

統計情報を見る眼

単位数	履修方法	配当年次
2	R or SR	2年以上



科目コード **DA3146** 担当教員 **平川 昌宏**

※この科目の会場スクーリングは隔年開講予定です。平成28年度の開講後は、平成30年度開講予定です。
※オンデマンド・スクーリングは平成28・29年度開講予定です。

■科目の内容

私たちは普段テレビや新聞、広告などを通してたくさんの統計情報に接しています。統計情報は物事を客観的に理解し、人に伝える上で非常に有効で便利なものです。さらに、統計情報を適切に「読みこなし」「使いこなす」ことによって、そこから有益な結論を論理的に導き出すことが可能になると考えられます。しかし、その反面、このように統計情報を適切に「読みこなし」「使いこなす」ことは必ずしも簡単なことではありません。さらに言えば、それができなければ、統計情報に「騙され」、知らず知らずのうちに真実とは異なる解釈に至ってしまったり、真実とは異なる印象を抱いてしまう危険性すらあると考えられます。

本科目では、まず、統計情報の種類やそれらの特徴について理解することを目的とします。さらに、それを解釈する際の統計的な思考法について学んでいくことを目的とします。以上を通して、統計情報に騙されることなく、それを「読みこなし」「使いこなす」ことができるよう「統計情報を見る眼」をより豊かなものにすることを目指します。

なお、統計情報は事象を数値的に処理したものであり、統計情報についての理解を深める上である程度の数学的な知識が必要になります。その点で、特に数学に苦手意識を持つ方には、履修することがはばかれる科目かもしれません。スクーリングでは、この点を考慮し、数学的な内容は必要最小限にとどめ、統計的な考え方の習得に焦点を当てていくこととします。

■到達目標

- 1) 記述統計と推測統計について説明できる。
- 2) 平均や度数といった統計情報の特徴、およびこれらを理解・使用する際の留意点について説明できる。
- 3) 相関について、さらには、相関関係と因果関係の違いについて説明できる。
- 4) 統計的検定の思考法について説明できる。
- 5) 示された統計情報、そして、そこでなされている主張について批判的に検討できる。

■教科書

飯田泰之著『考える技術としての統計学』日本放送出版協会、2007年

■在宅学習15のポイント

回数	テーマ	学習内容・キーワード	学びのポイント
1	統計的思考とは何か① ：統計詐欺の3つのパターン (第1章 第1節)	「統計詐欺の3つのパターン」について理解する。さらに、記述統計と推測統計という2種類の統計についてその内容を理解する。 キーワード：見せ方の嘘、選択の嘘、収集の嘘、記述統計、推測統計	「統計詐欺の3つのパターン」については、身近な具体例を探すことで、理解をより豊かなものにしてください。
2	統計的思考とは何か② ：演繹法と帰納法 (第1章 第2節：～p. 28 15行目)	演繹法と帰納法の違いについて理解した上で、帰納法と統計的思考の類似性について理解する。 キーワード：演繹法、帰納法	テキストに「(現実的な帰納法)は統計的な発想といってもよいでしょう」と述べられています。なぜそのように考えられるかを説明できるように学習を進めてください。また、統計情報の中から「ふつう」と「ふつうじゃない」を区別することの意義を考えてみてください。
3	統計的思考とは何か③ ：仮説演繹法 (第1章 第2節：p. 28 15行目～)	仮説演繹法とはどのような思考法か理解し、このような思考法に統計が果たす役割について理解する。 キーワード：仮説演繹法	「理論的に導いた全体(仮説)を確かめる手法として、統計を用いる」という発想は、とても大切になる考え方となります。具体的な内容は以降の学習に含まれますが、その前提として「仮説演繹法」の思考法について説明できるようにしておいてください。また第3節(実践に生かすデータ活用法)は、前回と今回の学習内容をより具体的に説明する内容になっていますので、一読しておいてください。
4	「平均」による情報縮約① ：平均とは (第2章 第1節)	算術平均の算出法とその性質について理解する。 キーワード：算術平均、偏差、やじるべえの支点	「ふつう」を示す値として、平均(特に算術平均)がよく用いられます。その性質についてテキストでは「やじるべえの支点」と表現されていますが、その理由について説明できるように学習を進めてください。
5	「平均」による情報縮約② ：平均が「ふつう」を表さない場合 (第2章 第2節：～p. 54 6行目)	算術平均の留意点(どのような時、算術平均が「ふつう」を表現しないか)を理解すると同時に、度数分布表(ヒストグラム)の重要性について理解する。 キーワード：度数分布表(ヒストグラム)	テキストに「人は(日本人は?) どうも『平均はこれこれです』といわれるのに弱い。」とあります。それゆえに、算術平均の留意点を念頭に置きながらこの値を解釈していくことが大切になると考えられます。算術平均の留意点にあてはまりそうな身近な具体例を探すことで、理解をより豊かなものにしてください。
6	「平均」による情報縮約③ ：様々な代表値 (第2章 第2節：p. 54 7行目～)	最頻値、中央値がどのような値か理解する。さらに平均値、最頻値、中央値の3つの値を見比べてデータを解釈することの重要性を理解する。 キーワード：最頻値、中央値、代表値	最頻値、中央値、平均値という3つの値から、ゼータ全体の特徴(ヒストグラムの形)をどのように推測できるのかという点について理解を進めてください。

回数	テーマ	学習内容・キーワード	学びのポイント
7	「平均」による情報縮約④ ：データの「ばらつき」を理解する (第2章 第3節)	データ全体のばらつきを示す値(四分位範囲, 四分位偏差, 分散, 標準偏差)の算出法, およびその性質について理解する。 キーワード: 四分位点, 四分位範囲, 四分位偏差, 外れ値, 分散, 標準偏差, 偏差値	データ全体のばらつきを示す値として, 特に「分散」や「標準偏差」は多く用いられています。さらに, 以後の学習でも重要となりますので, 理解を確実なものとしておいてください。また, データのばらつきに着目することの意義について説明できるように, 学習を進めてください。
8	「比較」による状況判断① ：比較する数値の単位をそろえる (第3章 第1節)	複数のデータが比較可能か判断する際に, その単位を確認することの重要性を理解する。 キーワード: フロー, ストック	私たちは, 複数のデータを比較する中で, 物事の優劣や変化について判断を行います。だからこそ, 「そもそもそれらのデータが比較可能か」を考慮することが大切だと考えられます。今回の学習ではこの点について考慮する際の観点の1つとして「単位」に着目することの重要性について学びを進めてください。また「単位がそろっていないまま比較がなされている」具体例を考えることで, 理解を豊かにしてください。
9	「比較」による状況判断② ：サイズあたりで比較を行う (第3章 第2節)	複数のデータが比較可能にするための調整法として, サイズあたりの調整を理解する。 キーワード: サイズあたりの調整	前回の学習に引き続き, 複数のデータが比較可能かを考慮する際の観点の1つとして, 「サイズあたりの調整が必要か」という点に着目することの重要性について学びを進めてください。また「サイズあたりの調整がなされていないまま不適切な比較がなされている」具体例を考えることで, 理解を豊かにしてください。
10	変数間の関係性を理解する①: 視覚化する (第5章 第1節: ~ p. 150 15行目)	変数間の関係を視覚化する方法として散布図を理解する。また様々な相関関係について理解する。 キーワード: 散布図, 正の相関関係, 負の相関関係, 無相関, 完全相関, 2次関数状の関係	変数間の関係性を数量化する前に, それを視覚化し, 関係性をおおまかにイメージすることが大切だと考えられます。その方法として散布図の理解を進めてください。また, 様々な相関関係について, 身近な具体例を探することで理解を豊かにしてください。
11	変数間の関係性を理解する②: 数量化する (第5章 第1節: p. 150 16行目 ~ p. 153 5行目)	相関係数から2つの変数の関係性についてどのような情報が読み取れるかを理解する。さらには, 相関係数を用いた分析を行う際の留意点を理解する。 キーワード: 相関係数	2つの変数の関係性を記述し理解する上で, 相関係数は非常に有用な値となります。その一方でテキストに「変数間関係を見るためにまず初めに散布図作りをやっておかないと, 手痛い見落としをすることがあります」と述べられています。どのような「見落とし」が生じうるのか説明できるように, 学習を進めてください。

回数	テーマ	学習内容・キーワード	学びのポイント
12	変数間の関係性を理解する②：相関関係を解釈する (第5章 第3節：～p.173)	相関関係と因果関係の違いについて理解した上で、相関関係を解釈する際の留意点について理解する。 キーワード：因果関係	テキストで述べられている「ナマズの動き」と「地震」の関係のように、因果関係はないがデータ上相関関係が見られる現象のことを「擬似相関」と言います。「擬似相関」の身近な具体例を考えることで、理解をより豊かにしてください。
13	部分から全体を解読する①：記述統計から推測統計へ (第4章 第1節)	推測統計の考え方について理解する。さらに、母集団（ユニバーサル、ユニバース）と標本（サンプル）という用語を理解すると同時に、サンプリングの重要性について理解する。 キーワード：推測統計、母集団、標本(サンプル)、中心極限定理、誤差、サンプル平均、母平均、正規分布	今回より学習内容が推測統計というより複雑な内容となります。まずは、推測統計の考え方、さらに推測統計を行う上でのサンプリングの重要性について理解を進めてください。中心極限定理については、複雑な内容ですが記述統計と推測統計をつなぐ重要な定理となりますので、そのイメージをつかんでもらえればと思います。
14	部分から全体を解読する③：検定の思考法1 (第4章 第3節：～p.141 14行目)	統計的な検定の思考法について理解する。 キーワード：検定、帰無仮説、対立仮説	「帰無仮説を否定できるかどうかを確率的に検討する」という検定の思考法は、少々複雑かもしれませんが、テキストでは幾つかの具体例を用いこの考え方について解説が試みられていますので、それらを参照しながら理解を進めてください。
15	部分から全体を解読する③：検定の思考法2 (第4章 第3節：p.141 15行目～)	統計的な検定の中で生じうる2つの誤りについて理解する。 キーワード：第一種過誤と第二種過誤	統計的な検定はあくまで、確率的な判断を行うものであり、誤る可能性を含んだものです。また、許容される「誤りの大きさ」は分析者が決めるものであり、その意味で分析者の考え方によって、検定結果も異なってくる可能性もあります。そのような検定の特徴について理解を進めてください。

■レポート課題

1 単位め	テキストの第1章を読み、「統計詐欺の3つのパターン」を具体例を挙げながらまとめてください。
2 単位め	<p>次の2つの課題1・2のうち、いずれか1つに回答してください。</p> <p>課題1 1つの変数の特徴を記述する際に用いられる値として算術平均があります。</p> <ol style="list-style-type: none"> 算術平均の特徴について述べてください。 1つの変数の特徴を記述する上で算術平均を用いる際の留意点について述べてください。さらに、算術平均がデータを代表する値として不適切な場合、そのデータの特徴を理解し、記述する際の方法について述べてください。 <p>課題2 2つの変数の関係を述べる際に相関という言葉が用いられます。</p> <ol style="list-style-type: none"> 相関の種類についてまとめてください。また、相関係数の特徴について述べてください。 相関関係を解釈する際の留意点について具体例を挙げて説明してください。その際、「因果関係」、「疑似相関」という用語を必ず用いること。

1単位め
アドバイス

テキストの第1章『統計的思考とは何か』では、真実とは異なる印象を抱かせる統計情報のパターンとして、3つのパターン（①見せ方の嘘、②選択の嘘、③収集の嘘）が述べられています。この3つのパターンそれぞれについて、テキストの内容を参考にしながら述べてください。その際、各パターンの具体例をそれぞれ挙げるようにしてください。

具体例については、まず、皆さんのまわりにある統計情報、あるいは、テキスト以外の文献等で紹介されている統計情報を見直し、その中で各パターンに該当するものがないかどうかを検討してください。このような検討を行った上で、それでも各パターンに該当する適切な具体例が見当たらない場合は、テキストで述べられているものについて述べても構いません。

2単位め
課題1
アドバイス

たとえば、「〇〇のテストの平均は△△点だった」など、平均（とりわけ算術平均）は私たちにとって非常になじみ深い統計情報です。それゆえに、その特徴や留意点を理解した上でつきあっていく必要があると考えられます。今回の課題については、テキストの第2章『「平均」による情報縮約』を学習した上で取り組むようにしてください。詳細なアドバイスは以下の通りです。

i) について

- ・テキストでは、4種類の平均について述べられています。レポート課題では、その中で算術平均について述べるようにしてください。
- ・テキストでは算術平均の特徴として「やじるべえの支点」という表現が用いられています。この点について詳しく述べるようにしてください。

ii) について

- ・主にテキスト第2章の第2節『平均の誤解を解く』に書かれている内容を参考にし、レポートを作成するようにしてください。
- ・他のデータに比べて極端に大きな値や小さな値を示すデータのことを「外れ値」と言います。算術平均を用いる際の留意点について述べる際、この用語を用いるようにしてください。

2単位め
課題2
アドバイス

ある事柄について予測を行う、あるいは原因をさぐる上で、2つの変数の関係を理解することは非常に有益だと考えられます。今回の課題については、テキストの第5章『「回帰分析」による関係把握』を学習した上で取り組むようにしてください。詳細なアドバイスは以下の通りです。

i) について

- ・主にテキスト第2章の第1節『関係性を発見する』の中の、『まずは視覚化する』（p. 148～）、『次に数量化する』（p. 150～）に書かれている内容を参考に、レポートを作成するようにしてください。
- ・相関係数の特徴については、この値から2つの変数の関係についてどのような情報を読み取ること

が可能なのかについて述べてください。

ii) について

- ・主にテキスト第2章の第3節『回帰分析の有効利用へ』の中の『因果と相関』（p. 171～）の中で、回帰分析の根本的な問題点として述べられている事柄を参考に、レポートを作成するようにしてください。
- ・「疑似相関」とは、実際には直接関係のない変数同士（たとえば、「地震」と「ナマズがさわぐ」）が、両者に共通する別の要因（「地殻変動」）の影響を受けて、見かけ上相関関係が生じる現象のことを言います。
- ・具体例としては、皆さんのまわりにある統計情報や一般的に言われている通説の中で、「疑似相関」であることが考えられるものを1つ挙げるようにしてください。

※この科目は「TFUオンデマンド」上で、担当教員によるレポート・アドバイスの動画を視聴することができます。

■科目修了試験 評価基準

科目修了試験100%。科目修了試験は論述式の試験を行う。在宅学習の15のポイントで記されているキーワードについての理解を問う問題となっている。

■参考図書

- 青木繁信著 『統計数字を読み解くセンス』化学同人, 2009年
- 上田尚一著 『統計グラフの賢い見方・作り方』講談社ブルーバックス, 1988年
- 大村 平著 『統計のはなし（改訂版）』日科技連, 2002年
- ジョエル ベスト著 『統計はこうしてウソをつく——だまされないための統計学入門』林大訳 白揚社, 2002年
- 谷岡一郎著 『データはウソをつく』筑摩書房, 2007年
- ダレル ハフ著 高木秀玄訳 『統計でウソをつく法』講談社ブルーバックス, 1968年
- 吉田寿夫著 『本当にわかりやすいすぐ大切なことが書いてあるごく初歩の統計の本』北大路書房, 1999年