

「特集 スポーツ統計科学の新たな挑戦」について

田村 義保¹・酒折 文武² (オーガナイザー)

ICT, GPS, センサー技術などの発達により, スポーツの世界においても選手やボールのトラッキングデータをはじめとして, これまで以上に大規模かつ複雑なデータが収集されるようになった. そして, 以前から活用されてきたボックススコアやスタッツ, あるいはプレイ・バイ・プレイの行動履歴データ, あるいは試合の動画等と合わせて, チーム編成や戦略評価, さらにリアルタイムな戦術判断へのさらなる活用が待たれている. 2019年のラグビーワールドカップ日本開催や2020年東京オリンピックなどビッグイベントを控え, 2014年には日本スポーツアナリスト協会が設立されるなど, その機運も高まっている.

欧米諸国では, 野球, サッカーやアメリカンフットボール, バスケットボールなどを始めとしたメジャーなスポーツにおいて先進的にデータ収集や高度な統計分析が行われ, 科学的な見地から戦術解析や選手評価がなされてきた. また, ASAでは1992年より毎回スポーツ統計のセッション(SIS, Statistics in Sports)が生まれ, 数々のシンポジウム開催やJQAS(Journal of Quantitative Analysis in Sports)とJSA(Journal of Sports Analytics)の発行, さらにJASA等の論文誌への論文投稿も多く見られるなど, 学術的な場においてもこうした研究は盛んに行われており, 統計学の専門家による積極的な関与が見られる.

一方, 日本では, スポーツの場における統計学の活用は必ずしも十分であるとはいえなかった. このような状況を鑑みて, 2009年から日本統計学会でスポーツ統計分科会の活動を開始した. さらに, 2011年からは統計数理研究所において共同研究を開始するとともに, 裾野の拡大とこの分野の研究推進を目指して「スポーツデータ解析コンペティション」を開催してきた. コンペティションでは, 日本プロ野球・アメリカメジャーリーグ・Jリーグ・Bリーグのデータについてデータスタジアム株式会社から提供を受け, 中央大学や立教大学の協力を得て, 大学院生・大学生を含むコンペ参加者へのデータ提供, 発表の審査, 優秀者発表会を行っている. 成果の一部は統計数理研究所共同研究リポート「スポーツデータ解析における理論と事例に関する研究集会」として発行してきている.

このコンペティションは, 大学院生や大学生のデータ解析経験の場の提供という意味で, データサイエンス教育・データサイエンティスト育成の役割も担っている. スポーツを含むあらゆるビジネス領域におけるデータサイエンティストの必要性とその不足を受け, 2016年12月に文部科学省は「数理及びデータサイエンスに係る教育強化」の拠点校として, 北海道大学, 東京大学, 滋賀大学, 京都大学, 大阪大学, 九州大学の6校を選定している. また, 平成29年度「データ関連人材育成プログラム」の取組機関として2017年8月に東京医科歯科大学, 電気通信大学, 大阪大学, 早稲田大学を代表機関とする4組織を選定している. さらに, 2017年4月からは滋賀大学データサイエンス学部で, 大学としては日本で初めてのデータサイエンティスト教育が始まっている. 今後は, こうした機関におけるデータサイエンス教育とも協力し,

¹ 統計数理研究所: 〒190-8562 東京都立川市緑町10-3

² 中央大学 理工学部: 〒112-8551 東京都文京区春日1-13-27

コンペティションのさらなる活性化が期待される。

コンペティションを含む活動を通じて、スポーツ統計科学としての発展と、スポーツ業界やスポーツ現場への貢献や交流も進んできた。本特集は、こうしたスポーツ統計科学の現状や今後を広く知らせるために計画した。

本特集には、野球、バレーボール、サッカーと 3 つの競技に関わる計 9 編の論文が掲載されている。6 編が原著論文、3 編が研究ノートである。これらの論文の一部はスポーツデータ解析コンペティションの成果を論文としてまとめたものである。4 編は野球に関する論文であり、永田論文、酒折論文、中村論文はアメリカのメジャーリーグ (MLB) に関する分析、阿部論文は日本プロ野球に関する分析である。同じ野球ではあるが、永田論文と酒折論文は PITCHf/x と呼ばれる投球のトラッキングデータに基づくものであるのに対し、中村論文と阿部論文ではプレイ・バイ・プレイのデータを用いている。また、小中論文はバレーボールの試合結果からのレーティングに関する研究である。そして 4 編がサッカーに関する論文である。これらは J リーグのトラッキングデータやボールに関係するプレイ・バイ・プレイのデータを用いている。なお、これらのデータのうち、メジャーリーグの PITCHf/x データとバレーボールの試合結果のデータはインターネットを通じて自由に入手可能である。また、野球やサッカー、そしてそれ以外の競技においても、ある程度集約されたデータであれば多くのものは入手可能である。スポーツに興味のある統計学研究者のかたがたのさらなる参入も期待したいと思っている。

最後に、この特集「スポーツ統計科学の新たな挑戦」の査読者の方々、並びに編集担当の方々、本特集でもデータ活用を許諾いただけたデータスタジアム株式会社に、この場をお借りして感謝を申し上げたい。