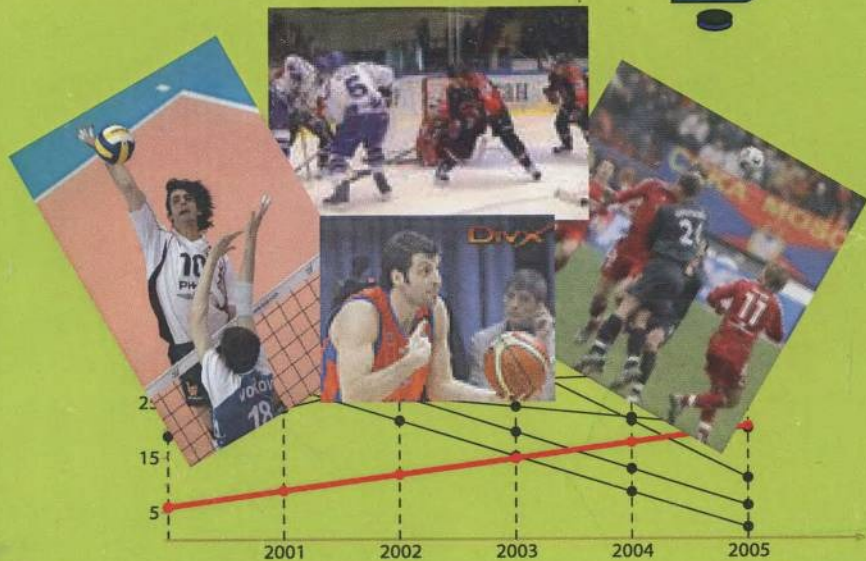
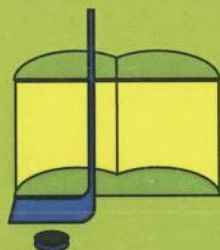




В.В. Афанасьев  
И.Н. Непряев

# Математическая статистика в командных видах спорта



Министерство образования и науки Российской Федерации  
Ярославский государственный педагогический университет  
имени К.Д. Ушинского

*В. В. Афанасьев, И. Н. Непряев*

**Математическая  
статистика  
В  
командных  
видах спорта**

Ярославль  
2006

ББК 75 в 631.8 + 22.172

УДК 519.22; 796:311

А 941

Афанасьев В. В., Непряев И. Н. Математическая статистика в командных видах спорта. Ярославль: Издательство ЯГПУ им. К.Д. Ушинского, 2006. 120с.

Книга предназначена для студентов и преподавателей факультетов физической культуры, для всех любителей, функционеров и организаторов спорта. В ней изложены основные подходы к использованию методов математической статистики к анализу тренировочного процесса и эффективности проведения спортивных соревнований.

Настоящую работу можно рассматривать и как один из подходов к изложению курса метрологии. Предлагаемые статистические идеи переносимы на индивидуальные виды спорта и другие области соревнований и проявлений человека и коллектива.

Рецензенты: А. П. Данилов, заслуженный тренер России.

М. Н. Жуков, доктор педагогических наук.

ISBN 5-87555-234-4

© В. В. Афанасьев

© И. Н. Непряев

## Содержание

Введение .....	5
Глава I. О спорте в цифрах .....	7
§1. Вариационный и статистический ряд спортивных результатов .....	7
§2. Выборочные характеристики положения вариационного ряда спортивных достижений .....	18
§3. Выборочные характеристики рассеивания вариационного ряда и их использование в спорте .....	24
Глава II. Корреляция и конкордация в спорте .....	27
§4. Ранговая корреляция в хоккее .....	27
4.1. Командные результаты в регулярном чемпионате и в плей-офф .....	29
4.2. Стабильность последних чемпионатов России .....	30
4.3. Личные результаты игроков «Локомотива» .....	32
4.4. Лучшие игроки чемпионата России .....	34
4.5. Итоги чемпионатов мира по хоккею .....	36
4.6. Анализ предсезонной подготовки .....	37
4.7. А судьи кто? .....	41
§5. Корреляционные матрицы по футболу .....	42
5.1. Стабильность чемпионатов России по футболу .....	42
5.2. Чемпионаты ведущих европейских стран .....	47
5.3. Мировые футбольные державы .....	51
§6. Конкордация и стабильность чемпионатов европейских стран .....	53
6.1. Итоги последних пяти чемпионатов четырех европейских стран .....	54
6.2. Итоги последнего чемпионата четырех европейских стран .....	58
6.3. Текущие результаты последних чемпионатов России по хоккею и футболу .....	63
6.4. Конкордация в хоккее .....	67
§7. Коэффициент корреляции Пирсона .....	69
7.1. Бюджеты клубов и их результаты .....	71
7.2. Влияние первого гола на результат .....	74
7.3. Совмещенные ковариационные графы .....	76

Глава III. Бюджеты клубов и их спортивные достижения .....	82
§8. Самый популярный вид спорта .....	82
§9. Бюджеты хоккейных клубов и их результаты .....	87
§10. А как дела в волейболе? .....	90
Глава IV. Линии регрессии и прогнозы в спорте .....	92
§11. Прямые регрессии и прогнозы в хоккее .....	93
11.1. Прогноз результативности хоккеистов.....	93
11.2. Российские чемпионаты по хоккею и их прогнозы .....	106
11.3. Что мы ожидаем от чемпионата мира? .....	107
§12. Прогнозы и надежды российского футбола.....	108
12.1. Результативность российских футболистов.....	109
12.2. Российские чемпионаты по футболу .....	110
12.3. Прогнозы на предстоящий чемпионат мира в Германии .....	114
Список использованной литературы .....	116

## Введение

Применение методов математической статистики в спортивных играх, как правило, ограничивается указанием набранных очков, числом побед, ничьих, поражений, разностью забитых и пропущенных шайб (или мячей). Приводится сравнение результатов отдельных команд или игроков в двух или нескольких первенствах, но не анализируются совокупные результаты всех команд или игроков по их амплуа. Этот пробел пытаемся устранить, рассматривая и анализируя результаты в командных видах спорта.

С этой целью рассматриваются итоги чемпионатов России и ведущих европейских стран по основным видам спорта – футболу, хоккею, волейболу. Через вычисление коэффициентов корреляции и конкордации определяется устойчивость или изменчивость результатов той или иной лиги за последние годы. Такой подход позволяет организаторам первенств своих стран вносить изменения в их регламент и делать это статистически обоснованно.

В работе анализируются командные и личные результаты последних хоккейных чемпионатов мира и итоги регулярного чемпионата и плей-офф суперлиги. Вычислены основные их характеристики и предлагаются вероятностно-статистические интерпретации, которые могут быть полезны руководителям команд и действующим хоккеистам. Найденные результаты могут использоваться в организации учебно-тренировочного процесса, а также для комплектования команд в период заявок и дозаявок.

Эта книга еще и возможность поговорить о спорте на языке цифр, статистического анализа с преподавателями, учителями физической культуры и спортивными функционерами.

Методами математической статистики можно устанавливать перспективность спортсменов, условия, наиболее благоприятные для тренировок, их эффективность.

Анализируя результаты нескольких турниров или ежегодных чемпионатов, можно на основе линий регрессии делать вероятностные прогнозы на успешность выступления той или иной команды, прогнозировать результативность игрока в наступающем первенстве. Результат будет тем точнее, чем точнее будут соблюдаться правила и условия предыдущих чемпионатов. Так, предполагается, что хоккейный нападающий выступает в той же тройке, состав команды и ее руководство (президент, тренерский штаб и менеджеры) не претерпели больших изменений.

Вероятностно-статистический прогноз может быть использован старшим тренером и руководством клуба для оценки выступления как отдельного спортсмена, так и команды в целом. Президенту клуба при подготовке нового контракта с игроком целесообразно поинтересоваться статистически обоснованными возможностями спортсмена, и в этом случае финансовые условия его контракта будут математически обоснованы, и такой подход позволит упорядочить размер и продолжительность рассматриваемых контрактов.

Не следует только предлагаемый подход использовать прямолинейно и однозначно. Разумно статистические данные рассматривать с учетом психологических, физиологических и других факторов.

Целью предлагаемой работы является формирование у специалистов устойчивого желания измерять и сравнивать командные и личные показатели спортсменов на соревнованиях и тренировках, анализировать их, давать обоснованные рекомендации по совершенствованию учебно-тренировочного процесса. Будущему учителю физической культуры следует развивать стремление к измерению и фиксации функциональных показателей школьников в процессе их занятий на уроках физкультуры или занятиях в спортивных секциях.

Организаторам спортивных соревнований и функционерам спорта в работе предлагаются вероятностно-статистические методы оценивания справедливости финансовых и других условий проведения первенств.

Любители спорта найдут в книге необычный подход к рассмотрению и анализу большого множества спортивно-статистических показателей и достижений команд и спортсменов в самых популярных видах спорта – футболе, хоккее и волейболе.

Всем вместе предлагается занимательным образом соединить, казалось бы, далекие друг от друга математику и спорт. Авторы стремились использовать минимально необходимый и достаточно простой математический аппарат. Овладение предлагаемыми вероятностно-статистическими методами позволит лучше и глубже понять возможности человеческого организма, соотнести уровни финансирования и спортивные результаты, сделать прогнозы на спортивные достижения в далекой и близкой перспективе.

Предложенный подход может быть продолжен к анализу не только в командных видах спорта, но и в личных, таких, как легкая и тяжелая атлетика, плавание, конькобежный или лыжный виды спорта, и другие. Грамотное использование вероятностно-статистических методов позволит планировать, а при необходимости и корректировать организацию учебно-тренировочного процесса.

## Глава I. О спорте в цифрах

### §1. Вариационный и статистический ряд спортивных результатов

В математической статистике исследуются утверждения, которые могут быть сделаны на основе измерений некоторой величины. Если выборка объема  $n$  содержит  $k$  различных элементов  $x$ :  $x_1, x_2, x_3, \dots, x_k$ , причем  $x_i$  встречается  $m_i$  раз, то число  $m_i$  называют частотой элемента  $x_i$ . Также рассматриваются величины  $f_i = \frac{m_i}{n}$  – относительные частоты элементов  $x_i$ .

Данные измерений записывают в виде вариационного или статистического ряда. Вариационным или статистическим рядом называется таблица, первая строка которой содержит в порядке возрастания элементы  $x_i$ , а вторая – их частоты  $m_i$  (относительные частоты  $f_i = \frac{m_i}{n}$ ).

Рассмотрим снайперские качества нападающих команды «Локомотив» (Ярославль) по итогам регулярного чемпионата.

#### Результативность нападающих команды «Локомотив» по итогам регулярного чемпионата 2004 - 2005 гг.

Таблица 1

№	Игрок	Голы
1	Михнов А.	0
2	Галимов А.	1
3	Руденко К.	1
4	Карповцев А.	2
5	Науров В.	2
6	Самылин В.	2
7	Жуков С.	3
8	Яшин А.	3
9	Рязанцев А.	4
10	Антропов Н.	4
11	Васильев А.	5
12	Горохов И.	5
13	Шафигулин Г.	5
14	Рахунке К.	6
15	Мерфи К.	6
16	Красоткин Д.	7
17	Швидкий Д.	7



18	Королев И.	8
19	Непряев И.	10
20	Власенков Д.	10
21	Бут А.	12
22	Счастливым П.	15
23	Ткаченко И.	15
24	Антипов В.	18

Данные запишем в виде следующего вариационного ряда.

### Вариационный ряд

Таблица 2

$x_i$	0	1	2	3	4	5	6	7	8	10	12	15	18
$m_i$	1	2	3	2	2	3	2	2	1	2	1	2	1

$$n = \sum m_i = 24$$

Эти же данные можно записать в виде статистического ряда.

### Статистический ряд

Таблица 3

$x_i$	0	1	2	3	4	5	6	7	8	10	12	15	18
$f_i$	0,042	0,083	0,125	0,083	0,083	0,125	0,083	0,083	0,042	0,083	0,042	0,083	0,042

$$f_i = \frac{m_i}{n}, \quad \sum_{i=1}^{13} f_i = 1$$

Полученный статистический материал  $x_1, x_2, x_3, \dots, x_k$  наблюдений представляет собой первичные данные и величины, подлежащие статистической обработке. Статистические данные оформляются в виде таблицы, графика, гистограммы и т.д.

Полигоном частот /относительных частот/ выборки называется ломаная с вершинами в точках  $(x_i, m_i)$  или  $(x_i, f_i)$  для полигона относительных частот.

Для построения полигона частот обозначим на оси абсцисс возможные значения признака  $x_i$ , а на оси ординат соответствующие частоты  $m_i$  и полученные точки с координатами  $(x_i, m_i)$  соединим отрезками.

Размах варьирования (разность наибольшего и наименьшего значения признака) в нашем случае:  $R = x_{max} - x_{min} = 18 - 0 = 18$ .

Построим полигон частот для приведенных данных:

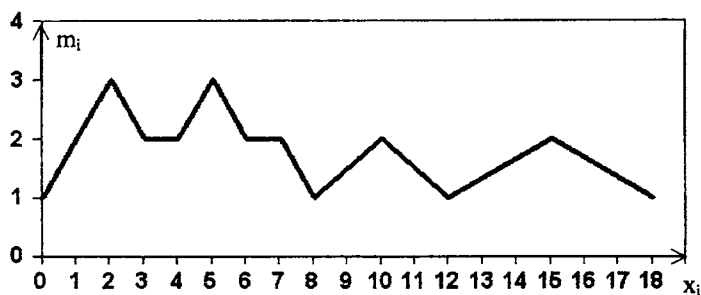


Рис. 1. Полигон частот

Чтобы построить полигон относительных частот, отложим на оси абсцисс значения  $x_i$ , а на оси ординат – относительные частоты  $f_i$ , после этого последовательно соединим полученные точки отрезками.

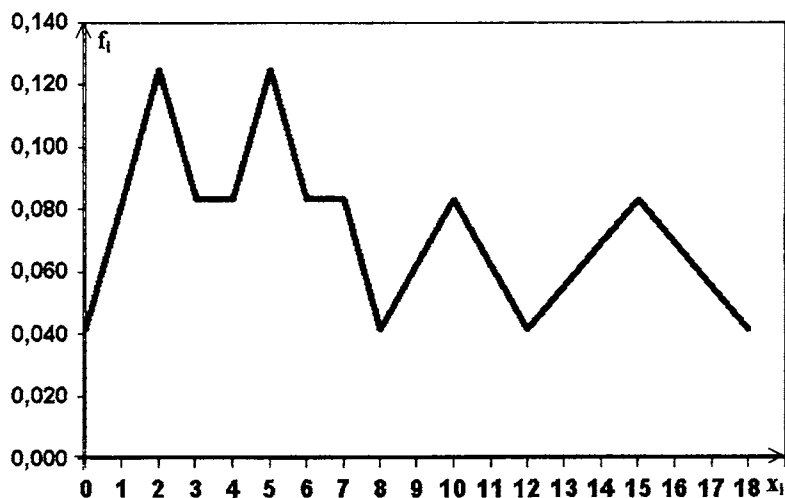


Рис. 2. Полигон относительных частот

С помощью построенных графиков мы наглядно представили распределение количества голов, забитых теми или иными игроками.

Изменение признака можно оценить через размах варьирования (в нашем примере  $R = 18 - 0 = 18$ ), который показывает различия между лучшим и худшим снайперами. В данном случае между В. Антиповым и А. Михновым.

В спорте статистические методы применяются не только для анализа результативности, но и для анализа других качественных характеристик спортсменов и команд. В качестве примера рассмотрим вес игроков команды «Локомотив» (Ярославль).

### Вес игроков команды «Локомотив»

Таблица 4

Игрок	Вес (кг)
Рязанцев А.	82
Карповцев А.	102
Рахунк К.	97
Мерфи К.	85
Жуков С.	87
Васильев А.	90
Красоткин Д.	92
Горохов И.	90
Михнов А.	99
Бут А.	91
Непряев И.	85
Шафигулин Г.	79
Счастливый П.	92
Ткаченко И.	85
Швидкий Д.	95
Королев И.	88
Науров А.	86
Галимов А.	86
Самылин В.	97
Антипов В.	84
Руденко К.	80
Власенков Д.	95
Яшин А.	96
Антропов Н.	104

В этом случае для анализа удобно использовать гистограмму выборки. Для этого разобьем диапазон изменения веса на интервалы и для каждо-

го из них рассмотрим, сколько игроков ( $m_i$ ) имеют вес в данном диапазоне ( $h_i$ ).

### Распределение игроков команды «Локомотив» по весу

Таблица 5

Вес	Частота $m_i$	Плотность частот $\frac{m_i}{nh_i}$
75-80	1	0,004
80-84	2	0,014
84-90	8	0,056
90-95	5	0,042
95-100	6	0,05
100-104	2	0,021

Построим на оси абсцисс заданные интервалы и проведем над этими интервалами отрезки, параллельные оси абсцисс и находящиеся на расстояниях, равных соответственным плотностям относительной частоты. Гистограмму можно «сгладить» и по полученному графику сделать предположение о характере распределения.

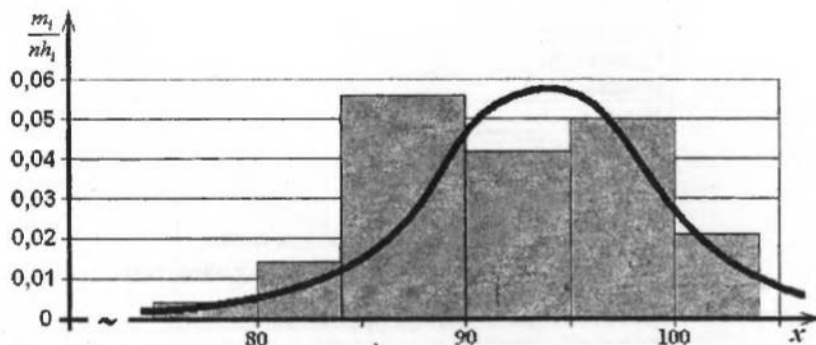


Рис. 3. «Сглаженная» гистограмма выборки

Аналогичным образом можно рассмотреть результативность игроков команды «Локомотив»:

Результативность игроков команды «Локомотив»

Таблица 6

Игрок	Количество голов
Рязанцев А.	4
Карповцев А.	2
Рахунек К.	6
Мерфи К.	6
Жуков С.	3
Васильев А.	5
Красоткин Д.	7
Горохов И.	5
Михнов А.	0
Бут А.	12
Непряев И.	10
Шафигулин Г.	5
Счастливый П.	15
Ткаченко И.	15
Швидкий Д.	7
Королев И.	8
Науров А.	2
Галимов А.	1
Самылин В.	2
Антипов В.	18
Руденко К.	1
Власенков Д.	10
Яшин А.	3
Антропов Н.	4

Распределение количества шайб,  
заброшенных игроками команды «Локомотив»

Таблица 7

Количество го- лов	Частота $m_i$	Плотность частот $\frac{m_i}{nh_i}$
0-5	13	0,108
5-10	7	0,058
10-15	3	0,025
15-18	1	0,014

Построим гистограмму выборки по рассмотренным данным:

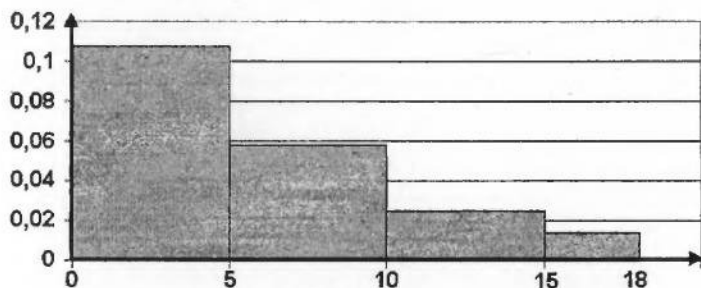


Рис. 4. Гистограмма распределения количества заброшенных шайб игроками команды «Локомотив»

Из способа построения гистограммы следует, что полная ее площадь равна единице, что позволяет «сглаживать» гистограмму и по ней приближенно находить классический непрерывный закон распределения.

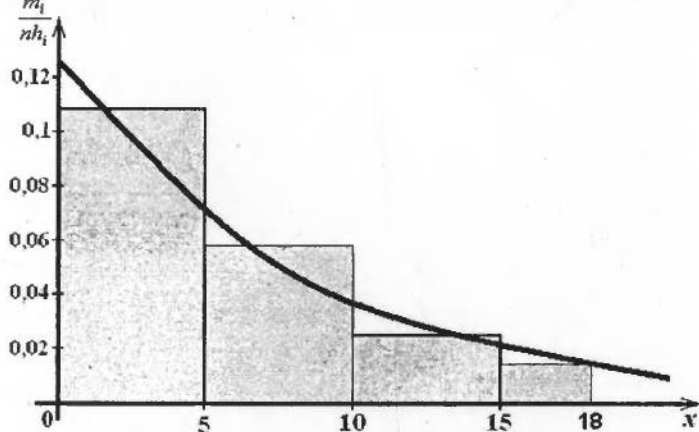


Рис. 5. «Сглаженная» гистограмма распределения количества заброшенных шайб игроками команды «Локомотив»

Анализ «сглаженной» гистограммы является удобным способом предположить вид дифференциальной функции распределения. В данном случае распределение количества заброшенных шайб (дискретный

закон распределения) напоминает показательный (непрерывный) закон распределения. Ранее рассмотренное распределение игроков по весу (рис. 3) напоминает непрерывный нормальный закон распределения.

Также «сглаженную» кривую можно построить и по количеству набранных игроками команды очков (по системе «гол+пас»):

**Количество очков, набранных игроками команды  
«Локомотив»**

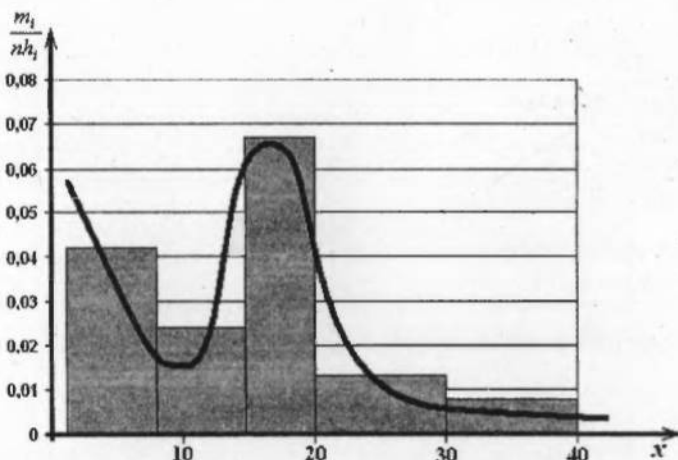
Таблица 8

<b>Игрок</b>	<b>Количество очков</b>
<b>Рязанцев А.</b>	19
<b>Карповцев А.</b>	7
<b>Рахунек К.</b>	14
<b>Мерфи К.</b>	12
<b>Жуков С.</b>	4
<b>Васильев А.</b>	17
<b>Красоткин Д.</b>	17
<b>Горохов И.</b>	16
<b>Михнов А.</b>	9
<b>Бут А.</b>	34
<b>Непряев И.</b>	20
<b>Шафигулин Г.</b>	11
<b>Счастливый П.</b>	30
<b>Ткаченко И.</b>	30
<b>Швидкий Д.</b>	18
<b>Королев И.</b>	28
<b>Науров А.</b>	4
<b>Галимов А.</b>	2
<b>Самылин В.</b>	6
<b>Антипов В.</b>	40
<b>Руденко К.</b>	1
<b>Власенков Д.</b>	19
<b>Яшин А.</b>	6
<b>Антропов Н.</b>	19

**Распределение количества очков,  
набранных игроками команды «Локомотив»**

Таблица 9

Количество очков	Частота $m_i$	Плотность частот $\frac{m_i}{nh_i}$
1-8	7	0,042
8-15	4	0,024
15-20	8	0,067
20-30	3	0,013
30-40	2	0,008



**Рис. 6. «Сглаженная» гистограмма распределения количества заброшенных шайб игроками команды «Локомотив», которая не является классическим законом распределения.**

Построим гистограмму забитых клубами суперлиги голов за 60 матчей (т.е. по окончании регулярного чемпионата).



**Количество голов, забитых командами суперлиги  
в сезоне 2004 - 2005 гг.**

Таблица 10

<b>Команды</b>	<b>Количество забитых голов</b>
"Металлург" Мг.	190
"Авангард"	178
"Динамо"	173
"Ак - барс"	171
ЦСКА	154
"Локомотив"	151
"Нефтехимик"	138
"Лада"	137
"Северсталь"	132
СКА	130
"Металлург" Нк.	129
"Химик"	129
"Салават Юлаев"	113
"Сибирь"	93
"Спартак"	86
"Молот – Прикамье"	77

Вид гистограммы зависит от способа разбиения диапазона значений на интервалы.

**Распределение количества голов, забитых командами  
суперлиги  
(первый способ разбиения)**

Таблица 11

<b>Количество голов</b>	<b>Частота <math>m_i</math></b>	<b>Плотность частот <math>\frac{m_i}{nh_i}</math></b>
77-128	4	0,005
129-136	4	0,031
137-172	5	0,009
173-190	3	0,011

**Распределение количества голов, забитых командами  
суперлиги (второй способ разбиения)**

Таблица 12

Количество голов	Частота $m_i$	Плотность частот $\frac{m_i}{nh_i}$
77-92	2	0,008
93-136	6	0,009
137-172	5	0,009
173-190	3	0,011

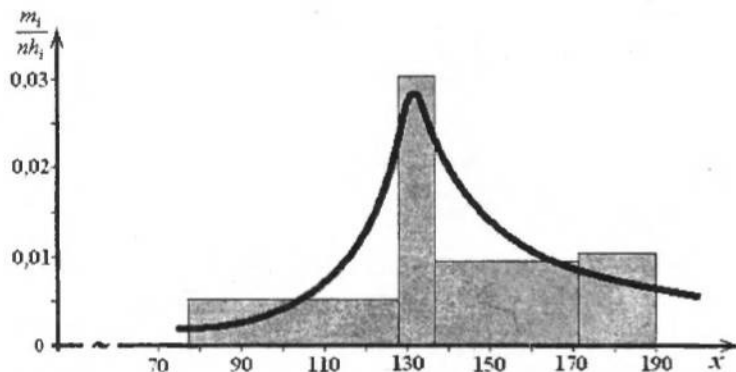


Рис. 7. «Сглаженная» гистограмма распределения количества голов, забитых командами суперлиги (первый способ разбиения)

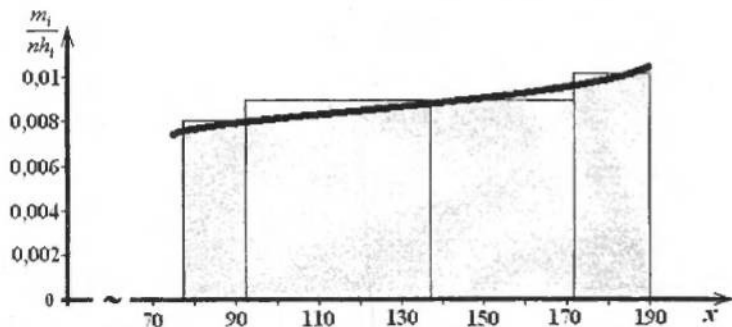


Рис. 8. «Сглаженная» гистограмма распределения количества голов, забитых командами суперлиги (второй способ разбиения)

Второй способ разбиения, который приближен к равномерному закону распределения.

Из способа построения гистограммы следует, что полная ее площадь равна единице, что позволяет «сглаживать» гистограмму и по ней приближенно находить классический непрерывный закон распределения.

## §2. Выборочные характеристики положения вариационного ряда спортивных достижений

Основными характеристиками положения вариационного ряда являются выборочное среднее, медиана  $M_e$  и мода  $M_o$ , каждая из которых может являться и определенным спортивным показателем.

### Выборочное среднее

Выборочным средним называется величина  $\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i m_i$ , по

которой можно судить о средних достижениях команд и отдельных игроков.

Рассмотрим результаты команд суперлиги по итогам регулярного чемпионата 2004 – 2005 гг. по хоккею.

### Положение команд. Чемпионат России 2004 – 2005 гг.

Таблица 13

№	Команды	В	ВО	Н	ПО	П
1	"Динамо"	35	5	7	4	9
2	"Лада"	35	1	11	0	13
3	"Металлург" Мг	34	2	5	4	15
4	"Ак - барс"	34	3	5	1	17
5	"Локомотив"	29	2	12	2	15
6	"Авангард"	29	3	10	1	17
7	"Металлург" Нк	26	1	12	2	19
8	"Нефтехимик"	26	2	7	2	23
9	"Химик"	25	1	6	0	28
10	ЦСКА	21	3	9	3	24
11	"Северсталь"	21	2	10	3	24
12	СКА	22	1	5	3	29

13	"Салават Юлаев"	19	1	7	2	31
14	"Сибирь" -	12	2	12	2	32
15	"Спартак"	10	0	10	2	38
16	"Молот - Прикамье"	5	2	4	0	49

Выборочное среднее в данном случае можно найти, подсчитав среднее арифметическое соответствующих показателей по всем командам суперлиги.

Выигрыши

$$x^* = \frac{1}{16}(35+35+34+34+29+29+26+26+25+21+21+22+19+12+10+5) = 24$$

Выигрыш в овертайме

$$x^* = \frac{1}{16}(5+1+2+3+2+3+1+2+1+3+2+1+1+2+0+2) = 2$$

Ничьи

$$x^* = \frac{1}{16}(7+11+5+5+12+10+12+7+6+9+10+5+7+12+10+4) = 8$$

Поражения в овертайме

$$x^* = \frac{1}{16}(4+0+4+1+2+1+2+2+0+3+3+3+2+2+2+0) = 2$$

Поражения

$$x^* = \frac{1}{16}(9+13+15+17+15+17+19+23+28+24+24+29+31+32+38+49) = 24$$

Используем полученные средние значения для сравнения этих показателей с результатами местной команды «Локомотив».

По данным значениям строим сравнительную диаграмму клубов суперлиги и ярославского «Локомотива».



Рис. 9. Сравнительная диаграмма команд суперлиги и команды «Локомотив» (Ярославль).

По диаграмме видно, что у «Локомотива» число выигрышей и ничьих больше, чем в среднем по лиге, а поражений – меньше. Это свидетельствует о том, что «Локомотив» показал результаты выше средних в суперлиге.

### Медиана (М<sub>е</sub>)

Рассмотрим данные о результативности одного из ведущих нападающих команды «Локомотив» Владимира Антипова в сезоне 2004 - 2005 гг.

### Результативность В. Антипова в регулярном чемпионате 2004 – 2005 гг.

Таблица 14

№	Месяц-год	Количество голов (Г)	Количество очков (Г + П)
1	09.2004	8	13
2	10.2004	2	9
3	11.2004	1	3
4	12.2004	2	5
5	01.2005	3	5
6	02.2005	2	3
7	03.2005	0	2

**Количество голов, забитых В. Антиповым  
в регулярном чемпионате 2004 – 2005 гг. (по месяцам)**

Таблица 15

№	1	2	3	4	5	6	7
Количество голов	8	2	1	2	3	2	0

Медианой ( $M_e$ ) называется вариант  $x_i$  такой, что

$$\sum_{i=1}^l m_i \geq \frac{n}{2} \quad \text{и} \quad \sum_{i=l}^k m_i \geq \frac{n}{2}.$$

Медиана обладает тем свойством, что сумма абсолютных величин отклонений вариантов от медианы меньше, чем от любой другой величины (в том числе и от выборочного среднего). В рассмотренном случае  $M_e = 2$ , так как

$$8 + 2 = 10 \geq \frac{18}{2} \quad \text{и} \quad 2 + 1 + 2 + 3 + 2 = 10 \geq \frac{18}{2}.$$

Это говорит о том, что В. Антипов за первые два месяца чемпионата забил столько голов, сколько и за оставшиеся 5 месяцев.

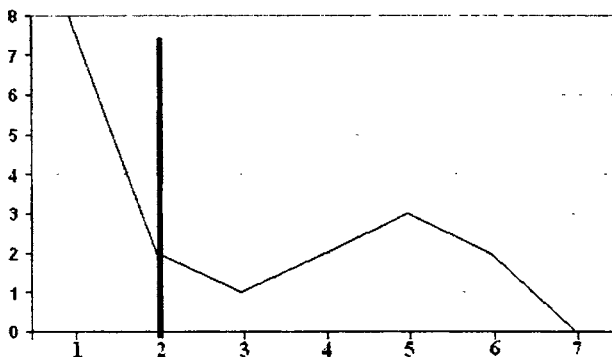


Рис. 10. Количество забитых голов по месяцам  
регулярного чемпионата 2004 – 2005 гг. и медиана

Аналогично можем рассмотреть количество набранных В. Антиповым очков (по системе «гол + пас») в регулярном чемпионате 2004 – 2005:

Количество очков, набранных В. Антиповым в регулярном чемпионате 2004 – 2005 гг. (по месяцам)

Таблица 16

№	1	2	3	4	5	6	7
Количество очков	13	9	3	5	5	3	2

Значение медианы:  $M_e = 2$ , т.к.  $13 + 9 = 22 \geq \frac{40}{2}$  и  $9 + 3 + 5 + 5 + 3 + 2 = 27 \geq$

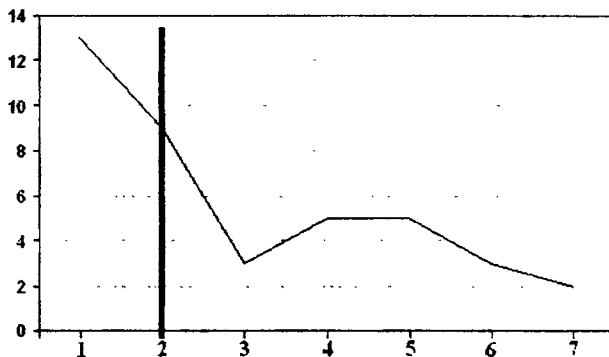


Рис. 11. Количество набранных очков по месяцам регулярного чемпионата 2004 – 2005 и медиана

Полученные медианы для количества голов и очков у В. Антипова совпадают и показывают, что его результативность за первые два месяца чемпионата была почти такая же, как и за оставшиеся 5 месяцев. На это следует обратить внимание и самого игрока, и тренерского штаба команды.

### Мода ( $M_o$ )

Модой ( $M_o$ ) называется вариант, наиболее часто встречающийся в данном вариационном ряду.

Рассмотрим для примера итоги регулярного чемпионата России команд суперлиги 2004 - 2005 годов.

### Положение команд. Чемпионат России 2004 – 2005 гг.

Таблица 17

№	Команды	В	ВО	Н	ПО	П
1	"Динамо"	35	5	7	4	9
2	"Лада"	35	1	11	0	13

3	"Металлург" Мг	34	2	5	4	15
4	"Ак – барс"	34	3	5	1	17
5	"Локомотив"	29	2	12	2	15
6	"Авангард"	29	3	10	1	17
7	"Металлург" Нк	26	1	12	2	19
8	"Нефтехимик"	26	2	7	2	23
9	"Химик"	25	1	6	0	28
10	ЦСКА	21	3	9	3	24
11	"Северсталь"	21	2	10	3	24
12	СКА	22	1	5	3	29
13	"Салават Юлаев"	19	1	7	2	31
14	"Сибирь"	12	2	12	2	32
15	"Спартак"	10	0	10	2	38
16	"Молот- Прикамье"	5	2	4	0	49

Найдем  $M_0$  для всех команд по выигрышам в овертайме, составив для этого соответствующий вариационный ряд.

Таблица 18

В.О.	0	1	2	3	4	5
	1	4	6	3	0	1

Из этого следует, что самым модным результатом является два выигрыша в овертайме, которые встречались в шести играх.

Аналогично найдем  $M_0$  для числа поражений в овертайме.

Таблица 19

П.О.	0	1	2	3	4
	3	2	6	3	2

Из этого следует, что самым модным, или чаще всего встречающимся, результатом являются два поражения в овертайме.

Определим моду для числа зафиксированных ничьих в регулярном чемпионате России по хоккею.

Таблица 20

Н	4	5	6	7	9	10	11	12
	1	3	1	3	1	3	1	3



В этом примере не будет однозначного ответа, так как модой здесь являются четыре варианта – это пять ничьих, семь, десять или двенадцать.

### §3. Выборочные характеристики рассеивания вариационного ряда и их использование в спорте

Для изучения вариационных рядов недостаточно характеристик положения, поэтому вводят еще характеристики рассеивания. Главной характеристикой рассеивания является дисперсия, которая вычисляется по формуле  $D^*[X] = \frac{1}{n} \sum_{i=1}^k (x_i - x^*)^2 m_i$ , а положительный квадратный корень из дисперсии называется средним квадратическим отклонением  $\sigma^*[X] = +\sqrt{D^*[X]}$ .

Ошибка репрезентативности выборочной средней выражается формулой  $S_x = \sqrt{\frac{\sum (x_i - x^*)^2}{n(n-1)}}$ , а коэффициент вариации

$$C_s = \frac{S_x}{x^*} * 100\%$$

показывает близость выборочной средней к генеральному параметру. Показатель  $C_s$  считается в статистике удовлетворительным, если варьирует в пределах до 10%.

Оценим относительную равноценность участников чемпионата команд суперлиги и НХЛ в зависимости от размера их бюджета.

#### Бюджеты некоторых клубов суперлиги в 2004 г.

Таблица 21

№	Команды	Бюджет команды (млн.\$)
1	"Динамо"	20
2	"Лада"	18
3	"Металлург" Мг.	24
4	"Ак – барс"	50
5	"Локомотив"	19
6	"Авангард"	35
7	"Химик"	20
8	ЦСКА	17
9	"Северсталь"	23
10	СКА	10
11	"Спартак"	8

Найдем средний бюджет для клубов суперлиги:

$$x^* = \frac{1}{11} (20+18+24+50+19+35+20+17+23+10+8) \approx 22$$

Далее найдем сумму квадратов разности бюджетов и среднего бюджета

$$(20-22)^2+(18-22)^2+(24-22)^2+(50-22)^2+(19-22)^2+(35-22)^2+(20-22)^2+(17-22)^2+(23-22)^2+(10-22)^2+(8-22)^2 = 2^2+4^2+2^2+28^2+3^2+13^2+2^2+5^2+1^2+12^2+14^2 = 1365$$

Разделим полученную сумму на  $n(n-1)$ , в данном случае  $11 \cdot 10$ , тогда

$$S_x = \sqrt{\frac{1365}{110}} \approx 3,51$$

откуда находим коэффициент вариации

$$C_s = 3,51/22 \cdot 100\% = 0,15 \cdot 100\% = 15\%$$

Также посчитаем коэффициент вариации для Национальной хоккейной лиги.

#### Бюджеты некоторых клубов НХЛ в 2003 г. (до локаута)

Таблица 22

№	Команды	Бюджет команды (млн.\$)
1	Детройт "Ред Уингз"	79
2	Сан-Хосе "Шаркс"	35
3	Бостон "Брюинз"	46
4	Филадельфия "Флайерз"	65
5	Колорадо "Эвеланш"	61
6	Нэшвилл "Предэйторс"	23
7	Лос-Анджелес "Кингз"	46
8	Атланта "Трэшерз"	27
9	Нью-Йорк "Рейнджерс"	77
10	Вашингтон "Кэпиталс"	51
11	Питсбург "Пингвинз"	27

Найдем средний бюджет для клубов НХЛ:

$$x^* = \frac{1}{11} (79+35+46+65+61+23+46+27+77+51+27) \approx 48$$

Далее найдем сумму квадратов разности бюджетов клубов и среднего

$$(79-48)^2+(35-48)^2+(46-48)^2+(65-48)^2+(61-48)^2+(23-48)^2+(46-48)^2+(27-48)^2+(77-48)^2+(51-48)^2+(27-48)^2 = 31^2+13^2+2^2+17^2+13^2+25^2+2^2+21^2+29^2+3^2+21^2 = 961+169+4+289+169+625+4+441+841+9+441 = 3953$$

Разделим полученный результат на  $n(n-1)$ , в данном случае  $11 \cdot 10$  и получим:

$$S_x = \sqrt{\frac{3953}{110}} \approx 5,9$$

Следовательно, коэффициент вариации

$$C_v = 5,9/48 \cdot 100\% = 0,12 \cdot 100\% = 12\%$$

Если сравнить два показателя, то можно сделать вывод, что и в суперлиге и в НХЛ бюджеты распределены не совсем справедливо, но справедливость распределения бюджетов в суперлиге меньше, чем в НХЛ.

Заметим неслучайность локаута в НХЛ и потери целого сезона, а следовательно, и прибыли клубов. Нам неизвестно, какие расчеты проводились в НХЛ, но сама экономическая ситуация подсказала необходимость внесения изменений в правила функционирования всей лиги и каждого клуба в отдельности. Введение ограничений на бюджеты клубов приведут к более справедливым условиям проведения хоккейного турнира в НХЛ, что установлено в этом параграфе. К такому результату рано или поздно придут и в хоккейной федерации России. Предпосылки к этому появились уже в 2005 году, когда экономический крах настиг «Ладу» и «Молот – Прикамье». А результат, который при этом показывает команда «Лада», говорит еще и о явно завышенных зарплатах в суперлиге.

## Глава II. Корреляция и конкордация в спорте

Понятие **корреляции** является одним из основных понятий теории вероятностей и математической статистики, оно было введено Гальтоном и Пирсоном. Закон природы или общественного развития может быть представлен описанием совокупности взаимосвязей. Если эти зависимости стохастичны, а анализ осуществляется по выборке из генеральной совокупности, то данная область исследования относится к задачам стохастического исследования зависимостей, которые включают в себя корреляционный, регрессионный, дисперсионный и ковариационный анализы.

Наряду с линейным коэффициентом корреляции для измерения тесноты между двумя коррелируемыми признаками часто используются менее точные, но более простые по расчету коэффициенты корреляции рангов (или **ранговые коэффициенты корреляции**) Спирмена и Кендалла. Оба показателя, названные именами английских ученых, предложивших эти коэффициенты, основаны на корреляции не самих значений коррелируемых признаков, а их рангов.

Если число ранжируемых признаков больше двух, то для измерения тесноты связи между ними можно использовать предложенный М. Кендаллом и Б. Смитом коэффициент конкордации (**множественный коэффициент ранговой корреляции**).

### §4. Ранговая корреляция в хоккее

В качестве измерителей степени тесноты парных связей между количественными переменными будем использовать коэффициент ранговой корреляции. Пусть объекты генеральной совокупности обладают двумя качественными признаками и выборка объема  $n$  содержит независимые объекты, которые будем располагать (ранжировать) в порядке ухудшения качества. Для оценки степени связи признаков вводят коэффициенты ранговой корреляции Спирмена и Кендалла (см., например, [2] С.185-190). Рассматривая ранги  $x_1, x_2, \dots, x_n$  как возможные значения случайной величины  $X$ , а  $y_1, y_2, \dots, y_n$  – как возможные значения случайной величины  $Y$ , можно вычислить выборочный коэффициент корреляции Спирмена: 
$$r_s = 1 - \frac{6 \sum d_i^2}{(n-1) \cdot n \cdot (n+1)},$$
 где  $d_i = x_i - y_i$  (разность соответствующих рангов). Предпочтение отдаем коэффициенту Спирмена, поскольку он дает более точный результат.

Заметим, что при равных показателях у нескольких участников им присваивается один общий ранг, равный среднему арифметическому соответствующих возможных мест.

Коэффициент корреляции может принимать значение от минус единицы до плюс единицы, причем при  $r = 1$  связь прямая, а при  $r = -1$  – обратная.

Считают, что если:

$|r| \leq 0,3$ , то связь слабая;  
 $0,3 < |r| \leq 0,7$  – связь средняя;  
 $0,7 < |r| \leq 1$  – связь сильная.

В том случае, когда исследуется связь между несколькими признаками, корреляцию называют множественной, и она задается всеми коэффициентами  $r_{ij}$  парных корреляций, которые записывают в корреляционную матрицу:

$$(r) = \begin{pmatrix} 1 & r_{12} & r_{13} & \dots & r_{1n} \\ & 1 & r_{23} & \dots & r_{2n} \\ & & 1 & \dots & r_{3n} \\ & & & & M \\ & & & & 1 \end{pmatrix}$$

Заметим, что матрица  $(r)$  является треугольной, поскольку  $r_{ij} = r_{ji}$  и нет смысла их повторять дважды.

### Анализ итогов чемпионата России 2004 - 2005 годов

Приведем итоговую таблицу регулярного чемпионата

Таблица 23

№ п/п	Название команды	И	В	ВО	Н	ПО	П	Ш	О
1.	"Динамо"	60	35	5	7	4	9	179-106	126
2.	"Лада"	60	35	1	11	0	13	140-86	118
3.	"Металлург Мг"	60	34	2	5	4	15	193-124	115
4.	"Ак – барс"	60	34	3	5	1	17	174-113	114
5.	"Локомотив"	60	29	2	12	2	15	159-104	105
6.	"Авангард"	60	29	3	10	1	17	182-148	104
7.	"Металлург Нк"	60	26	1	12	2	19	129-126	94
8.	"Нефтехимик"	60	26	2	7	2	23	141-130	91
9.	"Химик"	60	25	1	6	0	28	129-157	83
10.	ЦСКА	60	21	3	9	3	24	156-147	81
11.	"Северсталь"	60	21	2	10	3	24	138-144	80

12.	СКА	60	22	1	5	3	29	133-169	76
13.	"Салават Юлаев"	60	19	1	7	2	31	114-156	68
14.	"Сибирь"	60	12	2	12	2	32	97-138	54
15.	"Спартак"	60	10	0	10	2	38	89-164	42
16.	"Молот-Прикамье"	60	5	2	4	0	40	81-222	23

Проанализируем показатели команд по числу выигрышей, выигрышей в овертайме, ничьим, разности шайб и добавим результаты плей-офф.

Полученные результаты проранжируем по ухудшению признака и представим полученные ранги в следующей таблице:

Таблица 24

Место (1)	Команды	В (2)	ВО (3)	Н (4)	ПО (5)	Ш (6)	Плей-офф	
							место	шайбы
1.	"Динамо"	1,5	1	10	15,5	1	1	1
2.	"Лада"	1,5	3	4	2	5	2	2
3.	"Металлург" Мг	3,5	7,5	14	15,5	2	5	5
4.	"Ак – барс"	3,5	3	14	4,5	3	6	4
5.	"Локомотив"	5,5	7,5	2	8,5	4	3	3
6.	"Авангард"	5,5	3	6	4,5	6	4	7
7.	"Металлург" Нк	7,5	13	2	8,5	9	7	6
8.	"Нефтехимик"	7,5	7,5	10	8,5	7	8	8
9.	"Химик"	9	13	12	2	11		
10.	ЦСКА	11,5	3	8	13	8		
11.	"Северсталь"	11,5	7,5	6	13	10		
12.	СКА	10	13	14	13	12		
13.	"Салават Юлаев"	13	13	10	8,5	14		
14.	"Сибирь"	14	7,5	2	8,5	13		
15.	"Спартак"	16	16	6	8,5	15		
16.	"Молот-Прикамье"	6	7,8	11	2	16		

#### 4.1. Командные результаты в регулярном чемпионате и в плей-офф

##### 4.1.1. Связь между местом в чемпионате и в плей-офф.

Воспользуемся формулой Спирмена, предварительно вычисляя квадраты разностей соответствующих рангов.

$$\sum a_i^2 = (3-5)^2 + (4-6)^2 + (5-3)^2 + (6-3)^2 = 21, \quad n=8$$

$$r_s = 1 - \frac{6 \cdot 21}{7 \cdot 8 \cdot 9} = 0,75.$$

Связь сильная и прямая, хотя и достаточно далекая от полной прямой связи, что говорит о некоторых расхождениях в итогах регулярного чемпионата и плей-офф.

#### 4.1.2. Результативность в чемпионате и в плей-офф.

$$\sum d_i^2 = (2-5)^2 + (5-2)^2 + (4-3)^2 + (3-4)^2 + (7-6)^2 + (6-8)^2 + (8-7)^2 = 26$$

$$r_s = 1 - \frac{6 \cdot 26}{7 \cdot 8 \cdot 9} = 0,69$$

Связь уже средняя.

4.1.3. Найдем корреляционную матрицу связи между местом, выигрышами, выигрышами в овер-тайме, ничьими, поражениями и разностью шайб в регулярном чемпионате.

$$(r_s) = \begin{matrix} M \\ B \\ BO \\ H \\ ПО \\ Ш \end{matrix} \begin{pmatrix} 1 & 0,99 & 0,61 & 0,05 & -0,08 & 0,96 \\ & 1 & 0,68 & -0,03 & 0,21 & 0,79 \\ & & 1 & 0,13 & 0,03 & 0,69 \\ & & & 1 & 0,23 & 0,01 \\ & & & & 1 & -0,23 \\ & & & & & 1 \end{pmatrix}$$

По полученной корреляционной матрице заключаем, что существует сильная прямая связь между местами команд и числом выигрышей (и разностью шайб); между разностью шайб и выигрышами. Можно считать практически независимыми места команд и количеством ничьих (поражений в овер-тайме); между выигрышами и ничьими, между выигрышами и поражениями в овер-таймах.

## 4.2. Стабильность последних чемпионатов России

Приведем проранжированную уже таблицу последних трех регулярных чемпионатов и добавим еще места команд 2004 - 2005 гг. по результатам домашних и гостевых игр, по результатам игр только с командами верхней половины турнирной таблицы.

Таблица 25

М (1)	Команды	2003- 2004 (2)	2002- 2003 (3)	Дома (4)	В гос- тах (5)	Случ- шим (6)	С худ- шим (7)
1	"Динамо"	6	7	3	1	1	2
2	"Лада"	2	5	1	3	2	4
3	"Металлург" Мг.	1	6	2	5	6	1
4	"Ак – барс"	5	4	5	2	4	3
5	"Локомотив"	7	1	6	4	3	8
6	"Авангард"	3	2	4	6	5	5
7	"Металлург" Нк.	4	9	7	8	9	7
8	"Нефтехимик"	8	11	9	7	11	6
9	"Химик"	12	16	8	12	12	9
10	ЦСКА	10	10	12	9	7	12
11	"Северсталь"	13	3	11	10	10	11
12	СКА	14	12	10	11	8	13
13	"Салават Юлаев"	9	8	13	13	15	10
14	"Сибирь"	11	13	14	14	14	14
15	"Спартак"	15,5	14	15	15	13	16
16	"Молот – Прикамье"	15,5	15	16	16	16	15

4.2.1. Итоги последних трех регулярных чемпионатов России представим в виде полученной матрицы ранговых корреляций:

$$(r) = \begin{pmatrix} 1 & 0,86 & 0,65 \\ & 1 & 0,66 \\ & & 1 \end{pmatrix}$$

Поскольку коэффициент корреляции  $r_{12} = 0,86$ , то последние два чемпионата характеризуются достаточно большой степенью стабильности, а сравнения уже с чемпионатом трехлетней давности, как и сами чемпионаты 2002 – 2003 гг. и 2003 – 2004 гг., годов, свидетельствуют о значительных их расхождениях.

4.2.2. О стабильности выступления команд в последнем чемпионате будем судить по вычисленной матрице ранговых корреляций между общим количеством очков (О) и очков, набранных в домашних (Д) и гостевых (Г) играх, в играх с лучшими (Л) и худшими (Х) командами.

$$(r_s) = \begin{matrix} O \\ Д \\ Г \\ Л \\ Х \end{matrix} \begin{pmatrix} 1 & 0,97 & 0,96 & 0,9 & 0,94 \\ & 1 & 0,91 & 0,86 & 0,93 \\ & & 1 & 0,93 & 0,91 \\ & & & 1 & 0,74 \\ & & & & 1 \end{pmatrix}$$

По элементам матрицы можно судить о достаточно высокой степени стабильности последнего хоккейного сезона. Исключение состав-



ляет только средняя связь ( $r = 0,74$ ) между результатами команд с лучшими и худшими командами, которая может быть объяснена, в частности, тем что такую классификацию команды получили уже по итогам чемпионата.

Обратим внимание тренерских штабов команд СКА, ЦСКА, «Локомотива» и «Салавата Юлаева», результаты которых значительно лучше в играх с более сильными соперниками, что может быть резервом в организации учебно-тренировочного процесса при подготовке (особенно психологической) к следующему сезону.

### 4.3. Личные результаты игроков «Локомотива»

4.3.1. Результаты нападающих «Локомотива» в плей-офф и в чемпионате.

Найдем матрицу ранговых корреляций результативности нападающих и ими набранных штрафных минут в плей-офф и в регулярном чемпионате. С этой целью приведем статистические данные, которые проранжируем.

Таблица 26

Место	Игрок	И	Гол+пас	Ранг	Нападающие		Штрафники			
					очки	ранг	плей-офф		регулярный	
							мин.	ранг	мин.	ранг
1	Яшин А.	9	10(3+7)	1	6/10	7	10	8	14	8
2-3	Антропов Н.	9	7(3+4)	2,5	19/20	6	18	10	44	7
2-3	Королев И.	9	7(1+6)	2,5	28	5	2	2,5	26	1
4	Бут А.	8	6(3+3)	4	33	2	0	1	58	6
5	Ткаченко И.	9	5(2+3)	5	30	3	8	6	30	3
6	Счастливый П.	9	4(1+3)	6	29	4	2	2,5	28	2
7	Власенков Д.	9	2(2+0)	7	18	9	6	6	36	4
8-10	Крюков А.	7	1(1+0)	9	17	10	4	4	44	5
8-10	Непряев И.	9	1(1+0)	9	20	8	16	9	73	10
8-10	Антипов В.	7	1(0+1)	9	39	1	4	5	65	9

Обратим внимание на то, что А. Яшин и Н. Антропов в регулярном чемпионате сыграли только 10 и 26 матчей соответственно против почти 60 матчей у других, поэтому их рейтинги находились с этим учетом. Вычисления опускаем и приводим получившуюся корреляционную матрицу.

$$(r_s) = \begin{pmatrix} 1 & 0,17 & -0,05 & 0,26 \\ & 1 & 0,37 & 0,05 \\ & & 1 & 0,53 \\ & & & 1 \end{pmatrix}$$

Выделим слабую связь ( $r_s=0,17!$ ) между очками, набранными нападающими «Локомотива» в плей-офф и в регулярном чемпионате, на которую тренерскому штабу команды необходимо обратить внимание при подготовке к следующему хоккейному сезону. Практически независимы показатели результативности и нарушений правил у нападающих ( $|r_s|=0,05$ ), существует средняя связь ( $r_s = 0,53$ ) между нарушениями в плей-офф и в чемпионате, что может свидетельствовать об отсутствии особого ажиотажа среди нападающих в играх плей-офф по сравнению с регулярным чемпионатом.

#### 4.3.2. Результаты нападающих и защитников «Локомотива» в регулярном чемпионате.

Приведем отдельные результаты нападающих и защитников и проранжируем их.

Таблица 27

Игрок	Место-ранг	Гол+пас	Ранг по голам	Ранг по передачам
Антипов В.	1	39(18-21)	1	1,5
Бут А.	2	33(12+21)	4	1,5
Ткаченко И.	3	30(15+15)	2,5	4,5
Счастливый П.	4	29(15+14)	2,5	6
Королев И.	5	28(8+20)	6,5	3
Непряев И.	6	20(10+10)	5,5	8
Антропов Н.	7	19(4+15)	11	5
Власенков Д.	8,5	18(10+8)	5,5	10
Швидкий Д.	8,5	18(7+11)	9	7
Крюков А.	10	17(8+9)	6,5	9
Шафигулин Г.	11	10(5+5)	10	11
Самылин В.	12	6(2+4)	12	12
Галимов А.	13	2(1+1)	13	13

Таблица 28

Игрок	Место-ранг	Гол+пас	Ранг по голам	Ранг по передачам
Рязанцев А.	1	19(4+15)	6	1
Красоткин Д.	2,5	17(7+10)	1	4
Васильев А.	2,5	17(5+12)	4,5	2
Горохов И.	4	16(5+11)	4,5	3
Рахунке К.	5	14(6+8)	2,5	5

Murphy С.	6	12(6+6)	2,5	7
Stepanek M.	7	7(0+7)	9	6
Карповцев А.	8	6(2+4)	8	8
Жуков С.	9	4(3+1)	7	9

Сравним корреляционные матрицы связей набранных очков, голов и передач:

для нападающих	для защитников
очки $\begin{pmatrix} 1 & 0,87 & 0,93 \\ \text{голы} & 1 & 0,68 \\ \text{передачи} & & 1 \end{pmatrix}$	очки $\begin{pmatrix} 1 & 0,53 & 0,95 \\ \text{голы} & 1 & 0,37 \\ \text{передачи} & & 1 \end{pmatrix}$

Из приведенных корреляционных матриц видно, что и у нападающих, и у защитников самая сильная прямая связь между очками и передачами, а у нападающих связь между очками и между голами и передачами значительно существенней, чем у защитников (сравни соответствующие коэффициенты корреляции 0,87 и 0,53, а также 0,68 и 0,37).

#### 4.4. Лучшие игроки чемпионата России

Представим результаты лучших игроков регулярного чемпионата и плей-офф и проранжируем их по числу набранных очков, голов и передач.

Таблица 29

№ п/п	Игрок	Очки			Голы			Передачи			Плей – офф				
		Очки	Ранг	Голы	Ранг	Голы	Ранг	Передачи	Ранг	Очки	Ранг	Голы	Ранг	Передачи	Ранг
1.	М.Сушинский ("Авангард")	55	1	18	7,5	37	1	12	1	5	1	7	1		
2.	А.Кайгородов (Магнитогорск)	49	2	15	9	34	2	2	6	0	8	2	3,5		
3.	А.Морозов ("Ак – барс")	47	3	20	4	27	4	0	8,5	0	8	0	8		
4.	Д.Затонский ("Авангард")	46	4	23	1	23	6	6	2	3	3	3	2		
5.	П.Роса ("Динамо")	44	5	21	3	23	6	5	3	3	3	2	3,5		
6.	А.Прокопьев ("Авангард")	43	6	14	10	29	3	3	4,5	3	3	0	8		
7.	И.Ковальчук ("Ак – барс")	42	7	19	5,5	23	6	0	8,5	0	8	0	8		

8.	А.Фролов (ЦСКА, "Динамо")	40	9	22	2	18	10	3	4,5	2	5	1	5
9.	Н.Жердев (ЦСКА)	40	9	19	5,5	21	8	0	8,5	0	8	0	8
10.	В.Антипов ("Локомотив")	40	9	18	7,5	22	9	0	8,5	0	8	0	8

4.4.1. Сравним результаты лучших игроков в регулярном чемпионате и в плей-офф.

Регулярный чемпионат

$$\begin{matrix} \text{очки} \\ \text{голы} \\ \text{передачи} \end{matrix} \begin{pmatrix} 1 & -0,13 & 0,89 \\ & 1 & -0,52 \\ & & 1 \end{pmatrix}$$

Плей-офф

$$\begin{matrix} \text{очки} \\ \text{голы} \\ \text{передачи} \end{matrix} \begin{pmatrix} 1 & 0,92 & 0,88 \\ & 1 & 0,72 \\ & & 1 \end{pmatrix}$$

И в регулярном чемпионате, и в плей-офф самая сильная связь между очками и передачами (0,89 и 0,88), которая почти такая же и у игроков «Локомотива». Удивительный результат получен для связи между очками и голами в регулярном чемпионате ( $r = -0,13$ ), что свидетельствует о независимости набранных очков и голов для лучших игроков, а в плей-офф для этих же игроков эта связь уже очень сильная ( $r = 0,92$ ). Заметим большую схожесть матрицы ранговых корреляций для нападающих «Локомотива» в регулярном чемпионате и лучших игроков чемпионата в плей-офф (чтобы это значило?).

4.4.2. Сравним соответствующие результаты лучших игроков в регулярном чемпионате и в плей-офф.

а) *Набранные очки (голы + передачи).*

Найдем сумму квадратов разности рангов, набранных очков в чемпионате и в плей-офф.

$$\sum d_i^2 = (2-6)^2 + (3-8,5)^2 + (4-2)^2 + (6-4,5)^2 + (7-8,5)^2 + (9-4,5)^2 + (9-8,5)^2 + (9-8,5)^2 = 77,5$$

Откуда  $r = 1 - \frac{6 \cdot 77,5}{9 \cdot 10 \cdot 11} \approx 0,53$  и связь прямая, хотя и средняя.

б) *Заброшенные шайбы.*

$$\sum d_i^2 = 134 \text{ и } r \approx 0,19 - \text{связь слабая.}$$

в) Передачи, с которых забрасывали шайбы.

$$\sum d_i^2 = 95,5 \text{ и } r \approx 0,42 - \text{связь прямая и средняя.}$$

Полученные результаты свидетельствуют о средней связи результатов лучших хоккеистов в регулярном чемпионате и в плей-офф, а по заброшенным шайбам эта связь и вовсе слабая, что должно бы насторожить тренеров команд.

#### 4.5. Итоги чемпионатов мира по хоккею

Проанализируем результаты последних пяти чемпионатов мира, которые приведены в следующей таблице:

Таблица 30

М	Команды	ЧМ-2005		ЧМ-2004		ЧМ-2003		ЧМ-2002		ЧМ-2001	
		место/ранг	место	ранг	место	ранг	место	ранг	место	ранг	
1	Канада	2	1	1	1	1	6	6	5	5	
2	Швеция	4	2	2	2	2	3	3	3	3	
3	Словакия	5	4	4	3	3	1	1	7	7	
4	Чехия	1	5	5	4	4	5	5	1	1	
5	Финляндия	8	6	6	5	5	4	4	2	2	
6	США	6	3	3	13	13	7	7	4	4	
7	РОССИЯ	3	10	10	7	7	2	2	6	6	
8	Германия	15	9	9	6	6	8	8	8	8	
9	Швейцария	7	8	8	8	8	10	10	9	9	
10	Латвия	9	7	7	9	9	11	11	13	12	
11	Австрия	16	11	11	10	10	12	12	11	11	
12	Украина	11	14	14	12	12	9	9	10	10	
13	Белоруссия	10	18	16	14	14	17	14	14	13	
14	Дания	14	12	12	11	11	18	15	22	16	
16	Словения	13	17	15	15	15	13	13	17	14	
17	Казахстан	12	13	13	17	16	21	16	21	15	

Из таблицы видно, что стабильность результатов у команд различна: так, сборная Швеции до последнего чемпионата была всегда в числе призеров, в то время как сборная России то поднималась до второго места, а то опускалась до десятого. Попытаемся определить стабильность всех 16 команд-финалистов последнего чемпионата мира за последние пять лет. С этой целью построим корреляционную матрицу, сравнивая результаты всех команд за два года.

В качестве иллюстрации найдем коэффициент ранговой корреляции последнего чемпионата и ЧМ-2004. Для этого найдем сумму квадратов разностей соответствующих мест команд в эти годы:

$$\sum a_i^2 = (2-1)^2 + (4-2)^2 + (5-4)^2 + (1-5)^2 + (8-6)^2 + (6-3)^2 + (3-10)^2 + (15-9)^2 + (7-8)^2 + (9-7)^2 + (16-11)^2 + (11-14)^2 + (10-16)^2 + (14-12)^2 + (13-15)^2 + (12-13)^2 = 204.$$

Используя формулу Спирмена, найдем коэффициент корреляции:

$$r_s = 1 - \frac{6 \cdot 204}{15 \cdot 16 \cdot 17} = 0,7.$$

Аналогично вычисляем другие коэффициенты ранговой корреляции и получаем корреляционную матрицу:

$$r_s = \begin{matrix} \text{ЧМ-01} \\ \text{ЧМ-02} \\ \text{ЧМ-03} \\ \text{ЧМ-04} \\ \text{ЧМ-05} \end{matrix} \begin{bmatrix} 1 & 0,839 & 0,747 & 0,780 & 0,744 \\ & 1 & 0,797 & 0,717 & 0,732 \\ & & 1 & 0,788 & 0,625 \\ & & & 1 & 0,7 \\ & & & & 1 \end{bmatrix}$$

Откуда видно, что наиболее тесная связь ( $r_s=0,839$ ) между результатами ЧМ-2001 и ЧМ-2002, а кроме этого во всех соседних чемпионатах за исключением двух последних ( $r_s=0,8$ ), что объясняется устойчивостью (всего-то на два сезона) состава команд и тренерских штабов национальных сборных. Особенность сравнения итогов двух последних чемпионатов мира объясняется прежде всего локаутом в НХЛ в 2004 - 2005 гг., а отсюда участием практически всех сильнейших хоккеистов в сборных России, Чехии, Словакии, с одной стороны, и отсутствием игровой практики многих профессионалов сборных США и Канады - с другой.

Обратим внимание на то, что при парной корреляции в данном случае возможно применение только коэффициента ранговой корреляции, поскольку чемпионаты мира проводятся по смешанной схеме с использованием и круговой системы, и игр с выбыванием.

#### 4.6. Анализ предсезонной подготовки

В течение предсезонной подготовки 2004 - 2005 года в команде «Локомотив» проводились контрольные измерения физической подготовки игроков. Показатели для оценки физической готовности игроков были разработаны финским тренером Хейкиле. Результаты представлены в таблице.

**Результаты тестов физической подготовки игроков  
(июнь 2004 г.)**

Таблица 31

Общая оценка	Фамилия игрока	Год рожд.	Гибкость		40 м бег		1-нога/лев		1-нога/прав		2х300 м бег		Вес		Рост	
3,154	Рязанцев А.	80	27	6	5,66	1	4,02	2	4,26	1	96,2	3	93	8	182	5
2,923	Жуков С.	75	20	4	5,89	1	4,6	1	4,68	1	98,7	2	86,4	6	192	8
4,200	Васильев А.	77	25	5									90,2	8	183	5
3,231	Красоткин Д.	71	20	4	5,82	1	4,66	1	4,65	1	97,2	2	91	8	182	5
5,000	Горохов И.	77	20	4	5,47	2	3,96	3	3,83	4	91,6	4	87	6	181	4
3,385	Бут А.	80	24	5	5,6	1	4,17	1	3,97	3	94,1	3	90	8	187	7
5,231	Непряев И.	82	24	5	5,33	3	3,63	6	3,71	5	91,2	4	86	6	188	8
3,462	Шафигулин Г.	85	9	1	5,38	3	3,81	4	3,82	4	97,8	2	91	8	190	8
2,875	Ткаченко И.	79	27	6	5,62	1	4,23	1	4,23	1	98,7	2	81,8	3	180	4
3,615	Швидкий Д.	80	28	6	5,45	2	4,29	1	4,29	1	91,8	4	94	8	183	5
4,462	Науров А.	85	25	5	5,18	5	3,57	7	3,94	3	90,7	5	85,8	8	182	5
4,000	Галимов А.	85	21	4	5,37	3	3,86	4	4,08	2	89,9	5	77,4	1	182	5
3,071	Самылин В.	72	23	5	5,71	1	4,28	1	4,4	1	98	2	95	8	185	6
2,769	Антипов В.	78	25	5	5,61	1	4,48	1	4,19	1	97,6	2	88,4	7	180	4
3,692	Власенков Д.	78	30	7	5,43	2	4,39	1	4,41	1	98,7	2	94,5	8	183	5
4,077	Крюков А.	82	24	5	5,77	1	4,21	1	4,04	2	91,8	4	90	8	193	8

Таблица 31 (продолжение)

Фамилия игрока	жир %	живот	сквод	скамья								30" сквод	30" бег			
				1	2	3	4	5	6	7	8					
Рязанцев А.	15,6	1	23	5	120	3	1,3	1	110	4	1,2	2	18	3	6150	2
Жуков С.	12	5	10	1	120	3	1,4	2	90	1	1,0	1	17	2	6170	2
Васильев А.	15,2	1	16	2					90	1	1,0	1				
Красоткин Д.	17	1	25	5	130	5	1,4	3	100	2	1,1	1	13	1	6650	5
Горохов И.	10	8	22	4	140	7	1,6	7	120	6	1,4	6	25	6	6820	6
Бут А.	9	4	20	4	120	3	1,3	1	95	1	1,1	1	17	2	6150	2
Непряев И.	6,8	4	25	5	130	5	1,5	5	105	3	1,2	3	26	7	6600	5
Шафигулин Г.	11,9	5	16	2	110	1	1,2	1	100	2	1,1	1	12	1	6550	5
Ткаченко И.	12	5														
Швидкий Д.	8	4	18	3	120	3	1,3	1	110	4	1,2	2	19	3	6830	6
Науров А.	8,6	4	30	7	120	3	1,4	2					17	2	6600	5
Галимов А.	5	4	23	5	120	3	1,6	6	95	1	1,2	3	17	2	7200	8
Самылин В.	17	1	16	2	130	5	1,4	2	110	4	1,2	2	21	4	6150	2
Антипов В.	12,9	3	19	3	100	1	1,1	1	90	1	1,0	1	12	1	6700	6
Власенков Д.	12	5	18	3	130	5	1,4	2	105	3	1,1	1	22	5	6150	2
Крюков А.	11	7	20	4	110	1	1,2	1	95	1	1,1	1	23	5	6750	6

За каждый вид выставлялась оценка по 8 балльной системе. Далее вычислялся общий коэффициент физической подготовки хоккеистов как среднее арифметическое спортивных показателей.

Определим связь между полученными коэффициентами физической подготовки и количеством проведенных хоккеистами игр в чемпионате. С этой целью проранжируем показатели и вычислим коэффициент ранговой корреляции Спирмена (см. табл. ).

Таблица 31 (продолжение)

Фамилия игрока	Общий коэффициент физической подготовки	Ранг	Количество игр	Ранг
Рязанцев А.	3,154	12	53	9
Жуков С.	2,923	14	54	8
Васильев А.	4,200	4	60	2,5
Красоткин Д.	3,231	11	60	2,5
Горохов И.	5,000	2	52	10,5
Бут А.	3,385	10	60	2,5
Непряев И.	5,231	1	56	7
Шафигулин Г.	3,462	9	46	13
Ткаченко И.	2,875	15	59	5
Швидкий Д.	3,615	8	52	10,5
Науров А.	4,462	3	11	16
Галимов А.	4,000	6	41	14
Самылин В.	3,077	13	39	15
Антипов В.	2,769	16	58	6
Власенков Д.	3,692	7	51	12
Крюков А.	4,077	5	60	2,5

Предварительно найдем сумму квадратов разности, соответствующих количеству игр, проведенных игроками, и общим коэффициентом физической подготовки.

$$\sum d_i^2 = (9-12)^2 + (8-14)^2 + (2,5-4)^2 + (2,5-11)^2 + (10,5-2)^2 + (2,5-10)^2 + (7-1)^2 + (13-9)^2 + (5-15)^2 + (10,5-8)^2 + (14-6)^2 + (15-13)^2 + (6-16)^2 + (12-7)^2 + (2,5-5)^2 = 758,5$$

Используя формулу Спирмена, найдем коэффициент корреляции

$$r_s = 1 - \frac{6 \cdot 758,5}{15 \cdot 16 \cdot 17} = -0,11$$

который получился по модулю близким к нулю. Откуда видно, что количество сыгранных игр в регулярном чемпионате практически не зависит от физической подготовки игроков. Можно предположить, что на количество сыгранных игр также влияют такие показатели, как технико-тактические навыки игроков, игровое мышление на площадке, травмы, полученные в результате игр и тренировок, амплуа игроков, селекционная работа в дозаявочный период, стратегические задачи тренерского состава.

Проанализируем зависимость количества проведенных игр в регулярном чемпионате, оценки физической подготовки, очков, набран-



ных в регулярном чемпионате и полезности игроков. Составляем таблицу и проставляем ранги по всем этим показателям (см. табл.)

Таблица 32

Фамилия игрока	Общий коэффициент физической подготовки	Ранг	Количество игр	Ранг	Количество набранных очков	Ранг	Полезность	Ранг
Рязанцев А.	3,154	12	53	9	19	5,5	+19	2
Жуков С.	2,923	14	54	8	4	14,5	+5	13
Васильев А.	4,200	4	60	2,5	17	9	+10	8,5
Красоткин Д.	3,231	11	60	2,5	17	9	+21	1
Горохов И.	5,000	2	52	10,5	16	11	+15	3,5
Бут А.	3,385	10	60	2,5	34	2	+10	8,5
Непряев И.	5,231	1	56	7	20	4	+7	11,5
Шафигулин Г.	3,462	9	46	13	11	12	+7	11,5
Ткаченко И.	2,875	15	59	5	30	3	+9	10
Швидкий Д.	3,615	8	52	10,5	18	7	+11	6,5
Науров А.	4,462	3	11	16	4	14,5	+2	14
Галимов А.	4,000	6	41	14	2	16	-3	15
Самылин В.	3,077	13	39	15	6	13	-4	16
Антипов В.	2,769	16	58	6	40	1	+15	3,5
Власенков Д.	3,692	7	51	12	19	5,5	+14	5
Крюков А.	4,077	5	60	2,5	17	9	+11	6,5

Опуская расчеты коэффициентов ранговой корреляции, запишем полученные коэффициенты в корреляционную матрицу, где И – количество игр, проведенных в чемпионате, Ф – оценка физической подготовки, О – количество очков, набранное в регулярном чемпионате, П – полезность игроков по системе «гол+пасс»

$$r_s = \begin{matrix} И \\ Ф \\ О \\ П \end{matrix} \begin{bmatrix} 1 & -0,11 & 0,58 & 0,51 \\ & 1 & -0,27 & -0,08 \\ & & 1 & 0,55 \\ & & & 1 \end{bmatrix}$$

Откуда видно, что наиболее тесная связь ( $r_s=0,58$ ) между количеством набранных очков и количеством игр и ( $r_s=0,55$ ) между полезностью и набранными очками. Т.е. чем больше проведено игр, тем выше вероятность набрать большее количество очков. И чем больше набрано очков, тем выше полезность игрока.

#### 4.7. А судья кто?

Специальный выпуск журнала «Локо - вперед!» № 25 (2005/2006 гг.) предлагает информацию о главных арбитрах, которые обслуживали матчи с участием «Локомотива» в регулярном чемпионате сезона 2005 – 2006 гг.

В колонках, следующих за фамилией рефери, последовательно отражено количество игр «Локомотива», на которых он работал в роли главного арбитра, количество побед, ничьих и поражений нашей команды, соотношение штрафных минут, выписанных данным судьей игрокам «Локомотива» и их соперникам, и, наконец, количество очков, набранных «Локомотивом» в матчах, обслуживаемых этим судьей. Добавим к этой традиционной статистике разность штрафных минут и проанализируем ее (по возрастанию разности, что соответствует ухудшению признака). Проранжируем также набранные очки и добавим в таблицу квадраты разностей  $d_i^2$  соответствующих рангов.

Таблица 33

Судья	И	В	Н	П	Штраф	Разность	Ранг	Очки	Ранг	$d_i^2$
Кадыров (Уфа)	5	3	1	1	76–76	0	9	10	1	64
Поляков (Москва)	4	3	0	1	42–60	-18	1	9	2,5	2,25
Раводин (Москва)	3	3	0	0	69–80	-11	3	9	2,5	0,25
Буланов (Москва)	4	2	2	0	102–75	27	17	8	4	169
Захаров (Москва)	3	2	1	0	83–83	-1	8	7	7,5	0,25
Цыплаков (С.-Петербург)	3	2	1	0	48–58	-10	4	7	7,5	12,25
Карабанов (Москва)	3	2	1	0	43–22	21	16	7	7,5	72,25
Горский (Саратов)	4	2	1	1	62–60	2	10,5	7	7,5	9
Кулаков (Тверь)	4	2	1	1	56–60	-4	5,5	7	7,5	4
Черенков (Москва)	5	2	0	3	82–80	2	10,5	7	7,5	9
Гусев (Серов)	2	1	1	0	46–40	6	12,5	4	11,5	1
Семенов (Москва)	2	1	0	1	34–28	6	12,5	4	11,5	1
Рантала (Финляндия)	1	1	0	0	20–34	-14	2	3	13,5	132,25
Бирюков (Москва)	1	1	0	0	14–6	8	14	3	13,5	0,25
Капралов (Москва)	1	0	0	1	22–8	14	15	0	16	1
Минар (Чехия)	1	0	0	1	33–37	-4	5,5	0	16	110,25
Антропов (Екатеринбург)	2	0	0	2	16–18	-2	7	0	16	81

$\Sigma=48$

Найдем коэффициент ранговой корреляции набранных «Локомотивом» очков и разностью штрафных минут в матчах, обслуживаемых этими судьями.

$$r_s = 1 - \frac{6 \cdot 669}{16 \cdot 17 \cdot 18} \approx 0,18$$

По полученному коэффициенту корреляции можно сделать вывод о том, что набранные «Локомотивом» очки практически не зависят ( $r_s$ ,

близок к нулю) от «выданных» командам штрафных минут, т.е. от отношения судей к нашей команде.

Интересно было бы изучить аналогичный вопрос об «удобности» судей и к другим командам суперлиги, что читатель сможет аналогично проделать сам при наличии традиционной статистики по другим командам.

## §5. Корреляционные матрицы по футболу

В предыдущем параграфе мы уже использовали корреляционные матрицы (наряду с коэффициентами ранговой корреляции) для анализа хоккейных турниров. Поскольку по итогам спортивных турниров приводится не только основной результат каждой команды, но и несколько других показателей, составляющих успеха или неудачи, то нам кажется разумным использовать корреляционные матрицы. По элементам такой матрицы можно судить о степени связи тех или иных показателей с общим результатом и их парной связи. Анализ полученных корреляционных матриц может дать много пищи для размышлений специалистам, менеджерам и самим спортсменам.

### 5.1. Стабильность чемпионатов России по футболу

Рассмотрим итоги чемпионата России по футболу в 2005 году. Для удобства вычисления ранговых коэффициентов корреляции проанализируем результаты, показанные командами дома и в гостях, с командами из верхней (с лучшими) и нижней (с худшими) половины турнирной таблицы.

#### Итоговая таблица чемпионата России по футболу 2005 г. Премьер-лига

Таблица 34\*

Место (1)	Команды (2)	Дома (3)	В гостях (4)	С лучшими (5)	С худшими (6)	Бюджет	
						млн. \$ (7)	ранг (8)
1	ЦСКА	1	3	1	1	35	3,5
2	"Спартак"	4,5	1,5	6	2	50	1
3	"Локомотив"	4,5	1,5	3	4	35	3,5
4	"Рубин"	2	5,5	2	10	10	11,5
5	"Москва"	6	4	4	7	20	6,5
6	"Зенит"	3	7	5	6	20	6,5
7	"Торпедо"	7,5	5,5	10	3	10	11,5
8	"Динамо"	11	8	15	5	45	2
9	"Шинник"	7,5	10	9	9	10	11,5

10	"Томь"	10	9	8	11	8	15
11	"Сатурн"	10	11,5	12	12	22,5	5
12	"Амкар"	10	11,5	7	14	8	15
13	"Ростов"	13	13	16	8	10	11,5
14	"Крылья Советов"	10	15	11	13	12,5	9
15	"Алания"	15	16	13	15	8	15
16	"Терек"	16	14	14	16	17,5	8

\* Бюджеты клубов взяты по версии «Комсомольской правды», 8-15 дек.2005, с.34.

О стабильности выступления команд в последнем чемпионате будем судить по вычисленной из таблицы № 37 матрице ранговых корреляций между общим количеством очков ( $O$ ) и очков, набранных в домашних ( $D$ ) и гостевых ( $G$ ) играх, в играх с лучшими ( $L$ ) и худшими ( $X$ ) командами и с учетом бюджетов клубов ( $B$ ):

$$(r_s) = \begin{matrix} O \\ D \\ G \\ L \\ X \\ B \end{matrix} \begin{pmatrix} 1 & 0,92 & 0,97 & 0,8 & 0,86 & 0,57 \\ & 1 & 0,84 & 0,89 & 0,73 & 0,46 \\ & & 1 & 0,77 & 0,83 & 0,58 \\ & & & 1 & 0,45 & 0,21 \\ & & & & 1 & 0,62 \\ & & & & & 1 \end{pmatrix}$$

Полученная подматрица (сравнения чисто спортивных показателей) очень похожа по результатам на аналогичную матрицу ранговых корреляций в хоккее. Так, самый большой коэффициент корреляции ( $r = 0,97$ ) получен между общим количеством очков и гостевыми играми, а в хоккее он почти такой же ( $r = 0,96$ ) и немного уступает коэффициенту между общими результатами и домашними играми ( $r = 0,97$ ).

Также можно видеть, что наименьшим коэффициентам корреляции соответствует соотношение результатов команд с лучшими и худшими командами. Причем в футболе этот коэффициент равен только 0,45 и соответствует слабой связи, что можно объяснить прежде всего психологическими факторами.

Обратим внимание и на то, что ни один спортивный показатель не коррелирует сильно с бюджетами клубов (см. последний столбец матрицы). Этот факт, с одной стороны, должен насторожить руководителей команд и особенно таких, как «Динамо», «Терек», «Сатурн», а с другой – убедить владельцев клубов отказаться от безудержной гонки по увеличению ежегодных бюджетов клубов. Да и Федерация футбола России, как, впрочем, и хоккея тоже, на наш взгляд, должна внести ограничения на величину бюджетов. Заметим еще, что размер финансирования команд слабее всего коррелирует с результатами домашних игр ( $r = 0,46$ )

и с показателями, показанными против команд из верхней половины турнирной таблицы ( $r = 0,21$ , и только!).

Аналогично найдем матрицы ранговых корреляций по итогам чемпионатов России в предыдущие два года.

**Итоговая таблица чемпионата России по футболу 2004 г.  
Премьер-лига**

Таблица 35

Место	Команды	С лучшими			С худшими		
		Кол-во игр	Очки	Ранг	Кол-во игр	Очки	Ранг
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	"Локомотив"	14	32	1	16	29	7
2	ЦСКА	14	23	2,5	16	37	1
3	"Зенит"	14	23	2,5	16	33	4
4	"Крылья Советов"	14	21	4	16	35	3
5	"Торпедо"	14	18	6,5	16	36	2
6	"Шинник"	14	18	6,5	16	26	8
7	"Сатурн"	14	11	9	16	30	6
8	"Спартак"	14	9	13	16	31	5
9	ФК "Москва"	16	20	5	14	20	9,5
10	"Рубин"	16	14	8	14	19	12
11	"Амкар"	16	12	10	14	18	14,5
12	"Ростов"	16	11	11,5	14	18	14,5
13	"Динамо"	16	10	14	14	19	12
14	"Алания"	16	8	16	14	20	9,5
15	"Кубань"	16	9	15	14	19	12
16	"Ротор"	16	11	11,5	14	11	16

$$(r_{\rho}) = \begin{matrix} L \\ X \end{matrix} \begin{bmatrix} 1 & 0,92 & 0,83 \\ & 1 & 0,58 \\ & & 1 \end{bmatrix}$$

**Итоговая таблица чемпионата России по футболу 2003 г.  
Премьер-лига**

Таблица 36

Место	Команды	С лучшими			С худшими		
		Кол-во игр	Очки	Ранг	Кол-во игр	Очки	Ранг
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	ЦСКА	14	21	2,5	16	38	1,5
2	"Зенит"	14	27	1	16	29	6
3	"Рубин"	14	20	5	16	33	4
4	"Локомотив"	14	14	7	16	38	1,5
5	"Шинник"	14	19	6	16	28	7
6	"Динамо"	14	13	9	16	33	4
7	"Сатурн"	14	12	10	16	33	4
8	"Торпедо"	14	21	2,5	16	22	13

9	"Крылья Советов"	16	22	3,4	14	20	11
10	"Спартак"	16	14	8	14	22	8,5
11	"Ростов"	16	12	12	14	22	8,5
12	"Ротор"	16	12	12	14	20	11
13	"Алания"	16	12	12	14	19	14,5
14	ФК "Москва"	16	10	15	14	19	14,5
15	"Уралан"	16	8	16	14	20	11
16	"Черноморец"	16	11	14	14	13	16

$$(r_{xy}) = \begin{matrix} O \\ L \\ X \end{matrix} \begin{bmatrix} 1 & 0,85 & 0,86 \\ & 1 & 0,48 \\ & & 1 \end{bmatrix}$$

Полученные корреляционные матрицы обладают одной закономерностью – незначительным отличием сильной прямой связи между общим результатом и домашними или гостевыми показателями и слабой зависимостью результатов дома и в гостях, что, впрочем, не является неожиданным.

Рассмотрим результаты первенства России по футболу в чемпионате рангом ниже премьер-лиги – в первом дивизионе. Здесь есть свои особенности как в организации футбольного хозяйства, так и в статистической обработке результатов, поскольку обмен происходит и с высшими и с низшими классами.

**Таблица первенства России по футболу в первом дивизионе**

Таблица 37

2005 г.	Команды	2004г.		Домашние игры		Гостевые игры	
		Место	Ранг	Очки	Ранг	Очки	Ранг
1	"Луч-Энергия"	14	13	55	1	49	2
2	"Спартак" НЛ	12	11	54	2	48	4
3	КАМАЗ	4	3	49	4	45	3
4	"Химки"	5	4	44	7	44	1
5	"Кубань"	-	1	52	3	38	5
6	"Динамо" Мх	11	10	49	5	32	6
7	"Урал"	-	19	47	6	33	7
8	"Орел"	6	5	42	9	36	11
9	"Спартак" Чл	-	19	42	8	31	13
10	"Чкаловец-1936"	-	19	33	15	35	8
11	"Анжи"	8	7	36	13	17	14

12	"СКА-Энергия"	7	6	32	16	31	10
13	"Динамо" Бр	15	14	37	11	30	17
14	"Волгарь-Газпром"	-	19	41	10	20	18
15	"Локомотив"	10	9	34	14	26	15
16	"Авангард"	-	19	28	17	26	12
17	"Факел"	-	12	37	12	15	19
18	"Металлург-Кузбасс"	16	15	22	20	29	9
19	"Амур"	-	19	21	21	13	16
20	"Сокол"	3	2	27	18	11	21
21	"Металлург"	9	8	23	19	21	22
22	"Петротрест"	-	19	17	22	13	20

Заметим, что при ранжировании результатов 2004 года ранг номер один присвоен «Кубани», выбывшей из премьер-лиги вместе с прекратившим существование «Ротором». Семи командам, пополнившим дивизион из второго дивизиона, присвоен ранг, равный среднему арифметическому номеров с 16 по 22, под которыми они и вошли в более высокий класс.

Определим коэффициент ранговой корреляции результатов первенства России по футболу 2004 и 2005 годов в первом дивизионе:

$$\sum d_i^2 = 12^2 + 9^2 + 16 + 16 + 12^2 + 9 + 81 + 64 + 16 + 36 + 1 + 25 + 36 + 9 + 25 + 9 + 18^2 + 13^2 + 9 = 1214$$

$$r_s = 1 - \frac{6 \cdot 1214}{21 \cdot 22 \cdot 23} = 0,315 \text{ и получается, что существует слабая}$$

прямая связь между результатами первенств по футболу в 2004 и 2005 году в первом дивизионе.

Аналогично вычисляем все другие коэффициенты парных ранговых корреляций футбольных результатов в первом дивизионе в 2005 году, которые занесем в корреляционную матрицу.

$$(r_s) = \begin{matrix} O \\ D \\ \Gamma \end{matrix} \begin{bmatrix} 1 & 0,93 & 0,89 \\ & 1 & 0,71 \\ & & 1 \end{bmatrix}$$

Откуда видно, что наиболее согласованы общий итог первенства в первой лиге и результаты домашних игр ( $r = 0,93$ ), а менее – результаты, показанные в домашних и гостевых играх ( $r = 0,71$ ), хотя этот коэффициент корреляции значительно больше, чем для первенства в премьер-лиге.

## 5.2. Чемпионаты ведущих европейских стран

Рассмотрим итоги последних пяти чемпионатов ведущих европейских стран и России, которая в очередной раз не попала на чемпионат мира и поэтому не является ведущей футбольной державой, к сожалению. Изложим результаты кратко и единообразно. Начинаем с таблицы проранжированных результатов последних чемпионатов своих стран, по которой находим все парные коэффициенты ранговой корреляции Спирмена. Полученные данные заносим в корреляционную матрицу, а технические выкладки опустим.

### Германия

Таблица 38

№	Команды	Ранг				
		2005	2004	2003	2002	2001
1	"Бавария"	1	2	1	3	1
2	"Шальке"	2	7	7	5	2
3	"Вердер"	3	1	6	6	7
4	"Герта"	4	12	5	4	5
5	"Штутгарт"	5	4	2	8	15
6	"Байер"	6	3	15	2	4
7	"Боруссия Д"	7	6	3	1	3
8	"Гамбург"	8	8	4	11	13
9	"Вольфсбург"	9	10	8	10	9
10	"Ганновер"	10	14	11	21,5	21,5
11	"Майнц"	11	21,5	21,5	21,5	21,5
12	"Кайзерслаутерн"	12	15	14	7	8
13	"Арминия"	13	21,5	16	21,5	21,5
14	"Нюрнберг"	14	21,5	17	15	21,5
15	"Боруссия М"	15	11	12	12	21,5
16	"Бохум"	16	5	9	21,5	18
17	"Ганза"	17	9	13	14	12
18	"Фрайбург"	18	13	21,5	16	6
19	"Айнтрахт"	21,5	16	21,5	21,5	17
20	"Мюнхен 1860"	21,5	17	10	9	11
21	"Кельн"	21,5	18	21,5	17	10
22	"Коттбус"	21,5	21,5	18	13	14
23	"Ст.Паули"	21,5	21,5	21,5	18	21,5
24	"Унтерхакинг"	21,5	21,5	21,5	21,5	16

$$(r_s) = \begin{matrix} 2001 \\ 2002 \\ 2003 \\ 2004 \\ 2005 \end{matrix} \begin{pmatrix} 1 & 0,807 & 0,475 & 0,610 & 0,516 \\ & 1 & 0,691 & 0,655 & 0,675 \\ & & 1 & 0,773 & 0,773 \\ & & & 1 & 0,714 \\ & & & & 1 \end{pmatrix}$$



## Италия

Таблица 39

№	Команды	Ранг				
		2005	2004	2003	2002	2001
1	"Ювентус"	1	3	1	1	2
2	"Милан"	2	1	3	4	6
3	"Интер"	3	4	2	3	5
4	"Удинезе"	4	7	6	14	12
5	"Сампdoria"	5	8	25,5	25,5	25,5
6	"Палермо"	6	25,5	25,5	25,5	25,5
7	"Мессина"	7	25,5	25,5	25,5	25,5
8	"Рома"	8	2	8	2	1
9	"Ливорно"	9	25,5	25,5	25,5	25,5
10	"Лацио"	10	6	4	6	3
11	"Лечче"	11	10	25,5	16	13
12	"Кальяри"	12	25,5	25,5	25,5	25,5
13	"Реджина"	13	14	14	25,5	14
14	"Сiena"	14	13	25,5	25,5	25,5
15	"Кьево"	15	9	7	5	25,5
16	"Болонья"	16	12	11	7	10
17	"Фиорентина"	17	25,5	25,5	17	9
18	"Парма"	18	5	5	10	4
19	"Брешиа"	19	11	9	13	8
20	"Аталанта"	20	25,5	13	4	7
21	"Перуджа"	26,5	15	10	8	11
22	"Модена"	26,5	16	15	25,5	25,5
23	"Эмполи"	26,5	17	12	25,5	25,5
24	"Анкона"	26,5	18	25,5	25,5	25,5
25	"Пьяченца"	26,5	25,5	16	12	25,5
26	"Комо"	26,5	25,5	17	25,5	25,5
27	"Торино"	26,5	25,5	18	11	25,5
28	"Верона"	26,5	25,5	25,5	15	15
29	"Венеция"	26,5	25,5	25,5	18	25,5
30	"Венеце"	26,5	25,5	25,5	25,5	16
31	"Наполи"	26,5	25,5	25,5	25,5	17
32	"Барии"	26,5	25,5	25,5	25,5	18

$$(r_s) = \begin{matrix} 2001 \\ 2002 \\ 2003 \\ 2004 \\ 2005 \end{matrix} \begin{pmatrix} 1 & 0,743 & 0,678 & 0,634 & 0,401 \\ & 1 & 0,791 & 0,611 & 0,339 \\ & & 1 & 0,771 & 0,385 \\ & & & 1 & 0,623 \\ & & & & 1 \end{pmatrix}$$

## Испания

Таблица 40

№	Команды	Ранг				
		2005	2004	2003	2002	2001
1	"Барселона"	1	2	6	4	4
2	"Реал М"	2	4	1	3	1
3	"Вильярреал"	3	8	15	15	7
4	"Бетис"	4	9	8	6	25
5	"Эспаньол"	5	16	17	14	9
6	"Севилья"	6	6	10	8	25
7	"Валенсия"	7	1	5	1	5
8	"Атлетик Б"	8	5	7	10	14
9	"Депортиво"	9	3	3	2	2
10	"Малага"	10	10	14	9	8
11	"Атлетико М"	11	7	11	25	25
12	"Сарагоса"	12	13	25	20	15
13	"Хетафе"	13	25	25	25	25
14	"Реал С"	14	15	2	13	13
15	"Осасуна"	15	12	12	17	17
16	"Расинг"	16	17	16	25	19
17	"Мальорка"	17	11	9	16	3
18	"Леванте"	18	25	25	25	25
19	"Нумансия"	19	25	25	25	20
20	"Альбасете"	20	14	25	25	25
21	"Вильядолид"	25	18	13	12	16
22	"Сельта"	25	19	4	5	6
23	"Мурсия"	25	20	25	25	25
24	"Уэльва"	25	25	18	25	25
25	"Алавес"	25	25	19	7	10
26	"Р.Вальекано"	25	25	20	11	12
27	"Лас Пальмас"	25	25	25	18	11
28	"Тенерифе"	25	25	25	19	25
29	"Овьедо"	25	25	25	25	18

$$(r_s) = \begin{matrix} 2001 \\ 2002 \\ 2003 \\ 2004 \\ 2005 \end{matrix} \begin{pmatrix} 1 & 0,706 & 0,582 & 0,461 & 0,347 \\ & 1 & 0,770 & 0,613 & 0,481 \\ & & 1 & 0,762 & 0,572 \\ & & & 1 & 0,829 \\ & & & & 1 \end{pmatrix}$$

Россия

Таблица 41

№	Команды	Ранг				
		2004	2003	2002	2001	2000
1	"Локомотив"М	1	4	2	2	2
2	ЦСКА	2	1	1	7	8
3	"Зенит"	3	2	10	3	7
4	"Крылья Советов"	4	9	5	5	14
5	"Торпедо"	5	8	4	4	3
6	"Шинник"	6	5	6	19,5	19,5
7	"Сатурн"	7	7	7	6	9
8	"Спартак"	8	10	3	1	1
9	ФК "Москва"	9	14	14	14	17
10	"Рубин"	10	3	19,5	19,5	19,5
11	"Амкар"	11	19,5	19,5	19,5	19,5
12	"Ростов"	12	11	11	11	12
13	"Динамо" М	13	6	8	9	5
14	"Алания"	14	13	12	12	10
15	"Кубань"	15	19,5	19,5	19,5	19,5
16	"Ротор"	16	12	9	10	11
17	"Сокол"	19,5	19,5	16	8	19,5
18	"Анжи"	19,5	19,5	15	13	4
19	"Факел"	19,5	19,5	19,5	15	13
20	"Локомотив" НН	19,5	19,5	19,5	19,5	15
21	"Уралан"	19,5	13	13	19,5	16
22	"Черноморец"	19,5	19,5	19,5	16	6

$$(r_s) = \begin{matrix} 2000 \\ 2001 \\ 2002 \\ 2003 \\ 2004 \end{matrix} \begin{pmatrix} 1 & 0,724 & 0,580 & 0,3671 & 0,299 \\ & 1 & 0,765 & 0,511 & 0,587 \\ & & 1 & 0,726 & 0,752 \\ & & & 1 & 0,822 \\ & & & & 1 \end{pmatrix}$$

### 5.3. Мировые футбольные державы

Изучим зависимость двух спортивных показателей (рейтинг ФИФА и количество клубов) и двух экономических (количество населения и ВВП на душу населения) для 32 стран, завоевавших путевки на мундиаль 2006 г.

#### Финалисты чемпионата мира по футболу 2006 г.

Таблица 42\*

Группа	Команды	Рейтинг ФИФА		Кол-во клубов		Население		ВВП на душу населения	
			ранг	тыс.	ранг	млн.	ранг	тыс.дол.	ранг
A	1. Германия	16	14	26,7	3	82,4	5	26,6	6
	2. Коста-Рика	21	17	0,128	31	3,9	31	8,5	21
	3. Польша	23	18	7,76	8	38,6	14	9,51	17
	4. Эквадор	37	24	0,17	28	13,7	20	3,1	27
B	5. Англия	9	9	42,0	1	50,1	9	25,3	9
	6. Парагвай	30	21	1,5	17	6,0	28	4,2	26
	7. Тринидад Тобаго	51	30	0,135	30	1,1	32	9,5	18
	8. Швеция	14	12	3,23	11	8,9	26	25,4	8
C	9. Аргентина	4	4	3,06	12	38,7	13	10,2	16
	10. Кот-Д'Ивуар	41	26	0,2	27	17,0	18	1,5	31,5
	11. Сербия и Черногория	47	27	2,82	13	10,6	22	2,37	28
	12. Голландия	3	3	4,05	10	16,1	19	26,9	5
D	13. Мексика	7	7	1,49	18	104,9	4	9,0	19
	14. Иран	19	15	2,54	14	68,3	6	7,0	23
	15. Ангола	93	32	0,1	32	10,8	21	2,1	29,5
	16. Португалия	10	10	2,53	15	10,1	24	18,0	13
E	17. Италия	12	11	16,13	6	58,1	8	25,0	10
	18. Гана	50	29	0,25	25	20,5	16	2,1	29,5
	19. США	8	8	1,69	16	290,3	1	37,6	1
	20. Чехия	2	2	4,17	9	10,2	23	15,3	14
F	21. Бразилия	1	1	12,9	7	182,0	2	7,6	22
	22. Хорватия	20	15	1,19	20	4,4	30	8,8	20
	23. Австралия	49	28	0,25	25	19,7	17	27,0	4
	24. Япония	15	13	19,1	5	127,1	3	28,0	3
G	25. Франция	5	5	19,8	4	60,2	7	25,7	7
	26. Швейцария	36	23	1,45	19	7,3	27	31,7	2
	27. Южная Корея	29	20	0,66	22	48,3	10	19,4	12
	28. Того	56	31	0,25	25	5,1	29	1,5	31,5
H	29. Испания	6	6	33,6	2	40,2	12	20,7	11
	30. Украина	40	25	1,09	21	48,0	11	4,5	25
	31. Тунис	28	19	0,55	23	9,9	25	6,5	24
	32. Саудовская Аравия	32	22	0,153	29	24,3	15	10,5	15

\*Показатели таблицы взяты и обработаны из ежемесячной газеты «Футбол», №49 (80), 2005.-с.12-20.

Проранжируем для стран-финалистов чемпионата мира рейтинг ФИФА ( $\Phi$ ), количество клубов ( $K$ ), население ( $H$ ) и валовый внутренний продукт на душу населения ( $B$ ) и, вычислив предварительно суммы квадратов разностей соответствующих рангов, найдем искомую матрицу ранговых корреляций.

$$(r_s) = \begin{matrix} \Phi \\ K \\ H \\ B \end{matrix} \begin{pmatrix} 1 & 0,74 & 0,46 & 0,63 \\ & 1 & 0,52 & 0,5 \\ & & 1 & 0,71 \\ & & & 1 \end{pmatrix}$$

По матрице видно, что существует сильная прямая связь между рейтингом ФИФА и количеством клубов ( $r = 0,74$ ), а также между населением и ВВП на душу населения ( $r = 0,71$ ). Остальные пары имеют средний силы корреляции, особенно удивительно это для двух пар – между рейтингом ФИФА и количеством клубов ( $r = 0,46$ ) и между численностью населения и количеством футбольных клубов в стране ( $r = 0,5$ ).

Полученные данные для финалистов ЧМ – 2006 приводятся впервые и могут быть полезны футбольным специалистам как до чемпионата, так и после проведения с учетом его результатов и дополнительной статистики обработки.

Вызывает интерес сравнение полученных данных по финалистам ЧМ-2006 и коэффициентов УЕФА, по результатам выступления клубов в еврокубках («Советский спорт» 12.12.05, с.2).

### Результаты участия в еврокубках в 2005 г.

Таблица 43

№	Страна	Коэффициент УЕФА		Кол-во клубов			Рейтинг ФИФА		Население		ВВП на душу населения	
			ранг			ранг		ранг	млн.	ранг	тыс.\$	ранг
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	Испания	66,75	1	7	6		6	4	40,2	8	20,7	8
2	Италия	62,16	2	7	6		12	8	48,1	5	25,0	7
3	Англия	59,92	3	7	5		9	5	50,1	6	25,3	5
4	Франция	48,03	4	8	6		5	3	60,2	4	25,7	4
5	Германия	45,74	5	8	6		16	9	82,4	2	26,6	3
6	Португалия	43,04	6	6	1		10	6	10,1	14	18,0	9
7	Голландия	40,49	7	6	4		3	2	16,1	10	26,9	2
8	Греция	32,08	8	6	0		17	10	10,6	11	8,3	12
9	Бельгия	30,25	9	4	1		52	14	10,3	12	29,2	1
10	Шотландия	29,88	10	4	1		61	15	5,1	15	25,2	6
11	Россия	29,58	11	4	2		34	12	144,5	1	9,8	11
12	Украина	26,35	12	4	1		40	13	48,0	7	4,5	15
13	Турция	26,17	13	4	0		11	7	68,1	3	5,5	14
14	Чехия	26,08	14	4	1		2	1	10,2	13	15,3	10
15	Румыния	24,79	15	3	2		27	11	22,2	9	5,8	13

Найдем матрицу ранговых корреляций двух показателей по результатам еврокубков – коэффициент УЕФА ( $У$ ) и количества клубов, участвующих в еврокубках ( $К$ ), а также рейтинга ФИФА ( $\Phi$ ), населения ( $Н$ ) и ВВП на душу населения ( $В$ ).

$$(r_s) = \begin{matrix} У \\ К \\ \Phi \\ Н \\ В \end{matrix} \begin{pmatrix} 1 & 0,65 & 0,34 & 0,41 & 0,57 \\ & 1 & 0,29 & 0,47 & 0,48 \\ & & 1 & 0,05 & 0,19 \\ & & & 1 & -0,09 \\ & & & & 1 \end{pmatrix}$$

По полученной матрице видно, что почти поровну присутствует слабая и средняя связь показателей. Как и следовало ожидать, сильнее всего коррелирует ( $r=0,65$ ) коэффициент УЕФА и количество клубов – участников еврокубков. Обратим внимание на слабую связь ( $r = 0,34$ ) между коэффициентами УЕФА и рейтингом ФИФА, что свидетельствует о малой корреляции достижений сборных своих стран и результатами, показанными их представителями в европейских кубках.

### §6. Конкорданция и стабильность чемпионатов европейских стран

Для определения зависимости результатов двух первенств использовали коэффициент ранговой корреляции, а нескольких – корреляционные матрицы. В предыдущем параграфе рассматривали итоги последних чемпионатов ведущих европейских футбольных стран и России. Для полноты картины и определения степени зависимости результатов по совокупности всех пяти чемпионатов будем использовать **множественный коэффициент ранговой корреляции**, или, иначе говоря, **коэффициент конкорданции**:

$$W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1)},$$

где  $S$  – сумма квадратов отклонений суммы  $m$  рангов от их средней величины,

$m$  – число ранжируемых признаков,

$n$  – число наблюдений (количество команд).

Эта формула применяется в тех случаях, когда ранги по каждому признаку не повторяются. Если имеются связанные ранги, то коэффициент конкорданции рассчитывается с учетом числа таких повторяющихся (связанных) рангов по каждому фактору:

$$W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1) - m \sum_1^m (t^3 - t)},$$

где  $t$  – число одинаковых рангов по каждому признаку.

В обозначенных формулах главное вычисление сводится к нахождению  $S$ , которую можно найти двумя способами. Обозначив через  $R_{ij}$  ранг  $i$ -ого фактора у  $j$ -й единицы, ранжируем каждый из  $m$ -факторов, а затем находим сумму рангов по каждой строке и общую сумму.

$$\text{Тогда } S = \sum_1^n \left( \sum_1^m R_{ij} \right)^2 - \frac{\left( \sum_1^n \sum_1^m R_{ij} \right)^2}{n} \quad \text{или}$$

$$S = \sum_1^n \left( \sum_1^m R_{ij} - T \right)^2, \quad \text{где } T = \frac{\sum_1^n \sum_1^m R_{ij}}{n} - \text{средняя величина суммы рангов.}$$

гов.

В следующих вычислениях коэффициента конкордации  $W$  мы будем использовать и первый способ нахождения  $S$ , и второй

### 6.1. Итоги последних пяти чемпионатов четырех европейских стран

Вычислим коэффициенты конкордации для футбольных первенств Германии, Италии, Испании и России. Поскольку регламент соревнований в каждой стране допускает обмен команд с более слабой лигой, то входящим в элитный дивизион присваивается один и тот же ранг, а для вычисления  $W$  используем вторую формулу с наличием связанных рангов.

#### Коэффициент конкордации германского чемпионата

Таблица 44

Ранг	Команды	Ранг					$\sum (R_{i,j})$	$\sum (R_{i,j})^2$
		2005	2004	2003	2002	2001		
1	Бавария	1	2	1	3	1	8	64
2	Шальке	2	7	7	5	2	23	529
3	Вердер	3	1	6	6	7	23	529
4	Герта	4	12	5	4	5	30	900
5	Штутгарт	5	4	2	8	15	34	1156
6	Байер	6	3	15	2	4	30	900
7	Боруссия Д	7	6	3	1	3	20	400
8	Гамбург	8	8	4	11	13	44	1936
9	Вольфсбург	9	10	8	10	9	46	2116
10	Ганновер	10	14	11	21,5	21,5	78	6084
11	Майнц	11	21,5	21,5	21,5	21,5	97	9409
12	Кайзерслаутерн	12	15	14	7	8	56	3136
13	Арминия	13	21,5	16	21,5	21,5	93,5	8742,25

14	Нюрнберг	14	21,5	17	15	21,5	89	7921
15	Боруссия М	15	11	12	12	21,5	71,5	5112,25
16	Бохум	16	5	9	21,5	18	69,5	4830,25
17	Ганза	17	9	13	14	12	65	4225
18	Фрайбург	18	13	21,5	16	6	74,5	5550,25
19	Айнтрахт	21,5	16	21,5	21,5	17	97,5	9506,25
20	Мюнхен 1860	21,5	17	10	9	11	68,5	4692,25
21	Кельн	21,5	18	21,5	17	10	88	7744
22	Коттбус	21,5	21,5	18	13	14	88	7744
23	Ст.Паули	21,5	21,5	21,5	18	21,5	104	10816
24	Унтерхакинг	21,5	21,5	21,5	21,5	16	102	10404
							1500	114446,5

$$S = \sum (R_{i,j})^2 - \frac{(\sum R_{i,j})^2}{n}; S = 114446,5 - \frac{(1500)^2}{24} = 20696,5;$$

$$W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1) - m \sum_1^m (t^2 - t)}; \text{ где: } S = 20696,5; m = 5; n = 24; t = 6.$$

$$W = 0,731$$

### Коэффициент конкордации итальянского чемпионата

Таблица 45

Ранг	Команды	Ранг					$\sum (R_{i,j})$	$\sum (R_{i,j})^2$
		2005	2004	2003	2002	2001		
1	Ювентус	1	3	1	1	2	8	64
2	Милан	2	1	3	4	6	16	256
3	Интер	3	4	2	3	5	17	289
4	Удинезе	4	7	6	14	12	43	1849
5	Сампdoria	5	8	25,5	25,5	25,5	89,5	8010,25
6	Палермо	6	25,5	25,5	25,5	25,5	108	11664
7	Мессина	7	25,5	25,5	25,5	25,5	109	11881
8	Рома	8	2	8	2	1	21	441
9	Ливорно	9	25,5	25,5	25,5	25,5	111	12321
10	Лацио	10	6	4	6	3	29	841
11	Лечче	11	10	25,5	16	13	75,5	5700,25
12	Кальяри	12	25,5	25,5	25,5	25,5	114	12996
13	Реджина	13	14	14	25,5	14	66,5	4422,25
14	Сиена	14	13	25,5	25,5	25,5	103,5	10712,25
15	Кьево	15	9	7	5	25,5	61,5	3782,25
16	Болонья	16	12	11	7	10	56	3136
17	Фиорентина	17	25,5	25,5	17	9	94	8836
18	Парма	18	5	5	10	4	42	1764
19	Брешиа	19	11	9	13	8	60	3600



20	Аталанта	20	25,5	13	4	7	69,5	4830,25
21	Перуджа	26,5	15	10	8	11	70,5	4970,25
22	Модена	26,5	16	15	25,5	25,5	108,5	11772,25
23	Эмполи	26,5	17	12	25,5	25,5	106,5	11342,25
24	Анкона	26,5	18	25,5	25,5	25,5	121	14641
25	Пьяченца	26,5	25,5	16	12	25,5	105,5	11130,25
26	Комо	26,5	25,5	17	25,5	25,5	120	14400
27	Горино	26,5	25,5	18	11	25,5	106,5	11342,25
28	Верона	26,5	25,5	25,5	15	15	107,5	11556,25
29	Венеция	26,5	25,5	25,5	18	25,5	121	14641
30	Венечце	26,5	25,5	25,5	25,5	16	119	14161
31	Наполи	26,5	25,5	25,5	25,5	17	120	14400
32	Бари	26,5	25,5	25,5	25,5	18	121	14641
							2621	256393

$$S = \sum (R_{i,j})^2 - \frac{(\sum R_{i,j})^2}{n}; \quad S = 256393 - \frac{(2621)^2}{32} = 4171,72;$$

$$W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1) - m \sum_1^m (t^3 - t)}$$

где  $S = 4171,72$ ;  $m = 5$ ;  $n = 32$ ;  $t = 12$ ;  $W = 0,616$

### Кoeffициент конкордации испанского чемпионата

Таблица 46

Ранг	Команды	Ранг					$\sum (R_{i,j})$	$\sum (R_{i,j})^2$
		2005	2004	2003	2002	2001		
1	Барселона	1	2	6	4	4	17	289
2	Реал М	2	4	1	3	1	11	121
3	Вильярреал	3	8	15	15	7	48	2304
4	Бетис	4	9	8	6	25	52	2704
5	Эспаньол	5	16	17	14	9	61	3721
6	Севилья	6	6	10	8	25	55	3025
7	Валенсия	7	1	5	1	5	19	361
8	Атлетик Б	8	5	7	10	14	44	1936
9	Депортиво	9	3	3	2	2	19	361
10	Малага	10	10	14	9	8	51	2601
11	Атлетико М	11	7	11	25	25	79	6241
12	Сарагоса	12	13	25	20	15	85	7225
13	Хетафе	13	25	25	25	25	113	12769
14	Реал С	14	15	2	13	13	57	3249

15	Осасуна	15	12	12	17	17	73	5329
16	Расинг	16	17	16	25	19	93	8649
17	Мальорка	17	11	9	16	3	56	3136
18	Леванте	18	25	25	25	25	118	13924
19	Нумансия	19	25	25	25	20	114	12996
20	Альбасете	20	14	25	25	25	109	11881
21	Вальядолид	25	18	13	12	16	84	7056
22	Сельта	25	19	4	5	6	59	3481
23	Мурсия	25	20	25	25	25	120	14400
24	Уэльва	25	25	18	25	25	118	13924
25	Алавес	25	25	19	7	10	86	7396
26	Р. Вальекано	25	25	20	11	12	93	8649
27	Лас Пальмас	25	25	25	18	11	104	10816
28	Генерифе	25	25	25	19	25	119	14161
29	Овьедо	25	25	25	25	18	118	13924
							2175	196629

$$S = \sum (R_{i,j})^2 - \frac{(\sum R_{i,j})^2}{n} \quad S = 196629 - \frac{(2175)^2}{29} = 33504$$

Коэффициент конкордации вычисляем по формуле, учитывающей наличие одинаковых рангов по каждому признаку.

$$W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1) - m \sum_1^m (t^3 - t)}$$

$$W = \frac{12 \cdot 33504}{25 \cdot 28 \cdot 29 \cdot 30 - 5 \cdot (9^3 - 9)} = 0,680$$

где  $m$  – число ранжируемых признаков,  $n$  – число наблюдений,  $t$  – число одинаковых знаков по каждому признаку.

В нашем случае  $m = 5$ ,  $n = 29$ ,  $t = 9$ . Получившийся коэффициент конкордации выше среднего свидетельствует о достаточно большой степени стабильности чемпионатов по футболу в Испании, особенно учитывая достаточно большой обмен между лигами.

### Коэффициент конкордации российского чемпионата

Таблица 47

Ранг	Команды	Ранг					$\sum (R_{i,j})$	$\sum (R_{i,j})^2$
		2004	2003	2002	2001	2000		
1	"Локомотив"	1	4	2	2	2	11	121

2"ЦСКА"	2	1	1	7	8	19	361
3"Зенит"	3	2	10	3	7	25	625
"Крылья Советов"	4	9	5	5	14	37	1369
5"Торпедо"	5	8	4	4	3	24	576
6"Шинник"	6	5	6	19,5	19,5	56	3136
7"Сатурн"	7	7	7	6	9	36	1296
8"Спартак"	8	10	3	1	1	23	529
9ФК "Москва"	9	14	14	14	17	68	4624
10"Рубин"	10	3	19,5	19,5	19,5	71,5	5112,25
11"Амкар"	11	19,5	19,5	19,5	19,5	89	7921
12"Ростов"	12	11	11	11	12	57	3249
13"Динамо" М	13	6	8	9	5	41	1681
14"Алания"	14	13	12	12	10	61	3721
15"Кубань"	15	19,5	19,5	19,5	19,5	93	8649
16"Ротор"	16	12	9	10	11	58	3364
17"Сокол"	19,5	19,5	16	8	19,5	82,5	6806,25
18"Анжи"	19,5	19,5	15	13	4	71	5041
19"Факел"	19,5	19,5	19,5	15	13	86,5	7482,25
20"Локомотив"	19,5	19,5	19,5	19,5	15	93	8649
21"Уралан"	19,5	15	13	19,5	16	83	6889
22"Черноморец"	19,5	16	19,5	16	6	77	5929
						1262,5	87130,75

$$S = \sum (R_{i,j})^2 - \frac{(\sum R_{i,j})^2}{n} \quad S = 87130,75 - \frac{(1262,5)^2}{22} = 14680,47$$

$$W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1) - m \sum_1^m (t^3 - t)}$$

где  $S = 14680,47$ ;  $m = 5$ ;  $n = 22$ ;  $t = 6$ ;  $W = 0,677$

## 6.2. Итоги последнего чемпионата четырех европейских стран

Найдем множественные коэффициенты ранговой корреляции и корреляционные матрицы общих результатов и успехов в домашних и гостевых играх рассматриваемых футбольных стран.

Коефициент конкурдации чемпионата по футболу Германия 2004 - 2005

Таблица 48

Всего (O)				Дома (Д)				В гостях (Г)				
Ранг	Команды	Очки	Разница	Ранг	Разница	Ранг	Разница	Ранг	Разница	Ранг	Разница	
77	4444-14	1	0	0	3331-19	1	0	0	0	0	0	
63	3533-24	4	-2	4	2823-22	3	-1	1	1	1	1	
58	3133-15	6	-3	8	2895-22	2	1	1	4	16	11	
58	3234-13	5	-1	1	2825-18	4	0	0	1	1	13	
58	3894-15	3	2	4	2020-25	8	-3	8	-5	25	16	
57	3942-18	2	4	16	1823-26	10	-4	16	-8	64	18	
55	2824-18	9	-2	4	2823-26	5	2	4	4	16	21	
48	3135-20	11	-3	9	2328-28	6	2	4	5	25	25	
48	3135-20	7	2	4	1714-31	11	-2	4	-4	16	27	
45	2821-19	12	-2	4	1913-17	9	1	1	3	9	31	
43	3028-21	8	3	9	1322-34	14	-3	8	-6	36	33	
42	2820-21	13	-1	1	1823-31	12	0	0	1	1	37	
40	2421-21	14	-1	1	1816-28	13	0	0	1	1	40	
38	1825-25	16	-2	4	2030-28	7	7	49	8	81	37	
38	2828-21	10	5	25	76-30	17	-2	4	-7	49	42	
35	2330-29	15	1	1	1217-39	16	0	0	-1	1	47	
30	1717-31	17	0	0	1314-34	15	2	4	2	4	49	
18	1218-31	18	0	0	612-44	18	0	0	0	0	54	
												513
												18433

$$O \begin{pmatrix} 1 & 0,901 & 0,891 \\ 0,901 & 1 & 0,643 \\ 0,891 & 0,643 & 1 \end{pmatrix}$$

$$\sum_{O \neq D} d_i^2 = 96 \quad \sum_{D \neq G} d_i^2 = 346 \quad \sum_{D \neq G} d_i^2 = 106 \quad r = D \quad \tau_{OD} = 1 - \frac{6 \cdot 96}{17 \cdot 18 \cdot 19} = 0,901 \quad \tau_{OG} = 0,891 \quad \tau_{DG} = 0,643$$

$$S = \sum (R_{ij})^2 - \frac{(\sum R_{ij})^2}{n} \quad S = 18433 - \frac{(513)^2}{18} = 3812,5; \quad W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1)} \quad \text{где } S = 3812,5 \quad m = 3 \quad n = 18; \quad W = \frac{12 \cdot 3812,5}{3^2 \cdot 17 \cdot 18 \cdot 19} = 0,87.$$



Коефициент конкордации чемпионата по футболу Испании 2004 - 2005

Таблица 50

Всего (O)				Дома (D)				В гостях (Г)							
Ранг	Команда	Очки	Разница	Ранг	Дома	Очки	Разница	Ранг	Гостях	Очки	Разница	Дома	Гостях	$\sum R_{ij}^2$	$\sum R_{ji}^2$
1	Барселона	84	4840-12	3	-2	4	3883-17	1	0	0	2	4	5	25	25
2	Реал М	80	4643-12	1	1	1	3428-20	2	0	0	1	1	5	5	25
3	Вильярреал	65	4641-10	2	1	1	1928-27	9	-6	36	-7	49	14	196	196
4	Челси	62	4136-22	5	-1	1	2126-28	6	-2	4	-1	1	15	225	225
5	Эспаньол	61	4134-18	4	1	1	2020-28	8	-3	9	-4	16	17	289	289
6	Севилья	60	3925-19	11	-5	25	2519-22	4	2	4	7	49	21	441	441
7	Валенсия	58	3831-17	7	0	0	2023-22	7	0	0	0	0	21	441	441
8	Атлетик В	51	3739-24	8	0	0	1420-30	12	-4	16	-4	16	28	784	784
9	Реал Мадрид	51	2925-29	16	-7	49	2621-21	3	6	36	13	169	28	784	784
10	Малага	51	2819-24	15	-5	25	2321-24	5	5	25	10	100	30	900	900
11	Атлетика М	50	3928-13	6	5	25	1112-21	17	-6	36	-11	121	34	1156	1156
12	Сарагоса	50	3635-25	10	2	4	1417-32	13	-1	1	-9	9	35	1225	1225
13	Хетафе	47	3724-12	9	4	16	1015-34	18	-5	25	-9	81	40	1600	1600
14	Реал С	47	3121-24	13	1	1	1626-32	10	4	16	3	9	37	1369	1369
15	Осасуна	46	3326-24	12	3	9	1318-41	16	-1	1	-4	16	43	1849	1849
16	Леванте	44	3025-23	14	2	4	1416-35	14	2	4	0	0	44	1936	1936
17	Мальорка	39	2328-31	18	-1	1	1614-32	11	6	36	7	49	46	2116	2116
18	Леванте	37	2419-20	17	1	1	1320-38	15	3	9	2	4	50	2500	2500
19	Тумансия	29	2119-23	19	0	0	811-38	20	-1	1	-1	1	58	3364	3364
20	Альбасете	28	1917-22	20	0	0	916-34	19	1	1	1	1	58	3364	3364
														630	24706

$$\sum_{о,г} d_i^2 = 168; \sum_{о,г} d_i^2 = 260; \sum_{д,г} d_i^2 = 696 \quad r_s = D \begin{pmatrix} 0 & 1 & 0,874 & 0,804 \\ 1 & 0,477 & 1 & 0,477 \\ 1 & 0,477 & 1 & 0,477 \\ 1 & 0,477 & 1 & 0,477 \end{pmatrix}$$

$$r_{o,g} = 1 - \frac{6 \cdot 168}{19 \cdot 20 \cdot 21} = 0,874; \quad r_{o,r} = 0,805; \quad r_{g,r} = 0,477;$$

$$S = \sum (R_{ij})^2 - \frac{(\sum R_{ij})^2}{n}; \quad S = 24706 - \frac{(630)^2}{20} = 4861; \quad W = \frac{12S}{m^2(n-1) \cdot n \cdot (n+1)} \quad \text{где } S = 4861 \quad m = 3 \quad n = 20; \quad W = \frac{12 \cdot 4861}{3^2 \cdot 19 \cdot 20 \cdot 21} = 0,812.$$

## Кoeffициент конкордации чемпионата по футболу Россия 2004

Таблица 51

Всего (O)			Домаш (Д)			В гостях (Г)									
Ранг	Команды	Очки	Очки	Разница	Ранг	Дол.	До <sub>г</sub> <sup>2</sup>	Очки	Разница	Ранг	До <sub>г</sub> <sup>2</sup>	До <sub>г</sub> <sup>2</sup>	ΣR <sub>ij</sub> <sup>2</sup>	ΣR <sub>ij</sub>	ΣR <sub>ij</sub> <sup>2</sup>
1	Локомотив М	61	2819-6	5	-4	16	3225-13	1	0	0	4	16	7	49	49
2	ЦСКА	60	3025-11	4	-2	4	3028-11	2	0	2	4	8	8	64	64
3	Зенит	56	2824-21	6	-3	9	3031-16	3	0	3	9	12	144	144	
4	Ю.Советов	56	3228-19	2	2	4	2422-22	4	0	2	4	10	100	100	
5	Торпедо	54	3228-15	1	4	16	2225-22	5	0	4	16	11	121	121	
6	Шинник	44	3119-12	3	3	9	1310-17	9	-3	9	6	36	18	324	324
7	Сатурн	41	1920-16	12	-5	25	2217-14	6	1	1	6	36	25	625	625
8	Спартак	40	2122-20	9	-1	1	1921-24	7	1	1	2	4	24	576	576
9	ФК Москва	40	2525-21	7	2	4	1513-18	8	1	1	1	1	24	576	576
10	Рубин	33	2423-14	8	2	4	99-17	14	-4	16	-6	36	32	1024	1024
11	Амкар	30	2015-15	11	0	0	1012-27	13	-2	4	-2	4	35	1225	1225
12	Ростов	29	1917-18	13	-1	1	1011-24	12	0	0	1	1	37	1369	1369
13	Динамо М	29	2017-15	10	3	9	910-23	15	-2	4	-5	25	38	1444	1444
14	Алания	28	1614-24	15	-1	1	1214-28	10	4	16	5	25	39	1521	1521
15	Кубань	28	1814-17	14	1	1	1012-25	11	4	16	3	9	40	1600	1600
16	Ротор	22	1418-26	16	0	0	810-27	16	0	0	0	0	48	2304	2304
													408	13068	13068

$$\sum_{O,Г} d_i^2 = 104; \quad \sum_{O,Г} d_i^2 = 68; \quad \sum_{Д,Г} d_i^2 = 226 \quad r_s = D \begin{pmatrix} O & 1 & 0,874 & 0,900 \\ 1 & 0,668 & 1 & 1 \\ Г & 1 & 1 & 1 \end{pmatrix}$$

$$r_{o,г} = 1 - \frac{6 \cdot 104}{15 \cdot 16 \cdot 17} = 0,847; \quad r_{o,r} = 0,9; \quad r_{г,r} = 0,668;$$

$$S = \sum (R_{ij})^2 - \frac{(\sum R_{ij})^2}{n}; \quad S = 13066 - \frac{(408)^2}{16} = 2662; \quad W = \frac{12,5}{m \cdot (n-1) \cdot n \cdot (n+1)} \quad \text{где } S = 2662 \quad m = 3 \quad n = 16; \quad W = \frac{12 \cdot 2662}{3^3 \cdot 15 \cdot 16 \cdot 17} = 0,870.$$

Итоги последних чемпионатов своих стран по футболу, рассмотренные здесь с позиции показанных результатов, во всем первенстве и в домашних и гостевых матчах показали наименьшую согласованность этих результатов в итальянском первенстве (коэффициент конкордации равен только 0,64). Наиболее согласованы итоги последних чемпионатов в Германии и в России, для которых множественный коэффициент ранговой корреляции получился равным 0,87.

### 6.3. Текущие результаты последних чемпионатов России по хоккею и футболу

6.3.1. В этом разделе рассмотрим положение дел в российском хоккее (по состоянию на 27.11.2005), для этого определим степень зависимости текущих результатов команд в регулярном чемпионате и показатели в домашних и гостевых играх. Изучим и влияние на успехи команд приобретения новых игроков в период заявок и дозаявок, а также наличие в командах опытных игроков (старше 30 лет).

1. Вычислим коэффициент конкордации соответствия общих набранных очков и показанных результатов в домашних и гостевых играх. Поскольку присутствуют одинаковые ранги в домашних результатах команд, то используем следующую формулу:

$$W = \frac{12 \cdot S}{m^2 \cdot (n-1) \cdot n(n+1) - m \sum (t^2 - t)}$$

В нашем случае  $S = 18503 - \frac{(513)^2}{18} = 3883$  и тогда

$$W = \frac{12 \cdot 3883}{9 \cdot 17 \cdot 18 \cdot 19 - 3(4^3 - 4)} = \frac{2 \cdot 3883}{3 \cdot 17 \cdot 19 - 30} = \frac{7766}{8691} \approx 0,89$$

Следовательно, текущие результаты суперлиги (по состоянию на 27.11.2005) по хоккею и результаты дома и в гостях имеют сильную прямую связь.

2. Найдем связь между результатами команд и наличием в команде возрастных хоккеистов (старше 30 лет). С этой целью вычислим коэффициент ранговой корреляции Спирмена.

$$\begin{aligned} \sum d_i^2 &= 6,25 + 196 + 0,25 + 121 + 144 + 2,25 + 16 + 2,25 + 4 + 4 + 30,25 + 1 + 90,25 + 16 \\ &\quad + 256 + 36 = 925,5 \\ r_s &= 1 - \frac{6 \cdot 925,5}{17 \cdot 18 \cdot 19} = 0,955 \end{aligned}$$



**Чемпионат России по хоккею (2005 - 2006 годы)**  
**Положение команд**

Ме-сто	Команды	Оч-ки	Дома		В гостях			$\sum R_j$	$(\sum R_i)^2$	Кому за 30 лет			Заявка			С лозья-кой			
			% очков	Ран г	% очков	Ран г	Ран г			Кол-во	Ран г	Кол-во	Ран г	Кол-во	Ран г	Кол-во	Ран г	Кол-во	Ран г
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18		
1	"Металлург" (Магнитогорск)	64	95,24	1	61,54	3	5	25	1	3,5	10	16	9	17					
2	"Авангард" (Омск)	55	64,10	4	66,67	1	7	49	4	2	11	8	16	6,5					
3	"Локомотив" (Ярославль)	50	62,22	5	56,41	4	12,5	144	-	17	3	18	7	18					
4	ЦСКА (Москва)	50	66,67	2,5	51,28	6	12,5	156,25	2	3,5	10	16	9	16					
5	"Ак - барс" (Казань)	47	66,67	2,5	50,00	9	16,5	272,25	1	16	4	11,5	14	13,5					
6	"Лада" (Тольятти)	46	61,90	6	51,28	7	19	361	8	18	-	16	9	11					
7	"Салават Юлаев" (Уфа)	43	42,86	12,5	64,10	2	21,5	462,25	4	5,5	9	10	15	8,5					
8	"Динамо" (Москва)	43	52,38	7	53,85	5	20	400	9	12	6	5	18	1,5					
9	"Химик" (Моск. обл.)	41	46,15	11	51,11	8	28	784	6	7,5	8	3,5	19	4					
10	"Спартак" (Москва)	39	52,38	8	40,48	11	29	841	-	12	6	8	16	12					
11	"Северсталь" (Череповец)	38	51,28	9	42,86	10	30	900	4	9	7	2	21	4					
12	СКА (Санкт-Петербург)	32	48,72	10	30,95	15	37	1369	5	12	6	1	22	1,5					
13	"Сибирь" (Новосибирск)	32	40,48	14	38,46	13	40	1600	5	7,5	8	13	13	10					
14	"Нефтехимик" (Нижнекамск)	30	33,33	17	40,48	12	43	1849	3	15	5	14	11	15					
15	ХК МВД (Тверь)	30	38,10	15	33,33	14	44	1936	4	5,5	9	8	16	6,5					
16	"Металлург" (Новокузнецк)	23	33,33	17	23,08	16	49	2401	2	12	6	6	17	8,5					
17	"Витязь" (Чехов)	22	33,33	17	20,00	17	51	2601	6	1	14	3,5	19	4					
18	"Молог-Прикамье" (Перьмь)	21	42,86	12,5	7,14	18	48,5	2352,25	1	12	6	11,5	14	13,5					
							513	18503											
							$\Sigma$												

Таблица 52

Получили удивительный результат о почти полной корреляции достижений команд и наличии определенного количества возрастных хоккеистов. Это тем более удивительно, что у ряда команд эти показатели прямо противоположны. Так как у «Локомотива», находящегося на 3-м месте, только три (17-й ранг) хоккеиста старше 30 лет. «Лада», занимающая 6-е место, вообще не имеет возрастных хоккеистов. С другой стороны, «Витязь» из г.Чехова имеет наибольшее количество возрастных хоккеистов и в то же время занимает только предпоследнее место.

3. Вычислим теперь коэффициент корреляции между достижениями команд и количеством заявленных (с учетом дозаявок) новых хоккеистов.

$$\sum d_i^2 = 256 + 20,25 + 225 + 144 + 72,25 + 25 + 2,25 + 42,25 + 25 + 4 + 49 + 110,25 + 9 + 1 + 72,25 + 56,25 + 169 + 20,25 = 1283,25$$

$$r_s = 1 - \frac{6 \cdot 1283,25}{17 \cdot 18 \cdot 19} \approx -0,32$$

Поскольку получили, что  $-0,7 < r_s < -0,30$ , и это означает наличие **обратной**, средней силы связи между результатами команд и количеством приобретенных хоккеистов. Откуда следует вывод о неразборчивости многих приобретений, которые не дают положительного эффекта. Этот вывод подтверждается, с одной стороны, достижениями таких команд, как «Металлург» (Магнитогорск), «Локомотив», ЦСКА, которые не были активны на трансферном рынке, а с другой – большими приобретениями «Динамо», СКА, «Химик», «Северсталь», «Витязь» и слабыми их спортивными показателями.

6.3.2. Вновь вернемся к итогам последнего чемпионата России по футболу и вычислим коэффициент конкордации для общих итогов и результатов, показанных командами в домашних или гостевых играх.

### Чемпионат России по футболу 2005 г. Премьер – лига. Итоговая таблица

Таблица 53

Место 2005г	Команды	Очки	Дома		В гостях		$\sum_j R_{ij}$	$(\sum_j R_{ij})^2$	Дублеры	
			Очки	Ранг	Очки	Ранг			Очки	Место
1	2	3	4	5	6	7	8	9	10	11
1	ЦСКА	62	39	1	23	3	5	25	70	1
2	"Спартак"	56	30	4,5	26	1,5	8	64	46	6
3	"Локомотив"	56	30	4,5	26	1,5	9	81	50	5

4	"Рубин"	51	33	2	18	5,5	11,5	132,25	29	13
5	"Москва"	50	29	6	21	4	15	225	64	2
6	"Зенит"	49	32	3	17	7	16	256	59	3
7	"Торпедо"	45	27	7,5	18	5,5	20	400	42	8
8	"Динамо"	38	23	11	15	8	27	729	56	4
9	"Шинник"	38	27	7,5	11	10	26,5	702,25	28	14
10	"Томь"	37	23	10	14	9	30	900	43	7
11	"Сатурн"	3	23	10	10	11,5	33,5	1122,25	41	9
12	"Амкар"	33	23	10	10	11,5	34,5	1190,25	27	15
13	"Ростов"	31	22	13	9	13	39	1521	31	12
14	"Крылья Советов"	29	23	10	6	15	40	1600	22	16
15	"Алания"	23	18	15	5	16	46	2116	36	10
16	"Терек"	20	12	16	8	14	46	2116	33	11
$\Sigma$									407	12955

Определим связь между результатами основных и дублирующих составов. С этой целью вычислим коэффициент ранговой корреляции Спирмена.

$$\sum d_i^2 = 16 + 4 + 81 + 9 + 9 + 1 + 16 + 25 + 9 + 4 + 9 + 1 + 4 + 25 + 25 = 238$$

$$\text{Тогда } r_s = 1 - \frac{6 \cdot 238}{15 \cdot 16 \cdot 17} = 0,65 \text{ и связь прямая и средняя.}$$

Найдем коэффициент конкордации для определения соответствия общих результатов чемпионата и результатов в гостевых и домашних матчах. Используем формулу коэффициента конкордации с наличием одинаковых рангов:

$$W = \frac{12 \cdot S}{m^2 \cdot (n+1) \cdot n(n+1) - m \sum (t^3 - t)}$$

$$\text{В нашем случае } S = 12955 - \frac{407^2}{16} = 2602 \text{ и}$$

$$\text{тогда } W = \frac{12 \cdot 2602}{9 \cdot 15 \cdot 16 \cdot 17 - 3[(9^3 - 9) + (5^3 - 5)]} = \frac{2502}{3 \cdot 15 \cdot 4 \cdot 17 \cdot 210} \approx 0,913.$$

Итоги чемпионата России по футболу в премьер-лиге и результаты в домашних и гостевых играх имеют сильную связь, поскольку коэффициент достаточно близок к единице.

#### 6.4. Конкордация в хоккее

В этом разделе коэффициент конкордации используем для иллюстрации стабильности показателей лиги в целом (а не отдельной команды) за несколько сезонов российских первенств 2001 - 2005 гг. и СССР 1980 - 1984 гг. С помощью коэффициента конкордации мы сможем также проиллюстрировать стабильность игры лучших бомбардиров лиги за последние четыре сезона.

Рассмотрим результаты последних четырех регулярных чемпионатов России по хоккею и найдем степень зависимости их результатов.

#### Результаты регулярных сезонов 2001 – 2005 гг.

Таблица 54

Наименование команды	2001-2002	Ранг	2002-2003	Ранг	2003-2004	Ранг	2004-2005	Ранг	$\sum_j R_{ij}$	$\left(\sum_j R_{ij} - T\right)^2$
	Занимаемое место		Занимаемое место		Занимаемое место		Занимаемое место			
"Динамо"	7	7	7	7	6	6	1	1	21	169
"Лада"	4	4	5	5	2	2	2	2	13	441
"Металлург" Мг	5	5	6	6	1	1	3	3	15	361
"Ак-барс"	2	2	4	4	5	5	4	4	15	361
"Локомотив"	1	1	1	1	7	7	5	5	14	400
"Авангард"	3	3	2	2	3	3	6	6	14	400
"Металлург" Нк	12	10	9	9	4	4	7	7	30	16
"Нефтехимик"	13	11	12	11	8	8	8	8	38	16
"Химик"	-	16	-	16	12	12	9	9	53	361
ЦСКА	17	14	10	10	10	10	10	10	44	100
"Северсталь"	6	6	3	3	13	13	11	11	33	1
СКА	16	13	13	12	14	14	12	12	51	289
"Салават Юлаев"	9	8	8	8	9	9	13	13	38	16
"Сибирь"	-	15	14	13	11	11	14	14	53	361
"Спартак"	11	9	15	14	-	15	15	15	53	361
"Молот-Прикамье"	15	12	17	15	-	16	16	16	59	625
Сумма									544	4278

Средняя величина суммы четырех рангов равна  $34 = (544/16)$ , то  $S=4278$ . Учитывая, что  $t=0$ , получаем

$$W = \frac{12S}{m^2(n-1)n(n+1) - m \sum_1^n (t^3 - t)} = \frac{12 \times 4278}{4^2 \times 15 \times 16 \times 17} = 0,786$$

Команды суперлиги за четыре последних сезона выступают достаточно стабильно, о чем свидетельствует полученный коэффициент конкордации, равный 0,786. Однако, чтобы это утверждение не было ошибочным,  $W$  проверяют на значимость.

Рассмотрим результаты чемпионатов СССР 1980 - 1984 гг.

## Результаты чемпионатов СССР 1980 - 1984 гг.<sup>1</sup>

Таблица 55

Команды	1980-1981	Ран г	1981-1982	Ран г	1982-1983	Ранг	1983-1984	Ран г	$\sum_j R_{ij}$	$\left(\sum_j R_{ij}\right)^2$
	Занимаемое место		Занимаемое место		Занимаемое место		Занимаемое место			
ИСКА	1	1	1	1	1	1	1	1	4	16
"Спартак" М.	2	2	2	2	2	2	2	2	8	64
"Динамо" М.	3	3	3	3	3	3	4	4	13	169
"Трактор" Чел.	4	4	6	6	-	10,5	9	9	29,5	870
"Динамо" Р.	5	5	8	8	5	5	8	8	26	676
"Сокол" К.	6	6	5	5	4	4	5	5	20	400
"Крылья Советов" М.	-	9,5	-	10,5	7	7	7	7	34	1156
"Торпедо" Г.	-	9,5	4	4	6	6	6	6	25,5	650
"Химик" Восресенск	-	9,5	-	10,5	8	8	3	3	31	961
СКА	-	9,5	7	7	-	10,5	10	10	37	1369
"Салават Юлаев"	-	9,5	-	10,5	-	10,5	-	11,5	42	1764
"Динамо" Минск	-	9,5	-	10,5	-	10,5	-	11,5	42	1764
<b>Сумма</b>									<b>312</b>	<b>9859</b>

Для расчета коэффициента конкордации используем другой метод вычисления S, идея которого отражена в двух последних столбцах таблицы:

$$S = \sum_i \left( \sum_j R_{ij} \right)^2 - \frac{\left( \sum_i \sum_j R_{ij} \right)^2}{n}$$

$$S = 9859 - \frac{312^2}{12} = 1747$$

Тогда

$$W = \frac{12 \times 1747}{4^2 \times 15 \times 16 \times 17 - 4[6 \times (36 - 1) + 4(16 - 1) + 4(16 - 1) + 2(4 - 1)]} = 0,328$$

Один из выводов, очевидных после сравнения коэффициентов конкордации чемпионатов России 2001 – 2005 гг. и чемпионатов СССР 1980 – 1984 гг., состоит в том, что в России, как и в СССР, есть команды, борющиеся из года в год за золотые медали (насчитывается при-

<sup>1</sup>Энциклопедия «Хоккей», «Терра Спорт», издание 2000 г.

мерно 6 таких команд), а также команды, которые с таким же постоянством борются за выживание в нижней части турнирной таблицы.

Далее рассмотрим стабильность выступления лучших игроков в регулярных чемпионатах за последние четыре сезона на примере лучших бомбардиров суперлиги.

### Результаты лучших бомбардиров сезонов 2001 – 2005 гг.<sup>2</sup>

Таблица 56

Фамилия игрока	2001-2002	Ранг	2002-2003	Ранг	2003-2004	Ранг	2004-2005	Ранг	$\sum Ri$	$(\sum Ri)^2$
	Кол-во очков (г+п)		Кол-во очков (г+п)		Кол-во очков (г+п)		Кол-во очков (г+п)			
Сушинский М.	51	1	34	6	61	1	55	1	9	81
Затонский Д.	41	4,5	36	5	48	4	46	2	15,5	240,25
Антипов В.	18	10	30	9	25	9	40	4	32	1024
Язин А.	43	2	37	3,5	53	3	24	9	17,5	306,25
Епанчинцев В.	41	4,5	40	1	38	5	28	8	18,5	342,25
Прокопьев А.	32	7	37	3,5	57	2	45	3	15,5	240,25
Квартальнов Д.	42	3	29	10	29	8	15	10	31	961
Калужный А.	29	8	39	2	31	6	30	7	23	529
Бут А.	26	9	32	7	22	10	39	5	31	961
Чупин А.	35	6	31	8	30	7	37	6	27	729
<b>Сумма</b>									<b>220</b>	<b>5414</b>

$$S = 5414 - \frac{220^2}{10} = 574$$

$$W = \frac{12 \times 574}{16(1000 - 10) - 4 \times 12} = 0,436$$

Вышеперечисленные игроки выступают нестабильно в последних чемпионатах России, т.к. коэффициент конкордации равен 0,436.

### §7. Коэффициент корреляции Пирсона

Понятие корреляции является одним из основных понятий теории вероятности и математической статистики. Оно было введено Гальтоном и Пирсоном. Закон природы или общественного развития может быть представлен описанием совокупности взаимосвязей. Если эти зависимости стохастичны, а анализ осуществляется по выборке из генеральной совокупности, то данная область исследования относится к задачам стохастического исследования зависимостей, которые включают в себя корреляционный, регрессионный, дисперсионный и ковариационный анализы.

<sup>2</sup> [www.metallurg.ru](http://www.metallurg.ru) – официальный сайт хоккейного клуба «Металлург Магнитогорск»

Предложенный в предыдущих параграфах подход к рассмотрению командных и личных итогов розыгрышей по хоккею и футболу продолжим здесь с использованием коэффициента корреляции Пирсона. Коэффициент ранговой корреляции, широко используемый ранее, обладает, с одной стороны, простотой применения, но с другой – не учитывает возможных существенных различий в показателях (очков, голов, разности шайб и т. д.) даже у соседних по рангу команд или игроков. Пусть при проведении некоторого опыта наблюдаются две случайные величины  $X$  и  $Y$ , причем одно и то же значение  $x$  встречается  $n_x$  раз, значение  $y$  встречается  $n_y$  раз, одна и та же пара чисел  $(x, y)$  наблюдается  $n_{xy}$  раз. Все данные записываются в виде таблицы, которую называют корреляционной.

**Выборочная ковариация**  $k(X, Y)$  величин  $X$  и  $Y$  определяется формулой

$$k(X, Y) = \frac{1}{n} \sum (x_i - x^*)(y_j - y^*)n_{ij},$$

где  $n = \sum n_{ij}$ , а  $x^*, y^*$  – выборочные средние величин  $X$  и  $Y$ . При небольшом количестве экспериментальных данных  $k(X, Y)$  авторы предлагают находить [2], как полный вес ковариационного графа:

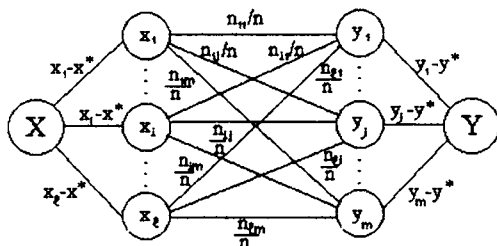


Рис. 12. Ковариационный граф.

Ковариация  $k(X, Y)$  характеризует не только степень зависимости случайных величин (ковариация двух независимых случайных величин равна нулю), но также их рассеивания вокруг точки  $(x^*, y^*)$ .

**Выборочный коэффициент корреляции** находится по формуле

$$r(X, Y) = \frac{k(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{\sum n_{xy}xy - x^*y^*}{n\sigma_x \cdot \sigma_y},$$

где  $\sigma_x^*$ ,  $\sigma_y^*$  – выборочные средние квадратические отклонения величин  $X$  и  $Y$ . Выборочный коэффициент корреляции  $r(X, Y)$  показывает тесноту линейной связи между  $X$  и  $Y$ : чем ближе  $|r(X, Y)|$  к единице, тем сильнее линейная связь между  $X$  и  $Y$ . Случайные величины  $X$  и  $Y$  называются некоррелированными, если  $r(X, Y) = 0$ .

### 7.1. Бюджеты клубов и их результаты

Оценим с помощью выборочного коэффициента корреляции степень линейной связи между бюджетами клубов и набранными очками в ходе регулярного хоккейного чемпионата 2005 – 2006 гг.

#### Бюджеты хоккейных клубов и результаты двух кругов 2005 - 2006 гг.\*

Таблица 57

Место	Клуб	Бюджет в млн.\$	Очки
1	"Металлург" (Мг)	26	83
2	"Локомотив"	24	65
3	"Ак – барс"	35	63
4	"Авангард"	30	62
5	"Лада"	9	57
6	ЦСКА	15	56
7	"Динамо"	24	54
8	"Химик"	15	54
9	"Спартак"	9	50
10	"Салават Юлаев"	10	49
11	"Северсталь"	15	44
12	СКА	15	44
13	"Сибирь"	10	42
14	"Нефтехимик"	12	41
15	ХК МВД	12	24
16	"Витязь"	15	28
17	"Металлург" (Нк)	12	27
18	"Молот-Прикамье"	6	25

\*Бюджеты клубов даны по версии «Комсомольской правды» от 8 - 15 декабря 2005 г. - с. 34.



Пусть случайная величина  
 $Y = \{\text{число очков, набранных командами суперлиги}\},$   
 $X = \{\text{величина бюджетов клубов хоккейной суперлиги}\}.$

Тогда из таблицы 57 находим основные характеристики двумерной случайной величины  $(X, Y)$ .

1. Найдем выборочные средние случайные величины  $X$  и  $Y$ :

$$x^* = \frac{1}{18}(26 + 24 + \dots + 12 + 6) = 17 (\text{млн.дол.})$$

$$y^* = \frac{1}{18}(83 + 65 + \dots + 27 + 25) = 48,78 \approx 49 (\text{очков})$$

2. Определим выборочную ковариацию двумерной случайной величины  $X$  и  $Y$ :

$$k(X, Y) = \frac{1}{18}[(26-17) \cdot (83-49) + (24-17) \cdot (65-49) + \dots + (6-17) \cdot (25-49)] \approx 76,17$$

3. Вычислим выборочные дисперсии с. в.  $X$  и  $Y$ :

$$D_X = \frac{1}{18}[(26-17)^2 + (3-17)^2 + \dots + (12-17)^2 + (6-17)^2] = 60,5$$

$$D_Y = \frac{1}{18}[(83-49)^2 + (65-49)^2 + \dots + (27-49)^2 + (25-49)^2] \approx 211,4$$

4. Найдем выборочный коэффициент корреляции Пирсона для числа очков, набранных командами суперлиги по итогам двух кругов регулярного чемпионата, и величиной бюджетов хоккейных клубов:

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D_X \cdot D_Y}} \approx \frac{76,17}{\sqrt{60,5 \cdot 211,4}} \approx 0,674$$

Полученный коэффициент корреляции свидетельствует о существовании прямой, но только средней силы связи между бюджетами хоккейных клубов и их текущими спортивными результатами. Подтверждением этого факта могут являться достаточно высокие места команд «Лада», ЦСКА, «Спартак», «Салават Юлаев», которые имеют сравнительно малые финансовые возможности. Наоборот обстоят дела в этом сезоне у «Авангарда», «Динамо», «Витязя», СКА и «Северстали».

Рассмотрим итоги только что закончившегося сезона 2005 г. премьер-лиги по футболу. Найдем корреляционную матрицу парных корреляций случайных величин:

$X = \{\text{число очков, набранных командами в сезоне 2005 г.}\},$

$Y = \{\text{величина бюджетов клубов премьер-лиги}\},$

$Z = \{\text{число легионеров футбольных клубов из дальнего и ближнего зарубежья}\}.$

**Бюджеты футбольных клубов и итоги чемпионата 2005 г.**

Таблица 58

Место 2005 г.	Клуб	Бюджет в млн. \$	Оч- ки	Легио- неры
		(X)	(Y)	(Z)
1	ЦСКА	35	62	11
2	"Спартак"	50	56	15
3	"Локомотив"	35	56	11
4	"Рубин"	10	51	15
5	"Москва"	20	50	12
6	"Зенит"	20	49	13
7	"Торпедо"	10	45	12
8	"Динамо"	45	38	22
9	"Шинник"	10	38	16
10	"Томь"	8	37	13
11	"Сатурн"	22,5	33	15
12	"Амкар"	8	33	11
13	"Ростов"	10	31	17
14	"Крылья Со- ветов"	12,5	29	21
15	"Алания"	8	23	23
16	"Терек"	17,5	20	11

Вычислим числовые характеристики с.в.  $X, Y, Z$ , как и в предыдущем случае.

$$1. x^* = \frac{1}{16}(35 + 50 + \dots + 8 + 17,5) = 20,09 \approx 20(\text{млн.дол.})$$

$$y^* = \frac{1}{16}(62 + 56 + \dots + 23 + 20) = 37(\text{очков})$$

$$z^* = \frac{1}{16}(11 + 15 + \dots + 23 + 11) = 14,88 \approx 15(\text{игроков})$$

2. Найдем три попарные ковариации трехмерной случайной величины  $(X, Y, Z)$ .

$$k(X, Y) = \frac{1}{16}[(35 - 20) \cdot (62 - 37) + (50 - 20) \cdot (56 - 37) + \dots + (17,5 - 20) \cdot (20 - 37)] \approx 76,6$$

$$k(X, Z) = \frac{1}{16}[(35-20) \cdot (11-15) + (50-20) \cdot (15-15) + \dots + (17,5-20) \cdot (11-15)] \approx 2,1$$

$$k(Y, Z) = \frac{1}{16}[(62-37) \cdot (11-15) + (56-37) \cdot (15-15) + \dots + (20-37) \cdot (11-15)] \approx -26,38$$

3. Вычислим выборочные дисперсии

$$D_x = \frac{1}{16}[(35-20)^2 + (50-20)^2 + \dots + (8-20)^2 + (17,5-20)^2] \approx 179,7$$

$$D_y = \frac{1}{16}[(62-37)^2 + (56-37)^2 + \dots + (23-37)^2 + (20-37)^2] \approx 158,7$$

$$D_z = \frac{1}{16}[(11-15)^2 + (15-15)^2 + \dots + (23-15)^2 + (11-15)^2] = 14,25$$

4. Вычислим коэффициент корреляции Пирсона.

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D_x \cdot D_y}} = \frac{76,6}{\sqrt{179,7 \cdot 158,7}} = 0,45$$

$$r(X, Z) = \frac{2,1}{\sqrt{179,7 \cdot 14,25}} = 0,04$$

$$r(Y, Z) = \frac{-26,38}{\sqrt{158,7 \cdot 14,25}} = -0,55$$

5. Из полученных коэффициентов корреляции Пирсона составим корреляционную матрицу

$$(r) = \begin{matrix} X & \begin{pmatrix} 1 & 0,45 & 0,04 \\ & 1 & -0,55 \\ & & 1 \end{pmatrix} \\ Y \\ Z \end{matrix}$$

Обратим внимание на то, что показатели бюджетов клубов и их легионеров в 2005 году были практически независимы друг от друга ( $r=0,04$  близко к нулю), а число очков и легионеров команд имеет обратную связь, хотя и средней силы ( $r = -0,55 < 0$ ).

## 7.2. Влияние первого гола на результат

Проанализируем статистику Акселя Вартапянца о влиянии первого гола на результаты матчей, предложенную в «Спорт - экспрессе» (прил. «Футбол», № 144 от 23.12.2005 г., с.3).

Как влиял первый гол на результат матчей?

Таблица 59

Команды	1-й забитый гол						1-й пропущенный гол					
	К	В	%	Н	НП	П	К	В	%	Н	НП	П
ЦСКА	18	17	94,4	1	100,0	-	9	1	11,1	4	55,5	4
"Спартак"	19	14	73,7	3	89,5	2	9	2	22,2	3	55,5	4
"Локомотив"	13	12	92,3	1	100,0	-	8	2	25,0	4	75,0	2
"Рубин"	17	14	82,4	2	94,1	1	9	-	0	3	33,3	6
"Москва"	17	13	76,5	3	94,1	1	9	1	11,1	1	22,2	7
"Зенит"	17	13	76,5	4	100,0	-	9	-	0	2	22,2	7
"Торпедо"	11	10	90,9	-	90,9	1	16	2	12,5	6	50,0	8
"Динамо"	17	11	64,7	-	64,7	6	11	1	9,1	-	9,1	10
"Шинник"	10	7	70,0	2	90,0	1	14	2	14,3	3	35,7	9
"Томь"	9	6	66,7	1	77,8	2	13	3	23,1	1	30,8	9
"Сатурн"	13	7	53,8	4	84,6	2	12	1	8,3	-	8,3	11
"Амкар"	11	5	45,5	5	90,9	1	12	2	16,6	-	16,6	10
"Ростов"	10	6	60,0	4	100,0	-	18	2	11,1	1	16,6	15
"Крылья Советов"	11	7	63,6	3	90,9	1	17	-	0	3	17,6	14
"Алания"	6	3	50,0	2	83,3	1	20	2	10,0	2	20,0	16
"Терек"	7	5	71,4	-	71,4	2	20	-	0	2	10	18
<b>ИТОГО:</b>	<b>206</b>	<b>150</b>	<b>72,8</b>	<b>35</b>	<b>89,8</b>	<b>21</b>	<b>206</b>	<b>21</b>	<b>10,2</b>	<b>35</b>	<b>27,2</b>	<b>150</b>

Рассмотрим четыре случайные величины:

$X = \{\text{число первых забитых голов}\},$

$Y = \{\text{число выигрышей при первом забитом голе}\},$

$Z = \{\text{число первых пропущенных голов}\},$

$U = \{\text{число выигрышей при первом пропущенном голе}\}.$

Вызывает интерес изучение степени корреляции двух пар случайных величин – это  $(XY)$  и  $(ZU)$ .

1. Найдем выборочные средние случайных величин

$x^* = 12,875 \approx 13 = z^*$ , поскольку среднее число первых забитых голов равно среднему числу первых пропущенных голов.

$$y^* = 9,375 \approx 9,4;$$

$$u^* = 1,3125 \approx 1,3.$$

2. Вычислим выборочные дисперсии всех четырех случайных величин:

$$D_x = \frac{1}{16}(25 + 36 \cdot 2 + 16 \cdot 5 + 4 \cdot 3 + 9 \cdot 2 + 49) = 16;$$

$$D_y = \frac{1}{16}(57,76 + 21,16 + \dots + 40,96 + 29,16) = 16,61;$$

$$D_z = \frac{1}{16}(16 \cdot 6 + 25 \cdot 2 + 9 \cdot 4 + 49 \cdot 2) = 16,25;$$

$$D_u = \frac{1}{16}(0,094 \cdot 4 + 0,49 \cdot 7 + 2,89) = 0,42.$$

3. Определим выборочные ковариации рассматриваемых нами пар случайных величин:

$$k(X, Y) = \frac{1}{16}[(18-13) \cdot (17-9,4) + (19-13) \cdot (14-9,4) + \dots + (7-13) \cdot (5-9,4)] \approx 14,66$$

$$k(Z, U) = \frac{1}{16}[(9-13) \cdot (1-1,3) + (9-13) \cdot (2-1,3) + \dots + (20-13) \cdot (0-1,3)] \approx -0,9$$

4. Найдем интересующие нас коэффициенты корреляции Пирсона.

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D_x \cdot D_y}} = \frac{14,66}{\sqrt{16 \cdot 16,61}} \approx 0,9$$

$$r(Z, U) = \frac{k(Z, U)}{\sqrt{D_z \cdot D_u}} = \frac{-0,9}{\sqrt{16,25 \cdot 0,42}} \approx -0,34$$

Откуда следует значительная прямая корреляция ( $r=0,9$ ) между первым забитым голом и общим успехом в матче и средней силы обратная связь ( $r = -0,34$ ) между первым пропущенным голом и положительным исходом матча.

### 7.3. Совмещенные ковариационные графы

Предложение о геометрической интерпретации нахождения важнейшей характеристики случайных величин было дано в [2]. В этом разделе впервые предлагается для вычисления коэффициента корреляции Пирсона использовать сразу два совмещенных ковариационных графа.

За месяц до окончания регулярного чемпионата по хоккею, накануне XX зимних Олимпийских игр, пять команд гарантировали себе место в первой восьмерке, а следовательно, получили преимущество своей площадки, по крайней мере, в стартовом раунде плей-офф. Если указанная пятерка продолжит выступать на том же уровне, то мы уже можем стать свидетелями «разборок» внутри этой самой «пятерки избранных», это видно из таблицы их очных встреч. Добавим в таблицу еще и общее число набранных командами очков (по состоянию на 30.01.06).

Таблица 60

	Команды	И	В	ВО	Н	ПО	П	Ш	О	Общие очки
1.	"Металлург" Мг	8	5	1	0	0	2	26-13	17	108
2.	"Ак - барс"	8	4	1	1	1	1	22-16	16	81
3.	"Авангард"	9	3	1	0	1	4	13-22	12	83
4.	"Локомотив"	8	3	0	1	1	3	20-20	11	85
5.	ЦСКА	9	2	0	0	0	7	23-33	6	77

Найдем выборочные средние случайных величин:  $X = \{\text{число набранных очков со всеми командами}\}$  и  $Y = \{\text{число набранных очков друг с другом}\}$ .

$$x^* = \frac{1}{5}(108 + 81 + 83 + 85 + 77) = 86,8$$

$$y^* = \frac{1}{5}(17 + 16 + 12 + 11 + 6) = 12,4$$

Используя ковариационный граф, вычислим коэффициент корреляции Пирсона между их общими результатами и результатами, показанными ими друг с другом.

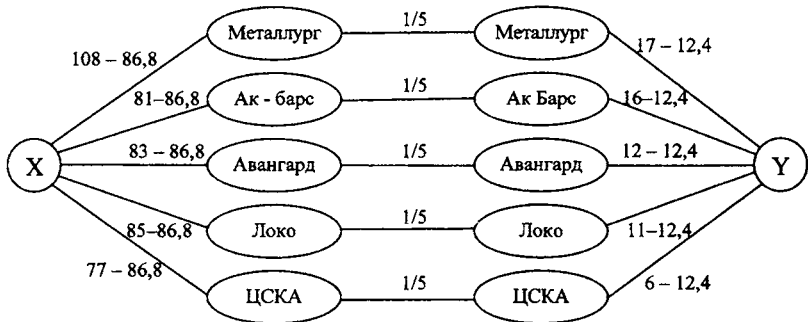


Рис. 13.

По ковариационному графу вычислим ковариацию и дисперсии:

$$k(X, Y) = \frac{1}{5}[(108 - 86,8)(17 - 12,4) + \dots + (77 - 86,8)(6 - 12,4)] = 28,68;$$

$$D_x^* = \frac{1}{5}(449,44 + 33,64 + 14,44 + 3,24 + 96,04) = 119,36;$$

$$D_y^* = \frac{1}{5}(21,16 + 12,96 + 0,16 + 1,96 + 43,56) = 15,96.$$

Тогда коэффициент корреляции Пирсона

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D_x^* \cdot D_y^*}} = \frac{28,68}{43,65} \approx 0,657.$$

Это означает, что общие результаты, показанные лучшей пятеркой команд суперлиги, коррелируют средне с их результатами между собой.

В этом же выпуске журнала «Локо - вперед!», № 23 (2005/2006 гг.), приводится таблица наиболее результативных российских команд в новом веке:

Таблица 61

Команды	99/00	00/01	01/02	02/03	03/04	04/05	05/06	Всего
"Металлург" Мг	175	205	166	124	204	206	150	<b>1230</b>
"Авангард"	136	163	184	188	203	208	107	<b>1189</b>
"Ак - барс"	190	146	175	166	179	180	120	<b>1156</b>
"Локомотив"	113	138	195	192	158	182	144	<b>1122</b>
"Динамо"	170	92	132	133	134	211	115	<b>987</b>

Аналогично предыдущим вычислениям найдем коэффициенты корреляции между числом заброшенных шайб за неполные 7 сезонов ( $X$ ) и числом шайб, заброшенных в первом сезоне XXI века ( $Y_1$ ), а также числом заброшенных шайб в не закончившемся еще последнем первенстве ( $Y_2$ ).

$$x^* \approx 1149, y_1^* \approx 157, y_2^* \approx 127.$$

Совместим ковариационные графы для пар случайных величин  $(X, Y_1)$  и  $(X, Y_2)$  следующим образом:

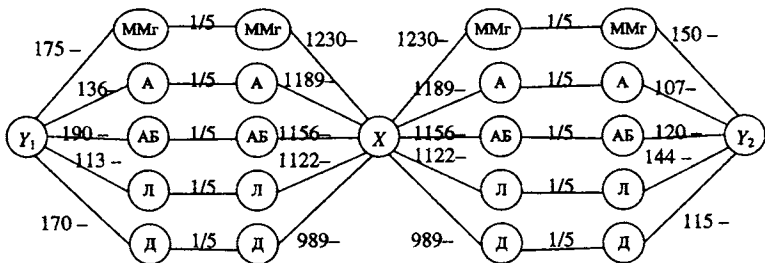


Рис. 14.

Выборочные ковариации находим по весу каждого из двух ковариационных графов:

$$k(X, Y_1) = \frac{1231}{5}; \quad k(X, Y_2) = \frac{6099}{5}.$$

По графам удобно вычислять и дисперсии трех случайных величин:

$$D_x^* = \frac{1}{5}(6561 + 1600 + 49 + 729 + 26244) = \frac{35183}{5}; \quad D_{Y_1}^* = \frac{3959}{5}; \quad D_{Y_2}^* = \frac{1411}{5}.$$

Найдем теперь искомые коэффициенты корреляции Пирсона

$$r(X, Y_1) = \frac{1231}{\sqrt{3959 \cdot 35183}} \approx 0,104; \quad r(X, Y_2) = \frac{6099}{\sqrt{1411 \cdot 35183}} \approx 0,866.$$

Откуда следует, что общая результативность за 7 сезонов лучших российских команд очень слабо связана с результативностью в сезоне 1999/2000 гг. и сильно коррелирует с их результативностью в последнем сезоне 2005/2006 гг.

Обратим внимание на то, что совмещенные ковариационные графы удобны и для нахождения коэффициента корреляции Пирсона случайных величин  $Y_1$  и  $Y_2$ . Вычислим выборочную ковариацию  $k(Y_1, Y_2)$  как бы «транзитом», через ковариационные графы двумерных случайных величин  $(Y_1, X)$  и  $(X, Y_2)$ .

$$k(Y_1, Y_2) = \left[ 18 \cdot \frac{1}{5} \cdot 23 + (-21) \cdot \frac{1}{5} \cdot (-20) + 33 \cdot \frac{1}{5} \cdot (-7) + (-44) \cdot \frac{1}{5} \cdot 17 + 13 \cdot \frac{1}{5} \cdot (-12) \right] = \frac{-301}{5}$$

$$\text{Тогда коэффициент корреляции } r(Y_1, Y_2) = \frac{-301}{\sqrt{3959 \cdot 1411}} \approx -0,13.$$

В качестве мер связи более чем двух случайных величин используются коэффициенты парной корреляции, которые удобно интерпретировать на графе. **Корреляционным графом** назовем множество вершин, обозначающих переменные, соединенных попарно ребрами с проставленными рядом соответствующими им коэффициентами корреляции.

В нашем случае корреляционный граф выглядит следующим образом:

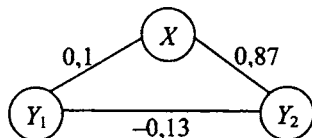


Рис. 15.



Проанализируем далее результаты команд за тот же временной отрезок, которые наиболее преуспели в оборонительных результатах. Интересующийся хоккеем болельщик назовет прежде всего тольяттинскую «Ладу». Действительно, «Лада» остается единственной командой суперлиги, пропустившей менее 100 шайб в каждом из семи последних регулярных чемпионатов.

Таблица 62

Команды	99/00	00/01	01/02	02/03	03/04	04/05	05/06	Всего
"Лада "	71	89	77	72	95	86	84	<b>574</b>
"Локомотив"	59	84	80	90	123	104	91	<b>631</b>
"Динамо"	62	105	98	102	108	106	110	<b>691</b>
"Ак - барс"	82	84	88	106	122	113	102	<b>697</b>
"Металлург" Мг	96	96	125	101	129	124	65	<b>736</b>

Пусть  $X = \{\text{число пропущенных шайб за 7 сезонов}\}$ , а  
 $Y_1 = \{\text{число пропущенных шайб за первый сезон XXI века}\}$  и  
 $Y_2 = \{\text{число пропущенных шайб за последний сезон}\}$ .

Найдем выборочные средние и дисперсии указанных случайных величин:

$$x^* = \frac{1}{5}(574 + 631 + 691 + 697 + 736) = 665,8 \approx 666, \quad D_x^* = \frac{16175}{5};$$

$$y_1^* = \frac{1}{5}(71 + 59 + 62 + 82 + 96) = 74, \quad D_{y_1}^* = \frac{846}{5};$$

$$y_2^* = \frac{1}{5}(84 + 91 + 110 + 102 + 65) = 90,4 \approx 90, \quad D_{y_2}^* = \frac{1162}{5}.$$

Построим совмещенный ковариационный граф случайных величин  $X$ ,  $Y_1$ ,  $Y_2$ .

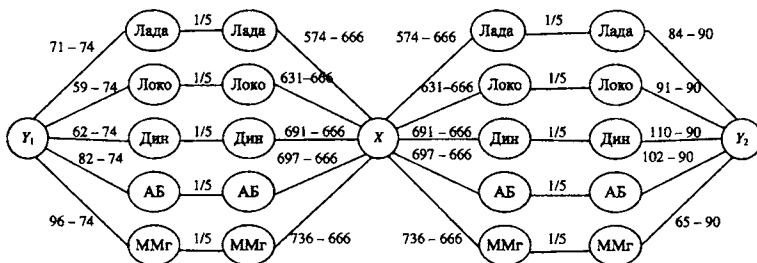


Рис. 16.

По совмещенному ковариационному графу вычислим три выборочные ковариации:

$$k(X, Y_1) = \frac{1}{5}[(-3) \cdot (-92) + (-15) \cdot (-35) + (-8) \cdot 25 + 8 \cdot 31 + (-22) \cdot 70] = \frac{-691}{5}$$

$$k(X, Y_2) = \frac{1}{5}[(-92) \cdot (-6) + (-35) \cdot 1 + 25 \cdot 20 + 31 \cdot 10 + 70 \cdot (-25)] = \frac{-425}{5};$$

$$k(Y_1, Y_2) = \frac{1}{5}[(-3) \cdot (-6) + (-15) \cdot 1 + (-8) \cdot 25 + 8 \cdot 10 + (-22) \cdot 70] = \frac{-643}{5}.$$

Найдем коэффициенты корреляции Пирсона и построим корреляционный граф:

$$r(X, Y_1) = \frac{-691}{\sqrt{16175 \cdot 846}} = -\frac{691}{3699} \approx -0,19;$$

$$r(X, Y_2) = \frac{-425}{\sqrt{16175 \cdot 1162}} = -\frac{425}{4335} \approx -0,1;$$

$$r(Y_1, Y_2) = \frac{-643}{\sqrt{846 \cdot 1162}} = -\frac{643}{992} \approx -0,65.$$

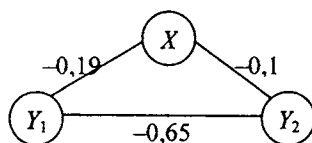


Рис. 17.

По корреляционному графу видно, что число пропущенных рассматриваемыми командами шайб за семь сезонов практически не зависит как от пропущенных шайб за первый, так и за последний сезон. Пропущенные же шайбы в сезоне 1999 - 2000 гг. имеют средней силы обратную связь ( $r = -0,65$ ) с пропущенными шайбами в последнем сезоне. Предлагаемый подход может быть продолжен и для анализа большего числа показателей, и для анализа других статистических показателей.

### Глава III. Бюджеты клубов и их спортивные достижения

Существуют два крайних расхожих утверждения. По первому – деньги в спорте, да и в жизни в целом, решают все, а по второму – деньги ничего не решают. В этой главе, используя методы математической статистики, находим и анализируем соответствие бюджетов футбольных, хоккейных и волейбольных клубов их спортивным результатам. Причем анализируются не отдельные клубы, а лиги в целом. Исследование проводилось по единой схеме. Определялся коэффициент вариации бюджетов клубов, по которому можно судить о репрезентативности выборки, другими словами в нашем случае – о финансовой справедливости начальных условий проведения того или иного чемпионата. После этого вычислялся коэффициент корреляции Пирсона или ранговый коэффициент Спирмена (см. [2], § 35, 36) бюджетов клубов и их спортивных достижений. В том случае, когда рассматривалось несколько спортивных результатов, то определялся коэффициент конкордации.

#### §8. Самый популярный вид спорта

В этом параграфе определим зависимость бюджетов российских и европейских футбольных клубов от их спортивных результатов за последние два сезона. Запишем в таблицу 63 итоги последнего закончившегося чемпионата России по футболу и первого круга 2005 года, проанализировав полученные результаты и бюджеты. Напомним, что, как и в § 4, командам, набравшим одинаковое число очков, присваивается один и тот же ранг, равный среднему арифметическому их возможного деления мест.

#### Турнирные таблицы и бюджеты клубов Российской футбольной премьер-лиги за 2004 год и первый круг 2005 года

Таблица 63

Итог 2004 года								2005 год		
М	Команды	Всего			В гостях		Бюджет 2004 г.		Бюджет 2005 г. (млн.\$)	Итог 1-го круга (Очки)
		очки	очки	ранг	очки	ранг	(млн.\$)	ранг		
(1)				(2)		(3)		(4)		
1	"Локомотив"	61	28	5	33	1	30	3	50	33
2	ЦСКА	60	30	4	30	2,5	50	2	50	31
3	"Крылья Советов"	56	32	1,5	24	4	18	6,5	23	19

4	"Зенит"	56	26	6	30	2,5	20	5	25	25
5	"Торпедо"	54	32	1,5	22	5,5	12	10	12	18
6	"Шинник"	44	31	3	13	9	11	11,5	12	16
7	"Сатурн"	41	19	12,5	22	5,5	22	4	30	19
8	"Спартак"	40	21	9	19	7	52	1	50	25
9	"Москва"	40	25	7	15	8	18	6,5	15	23
10	"Рубин"	33	24	8	9	14,5	11	11,5	14	26
11	"Амкар"	30	20	10,5	10	12	5	16	6	15
12	"Ростов"	29	19	12,5	10	12	9	13	11	10
13	"Динамо" (М)	29	20	10,5	9	14,5	15	8	55	22
14	"Алания"	28	16	15	12	10	7	14	8	14
15	"Кубань" (Терек)	28	18	14	10	12	13	9	30	9
16	"Ротор" (Томь)	22	14	16	8	16	6	15	8	15

8.1. Найдем матрицу ранговых корреляций Спирмена (как это делали в [8]) между общими результатами команд, результатами, показанными этими же командами дома и в гостях, и бюджетами клубов в 2004 году.

$$(r_s) = \begin{matrix} \text{общие} \\ \text{дома} \\ \text{гостях} \\ \text{бюджет} \end{matrix} \begin{pmatrix} 1 & 0,87 & 0,9 & 0,67 \\ & 1 & 0,66 & 0,44 \\ & & 1 & 0,73 \\ & & & 1 \end{pmatrix}$$

По матрице видно, что самая слабая связь существует между результатами команд, показанными дома, и бюджетами клубов, коэффициент корреляции равен только 0,44 (!) По всей вероятности, эта слабая зависимость обусловлена психологией игроков и тренеров команд, степенью их ответственности перед своим зрителем. Обратим внимание и на то, что только средней силы прямая связь ( $r = 0,67$ ) между бюджетами клубов и их общими итоговыми результатами в чемпионате 2004 года по футболу. Иллюстрацией достаточно средней связи между спортивными результатами клубов и их бюджетами служит тот факт, что у «Шинника», «Амкара», «Торпедо» и «Крыльев Советов» результаты на футбольном поле более чем на три ранга лучше их бюджетов (честь и хвала этим клубам), но зато у такого же количества команд («Спартак», «Кубань», «Динамо» и «Сатурн») аналогичные «достижения» прямо

противоположные. Причем, что интересно, «Спартак» на следующий сезон сократил свой бюджет и при этом занимает лидирующие позиции после первого круга, а «Динамо» и «Сатурн» значительно увеличили бюджеты и практически провалили первую половину чемпионата 2005 года, а по его ходу заменили своих главных тренеров, известных специалистов Романцева О. И. и Тарханова А. Ф. Отсюда особенно явно видно, что на спортивные достижения не прямопропорционально влияют бюджеты команд, а в некоторых случаях это влияние является просто обратным.

8.2. Для определения степени зависимости одновременно всех четырех рассмотренных показателей 2004 года вычислим коэффициент конкордации (множественный коэффициент ранговой корреляции) и сравним его с коэффициентом конкордации трех спортивных достижений команд (без учета бюджетов клубов).

Составим расчетную таблицу для определения коэффициента конкордации.

Таблица 64

Сумма четырех рангов $\sum_{i=1}^{16} R_{ij} / (1), (2), (3), (4) /$	$(\sum R_{ij} - T)^2$	Сумма трех рангов $\sum_{i=1}^{16} R_{ij} / (1), (2), (3) /$	$(\sum R_{ij} - T)^2$
10	576	7	242,25
10,5	552,25	8,5	289
15	361	8,5	289
17,5	272,5	12,5	169
22	144	12	182,25
28,5	30,25	18	56,25
29	25	25	0,25
25	81	24	2,25
30,5	12,5	24	2,25
44	100	32,5	49
49	225	33	56,25
49,5	240,25	36,5	121
46	144	38	156,25
53	361	39	182,25
50	256	41	240,25
65	441	48	650,25
$S_1$	3721,75	$S_2$	2787,75

Коэффициент конкордации находится по формуле:

$$W = \frac{12S}{m^2 \cdot (n-1) \cdot n(n+1)}$$

где  $S$  – сумма квадратов отклонений  $m$  рангов от их средней величины  $T$ ;

$m$  – число ранжируемых признаков (в нашем случае 4 и 3);

$n$  – число ранжируемых единиц (число команд).

В нашем случае:

$$а) n = 16, m_1 = 4, T_1 = \frac{1}{16}(1+2+\dots+16) \cdot 4 = 34, \quad W_1 = \frac{12 \cdot 3721,75}{16 \cdot 15 \cdot 16 \cdot 17} \approx 0,684.$$

$$б) n = 16, m_2 = 3, T_2 = \frac{1}{16}(1+2+\dots+16) \cdot 3 = 25,5 \text{ и}$$

$$W_2 = \frac{12 \cdot 2787,75}{9 \cdot 15 \cdot 16 \cdot 17} \approx 0,911.$$

Коэффициент конкордации  $W$  может принимать значение от 0 до 1. Полученные значения  $W_1 = 0,684$  и  $W_2 = 0,911$  позволяют сделать вывод о средней зависимости между четырьмя рассматриваемыми показателями (три спортивных и бюджет) и сильной зависимости трех только спортивных достижений. Отсюда можно сделать вывод о более сильной связи чисто спортивных показателей по сравнению со смешанными показателями (спортивные результаты и бюджеты клубов).

8.3. Найдем степень зависимости бюджетов клубов от результатов команд премьер-лиги в первом круге чемпионата 2005 года. С этой целью вычислим коэффициент корреляции Пирсона:

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D[X] \cdot D[Y]}},$$

где дисперсии  $D[X] \approx 253,88$  и  $D[Y] \approx 43,81$ , а ковариация

$$\begin{aligned} k(X, Y) &= M[(X - \bar{x}) \cdot (Y - \bar{y})] = \frac{1}{16} \sum_{i=1}^{16} (x_i - \bar{x})(y_i - \bar{y}) = \\ &= \frac{1}{16} [(50 - 24,9)(33 - 20) + (50 - 24,9)(31 - 20) + \dots + (8 - 24,9)(15 - 20)] \approx 69,51. \end{aligned}$$

Исходим из того, что средний бюджет клубов  $\bar{x} = 24,9$ , а среднее число очков  $\bar{y} = 20$ .

$$\cdot \cdot \cdot \text{Тогда } r(X, Y) = \frac{69,51}{\sqrt{253,9 \cdot 43,8}} \approx 0,66 \text{ и}$$

коэффициент корреляции практически совпадает с ранговым коэффициентом Спирмена (см. п. 8.1.) для соответствующих показателей по итогам предыдущего чемпионата, что свидетельствует о некоторой устойчивости зависимости футбольных результатов от финансирования российской премьер-лиги.

8.4. Найдем, насколько разнятся бюджеты «богатых» и «бедных» клубов российской футбольной премьер-лиги. Для этой цели используем коэффициент вариации  $C_s$ , который вычисляется по формуле

$$C_s = \frac{S_x}{\bar{x}} \cdot 100\% = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n(n-1)}} \cdot \frac{100\%}{\bar{x}}$$

Подсчитав предварительно средние бюджеты  $\bar{x}^1$  и  $\bar{x}^2$  клубов в 2004 и 2005 годах ( $\bar{x}^1 = 19,3$  и  $\bar{x}^2 = 24,9$ ), найдем коэффициенты вариации в эти годы:

$$C_s^1 = \sqrt{\frac{(30-19,3)^2 + (50-19,3)^2 + \dots + (6-19,3)^2}{16 \cdot 15}} \cdot \frac{100\%}{19,3} \approx 18,5\%;$$

$$C_s^2 = \sqrt{\frac{(50-24,9)^2 + (50-24,9)^2 + \dots + (8-24,9)^2}{16 \cdot 15}} \cdot \frac{100\%}{24,9} \approx 17\%.$$

Что свидетельствует о почти одинаковой и достаточно большой степени отклонения бюджетов клубов от своих средних значений. Показатель  $C_s$  считается в статистике удовлетворительным, если он варьирует в пределах до 10%.

8.5. Определим зависимость спортивных результатов от бюджетов лучших европейских футбольных клубов. С этой целью используем данные одной из известных аудиторских английских компаний, которая провела исследование доходов футбольных клубов в 2004 году и получила следующие результаты (в млн. фунтов стерлингов):

Таблица 65

№	Клуб	Доход	Рейтинг
1.	«Манчестер Юнайтед» (Англия)	171,5	2152
2.	«Реал» (Испания)	156,3	2205
3.	«Милан» (Италия)	147,2	2380
4.	«Челси» (Англия)	143,7	2272
5.	«Ювентус» (Италия)	142,4	2276
6.	«Арсенал» (Англия)	115,0	2300
7.	«Барселона» (Испания)	112,0	2148
8.	«Интер» (Италия)	110,3	2104
9.	«Бавария» (Германия)	110,1	2146
10	«Ливерпуль» (Англия)	92,3	1823
	Средние значения	130,1	2181

В таблицу внесены и спортивные рейтинги этих клубов. Для определения степени зависимости бюджетов самых богатых клубов Европы и их спортивных достижений в 2004 году вычислим коэффициент корреляции Пирсона:

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D[X] \cdot D[Y]}},$$

где ковариация

$$k(X, Y) = \frac{1}{10} \sum_{i=1}^{10} (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{10} [(171,5 - 130,1) \cdot (2152 - 2181) + (156,3 - 130,1) \cdot (2205 - 2181) + \dots + (92,3 - 130,1) \cdot (1823 - 2181)] \approx 1979,5,$$

а дисперсии величины доходов и рейтингов следующие:

$$D[X] = \frac{1}{10} [(171,5 - 130,1)^2 + (156,3 - 130,1)^2 + \dots + (92,3 - 130,1)^2] \approx 580,55,$$

$$D[Y] = \frac{1}{10} [(2152 - 2181)^2 + (2205 - 2181)^2 + \dots + (1823 - 2181)^2] \approx 20889,2.$$

Тогда  $r(X, Y) = \frac{1979,5}{\sqrt{580,55 \cdot 20889,2}} \approx 0,568$ , что указывает на наличие

только средней прямой связи между бюджетами самых богатых футбольных клубов Европы и их спортивными достижениями, выраженными в рейтингах.

8.6. Коэффициент вариации бюджетов 10 европейских клубов находим, как и в п. 8.4. для российских клубов.

$$C_x = \sqrt{\frac{580,55}{9}} \cdot 100\% \approx 6,17\%, \text{ что свидетельствует о репрезентативности}$$

выборки десяти европейских футбольных грантов, а их спортивные результаты в большей степени, чем в России, соответствуют их спортивному уровню, а не финансированию.

## §9. Бюджеты хоккейных клубов и их результаты

Рассмотрим хоккейные клубы суперлиги, у которых обнародованы их бюджеты. Внесем в таблицу 66 бюджеты этих клубов и набранные ими очки в регулярном чемпионате 2004 - 2005 гг.



Таблица 66

№	Команды	Бюджет команды	Набранные очки
		(млн.\$)	
		X	Y
1	"Динамо"	20	126
2	"Лада"	18	118
3	"Металлург" Мг.	24	115
4	"Ак – барс"	50	114
5	"Локомотив"	19	105
6	"Авангард"	35	104
7	"Химик"	20	83
8	ЦСКА	17	81
9	"Северсталь"	23	80
10	СКА	10	76
11	"Спартак"	8	42
	<b>Средние значения</b>	22	95

### 9.1. Коэффициент вариации бюджетов клубов суперлиги.

$$C_s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n(n-1)}} \cdot \frac{100\%}{\bar{x}} =$$

$$= \sqrt{\frac{(20-22)^2 + (18-22)^2 + \dots + (8-22)^2}{11 \cdot 10}} \cdot \frac{100\%}{22} \approx 15\%,$$

что свидетельствует о достаточно большом разбросе бюджетов хоккейных клубов России, но зато несколько меньшем, чем у футбольных, что может считаться положительным отличием хоккейного чемпионата России от футбольного.

### 9.2. Корреляция между бюджетами клубов и набранными ими очками.

Вычислим коэффициент Пирсона по формуле

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D[X] \cdot D[Y]}},$$

где ковариация  $k(X, Y) = \frac{1}{11} \sum_{i=1}^{11} (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{11} [(20-22) \cdot (126-95) + (18-22) \cdot (118-95) + \dots + (8-22) \cdot (42-95)] \approx 143,6,$

$$\text{дисперсии } D[X] = \frac{1}{11} [(20-22)^2 + (18-22)^2 + \dots + (8-22)^2] \approx 124,1;$$

$$D[Y] = \frac{1}{11} [(126-95)^2 + (118-95)^2 + \dots + (42-95)^2] \approx 560,4.$$

$$\text{Тогда } r(X, Y) = \frac{143,6}{\sqrt{124,1 \cdot 560,4}} \approx 0,545, \text{ что показывает наличие средней}$$

силы связи между бюджетами хоккейных клубов суперлиги и их результатами в регулярном чемпионате.

9.3. Рассмотрим еще один пример, сопоставив очки, набранные командами НХЛ по итогам сезона 2003 - 2004 года (заметим, что в 2004 -2005 гг. в НХЛ был локаут), с бюджетами этих команд (и их вариацию).

### Бюджеты клубов и очки, набранные командами НХЛ

Таблица 67

№	Команды	Бюджет команды (млн. \$)	Набранные очки
1	"Детройт Ред Уингз"	79	109
2	"Сан-Хосе Шаркс"	35	104
3	"Бостон Брюинз"	46	104
4	"Филадельфия Флайерз"	65	101
5	"Колорадо Эвеланш"	61	100
6	"Нэшвилл Предэйторс"	23	91
7	"Лос-Анджелес Кингз"	46	81
8	"Атланта Трэшерз"	27	78
9	"Нью-Йорк Рейнджерс"	77	69
10	"Вашингтон Кэпиталс"	51	59
11	"Питсбург Пингвинз"	27	58
	Средние значения	49	87

Вычислим коэффициент корреляции

$$r(X, Y) = \frac{k(X, Y)}{\sqrt{D[X] \cdot D[Y]}} =$$

$$= \frac{\frac{1}{11} [30 \cdot 22 + (-14) \cdot 17 + \dots + (-22) \cdot (-29)]}{\frac{1}{11} \sqrt{[30^2 + (-14)^2 + \dots + (-22)^2] \cdot [22^2 + 17^2 + \dots + (-29)^2]}} \approx 0,27$$

Найденное значение выборочного коэффициента корреляции ( $r = 0,27$ ) показывает, что в данном случае между бюджетами клубов и результатами команд (набранными в ходе регулярного чемпионата очками) существует слабая прямая связь. Такой результат стимулирует зрителей, поскольку свидетельствует о непредсказуемости большинства матчей, и обусловлен всей системой организации чемпионата НХЛ и в первую очередь порядком проведения ежегодных драфтов хоккеистов (когда слабейшие клубы предыдущего чемпионата имеют преимущество в выборе новичков).

Определим также коэффициент вариации бюджетов клубов НХЛ.

$$C_s = \sqrt{\frac{(79 - 49)^2 + (35 - 49)^2 + \dots + (27 - 49)^2}{11 \cdot 10}} \cdot \frac{100\%}{49} \approx 11,44.$$

Если сравнивать полученные коэффициенты вариации, то можно сделать вывод, что и в суперлиге, и в НХЛ бюджеты клубов распределены не совсем справедливо, и справедливость распределения бюджетов в суперлиге меньше, чем в НХЛ.

Заметим, что недавно закончившийся в НХЛ локаут и достигнутое соглашение об ограничении бюджетов клубов приведет и к снижению их коэффициента вариации, а следовательно, спортивные результаты будут в большей степени зависеть от мастерства хоккеистов и их тренеров, от организации учебно-тренировочного процесса.

### §10. А как дела в волейболе?

Приведем таблицу результатов сезона 2004 - 2005 гг. волейбольной суперлиги, бюджеты команд и их ранги.

Таблица 68

Место	Команды	Бюджет (млн.\$)	Ранг
1	«Локомотив-Белогорье» (г. Белгород)	10,0	2
2	«Динамо» (г. Москва)	15,0	1
3	«Динамо» (г. Казань)	5,0	5
4	«Искра» (г. Одинцово)	5,5	4
5	«Луч» (г. Москва)	9,0	3
6	«Локомотив-Изумруд» (г. Екатеринбург)	4,5	6
7	«Нова» (г. Новокуйбышевск)	3,0	8
8	«ЗСК-Газпром» (Сургутский район)	3,2	7

9	«Нефтяник-Башкортостана» (г. Уфа)	2,5	10
10	«Нефтяник» (г. Ярославль)	0,8	13
11	«Факел» (г. Новый Уренгой)	2,6	9
12	«Локомотив» (г. Новосибирск)	1,5	11
13	«Спартак» (г. С.-Петербург)	0,8	13
14	МГТУ (г. Москва)	0,8	13

10.1. Найдем средний бюджет клубов суперлиги:

$$\bar{x} = \frac{1}{14} (10+15+5+5,5+9+4,5+3+3,2+2,5+0,8 \cdot 3+2,6+1,5) \approx 4,6,$$

тогда коэффициент вариации

$$C_s = \sqrt{\frac{\sum (x_i - 4,6)^2}{14 \cdot 13} \cdot \frac{100\%}{4,6}} \approx \frac{1,16}{4,6} \cdot 100\% \approx 25,2\%,$$

который оказывается самым большим из рассматриваемых и указывает на очень большой разброс в финансировании российских волейбольных клубов.

10.2. Найдем коэффициент ранговой корреляции  $r_s$  Спирмена между местами команд и их бюджетами, предварительно вычислив сумму квадратов соответствующих рангов ( $\sum d_i^2 = 28$ ).

$$r_s = 1 - \frac{6 \sum d_i^2}{13 \cdot 14 \cdot 15} = 1 - \frac{6 \cdot 28}{13 \cdot 14 \cdot 15} \approx 0,94$$

Получили, что волейбольное первенство в России самое несправедливое ( $C_s = 25\%$ !) и самое предсказуемое ( $r_s = 0,94$  и близко к единице, что означает решающую роль денег). Говоря проще можно утверждать, что первенство России по волейболу в суперлиге можно не проводить, а медали выдавать в соответствии размерами бюджетов клубов.

Предложенный подход может распространяться и на другие виды спорта, а также следует регулярно отслеживать рассмотренные показатели и по ним, возможно, спортивным федерациям вносить коррективы в организацию первенств и вводить общие правила ведения хозяйства спортивных клубов.

## Глава IV. Линии регрессии и прогнозы в спорте

В спорте, как и в других видах деятельности человека (или машины), часто возникает желание предсказать ожидаемый результат. Проводится масса лотерей, спортпрогнозов, делаются ставки на предсказания результатов спортивных соревнований, происхождение тех или иных событий, достижений. Умение прогнозировать является ценным качеством специалиста и часто мотивируется наличием у него развитой интуиции. Не принижая значения этого качества, считаем все же более важным для специалиста умение получать и перерабатывать соответствующую информацию. Чем больше получено информации и она является объективной, чем грамотней специалист, который умеет отбирать самую ценную информацию, тем больше вероятность совпадения прогноза с дальнейшим результатом.

Совпадение прогноза и результата не самоцель, если это не лотерея. Более важна в жизни частота и близость полученного результата и прогноза, что позволяет ставить реальные задачи и стремиться к их достижению. Надо понимать, что чудеса в жизни, да и в спорте, в частности, встречаются крайне редко. Поэтому все ожидания российских футбольных болельщиков увидеть сборную нашей страны призером чемпионата мира или Европы не более чем несбыточная мечта, поскольку отсутствуют какие-либо систематические достижения как на клубном уровне, так и тем более на уровне сборных.

Имидж команды, ее мощь и сила – результат долгой и кропотливой работы тренеров, администраторов, президентов клубов и, разумеется, самих игроков. В каждой команде есть свои «звезды», любимцы публики, зарабатывающие свой авторитет не только числом забитых мячей или шайб, но и красивой игрой. Такие лидеры не только приносят очки, но и способствуют поддержанию боевого духа команды, стимулируют молодых, менее опытных игроков. Поэтому задача президента клуба – пригласить в свою команду самых результативных, сильных и харизматичных игроков. Но ведь игрок – человек, а не машина, и такие факторы, как везение, удача, настроение, обстановка в семье, и многие другие влияют на его игру. Подписав контракт с самым результативным игроком прошедшего сезона, нельзя быть уверенным на все 100%, что он покажет такие же результаты и в новом сезоне.

Вероятностно обоснованными являются прогнозы, построенные на анализе статистических данных, которые были показаны на протяжении нескольких лет или в ряде замеров или испытаний. Наиболее подходящие результаты дают линии регрессии, построенные из установленной корреляционной зависимости случайных величин.

**Корреляционной зависимостью**  $Y$  от  $X$  называют функциональную зависимость условной средней  $y_x^*$  от  $x$ .

$y_x^* = f(x)$  представляет **уравнение регрессии**  $Y$  на  $X$ , а  $x_y^* = \varphi(y)$  – уравнение регрессии  $X$  на  $Y$ .

Корреляционная зависимость может быть линейной и криволинейной. В случае линейной корреляционной зависимости выборочное уравнение прямой линии регрессии  $Y$  на  $X$  имеет вид:

$$y_x^* - y^* = \frac{k(X,Y)}{D_x^*} (x - x^*),$$

где  $x^*, y^*$  – выборочные средние величины  $X$  и  $Y$ ,  $k(X,Y)$  – выборочная ковариация этих случайных величин, а  $D_x^*$  – выборочная дисперсия величина  $X$ .

## §11. Прямые регрессии и прогнозы в хоккее

Построим по результатам последних четырех сезонов прямые регрессии числа очков некоторых известных российских хоккеистов и мест ведущих команд Суперлиги. По полученным прямым можно делать вероятностные прогнозы на следующие сезоны по результативности хоккеистов и выступления команд в чемпионатах, что может быть весьма полезно для специалистов и менеджеров. Сравнение показанных результатов с прогнозами позволяет делать вывод о прогрессе или регрессии как самих хоккеистов, так и команд в целом.

### 11.1. Прогноз результативности хоккеистов

#### Результативность некоторых хоккеистов по итогам четырех последних чемпионатов России

Таблица 69

№	Игрок	Команда	Очки (гол + пас) сезона			
			2001-2002	2002-2003	2003-2004	2004-2005
1	Сушинский М.	"Авангард"	66	41	69	67
2	Затонский Д.	"Авангард"	48	42	52	52
3	Антипов В.	"Локомотив"	23	37	25	40
4	Непряев И.	"Локомотив"	11	11	10	21
5	Коваленко А.	"Локомотив" "Авангард"	54	39	34	21
6	Зиновьев С.	"Спартак" "Ак - барс"	29	33	14	38
7	Королев И.	"Локомотив"	34	9	22	8
8	Буцаев В.	"Динамо"	27	37	16	11
9	Шахрайчук В.	"Динамо"	25	8	21	25

Покажем на примере результативности Максима Сушинского, как находится уравнение прямой регрессии очков по годам (1(2001 - 2002), 2, 3, 4) их набирания.

Средний показатель набранных очков  $y^*$  у Максима составил

$$y^* = \frac{1}{4}(66 + 41 + 69 + 67) = 60,75,$$

а среднее число рассматриваемых годов  $x^* = \frac{1}{4}(1 + 2 + 3 + 4) = 2,5$ .

Построим предложенный в ([2] с.178) ковариационный граф:

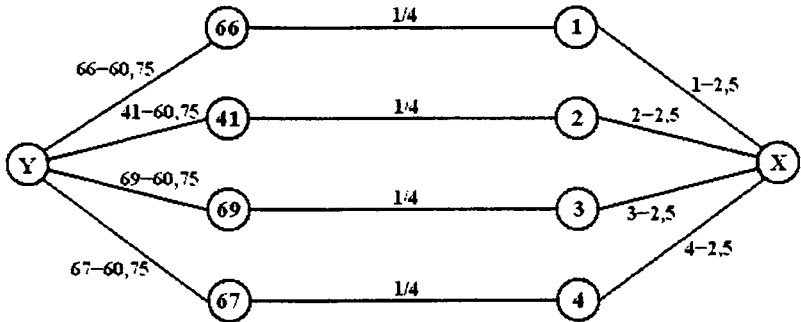


Рис.18. Ковариационный граф.

Ковариацию находим как вес всего ковариационного графа:

$$k(x,y) = \frac{1}{4} [5,25 \cdot (-1,5) + (-19,75) \cdot (-0,5) + 8,25 \cdot 0,5 + 6,25 \cdot 1,5] = 3,875.$$

Дисперсию тоже можно вычислить по графу:

$$D_x^* = \frac{1}{4} [(1-2,5)^2 + (2-2,5)^2 + (3-2,5)^2 + (4-2,5)^2] = 1,25.$$

Тогда уравнение прямой регрессии Y на X будет следующим:

$$y_x^* - 60,75 = \frac{3,875}{1,25} (x - 2,5) \text{ или } y_x^* = 3,1x + 53.$$

На следующий год ( $x = 5$ ) можно ожидать, используя полученную прямую регрессии, что Сушинский наберет

$$y_x^* = 3,1 \cdot 5 + 53 \approx 69 \text{ очков.}$$

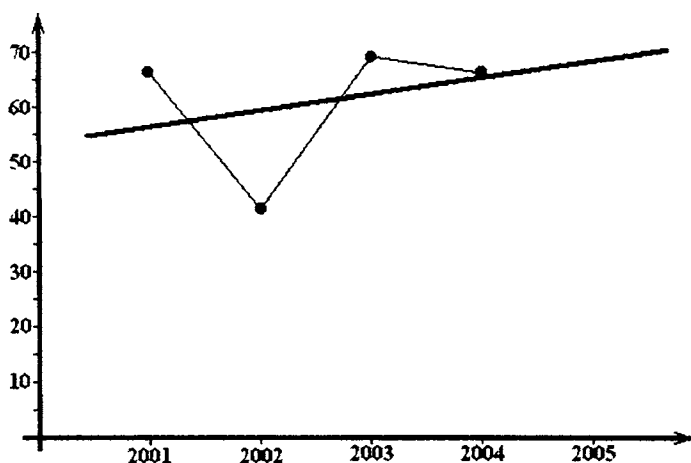


Рис. 19. Результативность М. Сушинского и прямая регрессии результативности по четырем последним первенствам России.

Аналогично находим и уравнения прямых регрессий для других перечисленных хоккеистов и ожидаемые количества очков в следующем сезоне 2005 - 2006 г.г.

### Уравнения прямых регрессии и ожидаемые результаты некоторых хоккеистов

Таблица 70

Игрок	Уравнение прямой регрессии	Ожидаемый результат на следующий сезон
Затонский Д.	$y_x = 2,2x + 43$	$y_{x=5} = 54$ (очка)
Антипов В.	$y_x = 5x + 18,8$	$y_{x=5} = 44$ (очка)
Непряев И.	$y_x = 2,9x + 6$	$y_{x=5} = 21$ (очко)
Коваленко А.	$y_x = -10,4x + 63$	$y_{x=5} = 11$ (очков)
Зиновьев С.	$y_x = 0,8x + 26,5$	$y_{x=5} = 31$ (очко)
Королев И.	$y_x = -6,5x + 34,5$	$y_{x=5} = 2$ (очка)
Буцаев В.	$y_x = -6,7x + 39,5$	$y_{x=5} = 6$ (очков)
Шахрайчук В.	$y_x = -2x + 30$	$y_{x=5} = 20$ (очков)



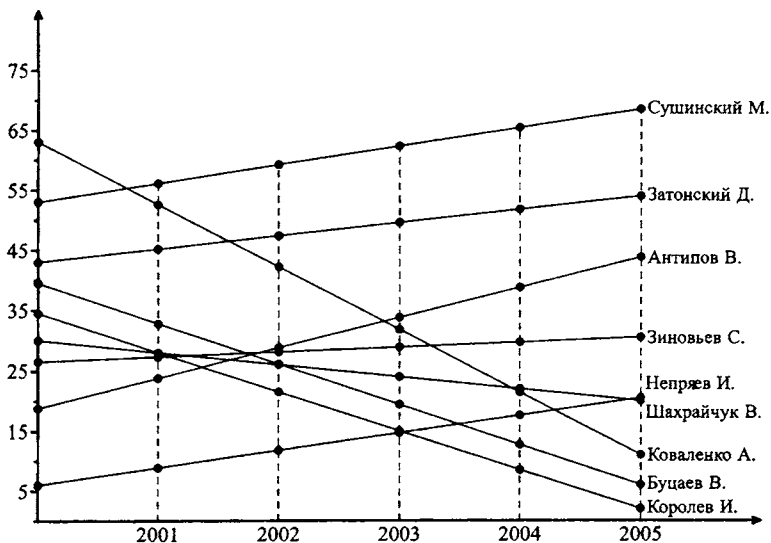


Рис. 20. Прямые регрессии набранных очков нападающих по четырем последним первенствам России.

По представленным графикам хорошо видны тенденции результативности нападающих. Самая хорошая ожидаемая перспектива у Владимира Антипова, поскольку прямая регрессии имеет самый большой угловой коэффициент ( $k = 5$ ). Достаточно высокие темпы результативности имеет Сушинский М. ( $k = 3,1$ ) и Непряев И. ( $k = 2,9$ ). Наихудшие тенденции за последние четыре года наметились у Коваленко А. ( $k = -10,4 < 0$ ), Буцаева В. ( $k = -6,7$ ) и Королева И. ( $k = -6,5$ ).

Построим прямые регрессии очков российских хоккеистов (защитников) по итогам последних четырех чемпионатов России, и перенесем полученные графики на соответствующие фотографии известных хоккеистов (см. стр. 97-104)



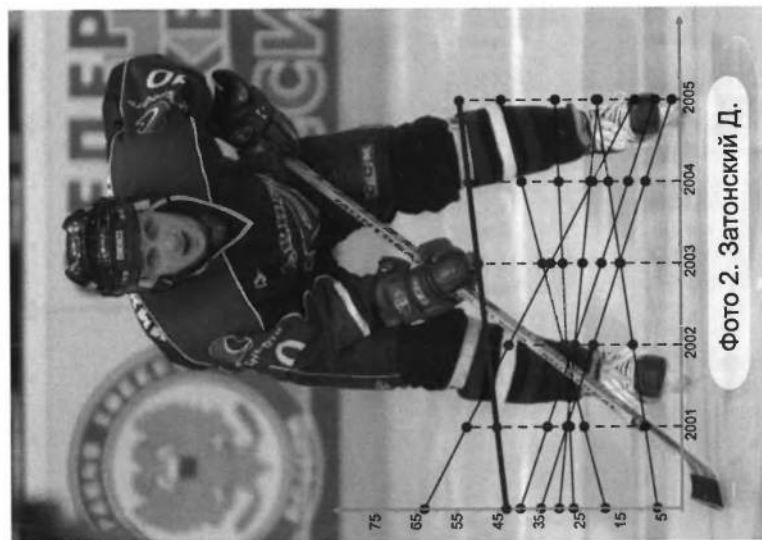
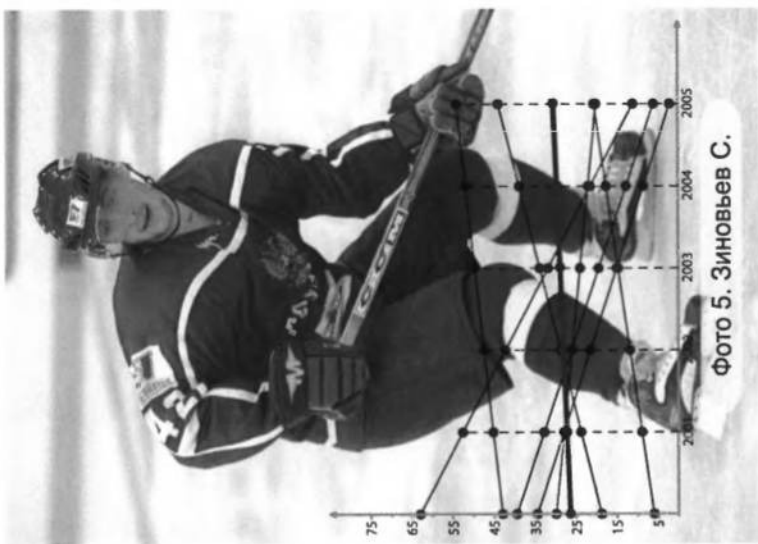
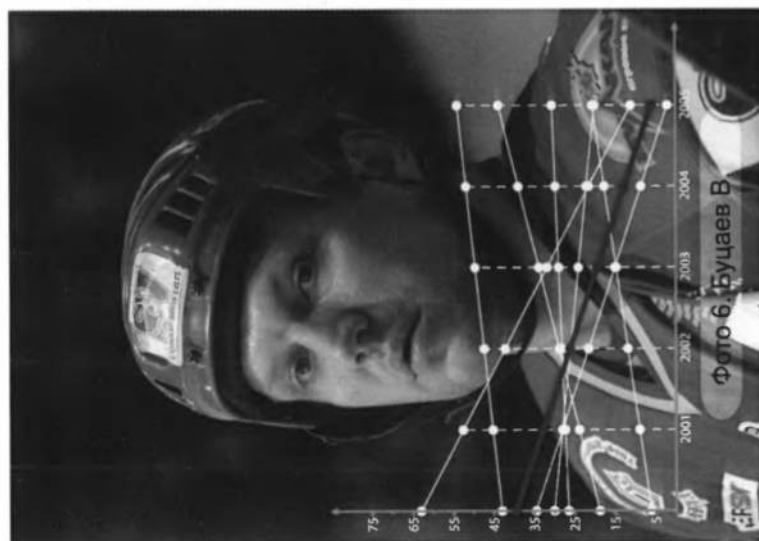


Фото 2. Затонский Д.

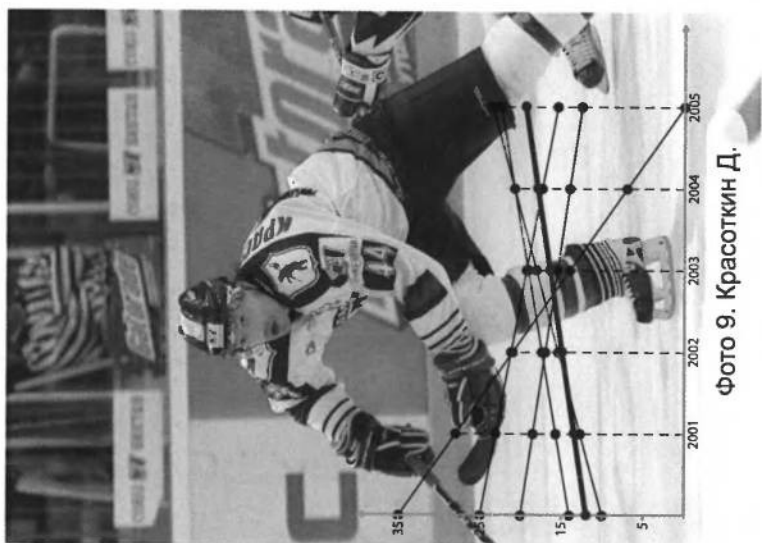
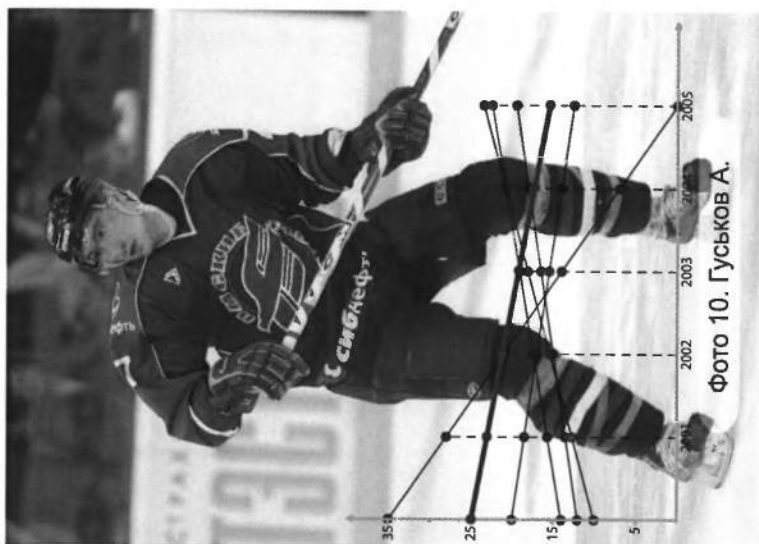


Фото 3. Антипов В.









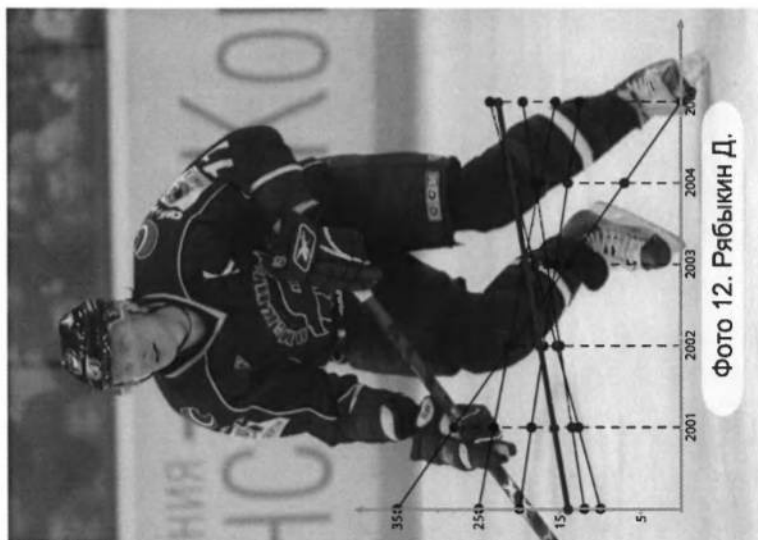


Фото 12. Рябыкин Д.

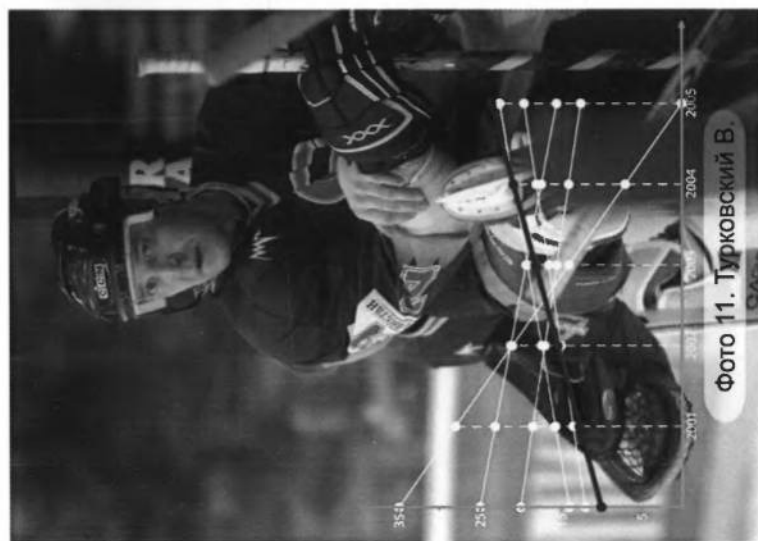
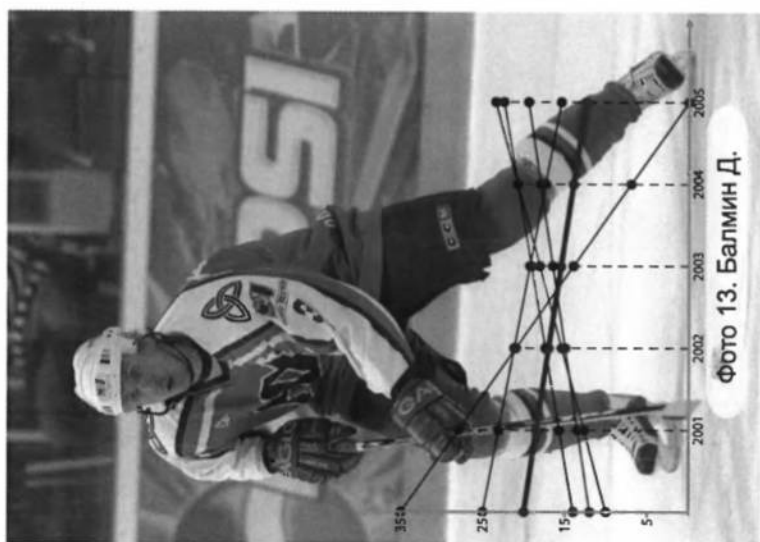


Фото 11. Турковский В.





**Результативность некоторых защитников  
по итогам четырех последних чемпионатов России**

Таблица 71

№	Игрок	Команда	Очки (гол + пас) сезона			
			2001-2002	2002-2003	2003-2004	2004-2005
1	Климентьев С.	"Металлург" Мг.	30	14	21	4
2	Гуськов А.	"Локомотив"	19	27	20	15
3	Балмин Д.	"Ак – барс"	18	16	19	12
4	Красоткин Д.	"Локомотив"	16	10	22	17
5	Рябыкин Д.	"Авангард"	10	26	19	18
6	Турковский В.	"Северсталь"	11	18	18	20

**Уравнения прямых регрессии и ожидаемые результаты  
некоторых защитников**

Таблица 72

Климентьев С.	$y_x = -7,1x + 35$	$y_{x=5} = 0$ (очков)
Гуськов А.	$y_x = -1,9x + 25$	$y_{x=5} = 16$ (очков)
Балмин Д.	$y_x = -1,5x + 20$	$y_{x=5} = 13$ (очков)
Красоткин Д.	$y_x = 1,5x + 12$	$y_{x=5} = 20$ (очков)
Рябыкин Д.	$y_x = 1,7x + 14$	$y_{x=5} = 23$ (очка)
Турковский В.	$y_x = 2,7x + 10$	$y_{x=5} = 24$ (очка)

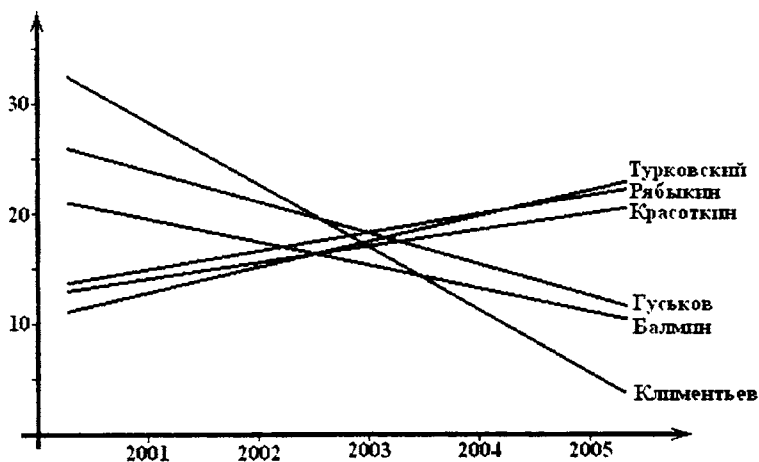


Рис. 21. Прямые регрессии набранных очков защитников по четырем последним первенствам России.

По представленным графикам видны тенденции результативности защитников. Самая благоприятная перспектива у Турковского Василия, поскольку прямая регрессии имеет самый большой коэффициент ( $k = 2,7$ ). Почти одинаковые перспективы у двух Дмитриев - Рябыкина и Красоткина, у которых прямые регрессии имеют соответственно коэффициенты  $k = 1,7$ ;  $k = 1,5$ . Наихудшие тенденции наметились у Клементьева С. ( $k = -7,1$ ) и у Гуськова А. ( $k = -2,3$ ).

### 11.2. Российские чемпионаты по хоккею и их прогнозы

Теперь построим прямые регрессии занятых мест командами Российской хоккейной Суперлиги в регулярных чемпионатах России за последние четыре сезона.

#### Результаты регулярных чемпионатов России за последние четыре сезона

Таблица 73

№	Команды	Место в сезоне			
		2001-2002	2002-2003	2003-2004	2004-2005
1	"Динамо"	7	7	6	1
2	"Лада"	4	5	2	2
3	"Металлург" Мн	5	6	1	3
4	"Ак – барс"	2	4	5	4
5	"Локомотив"	1	1	7	5
6	"Авангард"	3	2	3	6
7	"Металлург" Нк	12	9	4	7
8	"Нефтехимик"	13	12	8	8
9	ЦСКА	17	10	10	10

#### Уравнения прямых регрессии и ожидаемые результаты команд

Таблица 74

Команда	Уравнение прямой регрессии	Ожидаемый результат на следующий сезон
"Динамо"	$y_x = -1,9x + 10$	$y_{x=5} = 0,5 \approx 1$ (место)
"Лада "	$y_x = -0,9x + 5,5$	$y_{x=5} = 1$ (место)
"Металлург" Мн.	$y_x = -1,1x + 6,5$	$y_{x=5} = 1$ (место)
"Ак – барс"	$y_x = 0,7x + 2$	$y_{x=5} = 5,5 \approx 6$ (место)
"Локомотив"	$y_x = 1,8x - 1,5$	$y_{x=5} = 7,5 \approx 8$ (место)
"Авангард"	$y_x = x + 1$	$y_{x=5} = 6$ (место)
"Металлург" Нк.	$y_x = -1,4x + 11,5$	$y_{x=5} = 4,5 \approx 5$ (место)
"Нефтехимик"	$y_x = -1,9x + 15$	$y_{x=5} = 5,5 \approx 6$ (место)
ЦСКА	$y_x = -2,1x + 17$	$y_{x=5} = 6,5 \approx 7$ (место)

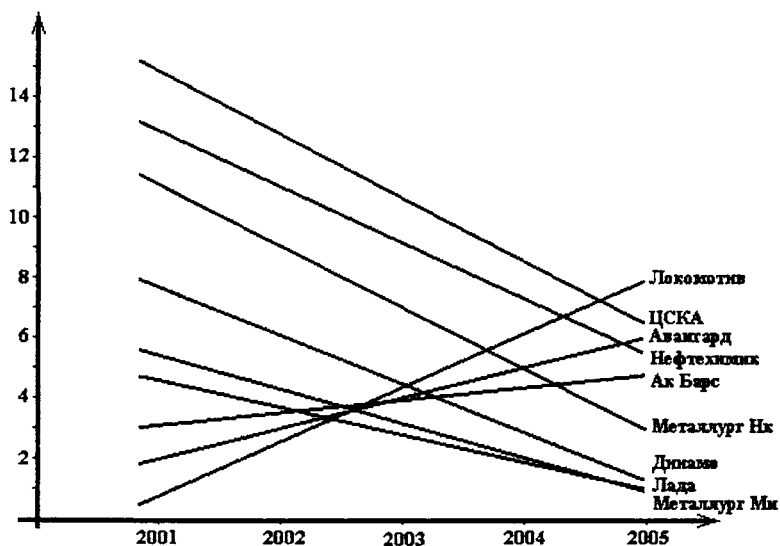


Рис. 22. Прямые регрессии мест команд по четырём последним первенствам России.

По представленным графикам самая благоприятная перспектива на новый сезон наблюдается у команд «Металлург Мг», «Динамо» и «Лада».

Следует уточнить, что прогнозы сделаны на основе статистических данных за последние четыре сезона и не учитывают возможные изменения в содержании клуба, его финансирования, новые приобретения и возможные форс-мажорные обстоятельства. Так, известные и печальные изменения в отношении тольяттинского АвтоВАЗа к своей хоккейной команде «Лада», конечно же, отразятся на ожиданиях и болельщиков команды, и специалистов. А знали бы владельцы клуба его возможности, в том числе и вероятностно-статистические!

### 11.3. Что мы ожидаем от чемпионата мира?

Очередной чемпионат мира по хоккею состоится в мае 2006 года в Латвии. В ожидании любого значимого соревнования болельщики делают прогнозы, а затем сверяют их с реальными результатами. Проанализируем выступления ведущих хоккейных дружин на последних пяти чемпионатах мира, сделаем прогноз и мы. Построим прямые регрессии

команд по результатам таблицы № 30 из главы II. Поскольку берется статистика последних пяти чемпионатов, то уравнение прямых регрессии будет иметь вид:

$$y_x^* - y^* = \frac{k(X, Y)}{2}(x-3),$$

где  $y^*$  - среднее место команды за последние 5 лет;

$y_x^*$  - ожидаемое место в год  $x$  и  $x^*=3$ , а  $D_x^*=2$ .

Сведем полученные результаты в следующую таблицу.

Таблица 75

№	Команда	Уравнение прямой регрессии	Ожидаемый результат на ЧМ-2006
1	Канада	$y_x^* = -1,1x + 6,3$	$y_{x=6}^* = -1,1 \cdot 6 + 6,3 \approx 1$ (место)
2	Швеция	$y_x^* = 0,1x + 3,1$	$y_{x=6}^* = 3,7 \approx 4$
3	Словакия	$y_x^* = -0,1x + 4,3$	$y_{x=6}^* = 3,7 \approx 4$
4	Чехия	$y_x^* = 3,2$	$y_{x=6}^* = 3,2 \approx 3$
5	Финляндия	$y_x^* = 5$	$y_{x=6}^* = 5$
6	Россия	$y_x^* = 0,2x + 5$	$y_{x=6}^* = 6,2 \approx 6$
7	США	$y_x^* = 6,6$	$y_{x=6}^* = 6,6 \approx 7$
8	Швейцария	$y_x^* = -0,6x + 10,2$	$y_{x=6}^* = 6,6 \approx 7$
9	Латвия	$y_x^* = -1,2x + 13,4$	$y_{x=6}^* = 6,4 \approx 6$
10	Украина	$y_x^* = 1,4x + 7$	$y_{x=6}^* = 15,4 \approx 15$
11	Словения	$y_x^* = -0,4x + 13,8$	$y_{x=6}^* = 11,4 \approx 11$
12	Белоруссия	$y_x^* = -0,7x + 14,7$	$y_{x=6}^* = 10,5 \approx 11$

## §12. Прогнозы и надежды российского футбола

Рассмотрим аналогично предыдущему параграфу статистику как отдельных футболистов, так и команд в целом.

Представляют интерес перспективы игроков российской футбольной премьер-лиги. Построим прямые регрессии для известных бомбардиров РФПЛ, используя статистику забитых ими голов за последние годы.

**12.1. Результативность российских футболистов  
Бомбардиры РФПЛ за последние пять сезонов**

Таблица 76

№	Фамилии игроков	Количество забитых голов				
		2001	2002	2003	2004	2005
1	<b>Кержаков А.</b> («Зенит»)	6	14	13	18	7
2	<b>Есипов В.</b> («Сатурн»)	11	7	11	10	4
3	<b>Павлюченко Р.</b> («Спартак»)	5	4	10	10	11
4	<b>Семшов И.</b> («Торпедо»)	6	11	6	9	12
5	<b>Кириченко Д.</b> («Москва»)	13	15	5	9	14
6	<b>Спивак А.</b> («Зенит»)	3	5	6	10	6
7	<b>Лебеденко И.</b> («Локомотив»)	0	6	3	9	6

**Уравнения прямых регрессии и ожидаемое количество голов в следующем сезоне**

Таблица 77

Фамилии игроков	Уравнение прямой регрессии	Ожидаемое количество голов на следующий сезон
<b>Кержаков А.</b>	$y_x^* = 0,6x + 9,8$	$y_{x=6}^* = 13,2 \approx 13$
<b>Есипов В.</b>	$y_x^* = -1,1x + 11,9$	$y_{x=6}^* = 5,3 \approx 5$
<b>Павлюченко Р.</b>	$y_x^* = 1,8x + 2,6$	$y_{x=6}^* = 13,4 \approx 13$
<b>Семшов И.</b>	$y_x^* = x + 5,8$	$y_{x=6}^* = 11,8 \approx 12$
<b>Кириченко Д.</b>	$y_x^* = -0,4x + 12,4$	$y_{x=6}^* = 10$
<b>Спивак А.</b>	$y_x^* = 1,1x + 2,7$	$y_{x=6}^* = 9,3 \approx 9$
<b>Лебеденко И.</b>	$y_x^* = 1,5x + 0,3$	$y_{x=6}^* = 9,3 \approx 9$

Для лучшей иллюстрации полученных результатов изобразим линии регрессии на координатной плоскости.

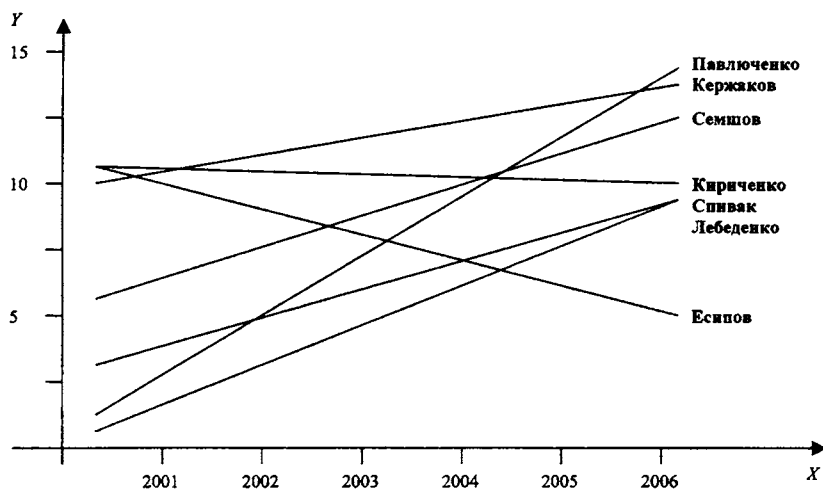


Рис.23.

По представленным графикам самая благоприятная перспектива наблюдается у Павлюченко, так как его прямая регрессии имеет самый большой условный коэффициент ( $k = 1,8$ ). Положительные тенденции наметились у Лебеденко, Спивака, Семшова и Кержакова. По графикам Кириченко и Есипова предсказывается снижение их результативности, угловые коэффициенты соответствующих прямых регрессии отрицательные.

Теперь построим прямые регрессии для команд российской премьер-лиги.

### 12.2. Российские чемпионаты по футболу

По результатам чемпионатов России за последние пять лет построим для каждой команды прямые регрессии их мест в зависимости от года соревнований.

Результаты чемпионатов России по футболу за 2001 - 2005 гг.

Таблица 78

№ п/ п	Команда	Место в сезоне				
		2001	2002	2003	2004	2005
1	"Спартак" (Москва)	1	3	10	8	3
2	"Локомотив" (Москва)	2	2	4	1	2
3	"Зенит" (С-Пб)	3	10	2	3	6
4	"Торпедо" (Москва)	4	4	8	5	7
5	"Крылья Советов" (Самара)	5	5	9	4	14
6	"Сатурн" (Московск. обл.)	6	7	7	7	11
7	ЦСКА (Москва)	7	1	1	2	1
8	"Динамо" (Москва)	9	8	6	13	9
9	"Шинник" (Ярославль)	-	6	5	6	8
10	"Алания" (Владикавказ)	12	12	13	14	15
11	"Ростов" (Ростов-на-Дону)	11	11	11	12	13

В таблице приведены показатели команд, которые участвовали во всех чемпионатах России, и ярославской команды «Шинник». Найдем прямые регрессии по результатам, показанным командами в последних пяти первенствах России.

Покажем поиск линий регрессии на примере ведущих футбольных клубов России. Примем за случайную величину  $X = \{\text{год чемпионата в XXI веке}\} = \{1, 2, 3, 4, 5\}$ , а за  $Y = \{\text{место команды в соответствующем чемпионате России}\}$ . Вычислим, во-первых, основные характеристики случайной величины  $X$ .

Выборочная средняя  $x^* = \frac{1}{5}(1+2+3+4+5) = 3$  и

выборочная дисперсия  $D_x = \frac{1}{5}[(1-3)^2 + (2-3)^2 + (4-3)^2 + (5-3)^2] = 2$ .

Во-вторых, выбираем одну из команд и вычисляем для ее показателей выборочную среднюю  $y^*$  и выборочную ковариацию  $k(X, Y)$ .

Тогда уравнение прямой регрессии находим из следующего соотношения:

$$y_x^* - y^* = \frac{k(X, Y)}{2} \cdot (x - 3).$$

Найдем уравнения прямых регрессии для ведущих российских футбольных клубов.



1. «Спартак»  $y^* = \frac{1}{5}(1+3+10+8+3) = 5.$

Выборочную ковариацию вычисляем по ковариационному графу.

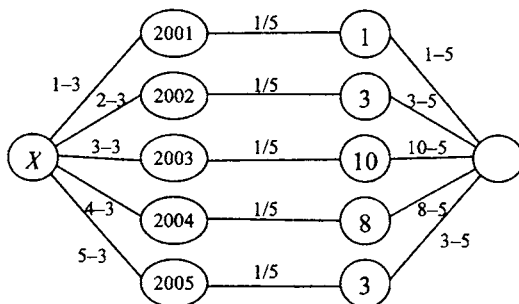


Рис. 24.

$$k(X, Y) = \frac{1}{5}[(-2) \cdot (-4) + (-1) \cdot (-2) + 0 + 1 \cdot 3 + 2 \cdot (-2)] = \frac{9}{5}.$$

Искомое уравнение регрессии имеет вид:

$$y_x^* - 5 = \frac{9}{5 \cdot 2} \cdot (x - 3) \quad \text{или} \quad y_x^* = 0,9x + 2,3.$$

Тогда в 2006 году, если сохранится тенденция последних пяти лет, можно ожидать от «Спартак» следующий результат:

при  $x=6$   $y_{x=6}^* = 0,9 \cdot 6 + 2,3 = 7,7 \approx 8(\text{место}).$

Такой результат вряд ли устроит как болельщиков «народной» команды, так и ее руководство.

2. «Локомотив»  $y^* = \frac{1}{5}(2+2+4+1+2) = 2,2$

$$k(X, Y) = \frac{1}{5}[(-2) \cdot (-0,2) + (-1) \cdot (-0,2) + 0 + 1 \cdot 1,8 + 2 \cdot (-0,2)] = 0,4$$

$$y_x^* - 2,2 = \frac{0,4}{2}(x - 3) \quad \text{или} \quad y_x^* = 0,2x + 1,6$$

при  $x=6$   $y_{x=6}^* = 2,8 = 7,7 \approx 3(\text{место}).$

3. «Зенит»  $y^* = \frac{1}{5}(3+10+2+3+6) = 4,8$

$$k(X, Y) = \frac{1}{5}[(-2) \cdot (-1,8) + (-1) \cdot 5,2 + 0 + 1 \cdot (-1,8) + 2 \cdot 1,2] = -0,2$$

$$y_x^* - 4,8 = \frac{-0,2}{2}(x - 3) \quad \text{или} \quad y_x^* = -0,1x + 5,1$$

при  $x=6$   $y_{x=6}^* = 4,5 \approx 5(\text{место}).$

Аналогично находим остальные уравнения прямых регрессии. Исключение из рассматриваемых команд составляет команда «Шинник», которая в 2001 году выступала в первой лиге. Поэтому здесь возможны два подхода. Первый – приписать «Шиннику» 17-е место (ранг) в 2001 году или второй – брать только четыре последних сезона. Рассмотрим второй подход. Тогда

$$x^* = \frac{1}{4}(2 + 3 + 4 + 5) = 3,5, \quad D_x^* = \frac{1}{4}(2,25 + 0,25) \cdot 2 = 1,25,$$

$$y^* = \frac{1}{4}(6 + 5 + 6 + 8) = 6,25,$$

$$k(X, Y) = \frac{1}{4}[(2 - 3,5) \cdot (6 - 6,25) + (3 - 3,5) \cdot (5 - 6,25) + (4 - 3,5) \cdot (6 - 6,25) +$$

$$+ (5 - 3,5) \cdot (8 - 6,25)] = \frac{2,75}{4}, \quad y_x^* - 6,25 = \frac{2,75}{4 \cdot 1,25}(x - 3,5) \quad \text{или} \quad y_x^* = 0,55x + 4,325$$

при  $x=6 \quad y_{x=6}^* = 7,625 \approx 8$  (место)

Полученные уравнения прямых регрессии и ожидаемые результаты команд в 2006 году занесем в следующую таблицу.

Таблица 78

№ п/п	Команды	Уравнение прямой регрессии	Ожидаемый результат на следующий сезон
1	"Спартак" (Москва)	$y_x^* = 0,9x + 2,3$	$y_{x=6}^* = 7,7 \approx 8$
2	"Локомотив" (Москва)	$y_x^* = 0,2x + 1,6$	$y_{x=6}^* = 2,8 \approx 3$
3	"Зенит" (С.-Пб)	$y_x^* = -0,1x + 5,1$	$y_{x=6}^* = 4,5 \approx 5$
4	"Торпедо" (Москва)	$y_x^* = 0,7x + 4,3$	$y_{x=6}^* = 8,5 \approx 9$
5	"Крылья Советов" (Самара)	$y_x^* = 1,7x + 2,3$	$y_{x=6}^* = 12,5 \approx 13$
6	"Сатурн" (Московск. обл.)	$y_x^* = x + 4,6$	$y_{x=6}^* = 10,6 \approx 11$
7	ЦСКА (Москва)	$y_x^* = -1,1x + 8,7$	$y_{x=6}^* = 2,1 \approx 2$
8	"Динамо" (Москва)	$y_x^* = 0,5x + 7,5$	$y_{x=6}^* = 10,5 \approx 11$
9	"Шинник" (Ярославль)	$y_x^* = 0,55x + 4,3$	$y_{x=6}^* = 7,6 \approx 8$
10	"Алания" (Владикавказ)	$y_x^* = 0,8x + 10,8$	$y_{x=6}^* = 15,6 \approx 16$
11	"Ростов" (Ростов-на-Дону)	$y_x^* = 0,4x + 10,4$	$y_{x=6}^* = 12,8 \approx 13$

Заметим, что сделанные прогнозы учитывают сложившиеся тенденции развития российского футбола и не могут учитывать амбиции и профессионализм президентов клубов, финансовые возможности старых и, быть может, новых их владельцев.

### 12.3. Прогнозы на предстоящий чемпионат мира в Германии

Чемпионаты мира по футболу в отличие от хоккейных проводятся один раз в четыре года, и им предшествуют отборочные раунды по континентальным зонам, которые надо успешно пройти. Сборная России в очередной раз оставила своих болельщиков без футбольного праздника, что свидетельствует о низком уровне российского футбола, о падении его престижа, к сожалению. Тем не менее попытаемся сделать прогноз на предстоящий чемпионат мира в Германии, который будет основан на результатах, показанных ведущими футбольными клубами на последних пяти мировых форумах. Понятно, что за эти 20 лет сменилось несколько поколений футболистов, да и большинство наставников команд, но сохранились традиции, постановка футбольных хозяйств и спортивной инфраструктуры той или иной страны. Из этого мы и исходим, вычисляя уравнения прямых регрессии достижений сборных в проводимом первенстве.

Заполним таблицу результатов ведущих футбольных сборных, при этом команде, вышедшей в четвертьфинале, присвоили пятое место (на самом деле 5 - 8 место), считая четырех неудачников одной четвертой одинаково мастеровитыми. Аналогично даем 9-е место неудачникам одной восьмой финала и 17-е – каждой команде, не вышедшей из группы.

Таблица 79

№	Команды	ЧМ-1986	ЧМ-1990	ЧМ-1994	ЧМ-1998	ЧМ-2002	Среднее место $x^*$
1	Аргентина	1	2	9	5	17	6,8
2	Бразилия	5	9	1	2	1	3,6
3	Италия	9	3	2	5	9	5,6
4	Германия	2	1	5	5	2	3,0
5	Англия	5	4	17	9	5	8,0
6	Испания	5	9	5	17	5	8,2

Для нахождения уравнений прямых регрессии пронумеруем чемпионаты мира по порядку их следования натуральными числами 1, 2, 3, 4, 5 ( $= x$ ). Учитывая, что  $x^* = 3$ , а  $D_x^* = 2$ , получаем искомые уравнения

$$y_x^* - y^* = \frac{k(X, Y)}{2}(x - 3),$$

которые записываем в таблицу.

Таблица 80

№	Команда	Уравнение прямой регрессии	Ожидаемый результат на следующий ЧМ
1	<b>Аргентина</b>	$y_x^* = 3,5x - 3,7$	$y_{x=6}^* = 17,3 \approx 17$ (место)
2	<b>Бразилия</b>	$y_x^* = -1,5x + 8,1$	$y_{x=6}^* \approx 1$
3	<b>Италия</b>	$y_x^* = 0,2x + 5$	$y_{x=6}^* = 6,2 \approx 6$
4	<b>Германия</b>	$y_x^* = 0,4x = 1,8$	$y_{x=6}^* = 4,2 \approx 4$
5	<b>Англия</b>	$y_x^* = 0,5x + 6,5$	$y_{x=6}^* = 9,5 \approx 10$
6	<b>Испания</b>	$y_x^* = 1,6 + 3,4$	$y_{x=6}^* = 13$

Полученные результаты носят вероятностно-статистический характер, и их разумно использовать тренерам и руководителям федераций при оценке выполнения поставленных задач.

Предложенный в главе подход может быть полезен спортивным функционерам, тренерским кадрам, учителям физической культуры, любителям спорта при анализе спортивных показателей и достижений, при планировании и проведении спортивно-массовой работы со школьниками и молодежью. Сравнение показанных результатов с прогнозируемыми позволяет и самому спортсмену оценивать свою спортивную подготовку, да и психологическую устойчивость, вносить корректировку в организацию тренировочного процесса.

## Список использованной литературы

- [1] Аулик И. В. Определение физической работоспособности в клинике и спорте. – М.: Медицина, 1990. – 192 с.
- [2] Афанасьев В. В. Теория вероятностей в вопросах и задачах. // Учебное пособие. – Ярославль: Изд-во ЯГПУ, 2004. – 250 с.
- [3] Афанасьев В. В. Дидактический модуль курса стохастики (математическая статистика). // Учебное пособие. – Ярославль: Изд-во ЯГПУ, 2003. – 56 с.
- [4] Афанасьев В. В. Вероятностные игры. М.: «Математика». 2005, № 14, С. 35 - 38.
- [5] Афанасьев В. В. Теория вероятностей. // Учебное пособие для студентов высших учебных заведений. – М.: Изд-во ВЛАДОС-ПРЕСС, 2006. – 352 с.
- [6] Афанасьев В. В. Экспериментальное исследование творческой активности студентов в процессе обучения математике. / В. В. Афанасьев, Е. И. Смирнов. // Ярославский педагогический вестник, 1996, №3(6), С.110 -115.
- [7] Афанасьев В. В., Викулов А. Д. Физические упражнения и здоровье. // Тезисы доклада международной научной конференции «Технологии- 2005», Анталия (Турция). Изд-во «Академия естествознания», 2003.
- [8] Афанасьев В. В. Непряев И. Н. Математическая статистика в спорте. // Ярославский педагогический вестник, 2005, № 2(43), С.108 - 113.
- [9] Афанасьев В. В., Непряев И. Н. Бюджеты клубов и их спортивные достижения. // Ярославский педагогический вестник, 2005, № 3(44), С.10 - 17.
- [10] Афанасьев В. В., Непряев И. Н. Линии регрессии и прогнозы в спорте. // Ярославский педагогический вестник, 2006, № 1(46), С.81-90
- [11] Афанасьев В. В., Непряев И. Н. Совмещенные ковариационные графы. // Материалы конференции «Чтения Ушинского». – Ярославль: Изд-во ЯГПУ, 2006.
- [12] Барышева М. И., Баканова И. А., Муравьев А. В. Изменение скорости и объема движений кисти ведущей руки у младших школьников с нарушением письма. // Физкультура. Спорт. Здоровье. Материалы конференции «Чтения Ушинского». – Ярославль: Изд-во ЯГПУ, 2005, С. 158 -162.
- [13] Берман Р. Е., Воган В. К. Руководство по педиатрии. – М.: Медицина, 1987. – 640 с.

[14] Благущ П. К теории тестирования двигательных способностей. – М.: Физкультура и спорт, 1972. – 188 с.

[15] Боровков А. А. Математическая статистика. Оценка параметров. Проверка гипотез. / А. А. Боровков. – М.: Наука, 1984. – 442 с.

[16] Булгакова Н. Ж. Отбор и подготовка юных пловцов. – М.: Физкультура и спорт, 1978. – 212 с.

[17] Бурухин С. Ф., Горбачев М. С. Динамика развития физической подготовленности школьников в возрасте 10 - 13 лет на уроках физической культуры. // Физкультура. Спорт. Здоровье. Материалы конференции «Чтения Ушинского». – Ярославль: Изд-во ЯГПУ, 2004, С.14 - 17.

[18] Викулов А. Д., Сулейманов В. Н., Турчанинов С. Ю. Физиологические механизмы адаптации организма младших школьников в условиях режима повышенной двигательной активности. // Ярославский педагогический вестник, 2005, № 3(44), С. 84 - 87.

[19] Воскресенский А. Д., Вентцель И. Д. Статистический анализ сердечного ритма и показателей гемодинамики в физиологических исследованиях. // Проблемы космической биологии. – М.: Наука, 1976, т.26.

[20] Гласс Дж. Статистические методы в педагогике и психологии. / Дж. Гласс, Дж. Стенли. – М.: Педагогика, 1976. – 495 с.

[21] Гмурман В. Е. Теория вероятностей и математическая статистика. / В. Е. Гмурман. 7-е изд. – М.: Высшая школа, 2000. – 479 с.

[22] Годик М. А. Контроль тренировочных и соревновательных нагрузок. – М.: Физкультура и спорт, 1979. – 136 с.

[23] Годик М. А. Спортивная метрология. – М.: Физкультура и спорт, 1988. – 192 с.

[24] Гушин А. Г., Усилов А. А., Михайлов П. В. Адаптационные изменения у юных спортсменов с разным уровнем физической работоспособности. // Актуальные проблемы адаптации организма в норме и патологии: Материалы международной научной конференции. – Изд-во "Ремдер", 2005, с. 143 - 144.

[25] Донской Д. Д., Зацюрский В. М. Биомеханика. – М.: Физкультура и спорт, 1979. – 264 с.

[26] Зацюрский В. М. Кибернетика. Математика. Спорт. – М.: Физкультура и спорт, 1969. – 198 с.

[27] Зацюрский В. М. Основы спортивной метрологии. – М.: Физкультура и спорт, 1979. – 196 с.

[28] Иванов В. В. Комплексный контроль в подготовке спортсменов. – М.: Физкультура и спорт, 1987. – 216 с.

[29] Карпман В. Л., Белоцерковский З. Б., Гудков И. А. Тестирование в спортивной медицине. – М.: Физкультура и спорт, 1988. – 135 с.

- [30] Кендэл М. Ранговая корреляция. – М.: Статистика, 1975.–214с.
- [31] Кузнецов Д. Ю. Статистические методы анализа результатов психолого-педагогических исследований. // Ярославский педагогический вестник, 2000, № 2 (24), С. 129 - 135.
- [32] Масальгин Н. А. Математико-статистические методы в спорте. – М.: Физкультура и спорт, 1974. – 151 с.
- [33] Масленникова Ю. Л. Исследование состояния сосудистого русла у лиц со сниженным уровнем двигательной активности. // Физкультура. Спорт. Здоровье. Материалы конференции «Чтения Ушинского». – Ярославль: Изд-во ЯГПУ, 2005, С. 99 - 107.
- [34] Михайлов П. В., Муравьев А. В., Муравьев А. А. Сравнительный анализ физического развития и аэробной работоспособности юных хоккеистов и футболистов 14 - 15-летнего возраста. // Физкультура. Спорт. Здоровье. Материалы конференции «Чтения Ушинского». – Ярославль: Изд-во ЯГПУ, 2005, С. 107 - 111.
- [35] Муравьев А. А., Усилов А. А., Михайлов П. В., Муравьев А. В. и др. Показатели физического развития и функционального состояния организма юных футболистов 11 - 17 лет. // Физкультура. Спорт. Здоровье. Материалы конференции «Чтения Ушинского». – Ярославль: Изд-во ЯГПУ, 2004, С. 37 - 45.
- [36] Муравьев А. В., Зайцев Л. Г., Муравьев А. А., Бурухин С. Ф., Чопоров С. В., Баканова И. А. Оптимальный гематокрит в норме и патологии. // Материалы международной конференции «Гемореология в микро- и макроциркуляции». – Ярославль: Изд-во ЯГПУ, 2005, с.17.
- [37] Муравьев А. В., Маймистова А. А., Буланова С. В., Муравьев А. А., Чопоров С. В. Сравнительная гемореологическая эффективность пентоксифиллина (трентала) разных фирм. // Материалы международной конференции «Гемореология в микро- и макроциркуляции». – Ярославль: Изд-во ЯГПУ, 2005, С.23.
- [38] Начинская С. В. Математическая статистика в спорте. – М.: Здоровье, 1978. – 144 с.
- [39] Начинская С. В. Основы спортивной статистики. – М.: Высшая школа, 1978. – 192 с.
- [40] Непряев И. Н. Ранговая корреляция в хоккее. // Вестник студенческого научного общества. – Ярославль: Изд-во ЯГПУ, 2005, № 3 (3), С. 151 - 157.
- [41] Непряев И. Н. Конкордация в хоккее. // Физкультура. Спорт. Здоровье. Материалы конференции «Чтения Ушинского». – Ярославль: Изд-во ЯГПУ, 2006.

[42] Огородникова Л. А. Развитие тактической одаренности в игровых видах спорта. // Социальная психология XXI столетия. т. 2. – Изд-во МАПН, 2005.

[43] Огородникова Л. А. Некоторые социально-психологические проблемы современного спорта. // Человеческий фактор: Социальный психолог. Вып. 1(9), 2005.

[44] Огородникова Л. А. Межпредметные проблемы в преподавании спортивной психологии. // Высшая школа на современном этапе: психология преподавания и обучения. т. 2. – М.: Изд-во "Российское психологическое общество", 2005.

[45] Плохинский Н. А. Биометрия. – М.: Изд-во МГУ, 1970. – 368с.

[46] Садовский Л. Е., Садовский А. Л. Математика и спорт. – М.: Наука, 1985. – 191 с.

[47] Сонькин В. Д., Корниенко И. М., Тамбовцева Р. В. Основные закономерности и типологические особенности роста и физического развития. // Физиология развития ребенка: теоретические и прикладные аспекты. - М., 2000. - 31 с.

[48] Урбах В. Ю. Статистический анализ в биологических исследованиях. – М.: Медицина, 1975. – 295 с.

[49] Уткин В. Л. Измерения в спорте (введение в спортивную метрологию). – М.: ГЦОЛИФК, 1978. – 138 с.

[50] Astrand P. O. Do we need physical conditioning? // J. Physical Educat. – 1972. Vol.2. P.129 - 136.

[51] Alder H. Z., Roessler E. B. Introduction to Probability and Statistics. – San Francisco, 1975. – 426 p.

[52] Brown F. L., Amos I. R., Mink O. G. Statistical Concepts: A Basic Program, Second Edition. – New York, Evanston, San Francisco, London, 1975. – 140 p.

[53] Hamming R. H. The Art of Probability for scientists and engineers. – Redwood City, California etc., 1991. – 340 p.

[54] Marcus M. A survey of finite mathematics. – Univ. of California, Boston, New York, Geneva, 1969. – 486 p.

[55] Matherow G. Estimating and Choosing: An Essay on Probability in Practice. Springer. – Verlag, Berlin, Heidelberg, 1989. – 141 p.

[56] Romano T. R., Siegel A. F. Counterexamples in probability and statistics. 1986. – 303 p.

[57] Swinscow T. D.V. Statistics at Square One. – British Medical Journal, London, 1978. – 85 p.



**В. В. Афанасьев, И. Н. Непряев**

**Математическая статистика  
в командных видах спорта**

Редактор Л.К.Шереметьева

Подписано в печать 27.04.2006.  
Формат 60x84. Объем 7,5 п.л. Тираж 500. Заказ 856

Издательство Ярославского государственного педагогического  
университета имени К.Д. Ушинского (ЯГПУ)  
150000, Ярославль, Республиканская ул. 108

Типография ЯГПУ  
150000, г. Ярославль, Которосльская наб., 44  
Тел.: (4852) 32-98-69, 72-64-08



**Афанасьев Владимир Васильевич** - доктор педагогических и кандидат физико-математических наук, профессор, академик Российской академии естественных наук, автор более 150 статей и 10 книг по теории и методике обучения математике.

**Непряев Иван Николаевич** - мастер спорта по хоккею. Чемпион мира среди молодежи, бронзовый призер чемпионата мира, двукратный чемпион России, аспирант Ярославского государственного педагогического университета им. К.Д. Ушинского, автор 6 статей.

---