

統計検定直前対策講座
1級 統計数理
(シグマインベストメントスクール)

茨城大学 工学部 情報工学科
新納浩幸

自己紹介

新納 浩幸 (しんのう ひろゆき)

hiroyuki.shinnou.0828@vc.ibaraki.ac.jp

- ・ 茨城大学工学部情報工学科 教授
- ・ 専門は「**自然言語処理**」

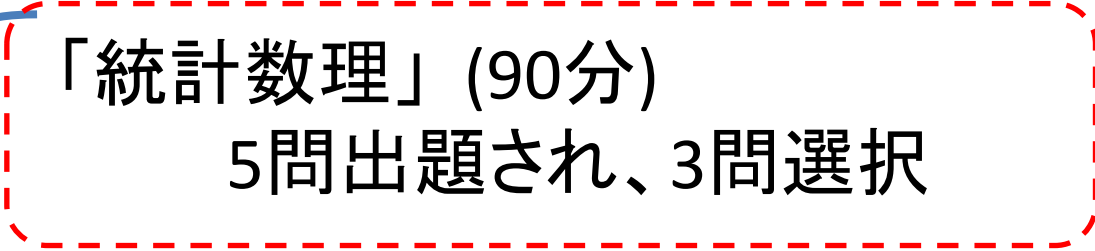
人工知能の中心アプリケーション
手法は機械学習、統計学がベース
(例) 統計的機械翻訳

- ・ 著書「数理統計学の基礎」(森北)、
「Rで学ぶクラスタ解析」(オーム社) など
- ・ 「統計学」の非常勤講師を20年近く継続中


統計検定 1級

本講座の対象

<http://www.toukei-kentei.jp/about/grade1.html>



「統計数理」(90分)
5問出題され、3問選択



「統計応用」(90分)
4分野「人文科学」、「社会科学」、
「理工学」、「医薬生物学」
から1分野選択
各分野から 5問出題され、3問選択

「数理統計」出題範囲

http://www.toukei-kentei.jp/about/pdf/grade1_hani150508.pdf

統計検定1級出題範囲表(統計数値)

大項目	小項目	ねらい	項目(学習しておくべき用語)例
確率と確率変数	事象と確率	確率と確率分布に関する基礎的な事項を理解し、種々の場面に応じた確率計算が正しくできる。	確率の計算、統計的独立、条件付き確率、ベイズの定理、包除原理
	確率分布と母関数		確率関数、確率密度関数、累積分布関数、生存関数、危険率、同時分布、周辺分布、条件付き分布 確率母関数、モーメント母関数(積率母関数)
	分布の特性値	分布の各種特性値の意味を理解すると共に、特性値の値から分布の形状が推測できる。	モーメント、期待値、分散、標準偏差、歪度、尖度、変動係数、パーセント点、中央値、四分位数、範囲、四分位範囲、最頻値、共分散、相関係数、偏相関係数
	変数変換	変数変換後の分布が導出できる。	変数変換、確率変数の線形結合
	極限定理と確率分布の近似	確率分布の極限的な性質を理解すると共に、分布の近似に応用できる。	大数の弱法則、中心極限定理 二項分布の正規近似とポアソン近似、少数法則、連続修正
	種々の確率分布	離散型分布	基本的な離散型分布を理解すると共に、各種の確率計算ができる。
連続型分布		基本的な連続型分布を理解すると共に、各種の確率計算ができる。	一様分布、正規分布(ガウス分布)、指数分布、ガンマ分布、ベータ分布、コーシー分布、対数正規分布、ワイブル分布、ロジスティック分布、多変量正規分布
標本分布		標本分布を理解し、応用に用いることができる。	t分布、カイ二乗分布、F分布
統計的推測(推定)	母集団と標本・統計量	尤度などの統計的推測に重要な役割を果たす概念を理解すると共に、パラメータの推定法の原理を知り、推定量の良さを数学的に立証できる。また、区間推定とは何かを理解し、信頼区間の性質を正しく述べることができる。	十分統計量、ネイマンの分解定理、順序統計量
	尤度と最尤推定		尤度関数、対数尤度関数、有効スコア関数、最尤推定
	各種推定法		モーメント法、最小二乗法、線形推定(BLUE)、その他の手法
	点推定量の性質		不偏性、一致性、十分性、有効性、推定量の相対効率
	モデル評価基準		カルバック・ライブラー情報量、情報量規準AIC、クロスバリデーション
	漸近的性質など		グラメル・ラオの不等式、フィッシャー情報量(1次元)、最尤推定量の漸近正規性、デルタ法
統計的推測(検定)	区間推定		信頼係数、信頼区間の構成、被覆確率
	検定の基礎	統計的検定の原理を理解し、種々の最適化で検定が構成でき、その性質を数学的に立証できる。特に正規分布に関する検定を正しく理解すると共に、そのほかの代表的な分布に関する検定ができる。	仮説、検定統計量、P値、有意水準、棄却域、第一種の過誤、第二種の過誤、検出力(検定力)、検出力曲線
	検定法の導出		ネイマン・ピアソンの基本定理、尤度比検定、ワルド型検定、スコア型検定
	正規分布に関する検定		平均値と分散に関する検定、複数の平均に関する検定
種々の検定法		二項分布・ポアソン分布など基本的な分布に関する検定、適合度の検定、ノンパラメトリック検定	

データ解析法の考え方と各種分析手法	分散分析	データ解析法の中でも重要な位置を占める分散分析と回帰分析について正しく理解し、応用することができる。	一元配置分散分析、二元配置分散分析、交互作用、共分散分析、多重比較
	回帰分析		線形単回帰、線形重回帰、最小二乗推定、回帰の分散分析、重相関係数、決定係数、残差、変数変換、平均への回帰(回帰効果)
	分割表の解析	実際問題で遭遇する分割表の解析ならびにノンパラメトリックな方法について理解し、実践することができる。	カイ二乗検定、フィッシャー検定、マクネマー検定、イエーツの補正
	ノンパラメトリック法		符号検定、ウィルコクソン順位和検定(マン・ホイットニーU検定)、ウィルコクソン符号付き順位和検定、順位相関係数
	不完全データ	不完全データの分析について理解すると共に、コンピュータを用いたシミュレーションができる。モデル構築に役立てる。	欠測(欠損)、打ち切り、トランケーション
	シミュレーション		乱数、モンテカルロシミュレーション、MCMC、ブートストラップ
ベイズ法		事前分布、事後分布、階層ベイズモデル、ギブスサンプリング	

かなり広範囲、深い内容、完全に学部授業のレベルを超えている

過去問から感じること(1)

非常に難問だけど、本質的な力や理解を試そうとしている、枝葉末節の知識は問われないように感じる・・・

(例)

2012年 問1、2013年 問1、2014年 問1



各年度の問1はサービス問題、密度関数とは何かを理解していれば容易

2013年 問4



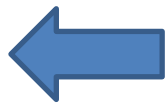
仮説検定の考え方さえわかっていたら容易

などなど

過去問から感じること(2)

数理統計として重要な項目を出題しているように感じる...

データ解析法は出題されていない、
検定は尤度比検定、適合度検定、平均の検定
推定は最尤法
分布は一様分布、二項分布、正規分布、
カイ2乗分布、ガンマ分布、ベータ分布



大事な部分だけでも、十分理解できていれば
なんとかなる感じ

大まかな問題構成

問題1 サービス問題(単なる数学の問題)

問題2 変数変換、2次元確率変数、分布

問題3 最尤法を中心とした推定

問題4 なんだかわからない問題...

問題5 検定

問題1, 2 を選択し、もう1問はできそうなもの

本日のスケジュール

1コマ目：10:00 から 11:30 (90分)

10分休憩

2コマ目：11:40 から 13:00 (80分)

1時間昼休み

3コマ目：14:00 から 16:00 (120分)

10分休憩

4コマ目：16:10 から 18:00 (110分)

4コマある

重要テーマ順に解説

(1) 変数変換

(2) モーメント母関数

(3) 最尤法

(4) 適合度検定
尤度比検定



主に1コマ目と2コマ目で
解説



主に3コマ目で解説

4コマ目は残りの
問題のポイントを解説

スケジュール詳細(1)

1コマ目(10:00 から 11:30) + 2コマ目(11:40 から 13:00)

スライドにより、「変数変換」「モーメント母関数」解説

過去問 2012-問2、2014-問2 を解説

残りの時間で

過去問 2012-問1、2013-問1、2014-問1 を解説

スケジュール詳細(2)

3コマ目(14:00 から 16:00)

スライドにより、「最尤法」「検定」解説

過去問 2013-問5、2014-問5 を解説

残りの時間で

過去問 2013-問4 (検定の考え方の確認)

過去問 2012-問4、2014-問3 を解説

(検定の問題だがポイントのみ)

スケジュール詳細(3)

4コマ目(16:10 から 18:00)

スライドにより、「重要公式」「雑多なポイント」解説

残された過去問を重要度順に解説

2013-問2 (条件付き確率)

2012-問3 (最尤法)

2013-問3 (推定量の性質)

} やや重要

2012-問5 (条件付き確率の平均、分散)

← 公式知っていれば容易

2014-問4 (複数群の平均に関するF検定、難問)

← 多分、試験時間内で解くのは無理