12. Suler J. The Basic Psychological Features of Cyberspace // Suler J. The Psychology of Cybespace. http://www.rider.edu/~suler/psycyber/basicfeat.html.

В.В. Афанасьев, И.Н. Непряев

Линии регрессии и прогнозы в спорте

Предложенная статья является продолжением предыдущих работ авторов [1], [2], посвященных применению методов математической статистики 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 чем более объективной она является, 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 важны 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, 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 команд.

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 квадратичная этих случайных величин, а D_x^* – выборочная дисперсия величины X.

§1. 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 успешности выступлений 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.

1.1. Прогноз результативности хоккеистов

Проанализируем результаты, показанные ведущими российскими хоккеистами в последних чемпионатах Суперлиги.

Результативность некоторых нападающих по итогам четырех последних чемпионатов России

Таблица 1

№	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 М. Сушинского, как находится уравнение прямой регрессии очков по годам (1(2001-2002),2,3,4) их набирания.

Средний показатель набранных очков y^* у него 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 [3. С. 178] ковариационный граф:

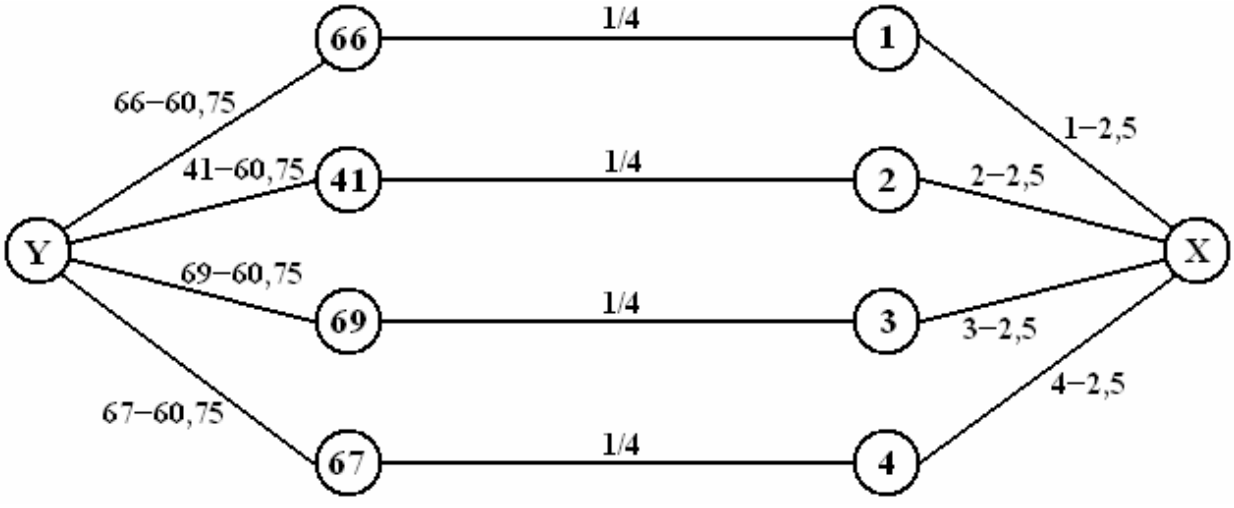

Рис. 1. Ковариационный граф

Выборочную ковариацию находим как вес всего ковариационного графа:

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

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 $y^*_x = 3{,}1 \cdot 5 + 53 \approx 69$ очков

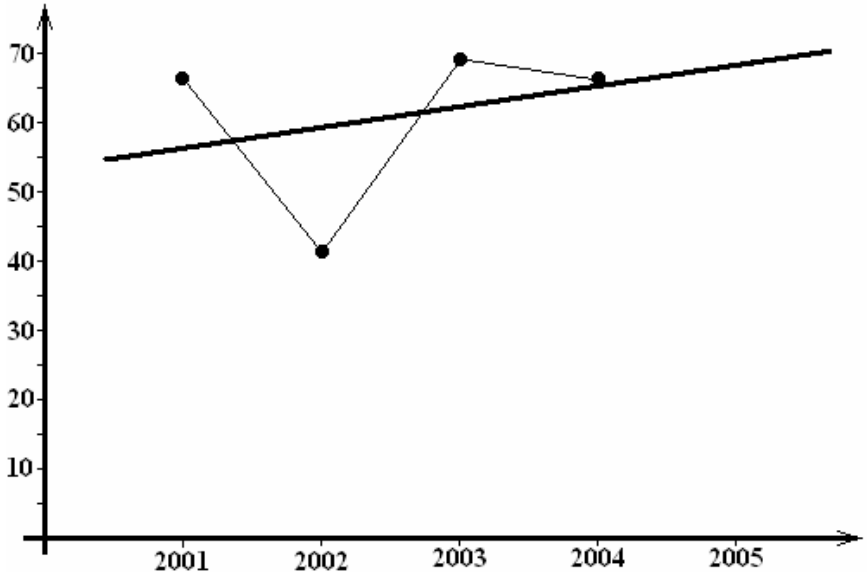

Рис. 2. Результативность М. Сушинского и прямая регрессии результативности по четырем последним первенствам России

Аналогично находим и уравнения прямых регрессий для других перечисленных хоккеистов и ожидаемые количества очков в сезоне 2005-2006 г., которые занесем в следующую таблицу.

Уравнения прямых регрессии и ожидаемые результаты некоторых хоккеистов

Таблица 2

Игрок	**Уравнение прямой регрессии**	**Ожидаемый результат на следующий сезон**
Сушинский М.	$y^*_x = 3{,}1x + 53$	$y^*_{x=5} = 69$ (очков)
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.	$y^*_x = -2x + 30$	$y^*_{x=5} = 20$ (очка)

Для лучшей наглядности построим на координатной плоскости (x,y) полученные прямые регрессии.

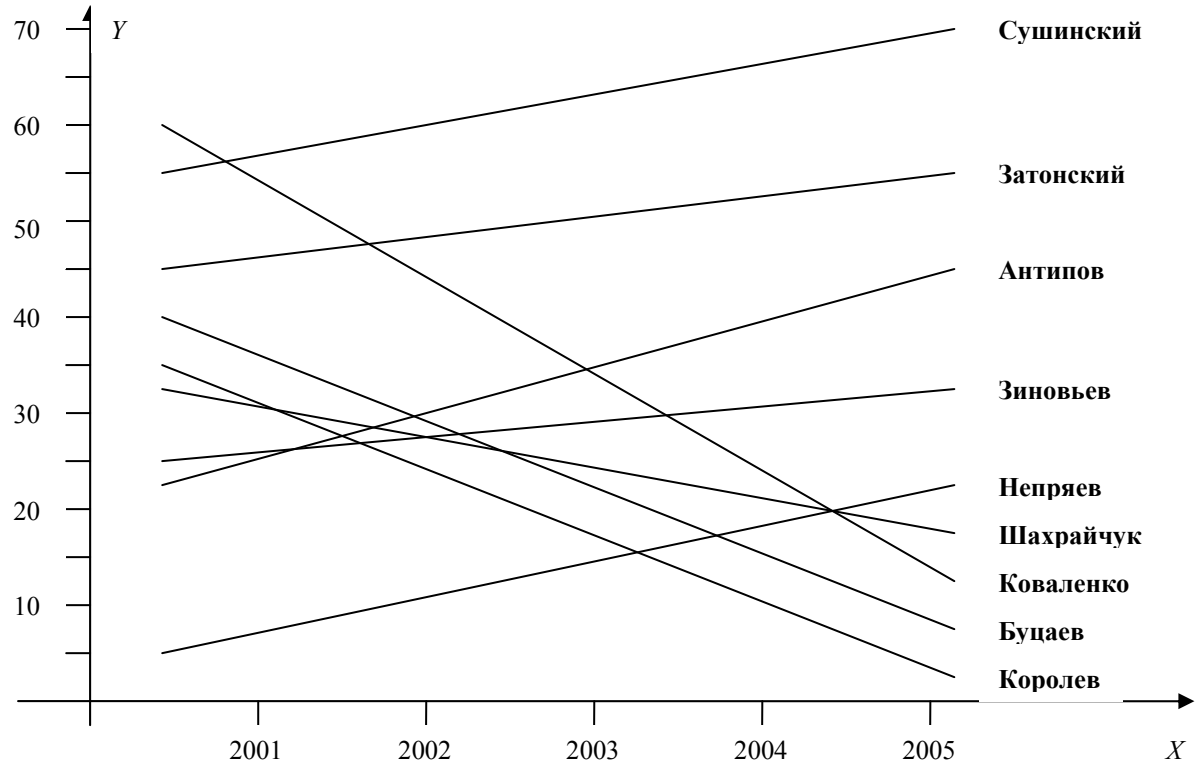

Рис. 3. Прямые регрессии набранных очков нападающих по четырем последним первенствам России

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

Аналогично построим прямые регрессии очков российских защитников по итогам последних четырех чемпионатов России.

Результативность некоторых защитников по итогам четырех последних чемпионатов России

Таблица 3

№	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 4

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)

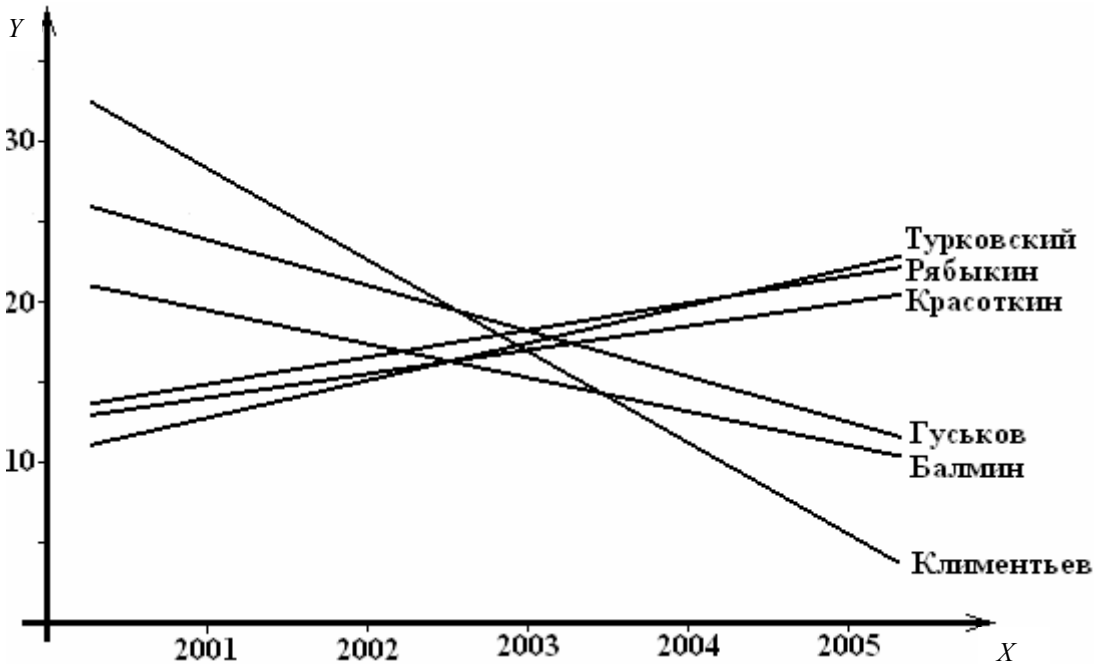

Рис. 4. Прямые регрессии набранных очков защитников по четырем последним первенствам России

По представленным графикам видны тенденции результативности защитников. Самая благоприятная перспектива у В. Турковского, поскольку прямая регрессии имеет самый большой угловой коэффициент (k = 2,7). Почти одинаковые перспективы у двух Дмитриев – Рябыкина и Красоткина, у которых прямые регрессии имеют соответственно коэффициенты 1,7 и 1,5. Наихудшие тенденции наметились у С. Клементьева (k = – 7,1) и у А. Гуськова (k= – 2,3), в этом случае уже игрока омского «Авангарда».

1.2. Российские чемпионаты по хоккею и их прогнозы

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

Результаты регулярных чемпионатов России за последние четыре сезона

Таблица 5

№	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

Наметившиеся тенденции команд достаточно хорошо просматриваются по полученным прямым регрессии.

Уравнения прямых регрессии и ожидаемые результаты команд

Таблица 6

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)

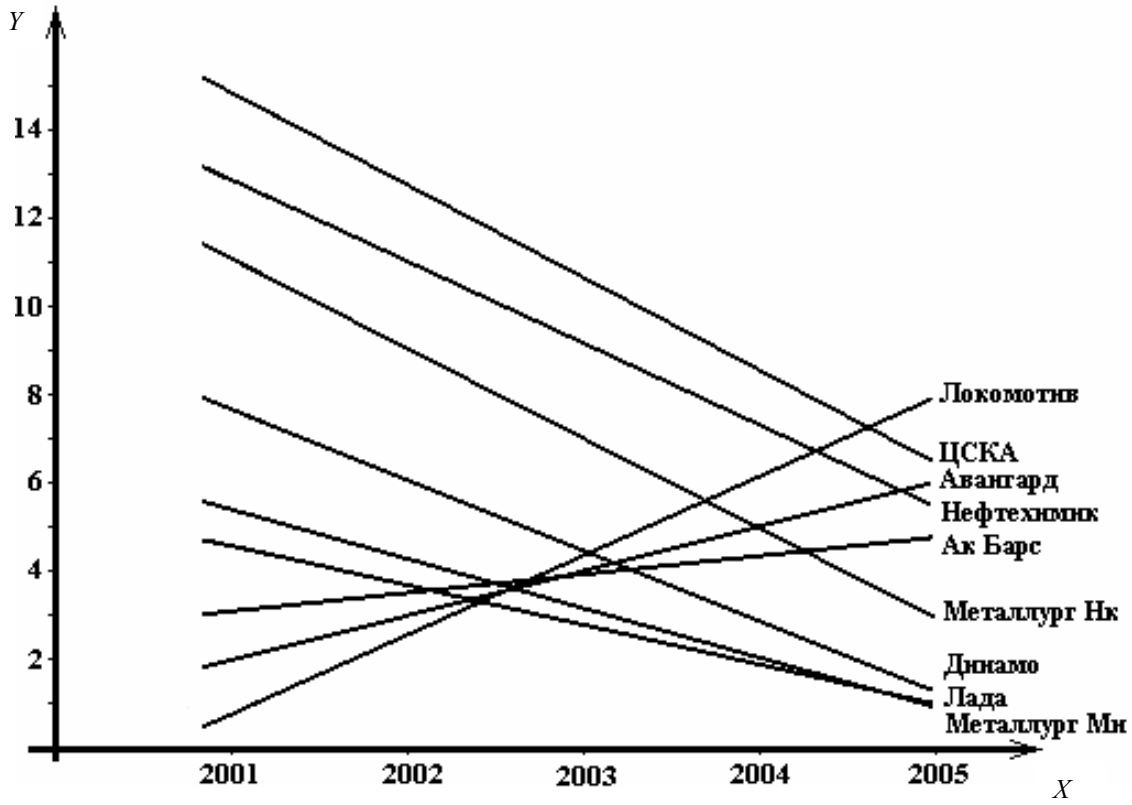

Рис. 5

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 возможных изменений в содержании клуба, его финансирования, новых приобретений и возможных форс-мажорных обстоятельств. Так, известные и печальные изменения отношения тольяттинского Автоваза к своей хоккейной команде «Лада», конечно же, изменят ожидания и болельщиков команды, и специалистов. А знали бы владельцы клуба возможности клуба, в том числе и вероятностно-статистические?!

§2. Прогнозы и надежды российского футбола

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

Представляет интерес проследить перспективы игроков Российской футбольной Премьер-Лиги. Построим прямые регрессии для известных бомбардиров РФПЛ, используя статистику забитых ими голов за последние годы.

2.1. Результативность российских футболистов

Бомбардиры РФПЛ за последние пять сезонов

Таблица 7

№	Фамилии игроков, команда	Количество забитых голов				
		2001	2002	2003	2004	2005
1	Кержаков А. («Зенит»)	6	14	13	18	7

2	Есипов В. («Сатурн»)	11	7	11	10	4
3	Павлюченко Р. («Спартак»)	5	4	10	10	11
4	Семшов И. («Торпедо»)	6	11	6	9	12
5	Кириченко Д. («Москва»)	13	15	5	9	14
6	Спивак А. («Зенит»)	3	5	6	10	6
7	Лебеденко И. («Локомотив»)	0	6	3	9	6

Уравнения прямых регрессии и ожидаемое количество голов в следующем сезоне

Таблица 8

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

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

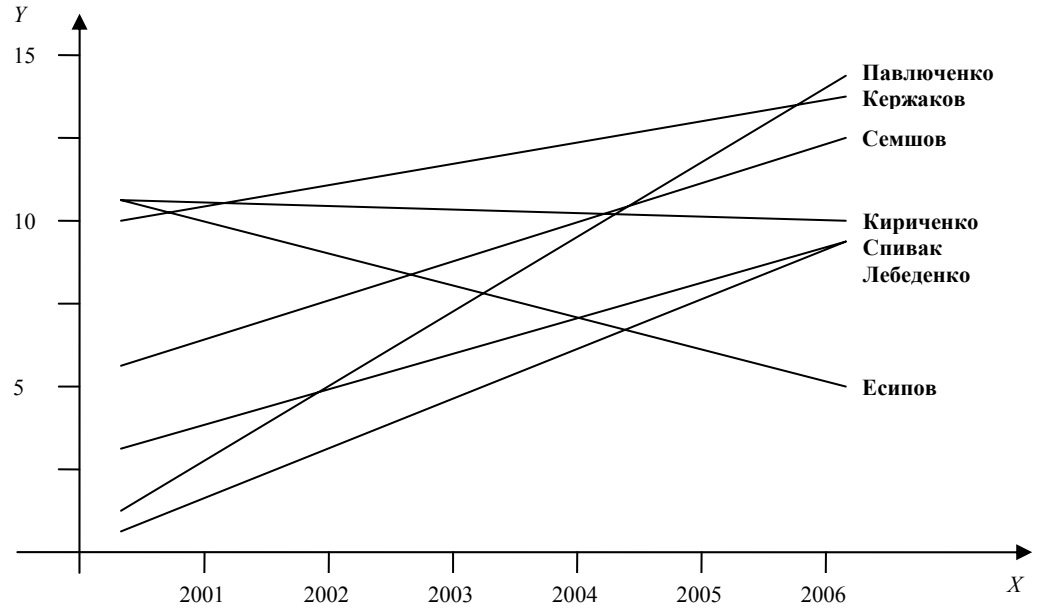

Рис. 6

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 и у Лебеденко, Спивака, Семшова и Кержакова. По графикам достижений Кириченко и Есипова предсказывается снижение их результативности, угловые коэффициенты соответствующих прямых регрессии отрицательные.

Теперь построим прямые регрессии для команд российской Премьер-лиги.

2.2. 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 9

№ 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.

Примем за случайную величину X = {год чемпионата в XXI веке}={1, 2, 3, 4, 5}, а за Y = {место команды в соответствующем чемпионате России}. Вычислим, во–первых, основные характеристики случайной величины X.

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

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

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 уравнение прямой регрессии для московского «Спартака».

Найдем среднее место команды

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

Выборочную ковариацию вычисляем по ковариационному графу.

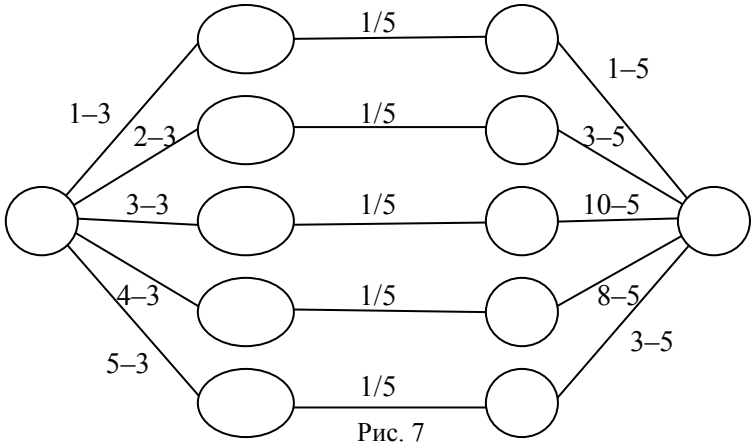

Рис. 7

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

Искомое уравнение регрессии имеет вид $y_x^* - 5 = \frac{9}{5\cdot 2}\cdot(x-3)$ или $y_x^* = 0{,}9x + 2{,}3$.

Тогда в 2006 году, если сохранится тенденция последних пяти лет, можно ожидать от «Спартака» следующего результата:

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

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

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 место (ранг) в 2001 году или второй – рассматривать только четыре последних сезона. Рассмотрим второй подход. Тогда

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

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

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

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

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

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

Таблица 10

№ п/п	Команда	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.

Предложенный в работе 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.

Библиографический список

1. Афанасьев В.В., Непряев И.Н. Методы математической статистики в спорте//Ярославский педагогический вестник. 2005. № 2 (43). С. 108-113.
2. Афанасьев В.В., Непряев И.Н.Бюджеты клубов и их спортивные достижения // Ярославский педагогический вестник. 2005. № 3 (44). С. 10-17.
3. Афанасьев В.В. Теория вероятностей в вопросах и задачах: Учебное пособие. Ярославль: Изд-во ЯГПУ, 2004. 250 с.