

インターネット環境における統計科学普及の ための教育用サイト構築の試み

渡辺 美智子¹・末永 勝征²・竹内 光悦³

宿久 洋⁴・山口 和範³・浅野 長一郎⁵

(受付 2001 年 4 月 11 日; 改訂 2001 年 9 月 26 日)

要 旨

コンピュータ・ネットワークが急速に普及し膨大な量のデータが電子的に蓄積されている現在、その有効利用と統計分析処理が課題になっている。その結果、学生はもとより一般社会人に対して、適切な統計分析のために必要とされる教育内容や技術指導の具体的な方法論の研究が必要とされている。一方、近年のインターネットの普及により、統計教育に関してもパーソナル・コンピュータ(PC)上でのマルチメディアや動的グラフの利用、各種のデータを提供する Web サイトへのリンクなどの諸機能を活用した新しい教材開発の道が開けている。

こうした現状を背景として、筆者等のグループは、統計教育および統計科学の普及を効果的に支援するためのツールとして、幾つかの教育用 Web サイトを構築した。具体的には、インターネット上でインタラクティブに統計基礎教育を行う「統計学習用インタラクティブ・テキスト:ITLS(Interactive Text for Learning Statistics)」, 検索機能を備える統計科学関連書籍についての「電子図書システム:EBSA(Electronic Books for Statistical Analysis)」, それに関連した「統計専用検索エンジン:Integrated Search Page」システム, そして表計算ソフトウェア Excel などのインターフェイスを介して呼び出し可能な DLL (Dynamic Link Library) や DLL を利用した分析マクロを Web 上で提供する「教育用統計分析プログラムの公開サイト:DLLSA (Dynamic Link Library for Statistical Analysis)」である。本稿では、これらの特徴と主な機能を紹介し、また、これらのサイトの統合化に関する可能性およびインターネットの利用環境を活用した統計教育のための今後の組織的な展開が、統計科学の普及と裾野の拡張に必要なことを論じた。

キーワード: オンライン・インタラクティブ・テキスト, 統計科学, 電子図書システム, 統計専用検索エンジン, 統計分析用 DLL, 教育用 Web サイト。

¹東洋大学 経済学部: 〒112-8606 東京都文京区白山 5-28-20

²純心女子短期大学: 〒890-8525 鹿児島市唐湊 4-22-1

³立教大学 社会学部: 〒171-8501 東京都豊島区西池袋 3-34-1

⁴鹿児島大学 理学部: 〒890-0065 鹿児島市郡元 1-21-35

⁵創価大学 工学部: 〒192-8577 東京都八王子市丹木町 1-236

1. はじめに

1.1 ネットワーク社会における統計教育の必要性

脱工業化から情報化・高度情報化社会を経て、1990年代後半以降のIT革命がもたらしたコンピュータ・ネットワーク社会では、膨大な電子化データの蓄積とその分析処理が大きな課題として取り上げられている。例えば、企業対企業および企業対消費者間でのトランザクション処理がコンピュータ・ネットワークを介してオンラインで記録される現在、諸種の受発注情報システム EOS (Electronic Ordering System) や販売時点情報管理システム POS (Point of Sales) によるデータベースを保有する企業も多く、それらのデータベースから探索的に情報を発掘する分析技法への期待は大きい。

また、官公庁・財団・民間企業が提供する金融・経済・社会諸事象に関する統計データベースの Web 上での公開は一般化しており、山田 (2000) の調査では、とくに民間企業が作成し定期的に更新する統計データ提供のサイト数は 1998 年 9 月時点の 233 から 2000 年 5 月時点で 512 と倍増していることが示されている。この傾向は今後も顕著に続くであろう。更に、官庁統計資料の公開においても、大学・経団連関係の強い要望から従来のマクロベースのデータ提供に加えて、ミクロ形式でのデータベース公開が検討されている (永山 (1999), 経済団体連合会 (1999))。

他方、濱砂 (1997) などで指摘されているように、諸種の経済・社会調査結果のネットワークを介してのミクロデータの公開と分析利用の一般化が先行する欧米諸国では、コンピュータによる情報収集・情報分析能力の差がデジタルデバイドとして社会問題化している。

このようにネットワーク化が急速に加速し電子化されたデータが氾濫する社会環境にあって、

- ・自発的な問題設定能力
- ・関連する資料・データなど情報の迅速な収集能力
- ・統計分析を含めた資料・データの的確な分析能力
- ・結果の効果的な情報発信能力

等を有する人材確保への社会的要請は高く、学生はもとより行政職員や企業人を対象とした統計分析を主とする人材教育のあり方が重要視されている。

1.2 インターネットを介した教育リソースの現状と本研究の位置付け

統計学はひろく多様な分野の実証分析に応用される学際的な基礎学問分野であるが、従来の統計学の講義形態は数理的解説が主であり、データ解析ツールとしての実践的な指導が伴わない点が少なからず問題であった。しかし、近年のインターネットの普及は、マルチメディアや動的グラフ、データ提供サイトとのホットリンクなどの活用を通して、受動学習から能動型・参加型学習へ、また理論学習から効率的な実践学習への移行を容易にする新メディアの教材開発を可能にしている。同時に、大学を始めとする教育・研究機関においても、ネットワーク環境の整備に伴い、学問分野を問わず、通常大学の講義にインターネット上の情報資源を組み入れることが従来に比べて比較的容易になってきている。そのため、インターネット上で対話型およびマルチメディアの特性を駆使した多様な教育資源の開発が、個人や研究施設単位で進行している。

統計・計量経済系の専門教育に関しても、海外では、JAVA applet 教材やグラフィックスを多用した動的な教育資料が Müller (1998), de Leeuw (1997), the GASP homepage (West (1997)) 等によって公開されている。また、1999 年度に開催された国際統計協会 (ISI) の第 52 回会議でも、「統計普及のための新技術」と題した特別セッションが設けられ、Web を利用した教育技術および教材改革の必要性が論じられた。2000 年度の統計計算に関する欧州国際会議

(COMPSTAT2000) においても、スペインの C. del Campo (2000) の統計学習ゲーム、P. Munoz (2000) のイントラネットを活用した統計コースの開設、ベルギーの P. Derius (2000) の統計概念の解説のための JAVA applet 集の共同利用など、統計教育に特化した会議でなかったにもかかわらず、新技術の下での統計教育教材開発の論文発表が目立っていた。また同年、国際数学教育者会議 (ICME-9) で招待講演として紹介された H-J. Mittag (2000) の統計・計量経済学のためのマルチメディア教材は、視覚に訴える効果的な動的グラフが評価され、2000 年度欧州学術ソフトウェア賞「数学部門」2 位を受賞している。このソフトウェアはインターネット上の通信教育と絡めて州教育委員会の下で組織的に開発されており、既に欧州で最大規模の通信大学であるドイツ Hagen 大学において、インターネット講義用教材として約 16000 人の受講実績とドイツ国内での他のおよそ 3000 校での採用実績を持つ。この分野で最も先行する北米での事例は、Mulkar (2000) に詳細にまとめられている。例えば、現実のデータ解析事例集や統計解析の概念を示すシミュレーションアプレットが、Rice 大学の “Virtual Lab in Statistics”⁽¹⁾、Carnegie Mellon 大学の “The Data and Story Library (DASL)”⁽²⁾、South Carolina 大学の “GASP (Globally Accessible Statistical Procedures)”⁽³⁾、Wisconsin 大学の Charles Stanton による “Java Demos for Probability and Statistics”⁽⁴⁾ などのそれぞれの Web ページで参照できる。また、NWP Associate, Inc. の “Statlets”⁽⁵⁾ においては、“User Manual” 中の “Glossary” という Web ページにおいて、統計用語の解説が参照できる。その他、Colorado 大学の Gary McClelland による “Data Analysis”⁽⁶⁾ や Pennsylvania State 大学の Dennis Roberts による “Minitab Worksheet Files”⁽⁷⁾ の Web ページおよび Dartmouth 大学の “Chance Database” における “Related Internet Sources”⁽⁸⁾ の Web ページでは、データ解析演習用のデータが提供されている。

上記の Web 上の教材は、とくに非英語圏にある欧州各国の場合、対応する母国語で開発されたものが多い。その背景には、言語(外国語としての英語)の障壁を設けることは、利用者にとって教材の難度を高めるだけで必ずしも各国内での基礎統計の普及には効果的ではないとの共通の認識がある。したがって、日本でも日本語を主体とした相応の統計教育関連サイトの開発が望まれる。現状では、青木(群馬大)⁽⁹⁾、阿部(高崎経済大)⁽¹⁰⁾、大西(上智大)⁽¹¹⁾ らをはじめとする複数の Web ページにおいて、体系的な基礎統計の解説がオンラインで利用できるようになりつつあるが、欧米諸国に比較すると、内容量や適用範囲で立ち遅れている事実是否めない。しかし、現在のインターネット人口や今後の利用者数の増加を見込めば、学生・一般社会人を問わず、利用者の各応用分野や知識レベルに応じたインターネット上での多様な統計関連の教育教材の必要性和需要はさらに高まることは自明であり、個別の取り組み以上に何らかの組織的な取り組みが期待される。

上記の視点の下に著者等のグループは、統計教育および統計科学の普及を効果的に支援するためのツールとして 3 つの教育用 Web サイトを構築した。具体的には、インターネット上でインタラクティブに統計基礎教育を行う「統計学習用インタラクティブ・テキスト: ITLS (Interactive Text for Learning Statistics)」, 検索機能を備える統計科学関連書籍についての「電子図書システム: EBSA (Electronic Books for Statistical Analysis)」, そして表計算ソフトウェア Excel などのインターフェイスを介して呼び出し可能な DLL (Dynamic Link Library) や DLL を利用した分析マクロを Web 上で提供する「教育用統計分析プログラムの公開サイト: DLLSA (Dynamic Link Library for Statistical Analysis)」である。これらに加えて「統計専用検索エンジン: Integrated Search Page」のページを作成した。これらはすべてインターネット上で一般公開しており、既に、一般社会人・学生・研究者等の利用頻度も高い。本稿では、これらのサイトの概要と今後の取り組みを併せて紹介する。

2. 統計学習用インタラクティブテキスト ITLS (Interactive Text for Learning Statistics)

2.1 プロジェクトの背景

本プロジェクトの前身は、統計学関連科目を担当する複数の私立大学文系学部教員によって実践型の統計分析や計量経済分析の教育を効果的に実現するために始められた教育用 Web サイトの共同開発および共同利用である。1999 年プロジェクト発足当時は、私立大学におけるコンピュータ・ネットワークの基盤整備が進んだとはいえ、大教室での多人数の学生を対象とした講義は避けられない状況であった。従って、そこではすべての教材のオンライン化を図ることは望ましい選択肢とは限らず、従来の「紙」を媒体としたテキストと PC ディスプレイ上でのオンライン型テキストの双方の利点と欠点を十分に比較した上で、これらの媒体の利点を最大限に生かした教材の開発や講義の形態を取り入れることがより現実的であった。したがって前身プロジェクトではとくに、従来型の紙媒体テキストの刊行を行い、次に、そのテキストと連動した相互補完的な Web サイトの構築という手順を選択した。共著テキストとして「インターネット時代の数量経済分析法——基礎からニューフロンティアまで——」(菊地・渡辺(1999))を多賀出版から 1999 年 4 月に刊行し、経済データの分析を具体的に始める際に必要となる基礎的な道具の解説に加えて、近年、実務面での関心が高いファイナンス・マーケット予測のための GARCH モデルや信用リスク評価のためのニューラルネットワーク分析法などの概説を与えた。

インターネット上に展開される各種教材の利用により、学生がコンピュータの操作を通して能動的に学習プロセスに参加できるような電子テキストの利用環境が得られる。同時に、従来のような受動的な講義とは別の方向から、学生の学習意欲の向上や持続性が期待できる。したがって、慣習的に文系学生が敬遠しがちなこの分野の学生参加型・実践型教育が可能となった(渡辺 他(1999))

ただし、この前身プロジェクトでは、刊行書籍との著作権規定のため、書籍内容に準じるような解説を Web に置くことができないなどの制約があった。そこで、刊行書籍との連携という形態から離れ、新たに統計学習用のインタラクティブ・テキストの開発と公開を主目的とする ITLS プロジェクトを計画した。その際、全体のデザインの統一や JAVA applet による動的グラフの追加、講義用スライドの整備充実、検索機能の拡張などを行い、学生および社会人の自習用や教員の講義資料としての利用に供してきた。

2.2 ITLS の内容と機能

ITLS の現段階での構成は、以下のカテゴリーに要約される(図 1)。

- 1) 情報検索&リンク集：インターネットの一般的な機能の説明と現実の経済・社会諸資料(統計データ)の所在の検索方法やデータダウンロードの方法のスライドによる解説。
- 2) 統計と Excel：表計算ソフトウェア Excel の基本操作のスライドによる解説。
- 3) データ分析(記述編・推測編)：統計学の基礎概念のスライドによる解説。
- 4) 回帰分析基礎理論、時系列分析と季節調整、多変量解析法の講義スライド。
- 5) Statistica での統計分析：統計分析専用ソフトウェアへの導入。
- 6) JAVA applet 集：JAVA による動的な統計グラフの提供。
- 7) 教育用分析データ集：Excel 形式(xls) およびカンマ区切り形式での統計分析実習用の課題データの提供。

具体的には、実践的なデータ解析技術への理解を促すため、上記の各カテゴリーで、キーワード検索やキーワードの補足説明を与える、講義スライドの閲覧とダウンロードができる、参照

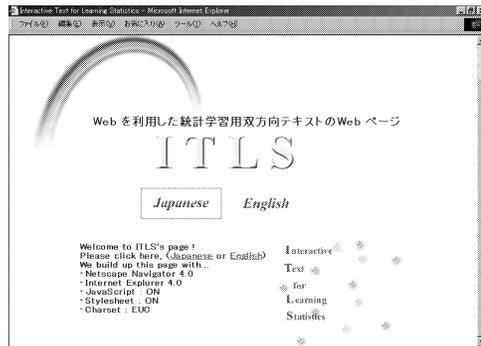


図 1. ITLS サイトのトップページ。



図 2. 内容項目別統計資料取得サイトの一覧。

サイトへのリンクを設ける，練習問題データベースを提供する，Q&A コーナーを設けるなど，従来のテキストだけではカバーできなかった機能の提供を行っている．また，統計学の概念を直感的に理解してもらえるように，活字や数式に頼らない平易な文体と短い文章，カラー，ハイパーリンク，画像や動的なグラフなどを多く含めるよう工夫している．一方，後述の電子図書システムを用いることで，さらに詳細に統計科学に関する知識取得・理解が徹底されるという効果が期待できる．以下に，いくつかの特徴的な機能・内容を紹介する．

① 実際の社会経済関連データ所在サイトへのリンク

実践型の統計学および計量経済分析手法の習得にあたっては，現実のデータを材料として分析手法を指導することが効果的である．その際，学習者の個別的な興味に合わせた分析題材を提供できるような環境設定を行うことが，分析の動機付けや学習意欲の継続維持に効果的であろう．しかもこのことが，学習者に対して，問題発見型および解決型のデータ解析技術を指導する手がかりを与えると考えられる．幸い，具体的なデータがフリーでダウンロードできるサイトは国内外に多く存在し，実際に URL 一覧を掲載している書籍やサイトもある．ここでは，利用者が想定するデータの種別をたどることで，そのデータを掲載するサイトへ障壁なくアクセスできるようにしている．具体的には，データの種別を項目別に中分類から小分類へと閲覧可能となるように分類し，それを経由して必要な取得サイトに導く「統計データ取得サイト」のページ(図 2 参照)を設けている．このページ上では，統計データの取得が可能で 100 前後のサイトへのリンクが設定されている．現在，新たに 200 以上のデータ提供サイトを整理し，内容充実に努めている．

② キーワードリスト

ITLS での先述の各カテゴリーにおいて，その内容の理解を助けるためのキーワードリストを掲載している．また，各キーワードは，その意味の説明を視覚的に表示するスライドにリンクされている(図 3 参照)．

③ 参照 URL の概略表示とハイパーリンク

各カテゴリーについて，参照関係にある他のサイトの一覧 (URL) を用意し，そのサイトの概略を示すとともに，実際にリンクも張っている．

④ PC 操作の画面による指導

統計手法の具体的な習得には，表計算ソフトを始め諸種の分析ソフトを併用する分析実習が効果的である．ここでは，課題別に，できるだけ操作段階ごとに画面を表示し，手順の解説を付けた．学習者にとって，PC のディスプレイ上でソフト操作の解説画面を直接参照しながら

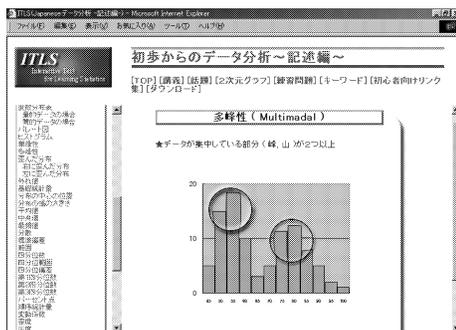


図 3. キーワードリスト(左)と解説画面(右).

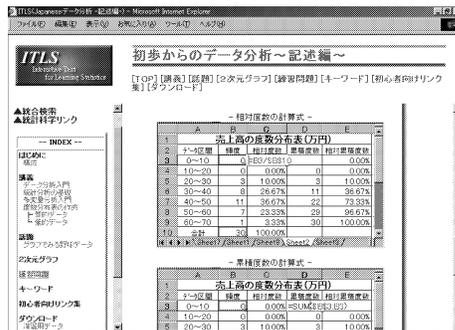


図 4. Excel 操作の指導画面.

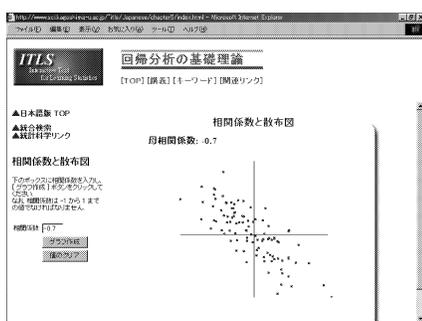


図 5. 相関係数(入力)と散布図(出力).

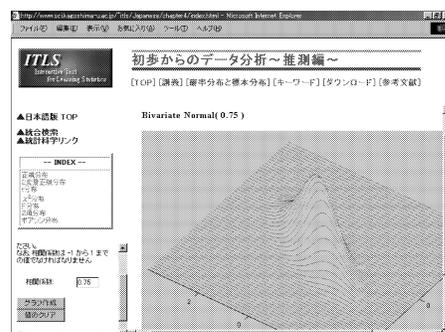


図 6. 相関係数(入力)と2次元正規分布(出力).

実際の分析操作が進められる利用環境を提供した(図4).

⑤ 対話型の動的な統計グラフ

数学を不得手とする学習者にとって、紙媒体のテキストのみでは例示の数に制限があるため、統計グラフとそのグラフの形状を決定するパラメータとの関係を理解することは難しい。ITLSでは、学習者がパラメータの値を任意に入力指定すると、出力結果のグラフの形状が即応的に変化する、いわゆる動的なグラフ機能を提供している。これらの会話型のグラフは、学習者に、数式を用いることなくパラメータの役割を、視覚的、直感的に理解させるのに有効である。以下に、ITLSで提供している代表的な動的グラフをいくつか紹介する(図5から図8)。

⑥ 講義スライドの閲覧とダウンロード

講義用のスライドをオンラインで閲覧できることに加えて、これらは自由にダウンロードできるので、学習者は各自のPCに保存し自習用に用いることができる。また、教員の利用者は、各自の講義に合わせて再編集し利用することも可能である。現在までに、記述統計、推測統計、多変量解析等について、10数種の講義用スライドのファイル(PowerPoint形式)が利用可能である。

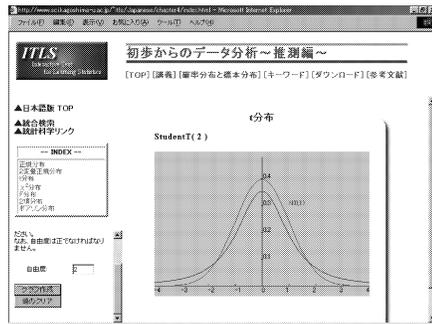


図 7. 自由度(入力)と t 分布(出力).

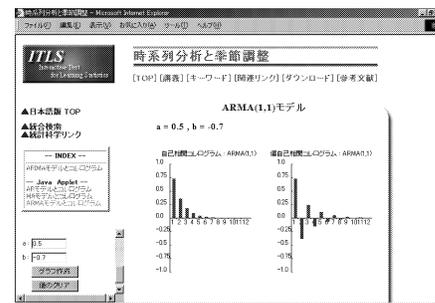


図 8. 係数(入力)に応じて変化するコレログラムと偏コレログラム.

⑦ 練習問題の提供とデータベース化

ITLS では、できるだけ多くの練習問題を掲載するよう準備中である。また、将来的には、問題の難易度の提示やデータベース化に取り組む予定である。

⑧ Q&A コーナー

ITLS の利用者に対して、メールでの質問・問い合わせを許している。回答に関しては、質問を整理し、Q&A コーナーのページに掲載する方式をとっている。

⑨ 統計用語専用のキーワード検索エンジン

ITLS の各ページで、本サイト内でのキーワードの検索ができる機能を提供している。また、ITLS 内に限らず統計科学系サイトに限定してキーワード検索を行うページへのリンクも、ITLS 内の各ページに置いているので、随時、用語や概念の解説を広範囲の専門サイトから得ることが可能である。なお、検索エンジンに関しては、4 章で更に詳しく述べる。

⑩ オンライン利用者アンケート利用者の ITLS に対する評価や希望事項を取得できるように、常設の利用者アンケート調査の Web ページを用意している。そこでは、通常の実験型設問に加えて、本サイトへの感想や意見を記入できるような自由記述形式の設問も用意した。

この他、PC 操作に不慣れな初心者向けに、基本操作(キーボード入力からファイルの保存方法まで)や WWW の閲覧・検索方法、メール設定、ワードプロセッサ(マイクロソフト WORD)の操作、簡単なホームページ(HP)の作成方法等に関して、操作画面を付けて平易に解説したページを用意し、コンピュータ・リテラシー教育との連動を図っている。これらを含めて現在、約 330 の講義スライドと 120 前後のキーワード解説画面、17 種の JAVA applet による対話型の動的な統計グラフが提供されている。

このように、常に最新情報に更新され蓄積された教材を、インターネットを介して公開することで、だれもが使用場所や時間の制約を受けずに自由に学習に活用できる。同時に、大学や企業の統計教育担当者にとっては、講義資料の共有化という利点もある。将来的には、利用教員間での自由な意見交換の場をネット上に設置する必要が出てくるであろう。

2.3 今後の課題

(1) 国内および国際相互協力

本サイトの開設は、統計科学の知識習得を望む利用者を対象とすると同時に、教員向けの講義用の教材提供サービスを行うことも意図している。このため、内容が豊富で相互利用可能な

コンテンツが必要となるが、筆者等プロジェクトメンバーのみでその作成や提供を行うことは効率的ではない。そこで、本サイトとの相互リンクによる教材提供への協力を仰ぐために、今後も国内・国際学会での論文報告を通して、このプロジェクトの有用性を広く示していく必要がある。また、国際的な相互協力も重要である。例えば、Hagen 大学で開発中の動画や JAVA applet などのマルチメディア・データベースの国際版⁽¹²⁾には、英語、ドイツ語、スペイン語など言語別のマルチメディア・アプレットが登録され公開されている。本プロジェクトでも、作成した JAVA applet の英語およびドイツ語版や Mittag (2000) で作成されたアプレットの日本語版などを作成・公開し、国際的な資源の共有化を進めている。

(2) ミラーサイト

画像や動的なグラフを多用し、また音声解説を付けるなどのマルチメディア化を進めていくと、利用回線の送受信時の負荷が増大し、応答が鈍くなることは否めない。しかし、将来的には大学間ネットワークの整備・改善が期待されるので、それに伴ったコンテンツの設計指針の方向も見えてこよう。また、講義利用時における一時的なアクセスの集中化現象などの当面の問題については、本システムの利用申し出のあった機関のサーバに対してコンテンツを提供し、このサーバをミラーサイトとして運用稼動する方式を用いることで、部分的には解決するであろう。学生の自宅自習用としては、CD-ROM 形式での配布を検討している。

3. 統計科学のための電子図書システム EBSA (Electronic Books for Statistical Analysis)

3.1 システムの目的

過去に統計科学、なかでも理論に関わる分野で専門的學術書もしくは諸種の応用分野への統計解説書として使用されてきたが、現在では絶版もしくは再版予定のない貴重な書籍が、数多く見受けられる。これらの書籍は、年月の経過に関わりなく普遍的かつ有用であり、統計科学に関する知識の宝庫である。しかし、既にその多くが実際に利用可能な状態にはなく、入手不能となっている。著者等のグループでは、このように絶版等で入手困難となった統計科学の理論および応用に関する書籍について、著作権者(該当する書籍の著者もしくは著者のご遺族等)および出版社から正式な許可を得たうえで、それぞれを電子化し、EBSA (Electronic Book for Statistical Analysis) の名称の Web サイト上で公開している。

表 1 は、現在、Web サイト“EBSA”に登録されている図書一覧である。EBSA では、各書籍の索引語から該当ページへリンクが張られている。また、EBSA 上のキーワード検索エンジンを通して登録されている索引語を検索することができる。現段階で電子化作業を終えた書籍の索引語の総数は 3017 個を数える。ここには重複もあるが、かなりの数の統計科学専門用語に関する解説がインターネット上で得られることになる。この意味で、本サイトは統計科学関連のオンライン辞書もしくは知識ベースとしても利用可能である。

図 9 は、登録電子書籍の表紙イメージ、図 10 は、EBSA サイトのトップページである。図 11 は、電子書籍「質的データの多変量解析(松田紀之著)」のトップページで、右側に目次と索引語検索用のメニューがある。図 12 は、同書籍の索引のページである。ページ番号にはすべて該当ページの画像ファイルとのリンクが設定されている。

EBSA のサイトは、Yahoo! ポータル・サイトに登録されており、「自然科学と技術」-「数学」-「統計学」のカテゴリの中で、「統計科学のための電子図書システムの Web ページ」として参照できる。本サイトは、2000 年 5 月の公開から 2001 年 4 月現在で、約 6800 のアクセスカウントを持つ。

表 1. 電子図書システム収録書簡リスト(2001. 4. 9 現在).

書名	著者	出版社	発行年	索引語数	PDF 容量
因子分析法通論	浅野長一郎	共立出版	1972 年	349	56.3MB
確率論及統計論	伏見康治	河出書房	1942 年	作成中	作成中
実験計画法	増山元三郎	岩波書店	1956 年	作成中	41.0MB
実験計画法講義 I 基礎編 (1)	北川敏男	培風館	1955 年	614	52.6MB
実験計画法講義 II 基礎編 (2)	北川敏男	培風館	1956 年	319	43.6MB
質的情報の 多変量解析	松田紀之	朝倉書店	1988 年	187	27.5MB
少数例のまとめ方	増山元三郎	河出書房	1953 年	753	16.9MB
推計学への道 統計学を越えて	増山元三郎	東京大学出版部	1950 年	作成中	作成中
調査統計入門	船津好明	共立出版	1977 年	244	16.0MB
続調査統計入門	船津好明	技興社	1986 年	作成中	作成中
多変量解析論	塩谷実 浅野長一郎	共立出版	1967 年	作成中	22.7MB
調査の技術	浅井晃	日科技連出版社	1992 年	313	44.3MB
統計学の三十年 —わが師わが友—	北川敏男	共立出版	1969 年	索引無し	25.1MB
統計学の認識 基盤と方法	北川敏男	白揚社	1948 年	作成中	69.0MB
21世紀の食糧・農業	奥野忠一	東京大学出版会	1975 年	索引無し	46.1MB
農業実験計画法小史	奥野忠一	日科技連出版社	1994 年	529	56.9MB
モンテカルロ法	宮武修 中村隆	日刊工業新聞社	1966 年	128	28.4MB



図 9. 登録書籍の表紙イメージ.

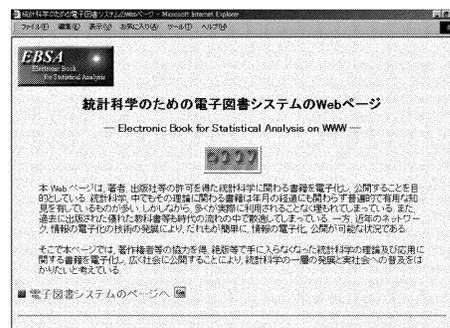


図 10. EBSA トップページ.

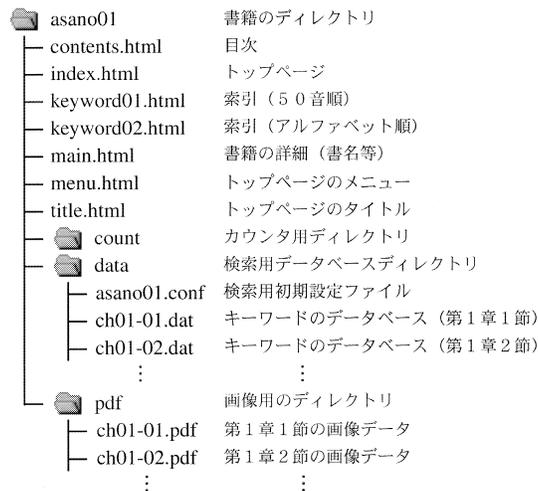


図 13. ファイルのディレクトリ構成(例:「因子分析法通論」)

表 2. 電子化作業環境.

機種	Power Macintosh G3
OS	Mac OS J1-8.1
スキャナ	UMAX Astra 1220S
ソフト	Adobe Photoshop 5.0
	Adobe Acrobat 3.0J
	VistaScan V2.43

レースケールを使用しており、また解像度を 600 dpi としている。

3.3 EBSA における検索システムの概要

索引語からの検索を可能にするため、各書籍に対して図 14 で示す内容の検索用初期設定ファイル (*.conf) を作成している。

例えば、書籍「因子分析法通論」に対する検索用初期設定ファイルは、ファイル名「asano01.conf」として、図 15 のような内容になっている。

また、書籍内の各節を格納した検索用データファイル (*.dat) は、キーワード(索引語)とそれに対応するページのリストから構成されている。例えば、「因子分析法通論」の第 1 章第 3 節に対応するデータファイル (ch01-03.dat) の内容は、図 16 の通りである。

以上の「検索用初期設定ファイル」と「検索データファイル」の準備の下に、実際の検索は以下のプロセスに従って実行される。

- (1) 検索キーワードの入力, 検索対象(書籍名)が選択される。
- (2) 選択された書籍の「data ディレクトリ」内にある「*.dat」ファイルにより検索が実行される。その際、ヒットしたファイル名と該当ページ番号により、「初期設定ファイル」(*.conf) から目的のファイルの当該するページが探される。

```

書籍名
検索用データファイル名 1 : 第 1 章 : 章名 : 章の最初のページ番号
検索用データファイル名 2 : 第 2 章 : 章名 : 章の最初のページ番号
...
検索用データファイル名 1 0 : 第 1 節 : 節名 : 節の最初のページ番号
...

```

図 14. 検索用初期設定ファイル (* .conf).

```

因子分析法通論
ch01:1 章:概論:1
ch01-01:1.1:因子分析法の認識:1
ch01-02:1.2:因子分析法の由来と適用分野:2
ch01-03:1.3:因子分析法の型:8
ch01-04:1.4:因子分析法の心構え:10
ch01-05:1.5:因子分析法の適用技法:13
ch01-06:1.6:因子分析法の計算プログラミング:16
ch02:2 章:成因分析法:19
ch02-01:2.1:成因分析法の立場:19
...
chA::付録:427
refer::参考文献:460
index::索引:477

```

図 15. 検索用初期設定ファイル(例) (asano01.conf)

(3) 当該ページとキーワードからなる HTML ファイルが自動的に作成される(図 17 参照). EBSA 内の検索システムにより自動的に作成された対象キーワードが含まれる節とページのリストを表示するための HTML ファイル上では, 利用者が実際に閲覧したい節あるいはページをクリックすると Acrobat Reader により当該ページが表示されるように, 書籍内容の画像ファイルとの間にリンクが張られている.

また, 検索キーワードの入力様式に関しては, アスタリスク(*)を指定する前方および後方,

完全規定の構造模型:8
正準分析法:9
正準相関分析法:8, 9
双因子解法:8, 9
対外的因子分析法:9
対内的因子分析法:9
多因子解法:8, 9
二因子解法:8, 9
不完全規定の模型:8

図 16. データファイル(ch01-03.dat)の内容 .

中間の一致検索機能(いわゆるワイルドカードの利用)や括弧を含めたアンド検索とオア検索およびノット検索機能など、通常の正規表現による検索が可能である。図 18 は、書籍内検索の例で(左側に表示)、キーワードとして「ロジット*」を入力して得られた検索結果を示している(右側に表示)。前方一致検索の結果、「ロジット変換」、「ロジットモデル」などのキーワードに対応した書籍が検索され、その書籍内のページ番号が出力されている。またここで、それぞれのページ番号は該当ページと直接リンクする。

図 19 は、複数の書籍内容を検索対象とした場合の例とその出力画面である。図 20 は、図 19 の画面において、登録済みの全書籍を対象として、キーワード「正規*」を入力し前方一致検索を行った結果である。書籍別に入力キーワードに対してヒットした索引語と、それに対応する該当ページ番号が出力されている。この機能により、同一キーワードに対して、複数の書籍に含まれる説明内容を同時に参照することができる。

3.4 今後の課題

現在の EBSA が備える検索システムでも、検索対象が数冊の単位であれば検索に要する時間をさほど気にすることなく利用できるが、検索対象とする書籍数が増えれば検索時間の増大という問題が生じる。これを改善する一つの方法として、Oracle 等のデータベース・エンジンを利用することが考えられる。また将来に向けて、システムの強化・改善を考えると、大量データを安全に保存し、高速なファイルアクセスを可能にするための大規模な記憶装置や高速バックアップ・システムの導入を検討する必要がある。

今後、出版社の都合で絶版の恐れがある学術専門書に関して、それらに含まれる専門知識の風化を防ぐ意味からも、ここに示した EBSA システムの継続的な管理・運営を可能にするような組織的な取り組みが必要である。そのためには、国立情報学研究所や統計関連諸学会、大学等が保有するサーバへのコンテンツ移植作業や相互利用システムの構築などを進める必要がある。また、EBSA システムでは、先ず絶版書を対象としてプロジェクトを進めてきたが、最新の統計諸手法の解説文書なども著者の意向を確認したうえで、従来の出版形態に頼らず、本システムを通して、利用者に迅速に広く公開することが可能である。このような方向も今後の研

```

<html>
<head>
<title>キーワード検索(全書籍対象)</title>
<meta http-equiv="Content-Type" content="text/html; CHARSET=euc">
</head>
<body bgcolor="#ffffff" background="/ebsa/images/back.gif">
<base font="4">
<pre>
<span style="font-size:20px;font-weight:bold"> 検索結果 </span>   <small>キーワード(ページ)</small>
    1.2 <a href="/ebsa/asano01/pdf/ch01-02.pdf" target="_top">
        因子分析法の由来と適用分野</a>
            二因子解法
<a href="/ebsa/asano01/pdf/ch01-02.pdf#page=1" target="_top">2</a>
    1.3 <a href="/ebsa/asano01/pdf/ch01-03.pdf" target="_top">
        因子分析法の型</a>
            対外的因子分析法
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=2" target="_top">9</a>
            二因子解法
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=1" target="_top">8</a>
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=2" target="_top"> 9</a>
            多因子解法
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=1" target="_top">8</a>
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=2" target="_top"> 9</a>
            双因子解法
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=1" target="_top">8</a>
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=2" target="_top"> 9</a>
            対内的因子分析法
<a href="/ebsa/asano01/pdf/ch01-03.pdf#page=2" target="_top">9</a>
...
        以上<span style="color:red"> 32</span> 枚のファイルが存在します.
<span style="font-size:20px;font-weight:bold">----- 検索終了 -----</span>
</pre>
</body>
</html>

```

図 17. 例(「因子分析法通論」でキーワード“因子”とした場合の検索結果(HTML ファイル)).

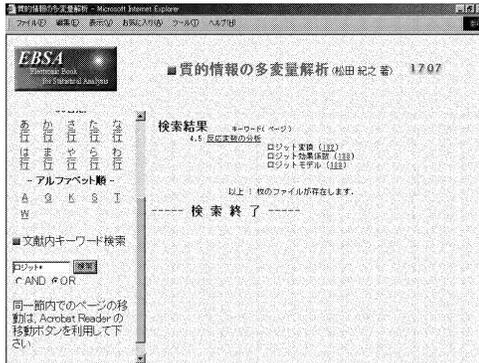


図 18. 書籍内キーワード検索結果(例).

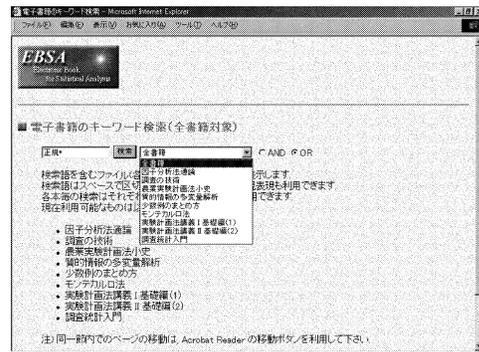


図 19. 複数の書籍を対象とした検索ページ.

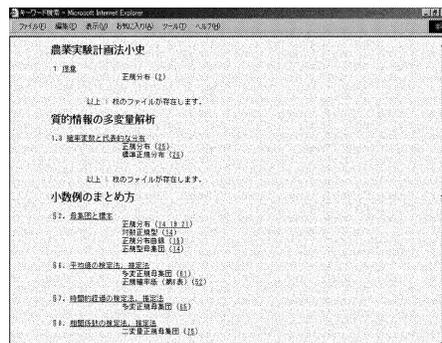


図 20. 複数書籍内容からの検索結果.

究課題としたい。

4. 統計専用の統合検索エンジン (Integrated Search Page)

現在、インターネット上で検索エンジンを辞書のように用いて、検索対象としたキーワードに関する諸種の情報収集を行う方式が一般化しつつある。しかし、既存のロボット型検索エンジンを利用した場合、利用者の要求に合致しない不適切なサイトまで検索対象としてサーチするなど、検索効率や検索内容の妥当性に関する問題が生じている。これを回避する手段として、専門領域に特化したサイトの絞り込みを行い、検索対象範囲を限定したうえで、ロボット型検索サービスを提供する方式が考えられる。

上記の視点から、著者等のプロジェクトでは、まず、検索対象を Web テキスト“ITLS”および電子図書システム“EBSA”との両者の範囲に限定した検索エンジンを用意した。次に、本プロジェクト以外の、統計科学関連の学習ページを提供する数種のサイトを選定し、“ITLS”および“EBSA”に加えてこれらのすべてを検索対象とする検索エンジンを用意し、統計関連キーワードの検索のための Web ページとして、両方の検索エンジンを「統合検索ページ (Integrated Search Page)」で提供している。

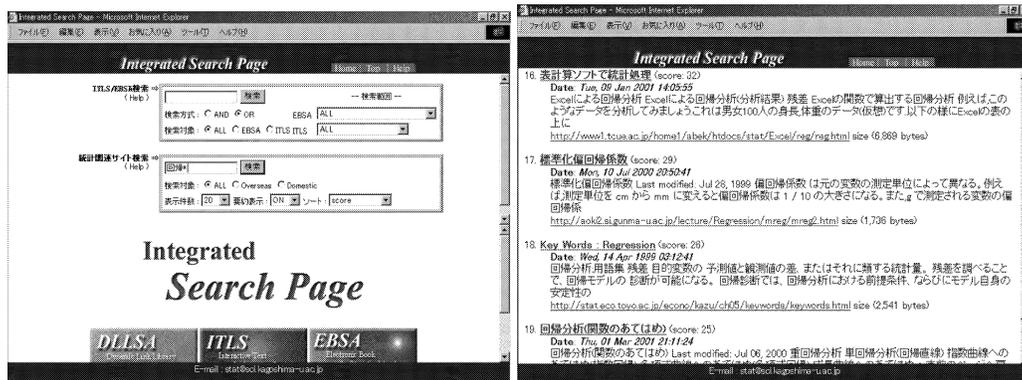


図 21. 統計科学系専用検索のページ(2 段目).

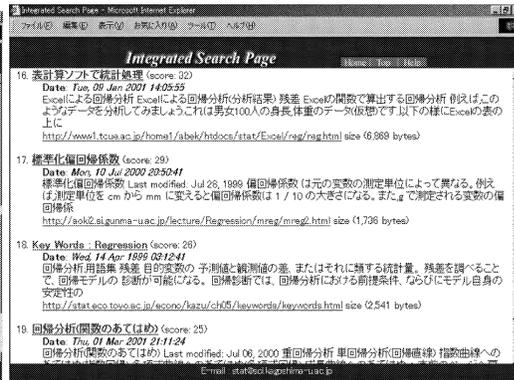


図 22. 統計科学系検索エンジンでの出力結果(一部).

ここでの全文検索システムは、フリーソフトウェア Namazu v1.3.0.6⁽¹³⁾ によって構築されている。また、検索対象とした統計科学関連サイトの情報は、ミラーリング機能を利用して本プロジェクトのサーバに定期的に写し取り、随時、索引データベースを更新している。検索キーワードの指定方法に関しては、既に述べた電子図書システム“EBSA”における場合と同様に、アスタリスク(*)を指定する前方および後方、中間の一致検索機能が利用できる。また、括弧を含めたアンド検索とオア検索およびノート検索機能など通常の正規表現による検索も可能である。

図 21 は、統合検索のための Web ページである。上段の検索ボックスは、“EBSA” and/or “ITLS” を対象としており、下段の検索ボックスは、国内および国外の統計科学関連サイトを対象としたものである。図 22 は、図 21 において、入力キーワードを「回帰*」とした場合の検索結果の一部である。

ただし、本来、目標としたい効率的な知識獲得のための検索を考えると、入力した用語の単なる検索だけでは不十分で、入力した用語と関連する用語も検索結果として自動的に表示する、ある種のシソーラス機能を備えていることが望ましい。こうした統計科学用語専用のシソーラスの開発は、今後の研究課題としたい。

5. 教育用統計分析ツールの Web 上での提供 (DLLSA: Dynamic Link Library for Statistical Analysis)

大学を始めとする多くの教育機関では、一般に、統計学の教育に入る前に、既に情報リテラシー教育の初段階で、Microsoft-Excel(以後、Excel と記述)に代表される表計算ソフトウェアの操作を教えている。また、Excel の使用経験を持つ社会人も多い。加えて、学生および社会人が購入しているほとんどの PC には、Excel がプレインストールされている。したがって、統計的なデータ分析を初めて学ぶ利用者には、Excel 上で統計分析を実行できる環境があれば便利である。ただし、Excel は本来、統計分析のための専用ソフトではないため、分析ツールに簡単な分析マクロがあるものの多変量解析や時系列分析の諸手法などを含まず、また、含まれている手法にしても誤解を招く操作上の日本語訳が付されているなど、問題点も挙げられる。

上記の視点から、Web サイト DLLSA (Dynamic Link Library for Statistical Analysis) を設けた。DLLSA では、諸種の統計分析法の計算プログラムを Windows システムの下で動作する DLL (Dynamic Link Library) として用意し、それに Excel 上で機能するユーザ・インター

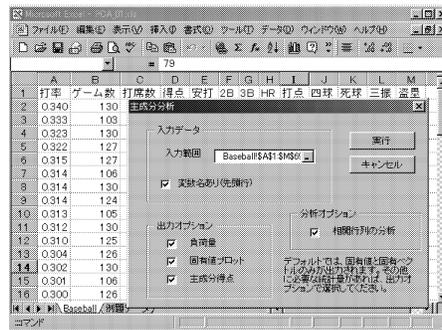


図 23. 主成分分析を実行するための Excel インターフェイス .

フェ - スを添えて提供している . 具体的には , 本サイトから各統計手法を分析マクロとして組み込んだ Excel シートを利用者の PC 上にダウンロードするだけで , 通常の Excel には含まれていない統計手法が容易に実行できる .

現在 , 多変量分析法を中心に , 主成分分析 , 計量多次元尺度法 I・II , 潜在クラス分析 , クラスター分析の諸手法を用意している . 図 23 は , Excel 上で主成分分析を実行するために提供しているインタ - フェイスである . 通常の Excel 操作と同じ感覚で高度な統計分析が可能となる .

6. 今後の取り組み

本報告で紹介したような各種のサイトは , 今後 , 統計教育や統計科学普及活動を進めるうえでの有用なツールとして位置付けられる . ここでは , 筆者等のグループで開発を進めてきた , インターネット上で利用できる各種システムや Web サイトについて議論した . 即ち , 統計学習用インタラクティブ・テキスト (ITLS) , 電子図書システム (EBSA) , 統計専用検索エンジン (ISP) , 統計教育用分析ツール (DLLSA) について , それらの開発意図や設計指針 , 主な機能を俯瞰的に述べた .

従来は , 統計科学普及の一つの手段として , 統計分析ソフトウェアの開発があったように思われる . しかし , PC が大衆化し , しかも表計算ソフトウェアのような平易なデータ処理ツールが普及した現在 , 統計科学に関する基礎教育のあり方の見直しや新たな教材開発の進むべき方向が問われており , それに関連した方法論を期待する声も高まっている . とくに , インターネットが新しいメディアかつコミュニケーション・ツールとして成長を続けている現状を見ると , インターネット上で取得可能な統計科学に関する知識や情報の充実を図ることは , 今後の統計教育や普及活動を効果的に進めることに繋がると考えられる .

このようなことから , 実践的な統計基礎教育を効果的に進めるためには , 広範な応用分野に接点を持つ統計科学の特質を活かした , 統計教育に適した統計データベース , 統計教育ソフトウェア , そして手軽に入手可能な分析ソフトウェア , オンライン統計辞書 , 解析事例データベースなどが有機的に結びついた統合システムが必要となる . 本稿で紹介した各サイトは , 統合化に向けての最初の試みであり , 今後この種のシステムが進むべき方向を示唆するものとしての意義は大きいと考えている . 本稿で示した Web サイトは , 現在 2 ヶ所のサーバ上で管理運用されている . これにアクセスするための URL は , それぞれ

<http://stat.eco.toyo.ac.jp/~stat/>(東洋大学)

<http://www.sci.kagoshima-u.ac.jp/~stat/>(鹿児島大学)

である .

大学間ネットワークシステムの強化に伴い, 統計教育の場に限らず, 今後も様々な分野で電子化教材の蓄積及びそれらの共有化が求められるであろう. 本研究がその一助になれば幸いである .

本研究は, 電子テキスト開発に対して, 一部, 平成 12 年度東洋大学井上研究助成を, また, 分析ソフトウェア開発に対して, 一部, 平成 12 ~ 平成 13 年度文部科学省科学研究費展開研究 (B)(研究課題番号 12558022) の助成を受けている .

注釈

- (1) <http://www.ruf.rice.edu/~lane/rvls.html>
- (2) <http://lib.stat.cmu.edu/DASL/>
- (3) <http://www.stat.sc.edu/rsrch/gasp/>
- (4) <http://www.math.csusb.edu/faculty/stanton/m262/>
- (5) <http://www.statlets.com/usermanual/glossary.htm>
- (6) <http://psych.colorado.edu/~mcclella/statistics.html>
- (7) <http://roberts.ed.psu.edu/users/droberts/mtbfiles.htm>
- (8) <http://www.dartmouth.edu/~chance/RelatedSources/sources.html>
- (9) <http://aoki2.si.gunma-u.ac.jp/lecture/>
- (10) <http://www1.tcue.ac.jp/home1/abek/htdocs/stat/>
- (11) <http://econom01.cc.sophia.ac.jp/>
- (12) <http://mmk.fernuni-hagen.de/>
- (13) <http://www.namazu.org/>

参 考 文 献

- de Leeuw, J. (1997). The UCLA statistics textbook and module, *Bulletin of the International Statistical Institute*, Book 2, 55-58.
- del Campo, C. (2000). A game to learn statistics, *Proceedings of COMPSTAT2000*, 169-170.
- Derius, P. (2000). A collection of applets for visualizing statistical concepts, *Proceedings of COMPSTAT2000*, 253-258.
- 経済団体連合会 (1999). わが国官庁統計の課題と今後の進むべき方向, 1-27.
- 菊地正佳, 渡辺美智子 (1999). 『インターネット時代の数量経済分析法』, 多賀出版, 東京 .
- 濱砂敬郎 (1997). 諸外国における調査個票(マイクロデータ)の多角的結合的利用の経験, 統計情報活用のフロンティアの拡大の総括的研究(平成9年度文部省科研費報告書, 松田芳郎 編), 148-162.
- Mittag, H-J. (2000). Multimedia and multimedia databases for teaching statistics, *CD-ROM of Proceedings of ICME-9 (Invited Paper)*.
- Mulkar, M. (2000). Internet resources for AP statistics teachers, *Journal of Statistics Education*, 8(2), (<http://www.amstat.org/publications/jse/secure/v8n2/mulekar.cfm>).
- Müller, M. (1998). Computer-assisted statistics teaching in network environments, *Proceedings of COMPSTAT 1998*, 77-88.

- Munoz, P. (2000). Information technologies in an advanced course, *Proceedings of COMPSTAT2000*, 233–234.
- 永山貞則 (1999). 日本でのマイクロデータ公開の展望, 統計情報活用のフロンティアの拡大(文部省科研費報告書, 松田芳郎 編), 10–15.
- 山田 茂 (2000). 民間作成統計のホームページへの収録状況, 第 68 回日本統計学会講演報告集, 228–229.
- 渡辺美智子, 門間麻紀, 櫻井尚子, 橋本紀子, 浅野美代子, 袴道 守, 井上達紀, 山口和範 (1999). 統計・計量経済学のためのインタラクティブなインターネット教材の共同開発, *情報教育方法研究*, 2(1), 37–42.
- West, R. W. (1997). Statistical applications for the World Wide Web, *Bulletin of the International Statistical Institute*, Book 2, 7–10.

A Proposal for Educational Web Sites for Disseminating Statistical Knowledge on Internet Environment

Michiko Watanabe

(Faculty of Economics, Toyo University)

Katsumasa Suenaga

(Science of Living Department, Kagoshima Immaculate Heart College)

Akinobu Takeuchi

(College of Social Relations, Rikkyo University)

Hiroshi Yadohisa

(Faculty of Science, Kagoshima University)

Kazunori Yamaguchi

(College of Social Relations, Rikkyo University)

Chooichiro Asano

(Faculty of Engineering, Soka University)

The large-scale growth of the Internet as a large-scale communication media has increased information about statistical science and promoted education and efficient dissemination of statistics. This paper outlines the functions of statistical educational Web sites that our project members have been developing as a collaboration research. The first one is the ITLS (Interactive Text for Learning Statistics) site, which is an interactive online textbook for statistics by multimedia tools like JAVA applets. Our newly styled textbook is designed for Web pages and includes a lot of dynamic graphics and images. The second one is the EBSA (Electronic Books for Statistical Analysis) site, which is an electronic book reading system with special searching tools. The third is the DLLSA (Dynamic Link Library for Statistical Analysis) site, which provides various statistical analysis programs that can be called in several application softwares for the Windows system, e.g. Microsoft-Excel. The DLLSA site also provides interface programs over Microsoft Excel. We mention the special searching system only for keywords on statistics.

These educational sites on the Internet have gradually extended both the range of persons who want to study and the chance for studying statistics under every circumstance.

Key words: Online interactive text , statistical science , electronic book reading system, search engine for statistics, dynamic link library for statistical analysis, educational Web site.