

京都大学  
データ科学イノベーション  
教育研究センターの取組

京都大学

情報学研究科

国際高等教育院 附属

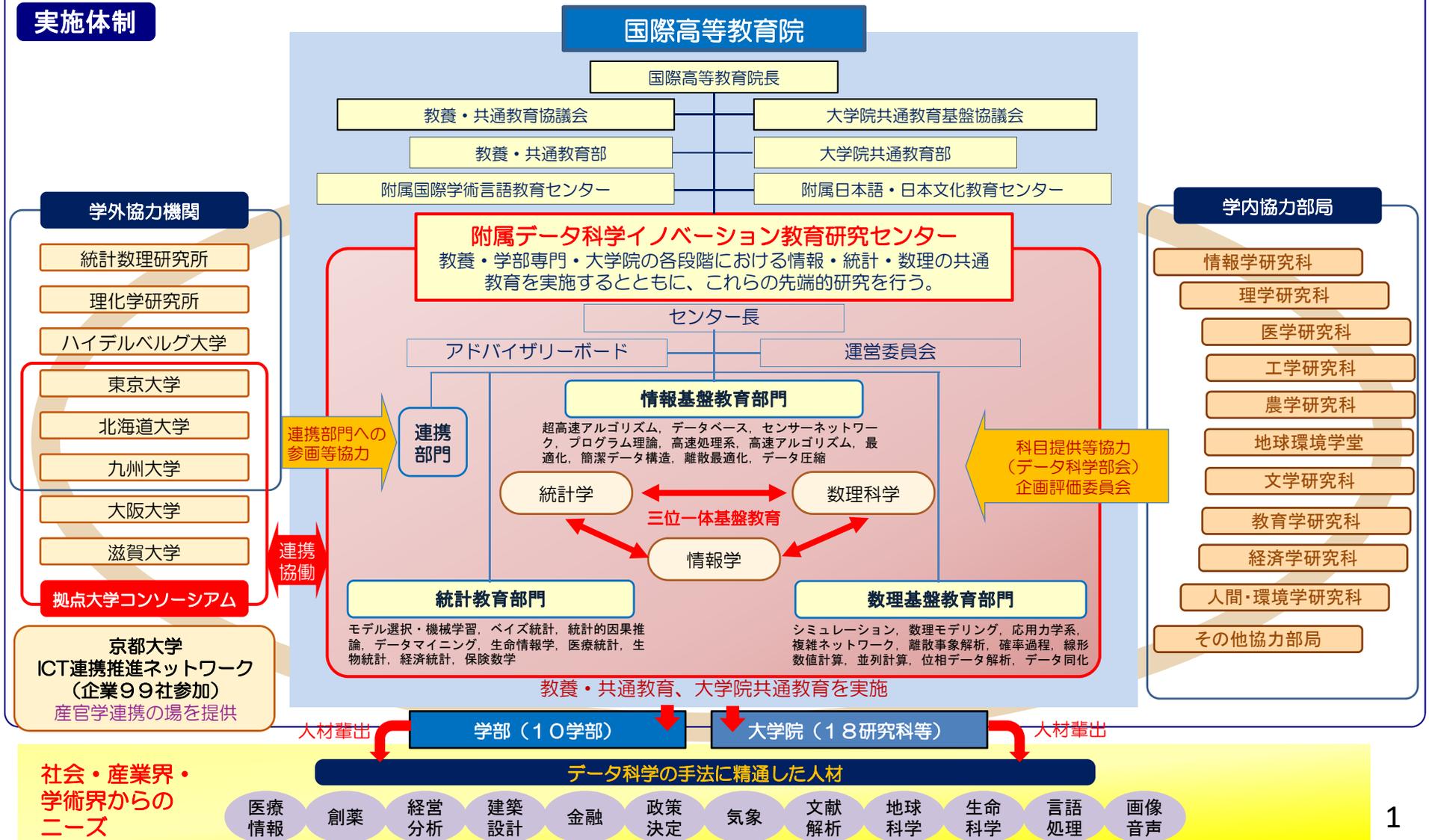
データ科学イノベーション教育研究センター企画評価委員

山本章博

## 事業概要

論理力の涵養を根幹とした21世紀の基礎教養としての情報・統計・数理の基盤教育を担う全学的な組織体制を整備し、直近の課題であるデータ科学者の養成とともに、中長期的には第4次産業革命をトップレベルで支える人材育成を行う。

## 実施体制



# センターのミッション

- 数理・データサイエンスの全学的な教育（一般教育・専門基礎教育等）の実施，カリキュラムの設計・教材作成等
- 多方面にわたる応用展開を念頭に新たな価値の創出ができる人材育成に向けた教育の実施（例えば，数理・データサイエンス分野と文系分野を含む様々な他分野との連携等）
- 全国的なモデルとなる標準カリキュラムの作成・普及（センターでコンソーシアムを形成し，協働して取り組む）
- 数理・データサイエンスと社会とのつながりについてもって教えることができる教員の養成（FD等の充実）
- 地域や分野における拠点として，取組成果の他大学への展開・波及
- 大学，産業界及び研究機関等と連携したネットワークを形成し，実践的な教育の実施

附属データ科学イノベーション教育研究センター

www.z.k.kyoto-u.ac.jp/introduction/innovative-research-and-education-in-data-science

京都大学 国際高等教育院 English

全学共通科目への意見箱 事務局お問い合わせ サイトマップ

新入生の方へ 在学生・学内の方へ 学外の方へ 検索

国際高等教育院について 教養・共通教育とは 施設案内・アクセス Q&A

国際高等教育院について

- ・ 教育院長のあいさつ
- ・ 教育院のポリシー、ミッション
- ・ 組織体制
- ・ 教員名簿
- ・ 主な活動と取り組み
- ・ 英語で学ぶ教養・共通科目
- ・ 広報誌
- ・ 点検・評価

トップページ > 国際高等教育院について > 附属データ科学イノベーション教育研究センター

## 附属データ科学イノベーション教育研究センター

データ科学イノベーション教育研究センターは、2017年4月、論理力の涵養を根幹とした21世紀の基礎教養としての情報学・統計学・数理科学に関する教育及びこれに必要な調査研究等を行うために、国際高等教育院に新たに設置された組織です。

- 概要
- 教員名簿

**2017年4月1日設立**



取組内容・成果指標との関連性

論理力の涵養

教育の課程・スキルのレベル・習得する能力・人材像

高度専門教育  
(博士後期課程)  
使いこなせたら  
創り出せ

専門教育  
(修士課程)  
知るほどに、使うほどに  
楽しくなる

専門基礎教育  
(学部2・3回生)  
使いながら  
深く知る

全学基礎教育  
(学部1・2回生)  
楽しみながら  
基本を知る

高大連携  
未来に向けて  
夢を見る

レベル3+ : 世界トップレベルの研究・革新的トレンドの創生指導的研究者

Wisdom & Innovation

レベル3 : 知識から知恵を見出し、創造につなげる力  
研究者・次世代育成者(棟梁\*1)  
大規模データ解析を活用する研究者

Knowledge

レベル2 : 情報から法則性などの知識を取り出す力  
データ分析技術者(独り立ち\*1)  
業務例: 大規模データの解析・モデル開発

Information

レベル1 : データから分布などの情報を抽出する力  
データ処理技術者(見習い\*1)  
業務例: データ分析の前処理や基礎分析

Data

レベル0 : 基本リテラシー  
情報・統計・数理に関する一般常識・基本知識の獲得

Dream

レベル-1 : 動機付け  
様々な例を通して面白さを垣間見る

大学の取り組みに協力

\*1: データサイエンティスト協会の定義より  
<http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

先端研究におけるデータ科学の応用の実践  
企業・連携研究機関との共同研究

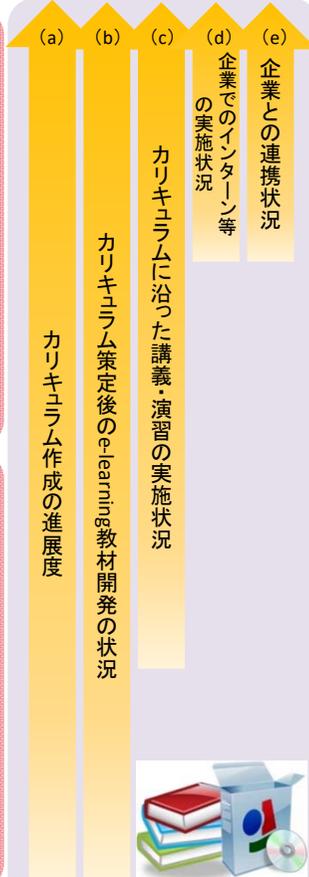
専門的な理論  
実問題への応用  
科目例: 統計的機械学習  
データ分析応用演習  
数理モデリング

汎用的な専門知識の獲得  
事例集による基礎演習  
科目例: データ処理基礎演習  
データベース基礎  
シミュレーション基礎

基本知識の習得  
事例集を中心とした入門  
科目例: 統計入門  
プログラミング基礎  
線型代数学

高校数学の発展としてのデータサイエンス  
取組例: ELCAS, 特色入試  
オープンキャンパス

全国的なモデルとなる標準カリキュラム・教材開発への貢献



統計科学副学位プログラム

博士後期課程・修士課程の大学院生に対して本来の専門に加えて、データ科学に関する**一定の科目群**からなるプログラムを履修する場合副学位として「統計科学副学位プログラム修了証」を授与

本来の専門分野での学位に加えてデータ科学への深い知見を有することで、企業におけるデータ科学者としての活躍などキャリアパスの多様化を図る。

連動

社会人・博士人材への付加教育

主に科目等履修生として、相当するレベルの科目を履修。**一定の科目群**の履修者に対し修了証を授与

博士学位取得者への教育

(レベル3相当)

理系・文系を問わず一定の数理的知識、計算スキル、データからの知識を獲得する訓練を受けて博士学位を修得した者を対象に、職務上必要とされる高度なデータ科学的手法や先端的研究内容を教授

修士課程修了者への教育

(レベル2相当)

個人の既取得スキルに応じて、職務上必要とされるデータ科学的手法を教授

成果指標との関連性

地域における拠点として他大学等への展開・普及

# 目次

- センター設立までのデータ科学教育
  - 全学共通科目「数理統計」「統計入門」
  - 学部入試と「統計」
  - 大学院研究科横断型教育プログラム
  - 学部専門基礎教育
- データ科学イノベーション教育研究センター
  - 国際高等教育院附属センター
  - 基本コンセプト
  - 大学院研究科横断型教育プログラムの刷新
- センターの目指すところ
  - 情報・統計・数理の「三位一体」
  - 情報学研究科・医学研究科との連携
  - 課題

# センター設立までの データ科学教育

# 京都大学における教育制度の枠組

カテゴリ	対象学生	主な実施部局
教養・共通教育	主に学部1-2回生	国際高等教育院
専門基礎教育	主に学部2-4回生	各学部
専門教育	修士(博士前期)課程	各研究科
高度専門教育	博士後期課程	各研究科

カテゴリ	対象学生	主な実施部局
全学共通科目	学部1-4回生	国際高等教育院
研究科横断型プログラム	大学院	各研究科

# 全学共通科目「数理統計」

- 自然科学科目群 主として2回生 理系学生向
- 「確率論基礎」ならびに「微分積分学(講義・演義)A,B」および「線形代数学(講義・演義)A,B」, または「微分積分学A,B」および「線形代数学A,B」の内容を既知とする。

## 内容

### 1. 標本論

母集団と標本、無作為抽出、層別無作為抽出、母平均、母分散、標本平均、標本分散、不偏分散、統計的推測の考え方

### 2. 数理統計に現れる分布

正規分布、カイ二乗分布、F分布、t分布

### 3. 推定

点推定、区間推定、信頼係数、正規母集団の平均・分散の推定、2標本の平均差・分散比の推定など

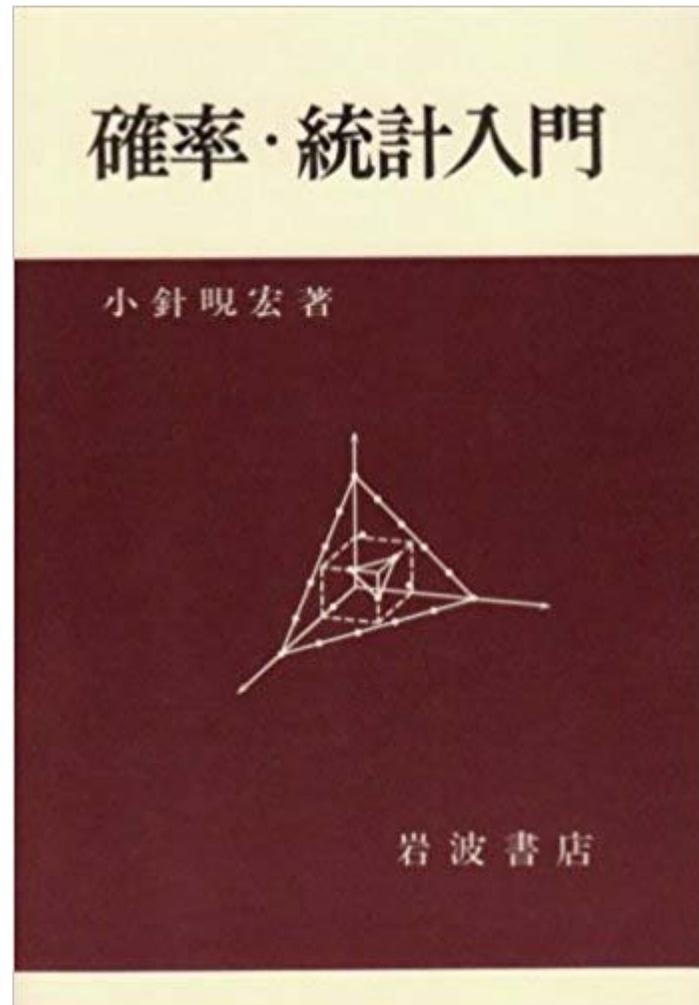
### 4. 仮説検定

帰無仮説、対立仮説、有意水準、第一種の過誤と第二種の過誤、母数に対する仮説の検定(正規分布の平均、分散など)、適合度検定、分散分析の考え方

### 5. 回帰分析(時間の都合により省略することがある。)

# 全学共通科目「数理統計」

• 1973/5/31 刊行



# 全学共通科目「統計入門」(内容)

- 自然科学科目群 全回生 全学生向
- 主に文系の学生が高校で履修したレベルの数学の知識を必要とする。
  - 統計ならびに統計学に関する基本的な考え方を中心に講義することで、より発展的な統計・統計学の学習への礎となることを目指す。
  - 統計分析手順の機械的な利用や解釈だけを講義するのではなく、その基礎となる考え方を学ぶことを目指している。しかし、統計学的命題について、厳密な数学的証明は避け、あくまで統計・統計学のエンドユーザーとして必要とされる直感的な理解を目指す。
  - 講義中に実施する小テストとレポート(JMPIによる演習課題)によって、講義で解説した基本的概念・原理の理解度、統計データの収集・集計・分析・解釈についての応用力を評価する。

## 内容

- 量的データの確認と要約, 質的データの確認と要約
- 二元分割表とカイニ乗検定, フィッシャーの正確検定, リスク比とオッズ比, リスク差の検定・推定
- 検定・推定と標本規模の関係, 確率分布と極限定理, 平均の差の検定, 分散分析, 相関, 回帰

# 全学共通科目「統計入門」(実施体制)

- 平成26年度 国際高等教育院に**統計教育特別部会**を設置して設計
  - 社会統計を専門とする教員, 医療統計を専門とする教員, 各学部代表教員から構成
- 平成27年度より実施, 前期・後期各2クラス
  - 全身の「社会統計学」は前期・後期各1クラスのみ
- 平成28年度から前期・後期各3クラス開講
  - 平成28年度 前期各クラス履修者数  
160名(医学部医学科クラス指定, **修士1名含む**)  
94名(人間健康1回クラス指定, **修士4名含む**)  
194名(人間健康2回クラス指定, 工学部80名含む)
- 情報学研究科**知能情報学専攻**教員6名が担当
  - 平成28年度から 神谷教授(脳科学), 鹿島教授(機械学習), 黒橋教授(自然言語処理)など

# 入試と「統計」

## 京都大学工学部の例

統計に関する内容は、京都大学入試でも 20 年にわたり明示的に除外されている。(京都大学事務本部教育推進・学生支援部入試企画課にご協力いただきました。)

### 2015–2017 年度入試 (2009 年公示, 2012 年度から実施の指導要領に準拠)

数学 I, 数学 II, 数学 III, 数学 A, 数学 B

- 数学 I, 数学 II, 数学 III, 数学 A: 全範囲
- 数学 B: 「数列」, 「ベクトル」 ( 「確率分布と統計的な推測」 を除外。 )

# 入試と「統計」

## 京都大学工学部の例

2006–2014 年度入試 (1999 年告示, 2003 年度から実施の指導要領に準拠)

数学 I, 数学 II, 数学 III, 数学 A, 数学 B, 数学 C

- 数学 II: 「微分・積分の考え」においては一般の多項式を扱うこととし, 立体の体積を含む.
- 数学 B: 「数列」「ベクトル」「数値計算とコンピュータ」の3分野を出題範囲とする. ただし, 「ベクトル」は直線・平面の方程式を含むものとし, 「数値計算とコンピュータ」はプログラミングを除く. (「統計とコンピュータ」を除外.)
- 数学 C: 「行列とその応用」「式と曲線」「確率分布」の3分野を出題範囲とする. ただし, 「確率分布」は「確率の計算」のみを範囲とする. (「統計処理」を除外.)
- 数学 III: 簡単な微分方程式及び曲線の長さを含む.

# 入試と「統計」

## 京都大学工学部の例

1997–2005 年度入試 (1989 年告示, 1994 年度から実施の指導要領に準拠)

数学 I, 数学 II, 数学 III, 数学 A, 数学 B, 数学 C

- 数学 A: 「数と式」, 「数列」の 2 分野から出題する. (「平面幾何」, 「計算とコンピュータ」を除外.)
- 数学 B: 「ベクトル」, 「複素数と複素数平面」の 2 分野から出題する. (「確率分布」, 「算法とコンピュータ」を除外.)
- 数学 C: 「行列と線形計算」, 「いろいろな曲線」の 2 分野から出題する. (「数値計算」, 「統計処理」を除外.)

# 入試と「統計」

## 京都大学工学部の例

1985–1996 年度入試 (1978 年告示, 1982 年度から実施の指導要領に準拠)

数学 I, 代数・幾何, 基礎解析, 確率・統計, 微分・積分

※ 文系学部に対しては, 「確率・統計のうち, 資料の整理, 統計的な推測の 2 分野からは出題しない。」旨の注記あり.

# 入試と「統計」

## 京都大学工学部の例

1977–1984 年度入試 (1970 年告示, 1973 年度から実施の指導要領に準拠)

数学 I, 数学 IIB, 数学 III

※ 注記なし (!).

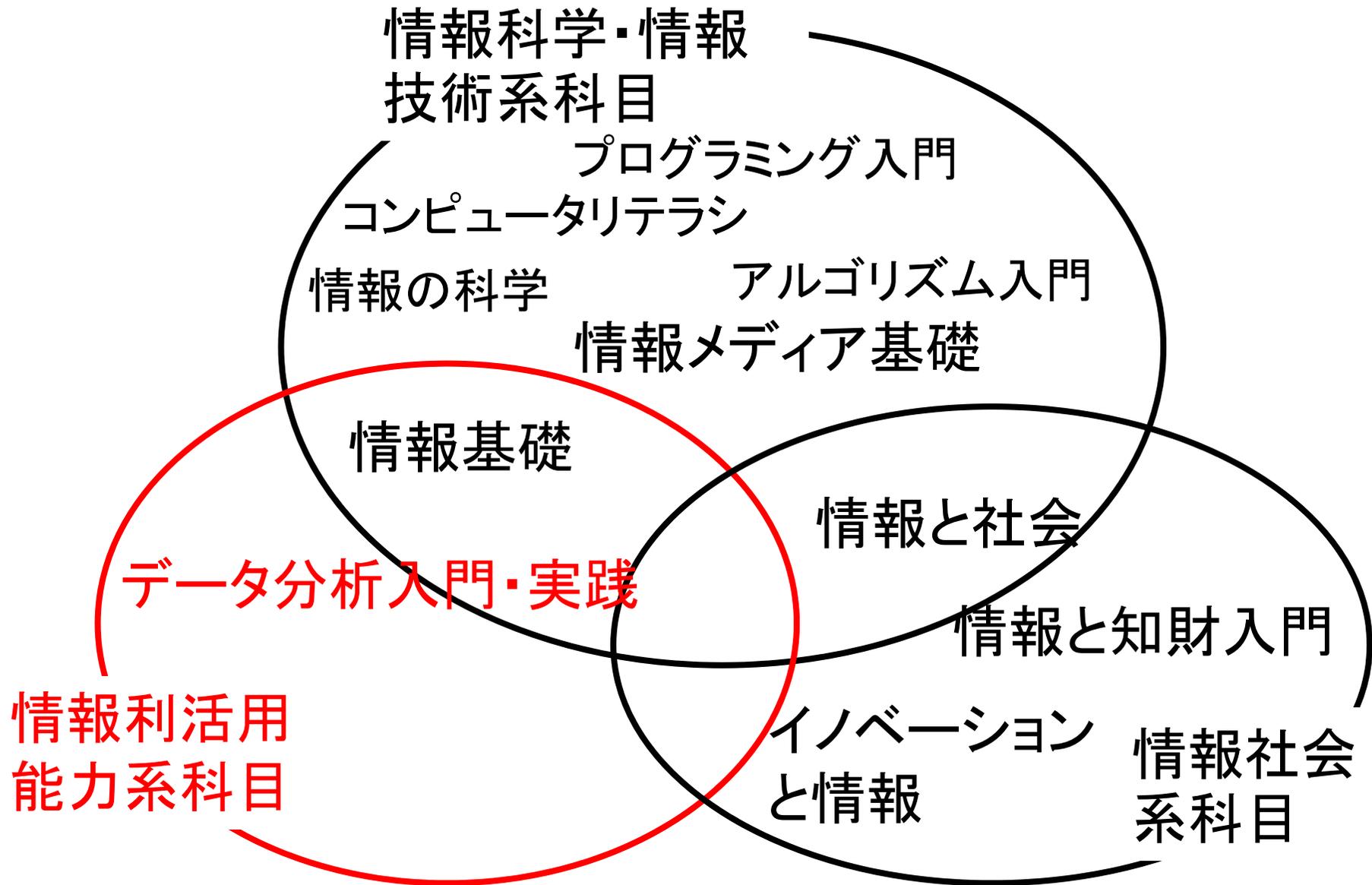
# 大学院研究科横断型教育プログラム

- 対象：大学院学生
- 大学院生が広い視野を持ち、新しい学問領域を創造できるような能力（俯瞰力と独創力）を備え、新たなブレークスルーを生むきっかけとなることを期待して、平成21年度から実施
  - **Aタイプ(研究科開講型)**:研究科で開講される大学院の授業から、「他研究科学生の履修・聴講に配慮した」、「多くの専門分野の共通基盤となりうる」横断型教育にふさわしい科目の提供を受け、広く他研究科大学院生の受講を促す。
    - 自然科学総合科目群, 人文科学総合科目群, 社会科学総合科目群, 環境・生命・医療科目群, **情報活用・計算科学科目群**, マネジメント・キャリア・研究者倫理科目群, 国際性・コミュニケーション科目群, 社会実装・イノベーション科目群
  - **Bタイプ(研究科開講型)**:研究科・研究所等から提供された5コマ単位の授業で、バラエティに富んだ内容の(双方向型の)授業を実施。単位付与されない。

# 高度情報教育基盤ユニット

- 平成21年度～平成25年度 概算要求  
情報学研究科附属情報教育推進センター設置
- 平成26年度～平成30年度 概算要求  
高度情報教育基盤ユニットとして発展
  
- 全学共通教育カリキュラムの設計
  - 学部レベル, 大学院(研究科横断)レベル
- 教材等の作成
  - 学部・大学院計10科目の電子教科書, スライド, 演習ソフト
- 講義室のIT化
  - アクティブ・ラーニング対応設備の導入と活用

# 全学共通教育における情報科学の枠組み(現在)



# 「情報分析・管理論」同演習」 「データ分析入門」「同実践」

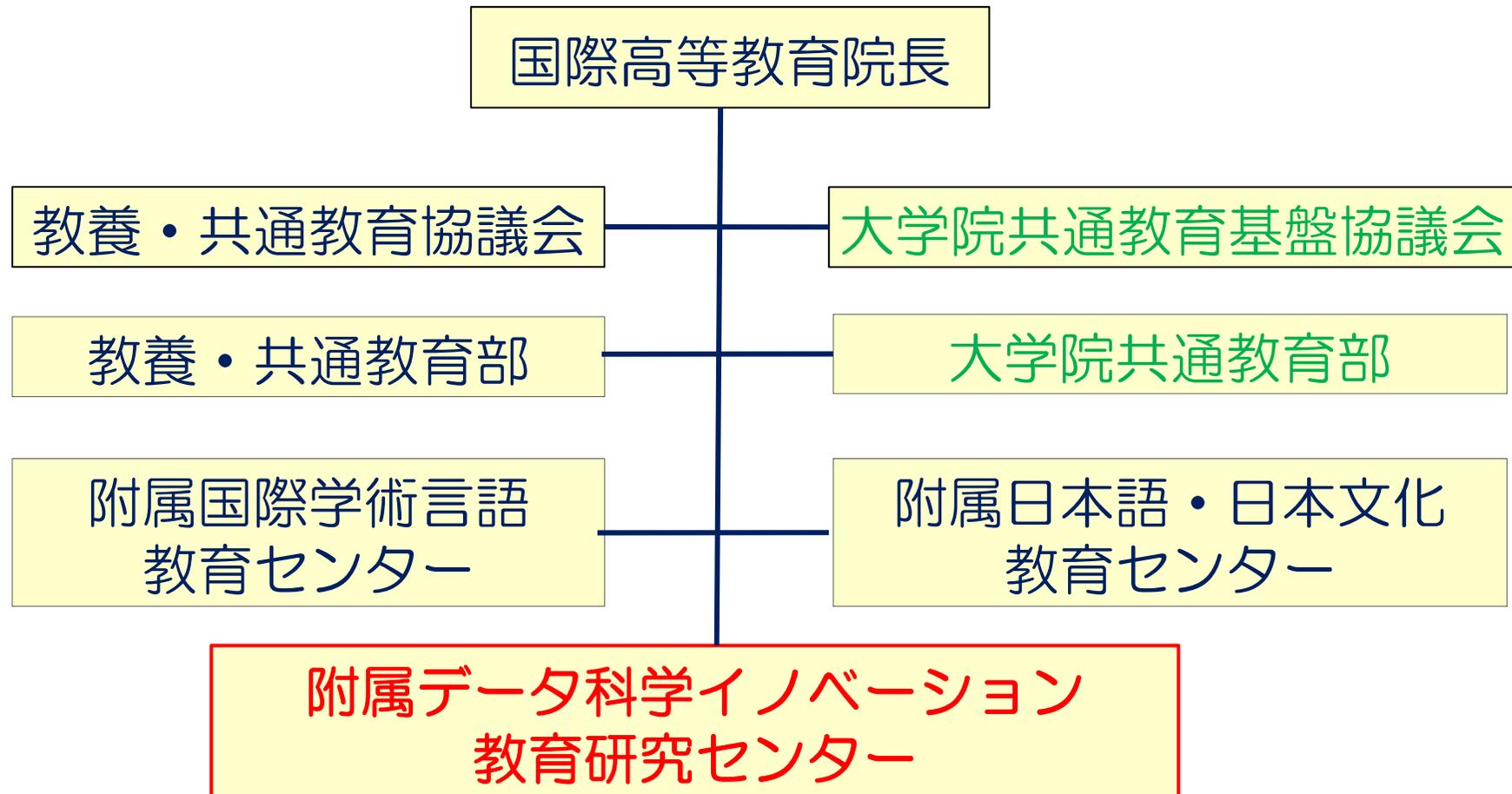
文理横断型 情報活用・計算科学科目群

- 文理の分野を問わず最先端の研究を進めていくため不可欠である**大規模データの分析・管理技術**について、多くの分野で利用されているトピックを精選して講述する。特に、具体的な問題に対してどのようにそれらの技術が適用できるかを学ぶことで技術の基本的な考え方を理解し、各自の分野で応用が可能になることを目指している
- **演習と対で受講することを原則**とする。
- プログラミングを伴う演習を受講する意欲があれば、プログラミング経験は問わない。**各自のノートPC等の持ち込みを前提**としているが、**貸出用のPCも用意**している。内容
- 問題のモデル化と問題の解き方
- データマイニング, データベース, 情報検索
- 情報の可視化と対話技術
- 仮説検定, 回帰分析, 主成分分析

# データ科学イノベーション 教育研究センター

# 国際高等教育院附属センター

国際高等教育院: これまでは学部の共通教育の設計と実施を担当



# 基本コンセプト

## 俯瞰的な視野を持つ人材の育成

直近の課題：データ科学の専門家の養成

中長期の課題：第4次産業革命をトップレベルで支える  
幅広い人材の育成

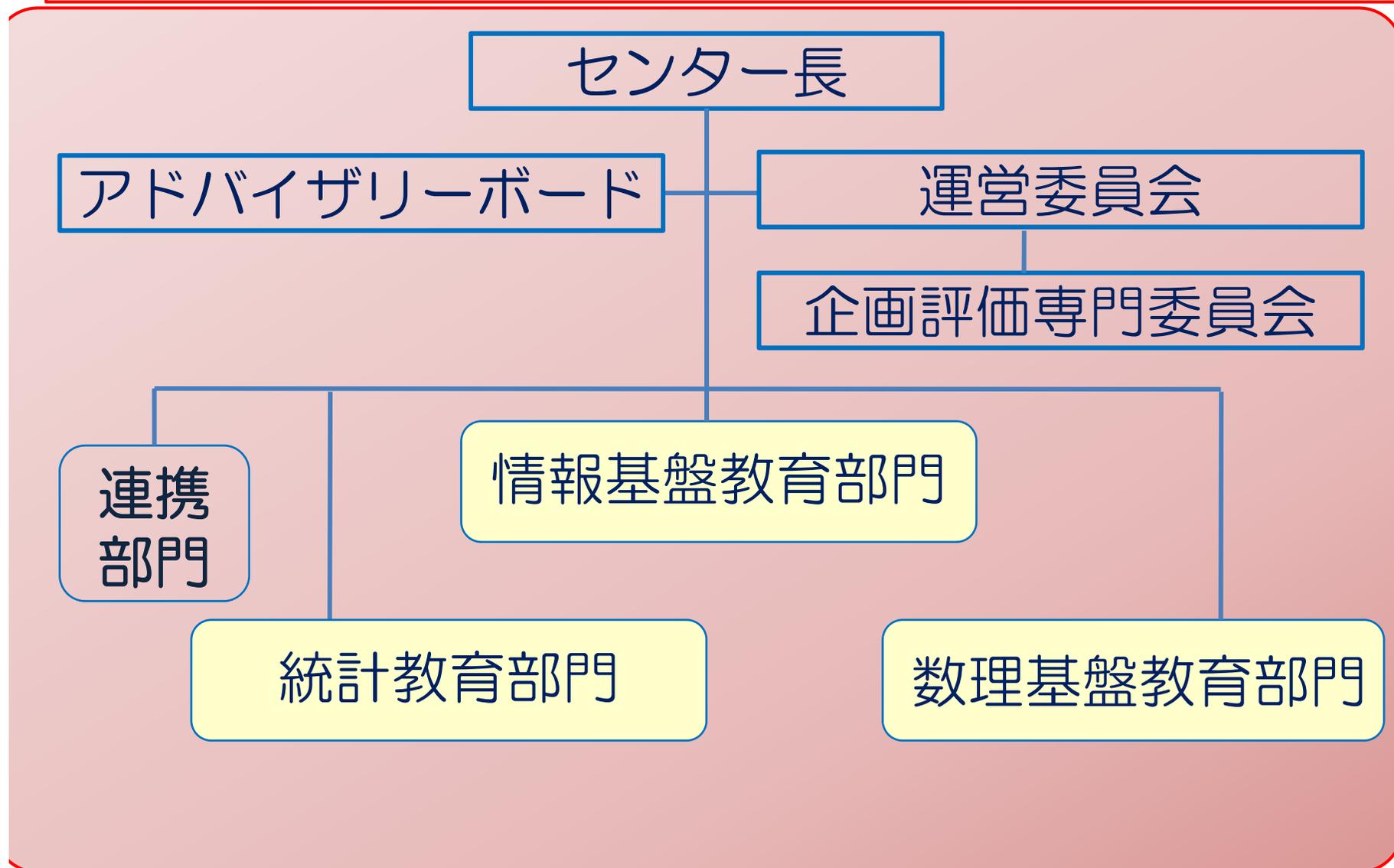


京都大学の伝統の下での新時代の人材育成

しっかりした基礎の上の応用、自学自習、論理  
力の涵養

# データ科学イノベーション教育研究センター

教養・学部専門・大学院の各段階における情報・統計・数理の共通教育を実施するとともに、これらの先端的研究を行う



# 体制

- センター長は国際高等教育院長が兼務
- 企画評価専門委員会
  - 国際高等教育院
    - 情報学研究科
    - 文学研究科
    - 理学研究科
    - 教育学研究科
    - 医学研究科
    - 経済学研究科
    - 工学研究科
    - 人間・環境学研究科
    - 農学研究科
    - 地球環境学堂
- センターが雇用する特定教員
  - 教授1名，講師5名，助教1名
  - 各教員は他研究科を兼務

# 体制

- センター長は国際高等教育院長が兼務
- 企画評価専門委員会  
国際高等教育院

全学に散らばるデータ科学に関する  
教育研究を束ねる

農学研究科

地球環境学堂

- センターが雇用する特定教員  
教授1名，講師5名，助教1名  
– 各教員は他研究科を兼務

# 体制

- センター長は国際高等教育院長が兼務
- 企画評価専門委員会  
国際高等教育院

科目の立案, 設計, 調整を  
国際高等教育院データ科学部会と協力

農学研究科

地球環境学堂

- センターが雇用する特定教員  
教授1名, 講師5名, 助教1名  
– 各教員は他研究科を兼務

# 大学院研究科横断教育プログラムの刷新

- 国際高等教育院の内部に，大学院共通・横断教育基盤と企画評価専門委員会を設置
- 大学院共通科目の分類案
  - 社会適合科目  
例: 研究倫理, 知的財産, アントレナーシップ
  - 情報テクノサイエンス科目  
例: 統計学, データ分析, 情報処理, 情報セキュリティ
  - コミュニケーション科目  
例: 英語プレゼンテーション演習, 英語論文執筆
  - ...

# 科目の刷新:平成30年度から

学位プログラム

論理力の涵養

プログラムのレベル・習得する能力・人材像

高度専門教育  
(博士後期課程)  
使いこなせたら  
創り出せ

専門教育  
(修士課程)  
知るほどに、使うほどに  
楽しくなる

専門基礎教育  
(学部2・3回生)  
使いながら  
深く知る

全学基礎教育  
(学部1・2回生)  
楽しみながら  
基本を知る

高大連携  
未来に向けて  
夢を見る

研究・革新的トレンドの創生  
指導的研究者

Wisdom & Innovation

レベル3:知識から知恵を見出し、  
創造につなげる力  
研究者・次世代育成者(棟梁\*1)  
大規模データ解析を活用する研究者

Knowledge

レベル2:情報から法則性などの  
知識を取り出す力  
データ分析技術者(独り立ち\*1)  
業務例:大規模データの解析・モデル開発

Information

レベル1:データから分布などの  
情報を抽出する力  
データ処理技術者(見習い\*1)  
業務例:データ分析の前処理や基礎分析

Data

レベル0:基本リテラシー  
情報・統計・数理に関する  
一般常識・基本知識の獲得

Dream

レベル-1:動機付け  
様々な例を通して面白さを垣間見る

大学の取り組みに協力

の定義より  
<http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

先端研究におけるデータ  
科学の応用の実践  
企業・連携研究機関  
との共同研究

専門的な理論  
実問題への応用  
科目例:統計的機械学習  
データ分析応用演習  
数理モデリング

汎用的な専門知識の獲得  
事例集による基礎演習  
科目例:データ処理基礎演習  
データベース基礎  
シミュレーション基礎

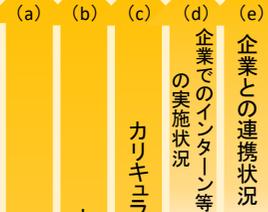
基本知識の習得  
事例集を中心とした入門  
科目例:統計入門  
プログラミング基礎  
線型代数学

高校数学の発展としての  
データサイエンス  
取組例:ELCAS, 特色入試  
オープンキャンパス

統計科学副学位プログラム

学士課程教育プログラム

への貢献



カリキュラム作成の進展度

カリキュラム策定後の e-learning 教材開発の状況

カリキュラムに沿った講義・演習の実施状況

企業でのインターン等の実施状況

企業との連携状況



成果指標との関連性

地域における拠点として  
他大学等への展開・普及

博士後期課程・修士課程の大学院生に対して本来の専門に加えて、データ科学に関する**一定の科目群**からなるプログラムを履修する場合副学位として「統計科学副学位プログラム修了証」を授与

本来の専門分野での学位に加えてデータ科学への深い知見を有することで、企業におけるデータ科学者としての活躍などキャリアパスの多様化を図る。

連動

社会人・博士人材への付加教育

主に科目等履修生として、相当するレベルの科目を履修。  
**一定の科目群**の履修者に対し修了証を授与

博士学位取得者への教育

(レベル3相当)

理系・文系を問わず一定の数理的知識、計算スキル、データからの知識を獲得する訓練を受けて博士学位を修得した者を対象に、職務上必要とされる高度なデータ科学的手法や先端的研究内容を教授

修士課程修了者への教育

(レベル2相当)

個人の既取得スキルに応じて、職務上必要とされるデータ科学的手法を教授

# 科目の刷新:平成30年度から

論理力の涵養

レベルの習得する能力・人材像

高度専門教育  
(博士後期課程)  
使いこなせたら  
創り出せ

専門教育  
(修士課程)  
知るほどに、使うほどに  
楽しくなる

専門基礎教育  
(学部2・3回生)  
使いながら  
深く知る

全学基礎教育  
(学部1・2回生)  
楽しみながら  
基本を知る

高大連携  
未来に向けて  
夢を見る

研究・革新的トレンドの創生  
指導的研究者

Wisdom & Innovation

レベル3:知識から知恵を見出し、  
創造につなげる力  
研究者・次世代育成者(棟梁\*1)  
大規模データ解析を活用する研究者

Knowledge

レベル2:情報から法則性などの  
知識を取り出す力  
データ分析技術者(独り立ち\*1)  
業務例:大規模データの解析・モデル開発

Information

レベル1:データから分布などの  
情報を抽出する力  
データ処理技術者(見習い\*1)  
業務例:データ分析の前処理や基礎分析

Data

レベル0:基本リテラシー  
情報・統計・数理に関する  
一般常識・基本知識の獲得

Dream

レベル-1:動機付け  
様々な例を通して面白さを垣間見る

大学の取り組みに協力

の定義より  
<http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

研究所の協力による  
Spring School, Summer School

データ科学:理論から実用へA, B(集中)  
統計数理研究所との共同  
データ科学:理論から実用へC, D, E(集中)  
情報分析・管理論, 同演習

総

汎用的な専門知識の獲得事例集による基礎演習

データ科学基礎・演習I, II(前期2, 後期2)

総

続統計入門(前期1, 後期1)  
統計入門(前期5, 後期5)  
統計入門(文系向 前期1, 後期1)  
データ科学のための数学入門I, II

高校数学の発展としてのデータサイエンス

ELCAS, 特色入試, オープンキャンパス

統計科学副学位プログラム

学士課程教育プログラム

科目名	概要
統計入門	<p>データ科学が広範な分野に関わることから、この講義では、より多くの学生が統計学の基本的な考え方を理解することを目標にしている。数学的な厳密な理論を理解することよりも、データを扱うエンドユーザーとしてデータの性質に応じた適切な分析方法が選択できるようになることを目指している。そのため、生活の身近な話題についての応用例を数多く紹介し、自然な形で統計的思考法が身に着くようにする。さらに統計解析ソフトを自習形式で取り入れ、実際的なデータ処理を通じた感覚的な理解も深めてゆく。</p>
続・統計入門	<p>「統計入門」で統計的思考の基礎を学習した後、さらに広く知識を身に付けたい者を対象に、「統計入門」では扱いきれなかった話題、例えば分散分析や回帰分析といった話題や、近年重要性が増している因果推論についても紹介する。「統計入門」と同様に、数学的な厳密な理論よりもデータや事例を通じた直感的な理解を目指す。</p>
数理統計	<p>統計学は様々な数値データの処理手法を提供しているが、ややもするとその手法の意味を理解することなく機械的な計算に陥ってしまいがちである。手法の意味するところを理解して初めて得られた解析結果を適切に利用することが可能となる。数理統計学は確率モデルに基づいた数理的手法による統計学であり、この講義では推定や検定などのデータ処理手法を裏付ける数学的な論理の構造を理解することにも力点が置かれる。</p>

科目名	概要
データ分析基礎	<p>データ科学が広範な分野に関わることから、この講義では、より多くの学生が統計学の基本的な考え方を理解することを目標にしている。数学的な厳密な理論を理解することよりも、データを扱うエンドユーザーとしてデータの性質に応じた適切な分析方法が選択できるようになることを目指している。そのため、生活の身近な話題についての応用例を数多く紹介し、自然な形で統計的思考法が身に着くようにする。さらに統計解析ソフトを自習形式で取り入れ、実際的なデータ処理を通じた感覚的な理解も深めてゆく。</p>
データ分析演習Ⅰ,Ⅱ	<p>大規模なデータの分析の演習を通じてデータ分析の基礎の体得を目標としている。主に利用されるデータ分析手法や統計解析ソフトは学術分野によって異なるため、履修者が今後専門とする分野や興味に合わせてできるだけ幅広い選択ができるように、異なる内容の演習科目を揃えている。</p>
数理・データサイエンスのための数学入門Ⅰ,Ⅱ	<p>データ科学は統計学、数理科学、情報学の融合した学問であるため、エンドユーザーとして手法を学ぶだけでも、最低限の数学的知識が必要である。この講義ではデータ科学を学ぶための基礎としての数学の中から、Ⅰでは線形代数、距離空間、統計的推定について、Ⅱでは微分積分の内容を総合的に講述する。数学的な理論の完全な体系ではなく、データ科学への応用を重視する内容とする。さらに、文系学生が高校での数学Ⅲの知識を仮定しないように配慮する。</p>

吉野 章 | 京都大学 大学院 - 地 X 京都大学国際高等教育院 附属 X +

ds.k.kyoto-u.ac.jp/spring\_school\_2018/

京都大学 国際高等教育院 附属データ科学イノベーション教育研究センター

# Data Science Spring School 2018

HOME 開催主旨 講演・講義 プログラム アクセス LANGUAGE: EN

京都大学 国際高等教育院 附属データ科学イノベーション教育研究センターは、論理力の涵養を根幹とした21世紀の基礎教養としての情報学・統計学・数理科学に関する教育及びこれに必要な調査研究等を行うために、国際高等教育院に新たに設置された組織です。本センターが主催するData Science Spring School 2018では、データ科学研究に関わる著名な研究者の講演および演習形式の講義を提供します。

## WHAT'S NEW?

• (2018年3月8日 更新) 懇親会の開始時刻が18時30分に変更になりました

fonts.gstatic.com からデータを転送しています...

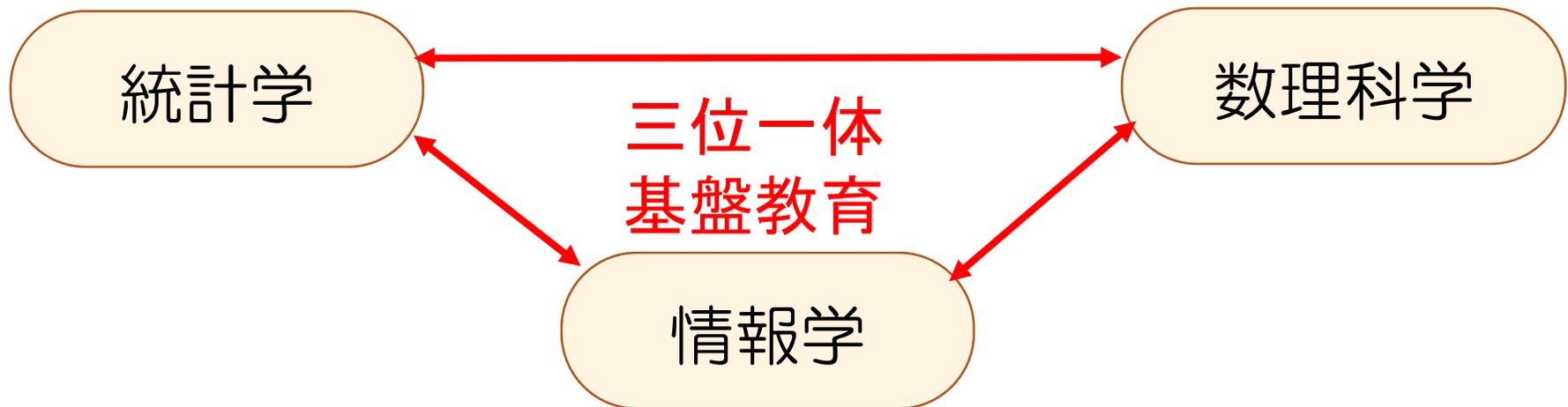
# センターの目指すところ

# 基本コンセプト

## 俯瞰的な視野を持つ人材の育成

直近の課題：データ科学の専門家の養成

中長期の課題：第4次産業革命をトップレベルで支える  
幅広い人材の育成



京都大学の伝統の下での新時代の人材育成

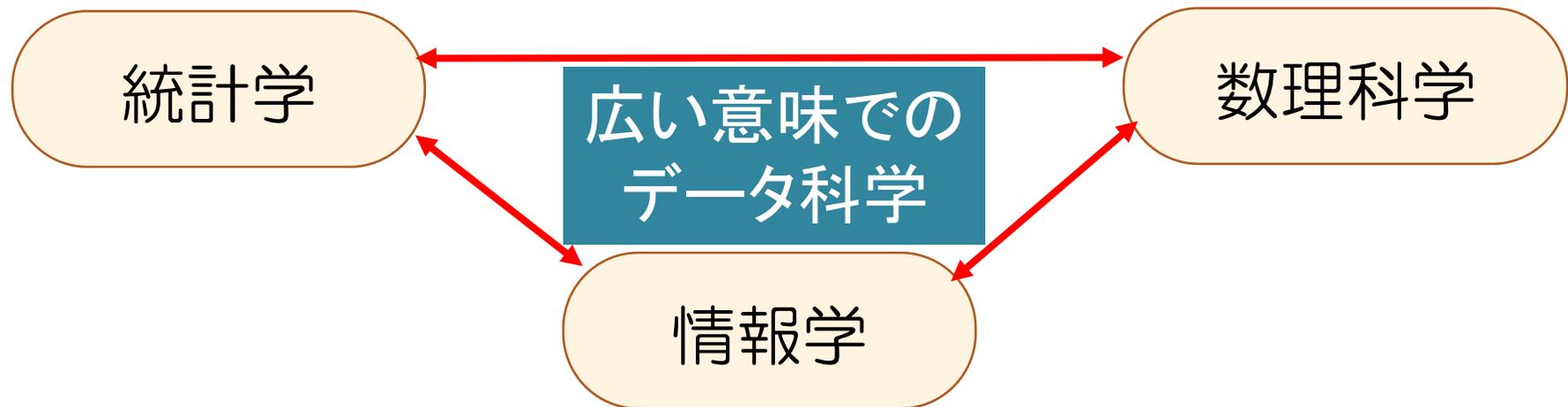
しっかりした基礎の上の応用、自学自習、論理  
力の涵養

# 基本コンセプト

## 俯瞰的な視野を持つ人材の育成

直近の課題：データ科学の専門家の養成

中長期の課題：第4次産業革命をトップレベルで支える  
幅広い人材の育成

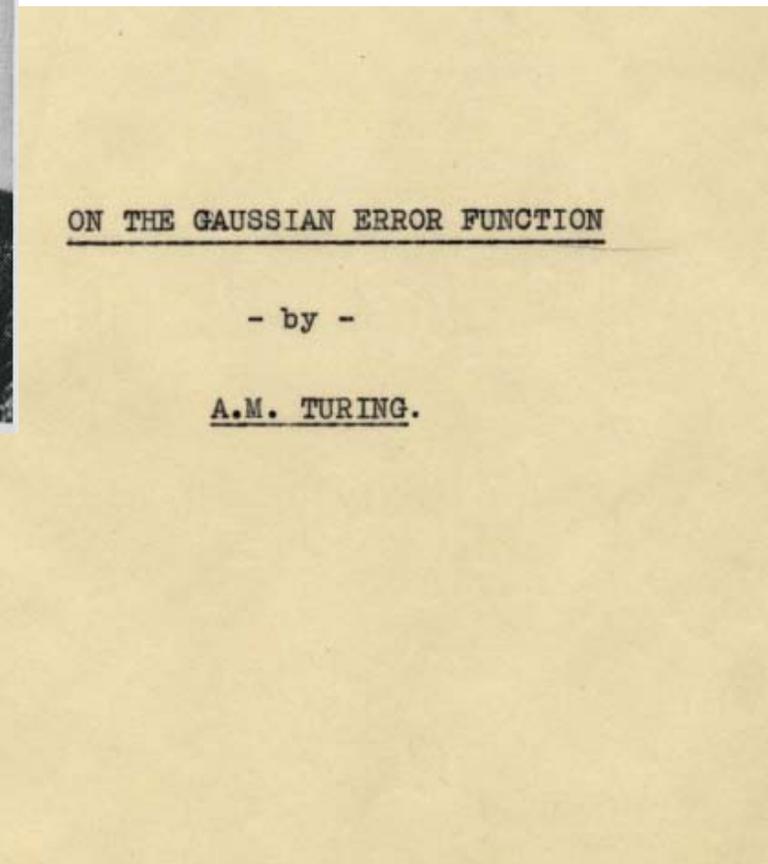
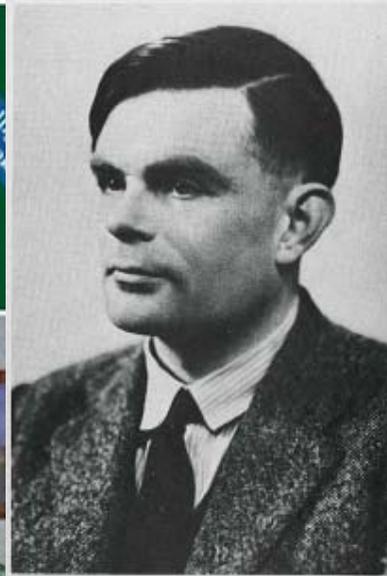
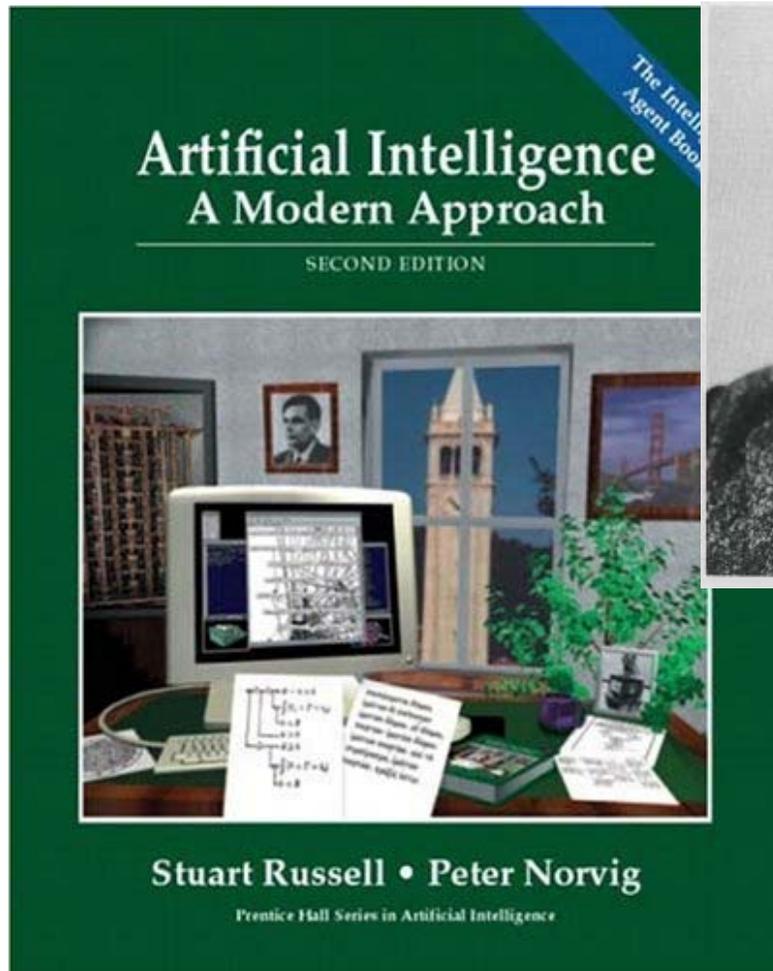


京都大学の伝統の下での新時代の人材育成

しっかりした基礎の上の応用、自学自習、論理  
力の涵養

## Alan Turing (1912-1954)

- 今日の**計算機**のモデルとなる抽象的な機械を考案した.
- **数学者**として, **中心極限定理**の証明を試みた.
- エニグマの解読に**ベイズ推定**を利用した



# 帰納と演繹[広辞苑より]

【帰納】(induction) 推理および思考の手続の一。個々の**具体的**事実から**一般的**な命題ないし法則を導き出すこと。特殊から普遍を導き出すこと。↔**演繹**

【帰納的推理(論)】(inductive inference)

**演繹的に妥当ではないが**、前提が結論に対して一定の支持を与える推論。広義には、演繹的に妥当ではないが正当であるとみなされる推論のすべてをいう。**狭義には**、個々の**具体的**事実から**一般的**な命題ないし法則を導き出す推論をいう。

【演繹】(deduction)推論の一種。前提を認めるならば、結論もまた必然的に認めざるをえないもの。数学における証明はその典型。↔**帰納**

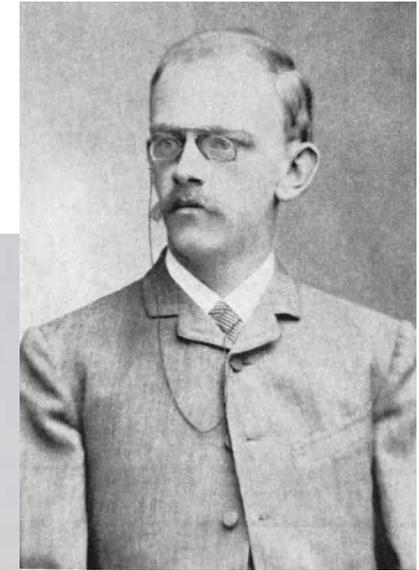
# 計算代数と教師無し学習

- $R$  : 単位的可換環

Prop[Stephan and Venstov 98].  $T$ がイデアル  $I$  の有限証拠集合  $\Leftrightarrow T$ が  $I$  の有限基底.

Cor.  $R$  のイデアル全体が正データから極限同定可能.  
 $\Leftrightarrow R$  が有限生成.  
 $\Leftrightarrow R$  がNoether環.

# Hilbertの原論文



## 16. Über die Theorie der algebraischen Formen<sup>1</sup>.

[Mathem. Annalen Bd. 36, S. 473—534 (1890).]

### I. Die Endlichkeit der Formen in einem beliebigen Formensysteme.

Unter einer algebraischen Form verstehen wir in üblicher Weise eine ganze rationale *homogene* Funktion von gewissen Veränderlichen, und die Koeffizienten der Form denken wir uns als Zahlen eines bestimmten Rationalitätsbereiches. Ist dann durch irgend ein Gesetz ein System von unbegrenzt vielen Formen von beliebigen Ordnungen in den Veränderlichen vorgelegt, so entsteht die Frage, ob es stets möglich ist, aus diesem Formensysteme eine endliche Zahl von Formen derart auszuwählen, daß jede andere Form des Systems durch lineare Kombination jener ausgewählten Formen erhalten werden kann, d. h. ob eine jede Form des Systems sich in die Gestalt

$$F = A_1 F_1 + A_2 F_2 + \dots + A_m F_m$$

bringen läßt, wo  $F_1, F_2, \dots, F_m$  bestimmt ausgewählte Formen des gegebenen Systems und  $A_1, A_2, \dots, A_m$  irgendwelche, dem nämlichen Rationalitätsbereiche angehörige Formen der Veränderlichen sind. Um diese Frage zu entscheiden, beweisen wir zunächst das folgende für unsere weiteren Unter-

# Hilbertの原論文

suchungen grundlegende Theorem:

Theorem I. *Ist irgend eine nicht abbrechende Reihe von Formen der  $n$  Veränderlichen  $x_1, x_2, \dots, x_n$  vorgelegt, etwa  $F_1, F_2, F_3, \dots$ , so gibt es stets eine Zahl  $m$  von der Art, daß eine jede Form jener Reihe sich in die Gestalt*

$$F = A_1 F_1 + A_2 F_2 + \dots + A_m F_m$$

*bringen läßt, wo  $A_1, A_2, \dots, A_m$  geeignete Formen der nämlichen  $n$  Veränderlichen sind.*

Die Ordnungen der einzelnen Formen der vorgelegten Reihe sowie ihre Koeffizienten unterliegen keinerlei Beschränkungen. Denken wir uns die letzteren als Zahlen eines bestimmten Rationalitätsbereiches, so dürfen wir annehmen, daß die Koeffizienten der Formen  $A_1, A_2, \dots, A_m$  dem nämlichen

---

<sup>1</sup> Vgl. die vorläufigen Mitteilungen des Verfassers: „Zur Theorie der algebraischen Gebilde“, Nachr. Ges. Wiss. Göttingen 1888 (erste Note) und 1889 (zweite und dritte Note). Dieser Band Abh. 13 bis 15.

# Hilbert's original paper

- Hilbertは, identification in the limit from texts をもって基底定理の“(構成的な)証明”と主張し, Gordanから“これは数学ではない、神学である”と非難される.



# 比較表

$H$	ある分布族 $\{\theta \in \Theta \mid D(\theta)\}$	多項式の 有限集合全体
$h$	未知パラメータ $\theta$	$\{f_1, f_2, \dots, f_n\}$
$D$	標本	$g_1 f_1 + g_2 f_2 + \dots + g_n f_n$
Loss	$\ \theta - \hat{\theta}\ $	生成イデアルの 差
argmin	推定アルゴリズム $\hat{\theta}$	Buchberger アルゴリズム
推論の正当 性	統計的一致性	格子点集合の 極小元は有限個

# 比較表

$H$	ある分布族	多項式の
$h$	<p style="text-align: center;">三位一体 基盤</p>	
$D$		
Loss		
argmin		
推論の正当性	統計的一致性	格子点集合の極小元は有限個

# 情報学研究科との連携

## 情報学研究科「データ科学」関連科目リスト

多量のデータを集め、コンピュータで解析して、意味のある情報を引き出す「データ科学」が注目されています。データ科学の観点から実際の問題に取り組むにあたっては、データ解析の問題をどのように定式化するか、定式化された問題をどのように解くか、さらにはコンピュータをはじめとする情報基盤をいかに活用して問題を解くかといった課題があり、これらの課題を適切に考察する必要があります。一方で、情報学研究科が掲げる「情報学」の対象としては、「情報の生成、認識、表現、収集、組織化、最適化、変換、伝達、評価、制御」が挙げられており、データ解析にあたっての上記の課題は、まさに情報学の対象に含まれるものだけということが出来ます。

情報学研究科では、データ科学について学ぶことができる科目が多数開講されています。それらの科目を一覧できる資料を作成しました。データ解析の問題の定式化に際して基礎となる統計科目群、対象の適切なモデリングや効率的な解法について議論する数理基盤科目群、データ解析に欠かせない情報処理に関する主題を扱う情報基盤科目群などに区分して、データ科学に関連する内容を含む科目群を提示しています。これらの科目群をバランスよく学修することで、データ解析にもとづく意思決定を支援する以下のようなスキルが身につけられると期待しています。

- (1) 数理統計学や機械学習の知識にもとづいた、データ解析の問題の適切な定式化。
- (2) 問題の定式化の際に必要な、対象となる現象の適切なモデル化。
- (3) モデルにもとづく現象の再現、予測に有用なシミュレーション技術。
- (4) 定式化された問題を効率的に解くための手法やアルゴリズムの設計、選択。
- (5) ビッグデータの解析を効率よく行うためのコンピュータシステムの活用。

データ科学の観点から自らの研究をさらに深めようとお考えの皆さん、さらにはより広くデータ科学に関心をお持ちの皆さんが、履修計画を立て学修を進める際の手引きとして、この資料を活用してください。

## 情報学研究科「データ科学」関連科目リスト

- ・各科目の詳細についてはシラバスを参照してください。
- ・備考欄に「E」と表記されている授業科目は、英語だけでも修得できる科目です。
- ・備考欄に「横断」と表記されている授業科目は、研究科横断型プログラムへ提供している科目です。
- ・○印の科目は隔年講義で奇数年度に開講されます。□印の科目は隔年講義で偶数年度に開講されます。

### 【統計科目群】

#### \* 統計的機械学習

これらの科目を学ぶことで様々な統計的データ分析手法について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
統計的システム論	システム科学専攻	専攻専門科目	下平	前期		
統計的学習理論	知能情報学専攻	専攻専門科目	鹿島	前期		E
適応システム論	システム科学専攻	専攻専門科目	田中(利)	前期		
計算知能システム論	システム科学専攻	専攻専門科目	田中(利)・上田(修)	集中		
非線形物理学特論 I	先端数理科学専攻	専攻専門科目	伊庭(統計数理研究所)	後期		
データ科学:理論から実用へA	情報学研究科	その他	島谷(統計数理研究所)	集中		横断
データ科学:理論から実用へB	情報学研究科	その他	中野(統計数理研究所)	集中		横断

#### \* 統計的データ分析

これらの科目を学ぶことで様々な分野での統計的データ分析について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
情報システム分析論	社会情報学専攻	専攻基礎科目	守屋・吉川(正)・大手・松井[経]	前期	*	E
金融工学	数理工学専攻	専攻専門科目	野崎・瀬古	集中		
論理生命学	システム科学専攻	専攻専門科目	石井・大羽	前期		
物理統計学特論	数理工学専攻	専攻専門科目	梅野	前期		

# 情報学研究科との連携

## 【数理基盤科目群】

### \* 数値計算

これらの科目を学ぶことで様々な数値計算手法について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
計算科学入門	情報学研究科	研究科共通科目	中村(佳)・木村・関戸・藤原(宏)・中尾 他	前期	*	横断
○計算力学特論A	先端数理科学専攻	専攻専門科目	西村	後期		
□計算力学特論B	先端数理科学専攻	専攻専門科目	未定	後期		
応用数理学特論 I	先端数理科学専攻	専攻専門科目	原田(健)	後期		
○応用数理学通論A	先端数理科学専攻	専攻基礎科目	磯	前期		
非線形物理学特論 I	先端数理科学専攻	専攻専門科目	伊庭(統計数理研究所)	後期		
○非線形力学特論A	先端数理科学専攻	専攻専門科目	簡	後期		
システム解析通論	数理工学専攻	専攻基礎科目	中村(佳)・太田	後期		

### \* 最適化

これらの科目を学ぶことで様々な最適化の手法について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
計画数学通論	数理工学専攻	専攻基礎科目	永持・山下・Shurbevski	後期		
最適化数理特論	数理工学専攻	専攻専門科目	山下	前期		
離散数理特論	数理工学専攻	専攻専門科目	永持	後期		
離散アルゴリズム理論	通信情報システム専攻	専攻基礎科目	牧野	前期		

### \* 数理モデル(基盤)

これらの科目を学ぶことで様々な分野の基礎となる数理モデルについて学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
○非線形物理学通論A	先端数理科学専攻	専攻基礎科目	青柳	前期		
□非線形物理学通論B	先端数理科学専攻	専攻基礎科目	未定	前期		
数理物理学通論	数理工学専攻	専攻基礎科目	柴山・五十嵐(顕)	後期		
情報システム特論	システム科学専攻	専攻専門科目	増山	後期		
システム解析通論	数理工学専攻	専攻基礎科目	中村(佳)・太田	後期		
統合動的システム論	システム科学専攻	専攻専門科目	大塚	後期		
応用数理学特論I	先端数理科学専攻	専攻専門科目	原田(健)	後期		
○数理学特論A	先端数理科学専攻	専攻専門科目	田口	後期		
○非線形解析特論A	先端数理科学専攻	専攻専門科目	白石	後期		
物理統計学特論	数理工学専攻	専攻専門科目	梅野	前期		

### \* 数理モデル(応用)

これらの科目を学ぶことで様々な分野での数理モデルの応用について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
応用数理工学特論A	数理工学専攻	専攻専門科目	黒田(正)	集中		
応用数理工学特論B	数理工学専攻	専攻専門科目	山本(彰)・佐藤(達)	集中		
情報分析・管理論	情報学研究科	研究科共通科目	浅野・加藤(誠)	前期 後期	*	横断
金融工学	数理工学専攻	専攻専門科目	野崎・瀬古	集中		

## 【情報基盤科目群】

### \* 情報学基礎

これらの科目を学ぶことでデータ処理の基礎について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
情報学演習 I	情報学研究科	研究科共通科目 (デザイン学科目)	大島	前期集中	*	
情報学演習 II	情報学研究科	研究科共通科目 (デザイン学科目)	大島	後期集中	*	
計算科学入門	情報学研究科	研究科共通科目	中村(佳)・木村・関戸・藤原(宏)・中尾 他	前期	*	横断
計算科学演習A	情報学研究科	研究科共通科目	木村・関戸	前期	*	横断
アルゴリズムと情報学入門	通信情報システム専攻	専攻基礎科目	Le Gall	前期		E
ビッグデータの計算科学	情報学研究科	その他	中村(佳)・木村・関戸・小山田	後期		横断
計算科学演習B	情報学研究科	その他	中島・木村・関戸・深沢	夏期集中	*	横断
スーパーコンピューティング特論	システム科学専攻	専攻専門科目	中島・深沢	後期		

### \* 機械学習・パターン認識

これらの科目を学ぶことでデータの分類・識別について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
計算論的学習理論	知能情報学専攻	専攻専門科目	山本(章)	後期		E
パターン認識特論	知能情報学専攻	専攻専門科目	河原(達)・川嶋・吉井・Liang	前期		E
音声情報処理特論	知能情報学専攻	専攻専門科目	河原(達)・加藤(宏)・吉井・糸山・秋田	後期		E
言語情報処理特論	知能情報学専攻	専攻専門科目	黒橋・河原(大)・森(信)	前期		E
コンピュータビジョン	知能情報学専攻	専攻専門科目	川嶋・延原・Liang	後期		E
生命情報学基礎論	知能情報学専攻	専攻基礎科目	阿久津・細川・前川(真)・林田・田村	後期		

### \* 情報システム

これらの科目を学ぶことでデータの分析・管理と検索について学習できます

科目名	提供専攻等	科目区分	担当教員	開講期	演習的内容を含む科目	備考
情報分析・管理論	情報学研究科	研究科共通科目	浅野・加藤(誠)	前期 後期	*	横断
情報分析・管理論演習	情報学研究科	研究科共通科目	加藤(誠)・浅野	前期 後期	*	横断
情報システム設計論	社会情報学専攻	専攻基礎科目	吉川(正)・田島・Jatowt・山本(岳)・石田・松原・清水	前期	*	E
情報システム論実習	社会情報学専攻	専攻基礎科目	大手・大島・松井[経]・馬・松原・守屋・山本(岳)・林(冬)	前期	*	E
情報システム分析論	社会情報学専攻	専攻基礎科目	守屋・吉川(正)・大手・松井[経]	前期	*	E
分散情報システム	社会情報学専攻	専攻専門科目	吉川(正)・馬	後期		E
情報組織化・検索論	社会情報学専攻	専攻専門科目	吉川(正)・山本(岳)	後期		



# 科目の刷新:平成30年度から

論理力の涵養

レベル3: 研究・革新的トレンドの創生  
指導的研究者  
のレベル・習得する能力・人材像

研究・革新的トレンドの創生  
指導的研究者

の定義より  
<http://www.datascientist.or.jp/news/2014/pdf/1210.pdf>

## Wisdom & Innovation

高度専門教育  
(博士後期課程)  
使いこなせたら  
創り出せ

レベル3: 知識から知恵を見出し、  
創造につなげる力  
研究者・次世代育成者(棟梁\*1)  
大規模データ解析を活用する研究者

研究所の協力による  
Spring School, Summer School

## Knowledge

専門教育  
(修士課程)  
知るほどに、使うほどに  
楽しくなる

レベル2: 情報から法則性などの  
知識を取り出す力  
データ分析技術者(独り立ち\*1)  
業務例: 大規模データの解析・モデル開発

データ科学:理論から実用へA, B(集中)  
統計数理研究所との共同  
データ科学:理論から実用へC, D, E(集中)  
情報分析・管理論, 同演習

## Information

専門基礎教育  
(学部2・3回生)  
使いながら  
深く知る

レベル1: データから法則性などの  
情報を抽出する力  
データ処理技術者(見習い\*1)  
業務例: データ分析の前処理や基礎分析

汎用的な専門知識の獲得事例集による基礎演習

データ科学入門・実践I, II(前期2, 後期2)

## Data

全学基礎教育  
(学部1・2回生)  
楽しみながら  
基本を知る

レベル0: 基本リテラシー  
情報・統計・数理に関する  
一般常識・基本知識の獲得

続統計入門(前期1, 後期1)  
統計入門(前期5, 後期5)  
統計入門(文系向 前期1, 後期1)  
データ科学のための数学入門I, II

## Dream

高大連携  
未来に向けて  
夢を見る

レベル-1: 動機付け  
様々な例を通して面白さを垣間見る  
大学の取り組みに協力

高校数学の発展としてのデータサイエンス

ELCAS, 特色入試, オープンキャンパス

統計科学副学位プログラム

学士課程教育プログラム



# 課題2

## 統計科学副学位プログラム

博士後期課程・修士課程の大学院生に対して本来の専門に加えて、データ科学に関する**一定の科目群**からなるプログラムを履修する場合副学位として「統計科学副学位プログラム修了証」を授与

本来の専門分野での学位に加えてデータ科学への深い知見を有することで、企業におけるデータ科学者としての活躍などキャリアパスの多様化を図る。



## 社会人・博士人材への付加教育

主に科目等履修生として、相当するレベルの科目を履修。**一定の科目群**の履修者に対し修了証を授与

### 博士学位取得者への教育（レベル3相当）

理系・文系を問わず一定の数理的知識，計算スキル，データからの知識を獲得する訓練を受けて博士学位を修得した者を対象に，職務上必要とされる高度なデータ科学的手法や先端的研究内容を教授

### 修士課程修了者への教育（レベル2相当）

個人の既取得スキルに応じて，職務上必要とされるデータ科学的手法を教授

# まとめ

- 京都大学データ科学イノベーション教育研究センター
  - 京都大学におけるこれまでのデータ科学教育  
「統計入門」、高度情報教育基盤ユニット
  - 国際高等教育院附属センター  
大学院研究科横断型教育プログラムの刷新
  - 情報・統計・数理の「三位一体」  
情報学研究科・医学研究科との連携
  - 課題