

学習者参加型確率分布シミュレーション教材開発

木村 清

尚絅学院大学総合人間科学部

概要: 統計量の確率分布の概念の理解を促す目的で、表計算ソフト(Excel)と Web を用いた教材を作成した。これは、教室内の各学習者の Excel のワークシート上で得られたシミュレーションによる統計量をサーバーに送信する一方で、教師用パソコンから Web にアクセスすることでそれらのヒストグラムすなわち確率分布を表示することができるものである。実際に授業で使用したところ、学習者の参加意識も高まり有効な教材であることが確認できた。

キーワード: コンピュータ利用教育 シミュレーション 確率分布 学習者参加

Development of Learners' Participation Type Educational Materials using Probability Distribution Simulation

Kiyoshi Kimura

Faculty of Comprehensive Human Sciences, Shokei Gakuin University

Abstract: I developed education materials with the spreadsheet application (Excel) and Web-based system which aim at understanding of the concept of probability distribution. This system displays the histogram of statistics values which are calculated by each learners' Excel simulation and sent to the server. The result of questionnaire on the real classes suggests that this system helps students improve their sense of participation and get the concept of probability distribution.

1. はじめに

推測統計では、標本の値を使って計算された統計量を確率変数としてとらえ、その分布の特性、すなわち統計量の確率分布の性質を利用して種々の推定・検定を行う。このような統計的手法の学習の際には、確率変数や確率分布の概念を理解することが特に重要である。数学的扱いを専門としない学習者に対してこれらの理解を促すには、具体的な確率的現象面からアプローチする方法が求められる。そのためにコンピュータ・シミュレーションを利用することは一つの有効な手段であることはいうまでもない。

筆者はかつて表計算ソフト Excel のマクロを使用したシミュレーション教材をいくつか作成してきた。今回、それらを対面授業の際により有効に使用できるよう、学習者参加型の教材に発展させた。

本稿では、はじめに従来の教材について述べ、その問題点を指摘した上で、今回開発した教材

のねらいについて述べる。その後、開発したシステムの概要および、授業実施例とその後の学習者のアンケート結果について報告する。

2. 従来例と問題点

Excel のワークシート上で統計量の確率分布のシミュレーションを行う教材を提供することは、確率分布の概念の理解と、Excel 上での統計量の計算方法の理解を同時に促すことになり意義のあることと考えられる。

ここでは従来例として、Excel のマクロプログラムを用いたサイコロシミュレーション教材を例示し、その概要と問題点について述べる。

このシミュレーションでは、正常なサイコロを 60 回振って出た目の分布からカイ 2 乗の値(統計量の実現値)を計算する。(画面の様子は教材の改善例として示した図 5 とほぼ同一である)。動作の概略を以下に示す。

試行ボタンがクリックされると、マクロプログラムにより 60 回分の試行結果が標本データとして得られる。ワークシートにはあらかじめ

数式が設定されており、標本データが更新されると、1 から 6 の目の度数が集計・更新される、と同時にカイ 2 乗の値も更新される。すなわち試行ボタンをクリックするたびに(正常な)サイコロを 60 回振ったときの統計量(カイ 2 乗の実現値) が得られる。

この教材のねらいは、以下のようなものであった。

- ・カイ 2 乗の計算を表計算上で確認できること
- ・正常なサイコロの場合でも統計量の実現値にかなりのばらつきが生ずるが、大幅にずれることは滅多にない、という確率分布のイメージを学習者に持たせること
- ・正常かどうかわからないサイコロを振って、カイ 2 乗値が極端に大きな値になった場合には、正常なサイコロであるという仮説を捨てることもある、という仮説検定の基本的考え方に結びつけること

2.1 従来の授業実施例

昨年度までの実際の授業では、試行ごとに計算された統計量がどの辺の値をよく取るかを、学習者に何度も試行させて観察させた。その際特異な値が得られた学習者には、そのときの標本データを発表させるなどした。

またこのシミュレーションを体験させたあとで、カイ 2 乗検定の例題(正常なサイコロかを検定する例題)を行っていた。

2.2 従来例の問題点

本来であれば、統計量の分布(実験途中のヒストグラム)をリアルタイムで表示させることで、より理解が深まるのであるが、実際問題として一つの画面内に標本データ、表計算エリアの他に、計算された統計量のリストやヒストグラムを収めるのが困難であった。

また、仮に各自の手元でリアルタイムにヒストグラムを表示したとして、教師がスクリーンに説明用に表示するヒストグラムの形と大きく異なることもあり、学習者に余計な混乱を引き起こしかねない、という問題があった。加えて動的に変化するデータを用いて適切なヒストグラムを表示するマクロを盛り込むことで、教材(Excel のファイル)の保守性が低下してしまうことも懸念材料であった。

あるいは、学習者の手元に数多くの試行結果

を蓄積し、Excel の操作の演習を兼ねて各自にヒストグラムを書かせることも可能であった。しかし、Excel の操作に充分習熟していない学習者は、本来習得すべき確率分布の概念よりもヒストグラム作成の操作に気を取られてしまうこと、何より時間がかかってしまうことが問題であった。

3. 開発した教材の概要とねらい

前節で述べたねらいを達成するためには、シミュレーションによって得られた統計量のヒストグラムを提示し、その傾向と確率分布を示すことが不可欠である。そこで、**図1**に示すように、①個々の学習者が行ったシミュレーションの結果を②サーバーに送信し、③教師用 PC でヒストグラムを表示、教室前方のスクリーンに映し出す方式を考案し実現した。

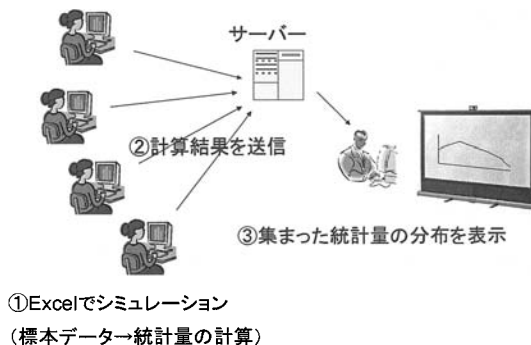


図1 全体構成と動作の流れ

これにより、次の効果が期待できる。

- ・教室前方のスクリーンに映し出すことで、授業時に、学習者の注意を一点に集めることができる
- ・学習者のシミュレーション実験への参加意識を高められる
- ・学習者個々の画面と教室前方のスクリーンとに提示する場所を分けることで、1)各学習者の場所で独立して生じている確率的現象と、2)それらを集めて観察することで一定の傾向が見いだせるという2つの側面を分かり易く提示することができる。

このように、実験への参加意識が高められた状況で、確率分布に関する基本概念を現象面から教室全体で共有することができるようになることが期待される。

4. 開発した教材の概要

図 2 にシステムのブロック図を示す。

このうち、標準データ発生マクロと統計量の計算ワークシート部分については従来教材例とほぼ同一のため、ここでの説明は省略する。

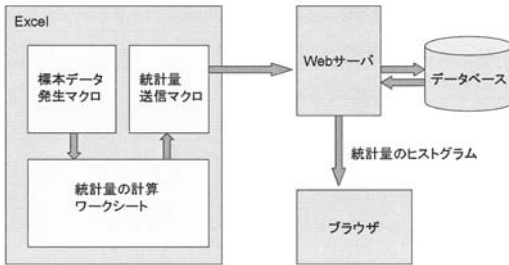


図 2 システムのブロック図

4.1 統計量送信マクロ

このマクロでは、ワークシートの所定のセルに計算された統計量を、POST メソッドを使用して Web サーバーに送信する。その際、以下の情報を付して送信する。

- 1) シミュレーションの種類を識別 ID
- 2) 授業クラスを識別するセッション名
- 3) 学習者を識別するユーザ名

ここで 1)の識別 ID は対応する統計量送信マクロの中で直接指定している。2)のセッション名は、本教材を複数の教室で同時に使用した際に、当該クラスの結果だけを抽出するためのものである。教材使用開始時に教員が任意のセッション名を示して、学習者にワークシートの所定のセルに入力させる。また 3)のユーザ名は、自分のデータが確かに送られていることを確認するためのものである。システム上、ユーザ管理は行っておらず、単にワークシートの所定のセルに入力された値を使用しているだけである。

4.2 Web とデータベース

個々の学習者の統計量送信マクロから送られたデータは、前述の 3 つの識別情報にタイムスタンプを付して順次データベースに格納される。Web サーバーには Apache、スクリプト言語は PHP、データベース・システムには PostgreSQL を使用している。

既存セッション

参加したセッション名をクリックすると、そこに集められたデータ、分布が表示されます。

(新規セッションでも、データが一件も送信されていない場合は以下のリストには表示されません)

No.	セッション名	最新データ日時	データ数
1	1108ef	2007-11-08 17:41:47	969
2	1108	2007-11-08 17:31:23	69
3	1108EF	2007-11-08 17:30:08	54
4	1180ef	2007-11-08 17:30:06	4
5	1108ab	2007-11-08 17:27:29	707
6	1108	2007-11-08 17:25:09	50
7	class1	2007-11-08 17:21:27	33
8	1108cd	2007-11-08 17:08:51	276

図 3 セッションリスト画面例

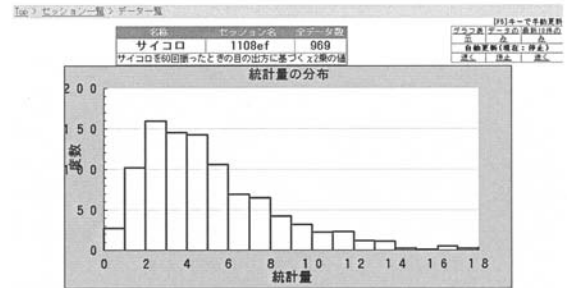


図 4 実際に得られたヒストグラムの例

ヒストグラムを描画する際は、所定のシミュレーション識別 ID とセッション名により目的の統計量の集団を抽出し、PHP のグラフ作成ライブラリ (Jpgraph) を用いて描画している。

なお、統計量のテーブルとは別に、シミュレーション識別 ID をキーとして、説明文とヒストグラム描画のパラメータ (級範囲と級幅) を格納するプロパティ・テーブルを用意している。

教師がブラウザで所定の URL にアクセスし、シミュレーションのタイプを選択した後、画面上的リスト (図 3) から自分の授業のセッション名をクリックするだけで、その時点のヒストグラムが表示される (図 4)。グラフ表示の自動更新の頻度は画面右上のメニューから選択できるようにしている。

4.3 シミュレーション画面例

図 5 にサイコロのカイ 2 乗検定を想定したシミュレーション画面の例を示す。

後の授業や演習でカイ 2 乗検定を行う場合の表計算部分を中央に配置し、マクロによる標準抽出 (試行) 部分を右下に配置している。以上

	A	B	C	D	E	F	G	H	I	J	K	L
1	セッション名	class1	←授業で教員から指示された記号・番号を入れます									
2	送信者名	guest	←誰が送信したかを示すものです。本名でなくてもかまいません									
3												
4												
5	サイコロを60回振るシミュレーションの集計結果											
6	サイコロの目	1	2	3	4	5	6	計				
7	観察値	4	9	14	13	8	12	60				
8	期待値	10	10	10	10	10	10	60				
9	期待との差	-6	-1	4	3	-2	2					
10	(期待との差)の2乗	36	1	16	9	4	4					
11	(期待との差)の2乗/期待	3.6	0.1	1.6	0.9	0.4	0.4	7	←統計量	サイコロを60回振る		
12										回	出た目	
13								$p=0.2206$	統計量を送信	1	5	
14									送信しました	2	3	
15										3	6	
16										4	6	
17									集計結果Web	5	1	
18										6	2	
19										7	2	
20										8	4	
21										9	4	
22										10	3	
23										11	3	

図5 サイコロシミュレーションの例(部分)

は従来の教材を引き継いでいる部分である。

今回新たに追加したのは、画面上では、左上のセッション名・ユーザ名入力部とデータ送信ボタン(表計算部分で得られた統計量の近くに配置)だけである。

5. 実施

開発した教材は表1に示した授業で用いた。

表1 授業実施

授業科目	「健康栄養情報演習」(1年後期3単位) 管理栄養士資格過程必修
授業形態	2コマ連続(講義+PC演習)
実施日	2007年11月8日(木)第4, 5時限
当日受講者	87名
講義(前半100分) (合同クラス)	・標本調査と確率 ・確率分布 ・正規分布表演習
演習(後半80分) (3クラス分割)	・二項分布シミュレーション ・標本平均シミュレーション ・カイ2乗分布シミュレーション

このうち、本教材を用いたのは上記演習最後の「カイ2乗分布シミュレーション」の部分で、所要時間は約15分であり、その流れは以下の通りであった。

- 1) Excelファイルのダウンロード
- 2) セキュリティレベルの確認
- 3) システムとシミュレーションの説明
- 4) セッションID、ユーザIDの入力
- 5) 初めの1回のデータ送信
- 6) 説明
- 7) 各自20回程度実行、送信
- 8) 結果の観察、説明

なお、今回の授業では前述のサイコロのシミュレーション(自由度5のカイ2乗分布)の他、

コイン投げのシミュレーション(自由度1のカイ2乗分布)も用意し、自由度によりヒストグラムの形が大きく異なることも確認した。

6. 評価と課題

6.1 授業後の評価

授業後に簡単なアンケート調査を行った。

まず、全体として「実習室で各自が確率分布のシミュレーションを体験したことは、確率分布に対する理解を深めるのに有効であったか」の問いに対しては以下のような回答が得られた。

そう思う	35	(40.2%)
ややそう思う	48	(55.2%)
あまりそうは思わない	2	(2.3%)
無回答	2	(2.3%)
総計	87[人]	(100.0%)

これは演習の時間全体を通しての(他の教材の使用も含んだ)回答であるが、全体の95%の学生は、シミュレーションが確率分布の理解に有効と感じたことがわかった。

次に本教材について尋ねた質問については図6のような結果が得られた。

特に評価が高いのは(1)(2)(5)の3項目であった。簡単でおもしろく、参加意識が持てたという評価は、本教材が授業の「つかみ」としても有効に機能したことを示している。

その他の項目の評価は前記3項目より相対的に若干低いものの、「ややそう思う」まで含

めてほぼ7～8割が肯定的な評価をしていることが分かった。

一方教員側の操作であるが、前述のように今回の授業実施に際しては3つの実習室に分かれて同時に行った。筆者以外の授業担当者2名に対しては、事前に教材のExcelファイルとA4版3ページの操作説明書をメールで送付しておいた。この教材の使い方について各担当者への口頭の説明や、合同のリハーサルは行わなかった。しかし実際、特に問題なく授業を進めることができた。

以上、学生側の評価、教師側の操作の簡便さなどを総合すると、本教材の当初のねらいはほぼ達成できたと考えられる。

6.2 拡張性について

本教材のシステムはブロック図(図2)および画面例(図5)に示したように、標本値から統計

量を計算する部分は通常の表計算(ワークシート)と何ら変わりはない。他の統計量のシミュレーションに拡張する場合は、その統計量の計算用ワークシートに、標本データ発生マクロと統計量送信マクロを付加すればよい。

一方、Web側の設定については、シミュレーションの種類に応じたプロパティ・テーブルにレコードを追加するだけで対応できる。

このような点で他の統計量への拡張は比較的容易であると考えられる。本稿執筆時点で以下の4つのシミュレーションを一つのファイルに収めたものを作成している。

- ・ 標本平均の分布
 - 正規分布
 - t分布
- ・ カイ2乗分布
 - 自由度1
 - 自由度5

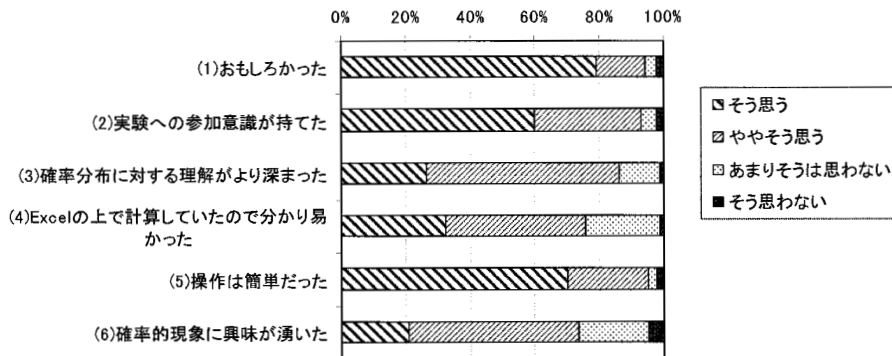


図6 本教材に対する学生の反応

6.3 インターネットでの利用

本教材は 1)教育現場でも普及している Excel を用いること、2)インターネットからも利用できる Web の 3 層システムを使用していること、3)教員側操作も簡便なことから、筆者の担当授業だけでなく、他校の授業でも利用可能であると考えられる。(当分の間、暫定的に以下の URL から利用可能としている

http://www.shokei.ac.jp/isc/~kimura/exp_p_dis/sc/expList.php)

6.4 セキュリティと保守性について

今後はサーバー回りのセキュリティについて検討する必要がある。ただし、機能やデータ構造が非常に単純であるため、文字数の制限など、比較的簡単な方法で対応することを考えている。

データベースの保守については、現在手作業

で行っているが、対面授業でのみ使用すると割り切ること、たとえば、授業日時から一定時間経過したセッションデータを自動削除するなどすることで、保守の手間を軽減することができると考えている。

7. まとめ

確率分布の概念形成に有効な学習者参加型の教材を作成し実際の授業で使用し、ほぼ所望の目的が達成できた。使用法も簡便であり、学習者の参加意識も高く、授業の「つかみ」としても有効に機能することが分かった。今後、他校からの利用の可能性を視野に入れ、セキュリティと保守性の改善を行い、カイ 2 乗以外の統計量への拡張を検討したい。