



質保証委員会

平成 24 年度(2012)活動報告書

文部科学省 大学改革推進等補助金
大学間連携共同教育推進事業 平成 24 年度採択
「データに基づく課題解決型人材育成に資する統計教育質保証」

まえがき

統計教育大学間連携ネットワーク質保証委員会は、連携の活動開始と同時に発足した各種委員会の一つである。質保証委員会のミッションは、外部評価委員会が定める指針に準拠して統計教育のカリキュラム編成上の参照基準の改訂作業を進め、統計教育の枠組みをカリキュラム策定委員会に提示することにある。

本報告書は、質保証委員会の平成 24 年度の活動をまとめたもので、以下の構成となっている。最初に、質保証委員会のミッションおよびそのメンバーを紹介し、平成 24 年度に開催された 4 回の議事録を収録した。これにより、質保証委員会での議論の内容が理解される。次に、委員会報告の一環として、質保証委員会のミッションである参照基準の初めの部分を資料として収録した。これは参照基準の第 1 版として平成 22 年に策定されたものであり、本委員会にて改訂を行うとするものであるが、参照基準のポリシー自体はそう変得る必要がないことが、委員会にて合意されている。

次に、質保証の議論を深める上で参考にした 4 編の海外の文献の日本語訳を掲載した。それらは掲載順に、英国高等教育質保証機構 (QAA) が出版した「数学・統計学・オペレーションズリサーチ分野の専門分野の評価基準書」、The American Statistician に掲載された論文「事例ベース ビジネス統計コースの展開」、2010 年 7 月にスペインのバルセロナで開催された EDULEARN10 での発表論文「大学院生の統計的思考の育成：統計的思考の学習環境モデルの評価」、および米国統計学会が公表した「統計教育における評価と指導方法に関するガイドライン」である。なお、これらの翻訳にはまだ不正確な部分が残されているので、参考にされる際にはその点に留意し、出来れば直接、原文に当たられたい。

日本における大学教育の質保証は、議論が始まったばかりという感がある。上記に示したような諸外国の例を参考にしながら、日本における統計教育の質保証の議論を深め、具体的な手順を示すことが本委員会に与えられた使命である。その実現のため、大学間連携ネットワークに設置された他の委員会との連絡を密にした上で、今後ともメンバー一同努力していく所存である。

平成 25 年 3 月

質保証委員会委員長
成蹊大学理工学部教授
岩崎 学

目 次

平成 24 年度質保証委員会報告	1
質保証委員会メンバー	2
質保証委員会議事録（第 1 回～第 4 回）	3
統計学分野の教育課程編成上の参照基準	7
数学・統計学・オペレーションズリサーチ分野の評価基準書	12
1. 序論	15
2. MSOR の特性と範囲	21
3. 知識、理解と技能	29
4. 指導、学習と評価	34
5. 評価基準書	37
付録 A, B	40
事例ベース ビジネス統計コースの展開	42
1. はじめに	42
2. 事例の使用手法：二つのシナリオ	44
3. 学生のためのヒント	49
4. 新インストラクターに役立つヒント	50
5. リソース	52
6. おわりに	55
大学院生の統計的思考の育成：統計的思考の学習環境モデルの評価	58
1. 導入	59
2. 方法論	61
3. 議論と結論	66
統計教育における評価と指導方法に関するガイドライン	68
序論	69
入門コースの学生の到達目標：統計学の教育を受ける意義	72
提案	74
付録 A	84
付録 B	94
付録 C	109
付録 D	112

平成 24 年度質保証委員会報告

質保証委員会委員長
成蹊大学理工学部教授
岩崎 学

統計教育大学間連携ネットワーク質保証委員会は、ネットワーク運営委員会の下部組織として、外部評価委員会およびカリキュラム策定委員会と共に連携の活動開始と同時に発足した。質保証委員会のミッションは、外部評価委員会が定める指針に準拠して統計教育のカリキュラム編成上の参照基準（以後、参照基準と呼ぶ）の改訂作業を進め、統計教育の枠組みをカリキュラム策定委員会に提示することにある。

また、各大学が実施する教育に対して、ファカルティデベロップメント (Faculty development = FD) 活動の活性化およびカリキュラムの認証・評価の枠組を検討し、さらに、統計検定との連携を具体化する作業として、ステークホルダーである統計関連学会と協力しながら、財団法人統計質保証推進協会が実施する統計検定を拡充した「統計教育達成度評価システム」の開発に着手することも視野に入れている。

質保証委員会が改訂する参照基準は、今後の我が国における統計教育の重要性を反映し、国際通用性を確保するためにも重要な基準となるものであり、この基準によって我が国の統計教育の質を高める基礎の確立が期待される。また、各大学に対する FD 活動の活性化は、統計学に関する教育効果を高める上で有効であることは疑いない。

さらに、統計検定を拡充した統計教育達成度評価システムが開発されれば、連携大学に限らず、広く我が国の統計教育の質を評価する客観的な方法が提供できることになり、いずれも教育に対する多大な効果が期待できる。

質保証委員会は、連携 8 大学の委員に加え、ステークホルダーのうち統計関連学会連合のメンバーである 6 学会（応用統計学会、日本計算機統計学会、日本計量生物学会、日本行動計量学会、日本統計学会、日本分類学会）が協力して設置した統計教育推進委員会、および日本統計学会の内部に設置している質保証委員会および統計教育委員会から推薦された委員から構成されている。

質保証委員会の平成 24 (2012) 年度の活動メンバーは次ページに記載の 22 名である（委員長は日本統計学会から推薦の岩崎学成蹊大学教授が第 1 回委員会に於いて、委員間の互選により選出された）。いずれの委員も任期は連携プロジェクト終了時までとされているが、各大学もしくは学会の事情により交代する場合には、交代した委員の任期は前任者の残余機関となる。

平成 24 年度の質保証委員会は 4 回開催された。次ページ後以降にそれらの各委員会の議事録を収載する。

質保証委員会メンバー

<連携大学からの推薦者>

1. 東京大学：駒木文保（大学院情報理工学系研究科・教授）
2. 大阪大学：竹内恵行（大学院経済学研究科・准教授）
3. 総合研究大学院大学：三分一史和（複合科学研究科統計科学専攻・准教授）
4. 青山学院大学：荒木万寿夫（経営学部・教授）
5. 多摩大学：大森拓哉（経営情報学部・教授）
6. 立教大学：小野寺剛（社会情報教育研究センター・助教）
7. 早稲田大学：西郷 浩（政治経済学術院・教授）
8. 同志社大学：鄭 躍軍（文化情報学部・教授）

<統計関連学会連合からの推薦者>

1. 応用統計学会：田栗正章（中央大学大学院理工学研究科データ科学副専攻・客員教授）
2. 応用統計学会：福井武弘（総務省統計研修所・所長）
3. 日本計算機統計学会：水田正弘（北海道大学情報基盤センター・教授）
4. 日本計量生物学会：松山 裕（東京大学大学院医学系研究科・准教授）
5. 日本行動計量学会：植野真臣（電気通信大学大学院情報システム学研究科・准教授）
6. 日本計量生物学会：浜田知久馬（東京理科大学工学部経営工学科・教授）
7. 日本計量生物学会：三中信宏（農業環境技術研究所生態系計測研究領域・上席研究員）
8. 日本計量生物学会：岸野洋久（東京大学大学院農学研究科・教授）
9. 日本統計学会：岩崎 学（成蹊大学理工学部情報科学科・教授）
10. 日本分類学会：竹内光悦（実践女子大学人間社会学部・准教授）

<日本統計学会統計教育委員会からの推薦者>

1. 日本統計学会：椿 広計（情報・システム研究機構統計数理研究所・副所長）
2. 日本統計学会：櫻井尚子（東京情報大学総合情報学部環境情報学科・教授）
3. 日本統計学会：和泉志津恵（大分大学工学部・准教授）

第1回統計教育大学間連携ネットワーク質保証委員会 議事録

日時：2012年11月17日（土）15:00～20:00

場所：青山学院大学17号館6階603室

出席：田栗正章，福井武弘，水田正弘，松山 裕，岩崎 学，竹内光悦，和泉志津恵，駒木文保，竹内恵行，三分一史和，荒木万寿夫，小野寺剛，西郷 浩，宿久 洋（鄭躍軍委員代理）。

本委員会に先立つ14:00-15:00に同所にて，統計教育大学間連携ネットワーク質保証委員会およびカリキュラム検討委員会の合同委員会が行われ，成蹊大学の岩崎学氏が出席委員の互選により質保証委員会の委員長に選出された。

議 事

1. 質保証委員会の活動について

岩崎委員長より，質保証委員会のメンバーが，統計関連学会統計教育推進委員会，連携8大学および日本統計学会統計教育委員会のそれぞれから選出された委員から組織されているとの紹介があった後，統計教育大学間連携ネットワーク（大学間連携共同教育推進事業）における質保証委員会の活動内容について説明があった。

続いて田栗委員より，統計関連学会連合統計教育推進委員会の従来の活動経緯および課題について説明があった。議論の後，質保証委員会のミッションとして平成22年8月に公表された「統計学分野の教育課程編成上の参照基準」（以下「参照基準」と略）の改訂作業を当面の活動の中心に据えることが確認された。

2. ワーキンググループ (WG) の設置

今回はWGは設置せず，今後の活動の中で必要に応じ設置を検討することとした。

3. 今後の取組方針

まず参照基準の改訂に向けて検討を進めることとし，各委員が次回委員会までに自分の専門分野とその周辺分野について現行参照基準を読み込み，改訂の方向性についての検討を行うことが了承された。

4. その他

岩崎委員長より，委員会の運営に関する事務手続等について説明があった。

5. 次回開催

2012年12月15日（土）10:00 - 12:00に第2回委員会を青山学院大学において開催することとした。

なお，本委員会終了後20:00まで，本委員会の委員ならびにカリキュラム策定委員会のワーキンググループの委員間で，今後の委員会活動についての話し合いが行われた。

第2回統計教育大学間連携ネットワーク質保証委員会 議事録

日時：2012年12月15日（土）10:00～12:00

場所：青山学院大学 17号館 6階 602教室

出席：田栗正章，水田正弘，松山 裕，岩崎 学，櫻井尚子，駒木文保，竹内恵行，三分一史和，荒木万寿夫，小野寺剛，西郷 浩，鄭 躍軍。

（保科架風：書記）

配布資料：第1回統計教育大学間連携ネットワーク質保証委員会議事録。

議 事

1. 第1回議事録の確認

資料の前回議事録の確認を行った。統計関連学会連合統計教育推進委員会委員長に田栗委員が就任したことが岩崎委員長から報告された。

2. 質保証委員会の活動

2.1 委員の追加に関する合意

日本計量生物学会から浜田知久馬（東京理科大学），岸野洋久（東京大学），三中信宏（農業環境技術研究所）の3名を委員に推薦したいとの要望があり，これを了承し，代表校である青山学院大学に手続きの確認をとることとなった。また，今後委員を追加する際には，6学会・8大学からの推薦，または委員長推薦によって行うこととなった。任期については岩崎委員長が改めて確認をとり，次回委員会にて報告することとなった。

2.2 参照基準改訂の具体的作業

スケジュールの問題から，参照基準は冊子だけでなく Web 掲載の形での改訂も選択肢に入れて作業を進めることとなった。

現行参照基準に対して各委員から意見が出され，理念，目標，具体的な教育内容・評価方法の枠組みの中で必要最小限／発展的な内容を書き分けていき，その作業の過程で分野分けについて検討していくという方向で改訂を行うこととなった。また，各委員の専門分野に則って担当を決め，次回委員会までに参照基準に含む内容や周辺分野の分野分けなどについて検討を行うこととなった。

2.3 その他の案件

大学以前の統計教育を大学・社会人にいかに繋げていくかについて考えていくことを質保証委員会の活動に含め，今後議論していくことが了承された。

議事3. 次回開催について

2013年1月12日（土）に第3回委員会を開くこととなった。

第3回統計教育大学間連携ネットワーク質保証委員会 議事録

日時：2013年1月12日（土）15:00～20:00

場所：青山学院大学 青山キャンパス 総研ビル 3階第11会議室

出席：田栗正章，福井武弘，水田正弘，植野真臣，岩崎 学，竹内光悦，櫻井尚子，駒木文保，西郷 浩，鄭 躍軍，岸野洋久。

（保科架風；書記）

配布資料：開催通知，第2回委員会議事録（掲載版/詳細版），参照基準改定案（1，2章）
＜田栗＞，参照基準（第1版）1，2，3章，統計検定レベル表1～4級，参照基準（経営学）メモ＜竹内恵行＞，AP Statistics 目次，RSS シラバス (Higher Certificate Module 1～6)。

議 事

1. 第2回議事録の確認

資料の前回議事録の確認を行った。また，日本計量生物学会から推薦された3名の委員の追加が運営委員会で認められたこと，さらに各委員の任期が5年間であることが岩崎委員長から報告され，任期途中で委員を交代する際は交代する委員の推薦母体から新たな人を推薦するということが確認された。

2. 運営委員会，その他の委員会からの報告と配布資料の説明

岩崎委員長によって配布資料の説明が行われた。資料に関連し，アメリカの高校生向けの大学単位先取り制度である AP statistics の内容が2014年から高校で必修となることが報告された。

3. 参照基準の改訂の具体的作業

田栗委員から参照基準の1，2章の改訂案についての説明が行われた。各委員から大学基礎科目の改訂案が必要との意見が出され，まず岩崎委員長，田栗委員，竹内光悦委員，櫻井委員が大学基礎科目の参照基準案を早期に作成し，また，次回委員会までに各分野における具体的な作業と分野の分類の検討を行うこととなった。諸外国の情報が欲しいとの意見が出され，カリキュラム策定委員会で現在まとめている諸外国での統計教育のレビューを参考にすることとなった。

4. 次回開催について

2013年3月のカリキュラム策定委員会の日程を考慮に入れながら日程を調整することとなった。

第4回統計教育大学間連携ネットワーク質保証委員会 議事録

日時：2013年3月16日（土）15:00～20:00

場所：青山学院大学 青山キャンパス 17号館 3階 17302教室

出席：田栗正章，福井武弘，水田正弘，松山 裕，浜田知久馬，岸野洋久，三中信宏，岩崎学，竹内光悦，駒木文保，竹内恵行，三分一史和，荒木万寿夫，大森拓哉，小野寺剛，西郷 浩，鄭 躍軍
（保科架風；書記）

欠席：植野真臣，椿 広計，櫻井尚子，和泉志津恵

配布資料：第3回委員会議事録（掲載版/詳細版），質保証委員会打ち合わせ資料，参照基準改定案（1，2章）＜田栗＞

（注）15:00～16:00は質保証委員会・カリキュラム策定委員会の合同委員会として開催された。

議 事

1. 第3回議事録の確認及び質保証委員会打ち合わせ内容の確認

資料の前回議事録の確認を行った。また、3月6日の質保証委員会打ち合わせの結果が岩崎委員長より報告された。

2. 参照基準の分野分け

質保証委員会打ち合わせでの検討結果を受け、分野分けの議論を行った結果、参照基準の第2版では「人文科学」、「法学・政治学（政策科学）」、「社会学」、「経済学」、「経営学」、「数理科学」、「情報科学」、「総合理工学」、「生命科学」、医歯薬学、「生活・健康科学」の12分野とすることとなった。

3. 参照基準の3.1節の討議

本委員会に先立って開催された合同委員会の資料中の参照基準第2版（案）3.1節を参照しながら内容の確認を行い、基礎的内容の範囲等について討議をし、さらに文章の検討を継続することとなった。

4. 参照基準改定の方針

タイムスケジュールとして、ゴールデンウィーク明けには参照基準第2版の第1稿を作り、夏までには完成を目指すということが確認された。また、各分野の担当の確認が行われた。

5. その他

2013年度は参照基準第2版の完成を質保証委員会の目標とすることが確認された。次回開催を含め2013年度のスケジュールについては4月の運営委員会を受けて決定する。

統計学分野の教育課程編成上の参照基準

質保証委員会の改訂対象となる参照基準は、平成 22 年 8 月 20 日に、統計関連学会連合理事会ならびに統計教育推進委員会の名で公表された。ここではその参照基準の要約、第 1 章及び第 2 章を再掲する。

質保証委員会では、第 3 章以降の各分野の切り分けを見直し、新たに必要と思われる分野を追加すると共に、各分野における基準のより一層の明確化を企図して改訂を進めることになっている。具体的には、第 3.1 節の大学基礎課程における参照基準を最初に見直し、それに基づいた形で各専門分野における統計教育の参照基準を策定する。

要 約

21 世紀の知識創造化社会において、統計学は多くの分野で役立つと認識されるようになった。特に、コンピュータの急速な進展に伴う技術革新が目覚ましい社会においては、複雑で大規模なデータや非線形でダイナミックな現象に対処可能な、新たな統計学を創成する必要がある。

現代社会におけるこれらの重要な課題に対処するため、諸外国においては研究者育成のための統計教育システムはもちろんのこと、初等教育段階からの広汎な統計リテラシー涵養のための教育システムも整備されつつある。

これに対して、現時点での日本の統計教育のシステムは、極めて脆弱と言わざるをえない。しかし、平成 20～21 年に公表された新学習指導要領では、小学校 3 年生の「算数」から高等学校の「数学 B」までの、各学年で履修する“統計”の内容は、履修の順序や“確率”との関係も含めて、かなりよく考えて策定されている。したがって、それとの円滑な接続を図るためにも、高等教育段階における統計教育について検討を行い、統計関連学会連合として一定の基準を示すことは、喫緊の課題である。そこでここでは、大学の学士課程における統計教育について考えるが、初等・中等教育や大学院教育との接続についても考慮しながら検討を進める。

近年の、不確定性を伴う複雑で大規模な情報に対処する素養を育成するためには、教養としての統計リテラシー教育、将来統計手法を利用する学生のための統計教育、将来統計を職業とする学生のための統計教育等について考える必要がある。そこで、これらの教育における「統計学分野の教育課程編成上の参照基準」（以下、「参照基準」）を策定する。

さて、ここで考える「参照基準」とは、“各大学の教育課程編成に当たって、学生に求める価値観・倫理観や基本的な素養（知識・能力・スキル）を教育目標として定め、そのために必要な学習内容・学習方法を具体的に検討する際に参照されるべき基準”と考える。その際、統計学分野に関連する具体的な職業生活を想定し、それを支える基礎を如何に培うかという観点も視野に入れて検討を行う。

統計学は分野横断的な性質を有しており、近年では程度の差こそあれ、どの学術分野に

においても必須の学問になりつつある。また、社会においても、統計学はさまざまな分野で貢献している。このような状況を考慮し、「参照基準」の策定に際しては、大学基礎科目や統計学を専門とする課程に対する基準だけでなく、統計学に関わりをもつ分野における「参照基準」も策定する。

また、参照基準の具体的な構成要素については、統計教育の分野の広汎性や、現行の統計教育システムの脆弱性等を考慮し、ある程度具体的に示すことが必要かつ有益と考えた。具体的には、(i) 当該分野の理念、(ii) 到達目標（身に付けるべき知識・能力・スキル）、(iii) 目標を達成するための教育内容・評価方法の例、の3項目とする。

本「参照基準」は、次のような構成になっている。まず第1節では、策定に際しての基本的考え方をまとめた。第2節では、参照基準の基礎となる“統計学の考え方・ポイント”について記述した。すなわち、“統計学の本質は、帰納的推論の中に演繹的論理の過程を導入することにより科学的な結論を導く点にある”ことを述べた。そして、演繹的論理のみの教育、データの記述の域を出ない教育では十分ではないことを述べた。最後の第3節は、統計学に関わりをもつ8つの分野に焦点をあて、上述した3項目からなる「参照基準」を与えた。ここで、8つの分野とは、1.大学基礎科目分野、2.心理学・教育学分野、3.経済学分野、4.社会学分野、5.経営学分野、6.数理科学分野、7.工学分野、8.医学・薬学分野、である。

この「参照基準」は、各大学・大学院の教育課程編成の際や、社会での各種教育プログラムのデザイン・実施・評価の際、また統計の各分野の学位が表象する性質・基準について知りたいと思った際等々に、有効に活用できるものとする。

1. 「統計学分野の教育課程編成上の参照基準」策定に際しての基本的考え方

21世紀の知識創造化社会において、統計学は、データの収集・情報の抽出・帰納的推論・科学的決定等を必要とするあらゆる分野で役立つと認識されるようになり、急速な発展を遂げつつある。特に、20世紀後半からのコンピュータの急速な進展に伴う技術革新が目覚ましい社会においては、ゲノムデータ・画像データ・自然／人工災害データのような、複雑で大規模なデータを解析するための新たな統計学を創成する必要がある。さらに、病気の診断／予後予測・金融工学等の問題に対しては、非線形でダイナミックな現象に対処可能な方法を開発する必要もある。

これらの問題は現代社会における重要な課題であり、諸外国においては活発に研究が行われている。また、そのために必要となる研究者育成のための統計教育システムはもちろんのこと、初等教育段階からの統計リテラシー涵養のための教育システムも整備されつつある。すなわち、多くの諸外国においては、社会における統計的考え方・公的統計の重要性や、統計分析の必要性が広く認知されている。

これに対して日本では、生命科学や経営学等の分野において顕著なように、現時点での

統計教育のシステムは極めて脆弱と言わざるをえない。このような状況を打開するためには、初等教育から高等教育に至るまでの統計教育を一貫して考え、そのために必要な教育システムを整備することが喫緊の課題である。ここでは、大学の学士課程における統計教育について考えるが、初等・中等教育や大学院教育との接続についても考慮しながら検討を進める。

さて、ここで言う「統計学分野の教育課程編成上の参照基準」（以下、「参照基準」）とは、“各大学の教育課程編成に当たって、学生に求める価値観・倫理観や基本的な素養（知識・能力・スキル）を教育目標として定め、そのために必要な学習内容・学習方法を具体的に検討する際に参照されるべき基準”と考える。参照基準の具体的な構成要素については下で述べるが、“基本的な素養”については、統計学分野に関連する具体的な職業生活を想定しながら、職業人として具備すべき専門知識や倫理等も含めて、長期にわたる職業生活を支える基礎を如何に培うかという観点も視野に入れて、参照基準の策定を行った。

このような基本的な考え方に基づいて策定された「参照基準」は、各大学・大学院の教育課程編成の際に参照してもらうことはもちろんであるが、社会における各種の教育プログラムのデザインや、その実施・評価に関わる人々の役に立てってもらうことも企図している。また、これから大学で学ぼうとする人々や、卒業生を雇用しようとする人々が、専門分野の学位が表象する性質と基準について知りたいと思った場合にも、役に立つと思われる。しかしこれは、教育課程編成における各大学の自主性・自立性を縛るものではなく、各大学はそれぞれの考えに基づいて取捨選択を行うことが重要である。

「参照基準」の策定は、基本的には、以上のような考え方に基づいて進められた。しかし、統計学分野に関しては、他の多くの学術分野とは多少異なる要素を考慮する必要がある。周知のように、統計学は分野横断的な性質を有しており、特に近年のグローバル化社会・情報化社会においては、程度の差こそあれ、どの学術分野においても必須の学問分野になりつつある。さらにまた、社会の中においても、統計学はさまざまな分野で貢献している。例えば、社会調査士の資格試験においては、統計学は必須の科目であり、また製薬会社には、**professional biostatistician** を置くことが要請されている。

このような状況に鑑みた場合、統計学分野の「参照基準」は、大学学士課程における教育課程編成の範囲内で考えるだけでは不十分であり、上述したように大学院や社会との接続を念頭においた「参照基準」を策定する必要がある。また、学士段階における教育課程編成のための「参照基準」の策定についても、大学基礎科目や統計学を専門とする課程に対する基準だけでなく、統計学に関わりをもつ分野における「参照基準」を策定する必要がある。

以上をふまえて、第3節では、大学基礎科目としての統計教育に加えて、統計学とかなり深い関わりをもつと考えられる人文科学、社会科学、自然科学の各分野における教育課程編成のための「参照基準」を策定した。具体的には、次の分野についての基準を策定したが、これ以外にも、例えば理学一般、環境科学、農学のように、統計学と関係が深い重

要な分野もある。

1. 大学基礎科目分野
2. 心理学・教育学分野
3. 経済学分野
4. 社会学分野
5. 経営学分野
6. 数理科学分野
7. 工学分野
8. 医学・薬学分野

また、参照基準の具体的な構成要素については、以下のような項目とした。ここで、統計教育の分野の広汎性や、現行の統計教育システムの脆弱性等を考慮した場合、学習内容はある程度具体的に示すことが必要かつ有益と考えたため、かなり具体的な例示を取り入れている。

- (i) 当該分野の理念
- (ii) 到達目標（身に付けるべき知識・能力・スキル）
- (iii) 目標を達成するための教育内容・評価方法の例

なお、統計調査や実験により収集された情報の保護についての倫理規定（特に人間を対象として取得したデータの利用規程等）、これらの情報の二次的利用のためのデータ提供についての倫理規定、統計法の統計情報保護に関する規定等については、それぞれの分野において、適切に教育される必要がある。

2. 統計学の様々な分野における参照基準の基礎となる“統計学の考え方・ポイント”

現代社会においては、統計的なものの考え方や統計分析の手法が極めて重要となっている。すなわち、近年の情報化社会においては、我々はさまざまな不確実性に直面するため、多くの統計的判断を行うことが必要となる。つまり、我々はさまざまな場面において、リスクを最小にする迅速な決断を迫られることが少なくない。これに対処するためには、「読み、書き、ソロバン（計算）」に加えて、一人一人が統計的な考え方（統計リテラシー）をもつことが不可欠である。

このような素養を涵養するためには、初等・中等教育段階から現実のデータに日常的に接し、不確実性の概念について繰り返し訓練することによってはじめてその感覚が養えるものである。すなわち、実際のデータを素材に、教科横断的にそれを十分活用するような教育が行われる必要がある。平成21年3月に公表された新たな高等学校学習指導要領においては、必修科目の「数学Ⅰ」に“データの分析”という項目が規定されている。また、この準備のために、小学校では“数量関係”の領域の中で第3学年から、また中学校では“資料の活用”の領域の中で“統計”の学習を行うことが、それぞれの学習指導要領に規定されている。

このように、今回公表された新学習指導要領では、小学校3年生の「算数」から高等学校の「数学B」までの、各学年で履修する“統計”の内容は、履修の順序や“確率”との関係も含めて、かなりよく考えて策定されている。これは、“統計的素養”が現代に生きる者にとって必須のリテラシーであることが認識されつつあることの現れと考えられよう。

では、高等教育段階における統計教育はどうあるべきだろうか。近年の複雑な大規模情報に対処する素養を育成するためには、大学基礎教育としての統計リテラシー教育、将来統計手法を利用する学生のための統計教育、将来統計を職業とする学生のための統計教育等について考える必要がある。これらはいずれも大学・大学院で行われるべき教育である。現状では、自然科学や工学をはじめ人文科学・社会科学等の研究者が、必要に応じて教育を行っているが、その内容は多様で、必ずしも十分な統計教育が実施されているとはいえない。この解決のためには、統計学の専門家が、それぞれの分野の研究者と連携して教育にあたるのが重要である。

さて、統計学の考え方の大きな特徴は、その科学的推論の第一段階において、帰納的推論を行うことにある。すなわち、与えられたデータに基づいて、仮説やモデルのいくつかを選び出す規則を作り出すことである。さらに、そのような規則によって特定の仮説が選ばれたときの不確実性の程度を計算し、誤った決定の割合または誤りによる損失を最小にするような規則を見つけ出す（不確実性の数量化）。この後者の過程、すなわち問題を最適な決定を行う問題として定式化してしまった段階からは、演繹的推論に基づいて確率計算や数理的な解析を行うことになる。

このように、統計学の本質は、帰納的推論の中に演繹的論理の過程を導入することにより科学的な結論を導く点にある。しかるに現在の大学における統計教育では、演繹的論理のみの教育、またはデータの記述の域を出ない教育しか行われていない例が少なくなく、統計学の本質をよく理解した教育・研究者による教育が必要とされている。

以上、大学の学士課程における教育課程編成のための「参照基準」策定の基礎となる“統計学の考え方・ポイント”について述べてきたが、最後に社会における統計教育について、簡単にふれておく。まず、前述した社会に対する統計リテラシー教育のためには、新聞・テレビ等のマスコミにより伝達される統計情報を正しく解釈し、それを偏見なしで紹介できる記者・科学コミュニケーター等の育成が必要であり、大学教育においてもこのことを意識すべきである。次に、もう1つの社会に対する広い意味での統計教育は、産官学の全分野の研究者との共同研究・技術開発である。第1節の冒頭で述べたように、統計学はかつてないほどの広がりを見せており、社会のあらゆる分野において統計学の重要性は今後ますます増大すると考えられるが、このための方法論を提供することは統計学の使命である。

数学・統計学・オペレーションズリサーチ*

Subject benchmark statement: Mathematics, statistics and operational research

まえがき

専門分野における評価基準書 subject benchmark statements（以降，単に「評価基準書」という）は，大学などの学術機関が，特定の学問分野や領域における教育プログラムの性格と特性を説明する手段を提供している．それはまた，学習者が学位取得などの最終評価に値するような能力と達成度に関する標準的なレベルに関する基準をも与えるものである．

評価基準書は，同時期に策定された他の基準書とともに，優等学士号⁽¹⁾ (bachelor's degree with honours) の要件についても言及している．また，評価基準書の中には，修士号授与に関する指針を提示しているものもある．

評価基準書は，様々な目的に使用される．たとえば，大学などの高等教育機関 (higher education institutes = HEIs) が，ある学問分野の新規のプログラムを企画し展開する際の重要な外部参考資料ともなる．評価基準書は，プログラムにおける学習成果と達成度を明確に規定するための基本的なガイダンスを提供するが，その学問分野における詳細なカリキュラムを指定するものではない．

評価基準書は，高等教育機関がその内部で教育の質の保証を追求する際の支援も行う．基準書の利用により，ある特定のプログラムで指定される学習成果を，基本的に合意される基準に照らし合わせて審査，評価することができる．それはまた，プログラムを新規に作成したりあるいは既存のプログラムを改訂したりする際に，それらが柔軟でかつ革新的であるための議論の刺激にもなる．高等教育機関における教育プログラムの企画，遂行ないしは再評価の支援のために，この基準書の使用により，教育水準とその質に対し，制度上負うべき責任を強調する動きを支援するものともなる．

評価基準書は，これから学習しようとする学生や，彼らを雇用しようとする雇用者が，その学問分野では，どのような教育がなされ，どのような能力の獲得がなされるのかという情報を得るために用いられることもあろう．

この基準書で示された内容と，それぞれの学問分野の教育機関や専門家により個別に作られた基準との間の関係の吟味は，個々の高等教育機関がその内容を詳細に考える上で重要な手掛かりとなる．

本評価基準書は，2002年に作成された基準書の改訂版である．これは，本年出版された全

* JINSE の責任において翻訳した．英語版原典の翻訳・転載については，The Quality Assurance Agency for Higher Education より許可を得ている．© The Quality Assurance Agency for Higher Education 2007

学問分野における評価基準書の一部として、高等教育質保証機構 (QAA) が改訂作業を実施した。その学問分野のために選ばれた専門家グループにより、当初の基準書の内容の精査と改訂作業が行われた。ここで提示した改訂評価基準書は、多くの専門家と利害関係者グループによる十分な協議を経たものである。

QAA は、この評価基準書、ならびに関連した学問分野におけるその分野の評価基準書を出版、配布している。

障害者平等義務法 (disability equality duty = DED) が 2006 年 12 月 4 日に施行された。

DED⁽²⁾ は、高等教育機関を含む国家諸機関に、障害者の雇用機会均等などの問題への先駆的な取り組みを要請している。その義務は、障害者差別禁止法の個人の権利を補っており、総じて障害者のための公共サービスや行政の成果の向上を目的としている。高等教育機関はそのような義務が確実に実行する責任を担っている。

人権平等委員会⁽³⁾ は、高等教育機関の義務の実施準備のための手引き⁽⁴⁾を発刊し、義務の推進方法の説明に役立つガイダンスを提供している。評価基準書がその一部であるところの学術的な基礎基盤⁽⁵⁾の構成を行う際には、高等教育機関はこの手引書を参照することが奨励されている。

評価基準書に加えて高等教育機関を支援するための追加情報は、16 歳以降の教育と関連サービスの提供者のための実施規約 (改訂版)⁽⁶⁾ と、高等教育の平等性と多様性の促進のために確立された平等促進ユニット (Equality Challenge Unit)⁽⁷⁾の中にも見ることができる。

(1) これはスコットランド単位と資格制度 (level 10)、ウェールズ単位資格 (level 6) と同等のものである。

(2) イングランド、スコットランド、ウェールズにおける

(3) 2007 年 10 月 1 日に機会平等委員会、人種差別委員会、障害者権利委員会が合併し、新たな人権平等委員会が発足した。

(4) 高等教育と障害者平等義務手引書で、イングランド、スコットランド、ウェールズにおける学長、副総長、理事会、高等教育大学の部長と HEIs のための手引き書。

www.equalityhumanrights.co/en/forbusinessesandorganisation/publicauthorities/disabilityequality/pages/disability.aspx 参照

(5) 学術的基礎構造とそこにおける評価基準書の役割

www.aqq.ac.uk/academicinfrastructure 参照

(6) 16 歳以降の教育と関連するサービスの提供者のための実施規約 (改訂版)

www.equalityhumanrights.com/en/publicationsandresources/Disability/Pades/Education.aspx 参照

(7) 平等課題ユニット www.ecu.ac.uk 参照

序文

2002年に、高等教育質保証機構(QAA)により、数学・統計学・オペレーションズリサーチ(Mathematics, Statistics and Operational Research =MSOR)分野の学部学生のための評価基準書が出版された。2006年にその改訂に着手するため、QAAは数学・統計学・オペレーションズリサーチの高等教育学術センターに、その分野の専門家グループの選出を要請した。

初版の科目評価基準書は、広くMSOR分野に好意的に受け入れられた。そこでは、包括的で実施可能な方法で、成績優秀者レベルのMSORプログラムが記述されることが認められた。この評価基準書はMSOR分野の関係者およびそれ以外の基準書の使用者に役立つものであった。改訂に先立ってQAAが行った予備調査では、初版の基準書はあまり改訂される必要はないであろうというのが大方の一致した見解であった。

評価基準書改訂グループはこの見解に基づき、2002年の初版にほぼ従って基準書の改訂に着手したが、実際のところほとんど改訂する必要を認めなかった。行われた変更は、記述内容のさらなる明確化と、実情に合わせた形での何ヶ所かの部分の内容の最新化であった。

2002年の初版では、基準書の守備範囲を数学の修士(MMath)相当の教育プログラムには広げなかった。改訂グループは、その審議の時点では基準書をこれらのプログラムを含むところまで拡張する価値があると認識はしたが、その拡張に関し、文案を起草し、公に意見を求めることはかなりの作業規模となることも理解した。したがって、本基準書の出版の遅れがないように、これまでと同様の範囲にとどめた形で基準書の改訂を行い、MMathタイプの教育プログラムを含ませるために、付録の形で現在の基準書に付すことにした。改訂グループは2008年にこの作業が完成するものと予測した。

改訂グループは、初版の評価基準書作成グループの卓越した作業、ならびにQAAの改訂作業への貢献に対し、心よりの謝意を表す。

2007年6月

1. 序論

1.1 数学・統計学：オペレーションズリサーチ (MSOR) の学問分野では、各学問の守備範囲が非常に広いことから、それぞれに共通するプログラムを探すことはほとんど不可能である。このグループの中に含まれている教育プログラムの多くは、グループ内における科目群の間よりも、それ以外の学問分野における教育プログラムと共有している部分が多い。統計学の解釈面を重視したプログラムの中におけるよりも、物理学や工学の方が、より多く数学的な内容を含んでいたりもするし、ビジネスの分野のほうが数学よりもよりオペレーションズリサーチの内容を含んでもいよう。

1.2 それにもかかわらず、統計学とオペレーションズリサーチは、従来から数理科学の一員とみなされ、プログラム間では重ならないとしても、個々の科目内容の間ではかなりの重複がある。この3つの学問分野は歴史的には結びついており、同じ学部内で教えられていることも多い。

1.3 これらの学問分野を一つの評価基準書で扱うことは可能であるかもしれないが、本書の使用者は、この領域に配置されるプログラム間にはきわめて大きな多様性があることを念頭に置く必要がある。さらに、この基準書の領域に明らかに入ると考えられるプログラムが数多くあるばかりでなく、部分的にこの範囲に入ると思われるプログラムもまた多くある。この基準書は適切な指標を提示するものではあるが、重要なことは、どの基準をどの特定プログラムに適用するかを決定するのは、常に使用者である個々の教育機関であるということである。

1.4 プログラムが焦点をあてている学問分野に基づく分類に加え、プログラムのスタイルとして考えられる重要な分類もある。ある教育プログラムは、その中で扱っている内容の背後にある理論的な側面を重要視し、問題解決のための抽象的かつ一般的な方法の導出や確立に重きを置くこともある。このようなプログラムでは、主として方法論の教授と学習が主たる内容を占め、一般的な理論や手法が多くの分野に応用を持つという観点からの教育がなされる。本基準書では、これを便宜上「理論ベースのプログラム」と呼ぶ。他のプログラムでは、数学的な結果はそれとして受け入れ、その性質と特徴および限界を知った上で、実際問題への適用に重きを置く場合もある。実際問題への応用を主眼とするという意味で、ここではこれを便宜上「実用ベースのプログラム」と呼ぶ。

1.5 完全に理論ベースのみ、あるいは実用ベースのみといった教育プログラムはほとんどないであろう。大部分のプログラムは、両者のアプローチの要素を持ち合わせており、両極端のプログラムの間にわたる広範囲のプログラムのスペクトラムのいずれかに位置する

ものとなる。プログラムの特性は必ずしもそのタイトルに反映されているとは限らず、プログラムの目的と目標こそが、スペクトラムの中でその位置を明らかにさせるのである。重要なことは、現存する全てのタイプのプログラムが、それぞれ違った形で教育上の価値を有するということである。このことは、この科目分野におけるプログラムの多様性を示す一面であろう。

1.6 'MSOR' という用語は、本基準書では特定の学問領域として使われる。この語は標準的な頭字語ではなく、ここでの目的にのみ使われることに注意する。つまりそれは、この基準書でカバーされる MSOR の分野を意味し、特段他の意味はない。「数理科学」という語は様々なタイプのプログラムのタイトルに現れ、なじみの深いものであることから、本領域を数理科学と表現すべきという考え方もあろうが、その語は、たとえば OR は理論物理よりも数理的なのかといった新たな問題も引き起こす恐れもあるため、ここでは特に MSOR なる語を用いている。

1.7 MSOR は、いろいろな意味で特別な性質をもつ学問分野である。学習および教授という観点からみると、この分野の特徴はまず累積的であるという点に現れる。またそれは長い歴史を持つある意味成熟した学問分野であり、これまでの多くの学問的成果の上に立ちさらに発展を遂げている分野であるということが出来る。このことは他の学問分野にも当てはまるのであるが、それでもこの特徴は MSOR プログラムにおける際立った特質であるといえる。このことは、学習成果の評価が、個々のモジュールに焦点をあててというよりむしろ、科目群全体を網羅した戦略レベルで設定される必要があることを意味し、優秀な学習者は、個々のモジュールレベルよりむしろ、プログラム全体でのレベルで評価されねばならない。すでに熟知されている特徴ではあるが、MSOR の評価得点は広い範囲で分布していて、これを考慮して学習者の習熟後が総合評価されねばならない。この事項に関しては、段落 2.1 から 2.4 と 4.5 で改めて述べる。

用語の説明

1.8 本評価基準書で用いる「プログラム」という語は、学部課程の修了のような最終的な評価を目指す学習者が従う全学習計画案を意味する。「モジュール」という語は、そのプログラムを構成する授業の個々のユニットを意味する。あるプログラムでは、それぞれが各学年内で展開されるよう綿密に築かれたモジュールの連鎖で構成されており、その多くが、おそらく最終学年は除いて必修となっているかもしれない。またあるプログラムでは、履修スケジュールや先修条件を満たした上で、しかるべき指導者の指導をもとに、学習者がいくつかのモジュールを自分なりに自由に選択して組み合わせることにより構成することもある。プログラムの大多数はその両極のプログラム（必修と自由選択）の要素を持ち合わせており、多くの場合、そのような広いスペクトラムでの選択が可能ないようにデザイ

ンされている。

1.9 本評価基準書では、ほとんどのプログラムの継続期間を3年としているが、例外として4年のものもある。3年は、通常的全日制学生が必要とする長さである。MSORプログラムの中には、通信教育を含むパートタイム様式のものもある。学習者がこの種のプログラムを完了するには、より多くの時間がかかるのが通常で、時として（上限はあるもの）かなりの時間がかかることもある。

1.10 本基準書はイギリス (UK) 全土に適用するものであるが、スコットランドの伝統的教育は英国の他の地域とは多少異なっていることを忘れてはならない。とりわけ、スコットランドでは全日制学生は4年で学位取得プログラムを完了することが標準であるが、それはほぼ他の地域の3年間の優等学位と同等である。通常、1年次のスコットランドのプログラムは、他地域のそれと比べ、広範にわたるかもしくはややレベルが低いものとなっている。

1.11 この基準書における、学科 (department(s)) とは、当該教育機関におけるサブユニットで、プログラムのデザイン、実施および管理を担う単位を指す。また、department に相当する語として division や school が使われることもある。

1.12 MSOR は非常に幅のある学問分野で、その中で相互に関係し合うが、また多くの他の学問分野とも相互に関係している。MSOR の学問分野の理解のためには、それらの相互関係について議論する必要がある。このためには、最初に MSOR の構成要素である3つの分野の性質が考慮されなければならない。

数学

1.13 数学は、古代ギリシャから初期文明を含む様々な文化を経た起源をもつ、それ自体大きな知的学問体系である。数学の起源は、測量学、機械の構造、商業といった、実際的な問題を解くための方法を系統的に発達させたことにある。ある特定の状況の細部を排除し、抽象的な方法論に特化して進化発展した数学は、その結果、幅広い応用分野と、多くの異なった問題の本質的かつ共通な特質を明瞭化するに至った。一般化と抽象化が数学の重要な特徴である。このことが抽象的存在に関する命題の論理的な証明へと数学を導いた。こうして数学は緻密で隙のない議論とともに発達し、それは、明確な仮定から導かれた確実で信頼できる結論をともなう論理的推論を含む科学へと発展した。その結果、数学は人間の思考の発達に際立った貢献を果たしている。初期の数学が依然として今日的な価値を持つ一方で、急速な発達をとげつつある学問でもある。たとえば、アンドリュー・ワイルズ (Andrew Wiles) の論文「フェルマーの最終定理 (Fermat's Last Theorem)」のような印象深い出来事が、数学は発展し続けるものという事実を人々の意識上に載せた。

1.14 数学の抽象的な研究は、それ自体知的な価値の探求であり、興奮と美を包含する世界への扉を開く。しかし、その主題は現実世界との関わりを失ってははいない。数学の目的の一つは、多くの分野で共通するパターンの発見であり、抽象的で一般的な関係の構築こそが、数学がほとんどすべての学問に応用可能であることの所以である。実際、数学は時としてパターン発見の研究分野と定義付けられたりもする。数学の研究の焦点は、パターンであれば何でもいいというものではもちろんなく、抽象的な対象間の正しい関係を映すパターンに当てられる。自分たちを取り巻く世界を理解するため、われわれは概してパターンを探すものである。パターンを見つけたり、見つけたと思ったりした時、われわれは数学に照らし合わせて、確認したパターンについてもっと知ろうとし、また単に実際に観察したもののみならずその背後にあるより一般的なパターンの存在に思いをはせるのである。例えば、幾何学を考えてみよう。幾何学 (geometry) という言葉は、本来は土地を測ることを意味したが、この事実から、この学問が、2500 年前の人々の営みにその端を発し、単なる測量を超えたより普遍的な構造の探求に発展してきたことがうかがわれるのである。

1.15 数学には極めて広い適用の可能性がある。数学は大半の科学技術の基礎であるばかりでなく、どんな学問であろうとも、モデル構築を要求するほとんど全ての状況に対してもその基礎となるのである。ここ数十年、数学の使用の爆発的成長が、従来の基盤である科学技術、エンジニアリング分野の外部で起こっている。このことは、おそらくまだ一般の人々にはよく知られていないのだが、例えば、コンピュータテクノロジーや CT スキャナーのような医療技術の下支えに、また、電子金融取引に必要な暗号手法や、宇宙探査ロケットの制御といった事柄の背後に数学が在るのである。

統計学

1.16 統計学の歴史は、数学ほどは長くない。その起源をたどると、二つの螺旋構造が分かる。一つ目は 17 世紀中頃の賭け事についての議論で、それが確率論を構築する最初の試みへとつながった。二つ目は、今日の公的統計と呼ばれるものであり、19 世紀にその類の集計が次第に現実のものとなった結果、データの表示、分類、解釈において新しい発展が見られた。今日的に見れば非常に初歩的なテクニックの記述統計学でしかないというものの、当時としてはかなり洞察力のあるもので、その活用を通して、多くの公共政策における目覚ましい進歩が見られた。ロンドンのコレラ発生の主な原因となったポンプの特定や、フローレンス・ナイチンゲールによる今日の広範な医療統計の先例などが含まれている。

1.17 確率論は今日、それ自身固有の研究分野と種々の応用分野において発展を遂げている。統計学は、限られたデータからその背後にある構造に関する推論を行う過程の中で、確率論を利用してきた。数学の箇所ですでに触れたが、この推論の課程はある意味でパ

ターンを探すことに他ならない。それは、データを収集、分析、解釈する全過程を包含し、今日では特に、観察研究と実験研究のためのデザインがより大きな関心事となっている。数学的な考え方と方法論が、現実問題におけるランダム性、偶然性、変動性、リスクを含む問題を解決するために用いられ、この理論と応用の相互作用こそが統計学の大きな特徴となっている。現実世界の大部分の事柄はこれらの特徴を有することから、特に医療、品質管理、マネジメントなどに代表されるように、自然科学、社会科学、ビジネス、経済などの多くの分野において、統計学の役割は増大しつつある。統計学は、限られたデータから得られる情報に基づいて決定がなされなければならないような、どのような場面にも適用される。特に近年の近年のコンピューティングパワー（計算力）の進歩により、統計的方法論はますます応用分野を広げつつある。

オペレーションズリサーチ

1.18 OR は、20 世紀に起こった比較的新しい学問分野である。第二次世界大戦中の活動の組織化に各分析手法の起源が多く見られる。それは、きわめて数学的な要素を持つ複雑な最適化プロセスから、あまり数学的ではないがある種の問題解決の方法まで、幅広い範囲にわたって焦点を当てている。その応用は、産業、ビジネス、商業など、また政府関係では、健康、社会福祉サービス、軍事といった多方面にわたり、重要に活用されている。それは現代における、重要な定量的な管理ツールの一つになった。

1.19 OR は多分、学問分野としては数学や統計学よりあまり注目されてはいない。しかし、OR ではモデル構築がきわめて重要な位置を占め、その面において数学と統計学に深くその基礎をおいている。この分野での問題解決と意思決定の側面は幅広い基盤を持ち、それらはさまざまな数学以外の学問においても重要な役割を果たす。例えば、産業心理学、社会学はとりわけ重要な分野であろう。このため、OR の教育プログラムは比較の実用ベースなものであることが多い。

1.20 「オペレーションズリサーチ」という語は次第になじみのあるものになってきているが、多くの教育機関では、これを「マネジメントサイエンス（経営科学）」のような、他のタイトルをその教育プログラムにつけていることに注目したい。このようなタイトルは、数学的内容は少なく非常に実用ベースのプログラムであることを表示している。それらのプログラムは、その構成方針の理由により、MSOR の評価基準書の範囲内には完全には収まらないかもしれない。

MSOR 内の関係

1.21 MSOR 内の 3 つの学問分野は互いに複雑に関係している。数学が基本領域であることは言うまでもない。他の 2 つは、実のところ数学自身そうなのであるが、数学をベース

にした問題解決やモデル構築のプロセスに焦点をあてる側面をもつ。さらに、これらのプロセスを通じ新たな問題が MSOR にフィードバックされ、時には新しい研究分野さえ生み出すのである。

1.22 統計学と OR は非常に数学的な視点から教授されることが多いが、両者は方法論中心で応用重視の視点からのアプローチもされる。このことが、この基準書が言及する理論ベースのプログラムと実用ベースのプログラムの差異を強く浮き彫りにする。

他学問との関係

1.23 MSOR の広範な適用可能性のため、それは物理学、化学、天文学、エンジニアリング、コンピュータサイエンス、経済学、会計学、保険数理、財務など多くの他の科目と様々なレベルでリンクしている。これら外部の学問分野との関係は、時には MSOR 内での関わり以上に強いものもある。これらのつながりの深さを示す例として、MSOR の専門家たちが、しばしば他の分野においてもその分野の専門家として認知されているということがあげられる。プログラムデザインの点では、これら他分野のプログラムの中には MSOR に非常に依存しているものがあり、それらの分野における評価基準書は、少なくともその一部において MSOR 科目の評価基準書と密接な関連を持っている。

MSOR 卒業生の職業

1.24 MSOR のプログラムを卒業した学生は、きわめて多様な職業の選択が可能となる。雇用主は、このプログラムの卒業生たちが、数式的、論理的思考と、かれらの特質である分析的な問題解決へのアプローチに精通しているであろうという理由で、彼らの知的能力と厳密な論理展開力を非常に高く評価する。

1.25 MSOR のスキルが明白に認識されている分野では、MSOR の卒業生は容易に仕事が探せるであろう。例えば、主要な産業技術の中心における数学者の仕事、医薬分野での多くの専門的な統計学者の仕事などがあげられる。もちろん、伝統的部門での指導職や学術的な調査の仕事も、極めて重要なものである。また、その他の仕事の機会も、大企業から小さな機関まで、ビジネス、商業、公共・民間セクターなどの業界全体に渡って多く存在する。雇用主は、MSOR 卒業生のユニークな知識と特殊な技能（科目評価基準書 3.21 から 3.24, 3.25 から 3.27 参照）に、高い価値を見出している。

1.26 MSOR の諸学会と専門家団体は協力して卒後キャリアに関するウェブサイトを開設している。 www.mathscareers.org.uk と www.rss.org.uk/careers と www.orsoc.org.uk/careers を参照されたい。

1.27 MSOR の教育プログラムの多様性の持つ価値は、その中におけるどのようなプログラムを履修したかを問わず、卒業生たちがすぐに適当な職を見つけられるという事実からも証明される。プログラムはこれまで社会の急速な変化に素早く対応してきている。その良い例として、ファイナンス分野で数学への関心が急激に高まっており、それはとりわけ重要な雇用分野になっていることがあげられる。多くの教育機関は、そのような発展に応じて、新しいモジュール、時には全く新しいプログラムを提供している。いずれにせよ、MSOR 評価基準書は、そのような柔軟性と革新をいささかも抑制するものではない。

2. MSOR の特性と範囲

MSOR の累積的特質

2.1 MSOR に含まれる学問分野における教授科目は、多くの場合累積的であるという特徴を持つ。何が教えられ、何が学習されるかは、かなりの部分、それ以前に学習した内容に依存する。このことは、他の学問に比して MSOR に際立って見られる特徴である。MSOR プログラムは、それまでに扱われた教科内容を考慮し、その上への積み重ねとして科目を配置するという形式でデザインされなければならない。初級、中級に相当する分野がカバーされるまで、純粋数学の上級領域が扱われてはならない。応用分野の授業科目の開発においては、その分野での特質を吟味したうえで、要求されている方法と技術がそれまでに扱われていることを常に確認する必要がある。このことは、実用ベースのプログラムより理論ベースのプログラムへの制約となるかもしれないが、実用ベースのプログラムであっても、あることが教授される前にそれに関係した事柄が前もって理解されていることが必要である。

2.2 MSOR の累積的な特性は、プログラムをどこから開始し、そして、このプログラムの卒業生が到達すると予想されるレベルの設定に大きく影響を与える。異なるプログラムは異なるエントリー基準でデザインされるが、MSOR プログラムを学びたいと思う学生の多様性を考えれば、それは当然そうあるべきである。総じて、科目分野全般に渡り、異なる知識レベル、スキル、能力が身に付けられなければならない。

2.3 前もってその分野に関する知識のある学習者や、たとえば他の国などから移ってきたような、他の教育機関から移動してきた学習者のケースに対応して、既習の知識に合わせるという問題が生じる。正式な資格認定までは至らないにせよ、学生が今後の学習内容に対して必要とされる知識をどの程度まで持ち合わせているかを入学選考担当者は判断することが必要である。このように、規則内での配慮はもちろん、ある種の柔軟性が求められるのである。

2.4 MSOR の累積的特性は、さらなる重要な結果を持ち合わせている。これは、学習内容

が身に付くまでに非常に時間がかかるものであり、学習者はそれを学んだ後しばらく時間が経ってから初めて、それが何を意味するのかが完全に理解されたりするのである。しかし、もしあるトピックを学習した直後に、その内容に関する評価が行われるのであれば、学習内容の吸収の過程が完了していないので、課題の習得や評価のねらいと目標の達成を確かめることは難しいと言えよう。このことの重要性は、戦略計画レベルでの学習成果の設定について論じた 1.7 項において述べたところである。この事実は、プログラムの評価者がプログラムの評価の際に認識しておかなくてはならない。また同じ理由で、MSOR のプログラムを、容易に評価できるねらいと目標にのみ制限してはならない。

MSOR プログラムの特質

2.5 MSOR 分野の特別な性質、プログラムが提供するものの広範性と多様性について、再度強調しておく。

2.6 MSOR プログラムは、その多岐にわたる異なった対象と広い学問的背景によってその学習者を魅了するであろう。すでに指摘したように、雇用の可能性はあらゆるところに存在している。プログラムが内包する多様性は計り知れないほど有益であり、硬直化した基準で測られる類のものではない。したがって、MSOR 科目評価基準書は、決して国全体あるいはすべての学問領域に普遍的に適用されるような種類のものではないという QAA の見解と一致するものである。

2.7 提供されるプログラムの多様性は、そのプログラムが理論ベースと実用ベースの間の極めて広範囲のスペクトラムに位置する場所や、プログラム内での科目の選択性の幅に起因する。あるプログラムでは、ある学問分野における特定の内容が非常に深い形で開発され、別のプログラムでは、幅広い領域をカバーする広範な活動に向かって開発が進められている。きわめて理論ベースなプログラムの卒業生は、同じトピックを実用ベースのプログラムで学習した卒業生に比べ、その分野固有の知識と理解をより十分に習得するであろう。また、その一部が他の学問分野と密接に結びついているのが特徴の MSOR プログラムにおいては、その学問分野における知識と理解がより強調されるであろう。

2.8 多様性のさらなる重要な源は、多くの場合、教授スタッフの学術上の専門的な興味に存する。学部生のための MSOR プログラムは、知識の開拓者レベルへの到達は期待していないが、研究活動を活発に行っているその分野での専門家から教えることは、学習者にとって刺激的な経験であるに違いない。MSOR プログラムが提示する教材の選択は、もちろんそれが主に教育的価値で決められるとはいえ、しばしば教授スタッフの研究内容と専門的な興味によって影響される。もちろん、これらはプログラムデザインの背景にある主要な要因ではないが、活気に満ち、知的冒険心を刺激する学習経験を提供するという

点で重要である。

2.9 その多様性にもかかわらず、全ての MSOR プログラムの基本的な特徴は特定され得る。全てのプログラムは、学習者に基本知識と考え方の確実な理解と、ある分野に固有な知識の習得を目的としている。理論ベースのプログラムであろうと実用ベースのプログラムであろうと、教育内容の正確な理解と、厳密な方法による問題解決の方法を繰り返し教えることが強調される。これによって、MSOR 自身あるいは他の応用分野で、挑戦的な問題解決をやり遂げることができるのである。また、卒業生が学んだ分析力を通じて能力をさらに開発する可能性についても強調されている。

2.10 多くの MSOR プログラムは、とりわけ最終学年において、核となる本質的な学習内容を特定化するとともに、多くの選択肢をも提供する。これについてはセクション 3 で詳しく述べる。本質的にモジュール式のプログラムであっても、通常、基本となる必修モジュールのセットは提供されている。

2.11 MSOR プログラムは、3つのそれぞれの分野で、単一専攻課程プログラムを提供している。これらの中では通常、数学プログラムに最も多い学習者がいる。統計学のみプログラムではそこでの学習者は比較的少なく、OR のみではさらに少なくなる。個々の分野での学習者の数は、プログラムのタイトルと定義の多重性のため、大学入学者にかかわる出版物や高等教育統計機構によってもたらされる統計数字から、読みとることは難しい。

2.12 数学の単一専攻課程プログラムは、必修ではないのだが、ほとんど常に統計学のモジュールを含んでいる。OR に関しても、しばしば選択されるモジュールが見られる。他の2教科の単一専攻課程プログラムは、ほとんどの場合数学を含んでいる。数学は、これら2つの領域においても重要な位置を占め、特に理論ベースの統計プログラムでは典型的にそうになっている。その一方で、数学が全体のほんの一部分の割合しか占めないプログラムもある。これは、実用ベースの OR プログラムに典型的にみられる。

2.13 MSOR のプログラムの中には、MSOR のうちの2つの分野、あるいは3つの独立した分野の学習を結びつけるものもある。例として、純粋数学、応用数学、統計の幅広い範囲のプログラムがある。プログラムのタイトルからは、どのタイプのものか明らかでないものもある。例えば、単に「数学」とタイトルされるプログラムが、幅広いアプローチを示すものとして使われたり、またある教育機関では、数学そのものに深く根ざしたプログラムを示すものとして使われたりするるのである。プログラムの性質は、タイトルではなくその目的と目標から明らかになるのである。

2.14 プログラムが、単独での学位を目指すものかあるいは複数の分野のジョイントでの

学位を目指すものであるかは、教育機関がどのような構成になっているか、あるいは全ての教育指導が同じ学科あるいは学部で行われるかに依存する。評価基準の目的では、MSOR 評価基準書は、単独での学位取得プログラムでもジョイントでの学位取得プログラムでも同様に適用できる。

2.15 MSOR プログラムは、その中で提供されているモジュール以外のものとして、他の学問分野における教育プログラムのモジュールのうちのいくつかを履修する機会を提供する。一般的な例としては、会計学、経済、外国言語のモジュールがある。そのような他の領域のモジュールの選択は、卒業後の職業選択に資するという目的だけではなく、それらの中に数学的な分析を必要としている問題が内在していることに起因していて、そのこと自身、MSOR の分野の学習においてきわめて重要な側面を有している。これらのモジュールは、その分野における学部学科で教えられ、その場合、それは当該分野の専門家に教えられるものと同じモジュールであることも、そうでない場合もある。一般論として、MSOR 評価基準書は、この種の「外部」モジュールまで及ぶことはできない。ただし、「外部」モジュールがプログラムの特質全体に影響を与える限りにおいてはこの限りではない。

4年プログラム

2.16 3年プログラムが基準であるイングランド、ウェールズ、北アイルランドで実施されている学問の伝統に触れた 2.17 から 2.19 で 4年プログラムについて言及した。1.10 で述べたように、スコットランドの伝統は幾分異なっており、4年プログラムが基準である。2.17 から 2.19 で論じるプログラムに必要な年限については、スコットランドでは1年加算する必要がある。

2.17 4年間の数学修士号 (MMath) のタイプのプログラムは、1992年の数学者の委員会の審議の結果開発され、その後政府へ説明された。大学での数学の内容の変更により、大学院の学問に備える十分な準備ができるプログラム、あるいは、多数の学生に厳しい要求をしないで、卒業後の仕事内容が要求するのと同程度の数学的な知識と技術を与えるプログラムをデザインすることは困難だ、ということが主張された。その結果、このタイプ（現在、通常、総合修士プログラムと呼ばれるもの）のプログラムの確立が合意され、学部生のプログラムと同じ方法で開発された。同様のプログラムが、物理、化学、エンジニアリングにみられる。MSOR プログラムでは MMath や MSci というタイトルが一般的だが、例えば MMathSat のような他のタイトルのものもいくつかある。そのようなプログラムは現在幅広く提供されている。それらはもともと研究の広さよりも、深さを追及してデザインされた。MSOR 評価基準書は、明らかにそれらと関連しているが、現在のところそれらを含んではデザインされていない。それらを含む評価基準書の付録が開発されるであろう。

2.18 4年プログラムで全く異なっていたものはサンドイッチプログラムで、そこでは学習者は、監督者のもと、産業、ビジネス、商業の専門的な部署で、インターンとして1年（時に2回の分割した期間を）費やす。大学の中には、学習者が希望すればそのような機会を提供し、あるプログラムでは学習者にそのような経験を積むよう要求するものもある。その場合、学生のインターンシップのための基礎となるインフラの整備とチューターによるサポートが必要となる。これらのプログラムの学問的構成要素は、通常、学習者がこのようなインターンシップを受けるか受けないかによって全く影響は受けないので、MSOR評価基準書は、そのまま適用することができる。

2.19 これと似た状況が、学習者が海外（一般にヨーロッパ）の教育機関で1年過ごした場合である。そのようなプログラムは、MSOR科目として共有でき、その場合は、2.30から2.33で述べられる外国語学習によるジョイントプログラムのための評価アレンジメントに対応する。そのようなものは、MSOR科目で比較的小さな構成の言語トレーニングのプログラムであることもある。この状況は、2.15で述べられたものと似ている。教育機関は、国が異なることによりすべての教育課程が同じにはなり得ないことを認識しつつ、海外で受けるMSORの内容の学習と、UKで受ける残りの部分のMSORプログラムがうまくかみ合うような特別な配慮を確実に行わなければならない。

数学に焦点をあてたプログラム

2.20 典型的な数学のプログラムには、連続数学、離散数学、論理的推論、数学的な問題解決、数学モデリングなどがある。厳密な数学的証明は、理論ベースのプログラムでは非常に重要である。多くのプログラムは、確率論と統計学の基本を含む。典型的なプログラムの要素については3.8から3.20までで述べる。これらのプログラムは、単一の独立したモジュールというより、モジュールのセットで構成されることが多い。

2.21 数学理論、方法、技術の応用は、数学の他の分野か、あるいは明らかに異なった応用分野や、多くの場合その両者に渡って探求されるであろう。そのような分野とは、力学、統計学、オペレーションズリサーチ、物理学、天文学、化学、生物学、エンジニアリング、財政学、経済学、保険学などである。プログラムはその目的と対象に従って、ある特定の分野に焦点をあてることになる。

2.22 応用分野でのモジュールは、その科目を展開する別の学部で教えられることもあろうが、そこでは、それはその分野の専門家が教えるモジュールと同じものか、あるいは異なるものであるかもしれない。しかしながら多くの場合、指導は数学の学部で行われる。一つの例は数理物理で、これは英国の数学科において多くの数学の研究者が取り組んでい

る科目である。数学の学部で教えられるモジュールは、そうである必要はないにもかかわらず、本質と形式において、他の学部で教えられるものと異なる傾向にある。数理物理の例は、多くの点で数学は、MSOR 科目評価基準書に含まれる OR に対してと同じくらい、違う評価基準書がある物理学に近いという事実を物語っている。

統計学に焦点を当てたプログラム

2.23 典型的な統計学のプログラムについては 2.20 で述べた。それらはまた、データに基づく研究、サンプルからの推論を行うための確率ベースのモデル化、統計理論、統計の様々な分野への応用、および得られた結果を的確に伝えるコミュニケーション手段などにより、事象のもつ不確定性を理解しかつそれらを管理するための方策を含んでいる。

2.24 統計学の分野に見られるプログラム要素は、数学分野でのプログラムに加え、データの視覚化、推論、尤度、線形および非線形のモデリング、実験計画、確率過程と時系列、ベイズ流の推論、コンピュータ統計学、および統計解析専用のコンピュータパッケージの使用などを含む。

2.25 統計学のプログラムは、多くの応用分野からの問題を扱うことで特徴付けられる。その範囲は、生物学、化学、医学、薬学、エンジニアリング、地理学、考古学、環境科学、保険学、経済学、経営学、法律など様々である。これらの分野では、しばしば各学部で教授される独立したモジュールも利用可能である。

オペレーションリサーチズに焦点を当てたプログラム

2.26 OR におけるプログラムには、多くの異なったタイトルが付けられている。OR という言葉がタイトルに付けられている場合もあるし、他のタイトルとしては、ビジネス決定理論、ビジネスシステムモデリング、経営科学、などがある。最後のものについては、タイトルだけでは、MSOR のものであるかどうか判断できないかもしれない。

2.27 OR のプログラムは本来、複雑な事象のモデルの確立に関連しているが、それらの事象では解析的な解が得られにくいものが多く存在する。本来の問題の文脈の中における結果の解釈がとりわけ重要である。数学的および非数学的な側面の両者がモデリングには必要である。数学的な側面では、線形および非線形計画法、整数計画法、動的計画法などのような数理計画法が中心的话题となる。また 2.20 に関連して、数学の領域（特に離散数学）が、また、2.23 及び 2.24 に関連して統計学の領域も含まれる傾向にある。

2.28 プログラム要素は、OR 分野のなかでプログラムが焦点を当てているものに拠る。それらは、数理計画法、組み合わせ分析、グラフとネットワーク、スケジューリングと逐次

モデル，ゲーム理論，決定分析，決定サポートシステム，在庫管理モデル，シミュレーションモデリング，発見的方法論，ソフトシステムメソッドなどからなるであろう。

2.29 統計学の場合と同様，OR のプログラムは，さまざまな種類の応用分野に関係するところに特徴がある．分野としては，製造と生産，健康関連分野，企業戦略プランニング，輸送，物流と拠点配置，など他にも多い．これらの分野では，しばしばそれら特有の分野で教授される独立したモジュールも利用することができる．

ジョイントプログラム

2.30 MSOR における授業科目は，ジョイントプログラムとして構成されることも多い．実際，2.13 で述べられたように，異なる MSOR 科目の要素をつなぎながら，MSOR の中に全体を位置させるプログラムが多い．

2.31 他の学問分野とのジョイントプログラムに関しては，他の学問分野におけるプログラムと同様，MSOR 科目と他の学問との構成比は，多対少あるいは 50 対 50 である（時に，少対多の割合であることもある）．そのようなプログラムの評価基準は，すでに他の場所で述べた特別の問題を提起している．それらのうち MSOR の部分は，MSOR 評価基準書に入る．MSOR の内容が全体より少ない（時には，かなり少ない）プログラムであっても MSOR の部分は，MSOR 評価基準書により評価されるであろう．

2.32 モジュールからなるプログラムでは，より複雑な問題が起こる．そこでは，学習者はできる限り選択可能な広範囲のモジュールによりプログラムを組み立てる．そのようなプログラムが MSOR 要素を含んでいて，その MSOR の内容が全体よりはるかに少ない割合のプログラムにも，MSOR 評価基準書は適応可能である．

2.33 ジョイントプログラムの命名については注意が必要である．場合によっては，2つ（あるいは2つ以上）の学問分野が互いをサポートしあうよう計画的に開発されているものがある．MSOR の例でいえば，数学と統計学のプログラムがそうである．その他の場合では，授業科目は互いに並行又は独立して展開される．数学と言語のプログラムはこのタイプにあたる．教育機関の中には，後者のような場合はジョイントプログラムとしていて，前者のタイプは2科目の単独専攻課程としているものもある．いずれにせよ，この2つのタイプのプログラムは，同一のあるいはよく似たタイトルを持っていることが多い．他の多くの領域でそうであるように，プログラムの特性はその目的と目標によって明らかにされるのである．

他のプログラムにおける MSOR 科目

2.34 MSOR 科目は、それらが、他の分野のプログラムに必然的に現れるという点で、全ての学問の中で独特である。例えば、数学はエンジニアリングあるいは自然科学において鍵となる役割を果たし、統計学は環境分野や社会科学に広く使われ、統計学とオペレーションズリサーチは、ともに経営学にとって重要である。他のプログラムにおける MSOR 科目のスタイルと表現法は、通常、MSOR プログラムのそれとは異なる。多くの場合、MSOR 科目は、MSOR 学部の教員によって教授されるが、一方でそのプログラムが基礎を置く学部によって教えられることもある。他のプログラムは、それらの中における MSOR の位置づけに注意をはらい、それが適切である場合には MSOR の要素は当該分野の評価基準書の中に組み込まれている。しかし一般に、MSOR 評価基準書は、他分野での学習評価などには適用されない。

2.35 MSOR 科目は一般に、高等教育プログラムへのエントリー資格を持たないような学習者への準備プログラムとして用いられることもある。繰り返しになるが、MSOR 評価基準書はそのような場合には直接適用できないが、そのようなプログラムのデザインの際には参考として訳に立つであろう。

MSOR プログラムの機関認定

2.36 王立統計学会 (The Royal Statistical Society = RSS) は、その制定する基準に照らし合わせて、教育プログラムに対する認証評価を行っている。認証評価を受けたプログラムの卒業生は、個人からの申請によって申請によって自動的に Graduate Statistician となる資格を有する。認証評価されていないプログラムの卒業生であっても、そのプログラムにおいて統計学の比率が十分にあれば、個人の申請により Graduate Statistician として認可される可能性がある。Graduate Statistician であり、その後適当な実務上の経験を加えることにより Chartered Statistician の地位が与えられる。

2.37 数学応用協会 (The Institute of Mathematics and Its Applications) は、その制定する要求事項に合致した場合、MSOR のプログラムを認証評価し、Chartered Mathematician の地位を与えている。このプログラムの卒業生は、申請により Graduate Membership の地位が与えられ、さらにプロフェッショナルとしての経験を重ねることにより Corporate Membership が与えられる。

2.38 MSOR プログラムの卒業生は、適切な科目を履修し、それらに対し十分に高い基準で合格すれば、一般に、他の多くの職能団体、例えば会計事務や保険数理士のような団体の試験を一部免除されるという優遇措置もある。

3. 知識、理解と技能

序論

3.1 セクション2で述べたように、この基準書でカバーされる MSOR の一般科目分野は非常に広範である。したがって、この分野の卒業生に期待される知識と技術は同様に幅広い範囲に渡る。

3.2 そのような卒業生は、MSOR 内の特定の分野に固有の知識と技術を持ち合わせている。この基準書では、そのような固有の知識と技術を、分野固有 (subject-specific) と呼ぶ。高レベルの学習内容の知識と技術は、卒業生間で必然的に異なるものとなる。というのも、当該学問分野内でどのような科目群を履修したかは、その分野の卒業生間でかなり異なっているからである。この多様性は MSOR 分野の当然の特徴であり、歓迎されるものであって決して制限されてはならない。さらに、教育プログラムの開発段階においては、新しい研究分野の包含は、学問領域のダイナミズムをさらに推し進めるという意味で必要不可欠なことである。しかしながら、ある程度の範囲内ではあるが、プログラムの卒業生が習得した分野固有の知識と技術を識別することは可能である。

3.3 多様性は、その上にそれらが確実に展開されるための基礎を必要とする。この基準書でカバーされる連続的な科目群においては、ある科目群の下に基礎的な科目群を必要とする、というような階層構造が必要である。この階層構造はプログラム全体に渡って存在する。基礎のうちの大部分は、通常プログラムの初期段階に配置されることが多いが、必ずしもすべてがそうであるとは限らないということを知っておく必要もある。同様に、プログラムの初期段階がすべて基礎のために費やされるというものでもない。多くのプログラムでは、必要とされる知識が整っているという条件の下で、より発展的な内容や応用的な内容が、プログラムの初期の段階で始められることもある。

3.4 上記の理由で、MSOR 評価基準書がカバーするプログラムの全卒業生のために、分野固有の知識の総合的なリストを構築しようとすることは意味がない。そのようなリストは、あまりにも規範的に過ぎ、既存のプログラムに不必要で望まない修正を強要し、何ら意味ある恩恵を与えないであろう。

3.5 しかしながら、ある種の知識体系は、1年間の学習内容全般に渡って広がる水平方向のテーマと、数年にまたがって続く垂直方向のテーマとして特徴付けられるかもしれない。この種の捉え方により、3.8 から 3.20 で述べられる、MSOR 分野の卒業生が習得しているべき分野固有の知識と理解の内容を大局的にみることができる。

3.6 MSOR 分野の卒業生は、より一般的な高い技術を所有している。その例として、彼ら

には、高度な数学の基礎知識があり、プログラムの卒業生の多くは、コンピューティングの応用に精通しているであろう。

3.7 したがって、下記のセクションに記載した内容は、MSOR 評価基準書が適用される全てのプログラムに当てはまる。「卒業生」という言葉は、MSOR 内の特定の科目分野や、プログラムのスタイルによる条件付けがないならば、そのようなプログラムからの卒業生すべてを指して使われる。

分野固有の知識と理解

一般原則

3.8 MSOR の全卒業生は、専攻分野にふさわしい数学的方法論と技術の知識および理解と、MSOR に展開される多く科目群の中の特定の範囲のものに関する学習成果を持ち合わせる。加えて、プログラムの卒業生の多くは、自分の専攻科目の応用対象として最低限1つの応用分野を持ち、そこで用いられる理念や手法の適切な理解のために必要な努力を行っている。応用分野の性質と何がそこで学習されるかは、プログラムが理論ベースであるかあるいは実用ベースであるかによって異なるものとなる。プログラムが焦点を当てている対象によって、卒業生の獲得した知識・技術は異なるのである。

方法と技術

3.9 全卒業生は、彼らのプログラムにふさわしい数学的方法論と技術、およびそれらを使う能力を有している。全てのプログラムの共通点は、解析学と基礎的な線形代数を含むことである。他の方法論と技術は、そのプログラムの要求事項によって開発され、さらにその要求事項は、学生の到達レベルの決定にも大きく関わる。

3.10 例として、OR のプログラム卒業生は、条件付き最適化の知識と、資源配分の問題への応用、さまざまな意思決定プロセスのモデルに関するかなりの知識を持っている。その一方で、物理学やエンジニアリングにおける数学の応用に関するプログラムの卒業生は、微分方程式を扱うための方法に関する相応の深い知識を持っている。

3.11 これらの例は、MSOR プログラムのスペクトラムの中で特に離れているものを選んだのだが、それらがカバーする方法と技術は、互いに無関係なものではないということを強調したい。全てのプログラムは、数学の分野に関連する方法と技術を扱う。そして個々のプログラムは、それ自身の要求の度合いに応じて開発される。

数学の分野

3.12 理論ベースのプログラムの卒業生は、MSOR 専門分野のさまざまな範囲から得た知

識を持ち、各専門分野における主要な成果を理解している。すべての専門分野に共通な科目群として、代数学、解析学、幾何学、数論、微分方程式、連続体力学、数理物理学、確率論、統計学などがあげられるが、他にも多数ある。これらの科目群の知識と理解が、数学的な定式化と問題解決を通じて、多くの問題に対する数学的方法と技術に関する知識と理解をサポートする。

3.13 実用ベースのプログラムの卒業生も、MSORの専門分野の範囲から得た知識と理解の成果を持ち合わせている。しかしその知識は、モデルの数学的背景を深く理解するというより（時にはこの部分もカバーされているのだが）、モデルそのものの理解と、それがいつどのように適用されるかの理解をサポートするために、プログラムがデザインされるのが一般的である。

数学的思考と論理的プロセス

3.14 卒業生は、仮定とそれがどこで使われるかの重要性、ならびに仮定が成り立たない場合に起こり得る結果の重要性を理解するであろう。これは、仮定の妥当性と推論の妥当性との間の差異の理解を含んでいる。卒業生はまた、問題解決において用いられる数学的な理論や方法の展開における、一般化と抽象化のもたらす影響を完全に理解している。理論ベースのプログラムは、論理数学的な根拠と演繹的な推論の役割を強調する傾向にあり、しばしば、数学的な証明の手順を含む。実用ベースのプログラムは、問題解決に対して体系的な数学分析的アプローチを使用し、そのアプローチの理解を強調する。

3.15 この表題中にある知識と理解は、さまざまなプログラムにおける種々の活動、たとえばより高度な純粋数学における公理的アプローチや数学的モデリングの一般的な役割りに対し、有益な情報を提供する。

数値解析と数学計算

3.16 全卒業生は、プログラムが要求するレベルで、数学的な近似の過程とその限界についての知識と理解を持ち合わせている。

3.17 実用ベースプログラムの全卒業生と、理論ベースプログラムの卒業生の多くは、多くの場合複数の種類のコンピュータパッケージの利用を経験しているので、ある程度の数学的な計算の知識を有し理解している。彼らは取り上げられている問題に対するパッケージの妥当性を理解し、そのパッケージの基礎となる数学的なアルゴリズムの特質に気づいている。

モデリング

3.18 モデリングとは、適切な表記法を用いることにより、数学的もしくは統計的な手順で問題を定式化するプロセスである。全卒業生は、このプロセスの知識と理解を習得している。通常、問題は少なくとも一つの応用分野からのものであるが、MSOR 内の他の分野からのものもある。

3.19 実用ベースプログラムの全卒業生と、理論ベースプログラム卒業生の多くは、モデリングのテクニックのための条件と限界、モデルの実証と改訂の必要性に関する知識と理解を習得している。加えて、根底にある問題を分析し、可能な限りの解決をするための方法論、およびモデルの修正による影響の大きさを評価するためのモデルの使い方や、分析結果の解釈の仕方について、当然のことながら十分理解している。

学習の深さ

3.20 全卒業生は、特定分野において、高いレベルの知識と理解を所有する。プログラムにおける高度な学習内容は、プログラムのタイトルに反映されている。例えば、統計学を含むタイトルのプログラムを卒業した者は、統計的推論と多くの応用統計の本質的な理論の知識をかなりの部分を習得かつ理解している。数学というようなタイトルのプログラムは、この科目におけるかなり広い領域にわたるかもしれないが、それにもかかわらず、そのようなプログラムの卒業生は、いくつかの特定のトピックスに関する深い知識を有する。

分野固有のスキル

3.21 MSOR 卒業生は、非常に幅広い活動との関連で開発された、分野固有のスキルを習得している。これらのスキルは、卒業後に専門職業の場で起こる新しい問題の解決の際に、あるいは数学を含む多くの専門分野における作業からなる高度な学問研究の場において、十分活用できるほど高いレベルまで上達している。

3.22 多くの分野固有のスキルが、全 MSOR 卒業生に期待されている。これらの大部分は、学位プログラムの途中のある段階で客観的な評価を受けるが、それらのうちのいくつかは必ずしも明確な評価の対象とはならないことを認識しなければならない。分野固有のスキルのうちのいくつかは、全ての数学活動にわたり、科目内容の多くの分野に焦点を当てた評価に反映されている。

3.23 全 MSOR 卒業生に期待されている分野固有のスキルの多くは、問題が MSOR 内部から発生するのか、あるいは異なる応用分野から起こるのかという、問題ベースの科目分野としての MSOR の本質的特性に直接関係している。故に卒業生は、そのスキルを問題解決に適用することによって、鍵となる数学的な概念と成果を説明する能力を習得する。分析

と解決を容易にするため、またどのように数学的プロセス（ここでは解決の一部分のみをもたらしめているかもしれないという、妥当な理解を含んでいるのだが）をあてはめるかを把握するために、彼らは問題を理解し、問題の本質を抽象化し、記号の形で数学的に定式化できるのである。妥当な数学的プロセスを選択して適用し、仮説と結論の明確な識別を行い、論理的かつ数学的な議論を構築して発展させることができる。必要に応じて、数学的プロセスの支援と、詳しい情報の獲得のために、コンピュータと基本的な IT 設備を使い、それらから正確で明晰な数学的議論と結論を提示することができるのである。

3.24 MSOR の特定部門に焦点を当てたプログラムの卒業生は、これらの特定部門に関連した分野固有スキルを習得している。そのようなスキルの網羅的なリストの構築はあまり意味はないが、例として以下が考えられる。

- 純粋数学プログラムの卒業生は、論理的な推論と一般的な枠組みでの問題の解決、論理の明確な連鎖と抽象化に基づく問題設定とその解決に関するスキルを習得する。
- 物理学への数学の応用あるいは、理論物理学のプログラムの卒業生は、とりわけ、数学用語で物理理論を定式化する分野、解析的、数值的に方程式を解く分野、その解答の物理的解釈を与える分野に関連したスキルを習得する。
- 統計学プログラムの卒業生は、実験もしくは観察研究の計画と実施、ならびにそれらの結果得られるデータの分析に関わるスキルを習得する。
- OR プログラムの卒業生は、複雑な問題を最適化問題として定式化する技能、各種の問題が現れる文脈中での解決法の解釈に関連したスキルを習得する。

一般的スキル

3.25 MSOR 分野の卒業生は、MSOR 科目の学習経験により磨かれた、多くの一般的なスキルを獲得している。これらの科目は、MSOR 自身から起こる問題かあるいは応用分野からの問題かにかかわらず、本質的に問題解決のための学問である。したがって、卒業後の経験は、これまで学んだ数学の基礎学力と問題解決の分析的アプローチの基本精神に組み込まれるであろう。加えて、大部分の MSOR プログラムの重要な点は、ある一つの分野から得られる論理的知識を他の類似の分野に応用することである。応用の分野は、しばしばそれ自体がその分野における重要な研究のトピックであるが、学習プロセスの極めて重要な面は、一つの状況から他の状況に専門知識を移管したりすることで、多くの分野に共通した適応力を育むことである。

3.26 全 MSOR 卒業生に期待されている多くの基本スキルは、他のプログラムの卒業生のほうがより多く習得しているようなケースもある。分野固有のスキルの場合以上に、いくつかのものは必ずしも明確な評価を受けないということを認識しなければならない。

3.27 MSOR 卒業生は、書物、学術雑誌、インターネットなど、様々なメディアを使って、独力で学ぶ能力を含む、基本的学問スキルを所有するようになる。彼らはまた、問題の解決法とその結論の導出を通じて、辛抱強く貫徹して独力で問題に取り組むことができるようになり、時間の管理と自己管理の基本スキルを発達させる機会を持つようになる。彼らは、とりわけ新しい分野の新しい問題に取り組む迅速さを必要とする際には、柔軟性があり、問題を論理的に考察するために、また、分析的にアプローチするために、知識を一つの状況から他の状況へ変えることができる。彼らは、作業のあらゆる段階で、数学的基礎力の知識と推論に全面的に自信があり、数学的基礎力のスキルを高いレベルで発達させることができる。また、ワードプロセッサ、インターネットを使い、情報を得る能力などの基礎 IT スキルを持ち、これらのスキルを責任が取れるしっかりしたやり方で常に訓練し、成果の提示にあたっては参考文献などの情報源が適切に言及されるよう注意を払う。彼らは、議論に参加するため、理路整然と書くため、そして結果を明瞭に伝えるため、言語能力を含む基本的なコミュニケーション能力も持ち合わせる。必要に応じて、個人情報データを取り扱う際の機密・人権保護などに関して要求される敏感さを含む倫理的な知識をもつ。これら全ての能力は、一般的に MSOR 卒業生の雇用可能性を高める。なお、関連して 1.24 から 1.27 までを参照のこと。

4. 指導、学習と評価

序論

4.1 MSOR 科目の学生は、学習の過程で高度の概念的・抽象的考え方を要求されることで特徴づけられる。このことは、学習者に大きな試練と挑戦を与える一方で、その克服により相応の益も得られる。要求されるレベルの抽象性を有した上での事象の理解は、知的にきわめて困難であることも多くあるが、難しい考えを深く理解することは、しばしば努力に要した以上にふさわしい報酬を提供もしてくれる。新しい概念の習得は、それ自体満足に足る経験で、それを達成した学習者に将来同じような挑戦を受ける自信を与える。それはしばしば直接実際的な恩恵ももたらす。というのも、一度概念が理解されると類似の状況における問題解決法の習得は容易になるからである。

4.2 MSOR 卒業生は、抽象的論法、論理的推論および問題解決におけるかなりの能力を所有していることが期待され、そのため 1.24 から 1.27 で述べたように、多様な職業や専門職の雇用につくことが可能となる。教授と学習の方法は、単に抽象的な論法の能力ばかりでなく、学習者が独力で自己動機付けにより学習する能力、問題解決スキル、多くの分野の雇用で共通する知識とスキルをも要求する。

4.3 他のすべての学問と同じように、MSOR における教授法は、より広範な教育の実践における変化に伴って発展してきた。この状況は継続的であり、教育機関は変化する状況に

対応し続けることになる。たとえ革新が危険をはらむものであっても、また決して簡単にはやりとげられなくても、教育機関は明らかな長所を持つと認められる教育方法の開発を断念してはならない。その故、教育機関における判断と行動は、過度に細かく批判的であるべきではない。

4.4 基礎的な MSOR での教育を支援する形で多くのソフトウェアが開発され、また、Eラーニングの発達で、多様でかつ有用なソフトウェアが、MSOR の学習全般をサポートするようになった。加えて、コンピュータの使用と MSOR の教育支援のためのソフトウェアは、短時間のうちに広がりを見せた。この例として、MSOR の教育目的のための表計算ソフト、数式演算システム、高度数値解析と方程式の数値解析のための高機能のプログラム、データ分析とモデル構築のための統計パッケージ、OR プログラムの定式化と解析のための数学プログラムのパッケージの使用などがあげられる。

4.5 これらの様々な変化は、多様な学習指導と評価方法をもたらした。コンピュータの利用は、展開される教育プログラムのスタイル、科目、学習者のレベルや背景、学部の特別な興味関心、そして利用できるリソースによって著しく異なるかもしれない。

4.6 教育機関は、その教育目的に合った形での学習、指導、評価方法などを選択し採用する。学問分野の特性にもよるが、教員による通常の教室での講義における熱のこもった議論を見ることは、学習者の大きな利益になる。議論が拡張し発展する様を実際に見ることは必要で、またそのためには、教員が準備した教材を提示するためにかかなりの物理的なスペースを要する。一般に、黒板を使った講義は、現時点においても変わらぬ利点をもち続けている。そのような講義は通常、問題クラス、個別指導あるいは少人数のセミナーによってサポートされ、これらに対する人的サポートは、研究機関内の公式サポートセンター等によって提供される。加えて、教育機関が用いる方法としては、例えば、ワークショップ、グループワークやプロジェクトがある。どのような方法がとられるにせよ、学習者のニーズを反映させ、プログラムもしくはモジュールのねらいと学習の結果および教育の資源を反映させなくてはならない。学習、教授および評価の方法は定期的にレビューされなければならない。

4.7 前項 4.6 で顔の見える学習について述べたが、MSOR プログラムでは遠隔学習の利点も取り上げられなくてはならない。遠隔学習のためには、さまざまな機器や教授法および学習をサポートする人的配置などのインフラの整備が必要となる。遠隔学習で達成されるべき学習者の学習成果は、遠隔学習以外で行われる同じようなプログラムの成果と同様となるべきである。

教授法と学習

4.8 教育機関は、学習の機会と教授方法はプログラムとモジュールの目的と目標が適切に実現されるようにデザインされるべきだ、ということを常に念頭に置いている。それらは、学習者の知識と科目に対する熱意を育成し、学習のプロセスに従事し参加しているという意識を刺激する。それらは、深く学習を支援し、学習者が省察しかつ自らの学習に責任を持つことを支援し、彼らに学問的自信を与える。これらはもちろん、MSOR 評価基準書本来の一部として明確に認められている、というよりむしろ、教育機関の活動全てにわたる、非常に基本的な基準なのである。

4.9 MSOR 内の学習者は、学習内容に関する深い理解を達成するために、適切な課題をこなすことも加えたさまざまな学習活動を行う。遠隔学習でない限り、全てのプログラムは、様々な種類の講義と個人指導の両方を含む。これらの詳細な特性と、それに伴う種々のタイプの学習活動は、そのプログラムのスタイルを念頭に置きながら、MSOR 内の特定の学問分野の学習目標に照らし合わせた形で適切に決定されることになる。教育機関の中には、課外活動や付属するモジュールの配置によって、相応のスキルを開発させるような選択をするかもしれない。どのような学習活動と指導方法が使用されるにせよ、それらは学習者と望まれる学習目標の両方に対して適切であることが要求される。何が教えられ、それが実際どれほどよく学ばれ身に付けることができるかは、実施される方法そのものよりも重要なのである。

評価

4.10 教授法と学習（4.8 参照）に関しては、教育機関は、評価についての一般的な基準をもっているが、それは一般的に過ぎるため MSOR 評価基準書の一部としては認めることが困難である。これらの基準は、プログラムの妥当性、信頼性、公正性を評価するものであり、知識レベルとスキルの両面での評価を行うものである。また、学習と指導の目標に関する評価をサポートし、フィードバックの機会を与え、必要であれば外部評価（例えば、外部の評価者による評価）に委ねることもある。

4.11 学習活動に関して、学習者はプログラムやモジュールのねらいと学習成果、および評価される知識とスキルにより、さまざまな評価方法によって評価される。評価は、そこでのコンテキスト（文脈、状況）と目的に従って様々な形式によって実施され、学習者はそれら異なる評価形式のそれぞれに対し、それに見合った形で自らの理解度及び様々な能力を実証しなければならない。

4.12 MSOR 科目の評価は、数学的知識とその理解の程度にとどまるものではない。たとえば、多くのプログラムは、その目的と目標に照らし、身に付けた数学的な能力が種々

の応用分野において如何に活用されるか、統計的な推論ができるか、そして得られた成果を MSOR の文脈において的確に相手に伝えることができるかというコミュニケーション能力も評価の対象となる。

5. 評価基準書

5.1 MSOR 評価基準書でカバーされる基本的な科目分野は非常に幅広いことから、その領域の卒業生に期待される標準は、概して一般的なものにならざるを得ない。

5.2 MSOR のための評価基準は、「しきい値レベル」と「標準レベル」の2つによって定義される。

5.3 MSOR では、2つのレベル間の差異は、主として学習者の概念と技術に対する理解の深さ、学習者の知識の広さ、課題を遂行するために必要とするサポートとガイダンスの総量、学習者が解決しモデル化ができる問題の複雑さの程度、論理的な思考や証明の実行能力、それらを通して学習者が進歩した度合い、そして設備を利用して学習者が行う計算や操作の完成度による。

5.4 評価基準書の例えば「妥当な」という形容詞の解釈は、内部と外部の評価者の専門的判断による。

5.5 評価者の専門的判断は第一義的に重要である。それは、その評価が MSOR 卒業生の雇用可能性への知識と共に、プログラムにおけるこれまでの経験の蓄積に基づいているからである。

5.6 MSOR 分野における評価のための課題の設定には、かなりの専門的判断が伴う。

MSOR の筆記試験で問われる問題は、注意深く考えられなければならない。MSOR では、不適切な設問が課されてしまうと、学習者をその解答の結果で判定したくてもできなくなってしまう。設問は、レベルと内容が適切でなくてはならず、もしそうであれば、一定以上の能力を持ち、かつ準備を適切に行ってきた学習者であれば、MSOR の知識を使って与えられた時間内で問題を解くことができるであろう。設問は妥当な形式で出題され、曖昧さがあってはならない（あるケースでは、故意に曖昧に定義されていることもあるが）。このことは、もちろん全ての質問にあてはまるのだが、応用の場面で設定される問題解決を要求する設問であればなおさらである。

5.7 MSOR 科目の評価において、学習者が提出した答案の採点は、答えが一つに定まっているため特に専門的な判断が必要とされないと思われることがよくある。しかしたいい

の場合そうではない。例えば、学習者がモデル構築のプロセスについて説明しなければならない場合、あるいは、あるデータの分析結果を議論するとき、学習者の解答の専門的判断は、解答プロセスの評価に内在する本質的な部分なのである。学習者が提示した解答が、全く間違っているわけではないが若干の不備がある場合、あるいは明らかに正しい方法を使用しようとしているが成功していない場合、単位認定の際に、専門的な判断が要求されるのである。

5.8 MSOR の多くのタイプの評価では、周到に準備を行ないかつ能力も高い学習者に対する評価は通常、他の科目領域のものより高くなる傾向にある。とりわけ筆記試験ではそうである。完璧もしくはほぼ完璧な解答には、きわめて高い点数がつけられることが、他の学問領域より頻繁に生じるであろうし、逆に、MSOR における問題の特徴として、準備が十分でなかったり能力において劣っていたりする学習者にとっては、問題に全くといっていいほど手が付けられないといった事態も生じさせる。したがって、MSOR 分野の評価が、比較的広い範囲に分布することがよく見られ、これが学習者の能力の適切な相対評価につながるのである。教育機関は、MSOR の評価で見られる評価パターンが、他の教科のそれとはかなり異なるようだということが、MSOR での評価は、他の学問領域に比して、一般的な評価規則が適用され難いことを当然受け入れるであろう。

5.9 個々の学習者のパフォーマンスが、異なるモジュールによって非常に変化することは、この科目の固有の特徴である。学習者の中には、全てのモジュールに対してよい評価を得るものもいるかもしれない。しかし、最も優秀な学生でさえ、ある特定の MSOR の領域の理解が難しく、このことで、他では明らかによい成績をとるような評価結果（プロフィール）の中で、いくつかのモジュールにおいて低い評価しか得られないといったことが起こり得る。そのような場合は、学部学科の内外の評価者が、その学習者が優秀か否かを総合的に判断することもある。

5.10 同じように、なんとか到達のレベルに達しているような学習者は、しばしば多くのモジュールに不合格となり、わずかにいくつかのモジュールにのみ合格するような評価結果（プロフィール）となってしまうかもしれない。評価者は、そのような学習者が、その科目に対しある程度妥当な知識を提示し、許容され得るレベルへの達成にむけて前向きに努力していることを評価することにより、かなり低いレベルではあるが合格に値すると判断するかもしれない。

5.11 これらの理由で、MSOR 学習者の全般的な評価はしばしば、全体の平均点や複数の評価指標の組み合わせによってなされることが多い。そのような評価システムでは、全体的な達成度が学習者の成果として考慮される。個々のモジュールの合格は必ずしも全体の

パフォーマンスの評価には要求されないが、できればそうあるべきであるという見解は共有されている。5.8 で述べたように、MSOR では、一般的な枠組みでの評価規則は他に比べあまり適用しないし、それは専門家の判断を著しく侵害するものでもある。

しきい値レベル

5.12 5.1 および 5.11 で述べられた重要な点は、MSOR の「しきい値レベル」の評価基準を解釈するときに念頭に置かれなければならないということである。学習者は、全体としてそこで述べられている基準を満たさなくてはならないが、リストされている基準の一つ一つに合格する必要はない。

5.15 しきい値レベルに到達した卒業生の能力として、以下が要求される。

- 学習プログラムにおける基本的な知識を適切に理解していること
- 基本的な知識内での計算や操作をするスキルが適切なレベルであること
- この分野の核となる概念や原理をよく整備された文脈の中で応用し、そこで使われる技法やツールを適切に選択できること
- 論理的な推論の理解と、仮定と結論をきちんと示すこと
- 問題の数学的な定式化と適切な方法による解答の導出ができること
- 適度な正確性と明確性で、明瞭な推論と結論を示すこと
- 適切な基本スキルの提示
- 指導の下での専門的な業務の実施と、必要時の支援要請

標準レベル

5.14 5.1 および 5.11 で述べられた重要な点は、MSOR の「標準レベル」の評価基準を解釈するときに念頭に置かれなければならないということである。学習者は、全体としてそこで述べられている基準を満たさなくてはならないが、リストされている基準の一つ一つに合格する必要はない。

5.15 典型的なレベルに到達した卒業生の能力として、以下が要求される。

- 学習プログラムにおける主要な知識を適切に理解していること
- 主要な知識内での計算や操作をするスキルがある程度の高いレベルであること
- この分野の概念や原理全般を多少曖昧な状況においても応用し、そこで使われる技法やツールを適切に選択できること
- 論理的な推論を発展させ評価できること
- 問題の本質的な部分を抽象化させ、数学的な定式化と適切な方法による解答の導出ができること
- 正確性と明確性をもって明瞭な推論と結論を示すこと

- 適切な基本スキルの提示
- 独力での専門的な業務の実施と，必要時の支援要請

付録 A: 数学・統計学・オペレーションズリサーチ基準書の改訂メンバー

Professor Peter Giblin	Head of Departments of Mathematical Science in the UK
ピーター・ギブリン教授	英国数理科学学部長
Gerald Goodall	Royal Statistical Society
ゲラルド・グッダル	王立統計学会
Michael Grove	Higher Education Academy Subject Center for MSOR
マイケル・グローブ	高等教育機構 MSOR サブジェクトセンター
Dr Stuart Johns	Operational Research Society
スチュアート・ジョーンズ博士	オペレーションズリサーチ学会
Professor Duncan Lawson (chair)	Higher Education Academy Subject Center for MSOR
ダンカン・ローソン教授 (議長)	高等教育機構 MSOR サブジェクトセンター
Dr Niall Mackay	London Mathematical Society
ニール・マックアイ博士	ロンドン数学会
Professor Nigel Steele	Institute of Mathematics and its Applications
ニゲル・スティーアール教授	数学応用協会

付録 B: 数学，統計学，オペレーションズリサーチ基準書の策定メンバー

2002 年オリジナルの数学，統計学，オペレーションズリサーチ科目基準書による

Professor Rob Archbold	University of Aberdeen
ロブ・アーキボールド教授	アバディーン大学
Professor Russell Cheng	University of Southampton
ラセル・チェン教授	サウサンプトン大学
Professor Neville Davies	The Nottingham Trent University
ネビル・デイビス教授	ノッティンガム トレント大学
Dr John Erdos	King's College London
ジョン・アードス博士	キングスカレッジ ロンドン
Dr Judy Goldfinch	Napier University
ジュディ・ゴールドフィンチ博士	ネピア大学
Mr Gerald Goodall	Royal Statistical Society
ゲラルド・グッダール	王立統計学会
Mr Tony Palmer	De Montfort University
トニー・パーマー	ド・モントフォート大学
Professor Chris Robson (Chair)	University of Leeds

クリス・ロブソン教授	リーズ大学
Dr Stephen Ryrie	University of West England, Bristol
ステファン・リリー博士	ウエストイングランド, ブリストル大学
Professor Peter Saunders	King's College London
ピーター・ソーンダース教授	キングスカレッジ ロンドン
Dr Stephen Siklos	University of Cambridge
ステファン・シクロス博士	ケンブリッジ大学
Professor Joan Walsh	University of Manchester (retired)
ジョン・ウォルシュ教授	マンチェスター大学 (退職)

The quality Assurance Agency for High Education 高等教育質保証機構

Southgate House, Southgate Street, Gloucester, GL1 1UB, UK

Tel: 01452 557000, Fax: 01452 557070

Email comms@qaa.ac.uk

Web www.qaa.ac.uk

ISBN 978 1 84482 779 4

全ての QAA の出版物は当機構のサイト www.qaa.ac.uk にて入手可能

現在の出版物は下記にて入手可能

Linney Direct, Adamsway, Mansfield, NG18 4FN

Tel: 01623 450788, Fax: 01623 450481

Email qaa@linneydirect.com

登録チャリティー番号 1062746 及び SC037786

© The Quality Assurance Agency for Higher Education 2007:

www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/Subject-benchmark-statement-Mathematics-statistics-and-operational-research.aspx.

事例ベース ビジネス統計コースの展開
Developing Case-Based Business Statistics Courses

William C. PARR and Marlene A. Smith

The American Statistician, November 1998, Vol. 52, No. 4, 330-337

概要

これは、事例をベースとしたビジネス統計コースのガイドラインであり、具体的には事例ベースコースの利点と欠点、およびコース展開のために有益な資料を挙げ、事例ベース教育の指導者と授業にとってより良い特性を記したものである。

キーワード：事例，クラスディスカッション，コミュニケーションスキル，インタラクティブな教育，マネージメント，問題解決，統計教育

1. はじめに

ケーススタディは、多くのビジネスコース、(特にマネージメント、マーケティング、応用クラスなど)では使用されているものの、ケースメソッドはビジネス統計コースで広く採用されているものではない。統計学のクラスで事例の採用に消極的なのは、指導者がいろいろな種類のためらいや不確実性を感じているためと考えられる。トピックスを伝えるために、事例はどのように使われるのか。従来の方法と違い、何が事例を達成させるのか。事例ベースへの転換に、どれほどの努力が必要なのか、そして、どんなリソースが利用できるのか。著者自身による事例本位のビジネス統計コース開発の経験から、これらについて焦点をあててみた。

我々には、事例研究法の教育的有効性を無条件に受け入れる傾向がある。事例は、我々にとって役立つ。事例により学生は多様な解決方法を身に付けることが可能になる。第一に、事例はクラス的环境と教育のプロセスに影響をおよぼす。例えば、

- 事例を通して学生は伝統的でない問題解決法および自己啓発を学び、同時に論理的概念とその応用について、継続して理解を深めることができるようになる。
- 学生は事例を好むため、教室は生き生きと相互に作用し、刺激し合い、また、宿題や勉強を進んでやり、教室が和やかになる。
- ビジネスの学生は、他のクラスでケースメソッドを受けている場合もある。その場合も、このメソッドで彼らは楽しみを増し、より慣れていくであろう。

第二に、事例はカリキュラムの中で別のコースと統合でき、また、ビジネス社会の雇用ニーズに対応するものである。例えば、

*英語版原典の翻訳・転載については、*The American Statistician* より許可を得ている。
Copyright 1998 by The American Statistical Association. All right reserved.

- 企業は、グループで効果的に働ける卒業生を求めている。曖昧な事例の取り扱いについては、交渉やグループ意思決定等の練習のための貴重な機会を与え得る。
- 企業はうまくコミュニケーションのとれる卒業生を求めており、ケーススタディの最終成果はビジネスメモや口頭でのプレゼンテーションである。
- 事例分析の性質上、学生は事例の統計結果とビジネスの意思決定を結び付けなければならない。推奨されたビジネス上の意思決定や行動よりも、「最適回帰モデル」の答えは、経営者にとって役に立たないと見られることがある。

最後に、事例は現実を伝えるものである。学生たちは、職場の中で、統計的なメソッド応用能力への自信が高まったことを報告している

- 事例が多面的である場合、データの分析はしばしばつまらない計算の演習より、むしろプロセスを必要とする。それはマネージメントシミュレーションと並び、著者が意思決定のプロセスと、実際のプロセス開発の、よい処理法を学生に与えるために見出した最高のメソッドの一つである。
- 事例が、実在するデータを使って実際のビジネス問題を説明するとき、それはビジネス界での統計的手法の必要性を、より効果的に伝えるものとなる。

論文は、従来の統計コースについての苦悩で満たされている。例えば、Hogg (1991) が言うように、「我々の中には、統計学についての情熱と興奮を伝えられずに、統計研究の喜びと楽しみを失っている者がいる。これは、しばしば入門コースが重要視されない場合に起きる。教師の多くは無能で、向上のための努力が十分ではない。」 また Dallal (1990) によると、「統計の分野では、コースが役に立たない技術で終わってしまうため、挫折感を味わう学生がやたらに多い。そして、一生その勉強の興味を失ってしまう」。

たとえそうであっても、進歩はしている。最近の著者たちは、統計コース（特に導入コース）を、数学の特別コースとは違ったものとして扱うことが適切だと提案している。(Moore,1988; Moore and Roberts, 1989; Hogg, 1991; Watts, 1991; Radke-Sharpe, 1991; Snee, 1993); 応用されたデータの分析で、しばしば正しくない、または間違っている答えがあることを学生に認めさせること。(Watts,1991; Wild, 1994); 従来の講義形式以外の教室活動に従事させること。(Tanner, 1985; Snee, 1993; Bradstreet, 1996), そして、学生に統計結果を専門外の言葉で伝えるよう求めること。(Roberts, 1987; Gaudard and Hahn, 1991; Radke-Sharpe, 1991; Wild, 1994; Ledolter, 1995) 事例は非伝統型の教師を追求する効果的な手段である。

事例は、慣習的方式のはるか彼方にあるデータ分析の側面を実証している。例えば事例は、事例を解くためにどのメソッドを適用すべきか、ほとんど指示を与えることはない。いい加減なデータ（欠落、範囲外、誤記号さえ）が含まれている。データが必要な状況を説明し、独自の調査方法の工夫を学生たちに任せている。データ分析がもたらす経営上お

よび人道上の論点に取り組む。(例えば、統計的成果に疑念を持っていたり、データ収集を避けようとしたりする);そして、データの分析と報告で、時々生じる倫理的ジレンマを論証する。これらや、また、よく書かれている事例の特徴は、従来のコースで一般に見られたものより、データ駆動力型プロジェクトの、より現実的なシミュレーションを提供している。

ケースメソッドに不慣れな講師は、これらの目標を追求する時間が全くないと主張し、我々はそれを避けられないと考えている。これは結局学生に損な状況を与え、ひいては統計学の学問としての名声を脅かすものとなる。指導するメソッドの実践を学ぶ学生に、今は手助けをする時間がないというなら、いつその時間ができるのか。乏しい統計知識で、他者の指揮のもとで働きながら学ぶほうがいいのか？

医学生はまず死体で勉強し、それから熟練外科医の注視の中で手術を行う。海軍のパイロットはシミュレーターで訓練した後、経験を積んだ指導パイロットのもとで長時間の飛行をする。しかし、もし従来のビジネス統計コースのモデルを追従するなら、10代の若者には運転のヒントになるビデオを見せ、その後、車の鍵を渡してしまう。事例は、学生たちを彼ら自身の道に送り出す前に、データ分析が必要な、現実の、困難で曖昧な、そして有意義な場面を経験させる。

指導者はケースメソッドを自分で試した場合にのみ、その価値を認め始める。このことは、著者の双方が経験したことである。我々が事例ベース教育に取り組んだ時は、不安で、自信がなく、半信半疑だったが、かなりの年数をかけた後、この手法に大きな自信をもった。今ではそれを同僚に提唱している。

本稿は次のように編成されている。事例は、ビジネス統計コースで、学部から大学院レベルまでと入門からスペシャルトピックスコースまで、幅広く使われている。セクション2は、入門ビジネス統計コースと、大学院レベルの統計的プロセス制御コースの詳細である。事例によって、学生たちは数学だけではない他の方法に挑戦する。セクション3では、学生が事例ベースコースでどんな新しいスキルが必要なのか、また、それらのスキル展開のために、どのように支援ができるかを説明している。指導者は事例のために新しい一連のスキルが必要となる。セクション4は、不測の事態に何を予測し、いかに備えるかを述べている。セクション5には、コースの教材、セミナー、そして、コース展開のためのその他の資料がリストされており、最終的な考察は最後のセクションに記載する。

2. 事例の使用方法：二つのシナリオ

厳密には、事例とは何か。どのビジネス統計コースが最もケースメソッドに適しているか。教室で事例はどのように使われるか。これらの問題はテネシー大学と、デンバーのコロラド大学の、二つの継続している事例ベース統計コースの詳細な説明によって、このセクションに示している。

事例はあらゆる形と大きさとで現れ、時々、テキストの実例目的で使われる。事例は読者

に問題と解決を説明する。ここに、学生にシナリオとデータセットを与え、問題をよく考えさせ、解決策を練らして研究成果を報告させるための課題を用意する。従って、専門用語「ケース」は、実際的な、データ分析の実践的研究、そしてその分析に基づく結論のまとめと提案を表す。

2.1. ビジネス統計コース入門

デンバーのコロラド大学ビジネススクールは、何年間も学部と大学院の入門統計コースで、ケースメソッドを使っている。学部生と大学院生の成果に差はない。ここでは MBA/MS クラスが説明されているが、メソッドは学部と同様に入門統計コースが利用されている。学部のクラスで異なるアプローチがよりうまく働いた時、そのことは注目される。

本学の学生は年齢が高く、日中は働いていることが多い。大学院のクラスには、ほぼ同数の MBA と MS ビジネスの学生が在籍し、1 クラスは平均約 45 人である。大学院生が以前に統計コースを取ったかどうかわからない。学部のクラスは平均 40 人で、これらの学生は統計を学んでいない。

大学院レベルでは、従来の前期のトピックスには重きが置かれていない。ひとつかふたつの母集団からの推論では扱えない；確率は独立したトピックスとしては目立たず、推定法の基礎の実例が必要な時に、わずかに触れられただけである。（これらのトピックスは学部課程で今もカバーされている。）その代わりとして、大学院課程では、統一した方法論的テーマとして、回帰分析に重きがおかれている。コアとしての回帰で、学生の研究課題には、統計モデルの構築や解釈が関与する。時系列法（制御図を含む）、アンケートのデザインと分析、さらに他のトピックスもまた研究されている。実例となるシラバスは

<http://carbon.cudenver.edu/public/business/faculty/masmith.html>.

で入手可能である。

学生は、MINITAB、事例集、教科書または学内製の本の購入を求められる。学生は指導者の選択によって、16 週間の学期につき、だいたい 5 件から 10 件の事例を準備する。コース成績はこれらの事例のパフォーマンス、クラスでの事例ディスカッション、最終の研究課題、そしてクラスでのプレゼンテーションで判定される。クラス内でテストを行う指導者もいる。

（我々は、大学院レベルでは、従来の教科書をほとんど敬遠してきた。経験からして、従来の教科書が使われれば、学生の初歩的な情報源となる。事例の分析と提示をクラスの焦点にするため、従来の教科書の役割は重要視されない。）

事例の言葉の意味は 4 つのプロセスを含む：問題とデータセットの説明、分析、そして文書と口頭での結果の伝達。

1. 学生は背景説明と、機械可読形式のデータ セットを与えられる。学生がより成熟し、曖昧な状態の問題解決に慣れている大学院レベルでは、どのテクニックを適用するかについての指標が、ほとんど、または、全くない事例を選択する。事例で言明される問題は「最

後のスタッフミーティングで話し合われた問題点に焦点をあてて、レポートを準備せよ」のようなものであり、「利益と売上高の平均値と標準偏差を計算せよ」というようなものではない。入門統計学の学生は、方法論の選択に取り組みざるを得ない。クラスで何をやるべきかを彼らに教えたら、学生は自分自身の進め方に確信を持たなくなるだろう。学部生には更にガイダンスが必要になる。だが、学部生が方法論の問題と格闘するのも、非常に重要である。

(学生には方法論の問題をやる必要がある、というメッセージを強化するために、時々、最近の講義とは無関係な事例を割り当てる。先週の講義は時系列法を扱ったから、学生は、クロスセクションデータの実行結果を忠実に報告するかもしれない。学生は自分で適切な方法論を選ぶ責務があることを、すぐに理解する。これを最も効果的にするために、早期に、コースの間中、絶えずメッセージを送るべきである。)

2. データを分析する学生は、割り当てられる事例の性質が非常に曖昧で不確実なために、グループで学ぶことを奨める。

3. 学生は自分の分析の概要を文書で準備し、厳しいガイドラインが課される。

- ビジネスメモのフォーマットが必要で、プロトタイプは学期の初めに与えられる。
- メモには、提案と明確な限界についての記述が含まれている必要がある。分析するに十分でない場合；問題に戻って統計上の情報を明らかに関連付けなければならない。
- 統計などの専門用語は禁止され、統計的なバックグラウンドを持たない上司のため、レポートの準備が求められる。
- メモは標準的な文体と文法のルールに従わねばならず、スペルや文法の間違いで評価が下がる。
- 厳しいページ制限を強いられる。一般的にはダブルスペースでタイプした1～2ページ。

これらのガイドラインに従って準備されたレポートは、いくつかの目的にかない、学生は、職場にふさわしい標準文面でのコミュニケーションスタイルを学ぶ。学生に技術的な専門用語を一般の言葉に訳させることで、統計用語が本当に意味するものをしっかりと考えさせる。最後に、ページ制限は彼らの調査結果の要旨を慎重に考慮させる。彼らは無数のコンピュータのプリントアウトを吟味し、データから発見される重要なメッセージを捕らえなければならない。(ある審査員が指摘するように、学生には、多くの統計ソフトウェアパッケージからのコンピュータのプリントアウトは、その統計結果が、利用者向けの形式で表現されないことが、しばしばあると教えるべきである。彼らは自分の文書での分析と同様に、添付のアウトプットも注意深く管理すべきで、これは別のグラフィックパッケージを使用するか、統計的出力の広範囲な文書処理を必要とするかもしれない。)

4. 学生は、調査結果に関するクラスディスカッションを行う。ディスカッションはビジ

ネスミーティングの形をとり、学生は提案を求められ、分析を解説して、主張の正当性を示し、更にいくつかの目標が達成される。学生は口頭のプレゼンテーション技術を実践し、同じ問題への異なるアプローチを見る。また、論争が起きた時、彼らは自分のアプローチを正当化しなければならず、別のグループのアプローチや解釈が、彼らのものより優れていると認めざるをえない時もある。クラスディスカッションは、事例の難易度により、15分から1時間続く。

ケースを準備した学期の後、学生は独自の学習課題に着手する。最終的なコースプロジェクトは研究問題を限定し、データを集めて分析し、拡大したビジネスレポートで調査結果を報告する。コースプロジェクトは通常、多変量モデルの演習を含み、学生はこれらの最終プロジェクトを大きな自信をもって伝える。出発点となる初期のケースで、学生は、独自の研究に着手する能力に自信を持つ。これらのプロジェクトは、学期の終わりにクラスで発表される。

従来の統計入門コースを教えた人は、誰もがここで記述されているケーススタディアプローチのユニークさを認めるであろう。このコースは学生に、現実の問題を解決し、曖昧さに対処し、文書と口頭でのコミュニケーション技術を実践して、グループまたはクラスディスカッションやプレゼンテーションで他の人と交流することを求めている。事例解決のプロセスの間に、学生はどの技術を適用すべきか自分で決め、研究がもたらす結果を理解し、伝え、適切な補足説明をしなければならない。本質的に、ケーススタディメソッドは、我々のビジネスカリキュラムで無視されていたのけ者を、ビジネス学生の教育の統合された構成要素へと効果的に変えた。

2.2. 統計プロセス制御のMSコース（統計学 566）

ノックスビルのテネシー大学で、われわれは大学院コースで統計プロセス制御を教えている。そのコースは統計学のMSプログラムすべての学生が取っており、オペレーションズマネジメント、ロジステクス、または品質管理に強い関心を持つMBAの学生に、2年生の選択科目として取られている。

学生を混ぜた結果として年齢の双峰分布が生じ、統計学生のMSで、21歳～25歳、MBA生で、30歳～35歳となっている。1クラスは、20人から40人とさまざまである。統計プログラムのMSの学生は、一般的にいくつかの統計学の事前コースを終了しており、大学院の最初の学期で統計学 556 を取っている。MBAの学生は、しばしば5年から10年前に応用統計学の学部コースを取っている。そして、統合された最初の年のMBAプログラムの一部として、統計学の4時間コースと同等のものをもっており、基礎統計学、回帰（単・重）、統計的プロセス制御（基本チャート）、改良過程でのグループ化とSPCの使用）をカバーしている。

統計学 566 はプロセスの研究と、制御チャートの使用、そして関連するメソッドの改善に焦点をあてている。統計学 556 についての詳細は、ワールドワイドウェブ上で見られる。

(http://funnelweb.utcc.utk.edu/_wparr.)

クラスでカバーされる方法論的問題は、以下を含む；基礎的なチャート タイプの復習 (p,c,u,Xbar と R,Xbar と S,MR と X). プロセスの観点から、分散の構成要素の研究、サブグループ戦略、分析と計算上の区別、そして、その影響、許容差、分析能力、測定プロセスの研究、コントロールチャートと関連づけられたデータ、CUSUM チャート、EWMA、多変量制御チャート、シャイニン・メソッド (Shainin method) (多変量とプレコントロール)、制御チャート法の比較。

ほとんど例外なく全てのメソッドが、統計学 566 の特別な目的のためにインストラクターによって書かれたか、テネシー大学の Institutes for Productivity through Quality で使われたかのどちらかで、実際の事例を使いながら紹介されている。学生は、機械を使わずに大部分の分析をする。機械なしの計算が実行不可能や面倒な時は、指示は JMP (SAS 研究所の製品) を使って与えられ、それは、コントロールチャート分析の優れた能力を持っている。

コースの間に、学生は事例を利用していくつか異なることに従事する。例えば、

1. 文書の事例分析 一週間に約 1 回、学生は事例を与えられる。そのような事例は、バックグラウンドプロセスの解説とデータからなっており、2つのことが文書で求められる。1つは、1 ページのメモで、行動のための提案と、更なるデータ収集のための提案の可能性を要約し、経営者のために統計専門用語を使わずに書く。もう 1 つは、すべての計算結果と項目による、より詳細な分析の記事である。文書にした事例分析モデルの特別な変形は、5 つまたはそれ以上の段階に上げられるマルチフェイズケースによって提供される。これらは、第一段階事例で提供されるより、より幅広い窓口を通じて、問題解決とプロセス改善を例示する。時々、学生それぞれが事例分析を提出するが、そうでない時には、事例分析と記事を書くために 4~5 人のチームで作業する必要がある。(チームで事例分析をする場合の簡単なアドバイス；4~5 の要因によって評価のための時間を短縮することは大きなメリットであるが、最初に分析をする学生には十分な時間を与える。彼らは仕事を終えるために、おそらく、最低一週間の時間と、数回のチーム会議が必要だろう。

2. ケースディスカッション 学生は週に 1 回程度、次のクラス会議の討論用に準備するための事例を与えられる。これらの討論のために、文章資料を提出することはない。事例のクラス討議は、次のような質問をめぐって展開する傾向にある。(i) 何の分析が最も適切か、(ii) その結論は何か、(iii) 決定、実行、あるいは更なるプロセス研究のための提案は何か。ケースディスカッションは、コースの成績決定に含まれない。7 年以上の実証された多くの経験から、よく準備をし、活発に参加した学生はコースの成績もよい。

3. 誘導される事例 誘導される事例では、学生はクラスの前に事例を読み通すよう求められるが、詳細な文書は準備しなくともよい。その後、クラスでインストラクターが学生にケースツアーをさせる。クラスでリアル タイムに行われる分析を単純化するために、大量の中間計算がなされる。ケースガイドは新しいメソッドの説明にも利用する。統計学 566 には形式的な講義は非常に少ない。

事例を多く使うことで、このコースと伝統的な講義式のコースとの違いが実に明確になる。どんなに講義式のコースがクラス討論に馴染み易いとしても、インストラクターが、従来の役割から距離をおくと、学生は「よい回答」を出すことを、さらに自分のものにする。学生は講義ベースのコースの場合より、多くの新しい分析や問題を見つける。学生は、非常に積極的に反応する。ご意見コーナー（情報元：本コースホームページURL http://funnelweb.utcc.utk.edu/_wparr/Stat566.html）が示すように、彼らはコースに楽しみを見出している。MBA生とMS統計学専攻学生の混合クラスは、コースにとってはよい。MBA生は、すでにケースメソッドを熟知しており、同時に、統計専攻学生は大変しっかりとした統計学の学術的バックグラウンドをもって来るので、MBA生からケースメソッドとクラス討議について多くを学ぶ。

ケースメソッドが、無限に伸展するものではないことは明らかである。授業から考えてみると、例えば、ケースメソッドによる初年大学院生の統計理論の結果は、現在の、我々の想像を超えている。メソッドは、データで処理される経営問題として材料が存在するような、応用的なコースで、うまく作用する。

3. 学生のためのヒント

学生が統計コースを取るまでに、初期のケーススタディを経験しているかどうかはわからないが、彼らはおそらく、課せられたタスクの準備を全くしないだろう。これは特に、より伝統的な統計コースを取った学生の現実であるが、機転と励ましで適応させられる。

学生は事例分析のガイドラインを必要としている。そういった記事は多くある。大学のマネージメントかマーケティング部門に電話して、事例分析について学生に何を与えているか尋ねるとよい。コピーを借りられるかどうか、また、必要に応じて専門的研究ができるかどうかも尋ねてみる。大学院生はクラスの前に、おそらくケースメソッドの経験をしており、アプローチに対して満足するだろう。学部生にとって、ケースメソッドは全てが新しい経験であり、少なくとも最初のうちは、より多くの指導が必要である。事例の書き方やクラスディスカッションの方法を説明する時間をとる。

学生のライティングスキルが標準以下であるとわかったら、次のことが助けになる。最初に、学習ガイドを与える。他の事例文書の見本をつけて、書き方を明確にしたプリントを与える。コース教材の一部として、ビジネスライティングに関する参考資料を求め、作るかして、それらの材料からリーディングをシラバスに入れる。もし、学内にライティング研究室のような施設があるなら、コースの概略にそれを提示する。次に、これは最も重要なことであるが、彼らのライティングの善し悪しを直ちに、明確にフィードバックする。書面での報告書が必要な最初のレポートでこれを行えば、すぐに予想がたてられる。書面のフィードバックと事例の段階で、コミュニケーションと解釈がデータの分析と同じく重要であることを明らかにする。最初の事例の成績が文法や文体の間違いのせいで低かった場合は、学生の学業成績が劇的にレベルアップすることに注目している。

現在、多くのビジネスプログラムは、経営コミュニケーションコースを含んでおり、もし、そのようなコースがカリキュラムにあるならば、統計コースを必須科目に入れる。

学生は、ケーススタディに内在する曖昧さの扱いに難儀する 경우가よくある。曖昧さは様々な形でやってくる。ある事例は、データの分析が答えよりも多くの質問をもたらす。従来の統計コースを経験した学生は、「正しい答え」を出すことがタスクだと考えている。データセットの中には、それぞれ等しく良い洞察をもたらすものもある。彼らは、一つの「最高」に正しい答えを望む。コースは、経営問題を明らかにするために統計結果の使用を求めるので、数学は簡単に理解しても、統計結果の解釈や含意には力を尽くせない学生は、事例を生かせない。

(注意：評価を行う上で、この曖昧さが大学院生をかなり悩ます場合がある。その種の問題が起きるかもしれないことを、指導アシスタントとよく話しておくこと。さもなければ、自分を正当化する学生に、はなはだしい泣き言を聞かされることになる。)

教室で彼らについて話すことが、不安に打ちのめされた学生の助けになる。実生活の統計的洞察はプロセスを必要とし、1回限りの数学の演習ではないことを主張する。曖昧さがある場合、学生に「正しい」答えを与える気持ちを押さえること。代わりに、議論している取り組みの有利な点と不利な点の方向づけで、討論を導くこと。適切な時に生じた曖昧さは、クラス 討論と書面のフィードバックの両方で許容できるようにする。

最後に、学生は、席に座ってノートをとることに、幸せを感じていることがよくある。ケーススタディのクラスは従来の講義型コースよりインターアクティブな傾向にある。クラスの初日に、参加を期待しているというメッセージを送る。

4. 新インストラクターに役立つヒント

事例に基づくコースに移行するには、インストラクターとしての役割にいくつかの変化を必要とする。このセクションでは、2つの方向に沿って新インストラクターに役立つヒントを提供する。教室で異なることをする必要のあることと、舞台裏ですることである。

4.1. 教室内でのヒント

クラス規模はケース メソッドの成功の重要な指針となる。何人かのインストラクターが、大きなクラスでケースメソッドの成功した使用例を報告しており、著者はすでに95人以上で試した。より大きなクラスでは、通常、講義と朗読部分の両方を含んでいる。たとえば、週3回のミーティングのうち、1回は指導助手と一緒にして、復誦授業の間に事例をすると、都合がよい。これは、それにだけに頼らずにケースメソッドの実験が可能となる。小さなクラスでは、ケース討論をするのに素晴らしい環境を提供しているが、いつもとは限らない！15人~20人のより小さいクラスでは、熱心で、用心深い、そして熟考する学生が数名いなければ難しくなる。

我々の多くは数学が好きだったから統計学の分野に来たが、それは当然のことと言える。

本来備わっている、現実的で複雑なデータ分析の、非常な曖昧さによって、ケースメソッドは、学生に正しい答えの受け入れを押し付けられるかもしれない。それは、クラス討論の間中、学生の反応を意識的にランクづけしない努力が必要だ。次のようなフレーズの練習が役立つ。「私はそれに関してのあなたの推論を、完全に理解していない」や「なぜあなたが B 位置より優先して A 位置を選んだか、詳しく話してもらえませんか？」同様に、ケースディスカッションの終わりに、「その」正しい答えのために、自分に目を向けることを学生に期待しても、その衝動を避けること。当然、技術的に不正確なクラスで作られた発言は、どれも訂正したい。

もし、その事例のクラス討論の実行を決めたら、長い沈黙の時間に耐えることを学ぶ必要がある。これはなかなかぞっとする体験である。こちら側の最も危険な反応は、無言である。著者は時々次のようなことを試みる。突然の問いかけ、(自発的でない学生に対し、鋭い質問をする) 部屋の後ろに歩いて行って座り、質問して 2、3 分部屋を離れるか、あるいは、ひどく苛ついた態度で見回し、それから学生に取り掛からせる。

ケースメソッドではクラスのコントロールをあきらめなければならないことがあるが、これは難しい。しかし、自分の発言(全時間中の 75% 発言していたとする)が、授業全体のわずか 20% 足らずの内容しかとらえられていないと考えることが必要である。(クラスを録音してから、話しているインストラクターの他の誰かが時間を計りながら、少しの間、聴くことは役に立つ。) クラスは、与えられたディスカッションの何がストレスかを、ディスカッションによって解決する。クラスのコントロールは他の方法を通して発揮される；誰に呼びかけるか決める。よい事例を選ぶ、など。

4.2. クラスの他に

もし、文書ケース分析を考慮しているなら、これらは、解決した問題からなる普通の宿題より、評価にずっと時間がかかる。学生の事例文書は、質において大きく異なり、内容と文体の評価が求められ、指導助手のサポートは非常に役に立つ。評価を手伝っている学生は、一般的にそれを非常に役立つ経験と考えている。

優秀な採点者になることは難しい。ケースメソッドが求める採点時の問題は、とりわけ一定期間内に業績を上げないと終身雇用契約ができない人々には大きな問題である。成績をつける重い負担を緩和するための方法がある。第一に、学生に、事例の課題の一部、または全てをグループで作業させる。グループに関するよく知られた問題があるが、チームワークはビジネス教育の重要な要素になってきた。チームワークをよりスムーズに行わせる方法に関しては、マネジメント部門の誰かにヒントをもらえる。第二に、レポートの長さに厳しい制限をおく。がらくたの中から選び出すために、学生にページ制限を強いるだけでなく、評価の時間を短くすることもできる。著者の 1 人は、クラスで公然と、「私は最初のページしか読まない」と言っている。フォントサイズ、マージン、間隔を指定するのは便利なことであると気づくだろう。もしもこうした情報を提供しなければ、逆に聞か

れるだろう。(又は学生が、憶測を立てるのが役に立つものの、あなたの視力保持には非常によろしくないと思っているのに気づくだろう)。

文書ケースの評価に代わるものは、クラスディスカッション形式か、プレゼンテーション形式か、どちらかの口頭のクラス参加を評価することである。我々は両方の方法を試して、効果を認めた。いくつかの文書と口頭の組み合わせは、厄介な採点をなんとかできるようにしてくれる。

事例のクラス 討論は、時間がかかる。それは、進み具合を把握したり、前とは違う結論を導き出せている、との判断が難しいためである。適切な話題を扱うという問題は、特に入門ビジネス統計コースでは重要であるが、本稿の範囲を超えている。たとえそうであっても、組織内で先取りして、準備措置を取り始める。統計コースで、現在教えられているどんなトピックスが、流れに沿ったコースで必要なのか。どの程度の詳細情報が必要か。それは同僚にこれらの疑問を提示することになる。内容のカットの余地が無いことに気づき、同じトピックスでより効果的に教えることを考える。1つの例として、より関連のある数学的な計算を行うために、ソフトウェアに頼る必要がある。

もう1つの可能性は、より退屈な話題のいくつかを、後回しにすることだ。これは、必ずしも統計数学の授業を少なくするという意味ではない。むしろデータの分析が教室の第一の焦点であるのに対し、統計数学を2番目の位置に置くことを提案する。例をあげると、均一な確率分布を導入して、学生は講義を受けることなく、独自に正規分布の問題を読み解く。小テストは学生の参加を誘うために必要である。そうすることで、90分の講義を10分の小テストに代えた。

技術的な話題を後回しにする方法は、学生に十分な一連の短い材料を供給することである。その形式は、「A Short Note on the c Chart」又は「A Note on Analytical and Enumerative Methods」という、自分たち自身で読み、研究する形式である。解決すべき2、3の問題は、彼らが資料の理解を確かにする助けになる。

学生自身の時間を使ってテクニカルな技術を学ばせるこの方法は、ちょっとこわい。インストラクターとしては、学生が最初の授業を聞かずに学ぶことはあり得ず、また、学べないだろうと考えている。ある程度まで、これは真実である。講義のない教室なんてとても支持できない。どちらかといえば、日常的な作業は教室外に置いて、大事なデータ分析のディスカッションに教室の時間を使うことの方が、学生にとって非常に大きな価値があると考えている。

5. リソース

このセクションでは、事例ベースコースの展開を支援するための、さまざまなリソースを挙げた。リストは決して幅広いものではなく、我々自身の活動の中で有用と認めたものが含まれている。

5.1. 事例コレクション

次にあげるのは、第一段階か第二段階の学生のための、独立したケースブックのリストである。

- **Bryant and Smith (1995)** 50件の事例が2冊にまとめられており、1冊はビジネス統計コース向けに、もう1冊はより進んだ学生（学部の上級生もしくはマスターの1学期）向けになっている。新しい25件の事例の3冊目がこの秋に出る予定である。学生の入門書には、ライティング ビジネスメモに関する助言が載っている。ケースのクラス 討論の指導法を含む幅広い指導者ガイドである。

ほとんどのデータセットは実際のもので、事例の大部分はビジネス環境から得たものである。指導者ガイドにはデータディスクが付属しており、10以上の事例をチョイスしたオーダーメイド出版が可能である。

- **Carlson (1997)** この28件の事例は、実際のデータ セットと、ビジネスシナリオを使ったビジネス指向のものである。方法論の範囲は、記述統計から重回帰にわたる。これらの事例は学部のビジネスコースに適しているが、より高度な例は、MBAの学生の興味を引くかもしれない。また、これらの事例は統計的技術応用についての入門指導に役立つ。理論的背景と統計の基本原則は、時々事例の中にあり、データ ディスクには、指導マニュアルが付属している。オーダーメイド出版あり。

- **Chatterjee, Handcock, and Simonoff (1995)** この1冊に、初年度コースに必要なこと全般に渡る60以上の事例を含み、指導者版には、事例についてのコメントが載っている。学生版の事例のいくつかは完全に解決していて、学生は、コレクションにある、部分的、あるいは完全に解決した事例を試みる前に、自分で事例方法論を学ぶことができる。教育方法論で行われる随時の試みが、事例の技術的用語を明確にする。全てのデータ セットは現実のものである。2、3のデータ セットのみが、特にビジネス問題に適応している。学生版にはデータディスクが付属している。

- **Klimberg, Arnold, and Berger (1994)** 22の事例それぞれが、経営的な特色をもっている。名前と肩書きを持った人が、問題解決のために協力し合い、事例から異なる段階の指示をうけることになる。例えば、ある事例は柱状グラフを作るよう命じ、別の事例は単純にデータ分析をさせ、事例に存在する事理的をしばらくしながら、レポートの準備をするように指示する。指導者ガイドには事例の回答、授業ノート、MINITABのプリント、データ ディスクが含まれる。

- **Peters and Gray (1994)** 18の事例全てがビジネス応用に適応できる。事例ブックはより多くの指導ガイドを提供し、(特定の質問と、事例解決のための課題割り当て) Klimberg et al. (1994) の事例の様に、経営的な設定を含む。学生用の事例ブックには、データとディスクがついている。

運用管理と統計トピックスの両方を対象にした、ケース スタディ ブックがある。Bodily,

Carraway, Frey and Pfeifer (1996), Lapin and Whisler (1996) 統合定量法の指導者用に、回帰や予測のような統計トピックスと同様、シミュレーション、行列、在庫目録、輸送問題などの伝統的 OM トピックスを含んだ事例がある。

これらのケースブックに加え、筆者と Roberts (1991) は、学生の研究課題が豊富な事例研究の情報源になると見ている。Bryant and Smith (1995) のケースブックには、以前の学生の研究課題が多く載っており、現在、テネシー大学の統計プロセス制御コースでは、将来のクラスのために学生に適切な形式で事例を書かせている。多くの伝統的な教科書には、実際のビジネス場面からの実例が載っている。入門統計コースの適切な例としては Levine ラヴリン, Ramsey and Berenson (1995), Siegel (1997) など、多くがある。Leitnaker, Sanders and Hild (1996) には、かなり多くの事例があり、全てが様々に利用されるビジネス プロセス分析と改善のための管理図に焦点をあてている。Institutes for Productivity through Quality からの事例もテネシー大学に用意されているが、これらは現在、テネシー大学以外では利用できない。

5.2. 事例の教授とライティング研究会

年次大会は、ビジネススクールで統計学をより有効にすることをテーマにしている。一般的には、ビジネス統計コースを目標にした実際的訓練をとまなうケース メソッドによるティーチング セッションを含む。会議は通常 6 月に、国内の様々な研究所で主催され、Amstat News, the periodical of the American Statistical Association (1429 Duke Street, Alexandria, VA 22314)などで発表される。

決定科学協会 (Decision Sciences Institute) は、毎年 11 月の年次総会で、事例のライティングと指導のワークショップを提供している。半日セミナーは、初歩と経験のあるケースライターと指導者の両方に対応している。ワークショップは Decision Line (決定科学協会 (College of Business Administration, Georgia State University, Atlanta, GA 30303) の出版物)誌上で発表される。

おそらく、ケースティーチングの最良の情報源は、経験を積んだケース指導者を観察することだ。ほとんどのビジネス スクールには、このメソッドだけで教える指導者が少数いる。マーケティングやマネジメントの上級やシニアレベルのコースでは、しばしばケースメソッドが広範囲に利用される。教授に彼らの経験について話し、クラスで参観できるよう頼むと良い。その事例を教えている教授は、他の人と経験を分かち合うことを喜ぶ。

5.3. 他のリソース

多くの出版物がケースメソッドについて考察している。入門には、Christensen (1987) および Barnes, Christensen and Hansen (1994) がよい。出版目録にはクラスディスカッションの指導、学生の教室参加の実施、他の事例の関連問題を含むトピックスの引用が載っている。Christensen, Garvin, and Sweet (1991) は、初心のケース指導者のための、もうひとつの優れた

た情報資源である。ケース ベース教育の経験談を含んでおり、特に興味深いのは、ケースメソッドの、技術的材料の運用における非常に良い記事が、一つ含まれていることである。

ハンドブックや参考書は学生のビジネスメモ作成の助けになる。Baugh, Fryar and Thomas (1994) 及び Poor (1992) は、特にビジネスの学生を対象にしている。表題の長いリストの中の二つである。一般のライティングに関する参考書は、Strunk and White (1979) および Hodges, Horner, Webb, and Miller (1994) などがある。Siegel (1997) のビジネス統計学の教科書には、ビジネス レポートの書き方の章がある。Carlson (1997), Bryant and Smith (1995) のケーススタディ集には、書面のビジネスレポート作成の説明がある。

6. おわりに

事例ベースのビジネス統計コースを教える利点は何か？指導者と学生両方にとっての、より楽しい授業体験に加え、この方法はいくつかの重要な目標を達成し、学生は統計的方法で、仮想の、「フライトシミュレーター」タイプの経験をする。学生は書面と口頭のコミュニケーション技術を磨くことができる。統計的な意思決定と、その決定を他から擁護する体験から、能力と知識に自信を得る。学生はアクティブリスニングのような有益な技術を学ぶのである。

たとえそうでも、初期段階から、事例に基づくコースへの移行に関しては、多くの不安がある。事例指導には一般に認められる若干の欠点があり、それぞれについてコメントする。

- 学生たちが、うまく事例に参加するためには、書面と口頭でのコミュニケーション技術に不足がある。そうであっても、コースの期間中でなければ、彼らは将来、いつ上達するであろうか？
- 学生たちは完全にクラス討論に参加するには、消極的過ぎるかもしれない。これは、意見の違いを他人にはっきり言うことが無作法と考えられている、異なる文化の学生に特有の問題である。そういった場合、突然の質問は極めて効果的である。コースの間に突然質問される場合があると学生に最初の授業で注意を促す。もう一つのオプションは、彼らのコースレベルによってクラス参加を決めることである。
- 指導者は学生がケースディスカッションから何を学ぶかを悩む。よい考えや、議論で最も強い人の意見によって、議論から外れてしまう学生がいるのではないか？学生の作った主要なポイントをとらえるために、ホワイトボードか黒板を使うのは賢明な方法であり、議論から多少外れていても、インストラクターが経過を見守っているというメッセージを送るよい手段である。学生は、事例を中心とした授業のインストラクターが黒板に書くことを何でも忠実に記録する。インストラクターが、事例に基づく授業ではあまり話さないというまさにその事実によって、彼が1つ2つの簡潔な文以上のことを言うために前に進み出た時に、そのための時間が生まれるのだ。
- 事例とクラス 討論での学生の評価は極めて主観的である。これは、計算問題の宿題の評

価に慣れている我々数学者が持つ恐れである。特に事例ベース教育に比較的不慣れな指導者にとっては、誰がよく準備しているか、誰がクラスにより貢献しているか、ビジネスメモの採点をどう付けるかなどをつかむことが難しいが、これは経験で消える恐れである。すべての評価は平均値を明白にし、評価が学生全体で首尾一貫していなければならない。

コース開発は時間がかかり、神経の疲れる経験で、特に、従来の指導法に疑問をもつ場合がそうである。もし、事例法統計コースに向かうと決めたら、急がずに1クラスで1事例を試し、学生にフィードバックさせる。事例の効果について自身の反応を評価してみる。あなたがすぐに、ケースベースコース法を選択すること請け合いです。

本稿はケースメソッドの導入書であり、ケースディスカッションとライティングケースの指導について、更なる論文を準備中である。(Parr, 1997).

参考文献

- Barnes, L.B., Christensen, C.R. and Hansen, A.J. (1994), *Teaching and the Case Method: Text, Cases, and Readings*, Boston: Harvard Business School Press.
- Baugh, S.L., Fryar, M. and Thomas, D. (1994), *Handbook for Business Writing (2nd ed.)*, Chicago, IL: NTC Publishing Group.
- Bodily, S.E., Carraway, R.L., Frey, S.C., and Pfeifer, P.E. (1996), *Quantitative Business Analysis Casebook*, Chicago, IL: Richard D. Irwin, Inc.
- Bradstreet, T.E. (1996), Teaching Introductory Statistics Courses So That Nonstatisticians Experience Statistical Reasoning, *The American Statistician*, **50**, 69-78.
- Bryant, P.G. and Smith, M.A. (1995), *Practical Data Analysis: Case Studies in Business Statistics, vols. I and II*, Chicago, IL: Richard D. Irwin, Inc.
- Carlson, W.L. (1997), *Cases in Managerial Data Analysis*, Belmont, CA: Duxbury Press.
- Chatterjee, S., Handcock, M.S. and Simono, J.S. (1995), *A Casebook for A First Course in Statistics and Data Analysis*, New York: Wiley.
- Christensen, C.R. (1987), *Teaching and the Case Method: Text, Cases, and Readings*, Boston: Harvard Business School Publishing Division.
- Christensen, C.R., Garvin, D.A. and Sweet, A. (1991), *Education for Judgment: The Artistry of Discussion Leadership*, Boston: Harvard Business School Press.
- Dallal, G.E. (1990), Statistical Computing Packages: Dare We Abandon Their Teaching to Others? □ *The American Statistician*, **44**, 265-266.
- Gaudard, M. and Hahn, G.J. (1991), An Undergraduate Concentration in Applied Statistics for Mathematics Majors, *The American Statistician*, **45**, 115-120.
- Hodges, J.C., Horner, W.B., Webb, S.S. and Miller, R.K. (1994), *Harbrace College Handbook (12th ed.)*, Fort Worth, TX: Harcourt Brace College Publishers.

- Hogg, R.V. (1991), Statistical Education: Improvements are Badly Needed, *The American Statistician*, **45**, 342-343.
- Klimberg, R., Arnold, P. and Berger, P. (1994), *Cases in Business Statistics*, Boston: Allyn and Bacon.
- Lapin, L.L. and Whisler, W.D. (1996), *Cases in Management Science*, Belmont, CA: Duxbury Press.
- Ledolter, J. (1995), Projects in Introductory Statistics Courses, *The American Statistician*, **49**, 364-367.
- Leitnaker, M.G., Sanders, R.D. and Hild, C. (1996), *The Power of Statistical Thinking: Improving Industrial Processes*, Reading, MA.: Addison Wesley.
- Levine, D.M., Ramsey, P.P., and Berenson, M.L. (1995), *Business Statistics for Quality and Productivity*, Englewood Clis, NJ: Prentice-Hall, Inc.
- Moore, D.S. (1988), Should Mathematicians Teach Statistics? *The College of Mathematics Journal*, **19**, 3-35.
- Moore, T.L. and Roberts, R.A. (1989), Statistics at Liberal Arts Colleges, *The American Statistician*, **43**, 80-85.
- Parr, W.C. (1997), *Case Writing*, manuscript in preparation.
- Peters, L.H. and Gray, J.B. (1994), *Business Cases in Statistical Decision Making: Computer Based Applications*, Englewood Clis, NJ: Prentice-Hall, Inc.
- Poor, E. (1992), *The Executive Writer: A Guide to Managing Words, Ideas, and People*, New York: Grove Weidenfeld.
- Radke-Sharpe, N. (1991), Writing As a Component of Statistics Education, *The American Statistician*, **45**, 292-293.
- Roberts, H.V. (1987), Data Analysis for Managers, *The American Statistician*, **41**, 270-278.
- Roberts, H.V. (1991), *Data Analysis for Managers with MINITAB (2nd ed.)*, San Francisco, CA: The Scientific Press.
- Siegel, A.F. (1997), *Practical Business Statistics (3rd ed.)*, Burr Ridge, IL: Richard D. Irwin, Inc.
- Snee, R.D. (1993), What's Missing in Statistical Education? *The American Statistician*, **47**, 149-154.
- Strunk, W. and White, E.B. (1979), *The Elements of Style (3rd ed.)*, New York: Macmillan Publishing Co., Inc.
- Tanner, M.A. (1985), The Use of Investigations in the Introductory Statistics Course, *The American Statistician*, **39**, 306-310.
- Watts, D.G. (1991), Why Is Introductory Statistics Difficult to Learn? And What Can We Do to Make it Easier? *The American Statistician*, **45**, 290-291.
- Wild, C.J. (1994), Embracing the 'Wider View' of Statistics, *The American Statistician*, **48**, 163-171.

大学院生の統計的思考の育成：統計的思考の学習環境モデルの評価*

DEVELOPING POSTGRADUATE STUDENTS STATISTICAL THINKING IN UNIVERSITY: EVALUATION OF A STATISTICAL THINKING LEARNING ENVIRONMENT MODEL

Jing Sun, Nicholas Buys

Health Group and The Griffith Health Institute, Griffith University (AUSTRALIA)

j.sun@griffith.edu.au, n.buys@griffith.edu.au

要 旨

序論：統計教育は、実験研究に対して定量的な手がかりを実証するので、健康学科における研究者にとってはたいへん重要である。このプレゼンテーションでは、学生の統計的な推論や思考を発展させるように考案された、双方向の、大学における大学院レベルでの統計学課程のためのモデルを評価する。この研究のために使用されるモデルはSTLE（統計的思考の学習環境）と呼ばれ、学習の構成主義者理論によって支持されている。

構成主義者の学習に対する理論的取り組みでは、新しい知識や理解は既存の知識や考えに基づいており、経験の上に確立されているのだと主張している。私たちは行動から学んでいるのである。そして学ぶ際には、元々の知識はなくなる。つまり、新しい知識に統合される。最新の学習理論によると、良い教育実践は、学生に知識を構築するよう動機づける学習環境の構想で成り立っている。これには、学生が仲間と話し合ったり熟考したりするだけでなく、自分たちの学習についても考えたり推測したり、また熟考する機会を与えるような活動も含まれている。

方法：STLE のモデルは、学生の統計学に対する理解や統計的に考え判断する能力を高め、実際に統計技術を応用できるように開発された。STLE は教材、クラスにおける文化活動、討論、オンライン技術、教授法、そして評価が相互に作用する組み合わせとなっている。2010年のコーホートの90人の学生がこのモデルの6つの原則を評価するため研究に参加するよう勧められた。この原則とは、1. ツールと手順を提案することよりも、中心となる統計的考えを発展させることに重点を置く。2. 研究において仮説をたてたり検定したりする際には、学生を奨励するために、実際の、興味を与えるようなデータを使用する。3. 学生の推論や批判的な思考を発達させる手助けとなるような、クラスでの活動を使用する。4. 学生が自分たちの推論を検定し、分析データを調査し、彼らの統計学的な推論を発展させられる、適切な技術的道具の使用を統合する。5. 統計学的論点や、重要な統計的思考に重点を置いて支持されている論争が、話題にあがるようなクラス内での会話を促す。6. 指導計画と進展を見極めるだけでなく、学生が知っていることを認知し、彼らの統計

*英語版原典の翻訳・転載については、原著者より許可を得ている。

的思考の進歩をチェックできる評価を使用する。学生の研究や彼らのプロジェクトに関連した取り組みにおいて、学習や学習の応用を奨励するためには STLE 方法が有効であるということの評価するために、定性的と定量的、両方の取り組みをする混合型の研究方法が使用される。

結果：定性的インタビューと定量的調査の両方に回答した参加者たちは、STLE モデルが学生の批判的思考の発達を著しく高め、定量的な研究において学習と研究技術を高めたと述べている。

考察と結論：STLE モデルは統計学教育において効果的な方法であり、先に述べた 6 つの主義は、学生がデータを用いて推論したりテストしたりする、統計的推論を話し合ったり説明したりする、または、さらなる研究や取り組みにおいて統計学的原則を使用し実施することに重点を置くような授業を展開する際には重要な要素となる。

キーワード：教授；自発的学習；統計的思考；大学院統計学課程；構成主義

1 導入

大学院生への統計教育は、これが実験研究に対する定量的な手がかりを実証するため、健康学科の研究者にとっては極めて重要であった。健康関連の学科では、院生における定量的研究が、学生自ら学習を主導し、自立した研究計画を実行しなければならないという、課題を基にした学習環境にあることが多い。健康学科においては情報源が様々であり、学生に多角的な知識へのアクセスを提供する必要もあることから、統計学の講師が学生の研究の問題についてすべての質問に答えることは、めったにできないであろう。それに加えて、この情報時代の教育の目的は、学生に現実社会の問題解決するスキルを使えるよう準備させることである；しかし、教育はこの役割ができていないと批判され続けてきた。一般的に言われる理由は、大学における統計学の学習経験が現実の社会経験とはあまりにかけ離れているため、学生たちはこのふたつの環境 [1-4] の間でスキルを転換できないということである。現代の教育では、すぐに利用できる公式を提供することも、すべての健康学科の問題を解決できるような難問に統計学の手順を利用することさえも十分にできないのである。それゆえ、学生の学びは問題を基盤としたものであり、どうやって情報収集にとりかかるのかを教えるだけでなく、彼らに利用できる情報なのか熟考させるものであることがたいへん重要となってくる。これに基づくと、彼らが得たものは、ただの手順や情報以上のものになる。つまり、それは自分たちの研究や仕事で実行する能力へと応用できるように、学生の思考、行動、実行に情報を与え変化させる知識であるべきである。

問題を基盤とした統計学習は、統計学の概念を理解するという点で複雑性に満ちている。

したがって、学生が批判的思考をしていなければ、問題基盤の学習で与えられた意味解釈の過程は立証されないし、学習は妨げられる [5]。さらには、学習者として、後には常に技術を向上させる職業分野での専門家として、様々な状況に対してどのように反応するのか彼らが熟考することが重要であり、その結果彼らは統計学的手法を使いながら批判的思考や思慮深く、よく考えた上での行動を通して問題に取り組むことができるのである。

構成主義者の研究方法に従った学習に対する現代の見解では、学習には熟考と抽出を通して概念の構造を組み立てることが必要である。それぞれの学習者は自分の知識を構築しなければならないので、概念は言葉によって教師から生徒に伝えることはできない。学習は、学習者が積極的に概念 [2] の構築と組織に関わった時のみ起こるのである。統計学を教える際には、構成主義では、人はどう学ぶのかに私たちの注意を集める。そして、統計学の知識は、数学の問題や環境に積極的にかかわって生まれてくる疑問や取り組みに応じたモデルを形成している人々から結果が生じているのであって、ただ単に情報を取り込むことから生まれてくるものではないと述べている。私たちは行動から学ぶのである。学んだ時、以前得た知識は消えない。そして新しい知識と統合するのである。さらに、知識は学習者が経験し、知識 [2, 6] を構築する背景と環境に大きくかかわっている。言い換えれば、理解は経験によって形成されるのである。構成主義者は、実際の活動における認識の経験を力説する。本物であるということは、背景が現実社会の取り組みであるという意味ではない。むしろ、学習活動は実社会での慣例 [2] を表すようなタイプの取り組みを採用すべきである。

構成主義者の学習に対する取り組みが示すものは、良い教育実践とは、学生が知識を構築する動機づけとなるような学習環境を設計することにあるということである。これには、仲間と話し合ったり考えたりするだけでなく、自分たちの学習を検討し推論し熟考する多くの機会を学生に与える活動も含まれている。Jonassen, et al. [7] によると、教育の目的は、例えば批判的思考のような、より高いレベルの認知技術を教えること、そしてそれらの技術を応用する気質を育むことである。これを達成するための緻密な知力には、推論、判断、批判的思考、複雑な状況 [8, 9] の分析も含まれている。この教育目標を達成するために、大学院生の統計学課程教育に対して現代の学習理論が示唆しているものには、知識を構築し、推論し、学習の時に熟考するよう学生を刺激し、動機づけをし、影響を与える学習環境を設計することが含まれている。この教育の取り組みは、学生を引き入れ、自分たちの説明、評価、コミュニケーション、そしてこういった経験を理解するのに必要な統計モデルの応用を支持するような経験を生み出すことである。

この論文は、学生の統計学的推論と思考を発達させるよう設計された大学における、双方向の大学院レベルの統計学課程のためのモデルを評価する。この研究に使用されるモデル

は、「環境学習の統計的思考の学習環境」(STLE) と呼ばれ、学習における構成主義者理論のうえに構築されている。STLE モデルは、教育法、教材、クラスの活動、文化活動、討議、オンライン技術と評価が相互作用するように組み合わせられたものである。これには、仲間と話し合ったり考えたり、そして個別指導時間の取り組みに対する意見・感想を出すことで、学生に努力を促したりするだけでなく、考え推論し自分たちの学習について熟考する多くの機会を提供する活動もある。

教える過程では、学生の思考は研究の問題から始まり、指示によって導かれ、仲間のグループ活動に参加する機会と同様に自分たちの知識を構築する機会も与えられる。これは、学生の統計学の理解や、統計的に考え推論する能力、そして統計技術を実際に応用するために開発された。

この研究で主に提起されている問題は、STLE 教育法は学生の批判的思考を発達させ、研究に着手する能力を増加させるのかどうかである。

2 方法論

2.1 方法

統計学での学生の批判的思考を促進させ、研究に着手する能力を増大させ、統計学課程において社会的相互作用を促進させることについて、STLE の方法の有効性を評価するために縦断的デザインと混合型取り組みを、この研究では採用した。統計教育における STLE の方法実行の過程での学生の意見を調査するために、個別のインタビュー形式が使用された；そして調査は、STLE の方法に関するデータと、学生の批判的思考と研究を行う能力の増大との関係性について、そして社会的相互作用についてデータを集めるために利用された。

2.2 参加者

一流大学である大都市の大学の統計学課程の学習が終了するあたりで、学生たちから被調査母体が選ばれた。これらの参加者は、理学療法、医療科学、福祉事業、公衆衛生、看護、産科学、薬学、歯学を含めた保健指導の分野において優等生、あるいは大学院生であった。定性的データの収集においては手ごろな被調査母体を使用した。前期から 5 人の学生と現在の学期から 5 人の学生がインタビューに応じた。優等学生と大学院生たちは、個別に対応され、手書きでアンケート用紙に記入するよう依頼さるか、もしくは、サーベイ・モンキー・ウェブサイトのオンライン経由で記入を完了するよう e-メールで指示された。90 人の学生の被調査母体が得られた。全体で、定量性調査での回答は合計 34 人分の使用可能な回答があった。オーストラリアでは、ほとんどの地方の教育機関にはよくあることだが、回答者の 60 パーセントがオーストラリア人の国内在住者であり、40 パーセントは中国、インド、インドネシア、マレーシア、台湾、香港、スウェーデン、カナダ、サウジアラビア

など国々から来た留学生たちであった。学校中でも大体同じ位の数に相当しており、性別による割合もほぼ同じであった。

2.3 「統計的思考の学習環境」の設計

STLE モデルは、学生が統計学を理解し統計学的に考え推論する能力を高め、実際に統計技術を応用できるように開発された。STLE は、課程の教材、授業での活動文化、討論、オンライン技術、教授法と評価を相互作用できるように組み合わせたもので、このモデルには 6 つの原則がある：1. ツールと手順を提案することよりも、中心となる統計的思考を発展させることに重点を置く。2. 研究において仮説をたてたり検定したりする際には、学生を奨励するために、実際の、興味を与えるようなデータを使用する。3. 学生の推論や批判的な思考を発達させる手助けとなるような、クラスでの活動を使用する。4. 学生が自分たちの推論を検定し、分析データを調査し、彼らの統計学的な推論を発展させられる、適切な技術的道具の使用を集約する。5. 統計学的論点や、重要な統計的思考に重点を置いて支持されている論争が、話題にあがるようなクラス内での会話を促す。6. 指導計画と進展を見極めるだけでなく、学生が知っていることを認知し、彼らの統計的思考の進歩をチェックできる評価を使用する。

2.4 「統計的思考の学習環境」の測定

定性的なデータ収集においては、STLE の方法についての学生の意見や、彼らの批判的思考にそれがどの程度関係したのか、また、彼らの研究プロジェクトを始める能力がそれによってどの程度向上したのかを明らかにするため、幾分系統立てたインタビューの質問がされる。定量的データ収集においては、STLE と、学生の批判的思考と研究を始める能力の向上、そして社会的相互作用との間の関係を把握するために STLE 方式のアンケートが使われる。

これは、4 つのスケール（批判的思考教授法、問題基盤の課程設計、情報収集を基本とした課程内容、生徒の学習を刺激する方法）に沿った、STLE の取り組みを測る 20 項目の自己申告型のアンケートである。参加者は、6 点式 リカートスケールによって、それぞれの項目に自分たちの賛成の度合いを点数化する：1 確かにそうである 2 幾分反対である 3 少し反対である 4 少し賛成である 5 ある程度賛成する 6 大いに賛成である。この手段の有効性において、4 つのスケールの信頼性に対するクロンバックの アルファはそれぞれ、0.89,0.86,0.85 と 0.82 であった。この手段の有効性をさらに支持するため、20 個の質問によって説明される全体的な分散を検定するために因子分析が行われた。20 個の質問は STLE の取り組みに対して 78.18% の高レベルの分散を説明していることがわかった。高レベルの信頼性と有効性はそれ自体 STLE の取り組みを測るものとして有用であることを示してきた。そして、このようにして調査の質問に答えるため、現在の研究においても採用

された。STLE のスケールに加え、学生の批判的思考や研究に着手する能力の向上、そして社会的相互作用についての質問もまた、6 点式 リカートスケールで尋ねられた。

2.5 分析

STLE の取り組みに対する学生の意見のパターンに見識を得るために、学生の批判的思考や研究に着手する能力の向上、そして社会的相互作用において 3 つの質問が STLE の 4 つのスケールで行われた。学生たちが、この STLE 環境においてどのように批判的な思考を発達させているのか突き止めるため、STLE と、批判的思考、学習の発展、社会的相互作用の促進におけるそれぞれの不定の結果との関係を分析するために、回帰分析が使われた。

2.6 結果

回答者のスケール上の点数は、それぞれのスケールを構成する項目に対する平均点によって計算された。

表 1 は、総合的な点数と性別に分けた点数における、それぞれのスケールに対する平均値と偏差値の概観を示している。

表 1 : STLE における総合的点数と性別で分けた点数, 批判的思考, 学習, 社会的相互作用

	総合的点数	男子	女子
	M(SD)	M(SD)	M(SD)
批判的思考教授法	5.18(0.57)	5.16(0.58)	5.19(0.58)
問題基盤の課程設計	5.03(0.78)	4.70(0.89)*	5.27(0.61)
調査を基本とした課程内容	5.08(0.56)	5.09(0.51)	5.08(0.61)
学習を刺激する教授法	5.70(0.46)	5.64(0.50)	5.75(0.44)
批判的思考	5.26(0.71)	5.07(0.73)	5.40(0.68)
学習の発展	5.20(0.84)	5.07(0.91)	5.30(0.80)
社会的相互作用の獲得	4.91(1.02)	4.85(1.09)	4.95(0.99)

表 1 は、熟考方法、問題を基盤とする課程、調査を基本とする課程内容、学習意欲を刺激する方法の平均点において、一貫して高得点を取っていることを示している。学生の批判的思考の発達や研究に着手する能力の向上もたいへん高くなっている。これは、この課程で学生の社会的相互作用が 5 点以下のものの比較となっている。男子学生と女子学生とでは、問題を基盤とした課程において統計的に重大な相違があり、この要因が問題を基盤とした課程設計のスケールと、学生の批判的思考、研究に着手する能力の向上と社会的相互作用の向上との関係についての回帰モデルとして統制されるだろう。

表2はデータセットから生じた4つのスケールの相関関係を表している。

表2. STLEの取り組みの4つのスケール感の相関係数

	教授法	課程設計	課程内容	学生の動機づけ
教授法	1			
課程設計	0.62***	1		
課程内容	0.77***	0.64***	1	
学生の動機づけ	0.36***	0.47***	0.48**	1

** P < 0.01, *** P < 0.001

この表は4つのスケール間の中位から高い相関関係を示している；とりわけ批判的思考教授法と問題を基盤とした課程設計の間に高い相関関係が示されているように、特に注目すべきは、これらの数字が、この構成物が環境理論を学ぶ構成主義者にとって重要であるという理論上の仮定の裏付けとなる点である。；そして、批判的思考教授法と調査を基盤とした課程内容に示されるように；また、批判的思考教授法と学生の学習を動機づけする方法との間にあるものと比較されるように

表3は、STLEの4つの各スケールと、学生の批判的な統計学思考の発達、学習の向上、そして社会的相互作用との関係における、回帰分析の結果を表している。学生の批判的思考の発達と回帰分析における学習の向上と、STLEのそれぞれに対する取り組みとの関係を分析するために、STLEの基準のそれぞれの側面に対して合計点が使用された。0.05未満の確率レベル、0.20を越える回帰係数、そして0.40を越える相関関係は統計的に重要な影響を持っているとみなされ、その結果は以下の表に示されている。

表3 STLEの取り組みと学生の批判的思考、社会的相互作用の発達の間関係

変数	学生の批判的思考と定量的スキル				学生の学習と研究			
	回帰係数 (95%CL)	相関係数	P値	R ²	回帰係数 (95%CL)	相関係数	P値	R ²
教授法	0.99 (0.70-1.21)	0.79	0.001	61%	1.71 (0.86-1.47)	0.81	0.001	64%
課程設計	0.59 (0.34-0.84)	0.65	0.001	42%	0.73 (0.45- 1.02)	0.68	0.001	46%
課程内容	0.85 (0.51-1.20)	0.68	0.001	46%	1.08 (0.69-1.49)	0.71	0.001	51%
学生の 動機づけ	0.80 (0.32-1.27)	0.52	0.01	27%	0.86 (0.28-1.44)	0.47	0.01	22%

表3 続き

学生の社会文化的経験

変数	回帰係数(95%CI)	相関関係	P 値	R ²
教授法	0.67(0.14-1.19)	0.43	0.01	18%
課程設計	0.28(-0.19-0.4)	0.21	0.28	4%
課程内容	0.68(0.05-1.31)	0.37	0.04	15%
動機づけ	0.02(-0.78-0.82)	0.01	0.97	0

注. P<0.05 は統計的に有意な結果を示している.

上の表に示されるように、批判的思考教授法、問題を基盤とした課程設計、調査を基盤とした課程内容、そして学習を促す手法も含めて STLE の取り組みのあらゆる局面が、学生の批判的思考や研究に着手する能力の向上と著しく関連している。これは、STLE のそれぞれ 4 つのスケールと、この課程の学生の社会的相互作用の間関係を、低レベルから特に重要性のない関係まで比較したものである。

定性的データは、この取り組みは、批判的思考を明確に促進していると学生自身が感じているという、さらなる証拠を与えている：

・「私は使用されている取り組みは良いものだと思います。行動の仕方について、私は以前より好奇心が旺盛になったかもしれません。それに、物事に対して違った角度から取り組み、研究、特に定量的研究で何をすべきかさらに徹底的な理解ができたと思います。また、生徒の参加や、この課程の他の学生の参加を促すことについては、私たちの課程には様々な背景があり、この取り組みはたいへん興味深く、クラスでの討論も白熱します。それで、クラスの人たちはみな自分の研究計画に取り組み、それでみんなが学んだことを自分の計画に関連付けるのです。そして今度は問題を提起するのです。だから、とても面白いです。ひきつける授業でしたし、そのトピックについて誰もが実に知識欲旺盛なので多くの学生が参加しました。」

・「私にとって、定量的研究で使われたその取り組みは、定量的研究の技術に関する知識を学ぶのに、たいへん役に立ちました。また、私はこの知識と技術を、将来行う計画である研究プロジェクトで試してみるつもりです。」

定量的データに関する熟考はまた、STLE の取り組みが、生徒の学習を高め、研究と調査のスキルを向上させたことも示している。

・「はい。もちろん、そうです。申し上げたように、私はこの課程を取る前は研究に関する経歴はなにもありませんでした。そして、当然私の研究技術はとても向上しています。今ではどんな論文でもレポートでも読むことができますし、それを判断することができます。」

批評もできますし、会議で発言だってできます。結果の妥当性や信頼性も... そうなんです。とにかくこの取り組みは実際良いのです。」

・「はい、この取り組みによって私の研究スキルは目覚ましく改善されました。まず、物事を容易に分析できるという自信ができました。また、定量的研究の明確な理解ができましたし、問題をどのように処理するのかという、明確で簡潔なプランを企画する能力ができました。私が毎日学んできた技術を利用するため、これは本当に私にとって役立っています。」

・「これによって私の研究技術は確かに改善したと信じています。この課程に入る前は、わたしには研究の経歴は全くなかったため、定量的研究について何も知りませんでした。だから、もし自分が定量的研究をしたいと思えば、研究をどのように計画すべきか、データをどうやって集めるのか、そのデータをどう分析したらいいのか私にはわかりません。自分が研究技術を高められたという点で、私はこれが実に効果的であると思います。」

・「私は学ぶと同時に実践もしたので、私にとって、とてもためになると思います。」

3 議論と結論

現在のデータでは、STLE は大学院生の批判的統計思考が一般的に増加するのに役立ち、学習を高めたということを示している。思慮深い教授法、問題を基盤とした課程設計、そして調査に基づいた課程内容が、学生の批判的思考を著しく向上させ、学習と研究技術の発達を高めたということが見出された。前の項でも述べたとおり、この大学の課程の学習プログラムにおける健康関連の学科の優等生と大学院生は、この特定の統計学課程を取っている。課程設計と課程内容は、彼らの研究方法、大学院研究プロジェクトと課程の研究に特有であり、個人の研究と定量的スキルに関連のある側面をさらに論じている。学生の批判的思考はこれらの問題を基盤とした、また調査を基盤とした統計学習を通して著しく向上したことを示している。

STLE の取り組みと学生の批判的思考発達とのハイレベルな関係を比較すると、STLE の取り組みの4つのスケールと、学生の社会的相互作用との間の関係は低い。このことは、統計学の学習は、社会的相互作用 [1, 2] というよりもむしろ、統計学の背景と環境において起こるためなのかもしれない。これは、認識の取り込みという点において道理にかなっている。Jonassen, et al. [7] は、学生の思考は問題の研究に始まり、次に問題を熟考するための討論が続くと述べている；研究の問題提示、討論と熟考は、思考の連続で成り立っており、定量的研究においては、最新の研究から批判的思考と思慮深い行動へと導くと述べている。この研究では、STLE の取り組みによって、学生は統計的な考え方や経験に基づく分析方法について、意味を構築することができるということを明確にしている。このことは、中心となる学生の統計的概念、学生が研究で仮説を立てて検定するという取り組みができるような本物の動機づけができるようなデータセットの使用、学生が自分たちの推理を検定し分

析データを調査し彼らの統計学における思考や推論を発展させる統計ソフト使用の統合などを通して起こるものである；重要な統計的考えを熟考するクラスでの会話を促すこと；そして最終的に実際の調査データセットを学生が知りえたことを理解し、彼らの統計的思考の発達を見守るための評価へと組み入れること。このようなことから、STLE の取り組みはかなり効果的と言える。

参考文献

- [1] Brown JS, Collins A, Duguid P. Situated cognition and the culture of learning. *Educational Researcher*. 1989; **18** (1): 32-42
- [2] Duffy TM, Jonassen DH. Constructivism: New implications for instructional technology. In: Duffy TM, Jonassen DH, eds. *Constructivism and the Technology of Instruction: A Conversation*. Hillsdale, Nj: Lawrence Erlbaum Associates 1992:1-16.
- [3] Hiebert J, Carpenter TP, Fennema E, Fuson K, Human P, Murray H, et al. Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher*. 1996; **25** (4): 12-21
- [4] Sternberg ER. Teaching critical thinking: Are we making critical mistakes? *Phi Delta Kappa* 1985; **25** (4): 12-21
- [5] Song HD, Grabowski BL, Koszalka TA, Harkness WL. Patterns of instructional-designfactors prompting reflective thinking in middle-school and college level problem-based learning environments. *Instructional Science*. 2006; 34 (1): 63-87.
- [6] von Glasersfeld E. Introduction: Aspects of Constructivism. In: Fosnot C. T, ed. *Constructivism: Theory, Perspectives, and Practice*. New York Teacher College Press. 1996:pp. 3-7
- [7] Jonassen DH, Peck KL, Wilson BG. *Learning with Technology: A Constructivist Perspective*. Upper Saddle River, NJ: Merrill/Prentice Hall 1999
- [8] Savery JR, Duffy TM. Problem based learning: An instructional model and its constructivist framework. *Educational Technology*. 1995; **35** (5): 31-8
- [9] Williams B. Developing critical reflection for professional practice through problem-based learning. *Journal of Advanced Nursing*. 2001; 34 (1): 27-34
- [10] Greeno JG. On claims that answer the wrong questions. *Educational Researcher*. 1997; **26** (1): 5-17

統計教育における評価と指導方法に関するガイドライン

大学レポート*

Guidelines for Assessment and Instruction in Statistics Education (GAISE)

College Report

Martha Aliaga, George Cobb, Carolyn Cuff, Joan Garfield (chair), Rob Gould, Robin Lock, Tom Moore, Allan Rossman, Bob Stephenson, Jessica Utts, Paul Velleman, Jeff Witmer

概 略

アメリカ統計学会 (ASA) は、統計教育における評価と指導方法に関するガイドライン (GAISE) 事業に資金を提供した。この事業は2つのグループから成り立っており、1つは K-12 教育（訳注 高等学校までの教育）に重点を置き、もう1つは大学の入門コースに重点を置いている。このレポートはその大学グループによって作成された提案を示している。

レポートでは、大学の入門コースの歴史を手短に述べ、ジョージ・コブ氏による 1992 年のレポート⁽¹⁾ について概説している。コブ氏のレポートは、これらのコースを教える際の、一般的に認められた提案事項であると考えられている。入門コースの現在の状況として、入門コースの教育についての調査結果をまとめている。その後、私たちは、統計的リテラシーを持つことを目指して学生の到達目標を提示した。コブ氏のレポートで提案された内容を基本として、統計学入門コースを教えるための 6 つの提案を表した。この 6 つの提案は次のようなものである：

1. 統計的リテラシーを重視し、統計的思考力を育成する
2. 本物のデータを使う
3. 単なる手順に関する知識よりも、概念の理解に重点を置く
4. クラス内での活動的な学習を育む
5. 概念の理解とデータ解析を向上させるために情報機器を活用する
6. 学生の学びを向上させたり、見極めたりするような評価を用いる

レポートは、どのようにこれらの変革を行っていくのかという提案と、提案を詳細に例示するため、付録に数多くの例を載せて締めくくっている。

*英語版原典の翻訳・転載については、The American Statistical Association より許可を得ている。
Copyright 2010 by The American Statistical Association. All right reserved.

序論

GAISE プロジェクトは、K-12 カリキュラムにおける統計学の指導と評価、そして大学の統計学入門コースのための ASA 推奨のガイドラインを発展させるため、2003 年 ASA から選ばれたメンバーによる新しい試みで始められた助成金によって、資金が提供された。

私たちの大学課程のガイドラインに対する取り組みには、e-メールによる討論や個人的な小グループの集まりでの討論も含まれていた。私たちの討論は、現行の基準やガイドライン、統計の指導法や学習法に関する調査結果、統計的推論の根底にある重要な概念についての指導や評価に焦点を置く必要性についての最新の議論や提案を調べることから始まった。

入門コースの歴史と発展

現代の統計入門コースをずっとさかのぼっていくと、統計的手法についての初期の本が起源となっている。1925 年にその初版が発行された R.A.フィッシャー、「研究者のための統計的方法」は実践的な科学者たちを教育するのが目的だった。その 12 年後、ジョージ・スネデカーの「統計的方法」は、同じ内容を発展させたものあったが、対象とする読者に少し違いが見られた。フィッシャーの本より、スネデカーの本は、まだ学位を取得中の未来の科学者たちのための課程における教科書であった；統計学はかなり実務的で職業を主眼としているにもかかわらず、学術的な科目として確立され始めていた。1961 年までにはフレッド・モステラー、ロバート・ルーク、ジョージ・トーマスらによる、「確率の統計的な応用」(Probability with Statistical Applications) の出版に伴い、統計学はより幅広い学術的カリキュラムへと発展したが、ここでまた問題点が生じた。この早い時期には、統計学はその妥当性について、確率に大きくよりかからなければならなかった。

1960 年代の終わりごろから 1970 年代の初期の間に、ジョン・チューキーの探索的データ解析の考えが、カリキュラムに革命的と言ってよい、ひと組の変革をもたらした。つまり、ある種類のデータ解析の手法を確率に基づいたモデルから切り離すことで、データ解析が独立した知的活動としての地位を獲得し始めたこと、そして大きくて扱いにくい計算機につきっきりで何時間も過ごす必要なく、学生がデータ解析ができるように、間に合わせのデータツールを集めたものを導入したことである。コンピューターが後に、初等的な統計カリキュラムにおける「データ革命」を完全なものとしたが、チューキーの「探索的データ解析」(EDA) の考えが最初の技術的な躍進と、作り物の例を避ける新しい精神をもたらした。

大きな影響を及ぼす 2 つの本が、1978 年に現れた：デヴィッド・フリードマン、ロバート・ピサーニ、ロジャー・パーヴズによる「統計学」、とデヴィッド S. ムーアによる「統計：概念と論争」。この 2 冊の本の出版は、私たちが現代統計学入門コースとみなしているものの誕生を記している。

内容の進化が他の傾向と同時進行してきた。このうちのひとつは在籍者数の著しく、そ

して安定した伸びを示している点である。この点について、ふたつの統計がここにある。

- ・数理科学会議によると、2年制の大学では、統計学の登録者は、1970年の微積分学の登録者数の27%であったものが2000年の微積分学の登録者数の74%へと拡大した。
- ・統計学のAP(Advanced Placement)テストは1997年に初めて実施された。その最初の年にそれを受けた生徒は、7,500人いた。これは、APテストにおける他のどの教科の実施初年の受講者数よりも多かった。翌年には、15,000人の生徒がそのテストを受け、その翌年は25,000人、そのまた次の年は35,000人であり、2004年には、65,000人以上の生徒が統計学のAPテストを受けた。

コースの内容変化と登録者数のめざましい増加の両方によって、受講者の背景、興味、意欲が多様化し、拡散するという三番目の変化をもたらした。統計学は、農学と生物学の未来の科学者たちという狭いグループを対象としたスネデカーの本のような教科書を用いて教えられるコースから離れて、中学校から学士以降まで、あらゆるレベルの生徒に教えられる、そして様々な関心と目的を持つ、コースとなっていた。1940年代の教師は、スネデカーの「統計的方法」を使用し、たいていの学生は定量的スキルを身につけており、就職設計によって十分動機づけられていると考えていた。今日の統計学入門コースの教師は、異なった学生のグループを対象としている。ほとんどが若いうちから統計学を取っていて、高校でもますます増えている；差し迫った必要性から統計学にひきつけられた学生はほとんどおらず；定量的スキルのレベルは実に様々である。その結果、今日の教師は半世紀前のものより相当多くの動機づけや説明をすることへの取り組みに直面している。

統計学入門の、「何を、なぜ、誰が、いつ」が変化してきただけでなく、「どうやって」も変化してきた。ここ2、30年で、学生はどうやって統計学を学ぶのか、そしてどうやって私たち教師は彼らの学習を効果的にサポートできるのかということに重点を置く、並はずれたレベルの活動が見られてきた。

1992年のコブ氏のレポート

1991年春、ジョージ コブ氏は、数学の分野の重要な問題点を強調するため、アメリカ数学学会カリキュラム実行プロジェクト(MAA)の一部として、統計教育の電子メールグループを編成した。そのレポートはMAAの本、変革を求める声に注意せよ⁽²⁾に掲載され出版された。それには次の提案が含まれていた：

統計的思考力を力説せよ

どの入門コースにも、学生が統計的思考力の基本要素を学ぶ手助けとなる中心的な目標を掲げるべきである。上級コースの多くも、同じ基本要素をより明確に強調するような改善をすべきである。すなわち：

- ・データの必要性。個人的な決定を根拠（データ）に基づいて行う必要性と、根拠によらない仮定に基づいて行う推論に従って行動することに潜んでいる危険性を認識する

こと。

- ・データ作成の重要性。問題を明確に定式化することと、実際に適切な問題を扱っている良質なデータを手に入れることは、困難であり時間のかかることだということを認識すること。ほとんどの人は、自分でこの経験をするまで気がつかないようだ。
- ・変動のさまざまな原因。変動が至る所にあるということを認識すること。これは学問としての統計学の真髄であり、講義ではよく理解できない。経験する必要がある。
- ・変動の数量化と解釈。以下のことを考慮することで、変動は計測し説明できることを認識すること。(a)ランダム性と分布；(b)パターンと偏差（適合と残差）；(c)パターンに対する数学的モデル；(d)モデルとデータとの対話（診断）

データや概念を増やし、理論と方法を減らす

統計学のほとんどのコースは、理論と方法を少し減らして、データと概念をさらに重視することで改善できる。実行可能な最大の範囲まで、計算とグラフが自動化されるべきである。

活動的な学習を育む

一般的に統計学の教師は、講義で教える部分をずっと少なくし、プロジェクトや実験作業、そしてグループによる問題解決と討論活動のような、他の手段をもっと使用すべきである。伝統的な講義の環境であっても、学生に、より活動的に関わらせることは可能である。

3つの提案がたいへん幅広く応用されようとしている。（例えば、コースに微積分が必須条件なのか、学生が特定の統計的方法を学ぶことを期待されているかを気にかける必要があるのか、など）このコブ氏のレポートを作成した電子メールグループの仕事はレポートを完成させることで終わったが、そのグループのメンバーの多くが、大学生向けの統計学に関する ASA と MAA の合同委員会のメンバーとして活動し、特に普及と実施の活動に力を入れている。

統計学入門コースの最新の状況

コブ氏のレポートの出版後 10 年間にわたって、統計教育において多くの変革が実行された。近年では、多くの統計学者が、統計学入門を教えることを目的とした統計学教育における改革運動にかかわるようになった。⁽³⁾ そして国立科学財団は、この改革の実施面を企画する数多くのプロジェクトに資金を提供してきた。ムーア⁽⁴⁾ は、その改革の特徴を、内容（データ解析を増やし確率を減らす）、教育（講義を少なく、より活動的な学習に）、情報機器（データ解析と模擬実験）の変革という意味で述べている。

1998 年と 1999 年に、ガーフィールド⁽⁵⁾ は、入門コースがどのように教えられているのかわかりやすく、教育改革運動の影響を調査し始めるために、数学と統計学部出身の多くの統

計学指導者と心理学、社会学、商学、経済学部出身の少数の統計学指導者を調査した。この調査結果として、入門コースでは大きな変革が行われていて、変革の主な領域は情報機器の使用であり、そのコースの指導者は多くの時間を必要としていることがわかった。結果は意外にも学部間での違いは少なく、大きな違いは、グラフ電卓の使用の増加、2年制大学の数学科で教えられているコースにおいて、活動的な学習と新しい評価法、統計学部の指導者によるウェブ機器の使用の増加、また、なぜ変革を行う理由（より多くの数学指導者が統計教育からの提案に影響されている）に見られた。レポートでは、また、より情報機器の活用ができるように、さらに多くの変革がなされる予定であるということを一貫して述べている。

今日の統計入門コースは、実際多くの学問分野や学部にまたがって教えられているコースである。これらのコースを登録した学生たちは、様々な背景（例えば、数学、心理学など）や到達目標（例えば、研究プロジェクトで統計分析を行いたいと望んでいる学生もいれば、一般的な定量的推論で求められるスキルを身につけたい学生もいる。）を持っている。

昔のように、これらのコースの中には、大規模なクラスで教えられているものや、小規模なクラスで（あるいは新入生研修で）教えられているものもある。コンピューター実習室で教えられている学生もいれば、電卓しか使わないコースで教えられている学生もいる。あるいはまた、クラスメートや教員と対面しない遠隔地教育によるコースを取っている学生もいる。また、クラスの中には10週間の四半期で教えられるものもあれば、15週間の半期にわたって教えられるものもある。これらのクラスの1週間における時間数は3～6時間の範囲となっている。

今日の学生の到達目標は、概念の理解や統計的リテラシーと統計的思考力の育成に、より重点が置かれる傾向にあり、一連のツールと手続きを学ぶコースは少なくなっている。情報化の時代の中でデータを扱う必要性が伸び続けており、情報機器やソフトウェアの発展によって、ツールと手続きは、多くの人々にとって入手しやすく、使いやすくなっている。このように、手続きの仕組みを教える必要性の減少と、より多くの人々に基本的概念をしっかりと理解させることの重要性の増加によって、これらのツールを賢く使用し、理解することが必要となってきた。次のセクションに述べられているが、これらの新しい到達目標は、学生にとっての重要な学習目標の達成を支援するような、多くの統計学入門コースの再検討と改訂の必要性を強化するものである。

入門コースの学生の到達目標：統計学の教育を受ける意義

学生が統計的リテラシーをもち、データをうまく使いこなす消費者となるような教育に傾倒したコースを受け持っている人もいる。これは、芸術鑑賞のコースにすこし似ている。学生が統計解析の実施者になるような教育にかなり偏ったコースを持っている人もいる。これは、スタジオアートコースと類似している。ほとんどのコースは消費者と実施者を混

合したものであるが、その混合のバランスが、私たちが示すそれぞれの提案の重要性を決定づけるものとなる。

すべての統計入門コースにとって望ましい結果は、統計学の教育を受けた学生を生み出すということである。そしてそれは、学生が統計的なりテラシーを育成し、統計的に考えることができる能力を向上させるべきだということを意味する。

以下の目標は、そのような学生が何を知り、何を理解すべきか、ということを表している。この知識を身につけるには、統計的な技能を多少学ぶ必要があるが、特定の技能を身につけることは、それらを学ぶ経過の中で生まれる知識と同じくらい重要であるわけではない。それ故、私たちは特別な内容をカバーすることを勧めているわけではない。

学生は以下のことについて信念を持ち、理解すべきである：

- データは逸話に勝つ
- 変動は自然なことであり、予想できるものであり、量として表すことが可能である
- 無作為に標本を選ぶことで、調査や実験の結果を、標本の抜き取られた母集団へと一般化することができる。
- 比較に基づいた実験における無作為割り付けによって、原因と結果の結論を引き出すことができる
- 連関は因果とは限らない。
- 統計的な有意性は、必ずしも実用的な有意性を意味するわけではない。特に標本サイズの大きな研究では顕著である。
- 統計的に有意ではない差や関係を見つけたとしても、母集団において違いがないとか、関係がないということを必ずしも意味するものではない。特に、標本サイズの小さな調査では顕著である。

学生は次のことに気づくべきである：

- 調査や実験における偏りの共通した出所
- データが集められた方法に基づいて 統計的推論の結果を一般化できる母集団を決定する方法。
- データの集められた方法に基づいて、連関から因果関係を引き出せる場合を決定する方法
- 「正規」、「無作為」、「相関」のような言葉は、普通の使い方とは異なり、統計学において特定の意味を持っているかもしれないこと。

統計学が、問題の解決に取り組んでいるプロセスの部分を、学生は理解すべきである：

- データを得たり生みだしたりする方法
- データ解析の第一段階として、データをどうグラフ化するのかということと、興味の

ある問題に答えるのに十分な場合を知る方法

- データのまとめとグラフ表示を解釈する方法—疑問に答えることと条件を調べることの両方（統計的手順を正しく使うため）
- 統計的な推論の適切な利用方法
- 統計解析の結果の伝え方

学生は、以下の点を含めて統計的推論の基本的な考え方を理解すべきである：

- 標本分布の概念と、それをデータのサンプルに基づいて、どのように統計的推論に応用するのか（標準誤差のアイデアも含めて）
- 有意水準や p 値も含めた統計的有意の概念
- 信頼度や誤差限界の概念も含めた信頼区間の概念、

最後に、学生が知っておくべきこと：

- 背景を考慮して統計的な結果をどう解釈すべきか
- 統計的情報を含んでいるニュースの話や学術雑誌の記事を批判する方法。その際には、発表に何が欠けているのか、その情報を生みだすのに使用された研究の流れや方法も含まれる。
- いつ統計学者に支援を求めるべきか

提案

私たちは、コブレポート⁽⁶⁾に見られる3つの独創的な目標の中の考えを支持し、今日の実況を考慮しながらそれらを拡張させてきた。これらの提案の意図は、前のセクションで述べた学習の目標のリストを、学生が達成する手助けをすることである。

提案1：統計リテラシーを重要視し、統計的思考力を育成する

私たちは、統計的リテラシーを、基本的統計用語（例えば、統計用語と記号が何を表しているのかを知り、統計グラフが読み取ること）と統計の基本的考えを理解することとして定義する。統計リテラシーについては、Gal⁽⁷⁾、Rumsey⁽⁸⁾、Utts⁽⁹⁾を参照のこと。

統計的思考力は、統計学者が統計問題に取り組んだり解決したりするときに使う思考の形であると、定義されてきた。統計的思考力にはデータの必要性の理解、データ作成の重要性、変動の原因、変動の数量化と説明⁽¹⁰⁾として述べられてきた。私たちは以下の例と類推によって、統計学思考力を示す。

漏斗の例

上が幅広くあらゆる場合に対応し、また、下の方は狭く、対応できる部分が少ない漏斗を考えてみなさい。統計学者は現実的な問題解決者である。依頼人が問題（例えば、治療効

果はあるのか)を提示した時、統計学者はその問題に有効な、現実的な答えを得ようとする。話を伝えるには単純なグラフで十分である場合も多い。また、より詳細な図表によって、その問題に答えられる場合もある。そうでなければ、ある種の詳細な計算が必要になるかもしれない。全体な状況を単純化して、簡単な統計的検定によって、治療の効果はあるということを確認するかもしれない。もし、状況の単純化が難しければ、手の内にあるモデルの状況をもっと調べて、より厳密な統計的検定が使われるかもしれない。異なる統計学者たちは、与えられたデータセットに対して、すこし異なった解析をおこなうかもしれないが、主要な結論部分では大体同意するだろう。そして小さな点についてのみ、依頼人がこれらの点を気にかけるかどうか気になるだろう。もし問題に答える手順や標準的な方法がないのなら、統計学者はその時だけ新しいツールを開発するため新しい原理の開発を行うだろう。学生にある一定のツールや手順を示したり、どんな与えられた状況でも使用するべき最高の手順があったりするか、その手順のみが認められるという印象を与えるよりむしろ、この統計家が行う思考をモデルとするべきだ。

大工仕事との類似

第1週の大工仕事(統計学)コースでは、私たちは様々な種類のかんな(統計指標)の使い方を学んだ。第2週では、様々な種類ののこぎり(グラフ)の使い方を学んだ。その後、金づち(信頼区間)の使い方について学んだ。それから、様々なタイプの木材(統計的検定)の特徴について学んだ。コースの終わりまでには、大工仕事(統計学)の多くの面をカバーした。しかし、私が望んでいたのはテーブルの作り方(問題に答えるためのデータの集め方や統計解析の方法)を学ぶことであり、そのやり方を私は一度も学んでいなかった。私たちは、統計学の実践的な行動とは問題に答えるためにデータを集めて解析することだ、ということを学生に教えるべきである。

教師への提案：

- ✓ 具体的な例に取り組み、発想から推論まで、統計的問題を解く際にかかわってくる問題やプロセスを明らかにしながら、学生のための統計的思考のモデルとせよ。
- ✓ 情報機器を活用せよ。データを管理し、調査し、推論し、推論の過程の背後にある状況をチェックするために、情報機器を効果的に用いる方法を学生に示せ。
- ✓ 学生に、統計的思考力を育成し、それを活用するような練習課題を与えよ。この中には、結論のない問題やプロジェクトも含めるべきである。
- ✓ 学生にどのツールを使用するのかを教えたり、またそれをただ実行させたりするのではなく、適切な問題や情報機器を選ぶ練習課題をたくさん学生に与えよ。
- ✓ 学生の統計的思考力を評価し、フィードバックせよ。

付録にプロジェクトの例、活動、統計的思考力を育成し、評価するのに使用できる評価問題を示している。

提案 2：本物のデータを使用せよ

統計教育において、どのように、そしてなぜそのデータが作成されたり、収集されたりしたのかという問題を考えたり、解析結果を問題の背景と関連づけるたりすることを実際に行うためには、本物のデータを使うということが重要である。学生にとって関心のある実際のデータセットを使うこともまた、そのデータや適切な統計的検定について、学生に考えさせるのに良い方法である。アーカイブされたデータ、教室で生成されたデータ、模擬実験のデータも含めて、本物のデータには多くの種類がある。時には、ある特定の点（例えば、アンスコムデータの4つのデータセットが、同じ程度の相関関係であるのに、どのようにして著しく異なった散布図となっているのかということについて例示している）について例示するため、あるいは、ある特定の概念を評価するために、仮想的なデータセットが使用されることがあるかもしれない。一般的なデータ解析や調査のためではなく、この特定の目的のために作成された、あるいは現実的なデータのみを使うことが重要である。本物のデータを扱うという重要な局面は、学生が良い問題を定式化し、データがどのように作成されたのかに基づいて、それらの問題に適切に答えるためにデータが使えるようになる手助けをしている。

教師への提案

- ✓ ウェブのデータアーカイブや、教科書、ソフトウェアのパッケージ、教室での調査あるいは活動から、良い生のデータを探しなさい。もし、機会があれば、実践的な科学者から直接（学術雑誌やその人の所属する機関を通して）本物のデータを探し出さなさい。そのようなデータを使うとクラスを活気づけることができ、新しく発見されたデータを他の人に伝えることで、他の教師のための良いデータセットの蓄積を増やすことができる。データ要約ウェブサイト、学術雑誌の記事、世論調査のウェブサイト、教科書から、本物のデータに基づいた概略を探したり使ったりしなさい。
- ✓ 状況に合わせて、問題に答えるためにデータを用いたり、問題を定式化したりしなさい。
- ✓ データセットに使われた問題が学生の興味のあるものなのか確かめなさい。もし、誰も気にかけないような問題なら、それは入門クラスのデータセットとして良いとは言えない。（例：誰も聞いたことのない種の身体測定値）注：すべての学生にとって関心のあるデータセットはほとんどない。それゆえ、様々な背景からのデータセットを使うべきである。
- ✓ アンケートを作成したり、データを収集したりする前に、教室で生成したデータを使って、統計的問題を練り上げたり、データの解析の計画を立てたりしなさい。（例：様々な形のヒストグラムができ、関係を調査するための興味深いカテゴリカル変数を使いそうな質問をすること）。授業で学生から集められたデータは、学生にとって恥

ずかしいということがありうる情報を含んでいないこと、また学生のプライバシーが保たれるような情報であることが重要である。

- ✓ 大量のデータセットを入力するのに時間を費やすよりも、むしろ少量のデータセットあるいはデータセットの一部を使って、生のデータ入力を、学生に実習させなさい。大量のデータセットはコンピューターで使えるようにしなさい。
- ✓ コースの様々な部分で、変数の異なる部分集合を使用しなさい。しかし、全体を通して、同じデータセットを繰り返し用いなさい。(例：ふたつのグループを比較するために、平行箱ひげ図(ボックスプロット)を用い、それから同じデータに対して2標本t検定を適用すること。特性を調べるためヒストグラム(柱状図)を使用し、それから仮説検定のための条件を確認すること。)

付録には、課題、プロジェクト、テストなどでデータに使用する適切な方法(並びにあまり良くない方法)の例が示してある。

提案 3：単なる手順の知識よりも概念の理解を重要視しなさい

入門コースでは、教育内容として多くのことを含んでいるコースが多くあり、学生は、ただ表面上でしか理解していない、うまく統合できていない、すぐに忘れてしまうような知識の集めるだけで終わってしまう場合もある。もし学生が概念の重要性を理解していなければ、一連の手順を知っていることに、ほとんど価値がない。もし学生が重要な概念をよく理解していれば、特定の手順を容易に学べるだろう。学生の心の中には、手順の段階には注意を必要とすることが多すぎて、有用な教師でさえ概念の方向に向かって指導できていない。

手順よりも概念の方により注目することは、学生にとっても学部の要求としても政治的に難しいかもしれないということを認識しなさい。しかしながら、入門コースで概念の基礎がよくできている学生は、次の段階のコースで学習する研究方法、回帰、実験計画、あるいは統計手法を学習するための準備がよくできているだろう。

教師に対する提案：

- ✓ 様々な手法をカバーすることではなく、概念を発見することを本来の目的と考えなさい。
- ✓ 背景にあるアイデアにはあまり注目せず、多くの技能を扱うよりも、少ない技能を用いて例示することで学生が主要な概念の理解ができることを重要視すること。
- ✓ 核となる概念にもっと深い意味で焦点を当てるには、入門コースの内容を削減しなさい。多くのトピックをカバーすることよりも概念を重要視したシラバスの例は、付録に載せている。

結果の解釈をさらに強調させるため、決まりきった計算は情報機器を使って行いなさい。

数学という言葉は主要なアイデアを簡略に表現することができるが、概念の理解を高めるために公式を使用し、理解から遠ざけるような計算は避けなさい。例えば、 $s = \sqrt{\frac{\sum (y - \bar{y})^2}{n-1}}$ は、学生が、広がりやを測る尺度としての偏差基準の役割を理解するのに役立つ。しかし、 $s = \sqrt{\frac{\sum y^2 - (\sum y)^2/n}{n-1}}$ は計算のむずかしさを上回るような教育的な価値はない。

提案 4：教室での活動的な学習を育成する

教室において活動的な学習方法を用いることは、協同学習を促進する方法として価値のある方法であり、学生がお互いから学ぶことができる。活動的な学習によって、学生は重要な統計のアイデアを発見し、構成し、理解するようになり、統計的思考力を形作るようになる。活動によって学生が学習に従事することが多くなり、学習のプロセスを楽しくさせるという点で、さらに利点がある。活動的な学習方法のその他の利点は、学生に統計用語を使ってコミュニケーションを取らせ、チームで仕事することを学ぶことができる点にある。活動は教師に、生徒の学習に対する形式的ではない評価方法を提供し、生徒がどの位順調に学習を続けているかについて、教える側がフィードバックすることができる。教師は教材を教えるための活動の能力を過小評価してはいけないし、講義を過大評価してもいけないということが重要である。そういうわけで、大きな講義クラスでも活動を組み入れることについて、提案がされた。

活動的な学習に含まれるタイプは：

- ・グループや個別の問題解決の活動と討論
- ・実習室での活動（身体的な活動とコンピューターを使用した活動）
- ・現場で学生が生み出したデータに基づくデモンストレーション

教師に対する提案：

- ✓ 実際の問題の文脈に活動の基礎を置くこと。それゆえ、データは、「データを集めるために集める」のではなく、問題に答えるために集められるべきである（無論当然のことである）。
- ✓ 講義と活動、討論、実習を混合させること。
- ✓ 物理的な説明ができるコンピューターシミュレーションから始めなさい。（例、サイコロを振る、カードをきる）
- ✓ 学生からデータを集めること。（匿名で）
- ✓ データを分析する前に、活動としてデータを集める研究の結果について、学生が予測するよう奨励すること。このことは統計的手法の必要性を与える。（もし、すべての

結果が予想可能なら、データも統計も必要ないことになる)

- ✓ 学生にひとつひとつ手順の段階を踏ませるような活動をさせないで、データや問題について考えさせたり、討論させたりすること。
- ✓ 問題を明らかにし、学生に問題を通して学習させ、同じ授業時間内に活動を要約する時間が十分にあるか確認できるよう、前もって計画すること。次の授業時間に活動を完結させるのは困難である。次の授業時間の始めであっても、要点を繰り返したり、報告したりする時間があるか確認すること。
- ✓ 学生の実行したことや学習したことに対して、彼らに多くの意見・感想を与えること。
- ✓ 活動の重要な構成要素として評価も含めること。

大規模クラスの活動的な学習実施に対する提言：

- ✓ 大規模クラスでは、学生が作成するデータサイズが大きくなることを利用しなさい。
- ✓ 大規模クラスでは、学生にペアで活動させることが、大きなグループでさせるより簡単である場合がある。
- ✓ できれば、活動を行う際には、実習と討論は分けて実施しなさい。

提案 5：概念を発達させデータ解析に情報機器を利用しなさい

情報機器は統計学者の研究の方法を変えてしまった。そして、何をどうやって教えるのかも変えるべきである。例えば、正規分布の確率表のような統計表は、もはやp値を見つける際には必要なくなってきた。コンピューターを利用して計算できるようになっている。私たちは、コンピューター操作を教えるのではなく、むしろ、学生に結果の解釈や条件のチェックを重要視させながら、データ解析を行うために情報機器は使われるべきだと考えている。情報機器の利用は、また、学生が概念を可視化し、模擬実験によって抽象的な考えの理解を発展させる手助けとなるように利用すべきである。

ツールの中には、この2つのタイプの利用ができるものも多くある。その一方で、統計ソフトウェアパッケージとウェブ上のアプレットで補完的に用いる必要がある場合がある。使用されるツールは言うまでもなく、数値を計算するためのものとしてだけ見るのではなく、概念的なアイデアを調べ、学生の学習を高めさせる方法としても見るのが大切である。私たちは、情報機器の利用を、単なる計算（例：グラフ計算機に100個の数値を入力し、統計的指標を計算する）や誤った正確さ（小数点以下の桁数が多いところまで結果を出す）のために使うことに対して警告する。情報機器がすべて望まれた特徴を持っているわけではない。今後も新しいものが常に出現してくるであろう。

使える情報ツール

- ・ グラフ計算機

- ・統計パッケージ
- ・教育ソフト
- ・タブレット
- ・表計算ソフト（スプレッドシート）
- ・データアーカイブ、オンラインテキスト、データ解析ソフトを含めた WEB 上の情報源
- ・教室での反応システム

情報機器の使用法に対する教師への提言

- ✓ 大量の本物のデータセットを入手する
- ✓ 計算を自動化する
- ✓ 統計グラフを作ったり、適切に修正したりする
- ✓ 抽象的な概念を例証するために模擬実験をする
- ✓ 「もし...だったら、どうなる」といった問題を調べる
- ✓ レポートを作成する

情報ツールを選ぶ際に教師が考えるべきこと

- ・ データ入力の容易さ、多様なファイル形式からデータを取り込む能力
- ・ 対話能力
- ・ データ、グラフ、数値計算の間の強力なつながり
- ・ 特定の参加者にとって使用が簡単であること
- ・ 学生にとって入手可能で携帯できること

提 案 6： 学生の学びを調べ、向上させるために評価を用いる

教師が評価したものを学生は価値あるものとして受け入れる。そのため、到達目標に合わせた評価を行う必要がある。評価では、単に技能や手続き、計算した結果を重要視するのではなく、主要なアイデアを理解することを重要視すべきである。これは、コースの中で使われた評価（例：小テスト、中間考査、小規模なプロジェクト）や、総括的評価（コースの成績）で行われるべきである。有用で、タイミングよく行ったフィードバックが、学習を促すのに必要不可欠である。異なったタイプのコースの中では、評価の種類によって実用的なものもあれば、そうでないものもあるだろう。しかし、大規模クラスにおいても適切な評価を実行することは可能である。

評価のタイプ

- ・ 宿題
- ・ 小テストと定期試験
- ・ プロジェクト

- ・活動
- ・口頭での発表
- ・記述式のレポート
- ・詳細なレポート
- ・記事の批判

教師への提言

- ✓ コースの必須の構成要素として、評価を組み込みなさい。教師がクラスでしていることとうまく合った評価を行うことは、2週間前にクラスでやったことに焦点を置くことより、もっと効果的である。
- ✓ 学生の学びをもっと完全に評価をするために、さまざまな評価方法を使用する。
- ✓ メディアのニュース記事やグラフについて解釈したり、批判したりするような評価を用いることで、統計的リテラシーを評価する。
- ✓ 学生のプロジェクトや、決まりきった答えはない調査タスクのような評価を用いて統計的思考力を評価する。

大規模クラスにおける学生の評価に関する提案：

- ✓ 個別のプロジェクトの代わりに小グループによるプロジェクトを使用しなさい。
- ✓ 成績を出す前に、フィードバックを行い、プロジェクトを完全にするために仲間同士の評価を利用しなさい。
- ✓ グラフの適切な解釈を選んだり、妥当な統計手順を選んだりすることを重視している項目を使用しなさい。
- ✓ 学生の口頭発表のために討論の時間を活用しなさい。

実際に行うために

統計教育はフィッシャーとスネデガー以来、大きく発展してきた。さらに、世界中の統計学の教師たちは、概して現代的な取り組みや教授法を採用することに熱心である。それにもかかわらず、私たちが教え方を変えるのは必ずしも容易ではない。ある意味で、私たちはみんな教師であるとともに、学習者でもあり、少しヤドカリに似ている。成長するためには、慣れ親しんだ守りの殻をまず脱ぎ捨て、新しい、より大きなひとそろいの習慣と可能性に落ち着くまで、傷つきやすい時期を耐えなければならない。

私たちは、このレポートで多くのアイデアを提示してきた。最初はまず小さなステップを踏み、提案された方向に動くよう読者に忠告する。

小さなステップの例として以下のようなものがある：

- ・コースに活動を加える
- ・学生に小さなプロジェクトをさせる

- ・アプレット（小規模プログラム）を講義に組み入れする
- ・ソフトウェアの使用を学生に実演してみせる
- ・本物のデータセットの使用を増やす
- ・概念の理解により重点を置くために、最近扱おうとしている内容のリストからいくつかのトピックを削除する。

教育に関する哲学は教科書の選定に影響するだろうが、このレポートでの提案は教科書の選び方の提案ではない。それらは教え方に関するものである。入手できる情報源はたくさんある。MAA ノートの中の統計教育に関する巻、発展的な学部教育コンソーシアム (CAUSE) (causeweb.org)、統計関係討論リスト (www.lawrence.edu/fac/jordanj/isostat.html)、SIGMAA－MAA 内の統計教育グループ (www.pasles.org/sigmaastat)、そして ASA ウェブサイト、特に統計学教育センター (www.amstat.org/education)、そして統計教育セクション (www.amstat.org/sections/educ) などたくさんある。

未来に向けて GAISE する

これまで多くの改革がなされてきたが、統計学入門コースを改善する余地はまだたくさんある。さらに、小中高等学校で統計学の見方を学んで大学に入ってくる学生が増えているので、このコースは柔軟性を持たせ、変えやすいように調整しなければならない。AP のコースは統計教育の視野を変え続けている。ここでは、一般的な入門コースについて述べてきたが、ビジネス統計、数理統計のような他のコースについても考える必要があり、入門コースの次に行われるコースについてもその内容や到達目標として、入門コースで育成したことをベースとして、しっかりとした概念理解が築かれるものにしていくよう心を配らなければならない。

(脚注)

- (1) ジョージ・コブ 変革を求める声に注意せよ：カリキュラム実行に対する提案 (MAA 文書 No.22), 統計学教育の章, 3-43 ページ. アメリカ数学学会 1992 年 ワシントン DC
- (2) ジョージ コブ 変革を求める声に注意せよ：カリキュラム実行に対する提案 (MAA 文書 No.22), 統計学教育の章, 3-43 ページアメリカ数学学会. 1992 年 ワシントン DC
- (3) ジョージ コブ 統計教育再考；国立科学基金会議、統計教育ジャーナル、1 (1)、1993 年 URL <http://www.amstat.org/publications/jse/v1n1/cobb.html>
- (4) デヴィッド ムーア 新教育学と新しい内容：統計事例 国際統計学評論、65:123-165,1997
- (5) ジョアン ガーフィールド 統計学の教育改革の査定 国立科学基金への最終報告 2000 URL <http://education.umn.edu/EdPsych/Projects/Impact.html>
- (6) ジョージ コブ 変革への要望に耳を傾ける：カリキュラムの実行への提案 (MAA 記録

- No.22) 統計学教育の章 3-43 ページ 1992 年 ワシントン DC アメリカ数学学会
- (7) イッド ガル 大人の統計応用：意味・構成要素・責任 国際統計学批評 70:1-51 2002
- (8) D J ラムゼイ 統計学入門コースの目的としての統計応用力 統計教育ジャーナル 10(3) 2002 URL <http://www.amster.org/publications/jse/v10n3/rumsey2.html>
- (9) ジェシカ アッツ 教育を受けた国民は統計学と確率について何を知るべきか？ アメリカの統計学者 57 (2) : 74-79 2003
- (10) ジョージ コブ 変革への要望に耳を傾ける：カリキュラムの実行への提案 (MAA 記録 No.22) 統計学教育の章 3-43 ページ 1992 年 ワシントン DC アメリカ数学学会

付録 A： 活動と計画例

クラス活動の望ましい特徴

- 活動は現実社会をまねたものであるべきだ。それは「忙しい仕事」のように見えてはいけない。例えば、2項式の実験を行うためにコインやカードを使うのなら、それらが表す現実社会の2項式の実験を説明しなさい。z
- クラスは、活動をどのように行うのかについての決定に幾分関わるべきである。ステップの詳細な作り方に従うだけでは多くは学べない。
- クラスで決定したことは、授業で学んだ知識を必要とするものであるべきだ。例えば、学生が実験を計画しているなら、どうやって実験するのか直感的に決めるのではなく、授業で学んだ適切な実験計画の法則を考慮するべきだ。
- もし可能なら、学生が活動中に全体のプロセスがわかるように、活動に企画、データ収集と分析を含めるべきである。
- どのように活動を計画するのか討論するために学生をチームで取り組ませ、どのように行うのか討論するためにもう一度集合させるほうがよい場合もあるが、クラスで一緒に最初の計画やその他の決定に取り組ませた方がいい場合もある。それは、話し合う問題がどの程度難しいか、また、それぞれのチームが全く同じ方法で行う必要があるかどうかにもよる。
- 活動は、何が、なぜ行われているのかという概観で始まり、終わるべきである。
- 活動は楽しいものでなければいけない！

改善できる活動

ペプシ対コカコーラ活動

今日、私たちはペプシとコークのどちらの方がおいしいのかテストする。グループを4つに分けなさい。実験者として、グループからひとり選びなさい。注：もし自分が実験者でないなら、教室の前でできることは控えなさい。

1. 教室の前のテーブルの上に、ふたつの大きな炭酸の入ったびんがある。ひとつはペプシでもうひとつはコークのびんである。A, B とラベルのついたカップもある。実験者はテーブルの方に行って、コインを投げなさい。もし表が出たら、カップ A にペプシを注ぎ、コークはカップ B に注ぎなさい。どちらがどちらか覚えておきなさい。カップをチームに持って帰りなさい。
2. チームのメンバーに両方味見をさせなさい。どちらの方が好きか記録しなさい—カップ A なのかカップ B なのか。
3. 実験者はチームのメンバーに、より好まれたのはコークなのかペプシなのか、もう教えなさい。
4. 実験者はチームのメンバーひとりひとりに、このプロセスを一度ずつ繰り返しなさい。それから今度は、全部の生徒が一度はそれを行えるように、他のチームのひとり

が実験者に味見テストをこなさい。

5. クラス全体で集合する。教師が生徒のうち何人がコークの方を好きなのか尋ねる。
6. 教科書で公式を調べ、比率の信頼区間を求めなさい。クラスの中でコークの方が好きな生徒の比率を信頼区間で作図しなさい。
7. どちらの飲み物方がクラスでは好まれたか、仮説検定をこなさい。

中心極限定理活動

この演習の目的は、中心極限定理を証明することである。この定理が私たちに伝えているもの何なのか覚えておきなさい。それは、大きな被調査者母集団の意味は：

- ・ ベル型に近い
- ・ 母集団と同じ意味を持つ
- ・ 母集団基準の偏差 σ/\sqrt{n} に等しい標準偏差を持つ。

批評：テストは二重盲検法ではない。実験者がどの飲み物がどれなのかわからないようにする理由はない。最初に実験を設置する人は飲み物容器にラベルをつけたりはずしたりでき、それらを飲み物1、飲み物2と呼ぶことができる。それから飲み物はA,Bとラベルを貼られたカップに前もって用意される。提示する順番はそれぞれの味見をする人により、無作為に行うべきである。

批評：これは、少なくとも2つの理由から良い活動とは言えない。まず、それは全く現実社会の動機を持っていないし、統計は退屈だという神話を増強してしまう。次に、指示が完全すぎる。

学生側の調査の余地がない；彼らは単に従うべき「方法」を与えられているだけだ。

中心極限定理が当てはまることを証明するために、これらの指示に従ってください。

1. ペアに分かれなさい。各ペアでひとつサイコロを持つ。
2. 交替でサイコロを25回ずつ振るので50回振ることになる。上になった数字を毎回把握しておきなさい。
3. 結果のヒストグラムを書きなさい。サイコロの表面は等しく似かよっている。だから、ヒストグラムも均一の形になるはずである。そうなることを確かめなさい。
4. 50回振った中間値と標準偏差を求めなさい。
5. ひとつのサイコロを振った時の中間値と標準偏差は、それぞれ3.5と1.708である。自分たちの50回振った中間値は3.5に近いか？標準偏差は1.708に近いか？
6. クラス全体で集合しなさい。50回の中間値を持つはずである理論上の曲線を描きなさい。それが釣り鐘型で母集団の中間値と等しい中間値を持つということ覚えておきなさい。だから、この場合は3.5であり、標準偏差は $1.708/\sqrt{50} = .24$ になるはずである。
7. 各組のペアに、自分たちの50回の中間値を曲線上に印をつけさせなさい。中心極限定理を使って予想されることを考慮に入れると、それらが道理にかなっているかどうか注目しなさい。

この活動をどう改善すべきか？

統計学を基にした活動（シェーファーおよびその他）からの出典である「セントと中心極限定理」活動は、より指針に沿った中心極限定理を例証するための例を示している。他にも活動を基本とした統計学からの良い例がある：

- ・ 仮説検定活動（ひと組のトランプから無作為にカードを引き、常に同じ色を引く）への導入はうまく機能する。
- ・ グラフを変数に合わせるのは多くの議論と学習を生む。
- ・ ふぞろいの四角形が、当然のことながら標準になった。
- ・ 入門コースでは、無作為に抽出された返答は主要ではないが、それは統計学的な思考と幾分かかわりがある。

活動とプロジェクトの追加例

データ収集と分析： クラスのプロジェクト

ここで述べられているプロジェクトのようなプロジェクトの考えは、ロバート ワードロップの統計学：流動性の存在における学習（アイオワ デュビューク：ウィリアム C. ブラウン 1995）から由来している。これらのプロジェクトは、同様に認知心理学者、ダニエル カーネマンとエイモス トバルスキーによる研究に基づいている。

「司令官のジレンマ」のふたつの解釈について考えなさい。

第一版：優勢な敵の軍に脅され、司令官はジレンマに陥る。彼の部下である諜報部の将校たちが、ふたつ使える経路のうちひとつを使って、兵士たちを安全なところに導かない限り、彼らは待ち伏せされて捕えられ、彼らのうちの600人は死んでしまうだろうと言った。ひとつ目のルートを通れば200人は救える。ふたつ目のルートを通れば三分の一の確率で600人の兵士たちが助かり、三分の二の確率で誰も助からない。彼はどちらのルートを取るべきか？

第二版：優勢な敵の軍に脅され、司令官はジレンマに陥る。彼の部下である諜報部の将校たちが、ふたつ使える経路のうちひとつを使って、兵士たちを安全なところに導かない限り、彼らは待ち伏せされて捕えられ、彼らのうちの600人は死んでしまうだろうと言った。ひとつ目のルートを通れば400人は死ぬだろう。ふたつ目のルートを通れば三分の一の確率で兵士たちは誰も死なない、三分の二の確率で600人が死ぬ。彼はどちらのルートを取るべきか？

質問のふたつのバージョンは、ふたつの同じ答を持っている；両方とも同じ状況を述べている。ふたつの質問は言葉の使い方が違うだけである：ひとつは亡くなる命について述べ、もうひとつは助かる命について述べている。

この形式のひと組の質問は、ふたつの質問を「問題の論じ方」で比較する単純な無作為抽出の実験へと容易に導く。ひとそろいのテーマを形成し、無作為の数字の表を使ってそれらをふたつのグループに分類し、質問のひとつのバージョンをそれぞれのグループに割り当てなさい。結果は、 2×2 の数表にまとめられる。

質問	答え	
	A	B
バージョン1		
バージョン2		

データはふたつの比率を比較することによって分析できる。例えば、フィッシャーの直接確率検定か、連続性の補正とカイ二乗検定を使う。

ワードロップの本の中にある、集合1、2の練習問題にはこの構造について数多くの種類が載っており、それらの多くは学生によって行われる。ここに4つだけ簡略されたものがある。

- 歴史博物館の人に、歴史の本のある特定の主張の中で説得力のあるものを見つけたか尋ねてみなさい；主張は支持するデータの表がある／ない。
- 学生会の女性に、ある見知らぬ男性が近づいてきてお酒を飲みましょうと誘われたら行くかどうか尋ねなさい；男性は魅力的／魅力的ではないと表現される。
- ソフトクリームを注文している客に、普通のがよいか、ワッフルコーンがいいか尋ねなさい；ワッフルコーンは「自家製」である／「自家製」ではないと述べられる。
- 大学生に、(1)落ち込んでいる友達にカウンセリングを受けることを勧めるか？それとも、(2)自分が落ち込んだらカウンセリングを受けに行くか？

ふたつの答があるふたつのバージョンに基づくプロジェクトは、利点を多く持っている：

- ・ データ収集が、適切な時間内に完成できる。
- ・ 無作為抽出は、結果が正式な推論にふさわしいということを保証する。
- ・ 無作為抽出は、データ収集の可能性と分析の確率モデル使用とのつながりを明白にする。
- ・ 分析の方法は、比較的単純でわかりやすい。
- ・ 構成（ 2×2 数表）は、たいへん幅広く応用が利くものである。
- ・ 最後に、書式は無制限で、それによって学生は幅広いアプリケーションの範囲を得ることができ、テーマやペアの質問を選ぶ際に、そこから構想や独創性を活かす十分な機会が得られる。

チーム構成の関係についての質問

これらの助言は、教師向けのものである。学生向けの助言はプロジェクト4 チーム形成にある。

目標：学生に研究用の疑問を作らせ、それに答える手助けとなるように、データを収集し説明する経験を与えなさい。

与えるもの：(N=学生数； T=チーム数)

- ・ N枚の見出しカード、あるいはそれぞれのT色のカラー（またはボードのスペースを

使用；下を見なさい)

- ・ T枚 あるいは 2T枚の透明シートとペン(2T枚の理由となるステップ3を見なさい)
- ・ T個の 計算機

学生は4~6人のチームで取り組むべきである。プロジェクト例4 チーム形成を見なさい。

ステップ 1：各チームは、関連があるかどうか知りたいふたつの分類別の変数を明確に述べる。例えば、誰かが長男（長女）（あるいはひとりっこ）で、屋内の活動を好むか屋外の活動を好むか（最近の調査では、第一子は屋内の活動の方を好み、下の子供は屋外の活動の方を好むということを示している）；男性／女性と何かについての意見；クラス（上級生／下級生等）と、車を持っているかどうか等。時間内に容易に終わらせるため、ひとつの変数につき、ふたつの分類に制限したいと思うかもしれない。

データ収集には、ふたつの可能な方法がある一見出しカード（または紙片）あるいはボードを使用する。次の数段階はそれぞれ両方の方法で述べられるだろう。

ステップ 2：各チームは、チームの数の色の中から、ひとつの色の見出しカードを割り当てられる。例えば、チーム1は青かもしれないし、チーム2はピンクかもしれない、等々。

ボード：各チームに、疑問を書くためのチョークボードのスペースを割り当てなさい。

ステップ 3：各チームはクラス全体に自分たちのふたつの疑問を尋ねなさい。カード：チームは質問を透明シートに書いて皆に見せる。各チームが交替で教室の前に行く。学生はチームカラーに応じて答を見出しカードに書き、それらを集める。例えば、クラスの学生全員が、自分たちの答えをチーム1の質問にはブルーの見出しカードに、チーム2の質問にはピンクの見出しカードに、等々。ボード：チームのメンバーは、各学生が適切なセルに#の記号を書けるようになっている2方向の表とボードに、質問を書く。

プロジェクト2.2から採用、指導者の教材マニュアル、統計学に注意せよ

アツとヘッカード

注：これはひとつの分類的、そして定量的変数とふたつのサンプルの推測をする
時のために保持されたデータで行うこともできる。

ステップ 4：カード：各チームが質問をして学生たちが自分たちの答を書いた後、カードが集められ、それぞれのチームに渡される。例えば、チーム1は青のカードをすべて受け取るというように。ボード：クラス全員がボードの自分たちの割り当てのところに行き、#の記号を自分たちに合った表のセルのところに書く。

ステップ 5：各チームが集計し、要約し、自分たちの質問に対するデータのグラフを画面に表示する準備をする。結果は透明シートに書かれる。

ステップ 6：各チームが結果をクラスで発表する。

ステップ 7：そのデータが大きな母集団から無作為に抜きとったものだということにした
いのだったら、カイ二乗検定で独立したものとして扱う時のために、結果を取っておくこ
とができる。

プロジェクト 4： チーム形式

チームメンバー：

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

指 示：

1. クラスのメンバーに面白い関係があるかもしれないと思う分類の変数をふたつ考えな
さい。もしそうしたければ、例えば、GPA-高低（ $GPA \geq 3.0$ のような切り捨てを使って）
のように、定量的変数を分類の変数に変えてもよい。それぞれの変数は、割り当てられた
時間内に終わらせられるように、ふたつの分類を持つべきである。

2. どちらが説明変数なのか、どちらが反応変数なのか組み立てながら、自分の状況で意
味を通るなら、下のふたつの変数をリストに入れなさい。

- ・ 説明変数：
- ・ 反応変数：

3. 各チームがチョークボードに場所を割り当てられる。ひとりのチームメンバーがボー
ドのところに行って、自分たちのふたつの質問を書かなければならない。また、皆がそれ
に答えられるよう、#の記号を書きこむ四角のある「2方向」の表も書きなさい。

4. 全員がボードのところに行って、各チームの質問に対して適切な四角に#の記号を書
く。

5. 皆がボードのところに行ってデータを記入したら、自分のチームの質問に対して下の
表に合計を入力しなさい。また、それぞれの変数に対して何の分類なのか入力する。

反応変数

説 明 変 数	分類 1	分類 2	合 計
分 類 1			
分 類 2			
合 計			

6. 自分のチームの透明シートに表示できるように、適切な数とグラフの要約を作りなさい。必要があれば、発見したことの手短な要点を下と後ろに書く。

7. チームの中からひとりが、透明シートを使ってチームの結果をクラスで発表する。

8. このシートと透明シートを提出する。

ふたつの条件のもと、手先の器用さを比較する

これらの指示は、教師向けである。学生向けは「プロジェクト5 チーム形成」にある。

目標：学生に計画、遂行、分析、実験という経験をさせなさい。

与えるもの：(N=学生の数、 T=チームの数)

- ・ ふたつの色の違う乾燥豆がそれぞれ 30 粒くらい入ったボウル T 個
- ・ 2T 個の空の紙コップまたはボウル
- ・ T 個の ストップウォッチ あるいは秒針のある時計

話：ある会社にはふたつのタイプの小さな部品を仕分けするのを仕事としている従業員が大勢いた。従業員たちはけがを繰り返しがちであった。それで会社は、従業員たちが入れ替わったら、能力の高い人を使ったり、能力の低い人を使ったりして生産性で大きな損失が出るだろうかと考えた。(あるいは、もしゴム手袋を使っているのなら、話はそうなりうるだろう。健康上の理由で彼らは手袋を必要とするだろう) それゆえ、この比較をする実験を計画し、遂行し、ぶんせきするのである。学生たちは、2色の豆をボウルから色分けしてふたつの紙コップに移すのにどの位時間がかかるのか見るためにタイムを測る。比較は利き手とそうでない手を使った後に行われる。代替案は、例えば 30 秒というように時間を決めて、いくつの豆がその時間内に移せたか見るというものである。

ステップ 1：クラスで、どのように実験を行うのか話し合うこと。これは、まずチームで行う。下の提案をみること。

- ・ 処置は何か？実験の単位は何か？
- ・ 考慮すべき実験計画の方針は以下のとおりである。実験を計画し実行する際に、できるだけそれらを多く使用しなさい。どうしてそれらが使われるのか話し合いなさい。
 - つり合ったペアの作成と区画
 - 実験単位に対する処置法の無作為抽出、あるいは処置法の順序の無作為化
 - 目隠し試験、または二重目隠し試験
 - グループの制御
 - 偽薬

プロジェクト 12.2 から採用、指導者の教材マニュアル、統計学に注意せよ
アツツとヘッカード

注：変分は、熟練の人とそうでない人の代わりに、ゴム手袋をつけて仕事をさせるのと、つけないでさせること。その場合、ゴム手袋のN組のペアが必要になる。

➤ 学習の影響あるいは疲労

- ・ 興味の媒介変数は何か
- ・ どのような分析のタイプが適切か—仮説検定、信頼区間、または両方

各生徒が自分の手で一度タスクを完成させるように、クラスで決定すべきである。なぜこれが、クラスを半分に無作為に分けて、クラスの半分は利き手を使い、もう半分が利き手ではない方を使うより望ましいのか？ 順番はどのようにして決められるのか？ すべての生徒にそれは同じであるべきなのか？ 練習は許されるのか？ 目隠し試験や二重の目隠し試験は可能なのか？

ステップ 2：チームに分かれて実験を実行すること。

プロジェクト 5 チーム形成は、チームメンバーにタスクを割り当てるひとつの方法を示している。

ステップ 3：記述統計学と推論の準備

クラス全体で集合し、違いの幹葉図を描きなさい。この分析の必要条件が見合っているかどうか話し合いなさい。異常値はあったか？ もしそうなら、それらは説明がつくか？ その違いに対する中間値と標準偏差を、誰かに計算させなさい。

ステップ 4：推論

チームで集合させなさい。各チームは、中間値の相違に対して信頼区間を求め、仮想検定を実行させること。

ステップ 5：クラスで再集合し、結論について話し合うこと。

被調査者母体で実験を計画し分析する方法についての提案

プロジェクト 5

計画の問題：

- つり合ったペアの作成と区画：各学生はペアで、自らの手でタスクを一度行うように、使用されなければならない。
- 実験のユニットの処置の無作為抽出、あるいは処置の順序の無作為抽出：それぞれの学生に対してどちらの手を使うのかについて、無作為に抽出する。
- 目隠し実験か二重目隠し実験か：明らかに、学生はどちらの手が使われているのか知っているが、タイムキーパーは知る必要はない。
- グループ制御：この実験には関係がない。

- 偽薬：この実験には関係がない。
- 学習

効果あるいは疲労：学習効果はありそうなので、練習問題を数回分作りたくなるかもしれない。また、各生徒の両手の順番を無作為に抽出することが、このことに役立つだろう。

可能な計画の例：

学生にコインを投げさせる。表が出たら利き手から。裏が出たら利き手ではない方から。豆を分離させるのにどの位時間がかかるのか見るために、時間を測る。時間を計測する学生は、見ないことで、状況に対して目隠しをしていることにできる。

分析：

- 興味の要素は何か？

答え：それぞれの人の関心の確率変数を、以下の「手先の器用さの違い」であると定義する。

d = 利き手ではない方の手で、余計に必要なとした秒数

d = 聞き手ではない方の手でかかった時間－利き手でかかった時間

Define μ_d = 母集団中間値 手先の器用さの違い

- 帰無仮説と対立仮説は何か？●

$$H_0: \mu_d = 0$$

$$H_A: \mu_d > 0$$

(利き手でより速い)

- 信頼区間は適切か？●

そう、それは、従業員が利き手だと、どの位早く仕事を達成できるかについての情報を

$$\bar{d} \pm t^* \left(\frac{s_d}{\sqrt{n}} \right)$$

与える。信頼区間の公式は、 $\bar{d} \pm t^* \left(\frac{s_d}{\sqrt{n}} \right)$ となり、そこでは t^* が $df = n - 1$ で、臨

界 t 値であり、 s_d が点差の標準偏差となる。検定を行うために、 $t = \frac{\bar{d} - 0}{s_d / \sqrt{n}}$ を計算し、確率値を求めるために、臨界 t 値と比較しなさい。

プロジェクト 5：チーム形式

チームメンバー：

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

指 示：

チームで取り組む。各チームは、豆の入ったボウルと2つの空のカップを取る。豆をボウルから空のカップに、色分けして移す。どの位かかるか見るために、タイムを測る。これを、それぞれの手で各2回行う。それぞれの手で一度、順序は無作為に決める。

- これらの作業を計画する。望むなら、それぞれの回で作業を入れ替えてもよい。
 - 調整係 — 実演する
 - 無作為抽出係 — どちらの手から始めるか、人が代わるたびにコイントスで決める。
 - タイムキーパー — タスクを行うそれぞれの人のタイムを測るため、秒針のある時計を持っていなければならない。
 - 記録係 — 下の表に結果を記録する。
- 誰から始めるのか決める。無作為抽出係は、どちらの手で先に行うのか告げる。最初の人々が2番目の手に変える前に、それぞれの人が一度タスクを完成しておくべきである。そうすると、手を変える前に皆が休む時間ができる。
- タイムキーパーは、ボウルからカップに色分けしてひとつずつ豆を移す間、その人の時間を測る。
- 記録係は時間をメモし、表に記録する。
- これを各チームメンバーについて繰り返す。
- その後、それぞれの人が、1回目に使わなかった方の手で2回目を行う。
- それぞれの違いを計算する。

クラスの結果

ここにデータを記録する：

名 前	利き手ではない方の手のタイム	利き手のタイム	d = 違い 利き手ではない方 - 利き手

テストされ、予想される母集団特性値は：

信頼区間：

仮説検定 — 推測と結果

付録 B： 評価項目の例

難易度の、特質に関する問題と解説を伴った、評価項目の例をいくつか挙げる。

まず、難易度の本質に関する問題と解説を伴う、評価項目の例をいくつか取り上げる。

検定では使用を避ける評価項目：正/誤、背景や説明なしで計算のみ、計算や分析を入力するにはデータが多すぎる項目、定義や公式の記憶を試すだけの項目。

項目 1

ある教師は、前期で初歩的な統計の 2 セクションを教えた。午前 8 時のクラスの期末考査の中間値と標準偏差は、78 と 8 で、午後 4 時のクラスは 75 と 10 であった。これらの数字をよく調べると、より出来の良い学生は、恐らく午前 8 時の方に届け出していると教師は考える。それで彼女は、自分の学生の 2 つのグループの期末考査の点数の中間値は同じかどうか、テストすることにした。

項目 2

ある経済学者は、男性と女性の重役の月給の中間値を比較したい。彼は、それぞれから 10 の被調査者母体を無作為に取り、t 検定を行う。結果の確率値は 0.45 である。

1. 帰無仮説と対立仮説を述べなさい。
2. 統計的結論を出しなさい。
3. 統計学の学習をしていない人にもわかるような言葉で、自分の結論を述べなさい。

批評：教師はすべての母集団のデータを持っているので、統計的な推量は必要ない。

批評：質問は t 検定の必要条件を述べておらず、被調査母体の規模が小さい。条件はここでは、ほとんど守られていない。月給はほぼ確実に非対称である。

項目 3

次のうちどれが、確率値の限定をするのか？

- A. 無効仮説が当てはまるなら、それは無効仮説を拒否する可能性である。
- B. 無効仮説が当てはまるなら、それは無効仮説を拒否しない可能性である。
- C. 観察されたものと同じくらい、それは極端な観察データの可能性である。
- D. それは無効仮説が当てはまる可能性である。

評価項目を改善する方法を示す例

正/誤項目は、うまく書かれても、その話題について何も知識はなく、項目を正しくする可能性は常に 50% あるので、学生についての情報をあまり多く与えない。ある最新の取り組みは、項目を 3 つ、あるいはそれ以上の選択肢による、強制選択質問に変えることである。

項目 4

集合データの標準偏差の大きさは、中心がどこにあるのかによる。正しいか、正しくないか？

こう変える：

集合データの標準偏差の大きさは、どこに中心があるのかによるか？

- A. そうである。中間値が高ければ高いほど、標準偏差は高くなる。
- B. そうである。標準偏差を計算するために中間値を知る必要があるから。
- C. そうではない。標準偏差の大きさは、分布の位置に影響されない。
- D. そうではない。標準偏差の大きさは、中央値とどの位違うのかではなく、値がどの位異なっているのかしか測れないから。

批評：これらのうちどれも、完全に正しいわけではない。BとDの答えは明らかに間違っている。Aの答えは重要レベルである。Cの答えは継続すれば正しい（が、可能性は低い）だろう。「... あるいはさらに極端になるだろう、無効仮説が当てはまると仮定すれば。」

項目 5

+1の相関関係は、-1の相関関係よりも強い。正しいか、正しくないか。

このように書き変える：

教育研究ジャーナルに載っている最近の記事に、数学の成績と総合的数学の適性の間の相関関係は、+0.8であると報告している。また、数学の成績と数学の不安テストの間では-0.8の相関関係であると報告している。次のどの解釈が最も正しいか？

- A 相関関係+0.8は、相関関係-0.8より強い関係を示している。
- B 相関関係+0.8は相関関係-0.8とちょうど同じくらい強い。
- C どちらの方が相関関係が強いとは言えない。

背景は、現実の状況で、学生が統計学的考えを知ったり扱ったりする手助けとするのに重要である。

項目 6

XとYが深く関連づけられていることが、いったん証明されると、Xの変化がYに変化を引き起こすことを立証するためには、どんな形の研究をすることが必要になってくるのか？

背景が付け加えられる：

ある研究者は実験的薬と T₄ リンパ球細胞レベルの関係を、HIV/AIDS 患者について研究している。T₄ リンパ球は免疫システムの一部で、HIV 感染者にそれが低下したときに見られる。その 2 つの変数、薬の服用量、T₄ 細胞レベルが高く関連しているといったん証明されると、服用量の変化が T₄ 細胞レベルでの変化をもたらすということを証明するために、どんなタイプの研究が必要となるのか？

- A. 相関関係研究
- B. 制御実験
- C. 予想研究
- D. 調査

概念や解釈を損ない、学生にとって問題の重点となってしまうかもしれない、繰り返し、またはうんざりする検査の計算は避けるよう努める。

項目 7

一年目の教科課程のコースでは、メジャーリーグスポーツの放送権における、展開の経過を分析するために、ダーウィンの法則を適用するよう学生に求める、20 点のレポート問題を含んだ最終試験を使用した。コース内の 4 人の教授の採点に一貫性を調べるため、6 段階のレポートのサンプルが無作為にそれぞれの教授から選ばれた。点数は下の表にまとめられている。4 人の指導者の中間値における違いをテストするために、分散分析を行いなさい。

指導者	点数					
アフィンジャー	1 8	1 1	1 0	1 2	1 5	1 2
ビューリユー	1 4	1 4	1 1	1 4	1 1	1 4
クリアリー	1 9	2 0	1 5	1 9	1 9	1 6
ディーン	1 7	1 4	1 7	1 5	1 8	1 5

一年目の教科課程コース…（上と同じ）…。点数は、全体のサンプルに対する記述的統計学と一方向の分散分析の出力と共に、下の表にまとめられている。

記述的統計学

変数	N	Mean	Median	TrMean	StDev	SEMean
点数	24.00	15.00	15.00	15.00	2.92	0.60

一元配置分散分析

分散分析表 除外

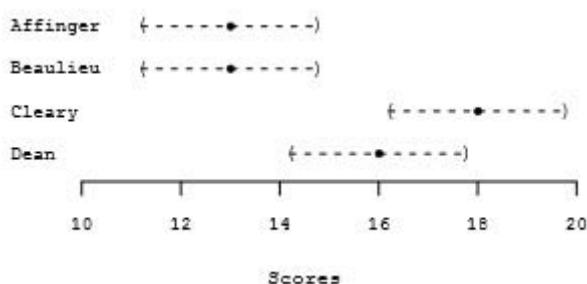
レベル	N	Mean	StDev
-----	---	------	-------

アフィンジャー	6	13.00	2.97
ビューリユー	6	13.00	1.55
クリアリー	6	18.00	2.00
ディーン	6	16.00	1.55

統合標準偏差 = 2.098

批評：上の質問形式には、結果を得るためにかなり計算機を使う必要があるし、解釈を問うということもあり得ない。下の改訂版はそれでもまだいくらか計算を必要とする（それは与えられる計算の出力量により調整できる）が、計算は比較的効率よくできる。与えられる計算の出力が何であるのかよくわかっている学生なら特に。

統合標準偏差に基づく中間値の各 95%信頼区間



1. 残念ながら、出力からの分散分析表がないので、それを構成するために上記の情報を使い、4人の指導者たちに割り当てられた平均点における大きな違いをテスト（5%レベルで）する。仮定と結論を必ず含むこと。残りの部分（または次の質問）を完成するのに必要である表の一部を作るのにてこずったら、筋の通った推量をするか援助（要点だけの料金で）を求めなさい。
2. 分散分析表ができたなら、アフィンジャー博士から与えられた平均点の、95%信頼区間を構成しなさい。注：解答はグラフが示すものと一致していること。

良い評価項目の追加例

項目 8

ひとりの学生が1学期中に教科書に費やす金額を、Yで表しなさい。ナンシーが統計学を心得ていて、秋の前期1と春の後期2がどの位なのか知りたいと仮定して、それを比較しなさい。特に、彼女は？1の平均と？2の平均に関心があると仮定しなさい。ナンシーはいくつか統計学コースを取っていて、信頼区間や仮定検定をどう解釈するかも含め、統計学をよく知っていると思定されるかもしれない。それぞれの学期から無作為にサンプルを抽出し、データを分析してレポートを書きなさい。次の4人の人たちのアドバイスを聞きな

さい：

1. ラッドはこう言う、「 $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$ の $\alpha = .05$ テストを行い、ナンシーに H_0 を認めるべきではないか伝えなさい。」
2. リンダはこう言う、「 $?_1 - ?_2$ の 95% 信頼区間を出しなさい。」
3. スティーブは言う、「 $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$ のテストを行い、ナンシーにそのテストから出た確率値を報告しなさい。」
4. グロリアは言う、「 \bar{y}_1 と \bar{y}_2 を比較しなさい。もし、 $\bar{y}_1 > \bar{y}_2$ なら、それから $\alpha = .05$ を使って $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 > \mu_2$ をテストし、ナンシーに、 H_0 を認めないかどうか言いなさい。もし、 $\bar{y}_1 < \bar{y}_2$ なら、それから $\alpha = .05$ を使って $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 < \mu_2$ をテストし、ナンシーに H_0 を認めるかどうか告げなさい。

4つのアドバイスを最も良いものから悪いものまでランクをつけ、なぜそのようにランクをつけたのか説明しなさい。つまり、なぜそれが他のものより良いのか説明しなさい。

項目 9

研究者たちは、ふたつの母集団から無作為に、被験者のサンプルを抽出し、そのデータにウィルコクソン・マン・ウィトニーテストを応用した；そのテストで、無指向性代替を使った確率値は.06. だった。以下に述べているものは、それぞれ正しいか正しくないか、またそれはなぜなのか答えなさい。

1. ふたつの母集団の分布が同じだという確率は6%である。
2. もしもふたつの母集団の分布が本当に同じなら、この研究者たちが観察した違いが同じくらい極端なこのふたつのサンプルの違いは、6%の確率でしか起こらない。
3. もし、このふたつの母集団を比較する新しい研究がされたら、 H_0 が再び拒否される確率は6%ある。
4. もしも $\alpha = .05$ で指向性代替が使われ、データが代替仮説に特定された方向で分けられたら、 H_0 は拒否されるだろう。

項目 10

月曜のCNNのウェブページ (www.cnn.com/HEALTH/9612/16/faith.healing/index.html) の記事は、このような文で始まる、「月曜に発表された調査によると、宗教的信仰は患者の治療を促すと、家庭医たちは圧倒的に信じている」と。後からその記事はこう述べる、「医学研究者たちは、宗教の利点はストレスを減らすことで免疫システムを助けるというとても簡単なことかもしれない」と。そしてハロルド・コーニング博士は、伝えられるところによると、こう述べている、「定期的に教会に通っている人は、うつ病にかかる割合が、あまり教会に行かない人の半分である」と。

コーニング博士が述べていることと、記事の示している、宗教的信仰と習慣がうつ病との闘いを助けるという主張を批判するために、統計学の用語を用いなさい。批評には次の単語のうち、いくつかを選ぶ必要があるだろう：

観察に基づく研究、目隠しの、二重目隠しの、精度、偏り、被調査者母体、擬似の、誤りを証明、因果関係、結合、無作為、有効な、信頼できる

項目 11

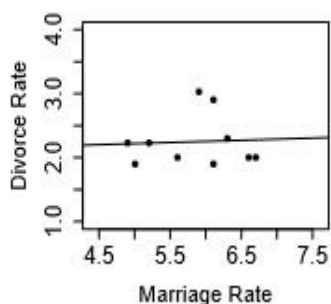
フランシスコ フランコ (98 年度クラス) は、100 個のハーシーキッス (アーモンド入り) の重さを測った。被調査母体の平均は 4.80 グラムであり、SD は .28 グラムであることを彼は発見した。この背景の状況で、平均の標本分布によって、中間値は何であるのか説明しなさい。

項目 12

ある園芸家は 3 つのタイプの豆の種、A タイプ、B タイプと C タイプの産出を比較したい。彼女は、A タイプの種を無作為に 3 つに分け、それぞれを庭の東の部分、中央部分、そして西の部分にいくらか植えた。その後、B タイプと C タイプも同様に行った。

1. この園芸家はどんな種類の実験計画を使用しているのか？
2. なぜ、この種の計画がこの状況で使用されるのか？ (背景の状況で説明せよ)

項目 13



この散布図は離婚率 y と婚姻率 x がどの程度、データ採集した 10 カ国の中で関連しているのかを示している。回帰線が図表に書きくわえられた。

1. アメリカはもともと集められた国々の中で 10 点の国ではない。たまたまアメリカの婚姻率が他の 10 カ国より高いというだけである。さらに、他の国々のパターンを考慮に入れると、アメリカの離婚率は予想より高い。アメリカを集合データに加えると、回帰線にどう影響するのか？ それはなぜだろう？
2. アメリカが集合データに加えられた後の散布図と回帰線について考えなさい。残余の図表の略図を書きなさい。三角形の図表に軸線を分類し、アメリカを特定しなさい。

項目 14

研究者たちはふたつの薬、フォルモテロールとサルブタモールを、運動から引き起こされる喘息の患者の治療法として、噴霧溶液で擬似薬とで比較したいと思った。患者たちは薬か擬似薬を飲み、運動をし、「強制呼気量」を測ることになっていた。30人の被験者が有効だった。(A.N. トソイとその仲間たち、*ヨーロッパ呼吸器ジャーナル* 3 (1990) : 235 ; ベリーを通じて、*統計学：ベイジアン*の相関関係に基づく)

1. これは実験なのか、観察的研究とすべきなのか？ その理由は？
2. この背景の状況では、擬似薬の効果は何か？
3. ここでの乱塊法計画 (RBD) の始め方を手短かに説明しなさい。
4. 乱塊法計画はどのように役立つのか？つまり、このような状況で、RBD を使用する主な利点は何か？

項目 15

私は、114 クラスから 8 人の生徒が、第 2 回目の試験 (4 月) の前の復習授業に出席していることに気がついた。これら 8 人の生徒たちの平均点は、復習授業に出席しなかった 21 人の生徒たちの平均点より低かった。この情報を復習授業の有効性における研究に、私が使いたいと考えていると想定しなさい。

1. これはどんな種類の研究か、観察的、あるいは実験的なのか？ それはなぜか？
2. ここで懸念されるのは、どんな種類のサンプリング誤差や標本抽出バイアスなのか？
3. (仮説) 私は友人のジョージに、出席した 8 人と出席しなかった 21 人のデータを渡した。ジョージは仮説検定を行うためにそのデータを使用した。仮説検定は意味があるのだろうか？もしそうなら、 H_0 は何か？そうでなければ、なぜなのか？

項目 16

次の 3 つの状況について、それぞれ自分が行う分析のタイプを述べなさい (例 : 1 標本 t 検定、回帰、独立性カイ二乗検定、カイ二乗適合度検定等)、もしあなたが未補正のデータをすべて持っているなら。そして自分が分析を行う変数の役割を明細に述べなさい、しかし、分析を実際に行ってはいけない。

1. エリザベス ラーンツ ('02 のクラス) は、13 人の学生の脈について運動の影響を測った。彼女は運動 (30 回のジャンピングジャックをする) の前後に脈を計測し、変化の平均は 55.1 で変化の SD は 18.4 だということを発見した。このデータをどう分析するか？
2. 3 つの HIV 治療法について、子供の HIV の進行を防ぐ効果がテストされた。薬 A を与えられた 276 人の子供たちのうち、259 人が生き、17 人が亡くなった。B の薬

を与えられた 281 人のうち、274 人の子供たちが生き、7 人が亡くなった。C の薬を与えられた 274 人の子供たちのうち、264 人が生き、10 人が亡くなった。このデータをどう分析するのか？

3. ある研究者は前腕の長さとうりとの関係に興味を持った。彼は、16 人の女性に前腕の長さとうりを測った。これらのデータをどう分析するか？

項目 17

4 つの標本から、私はデータデスクに、データを平行な構成で作図させた。それから、 $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ の検定を行い、 $\alpha = .05$ レベルで H_0 を認めなかった。また、 $H_0 : \mu_1 = \mu_2 = \mu_3$ も検定し、 $\alpha = .05$ レベルで H_0 を認めなかった。しかしながら、 $\alpha = .05$ を用いて $H_0 : \mu_2 = \mu_3$ を検定した時、 H_0 を認めた。同様に、 $\alpha = .05$ を用いて $H_0 : \mu_1 = \mu_4$ を検定したとき、 H_0 を認めた。

1. あなたの仕事は、データを平行な構成でグラフを描くことである。つまり、私がテストについて話したことに基づいており、データがどのように見えるのかという考えを持つべきである。グラフを描くのに、その考えを使いなさい。構想に加える三角形で標本の間値を示しなさい。
2. パート 1 でグラフ化した同じ標本の間値で、データを得ることは可能であるが、 $\alpha = .05$ レベルで $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ は認められるという仮定に対しては得られない。この背景のグラフを描きなさい。つまり、パート 1 から得た同じ標本間値 (三角形) を保ちなさい。しかし、 H_0 が認められるなら、データがどのように異なっているのか示しなさい。

項目 18

アトレイ チャーク ('02 クラス) は、12 の朝食シリアルは無作為標本についてデータを集めた。彼は X =繊維 (グラム/オンス)、 Y =価格 (セント/オンス) と記録した。データの散布図は、線形の関係を示している。適合回帰モデルは、

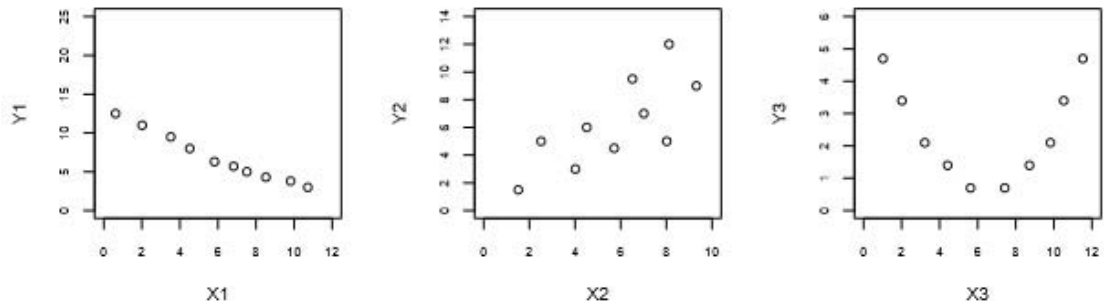
$$\hat{y} = 17.42 + 0.62x$$

サンプルの相関係数 r は、0.23 である。 b_1 の SE は、.81 で、また、 $s_{y|x} = 3.1$ である。

1. r^2 を求め、この問題の状況での r^2 を説明しなさい。
2. あるシリアルには 2.63 グラムの繊維 (オンスあたり) があり、17.3 セント (オンスあたり) の価格であると、仮定しなさい。このシリアルの残りはどうか？
3. この問題における状況で、 $s_{y|x} = 3.1$ の値について説明しなさい。つまり、 $s_{y|x} = 3.1$ であるというのは、どういう意味なのか？
4. この問題の状況では、「回帰の影響」によって何を意味しているのか説明しなさい。

項目 19

下のそれぞれの散布図におけるデータのサンプルの、相関関係を概算しなさい。



項目 20

整合実験では、ジョーズピザとドミノズピザの通常のチーズピザの味を比較する。それぞれの被験者は印のない 2 切れのピザ、各タイプから 1 切れ、無作為の順序で味見をし、どれを好むか述べる。この研究に参加している 50 人の被験者のうち、21 人がジョーズピザの方を好んでいる。

1. ドミノズピザよりジョーズピザの方を好む母集団の比率に対する 96%の信頼区間を求めなさい。
2. 望ましい信頼区間の誤差が 4%の場合、どの位の規模の母集団が必要か？

項目 21

心臓病学者 5 人に 1 人が、動脈硬化を防ぐため、一日に一錠アスピリンを飲んでいるという主張があった。この主張が事実だと仮定しなさい。もし 1,500 人の心臓病学者が無作為に選ばれたら、1,500 人のうち少なくとも 275 人が一日に一錠アスピリンを飲む確率はどのくらいか？

項目 22

散布図が、変数間の関係の概要を、視覚的に適切に示すのかどうか特定しなさい。それぞれの場合で、自分の推論を説明しなさい。

1. 血圧と年齢
2. 国の地域とより強い銃規制法についての意見
3. SAT (大学進学適性テスト) の言語と数学の点数
4. ハンドスパン (親指と他の指を開いた手幅) と性 (男性か女性か)

項目 23

次の文章は、それぞれいくつか統計学の分析を求める状況を述べている。それぞれに対して指示に従いなさい：

1. (下のリストより) 適用するのにふさわしい統計学の手順の名前をあげなさい。同

じ手順を 2 回以上使ってもよい。また、ひとつの質問に対して正しい答が複数あるものもあるかもしれない。

2. 問題の中では確率値も与えられる。その特定の問題に対して結論を導くのに、それを使いなさい。ただ H_0 を認めないだとか、あるいは H_0 を否認できないと言うだけではいけない。それ以上のことを述べなさい。(5%の重要度を想定すること)

信頼区間 (中間値 p の…)	通常の分布
母集団規模の決定	相関関係
中間値テスト	単一線状回帰
比率テスト	複合回帰
中間値の差 (ペアデータ)	二方式図表 (カイ二乗検定)
中間値の差 (ふたつの独立標本)	中間値の差に対する分散分析表
比率の差	中間値に対する二方式分散分析表

あなたが選ぶ統計学的手順：

A. 人類学者たちは同じ地域にふたつの埋葬地を発見した。彼らは、その地域にはいくつかの部族が暮らしていたこと、そしてそれらの部族は頭蓋骨の長さの違いによって分類されてきたことも知っている。彼らは、それぞれの埋葬地から発見され無作為抽出された頭蓋骨の標本を測り、そのふたつの埋葬地が異なる部族によって作られたのか決定したいと考えた。(確率値 = 0.0082)

B. ハワイの農園主協会では糖分をより多く含む果肉を生産するため、3つの新種のパイナップル (A、B、C と呼ぶ) を開発中である。それぞれの種類で 20 本 (全部で 60 本) が無作為に 2 エーカーの畑に分散された。収穫後、パイナップルの結果として糖分の含有量が測られた。そして生産高が各種類に対して記録された。3つの種類で、糖分の平均含有量に大きな差があるか? (確率値 = 0.987)

C. 研究者たちは、子供向けテレビ番組における暴力シーン調査モニター (VICTIM) に、土曜の朝のテレビ番組における暴力的演技描写の頻度について研究するよう依頼された。彼らは、12 週間にわたってこの時間帯に放送された 40 番組の無作為標本を抽出した。標本 40 番組のうち 28 番組が暴力描写シーンが過剰であると判定された。土曜朝のテレビ番組すべての母集団について述べるのに、彼らはどのようにこれらの情報を使うべきか?

D. 就職指導室では、最上級生の計画と彼らがどのように専門を関連づけていくのか、関心

を持っている。多数の学生が調査され、彼らの専攻（自然科学、社会科学、人文）や将来の計画（大学院、就職、未定）によって分類されている。専攻のタイプと将来の計画は関連しているか？（確率値 = 0.047）

E. 2年生マガジンでは15歳の無作為サンプルに、性交経験があるかどうか（はい、いいえ）尋ねた。男子と女子では回答に差があるかどうか見るものであった。（確率値 = 0.029）

F. ベトナム戦争中、死体の（殺された敵の数）が各軍の部隊から毎週報告された。これらの数の最後の桁はかなりでたらめだったと思われる。しかし、その数がでっちあげだったかもしれないという疑惑が持ち上がった。これをテストするために、死体のカウント数の大きな無作為標本が調べられ、最後の桁が0あるいは5だった頻度が記録された。でたらめな数をでっちあげている人たちがこれらの桁を使う頻度は、でたらめな可能性が示す頻度よりも低いと、心理学者たちは述べている。（すなわち、103の方が100とカウントするよりも「本当」らしく聞こえる。）もしデータが本物のカウント数なら、0か5で終わる数の比率は、約0.20となるだろう。（確率値 = 0.002）

G, シャーロック ホームズの冒険のひとつで、彼は犯行現場で犯人が残した足跡を発見し、それらの距離を測った。多くの人の見本をとり身長と歩幅の距離を測り、彼は容疑者の身長を予想できると自信を持って発表した。どうやって予想するのか？

項目 24

自宅保有者に売られるタイプのラドン検知器は、どの位正確なのか？この疑問に答えるため、大学の研究者たちは、ラドン1リットル当たり105ピコキュリーまでさらず特別な部屋に検知器を12個置いた。表示された探知機の数値は以下のものであった。ミニタブからの記述統計学のプリントアウトがそれに続く。

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	96.6	119.3	104.8	101.7

変数	N	Mean	Median	TrMean	StDev	SE Mean	Minimum	Maximum	Q1	Q3
数値	12	104.13	102.75	103.54	9.40	2.71	91.90	122.30	96.90	109.90

1. このタイプすべての探知機の中間値20という数値は、105という真値とは異なっているというもっともらしい証拠はあるのか？ $\alpha = .05$. で適切な仮説検定を行いなさい。
2. この問題と関連しているタイプ I エラーは何か？

3. この問題と関連しているタイプIIエラーは何か？
4. もし、探知機の数値が5ピコキュリー低い(105であるべきところを実際は100を示す)なら、タイプIIのエラーの確率は何か？

項目 25

アメリカ食品医薬品局 (FDA) の研究によると、コーヒー一杯には平均 115mg のカフェインが含まれており、カップごとの量は淹れ方によって 60 から 180mg の幅がある。自分の好きな淹れ方を使って、95%で 5mg 以内のカフェイン含有の中間値を概算するために、FDA の実験をやってみたくて仮定しなさい。このような問題では、母集団の標準偏差を範囲の 4 分の 1 まで概算することができる。何杯のコーヒーを淹れなければならないか？

項目 26

ある特定の薬を適用することで、禿げた男性に髪を修復できるとある広告では主張している。この主張を試すために使用する実験計画の概要を作りなさい。この実験で 20 人の禿げた男性に使用するお金があると想定しなさい。

項目 27

幼児の鉄分不足に関する研究では、様々な食事養生計画に従っている幼児のサンプルを比較した。ひとつのグループには母乳で育った幼児を含んでおり、一方他のグループの子供たちは鉄分の補強を含まない粉ミルクで育っている。ここに、12 カ月の時点での血中のヘモグロビンレベルに対する結果の概要がある。

グループ	N	\bar{X}	s
母乳	23	13.3	1.7
粉ミルク	19	12.4	1.8

子供たち(母乳で育ったものと粉ミルクで育ったもの両方)の血中のヘモグロビンレベルは普通に分散していると想定しなさい。観察された違いの統計学的重要性を決定するために、有意性検定を行いなさい。

項目 28

+0.4 の相関関係と-0.6 の相関関係では、どちらの方が強い直線関係を示しているか？自分の選択について簡潔に説明しなさい。

項目 29

内科医のグループが、医療診断テストに使われたテストと同じ入念なテストに対して、嘘発見器を受けさせた。もし 1,000 人の人たちが嘘発見器を受けさせられ、500 人は本当のこ

とを言い、500人は嘘をついたということになったら、嘘発見器は、本当のことを言っている人のうち約185人は嘘つきで、嘘つきのうち120人は本当のことを言っていたということを示すだろうということ、彼らは発見した。嘘発見器テストの応用では、その人が嘘つきであると示されるまで、各個人は本当のことを言う人だと推定される。この問題の背景でタイプIエラーは何か？この問題の背景でタイプIエラーの確率は何か？この問題の背景でタイプIIエラーは何か？この問題の背景でタイプIIエラーの確率は何か？

項目 30

聴覚学者たちは、高齢者のためのカナダのプログラムで、聴覚障害のある患者用のリハビリテーションプログラムを最近開発した。ある特定の高齢者ホームの住人から単純無作為に抽出された30人のサンプルと高齢者たちは、以下のようにコード化された感音性難聴の程度とタイプで診断された：1 = 正常の限度内で聞こえる 2 = 高周波難聴 3 = 軽度難聴 4 = 軽～中程度難聴 5 = 中程度難聴 6 = 中程度～重度難聴 7 = 重度～完全難聴。データは以下のとおりである：

6	7	1	1	2	6	4	6	4	2	5	2	5	1	5
4	6	6	5	5	5	2	5	3	6	4	6	6	4	2

1. データの箱ひげ図を描きなさい。
2. データについて適切な説明をきなさい。
3. このカナダのプログラムにおける、高齢者の難聴の中間値に対して95%の信頼区間を求めなさい。上記のデータの中間値と標準偏差は、それぞれ4.2と1.808である。その区間を説明きなさい。

項目 31

公益事業の会社では、農業関連の顧客は、ピークの時間の料金が異なるなら、この時間帯に電気を使わないようにするのかどうか知りたいと思った。顧客は標準料金のままいくのか、時間帯によって違う構造でいくのか、無作為に割り当てられた。ピーク時とそうでない時間帯の使用を記録する特殊なメーターが設置された；メーターを読み取った技術者は、各顧客がどちらの構造の料金を割り当てられているのか知らなかった。

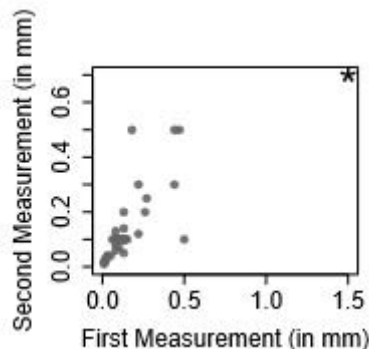
1. これは観察研究なのか、実験なのか？自分の意見を擁護きなさい。
2. 説明変数、反応変数は何か？
3. この取り組みにおける潜在的交絡変数をリストきなさい。
4. これは整合計画なのか？自分の答を弁護きなさい。

項目 32

学期の始めに、私たちは統計学の教科書のページ幅を測った。下は最初に測ったもの 対 二

番目に測ったものの分散図である。

1. この分散について述べなさい。
2. アステリスクがついている点とついていない点の相関関係を概算しなさい。



項目 33

レジャー調査ジャーナルにおける研究は、学業成績とレジャー活動との関係を調査した。159人の高校生の被調査母体のうち、それぞれの人が毎週いくつぐらいのレジャー活動に参加しているのか答えるよう依頼された。そのリストの中で、読書、作文、算数を含む活動は、「学習的レジャー活動」と分類された。結果のいくつかは下の表にある：

	中間値	標準偏差
GPA	2.96	0.71
レジャー活動の数	12.38	5.07
学習的レジャー活動	2.77	1.97

これらの数字に基づき（そして、GPAは0と4の間の値であり活動の数はマイナスにはなりえないということを理解して）、上記の変数それぞれに対して起こりうる歪みについて討論しなさい。

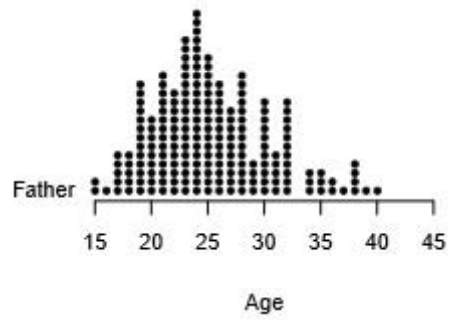
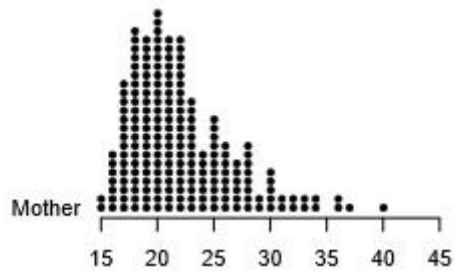
項目 34

イベントAとBは互いに素である。AとBが独立したものになりえるかどうか討論しなさい。

項目 35

200人の母親の被調査者母体と、200人の父親の被調査者母体が抽出された。第一子をもうけた時の母親の年齢と、第一子をもうけた時の父親の年齢が記録された。

1. 母親の年齢に対するデータについて説明しなさい。
2. 父親の年齢に対するデータについて説明しなさい。
3. 分布を比較しなさい。
4. ふたつの母集団を比較したい時、年齢の相関関係を調べるためにある提案がされる。これは良い提案か？それはなぜそうなのか、そうでないのか？



付録 C : 技術の使用例

この例は、現実社会の背景から始まり、トランプを使って学生に物理的に模擬実験させ、模擬実験を自動化するためにコンピューター技術を持ちこんでいる。

コカイン中毒の治療効果を調べる技術を基盤とした擬似体験

コカイン中毒の治療研究は、中毒患者がコカインから離れているのに効果があるふたつの薬を、比較する実験の結果を発表した (DM バーンズ、「コカイン中毒のサイクルを破れ」、サイエンス 241 号、1988、1029-1030 ページ) 治療を求めている 48 人のコカイン中毒者のグループは、無作為に 24 人ずつのふたつのグループに分けられた。ひとつのグループは、デシプラミンという新しい薬で治療され、一方他のグループはリチウムを与えられた。結果は下の表にまとめられ、その表では、元に戻らなかった患者をうまく治療されたとみなした。

	元に戻らず	元に戻る
デシプラミン	14	10
リチウム	6	18

私たちは、この特殊な実験において、デシプラミンのほうがリチウムよりもうまくいったということを観察したが、この改善は統計学的に重要だと結論づけてよいのだろうか。(つまり、薬が同じ位効果的で、それはデシプラミンのグループでもっとうまくいった事例がたまたまその位多く出る無作為の分け方だったとしたら、そのような大きな違いがでると予想しただろうか。) わたしたちは、模擬実験を通してこの疑問に答える。まずトランプを切ることを基本とした実演を実際に行う。それから治療グループの中毒患者の無作為な分け方で違いが見られるコンピューターの模擬実験を使って述べる。

物理的模擬実験

54 枚ひと組のトランプ (2 枚のジョーカーを含む) を取りあげ、黒いカード (スペードかクラブ) 6 枚を取り除く。残りのトランプはコカインの実験における被験者にマッチさせる。赤いカード全部とジョーカーは元に戻ってしまった患者を表し、20 枚の黒いカードはうまく治療できた患者を表している。このひと組のトランプをシャッフルし、それぞれ 24 枚ずつの山になるようにふたつに配ると、中毒患者の分け方を、成功がどちらの薬を飲んだかに依らないふたつの治療グループに分ける模擬実験ができる。これを行い、それぞれのグループを「成功」(黒いカード) と「元に戻る」(赤/ジョーカー) でカウントし、二方向の表に記入しなさい。

	成功	元に戻る
デシプラミン		
リチウム		

いったん表にひとつの数を記入すると、残りも埋められるということに気づきなさい。わかっているように、各治療グループに 24 人いて、20 人は後戻りしない、28 人は戻る（そういう理由から 2×2 の表にはたった 1 自由度しかない時々いうのである）。それでは内容を単純化するために、デシプラミングループの「元に戻らない」の数のように、ひとつのカウントを見てみよう。

トランプを再び、すべてシャッフルし、デシプラミングループに 24 枚配る。そして、黒いカードの枚数を数える。

デシプラミングループの「元に戻らない」の数 = _____

クラスで結果を構想の中にためておきなさい（デシプラミングループに分けられた 24 枚のカードの各無作為グループ黒いカードの数を数える）。実際の実験で観察された 14 の事例と同じくらい（またはもっと）大きな数は黒いカードではどの位頻繁にでてきたか？

もともとのデータの確率値は、どちらの薬も同じ位効果的だと想定して、14 以上の「元に戻らない」の事例がデシプラミングループに行くという無作為の割り当ての比率である。クラスの構想の中にあるデータを使って比率を概算しなさい。

コンピューター模擬実験

14 以上の元に戻らないケースをデシプラミングループに入れる無作為割り当ての比率をより正確に概算するために、コンピューター模擬実験に移る。

ふたつの列と 48 行でできている集合データ（オンラインで提供）から始める。最初の列（治療法）には「デシプラミン」値が最初の 24 行にあり、「リチウム」が残り 24 行にある。2 番目の列（結果）には「元に戻らない」値と「元に戻る」値が、コカイン実験よりもともとあった 2×2 の表にあるデータに合わせるために、書かれている。

結果がどの薬を摂取したのかによらない、治療グループの被験者の新しい無作為割り当てを表すために、コンピューターで「結果」の列の値を並べかえなさい。デシプラミンの治療グループの「元に戻らない」事例の数を数えなさい。そしてその結果をどこかに残しておきなさい。このプロセスを 1,000 回¹¹自動的に繰り返しなさい。

1,000 回の模擬実験の数の分布のヒストグラムかドットプロットを見なさい。デシプラミングループの 14 もの「元に戻らない」事例は、普通とは思えないか？

デシプラミングループに 14 以上の成功がある模擬実験の数を数えなさい（グラフから、あるいは実行可能なら、模擬実験カウント列を区分して）、そしてもともとのデータの確率値のもうひとつの近似値を得るため、1,000 ごとに分けなさい。成功した事例のより大きな数（14）が、偶然デシプラミングループに現れたというのは道理にかなっていると思えるか？あるいは、デシプラミンはコカイン中毒の治療にリチウムよりよく効くかもしれないと、結論づける方がより適切なのだろうか？

- ¹¹ 技術の二者択一：ここで最も難しい段階は、多くの無作為割り当てをすることである。例えばファソムのように、パッケージの中には全くそういった目的のために設計された使用が簡単なツールもある。他にもミニタブのように、前もって作っておき、ループ（繰り返し）で繰り返せるマクロスを使ってちょっとしたプログラミングができるものがある。いくぶん啓発的に欠ける模擬実験が、超幾何分布から無作為なデータの生成もできる統計学のパッケージで、行える。そうすると学生は物理的な無作為化とのつながりを見失ってしまうだろうが。最後に、やる気のある指導者は、必要な模擬実験を行ったり、結果を収集するのにアプレット構成する（ウェブ上で見つかるかもしれない）ことができる。

付録 D：裏付けのない実際的な現実のデータ例

未補正のデータ（勧められない）

下記のデータの最小 2 乗直線を求めなさい。X = 5 のときの Y を予測するのにそれを使いなさい。

X	1	2	3	4	6	8
Y	3	4	6	7	14	20

（より良いがまだ最善ではない）実際のデータ

下のデータはレストランの 6 つのテーブルの客の数と、食事後テーブルに残されたチップの多さを示している。テーブルで食事した人の数から、チップの額を予測するために、データを使って最小 2 乗直線を求めなさい。結果を使い 5 人で食事したテーブルのチップの額を予測しなさい。

食事した人	1	2	3	4	6	8
チップ	\$ 3	\$ 4	\$ 6	\$ 7	\$ 14	\$ 20

批評：作られたデータには何の背景もない（勧められない）。問題は、意義のある解釈の可能性がなく、ただ単に計算するだけということにある。

批評：問題を、より魅力的にする背景が加えられ、学生に実用的な統計学の使用を示している。

本物のデータ（勧められる）

下のデータは、小テストの点数（20 点中）と中間考査の成績（100 点中）を、前期でこのコースを取った 8 人の学生のサンプルについて示している。これらのデータを使い、小テストの点数から中間考査の点数を予想する最小 2 乗直線を求めなさい。小テストと中間考査は今期は同じ位の難易度であり、同じ直線関係を今年適用すると想定し、小テストで 17 点取った学生の中間考査で取る成績は何点が予想されるか？

小テスト	20	15	13	18	18	20	14	16
中間考査	92	72	72	95	88	98	65	77

批評：データは、そのコースを取っている学生にとって、興味がある実際の背景から取りあげている。

文部科学省

平成 24 年度大学間連携共同教育推進事業

「データに基づく課題解決型人材育成に資する統計教育質保証」

統計教育大学間連携ネットワーク

質保証委員会

平成 24 年度 (2012) 活動報告書

2013 年 3 月 25 日 発行

編 集 大学間連携共同教育推進事業連携大学：東京大学・大阪大学・総合研究大学院大学・
青山学院大学（代表校）・多摩大学・立教大学・早稲田大学・同志社大学

発 行 大学間連携共同教育推進事業連携大学：東京大学・大阪大学・総合研究大学院大学・
青山学院大学（代表校）・多摩大学・立教大学・早稲田大学・同志社大学

〒150-8366

東京都渋谷区渋谷 4-4-25

青山学院大学 大学間連携 G P 事務局

非売品・禁無断転載