	3,424	2,918	1,024	1,044	1,824					

разностей в «поле анализа» показывает возможное существование определенных статистических связей цен на золото в годы, что предшествовали финансовым кризисам между собой и с их началом.

2. Полученный в результате проведенного исследования статистический портрет финансового кризиса заслуживает внимания и более глубокого анализа (выходящего за рамки представленной темы работы) особенно в годы, когда статистический признак распознавания (рассчитанный в работе) $Pr(общ)$ превысит пороговое значение 3,66 и вероятность начала финансового кризиса может возрасти.

Выводы из данного исследования. В предложенной работе детально изложен разработанный автором корреляционно-регрессионный метод статистического анализа знаковых признаков переменных разностей в поле анализа. Конечной целью рассмотренного метода анализа есть составление статистического портрета исследуемого события.

Портрет в свою очередь дает возможность сделать вывод: принадлежит или не принадлежит исследуемый тренд к тому или иному исследуемому явлению с заданными показателями качества распознавания $R_{пр}$, $R_{л}$. Также метод дает широкие возможности выявления нелинейных зависимостей между отдельным событием и группой событий, определенной как эталонной, при проведении научных исследований в любых областях нашей жизни и деятельности. На примере исследования цен на золото в долларах США в годы, что предшествовали финансовым кризисам, что рассмотренный в работе, и в итоге составлением математического портрета финансового кризиса автор показывает, что при помощи метода можно решать круг разных задач, в том числе и финансовых. Конечной целью рассмотренного метода, безусловно, есть желание улучшения качества нашей жизни и деятельности в различных их областях.

Список литературы:

1. Корреляция [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/Корреляция>.
2. Т. Андерсон. Статистический анализ временных рядов. (Перевод с английского И.Г. Журбенко и В.П. Носко.) Москва: Издательство «Мир», 1976 г. – С. 76-94.
3. Goldprice. [Электронный ресурс]. – Режим доступа: <https://goldprice.prg/ru/gold-price.html>.
4. Топ крупнейших мировых финансовых кризисов. [Электронный ресурс]. – Режим доступа: <https://benefit.by/page/show/articles/2187>.

Крутік О.В.

Сквирська державна районна адміністрація

СТАТИСТИЧНЕ РОЗПІЗНАВАННЯ ЧИСЛОВОЇ ПОСЛІДОВНОСТІ МЕТОДОМ АНАЛІЗУ ЗНАКОВИХ ОЗНАК ЗМІННИХ РІЗНИЦЬ В ПОЛІ АНАЛІЗУ СИСТЕМАМИ З ШТУЧНИМ ІНТЕЛЕКТОМ

Анотація

У цій роботі автор пропонує складати статистичні портрети подій, лінійна кореляційна обробка числових характеристик яких не дає істотних результатів, розробленим ним методом кореляційно-регресійного аналізу. Як і в методі Фехнера, в пропонованому методі аналізуються знакові ознаки відхилень від середніх значень величин але не лише досліджуваних послідовностей але і їх змінних різниць в запропонованому автором полі аналізу. Новий метод значно розширює кількість аналізованих чинників, що покращує якісні характеристики кінцевого продукту дослідження – статистичного портрета події. Метод дає широкі можливості для виявлення статистичних залежностей між подіями при проведенні наукових досліджень в будь-яких сферах нашого життя і діяльності. Що автор і показує на прикладі дослідження цін на золото в доларах США в роки, що передували фінансовим кризам і у результаті складанням статистичного портрета фінансової кризи. Кінцевою метою розглянутого методу, безумовно, є бажання поліпшення якості нашого життя і діяльності в різних їх областях.

Ключові слова: статистичне розпізнавання, кореляційно-регресійний аналіз, знакові ознаки, змінні різниці, штучний інтелект, ціна золота, фінансові кризи.

Krutik A.V.

Skvyra State District Administration

**THE STATISTICAL RECOGNITION OF THE NUMBER SEQUENCE
BY THE MEANS OF ANALYZING THE SIGNIFICANT CHARACTERISTIC
OF THE VARIABLE DIFFERENCES IN THE ANALYSIS GROUND
WITH THE ARTIFICIAL INTELLIGENCE SYSTEMS**

Summary

In the work the author suggests to build the statistical portraits of the events by the means of the Correlation-Regression Analysis that was developed by him. Because the Line Correlation Processing of the numerical characteristic of the statistical portraits doesn't give the significant results. Just as in the Fechner Method in the suggested method we analyze the significant characteristics of the deviation from the mean values (mathematical expectations) but not just in the sequence under research but also in its variable difference in the suggested by the author ground of analysis. The new method is significantly extending the number of the factors we analyze which improves the qualitative characteristics of the final product of the research: the statistical portrait of the event. The method gives ample opportunities to educe the statistic dependences among the events during the scientific researches in any sphere of our life and activity. That is what the author is showing us using the example of the USA dollars gold price research in the years before the financial crises and as the result the author builds the statistical portrait of the financial crisis. Obviously, the final cause of the considered method is the desire to improve the quality of our life and activity in the different spheres.

Keywords: statistical recognition, correlation-regression analysis, significant characteristics, variable differences, artificial Intelligence, the gold price, financial crisis.