Ахметов Р. Ф. Особливості прогнозування результативності спортсменів як фактора підвищення ефективності навчально-тренувального процесу // Молода спортивна наука України. – 2007. – Вип. 11, т. 3. – С. 35–45.

ОСОБЛИВОСТІ ПРОГНОЗУВАННЯ РЕЗУЛЬТАТИВНОСТІ СПОРТСМЕНІВ ЯК ФАКТОРА ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ НАВЧАЛЬНО-ТРЕНУВАЛЬНОГО ПРОЦЕСУ

Ахметов Рустам

Житомирський державний університет імені Івана Франка

Анотація. ***Особливості прогнозування результативності спортсменів як фактора підвищення ефективності навчально-тренувального процесу. Ахметов Рустам. Житомирський державний університет імені Івана Франка.*** Мета дослідження – розробка програми прогнозування результатів спортсменів за даними їх спеціальних спортивних параметрів і отриманої функції лінійної регресії з урахуванням її дисперсії для підвищення ефективності навчально-тренувального процесу. Ставиться та вирішується основна задача прогнозу результативності спортсменів на базі статистичного факторного аналізу й експертного ранжування повної сукупності антропометричних, технічних і спеціалізованих параметрів. У роботі використовувалися такі методи: векторний аналіз у багатовимірному евклідовому просторі; теорія матриць, сингулярні числа та спектральні представлення; дисперсійний і факторний аналіз у математичній статистиці; функціонально-програмне забезпечення вирішення математичних задач на базі сучасного пакету прикладних програм типу Matlab. За допомогою векторного та матричного аналізу виділяється максимально можлива кількість інформативних параметрів. За допомогою дисперсійного та факторного аналізу вирішується завдання про мінімально достатнє число інформативних спортивних параметрів. За допомогою теорії багатовимірної лінійної регресії в

евклідовому просторі вдається вирішити завдання прогнозу на основі обмежених статистичних даних за віковими групами.

Ключові слова: прогнозування, лінійна регресія, апроксимація, ваговий вектор.

Аннотация. ***Особенности прогнозирования результативности спортсменов как фактора повышения эффективности учебно-тренировочного процесса. Ахметов Рустам. Житомирский государственный университет имени Ивана Франко.*** Цель исследования – разработка программы прогнозирования результатов спортсменов по данным их специальных спортивных параметров и полученной функции линейной регрессии с учетом ее дисперсии для повышения эффективности учебно-тренировочного процесса. Ставится и решается основная задача прогноза результативности спортсменов на базе статистического факторного анализа и экспертного ранжирования полной совокупности антропометрических, технических и специализированных параметров. В работе использовались такие методы: векторный анализ в многомерном евклидовом пространстве; теория матриц, сингулярные числа и спектральные представления; дисперсионный и факторный анализ в математический статистике; функционально-программное обеспечения решения математических задач на базе современного пакета прикладных программ типа Matlab. С помощью векторного и матричного анализа выделяется максимально возможное число информативных параметров. С помощью дисперсионного и факторного анализа решается задача о минимально достаточном числе информативных спортивных параметров. С помощью теории многомерной линейной регрессии в евклидовом пространстве удается решить задачу прогноза на основе ограниченных статистических данных по возрастных группах.

Ключевые слова: прогнозирование, линейная регрессия, аппроксимация, весовой вектор.

Summary. ***The Peculiarities of Prognosticating the Athletic Effectiveness as a Factor of Stepping up the Potency of a Training Process. Akhmetov Rustam. Zhytomyr Ivan Franko State University.*** The research target is to develop an athletic effectiveness prognostication program based on athletes' individual specific physical parameters and the deducted function of linear regression relying on the dispersion of training effectiveness improvement. The aim of the research is also to prognosticate athletic effectiveness relying on the statistic factor analysis and full scope of anthropologic technical and specialized parameters expert ranging. The following methods have been employed in the research: multidimensional Euclidian space vector analysis, matrix theory, singular numbers and spectral representation, mathematic statistics dispersion and factor analysis, functional program to solve mathematic problems based on modem package of MATLAB applied programs. The vector and matrix analyses enable to define the maximum possible number of informative athletic parameters. The dispersion and factor analyses enable to define the minimum efficient number of informative athletic parameters. The multidimensional Euclidian space vector analysis enables the age groups statistics prognostication.

Key words: prognostication, linear regression, approximation, weight vector.

Постановка проблеми. В останні роки українським легкоатлетам не вдається перемагати на великих міжнародних змаганнях. Цей факт стимулює фахівців не тільки підвищувати ефективність тренувального процесу, але і продовжувати розробку точності прогнозу результативності спортсменів, що значною мірою буде сприяти якісному відбору в цьому виді спорту. У зв’язку з цим досить важливою є розробка програми прогнозу результативності на базі деякої сукупності параметрів спортсменів.

Аналіз останніх досліджень. 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).

Мета дослідження – розробка програми прогнозування результатів майстрів спорту міжнародного класу зі стрибків у висоту за даними їх спеціальних фізичних параметрів і отриманої функції лінійної регресії з урахуванням її дисперсії.

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

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

де $\mathrm{N_p}$ – повне число спортивних параметрів, включаючи сам результат (Н). Повна множина Р-вимірних групувань з ($\mathrm{N_p}$-1) по Р дорівнює числу сполучень з ($\mathrm{N_p}$-1) по Р:

$$\vec{\mathrm{X}}_P \in U_{\vec{X}_P} = \{\vec{X}_P^{\alpha}, \alpha = 1,2,\ldots,C^P_{N_p-1}\},\ (1)$$

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

Інформативність різних Р-вимірних групувань $\vec{X}_P$ у завданнях прогнозування результативності буде також різною. Питання про вибір оптимальної сукупності найбільш інформативних параметрів з множини (1) при різних Р 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

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

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:

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

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:

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

де набір фізичних 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 $\sigma_\xi(T)$. Наскільки „вдало” отримана оцінка (4), – залежить від багатьох факторів і останнє слово тут за практикою (експериментальною апробацією). Проведена в роботі апробація моделі (4) показує, що вона практично цілком прийнятна. СКВ при цьому не перевищує 3-х сантиметрів, а прогнозований рекордний результат становить 250 см.

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

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

$$\mathrm{H}_0 + \sum_{n=1}^{P} \alpha_n X_n(t_1) = \overline{\mathrm{H}}(t_1)$$

$$\mathrm{H}_0 + \sum_{n=1}^{P} \alpha_n X_n(t_2) = \overline{\mathrm{H}}(t_2) \quad (5)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\mathrm{H}_0 + \sum_{n=1}^{P} \alpha_n X_n(t_N) = \overline{\mathrm{H}}(t_N),$$

де N – число вікових груп (у цій роботі N<18). Система (5) подається в матричному вигляді:

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

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

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

$$\vec{s}_M = \begin{pmatrix} \mathrm{H}_0 \\ \alpha_1 \\ \cdots \\ \alpha_P \end{pmatrix}_M = \begin{pmatrix} s_1 \\ s_2 \\ \cdots \\ s_M \end{pmatrix}, \quad M = P + 1, \qquad (7)$$

$$s_1 = \mathrm{H}_0\ , s_2 = \alpha_1, s_3 = \alpha_2, ..., s_M = \alpha_P,$$

матричну систему (6) подаємо також у стандартному вигляді:

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

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

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

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

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

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 виконання умови (9) статистичне вирішення задачі лінійної регресії подається у вигляді:

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

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

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

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

де Y_{NM}^{-} – псевдозворотна матриця [5]; Λ_{NN}^M – вектори лінійної оболонки з базисних векторів $\{\vec{Y}_N^m, m=1,2,...,M\}$; $\Lambda_{NN}^{M\perp}$ – 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):

$$5 \leq N \leq 8 \, . \; (12)$$

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:

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

Апробація алгоритмів прогнозу результативності стрибунів у висоту за різною кількістю інформативних параметрів

У програмі РЕГРЕСІЯ (cor2din.com) є такі розділи:

1. Виклик вихідних статистичних даних (файл g1_21_9.dat).
2. Шифр файлу: TN-M (x1, x2, …, xM) для річних періодів і TNd-M (x1, x2, …, xM) для піврічних періодів, де N – число вікових груп (річних або піврічних), за якими робиться прогноз на майбутнє; M – число інформативних параметрів (N≥ M+2).
3. Вибір M інформативних параметрів (з номерів 2-21).
4. Аналіз рангу регресійної матриці $Y_{N(M+1)}$ методом Грама-Шмідта.
5. Аналіз кореляції інформативних параметрів за роками.
6. Спектральний аналіз матриці Грама $Y^T Y$ розміром (M+1)*(M+1).
7. Оцінка точності перетворення матриці Грама.
8. Оцінка статистичних характеристик інформативних параметрів (середні, СКВ, кореляційна матриця).
9. Вирішення задачі лінійної регресії.
10. Оцінка дисперсії шуму (СКВ=s).
11. Прогнозування за межі обраних вікових груп, включаючи прогнозування рекордних результатів.

Висновки

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

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

де $\mathrm{H}_0, \alpha_1, \alpha_2, ..., \alpha_P$ – 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 фізичних параметрів (Р), як мінімум на дві одиниці:

$$NВГ \geq P+2\ .$$

Так, при кількості інформативних фізичних параметрів Р=3 потрібні середні значення більш, ніж по 5-ти річних вікових групах (10, 11, 12, 13 і 14 років) або більш, ніж по 6-ти піврічних вікових групах (10; 10,5; 11; 11,5; 12 і 12,5 років). При цьому можна зробити прогноз результативності не тільки на будь-який „внутрішній” момент часу t_0 ($10 \leq t_0 \leq 14$) або ($10 \leq t_0 \leq 12.5$), але й на майбутні моменти часу t_0>14 або t_0>12,5, включаючи прогноз рекордних результатів. Для цього достатньо в отриману формулу регресії підставити значення прогнозних середніх значень фізичних параметрів $\{\ \bar{x}_n(t_0), n = 1,2,..., P\ \}$:

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

Зокрема, при прогнозуванні за трьома параметрами (x_{12}, x_9, x_{21}) по 5-ти річних вікових групах (10, 11, 12, 13 і 14 років) отримана така регресійна функція:

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

де x_{12} – висота вильоту ЗЦТ; x_9 – швидкість вильоту ЗЦТ; x_{21} – стрибок вгору з трьох кроків. При цьому прогнозоване значення результату для майстрів спорту міжнародного класу складає 2,36 см, що відрізняється від їхнього середнього результату (2,33 см) усього на 3 см.

Література

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. Баландин В.И., Блудов Ю.М., Плахтиенко В.А. Прогнозирование в спорте. – М.: Физкультура и спорт, 1986. – 193 с.
3. Гантмахер Ф.Р. Теория матриц. – 4-е изд.– М.: Наука, 1988. – 552 с.
4. Крамер Г. Математические методы статистики: Пер. с англ. / Под ред. академика А.Н. Колмогорова. – М.: Мир, 1975. – 648 с.
5. Платонов В.Н. Общая теория подготовки спортсменов в олимпийском спорте. – К.: Олимпийская литература, 1997. – 583 с.
6. Плахтиенко В.А., Мельник В.Г. Прогнозирование в спорте. – Л.: ВДКИФК, 1980. – 79 с.
7. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.
8. Саати Т.Л. Взаимодействие в иерархических системах // 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.