特集にあたって

生田目　崇（専修大学）

今年も，本学会の先端マーケティング分析研究部会もメンバーである経営科学系研究部会連合協議会の主催する「データ解析コンペティション」の特集を組ませていただいた．

例年どおり査読委員会を立ち上げ，7月下旬に査読付き論文の投稿を締め切った．9篇の投稿があり，まずは順調に査読プロセスをスタートさせた．第1回目の査読で9篇のうち5篇が返戻と判定され，4篇が再査読となった．査読レポートとともに著者に再投稿を促し，10月中旬に再投稿を締め切った．4篇のうち1篇の著者より締め切り直前になって，再投稿したかったが事情があり再投稿できずという連絡があり再投稿は3篇となった．

実は，第1回目の査読結果で残り4篇になった時点で，すでに標準の6篇の記事に足りないことが判明したため，今回投稿されたのは，課題フリー部門のみであったこともあり，投稿のなかったコンペティションの課題設定部門の最優秀チームに寄稿の打診をしており，了承されていた．本誌における特集記事は6ページが基本であるが，ここ数年特集を組んでいて，特集論文は平均して6ページを超えることが経験的にわかっていたため，これで再投稿のうち2篇が採録されればページ数的には大丈夫と考えていた．

ただし，さらに修正が必要と判断されることも考えて，再投稿論文の査読結果をできればなるべく返していただきたいと，査読者には無理にお願いし，早期に査読結果を返していただいた．しかしそのすべてが「再査読」もしくは「返戻」として報告されたのである．本号2月号の入稿の締切は11月下旬であり，修正論文の再々投稿を促しても，再査読結果は間に合わない事態となった．そこで，さらに急きょ，課題フリー部門の最優秀チームにも寄稿をしていただくようお願いし，快く引き受けていただいた．

さて，問題はここまで査読の進んだ論文である．今のままでは，査読付き論文としては掲載できないため特集号として成り立たず，いまさら再々投稿してもらう時間もない．私が本誌の編集委員となった初期に編集委員長だった田口東先生が「白紙の機関誌を出す夢を見る」と言われていたが，まさにその心境であった．苦慮した結果，今回は再査読まで進んだ3編については通常の寄稿として扱わせていただき，いきさつと査読のポイントをまとめたコメンタリーを付けて掲載するという案をまとめた．本件については，再査読結果を返送する前に，査読者各位，編集委員長の松井先生らに相談し了承された．そのうえで，論文の著者には，正式な査読つき論文としては掲載できないが，上記事情があるので査読での指摘事項にできるだけ対応していただいたうえで，ぜひ寄稿として記事掲載をお願いしたいという連絡をしたところ，了承をいただけ，私個人としては内心ほっとした．

この結果，本特集はコンペティションの最優秀チームからの寄稿2篇，投稿論文のうち，再査読まで進んだ3篇，査読のコメントなどをまとめた記事1篇の合計6篇で構成できることになった．

このように，本号の成立にはいろいろな方々のご協力をいただいた．まずは何より，今回の寄稿者の方々には無理なお願いを受け入れていただいたことを感謝する．特に松本氏，西郷氏（課題設定部門最優秀チーム），鮭川氏ら（課題フリー部門最優秀チーム）には急きょの寄稿を快諾いただいた．査読にあたられたみなさまにも感謝申し上げる．査読プロセスを設けたにもかかわらず査読論文が掲載できなかったのは残念ではあるが，査読を厳格にしていただいている証左でもある．本学会でのマーケティング分野の研究論文が一定の質を保つためにはこうした査読者による真摯な査読は欠かせない．また，こうした特集を毎年組ませていただいている編集委員会にも御礼申し上げたい．本年度もコンペティションを開催しており，今後の研究の発展のためにも今回に懲りず，関係者には引き続きご協力を賜りたい．

さて，平成23年度のデータ解析コンペティションであるが，（株）ゴルフダイジェスト・オンライン（GDO）社にご協力いただき，同社のECサイトのアクセス・ログ，購買・予約データをご提供いただいた．また，従来どおり参加各チームが分析目的から設定する形式（課題フリー部門）だけでなく，第1回のコンペティション

以来，大変久々に事前に分析目標を定めた「課題設定部門」を設けた．また，課題設定部門についてはGDO社より優秀チームへ表彰をいただいた．参加チーム数は69チーム，延べ300名を超す参加をいただき過去最大規模となった．

データの概要は以下のとおりである．

＜課題フリー部門＞

提供データは，2010年7月1日～2011年6月末のおよそ1年間のアクセス・ログおよびECサイトでの受注データおよび会員属性である．今回のアクセス・ログはサイト登録者のログに限って抽出されている．アクセス・ログは日時やページ属性のほか，セッションID，検索キーワード，商品分類などが含まれ，また，コンバージョン（購買）が発生すれば，フラグとして与えられている．会員属性には，年齢や性別のほか，GDOで独自に与えているハンディキャップやメールマガジンへの登録の有無も含まれている．

＜課題設定部門＞

今回の分析課題は，顧客のセグメンテーションであり，入会後90日間のデータから，入会後翌年6月末までの累積の「来訪回数」「購買額」の2つの項目に関するセグメントを充てるというものである．具体的には，来訪回数（F: Frequency）については，F1：4回以下，F2：24回以下，F3：25回以上の3セグメント，購買額（M: Monetary Value）についてはM1：年間0円，M2：年間8,000円未満，M3：年間8,000円以上の3セグメントで，各訪問者が$3 \times 3 = 9$のセグメントのどこに入るかを予測する．

なお，図1のように優良度に応じたウェイト付けがなされ，ウェイト付けされた正解数の合計によって評価がされる．なお，予測が外れた場合におけるペナルティは課していない．

データ項目は，セッションごとの閲覧サマリにゴルフ場の予約とゴルフ関連用品の購買受注データ，および会員情報である．アクセス・ログは上記のとおりセッション（一回の閲覧）ごとに，集計されたデータであり，そのセッションでのページビュー数とともに，カテゴリごとのページビューおよび，コンバージョン（予約と購買）があればフラグが立てられている．予約データは予約日とプレー日，予約人数と金額が与えられる．受注については受注日と商品属性，受注数量・金額，ポイントの付与と利用についてのデータが含まれる．

なお，データはモデル構築用の学習用データと予測用の検証用データに分かれている． 学習用データは2010年7月に登録した会員が対象で，登録後1年間のデータが提供されている．つまり，1年後のセグメントのわかっているデータである．これに対して，検証用データは2010年8月の登録会員であり，検証用データについては，登録後3カ月のデータのみが提供され，各チームから予測スコアを提出していただき，そのスコアを求めた．図1は各セグメントとウェイトであり，優良顧客ほどウェイトを高くしている．

F3	M1F3 ×1.2	M2F2 ×2.0	M3F3 ×2.5
F2	M1F2 ×1.0	M2F2 ×1.8	M3F2 ×2.0
F1	M1F1 ×1.0	M2F1 ×1.5	M3F1 ×1.5
	M1	M2	M3

図1　顧客セグメントとウェイト

F3	3,664	1,724	725
F2	1,559	1,985	1,791
F1	1,121	1,039	2,955
	M1	M2	M3

図2　検証用データのセグメント別人数

図2は，検証用データにおける各セグメントの人数である．小売業においてはFとMは高い相関をもつ場合も多いが，今回のデータではFはサイト訪問，Mは実際の購買であり相関はかなり低い．図1と照らし合わせると，ウェイトの高いセグメントのほうが相対的にあてはまる人数は少なく，優良顧客を重点に正解を求めるよりも，広く正解を求めるほうがスコアは高くなる．実際，最優秀チームのモデル構築においては「バランスよく」当てることを目指したという．なお，最優秀チームのスコア率（全サンプル正解に対するスコア）は70.4%であったが，2位とはわずか0.01%という激戦であった．

ここ数年，参加者が増えそれとともに参加者の得意とする研究分野の広がりと深化が著しい．それに伴って，研究成果も増えている．その反面，みんなの目が肥えてきており，査読論文が採択されにくくなっているのも事実である．今後とも，レベルの高い研究発表を通じて，ORにおけるマーケティング研究に一石を投じることができれば幸いである．