



## 競馬データにみられる東西調教師群間の成績比較

|       |                                                                                                         |
|-------|---------------------------------------------------------------------------------------------------------|
| メタデータ | 言語: Japanese<br>出版者:<br>公開日: 2013-08-27<br>キーワード (Ja):<br>キーワード (En):<br>作成者: 野田, 明男<br>メールアドレス:<br>所属: |
| URL   | <a href="http://hdl.handle.net/10271/26">http://hdl.handle.net/10271/26</a>                             |

# 競馬データにみられる東西調教師群間の成績比較

野田 明男

(総合人間科学講座・数学)

## A Comparison of Results in the Horse Racing Data Between the East and West Groups of Trainers

Akio NODA

Integrated Human Sciences · Mathematics

**Abstract:** In order to see how predominant results the west group of trainers gained over the east group in the horse racing data [3] of JRA for the period from 1998 to 2005, the author aims at a quantitative study of explored data analysis, and follows C.Hirotsu by making use of the statistical approach discussed in Chapter 6 of his book [1]. Indeed, based upon the data books [3], we make out a lot of  $2 \times 6$  contingency tables with ordered categories, and compute various statistics in Tables A, B and C of section 3 that play significant roles in the theory of accumulated chi-square test. These tables then enable us to extract some conclusions in the last section, concerning the theme mentioned in the title of this report.

**Key words:** contingency table with ordered categories, methods of assigning scores to categories, accumulated chi-square test, maximal component of the cumulative chi-squared statistics.

## 1. はじめに

日本の中央競馬において、「関西馬(栗東所属の調教師の管理馬)が関東馬(美浦所属の調教師の管理馬)に比べて優位に立つ」と言われて久しいけれど、どれほどの優位性があるのか、探索的データ解析を実行して、定量的に評価したい。これがわれわれの研究目的である。対象とするデータは、1998年度から2005年度までの8年間にわたる調教師リーディング(東西別になっているもの)の成績である([3]参照)。この間調教師数はほぼ一定で、東は111と115の間で、西は108と110の間で推移している。

掲示板にあがる1着から5着までと着外の順序づけられた6項目に、レース成績は分類される。まず東西2つの調教師群の全員に対して、各年度毎に2×6分割表を作成する。次に、リーディング順位(1着の度数による順位；同数なら2着の度数に、それも同数なら3着の度数による)にわれわれは注目して、1位から  $m$  位(ここでは、 $m = 20, 30, 40, 60, 80$  の5つの層を選ぶ)までの東西調教師群に対して集計し、2×6分割表を各年度毎に作成する。こうして得られた48通りの2×6分割表に、広津先生のすぐれた本([1], [2])に学んで、そこに展開されている統計解析の手法を適用する。こうして、東西両群の間にどれほどの有意差があるのか、調べた結果を以下簡潔に報告する。なお、[1]の第5, 6章は、浜松医科大学大学院の授業において、詳しく説明した統計解析であり、現実の競馬データに適用すればどんな結果を産み出すのか、その切れ味に興味を抱いたのがこの報告書の端緒をなしている。(受講してくれた大学院生達に感謝申し上げます。)

## 2. 統計解析

われわれが[3]に記載されているデータに基づいて作成した2×6分割表(順序づけられた6項目をもつ)に対する統計解析として、3通りの方法が[1]の第6章に述べられている。すなわち、(a)スコア法(b)累積 $\chi^2$ 法(c)最大累積 $\chi^2$ 法である。(a)について、通常の(1)平均順位をスコアとするWilcoxon法に加えて、われわれは競馬界特有の賞金配分法に従って、(2)1着に20、2着に8、3着に5、4着に3、5着に2、着外に0というスコアを割り振る方法を併用する。東の調教師を第1群、西を第2群として、第2群の総スコア  $W_2$  を算出する。それを標準化した統計量  $\{W_2 - E(W_2)\} / \sqrt{V(W_2)}$  ( $= W$  とおく)は標準正規分布に従うので、5%、1%、0.1%の限界値は、一様性の両側検定の場合、それぞれ1.960, 2.576, 3.291である。なお、 $W$  の値は小数第4位を四捨五入し、スコア法(1)では負の方向に、(2)では正の方向に「西優位」の度合いが強まることに留意する。これらの数値は、3.結果の表Aに記される。

$W^2$  が自由度  $\nu = 1$  の  $\chi^2$  分布に従うのに対比して、(b)累積 $\chi^2$ 法は  $\nu > 1$  という小数自由度をもつ  $\chi^2$  分布(特殊なガンマ分布)によって確率評価されるのが特徴である。すなわち、 $\chi^{*2}$  という統計量([1]第6章の記号を踏襲する)は、分割表の周辺度数が定める定数  $d$  で割算すれば、 $\nu = 5/d$  として  $\chi^2(\nu)$  分布で近似される。今の場合、 $1.7 < \nu < 1.9$  の範囲になる。従って、 $\chi^2(\nu)$  分布の5, 1,

0.1%点はそれぞれ、5.396 ( $v = 1.7$ ) 5.598 ( $v = 1.8$ ) 5.796 ( $v = 1.9$ ), 8.505 ( $v = 1.7$ ) 8.745 ( $v = 1.8$ ) 8.980 ( $v = 1.9$ ), 13.00 ( $v = 1.7$ ) 13.28 ( $v = 1.8$ ) 13.55 ( $v = 1.9$ ) を補間して求めることができる。48通りの2×6分割表に対し、 $\chi^{*2}$ の値と ( $d, v$ ) の値は、次節の表Bで示される。

以上(a)と(b)が、東西2群の成績比較という課題にわれわれが実践した統計解析であるが、もう1つ補助的に(c)最大累積  $\chi^2$ 法を取りあげる。その際馬券に対応して、4着以降を1つにして、1着、2着、3着とそれ以外の4項目に短縮する。このとき、[1]の付表を参照すると、統計量max acc.  $\chi^2$ の5%と1%の限界値は、それぞれ5.500と8.480となっている。max acc.  $\chi^2$ の値は最大値が生じる場所  $j$  ( $j = 1, 2, 3$  のいずれか)とともに表Cで与えられる。

### 3. 結果

調教師リーディング1位から  $m$  位までの6つの層 ( $m = 20, 30, 40, 60, 80$  と全員を表す  $\infty$ ) は行にとり、1998年度から2005年度までの8つの年度は列にとって、算出された統計量の値を、6×8分割表の形の表A, B, Cにまとめる。すなわち、表Aは(a)スコア法による場合で、上段に(1)平均順位をスコアとして、下段に(2)20, 8, 5, 3, 2, 0のスコアを与えたときの標準化された統計量  $W$  の値を示す。表Bは(b)累積  $\chi^2$ 法による場合で、 $\chi^{*2}$ の値と  $d, v$ の値を上から順に記す。表Cは(c)最大累積  $\chi^2$ 法によって、max acc.  $\chi^2$ の値と最大値が生じる場所  $j$  を示す。

|          | 1998   | 1999   | 2000   | 2001    | 2002   | 2003    | 2004   | 2005   |
|----------|--------|--------|--------|---------|--------|---------|--------|--------|
| 20       | 2.302  | -0.652 | -4.688 | -2.724  | -2.295 | 0.629   | -0.869 | -2.472 |
|          | -1.120 | -0.174 | 3.447  | 2.399   | 1.574  | -0.908  | 1.767  | 2.544  |
| 30       | 1.011  | -2.125 | -4.241 | -3.456  | -3.566 | -1.896  | -3.937 | -0.713 |
|          | 0.303  | 0.624  | 2.997  | 2.887   | 2.545  | 1.369   | 3.968  | 1.392  |
| 40       | 0.229  | -2.675 | -4.432 | -4.385  | -3.243 | -4.115  | -3.701 | -0.734 |
|          | 0.117  | 1.452  | 3.006  | 3.452   | 2.565  | 3.476   | 3.533  | 0.743  |
| 60       | -1.752 | -3.879 | -4.628 | -6.545  | -3.623 | -5.794  | -4.876 | -2.403 |
|          | 1.737  | 2.202  | 3.452  | 5.703   | 3.189  | 4.829   | 4.511  | 2.231  |
| 80       | -2.727 | -5.311 | -5.653 | -7.884  | -6.109 | -7.633  | -6.459 | -4.957 |
|          | 2.559  | 3.434  | 4.411  | 6.676   | 5.159  | 6.035   | 6.070  | 4.251  |
| $\infty$ | -5.280 | -7.725 | -7.744 | -10.365 | -8.071 | -10.754 | -9.886 | -8.243 |
|          | 4.612  | 5.758  | 6.742  | 8.708   | 6.410  | 8.305   | 8.870  | 7.058  |

表A  $W$  の値

|          | 1998   | 1999    | 2000    | 2001    | 2002    | 2003    | 2004    | 2005    |
|----------|--------|---------|---------|---------|---------|---------|---------|---------|
| 20       | 22.084 | 2.984   | 72.293  | 26.322  | 17.634  | 3.428   | 8.266   | 21.243  |
|          | 2.753  | 2.775   | 2.778   | 2.783   | 2.813   | 2.782   | 2.795   | 2.769   |
|          | 1.816  | 1.802   | 1.794   | 1.797   | 1.777   | 1.797   | 1.790   | 1.805   |
| 30       | 10.334 | 18.353  | 55.866  | 38.509  | 40.537  | 11.326  | 63.367  | 5.156   |
|          | 2.749  | 2.749   | 2.761   | 2.764   | 2.785   | 2.766   | 2.793   | 2.776   |
|          | 1.819  | 1.819   | 1.811   | 1.809   | 1.795   | 1.807   | 1.790   | 1.801   |
| 40       | 6.297  | 23.849  | 62.640  | 57.764  | 33.316  | 55.234  | 51.348  | 1.685   |
|          | 2.740  | 2.745   | 2.743   | 2.759   | 2.770   | 2.762   | 2.775   | 2.768   |
|          | 1.825  | 1.822   | 1.823   | 1.812   | 1.805   | 1.810   | 1.802   | 1.806   |
| 60       | 10.658 | 45.170  | 66.170  | 145.603 | 42.003  | 106.753 | 86.569  | 20.717  |
|          | 2.732  | 2.734   | 2.719   | 2.752   | 2.762   | 2.756   | 2.740   | 2.753   |
|          | 1.830  | 1.829   | 1.839   | 1.817   | 1.810   | 1.815   | 1.825   | 1.816   |
| 80       | 23.490 | 86.306  | 97.333  | 208.041 | 114.715 | 180.305 | 152.726 | 83.145  |
|          | 2.748  | 2.718   | 2.714   | 2.739   | 2.746   | 2.744   | 2.726   | 2.740   |
|          | 1.820  | 1.840   | 1.843   | 1.826   | 1.821   | 1.822   | 1.834   | 1.825   |
| $\infty$ | 86.157 | 187.335 | 193.567 | 353.260 | 214.809 | 343.629 | 350.236 | 231.270 |
|          | 2.835  | 2.692   | 2.699   | 2.714   | 2.722   | 2.723   | 2.715   | 2.716   |
|          | 1.764  | 1.858   | 1.853   | 1.842   | 1.837   | 1.836   | 1.842   | 1.841   |

表B  $\chi^{*2}$ の値と  $d, v$  の値

|          | 1998   | 1999   | 2000   | 2001   | 2002   | 2003   | 2004   | 2005   |
|----------|--------|--------|--------|--------|--------|--------|--------|--------|
| 20       | 7.177  | 0.792  | 17.098 | 6.101  | 1.586  | 1.629  | 3.584  | 5.376  |
|          | (3)    | (1)    | (3)    | (3)    | (3)    | (2)    | (1)    | (1)    |
| 30       | 3.197  | 5.011  | 10.309 | 7.569  | 6.277  | 2.262  | 16.293 | 2.655  |
|          | (1)    | (3)    | (3)    | (3)    | (3)    | (3)    | (3)    | (1)    |
| 40       | 2.856  | 5.540  | 10.072 | 11.328 | 4.674  | 12.054 | 10.875 | 1.140  |
|          | (2)    | (3)    | (3)    | (3)    | (3)    | (3)    | (3)    | (1)    |
| 60       | 4.351  | 8.666  | 9.439  | 30.595 | 7.648  | 22.795 | 16.481 | 4.276  |
|          | (1)    | (3)    | (3)    | (3)    | (3)    | (3)    | (2)    | (2)    |
| 80       | 6.364  | 18.059 | 16.067 | 45.427 | 20.694 | 37.371 | 32.771 | 15.547 |
|          | (1)    | (3)    | (3)    | (3)    | (3)    | (3)    | (3)    | (2)    |
| $\infty$ | 14.555 | 41.924 | 38.149 | 77.208 | 42.023 | 63.443 | 76.134 | 43.881 |
|          | (3)    | (3)    | (3)    | (3)    | (3)    | (3)    | (3)    | (3)    |

表C max acc.  $\chi^2$ の値(およびj の値)

#### 4. まとめ

表A, B, Cがわれわれに教える特徴の第一は、東西の調教師全員 ( $m = \infty$ ) の成績を比較すれば、すべての年度にわたって0.1%よりもはるかに小さいP値を得ること、つまり西優位の著明な有意差が認められることである。 $m = 80$ に移っても、 $m = \infty$ における数値の動きに連動して、西優位を表す有意差がすべての年度にわたって存在する。しかしながら、 $m = 60$ になると、1999年から2004年の間は大きな有意差を示しているが、左端の1998年では差は有意でなくなり ( $m = 20$ での東優位が影響)、右端の2005年では1%と5%の間のP値となって、有意差は縮小する(後述するように、リーディング21位から40位までの東の調教師の活躍に負う)。

表から導かれる第二の主張は、 $m \geq 40$ の範囲では各年度、 $m$ の増加とともに西の優位性が単調に強まって行くことである。すなわち西の調教師達は、リーディングの中間層は言うに及ばず、81位以降の下位層も東に比べてはるかに充実し、粒揃いの成績を毎年コンスタントに残している状況が明瞭となる。他方、 $m \leq 40$ の3つの行をみれば、われわれの統計量は  $m \geq 40$  の場合と比べてより複雑な変化をみせる。これは、1着の度数に基づいて並べたリーディング効果(統計量は1着から5着までの各度数だけでなく、着外の度数にも敏感に反応する)によるのに加えて、次に論じる上位層の東西比較の結果から示唆されるように、年度毎の偶然変動が働くためと考えられる。

リーディング上位層の東西比較 ( $m = 20, 30, 40$  による層別分析)を行って、この報告書を終える。1998年の欄からは、東の優位を表す5%未満の有意差から始まり、 $m$ が増加するにつれて西優位が徐々に強まる様子がみてとれる。次の1999年では、有意差のない  $m = 20$  の状況から、 $m = 30, 40$ に移ると5%未満の西優位の有意差が複数の統計量で観察される。2000年から2002年までの3年間は、上位層においても西優位が際立つ期間として特筆される。さて、2003年では  $m = 20$  ではほんのわずかの東優位から始まるが、 $m = 30$ で逆転し、 $m = 40$ になると著しい西優位に至る。次の2004年では、 $m = 30$ で有意差がくっきりと現れる。そして最後の2005年では、 $m = 20$ で5%未満の有意差が認められるのに反して、 $m = 30$ と40では西優位を表す差は有意でなくなってしまう、 $m = 60$ になってやっと  $m = 20$ と同程度の有意差を回復する(21位から40位までの東の調教師達の成績がよいためである)。こうして、上位層だけに限って東西の成績を比較すれば、年度毎の偶然変動(運・不運の要素)を認めざるを得ない。西の優勢は否定し難いものの、東の力も(上位層に限れば)侮るわけにはゆかぬような年度推移にわれわれは直面する。もちろん、中間層、下位層も含めてながめれば、西優位は歴然としており、今後もこの傾向が続くものと予想される。

#### 謝辞

資料の整理と原稿の清書をお願いした鴨藤江利子さんに、心から御礼申し上げます。また、レフェリーから寄せられた示唆に富むコメントを糧として、より深い分析へと進むつもりであります。

## 参考文献

- [1] 広津千尋：医学・薬学データの統計解析，東京大学出版会，2004.
- [2] 広津千尋：離散データ解析，教育出版，1982.
- [3] レーシングファイル(中央競馬全レース成績書)，No.22～50，ケイバブック，1999～2006.