Рустам Ахметов

Житомирський державний університет

імені Івана Франка, м. Житомир

ФАКТОРНИЙ АНАЛІЗ НАЙБЫЛЬШ ІНФОРМАТИВНИХ ПАРАМЕТРІВ СТРИБУНІВ У ВИСОТУ

Постановка проблеми. В останні роки українським стрибунам у висоту не вдається перемагати на великих міжнародних змаганнях. Цей факт стимулює фахівців не тільки підвищувати ефективність тренувального процесу, але й продовжувати розробку точності прогнозу результативності спортсменів. У зв'язку з цим досить актуальним є виділення найінформативніших параметрів стрибунів у висоту для вирішення в подальшому завдання прогнозу результативності.

Аналіз останніх досліджень і публікацій. Проведений у деяких роботах [1; 4] детермінований аналіз повної сукупності параметрів спортсменів (антропометричних, технічних і спеціалізованих) розкриває їхній фізичний зміст і показує, що всі вони є важливими характеристиками, які в сукупності й визначають, у кінцевому рахунку, спортивний результат (цільову функцію). Однак, детермінований аналіз не відповідає на дуже істотне питання: а яким чином оцінювати кількісно ступінь впливу на результат окремих параметрів чи деякої групи параметрів? Оскільки конкретні значення параметрів залежать випадковим чином від конкретного спортсмена, остільки вони завжди мають деякий випадковий розкид, який можна описати методами математичної статистики [3-5]. При цьому особливе значення має факторний аналіз [2; 3].

Ціль статті. Розробити науково-обґрунтовану методику виділення найінформативніших параметрів стрибунів у висоту в задачах прогнозу їх результативності.

Результати досліджень

Векторні й матричні статистичні характеристики для групи спортсменів

Повна сукупність параметрів, включаючи і спортивний результат (Н), представляється у вигляді деякого N-мірного вектора $\vec{x}_N$ (матриці-стовпця):

$$\vec{x}_N^T = (x_1, x_2, ..., x_N),$$

де „Т" – операція матричного транспонування, $\vec{x}_N^T$ – рядок, $\vec{x}_N$ – стовпець. У цій роботі дослідження обмежується випадком N=21: x_1 = Н – спортивний результат (висота стрибка; називається також цільовою функцією (ЦФ)).

Антропометричні параметри ($x_2 ... x_7$): x_2 – зріст; x_3 – довжина гомілки; x_4 – довжина стегна; x_5 – окружність стегна; x_6 – окружність литкового м'яза; x_7 – вага.

Технічні параметри ($x_8 ... x_{14}$): x_8 – швидкість розбігу перед відштовхуванням; x_9 – швидкість вильоту ЗЦТ (у момент відриву); x_{10} – кут вильоту ЗЦТ; x_{11} – тривалість фази відштовхування; x_{12} – висота вильоту ЗЦТ; x_{13} – імпульс сили відштовхування.

Спеціалізовані параметри ($x_{14} ... x_{21}$): x_{14} – ступінь використання силових можливостей при відштовхуванні (%); x_{15} – біг на 30 м з високого старту (час, секунди); x_{16} – швидкість спринтерського бігу (10 м з ходу); x_{17} – стрибок угору з двох ніг із місця; x_{18} – стрибок у довжину з місця; x_{19} – потрійний стрибок із місця; x_{20} – стрибок угору з штовхової ноги (махом іншої); x_{21} – стрибок угору з трьох кроків.

Вектор спортивних параметрів (ВСП) $\vec{x}_N$ залежить від конкретного спортсмена m=1,2,…,M у групі з M спортсменів (у даній роботі M=12). Залежність ВСП від спортсмена (його номера) і від часу (віку) представляється у вигляді:

$$\vec{x}_N = \vec{x}_N^m(t),\ t = t_1, t_2, ..., t_L, t_0;\ m=1,2,\ldots,M,$$

$$t_n = 10 + (n-1), \quad n = 1,2,...,8,$$

де n – число вікових груп (у даній роботі n=8); t_0 – умовний вік провідних спортсменів. Для простоти залежність ВСП від часу поки-що опускається і вікова група цілком характеризується M-мірним набором N-мірних ВСП спортсменів і подається у вигляді прямокутної матриці X_{NM}, яка називається далі груповою параметричною матрицею (ГПМ):

$$X_{NM} = (\vec{X}_N^1 \vec{X}_N^2 ... \vec{X}_N^M) = (x_{nm})_{NM},\ x_{nm} = \vec{X}_N^m[n],$$

$$X_{NM} = \begin{pmatrix} x_{11}x_{12}\ldots x_{1M} \\ x_{21}x_{22}\ldots x_{2M} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_{N1}x_{N2}\ldots x_{NM} \end{pmatrix}, \qquad (1)$$

де N – число рядків; M – число стовпців матриці; x_{nm} – елементи матриці (n-а компонента (координата) вектора $\vec{X}_N^m$).

Виділяючи окремо спортивний результат $x_1 = H$, ВСП $\vec{x}_N$ можна представити також у блоковому вигляді:

$$\vec{x}_N = \begin{pmatrix} H \\ \vec{y}_{N-1} \end{pmatrix}, \quad \vec{y}_{N-1} = \begin{pmatrix} x_2 \\ x_3 \\ . \\ . \\ x_N \end{pmatrix}, \tag{2}$$

де $\vec{y}_{N-1}$ – (N-1) – мірний вектор фізичних параметрів (ВФП) спортсмена.

У рамках статистичної термінології будемо вважати, що кожний із параметрів x_n (для кожної вікової групи) є деякою випадковою величиною, а ВСП $\vec{x}_N$ – випадковим вектором. Статистичні характеристики ВСП визначаються шляхом арифметичного усереднення:

$$\vec{a}_N = \overline{\vec{x}}_N = \frac{1}{M}\sum_{m=1}^{M} \vec{X}_N^m, \quad a_n = \overline{x}_n = \frac{1}{M}\sum_{m=1}^{M} x_{nm}, \tag{3}$$

$$D(x_n) = \sigma_n^2 = \overline{\Delta x_n^2} = \overline{x_n^2} - \overline{x_n}^2, \quad \Delta x_n = x_n - \overline{x_n}, \tag{4}$$

$$\Phi_{nk} = \overline{x_n x_k} = \frac{1}{M}\sum_{m=1}^{M} x_{nm} x_{km}, \tag{5}$$

$$\Psi_{nk} = \overline{\Delta x_n \Delta x_k} = \Phi_{nk} - \overline{x_n}\,\overline{x_k}, \tag{6}$$

$$\Psi_{nk} = \sigma_n \sigma_k \rho_{nk}, \quad \rho_{nk} = \frac{\Psi_{nk}}{\sigma_n \sigma_k}, \tag{7}$$

де a_n, σ_n^2 – середні значення і дисперсії параметрів x_n ($\sigma_n = \sqrt{D(x_n)}$ – СКВ); Δx_n – флуктуації параметрів щодо середніх значень; Φ_{nk}, Ψ_{nk} – взаємні кореляції та коваріації параметрів x_n, x_k; ρ_{nk} – взаємні коефіцієнти кореляції ($/\rho/ \le 1$).

Відповідні кореляційні та коваріаційні матриці представляються в алгебраїчному вигляді:

$$\Phi_{NN} = \frac{1}{M} X_{NM} X_{NM}{}^T = \overline{\vec{X}_N^m \vec{X}_N^{mT}} = \frac{1}{M}\sum_{m=1}^{M} \vec{X}_N^m \vec{X}_N^{mT}, \tag{8}$$

$$\Psi_{NN} = \overline{\Delta\vec{X}_N^m \Delta\vec{X}_N^{mT}}\,, \tag{9}$$

де риска зверху означає арифметичне усереднення за номером m $(m = 1,2,\ldots,M)$, тобто статистичне усереднення по спортсменах у групі з рівномірним дискретним розподілом імовірностей $p_m = 1/M$.

Відзначимо, що вихідна ГПМ X_{NM} містить інформацію не тільки про зв'язок різних параметрів x_n між собою, але і ступеня „схожості” чи параметричної близькості спортсменів між собою в групі. Для цього досить розглянути близькість чи кореляцію векторів $\vec{X}_N^m$, оцінюючи скалярні добутки векторів [2]:

$$B_{mk} = \frac{1}{N}(\vec{X}_N^m, \vec{X}_N^k) = \frac{1}{N}\sum_{n=1}^{N} \vec{X}_N^m[n]\vec{X}_N^k[n]\,. \tag{10}$$

Матрицю скалярних добутків (МСД) можна представити через ГПМ X_{NM}:

$$B_{MM} = \frac{1}{N} X_{NM}^T X_{NM}\,. \tag{11}$$

Мірою параметричної близькості спортсменів у групі може служити алгебраїчна кореляція $\vec{X}_N^m$ векторів чи так званий косинус кута між векторами:

$$R_{mk} = \cos\varphi_{mk} = \frac{(\vec{X}_N^m, \vec{X}_N^k)}{//\vec{X}_N^m//*//\vec{X}_N^k//}\,, \tag{12}$$

$$//\vec{X}_N// = \sqrt{(\vec{X}_N, \vec{X}_N)} = \sqrt{\sum_{n=1}^{N} x_n^2}\,,$$

де $//\vec{X}_N//$ – норма вектора в N-мірному евклідовому просторі [2].

Багатомірний нормальний закон розподілу та кореляційний еліпсоїд вектора спортивних параметрів. Задача факторного аналізу

Нормальна щільність імовірності ВСП представляється в стандартному вигляді [2]:

$$W(\vec{X}_N / \bar{\vec{X}}_N, \Psi_{NN}) = \frac{1}{\sqrt{(2\pi)^N \det(\Psi_{NN})}} \exp\{-\frac{1}{2}(\Psi_{NN}^{-1}\Delta\vec{X}_N, \Delta\vec{X}_N)\},$$

де $\det(\Psi_{NN})$ – визначник коваріаційної матриці Ψ_{NN}.

Перетин щільності ймовірності визначає у просторі ВСП так званий кореляційний еліпсоїд:

$$W(\vec{X}_N /.) = const \quad \Rightarrow \quad (\Psi_{NN}^{-1}\Delta\vec{X}_N, \Delta\vec{X}_N) = const'. \tag{13}$$

Зокрема, у випадку незалежних параметрів x_n рівняння кореляційного еліпсоїда представляється у вигляді:

$$\left(\frac{x_1 - \bar{x}_1}{\sigma_1}\right)^2 + \left(\frac{x_2 - \bar{x}_2}{\sigma_2}\right)^2 + \ldots + \left(\frac{x_N - \bar{x}_N}{\sigma_N}\right)^2 = const'.$$

Відзначимо, що при належному виборі постійної $const'$ ВСП $\vec{X}_N$ знаходиться з високою ймовірністю всередині свого кореляційного еліпсоїда. У загальному випадку багатомірний кореляційний еліпсоїд характеризується своїми розмірами й орієнтацією, що визначаються в результаті рішення задачі про приведення квадратичної форми (13) до канонічного вигляду.

На відміну від більшості відомих робіт, у даній роботі факторний аналіз розглядається з позицій аналізу орієнтації та розмірів багатомірного кореляційного еліпсоїда повного вектора спортивних параметрів (ВСП). При цьому виділяється так званий принцип локалізації ВСП в обмежених підпросторах меншої розмірності, коли розміри кореляційного еліпсоїда в деяких головних напрямках стають нехтувано малими величинами. Потрібно, однак, підкреслити одну специфічну особливість статистичної обробки параметрів у малій групі спортсменів. Це принципова обмеженість числа спортсменів у групі (M = 12), що може призвести до великих відносних погрішностей середньоарифметичних оцінок невідомих статистичних середніх (при M = 12 вони складають 30-47%). У зв'язку з цим необхідно, насамперед, уточнити, а з якою основною метою оцінюються групові статистичні параметри? І який узагалі мають сенс „арифметичні" статистичні характеристики? У цій роботі основною метою є вирішення завдання прогнозу результативності за деякою сукупністю інформативних параметрів спортсменів у залежності від методики тренування. Тому на першому етапі досліджень питання впливу погрішностей арифметичних оцінок самих статистичних характеристик у цій роботі поки опускаються, а арифметичне усереднення розглядається просто, як аналог і окремий випадок статистичного усереднення (з рівномірним розподілом імовірності) для вирішення питань локалізації та факторного аналізу ВСП. Обґрунтуванням і критерієм корисності такого підходу є досить прийнятне для практики вирішення кінцевого завдання прогнозу результативності.

Сингулярні числа ГПМ і максимальне число найбільш інформативних параметрів спортсменів

У більшості випадків число аналізованих фізичних параметрів перевищує кількість спортсменів у групі: $N > M$.

У цьому випадку ранги симетричних матриць Φ_{NN} і B_{MM} збігаються на рівні M:

$$Rank\Phi_{NN} = RankB_{MM} = M \ .$$

Це випливає з того, що строкові та стовпцеві ранги довільних матриць збігаються [4]. Більше того, можна показати, що ненульові власні числа матриць $(X_{NM} X_{NM}^T)_{NN}$ і $(X_{NM}^T X_{NM})_{MM}$ збігаються та дорівнюють квадратам сингулярних чисел ГМП X_{NM} [2].

Таким чином, у випадку N=21 і M=12 серед двадцяти фізичних параметрів можна методами математичної статистики виділити для задач прогнозу не більш дванадцяти інформативних параметрів. У наступних науково-дослідних роботах представляється доцільним формувати об’єднані групи спортсменів з декількох автономних груп для забезпечення нерівності M>N. Тоді для задач прогнозу результативності можна використовувати всі N параметрів.

Висновки

Отримані результати дозволяють зробити такі висновки:

1. Задача факторного аналізу про виділення найбільш інформативних параметрів спортсменів означає, власне кажучи, розкриття області локалізації вектора фізичних параметрів (ВФП) у деякому обмеженому підпросторі повного багатомірного евклідового простору параметрів. При цьому базисом підпростору є набір перших „значимих” власних векторів коваріаційної матриці ВФП, які визначають орієнтацію кореляційного еліпсоїда ВФП. Власні

значення коваріаційної матриці ВФП визначають розмір кореляційного еліпсоїда, у якому локалізується ВФП.

2. Спектральний аналіз кореляційних матриць параметрів підтверджує теоретичний висновок про максимальне число інформативних параметрів, яке дорівнює числу спортсменів у групі (M = 12). При цьому спостерігається різке падіння власних чисел матриць, починаючи з номерів 4-7. Звідси випливає, що для завдань прогнозу ЦФ на першому етапі достатньо обмежитися трьома-шістьма найбільш інформативними параметрами: x_{12} (висота вильоту ЗЦТ); x_9 (швидкість вильоту ЗЦТ); x_{21} (стрибок угору з трьох кроків розбігу); x_5 (швидкість розбігу перед відштовхуванням); x_{15} (біг на 30 м з високого старту); x_{14} (ступінь використання силових можливостей при відштовхуванні).

Література:

1. Баландин В.И., Блудов Ю.М., Плахтиенко В.А. Прогнозирование в спорте. – М.: Физкультура и спорт, 1986. – 193 с.
2. Гантмахер Ф.Р. Теория матриц. – 4-е изд.– М.: Наука, 1988. – 552 с.
3. Крамер Г. Математические методы статистики/ Пер. с англ. Под ред. академика А.Н. Колмогорова. – М.: Мир, 1975. – 648 с.
4. Плахтиенко В.А., Мельник В.Г. Прогнозирование в спорте. – Л.: ВДКИФК, 1980. – 79 с.
5. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.

Анотація. У статті розглядаються актуальні питання виділення найбільш інформативних узагальнених параметрів стрибунів у висоту з деякої повної сукупності спортивних параметрів для вирішення в подальшому важливого завдання прогнозу результативності.

Annotation. The paper considers urgent issues of singling out the more important parameters of high-jumping to further prognosticate the performance.

Довідка про автора:

1. Ахметов Рустам Фагимович – кандидат педагогічних наук, професор, завідувач кафедри теорії і методики фізичного виховання, декан факультету фізичного виховання і спорту Житомирського державного університету імені Івана Франка, заслужений працівник фізичної культури і спорту України
2. 10014, м. Житомир, вул. Старий Бульвар, 10, кв. 44; тел.: 22-11-80 (дом.)