УДК 796.15.86

РАННИЙ ПРОГНОЗ РЕЗУЛЬТАТИВНОСТИ ПРЫГУНОВ В ВЫСОТУ

Р.Ф. Ахметов

Житомирский государственный университет имени Ивана Франко

Актуальность. В последние годы украинским прыгунам в высоту не удается побеждать на крупных международных соревнованиях. Данное обстоятельство стимулирует специалистов не только повышать эффективность тренировочного процесса, но и продолжать разработку точности прогноза результативности прыгунов в высоту, что в значительной мере будет способствовать качественному отбору в этом виде спорта. В связи с этим весьма актуальным является разработка программы прогноза результативности на базе некоторой совокупности параметров спортсменов.

Анализ последних исследований и публикаций. Прогнозирование основывается на использовании метода экстраполяции, Прогнозирование основывается на использовании метода экстраполяции, предполагающего распространение выводов, полученных из наблюдения над одной частью какого-либо явления, на другие его части [3; 6; 7]. В условиях спорта экстраполяция позволяет осуществить прогнозы роста результативности на основе изучения соответствующих закономерностей в предшествующие годы. Задачу прогноза результативности спортсменов можно решить на базе факторного анализа и динамики развития физических параметров и результатов на некотором ограниченном интервале времени (например, 10-13 лет) [1; 2; 7]. Для этого проводится линейная интерполяция результатов и физических параметров спортсменов между годовыми аттестационными периодами на более меньшие временные периоды – полугодовые и квартальные. Тогда в задачах синтеза линейной многомерной регрессии результативности представляется возможным использовать большее число информативных параметров.

Цель исследований – осуществить ранний прогноз результативности прыгунов в высоту по результатам анализа возрастных групп до 13 лет. Для этого спортивные результаты и значения усредненных физических параметров спортсменов линейно интерполировались на полугодовые или квартальные периоды.

Результаты исследований. В данной работе дается продолжение общего подхода [2] к частной задаче прогноза результативности прыгунов в высоту. Поскольку результаты и физические параметры спортсменов в группе имеют случайный разброс (дисперсию) [1], то, говоря о задаче прогноза результативности, имеет смысл рассматривать прогноз средней результативности $\overline{\mathrm{H}}(t)$, как функции средних по группе физических параметров $\vec{X}_P$, которые будем представлять в виде матрицы столбца:

$$\vec{\mathrm{X}}_P = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_P \end{pmatrix}, \ \mathrm{P}=1,2,\dots,\mathrm{N_p}-1;\ \mathrm{N_p} \geq 3,$$

где $\mathrm{N_p}$ – полное число спортивных параметров, включая сам результат (Н). Полное множество Р-мерных группировок из ($\mathrm{N_p}$-1) по Р равно числу сочетаний из ($\mathrm{N_p}$-1) по Р:

$$\vec{\mathrm{X}}_P \in U_{\vec{X}_P} = \{\vec{X}_P^{\alpha}, \alpha = 1,2,\dots,C_{N_p-1}^P\}, \quad (1)$$

$$C_{N_p-1}^P = \frac{(N_p-1)!}{P!(N_p-1-P)!}.$$

Так для прыгунов в высоту выделяется следующая полная совокупность спортивных параметров [2]:

РАСШИРЕННЫЙ ПЕРЕЧЕНЬ ИЗ 21 ПАРАМЕТРА ПРЫГУНОВ В ВЫСОТУ

1. Спортивный результат (высота) – Целевая функция.

АНТРОПОМЕТРИЧЕСКИЕ ПАРАМЕТРЫ (2-7)

2. Рост.
3. Длина голени.
4. Длина бедра.

5. Окружность бедра.
6. Окружность икроножной мышцы.
7. Вес.

ТЕХНИЧЕСКИЕ ПАРАМЕТРЫ (8-14)

(Регистрируемые и расчетные показатели технической подготовки)

8. Скорость разбега перед отталкиванием.
9. Скорость вылета ОЦТ (в момент отрыва).
10. Угол вылета ОЦТ.
11. Длительность фазы отталкивания.
12. Высота вылета ОЦТ.
13. Импульс силы отталкивания.
14. Степень использования силовых возможностей толчка (%).

СПЕЦИАЛИЗИРОВАННЫЕ ПАРАМЕТРЫ (15-21)

(Уровень спецфизподготовки)

15. Бег – 30 м (с).
16. Скорость спринтерского бега (10 м с хода).
17. Прыжок вверх в высоту с двух ног с места.
18. Прыжок в длину с места.
19. Тройной прыжок с места
20. Прыжок вверх с толчковой ноги (махом другой).
21. Прыжок вверх в высоту с трех шагов.

Информативность различных Р-мерных группировок $\vec{X}_P$ в задачах прогноза результативности будет также различной. Вопрос о выборе оптимальной совокупности наиболее информативных параметров из множества (1) при различных Р требует самостоятельных глубоких исследований в рамках отдельной НИР. В работе [2] предложен один из альтернативных вариантов решения задачи, который вполне приемлем с точки зрения точности прогноза. В первом приближении рассматривается задача линейного прогноза в рамках классической теории линейной регрессии (интерполяции) в математической статистике [5; 8]. Речь идет о нахождении аппроксимации

$$\overline{\text{H}} \cong \text{H}_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_P X_P \text{, (2)}$$

где $\text{H}_0, \alpha_1, \ldots, \alpha_P$ – неизвестные параметры регрессии, которые требуется оценить по данным некоторого количества возрастных групп. В более точной постановке приближенная линейная регрессия (2) представляется в виде:

$$\overline{\text{H}}(t) = \text{H}_0 + \alpha_1 X_1(t) + \alpha_2 X_2(t) + \ldots + \alpha_P X_P(t) + \xi(t),\ t \in T = (a,b) \text{, (3)}$$

где $\xi(t)$ – ошибка прогноза с нулевым средним ($M\xi(t) = 0$) и неизвестной дисперсией $\sigma_\xi^2 = M\xi^2$ (M – оператор математического ожидания – среднего). Если в результате решения задачи линейной регрессии на интервале времени T получены оценки неизвестных параметров регрессии:

$$\text{H}_0 = \text{H}_0^\wedge(T);\ \alpha_n = \alpha_n^\wedge(T), \quad n = 1,2,\ldots, P,$$

то прогнозное значение средней результативности вне этого интервала представляется в виде:

$$\overline{\text{H}}^\wedge(t_0) = \text{H}_0^\wedge(T) + \sum_{n=1}^{P} \alpha_n^\wedge(T) X_n(t_0), \quad t_0 > b\,, \qquad (4)$$

где набор физических параметров $\{X_n(t_0),\ \ n = 1,2,\ldots, P\}$ – задается на прогнозируемый момент времени t_0. При этом среднеквадратическая ошибка (СКО) прогноза оценивается величиной $\sigma_\xi(T)$. Насколько «удачно» получена оценка (4), – зависит от многих факторов и последнее слово здесь за практикой (экспериментальной апробации). Проведенная в работе [1] апробация модели (4) показывает, что она практически вполне приемлема. СКО при этом не превышает 3-х сантиметров, а прогнозируемый рекордный результат составляет 250 см. Зависимость (4) прогнозного значения результативности от времени (возрастной группы) называется далее прогнозной динамической характеристикой результативности (ПДХР). Как показано в работах[1-2], для расчета ПДХР требуется выполнить необходимое условие: $N \geq P - 2$, где N – объем временной выборки (число анализируемых возрастных групп). При этом точность прогноза возрастает с увеличением числа P используемых информативных спортивных параметров. Следовательно, для получения удовлетворительной точности прогноза результативности требуется и достаточно большой объем временной выборки N возрастных результатов и усредненных (по группе спортсменов) физических спортивных параметров. К настоящему времени наиболее распространена годовая регистрация результатов и физических параметров спортсменов (обычно после соревнований) в возрасте от 10 до 17 лет. Тогда объем временной выборки ограничивается величиной N_1=8 или N_2=9 (если регистрируются еще и результаты мастеров спорта международного класса). В связи с этим возможности раннего прогноза результативности, например, по результатам анализа в возрастных группах 10-12 (13) лет оказываются довольно ограниченными. Как показано в работах [1-2], достаточно удовлетворительный прогноз результативности прыгунов в высоту по трем важным информативным параметрам (X_{12}, X_9, X_{21}) возможен только при N=5 (возрасты от 10 до 14 лет) на период до 17 лет.

Нами разработана специальная модифицированная программа cor2din.pas в среде Turbo Pascal, которая позволяет расчетным способом линейной интерполяции увеличить объем временной выборки до N_d=17 (17=9*2-1). При этом в случае полугодовой интерполяции для 3-х мерной совокупности физических параметров (X_{12}, X_9, X_{21}) величина N=5 соответствует «пороговому» возрасту 12 лет. Однако, как показали расчеты, удовлетворительную точность прогноза удается получить не при N=5, а начиная с N=6, что соответствует пороговому возрасту 12,5 лет.

Матричное решение задачи линейной регрессии результативности по заданной совокупности наиболее информативных параметров.

Для оценки параметров регрессии $\mathrm{H}_0, \alpha_1, \ldots, \alpha_P$ составляется следующая система линейных алгебраических уравнений:

$$\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_1) = \overline{\mathrm{H}}(t_1)$$

$$\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_2) = \overline{\mathrm{H}}(t_2) \qquad (5)$$

………………………………….

$$\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_N) = \overline{\mathrm{H}}(t_N) ,$$

где N – число возрастных групп (в данной работе N<18). Система (5) представляется в матричном виде:

$$\mathrm{H}_0 \vec{1}_N + \sum_{m=1}^{P} \alpha_m \vec{X}_N^m = \vec{\overline{\mathrm{H}}}_N \Rightarrow \quad (6)$$

$$\vec{1}_N = \begin{pmatrix} 1 \\ 1 \\ .. \\ 1 \end{pmatrix}_N , \quad \vec{X}_N{}^m = \begin{pmatrix} X_m(t_1) \\ X_m(t_2) \\ \dots\dots\dots \\ X_m(t_N) \end{pmatrix} , \quad \vec{\overline{\mathrm{H}}}_N = \begin{pmatrix} \overline{\mathrm{H}}(t_1) \\ \overline{\mathrm{H}}(t_2) \\ \dots\dots\dots \\ \overline{\mathrm{H}}(t_N) \end{pmatrix} .$$

Вводя т.н. «сигнальный» регрессионный вектор (СРВ):

$$\vec{s}_M = \begin{pmatrix} \mathrm{H}_0 \\ \alpha_1 \\ \dots \\ \alpha_P \end{pmatrix}_M = \begin{pmatrix} s_1 \\ s_2 \\ \dots \\ s_M \end{pmatrix} , \quad M = P + 1 ,$$

$$s_1 = \mathrm{H}_0 , s_2 = \alpha_1, s_3 = \alpha_2, \dots, s_M = \alpha_P , \quad (7)$$

матричную систему (6) представляем также в стандартном виде:

$$\sum_{m=1}^{M} s_m \vec{Y}_N^m = \vec{\overline{\mathrm{H}}}_N \quad \Rightarrow \quad Y_{NM} \vec{s}_M = \vec{\overline{\mathrm{H}}}_N , \quad (8)$$

$$\vec{Y}_N^1 = \vec{1}_N, \vec{Y}_N^2 = \vec{X}_N^1, \dots, \quad \vec{Y}_N^M = \vec{X}_N^P , \quad Y_{NM} = (\vec{Y}_N^1 \vec{Y}_N^2 .. \vec{Y}_N^P) ,$$

где Y_{NM} – измеримая матрица наблюдений (ИМН); $\vec{\overline{\mathrm{H}}}_N$ – измеримый вектор средних результатов (ВСР).

Согласно общей теории линейной регрессии система (8) может быть решена, если она полностью определена или переопределена:

$$N \geq M + 1 = P + 2 \Rightarrow \quad Rank Y_{NM} = M . \quad (9)$$

Отметим, что величина (M+1) обусловлена тем, что в число неизвестных помимо M=P+1 неизвестных параметров регрессии необходимо включить также и неизвестное СКО σ_ξ. При выполнения условия (9) статистическое решение задачи линейной регрессии представляется в виде:

$$\vec{s}_M^{\wedge} = Y_{NM}^{-} \vec{\overline{\mathrm{H}}}_N , Y_{NM}^{-} = (Y_{NM}^T Y_{NM})^{-1} Y_{NM}^T , \quad (10)$$

$$(\sigma_\xi^2)^\wedge = \frac{1}{N-M} // \vec{\vec{\mathrm{H}}}_N^\wedge - \vec{\vec{\mathrm{H}}} //^2 = \frac{// \Lambda_{NN}^{M\perp} \vec{\vec{\mathrm{H}}}_N //^2}{N-M}, \quad (11)$$

$$\vec{\vec{\mathrm{H}}}_N^\wedge = Y_{NM} \vec{s}_M^{\,\wedge} = \Lambda_{NN}^M, \quad \Lambda_{NN}^M = Y_{NM} Y_{NM}^-, \quad \Lambda_{NN}^{M\perp} = \mathrm{I}_{NN} - \Lambda_{NN}^M,$$

$$Rank \Lambda_{NN}^M = M, \quad Rank \Lambda_{NN}^{M\perp} = N - M,$$

где Y_{NM}^- – псевдообратная матрица [4]; Λ_{NN}^M – вектор в линейную оболочку из базисных векторов $\{\vec{Y}_N^m, m = 1,2,\ldots,M\}$; $\Lambda_{NN}^{M\perp}$ – ортогональный вектор.

В данной работе наиболее точное решение получено в случае P=3 при различных N с учетом необходимого условия разрешения (9):

$$5 \le N \le 8. \quad (12)$$

Специфической математической особенностью задачи регрессии спортивного результата является то, что в силу довольно однородного состава групп столбцовые вектора ИМН Y_{NM} оказываются хотя и случайными, но с малым угловым расхождением относительно «единичного» вектора $\vec{1}_N$. Последнее обстоятельство требует жесткого контроля точности обращения матрицы Грама $(Y_{NM}^T Y_{NM})_{MM}$, т.к. в случае высокой угловой корреляции («схожести») векторов $\vec{Y}_N^m$ матрица Грама оказывается часто плохо обусловленной [4] с большим динамическим диапазоном собственных чисел в области малых величин. При этом точность обращения матрицы Грама с ростом размерности P>3 (числа учитываемых информативных параметров) начинает резко падать и дальнейшее увеличение размерности P не представляется возможным.

Отметим также, что в данной работе максимальное число возрастных групп с полугодовым периодом N_{max}=17.

Поэтому в силу условия (9) предельное число наиболее информативных параметров ограничивается величиной 15 (в работе [2] – она составляла 6).

Апробация алгоритмов прогноза результативности прыгунов в высоту по различному числу полугодовых возрастных групп

Программа РЕГРЕССИЯ (cor2din.com) содержит следующие разделы:

1. Вызов исходных статистических данных (файл g1_21_9.dat)

2. Шифр файла: TN-M (x1, x2, …, xM) для годовых периодов и TNd-M (x1, x2, …, xM) для полугодовых периодов, где N – число возрастных групп (годовых или полугодовых), по которым проводится прогноз на будущее; M – число информативных параметров ($N \geq M+2$)

3. Интерполирование значений физических параметров на полугодовые или квартальные периоды.

4. Выбор M информативных параметров (из номеров 2-21).

5. Анализ ранга регрессионной матрицы $Y_{N(M+1)}$ методом Грама-Шмидта.

6. Анализ корреляции информативных параметров по годам.

7. Спектральный анализ матрицы Грама $Y^T Y$ размером (M+1)*(M+1).

8. Оценка точности обращения матрицы Грама.

9. Оценка статистических характеристик информативных параметров (средние, СКО, корреляционная матрица).

10. Решение задачи линейной регрессии.

11. Оценка дисперсии шума (СКО=s).

12. Прогнозирование за пределы выбранных возрастных групп, включая прогноз рекордных результатов.

Выводы

1. Задача прогноза результативности спортсменов является задачей интерполяции средней (по возрастной группе) результативности ($\overline{\text{H}}$) в виде линейной комбинации средних значений наиболее информативных физических параметров спортсменов ($\bar{x}_1, \bar{x}_2, ..., \bar{x}_P$) с указанием точности (СКО) прогноза:

$$\overline{\text{H}} = \text{H}_0 + \alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2 + \ldots + \alpha_P \bar{x}_P + \xi, \quad \overline{\xi^2} = \sigma_\xi^2,$$

где $\text{H}_0, \alpha_1, \alpha_2, ..., \alpha_P$ – параметры регрессии; σ_ξ – СКО прогноза.

В условиях априорной неопределенности об СКО прогноза необходимым условием решения задачи прогноза является превышение числа используемых возрастных групп ($N_{\text{ВГ}}$) над числом используемых информативных физических параметров (P), как минимум на две единицы:

$$\text{NВГ} \geq \text{P}+2 .$$

Так, при числе информативных физических параметров P=3 требуются средние значения более, чем по 5-ти годовым возрастным группам (10, 11, 12, 13 и 14 лет) или более, чем по 6-ти полугодовым возрастным группам (10; 10,5; 11; 11,5; 12 и 12,5 лет). При этом можно дать прогноз результативности не только на любой «внутренний» момент времени t_0 $(10 \le t_0 \le 14)$ или $(10 \le t_0 \le 12.5)$, но и на будущие моменты времени t_0>14 или t_0>12.5, включая прогноз рекордных результатов. Для этого достаточно в полученную формулу регрессии подставить значения прогнозных средних значений физических параметров $\{\, \bar{x}_n(t_0),\, n = 1,2,..., P \,\}$:

$$\overline{\mathrm{H}}(t_0) \cong \mathrm{H}_0 + \alpha_1 \bar{x}_1(t_0) + \alpha_2 \bar{x}_2(t_0) + \ldots + \alpha_P \bar{x}_P(t_0) \quad (\pm\sigma_\xi).$$

В частности, при прогнозе по трем параметрам (x_{12}, x_9, x_{21}) по 5-ти годовым возрастным группам (10, 11, 12, 13 и 14 лет) получена следующая регрессионная функция:

$$H = 0.478 + 0.657\,x_{12} + 0.058\,x_9 + 0.806\,x_{21}\,, \quad \sigma = s = 0.9\,см.$$

где x_{12} – высота вылета ОЦТ; x_9 – скорость вылета ОЦТ; x_{21} – прыжок вверх с трёх шагов. При этом прогнозное значение результата для ведущих мастеров спорта составляет 2,36 см, что отличается от их среднего результата (2,33 см) всего на 3 см.

При прогнозе по трем параметрам (x_{12}, x_9, x_{21}) по 6-ти полугодовым возрастным группам (10; 10,5; 11; 11,5; 12 и 12,5 лет) получена следующая регрессионная функция:

$$H = 0.381 + 0.474\,x_{12} + 0.027\,x_9 + 1.369\,x_{21}\,, \quad \sigma = s = 0.2\,см.$$

ЛИТЕРАТУРА

1. Ахметов Р.Ф. Групповые статистические характеристики и факторный анализ многомерной совокупности параметров спортсменов в задачах прогноза результативности // Педагогіка, психологія та медико-біологічні проблеми фізичного виховання і спорту. – 2004. – № 6. – С. 91-104.

2. Ахметов Р.Ф. Прогноз результативности спортсменов на базе статистического факторного анализа и экспертного ранжирования полной

совокупности антропометрических, технических и специализированных параметров // Педагогіка, психологія та медико-біологічні проблеми фізичного виховання і спорту. – 2004. – № 7. – С. 82-95.

3. Баландин В.И., Блудов Ю.М., Плахтиенко В.А. Прогнозирование в спорте. – М.: Физкультура и спорт, 1986. – 193 с.

4. Гантмахер Ф.Р. Теория матриц. – 4-е изд.– М.: Наука, 1988. – 552 с.

5. Крамер Г. Математические методы статистики: Пер. с англ. / Под ред. академика А.Н. Колмогорова. – М.: Мир, 1975. – 648 с.

6. Платонов В.Н. Общая теория подготовки спортсменов в олимпийском спорте. – К.: Олимпийская литература, 1997. – 583 с.

7. Плахтиенко В.А., Мельник В.Г. Прогнозирование в спорте. – Л.: ВДКИФК, 1980. – 79 с.

8. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.

АНОТАЦІЯ

Розглядається простий алгоритм підвищення точності раннього прогнозу результативності спортсменів, оснований на розширенні кількості інформативних спортивних параметрів.

SUMMARY

The item features an elementary algorithm of stepping-up the accuracy of predicting athlets' performance efficiency, based on the increase of informative athletic parameters.