

[報 告] 計算科学・計算機科学人材育成のためのスーパーコンピュータ無償提供利用報告

## 応用動物科学系学生実験、動物の遺伝育種に関する基礎実験報告

— SAS を使った実験結果の統計分析演習 —

鈴木啓一

東北大学農学部・動物遺伝育種学分野

### 1. はじめに

東北大学農学部応用動物科学系では、動物科学に関する基礎的な実験手法を学ぶため毎年 3 年生を対象とした学生実験が一年間をとおして 7 つの分野が担当し行われています。この中で、動物遺伝育種学分野（担当：鈴木啓一）では、サイバーサイエンスセンターの「計算科学・計算機科学人材育成のためのスーパーコンピュータ無償提供」の制度を利用し、並列コンピュータ上でサービスされている統計処理アプリケーションソフトである SAS を使った統計処理教育を行っています。この学生実験では 15 回の実習を行い、はじめに 3 系統の近交系マウスを用い、分娩（0 日齢）から 8 週齢までの体重を週毎に学生に測定してもらいます。これらのマウスは生まれて 4 週齢から 8 週齢までの間は、高エネルギー・高蛋白質飼料と低エネルギー・低蛋白質飼料の 2 種類の飼料を給与され、性別に育成されます。得られたデータについて、SAS を使って分散分析や回帰分析を行い、体重や増体量に及ぼす系統、飼料、性の効果の検定を行い、コンピュータ処理に馴染むと同時に、統計処理方法の習得を目指しています。今年度も 34 名の 3 年生を対象とした実習を行ったのでその内容を報告致します。

### 2. データの取得

東北大学農学部動物飼育実験棟において、交配分娩した 3 近交系マウス (C57BL/6、C3H/He、BALB/c)、それぞれ 8~10 頭の雌が生んだ子供マウスの体重を分娩時から毎週 1 度測定します。3 週齢で離乳させ、4 週齢で雄雌それぞれ各系統 16 頭ずつを高エネルギー・高蛋白質と低エネルギー・低蛋白質の 2 グループに分け、8 週齢まで飼育します。それぞれの系統 32 頭(雄雌 16 頭ずつ) ずつなので合計 96 頭となります。毎週体重を測定し、それぞれの記録データを取ります。



### 3. データの解析

東北大学農学部コンピュータ演習室には35台のWindowsVista OSのコンピュータが設置されています。はじめに、Excelを使い飼料、性、系統毎にデータ入力を行います。また、メール設定、サイバーサイエンスセンター並列コンピュータへの接続のためのWinSCPとPutyの使い方を教えます。

### 4. SASプログラムを使った統計処理

農学部では、3年生を対象とした生物統計学(鈴木啓一担当)の講義が3年生の学生実験と並行して開講されています。14回の講義では、確率変数の分布、データの視覚化、点推定と区間推定、検定の考え方、母平均に関する検定、比率の検定、分散の検定、分散分析、分散分析の応用、一般線型モデル、ノンパラメトリック検定について教えます。

学生実験では、講義で修得した知識に基づき分散分析、多重検定を中心にSASを用い実際のデータ分析を行います。今年度は以下の日程で実習を行いました。

- 6月4日 メール設定、パソコンの使い方、Excelを使ったデータ入力
- 6月7日 SASの基本的な使い方として基本統計量、ヒストグラムの作成
- 6月8日 相関
- 6月9日 回帰
- 6月10日 ロジスティック回帰分析
- 6月11日 分散分析
- 6月14日 多重比較検定
- 6月15日 遺伝率
- 6月16日 課題まとめ
- 6月17日 課題発表準備
- 6月18日 課題発表準備
- 6月19日 課題発表



SASをランさせるためのコントロールファイルの書き方は予めTAが用意した上で、1週間で測定記録したデータを統計処理することができるようになり、さらに、全員を5つの班編制し、得られたデータから課題を設定しパワーポイントを使って班毎に発表をしてもりました。

今年度の課題発表タイトルは以下の通りでした。

1 班

題名：4 週齢時以降の各週齢時間の増体量に及ぼす影響について

2 班

題名：マウスの成長に及ぼす一腹の兄弟数について

3 班

題名：2005 年～2010 年の年度別にみるマウスの成長の違いについて体重、体長、尾長を解析し年度ごとの集団に違いがあるか比較する

4 班

題名：ロジスティック曲線から見る給与する飼料内容のマウスへの影響

5 班

題名：5 年間でみる!♂と♀どっちが太りやすいか??

## 5. SASを使った学生の感想

実質 2 週間にわたる SAS の統計処理を主体とした実験に対する学生の感想は以下の通りでした。

### 1. 今回の実験、実習は勉強になりましたか？

	1	2	3	4	5
マウスの飼育・扱い	22	11	1	0	0
Excelの使い方	25	6	1	1	1
SASでの統計解析	27	7	0	0	0
テーマごとの課題演習	17	12	4	0	1
発表準備	7	19	4	3	1

1：とても勉強になった、2：勉強になった、3：どちらともいえない、

4：あまり勉強にならなかった、5：意味がなかった

### 2. 統計解析について各テーマはどれくらい理解できましたか？

	1	2	3	4	5
SASの使い方	1	26	5	2	0
相関	1	21	10	1	1
回帰・ロジスティック回帰	2	20	10	1	1
分散分析・多重比較検定	1	19	10	3	1
遺伝率	1	6	15	11	1

1：理論まで完璧に理解した、2：だいたい理解した、3：どちらとも言えない、

4：あまりよく理解できなかった、5：まったく理解できなかった

今回の学生実験では、8 割以上の学生が勉強になったと感じ、統計解析に関しても 7 割以上の学生が大体理解したとの感想でした。生物統計学の講義では計算の演習を行うことはしておらず、一方、学生実験では自ら得たデータを使って実際の統計処理を高度な統計処理も可能な SAS プログラムを利用できたことは学生にとって大変理解しやすかったと思われます。これを機会に学生が 4 年生、さらに大学院に進学後、得られた研究データについて SAS を使って発表することに馴染んでくれることを期待したいと思います。