

データサイエンス

責任者・コーディネーター		情報科学科医学統計情報学分野 高橋 史朗 教授		
担当講座・学科(分野)		情報科学科 医学統計情報学分野、医歯薬総合研究所 医療開発研究部門、医歯薬総合研究所超高磁場 MRI 診断・病態研究部門		
担当教員		高橋 史朗 教授、西塚 哲 特任教授、山下 典生 准教授、小野 保 講師		
対象学年	1	区分・時間数 (1 コマ 2 時間計算)	講義	7 コマ 14 時間
期間	後期		演習	7 コマ 14 時間
			実習	0 コマ 0 時間

・学修方針（講義概要等）

近年、情報通信技術の発展によりビッグデータの扱いが容易となり、医療の現場においても ICT や AI への関心が急速に高まっている。特にゲノム医療、画像診断支援、診療・治療支援への活用は、国の重点施策に選定されるほどである。これらを支える主要な学問の一つがデータ科学・統計学であり、その重要性が以前に増して強く認識されている。医療関係者のみならず現代社会を生きるすべての人に、データ駆動型の思考法が強く求められている。本講義では、データ科学のリテラシーレベルであるデータの可視化と要約の方法、統計的推定法と検定法に関する基本的な知識を身につけることを目指す。さらに岩手県医療ビッグデータや公開データなどを用いた演習をとおして、データ駆動型思考法を涵養するとともに、地域医療の現状の把握と問題点の抽出・解決方法の考察・発信をとおして地域医療に関する見識を深めることも目指す。

・教育成果（アウトカム）

データ科学の基本的な知識および EZR など統計解析ソフトウェアの基本操作方法を習得することにより、第 2 学年以降の専門科目の実習・実験や臨床データを適切に扱え、読み、説明することができる。（ディプロマ・ポリシー：1, 2, 4, 5, 6, 8）

・到達目標（SBO）

1. 医療とデータ科学・AI の関わりについて例をあげて説明できる。
2. 必要な情報やデータをインターネットで収集できる。
3. 記述統計学について説明できる。
4. 統計解析ソフトウェアを用いてデータを視覚化・要約することができる。
5. 代表的な理論分布を説明でき、確率を計算することができる。
6. 点推定論、区間推定論、仮説検定論について説明することができる。
7. 統計解析ソフトウェアを用いて平均と割合の点推定値と区間推定値を求めることができる。
8. 統計解析ソフトウェアを用いて一標本 t 検定、二標本 t 検定、Welch の t 検定、多重比較を行うことができる。
9. 収集したデータを統計学的に解釈することができる。
10. グループワークで協働することができる。

・ 講義日程
【講義】

(矢) 東 1-A 講義室

クラス 月日	曜日	時限	講座(学科)	担当教員	講義内容/到達目標
C1 10/30	木	3	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】 ガイダンス, 医学研究と統計学: 講義・意見交換 1. データ科学を概説できる. 2. EBM における臨床研究とデータ科学の役割を概説できる. 【事前学修: 30 分】 教科書 19 章やインターネットを利用し, EBM・診療ガイドラインにおけるデータ科学の役割について調べる.
C2 10/28	火				
C1 10/30	木	4	医学統計情報学分野 医学統計情報学分野 医療開発研究部門 超高磁場 MRI 診断・病態解析部門	高橋 史朗 教授 小野 保 講師 西塚 哲 特任教授 山下 典生 准教授	【授業内容】 医療とビッグデータ: 講義・意見交換 1. 医療におけるビッグデータや AI の活用状況や最新研究を体感する. 【事後学修: 30 分】 EBM・診療ガイドラインにおけるデータ科学の役割についてレポートにまとめる.
C2 10/28	火				
C1 11/6	木	3	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】 記述統計学(1): 講義, 意見交換 1. データの種類に応じてデータの代表値・視覚化方法を説明できる. 2. 解析データセットの構造, 変数と種類を説明できる. 【事前学修: 30 分】 資料および教科書 p11-30 を通読し, まとめ, 不明点を明らかにする.
C2 11/4	火				
C1 11/6	木	4	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】 記述統計学(2): 演習 1. 統計解析ソフトウェア EZR を用いて配布データを視覚化し, 基本統計量を算出し, 分布の形状をレポートにまとめる. 【事後学修: 60 分】 講義内容を復習し, 課題をレポートにまとめる.
C2 11/4	火				

C1 11/13	木	3	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】正規分布と標本平均の分布，信頼区間：講義，意見交換 1. (標準)正規分布を説明できる。 2. Q-Q プロットを読むことができる。 3. 連続確率変数の標準化，確率計算を行える。 4. 推測統計学，調査法等を説明できる。 5. 標本平均の分布の形状，標準偏差と標準誤差の違いを説明できる。 6. t 分布について概説でき，t 分布表を読むことができる。 7. 標本平均の信頼区間を構成でき，結果を正しく解釈できる。 【事前学修：60 分】資料および教科書 p31-38, 44 - 52, 56-70 を通読し，まとめ，不明点を明らかにする。
C2 11/11	火				
C1 11/13	木	4	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】記述統計学(3)：演習 1. R を用いたプログラミングを行う。 2. 疑似乱数の平均値の分布を視覚化し，基本統計量を算出し，データ数等に応じた分布の形状の変化をレポートにまとめる。 3. 平均値の信頼区間を構成できる。 【事後学修：45 分】講義内容を復習し，課題をレポートにまとめる。
C2 11/11	火				
C1 11/20	木	3	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】二項分布，標本割合の分布，信頼区間：講義，意見交換 1. 比，割合，率の違いを説明できる。 2. 二項分布とその正規近似を説明でき，確率計算を行える 3. 標本割合の分布の形状を説明できる。 4. 標本割合の正確・近似の信頼区間を構成 【事前学修：30 分】資料および教科書 p39-44, 71-72 を通読し，まとめ，不明点を明らかにする。
C2 11/18	火				
C1 11/20	木	4	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】平均と割合の信頼区間：演習 1. R プログラムを行う。 2. 疑似乱数の割合の分布を視覚し，基本統計量を算出し，データ数等に応じた分布の形状の変化をレポートにまとめる。 3. 割合の信頼区間を構成できる。 【事後学修：60 分】講義内容を復習し，課題をレポートにまとめる。
C2 11/18	火				

C1 11/27	木	3	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】仮説検定(1)概論・一標本問題：講義，意見交換 1. 帰無仮説と対立仮説を説明できる。 2. 有意水準，棄却限界値，検定統計量を説明できる。 3. p 値を説明でき，正しく解釈できる。 4. 第 1 種の過誤，第 2 種の過誤，検出力を説明できる。 5. 一標本の母平均に関する検定を行うことができ，結果を正しく解釈できる。 6. 検定と信頼区間の関係を説明できる。 【事前学修：45 分】資料および教科書 p73-87 を通読し，不明点を明らかにする。
C2 11/25	火				
C1 11/27	木	4	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】一標本問題：演習，総合演習(1)：グループワーク 1. EZR を用いて一標本 t 検定を行う。 2. EZR を用いてグルコース負荷試験データの記述統計および推測統計を行う。 3. 岩手県医療ビッグデータ等を用いて，PI(E)CO で構造化した疑問に対するデータ解析を行う。 【事前学修：15 分】身の回りの健康課題に関する疑問を PI(E)CO で構造化する。 【事後学修：45 分】講義内容を復習し，課題をレポートにまとめる。
C2 11/25	火				
C1 12/4	木	3	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】仮説検定(2)二標本問題：講義，演習，意見交換 1. 対応のあるデータとないデータについて区別できる 2. 対応のある二標本の母平均の差の検定を行え，信頼区間を構成できる。 3. 対応のない二標本の母平均の差の検定を行え，信頼区間を構成できる。 4. EZR で解析することができ，結果を正しく解釈できる。 【事前学修：30 分】資料および教科書 p88-102 を通読し，まとめ，不明点を明らかにする。 【事後学修：60 分】講義内容を復習し，課題をレポートにまとめる。
C2 12/2	火				

C1 12/4	木	4	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】総合演習(2)：演習・グループワーク 岩手県医療ビッグデータなどの解析を行う。 【事前・事後学修：240分】：データのクリーニング，データ解析を実施し，発表資料を作成する。
C2 12/2	火				
C1 12/17	水	3	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】仮説検定(3)多群の比較：講義，演習，意見交換 1. 多重性の問題を説明できる。 2. 分散分析法を説明できる。 3. 多重比較法を説明でき，EZRで解析することができる。 【事前学修：30分】資料および教科書 p103-114 を通読し，まとめ，不明点を明らかにする。 【事後学修：30分】講義内容を復習し，課題をレポートにまとめる。
C2 12/15	月				
C1 12/17	水	4	医学統計情報学分野	高橋 史朗 教授 小野 保 講師	【講義内容】総合演習(3)：演習・グループワーク，プレゼンテーション 岩手県医療ビッグデータなどの解析結果をまとめて発表する。 【事前学修：150分】発表資料を完成させ，発表の準備を行う。 【事後学修：30分】グループワークから学んだこと，健康課題に対するデータの活用についてレポートを作成する。
C2 12/15	月				

・教科書・参考書等

教：教科書 参：参考書 推：推薦図書

	書籍名	著者名	発行所	発行年
教	生物統計学 標準教科書 第2版	寺尾 哲、森川敏彦	ムイスリ出版	2023
参	EZR でやさしく学ぶ統計学 改訂3版	神田善伸	中外医学社	2020
参	宇宙怪人しまりす 医療統計学を学ぶ	佐藤俊哉	岩波書店	2005
参	宇宙怪人しまりす 医療統計学を学ぶ 検定の巻	佐藤俊哉	岩波書店	2012
参	クリニカルトライアル よりよい臨床試験を志す人たちへ	ポコック	篠原出版	1989

・成績評価方法

【総括的評価】定期試験（MCQ35%，計算・記述 35%），提出課題（25%），グループワークにおける学生間相互評価（5%）に基づき総合的に評価する。

【形成的評価】授業態度，出欠，遅刻などにより態度評価を行う。講義ごとの課題で理解度を確認し，全員を対象に次講義開始時にフィードバックを行う。

・特記事項・その他

1. 演習では統計解析ソフトウェア（R EZR パッケージ）を各自の PC にインストールすること。インストールに関しては，初回講義時に指示する。なお、Mac 使用者かつインストールに失敗した者は事前の申し出により授業時間に限り大学所有の Windows PC を借用することができる（講義終了時に必ず返却。学外持ち出し不可）。
2. 講義資料および関連情報の提示，連絡は原則 WebClass で行う。
3. 演習では課題を課す。講義時間内に終わらなかった課題は事後学習とする。
4. 作成したりポートは期限までに WebClass にアップロードすること。
5. 定期試験については、試験後に解答および配点を WebClass に掲載する。持ち帰った問題用紙で自己採点するとともに、復習すること。希望者には、成績確定後に結果を開示する。

当該科目に関連する実務経験の有無 有

医師の資格を有する教員および脳画像研究等の専門家が、医学におけるデータサイエンスの実践例を説明する。

・授業に使用する機器・器具と使用目的

使用区分	機器・器具の名称	台数	使用目的
講義	ノート PC(MS Windows)	1	資料提示、プレゼン、統計問題計算、 演示
講義	教室付属 AV システム一式	1	資料提示、プレゼン
演習	ノートパソコン（各自）	150	