



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 49 за 2009 г.



Гавриков М.Б., Пестрякова Н.В.,  
Усков А.В., Фарсобина В.В.

Об оценках распознавания и  
структуре расположения  
изображений символов

**Рекомендуемая форма библиографической ссылки:** Об оценках распознавания и структуре расположения изображений символов / М.Б.Гавриков [и др.] // Препринты ИПМ им. М.В.Келдыша. 2009. № 49. 29 с. URL: <http://library.keldysh.ru/preprint.asp?id=2009-49>

**Ордена Ленина  
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ  
имени М.В. Келдыша  
Российской Академии наук**

**М.Б. Гавриков, Н.В. Пестрякова, А.В. Усков, В.В. Фарсобина**

**Об оценках распознавания и структуре  
расположения изображений символов**

**Москва – 2009**

**М.Б. Гавриков, Н.В. Пестрякова, А.В. Усков, В.В. Фарсобина. Об оценках распознавания и структуре расположения изображений символов.**

Препринт Института прикладной математики им. М.В. Келдыша РАН, 2009, 29 страниц, 11 рисунков, библиография: 22 наименования.

Для метода распознавания символов, основанного на полиномиальной регрессии, показано наличие закономерностей в выставлении оценок распознавания. Для каждого из символов найдено расположение правильно, неправильно распознанных изображений этого символа, а также символов, отличных от данного. Исследования проведены для печатных и рукопечатных цифр.

**M.B.Gavrikov, N.V.Pestryakova, A.V.Uskov, V.V.Farsobina. Estimations of recognition and structure of arrangement of images of characters.** Preprint, Inst. Appl. Mathem., Russian Academy of Sciences, 2009, 29 Pages, 11 Figures, 22 References.

For the character recognition method based on polynomial regression, presence of laws in exhibiting of estimations of recognition is shown. For each of symbols location of correctly recognized images of this symbol, and also symbols, distinguished from given is found. Studies are leaded for printed and hand-printed digits.

## **Введение**

Настоящая работа является очередной в серии [1 – 12], посвященной современному точному методу распознавания символов, основанному на регрессионном анализе [13 – 21]. Практика использования данного метода и сравнение его с другими способами распознавания показали, что он удовлетворяет высоким требованиям по качеству распознавания, быстродействию, монотонности оценок [2, 3].

При исследовании характеристик рассматриваемого метода возник вопрос, существуют ли закономерности в выставлении оценок распознавания?

В работах [4 – 8, 10] для рукопечатных цифр изучено поведение оценки при увеличении расстояния между растром (полиномиальным вектором) изображения символа и среднестатистическим растром (вектором) этого символа по соответствующей базе.

Для изображений каждой из рассматриваемых цифр (кроме «1») средняя оценка распознавания убывает монотонно по мере удаления от среднестатистического растра или полиномиального вектора этой цифры. Для векторов уровень шумов в указанных зависимостях существенно ниже, а, следовательно, степень монотонности выше, чем для растров. Для «1» средняя оценка сначала монотонно убывает, а затем монотонно увеличивается (также с упомянутыми погрешностями) и принимает максимальное значение 255 вблизи наибольшего удаления от среднестатистического растра (вектора).

Возникла гипотеза о том, что база единиц составлена из двух или более подбаз. При помощи численного эксперимента удалось разделить исходную базу единиц на две подбазы, каждая из которых имеет свой среднестатистический растр и вектор. Для них оценки распознавания и функции распределения ведут себя аналогично зависимостям, найденным для других символов.

В данной работе показано, что полученные для рукопечатных цифр свойства оценки выполняются и для печатных цифр.

В [4 – 8, 10] для среднестатистического растра (полиномиального вектора) каждой из рукопечатных цифр найдены диапазоны, в которых расположены растры (векторы) правильно, неправильно распознанных изображений этой цифры, а также цифр, отличных от данной. В настоящей работе получены аналогичные результаты для печатных цифр.

## 1. Метод распознавания

**1.1. Постановка задачи и методика решения.** Задача распознавания символов состоит в разработке алгоритма, позволяющего по растру изображения определить, какому символу из некоторого конечного множества с  $K$  элементами он соответствует. Представлением символа является растр, состоящий из  $N = N_1 \times N_2$  пикселей. Перенумеровав все пиксели растра, запоминаем в  $i$ -ой компоненте ( $1 \leq i \leq N$ ) вектора  $\mathbf{v} \in \mathbf{R}^N$  состояние  $i$ -го пикселя, а именно, значение на отрезке  $[0, 1]$  для рассматриваемых в работе серых растров.

Пусть для некоторого растра  $\mathbf{v}$  можно найти  $p_k(\mathbf{v})$  – вероятность того, что растр изображает символ с порядковым номером  $k$ ,  $1 \leq k \leq K$ . Тогда распознанным считается символ с порядковым номером  $k_0$ , где

$$p_{k_0}(\mathbf{v}) = \max \{p_k(\mathbf{v})\}, \quad 1 \leq k \leq K \quad (1)$$

Для решения задачи следует вычислить вектор вероятностей  $(p_1(\mathbf{v}), p_2(\mathbf{v}), \dots, p_K(\mathbf{v}))$ . Приближенные значения компонент вектора  $(p_1(\mathbf{v}), \dots, p_K(\mathbf{v}))$  будем искать в виде многочленов от координат  $\mathbf{v} = (v_1, \dots, v_N)$ :

$$p_k(\mathbf{v}) \cong c_0^{(k)} + \sum_{i=1}^N c_i^{(k)} v_i + \sum_{i,j=1}^N c_{i,j}^{(k)} v_i v_j + \dots, \quad 1 \leq k \leq K \quad (2)$$

Суммы в правых частях равенств (2) конечные и определяются выбором базисных мономов. А именно, если

$$\mathbf{x}(\mathbf{v}) = (1, v_1, \dots, v_N, \dots)^T$$

конечный вектор размерности  $L$  из выбранных и приведенных в (2) базисных мономов, упорядоченных определенным образом, то в векторном виде соотношения (2) можно записать так:

$$\mathbf{p}(\mathbf{v}) = (p_1(\mathbf{v}), \dots, p_K(\mathbf{v})) \cong \mathbf{A}^T \mathbf{x}(\mathbf{v}) \quad (2')$$

где  $\mathbf{A}$  – матрица размера  $L \times K$ , столбцами которой являются векторы  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(K)}$ . Каждый такой вектор составлен из коэффициентов при мономах соответствующей строки (2) (с совпадающим верхним индексом), упорядоченных так же, как в векторе  $\mathbf{x}(\mathbf{v})$ . Следовательно, приближенный поиск вектора вероятностей  $\mathbf{p}(\mathbf{v})$  сводится к нахождению матрицы  $\mathbf{A}$ .

Значение  $\mathbf{A}$  вычисляется приближенно в процессе обучения, используя содержащиеся в некоторой базе данных наборы пар векторов  $[\mathbf{v}^{(1)}, \mathbf{y}^{(1)}], \dots, [\mathbf{v}^{(J)}, \mathbf{y}^{(J)}]$ ,

$\mathbf{y}^{(j)}$ ] ( $\mathbf{v}^{(j)}$  образ символа с каким-либо номером  $k$  ( $1 \leq k \leq K$ ) и его базисный вектор  $\mathbf{y}^{(j)} = (0 \dots 1 \dots 0)$ , где 1 стоит на  $k$ -м месте,  $1 \leq j \leq J$ ):

$$\mathbf{A} \cong \left( \frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{x}^{(j)})^T \right)^{-1} \left( \frac{1}{J} \sum_{j=1}^J \mathbf{x}^{(j)} (\mathbf{y}^{(j)})^T \right) \quad (3)$$

При получении правой части (3) используется следующая рекуррентная процедура, где  $\mathbf{A}_0$  задана:

$$\mathbf{A}_j = \mathbf{A}_{j-1} - \alpha \mathbf{G} \mathbf{x}^{(j)} [\mathbf{A}_{j-1}^T \mathbf{x}^{(j)} - \mathbf{y}^{(j)}]^T, \quad \alpha = 1/J, \quad 1 \leq j \leq J \quad (4)$$

$$\mathbf{G} \equiv \mathbf{D}^{-1}, \quad \mathbf{D} = \text{diag}(E\{x_1^2\}, E\{x_2^2\}, \dots, E\{x_L^2\}), \quad E\{x_i^2\} \cong \frac{1}{J} \sum_{j=1}^J (x_i^{(j)})^2, \quad 1 \leq i \leq L$$

**1.2. Практическая реализация.** Изображения печатных и рукопечатных цифр представляли собой серые растры размера  $N = 256 = 16 \times 16$  (Рис.1).

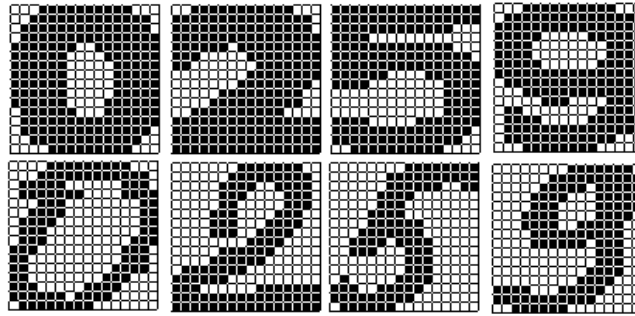


Рис. 1. Образы 16x16 печатных и рукопечатных цифр

Для рукопечатных цифр использовался следующий (длинный) вектор  $\mathbf{x}$ :

$$\begin{aligned} \mathbf{x} = & (1, \{v_i\}, \{v_i^2\}, \{(\delta v_i)_r\}, \{(\delta v_i)_r^2\}, \{(\delta v_i)_y\}, \{(\delta v_i)_y^2\}, \\ & \{(\delta v_i)_r^4\}, \{(\delta v_i)_y^4\}, \{(\delta v_i)_r(\delta v_i)_y\}, \{(\delta v_i)_r^2(\delta v_i)_y^2\}, \{(\delta v_i)_r^4(\delta v_i)_y^4\}, \\ & \{(\delta v_i)_r((\delta v_i)_r)_L\}, \{(\delta v_i)_y((\delta v_i)_y)_L\}, \{(\delta v_i)_r((\delta v_i)_y)_L\}, \\ & \{(\delta v_i)_y((\delta v_i)_r)_L\}, \{(\delta v_i)_r((\delta v_i)_r)_D\}, \{(\delta v_i)_y((\delta v_i)_y)_D\}, \\ & \{(\delta v_i)_r((\delta v_i)_y)_D\}, \{(\delta v_i)_y((\delta v_i)_r)_D\}) \end{aligned} \quad (5)$$

Короткий вектор (для печатных символов) составлен из элементов длинного вектора, записанных в первой строке (5):

$$\mathbf{x} = (1, \{v_i\}, \{v_i^2\}, \{(\delta v_i)_r\}, \{(\delta v_i)_r^2\}, \{(\delta v_i)_y\}, \{(\delta v_i)_y^2\}) \quad (6)$$

В (5) и (6) выражения в фигурных скобках соответствуют цепочкам элементов вектора, вычисляемым по всем пикселям растра (за исключением указанных ниже случаев). Через  $(\delta v_i)_r$  и  $(\delta v_i)_y$  обозначены конечные центральные разности величин  $v_i$  по ортогональным направлениям ориентации растра – нижние индексы  $r$  и  $y$  соответственно. Если имеется нижний индекс  $L$  (left) или  $D$  (down), то это означает, что соответствующие величины относятся к пикселю слева или снизу от рассматриваемого. Компоненты вектора  $\mathbf{x}$ , не имеющие индекса  $L$  или  $D$ , вычисляются для всех пикселей растра; с индексом  $L$  – кроме левых граничных; с индексом  $D$  – кроме нижних граничных пикселей. Вне растра считаем, что  $v_i = 0$  (используется при вычислении конечных разностей на границе растра).

Получаемые оценки могут выходить за рамки отрезка  $[0, 1]$  из-за того, что используемый метод является приближенным. Отрицательные значения искусственно обнулялись, а те, которые были больше 1, делались равными 1.

## 2. Распознавание символов обучающей базы

**2.1. Особенности обучения и распознавания.** При анализе программной реализации метода распознавания для печатных (рукопечатных) цифр как обучение, так и распознавание проводилось на одной и той же графической базе данных, содержащей изображения и коды символов. Это служит своего рода гарантом «чистоты эксперимента», поскольку результат распознавания на символьных последовательностях, «посторонних» для обучающей базы, может сильно отличаться для разных последовательностей. Причиной является большая или меньшая степень схожести баз обучения и распознавания. Данная проблема подробно рассматривалась в [12].

Для рукопечатных цифр при обучении и распознавании использовалась модификация длинного вектора  $\mathbf{x}$  – (5). После многократного обучения на базе в 174 778 элементов полученная матрица обеспечивает распознавание на той же базе 99,5% элементов (881 изображение распознано неверно). Для печатных цифр использовался короткий вектор  $\mathbf{x}$  – (6). После нескольких циклов обуче-

ния на базе в 5496 элементов распознается 99,95% элементов той же базы (ошибочно распознано 4 изображения).

Результатом распознавания образа является код символа и его целочисленная оценка, лежащая в диапазоне  $[0,255]$  (оценка 255 является наилучшей). Эта новая оценка получается следующим образом. В результате умножения оценки на 255 старый диапазон оценок  $[0,1]$  (см. п.1) переходит в новый  $[0,255]$ , после чего проводится дискретизация, а именно,  $[0,1] \rightarrow 1$ ,  $(1,2] \rightarrow 2, \dots$ ,  $(254,255] \rightarrow 255$ .

**2.2. Зависимость оценки распознавания от величины отклонения растра изображения символа от среднестатистического растра.** На рис.2а-11а представлены диаграммы зависимости средней оценки распознавания печатного символа («0», ..., «9») от величины отклонения между растрами его изображений и среднестатистическим растром этого символа по базе. На рис.2д-11д приведены аналогичные результаты, полученные для рукопечатных цифр.

Среднестатистический растр конкретного символа получаем следующим образом. Значение в каждом пикселе с номером  $i$  равно среднему арифметическому значений  $i$ -х пикселей по всем имеющимся в базе растрам рассматриваемого символа.

Расстояние между двумя растрами  $\mathbf{v}=(v_1, \dots, v_N)$  и  $\mathbf{u}=(u_1, \dots, u_N)$  определяем так: вычисляем модуль разности значений в  $i$ -х пикселях, затем суммируем по всем  $N$  пикселям:

$$\|\mathbf{v}-\mathbf{u}\| = \sum_{i=1}^N |v_i - u_i| \quad (7)$$

Диапазон отклонений между растрами распознанных верно изображений символа и среднестатистическим растром этого символа по рассматриваемой базе находится от минимального  $v\_true\_min$  до максимального  $v\_true\_max$ . В Таблице 1 приведены соответствующие значения для печатных, а в Таблице 2 для рукопечатных цифр. Заметим, что для большинства символов  $v\_true\_min_{печатн} < v\_true\_min_{рукопечатн}$  (исключением является цифра «5»),  $v\_true\_max_{печатн} < v\_true\_max_{рукопечатн}$  (кроме «3»).

Делим отрезок  $[v\_true\_min, v\_true\_max]$  (оси абсцисс на рис.2а-11а, рис.2д-11д) на 5 равных по длине частей – отрезок и 4 полуинтервала:  $[v\_true\_min, v\_true\_min + dv]$ ,  $(v\_true\_min + dv, v\_true\_min + 2dv]$ , ...,



$(v\_true\_min + 4dv, v\_true\_min + 5dv]$ , где  $dv = (v\_true\_max - v\_true\_min)/5$ .  
 Затем для совокупности изображений, попадающих в каждый такой участок, вычисляем среднюю оценку распознавания (оси ординат на рис. 2а-11а, рис.2д-11д).

Таблица 1 (растры, печатные цифры)

сим-вол	v_true_min	v_true_max	v_true_max/ v_true_min	v_false_min	v_false_max	$\gamma v\_min$	$\gamma v\_max$	P <sub>false</sub>
0	17,6	56,2	3,2			44,3	176,8	-
1	38,1	126,7	3,3	74,3	85,3	73,1	159,4	126,0
2	36,2	86,2	2,4			66,6	142,7	-
3	31,9	103,1	3,2	70,6	110,0	61,6	131,7	104,0
4	31,4	97,0	3,1			71,2	143,8	-
5	38,9	100,1	2,6			62,4	136,5	-
6	32,4	74,4	2,3			54,3	152,9	-
7	28,7	85,4	3,0			59,4	166,4	-
8	26,0	68,2	2,6			45,7	145,3	-
9	35,6	88,8	2,5			61,9	142,0	-

Таблица 2 (растры, рукопечатные цифры)

сим-вол	v_true_min	v_true_max	v_true_max/ v_true_min	v_false_min	v_false_max	$\gamma v\_min$	$\gamma v\_max$	P <sub>false</sub>
0	35,4	113,6	3,2	56,3	101,1	50,9	171,30	106,6
1	42,6	173,8	4,1	52,2	131,4	48,0	142,30	128,2
2	38,3	105,6	2,8	61,6	109,4	49,7	168,44	120,2
3	39,9	103,1	2,6	55,6	95,5	47,7	171,61	118,8
4	50,8	106,3	2,1	56,0	123,5	55,9	172,25	131,8
5	36,3	130,7	3,6	52,7	98,3	45,9	178,17	126,6
6	44,6	115,1	2,6	55,2	103,1	55,3	165,89	105,3
7	40,6	101,7	2,5	53,5	93,9	49,4	169,91	114,4
8	50,2	119,8	2,4	57,2	115,2	53,7	157,66	121,3
9	47,4	120,6	2,6	54,3	117,4	54,8	164,8	127,5

Напомним, что для рукопечатных цифр в работах [4 – 8, 10] отрезок  $[v\_true\_min, v\_true\_max]$  был поделен на 20 равных по длине частей. На соот-

ветствующих диаграммах по мере удаления растров от среднестатистического раstra средняя оценка распознавания имела следующую тенденцию: для всех символов (кроме «1») – к убыванию, а для «1» – сначала к убыванию, затем к нарастанию (принимая максимальное значение 255 при наибольшем удалении от среднестатистического образа). На эти закономерности накладывался высокий уровень шумов.

В настоящей работе ввиду существенно меньшего объема базы печатных цифр (по сравнению с рукопечатными) отрезок  $[v\_true\_min, v\_true\_max]$  делился на 5 частей как для печатных, так и для рукопечатных символов.

Для рукопечатных цифр (рис.2д – 11д) следует отметить, что при столь большом количестве изображений и малом числе отрезков деления в полученных зависимостях устраняются шумы (увы, совокупно с потерей точности). Сохраняются особенности немонотонной картины оценки распознавания для «1» и монотонное убывание для остальных цифр.

Для печатных цифр (рис.2а – 11а) поведение оценки распознавания сходно с тем, что наблюдалось в  $[4 - 8, 10]$  при делении отрезка  $[v\_true\_min, v\_true\_max]$  на 20 частей для отличных от «1» рукопечатных цифр: диаграммы имеют общую тенденцию к убыванию, но на эту закономерность накладывался высокий уровень шумов (с чем и связано нарушение монотонности для печатных «1» и «6»).

Заметим также, что в базе цифры «0» имеется всего лишь 55 изображений – значительно меньше объема баз остальных печатных цифр (от 282 до 967 изображений). Объяснение этого факта очевидно. Исходная база содержит как буквы, так и цифры. Среди цифр только указанная имеет неотличимую по начертанию букву, и изображения первой попали в базу второй и наоборот. Для реального распознавания это не имеет значения, поскольку при этом не делается различия между идентичными по написанию символами, а в проводимом исследовании данное обстоятельство оказалось существенным. Соответствующий уровень оценок (за исключением участка, наиболее удаленного от среднестатистического раstra) намного ниже, чем для остальных цифр: на самом левом участке для «0» средняя оценка равна 192,4, тогда как для остальных цифр оценки лежат в диапазоне от 241,9 до 251,5. Следует отметить, что все изображения цифры «0» распознаны безошибочно.

На рис.2б – 11б для печатных цифр и на рис.2е – 11е для рукопечатных приведены диаграммы – «дискретный» аналог функции распределения для распознанных верно изображений каждого из символов «0», ... , «9» в терминах растров. А именно, ось абсцисс такая же, как указано в предыдущих абзацах, а

по оси ординат отложено количество правильно распознанных изображений, попавших в каждую пятую часть отрезка  $[v\_true\_min, v\_true\_max]$ . Эти диаграммы аналогичны полученным в [4 – 8, 10] для рукопечатных цифр при делении отрезка  $[v\_true\_min, v\_true\_max]$  на 20 частей.

**2.3. Зависимость оценки распознавания от величины отклонения полиномиального вектора изображения символа от среднестатистического вектора.** На рис.2в – 11в представлены диаграммы зависимости средней оценки распознавания печатного символа («0», ..., «9») от величины отклонения между полиномиальным вектором  $x$ , построенным по его растру, и среднестатистическим полиномиальным вектором этого символа по базе. На рис.2ж-11ж приведены аналогичные результаты, полученные для рукопечатных цифр.

Среднестатистический полиномиальный вектор конкретного символа получаем следующим образом. Значение в каждой  $i$ -й компоненте вектора равно среднему арифметическому значений  $i$ -х компонент векторов, построенных по всем имеющимся в базе изображениям рассматриваемого символа.

Расстояние между двумя векторами  $v=(v_1, \dots, v_L)$  и  $u=(u_1, \dots, u_L)$  определяем так: вычисляем модуль разности значений в  $i$ -х компонентах, затем суммируем по всем  $L$  компонентам:

$$\|v-u\| = \sum_{i=1}^L |v_i - u_i| \quad (8)$$

Диапазон отклонений между полиномиальным вектором распознанного верно изображения символа и «среднестатистическим» вектором этого символа по рассматриваемой базе лежит на отрезке от минимального  $x\_true\_min$  до максимального  $x\_true\_max$ .

Делим отрезок  $[x\_true\_min, x\_true\_max]$  (оси абсцисс на рис.2в-11в, рис.2ж-11ж) на 5 равных по длине частей – отрезок и 4 полуинтервала:  $[x\_true\_min, x\_true\_min + xdv]$ ,  $(x\_true\_min + xdv, x\_true\_min + 2xdv]$ , ...,  $(x\_true\_min + 4xdv, x\_true\_min + 5xdv]$ , где  $xdv = (x\_true\_max - x\_true\_min)/5$ . Затем для совокупности изображений, имеющих полиномиальные векторы, попадающие в каждый такой участок, вычисляем среднюю оценку распознавания (оси ординат на рис. 2в – 11в, рис.2ж – 11ж).

Напомним, что для рукопечатных цифр в работах [5 – 8, 10] отрезок  $[x\_true\_min, x\_true\_max]$  был поделен на 20 частей, и на соответствующих диаграммах средняя оценка распознавания для каждого из рассматриваемых символов убывала монотонно (с некоторыми шумовыми погрешностями) по мере

удаления от среднестатистического вектора, а для «1» сначала монотонно убывала, а затем монотонно увеличивалась (также с некоторыми погрешностями) и принимала максимальное значение 255 на предпоследнем интервале удаления от среднестатистического вектора. Уровень шумов в этих зависимостях был существенно ниже, а, следовательно, степень монотонности выше, чем в аналогичных зависимостях для средней оценки распознавания при отклонении от среднестатистического растра.

Для применяемого в настоящей работе деления отрезка  $[x\_true\_min, x\_true\_max]$  на 5 частей в случае рукопечатных цифр (рис.2ж – 11ж) следует отметить, что (как и в терминах растров) большое количество изображений и малое число отрезков деления приводит к устранению шумов (с понижением точности) по сравнению с результатами предыдущего абзаца. Сохраняются особенности немонотонной картины оценки распознавания для «1» и монотонное убывание для остальных цифр.

В терминах полиномиальных векторов поведение оценки распознавания для печатных цифр (рис.2в – 11в) монотонное. Лишь для «0» наблюдается незначительный шум, что объясняется малым количеством соответствующих изображений в базе символов (как и более низкий уровень оценок по сравнению с остальными цифрами).

На рис.2г – 11г для печатных цифр и на рис.2з – 11з для рукопечатных приведены диаграммы – «дискретный» аналог функции распределения для верно распознанных изображений каждого из символов «0», ..., «9». А именно, ось абсцисс такая же, как указано в предыдущих абзацах, а по оси ординат отложено количество правильно распознанных изображений, попавших в каждую вышеописанную пятую часть отрезка  $[x\_true\_min, x\_true\_max]$ . Эти диаграммы аналогичны полученным в [5 – 8, 10] для рукопечатных цифр при делении отрезка  $[v\_true\_min, v\_true\_max]$  на 20 частей.

**2.4. Расположение правильно, неправильно распознанных и «чужих» символов в терминах растров и полиномиальных векторов.** Для неправильно распознанных образов символа диапазон отклонений между его растрами и «среднестатистическим» растром этого символа по рассматриваемой базе лежит на отрезке от минимального  $v\_false\_min$  до максимального  $v\_false\_max$ . Аналогично, диапазон отклонений между полиномиальным вектором неправильно распознанного изображения символа и «среднестатистическим» вектором этого символа по рассматриваемой базе находится на отрезке от минимального  $x\_false\_min$  до максимального  $x\_false\_max$ .

Для рукопечатных символов «0», ..., «9» согласно [6, 10] в таблице 2 приведены средние оценки неправильного распознавания  $P_{\text{false}}$  и значения величин  $v_{\text{false\_min}}$ ,  $v_{\text{false\_max}}$ , а в таблице 4 соответственно  $x_{\text{false\_min}}$ ,  $x_{\text{false\_max}}$ . Величина  $P_{\text{false}}$  приблизительно в два раза меньше, чем оценки правильного распознавания. Если сравнить  $v_{\text{true\_min}}$  ( $x_{\text{true\_min}}$ ) произвольного символа и  $v_{\text{false\_min}}$  ( $x_{\text{false\_min}}$ ) любого другого (или того же) символа, то  $v_{\text{true\_min}} < v_{\text{false\_min}}$  (аналогично  $x_{\text{true\_min}} < x_{\text{false\_min}}$ ).

Таблица 3 (векторы, печатные цифры)

сим-вол	$x_{\text{true\_min}}$	$x_{\text{true\_max}}$	$x_{\text{true\_max}}/x_{\text{true\_min}}$	$x_{\text{false\_min}}$	$x_{\text{false\_max}}$	$\gamma x_{\text{min}}$	$\gamma x_{\text{max}}$
0	77,6	216,0	2,8	-	-	169,8	515,2
1	128,8	351,2	2,7	244,1	280,9	255,7	475,5
2	143,2	311,1	2,2	-	-	245,2	454,7
3	125,8	354,8	2,8	255,1	369,6	225,8	420,0
4	122,7	321,1	2,6	-	-	239,2	465,2
5	151,6	315,2	2,1	-	-	230,4	423,8
6	131,5	260,1	2,0	-	-	208,6	464,7
7	114,2	296,5	2,6	-	-	218,9	480,4
8	106,3	265,5	2,5	-	-	173,7	467,2
9	138,3	311,3	2,3	-	-	224,6	453,7

Таблица 4 (векторы, рукопечатные цифры)

сим-вол	$x_{\text{true\_min}}$	$x_{\text{true\_max}}$	$x_{\text{true\_max}}/x_{\text{true\_min}}$	$x_{\text{false\_min}}$	$x_{\text{false\_max}}$	$\gamma x_{\text{min}}$	$\gamma x_{\text{max}}$
0	2004	5290	3,21	3002	5119	2186	7754
1	2416	7917	4,08	3046	6437	2698	6915
2	2237	5265	2,75	3491	5523	2860	7710
3	2276	4954	2,58	3026	4936	2697	7723
4	2798	5158	2,09	3135	5619	3148	7729
5	2104	6300	3,60	3142	4909	2558	8057
6	2416	5161	2,58	3375	4966	3004	7437
7	2324	5276	2,51	2913	5621	2788	7645
8	2679	5505	2,39	3335	5188	2901	6977
9	2559	5482	2,55	3076	5313	2964	7458

Ввиду безошибочного распознавания всех печатных символов, за исключением «1» и «3» (для каждого из них распознано неверно по два изображения), в таблице 1 для указанных двух символов приведены средняя оценка неправильного распознавания  $P_{false}$  и значения величин  $v_{false\_min}$ ,  $v_{false\_max}$ , а в таблице 3 соответственно  $x_{false\_min}$ ,  $x_{false\_max}$ . Величина  $P_{false}$  значительно меньше, чем оценки правильного распознавания. Если сравнить  $v_{true\_min}$  ( $x_{true\_min}$ ) некоторого символа и  $v_{false\_min}$  ( $x_{false\_min}$ ) для «1» или «3», то  $v_{true\_min} \ll v_{false\_min}$  ( $x_{true\_min} \ll x_{false\_min}$ ) – отношение второй величины к первой существенно больше, чем для рукопечатных символов.

Для произвольного символа  $G$  из рассматриваемых цифр «0», ..., «9» диапазон отклонений между растрами (векторами) изображений всех символов, кроме  $G$ , и «среднестатистическим» растром (вектором) цифры  $G$  по рассматриваемой базе находится от минимального  $\neg v_{min}$  до максимального  $\neg v_{max}$  (соответственно от  $\neg x_{min}$  до  $\neg x_{max}$ ).

Значения этих величин ( $\neg v_{min}$ ,  $\neg v_{max}$ ) для каждого из возможных  $G$ : «0», ..., «9» приведены в таблице 2 согласно [6, 10] (рукопечатные символы) и в таблице 1 (печатные символы). Аналогично ( $\neg x_{min}$ ,  $\neg x_{max}$ ) даны в таблице 4 согласно [6, 10] (рукопечатные символы) и в таблице 3 (печатные символы).

Для печатных цифр  $v_{true\_min}$  ( $x_{true\_min}$ ) произвольного символа меньше, чем  $\neg v_{min}$  ( $\neg x_{min}$ ) того же или какого-либо другого (!) символа. Для рукопечатных цифр выполняется закономерность, являющаяся частным случаем только что приведенной:  $v_{true\_min}$  ( $x_{true\_min}$ ) некоторого символа меньше, чем  $\neg v_{min}$  ( $\neg x_{min}$ ) того же (!) символа.

В рукопечатном случае при сравнении  $v_{false\_max}$  ( $x_{false\_max}$ ) некоторого символа и  $\neg v_{max}$  ( $\neg x_{max}$ ) произвольного символа имеем:  $v_{false\_max} < \neg v_{max}$  ( $x_{false\_max} < \neg x_{max}$ ). Для печатных символов также выполняются эти неравенства, однако неправильное распознавание наблюдалось лишь для двух символов.

Для каждого рукопечатного символа (кроме «1»), при сравнении  $v_{true\_max}$  ( $x_{true\_max}$ ) этого символа и  $\neg v_{max}$  ( $\neg x_{max}$ ) того же самого или любого другого символа (включая «1») выполняется:  $v_{true\_max} < \neg v_{max}$  ( $x_{true\_max} < \neg x_{max}$ ). При рассмотрении печатных символов указанные закономерности наблюдаются для всех символов.

Кроме того, для рукопечатных цифр любая из трех «минимальных» величин  $v_{true\_min}$ ,  $v_{false\_min}$ ,  $\neg v_{min}$  ( $x_{true\_min}$ ,  $x_{false\_min}$ ,  $\neg x_{min}$ ) некоторого символа меньше, чем любая «максимальная» величина  $v_{true\_max}$ ,  $v_{false\_max}$ ,  $\neg v_{max}$  ( $x_{true\_max}$ ,  $x_{false\_max}$ ,  $\neg x_{max}$ ) того же (или какого-

либо другого) (!) символа. Для печатных цифр выполняется закономерность, являющаяся частным случаем только что приведенной: любая из трех «минимальных» величин  $v\_true\_min$ ,  $v\_false\_min$ ,  $\neg v\_min$  ( $x\_true\_min$ ,  $x\_false\_min$ ,  $\neg x\_min$ ) некоторого символа меньше, чем любая «максимальная» величина  $v\_true\_max$ ,  $v\_false\_max$ ,  $\neg v\_max$  ( $x\_true\_max$ ,  $x\_false\_max$ ,  $\neg x\_max$ ) того же (!) символа.

### 3. Выводы

Итак, проведен сравнительный анализ характера распознавания печатных и рукопечатных цифр.

Качество распознавания первых намного выше, чем вторых. С целью улучшения распознавания рукопечатных символов предусмотрены два фактора. Объем базы обучения для каждого рукопечатного символа значительно больше, чем соответствующего печатного. Используемый как при обучении, так и при распознавании рукопечатных цифр полиномиальный вектор существенно длиннее, чем для печатных цифр (что, естественно, весьма ощутимо сказывается на ухудшении быстродействия).

Отмеченное стремление к увеличению объема базы каждой из рукопечатных цифр (по сравнению с соответствующей печатной) объясняет тот факт, что база рукопечатных единиц, как было продемонстрировано в предыдущих работах, оказалась составленной из двух подбаз. Это, в свою очередь, повлияло на наличие немонотонности в поведении оценки распознавания при удалении от среднестатистического растра (вектора) для данного символа по сравнению с монотонным характером для остальных рукопечатных цифр. Указанное обстоятельство не привело к заметному изменению величин оценки распознавания для «1» в сравнении с остальными цифрами.

В отличие от рукопечатных цифр, для всех печатных поведение оценок распознавания является монотонным. Однако возникла следующая ситуация: из-за перемешивания изображений нулей и буквы с таким же написанием объем базы цифры «0» оказался значительно меньше, чем для остальных печатных цифр. Это привело к понижению оценок распознавания изображений цифры «0». Тем не менее, оно оказалось безошибочным.

Для изображений каждой из рассматриваемых печатных цифр средняя оценка распознавания убывает монотонно по мере удаления от среднестатистического растра или полиномиального вектора этой цифры. В указанных зависимостях для растров наблюдаются шумы, которые практически отсутствуют в

терминах векторов. Следовательно, степень монотонности для векторов выше, чем для растров.

Уменьшение (от 20 до 5) числа частей деления отрезков  $[v\_true\_min, v\_true\_max]$  и  $[x\_true\_min, x\_true\_max]$  в случае рукопечатных цифр привело к устранению шумов в зависимостях для средней оценки распознавания при отклонении как от среднестатистического растра, так и от вектора.

Поведение оценки распознавания для всех печатных цифр и всех (кроме «1») рукопечатных аналогично.

Для рукопечатных (печатных) цифр как в терминах растров, так и в терминах полиномиальных векторов наблюдаются одинаковые закономерности в расположении правильно, неправильно распознанных и «чужих» символов. Для печатного и рукопечатного алфавитов имеется как сходство, так и различие.

Следует подчеркнуть, что для рассматриваемого набора рукопечатных цифр максимальное расстояние от среднестатистического растра (вектора) до начала отрезка правильного распознавания меньше, чем минимальное расстояние до начала отрезка неправильного распознавания (в случае печатных цифр намного меньше).

Кроме того, для каждой из рукопечатных цифр расстояние от среднестатистического растра (вектора) до начала отрезка правильного распознавания меньше, чем до начала отрезка с «чужими» цифрами. Для рассматриваемого набора печатных цифр максимальное расстояние от среднестатистического растра (вектора) до начала отрезка правильного распознавания меньше, чем минимальное расстояние до начала отрезка «чужих» цифр.

Для рассматриваемого набора рукопечатных цифр максимальное расстояние от среднестатистического растра (вектора) до конца отрезка неправильного распознавания меньше, чем минимальное расстояние до конца отрезка «чужих» цифр (аналогично для печатных цифр).

Для рассматриваемого набора рукопечатных цифр максимальное расстояние (кроме «1») от среднестатистического растра (вектора) до конца отрезка правильного распознавания меньше, чем минимальное расстояние (включая «1») до конца отрезка «чужих» цифр (для печатных цифр без каких-либо исключений).

Кроме того, для каждой из печатных цифр максимальное из расстояний от среднестатистического растра (вектора) до начал отрезков правильного, неправильного распознавания, а также «чужих» цифр меньше, чем минимальное из расстояний от среднестатистического растра (вектора) до концов отрезков правильного, неправильного распознавания, а также «чужих» цифр. Следова-



тельно, для каждого из символов имеется область с правильно, неправильно распознанными, а также чужими символами. Для рассматриваемого набора рукопечатных цифр максимальное из расстояний от среднестатистических растров (векторов) до начал отрезков правильного, неправильного распознавания, а также «чужих» цифр меньше, чем минимальное из расстояний от среднестатистических растров (векторов) до концов отрезков правильного, неправильного распознавания, а также «чужих» цифр. Следовательно, формально совместив в одной точке среднестатистические растры (векторы), имеем, что для совокупности символов имеется область с правильно, неправильно распознанными, а также чужими символами.

## Литература

- [1] *Гавриков М.Б., Пестрякова Н. В.* "Метод полиномиальной регрессии в задачах распознавания печатных и рукопечатных символов", //Препринт ИПМатем. РАН, М., 2004, №22, 12 стр.
- [2] *Гавриков М.Б., Пестрякова Н. В., Славин О.А, Фарсобина В.В.* "Развитие метода полиномиальной регрессии и практическое применение в задаче распознавания", //Препринт ИПМатем. РАН, М., 2006, №25, 21 стр.
- [3] *Гавриков М.Б., Мисюрёв А.В., Пестрякова Н.В., Славин О.А.* "Об одном методе распознавания символов, основанном на полиномиальной регрессии". Автоматика и Телемеханика. 2006, №2, с. 119-134.
- [4] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "О некоторых свойствах метода распознавания символов, основанного на полиномиальной регрессии", //Препринт ИПМатем. РАН, М., 2007, №69, 20 стр.
- [5] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "О некоторых свойствах оценки метода распознавания символов, основанного на полиномиальной регрессии", //Препринт ИПМатем. РАН, М., 2008, №7, 28 стр.
- [6] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "Анализ метода распознавания символов, основанного на полиномиальной регрессии ", //Препринт ИПМатем. РАН, М., 2008, №25, 28 стр.
- [7] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "Оценка распознавания символов для метода, основанного на полиномиальной регрессии", // Труды ИСА РАН под ред. чл.-корр. РАН, проф. Арлазарова В.Л. и д.т.н., проф. Емельянова Н.Е. М., «ЛИБРОКОМ», 2008, с. 194-219.
- [8] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "О среднестатистических растрах и векторах метода распознавания символов, основанного на полиномиальной регрессии", // Труды ИСА РАН под ред. чл.-корр. РАН, проф. Арлазарова В.Л. и д.т.н., проф. Емельянова Н.Е. М., «ЛИБРОКОМ», 2008, с. 220-232.
- [9] *Гавриков М.Б., Пестрякова Н. В., Усков А.В., Фарсобина В.В.* "Зависимость точности и оценки распознавания от степени различия между базами обучения и распознавания", // Труды ИСА РАН под ред. чл.-корр. РАН, проф. Арлазарова В.Л. и д.т.н., проф. Емельянова Н.Е. М., «ЛИБРОКОМ», 2008, с.233-240.
- [10] *Пестрякова Н.В.* "Структуры в распознавании". Информационные технологии и вычислительные системы. 2009, №1, С. 58-71.
- [11] *Пестрякова Н.В.* "Метод распознавания символов, основанный на полиномиальной регрессии", // Высокие технологии, фундаментальные исследова-

дования, образование: сборник трудов Седьмой международной научно-практической конференции «Исследование, разработка и применение высоких технологий в промышленности». 28-30.04.2009, Санкт-Петербург, Россия / под ред. А.П. Кудинова, Г.Г. Матвиенко.- СПб. Изд-во Политехн. ун-та, 2009. С.129-130.

[12] Гавриков М.Б., Пестрякова Н.В., Усков А.В., Фарсобина В.В. "Зависимость качества распознавания от взаимного расположения среднестатистических растров и векторов баз обучения и распознавания", //Препринт ИПМатем. РАН, М., 2009, №19, 28 стр.

[13] *Sebestyen G.S.* Decision Making Processes in Pattern Recognition, MacMillan, New York, 1962.

[14] *Nilson N. J.* Learning Machines, McGraw-Hill, New York, 1965.

[15] *Schürmann J.* Polynomklassifikatoren, Oldenbourg, München, 1977.

[16] *Schürmann J.* Pattern Classification, John Wiley&Sons, Inc., 1996.

[17] *Albert A.E. and Gardner L.A.* Stochastic Approximation and Nonlinear Regression // Research Monograph 42. MIT Press, Cambridge, MA, 1966.

[18] *Becker D. and Schürmann J.* Zur verstärkten Berücksichtigung schlecht erkennbarer Zeichen in der Lernstichprobe // Wissenschaftliche Berichte AEG-Telefunken **45**, 1972, pp. 97 – 105.

[19] *Pao Y.-H.* The Functional Link Net: Basis for an Integrated Neural-Net Computing Environment // in Yoh-Han Pao (ed.) Adaptive Pattern Recognition and Neural Networks, Addison-Wesley, Reading, MA, 1989, pp. 197-222.

[20] *Franke J.* On the Functional Classifier, in Association Francaise pour la Cybernetique Economique et Technique (AFCET), Paris // Proceedings of the First International Conference on Document Analysis and Recognition, St. Malo, 1991, pp.481-489.

[21] *Дж.Себер.* Линейный регрессионный анализ. М.:”Мир”, 1980.

[22] *Линник Ю.В.* Метод наименьших квадратов и основы математико - статистической теории обработки наблюдений. М.:”Физматлит”, 1958.

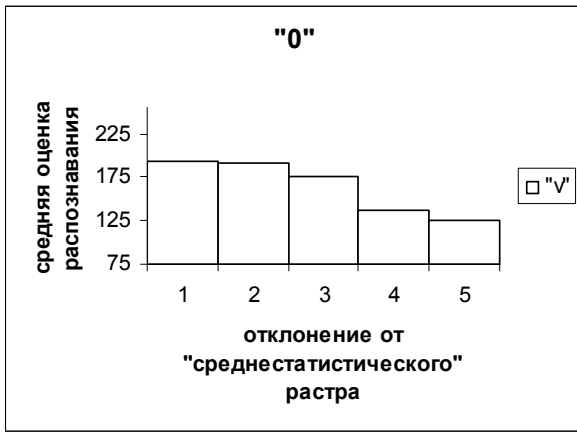


Рис. 2а

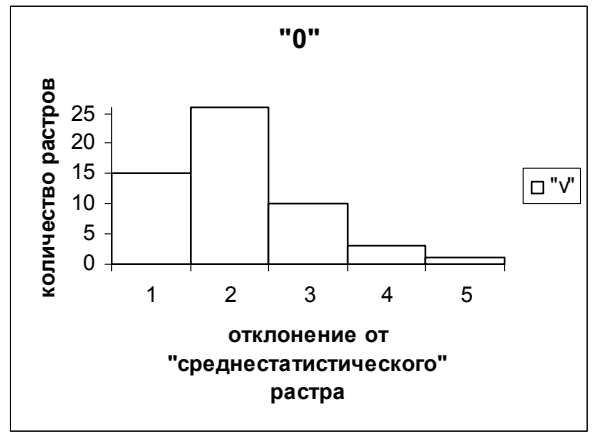


Рис. 2б

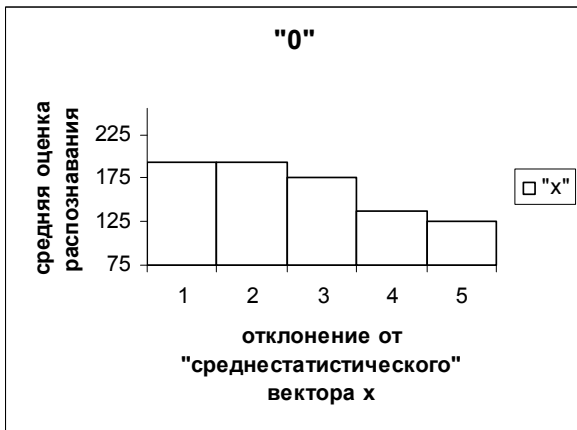


Рис. 2в

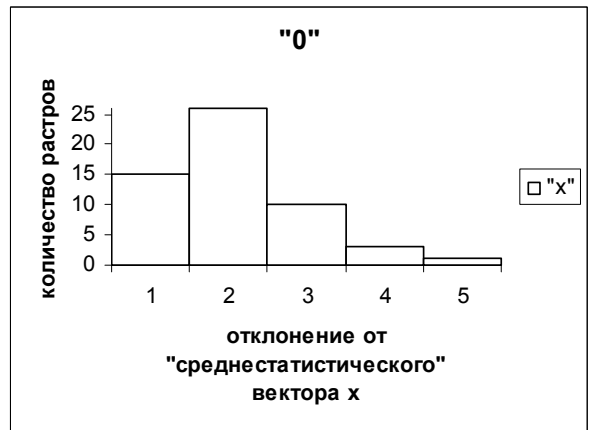


Рис. 2г

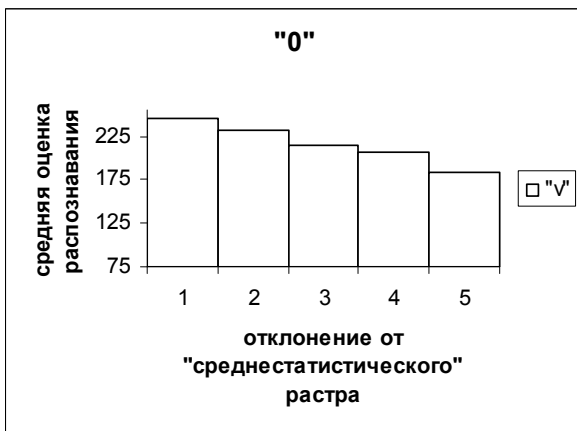


Рис. 2д

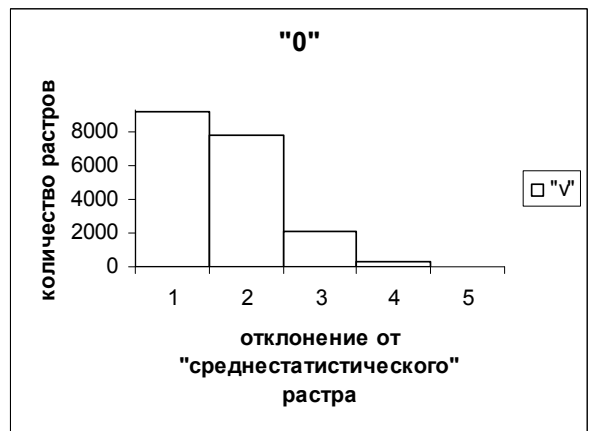


Рис. 2е

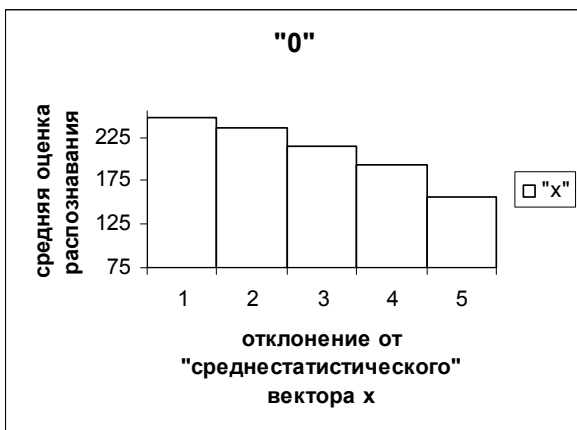


Рис. 2ж

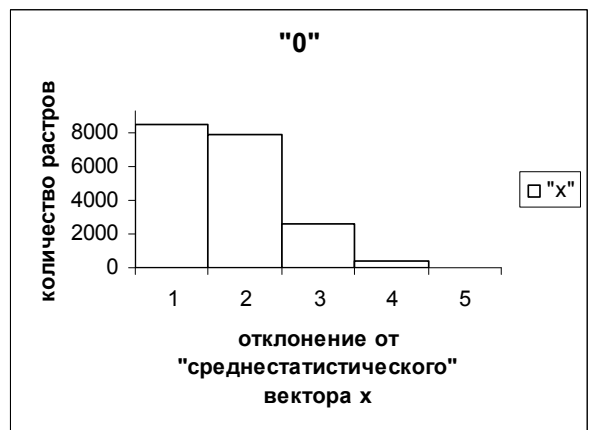


Рис. 2з

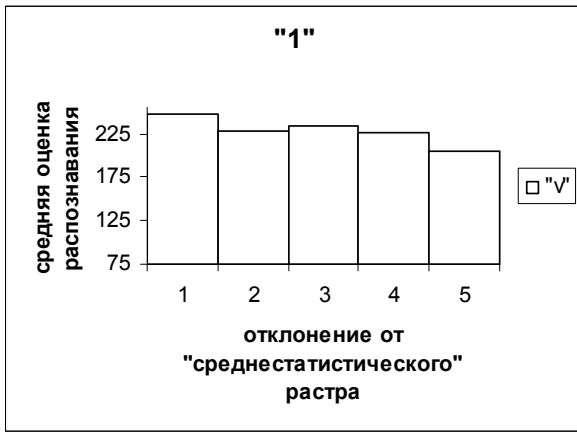


Рис. 3а

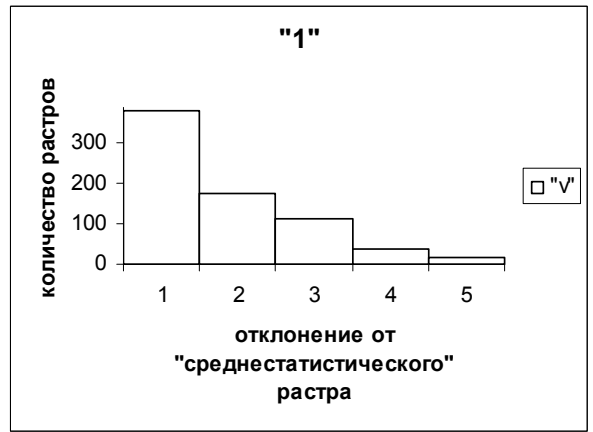


Рис. 3б

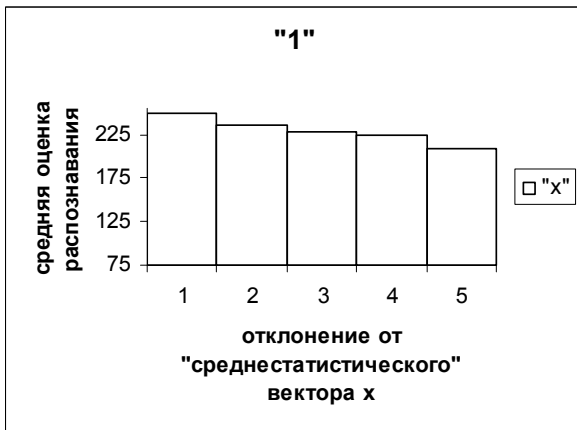


Рис. 3в

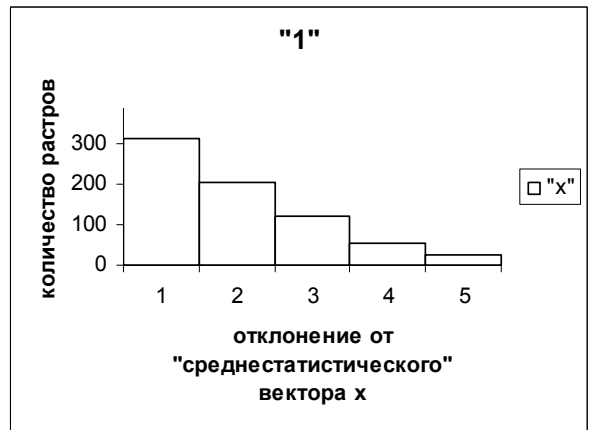


Рис. 3г

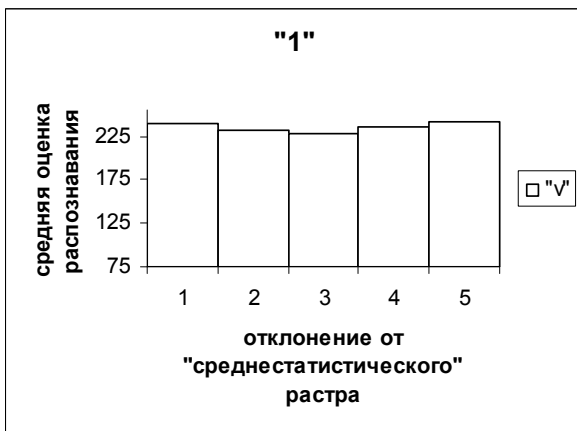


Рис. 3д

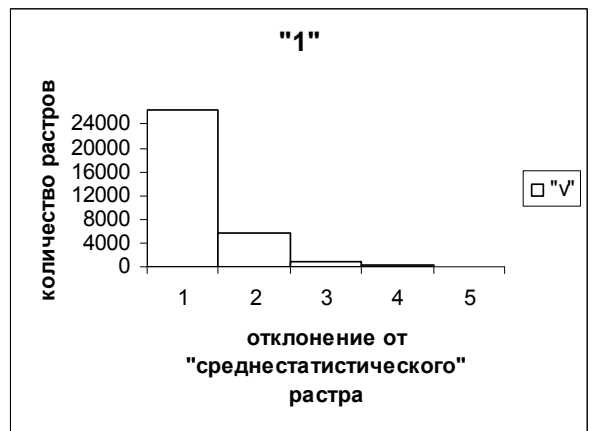


Рис. 3е

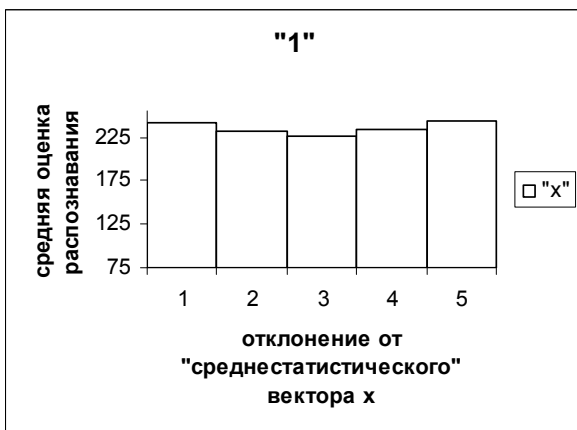


Рис. 3ж

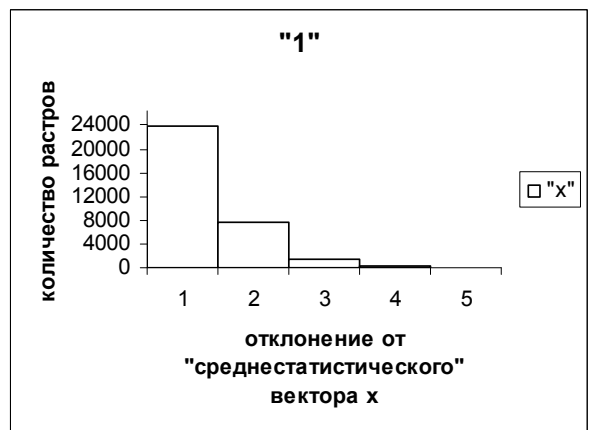


Рис. 3з

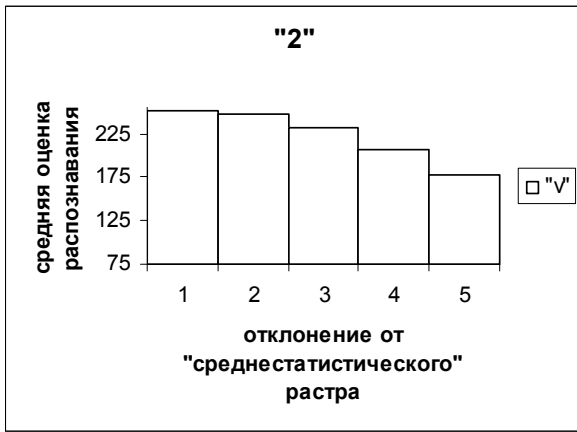


Рис. 4а

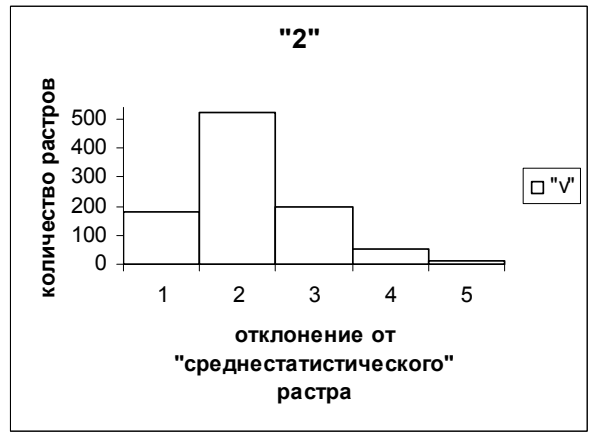


Рис. 4б

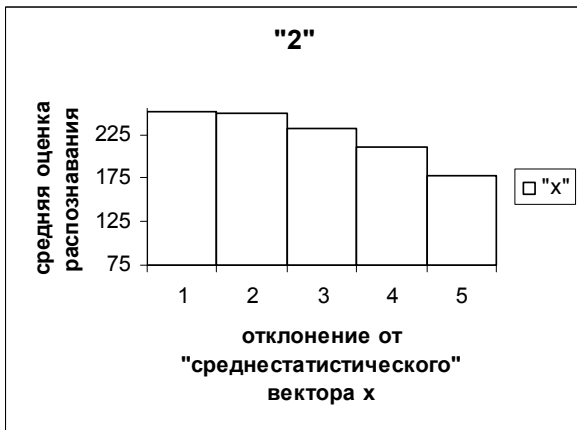


Рис. 4в

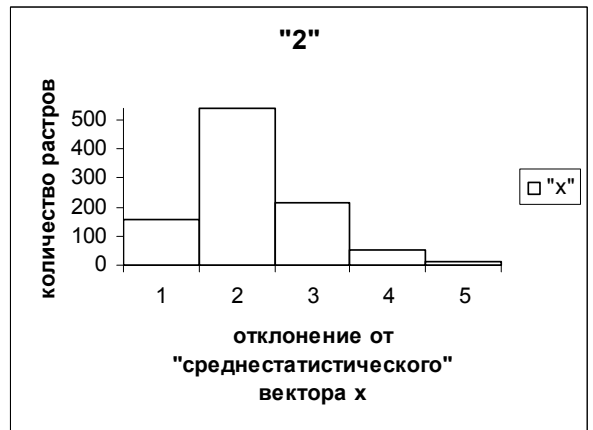


Рис. 4г

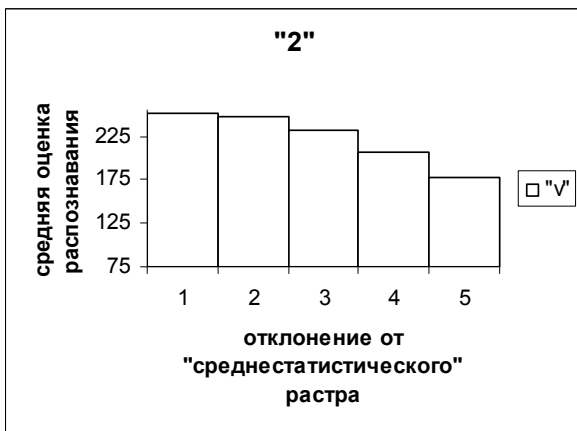


Рис. 4д

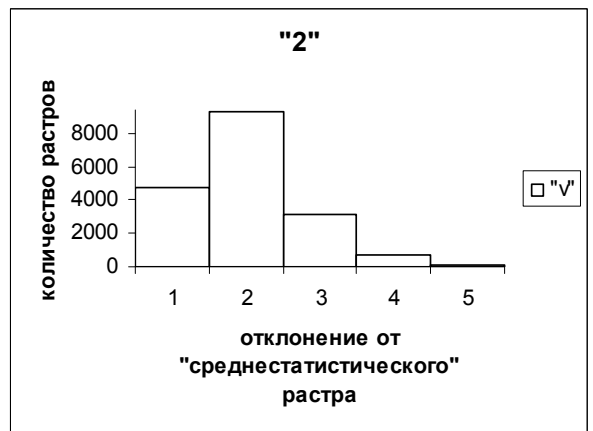


Рис. 4е

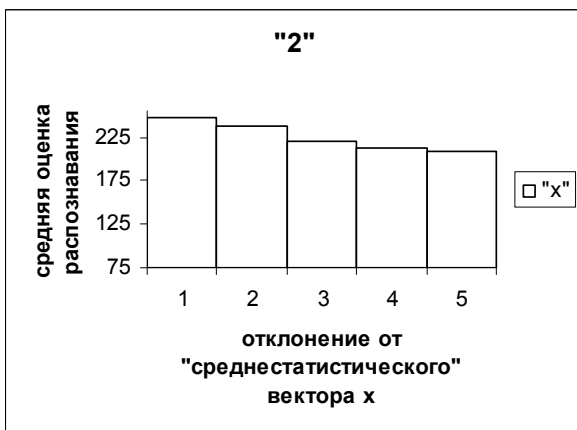


Рис. 4ж

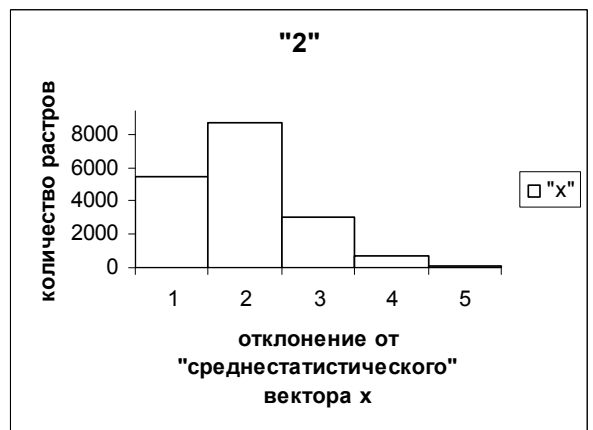


Рис. 4з

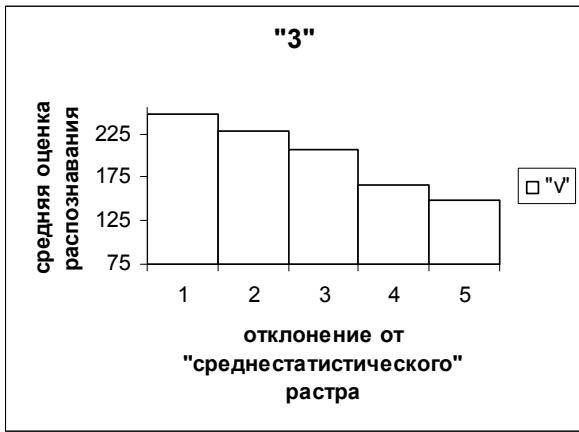


Рис. 5а

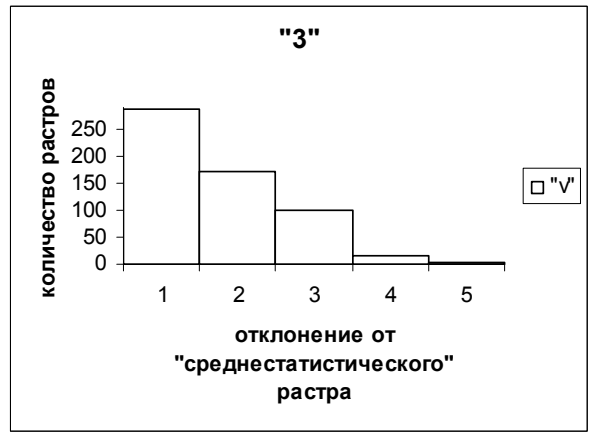


Рис. 5б

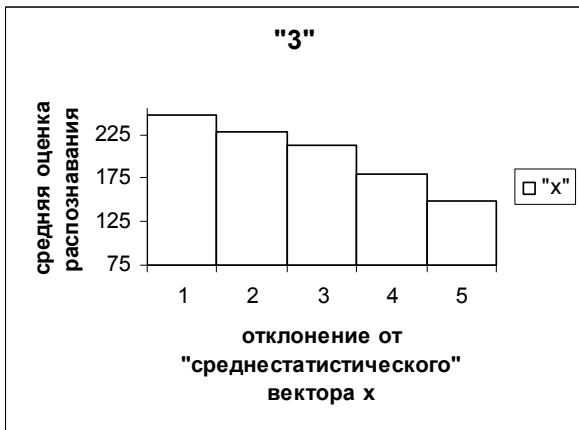


Рис. 5в

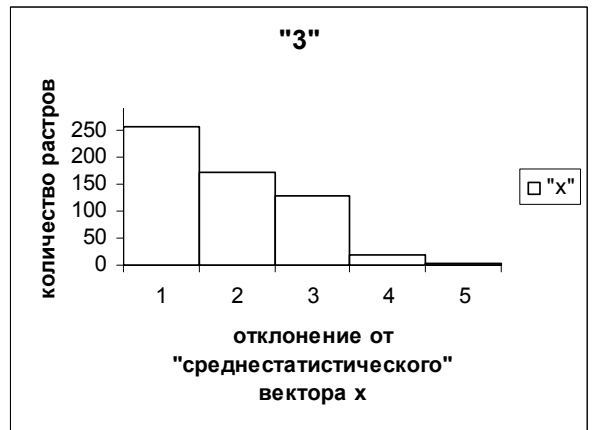


Рис. 5г

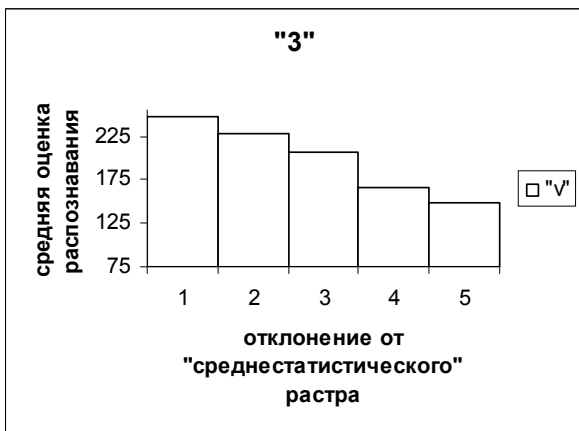


Рис. 5д

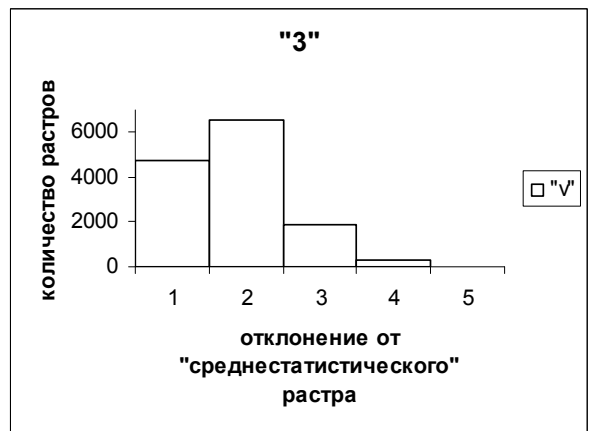


Рис. 5е

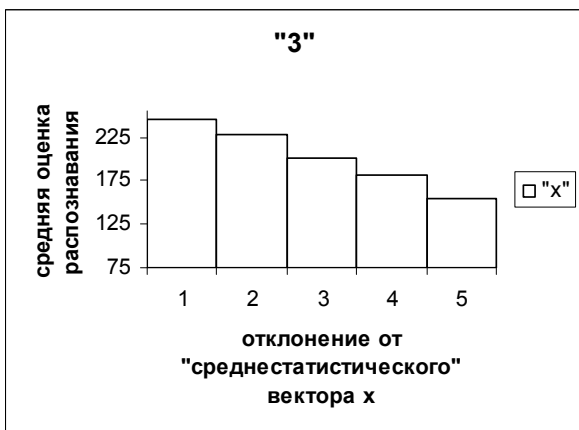


Рис. 5ж

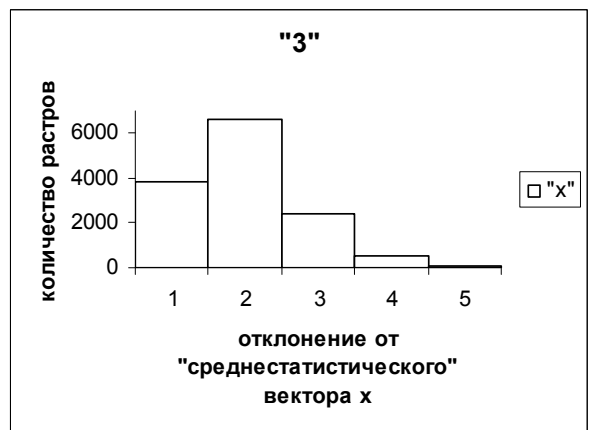


Рис. 5з

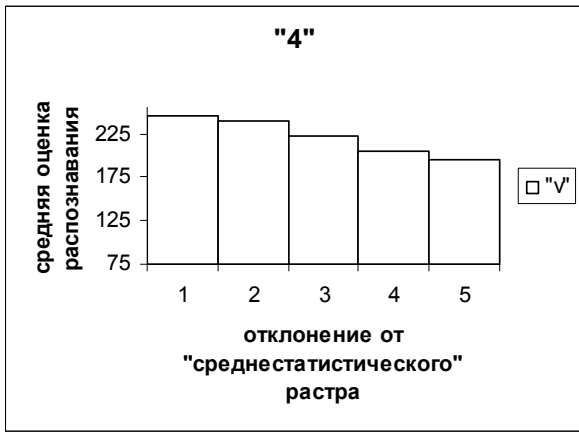


Рис. 6а

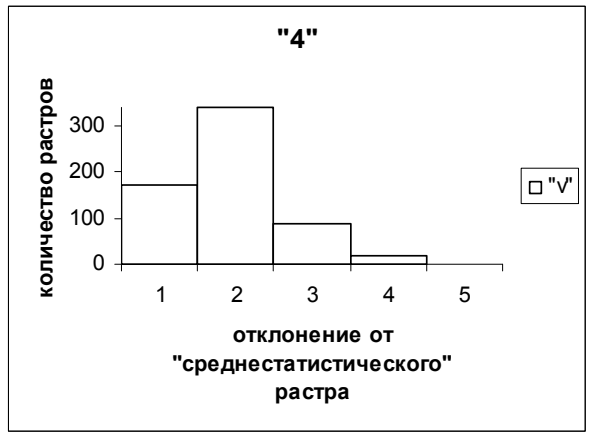


Рис. 6б

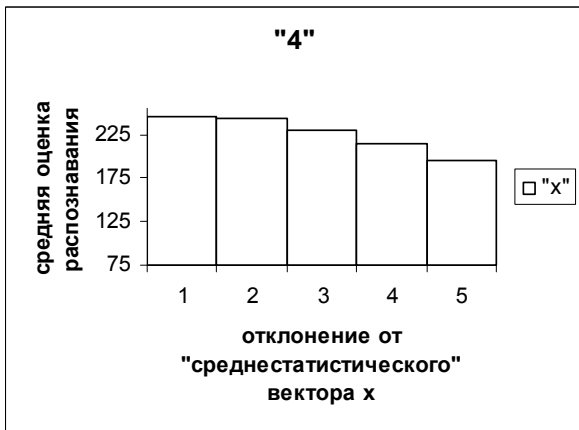


Рис. 6в

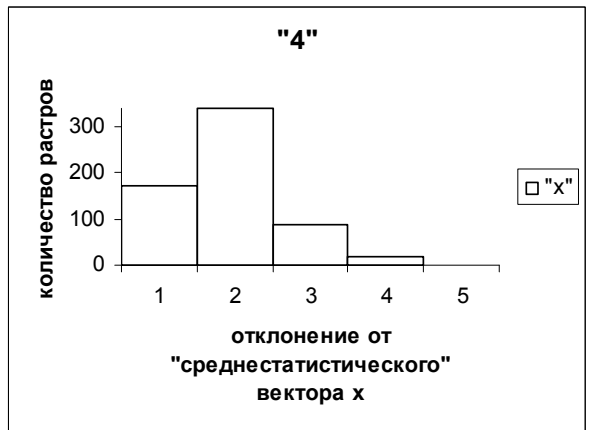


Рис. 6г

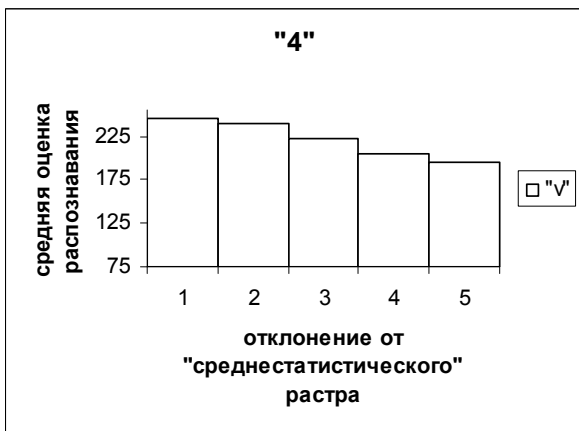


Рис. 6д

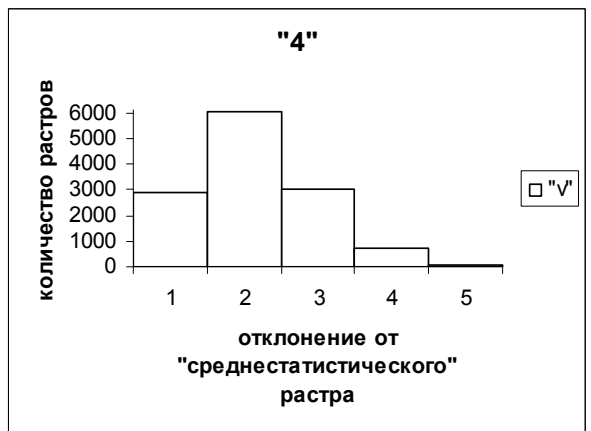


Рис. 6е

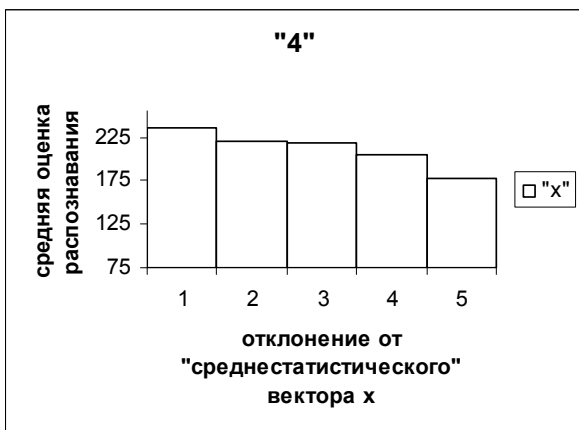


Рис. 6ж

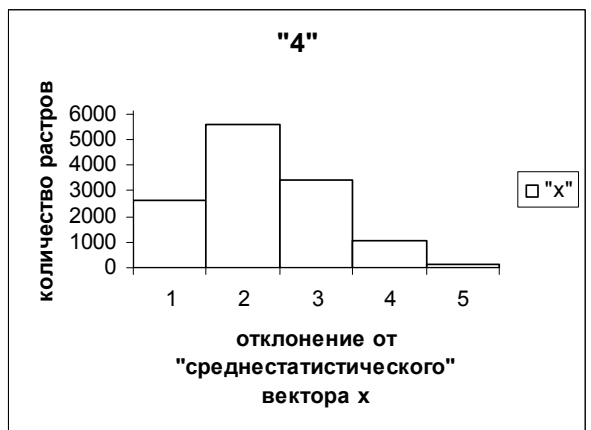


Рис. 6з



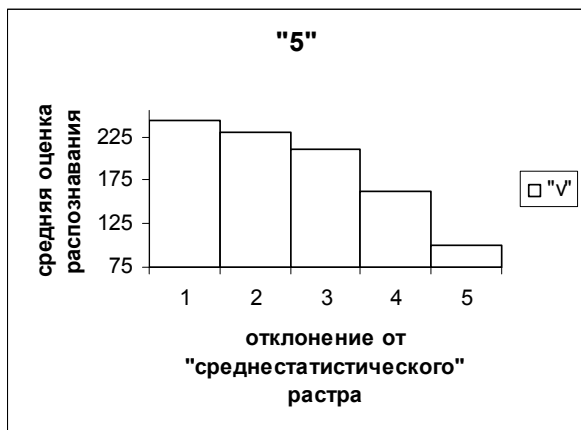


Рис. 7а

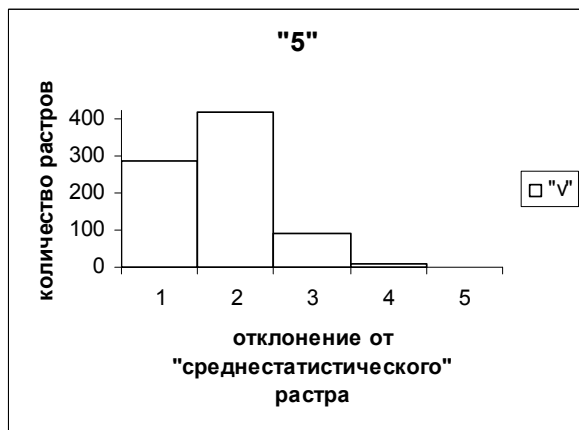


Рис. 7б

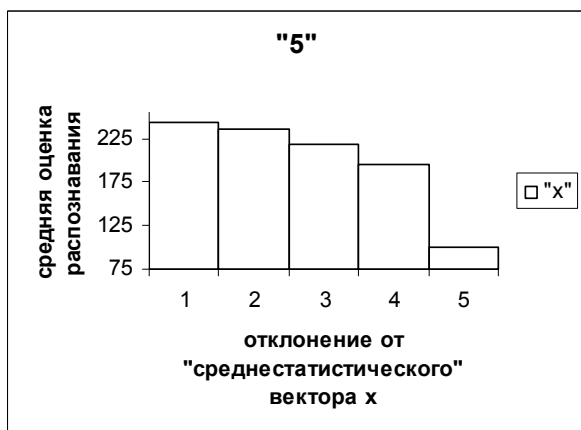


Рис. 7в

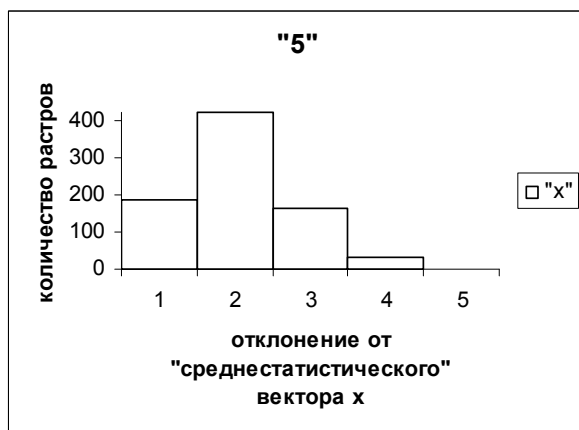


Рис. 7г

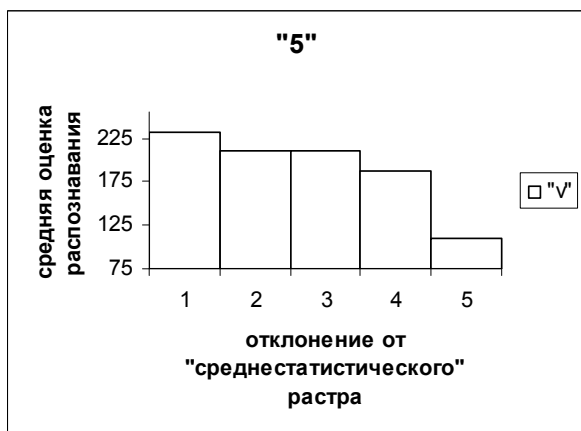


Рис. 7д

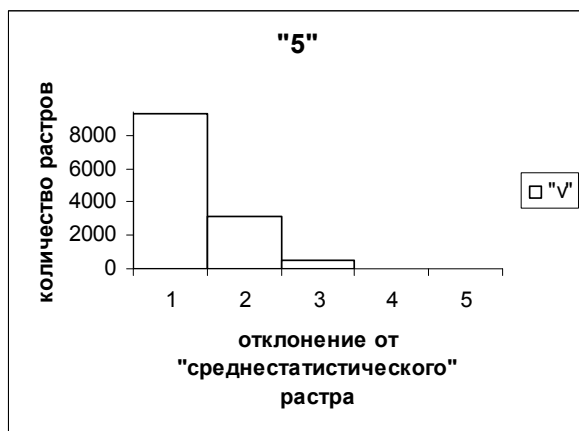


Рис. 7е

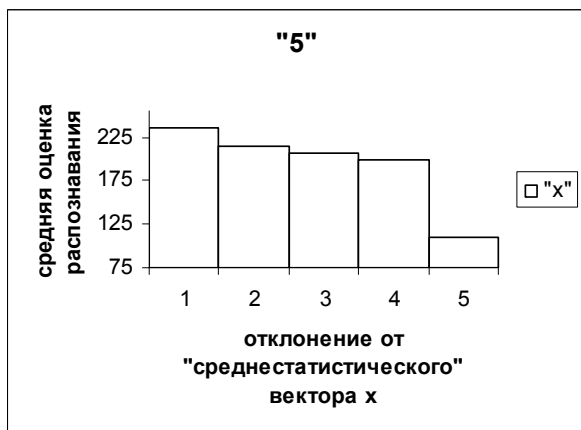


Рис. 7ж

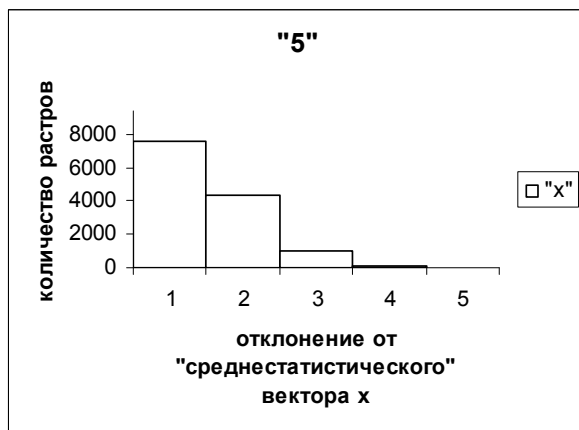


Рис. 7з

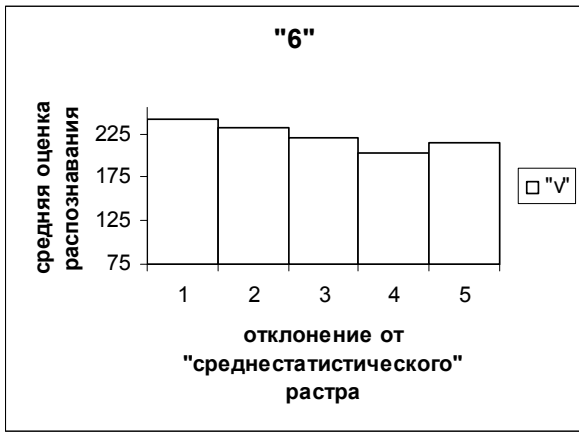


Рис. 8а

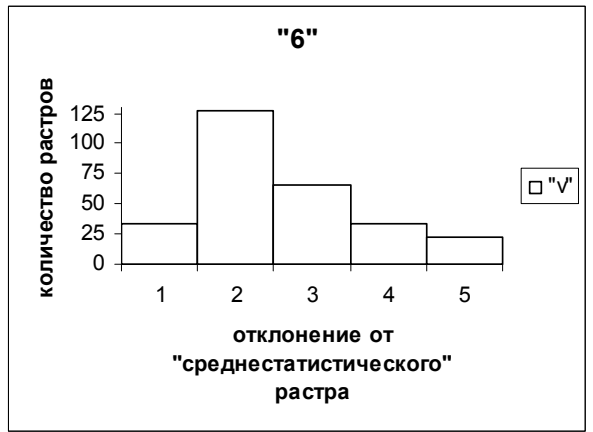


Рис. 8б

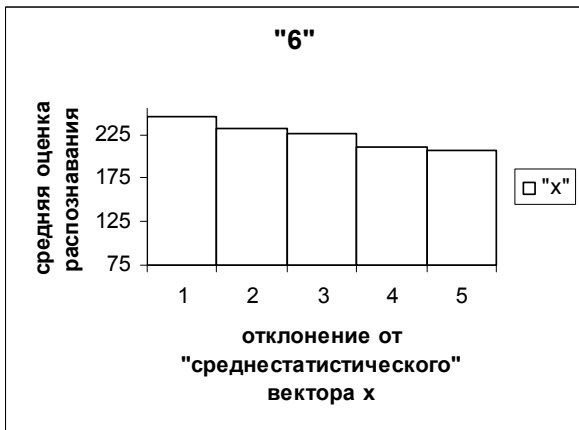


Рис. 8в

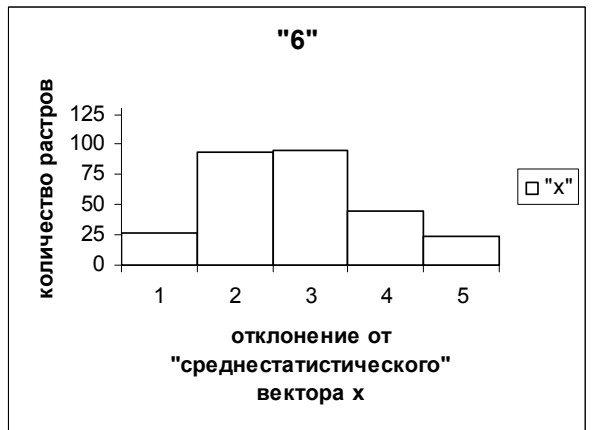


Рис. 8г

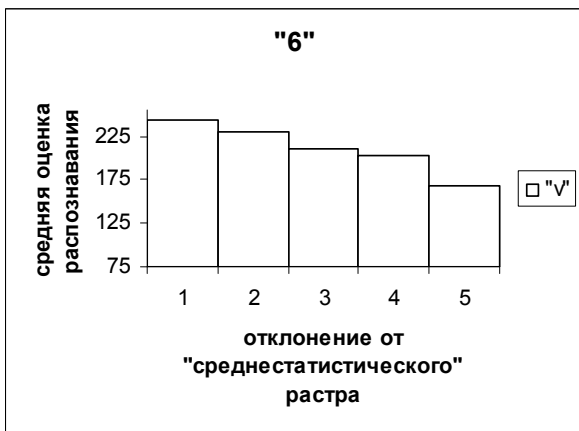


Рис. 8д

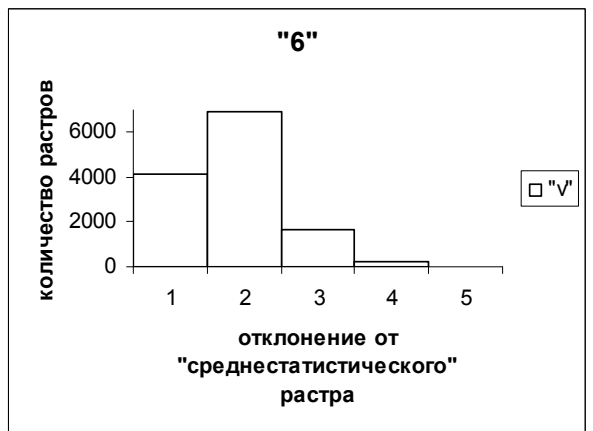


Рис. 8е

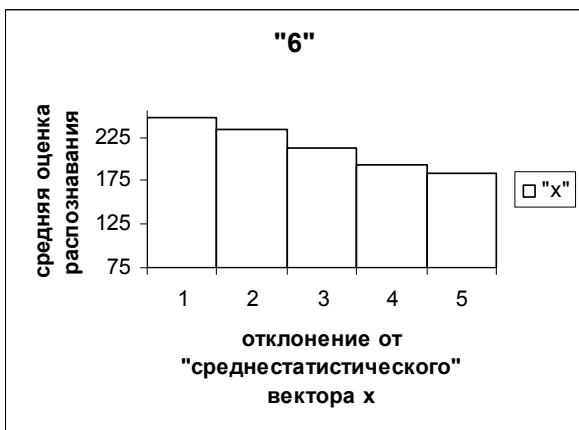


Рис. 8ж

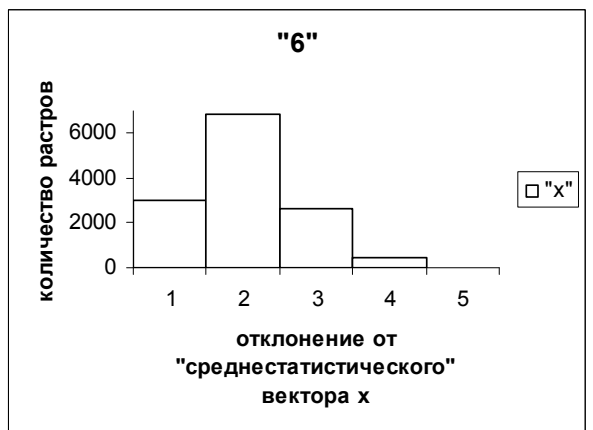


Рис. 8з

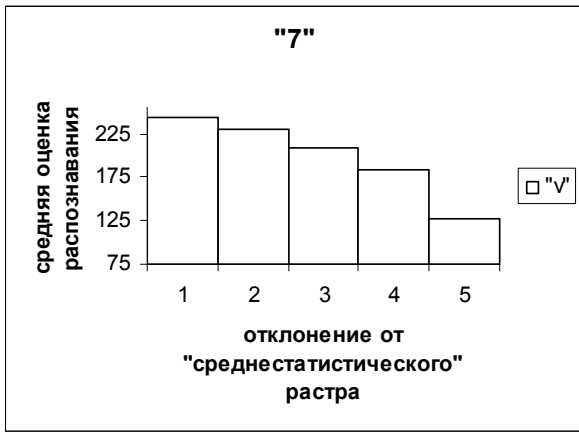


Рис. 9а

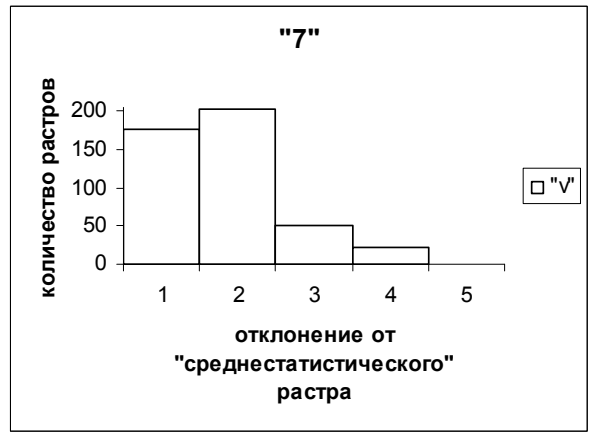


Рис. 9б

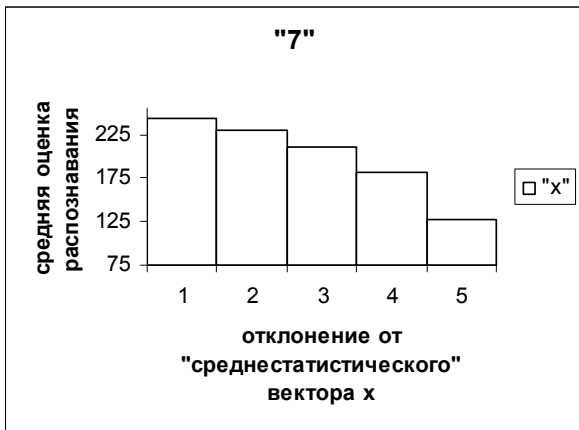


Рис. 9в

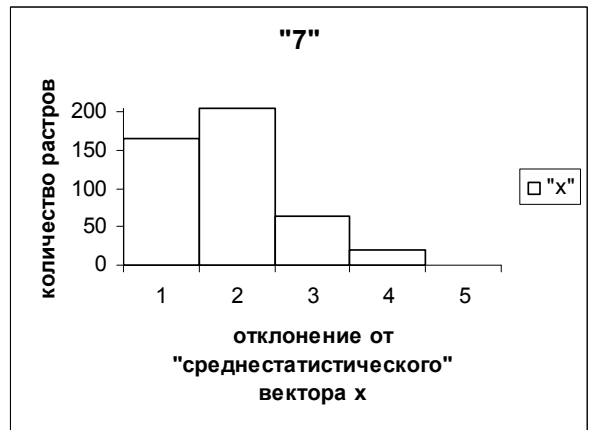


Рис. 9г

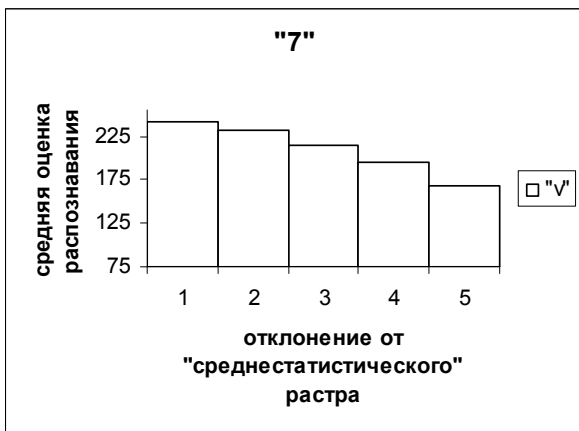


Рис. 9д

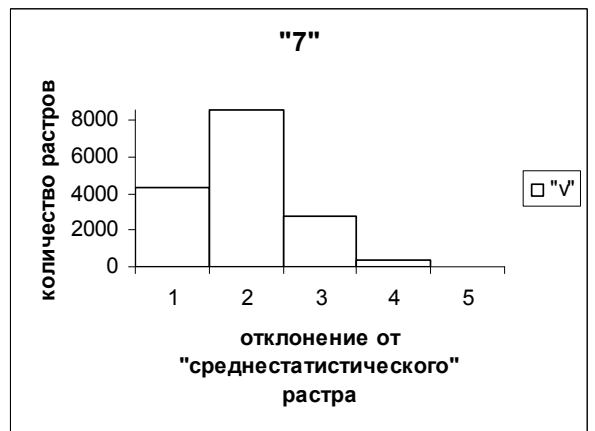


Рис. 9е

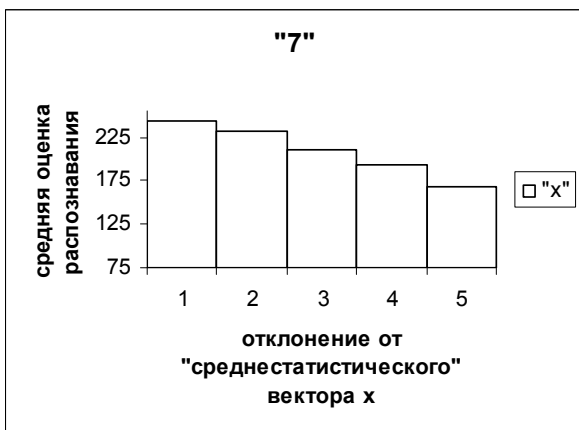


Рис. 9ж

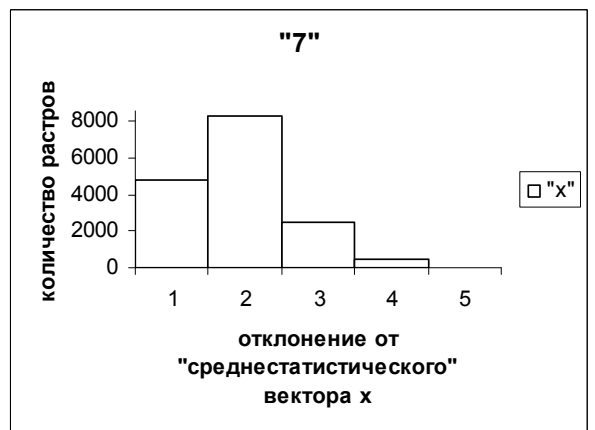


Рис. 9з

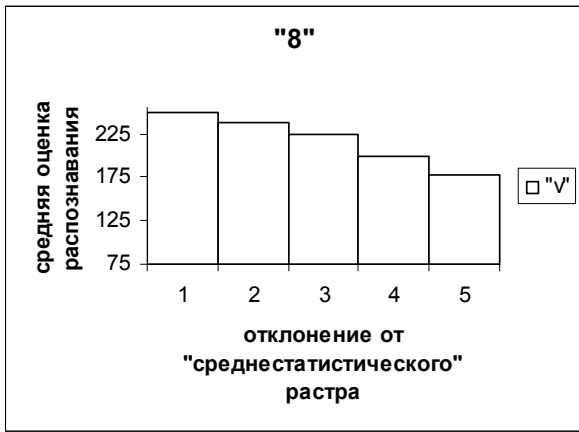


Рис. 10а

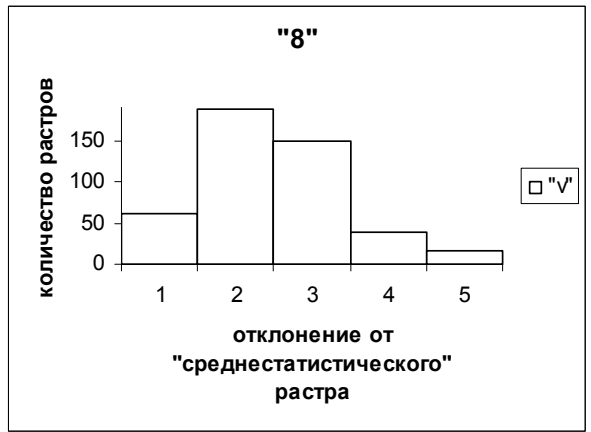


Рис. 10б

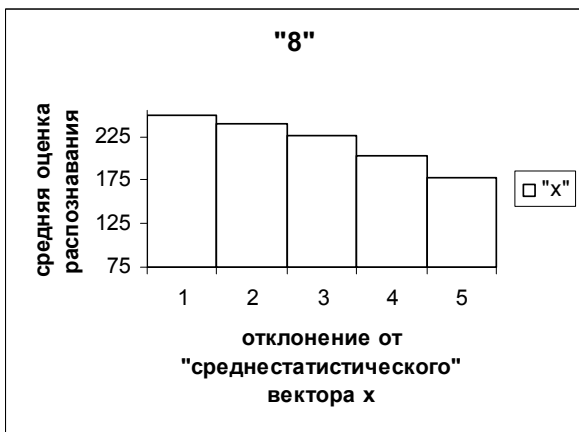


Рис. 10в

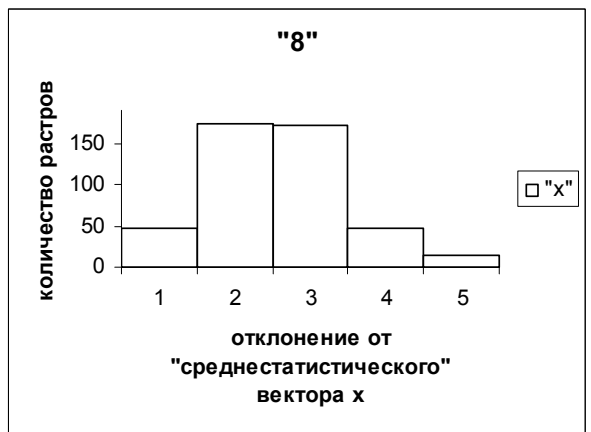


Рис. 10г

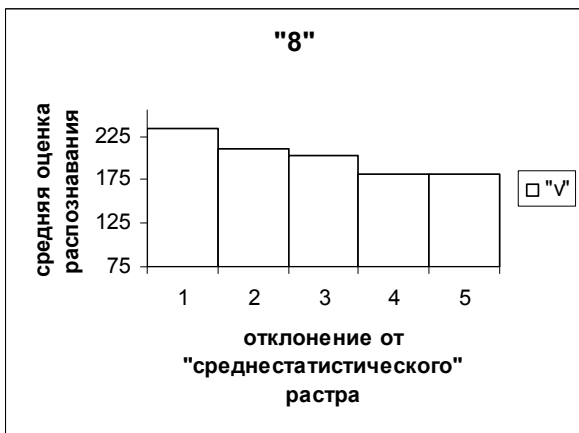


Рис. 10д

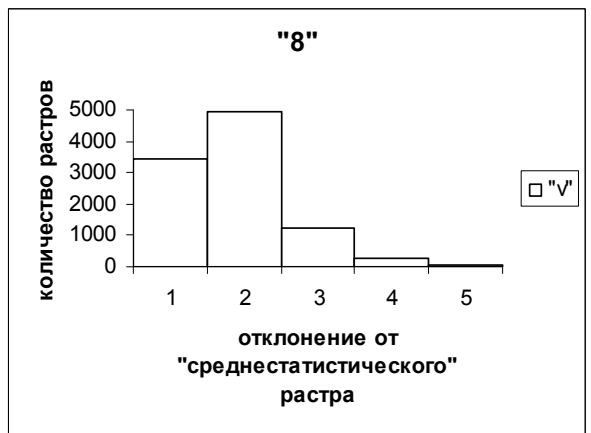


Рис. 10е

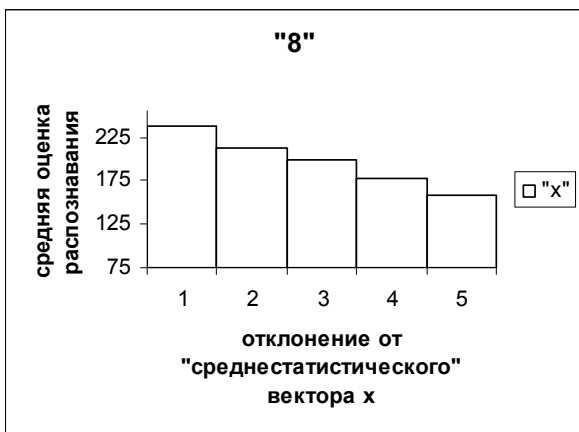


Рис. 10ж

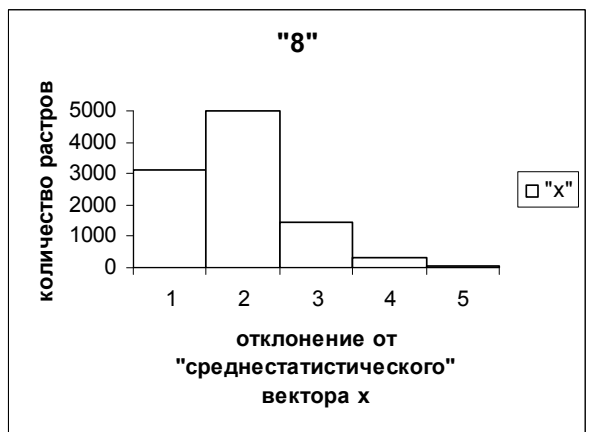


Рис. 10з

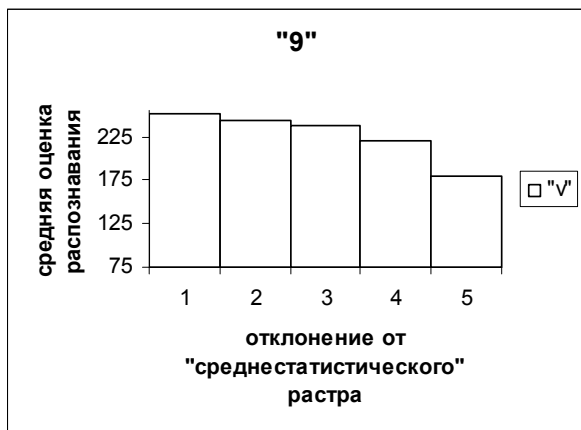


Рис. 11а

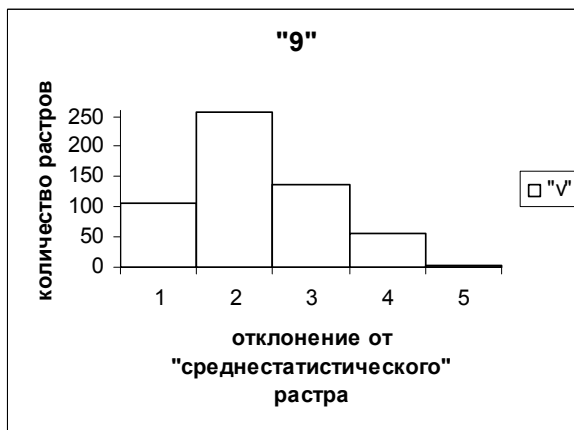


Рис. 11б

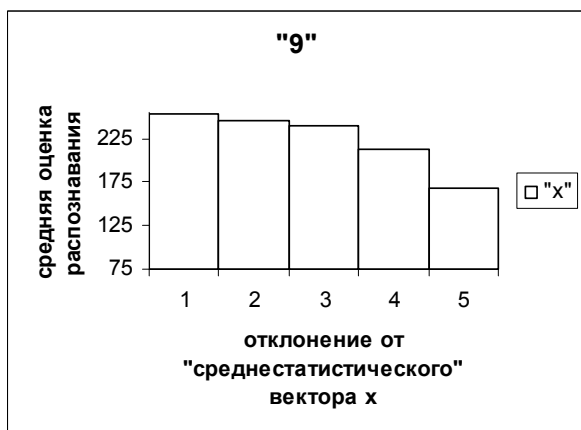


Рис. 11в

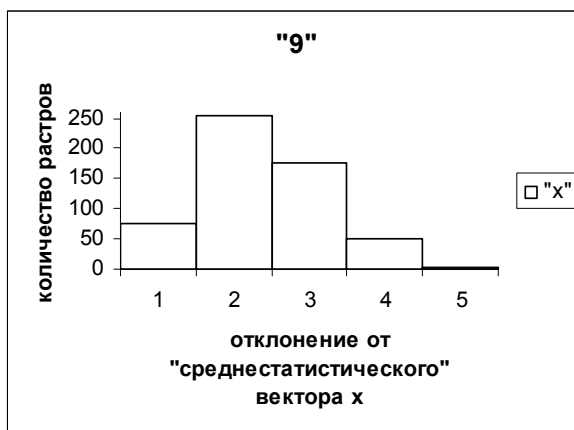


Рис. 11г

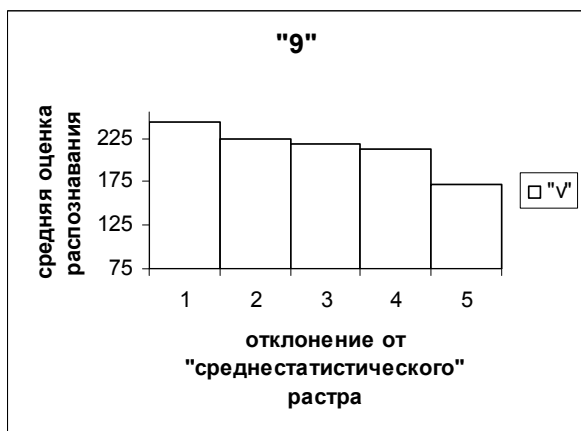


Рис. 11д

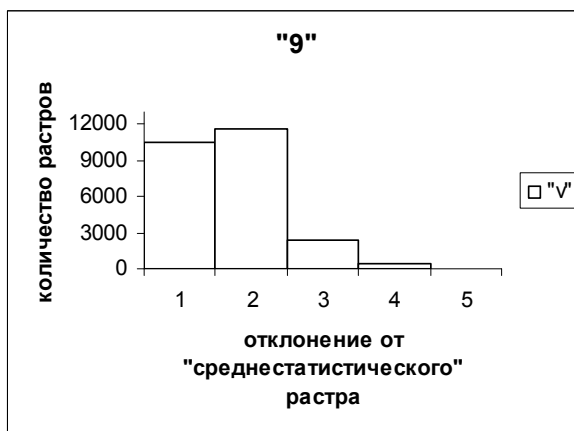


Рис. 11е

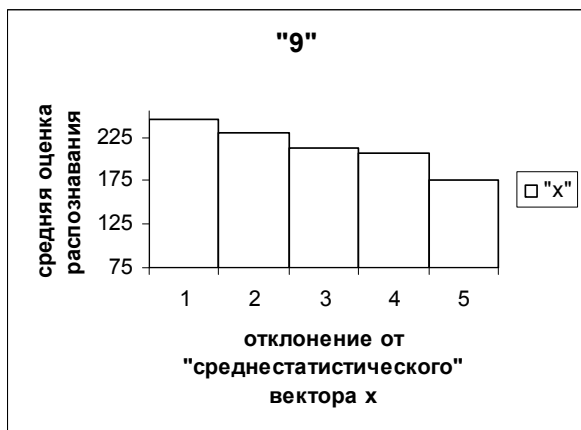


Рис. 11ж

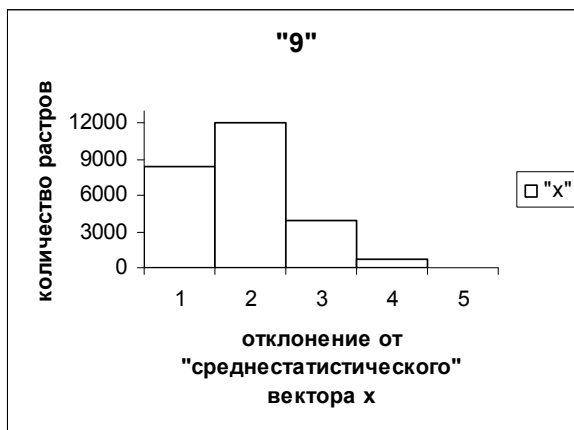


Рис. 11з