

中国武術の表演試合における問題点（2）

—5年後の追跡研究—

七 堂 利 幸※(武当派(拳)研究会)

I はじめに

前報告（1）では競技審判員の審判員間信頼性について解析した¹⁾。その結果は中国武術の表演競技の審判員間信頼性は、体操競技に比べて大変低いものであった。信頼性とは得点の正確さであり、信頼できない採点方法は妥当とされない²⁾。

ただ、審判方法の不備と選手エントリー順の公平性は別の問題である。審判方法が改善されなければ、せめて選手エントリー順の公平性は保たれなければならない。

1990年の前報告³⁾は、無作為に選手出場順を決めていたなかった。以後連盟はそれを改善したという。この選手エントリー順の公平性を中心に、1990年の前報告と同じ問題点が果たして5年経過して改善が見られたか追跡調査した。

II 方法

表演大会の審判方法は、原則として10点満点からの減点法（1/100の位まで表示）で五名の各審判員（審判長1、審判員4名）の最高点、最低点を除いてあとの三名の平均を出して優劣を決める方式を取っている。

(1) 選手出場順に得点分布図を描き、その経時的变化に一定の傾向がないか検討した。分布の傾向をみるトレンドの検定は前報告³⁾同様、ケンドール、スピアマン型の検定を行った⁵⁾。トレンド探索のため移動平均法も使った。

(2) また、得点系列に対しラン検定⁶⁾を行い、順位の無作為化を検討した。

(3) 前の人を基準に採点していないかどうか見るためコレログラムを検討した。分布グラフから

トレンドを探り、幹枝表示から度数分布を、箱型図から外れ値と分布の様相を見ることにより、採点の状況を推定した。

(4) 独立した採点方法をとっているか確認するため、予選で選ばれた選手を対象に、予選と決勝得点の平均値を比較検討した。

(5) 得点精度が改善されたかどうか、入手できた11回大会の17競技種目の標準偏差で前報告³⁾の結果（19種目）と比較した。

(6) 第1回大会から第6回大会までの優勝者得点を比較して、技術のレベルアップを見ようという商業誌があった⁷⁾。そこで本研究では、7回以後11回大会までの全優勝選手平均値の経時的変化を調べた。

以上の解析材料は、1994年の11回大会の資料を中心にして、1993年の第10回大会の予選と決勝のデータも使った。

ただ、使用したデータは、審判員間の信頼性をみた11回大会の38種目のうち20名以上のエントリー数があったものと、10回大会の成績は全種目入手できたので、これも併せて解析した。5%の有意水準のt値を考えた時（等分散）、自由度20前後以降まではt値は安定しないで⁴⁾、分布型も同定しにくい。出場選手が20人（組）未満の種目は、このサンプルサイズを考慮して、前報告同様対象外とした。

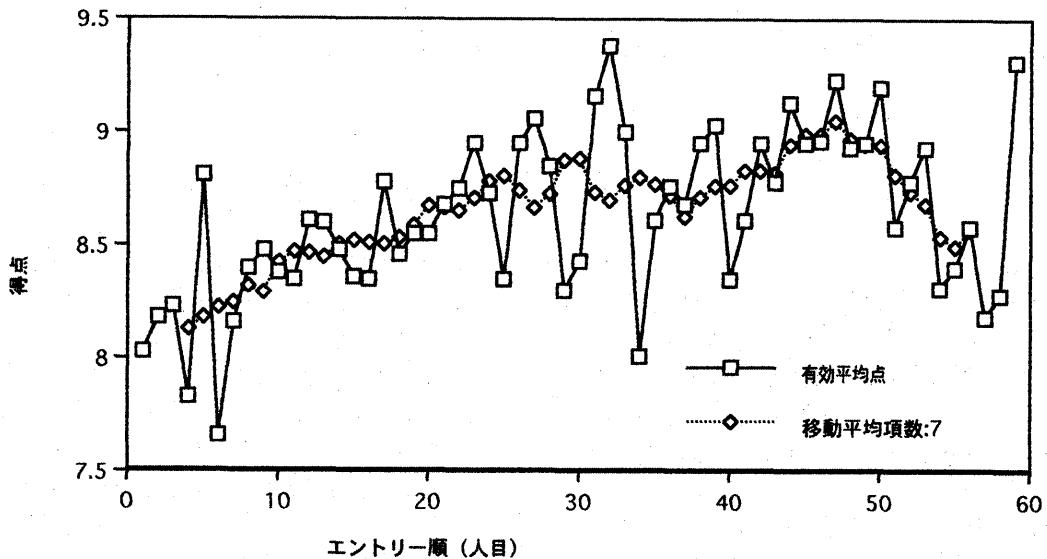
III 結果

調査対象から除外された種目は、出場選手が20人（組）未満の種目で、10回大会では48種目中16種目、11回では48種目中12種目であった。

①大阪で行われた1993年の男子長拳規定は、ラン検定で（有意水準 $p = 0.0008$ ）高度に有意となった（図1）。また、エントリー順が無作為でないことを表わす。これは顕著な上昇トレンドも見られた。この他では、1993年男子伝統器械（ $p =$

* 〒543 大阪市天王寺区上本町6-3-31-1111号

電話 06-779-0339



ラン検定 ($p = 0.0008$) でエントリー順序のランダム性
が否定される。

図1 男子長拳規定(1993)の経時的变化

0.0013), 1994年女子太極剣・刀 ($p = 0.0026$)
がエントリー順が無作為ではなかった。

②一方, 1994年東京で行われたアジア大会予選の結果は移動平均をとると, エントリー順に全体の約2/3までが高得点で, 残り約1/3以下が下降して低得点のグラフとなる。その高得点域の第何組という枠内には過去の大会(3年間)の優勝者や入賞者が, 別々に位置している。この傾向は男子長拳規定, 女子長拳規定, 女子楊式太極拳規定, 女子剣術規定(図2), 女子楊式・88式太極拳, 男子陳式太極拳規定, 男子刀術規定でも見られた。つまり, 7種目に同一傾向の時系列分布があった。

③1994年11回大会種目のトレンドの検定では, 男子総合太極拳(上側ケンドール, $p = 2.63\%$), 伝統器械(上側ケンドール, $p = 0.87\%$), 男子楊式太極拳規定(上側ケンドール, $p = 6.77\%$), 女子太極剣・刀(上側ケンドール, $p = 0.00\%$)に, 相変わらず全体的な上昇トレンドが見られた。

今回, 特異的なのは, 全体的な下降トレンドが, 1994年男子長拳規定(上側ケンドール, $p = 0.07\%$), 女子長拳規定(上側ケンドール, $p = 0.21\%$), 女子剣術規定(上側ケンドール, $p = 0.39\%$)で見

られたことである。続いて, 女子楊式・88式太極拳では上側確率はケンドール型で $p = 15.43\%$, 男子槍術規定では, 同じく $p = 15.46\%$ で, 下降傾向が認められた。

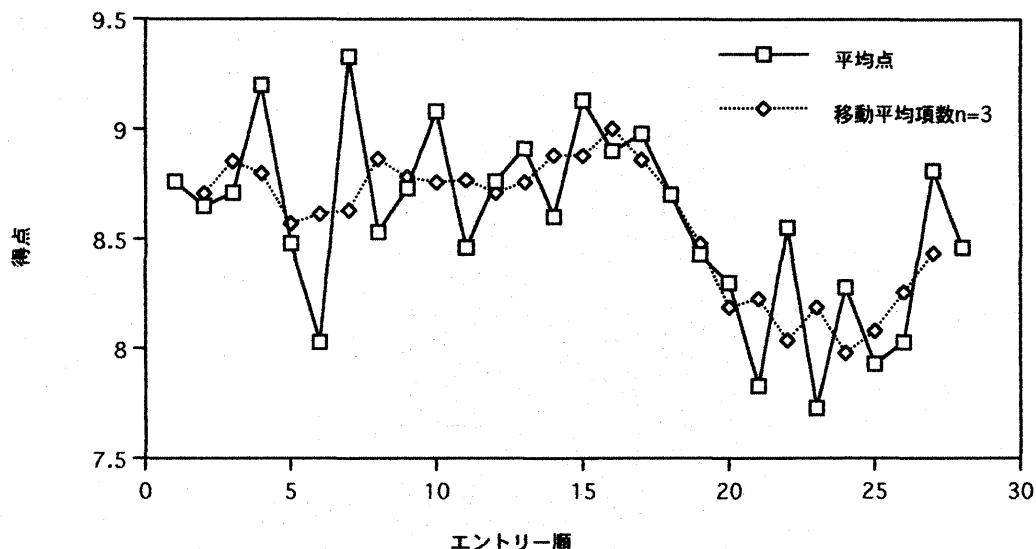
男子拳術A(形意拳・八卦掌・八極拳)は1993年, 1994年ともに同一枠内の三種目に合わせて部分的上昇トレンドが前報告³⁾と同様に発生している。

分類出来なかったその他の種目については, やはり過去の大会(3年間)の優勝者や入賞者に対し部分的な上昇トレンドが発生し, 幹枝表示で偏りのある傾向が認められる。またコレログラムより得点前後にまだ自己相関が認められる。

④図3のとおり, 同一選手での予選と決勝の得点の比較も行ったが, 決勝の平均得点が高かった。決勝平均得点と予選平均得点の差は, 図3のように標準偏差と比較出来るようにした。また標準偏差も決勝の方が大きかった。

⑤得点精度は5回大会19種目, 11回大会17種目の予選得点の標準偏差でみた。精度の平均値 $m \pm$ 標準偏差 SD を求めると, 前者は 0.286 ± 0.091 , 後者は 0.283 ± 0.115 となった。図4のように, 速く動く競技種目ほど, ゆっくり動く種目より精

七堂：中国武術の表演試合における問題点（2）



前から約2/3が高得点で残り約1/3が低得点で、1994年では7種目にこの傾向が認められた。
また、トレンドの検定の結果、下降トレンド ($p = 0.39\%$) が認められた。

図2 女子剣術規定（アジア大会予選、1994）の経時的变化

予選決勝得点比較

注目

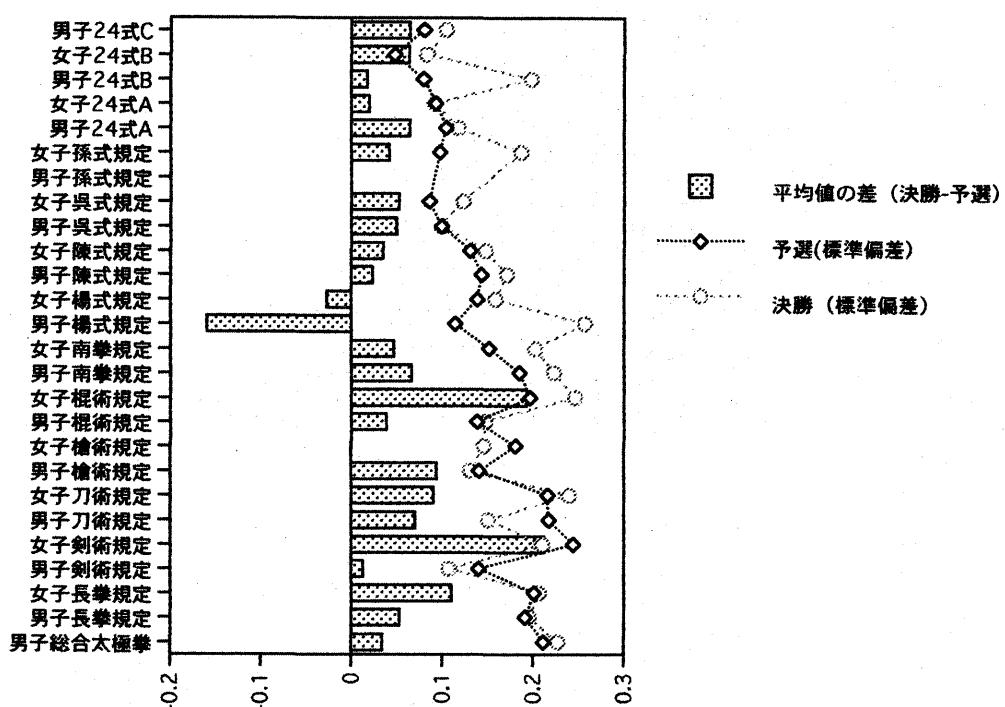
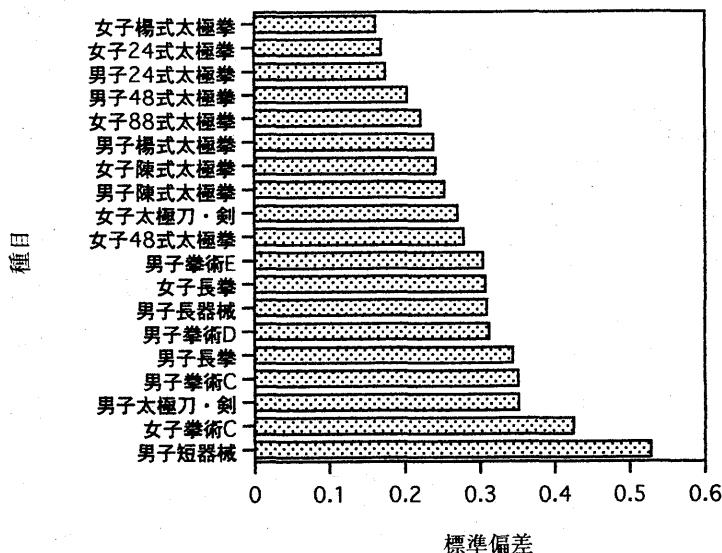
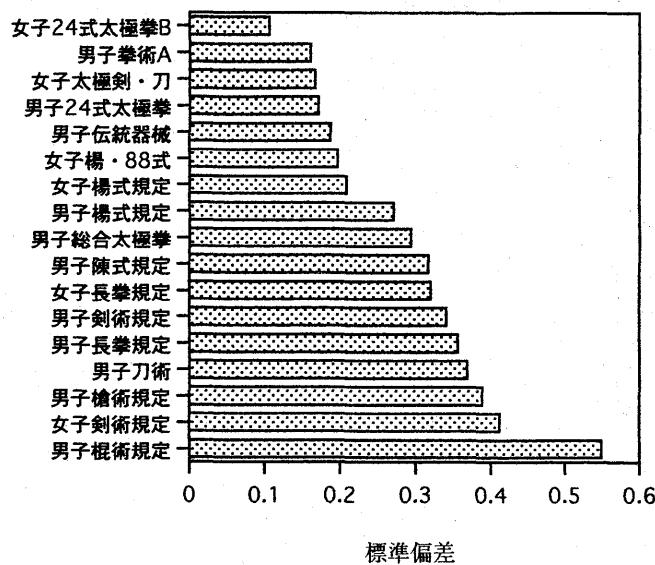


図3 平均値の差（決勝 - 予選得点）と標準偏差

第5回大会 (1988)



第11回大会 (1994)



種目名の不一致はあるが、精度の改善は読み取れず、動作の速い種目ほど、遅い種目に比較して精度は悪い。

図4 得点の測定精度

度は悪い傾向であった。

⑥優勝者得点を比較する場合、競技種目数を合わせるか、同一種目で比較する必要がある。7, 9, 11回大会の全優勝者得点を計算すると、m ± SD はそれぞれ、 9.034 ± 0.339 , 8.964 ± 0.262 , 9.023 ± 0.369 と、この4年間で優勝者得点は減少している。

IV 考察

①大阪市で行われた1993年の男子長拳規定は、エントリー順が無作為でないことが判明した。また顕著な上昇トレンドが見られた。この10回大会では地元開催地がエントリー順の決定など事務処理をした。実質的にその事務局に所属するA選手が、長拳を含めた3種目に最後尾にエントリーしている。無作為にエントリー順を決めているというには、3種目とも最後にエントリーというのは、偶然でないのではないかという疑問が当然湧く。今回エントリー順が無作為に行われていないことが統計的に証明され、この個別例への疑問を支持する結果となった。

この1993年長拳規定の他、1993年男子伝統器械、1994年女子太極剣・刀は、出場順が無作為ではないことがランク検定で証明されたが、この検定は系列の連を調べているだけで、この三種以外はエントリー順が無作為で公平であったということではない。他の原理の乱数検定⁸⁾で今回調べていないし、実施した方法で見つけることが出来なかったと言う意味である。

②得点の上昇トレンドが明白なものが4種目に前報告同様¹⁾見られた。全体的な上昇トレンドが見られなくとも、得点分布に一定の傾向があり、前回優勝・入賞者がピークに位置する。前回入賞者だけが高得点であれば問題ないが、その周辺に部分的な上昇トレンドや自己相関（前の選手得点が後の選手得点に影響する）も認められた。

③7種目の複数例に移動平均による同じ様な特異な分布パターンが見られた。偶然の一一致にしてはこの異常な7種目の同一パターンは多すぎよう。

全体的な上昇トレンドの存在を前報告³⁾で指摘した。これを改善するためには、後半に過去の入賞者をおくと上昇傾向は改善できない。そこで過

去の大会（3年間）の優勝者や入賞者を意図的に前半の枠内に配置すると、簡単にこのように7種目に同一傾向の時系列分布を作ることが可能である。これは出場順位のランダム性、公平性への疑問点である。

④今回、特筆すべきは、全体的な下降トレンドが、1994年の男・女の長拳規定、女子剣術規定で証明されたことである。前2/3が高く、後1/3が低いという特異な分布パターンをとったこの3種目は、結局全体的に下降している。続いて、女子楊式・88式太極拳や男子槍術規定でも有意水準に達しなかったが、同じ下降傾向にあった。上昇トレンドは後半に前回の高得点者がいるので、初めの点数を抑え、それを基準に前の点数に比較した採点方法を順次とったため生じたものであることは、1990年の報告³⁾でほぼ明らかである。では11回大会で異常な下降トレンドは何故生じたのか。前後の得点に自己相関がみられることから、前の人を基準に順次採点していることは、上昇トレンド発生の場合と同じと考えられる。全体的に見れば下降トレンドを呈していたが、前2/3の高得点は、上昇トレンドの原因同様、前回の入賞者を心理的に意識した採点の可能性が大きい。前回優秀な人はどこにエントリーされようと、その人が評価されるべきレベルであれば、その人だけ高得点になるはずで、特異な得点分布パターンは発生しにくい。これは上昇トレンドを抑えるため、前回入賞者を前半のしかも組単位の枠内に意図的に分けて配置させた可能性が高い。

過去、第1回大会で上昇トレンドのため後半有利で最初に作り上げられた入賞者ができる。その後、第2回大会から6回大会まで、前回入賞者を後半に意図的に置いた大会では上昇トレンドが発生した。それを私どもが指摘したので³⁾、前回入賞者を今度は前半に置いたため、このような予想外に異常な下降トレンドが発生したと考えるのが自然ではないだろうか。

⑤得点精度は5回、11回大会で、平均0.01の差があった。採点は1/100の位まで行われるため、これをわずかな精度の改善があったとみるのは早計である。総合太極拳を入れれば11回大会では17種目中11種目の規定種目が新設されており、種目

数が全く異なる。これだけの規定套路を導入しても0.01という改善しかなかったと見るべきであろう。

これはルールの不完全さと審判の訓練の不備を強く予想させ、審判員間の信頼性の低さを異なった面から支持している。採点方法が変わらないのであろうから、得点精度の改善はないのは当然であろう。

競技速度の速い種目ほど、速度の遅い種目より、精度が悪い傾向は前報告³⁾と変わらない。

⑥同一選手の予選と決勝の得点の比較も行ったが、決勝の得点が高かった。これも前報告³⁾と同じ結果で、決勝で全員が予選よりうまくなつたことになる。しかし、建て前である独立した採点法（減点法）をとっていれば、これは考えにくいくことである。依然として採点が前の選手と比較するという相対的な方法をとっているから、予選と違って決勝では最初の選手の基準点が変わり、その後に続く競技得点は変わる可能性がある。決勝では限定された人数で競うため、得点の幅が少なくてすみ最初の設定得点が高くなる。それで決勝は予選より得点が高くなる。これでは各個人の上達の度合が、得点で経年的に測れない。そして図3から分かるように、予選より決勝の方が点数のバラツキ（標準偏差）が大きい。また、報告(1)¹⁾から審判員間の信頼性も、予選より決勝が高くなつた。この結果から、決勝では審判員間の採点は予選より一致するが、得点精度が悪くなっているといえる。審判員間信頼性の比較的高かった¹⁾楊式太極拳規定、南拳規定も予選と決勝の得点差が起こっているから、同一選手の予選と決勝の得点差は、単に審判員間の信頼性が低いから起こっている現象とは考えにくい。

報告(1)¹⁾では、予選より決勝の方が、審判員間の信頼性が高かったが、これは決勝が小人数に絞り込まれた結果起りうる現象であろう。以上により、選手個人に独立した採点が行われていないことが強く疑われる。

⑦優勝者得点の推移（1回から6回）から、技術レベルアップを見ようとした商業誌があった⁷⁾。それに対しては前報告³⁾で、前の人に対して相対的方法で採点しているため、最初の選手得点の設

定、出場人数、トレンドの傾き次第で優勝得点が変化する可能性があることを指摘した。予選得点は毎年変動し、同一選手得点で、予選より決勝の方が競われる得点が高い。このような状況で毎年優勝者得点が決定しているので、毎年の優勝者得点だけを比較検討して、表演技術を論じるのは余りに軽率であろう。

また男女併せて第5回大会は25種目、6回は33種目、7回は49種目、11回大会は56種目で種目差の問題がある。優勝者得点を比較する場合、競技種目数を合わせるか、同一種目で比較する必要がある。

第6回から第11回まで、同一種目で全体的に優勝者の得点が上昇していることは5年間では無かった。また種目数が同じである第7・9・11回大会の全優勝得点の平均点は減少し、一見これは技術が年々レバーアップしていることになってしまふ。

幹枝表示で9、11回大会は低得点側で急峻な立ち上がりをみせたのは、得点の下限が抑えられ⁹⁾、優勝戦も最初の人を基準に採点していることが示唆される。

⑧報告(1)¹⁾と本報告の結果、5年後の1994年を経過しても、まだ中国武術の表演大会は競技として成立しているとは言い難い。套路検査員をおいても、規定套路を作っても改善は無かった。連盟がただ黙ってこの中国製ルールを施行している状況では、表演大会は競技としての成立は将来にわたっても容易でない。ルールを全面的に見直す必要がある。選手出場順位の公平性、審判員の信頼性、得点の妥当性ある競技を、連盟はもう一度根本から再構築する必要がある。しかし、このまま試合を継続するのであれば、せっかく今努力している選手のことを考え、せめて選手出場順位の公平性だけは確保すべきであろう。延べ千数百人の選手たちは練習し出場費を払っていることを忘れないようにしたい。そして、私どもが指摘した問題点を出場する選手に、説明し了解（Informed consent）をとる社会的責任が連盟にはあろう。

⑨ただエントリー順の無作為化の方法の詳細は公開すべきである。乱数を発生して無作為にしているというような粗い説明だけでは信用できない。

無作為化の実際を第三者に立ち合わせて公開し、あるいは順位の無作為化は第三者機関に委ねる方法を検討すべきである。連盟自身で問題点を改善出来ていないため、審判結果を万人が容易に監視出来るよう全面的な情報公開を求めたい。

V まとめ

5年経って中国武術表演競技の審判方法は改善されたかという問題で、今回の追跡調査の分析結果から次のことが判明した。

①ラン検定の結果、明らかに出場順位が無作為でないものがあった。依然として出場順位の公平性は疑問が残る。

②得点の上昇トレンドが4種目に見られ、特筆すべきことは今回下降トレンドが3種目に見られた。エントリー順の前約2/3が高得点、後約1/3が低得点という得点分布に一定の傾向があり、前に前回入賞者が位置する。この異常な同一パターンは7種目に及んだ。

③得点精度は標準偏差でみた。第5回、11回大会を比較した。11回大会では11/17種目に規定拳が導入されたりして大幅に種目が異なるため、そのまま比較はできないが、標準偏差の全平均値を比較すると精度の改善は見られたとは言い難い。速い動作速度の競技は遅い競技より、精度が悪い傾向は変わらない。

④同一選手の予選と決勝の平均得点の比較も行ったが、前報告同様、決勝の得点が高かった。これも各選手に独立した減点法がまだとられていない証拠となる。予選より決勝の方が標準偏差が大きく得点精度は悪かった。

⑤毎年の優勝者得点の経時的変化から、選手レベルの向上を測ろうとする安易な考えは問題がある。

⑥規定拳が採用されるようになり、5年経っても改善は見られなかった。表演大会は依然、採点

競技としては成立しているとは言い難い。出場順位の公平性、審判員間の信頼性、得点の妥当性のある審判方法を日本連盟はもう一度根本から再構築する必要がある。このまま試合を継続するなら、選手エントリーの公平性だけは確保すべきである。

⑦また、誰でもアクセスできる全面的な得点の情報公開を提案した。

統計ソフトはMINITAB、統計解析ハンドブックIV・ノンパラ編（共立出版）を使った。

文献

- 1) 七堂利幸：中国武術の表演試合の問題点（1）－審判員間の信頼性について－、武道学研究29(1): 26~33, 1996
- 2) Margaret J. Safrit (1980) : 体育アセスメントと評価、遊佐・永田・宮崎・青山訳、泰流社, 85 - 124, 1982
- 3) 七堂利幸、朝倉宏之、藤田誠：中国武術の型(表演)試合における審判方法の問題点、武道学研究, 22(3) : 57-64, 1990
- 4) 杉田暉道：臨床統計入門(1), 臨泌, 32(1), 41-46, 1978
- 5) 白旗慎吾：パソコン統計解析ハンドブックIX, 共立出版, 236 - 252, 336 - 343, 1987
- 6) S. ジーゲル：ノンパラメトリック統計学、藤本監訳、マクロウヒル, 238 - 246, 1983
- 7) 武術編集部：優勝得点にみる日本武術のレベルアップ、武術秋号、福昌堂, 46-47, 1989
- 8) 宮武修、脇本和昌：乱数とモンテカルロ法、森北出版, 36-50, 1978
- 9) 石川馨：品質管理入門、日科技連, 120 - 126, 1990

EXAMINATION OF JUDGING PROCEDURES FOR CHINESE MARTIAL ARTS FORM COMPETITIONS

— Part 2. Follow-up research after five years —

Toshiyuki SHICHIDO (Martial Arts Research Group)

In our last report we demonstrated the extremely low inter-rater reliability of the judging for Chinese martial arts form competitions. However, the level of fairness of the order entry for competitors is another problem. We compared this with the problems in our last report (Shichido, et al., 1990). The results are as follows;

- (1) The results of the run test proved that the order of appearance is not always random.
- (2) In four events there was a clear upward trend in scoring. Even though the upward trend overall was no longer visible, there was a definite tendency in the distribution of scores. This was seen in seven events. We must make special note of the fact that an unusual downward was observed in three events. There are doubts about the fairness of the order of entry of competitors.
- (3) The accuracy of scoring was seen in the standard deviation. We looked at three competitions (first, fifth, and eleventh). However, even though the number of events in each of these competitors was different, comparison of scoring accuracy between them shows no sign of improvement. The accuracy of judging for events with quick movements was also poor compared to that of events with slow movements.
- (4) We also compared the scores for the same competitor in both preliminary and final rounds. The avarage score and standard deviation of the final round was high compared to the preliminary round. This proves that the method scoring for each individual competitor is not independent.
- (5) The difference in average scores for the winners of the seventh through eleventh competitions shows a downward trend. However, it is impossible to measure the improvement in competitor's scoring simply from observing those of the winners.

There was no discernible improvement over the five years since the inclusion of required forms. It remains difficult to say the competition is organized as a scored match. At the very least, the Japan Wushu Taijiquan Federation should strive for fairness in order of appearance of competitors.