

2025年度 神戸大学 MBA(専門職大学院)

統計解析応用研究 シラバス

担当教員

分寺杏介(bunji@bear.kobe-u.ac.jp)

1. 授業のテーマと目的

現代社会において、データに基づいた意思決定は不可欠です。感度の高い方はすでに気づいているかもしれませんが、あらゆるビジネスの現場では、データドリブンな(データに基づいた)意思決定がこれまで以上に重要になっています。実際に近年では、企業や研究機関、公的機関、さらにはスポーツ、芸能に至るまで様々な分野において、大量に蓄積されたデータを分析し、次の課題を設定して戦略の策定からアクションの実施のプロセスに活かすケースが増えています。マーケティングを例に挙げてみると、オンラインショップでは「顧客はどのような購買パターンを持つのか」「カートに残っている商品をより購入してもらうためにはどうしたら良いか」、小売業では「どの商品をどの棚に置くべきか」「いつ値下げをするべきか」、広告業では「どのチャンネルに広告を出稿するべきか」などの様々な問いに対して、データに基づいた解を得て、さらに将来の経営活動や研究へ活用する動きが強まっているのです。ビッグデータや AI といった言葉が飛び交う今、データを読み解き、活用する力は、経営学分野に限らずあらゆる場面で皆さんの武器になるのです。インターネットを介した様々なサービスの中では、日夜、様々な情報(例えば消費者の好みや関心)が大量にデータとして蓄積されています。そうした情報を読み解き将来への備えを適切に実施できる人・企業ほど、様々な場面で優位に立てるようになってきている、という構造は、多くの方に理解頂けるでしょう。これは何も「(プロの)データサイエンティスト」に限った話ではありません。将来どのような職業に就いたとしても、データに基づき合理的な判断・説明・予測する力は必ずあなたの武器となるはずです。統計解析の手法はそういった「データをもとに根拠をもって判断し行動するスキル(エビデンス・ベースド・アプローチ、あるいはデータドリブンなど)」の最も基礎的なものの1つになります。また、このような考え方は企業の業務改善や問題解決のための各種プロジェクトでも一般的に用いられます。(以上うりぼーネットのシラバスより)

…ということで、理想的にはみなさんが自分自身で分析計画を立ててデータをあれこれ分析して報告書を作成してわかりやすく伝えて……とできるようになると良いのですが、いかんせん神戸 MBA には時間が無いようです。そして教員のリソースも足りません。そういうわけでこの講義では、統計学のごくごく基本的な概念、特に、限られたデータから全体像を推測し、意思決定に繋げるために用いられる「統計的推測」の考え方を学習します。「統計解析**応用**研究」という名前ですが、実際には基本も基本、学部レベルの(一部は高校数学でも必修になり始めている)内容の

お話になります。そういうわけで、現在データ分析バリバリの方や、学部時代統計学が得意だった方からするとかなりイージーな内容となります。そのような方は、ぜひ仲間をサポートしてあげてください。たぶん教員よりも学生同士のほうがフランクに質問などしやすいと思うので。

前半では、いまや高校数学でも扱われている平均値や分散、相関係数といった「記述統計量」の考え方を理解するところから始め、確率の基本、そして「母集団と標本」といった重要な概念を学びます。後半では、データから全体像を推測する「統計的推定」や、ある仮説が正しいかどうかをデータに基づいて判断する「統計的仮説検定」について学びます。

2. 授業の内容と計画

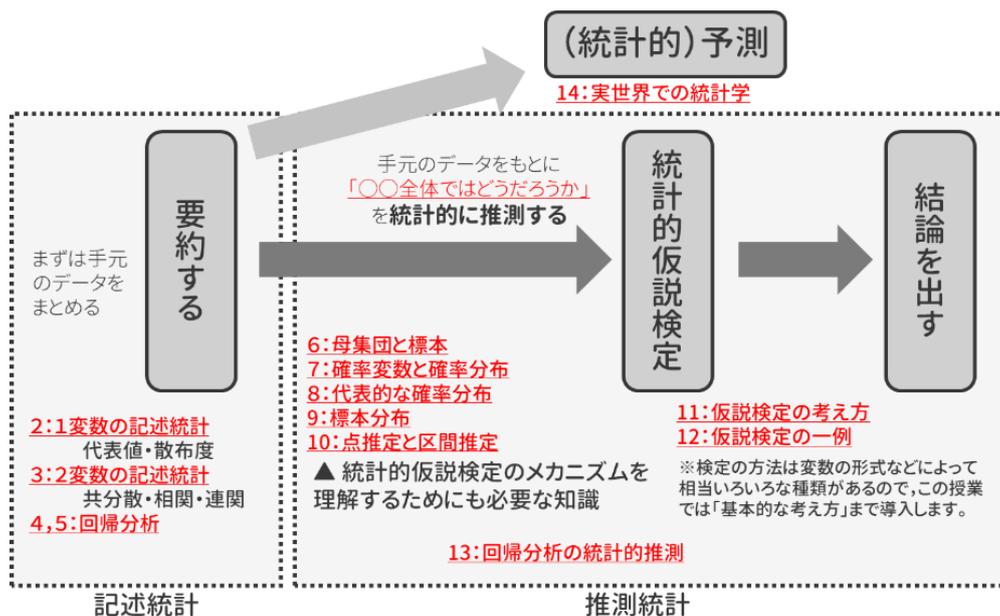
対面による講義を基本としますが、可能な限り Teams でのリアルタイム配信および講義動画の録画の共有を行い、ハイブリッド形式で実施する予定です。ただ過去の(他の講義の)履修者の様子を見てると、「録画があるからいつでも見れるし余裕のある時にやろう」と思っていると、大抵「余裕のある時」は来ないようです。MBA の場合、どうせ午後にも講義があるはずなので、外せない用事や体調不良でない限りは、できるだけ対面(少なくともリアルタイム配信)で参加し、学習のペースメーカーとすることをおすすめします。

なお、新たな未知なる感染症の流行など諸般の事情により授業形態が変更となる可能性もあります。変更がある場合は BEEF+にてお知らせします。

本講義の流れをざっくりと示したのが以下の図です。大きく分けると

- 前半(第 5 回まで):記述統計
- 後半(第 6 回から):推測統計

の 2 つのパートに分かれています。



【1日目:4月5日】

1. ガイダンス・データの性質

他の科目と比べたときの本講義の位置づけ、および統計学が何をしようとしているのか、について簡単にお話します。

余った時間で、統計学の視点からみた「データ」について少しだけお話します。

データ分析を料理に例えると、統計学は例えば「野菜の切り方の種類」や「加熱をするときの注意点」など、「調理方法」の講義に相当すると言えます。言うまでもなく、統計学の勉強だけを頑張ったところで、データサイエンティストになれるわけではありません。本講義で学ぶ内容を用いて良いアウトプットを生み出すためには、「良い食材の見分け方」や「各食材に適した調理法」などの知識が必要になります。ただ一方で、どんな野菜マニアであっても、仮に「加熱する」という工程を知らなければ作れる料理のレパートリはかなり限られてしまうでしょう。つまり、「良い」データ分析のためには、領域(・分野・業種)に固有の知識と、分析の方法に関する知識の両方が必要というわけです。そのように考えていくと、神戸 MBA のカリキュラムにおける本講義の位置づけが見えてくるでしょう。

また初回では、具体的に「統計学」という学問が、データを使って何をしようとしているのか、という大きな視点についてもお話します。これ以降の各回の内容は具体的なものになっていくわけですが、常に「結局のところ何がしたいのか」という視点を意識することで、理解が進むのではないかと期待しています。

2.1 変数の記述統計量

一つ一つの変数の特徴を簡潔に要約するために使用される統計量について学びます。

具体的には、平均値・中央値・最頻値などの代表値と、偏差や分散・標準偏差といった散布度の指標について見ていきます。

前半パートでは、まず手元にあるデータにどのような特徴があるかを考えていきます。例えば「平均値」などという指標は、大半の社会人であれば日常的に使うでしょうし、少なくとも知らない人はいないのではないかと思います。ただ平均値はあくまでも「データを単一の値で代表させる」方法の一つでしかないので、平均値の他にもよく用いられる「データを単一の値で代表させる」方法(中央値・最頻値)の考え方を簡単に紹介します。また同じように、「データの散らばりを単一の値で代表させる」方法(散布度)の考え方も紹介していきます。特に「分散」はこのあともいたるところで登場してくる最重要概念です。

【2日目:4月12日】

3. 2変数の記述統計量

前回に引き続き、2つの変数の関係性を表す指標として使用される統計量:共分散・相関係数および連関係数を紹介します。

データ分析では、多くの場合複数の変数の関係性を見ていくことになります。例えば「今月の店舗ごとの売上」という単一の変数だけを分析する場合、「平均でいくら、ばらつきはどの程度」といった程度の分析しかできませんが、ふつう関心があるのは「どんな店舗の売上が高いか」といった話のはずです。つまり「売上」と別の変数との関係性を考えてあげることで「〇〇が多い店舗ほど売上が高い傾向にある」といった議論ができ、戦略立案等につなげられるようになります。第3回では、2つの変数の関係性を表す最も基本的な統計量である「相関係数」をしっかりと理解することが目的です。

4. 回帰分析(1)

2変数の関係性を表すための方法である回帰分析の基本を見ていきます。ビジネスで用いられる分析手法の中でも最も基本的なもの、かつ最も広く用いられているものの一つではないかと思います。

先ほど述べた「〇〇が多い店舗ほど売上が高い傾向にある」といった関係性を知ることができると、次は「〇〇が△△である店舗では売上はどのくらいになると期待できるか」といった予測が可能になります。2変数の関係性を表すための方法として、相関係数と双壁をなす方法である「回帰分析」について、基本的な考え方から紹介していきます。回帰分析はかなり幅広く使われていて、正しく利用しようとするとかかなり奥深い分析手法なのですが、コンセプトは普段私達が日常的に使っているかもしれないヒューリスティックとも密接に関係しています。つまり、世の中で言われているほど難しい方法ではない、ということです(たぶん)。

【3日目:4月19日】

5. 回帰分析(2)

引き続き、回帰分析に関するいくつかの重要な話題および3変数以上の関係を扱うための方法について説明します。

前回の続きです。回帰分析では「〇〇が多い店舗ほど売上が高い傾向にある」という関係性を、複数の「〇〇」について同時に考えることも可能です。このように3変数以上を同時に扱う分析法は「多変量解析」とも呼ばれ、社会の中ではあらゆるところで様々な多変量解析が用いられています。世の中には数え切れないほどの変数がデータとして存在しているわけなので、2変数ごときでは何も見えてこないというわけですね。

6. 推測統計:母集団と標本

これ以降の内容に入る前に、推測統計を学ぶ上で非常に重要な「母集団」と「標本」という考え方を導入します。

ここから後半パートに入っていきます。統計学が用いられるとき、大抵の場合は調査対象をすべて観察・記録することはできない状況です。そこで、一部だけをサンプリングし、そのデータを使ってなんとかしようと考えていきます。推測統計では、手元のデータを「全てのあり得る候補からランダムに抽出されたもの」だと考えることで、データ=標本の背後にある「全てのあり得る候補」=母集団全体について統計的に推測を行っていくのです。この回では、推測統計の基本的なコンセプトを説明していきます。これが理解できないと、この先の計算などについて「なぜそんなことをやっているのか」がどんどん分からなくなっていくでしょう。

【4日目:4月26日】

7. 確率変数と確率分布

推測統計で必須の「確率」の考え方について、まずそもそも「確率」とはなにかを考えた後で、統計学では「確率」がどのように機能しているかについて、確率分布というものから見ていきます。

推測統計ではデータを「全てのあり得る候補からランダムに抽出されたもの」だと考えるわけですが、この「ランダムに」というのは、いわばくじ引きのようなものと言えます。例えば「当たり確率が $X\%$ のくじを N 回引いたときに K 回当たる確率」が理論的に計算できるのと同じ要領で、推測統計では「ある母集団から N 個のデータをサンプリングしたときに、標本での平均値がとる確率」といったものを計算していきます。そういうわけで、統計学においては「確率」というものが切っても切り離せない関係にあるのです。

8. 代表的な確率分布

実際の統計学において用いられることが多い代表的な(離散・連続)確率分布をいくつか紹介します。

とはいえ母集団というのは「神のみぞ知る」とても複雑なものです。究極的には「ある母集団から N 個のデータをサンプリングしたときにどうなるか」なんてことはわからないので、様々な仮定をおいて問題をシンプルに置き換えていきます。種々の仮定のもとでは、ある変数を取る値がどのような確率分布に従うかが明らかです。この回では、初歩的なレベルのデータ分析でよく出てくるいくつかの代表的な確率分布を紹介します。

【5日目:5月3日】

9. 標本分布

推測統計の目的である母数に関する推測のために必要な、標本統計量の分布の考え方を導入します。

確率分布について理解したところで、いよいよ「ある母集団から N 個のデータをサンプリングしたときに、標本での平均値が取る値」がどのような確率分布に従うか(=標本分布)を見ていきます。これがわかると、標本での統計量をもとにして「この標本統計量はどんな標本分布から得られたと考えるのが自然か」、そして「そんな標本分布はどんな母集団分布から生み出されたと考えるのが自然か」というロジックによって母集団についての推測を行うことができるようになります。

10. 統計的推定

データをもとに、母数を推定(点推定・区間推定)する方法およびその理論的基盤を学びます。

前回までに学んだいくつかの重要概念を組み合わせ、具体的に母集団の統計量(=母数)の推定の考え方を学んでいきます。実際に推定を行う場合には、例えば「母集団の平均値を推定する場合には標本平均をそのまま使えば良い」などルールのようなものはすでに明らかになっているのですが、ここではもう少し踏み込んで考えてみます。推定するために使用できる統計量の候補は色々と考えられる中で、どのような基準で選ばれた統計量が現在用いられているのか、ということも理解しておきたいところです。

また、サンプリング自体がランダムに行われる以上、標本統計量に基づく推定の結果もランダムに変動してしまいます。そのため、例えば標本平均をそのまま用いる代わりに、「標本平均の想定される変動幅」を考慮して推定結果を区間で出すこともよく行われます。この回では、この区間推定についても説明していきます。

【6日目:5月10日】

11. 統計的仮説検定の考え方

あらゆる統計的仮説検定の背景にある「帰無仮説・対立仮説」「有意水準」「検定統計量」などの概念を学びます。

世の中には数え切れないほどの種類の統計的仮説検定があります。しかし、初回でも簡単に説明するように、基本的な枠組みはすべて同じです。一例を挙げると、「母集団での平均値が〇〇である」という仮定(=帰無仮説)を一旦おいて、「その仮定が正しいとしたときに、標本平均はどのような確率分布に従うか」を考えます。こうすることで、実際に手元のデータから計算された標本平均から「帰無仮説が正しいとしたらそのような値が得られる確率はどの程度か」を計算できます。仮にこの確率が低ければ、最初に置いた帰無仮説が間違っていたのでは?と推測できるわけです。この回では、この統計的仮説検定の基本的な枠組みを紹介していきます。

12. 統計的仮説検定の例

仮説検定の(たぶん)最もベーシックなものをもとに、実際に仮説検定がどのように行われるかを学びます。

前回学んだ内容をもとに、具体的な統計的仮説検定の例をいくつか紹介していきます。内容はもちろんですが、複数の仮説検定において基本的な流れが同じであることをしっかりと感じ取ってください。

【7日目:5月17日】

13. 回帰分析の統計的推測

回帰分析モデルをもとに、「統計モデリング」の考え方を紹介した上で、回帰分析の統計的推測の方法を説明します。

前回に引き続き、統計的仮説検定および統計的推定の一例として、回帰分析における回帰係数のケースを見ていきます。ただ、この回で重要な内容は「統計モデリング」の考え方です。そもそもデータが発生するメカニズム(例えばある会社の利益がどのように決定するか)は人間の手には負えないほど複雑なものです。「8. 代表的な確率分布」のところにも記したように、統計学ではここに種々の仮定を置くことで、人間にも解釈できるレベルに落とし込んで様々な推論を行っています。回帰分析はこうした統計モデリングの最もシンプルな例の一つと見ることができます。実際のところ、回帰分析を統計モデリングの一つとして捉えなくても分析自体はできるのですが、この先社会で統計学を利用していく際に、この考え方が頭の中にあると視界がパッと明るくなるような、そんなコンセプトだと個人的には思っています。

14. 社会における統計学の利用

これまでの講義内容を元に、実際に社会科学ではどのように統計学が用いられているかの一例として、いくつかの発展的なトピックの理論的側面を簡単に紹介します。

最初にも書きましたが、本講義で扱ってきた内容は基本中の基本です。残念ながらこの講義の内容を完全に理解したとしても世の中のデータと対峙していくことはほぼ不可能と言っても良いでしょう(とはいえ、ここまでの内容ですらきちんと理解できている人はかなり少数なのですが…)。本当はもっと実戦的でもっと発展的な内容も紹介できると良いのですが、いかんせん時間が足りません。そこで最終回では、ここまで学んできた内容と結びつける形で、世の中ではどのような形で統計学が使われているのか、いくつかの発展的なトピック(一般化線形モデル・統計的学習・統計的因果推論・多変量解析・ベイズ統計)を少しだけでも紹介することで、今後の学習の指針を提供できればと考えています。一部の内容は「現代経営学応用研究(サーベイリサーチ)」へと繋がっていきます。

【8日目:5月24日】

15. 定期試験

定期試験はマークシート形式で実施します。申し訳ありませんが、不正行為の対策などが面倒なので、オンラインでの受験は認めません。必ず当日出席できるようにスケジュールの調整を済ませておいてください。なお、本科目における追試験は、[神戸大学経営学部のルール](#)に定められた事由に該当する場合に限り受け付けることとします。

また、筆記用具の他に「計算機能のみの電卓」の持ち込みを許可する予定です。一応頑張れば電卓が無くても解けないことは無いレベルの計算問題を出題しますが、絶対にあったほうが良いです。100均で買えるレベルで十分(ルートの計算までできれば十分、むしろ関数電卓まで行ってしまうと持ち込みできない)なので、早いうちに確認しておいてください。

3. 成績評価方法および成績評価基準

定期試験 80%、授業への参加度(各回に出す小テスト)20%で評価します。

定期試験では、いわゆる単語穴埋め問題というよりも、各概念の定義や意味を正しく理解しているかを問う問題、そして計算問題を出題する予定です。したがって、授業中でお話する概念や統計量について「きちんと理解する」「自分で計算できる」状態になることを目指してください。

4. 関連科目情報

定量的分析に関する講義としては、本講義の直後に(6月から)開講される「現代経営学応用研究(サーベイリサーチ)」があります。こちらでは、本講義で学んだ内容をもとに、より発展的な多変量解析の技法の紹介と、ソフトウェアによる演習を行っています。そういう意味では調理実習みたいなものかもしれません。違うかもしれません。

5. 事前・事後学修

事前学修では、講義資料の該当範囲に一通り目を通しておいてください。どちらかというと、事後学修(復習)に時間をかけるようにしたほうが良いかと思います。

本学では1単位あたりの学修時間を45時間としています。毎回の授業にあわせて事前学修・事後学修を行ってください。

詳細は初回でお話しますが、大学(院)における単位のルール上、授業を一回聞いただけで完全に理解できることは想定していません。具体的には各コマにつき180分程度の予習・復習を行うことまでは想定範囲内にあります。建前は(他の科目もすべて)そうなっているので、「授業を一回聞いただけではよくわからなかった」というのは残念ながら自己責任になってしまいます。

また講義期間中に数回、講義開始時間の前に MBA の先輩によるフォローアップセッションが予定されています。講義内容について不安のある方はぜひ参加をご検討ください。詳細は初回講義時に連絡します。

こちらについても初回に説明します。公式には今年度初めての取り組みです。

6. オフィスアワー・連絡先

- ・ 基本的にメールでの質問はいつでも受け付けます。
- ・ 授業の前後に直接質問しても構いません。
- ・ BEEF+から質問しても構いません。

いわゆるオフィスアワー(自由に相談に来て良い枠を決まった曜日・時間に設ける)はありません。ただ逆に言えば「いつでも質問は受け付けています」ということです。

7. 学生へのメッセージ

- ・ 統計の授業は自分で手を動かして計算することが非常に重要となります。座って話を聞いているだけで理解できるとは思わないので覚悟してください。

といいつつ、講義で扱う内容の分量的に、どうしても講義時間中に演習を十分に用意してあげることができません。ただ、講義資料に載っている計算の手順などは、写経でも良いので一度追いかけることをおすすめします。特に一つ一つの数式における各項の意味(例えば「なぜここで n で割るのだろうか?」「 $(x_i - \bar{x})$ って何を表していたっけ?」など)を考えていくと良いような気がします。講義資料でもできるだけそのような視点から話をしていきます。また、計算演習についてはもしかしたらフォローアップセッションで対応していただけるかもしれませんが未定です。

- ・ また「数学」と「統計」はかなり異なるものです。具体的に言うと、本講義で扱う内容における計算そのものはさほど難しくないのですが、それよりも「なぜその計算をしているのか」を理解するところがキモとなる科目です。なので、数学(計算そのものや、公式を覚えること)に苦手意識を持っている方も恐れることなく積極的に挑戦してみてください。

あえて厳しい言い方をすると、「こんなレベルでつまずいたらこの先苦労するぞ」という感じです。ただ本講義で伝えたい「考え方」はデータ分析の根底にあるようなものなので、「このレベルがきちんと理解できたらこの先もラクに理解できる様になるはず」という思いで講義を作っています。

8. 参考書

講義内容は、主に以下の参考書をもとに作成しています。本講義の内容をカバーしきれていないものもありますが、解説の数学レベルや説明の抽象度などが様々に異なるので、ぜひ自分に合った一冊を探してみてください。一部の本は、学内から電子書籍でアクセス可能となっています。

あくまでも私の主観ですが、上にあるものほど説明が平易である一方で、下にあるものほど数理的な説明が充実していると思います。

[基礎から学ぶ統計学 / 中原治：羊土社, 2022, ISBN:9784758121217](#)

[グラフィック 経営統計 / 森治憲：新世社, 2020, ISBN:9784883843152](#)

[コア・テキスト統計学\[第3版\] / 大屋幸輔：新世社, 2020, ISBN:9784883843077](#)

[基本統計学\[第5版\] / 宮川公男：有斐閣, 2022, ISBN:9784641165960](#)

[統計学入門 / 東京大学教養学部統計学教室：東京大学出版会, 1991,](#)

[ISBN:9784130420655](#)

[統計学 / 久保川達也・国友直人：東京大学出版会, 2016, ISBN:9784130629218](#)

「もっと易しい本」も「もっと専門性の高い本」も探せばいくらでもあると思います。上にあげたのはあくまでも「講義資料を作るときに参考にした本」です。文章での説明と図表での説明についてもどちらがわかりやすいかは人によります。レベル感だけでなく説明のスタイルも著者によってかなり異なるので、本屋や図書館などに行って「自分に合う一冊」を探すのがおすすめです。