

Девятченко Л.Д.

## МНОГОМЕРНЫЙ АНАЛИЗ СООТВЕТСТВИЙ ОРГАНИЗАЦИОННЫХ ПРИЗНАКОВ И РЕЗУЛЬТАТИВНОСТИ ТРУДА В ПРОКАТНОМ ПРОИЗВОДСТВЕ

В данной работе анализировались категоризованные данные готовой продукции широкополосного стана 2500 холодной прокатки ОАО ММК, производимой в условиях сменно-бригадной формы организации труда. Для измерения категорий организационно-производственных признаков использовались номинальные и порядковые шкалы. Статистические данные, представляющие собой частотные таблицы или так называемые таблицы сопряженности [1, 8], предопределили метод их обработки – анализ соответствий. Приложение в металлургии данного метода для обработки данных двухвходовых таблиц рассматривалось ранее [2].

Для многомерного анализа соответствий сбор первичной информации удобно было осуществить в формате бинарной (индикаторной) матрицы, элементами которой являются индикаторы 1 и 0, используемые для обозначения принадлежности или не принадлежности какого-либо объекта (наблюдения) к той или иной категории. В табл. 1 приведен фрагмент бинарной матрицы размера 273×16, предназначенной для изучения взаимосвязи организационных признаков (*C* – смены, *B* – бригады) на результативные признаки производства (*W* – уровни выпуска продукции, *S* – уровни брака, *N* – уровни выпуска несортной продукции). Для исследуемых признаков по результативности труда и смен использовалась трехуровневая градация, а для бригад – четырехуровневая. Общий объем многомерной выборки составил 273 смены (три календарных месяца наблюдений). Обработка данных выполнена в системе STATISTICA [4].

Индикаторная матрица – удобный формат для сбора информации, однако оказывается совершенно непригодной даже для обозрения собранного материала, особенно при большом объеме наблюдений. В многомерном анализе соответствий в качестве входного формата данных, предназначенных для обработки, используют так называемую [4] матрицу Берта (Burt table), которая представляет собой результат произведения исходной бинарной матрицы (обозначим ее  $X$ ) слева на транспонированную матрицу  $X^T$ , т.е.  $X^T X$  является матрицей Берта, основные свойства которой следующие:

- Матрица Берта  $X^T X$  – симметричная матрица порядка  $m$ , где  $X$  – бинарная матрица размера  $n \times m$ ,  $n \gg m$ ,  $n$  – число наблюдений,

$m$  – общее число категорий по всем  $K$  признакам.

- Ячейки матрицы  $X^T X$  содержат  $n_{ij}$  частоты, вычисляемые как скалярное произведение  $i$ -строки матрицы  $X^T$  и  $j$ -столбца матрицы  $X$ .
- Квадратная матрица  $X^T X$  состоит из  $K^2$  блоков (подматриц), образующихся в результате взаимного пересечения исследуемых признаков.
- Блоки, сформированные на главной диагонали матрицы Берта, – диагональные подматрицы со следом, равным  $n$ , где  $n$  – общий объем выборки.

Для многомерного индикаторного массива, фрагмент которого дан в табл. 1, получена матрица Берта размера 16×16 (табл. 2). Массив многовходовых данных, представленный в виде матрицы Берта и предназначенный для многомерного анализа соответствий, предполагает также свободный выбор переменных и для двухвходовых таблиц сопряженности.

Матрица Берта сама по себе даже без какой-либо математической обработки представляет интерес для исследователя. Любая строка (точно как и столбец) этой матрицы дает исследователю предварительную информацию о характере влияния на данную категорию какого-либо фактора каждой категории других факторов, учтенных в данном эксперименте. Как следует из данных табл. 2, напр., для бригады *B1* (4 строка матрицы Берта) за наблюдаемый период ( $n=273$  смены) было 67 выходов на работу, из них 20, 23 и 24 раза

Таблица 1

Фрагмент первичной информации в виде бинарной (индикаторной) матрицы размера 273×16

№ п/п	C			B				W			S			N		
	1	2	3	1	2	3	4	1	2	3	1	2	3	1	2	3
1	1	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0
2	0	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0
3	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	1
...	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
272	0	1	0	0	1	0	0	0	1	0	1	0	0	1	0	0
273	0	0	1	0	0	0	1	1	0	0	0	1	0	0	1	0

Таблица 2

Матрица Берта размера 16×16 для многомерного анализа соответствий точек-столбцов бинарной матрицы размера 273×16

	C1	C2	C3	B1	B2	B3	B4	W1	W2	W3	S1	S2	S3	N1	N2	N3
C1	91	0	0	20	21	26	24	26	30	35	31	24	36	30	27	34
C2	0	91	0	23	22	22	24	49	29	13	26	34	31	30	33	28
C3	0	0	91	24	26	20	21	16	32	43	34	33	24	31	31	29
B1	20	23	24	67	0	0	0	23	30	14	29	14	24	27	21	19
B2	21	22	26	0	69	0	0	19	21	29	28	23	18	22	21	26
B3	26	22	20	0	0	68	0	16	20	32	20	24	24	16	27	25
B4	24	24	21	0	0	0	69	33	20	16	14	30	25	26	22	21
W1	26	49	16	23	19	16	33	91	0	0	28	31	32	32	29	30
W2	30	29	32	30	21	20	20	0	91	0	32	28	31	31	28	32
W3	35	13	43	14	29	32	16	0	0	91	31	32	28	28	34	29
S1	31	26	34	29	28	20	14	28	32	31	91	0	0	33	26	32
S2	24	34	33	14	23	24	30	31	28	32	0	91	0	25	38	28
S3	36	31	24	24	18	24	25	32	31	28	0	0	91	33	27	31
N1	30	30	31	27	22	16	26	32	31	28	33	25	33	91	0	0
N2	27	33	31	21	21	27	22	29	28	34	26	38	27	0	91	0
N3	34	28	29	19	26	25	21	30	32	29	32	28	31	0	0	91

соответственно в 1, 2 и 3 смены. На эту бригаду выпадает в основном средний W2 уровень выпуска продукции (с частотой 30) и у нее чаще всего (29) самый низкий S1 уровень брака и самый низкий уровень N1 несоответствующей заказу продукции (частота 27). Для бригады B2, примерно при тех же условиях выхода на работу, наибольшая частота (29) приходится на высокий уровень W3 выпуска продукции и на низкий S1 уровень брака (частота 29), однако для нее характерен высокий уровень N3 несоответствующей заказу продукции (частота 29) и т.д.

Табл. 2 содержит 20 блоков симметричных относительно главной диагонали матрицы Берта или всего  $(K^2 - K)/2 = 10$  разных двухвходовых подтаблиц (число сочетаний из  $K$  элементов по 2). Каждая из этих двухвходовых таблиц сопряженности может представлять интерес для проверки  $H_0$ -гипотезы о независимости данной пары признаков с имеющимися у них категориями или уровнями. Для этого следует применить методику, рассмотренную ранее [2]. В качестве примера представим один из таких блоков в табл. 3 (соответствующий блок расположен в табл. 2 на пересечении 8–10 строк и 4–7 столбцов). Кстати, по свойствам матрицы Берта маргинальные частоты подтаблиц, определяемых как блок в матрице Берта, в  $i$ -строке и в  $j$ -столбце будут совпадать соответственно с диагональными  $(i, i)$  и  $(j, j)$  элементами матрицы Берта. В табл. 3 приведены частоты сопряженных признаков  $B$  и  $W$  с маргинальными частотами (нижняя строка и правый столбец соответственно).

Для проверки  $H_0$ -гипотезы о независимости многоуровневых признаков  $W$  и  $B$  следует вычислить ожидаемые частоты  $e_{ij} = \frac{n_{i0} \cdot n_{0j}}{n_{00}}$ , а затем статистику  $X^2_k = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$ , где для фактора  $B$  число уровней  $J=4$ , а для фактора  $W$  число уровней  $I=3$ .

По данным наблюдаемых частот  $n_{ij}$ , приведенным в табл. 3, и ожидаемых частот  $e_{ij}$ , вычисленных при независимости признаков  $W$  и  $B$ , эмпирический критерий  $X^2=21,1832$  с числом степеней свободы  $k=6$ , определяемых как  $k=(I-1)(J-1)$ .

Принимая во внимание, что на уровне значимости  $\alpha=0,05$  табличное [3] значение  $\chi^2_{6;0,05} = 12,59 < X^2 = 21,183$   $H_0$ -гипотезу о независимости признаков  $W$  и  $B$  отвергаем и принимаем альтернативную гипотезу  $H_1$  о значимой зависимости наблюдаемых признаков. Декомпозиция данного  $X^2$ -критерия представлена в табл. 4.

Данные табл. 4 показывают, что наибольший локальный вклад в отклонение от независимости, наряду с ячейкой  $(W_3, B_3)$ , дает ячейка  $(W_1, B_4)$ . Если теперь дополнить выявленные категории  $W_1$  и  $B_4$  противоположными категориями  $\bar{W}_1$  и  $\bar{B}_4$

Таблица 3 Категорированные данные по выпуску продукции и бригадной формы организации труда

Обозначенные уровни выпуска продукции	Б Р И Г А Д Ы				Сумма
	$B_1$	$B_2$	$B_3$	$B_4$	
$W_1$ – нижний	23	19	16	33	91
$W_2$ – средний	30	21	20	20	91
$W_3$ – высший	14	29	32	16	91
Сумма	67	69	68	69	273

Таблица 4 Декомпозиция  $X^2$ -критерия, вычисленного по данным табл. 3

Обозначенные уровни выпуска продукции	Б Р И Г А Д Ы				Сумма
	$B_1$	$B_2$	$B_3$	$B_4$	
$W_1$ – нижний	0,0199	0,6956	1,9607	4,3478	7,0241
$W_2$ – средний	2,6318	0,1739	0,3137	0,3913	3,5107
$W_3$ – высший	3,1094	1,5652	3,8431	2,1304	10,648
Сумма	5,7611	2,4347	6,1176	6,8695	21,183

$\bar{B}_4$ , то можем построить и анализировать полученную при этом простейшую дихотомическую таблицу  $2 \times 2$ , представленную в **табл. 5** вместе с результатами проверки  $H_0$ -гипотезы.

Таким образом, отклонение  $H_0$ -гипотезы подтверждает локальный вывод о существенном влиянии бригады  $B_4$  на низкий уровень  $W_1$  выпуска металлопродукции. Аналогично можно подтвердить и значимость влияния  $B_3$  на  $W_3$ .

Из 10 блоков матрицы Берта значительный интерес представляет также блок, связывающий уровни брака и последовательности смен по времени работы (ночная, дневная, вечерняя). Напр., нужно проверить  $H_0$ -гипотезу о незначимом влиянии работы в ночную смену ( $C1$ ) на высокий уровень брака ( $S3$ ).

Составим дихотомическую матрицу  $2 \times 2$ , общий вид которой для наблюдаемых и ожидаемых частот приведен в **табл. 6**.

*Комментарий к составлению табл. 6.* Из объема выборки  $n=273$  одинаковое число, равное 91, приходится на каждую смену, в том числе и на смену  $C1$ . В смену  $C1$  высокий уровень брака отмечался 36 раз. Значит, в смену  $C1$  не было высокого уровня брака  $91-36=55$  раз – это для категории  $\bar{S}3$ , а также, если по всем сменам 91 раз наблюдался высокий уровень брака  $S3$ , то не в смену  $C1$ , т.е. для категории  $\bar{C}1$  приходится  $91-36=55$  случаев  $S3$ . Число случаев  $\bar{S}3$ , т.е. когда высокий уровень брака не отмечался при работе двух других смен  $\bar{C}1$ , составляет  $273-91-55=127$ .

*Примечание.* Приведенные вычисления в компьютерной программе STATISTICA [4] выполняются автоматически путем ввода в бинарную матрицу для каждого индикаторного столбца дополнительного столбца с противоположной индикацией. В результате матрица Берта в своих размерах увеличивается вдвое, т.к. каждая категория в такой расширенной матрице как бы представляется в виде дихотомического фактора (свойство присутствует и свойство отсутствует).

Итак, возвращаясь к **табл. 6**, отметим, что расчетное значение  $X^2$ -статистики оказалось равным 2,303, что не позволяет отклонить  $H_0$ -гипотезу, т.к. табличное значение  $\chi^2$ -критерия с одной степенью свободы на уровне значимости  $\alpha=0,05$

Таблица 5

Двойная дихотомизация данных из табл. 3 для изоляции ячейки ( $W_1, B_4$ )

Категории	$\bar{B}_4$	$B_4$	Сумма	Комментарии
$W_1$	58	33	91	$X_1^2 = 8,727$ ; $p = 0,003$ ; $H_0$ отклонена
$\bar{W}_1$	146	36	182	
Сумма	204	69	273	

Таблица 6

Наблюдаемые и ожидаемые частоты таблицы сопряженности  $2 \times 2$ , составленной по данным матрицы Берта

Категории	Наблюдаемые частоты			Ожидаемые частоты		
	$S3$	$\bar{S}3$	Итог	$S3$	$\bar{S}3$	Итог
$C1$	36	55	91	30,004	60,666	91,000
$\bar{C}1$	55	127	182	60,666	121,334	182,000
Итог	91	182	273	91,000	182,000	273,000

Таблица 7

Двухходовая таблица сопряженности признаков В – бригады и S – уровни брака

Категории	S1	S2	S3	Итоги
B1	29	14	24	67
B2	28	23	18	69
B3	20	24	24	68
B4	14	30	25	69
Итоги	91	91	91	273

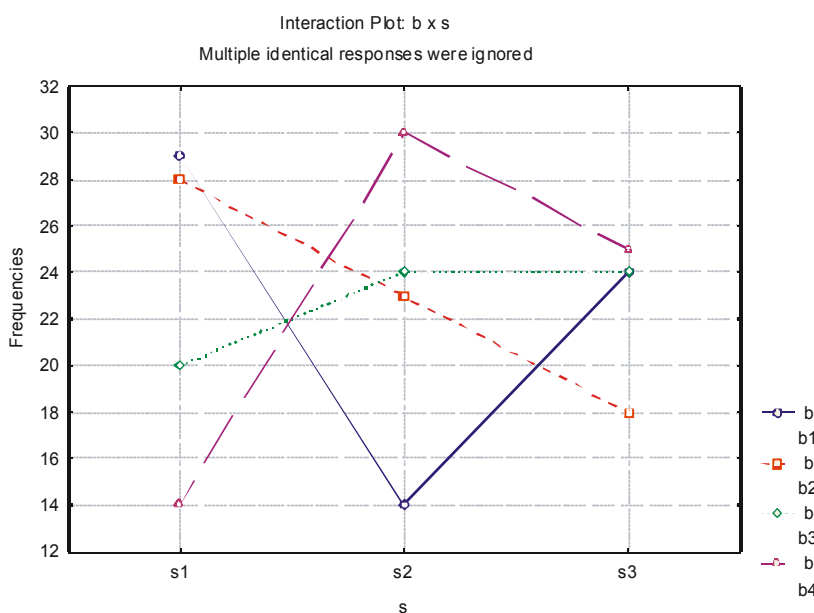


Рис. 1. Графики взаимосвязи сопряженных признаков в табл. 7

Таблица 8

Наблюдаемые частоты в таблице с тремя входами: В – бригады, W – уровни производства, С – рабочие смены

C	B	W			Всего по строкам
		W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	
C <sub>1</sub>	B <sub>1</sub>	7	9	4	20
	B <sub>2</sub>	7	3	11	21
	B <sub>3</sub>	3	8	15	26
	B <sub>4</sub>	9	10	5	24
Всего по столбцам		26	30	35	91
C <sub>2</sub>	B <sub>1</sub>	12	8	3	23
	B <sub>2</sub>	11	10	1	22
	B <sub>3</sub>	11	4	7	22
	B <sub>4</sub>	15	7	2	24
Всего по столбцам		49	29	13	91
C <sub>3</sub>	B <sub>1</sub>	4	13	7	24
	B <sub>2</sub>	1	8	17	26
	B <sub>3</sub>	2	8	10	20
	B <sub>4</sub>	9	3	9	21
Всего по столбцам		16	32	43	91

оказалось  $\chi^2_{1; 0,05} = 3,84 > 2,303$ .

Характер влияния отдельных бригад на уровни брака проявляется из данных подтаблицы (табл. 7), отвечающей блоку  $B \times S$  матрицы Берта.

По частотным данным табл. 7 построены соответствующие графики (рис. 1). Наиболее очевидный вывод, который следует из визуального анализа графиков на рис. 1: точка s2 для всех бригад, кроме b2, является как бы экстремальной. В этой точке для бригад b3 и b4 средний уровень брака достигает «Мах», а для бригады b1 – «Мин», при этом только в бригаде b2 с повышением уровня брака от s1 к s3 равномерно снижается частота его появления.

Теперь перейдем от частных задач, вытекающих из двухвходовых подматриц матрицы Берта, к задачам многомерного анализа соответствий, для которого упомянутая матрица служит входным форматом при обработке данных на ЭВМ.

Если число категорированных переменных  $M \geq 3$ , то сразу же возникают сложности в наглядности таких таблиц, т.к. таблица с размерностью  $I \times J \times \dots \times K$ , где  $I, J, \dots, K$  – числа категорий для соответствующих  $A_1, A_2, \dots, A_M$  факторов, входящих в многомерную таблицу

сопряженности, физически может быть представлена только в  $M$ -мерном пространстве. В простейшем случае, когда  $M=3$ , таблицу можно представить в виде прямоугольного параллелепипеда со строками, столбцами и слоями, где слои – это двухвходовые таблицы сопряженности. Воспользуемся такой формой представления частот (табл. 8) для анализа влияния сменно-бригадной формы организации труда на уровни выпуска продукции.

По данным табл. 8 для проверки  $H_0$ -гипотезы вычисляются ожидаемые частоты (табл. 9), а затем  $\chi^2$ -статистика, распределенная по ячейкам таблицы  $4 \times 3 \times 3$  (табл. 10).

Бесспорно, из анализа данных табл. 8–10 следует, что  $H_0$ -гипотеза на уровне значимости  $\alpha=0,05$  отвергается, т.к. при обозначенном числе степеней свободы  $df=4 \cdot 3 \cdot 3 - 1 = 35$  табличное [3] значение  $\chi^2_{35; 0,05} = 49,80185 < 77,27193$ .

Матрица Берта содержит пересечения каждой категории с любой другой категорией, в том числе и с их отрицанием (в случае расширения матрицы за счет дополняющих столбцов). Следовательно, если использовать метод декомпозиции  $\chi^2$ -критерия, то в каждой ячейке матрицы  $(i, j)$ -компонента  $\chi^2$ -критерия будет указывать на роль данного пересечения в формировании оценок общей сопряженности исследуемых признаков (их категорией).

В рассмотренном случае для  $M=3$  признаков, когда еще не теряется наглядность за счет послышного представления двухвходовых таблиц, стано-

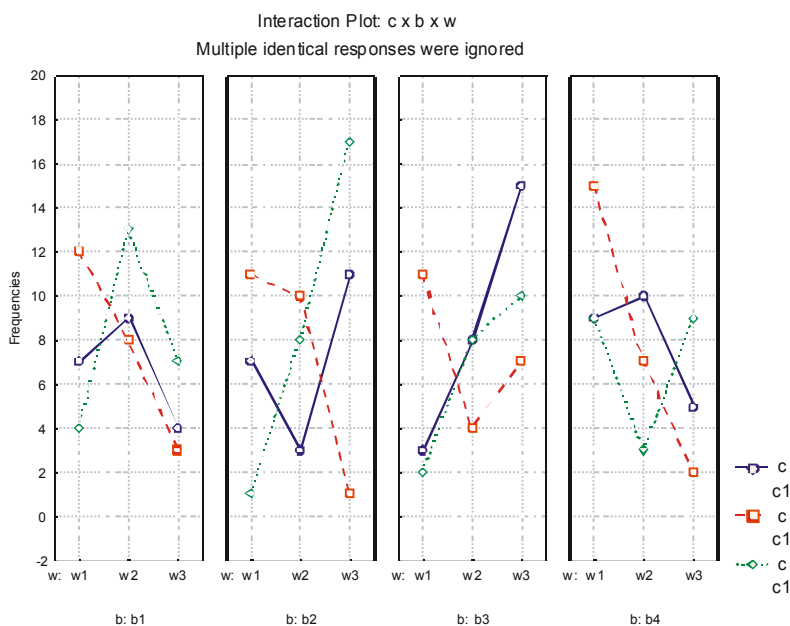


Рис. 2. Диаграммы для визуального анализа данных с тремя входами (факторы В, С и W)



вится очевидным, какие именно пересекаемые категории обеспечивают связь наблюдаемых признаков. Так, напр., данные **табл. 10** показывают, что наибольший вклад в  $\chi^2$ -статистику обеспечивает ячейка  $(C_3, B_2, W_3)$ , т.е. связь третьей смены, второй бригады и максимального уровня выпуска продукции существенно значима, значительный вклад в значимость выявляемых взаимосвязей дают ячейки  $(C_1, B_3, W_3)$ ,  $(C_2, B_4, W_1)$  и др.

Какой вид имеют сформировавшиеся взаимосвязи категорированных признаков, представленных в **табл. 8**, можно наблюдать на диаграмме **рис. 2**.

Оценка связи категорированных факторов, как отмечалось выше, является первейшей задачей анализа как двухвходовых, так и многовходовых таблиц сопряженности. Отметим, что мера связи, аналогичная парному коэффициенту корреляции Пирсона, определяющая положительное или отрицательное направление этой связи, не имеет смысла для факторов, измеряемых в номинальных шкалах, но не в порядковых шкалах. В случае дихотомических факторов (также и категорий, отражающих отсутствие и присутствие конкретного свойства) мера связи, используемая в порядковой (ранговой) статистике Кендалла, Спирмена, а также Гудмена–Краскала, становится вполне приемлемой [8].

Заметим, что изучаемые категорированные факторы (см. **табл. 8** и **рис. 2**) измерены в разных шкалах: фактор В – в номинальной шкале, а фак-

торы С и W – в порядковых шкалах. Следовательно, для категорий факторов С и W можно получить оценки корреляционных связей, используя подпрограммы [4], где реализованы соответствующие алгоритмы [7, 8]. В **табл. 11** приведена ранговая корреляция Спирмена (Spearman) для блока дихотомических переменных (категорий порядковых факторов С и W), а также статистические характеристики, в частности  $t(N-2)$ -статистика Стьюдента и **p-level** – уровень значимости для проверки  $H_0$ -гипотезы о независимости той или иной пары сопряженных признаков.

Как показывают данные **табл. 11**, статистически значимыми оказываются следующие пары категорий **c2 & w1**, **c2 & w3**, **c3 & w1** и **c3 & w3**, т.к. по умалчиванию значимой признается лишь  $t$ -статистика Стьюдента с  $N-2$  степенями свободы, если **p-level < 0,05**. При этом обратим внимание, что две пары дихотомических переменных (**c2 & w1**, **c3 & w3**) имеет значимую положительную корреляцию, а две других (**c2 & w3**, **c3 & w1**) – отрицательную, т.е. смена **c2** является как бы «переломной».

Если принять во внимание 3-уровневую категоризацию факторов С и W, то гипотеза о существенной значимости рассматриваемых факторов подтверждается на основе  $\chi^2$ -статистики Пирсона, т.к. эмпирическое значение  $\chi^2_4 = 34,9451 > \chi^2_{4; 0,05} = 9,4877$  при числе степеней свободы  $(3-1)(3-1)=4$  и уровне значимости

Таблица 9

Ожидаемые частоты в таблице с тремя входами:  
В – бригады, W – уровни производства,  
С – рабочие смены

C	B	W			Всего по строкам
		W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	
C <sub>1</sub>	B <sub>1</sub>	7,44	7,44	7,44	22,33
	B <sub>2</sub>	7,67	7,67	7,67	23,01
	B <sub>3</sub>	7,55	7,55	7,55	22,65
	B <sub>4</sub>	7,67	7,67	7,67	23,01
Всего по столбцам		30,33	30,33	30,33	91
C <sub>2</sub>	B <sub>1</sub>	7,44	7,44	7,44	22,33
	B <sub>2</sub>	7,67	7,67	7,67	23,01
	B <sub>3</sub>	7,55	7,55	7,55	22,65
	B <sub>4</sub>	7,67	7,67	7,67	23,01
Всего по столбцам		30,33	30,33	30,33	91
C <sub>3</sub>	B <sub>1</sub>	7,44	7,44	7,44	22,33
	B <sub>2</sub>	7,67	7,67	7,67	23,01
	B <sub>3</sub>	7,55	7,55	7,55	22,65
	B <sub>4</sub>	7,67	7,67	7,67	23,01
Всего по столбцам		30,33	30,33	30,33	91

Таблица 10

Распределение  $\chi^2$ -статистики  
по ячейкам таблицы 4x3x3,  
(использованы данные табл. 7 и 8)

C	B	W			Всего по строкам
		W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	
C <sub>1</sub>	B <sub>1</sub>	0,026021	0,327097	1,590538	1,943656
	B <sub>2</sub>	0,058527	2,843403	1,445750	4,347680
	B <sub>3</sub>	2,742053	0,026821	7,351324	10,12020
	B <sub>4</sub>	0,230626	0,707809	0,929452	1,867887
Всего по столбцам		3,057227	3,905131	11,31706	18,27942
C <sub>2</sub>	B <sub>1</sub>	2,794839	0,042150	2,649677	5,486666
	B <sub>2</sub>	1,445750	0,707809	5,800378	7,95394
	B <sub>3</sub>	1,445750	1,669205	0,040066	3,155021
	B <sub>4</sub>	7,005072	0,058527	4,191512	11,25511
Всего по столбцам		12,69141	2,477691	12,68163	27,85073
C <sub>3</sub>	B <sub>1</sub>	1,590538	4,155054	0,026021	5,771613
	B <sub>2</sub>	5,800378	0,014198	11,34927	17,16385
	B <sub>3</sub>	4,079801	0,026821	0,795033	4,901655
	B <sub>4</sub>	0,230626	2,843403	0,230626	3,304655
Всего по столбцам		11,70134	7,039475	12,40095	31,14177

$\alpha=0,05$ . Однако ранговый коэффициент корреляции  $\tau_K$  Кендалла [8], вычисленный для 3-уровневых порядковых факторов С и W, оказался, как и следовало ожидать статистически незначимым ( $\tau_K=0,08839$ ) из-за «переломного» характера категории с2 (дневная смена). Данный факт подтвержден опять же с помощью рангового коэф-

Таблица 11

**Ранговая корреляция Спирмена для исследуемого блока дихотомических переменных**

	Valid	Spearman	t(N-2)	p-level
c1 & w1	273	-0,071429	-1,17887	0,239482
c1 & w2	273	-0,005495	-0,09045	0,927995
c1 & w3	273	0,076923	1,27008	0,205147
c2 & w1	273	0,307692	5,32352	0,000000
c2 & w2	273	-0,021978	-0,36189	0,717715
c2 & w3	273	-0,285714	-4,90804	0,000002
c3 & w1	273	-0,236264	-4,00271	0,000081
c3 & w2	273	0,027473	0,45243	0,651324
c3 & w3	273	0,208791	3,51460	0,000516

Таблица 12

**Раздельное рассмотрение таблиц сопряженности уровней выпуска продукции W<sub>1</sub>, W<sub>2</sub>, W<sub>3</sub> для смен С<sub>1</sub>, С<sub>2</sub> и С<sub>3</sub>**

Отрицательная связь рангов					Положительная связь рангов				
Уровни	W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	Сумма	Уровни	W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	Сумма
С <sub>1</sub>	26	30	35	91	С <sub>2</sub>	49	29	13	91
С <sub>2</sub>	49	29	13	91	С <sub>3</sub>	16	32	43	91
Сумма	75	59	48	182	Сумма	65	61	56	182

Таблица 13

**Основные числовые характеристики, вычисленные по данным табл. 2, для многомерного анализа соответствий**

Номер размерности	Сингулярные числа $\sqrt{\lambda_i}$	Собственные числа $\lambda_i$	Процент от общей инерции $I_0 = 2,2$	Кумулятивный процент инерции $\sum I_{0i}$	$\chi^2$ -статистика
1	0,542768	0,294597	13,39078	13,3908	426,5197
2	0,512295	0,262446	11,92937	25,3201	379,9712
3	0,484770	0,235002	10,68192	36,0021	340,2378
4	0,463264	0,214613	9,75514	45,7572	310,7183
5	0,461318	0,212814	9,67338	55,4306	308,1140
6	0,444334	0,197433	8,97422	64,4048	285,8445
7	0,421788	0,177905	8,08660	72,4914	257,5725
8	0,417944	0,174677	7,93988	80,4313	252,8992
9	0,410215	0,168277	7,64894	88,0802	243,6322
10	0,385451	0,148572	6,75328	94,8335	215,1939
11	0,337139	0,113663	5,16649	100,0000	164,5618

фициента корреляции  $\tau_K$  Кендалла при соответствующей группировке категорий фактора С, представленной в табл. 12.

Коэффициент ранговой корреляции  $\tau_K$  Кендалла, вычисленный по данным левой части табл. 12, оказался отрицательным,  $\tau_K=-0,28665$ , а по данным правой части – положительным,  $\tau_K=0,40133$ . Асимптотическая стандартная ошибка, одинаковая для представленных в табл. 12 обоих сопряжений, равна 0,05. Следовательно, выдвинутые ранее предположения относительно наличия линейной корреляции с противоположными алгебраическими знаками подтверждены.

Таким образом, с помощью ранговой корреляции  $\tau_K$  Кендалла и  $r_S$  Спирмена получена оценка не только меры связи сопряженных признаков и их категорий, но и на направление действия этой связи, подобно коэффициенту корреляции Пирсона.

Следуя логике анализа двухвыходовых таблиц сопряженности [1, 2], далее следует определить собственные числа и соответствующие им собственные векторы, чтобы затем отобразить исходную информацию, содержащуюся в матрице Берта, в пространство выбранной размерности [5, 9]. В табл. 13 приведены вычисленные по данным табл. 2 основные числовые характеристики для выбора и построения удобной координатной системы выполняемых многомерных преобразований.

Использование матрицы Берта в многомерном анализе соответствий не позволяет исследователю выбрать максимальную размерность для точного воспроизведения заданного M-мерного пространства для отображения его в преобразованном ортогональном пространстве по следующим причинам.

Во-первых, не следует забывать, что основной целью анализа соответствий, тем более многомерного, является представление расстояний между точками-столбцами, изучаемых M категорий по всем K факторам в пространстве более низкой размерности.

Поэтому мы не стремимся к увеличению размерности отображаемого пространства даже в ущерб информационной ценности получаемого отображения.

Во-вторых, напомним, что собственные числа  $\lambda_i$   $i = 1, \dots, m$ ,  $m < M$ , располагаемые в порядке убывания  $\lambda_1 = \lambda_{\max}$ ,  $\lambda_2, \lambda_3, \dots, \lambda_m = \lambda_{\min}$ , последовательно характеризуют долю  $\lambda_i / \sum_{i=1}^{m < M} \lambda_i$  информационной ценности i-размерности отображения. Поэтому мы стремимся использовать для построения отображений только первые  $\lambda_i$ , более близкие к  $\lambda_{\max}$ .

В-третьих, в выбранном пространстве более низкой размерности в основном, пусть даже с искажением, будут сохраняться пропорции расстояний между точками-столбцами. Это позволяет в некоторой степени все же «распознать» сформированные *сгустки отдельных категорий*, пусть даже немного «размытыми» на границах их соприкосновений, в зависимости от выбираемых *i*-размерностей, используемых в ортогональных преобразованиях [6, 9].

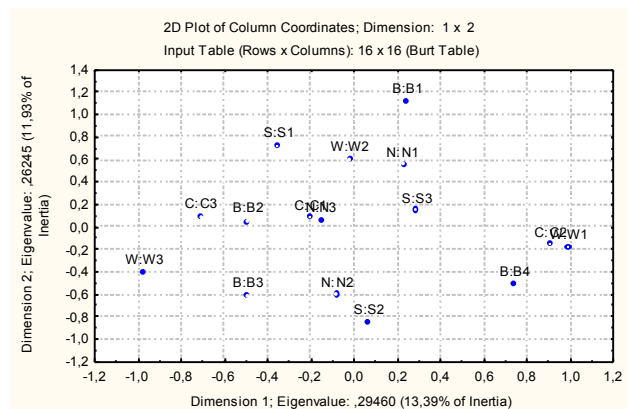
Таким образом, размерность выбираемого пространства для отображения информации, содержащейся в матрице Берта, как правило, равна двум (реже трем) для удобства *наглядного представления* информации по сопряженности изучаемых факторов.

На **рис. 3** представлено отображение информации, содержащейся в матрице Берта (см. **табл. 2**), в двумерное пространство с координатными осями, определяемыми на основе  $\lambda_1$  и  $\lambda_2$ , в системе аккумулируется 25,32% инерции. При этом на **рис. 3** просматриваются немного «размытыми» сформированные 3 сгустка точек. Возможно, что эти сгустки инициируются на фоне рабочих смен, однако смены хорошо разделены только для 1-размерности, для 2-размерности они практически находятся на одном уровне. Скорее всего, формирование групп инициируется определенным уровнем выпуска продукции, а также не последнюю роль в этом играют производственные бригады.

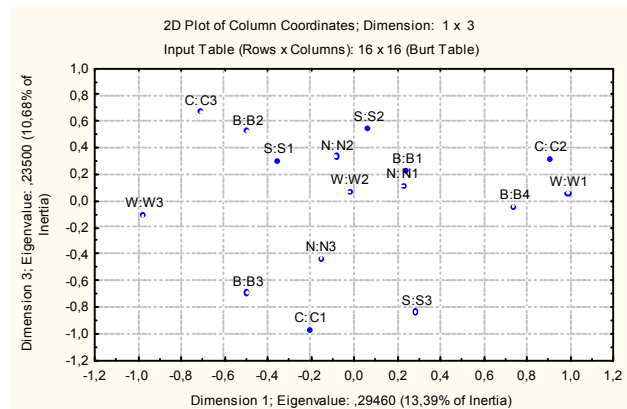
Примем во внимание слабую крутизну убывания собственных чисел матрицы Берта, отражающей специфику взаимосвязи данных по организации производства и показателям выпуска продукции, и представим исходную информацию опять же в двумерных ортогональных пространствах, но в других ракурсах. При этом будем использовать другие варианты сочетания собственных чисел. На **рис. 4** представлено отображение, основанное на  $\lambda_1$  и  $\lambda_3$ , (аккумулятивное 24,07% инерции), а на **рис. 5** – отображение, основанное на  $\lambda_2$  и  $\lambda_3$  (аккумулятивное 22,61% инерции).

Сравнивая во всех представленных на **рис. 3–5** ракурсах полученную информацию о «близости» друг к другу некоторых категорий, заметим, что формирование сгустков из отдельных категорий в основном носит устойчивый характер, однако некоторые точки (см. **рис. 3–5**) все же «перескакивают» из одного сгустка в другой в зависимости от используемого преобразования. Во всех трех ракурсах двумерных пространств очень близкими всегда оказываются группы точек: (C2,B4,W1); (C3,B2); (C1,S3). К этим группам примыкают и другие точки, однако констатировать это по отдельным двумерным отображениям затруднительно.

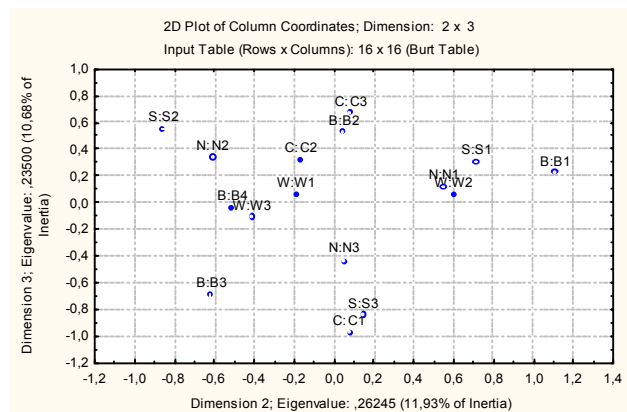
На **рис. 6** представлено трехмерное пространство, в основе построения которого использованы максимально возможные в данном случае собственные числа  $\lambda_1$ ,  $\lambda_2$  и  $\lambda_3$ . Система, построенная на их основе, аккумулирует более 36% инерции.



**Рис. 3.** Отображение исходной информации в двумерное пространство, определяемое на основе  $\lambda_1$  и  $\lambda_2$



**Рис. 4.** Отображение исходной информации в двумерное пространство, определяемое на основе  $\lambda_1$  и  $\lambda_3$



**Рис. 5.** Отображение исходной информации в двумерное пространство, определяемое на основе  $\lambda_2$  и  $\lambda_3$

3D Plot of Column Coordinates; Dimension: 1 x 2 x 3  
 Input Table (Rows x Columns): 16 x 16 (Burt Table)

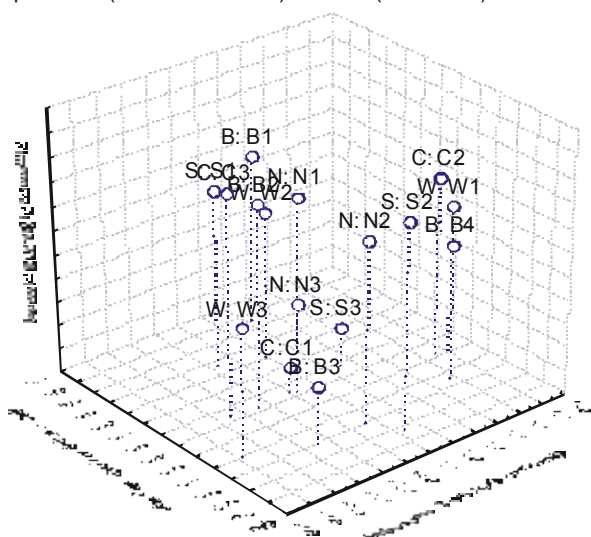


Рис. 6. Отображение исходной информации в трехмерное пространство, определяемое на основе  $\lambda_1, \lambda_2$  и  $\lambda_3$

Таблица 14

Оценка состава групп, объединяющих близкие по свойствам категории организационно-производственных признаков

Группа	Вариант формирования	Размерность	Выявляемый состав группы				
			C1,C3	B2,B3	W3		N3
Первая	$\lambda_1, \lambda_2$	1x2	C1,C3	B2,B3	W3		N3
	$\lambda_1, \lambda_3$	1x3	C1	B3	W3	S3	N3
	$\lambda_2, \lambda_3$	2x3	C1	B3		S3	N3
	$\lambda_1, \lambda_2, \lambda_3$	1x2x3	C1	B3	W3	S3	N3
Вторая	$\lambda_1, \lambda_2$	1x2	C2	B4	W1	S2	N2
	$\lambda_1, \lambda_3$	1x3	C2	B4	W1		
	$\lambda_2, \lambda_3$	2x3	C2	B4	W1,W3	S2	N2
	$\lambda_1, \lambda_2, \lambda_3$	1x2x3	C2	B4	W1	S2	N2
Третья	$\lambda_1, \lambda_2$	1x2		B1	W2	S1,S3	N1
	$\lambda_1, \lambda_3$	1x3	C3	B1,B2	W2	S1,S2	N1,2
	$\lambda_2, \lambda_3$	2x3	C3	B1,B2	W2	S1	N1
	$\lambda_1, \lambda_2, \lambda_3$	1x2x3	C3	B1,B2	W2	S1	N1

На рис. 6 уже более четко разделяются обозначенные ранее сгустки точек, которые характеризуют сходство отдельных категорий (проявляется общность интересов, коллективная ответственность, обобщенный уровень качества работы, потенциальные, моральные, интеллектуальные возможности и т.п.), объединяемых в однородные по свойствам не пересекаемые классы. Обобщение результатов по анализу соответствий и выявлению принадлежности категорированных

данных к определенным однородным по свойствам группам сведены в табл. 14.

Резюме по многомерному анализу соответствий. Сгустки точек, самопроизвольно объединяемых в непересекаемые классы, позволяют констатировать их схожесть, однородность присущих им свойств. В этом случае поиск меры связи уже тесно связанных категорий становится излишним.

Результаты, полученные с помощью многомерного анализа соответствий (см. рис. 6 и табл. 14), выявляют следующее:

- В условиях сменно-бригадной формы организации труда сформировались практически устойчивые 3 группы, связывающие категории признаков результативности труда (характеристики количества и качества выпускаемой продукции):
  - В смену С1 (ночную) превалирует максимальный уровень выпуска W3 продукции с максимальным уровнем некондиционной N3 и в большинстве случаев с максимальным уровнем брака S3, что наблюдается в основном в бригаде В3 (иногда в В2).
  - В смену С2 (дневную) превалирует минимальный W1 уровень выпуска продукции (изредка W3) и в основном со средними уровнями брака S2 и некондиционной продукции N2, что наблюдается в бригаде В4 (и очень редко в В1 и В2).
  - В смену С3 в основном наблюдается только средний уровень выпуска продукции W2 и низкие уровни брака S1 (изредка S2 и S3) и некондиционной продукции N1 (изредка N2), что сопутствует только бригадам В1 и чуть реже В2.
- Обнаружены значимые изменения в структуре качества труда бригад при переходе с ночной смены через дневную на вечернюю смену работы. В частности, наблюдается «перелом» (снижение – рост) по уровням выпуска продукции, при этом выпуск некондиционной и отбракованной продукции монотонно снижается.
- Для визуального анализа результатов – достигаемых уровней выпуска продукции (W1, W2, W3), а также характеристик качества выпускаемой продукции (некондиционной N1, N2, N3 и брака S1, S2, S3) по бригадам (B1, B2, B3, B4) в разные смены (C1, C2, C3) целесообразно использовать координатное пространство меньшей размерности, чем наблюдаемое многомерное пространство. При этом для уточнения выявляемых связей нелишне будет представление отображенной инфор-



мации в разных ракурсах, особенно в случаях, когда собственные числа матрицы связей изменяются практически монотонно.

- Для наиболее полного анализа информации, сосредоточенной в матрице Берта, наряду с многомерным анализом соответствий следует выполнять анализ двухвходовых таблиц сопряженности, а также простейших дихотомических

таблиц, особенно при уточнении направления связи при изучении порядковых переменных.

- Многомерный анализ соответствий может служить полезным инструментом для классификации номинальных и порядковых категорий наблюдаемых признаков.

#### Библиографический список

1. Аптон Г. Анализ таблиц сопряженности / Пер. с англ. и предисл. Ю.П. Адлера. М.: Финансы и статистика, 1982. 143 с.
2. Анализ сопряженности технологических нарушений и состояний аварийности на машинах непрерывной разливки стали / Л.Д. Девятченко, Л.Ш. Тимиргалеева // Вестник МГТУ им. Г.И. Носова. 2006. № 3. С. 15–23.
3. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. М.: Наука, 1983. 416 с.
4. Боровиков В. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. (+CD). СПб.: Питер, 2003. 268 с.
5. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. М.: Наука, 1984. 320 с.
6. Девятченко Л.Д. Признаки классификации. Введение в дискриминантный анализ: Учеб. пособие. Магнитогорск: МГТУ, 2001. 127 с.
7. Енюков И.С. Методы, алгоритмы, программы многомерного статистического анализа: Пакет ППСА. М.: Финансы и статистика, 1986. 232 с.
8. Кендалл М., Стьюарт А. Статистические выводы и связи. М.: Наука, 1976. 736 с.
9. Прикладная статистика. Классификация и снижение размерности / Айвазян С.А., Бухштабер В.М., Енюков И.С., Мещалкин Л.Д. М.: Финансы и статистика, 1989. 606 с.

УДК 658.562:621.771

Румянцев М.И., Шубин И.Г., Завалишин А.Н., Корнилов В.Л., Буданов А.П., Цепкин А.С., Пантелеева Н.А.

## ОЦЕНИВАНИЕ КАЧЕСТВА ТОНКИХ ГОРЯЧЕКАТАНЫХ ПОЛОС ДЛЯ ОПРЕДЕЛЕНИЯ ВОЗМОЖНОСТИ ЗАМЕЩЕНИЯ ХОЛОДНОКАТАНОЙ ЛИСТОВОЙ СТАЛИ ОБЩЕГО НАЗНАЧЕНИЯ\*

В настоящее время в мировой практике стала реальностью и продолжает усиливаться тенденция к производству тонких горячекатаных полос, часть из которых является самостоятельной товарной продукцией, а часть используется для замещения холоднокатаного металла. В типичном сортаменте холоднокатаной конструкционной стали более 30% составляют полосы толщиной 1,2–2,5 мм и выше (рис. 1, б). При этом в сортаменте широкополосных станов горячей прокатки (ШСП) доля полос толщиной 1,2–3,0 мм составляет около 58% (рис. 1, а). В качестве предпосылок указанной тенденции можно отметить повышение оперативности выполнения заказов и снижение издержек производства [1–3].

По нашему мнению [4], наиболее перспективным является применение тонких горячекатаных полос для замещения холоднокатаного листового

проката общего назначения. Анализ портфеля заказов ЛПЦ-5 ОАО «ММК» показал, что около 20% составляют холоднокатаные полосы толщиной 1,2–3,0 мм, поставляемые по ГОСТ 16523 с качеством отделки поверхности групп 2 и 3. При этом доля полос толщиной 1,5–1,8 мм составляет около 35% от этого количества (или около 7% от годового объема производства цеха), а полос толщиной 2,0–3,0 мм – примерно 50% (около 9% от годового производства) (рис. 2, а). До 80% металла в рассматриваемой сортаментной группе поставляется со свойствами категории вытяжки Г и около 16% – со свойствами категории вытяжки Н (рис. 2, б).

Для повышения объективности и достоверности оценок соответствия качества горячекатаных полос требованиям к холоднокатаному прокату разработали дифференциальные оценки единичных показателей качества, унифицированные с широко используемыми в системах менеджмента качества индексами и процесса [5]:

\* В работе принимали участие студентки Оплачко Т.В., Питренина М.А.