

統計学教育課程における コンピューターの使用

カート・ハインリックス

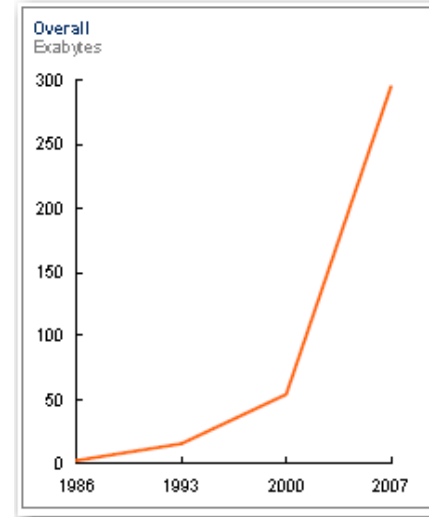


THE
POWER
TO KNOW.

序論 および 指針

- 統計学、コンピューター利用および社会
- コンピューター利用と統計学教育
 - 学生ユーザー
 - アクセスとインフラの傾向
 - 統計学教育におけるコンピューター役割と進化
- どのように学科課程にテクノロジーを統合するか
 - 使用するソフトウェアのタイプ
 - 教科書

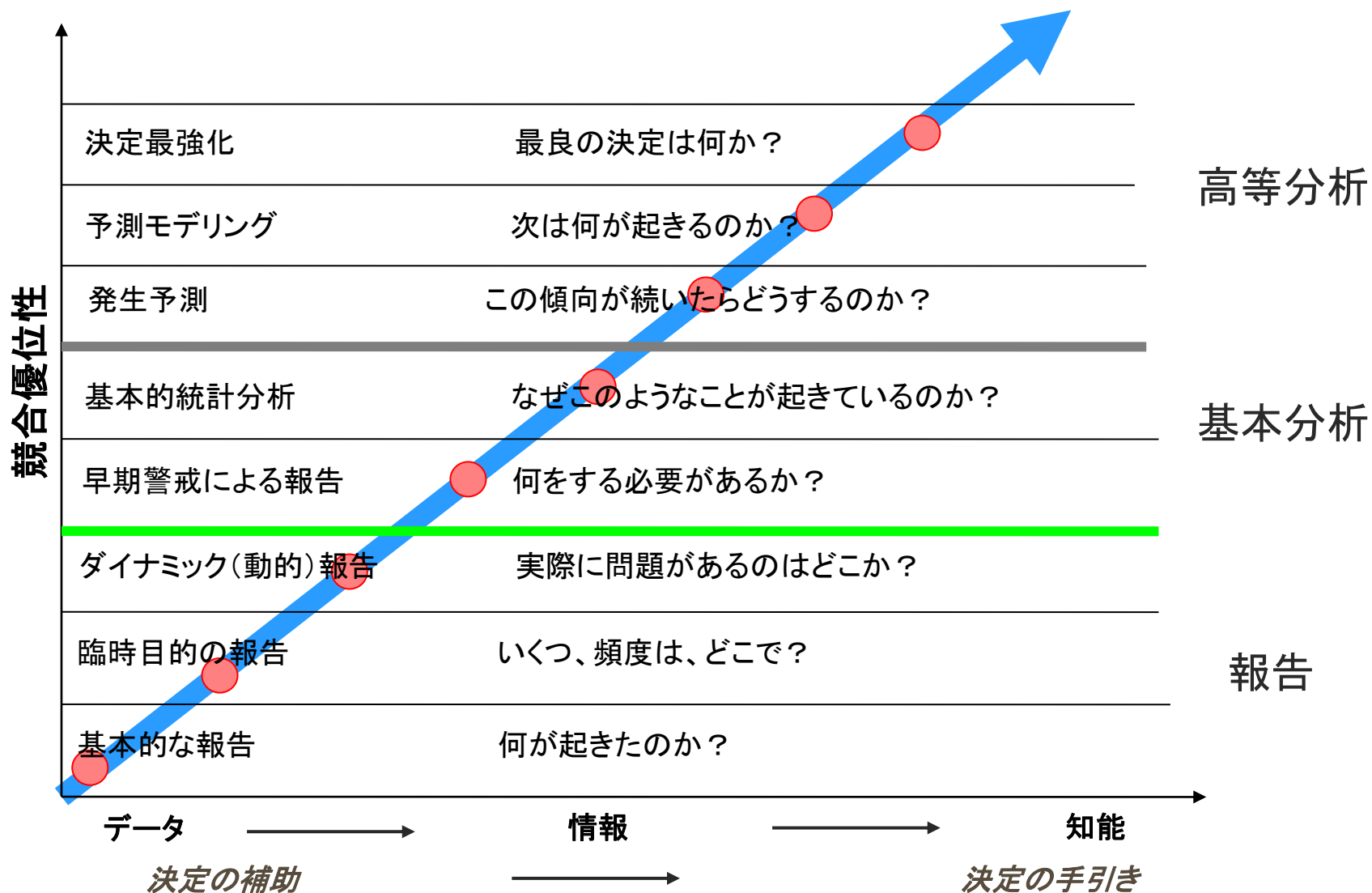
統計の技能は社会にとって重要か？



- データの増加 = ITの最大の課題
 - 回答した会社の47% (ガートナー)
- 「データを収集するのが目標ではなく、そこから洞察を得ることである」 (フォレスターリサーチ、デーブ・フランクランド)
- 「統計学は、データから見識を得る学問である」 (チャンス、2002)
- 分析能力を持つ企業は市場一般よりも業績が良い。
 - ↑アナリティック・インベストメント社はS&P500と比べ64%良い業績をあげた。(アクセンチュア)

分析学を使って成功する

SAS, 高等ビジネス分析学課程



統計の技能は社会にとって重要か？

- データの増加 = ITの最大の課題
 - 回答した会社の47% (ガートナー)
- 「データを収集するのが目標ではなく、そこから洞察を得ることである」(フォレスターリサーチ、デーブ・フランクランド)
- 分析能力を持つ企業は市場一般よりも業績が良い。
 - ↑アナリティック・インベストメント社はS&P500と比べ64%業績が良かった (アクセンチュア)
- 2018年までに分析能力を持つ人材の不足が深刻化する
 - 深い分析能力を持った人材が50%不足 (統計学修士、博士)
 - 決定を下すためにデータを用いる管理職は1,500万人
(マッキンゼー・グローバル・インスティテュート)

McKinsey Global Institute

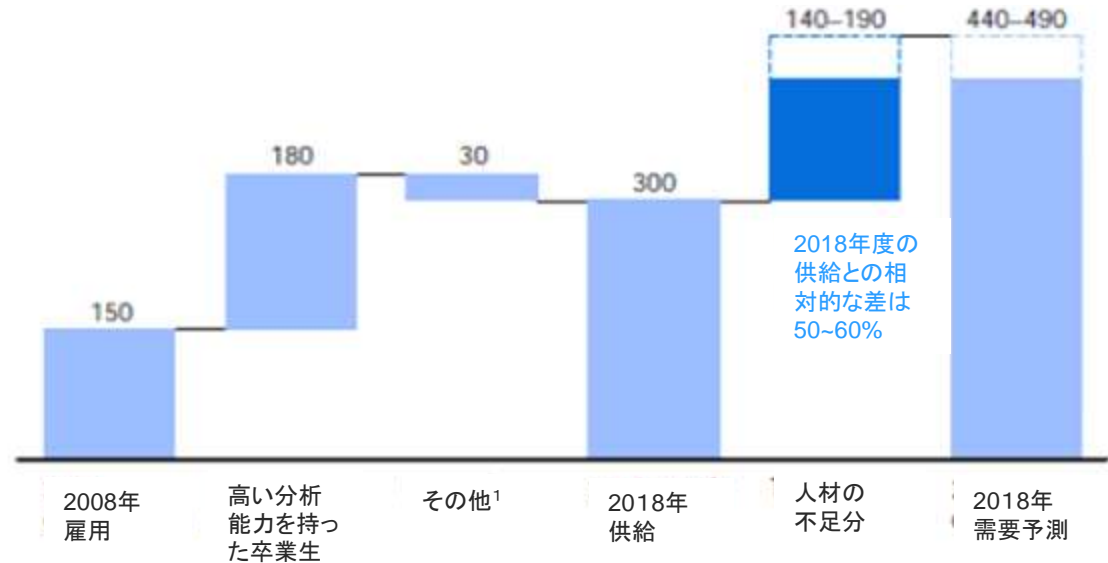


膨大なデータ: イノベーション、競争そして生産性における次の最前線

添付図表4

2018年までに、米国内における高い分析能力を持つ人材の需要は、その供給を50~60%上回る可能性がある

2018年までの高い分析能力をもつ人材の需要と供給
(千人)



¹ その他に供給に影響を及ぼしたのは、人員減少(-)、移民(+)、解雇した下位分析能力をもつ人材の再雇用などである。

出典: 米国労働統計局; 米国国勢調査; ダンアンドブラッドストリート; 企業インタビュー; マッキンゼー・グローバル・インスティテュート分析

統計学課程は二種類の聴講者を支える

1. サービスコースへの導入 (95% の統計課程の登録)
 - 種類 (一般、ビジネス、工学、生物学、心理学)
 - 課題 (難しい取組み) (学期制、大規模クラス、因習的なシラバス (概要))
 - 非常に大きなコースでは、助教などが教えることが多い
 - コンピューター演算: 使いやすさ (JMP、エクセル)

入門的「サービス」コース



統計学課程は二種類の聴講者を支える

1. 入門サービスコース(95%の統計課程の登録)

- 種類(一般、ビジネス、工学、生物学、心理学)
- 課題(難しい取組み)(学期制、大規模クラス、因習的なシラバス(概要))
- 非常に大きなコースでは、下位の教授が教えることが多い
- コンピューター演算: 使いやすさ(JMP、エクセル)

2. 先進課程(米国内の統計学を受講する学生の5%以下)

- 少人数のクラスを教授が最新の研究を踏まえ教えている
- コンピューターを使った演算: プログラミング(SAS、R)

入門的「サービス」コース



コンピューターへのアクセスとクラスでの利用

- 米国の教室の97%に一台以上のコンピューターがある
- 5人に1台の割合 (K-12、2011年米国教育省統計)
- 家庭におけるテクノロジーへのアクセス

- 統計学サービスコースでのコンピューター利用の範囲
 - 研究室や教室内での授業における利用
 - グループワークおよびチームプロジェクトでの利用
 - 生徒個人のノートパソコン

コンピューター利用と統計学教育

- コンピューターが統計教育を変えている
 - 以前の教え方
 - » 人工的 > 実際のデータ、テクニック中心 > データ中心
 - » 「計算の自動化は可能でそうになっていくべきであるように、図解も自動化が可能であり、そうになっていくべきである。」 (ムーア、1992年)
 - 2002年: 50% が授業でコンピューターを使用 (ガーフィールド)
 - 2011年: 76% (米国) が現在導入中、通常使用、常に使用
96%が何らかの使用を報告している (ハサード、2011年)

「コンピューターによる統計の目的は、数字ではなく、洞察を得ることである」 (パーゼン、1998年)

コンピューター利用と統計学教育

- コンピューターが統計教育を変えている
 - 以前の教え方
 - » 人工的 > 実際のデータ、テクニック中心 > データ中心
 - » 「計算の自動化は可能でそうになっていくべきであるように、図解も自動化が可能であり、そうになっていくべきである。」 (ムーア、1992年)
 - 以前に教えていたこと
 - » ブートストラッピング、データの可視化、データ管理
 - » 「教えられた内容は、計算できることによって、形になるのである」 (コップ、2007年)

従来型の制約とコースの特質

制約

- キヤノンについての前提
- 学科に根ざした必須科目(心理学、工学、ビジネス等)
- 14週間
- 生徒の数学と科目領域の履修歴
- 演算テクノロジーの使用可能性: 計算尺、計算機、コンピューター、一覧表(z, t, F...), アップレット
- 重要な実際データ(最近までかなり制限されていた)の入手やアクセスの可能性
- クラスの規模および課題の管理

コースおよび教科書の特徴

- 小さい標本および小さめの標本($n < 20$ および $30 < n < 100$)
- 人為的に作成されたデータおよび非無作為標本
- 計算に焦点をあてた指導
- 割当てを円滑にするための概算に焦点
- 生徒は試験で試されることが多い(数値の結果を求める)「能力」を示す。
- 「過程」に重点が置かれている
- 有意性の検定結果の解釈

計算能力に制約されない統計学入門、カーバー、2011年

ずっと昔のテクノロジーの役割



Entry is area A under the standard normal curve from -∞ to z(A)



Model 1
a
s
Model 1
a
s
b
e
Model 1
a
r

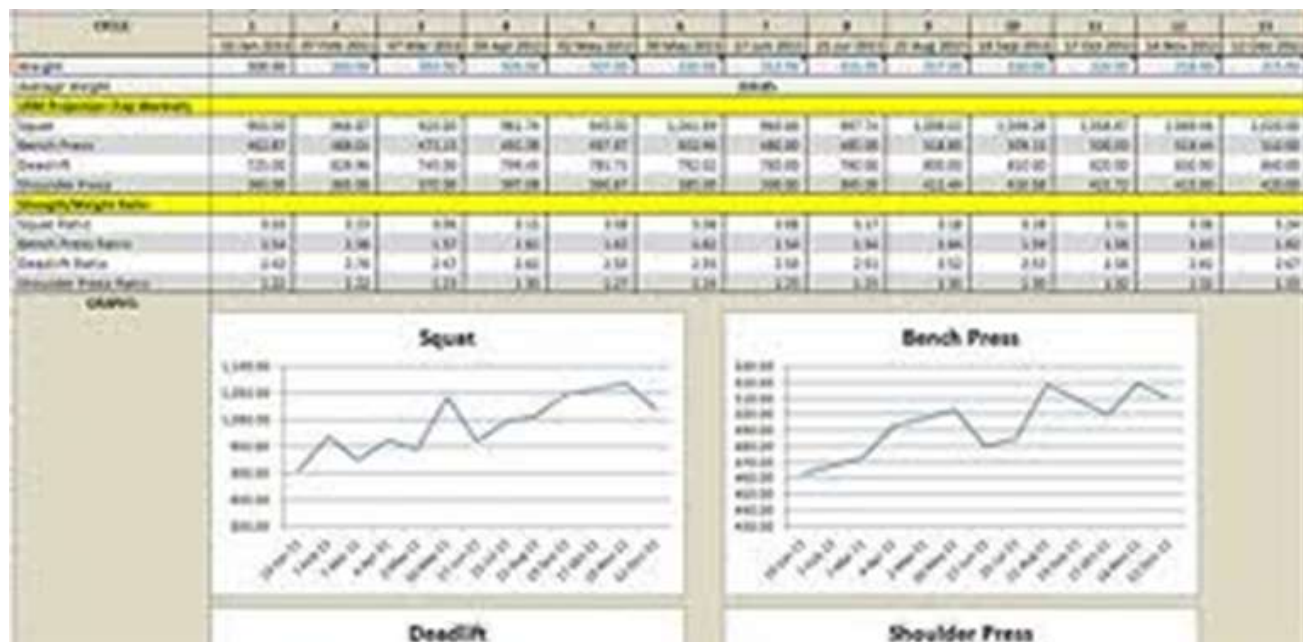
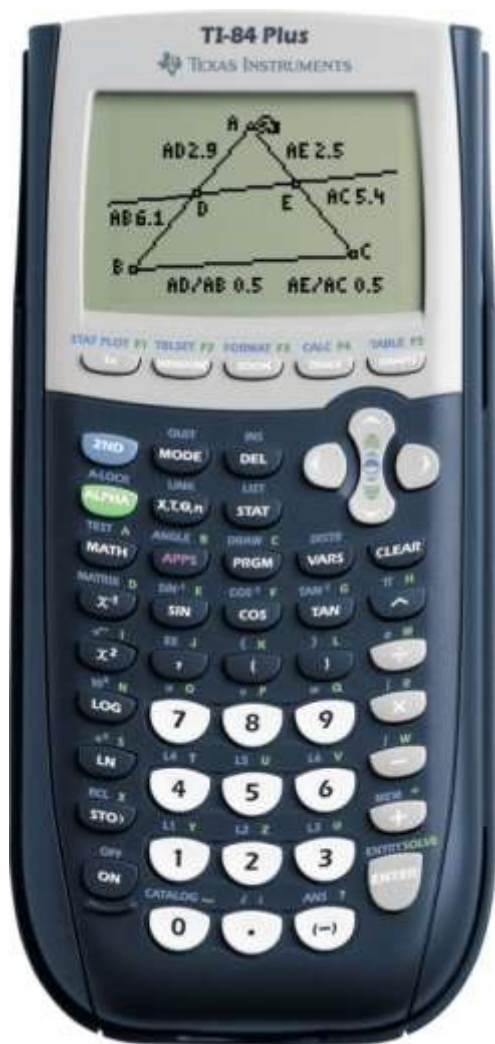
t	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5318	.5358
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5949	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7122	.7156	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7421	.7453	.7484	.7515	.7546
.7	.7577	.7608	.7638	.7668	.7697	.7726	.7755	.7784	.7812	.7841
.8	.7869	.7896	.7924	.7951	.7978	.8005	.8032	.8059	.8085	.8112
.9	.8138	.8164	.8189	.8214	.8238	.8262	.8286	.8310	.8334	.8358
1.0	.8381	.8405	.8428	.8450	.8473	.8495	.8517	.8538	.8559	.8580
1.1	.8601	.8621	.8641	.8661	.8681	.8700	.8719	.8737	.8755	.8773
1.2	.8790	.8808	.8826	.8843	.8861	.8878	.8895	.8912	.8929	.8946
1.3	.8962	.8978	.8994	.9010	.9026	.9042	.9057	.9072	.9088	.9103
1.4	.9119	.9134	.9149	.9164	.9179	.9193	.9208	.9222	.9236	.9251
1.5	.9265	.9279	.9292	.9306	.9319	.9332	.9345	.9358	.9371	.9384
1.6	.9397	.9409	.9421	.9433	.9445	.9456	.9467	.9478	.9489	.9500
1.7	.9511	.9521	.9531	.9541	.9551	.9561	.9570	.9579	.9588	.9597
1.8	.9606	.9615	.9625	.9634	.9643	.9651	.9659	.9667	.9675	.9683
1.9	.9691	.9698	.9706	.9713	.9720	.9727	.9734	.9741	.9747	.9754
2.0	.9761	.9767	.9773	.9779	.9784	.9790	.9796	.9801	.9806	.9811
2.1	.9816	.9821	.9826	.9831	.9836	.9841	.9846	.9850	.9854	.9859
2.2	.9863	.9868	.9873	.9877	.9881	.9885	.9889	.9893	.9897	.9901
2.3	.9904	.9908	.9912	.9916	.9919	.9923	.9926	.9929	.9932	.9935
2.4	.9938	.9941	.9944	.9947	.9950	.9953	.9956	.9958	.9961	.9963
2.5	.9965	.9967	.9969	.9971	.9973	.9975	.9977	.9978	.9979	.9980
2.6	.9981	.9982	.9983	.9984	.9985	.9986	.9987	.9988	.9989	.9990
2.7	.9991	.9991	.9992	.9992	.9993	.9993	.9994	.9994	.9995	.9995
2.8	.9995	.9995	.9996	.9996	.9996	.9997	.9997	.9997	.9997	.9998
2.9	.9998	.9998	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999
3.0	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

IS output is
y clear once
've gotten used
, though not
nera-ready."
nating is
e by template;
cing is automatic
often problematic.

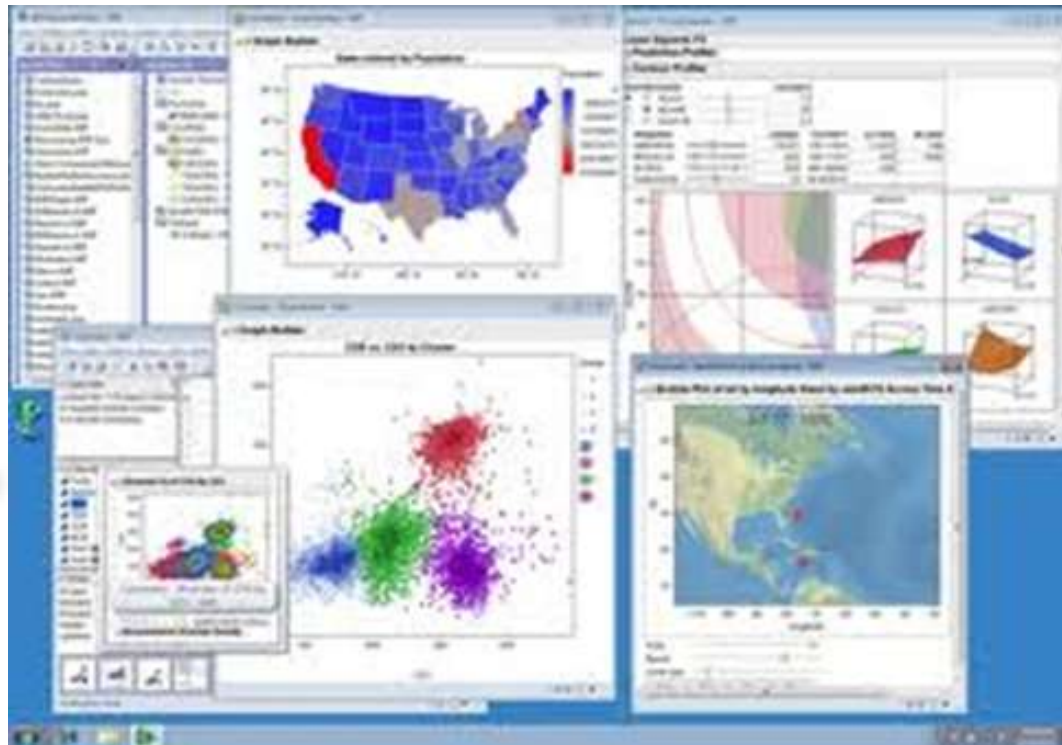
t	Sig.
1.453	.000
1.199	.001

year turnover

比較的現在に近い過去におけるテクノロジーの役割



現在および将来におけるテクノロジーの役割



統計学におけるコンピューター利用の目標と利点

- コンピューターは、コースにおいて興味深いことを実施できるようにしてくれる
- 目標は、以下を改善すること
 - データ、調査および発見への直接関与
 - » 積極的学習：91%の生徒が、講義よりも活動からの方が学びやすいと感じている。
(オーガスト、2002年)
 - 動機づけ、实际的、実践的そして楽しく
 - » 実際のデータは収集、入手、あるいはシミュレーションが可能である
 - » 72%の教授が コンピューターを使うことでその科目を面白くすることができると感じている。 (ハッサード、2011年)
 - 問題解決能力、コミュニケーション能力を発達させる
 - » プロジェクト&報告(レポート)

コンピューターを利用した統計を、コースにどのように取り入れるか

■ データの分析および可視化

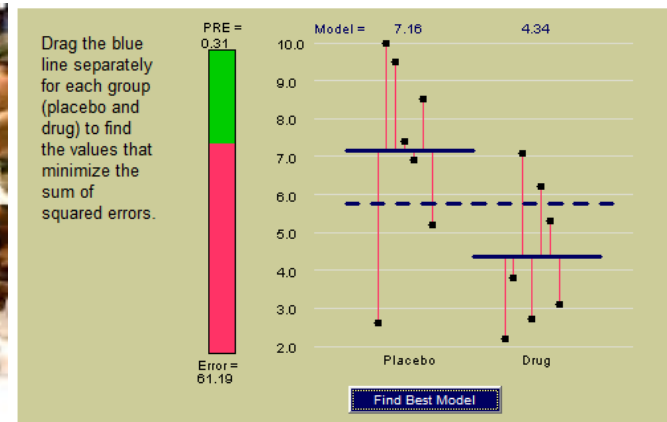
- 76% が通常あるいは常に使用している (ハッサード、2011年)

■ コンセプトの学習

- コンセプトを視覚的に伝える

■ その他の利用

- E-テキストブック、ウェブ・ポータル、クリッカー、コース管理



JMP (統計ソフトウェア) デモ

- EDA(探索的データ分析)、ダイナミックなグラフ作成、可視化
- デモ コンセプトツール

コンピューターの使用は、コースへのプラス または能率化につながるか？

出: 除外する事ができる項目

- 小さなデータベースを用いた手計算の大部分
- 偽データを用いたソフトウェアベースの計算全部
- ヒストグラム ビンズの定義づけ
- 初等確率の大部分
- t、z、F テーブル等の解釈
- 他の分布への正規近似
- Z(信頼)区間と有意差検定

入: 旧「上級」または「軽視されていた」項目

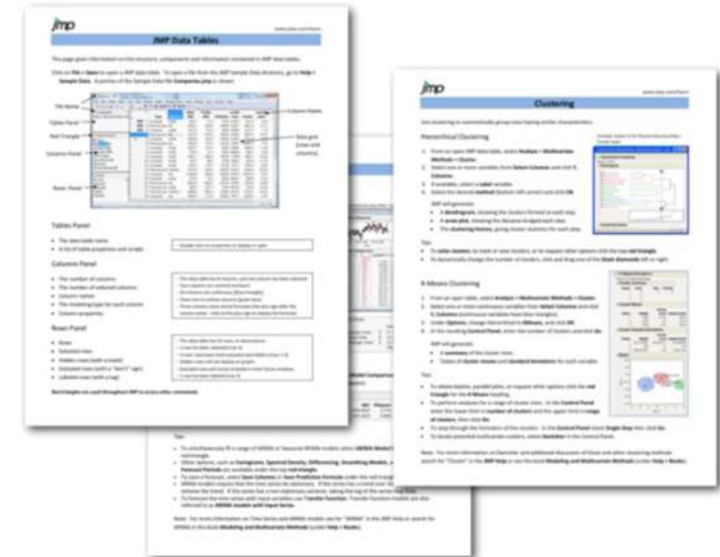
- データ準備、紛失データおよびデータ管理
- 統計調査についての作文
- 再抽出と並べ替え検定
- 標本加重を用いた抽出確率的非SRSデータの分析
- ノンパラメトリック手法
- 非線形モデル
- 多変数モデル

図2: 何が入って、何が出るのか

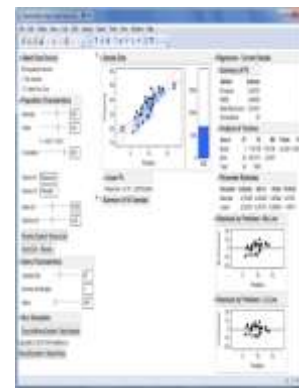
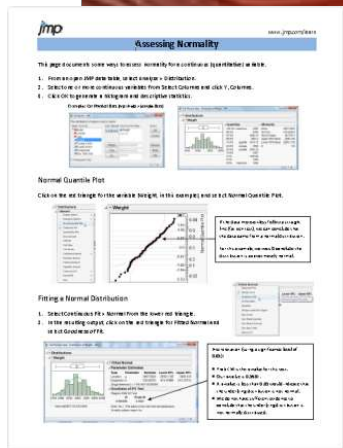
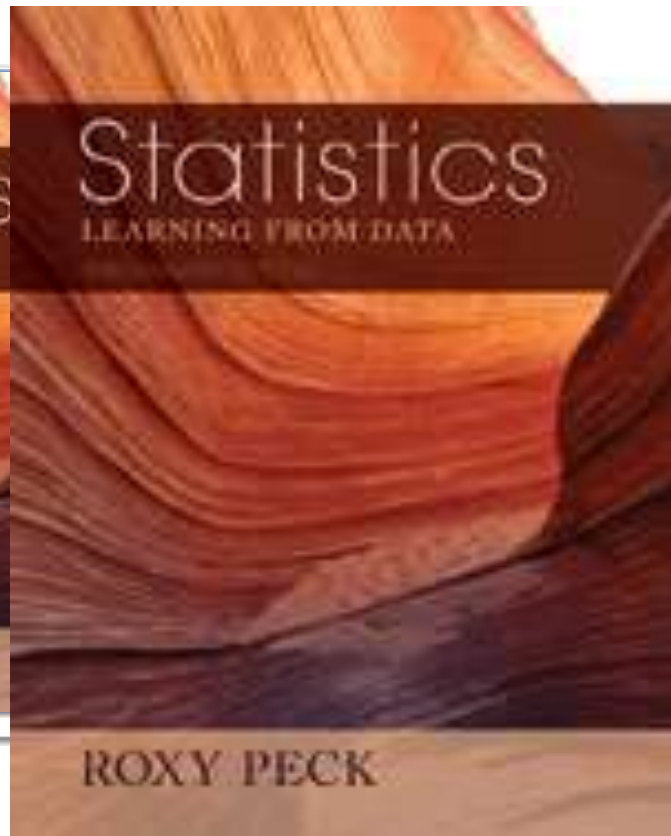
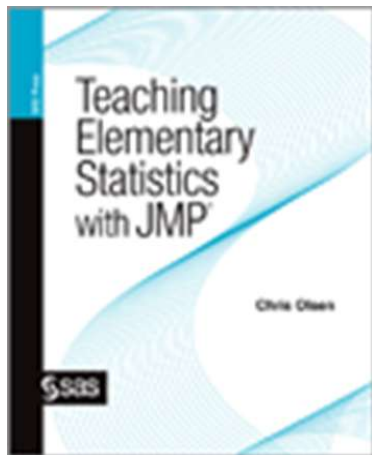
計算能力に制約されない統計学入門、カーバー、2011年

コンピューターをどのようにコースに取り入れるか:

- GAISE 提案を考慮する。(フランクリン&ガーフィールド、2006年)
- どのように、どの程度利用したいのかを決める。
 - テキストブックの統合:出力、指示、データ分析の実施
 - 補足的教材、資料



テキストブックからの導入多数



Case 9 - Direct Mail:
Regression and Forecasting

教室でのタブレット使用

- e-テキストブック、簡単、ダイナミック、(アップル、Kno)
 - 韓国は、20 億ドルをクラウドベースのe-テキストブックのホストに投資 — 3年間 (ITプロトコール 2012年)
 - 英国では、コンピューターを使う学生のうち22%がタブレットを使用すると予測
- 大学生- 2011年の所有率 = 7%、2012年 = 25%
 - 大学生の66%が、タブレットは学習にプラス効果があると考えている
 - 77% は、学習の質を大きく高めると考えている (ピアソン基金、2012年)



どのツールを選ぶか？

- GAISEは以下の内容を考慮するよう提案している：
(チャンス、その他 2007年)
 - データ入力およびインポートの簡単さ
 - 対話方式の能力
 - データ/グラフ/数値の間のダイナミックな連結
 - 聴衆にとっての使い勝手の良さ
 - 生徒にとっての入手可能性、移動可能性

イニシアチブ(率先的行動)と学習の機会

- **ジャーナル:** Technology Innovations in Statistics Education (統計学教育におけるテクノロジー革新)
- **協議会:** USCOTS、ICOTS、JSM
- **オンラインセミナー:** Causeweb
- **ワークショップ:** JMP ワークショップ、AP統計ワークショップ
- **STEM イニシアチブス**



ありがとうございました