

21世紀型ソフトスキルとしての統計思考力の育成
～科学的探究・問題解決・意思決定のための統計教育～

1. 指導要領改訂の背景
21世紀型ワークスキル
2. スキャンズレポート
アメリカ政府レポート
1990年(産業界&教育界を繋ぐ)
3. 問題解決力向上のための
統計教育
4. 分布の読み方・活用

東洋大学経済学部
渡辺 美智子
(統計数理研究所客員教授)

学習指導要領の主な改善事項

言語活動の充実

- 各教科等で批評、論述、討論などの学習

理数教育の充実

- 新しい科学的知見に対応
- **統計に関する内容を必修化**
- 知識・技能を活用・探究の重視
- 日常生活や社会との関連を重視

職業に関する教科・科目の改善

各種産業で求められる知識・技術・資質の育成

情報科「情報の科学」:**データに基づく問題解決**



新学習指導要領解説(改訂の経緯:冒頭)

- 21世紀は、「知識基盤社会」の時代
- 新しい知識・情報・技術が 政治・経済・文化をはじめ社会のあらゆる領域での活動の基盤として飛躍的に重要性を増す時代
- アイディアなど知識そのものや(それを創造する)人材をめぐる国際競争の激化

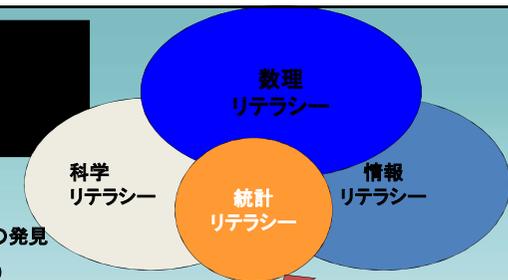
21世紀型ワークスキル:学習とイノベーション

S21世紀パートナーシッププログラム(2002)
Scansレポート(1992)

- ① 創造性と革新性 ② 批判的思考力と問題解決力
- ③ コミュニケーションとチームワーク (PISA2012)

21世紀(知識基盤社会)

科学技術の進展
高度な情報化社会
グローバル化



- ✓ 科学リテラシー
現象と現象間の規則性(因果)の発見
(科学的探究のための方法論)
- ✓ 情報リテラシー
問題解決のプロセス・データから有意な情報抽出(データマイニング)・
- ✓ 数理リテラシー
不確実性の数理モデル
(データのばらつき・分布として記述・読み取り活用)
- ✓ 社会科学との接点:政府統計, 経済統計, 金融統計... ,
計量心理学, 計量文献学, 計量政治学, 計量社会学, 計量経済学, ...

地球をスマートに！

イノベーションの鍵を握る数学
～金融からエンターテインメントまで

- データとテクノロジーの発展
- 「現在私たちは、多くの分野で膨大な量のデータを扱い、そのデータにリアルタイムでアクセスできます」
- 「さらに私たちは、従来よりも大規模な問題解決のための計算能力と、アルゴリズムを迅速に実行するために必要な速度を提供するネットワークを手に入れています」

サプライチェーン：生産予測・顧客認知・マーケティング・価格設定・輸送
公益事業：電力系統全体で電流と電圧を効果的に監視・予測
航空会社：業務・顧客管理(特定のフライトにおけるノショウの確率を予測)
病院・医療、ビデオ・ゲーム、人材管理、金融(株式の分析、リスクと価値の評価)、森林管理(植生タイプ、降雨量、火災発生確率)...

算数・数学科(統計内容) スパイラル、コンピュータの使用、実際のデータ、活用、議論

小学校「数量関係」					
第1学年	第2学年	第3学年	第4学年	第5学年	第6学年
*個数の数え上げ *絵グラフで表現 *絵グラフの読み取り 【質的データの分布】	*身の回りにある数量の分類整理 *簡単な表やグラフの作成 *表やグラフの読み取り	*資料(データ)の分類と整理 *表やグラフでの表現 *表やグラフの読み取り 棒グラフの読み方やかき方 【数量の大きさの比較】	*二つの観点からのデータの分類整理 *二元クロス表の特徴の読み取り *折れ線グラフのかき方と読み方 【二元クロス表】 【時系列、2変数の変化】 【質的データの分布】	*百分率 *データの分類整理 *円グラフや帯グラフ 【相対度数、累積度数】	*資料の平均 *度数分布表 *柱状グラフ(ヒストグラム) *起こり得る場合の数
中学校「資料の活用」			高校		
第1学年	第2学年	第3学年	数学I(データの分析)	数学A・数学活用	数学B
*ヒストグラムや代表値の必要性と意味 *ヒストグラムや代表値を用いて資料の傾向をとらえ説明する *平均値、中央値、最頻値、相対度数、範囲、階級 【量的データの分布の比較】 【コンピュータ、大規模データ】	*確率の必要性と意味への理解 *簡単な確率の計算 *不確定な事象の確率を用いた説明	*標本調査の必要性と意味を理解すること *簡単な場合について標本調査を行い、母集団の傾向をとらえ説明すること *全数調査	*統計の基本的な考え方や理解 *データの整理・分析・傾向の把握 *データの散らばり(四分位範囲(偏差)、分散及び標準偏差) *データの相関 散布図や相関係数	数学A 場合の数と確率 (ア) 確率とその基本的な法則 (イ) 独立試行と確率 (ウ) 条件付き確率 数学活用 データの分析	*標準偏差と母集団の標準偏差の平均、分散及び標準偏差を用いて確率分布の特徴をとらえる *二項分布、正規分布 *統計的な推測 (イ) 母集団と標本 (ウ) 統計的な推測の考え(ウ) 母平均の統計的な推測

科学的探究・問題解決・意思決定のプロセスを通して育成する統計的思考力

アメリカとカナダとニュージーランドの統計教育の比較

国	年齢	年齢														
		5	6	7	8	9	10	11	12	13	14	15				
アメリカ	統計グラフ(表・図)															
	基本統計量															
カナダ	統計グラフ(表・図)															
	基本統計量															
ニュージーランド	統計グラフ(表・図)															
	基本統計量															

1990年代以降
海外が進める統計教育改革
初等・中等・大学基礎教育

OECD 国際学力調査 PISA(生徒) PIACC(大学生・社会人)
数理リテラシー 「不確実性の数理」
科学リテラシー 「科学的探究」
読解力 「図・表・グラフの読み取り」

オーストラリアと日本の統計教育の比較

国	年齢	年齢														
		8	9	10	11	12	13	14	15							
オーストラリア	統計グラフ(表・図)															
	基本統計量															
日本	統計グラフ(表・図)															
	基本統計量															

▽球炎 新ポジション効果絶大

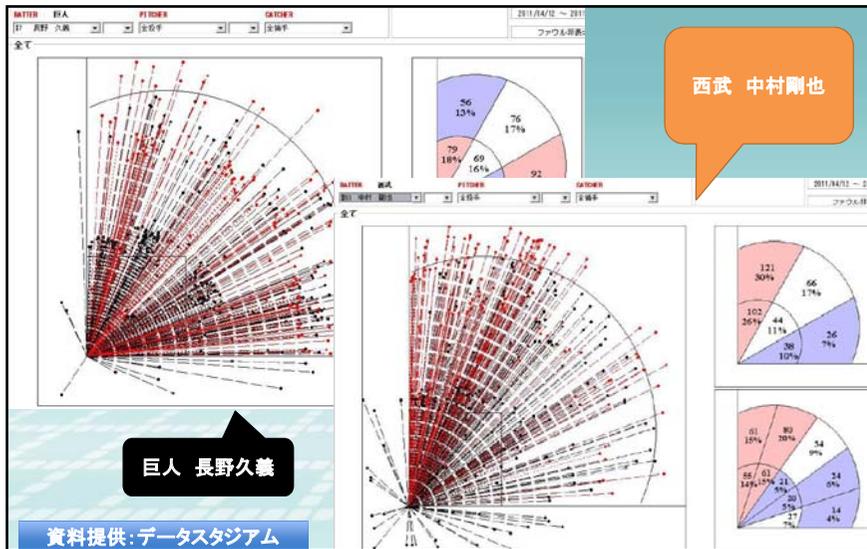
広島 ブラウン監督の野球の常識を覆す改革(2006.4)

守備の新ポジション...ベースボールの発祥から100年以上もかけてできた定位置を変える「大改革」

三塁手と一塁手が塁線を空けて守る守備位置
(米国から持ち込んだ)データに基づく

三塁手付近の打球は三塁線より三遊間側の方が多い。それなら塁線の打球は捨て、確率の高い三遊間側の打球をアウトにする発想。

「塁線を破られると長打になる」と考える日本の野球観では試行もできない。
1年を通して見れば必ず日の目を見る。



産業界と教育者との協働
統計的問題解決力
教育に関する産業界からの要望

科学技術・経済に関する国際競争の激化
 (人材育成)

21世紀型ワークスキルの議論
 米国 スキャンズレポート (1992年)
 イギリス デアリングレポート (1997年)
 OECD DeSeCoプロジェクト
 PISA型学力
 21世紀型スキル (ACT21)
 日本: 経産省
 社会人基礎力
 文科省
 学士力

**LEARNING
 A LIVING:
 A BLUEPRINT FOR
 HIGH PERFORMANCE**

★★★★★

**A SCANS REPORT FOR
 AMERICA 2000**

世界の自動車生産台数推移より

第3章 TOWARD A HIGH-PERFORMANCE FUTURE

日本の企業教育 (統計的問題解決) に学べ・・・

1980年代後半の日本経済発展のミラクルを
 指導したデミング博士

安かろう, ぼろかろう, から世界の
 "Made in Japan"へ
 同じ方法でアメリカも

1980年、NBC
 「If Japan can... Why can't we?
 (日本にできて、なぜ我々にできないのか?)
 ドキュメンタリーを放送
 1992年 産業界から =>

日本統計教育
 昭和21年アメリカ大統領府ライス使節団来日

- ① 日本の政府・民間統計の復活・整備
- ② 戦後の日本の民主主義を育成するためには、ものごとを客観的・合理的に認識し判断する(力)文化、国民の素養(リテラシー)の形成が基礎との勧告
- ③ 産業復興のための企業内統計教育
 統計的品質管理教育 PDCAサイクル
 トヨタ式「KAIZEN」・・・Fact Control
 If Japan can, why can't we・・・1980年・NBC
 統計的問題解決のサイクル

行政・教育・経営

行政評価
 教育評価

2010/10/320 09/01/05 数学教育実践研究会 4

**Dr. Deming:
 The American Who Taught the Japanese About Quality**

スキャンズ型のワークフォース・ノウハウ

初等・中等教育のカリキュラム改革
大学教育の質保証
(学士力・分野別学士力)

OECD PISA 16歳
中学生 学力調査

OECD PIACC
成人版学力調査

ITを活用した問題解決力

A national education-based assessment system
テストからアセスメントへ
"学びを学ぶ力"の自己チェック
国 や学協会での認証, 資格認定

RECOMMENDATIONS FOR THE "LEARNING A LIVING" SYSTEM
THE COMMISSION RECOMMENDS FULL IMPLEMENTATION OF THE FOLLOWING ACTIONS BY THE YEAR 2000:

Preventing Schools

- Workplace know-how (the SCANS foundation and workplace competencies) should be taught along the entire continuum of education, from kindergarten through college.
- Every student should complete middle school (about age 14) with an introduction to workplace know-how.
- Every student by about age 16 should attain initial mastery of the SCANS know-how.
- Every student should complete high school sufficiently proficient in the SCANS know-how to earn a decent living.
- All federally-funded programs for youth and adults should teach the SCANS know-how.

Restructuring Assessment

- A national education-based assessment system should be implemented that will permit institutions to certify the levels of the SCANS competencies that their students have in.
- Public and private employers should define requirements for higher-level competencies.
- Employment-based assessments should permit diagnoses of individual learning needs.

社会・仕事のノウハウを学校教育へ
アカデミックな基礎基本 (only) から
問題解決的思考へ

目的を持って、体系的に問題を解決する訓練を徹底

サイエンス(1996年) 全米科学スタンダード

科学的探究 (Science Inquiry)

科学の本性 (Nature of Science)

科学の文法 (Grammar of Science)

科学とは現象を注意深く観察し、現象と現象間の関連性を見出し、法則として確立すること、
現象を繰り返し測定 = 不確実 (ばらつき)

産業界

統計的問題解決のサイクル

PDCAサイクル

- Plan
- Do
- Check
- Action

DMAIC

- Define the problem
- Measure the process
- Analysis the process
- Improve the process
- Control the process

日本の品質管理教育 (カイゼン)

ばらつきの標準化・管理・予測 : 分布の読み方

欧米の品質管理教育 (シックスシグマ)

欧米の学校教育へ

科学 (理科) ・ 数学 ・ 技術 ・ 社会 ・ 情報

産官学連携の教育の質保証のフレーム
恒常的な教育の質改善 (PDCAサイクル)
(Scansレポート 1992年)

雇用・人材育成
要求されているコンピテンシーの明示

産業界・社会 (Workplace)

教育機関 学校

アセスメント資格認証
学生がどのようなコンピテンシーを持っているのか

学協会の役割

- ① ガイドラインと評価指標の確立
- ② カリキュラムの提示
- ③ 教材・授業モデル・研修機会の提供
- ④ アセスメント

統計的問題解決: SQC

Statistical Problem Solving:

『統計と確率』は、従来に比べ相当重要な位置を与えられるべきである
Committee of Inquiry into Teaching of Mathematics in Schools, 1982, イギリス Mathematical (National Academy of Sciences)
National Council of Teachers of Mathematics, 1989, 2000 全米数学教師協議会
国際数学連合 (数学教育委員会) ・ 国際統計協会 (統計教育分科会) 合同会議, 2008

- The Cobb Report
 - 米国数学会のカリキュラムアクションプロジェクト
 - 1992年に報告書
 - 統計リテラシー (狭義) の教育から統計的思考力 (統計的問題解決力) 育成の教育へ

統計的探究のプロセスの概念を理解し、仕事 (研究) に活用する

統計的思考力: 科学技術推進の第3の腕 (The third arm)

1992年 AP Statistics (検定試験)
9 million dollars 教材開発

統計的(科学的)問題解決の枠組み: Data-based problem solving

課題(issue)からデータで解ける問題(problem)に

客観的評価指標 Y (outcome)の 設定

Yに関する現状分析(分布)

Yをコントロールするための要因Xの探索

XとYの関連性の分析(因果・連関・相関分析)

コントロールできるXを制御して,
目的である指標Yの改善(KAIZEN)を図る

アンケート

本日は、ありがとうございます。
当店ではよりご満足いただけるスーパー銭湯を目指して従業員一同、努力してまいりました。
より良い店づくりのために、お客様のご意見をお聞かせ下さいませよう、お願い申し上げます。
※該当するものに○を付けて下さい。ご氏名・ご住所の記入は不要です。

ご来店日時 月 日 時間 性別 ①男 ②女 年齢 才

1. お風呂が「大浴」はいかがでしたか?
①大変よい ②よい ③普通 ④悪い ⑤大変悪い
(理由:)

2. お食事はいかがでしたか? (料理・提供時間・接客・値段等)
①大変よい ②よい ③普通 ④悪い ⑤大変悪い
(理由:)

3. その他、利用されたサービスはいかがでしたか?
(ボディケア・あかすりエステ・ヘアカット・岩盤浴・その他) ※該当するものに○を付けて下さい。
①大変よい ②よい ③普通 ④悪い ⑤大変悪い
(理由:)

4. 従業員の接客態度はいかがでしたか?
①大変よい ②よい ③普通 ④悪い ⑤大変悪い
(理由:)

5. 清掃は行き届いているでしょうか? (館内・浴室・更衣・トイレ等)
①大変きれい ②きれい ③普通 ④汚い ⑤大変汚い
(理由:)

6. 他店で人気の入浴施設はございますか? ①ある ②ない
(店舗名:)
(理由:)

7. 当店を今後も利用したいと思われますか?
①ぜひ利用したい ②利用したい ③どちらともいえない ④利用したくない ⑤もう利用しない
(理由:)

8. その他、お気づきの点がありましたらお願いします。

条件付き
データ行列

質的変数と量的変数 質的データと量的データ

■ テレビ番組視聴のアンケート結果

性別	スポーツ	ドラマ
男	みる	みない
男	みる	みない
男	みる	みない
女	みない	みる
女	みる	みる
⋮	⋮	⋮

■ 証券会社の売上等のデータ

企業名	資本金(万円)	従業員数(人)	売上(百万円)
A証券(株)	20000	76	7277
B投資信託(株)	80000	87	16408
C証券(株)	13000	140	3970
D証券(株)	1215000	2789	97134
E証券(株)	25020	105	3972
⋮	⋮	⋮	⋮

■ 取引内容を示したデータ

担当者名	得意先名	業種名	商品名	平均単価	数量	金額
熊原	セルフ山上	ミニスーパー	ケチャップ	2304	3	6912
熊原	セルフ山上	ミニスーパー	烏龍茶	215	3	645
渡辺	サキコーポ	スーパー	ドライパン粉	2300	3	6900
渡辺	サキコーポ	スーパー	焼肉のタレ	169	3	507
渡辺	サキコーポ	スーパー	カップうどん	79	3	237
⋮	⋮	⋮	⋮	⋮	⋮	⋮

分析：位置確認+3視点

「異動」が「異動」にどういった影響を与えているのか分析する

どんな時でもまずは最初に「どういった状況の」元のデータと比べる

◆ 相対的に見る

○ 1 全量 (100%) で見たときの異動の割合から位置を確認する

異なる時系列に合わせて2つの視点を併用しつづける

- 比較する
 - グループ (異動) データの異動を比較する
- 関係を見る
 - 異動が対してある (異動) データと目的の関係をみる
- 傾向を見る
 - 異動の推移と傾向の変化を見る

箱ひげ図は分布を比較するグラフ

データの分析(2011/2/20)

現状把握

- Yが質的変数
- 度数分布, パレート図
- Yが量的データ
- 度数分布, ヒストグラム

比較

- X(原因)が質的変数
- Y(結果)が量的変数
- 層別ヒストグラム
- 並列箱ひげ図

関係

- X(原因)が量的変数
- Y(結果)が量的変数
- 散布図・相関関係・相関係数

トレンド

- X(原因)が時間経過
- Y(結果)が量的変数
- 時系列グラフ

18

Are You a Data Detective?

データに基づく知の検証と創出のサイクル

- Problem (身近な課題の明確化)
- Plan (調査・実験研究のデザイン)
- Data (データの収集とデータ表の作成)
- Analysis (データの分析、パターンの発見)
- Conclusion (最初の課題に対する結論と、新たな課題の提示)

統計教育の方法(企業教育・海外の学校教育)

全体の**問題解決のフレームワーク**を先に学び、どの場面で個々の統計の手法を用いばよいかという点を学ぶ

ニュージーランド

小学校から高校までの統計的課題探究, 統計リテラシー, 確率
対象データのレベルと使用する統計スキルを上げながら、毎学年で繰り返されている (数学の時間数の3分の1を占める)

UK: A Problem Solving Approach

イギリス: ナショナルガイドラインの問題解決のフレーム

Problem solving approach

We report back what we found – and compare it with what we expected.

Then we examine our data and make it easier to understand.

Then we collect suitable data.

First we decide what problem to solve and what data we need.

レポート・プレゼン(仮説の結果)

統計教育の方法(企業教育・海外の学校教育)

全体の**問題解決のフレームワーク**を先に学び、どの場面で個々の統計の手法を用いばよいかという点を学ぶ

問題データ(仮説(予想))

問題解決の枠組み

大学生の思考力育成教材(ベネッセ)

「何がしたいのか」「目的に何が関係しているのか」はつきりさせるんだね

感覚でなんとなく類いこもではなく、客観的に分析しないといけないのか

なるほど、確かに気持ちはいい方法を知りたいな

状況整理 (何がしたいか確認する)

分析: 位置確認+3視点 (「原因」が「目的」にどう影響を及ぼしているのか確認する)

判断 (客観的に結論を出す)

現状把握

- ◆ 目的は? どうしたいのかはつきりさせる
- ◆ 影響してそうな要因は? 関係しそうなモノを洗い出す
- ◆ 図やグラフで分かりやすくまとめる
- ◆ 客観的な判断をする

相関分析

- 比較する (グループ(資料)データの違いで特徴を比較する)
- 関係を見る (数値として表せる(初期)データと目的の関係を見る)
- 傾向を見る (初期の推移と傾向の変化を見る)

因果の仮説

批判的思考力

因果か? の議論
公平な比較?

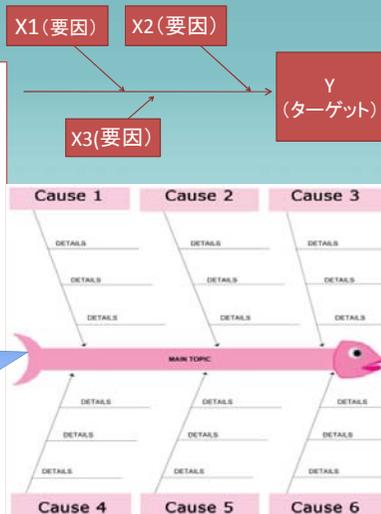
算数・数学科(統計内容)

スパイラル, コンピュータの使用, 実際のデータ, 活用, 議論

小学校「数量関係」					
第1学年	第2学年	第3学年	第4学年	第5学年	第6学年
* 個数の数え上げ * 絵グラフで表現 * 絵グラフの読み取り 【質的データの分布】	* 身の回りにおける数量の分類と整理 * 簡単な表やグラフの作成 * 表やグラフの読み取り	* 資料(データ)の分類と整理 * 表やグラフでの表現 * 表やグラフの読み取り 棒グラフの読み方やかき方 【数量の大きさの比較】	* 二つの観点からのデータの分類整理 * 二要素クロス表の特徴の読み取り * 折れ線グラフのかき方と読み方 【二要素クロス表】 【時系列, 2変数の変化】	* 百分率 * データの分類整理 * 円グラフや帯グラフ 【相対度数, 累積度数】 【質的データの分布】	* 資料の平均 * 度数分布表 * 柱状グラフ (ヒストグラム) * 起こり得る場合の数
中学校「資料の活用」			高校		
第1学年	第2学年	第3学年	数学I(データの分析)	数学A・数学活用	数学B
* ヒストグラムや代表値の必要性と意味 * ヒストグラムや代表値を用いて資料の傾向をとらえ説明する * 平均値, 中央値, 最頻値, 相対度数, 範囲, 階級 【量的データの分布の比較】(コンピュータ, 大規模データ)	* 確率の必要性と意味への理解 * 簡単な確率の計算 * 不確定な事象の確率を用いた説明	* 標本調査の必要性と意味を理解すること * 簡単な場合について標本調査を行い, 母集団の傾向をとらえ説明すること * 全数調査	* 統計の基本的な考え方の理解 * データの整理・分析・傾向の把握 * データの散らばり(四分位範囲(偏差), 分散及び標準偏差 * データの相関 散布図や相関係数	数学A 場合の数と確率 (7) 確率とその基本的な法則 (4) 独立な試行と確率 数学活用 データの分析	* 母集団と標本 確率変数の平均, 分散及び標準偏差を用いて確率分布の特徴をとらえる * 二項分布, 正規分布 * 統計的な推測 (7) 統計的な推測の考え(ウ)母平均の統計的な推
現状分析	要因図の作成	要因分析	質的×量的 量的×量的	層別ヒストグラム	並列箱ひげ図
				散布図	トレンド

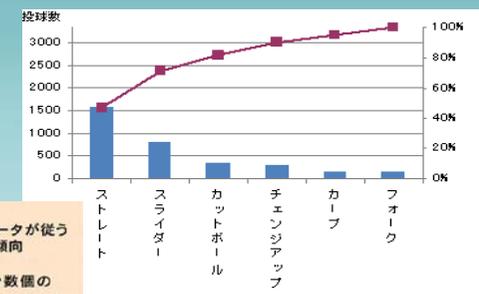
問題解決の道具

1. チェック・リスト (度数分布表)
2. パレート図
3. ヒストグラム
4. 特性要因図 (Fish bone's Diagram, Ishikawa's Diagram, Cause & Effect Diagram)
5. 層別 (グループに分けて比較)
6. グラフ&管理図
7. 散布図
8. 箱ひげ図



職場の問題の95%はこれらの道具で解決できる。

質的データの分布 (パレート図)



■ 全体の過半数のデータが従う中心的(標準的)な傾向
 ■ 標準から外れた少数個のデータの特性

球種名	投球数	累積投球数	割合%	累積割合%
ストレート	1560	1560	47.0%	47.0%
スライダー	814	2374	24.5%	71.5%
カットボール	339	2713	10.2%	81.7%
チェンジアップ	281	2994	8.5%	90.2%
カーブ	172	3166	5.2%	95.3%
フォーク	155	3321	4.7%	100.0%
総計	3321		100.0%	

data

条件付き分布 (確率)

データ分析は層別から

$$\text{特化係数} = \frac{\text{部分集団での比率}}{\text{全体集団での比率}}$$

たまたま見つけるから
見つけるべくして見つける

層別効果を検出する特化係数

ランナー状況別の球種に関する度数分布表

球種名	走者なし	1塁	2塁	1, 2塁	3塁	1, 3塁	2, 3塁	満塁	全体
ストレート	894	308	96	99	25	56	32	50	1560
スライダー	480	162	63	45	15	18	10	21	814
カットボール	208	65	21	16	2	10	4	13	339
チェンジアップ	158	53	27	19	4	8	6	6	281
カーブ	134	21	5	7	3	0	2	0	172
フォーク	81	27	17	7	5	7	5	6	155
総計	1955	636	229	193	54	99	59	96	3321

ランナー状況別の球種に関する度数分布表(%)

球種名	走者なし	1塁	2塁	1, 2塁	3塁	1, 3塁	2, 3塁	満塁	全体
ストレート	45.7%	48.4%	41.9%	51.3%	46.3%	56.6%	54.2%	52.1%	47.0%
スライダー	24.6%	25.5%	27.5%	23.3%	27.8%	18.2%	16.9%	21.9%	24.5%
カットボール	10.6%	10.2%	9.2%	8.3%	3.7%	10.1%	6.8%	13.5%	10.2%
チェンジアップ	8.1%	8.3%	11.8%	9.8%	7.4%	8.1%	10.2%	6.3%	8.5%
カーブ	6.9%	3.3%	2.2%	3.6%	5.6%	0.0%	3.4%	0.0%	5.2%
フォーク	4.1%	4.2%	7.4%	3.6%	9.3%	7.1%	8.5%	6.3%	4.7%
総計	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

層別効果を検出する特化係数

球種名	走者なし	1塁	2塁	1, 2塁	3塁	1, 3塁	2, 3塁	満塁	全体
ストレート	0.97	1.03	0.89	1.09	0.99	1.20	1.15	1.11	1.00
スライダー	1.00	1.04	1.12	0.95	1.13	0.74	0.89	0.89	1.00
カットボール	1.04	1.00	0.90	0.81	0.36	0.99	0.66	1.33	1.00
チェンジアップ	0.96	0.98	1.39	1.16	0.88	0.96	1.20	0.74	1.00
カーブ	1.32	0.64	0.42	0.70	1.07	0.00	0.65	0.00	1.00
フォーク	0.89	0.91	1.59	0.78	1.98	1.51	1.82	1.34	1.00
総計	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

データ分析は層別から

<ダルビッシュ有;2010年のランナー状況別にみた球種分布>

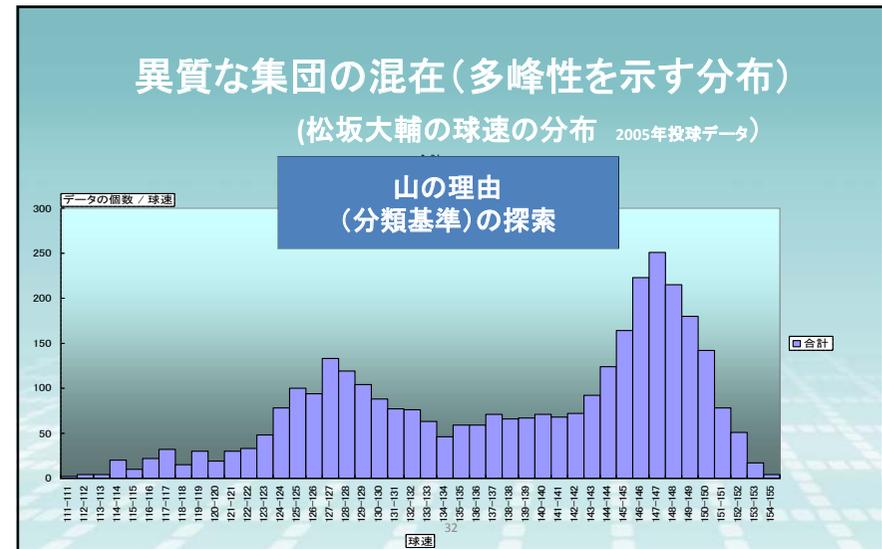
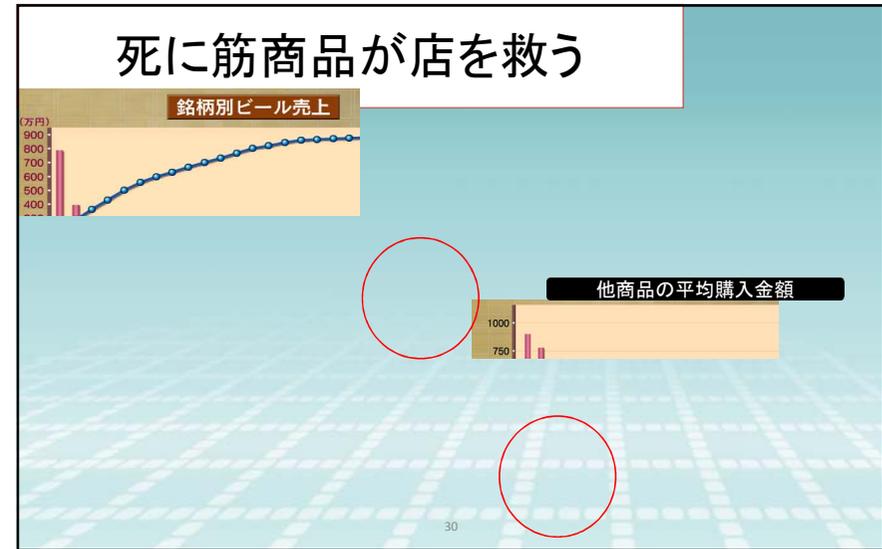
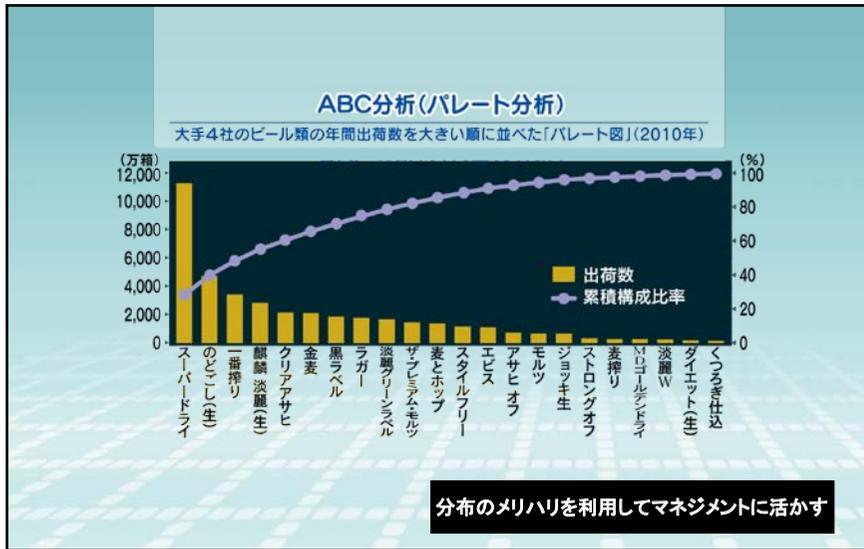
【割合】

走者状況	なし	一塁	二塁	三塁	一二塁	一三塁	二三塁	満塁	全体
ストレート	32.3%	25.1%	23.0%	42.2%	26.0%	29.5%	26.1%	39.6%	30.1%
スライダー	23.4%	23.0%	29.0%	25.3%	27.0%	17.0%	23.9%	20.8%	23.9%
カットボール	14.9%	24.5%	17.0%	12.0%	18.6%	20.5%	23.9%	22.6%	17.2%
チェンジアップ	9.7%	9.5%	3.9%	6.0%	5.6%	2.3%	0.0%	1.9%	8.3%
フォーク	7.2%	6.2%	7.8%	1.2%	11.2%	10.2%	2.2%	1.9%	7.1%
カーブ	6.0%	8.1%	10.2%	8.4%	7.4%	8.0%	13.0%	11.3%	7.1%
合計	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

【特化係数】

走者状況	なし	一塁	二塁	三塁	一二塁	一三塁	二三塁	満塁	全体
ストレート	1.07	0.83	0.76	1.40	0.86	0.98	0.87	1.32	1.00
スライダー	0.98	0.96	1.21	1.06	1.13	0.71	1.00	0.87	1.00
カットボール	0.87	1.43	0.99	0.70	1.08	1.19	1.39	1.32	1.00
チェンジアップ	1.17	1.14	0.47	0.72	0.67	0.27	0.00	0.23	1.00
フォーク	1.01	0.86	1.09	0.17	1.56	1.43	0.30	0.26	1.00
カーブ	0.84	1.14	1.44	1.19	1.05	1.12	1.83	1.59	1.00
合計	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

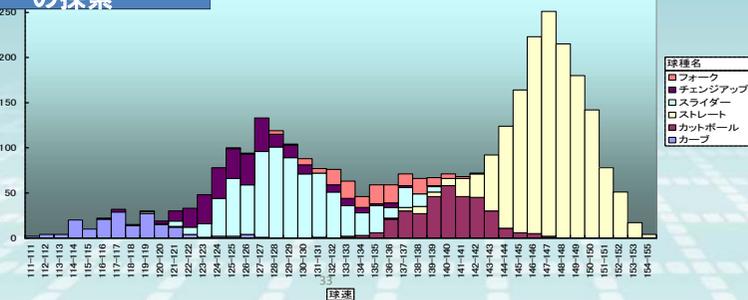
たまたま見つけるから
見つけるべくして見つける



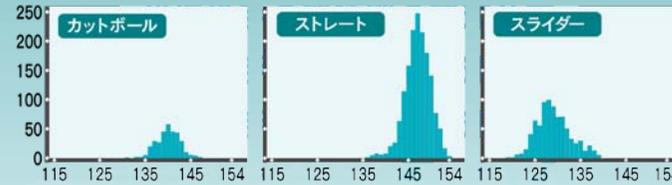
異質な集団の混在 (多峰性を示す分布)

(松坂大輔の球速の分布 2005年投球データ)

山の理由
(分類基準)
の探索

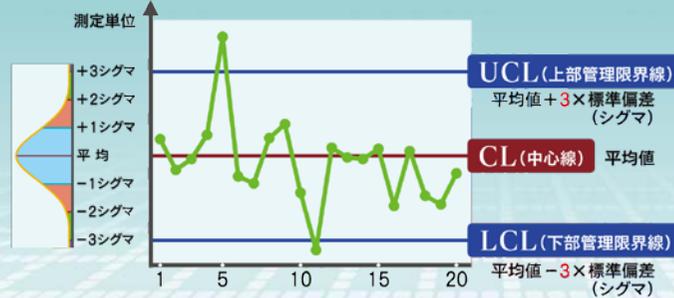


松坂投手の球種別の球速分布



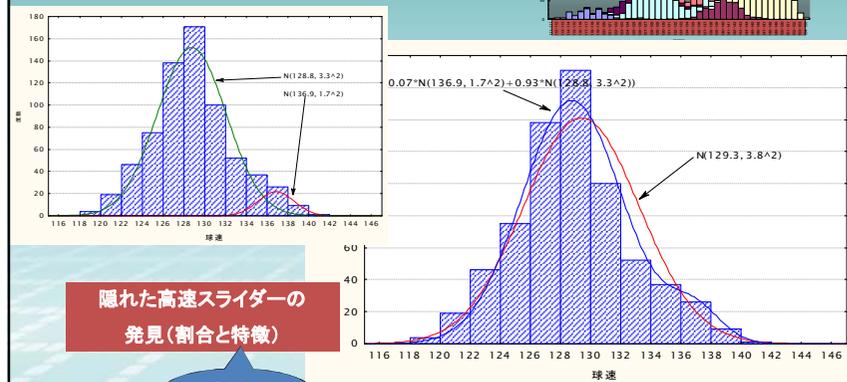
	データ(ケース)数	平均値 (km)	中央値 (km)	分散	標準偏差 (km)
ストレート	1560	147.1	147.0	7.92	2.8
カットボール	339	140.0	140.0	6.80	2.6
スライダー	814	129.1	129.0	14.05	3.7
スライダー(ダルビッシュ有)	764	125.3	125.0	22.18	4.7

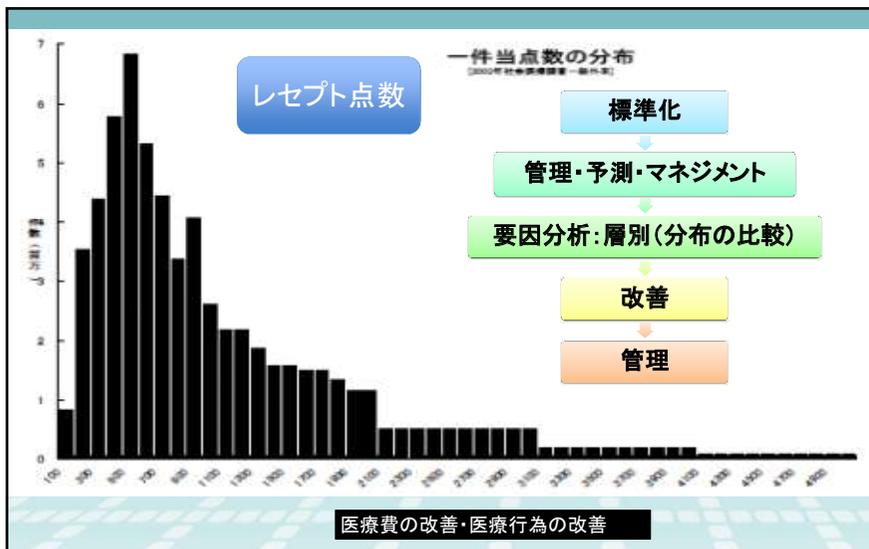
管理図(control chart)



潜在している異質性の発見

正規分布の混合を当てはめる





PISAの問題

地球の平均気温は上昇している。新聞や雑誌には、二酸化炭素排出量の増加が20世紀における温暖化の主因であるとする記事がよく転がっている。

図1 二酸化炭素排出量 (横軸: 1960~1990年)

図2 地球の平均気温 (横軸: 1860~1990年)

太郎さんは、この2つのグラフから、平均気温が上昇したのは二酸化炭素排出量が増加したためという結論を出しました。

太郎さんの結論はグラフのようなことを根拠としていますか。
(2) ともだちの花子さんは、「この結論を教える前に、温室効果に影響を及ぼす可能性のある他の要因が一定であることを確かめなければなりません」と言っています。花子さんが言おうとした要因を一つあげてください。

サッカーの4バックと3バック、どちらが？
データ: ジーコジャパン72試合の結果

システム	試合数	平均失点数
3バック	40	0.85
4バック	32	1.13

システム → 失点数
相手の実力

「喫煙」と「余命」の分析

- 過去の喫煙習慣のある人とない人の高齢者の余命を追跡調査で調べる場合、「喫煙」と「生命力」が交絡して、単純集計すると「喫煙」ありの場合が、「余命」が長くなることもある
- ある一定年齢まで生きているという条件で対象者を選ぶために「喫煙」グループに「生命力」が強い人が多く入ってしまう

データの分析(2011/4/23) 38

公平な比較? => 科学的探究

処置	サイズ	有効数	有効率
A	40	30	0.75
対照群	32	16	0.50

無作為(ランダムな)割り付け

システム	試合数	失点数	平均失点数
3バック	40	34	0.85
4バック	32	36	1.13

データの分析(2011/4/23) 39

ノーアウト一塁が発生したイニングの作戦と結果

バントの有無別の平均得点

送りバント	平均得点	得点確率	ケース数
しない	0.9407	0.4347	4810
する	0.8318	0.4240	993
全体	0.9221	0.4329	5803

バント無効!? 公平な比較?
科学的探究を通して問題解決をし、意思決定に繋げる統計思考力

データの分析(2011/4/23) 40

抽出された条件での送りバントの有無別の平均得点

条件探索が重要

選手	項目	範囲	得点		
打者	打点	≥ 12.045	送りバント	平均	ケース数
	打率	≥ 0.254		しない	0.5780
次打者	打点	≥ 33.075	する	1.0000	36
	打率	≤ 0.380	全体	0.6461	223
	打率	≥ 0.271			
投手	防御率	$2.106 \sim 3.294$			
	奪三振数	≥ 6.048			

2011/08/21 新課程勉強会 41

イギリス(ポーランドマス教材) : 政府 5億円
企業 4億円

対話型データに基づく問題解決教材

2011/08/21 新課程勉強会 42

分析 データ取得法 データ グラフ 分布 統計量 文脈

- 分析**
 - データを読む(グラフ・基礎統計)
 - データの背景を読む(文脈)
 - データとデータの間を読む(連鎖・相関)
 - 隠れたデータを読む(Big X)(因果に近づく)
- データ取得法**
 - 実験(条件制御) 無作為化割付け
 - 非実験データ(調整)
 - 調査(無作為抽出)
 - 観察
- データ**
 - カテゴリ(分類コード) 名目・順序尺度
 - 離散型
 - 連続型
 - 時系列
 - データ化
 - データ変換
 - 連続 \Rightarrow コード化
 - 変数変換
 - 層別
- グラフ**
 - 棒・折れ線・円グラフ
 - 複合グラフ
 - ドットグラフ
 - 幹葉グラフ
 - 箱ひげ図
 - ヒストグラム
 - (\Rightarrow 遊率密度関数)
 - 散布図
 - 時系列
- 分布**
 - 経験分布
 - 多峰性・単峰性
 - 外れ値
 - 対称性・非対称性
 - 理論分布
 - 確率分布
 - 標本分布
- 統計量**
 - 最小値 最大値
 - 種類値
 - 中央値 四分位数 5数要約
 - 範囲 四分位範囲
 - 平均 標準偏差
 - 相関
 - 回帰
 - 時系列
- 文脈**
 - 自分・家族
 - 身の回り・趣味
 - 自然
 - 地域
 - 国
 - 世界
 - 社会・経済・自然科学

2011/08/21 新課程勉強会 43

数理科学研究の推進 **数理科学教育の推進**

数理科学研究の重要性: 「21世紀の科学」(2004年)(米国国家科学技術会議)
自然・社会科学の対象の複雑性が増したこと
ICT化した計測技術による超大規模データセットへの取り組み

- ✓米国科学財団(NSF) 2004年度の重点領域 **数理科学(Mathematical Science)**
基礎数学と統計科学の融合
“巨大データに関する数学的・統計的挑戦”,
“不確実性の管理とモデリング”,
“複雑な非線形システムのモデリング”
- ✓NSF(2008年度から)
“データから知識へ”
“自然、人工物、社会システムにおける複雑性の理解”
- ✓文部科学省平成20年度の戦略目標
「多様で大規模な情報から『知識』を生産・活用するための基盤技術の創出」

指導要領の
今回の改訂のポイント
理数教育の推進
統計内容の充実

2011/08/21 新課程勉強会 44

統計思考 Statistical Thinking

データを扱う科学
(不確実性の数理, 発見科学)
Grammar of Science
Tools and theories for Data Analysis

国や社会の姿を映し出す「鏡」
<測る>
進むべき方向を示す「羅針盤」
<予測する>
経済や社会の内部構造に迫り、メカニズムを解明する「内視鏡」
<制御する>

『政府統計の構造改革』 公共財としての統計

学術・行政・経営などあらゆる場面の意思決定プロセスの科学化を支える
Fact Control
Evidence Based Practice, Policy Design

- 医療と医療
 - 遺伝学
 - 臨床試験
 - 疫学
 - 薬学
 - 医療の標準化
- 自然科学
 - 気象
 - 化学
 - 物理
- 環境
 - 農業
 - 自然保護
 - 森林破壊
 - 野生動物保護
- 社会科学
 - 保険・金融
 - アクチュアリー
 - 年金
 - 金融工学
 - 証券
- 行政
 - 企画立案
 - 行政評価
 - 政府統計
- ビジネス
 - 経済分析
 - 経営分析
 - マーケティング
 - 情報システム
 - ゲーム・ギャンブル

JSSC 統計検定 Japan Statistical Society Certificate

日本統計学会公式認定
資格を取って仕事を選べる自分に

統計検定 Japan Statistical Society Certificate

「統計検定」とは、統計に関する知識や活用力を評価する全国統一試験です。データに基づいて客観的に判断し、科学的に問題を解決する能力は、仕事や研究をするための21世紀型スキルとして国際社会で広く認められています。日本統計学会は、中高生・大学生・職業人を対象に、各レベルに応じて体系的に国際通用性のある統計活用能力評価システムを研究開発し、統計検定として資格認定します。

統計検定試験情報

試験日 2011年11月20日(日)

新着情報 | 試験情報 | 創設の経緯 | 問い合わせ(Q&A)

<http://www.toukei-kentei.jp/>