

ツールドフランスにおけるタイムの統計的分析

2004MM066 坂井要介

指導教員: 松田眞一

1 はじめに

私は、ツールドフランスで偉業を達成した元ロードバイク選手のランス・アームストロングが好きで、ツールドフランスに興味を持った。しかし、私はツールドフランスについてその選手以外何も知識がなかったので、ツールドフランスについて詳しく調べることにした。

1.1 ツールドフランスのルールについて

ツールドフランスは、自転車でのロードレース競技大会のひとつである。第1回は1903年に始まった。今回は2006年のデータを用いることにしたので103回目になる。毎年6月ごろに開催され、フランス、周辺国を含む3000kmを3週間以上で走り抜くステージレースである。参加選手のレベルの高さから、世界最大の自転車レースといわれている。そして、2006年の参加チーム数は20チーム(176人)ほとんどのレースがチーム戦であること、コースには平坦コース、山岳コースがあること、集団でゴールすると、集団でゴールした選手のタイムがすべて同タイムになることなどの知識を得ることができた。(web[1]参照)

2 データについて

ツールドフランスのレースのデータは、web[2]を参照した。データは各選手のレースごとのタイムを使用した。タイムは、時間、分、秒で表されているので、すべて秒にしてから用いることにした。また、同タイムの多いコース、コース1、2、3、5、6、9、14は除いて解析した。

3 主成分分析

チームの上位3人合計60人で解析した。データを基準化して解析した。チーム3人にした理由は、チーム総合成績はチームの上位3人のタイムを使用するためである。

3.1 主成分分析解析結果

第1主成分 「総合軸」寄与率0.34

第1主成分は、ほぼ全部マイナスになっていることから、レースごとの総合値を表していることが分かる。このことから、マイナスにいくほど総合力に優れていることがわかる。

第2主成分 「チーム力軸」寄与率0.15

コース4、7、10、11、13、15、19がプラスに働いている。コース8、12、16、17、18、20、21がマイナスに働いている。コース8、20がマイナスに働いていることに注目してほしい。この2つのコースは、タイムトライアルである。自力が必要であることがわかる。また、マイナスに働いているコースは、選手ごとのタイムの散らばりがある。コース12、16、17、18に関しては、他の山岳コースのコースの形状に関しては他との違いはないが、速度が遅いことがわかる。このことから、これらのレースは、スローペースでレースが行われ、最後にスパートしていることから

表 1: 主成分分析の結果

コース名	第1主成分	第2主成分	第3主成分
コース4	-0.170	0.267	0.258
コース7	-0.248	0.475	0.248
コース8	-0.326	-0.127	0.317
コース10	-0.246	0.413	0.299
コース11	-0.318	0.260	-0.274
コース12	-0.358	-0.133	-0.273
コース13	-0.118	0.088	-0.494
コース15	-0.370	0.177	-0.284
コース16	-0.318	-0.181	-0.246
コース17	-0.335	-0.273	0.119
コース18	-0.277	-0.339	-0.189
コース19	0.031	0.227	-0.129
コース20	-0.267	-0.295	0.328
コース21	-0.033	-0.172	0.088

タイムに散らばりができ、自力が必要であることがわかる。コース21は最後のレースなので、チームのことを考えず、自力で走るためマイナスに働いていることがわかる。このことから、プラスに働いているコースと、マイナスに働いているコースで、チーム力、自力にわけることができる。

第3主成分 「コースの形状軸」寄与率0.1

第3主成分は、コースの形状に関する成分である。共通点を考えてみると、コースが平坦であればあるほどプラスに働いていて、起伏が激しいほどマイナスに働いていることがわかった。さらに、平坦コースでも左右に蛇行しているコースは、プラスの値が小さいことがわかる。また、起伏が激しくても、左右にあまり蛇行していないコースは、マイナスの値が小さいことがわかった。結論としては、第3主成分は、平坦で、左右に蛇行していないコースになればなるほどプラスに働いていて、起伏があり、左右に蛇行しているほどマイナスに働いていることがわかる。

3.2 主成分分析のまとめ

今回の主成分分析で、コースに関して細かく区別をつけることができた。初めは、平坦コースは自力が必要なく、山岳コース、タイムトライアルでは自力が必要であると考えていた。しかし、平坦コース、山岳コース関係なく自力が必要であることがわかった。また、山岳コースでもチーム力が必要であることがわかった。

4 クラスタ分析(チーム)

チームの上位3人を抜き出して、合計60人でクラスタ分析を行った。解析には標準化ユークリッド距離、ワード法を用いた。

4.1 クラスタ分析結果

第1群 「非常に優秀群」

第1群では、平坦コース、山岳コースに関係なく非常に優秀な群であることが分かる。全部で14レースあるが12レースで平均タイムが一番になっている。

第2群 「スタミナを温存する山岳コース優秀群」

第2群では、後半の山岳コースに関しては、第1群の非常に優秀な群の次に優秀な成績をあげていることがわかる。しかし、前半の平坦コースに関しては、成績が悪いことがわかる。さらに、コース21の平坦コースでも優秀な成績をおさめていることがわかる。

第3群 「スタミナを温存しない平坦コース優秀群」

第3群では、前半の平坦コースに関しては、第1群の次に優秀な成績をあげていることがわかる。しかし、11レース以降の山岳コースに関しては、成績が悪いことがわかる。

第4群 「優秀群」

第4群では、第1群と比べて全体的に劣っていることが分かる。しかし、14レースを通じて平均が最下位になっていることが2回しかないことがわかる。

4.2 クラスタ分析(チーム)のまとめ

上位にきているチームは、1群の非常に優秀群の選手がいて、2群の山岳コース優秀群の選手がいるチームが上位を占めていることがわかる。次に、2群の山岳コース優秀群の選手がいるチームが占めている。次に、2群の山岳コース優秀群と、4群の優秀群の選手がいるチームが続いている。次に4群の優秀群のみの選手のチームが続いて、最後に3群の平坦コース優秀群の選手がいるチームが続いている。また、チーム総合タイムを重視しているチーム、個人の総合タイムを重視しているチームの2種類が存在することがわかった。

このように、チームの作戦や、どのような選手を獲得しているかがわかった。また、どのような選手を獲得しているチームが上位にきているかがわかった。

5 クラスタ分析(個人)

チームの上位5チームを抜き出して、合計41人で解析を行う。しかし、途中棄権の選手が7人いるので、34人でクラスタ分析を行った。解析には標準化ユークリッド距離、ワード法を用いた。

5.1 クラスタ分析結果

第1群 「あまり貢献できていない群」

第1群では、ほとんどのレースで平均が最下位になっていることがわかる。

第2群 「非常に優秀群」

第2群では、前半の平坦コース、後半の山岳コース、タイムトライアルすべてのレースでとても優秀な成績をおさめていることがわかる。

第3群 「優秀群」

第3群では、全体的に優秀な成績をおさめていることがわかる。しかし、第2群よりは優秀な成績をおさめていない。

さらにタイムトライアルであるコース8、20では優秀な成績をおさめられていない。

第4群 「タイムトライアル優秀群」

第4群では、全体的に見るといい成績をおさめていない。しかし、タイムトライアルに注目してほしい。コース8、20では、平均タイムが1位、2位になっている。

5.2 クラスタ分析(個人)のまとめ

上位5チームを解析した結果、チーム総合1位のチームは1群の選手がおらず、選手の役割がはっきりしていることがわかる。また、チームのタイムトライアル専門選手が多くいるチームと、少ないチームがあることがわかった。必ず山岳コースの控えの選手がいることがわかった。さらに、限ったレースで、いい成績をおさめている選手がいることがわかった。全部で20チームあるが、そのほとんどが9人のチームである。しかし、上位5チーム中で、3チームが8人以下の選手構成になっている。このことから、上位にきているチームは、他の下位のチームより、平坦コースも山岳コースもこなせる人がいるので少なくとも勝てることがいえる。

6 因子分析

データは主成分分析の時同様、チーム上位3人合計60人で解析した。

第1因子「山岳コース因子」第2因子「平坦コース因子」第3因子「自力因子」の3つの因子が抽出された。

6.1 因子分析のまとめ

個人成績上位5人の因子得点を見た結果、どの因子もマイナスに働いている因子が多くあることがわかる。総合2位になっている、Pereiro選手は平坦コースの第2因子でプラスに働いている。第3因子も同様である。しかし、他の4選手はすべてマイナスに働いていることから、どの因子でも得意といえるものがないことがいえる。このことから、得意な因子があるわけではなく、目立って活躍している因子がないが全体的に活躍できる選手が、上位にこられることがわかる。

7 おわりに

それぞれの分析で、いい結果が出たのは、ルール、コースについて調べることができたからである。しかし、主成分分析に関しては、累積寄与率が60%までしか解析できなかった。また、2007年のデータを用いて、2006年とのタイムの比較をするつもりだったが、データがなく、できなかった。今後は、まず主成分分析、因子分析をもっと詳しく解析する課題が残っている。さらに、2007年のデータを取得して、タイムを用いて比較していきたい。

参考文献

- [1] ツールドフランスジャージ&ルール:
home.netyou.jp/gg/tsururin/t04rule.html/.
- [2] Bicycling:TourdeFrance2006:
www.bicycling.com/tourdefrance/afp/tdf2006/flash/en/index.html/.