Особливості прогнозування результативності спортсменів як фактора підвищення ефективності навчально-тренувального процесу

Яворська Т. Є.
Житомирський державний університет імені Івана Франка

Анотації:

Розроблено програму прогнозування результатів спортсменів за даними їх спеціальних спортивних параметрів і отриманої функції лінійної регресії. Використано векторний та матричний аналіз для виділення максимально можливої кількості інформативних параметрів. Використано дисперсійний і факторний аналіз для вирішення завдання про мінімально достатнє число інформативних спортивних параметрів. Застосовано теорію багатовимірної лінійної регресії в евклідовому просторі для вирішення завдання прогнозу на основі обмежених статистичних даних за віковими групами.

Яворская Т. Е. Особенности прогнозирования результативности спортсменов как фактора повышения эффективности учебно-тренировочного процесса. Разработана программа прогнозирования результатов спортсменов по данным их специальных спортивных параметров и полученной функции линейной регрессии. Использован векторный и матричный анализ для выделения максимально возможного количества информативных параметров. Использован дисперсионный и факторный анализ для решения задачи о минимально достаточном числе информативных спортивных параметров. Применена теория многомерной линейной регрессии в евклидовом пространстве для решения задачи прогноза на основе ограниченных статистических данных по возрастным группам.

Yavorskaya T. The peculiarities of prognosticating the athletic effectiveness as a factor of stepping up the potency of a training process. The program of prognostication of results of sportsmen is developed from data of their special sporting parameters and got function of linear regression. A vectorial and matrix analysis is used for the selection of maximally possible amount of informing parameters. A dispersible and factor analysis is used for the decision of task about the minimum sufficient number of informing sporting parameters. The theory of multidimensional linear regression is applied in Euclidean space for the decision of task of prognosis on the basis of the limited statistical information on age-dependent groups.

Ключові слова:

прогнозування, лінійна регресія, апроксимація, ваговий вектор.

прогнозирование, линейная регрессия, аппроксимация, весовой вектор.

prognostication, linear regression, approximation, weight vector.

Вступ.

У системі спортивної підготовки велика роль приділяється прогнозу результативності спортсменів як фактору підвищення ефективності навчально-тренувального процесу. Ця проблема є актуальною, оскільки її вирішення забезпечує достатньо обґрунтовані передумови для прийняття управлінських рішень як у сфері спортивної підготовки, так і у сфері змагальної діяльності.

Прогнозування – розробка прогнозів у спорті – є формою конкретизації передбачення перспектив розвитку того чи іншого процесу або явища, характерного для спортивної діяльності [5]. Прогнозування тісно пов’язане з управлінням, тому що забезпечує досить обґрунтовані передумови для прийняття управлінських рішень як у сфері організації спорту, так і у сфері спортивної підготовки, змагальної діяльності.

Прогнозування ґрунтується на використанні методу екстраполяції, що припускає поширення висновків, отриманих зі спостереження над однією частиною якого-небудь явища, на інші його частини [2; 6; 8]. В умовах спорту екстраполяція дозволяє здійснити прогнози зростання результативності на основі вивчення відповідних закономірностей у попередні роки. Завдання прогнозування результатів спортсменів можна вирішити на базі факторного аналізу й динаміки розвитку фізичних параметрів і результатів на певному обмеженому інтервалі часу [1; 6]. Особлива увага приділяється „ранньому” прогнозуванню на період до 17 років за даними початкового періоду (10-13 років).

Дослідження проводилося згідно теми 2.3.5.1п «Удосконалення теоретико-методичних основ управління системою підготовки спортсменів швидкісно-силових видів спорту» Зведеного плану науково-дослідної роботи у сфері фізичної культури і спорту на 2006–2010 рр. Міністерства України у справах сім’ї, молоді та спорту. Номер держреєстрації 0108V008210.

Мета, завдання роботи, матеріал і методи.

Мета дослідження – розробити програму прогнозування результативності висококваліфікованих стрибунів у довжину з розбігу за даними їх спортивних параметрів і отриманої функції лінійної регресії з урахуванням її дисперсії.

Результати дослідження. Оскільки результати та параметри спортсменів у групі мають випадковий розкид (дисперсію), то, говорячи про завдання прогнозування результативності, має сенс розглядати прогноз середньої результативності $\bar{H}(t)$, як функції середніх по групі спортивних параметрів $\vec{X}_P$, що будемо представляти у вигляді матриці стовпця:

$$\vec{X}_P = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_P \end{pmatrix}, \text{P=1,2,…,N}_p\text{-1; N}_p\geq 3,$$

де N_p – повне число спортивних параметрів, включаючи сам результат (H). Повна сукупність P-вимірних групувань з (N_p-1) по P дорівнює числу сполучень з (N_p-1) по P:

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

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

Інформативність різних P-вимірних групувань $\vec{X}_P$ у завданнях прогнозування результативності буде також різною. Питання про вибір оптимальної сукупності найбільш інформативних параметрів (1) при різних P вимагає самостійних глибоких досліджень у рамках окремої НДР. У роботі запропоновано один з альтернативних варіантів вирішення завдання, який цілком прийнятний з погляду точності прогнозу. У першому наближенні розглядається завдання лінійного прогнозу в рамках класичної теорії лінійної регресії (інтерполяції) у математичній статистиці [4; 7-9]. Мова йде про вирахування апроксимації

$$\bar{\mathrm{H}} \cong \mathrm{H}_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_P X_P, \qquad (2)$$

де $\mathrm{H}_0, \alpha_1, \ldots, \alpha_P$ – невідомі параметри регресії, які потрібно оцінити за даними деякої кількості вікових груп. У більш точній постановці наближена лінійна регресія (2) представляється у вигляді:

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

де $\xi(t)$ – помилка прогнозу з нульовим середнім ($M\xi(t) = 0$) і невідомою дисперсією $\sigma_\xi^2 = M\xi^2$ (*M*– оператор математичного очікування – середнього). Якщо в результаті вирішення задачі лінійної регресії на інтервалі часу *T* отримані оцінки невідомих параметрів регресії:

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

то прогнозне значення середньої результативності поза цим інтервалом подається у вигляді:

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

де набір спортивних параметрів { $X_n(t_0), \quad n = 1,2,\ldots,P$ } – задається на прогнозований момент часу t_0. При цьому середньоквадратичне відхилення (СКВ) прогнозу оцінюється величиною $\sigma_\xi(T)$.

Матричне вирішення задачі лінійної регресії результативності за заданою сукупністю найбільш інформативних параметрів

Для оцінки параметрів регресії $\mathrm{H}_0, \alpha_1, \ldots, \alpha_P$ складається така система лінійних алгебраїчних рівнянь:

$$\begin{aligned} &\mathrm{H}_0 + \sum_{n=1}^{P} \alpha_n X_n(t_1) = \bar{\mathrm{H}}(t_1) \\ &\mathrm{H}_0 + \sum_{n=1}^{P} \alpha_n X_n(t_2) = \bar{\mathrm{H}}(t_2) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\mathrm{H}_0 + \sum_{n=1}^{P} \alpha_n X_n(t_N) = \bar{\mathrm{H}}(t_N) \end{aligned} \qquad (5)$$

де N – число вікових груп (у цій роботі N<8). Система (5) подається в матричному вигляді:

$$\mathrm{H}_0 \vec{\mathrm{I}}_N + \sum_{n=1}^{P} \alpha_n \vec{X}_N^n = \vec{\bar{\mathrm{H}}}_N \Rightarrow \qquad (6)$$

$$\vec{\mathrm{I}}_N = \begin{pmatrix} 1 \\ 1 \\ .. \\ 1 \end{pmatrix}_N, \quad \vec{X}_N{}^n = \begin{pmatrix} X_n(t_1) \\ X_n(t_2) \\ \ldots\ldots \\ X_n(t_N) \end{pmatrix}, \quad \vec{\bar{\mathrm{H}}}_N = \begin{pmatrix} \bar{\mathrm{H}}(t_1) \\ \bar{\mathrm{H}}(t_2) \\ \ldots\ldots \\ \bar{\mathrm{H}}(t_N) \end{pmatrix}$$

Вводячи так званий „сигнальний” регресійний вектор (СРВ):

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

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

матричну систему (6) подаємо також у стандартному вигляді:

$$\sum_{n=1}^{M} s_n \vec{Y}_N^n = \vec{\bar{\mathrm{H}}}_N \quad \Rightarrow \quad Y_{NM} \vec{s}_M = \vec{\bar{\mathrm{H}}}_N, \qquad (8)$$

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

де Y_{NM} – вимірна матриця спостережень (ВМС); $\vec{\bar{\mathrm{H}}}_N$ – вимірний вектор середніх результатів (ВСР).

Відповідно до загальної теорії лінійної регресії система (8) може бути вирішена, якщо вона цілком визначена чи перевизначена:

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

Відзначимо, що величина (M+1) обумовлена тим, що в число невідомих крім M=P+1 невідомих параметрів регресії необхідно включити також і невідому СКВ σ_ξ. При виконанні умови (9) статистичне вирішення задачі лінійної регресії подається у вигляді:

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

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

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

$$rang(\Lambda_{NN}^M) = M, \quad rang(\Lambda_{NN}^{M\perp}) = N - M,$$

де Y_{NM}^{-} – псевдозворотна матриця [5]; Λ_{NN}^M – вектори лінійної оболонки з базисних векторів { $\vec{Y}_N^m, m = 1,2,\ldots,M$ }; $\Lambda_{NN}^{M\perp}$ – ортогональний вектор; індекс вгорі „Т” означає операцію матричного транспонірування.

У даній роботі найбільш точне вирішення отримане у випадку P=3 при різних N з урахуванням необхідної умови припущення (9):

$$5 \leq N \leq 8. \qquad (12)$$

Розроблена спеціалізована програма cor2din.pas у оболонці Turbo Pascal. Специфічною математичною особливістю задачі регресії спортивного результату є те, що в силу досить однорідного складу груп стовпцеві вектори ВМС Y_{NM} є хоч і випадковими, але з малою кутовою розбіжністю відносно „одиничного” вектора $\vec{1}_N$. Ця обставина вимагає чіткого контролю точності перетворення матриці Грама $(Y_{NM}^T Y_{NM})_{MM}$, тому що у випадку високої кутової кореляції („схожості”) векторів $\vec{Y}_N^m$ матриця Грама є часто погано зумовленою [3] з великим динамічним діапазоном власних чисел в області малих величин. При цьому точність перетворення матриці Грама зі зростанням розмірності P>3 (числа інформативних параметрів, які враховуються) починає різко падати й подальше збільшення розмірності P не є можливим.

Відзначимо також, що в цій роботі максимальне число вікових груп з піврічним періодом N_{max}=8.

Тому в силу умови (9) граничне число найбільш інформативних параметрів обмежується величиною 6:

$$P \leq N - 2 \leq N_{max} - 2 = 8 - 2 = 6. \qquad (13)$$

Висновки

Задача прогнозу результативності спортсменів є задачею інтерполяції середньої (по віковій групі) результативності ($\bar{\text{H}}$) у вигляді лінійної комбінації середніх значень найбільш інформативних параметрів $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_P)$ із зазначенням точності (СКВ) прогнозу:

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

де $\text{H}_0, \alpha_1, \alpha_2, ..., \alpha_P$ – параметри регресії; σ_ξ – СКВ прогнозу.

В умовах апріорної невизначеності про СКВ прогнозу необхідною умовою вирішення задачі прогнозу є перевищення числа використовуваних вікових груп ($N_{ВГ}$) над числом використовуваних інформативних параметрів (P), як мінімум на дві одиниці:

$$N_{ВГ} \geq P+2.$$

Так, при кількості інформативних параметрів P=3 потрібні середні значення більш, ніж по 5-ти вікових групах (10, 11, 12, 13 і 14 років). При цьому можна зробити прогноз результативності не тільки на будь-який „внутрішній” момент часу t_0 $(10 \leq t_0 \leq 14)$, але й на майбутні моменти часу $t_0 > 14$, включаючи прогноз рекордних результатів. Для цього достатньо в отриману формулу регресії підставити значення прогнозних середніх значень інформативних параметрів $\{\ \bar{x}_n(t_0),\ n = 1,2,...,P\ \}$:

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

Подальший напрямок дослідження може бути спрямований на вдосконалення навчально-тренувального процесу спортсменок, які спеціалізуються у швидкісно-силових видах легкої атлетики.

Література

1. Ахметов Р.Ф. Групповые статистические характеристики и факторный анализ многомерной совокупности параметров спортсменов в задачах прогноза результативности / Р.Ф. Ахметов // Педагогіка, психологія та медико-біологічні проблеми фізичного виховання і спорту: зб.наук.праць за ред. проф. С.С. Єрмакова. — Харків: ХДАДМ (ХХПІ), 2004. – № 6. – С. 91-104.
2. Баландин В.И., Блудов Ю.М., Плахтиенко В.А. Прогнозирование в спорт / В.И. Баландин, Ю.М. Блудов, В.А. Плахтиенко. – М.: Физкультура и спорт, 1986. – 193 с.
3. Гантмахер Ф.Р. Теория матриц / Ф.Р. Гантмахер. – 4-е изд.– М.: Наука, 1988. – 552 с.
4. Крамер Г. Математические методы статистики / Г. Крамер. – М.: Мир, 1975. – 648 с.
5. Платонов В.Н. Общая теория подготовки спортсменов в олимпийском спорте / В.Н. Платонов. – К.: Олимпийская литература, 1997. – 583 с.
6. Плахтиенко В.А., Мельник В.Г. Прогнозирование в спорте / В.А. Плахтиенко, В.Г. Мельник. – Л.: ВДКИФК, 1980. – 79 с.
7. Пугачев В.С. Теория вероятностей и математическая статистика / В.С. Пугачев. – М.: Наука, 1979. – 496 с.
8. Саати Т.Л. Взаимодействие в иерархических системах / Т.Л. Саати // Техническая кибернетика. – 1979. – № 1. – С. 68–84.
9. Harman H.H. Modern factor analysis. – University of Chicago Press, 1960. Русский перевод: Современный факторный анализ. – М.: Статистика, 1972. – 516 с.
10. Lawley D.N., Maxwell A.E. Factor analysis as a statistical method. – Butterworths. – London, 1983. Русский перевод: Факторный анализ как статистический метод. – М.: Мир, 1987. – 413 с.

Надійшла до редакції 22.02.2010р.

Яворська Тетяна Євгенівна

tatiana-82@meta.ua

рекомендовано до друку:

Ахметов Р.Ф. (доктор наук з ФВ, проф.)