

因子分析を用いたサッカープレーデータによる疲労に関わる客観的評価

Objective evaluation of fatigue based on football play data using factor analysis

1W173071-7 助川 京輔
Keisuke SUKEGAWA

指導教員 河合 隆史 教授
Prof. Takashi KAWAI

概要：スポーツにおける「疲労」は、パフォーマンスにも多大な影響を及ぼすことから理解を深めることが重要である。しかしながら疲労というものは、“体が重い”や“動きが悪い”などとしばしば主観的で曖昧な表現をされる。このような表現では説得力に欠ける。そこで本研究では、数値化されないような「構成概念」を数値で表現することのできる因子分析を利用して、この「疲労」を数値で表現することを目指した。実際のプロサッカーの試合での選手プレーデータをもとに因子分析を行なった結果、「疲労」と命名しても問題のないような因子を抽出することができ、この因子が各観測変数（プレーデータ）に与えている影響を数値化することができた。さらにその影響の値（因子負荷）から因子命名の裏付けを行うこともできた。

キーワード：疲労, 客観的, 数値化, 因子分析

Keywords : fatigue, objective, quantification, factor analysis

1. 本研究の背景・目的

どのスポーツに関しても「疲労」と向き合うのは重要なことである。しかしながらこの「疲労」というものは主観的なものであり、何らかの数値で表現されるものではない。そこで本研究では、この「疲労」を試合中のプレーデータをベースに数値で客観的に表現することを目指す。

2. 本研究で用いる解析手法

本研究では、因子分析を解析手法として用いる。因子分析を用いる理由は、因子分析が具体的に測定できない、いわゆる「構成概念」を数値化できる手法であり、今回のテーマ「疲労」もこの構成概念にあたるものであると考えたためである。

3. 本研究の解析手順

本研究の解析手順は以下となる。

- (1) 使用データの選択
- (2) データ整理
- (3) KMO 測度による標本評価
- (4) 因子分析
- (5) 分析結果の評価

なお、解析ソフトとして R-Studio

(Version1.2.1335) および因子分析用ライブラリ

“psych”を使用する。

4. 解析結果

4.1. 使用データの選択

本研究では、データにある各選手の実出場時間 5 節

分の合計が基本的な試合時間 5 試合分を超えた GK を除く 78 選手を対象とした。この中から必要と考えられるプレーデータを抽出した（表 4.1.参照）。

また集約条件変数用に F_成功・縦位置・試合状態 ID・ハーフ開始相対時間を、ラベル付の用途で、試合 ID・選手 ID・チーム ID・背番号・ポジション ID も抽出した。

4.2. データ整理

はじめに、抽出された各プレーの試合状態 ID とハーフ開始相対時間をもとに新たな変数「timezone」を以下のように定義し追加する。

- 試合状態 ID=1 かつハーフ開始相対時間 2500 未満：timezone=1
 - 試合状態 ID=1 かつハーフ開始相対時間 2500 以上：timezone=2
 - 試合状態 ID=2 かつハーフ開始相対時間 2500 未満：timezone=3
 - 試合状態 ID=2 かつハーフ開始相対時間 2500 以上：timezone=4
- 以降「試合状態 ID」と「ハーフ開始相対時間」は「timezone」に集約された形となる。
- つづいて以下の 4 つの集約条件を設ける。

- 条件 1：ラベル付変数ごと
- 条件 2：timezone ごと
- 条件 3：縦位置ごと
- 条件 4：F_成功ごと

以上4つの集約条件を、ひとつあるいは複数用いて数種類のデータを作成する。

4.3. KMO 測度による標本評価

KMO 測度 (Kaiser-Meyer-Olkin の標本妥当性の測度) とは、データが因子分析に適切であるかどうかを判断するための基準のひとつである。4.2.で作成した数種類のデータに適応した結果、以下のデータを採用することとした。

<指標クリア条件とスコア>

- 条件 1+2+4 スコア: 0.64

4.4. 因子分析

本研究での因子分析では、初期解の推定法は最尤法、回転法は斜行プロマックス法を採用している..

4.4.1. 因子数決定

ここで、抽出する潜在因子数 (以降因子数) を設定する。今回は因子数決定の手法からスクリーテスト・ガットマン基準・VSS を併用して因子数を決定した結果、因子数2での因子分析を適応する。

4.4.2. 量的因子分析結果

表 4.1. 因子負荷行列

観測変数	因子 1	因子 2
F_ドリブル	0.02	0.16
F_パス	1.07	0.25
F_クロス	0.08	0.52
F_スルーパス	0.10	0.26
F_クリア	-0.21	-0.08
F_インターセプト	0.17	-0.11
F_ファウル	-0.19	0.14
F_被ファウル	-0.19	0.18
F_ボールロスト	-0.20	0.61
F_ボールゲイン	0.53	-0.21
F_ボールタッチ	1.05	0.19

表 4.1.は因子負荷行列(各観測変数とその因子にどれほど影響を受けているかを示す「因子負荷」を行列にしたもの)を示す。

表 4.2. 累積寄与率

	因子 1	因子 2
累積寄与率	0.76	1.00

表 4.2.に累積寄与率を示す。累積寄与率とは、その因子が全観測変数の分散をどれほど説明しているかの割合を累積したものである。

5. 考察

5.4.2 で示した解析結果のみでは疲労に関する因子を判断するのは難しい。そこで、集約条件変数 timezone ごとに各データを散布図にしたもので考察する。図 5.1.では最も疲労の度合いが変わるであろう timezone=1 のデータ群と timezone=4 のデータ群をプロットする。

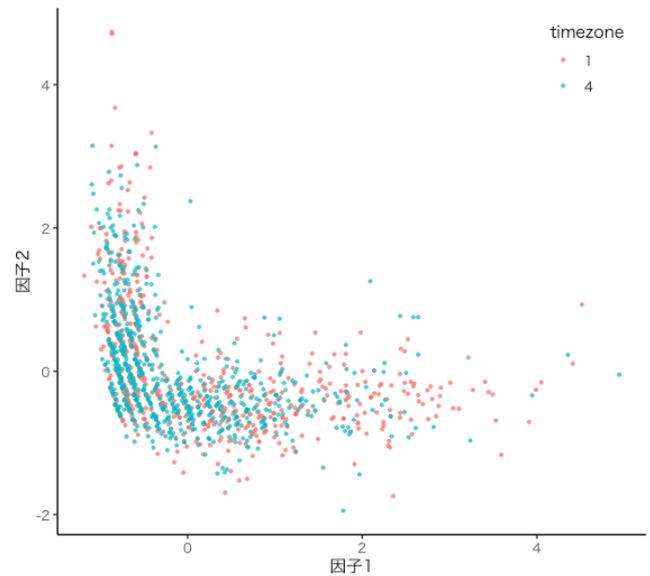


図 5.1. timezone ごとの散布図プロット

ここで timezone=1 のデータと timezone=4 のデータ間で2標本 t 検定 (ウェルチの検定) を行うと、因子 2 に関しては p 値が 0.633 であるが、因子 1 に関しては p 値が 2.203×10^{-5} であり違いが見られる。したがって因子 1 を“疲労”の因子と考える。ここで表 4.1.を見ると“F_クリア”や“F_ボールロスト”が負の値である。すなわち因子 1 が小さい (=疲労が大きい) ときにこの要素が強くなるわけである。これは主観的感覚でも試合終盤はボールロストなどのミスが増えたり、楽にボールをクリアしたりするというものとも一致している。したがって因子 1 を“疲労”と命名しても問題はないと考えられる。

6. 今後の展望

本研究では、因子分析を用いて「疲労」を抽出し、疲労が影響を与えている項目を数値で表現することができた。今後はより精度の高い疲労の因子抽出を目指し、これを応用して実際に「疲労値」にあたる予測値を算出することが最終目標となる。

7. 参考文献

豊田秀樹 (2012) 『因子分析入門』東京図書