

早稲田大学における統計システム入門セミナーの 現状と展望

齋藤 朗宏*

* 早稲田大学メディアネットワークセンター
akisaito@aoni.waseda.jp

海老原 崇**

** 早稲田大学大学院商学研究科
ebihara@mnc.waseda.ac.jp

1. はじめに

早稲田大学メディアネットワークセンターでは、前期2回、後期2回の毎年4回、"統計システム入門セミナー"と題して、統計システムの初学者を対象として、UNIX版SASの利用セミナーを実施している。

この統計システム入門セミナーの歴史は古く、記録として残っている限りでも、1997年には既にメディアネットワークセンターとしてセミナーを実施していたようである。現在では、本学でもWindows版のSASがいくつかの学部を導入され、また、当時利用されていたSASバージョン6から現在ではSASバージョン9へとアップデートされるなど、情報環境等は大きく変化している。そこで本稿では、この統計システム入門セミナーの歴史的経緯、現状を述べ、現在のセミナーの持つ問題点、今後の展望について議論する。

2. 統計システム入門セミナーの現状

2-1. 経緯

早稲田大学メディアネットワークセンターでは、学生並びに教職員向けの統計システムをUNIX上で利用出来るサービスとして、"statシステム"を提供している。文学部など一部学部のコンピュータ教室においては、Windows版のSAS、SPSSなどが利用可能であるが、全学向けに提供されているサービスは、現時点ではstatシステムのみである。

メディアネットワークセンターでは、statシステムを全学に向けて提供している立場から、統計システム入門セミナーを主催している。

2-2. statシステム

statシステムは、SunOS 5.8がインストールされているサーバ2台(stat11サーバ、stat12サーバ)

より構成されている。それぞれにはSAS(バージョン6.2, 8.2)やTSPなどのソフトウェアが導入されている。尚、SASに関しては、ターミナルからの処理が可能である点を考慮して、バージョン6.2を基本としている。

statシステムのアカウントは希望のあった学生、教職員にのみ発行しており、Web上での申し込みフォームを利用し、ほぼ即日の発行が可能となっている。ただし、セキュリティを考慮してID、パスワードの受け渡しは直接行っている。

statシステムのユーザはsshを利用してシステムにログインし、各種コマンドを実行する。2005年度まではtelnetを利用出来るサーバとsshを利用するサーバが併存していたが、セキュリティの観点からsshを利用するサーバのみに統一された。

学内コンピュータ教室からstatシステムのサーバにアクセスする場合には、telnetを利用していた当時はターミナルソフトウェアとしてはTera Term Pro、ftpクライアントとしてはFFFTPを利用していた。現在では、ターミナルソフトウェアとしてはPuTTY、sftpクライアントとしてはWinSCPを採用している。

2004年度1年間の利用申請件数は187件、2005年度1年間の利用申請件数は188件であり、ほぼ同数であった。授業としてstatシステムを利用しているのがメディアネットワークセンターの一部演習のみとなった影響もあり、学生、教職員あわせて5万人の中では、利用者はごく少数である。

2-3. 統計システム入門セミナーの現状

昨年度実績では、統計システム入門セミナーの受講者数は、全3回のセミナーを合計すると13名である(1回は参加者がなかったため中止)。一昨年度は全4回で21名であり、減少傾向にあるとみられ

るが、少ない時は参加者0名から多い時では13名と、人数の多い回と少ない回との差が大きいため、減少傾向については時期、曜日等その他の要因も考慮して検討する必要があるだろう。

また、各セミナーの参加総数は平均4～5名と決して多くはないが、新規利用者の1割近くは受講していることも事実であり、現段階では存在価値を否定することはできない。

2-4. 統計システム入門セミナーの内容

統計システム入門セミナーは、以下の内容から構成されている。

- ・ stat システムとは何か。
- ・ stat システム利用の流れ。
- ・ stat システムへのログイン。
- ・ vtsas の起動 (vtsas の画面については Figure 1 参照)。
- ・ vtsas の各種コマンド。
- ・ プログラム作成 (PROC PRINT を使用)。
- ・ プログラム作成と WinSCP を用いた外部データ読み込み。
- ・ 外部エディタを利用したプログラム作成とバッチ処理 (PROC UNIVARIATE を使用)。

上に示された通り、飽くまでも操作方法やデータ読み込みの方法、記述統計量の算出など、分析以前の基本的なデータ処理に必要な方法に内容は限定されている。

```

stat11.comp.waseda.ac.jp - PuTTY
+OUTPUT-----
コマンド =>

                                Print and Mean
                                20:25 Sat
                                -----
                                OBS    NAME    HEIGHT  WEIGHT
                                -----
                                1      Tanaka  172     83
                                2      Sato   165     52
                                3      Suzuki 180     75
+PROGRAM EDITOR-----
コマンド =>

00001 TITLE 'Print and Mean'; RUN;
00002 OPTIONS LS=75;
00003 DATA body;
00004 INPUT Name $ Height Weight;
00005 CARDS ;
00006 Tanaka 172 83

```

Figure 1: vtsas の画面

3. 統計システム入門セミナーの問題点

メディアネットワークセンターで開講している統計解析の授業(コンピュータによる統計解析)も含め、多くの統計学の授業では SAS などのソフトウェアの利用法以上に、何故その分析を行うのか、そのプロシジャを利用するにはどのような意味があるのかという点を重視している。

それに対して、統計システム入門セミナーは統計学を教える講義ではない。このことは、よりの絞ったセミナーを可能としている反面、セミナーの内容に対する制約ともなっている。

即ち、「何故」の部分の説明が様々な制約などから不可能であるため、そのような説明が不要な手法に限定せざるを得ず、学生がより求めている、自分の学習、研究に利用出来る分析手法の手順を学べるセミナーにはなっていないということである。

回帰分析を例に挙げるならば、主に REG, GLM, CALIS プロシジャが利用可能である。ただ回帰分析の結果を見るという目的であれば REG プロシジャが簡単であるが、線形モデルの枠組みで、分散分析と一括して学習するのであれば、GLM プロシジャを用いるのが適切である。また、SEM の枠組みで捉えるのであれば CALIS プロシジャだろう。

このようなとき、統計学の講義であるならば、講師の主義に応じて、講師の責任において使用するプロシジャを選択することができる。しかし、統計システム入門セミナーでは、こういった選択を必要とする手法にまで踏み込むことはできないのである。

4. 統計システム入門セミナーの展望

今後の展望としては、2通りの道が残されていると考えられる。1つが現状維持という道、もう1つがメディアネットワークセンターは飽くまでも情報提供に徹するという道である。

前者は、統計システム入門セミナーには本稿で挙げたような問題点はあるものの、一定の価値があるのも事実であるという考え方であり、後者は、何をどう分析するのかにはメディアネットワークセンターとしては一切踏み込まず、マニュアルの Web 上での公開などで利用者の利便性を図るべきであるという考え方である。