

データサイエンス

責任者・コーディネーター	情報科学科医用工学分野 高橋 史朗 教授		
担当講座・学科(分野)	情報科学科 医用工学分野、数学分野、医歯薬総合研究所 生体情報解析部門、超高磁場 MRI 診断・病態研究部門		
担当教員	高橋史朗教授、清水厚志教授、江尻正一教授、山下典生准教授、小野保講師		
対象学年	1	区分・時間数	講義 21 時間
期間	後期		

・学習方針（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を生きる人々に対してデータ駆動型の思考法が社会的に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。さらに受講生が興味をもつ公開データを収集・解析する実習をとおして、データ駆動型思考法や科学的探究心を涵養し、自ら問題を発見し、解決し、発信する力を養うことを目指す。

・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・実験や実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：4）

・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. EZR や Excel 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. EZR を用いて平均および割合の点推定値および区間推定値を求めることができる。
8. 収集したデータを統計学的に解釈することができる。

・ 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 10/27	木	3	医用工学分野 数学分野	高橋 史朗 教授 江尻 正一 教授	#01 ガイダンス. 医学研究と統計学 1. 統計学について概説できる。 2. EBM における臨床研究と統計学の役割について概説できる。 事前学習：教科書 19 章およびインターネットを利用して EBM・診療ガイドラインにおける臨床研究・統計学の役割について調べる。 事後学習：EBM・診療ガイドラインにおける臨床研究・統計学の役割についてレポートにまとめて Webclass で提出する。
C2 10/31	月				
C1 10/27	木	4	医用工学分野 生体情報解析部門 超高磁場 MRI 診断・ 病態解析部門	高橋 史朗 教授 清水 厚志 教授 山下 典生 准教授	#02 医療とビッグデータ 1. 医療におけるビッグデータ（ゲノムや画像）の活用状況や最新研究を体感する。 事後学習：講義を受けての自分の考えを 400～800 字でレポートにまとめて Webclass で提出する。
C2 10/31	月				
C1 11/10	木	3	医用工学分野 数学分野	高橋 史朗 教授 江尻 正一 教授	#03 記述統計学(1) 1. データの種類に応じてデータの代表値および視覚化方法を説明できる。 2. 解析用データセットの構造、変数とそのタイプを説明できる。 事前学習：WebClass にアップロードされた資料および教科書 p11-31 を通読する。各自の興味のある領域の公開されたデータセットを収集する。 事後学習；講義内容を復習する。
C2 11/7	月				
C1 11/10	木	4	医用工学分野	高橋 史朗 教授 小野 保 講師	#04 実習 1：記述統計学(2) 1.統計解析ソフトウェア EZR を用いて Fisher のアヤマメデータを視覚化し、基本統計量を算出することができる。 2. 各自が収集したデータに対して視覚化し基本統計量を算出してレポートにまとめる。 事後学習：結果をレポートにまとめて Weclass で提出する。
C2 11/7	月				

C1 11/17	木	3	医用工学分野 数学分野	高橋 史朗 教授 江尻 正一 教授	#05 正規分布 1. (標準)正規分布について説明できる. 2. Q-Q プロットを読める. 3. 確率変数の標準化を行い, 標準正規分布表から確率を求めることができる. 事前学習: WebClass にアップロードされた資料および教科書 p48-57 を通読し, わからないところを調べる. 事後学習: 講義内容を復習する.
C2 11/14	月				
C1 11/17	木	4	医用工学分野	高橋 史朗 教授 小野 保 講師	#06 実習 2: 記述統計学(3) 1. R を用いたプログラミング(乱数生成, 四則演算, 基本統計量の算出)を行う. 2. R/EZR を用いてヒストグラム, 箱ひげ図, Q-Q プロット作成し, 基本統計量を算出し, それらから分布の形状を読み取ることができる. 事後学習: アヤメデータ内の変数を 1 つ選び, 同様の記述統計を行い, リポートにまとめて Webclass で提出する.
C2 11/14	月				
C1 11/24	木	3	医用工学分野 数学分野	高橋 史朗 教授 江尻 正一 教授	#07 標本平均の分布と信頼区間 1. 全例調査と標本調査について説明できる. 2. 推測統計学 (点推定論と区間推定論) について説明できる. 3. 正規母集団からの標本平均の分布の形状と標準誤差を説明できる. 4. t 分布について説明でき, t 分布表を読むことができる. 4. 信頼区間を構成でき, 正しく解釈することができる. 事前学習: WebClass にアップロードされた資料および教科書 59-73 を通読し, わからないところを調べる. 事後学習: 講義内容を復習する.
C2 11/21	月				
C1 11/24	木	4	医用工学分野	高橋 史朗 教授 小野 保 講師	#08 実習 3: 標本平均の信頼区間 1. R プログラミン (標本平均の分布と標準誤差) を行う. 2. EZR を用いて 1 標本の平均の信頼区間を構成できる. 事後学習: 乱数を用いたシミュレーション結果と理論値を比較しリポートにまとめる. また, 各自が収集したデータの 1 変数について信頼区間をリポートにまとめる. いずれも Webclass で提出する.
C2 11/21	月				

C1 12/1	木	3	医用工学分野 数学分野	高橋 史朗 教授 江尻 正一 教授	#09 二項分布, 標本割合の信頼区間 1. 二項分布について説明できる. 2. 二項分布の正規近似について説明できる. 3. 中心極限定理に基づき割合の信頼区間を構成できる. 事前学習: WebClass にアップロードされた資料および教科書 p33-48, 74-75, 143-146 を通読し, わからないところを調べる. 事後学習: 講義内容を復習する.
C2 11/28	月				
C1 12/1	木	4	医用工学分野	高橋 史朗 教授 小野 保 講師	#10 実習 4: 標本割合の信頼区間 1. R プログラミング (中心極限定理) を行う. 2. EZR を用いて正規近似による 95% 信頼区間および Clopper-Pearson の 95% 信頼区間を構成できる. 事後学習: 乱数を用いたシミュレーション結果と理論値を比較しレポートにまとめる. がん臨床研究を例に奏効割合とその 95% 信頼区間を算出した結果をレポートにまとめる. いずれも Webclass で提出する.
C2 11/28	月				
C1 12/8	木	3	医用工学分野 数学分野	高橋 史朗 教授 江尻 正一 教授	#11 仮説検定(1): 概論・一標本問題 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る. 6. 検定と信頼区間の関係を説明できる. 事前学習: WebClass にアップロードされた資料および教科書 p77-87 を通読し, わからないところを調べる. 事後学習: 講義内容を復習する.
C2 12/5	月				
C1 12/8	木	4	医用工学分野	高橋 史朗 教授 小野 保 講師	#12 実習 5 1. EZR/Excel 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. 事後学習: グルコース負荷試験データの解析結果をレポートにまとめて WebClass で提出する.
C2 12/5	月				

C1 12/15	木	3	医用工学分野 数学分野	高橋 史朗 教授 江尻 正一 教授	#13 仮説検定(2)：二標本問題 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。 事前学習：WebClass にアップロードされた資料および教科書 p93-107 を通読し、わからないところを調べる。 事後学習：講義内容を復習する。
C2 12/12	月				
C1 12/15	木	4	医用工学分野	高橋 史朗 教授 小野 保 講師	#14 実習 6 1. EZR/Excel を用いて対応のある t 検定および二標本 t 検定を行う。 2. EZR を用いて薬剤に暴露されたラットから測定された逸脱酵素のデータの記述統計および推測統計を行う。 事後学習：逸脱酵素データの解析結果をレポートにまとめて WebClass で提出する。
C2 12/12	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	寺尾 哲、森川敏彦	ムイスリ出版	2016
参	EZR でやさしく学ぶ統計学	神田善伸	中外医学者	2015
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

学修達成度を、提出課題（40%）と定期試験の成績（60%）に基づき総合的に評価する。

・特記事項・その他

1. 実習では統計ソフトウェア R, EZR パッケージを利用するため, 各自の PC に EZR をインストールすること. インストールに関しては, 初回講義で指示する.
2. 取り組んだ課題は, 指示された期日までに WebClass へアップロードすること. 必要に応じて課題に対するフィードバックを Webclass メールや次回講義のはじめに行う.
3. 事前・事後学習には, 合わせて最低 1 時間 45 分を要する.

・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	関数電卓	1	統計問題計算、演示
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