

統計調査の学習・実施のためのシステムの作成

岡山理科大学・総合情報学部 森 裕一
東海大学・理学部 山本 義郎
鹿児島大学・理学部 宿久 洋

mori@soci.ous.ac.jp, yoshiro@yamamoto.name, yado@sci.kagoshima-u.ac.jp

1. はじめに

統計解析を行うには、解析手法の理論的学習だけでなく、目的に応じた手法を選択し、結果を解釈できる力の育成が必要である。これに加え、解析ソフトウェアの適切な利用とさまざまな角度からの考察が可能となる解析環境の提供も重要である。

これらに対して、筆者らは、主として教育利用を前提として、解析の目的に適したデータが自由に選べ、実際の解析を記述した事例（解析ストーリー）をもとに学習できるシステム **DoSS@d**（データ指向統計解析環境, Data-oriented Statistical System）を Web 上に構築している (Mori et al., 2003; 山本 他, 2003)。本システムは、データバンクとしての検索システムとそのデータをオンラインで解析できるシステムを同時に備えており、各種の学習コースを提供するサイトとなっている。

このシステムはすでに稼動しているが、データを収集する機能や集めたデータをデータベースへ登録する機能、あるいは、収集中のデータをその場で試行錯誤的に解析する機能などについて、現在、追加を検討中である。

一方、たとえば、街頭での調査や異地点間のデータを即時に集めたい場合などを考えると、街中や構内などのあらゆる場所で集めたデータをその場でアップロードし、そのデータをリアルタイムに解析できる機能がほしいところである。これらに対して、無線 LAN、インターネットへ接続可能な携帯電話システムなどのインフラの整備と、ノート PC への無線 LAN 機能の標準搭載など、すでに、モバイル環境は整ってきている。

そこで、本研究では、上記のような **DoSS@d** への機能追加（データ収集機能）とモバイル環境におけるデータ解析機能の整備（データのアップロードとリアルタイム解析）の 2 つの要求を満たすため、**DoSS@d** にデータをアップすると同時に解析が行える環境を実装することにする（全体の概念図は図 1 を参照）。これにより、データベースと、データ収集・リアルタイム解析の両面で、強力なシステムができあがることになる。

2. データ指向統計解析環境 **DoSS@d**

DoSS@d は、次の 3 つのサブシステムからなる。

- **DoDStat@d**（データ指向統計データベース, Data oriented Database of Statistics）は、実データのデータベースであると同時に、解析ストーリーのデータベースである。各データは、関連情報（データ

名、出典、ケース名、変数名・変数型など）と各種フォーマットのデータ本体からなり、対象分野や解析手法などで検索・選択が可能である。また、解析ストーリーには、過去そのデータが解析されたプロセスが記述されており、必要に応じて、そのストーリーに基づいた解析がオンラインで行えるようになっている（実際には、次の **DoAStat@d** へのパラメータ付きリンクである）。

- **DoAStat@d**（データ指向統計解析システム, Data oriented Analysis system of Statistics）は、サーバ側の統計エンジン（R と XploRe Quantlet Server）を利用した解析環境を提供する。ユーザは、**DoDStat@d** に格納されたすべてのデータを解析できるし、ユーザ保有のデータも解析できる。
- **DoLStat@d**（データ指向統計学習システム, Data oriented Learning system of Statistics）は、解析ストーリーに基づいて、統計を学習できるようにした Web ベースのシステムである。具体的には、**DoDStat@d** の解析ストーリーを学習目標に合わせて教育的に配列したもので、個々の解析ストーリーの学習時には、**DoAStat@d** のオンライン解析が利用できる。

したがって、**DoSS@d** は、次のような利用場面を想定している。

1. 統計教育にかかわる人が、自分の授業等の教材準備のために、本システムのデータベース **DoDStat@d** から、適したデータをさがす。
2. 統計教育にかかわる人が、自分の授業の中で、**DoLStat@d** のコースを使ったり、オンライン解析システム **DoAStat@d** を学生・生徒に利用させたり、課題として、**DoDStat@d** のデータを解析させたりする。
3. 統計を学習している学生・生徒が、統計的な問題の自習をするために、自習教材として **DoDStat@d** のデータや解析ストーリーを利用したり、**DoLStat@d** のコースを学習したり、あるいはオンライン解析システム **DoAStat@d** を利用したりする。
4. 統計解析を行う人が、解析手法などの評価用データとして、**DoDStat@d** のデータを利用する。

3. モバイル環境におけるデータ収集・解析システム

3.1 想定する場面

データ収集の場面として、ここで特に想定しているのは、街頭でのアンケートの回答を随時入力してその時点での傾向を把握したり、交通量調査で時系列の変化を考察したりする場合、あるいは、植物の

植生調査など離れた地点で計測したデータを一挙に集めて解析しようという場合である。これらの場面において、これまでのように、調査を終えてから個別に集まったデータを一まとめにして解析に移るのではなく、その場でデータをアップして、集計結果や解析結果を考察できるようにしようというものである。これらは、すでに、POS データや在庫管理の専用システムで実現されてきたが、無線 LAN などのモバイル環境が整ってきた現在では、DoSS@d にリアルタイムデータ収集・解析機能をつけ加えることは十分意義あることと考える。また、入力されたデータは、2 で述べた DoSS@d の利用場面でも活用できるので、新しい DoSS@d は、モバイル環境をも反映した、より強力なデータ提供サイトとなり得る。

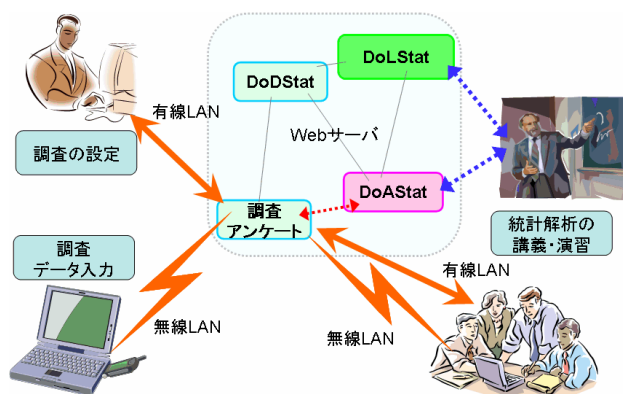


図 1 : DoSS@d の新サービス概念図

3.2 調査・アンケートデータ入力・解析システム

新たに実装する調査・アンケート作成・設定機能は、調査管理者が Web 上で、調査項目、回答方法（データの型、択一／複数選択、自由記述）などを設定して入力フォームを作成するもので、できあがった調査票により、Web 上でのデータ入力が可能となる。

これにより、調査担当者は無線 LAN が利用可能な場所で調査を実施し、得られたデータを直接入力したり、あるいは、調査結果を一まとめにして無線・有線 LAN で一括送信したりすることが可能になる（携帯電話による入力機能の実装も予定している）。

アップされたデータについては、DoAStat@d の解析機能を利用して、単純集計や基本的な解析ができるので、調査の途中で、データの変化を見たり、暫定的な解析結果を考慮しながら調査を継続したりできる。一方、終了した調査のデータは、DoDStat@d に登録されるので、DoSS@d の通常利用、すなわち、Web 上で DoAStat@d を用いて解析したり、講義でデータをダウンロードして、DoLStat@d の解析例を参考に、一般の統計解析パッケージ（SPSS や XploRe）を使って解析したりできる。

3.3 システムの実装

調査・アンケートデータ入力システムは、アンケートビルダー（株式会社ティス、http://www.tistis.com/tis/_enquetebuilder.html）を本システム用にカスタマイズしたもので、調査表作成、データ入力、要約解析の 3 つのモジュールをもつ（図

2）。各モジュールは、Web ブラウザを介して、PHP が制御を行う。入力されたデータは、MySQL により管理されるとともに、DoDStat@d への保存にあたっては、XML ファイルが生成される。解析は、DoSS@d の手続きに従い、DoAStat@d の解析エンジンが解析を行い、結果が利用者に返される。

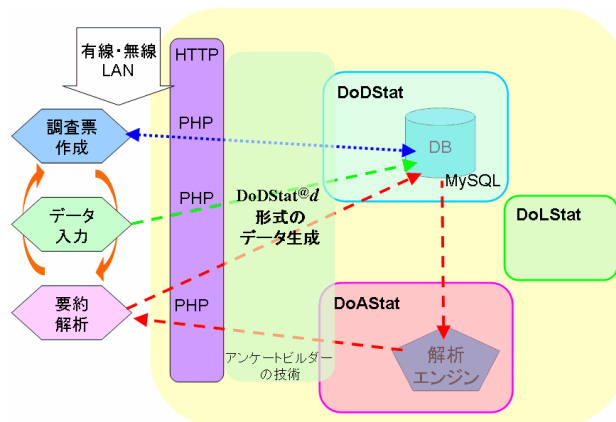


図 2 : データ入力・解析システムの構成

3.4 具体例

プロトタイプとして、授業評価アンケートを作成した（図 3）。学内の無線・有線 LAN 設備を利用して、どこからでも回答がアップできる。



図 3 : データ入力・解析システムの例（授業評価アンケート，上：入力フォーム，下：解析設定画面と集計結果）

註）本報告におけるモバイル統計の研究は、インテル(株) 協賛 CIEC 2004 年度プロジェクト事業として、補助を受けている。

参考文献

- 山本義郎・森裕一・宿久洋. (2003). 統計学習のためのオンラインシステムの作成. 2003 PC Conference 論文集, 51-52.
- Mori, Y., Yamamoto, Y., Yadohisa, H. (2003). Data oriented Learning System of Statistics based on Analysis Scenario/Story (DoLStat). Bulletin of the International Statistical Institute, 54th Session Invited Papers, Volume LX Two Books, Book 2, 74-77.