

まえがき

このテキストは、初心者の皆さんが、実証会計の初歩を学習するさいのお手伝いをするために書かれたものです。具体的には、実証会計の卒業論文を書きたいと思っている経済・経営系学部の学生の皆さん、実証会計のタームペーパーを書きたいと思っている MBA コースの学生の皆さんを、第一の読者として想定しています。もちろん、それ以外の方でも、同じような関心をお持ちであれば、このテキストをそれなりに面白く読んでいただけると思います。

最近では、企業財務のデータベースが大変よく整備されていますし、使い勝手のよい表計算ソフトも手ごろな価格で購入することができます。したがって、たとえ初歩的なものであっても、ファイナンスと統計分析の体系的で正確な知識さえあれば、初心者の皆さんでも実証会計のペーパーを書くことは、十分に可能です。このような可能性がせつかく身近にあるのですから、その可能性を生かさないのは本当にもったいないことです。このような思いが、このテキストを出版するに至った最も大きな動機です。1人でも多くの皆さんに実証会計に挑戦していただきたいと思います。このテキストが、そのような皆さんの学習に少しでも役立てば幸いです。

このテキストは、京都大学経済学部の私(藤井)のゼミナールで試行的に実施してきた実証会計の学習プログラムをもとにして執筆したものです。私のゼミナールでは毎年、2・3・4 回生が一緒になって共同研究を行っています。研究テーマは年ごとに変わりますが、会計情報の統計分析に依拠した会計研究という枠組みは、ゼミ生諸君の希望もあって、この数年、維持してきました。

最初、最も困ったのは、初心者向けの適当なテキストがないことでした。実証会計の定評あるテキストがないわけではないのですが、どれも研究者コースの大学院生かそれ以上のレベルをもった読者を対象としたもので、学部のゼミナールのテキストとして利用するには難しすぎました。しかし他方では、実証会計の共同研究に取り組むことをゼミナールのメインの目的としていましたから、ゼミ生諸君に実証会計の基礎を、少なくとも半期程度でマスターしてもらう必要があるという事情もありました。それまで実証会計についてはまったくの初心者だったゼミ生諸君に、実証会計の基礎を体系的かつ正確に理解してもらい、半年後にはオリジナルの研究が可能なレベルまでステップアップしてもらう。そのような必要に迫られて、学習プログラムを試行錯誤で組み立てました。それをこのたび、このようなテキストの形にして公表するのは、うえに書いたような思いがあったからです。第一の読者として学部や MBA コースの学生の皆さんを想定しているのは、このテキストのそうした生い立ちに由来するものです。

このテキストの実践編である第IV編の2つ章(第9章と第10章)は、私のゼミナールで実

際に行った学生諸君の共同研究の成果がもとになっています。ゼミ生諸君が書いた過年度の共同研究論文のなかから2編を選び、そのダイジェスト版を作成して収録したものです。ちなみに、ゼミ生諸君が書いた過年度の共同研究論文はすべて、京都大学学術情報リポジトリ(<http://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/24365/browse-date>)にアップロードされていますので、インターネットで閲覧が可能です。

繰り返しますと、このテキストは、自分で実際にオリジナルのペーパーを書きたいと思っている初心者の皆さんが、実証会計の初歩を学習するさいのお手伝いをするという目的に徹して執筆されたものです。したがって、このテキスト一冊で、実証会計のすべてをカバーすることはできません。とりあげているファイナンスの理論やモデルは、どれも基本的なものばかりです。統計分析については、最も初歩的で、その割には応用範囲の広い3つの手法に絞り込んで、解説しています。そしてまた、とくにお断りしておかなければならないのは、このテキストでおもな学習対象としているのは、市場アプローチの実証会計(レリバンス・スタディ)であるということです。もう1つの代表的な領域である契約アプローチの実証会計については、第1章などでその特徴を概説するにとどめています。初心者の皆さんには、市場アプローチの実証研究のほうが、とっつきやすいと判断したからです。私のゼミナールでも、どちらかといえば、市場アプローチの実証会計にウエイトをおいた共同研究を行っています。契約アプローチの実証会計についてもっと詳しく知りたいと思っている皆さんには、より専門的なテキストでの学習をお奨めします。

このテキストでは、現在最も普及している表計算ソフトの1つと思われる Excel の利用を想定して、統計分析の解説をしています。このテキストで取り上げている統計分析は基本的には、Excel で十分に可能です。しかし、Excel の通常の操作でカバーできる範囲には限界もあります。より発展的な統計分析が必要となる場合には、SPSS, EViews, TSP, S-PLUS, JMP などの統計専用ソフトを利用することをお勧めします。直接コマンドを打ち込む必要があるので少々ハードルが高くなりますが、R というフリーの統計専用ソフトもあります。このテキストでも、Excel 以外の統計専用ソフトを利用して計算した部分が、若干ですがあります。

統計分析に関する解説の正確性を期するために、私の大学院ゼミの旧学生である篠田朝也君(北海道大学大学院経済学研究科准教授)に共著者として、本書の作成に加わってもらいました。また、同じく私の大学院ゼミの旧学生である田中伸君(滋恵医療経営管理研究センター主席研究員)に、本書全体のチェックをしてもらいました。

最後に、本書の出版を快く引き受けてくださった中央経済社の山本時男社長、企画・編集の労をとってくださった竹内伸介氏に、御礼を申し上げます。

200■年■月

藤井秀樹

もくじ

まえがき

プロローグ 統計分析で見えてくる世界

第1章 実証会計とは何か

第I部 ファイナンスの基礎知識

第2章 会計とファイナンス

第3章 企業価値の評価モデル

第II部 統計分析の基礎知識

第4章 統計学を理解するための準備

4.1 平均と分散

4.2 母集団と標本集団

4.3 分布

4.4 信頼区間

4.5 検定

第5章 回帰分析

第6章 平均差の検定

第7章 カイ2乗検定

第III部 Excelの実習ガイダンス

第8章 Excelを使った統計分析の手ほどき

第IV部 実証会計の実例

第9章 実証会計の実例1—浮動株と株価形成に関する考察—

第10章 実証会計の実例2—高齢化の株価形成に与える影響—

エピローグ

索引

参考文献

用語

プロローグ 統計分析で見えてくる世界

0.1 はじめに

会計の知識は一応あるのだけれど、実証会計に必要なファイナンスや統計分析についてはまったくの素人なので、どうやって勉強していいのかわからない。周りの人たちがみんな統計を使った会計情報の分析をやっているの、自分もやりたいと思うが、数学が苦手なので内心諦めている。ご安心ください。そんな皆さんを、実証会計の世界に無理なくお連れするのが、このテキストの目的です。このテキストを読むだけで、実証会計で学部の卒業論文やMBAコースのタームペーパーが書けるレベルの知識が身につきます。

実証会計とはどのようなものかを、まずは体感してください。実証会計には、統計分析のほかにファイナンスの知識が必要です。しかし、ファイナンスは後回しにして、このプロローグでは、実証会計イコール統計分析というふうに単純化します。そして、できるだけ数式を使わずに、統計分析で見えてくる独特の世界を、皆さんにご紹介します。もちろん、このような単純化は好ましくないのですが、それは初心者の皆さんに実証会計の世界を体感して頂くための便宜と、ご理解ください。

0.2 統計分析の利点

まず、統計分析の利点を説明しておきましょう。

統計分析の第1の利点は、分析対象(これをサンプルといいます)に関する数量データを独特の方法で加工することによって、見た目では観察することのできないサンプル全体の傾向や性質を、数学的な厳密さで、はっきりと観察できるようにしてくれるという点にあります。このような分析によって、定性的で抽象的な議論を、定量的で具体的な議論に発展させることが可能になります。つまり、統計分析は、様々な現象についての私たちの理解を、より深いものにしてくれるのです。

統計分析の第2の利点は、統計的な計算数値(これを統計量といいます)が示す意義と限界について、分析者の間で共通の理解が存在するという点にあります。これは、統計分析においてはつねに、数量データがどのような条件で収集され、加工されたのかが明示されることに由来します。これらの条件さえ分かれば、誰でも、その統計量を後から検証することができ、また再現することができます。このような性質を、検証可能かつ再現可能といいます。統計量は、このような性質を持つために、現象の観察において、客観的な証拠と

して利用することができるのです。

つまり、統計分析によって、私たちは、ある現象についての観察を、より本質的で客観的なものにすることができるのです。とはいえ、統計分析さえやれば、手放しでそれが可能になるというわけではありません。統計分析を信頼に足るものとするためには、いろいろと留意しなくてはならないポイントがあります。しかし、そうしたポイントについては、次章以下でそのつどふれることにし、皆さんに統計分析を体感していただくためのお話を先に進めることにします。

0.3 統計分析のイメージ

それでは、実証会計の扉をノックしてみることにしましょう。

以下では、イメージがつかみやすく、様々な分析に応用が可能な、平均の差の検定、独立性の検定、回帰分析をとりあげることになります。数学的なことは気にせず、統計分析の大まかなイメージをとらえることに集中してください。

0.3.1 平均の差の検定

ある現象を観察しようとするときに、分析対象を2つのグループに分け、それぞれのグループの間に差があるかどうかを知りたいと思うことが、よくあります。そのようなときに、大きな助けとなるのが、平均の差の検定です。

平均の差の検定は、かんたんにいいますと、2つのグループに関する同種のデータが同じように分布しているかどうかを、平均値に着目してテストするものです。平均の差の検定は、正確には「2つの母平均の差の検定」と呼ばれ、統計分析では最も基本的な手法のひとつとされています。

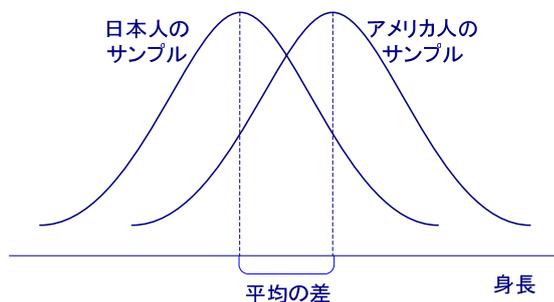
たとえば、アメリカ人と日本人を100人ずつランダムに抽出し、それぞれのグループの平均身長を比較するというケースを考えてみましょう。抽出されたサンプルの平均身長を計算したところ、アメリカ人のほうが日本人よりも4cm高いということが分かりました。さて、この4cmの差によって、両グループの間に「差がある」と主張することができるでしょうか。

このサンプルでは、たまたまアメリカ人のほうが日本人よりも平均身長が4cm高かっただけかもしれません。しかし、この4cmの差が、たまたま生じたものではないといいきれるほど大きな差であるということが分かれば、「アメリカ人のほうが日本人よりも一般に身長が高い」と主張できることになります。この例で求められているように、「2つのグループの平均の差がたまたま生じたものであるかどうか」を判定したいときに、平均の差の検定が威力を発揮しますⁱⁱ。

平均の差の検定では、「平均値に差があるかどうか」を、有意性という概念に依拠して判定しますⁱⁱⁱ。有意性とは、採取されたサンプルの間の差が、「偶然のものではない」といい

きれるだけの信頼性を備えているかどうかを言い表すための統計的な概念です。有意性は、この場合、「差の信頼性」といいかえても差し支えありません。統計分析ではふつう、「1%水準で有意」、「5%水準で有意」といった言い回しで、「差の信頼性」を表現します。この点の詳しいことは、第8章で学習します。

図表 0-1 2つの標本集団の平均の差のイメージ



両者の差はこのサンプルにおける偶然の差？
それとも、偶然ではないと言い切れるほどの差？

ここでは、4cmの差が「1%水準で有意」であるとします。これが意味することは、「このサンプルに限らず、アメリカ人と日本人からサンプルを繰り返し採取したとしても、アメリカ人の平均身長が日本人の平均身長よりも低くなるというようなことは、1%以下の確率でしか起こらない」と主張できるということです。つまり、「2つのグループの平均身長には差がある」と主張することには、99%以上の信頼性があるということが、平均の差の検定から分かるのです。平均身長の4cmという差は、それほど大きいということです。

注意しなければならないのは、平均の差の検定からいえることは、「平均身長に差がある」ということだけで、それ以上でも以下でもないということです。「だからアメリカ人のほうが日本人よりも幸せである」というようなことは、平均の差の検定からはいえません。

この点を、もっと分かりやすい別の例で、説明しておきましょう。2つの異なる出版社から出版されている数学のテキストAとBがあったとします。それぞれのテキストで学習した高校生が、数学の試験でとった得点の平均の差について調べたところ、テキストAで学習した高校生の平均点のほうがテキストBで学習した高校生の平均点よりも3点ほど高く、この差が1%水準で有意であったとします。このとき、「平均点で判定すれば、テキストAで学習した高校生の成績のほうが、テキストBで学習した高校生の成績よりも良い」ということが、99%以上の信頼性をもって主張できますが、統計分析でいえることは、そこまでです。つまり、「だからテキストAで学習したほうが良い」というような価値判断は、統計分析からは下せません。このような価値判断には「良い」と判断するための基準が必要

であり、その基準は、平均の差の検定には含まれていないのです^{iv}。実証会計では多くの場合、ファイナンスの理論が、そうした価値判断の基準を提供することになります。

では、平均の差の検定を、会計研究に応用すると、どのような分析が可能になるでしょうか。たとえば、2つの業界に属する企業の売上高利益率の平均に差はあるか、異なる会計方針を選択している企業の収益率の平均に差はあるか、異なる業績評価基準を採用している企業の資本利益率の平均に差はあるかなど、さまざまな分析が可能になると思いませんか。皆さんの関心にしたがって、様々な応用の可能性を考えてみてください。

0.3.2 独立性の検定

平均の差の検定では、2つのグループの「平均の差」に着目して、2つのグループに差があるかどうかを調べました。これに対して、独立性の検定では、2つのグループの特徴に関する「割合の差」に着目して、2つのグループが互いに独立しているかどうか(つまり2つのグループに差があるかどうか)を調べます。したがって、独立性の検定を行うときは、事前にサンプルをいくつかのカテゴリーに分類しておきます。

具体例で考えてみましょう。ワインの好みについて関東人と関西人の間に差があるかどうかを調べようとしています。そこで、関東と関西でランダムに抽出した成人200人ずつにワインが好きか嫌いかを尋ね、その結果を図表0-2のような表にまとめました。ここでは、関東人か関西人かという地域性を示すカテゴリーと、ワインが好きか嫌いかという好みを示すカテゴリーが、サンプルの分類基準になっています。このように複数のカテゴリーを組み合わせてサンプルを集計した表をクロス集計表(または分割表)といいます。独立性の検定では、このクロス集計表をもとにして、関東人であるか、関西人であるかによって、ワインの好みに差があるかどうかを調べることになります。

図表0-2 クロス集計表

	ワイン好き	ワイン嫌い
関東人	140人	60人
関西人	110人	90人

図表0-2によれば、関東人のワイン好きの割合は70%〔140人/(140+60)人〕であり、関西人のワイン好きの割合は55%〔110人/(110+90)人〕です。独立性の検定では、この関東人の70%という割合と、関西人の55%という割合の差が、「統計的に有意である」と主張するのに十分なほど大きな値であるかどうかを調べることになります。

具体的には、 χ^2 (カイ2乗)統計量を計算し、それを χ^2 分布と比較することで、さきの平均の差の検定と同じように、有意性の判定を行うことになります。このために、独立性の検定は、 χ^2 検定とも呼ばれています。 χ^2 検定の詳細は第9章で学習します。

さて、計算の結果、図表0-2で示された割合の差は「1%水準で有意」といえるほど大き

な差であることが分かりました^v。それが意味するのは、「関東人と関西人の間にワインの好みについて差がない」というようなケースは、100組のサンプルのうち1組のサンプルにおいてすら生じない」ということです^{vi}。つまり、99%以上の信頼性をもって、関東人と関西人の間にはワインの好みについて差があると主張することができるのです。この場合、関東人の方が関西人よりもワインを好む傾向が強いと主張することが、99%以上の信頼性をもって可能ということになります。

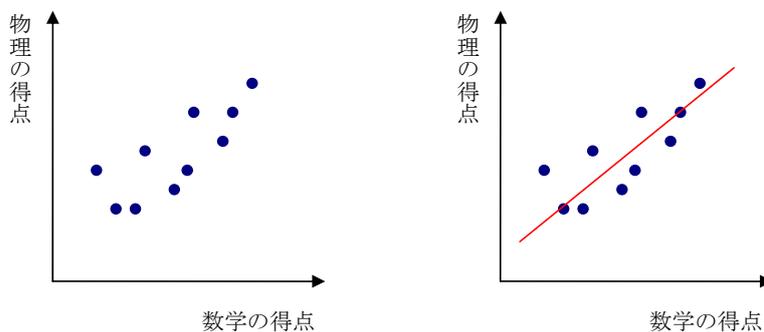
では、独立性の検定を会計研究に応用すると、どのような分析が可能になるでしょうか。たとえば、製造業と非製造業の間に事業資産の有税償却の実施状況について差はあるか、赤字企業と黒字企業の間定率法の採用状況について差はあるか、高収益企業と低収益企業の間減損会計の早期適用状況について差はあるかなど、さまざまな分析が可能になると思いませんか。独立性の検定も、会計研究において大きな威力を発揮しそうです。皆さんの関心にしたがって、様々な応用の可能性を考えてみてください。

0.3.3 回帰分析

対応関係がありそうな2種類のデータ群があるとします。このとき、一方のデータ群によって、他方のデータ群を説明したり、推定したりできないでしょうか。たとえば、ある個人の消費額をその人の所得額で説明できないか、ある個人の身長をその人の体重で推定できないかなどが、その具体例です。このような場合に役立つのが、回帰分析です。

具体例で考えてみましょう。大学生の数学の基礎学力が、数学の応用科目である物理の学力にどの程度影響を及ぼしているかを調べようとしています。そこで、ランダムに抽出した10人の大学生に、ある教室で同じ時間割に従って数学と物理の試験を受けてもらいました。その結果を集計し、数学の得点と物理の得点の散布図を描くと、図表0-3のようになりました。

図表 0-3 散布図と回帰直線



図表 0-3 の左の散布図をざっと眺めると、物理と数学の得点の間には正の相関がありそうだということが直感的に想像できます。正の相関とは、一方の変数が増加すれば、他方の変数も比例的に増加する関係をいいます。これとは逆に、一方の変数が増加すれば、他方の変数が比例的に減少する場合は、負の相関といえます。

このようなときに、直感的な想像ではなくて、数学的な厳密さで、物理と数学の得点の正の相関を観察することができれば、この調査の結果は大いに客観的なものになるでしょう。回帰分析は、図表 0-3 の右の散布図に示されるように、プロットされた点を一本の直線に要約することで、物理と数学の得点の関係を数量化する手法です。複数のプロットされた点を要約する直線を導き出すことを直線回帰といい、導き出された直線を回帰直線といいます。

回帰直線を求める代表的な方法として、最小 2 乗法があります。最小 2 乗法では、散布図にプロットされた各点と、当該各点から垂直方向にある直線上の点との距離の総和が最小になるような直線を求めることになります。詳しい計算方法は、第 7 章で学習します。こうして求められた回帰直線は、

$$(0.1) \quad Y_i = a + bX_i + \varepsilon_i$$

という計算式で表されることになります。(0.1)式は、回帰直線を表す式なので、回帰式と呼ばれます。この回帰式において、 Y は物理の得点を、 X は数学の得点を、 ε は誤差項を、それぞれ表しています。添え字の i は、ある特定の個人(i 君)を示す記号です。この回帰式の定数 a と係数 b を求めることができれば、数学の得点 X によって、物理の得点 Y を、数学的な厳密さで説明することが可能になります。そして、数学の得点 X の係数 b の値がプラスの値となれば、数学の得点 X と物理の得点 Y の間には正の相関関係が存在するということになります。

うえの説明で、誤差項という用語に引っかかった人がいるかもしれません。誤差項とは、回帰直線上の点と散布図上の実際の点の間のハズレの程度を示したものです。当たり前のことですが、数学の得点だけで、物理の得点の全てを説明することはできません。誤差項には、物理の得点を説明する数学の得点以外のすべての要素が集約されていると考えてもらえばよいでしょう。なお、うえの回帰式は、 X で Y を説明するという形になっていますので、統計分析では、 X を説明変数(または独立変数)、 Y を被説明変数(または従属変数)と呼びます。

回帰式については、当てはまりの良さ、いいかえれば回帰直線の説明力が問題になります。回帰直線の当てはまりの良さを示す指標に、決定係数と呼ばれるものがあります。これは通常 R^2 (アール・スクエア) と表記されます。決定係数は、0 から 1 の間の値をとり、1 に近いほどその直線の説明力は高いということになります。逆に 0 に近くなるほど説明力

は小さくなります。したがって、決定係数が高い値を示すような回帰直線が得られた場合、その回帰式には高い説明力があるということになります^{vii}。

ここまでは、2種類の変数を想定し、一方の変数で、他方の変数を説明するというお話をしてきました。つまり、1つの説明変数で、1つの被説明変数を説明するというケースです。このような回帰分析を、単回帰分析と呼びます。しかし、複数の説明変数によって、1つの被説明変数を説明する分析手法もあります。このような分析手法を重回帰分析といいます。たとえば、物理の得点 Z を、数学の得点 X と国語の得点 Y の2変数で説明するというような場合です。このような場合の重回帰式は、次のように表されることになります。

$$(0.2) \quad Z_i = a + bX_i + cY_i + \varepsilon_i$$

重回帰分析を行うときには、単回帰分析を行うときとは取り扱いの異なる点があり、いくつか出てきます。しかし、それらについては、第7章で学習することになります。

では、回帰分析を、会計研究に応用すると、どのような分析が可能になるのでしょうか。たとえば、当期に支出した広告宣伝費で次期の売上高を推計できるか、研究開発費の累計額で超過利益を説明できるか、利益の大小で株主価値（株価時価総額）の大小を説明できるかなど、さまざまな分析が可能になると思いませんか。回帰分析の応用範囲は無限といっても過言ではありません。皆さんの関心にしたがって、様々な応用の可能性を考えてみてください。

0.4 まとめ

このプロローグでは、実証会計イコール統計分析と単純化したうえで、統計分析で見えてくる独特の世界を皆さんに体感してもらうためのお話をしてきました。統計分析のイメージをつかんでもらえたでしょうか。ここで紹介した統計分析の手法は、ごく初歩的なものばかりです。しかし、それを、会計研究に上手に応用することで、驚くほど多くの発見や分析が可能になります。そして、それによって、皆さんは、学部や MBA コースの学生としては十分すぎるほど高度な議論を展開することができるようになるのです。

ここまでのお話で統計分析のイメージをつかんだ皆さんは、すでに様々な応用の可能性に思いをめぐらせているかもしれません。以下の章では、実際に実証会計を手がけるために必要な基礎知識を、順を追って、できるだけ分かりやすく、説明していきたいと思えます。

さて、それでは、いよいよ実証会計の扉を開いてみることにしましょう。

第1章 実証会計とは何か

この章のねらい

この章では、実証会計とは何かを、研究の基本的な手順に沿って、解説します。実証会計の全体像を大づかみに理解してください。実証会計における理論の役割を理解することがポイントです。実証会計は、研究課題の性質によって、市場アプローチによるものと、契約アプローチによるものに大別されます。

1.1 実証会計の現在

この章では、実証会計とは何かについて、ひと通り概説することにしたいと思います。皆さんに、実証会計の全体像を大づかみに理解してもらうことが目的です。

企業が公表した会計情報や財務情報等の統計分析を主要な一部として含む研究を今日、会計研究の領域では広く実証会計^{viii}と呼んでいます。なお、以下この章では、記述の簡略化のために財務情報等も含めて、たんに会計情報ということにします。

実証会計は、1960年代後半にアメリカで生まれた比較的新しいタイプの会計研究です^{ix}。しかしその後、実証会計は北米の会計研究者の間で急速に発展普及し、今日では会計研究の最も代表的な潮流となっています。図表 1-1 は、過去 5 年間(2000～2005 年)、会計研究の主要 3 ジャーナルに、何本の実証会計論文が掲載されたかを示したものです。ここに見られるように、主要 3 ジャーナルの掲載論文総数 481 本のうち、7 割以上の 344 本が実証会計論文で占められています。実証会計論文が圧倒的な多数を占めていることが分かります。図表 1-1 に示した市場アプローチと契約アプローチの違いについては、この章の 1.3 でふれることにします。

図表 1-1 主要 3 ジャーナルにおける実証会計論文の掲載状況

	掲載論文総数	実証会計論文		
		市場アプローチの実証会計論文	契約アプローチの実証会計論文	実証会計論文の合計
本数	481	205	126	344
(%)	(100.0)	(42.6)	(26.2)	(71.5)

(注) 1.主要 3 ジャーナルは、(1)Accounting Review, (2)Journal of Accounting Research, (3)Journal of Accounting and Economics です。

2.実証会計論文の合計には、2つのアプローチの実証会計論文以外のものが含まれています。

(出所)藤井[2007]164 頁。

日本では、アメリカよりやや遅れて、1970年代後半から実証会計が手がけられるようになりました^x。その後、アメリカほど急速ではありませんでしたが、実証会計は着実に発展普及してきました。たとえば、2005年の『会計』誌では、年間の掲載論文総数99本のうち、約2割の19本が実証会計論文となっています。そのほか近年では、乙政[2004]、須田編著[2004]、石塚編[2006]、石川[2007]といった研究業績も続々と刊行されています。アメリカと同じく、日本でも、実証会計の存在を抜きにして、会計研究の現状は語れないものとなっています。

1.2 実証会計の基本的な手順

実証会計の手順には、ごく大まかなものですが、標準的な流れがあります。それを整理すると、①問題意識と研究課題の提示、②先行研究のサーベイ、③モデルの提示とサンプルの収集、④統計分析の結果の提示と解釈、⑤結論というふうになります(図表1-2)。こうした手順に具体的な内容を与え、それを一つのまとまった研究計画として設計すること、またはそのように設計したものを、実証会計ではリサーチデザインといいます。

図表 1-2 実証会計の基本的な手順

① 問題意識と研究課題の提示(introduction)
② 先行研究のサーベイ(prior research)
③ モデルの提示とサンプルの収集(modeling and sampling)
④ 統計分析の結果の提示と解釈(empirical results and interpretation)
⑤ 結論(concluding remarks)

1.2.1 レポートと論文の違い

①の問題意識と研究課題の提示は、研究の出発点になる大変重要な部分です。どのような問題意識にもとづいて、何をどう明らかにしようとするのが、この部分で述べられます。研究課題を発見するためには、関連文献の精読を中心とした事前の学習が欠かせません。事前の学習がどれだけ周到になされているかで、研究の出来が決まるといっても過言ではありません。事前の学習には、十分な時間をとりたいものです。事前の学習の成果を、自分が取り組もうとする研究課題と関連させながらまとめたのが、②の先行研究のサーベイです。

研究課題を提示するときとくに留意する必要があるのは、自分がこれから手がけようとする実証会計のオリジナリティ(学術的貢献)は何かを明確にすることです。やや図式的な言い方をしますと、オリジナリティのないペーパーはレポート(essay)であり、オリジナリティのあるペーパーのみが論文(research paper)と見なされます。これは実証会計にかぎり

ません。会計研究に共通した評価基準です。しかし他のタイプの会計研究に比べると、実証会計では標準化された作業が多いために、先行研究のサーベイがとくに重視されることになるのです。

まったく新しい会計問題を、まったく新しい手法にもとづいて分析検討するというのが、実証会計の理想です。しかし、これは、専門の研究者でも、そう簡単にできることはありません。ほとんどの場合には、同種の会計問題を、同種の手法にもとづいて分析検討した先行研究が存在します。そうした条件のもとで自分の研究のオリジナリティをアピールするためには、先行研究で明らかにされた研究の到達点を確認しておく必要があります。先行研究の到達点に新しくつけ加えるものが、自分の研究のオリジナリティとなるからです。先行研究のサーベイを締めくくるときには、自分の研究のオリジナリティは何かをできるだけ明確に述べる必要があります。

専門の研究者が手がける実証会計の場合には、相当高いレベルのオリジナリティが要求されます。たとえば、研究課題のオリジナリティ、研究のベースとなる理論のオリジナリティ、モデルのオリジナリティ、分析手法のオリジナリティなどが、要求されます。

そのほか、サンプルのオリジナリティも、研究のオリジナリティとして認められます。たとえば、アメリカの1980年代のサンプルにもとづいて実施された実証会計を、アメリカの1990年代のサンプルにもとづいて追試したり(研究対象期間のオリジナリティ)、日本の1980年代のサンプルにもとづいて追試したりすること(研究対象国のオリジナリティ)も、オリジナリティの1つとして認められます。こうした研究は、専門の研究者が手がける実証会計ではプリテストと見なされることが多く、高い評価は与えられませんが、学部やMBAコースの学生の皆さんが卒業論文やタームペーパーとして手がける場合には、十分なオリジナリティを備えた研究として評価されます。卒業論文やタームペーパーでは、分析者の学習の成果を公表することが主たる目的となるからです。サンプルのオリジナリティを狙った研究は、初心者の方々には、実証会計への手頃な入り口となるでしょう。

1.2.2 理論モデルと実証モデル

研究課題が明らかにされると、それを分析作業として具体化することが、次の手順となります。それが、③のモデルの提示とサンプルの収集です。

たとえば、会計情報が投資家の意思決定にどのような影響を及ぼしているかを確かめることを、研究課題として取り上げたとします。この場合、「投資家の意思決定への影響」という大変漠然としたものが分析の対象となります。このような対象は、そのままの形では観察することができませんし、その程度を測ることもできません。そこで、「投資家の意思決定への影響」を象徴的に表していると考えられる何らかのデータを、「投資家の意思決定への影響」を数値化した変数として選択することが必要になってきます。このようにして選択された変数を、代理変数(proxy)といいます。「投資家の意思決定への影響」を観察しようとする場合にはふつう、証券価格とくに株価が代理変数として選択されます。

ファイナンスの理論にもとづいて定式化された企業評価モデルには、第 2 章で見るように将来キャッシュフローや将来利益を変数として組み込んだものが多いのですが、将来キャッシュフローや将来利益は理論上の数値であり、現実中存在するものではありません。したがって、そうした企業評価モデルを実証会計で用いる場合にも、数値として観察可能な何らかの変数を、将来キャッシュフローや将来利益の代理変数として選択することが必要になってきます。

このように、実証会計では、理論にもとづいて概念的に定式化されたモデル(これを理論モデルといいます)を、統計分析に利用可能なモデル(これを実証モデルといいます)に展開する必要があります。このような展開作業を、モデルの操作可能化(operationalization)といいます。

すでに述べましたように、入門的な実証会計では、モデルそれ自体のオリジナリティが問われることは、基本的にありません。しかし、代理変数の選択には、初心者でも色々な工夫をすることが可能な場合があります。そのような工夫を通じて実証会計のオリジナリティをアピールすることも、入門レベルでは意味のある 1 つの挑戦となります。第 11 章と第 12 章では、そうした工夫のいくつかを紹介したいと思います。

契約アプローチの実証会計では、仮説の検証という手順をとりますので、そのような手順にそくした形でモデルを提示する必要があります。具体的には、研究のストーリーラインにもとづいてある仮説が提示され、その仮説の真偽を検証するためのモデルが提示されます。そのアウトラインは、この章の 1.3 で述べることにします。

1.2.3 サンプルからのバイアスの排除

今日では企業財務のデータベースが大変よく整備されていますので、そうしたデータベースが揃っている大学や研究機関などでは、データベースから必要な会計情報を適宜ダウンロードすることによって比較的容易にサンプルを収集することができます。

サンプルの収集を行うさいに、とくに注意する必要があるのは、(1)条件の等しいサンプルを収集するということと、(2)分析者の恣意性を排除するということの 2 点です。

(1)については、(a)サンプル企業が属する業界を均質化する(たとえば「金融機関を除く」や「製造業に限定する」など)、(b)分析対象期間中に大規模な合併・分社化等を行った企業や決算期を変更した企業は除外する、(c) サンプル企業を 3 月決算企業に限定するといった作業が、通常行われます。(a)と(b)の理由は明白なので、とくに説明は必要ないと思います。(c)は、決算期の相違によって生じる可能性のあるバイアスを排除することを目的としています。観察の時期が異なれば、経済政策や為替相場などのマクロ環境要因が、証券市場や企業の会計行動に異なった影響を及ぼす可能性があります。そのような異なった影響を受けたサンプルをミックスして用いると、統計分析の結果に歪みを生じさせるおそれがあります。(c)は、こうした問題に対処するための手続です。

(2)も、上記の(c)と同じく、バイアスの排除を目的としたものです。分析に都合のよいサ

ンプルだけを恣意的に収集することによって、分析者は、統計分析の結果を自分の思うままにコントロールすることが可能となります。そのような実証会計が、研究の名に値しないのは、改めていうまでもありません。それは、実証会計において厳に慎まなければならないことです。(2)は、そうした恣意的な実証会計に陥らないための心がけです。

しかし、たとえば、フォーチュン誌のグローバル 500 にリストアップされた「世界の優良企業 500 社」のように、分析者の恣意性がまったく及ばない形で選択された企業グループをサンプルの母集団として採用することは、研究課題にもよりますが、許されています。第三者が作成したこのような企業リストを上手に活用することで、かえって幅のある実証会計が可能になる場合もあります。

以上のほかにも、サンプルが統計分析に適しているかどうかを確かめるために、チェックしておかなくてはならないポイントがいくつかありますが、少し専門的な内容になりますので、それは第 6 章でふれることにします。

1.2.4 ストーリーラインの組み立てと記述

モデルの提示とサンプルの収集が終われば、サンプルデータの統計分析は Excel などの市販の表計算ソフトを利用して、文字通り機械的に行うことができます。Excel を利用した実証会計の実習ガイドは、第 10 章で行います。

データの統計分析が終わると、次は、④の統計分析の結果の提示と解釈に移ります。統計分析の結果の提示は、実証会計の論文で、視覚的には最も強く印象に残る部分かもしれません。多くの細かな数字が、特殊な記号表現(たとえば補正 R^2 や***など)を交えて、表形式で示されるからです。慣れないうちは、そうした数字や記号表現が何を意味しているのか分からず、戸惑うことも多いでしょう。そのせいもあってか、会計情報の統計分析がすなわち実証会計であると考える人が、初心者の皆さんのなかには少なくないようです。事実、このテキストのプロローグでも、説明の便宜から、実証会計イコール統計分析という単純化を行いました。しかし、これまで述べてきましたように、統計分析の結果を提示するまでには、多くの時間と手順を費やさなくてはなりません。統計分析の結果の提示は、実証会計の非常に重要な部分ではありますが、研究の全体から見れば一部にすぎないのです。

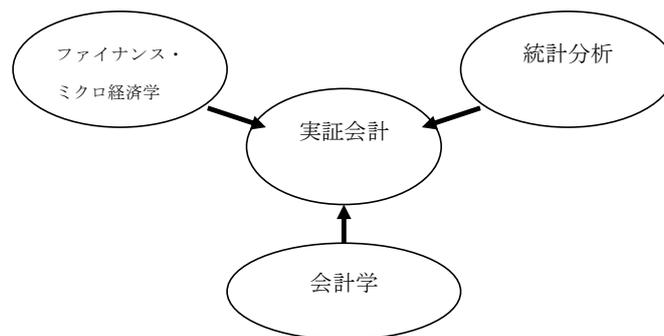
詳しいことは次の 1.3 で述べますが、市場アプローチの実証会計では、会計情報と市場の関係を明らかにすることが基本的な課題となります。他方、契約アプローチの実証会計では、会計実務のシステムティックな(つまり一定の傾向を持った)現象を説明し予測することが基本的な課題となります。こうした課題に取り組むためには、マイクロ経済学やファイナンスの理論にもとづいて、学術的に意味のあるストーリーライン(論理の流れ)を組み立てなくてはなりません。市場アプローチの実証会計ではファイナンスが、契約アプローチの実証会計ではマイクロ経済学が、おもに活用されます。

学術的に意味のあるストーリーラインの組み立てが必要なのは、他のタイプの会計研究

にも共通する点ですが、実証会計の場合には、ストーリーラインの記述に、さらに統計分析に固有の認識方法や考え方を取り入れることとなります。その点が、実証会計と他のタイプの会計研究の大きな違いです。したがって、実証会計では、マイクロ経済学やファイナンスの知識に加えて、統計分析の知識も必要となります。

以上を整理しますと、実証会計においては、会計情報をサンプルデータとして適切に取り扱うために会計学の知識が、研究のストーリーラインを組み立てるために、会計学の知識にくわえて、ファイナンスやマイクロ経済学の知識が、そのストーリーラインを実証可能な形で記述するために統計分析の知識が、それぞれ必要とされることとなります。これらの知識のうち、次章以下では、ファイナンスと統計分析を取り上げ、入門的な実証会計に最低限必要とされる基礎知識について解説していきます。

図表 1-3 実証会計を支える3つの要素



ストーリーラインの組み立てと記述に、話を戻しましょう。ストーリーラインの組み立てと記述は、実証会計の手順としては、モデルの提示とサンプルの収集に、密接に関連した作業です。研究のストーリーラインが組み立っていないければ、モデルを提示することも、サンプルを収集することもできないからです。つまり、どのようなモデルを用い、どのようなサンプルを収集するかは、研究のストーリーラインをどのように組み立てるかに依存して決まるのです。ストーリーラインの基礎には理論がありますから、その意味で、実証会計は理論依存的(theory laden)です。

だからこそ、統計分析の結果については、その研究の基礎になっている理論をふまえた解釈を行うことが必要不可欠となるのです。解釈とは、事実の観察結果を理論にフィードバックする作業です。統計分析の結果得られた数字の大小関係や統計学上の意味を述べるだけでは、実証会計とはいえません。そこで終わってしまえば、それは研究(research)ではなく、調査(search)となります。この点のもう少し具体的な解説は、次の1.3で行いたいと思います。

1.2.5 結論で述べること

⑤の結論では、実証会計全体の締めくくりを行います。結論は、研究を通じて明らかになった知見をまとめた部分ですから、そこで何か新しいことをつけ加えて述べるというようなことは、基本的にありません。と言うよりもむしろ、それ以前の部分で言及していない新しいことを結論として述べることは、論文のストーリーラインを破壊することになりますので、例外的な場合を除き、控えなくてはなりません。

ただし、研究の問題点や限界を、結論の一部として追加的に述べることはよくあります。完全な実証会計というものはありません。知見だけでなく、研究の問題点や限界も併せて明らかにしておくことは、研究の今後の発展のためにも大変有意義な作業となるでしょう。

1.3 市場アプローチと契約アプローチ

実証会計は、研究課題の性質によって、市場アプローチの実証会計と、契約アプローチの実証会計に大別されます。この節では、これら 2 つのアプローチについて、解説したいと思います。

1.3.1 市場アプローチの実証会計

市場アプローチの実証会計では、企業が公表する会計情報が市場の動きとどのような関係を持っているかを明らかにすることが、おもな研究課題とされます。言い換えると、市場アプローチの実証会計では、会計情報が投資家の投資意思決定との関連でどのような機能を果たしているかが、検証されることになります。実証会計では、会計のこのような機能を、意思決定支援機能と呼んでいます。

たとえば、会計情報の公表とともに株価の変化が生じたり、市場取引量が増減したりすれば、その会計情報にもとづいて株価の形成が行われたり、ポートフォリオの組替えが行われたと推定することができます。また、新しい会計基準にもとづいて公表された会計情報(たとえばキャッシュフロー情報)が、既存の会計情報(純利益情報)よりも、株価と強い相関関係を持っているとすれば、新しい会計情報は既存の会計情報よりも強い株価説明力(これを価値関連性 **value relevance** といいます)を持っていると推定することができます。

以上のような推定が実際にどの程度当てはまるかを、関連する会計情報の統計分析にもとづいて確かめるのが、市場アプローチの実証会計です。このために、市場アプローチの実証会計は、**レリバンс・スタディ(価値関連性研究)**と呼ばれることもあります。このテキストでおもに学習するのは、この**レリバンс・スタディ**です。

市場アプローチの入門的な実証会計では通常、ファイナンスの理論にもとづいて定式化された企業評価モデルが用いられます。そうしたモデルが用いられるとき、「**市場は完全で**

ある」ということが仮定されます。これは、ファイナンスの基本的な仮定です。

以上の解説を、もう少し具体的に整理すると、次のようになります。市場への影響を確かめたい会計情報を x とし、市場が受けた影響を表わす代理変数を y としますと、市場アプローチの実証会計では、

$$(1.1) \quad y=f(x)$$

という基本モデルを想定し、 x と y の相関関係を明らかにすることで、会計情報の市場への影響を検証します。 x と y の相関関係を明らかにするために最もよく利用されるのが、回帰分析です。ただし、相関関係を明らかにしただけでは、会計情報の市場への影響を検証したとはいえません。会計情報と市場の関係を理論によって説明し、相関関係を因果関係に展開したうえで、その関係を統計分析を通じて事実と付き合わせる必要があります。そこで実証会計を行うさいには(1.1)式をファイナンスの理論にもとづいて理論モデルに展開し、それをさらに統計分析やサンプル収集上の制約にもとづいて実証モデルに展開することになります。この点の詳しいことは、第7章で学習します。

1.3.2 契約アプローチの実証会計

契約アプローチの実証会計では、会計実務にシステムティックな現象が生じる理由を明らかにすることが、主な研究課題とされます。たとえば、有形固定資産の減価償却法の選択において、定額法を採用する企業グループと定率法を採用する企業グループに分離する傾向が観察されるとしましょう。そうすると、そのような傾向を生み出している理由は何かを解明することが、研究課題として取り上げられることとなります。そのさい、たとえば、「他の条件が等しければ、利益連動型報酬契約を株主と締結している経営者は定額法を愛好する」という仮説が設定され、関連する会計情報の統計分析にもとづいて、この仮説の真偽が検証されることとなります^{xi}。利益連動型報酬契約を株主と締結している企業では定額法を選択している事例が多いというシステムティックな傾向が直感的に観察されるときに、上記のような仮説が設定されます。そして、これが研究の実質的な出発点となります。

この事例では、定額法グループと定率法グループへの分離が生じた理由を、利益連動型報酬契約の有無によって説明できるかどうかを確かめることが、実証会計の課題とされています。このような場合、たとえば、利益連動型報酬契約の有無を基準として、定額法グループと定率法グループを区分することに統計的な意義があるかどうかを検証されることとなります。このような場合に利用されるのが、 χ^2 検定です。 χ^2 検定の結果、そのような区分には統計的な意義があるという結果が得られた場合には、上記の仮説を採択してもよいということになります^{xii}。

契約アプローチの実証会計では一般に、「すべての利害関係者は、自分の期待効用を最大

化するように行動する」ということが仮定されます。これはマイクロ経済学の基本的な仮定です。この仮定により、株主と経営者はともに、利益連動型報酬契約を締結するインセンティブを持つことが説明されます。つまり、こうした契約を締結することによって、株主は経営者の**モラルハザード**を抑止することができますし、経営者は**ボンディング費用**(信頼獲得費用)を節減することができます。そして、経営者は、会計方針の選択にさいして、それが利益連動型報酬契約で認められているかぎり、報告利益がより大きくなる会計手続(上記の例で言えば定額法)を偏好するという推定がなされます。上記の仮説は、こうした理論的推定にもとづいて設定されています。

この場合の契約とは、利己的に行動する複数の利害関係者が、それぞれの自由意思にもとづいて取り決めた利害調整のためのルールをさします。そうした契約は、法律上の契約とはかぎりません。契約アプローチの実証会計では、そのような広い意味での契約の取決と運用において会計がどのような役割を果たしているかが、解明されることとなります。実証会計では、会計のこのような機能を、契約支援機能と呼んでいます。

1.3.3 2つのアプローチについての簡単な締めくくり

以上のように、実証会計は研究課題の性質によって、大きく2つのアプローチに分けることができます。それぞれのアプローチでは、研究課題の性質が異なるために、採用される基礎理論や設定されるストーリーラインも異なってきます。

しかし、学術雑誌等に公表された実証会計論文のなかには、2つのアプローチを組み合わせたものも少なくありません。また、2つのアプローチは、市場メカニズムが適切に機能していることを仮定する点では、共通しています。実証会計を2つのアプローチに分けるのは基本的には、実証会計の特徴を概念的に整理するための便宜と考えてください。

この節では、市場アプローチと契約アプローチの特徴を駆け足で解説してきました。それぞれのアプローチの特徴をもっと詳しく知りたいと思う皆さんには、須田[2000]をお奨めします。これは、実証会計の専門書ですが、2つのアプローチの特徴が、理論的な背景も含めて大変詳しく丁寧に説明されています。

第1部 ファイナンスの基礎知識

第2章 会計とファイナンス

この章では実証研究に不可欠な要素である会計・ファイナンス・統計の概要を解説します。それぞれが実証研究に対してどのような役割を担っているのかを整理してください。また、ファイナンスの前提となっている諸理論についても言及します。

2.1 会計とファイナンスと統計

実証研究を行う上で大切な3つの要素があります。会計とファイナンスと統計です。

【図表2-1 実証研究に必要な3つの要素】



会計は実証研究に必要なデータを提供します。実証研究に用いる主要なデータである企業の会計情報は、一連の一般に公正妥当と認められた会計手続きによって作成されます。それらの会計情報は企業の財政状態および経営成績などの観点から、企業の実態を私たちの目の前に浮かび上がらせませす。企業の実態を示す主要なデータとしては、企業の資産・負債を示す貸借対照表 (Balanced sheet ,B/S)、収益と費用を示す損益計算書 (Profit and loss Statement ,P/L)、キャッシュの流入と流出を示すキャッシュフロー計算書等があります。

さて企業価値算定に用いられる会計情報は企業価値推定に役立ち、投資家の投資判断に有用なものでなければならない、とする考えを意思決定有用性アプローチといいます。会計の一連の手続きの結果として積み上げられた結果である会計情報は投資家によって有用だと認識され、投資の判断材料とされなければならないとする考え方です。この意思決定有用性アプローチにより、現在資本市場の理論 (=ファイナンス理論) と会計の有業研究が進展しています。実際に会計情報、特に利益情報が投資家の意思決定に役立っているのかを検証するために、第1章で述べた市場アプローチの実証研究が行われます。

コメント [TS1]: 契約理論 (エージェンシー理論などの経済学) がベースとなっている実証研究もあるので、資本市場アプローチの実証研究に限定している旨、記しておくべき。

コメント [s2]: Income Statement

コメント [TS3]: キャッシュ・フロー計算書 (中黒を入れる)

コメント [TS4]: 1行空け不要 (以下、指摘は省略)

コメント [TS5]: 有業?→融合か?

コメント [TS6]: 用語は統一「資本市場アプローチ」

資本市場の理論であるファイナンスは実証研究での計算の理論的枠組みを提供します。ファイナンス理論により、例えば会計情報をDDMやDCFのモデル式に代入することにより企業価値を算定することが出来ます。企業価値の理論モデルについては第3章で紹介します。

コメント [TS7]: 略語表記は、初出か？
オールソンモデルは落としておいてもよい
か？

統計は実証研究における計算手法を提供してくれます。例えば第5章で扱う回帰分析を行うことにより、会計情報とファイナンス理論とで導かれた企業価値と現実の株価の関係の強さについて恣意性を排除した数字で表すことが出来ます。しかし統計分析のみで判明するのは2つの変数の相関関係のみです。すなわち統計分析によって「AとBは何らかの関係がある」と主張することは出来ますが、統計分析のみでは「Aの結果Bとなる」もしくは「Bの結果Aとなる」といった因果関係を主張することは出来ません。そこで相関関係を因果関係にする論理（基準）を提供するのが先ほど述べたファイナンス理論です。例えば、利益を基に株価が決定されるという理論が存在するとした場合、統計分析によってある市場において利益と株価の高い相関関係が確認できたとすれば、「利益が株価を産む」（「株価が利益を産む」と解釈することは出来ないことに注意してください）という因果関係を市場に見出すことが出来ます。

コメント [TS8]: 企業価値の理論モデル→
企業価値評価の理論モデル

コメント [s9]: 計算手法ではなく検証手法
では？

コメント [TS10]: 日本語がこなれていない。

上記のように、会計とファイナンス、統計は密接に関係しており、実証研究を行う際にはこれら3つを理解することが大切になります。本書を通じてこれらの基本概念を解説していきます。

コメント [TS11]: 実証分析では、3者が密
接に関連している

2.2 効率的市場仮説

Fama[1970]が示した効率的市場とは、入手可能な情報を全て織り込んでおり、価格が新情報に即座に反応する市場のことです。効率的市場では、情報内容を正しく理解して、適切な意思決定を行う合理的な投資家によって構成されている市場を想定しています。

効率的な市場において株価は常に適正株価ということになります。つまり市場が効率的な場合は、情報が公開された瞬間に株価は適正株価へと移行するので、後でその情報を知りその情報を利用して投資を行ったとしても、超過利益を得ることは出来ないこととなります。

コメント [TS12]: 後で、とはいつ？その情
報、とはどういった情報？もう少し的確に。

そしてどのような情報が株価に織り込まれるのかについて、Fama[1970]は市場の効率性を3つに分類しました。

コメント [TS13]: 株価に影響を及ぼすと考
えられている情報が、株価にどのように織
り込まれていくのか、

・ 弱度の効率性（ウィークフォーム）

弱度の効率性とは、市場作られる株価は過去の価格系列を全て織り込んでいるというものです。もし市場が弱度の効率性を有している場合は、過去の株価系列のみを用いて超

コメント [TS14]: 市場作られる？

過利益を得ることは出来ない、ということになります。

- ・ 準強度の効率性 (セミストロングフォーム)

準強度の効率性とは、株価は企業が公表するあらゆる情報 (ディスクロージャー制度による公開情報) を公表後すぐに織り込むるものです。もし市場が準強度の効率性を有している場合、財務諸表を含むあらゆる公開情報の公開後は、それらを用いても超過利益を得られないこととなります。

コメント [TS15]: 織り込むる？

- ・ 強度の効率性 (ストロングフォーム)

強度の効率性とは、公開情報のほかに一般に公表されていない未公開情報をも株価が織り込んでいるとするものです。

コメント [TS16]: 全ての情報を織り込んで
いるということを明確に。

2.3 MM 理論

MM理論とは、Modigliani と Miller という2人の経済学者が提唱した資本構成と企業価値についての理論です。その結論は、「税金・取引費用が存在せず、企業は倒産しないならば、企業価値は資本構成によって変化しない」というものです。この主張はMMの第一命題と呼ばれています。パイ (企業価値) をどのように株主と債権者で分割しても、パイ自体の大きさは変わらないという主張です。

コメント [s17]: 論文を注につけるべきで
は

○ MM理論の前提

- ① 全ての投資家が、企業の将来利益の流列に対して同質の期待を持つ。
- ② 企業の将来利益の期待値は一定である。
- ③ 個人も企業も、安全利子率で無限に借入れ、貸し出しが出来る。
- ④ 資本市場は完全市場であり、取引コストはない。
- ⑤ 税金は存在しな

コメント [TS18]: しな「い」

MMの第一命題によれば「資本構成は企業価値に影響しない」ということなので、もし同質利益クラス、同質危険クラスに分類される2つの企業が存在し、負債比率が高い企業が負債比率が低い企業よりも企業価値が低評価の場合は、投資家はこの企業の株を買い、他方の企業の株を売ることにより裁定機会を得ることができます。裁定機会とはリスクを取らずに確実に収益を得ることのできるチャンスのことです。このような投資行動により2つの企業の株価 (企業価値) は収束していくこととなります。

コメント [TS19]: 日本語表現

第3章 企業価値の評価モデル

この章では、第2章で学んだファイナンスの知識を基に、企業価値を推定する方法を学びます。企業価値の推定には様々なモデルがありますが、本章ではその中でも特に重要なものを扱います。実際にモデルを使用し、企業価値の推定を行えるように理論的なバックグラウンドを理解してください。

3.1. 基本的な考え方

3.1.1 なぜ今企業価値か

昨今、**頃に**企業価値に対する関心が高まってきているように思われます。新聞や経済誌を開くと、「企業価値の向上」という謳い文句を目にしない日はありません。**最早**、「企業の価値を評価する」・「企業価値の向上に**勤める**」ことは**経済に関わるすべての人にとって常識と化している**と言えるでしょう。

しかし、この「企業価値」という言葉は非常に主観的に用いられることも多く、その理解度は決して高いとはいえません。それには2つの理由があると思われま

1つ目として、企業価値を理解するためには**キャッシュフロー**や**割引価値**などといった難解な専門用語が立ちはだかり、皆さんの理解を妨げるからです。

2つ目として、企業価値にはさまざまな捉え方があり、「これが企業価値である」と断定できるものではないことが挙げられます。**なぜなら、企業は多様なステークホルダー(利害関係者)に囲まれているから**です。たとえば**経営者**にとっての価値という捉え方もあれば、**株主**や**債権者**にとっての価値という捉え方もあります。さらに**従業員**・**顧客**・**国**・**地方公共団体**などもステークホルダーと考えることができ、「誰にとっての価値か」を考えなければいけないのです。

特に実証研究を行う上では、企業価値を数値として捉える必要があります。

たとえば、私たちは**企業価値と株価**の関係を統計的に分析するという研究を行いました。研究の始発点として、まずこの場合の**企業価値**とは何なのかを明確に定義し、定義にしたがって**数値化(定量化ともいいます)**する必要がありました。しかし、上述の通り**企業価値にはさまざまな捉え方があり、「誰にとっての価値か」ということを明確に定義せずに研究を進めた為、不要な遠回りを繰り返す羽目になりました。**

コメント [s20]: いま

コメント [TS21]: 読みにくい

コメント [TS22]: 読みにくい

コメント [TS23]: 努める

コメント [TS24]: 表現

コメント [TS25]: キャッシュ・フロー (中黒)

コメント [TS26]: 段落空けは不要 (以下同様; 指摘省略)

コメント [TS27]: 表現

コメント [TS28]: 「研究」という用語を振り回しすぎ?

また、資本市場アプローチでは、「投資家にとっての企業価値」という明確な定義がある。

「企業価値」にはさまざまな観点からの解釈の仕方があるので、分析を行う際には、「誰にとっての、どのような企業価値なのか」という点を明確にしておく必要があるということを描するにとどめればよいのでは?

このように、企業価値という概念や定義は研究を行う上で土台となるため、正確に理解することが不可欠です。次節では企業価値の定義を具体的に考えていきたいと思えます。

コメント [TS29]: 「正確」という表現が少し気になる。会計の実証研究における企業価値の把握に過ぎない。

3.1.2 企業価値の定義

前節で述べたように、企業価値の定義は様々です。そこでこの本では、株主にとっての価値を考えたいと思えます。

コメント [TS30]: 表現

例えば、あなたが企業 A という企業の株主であるとしましょう。あなたが企業 A の株式を所有していることで得られる利益とはなんのでしょうか。それは「ずばり」配当 (dividend) です。毎期企業 A から、所有している株式に見合ったお金が配当金としてやってきます。したがって、あなたにとってのこの企業の価値とは現在から将来に渡って得られる配当の合計なのです。(Capital Gain についてはコラム参照)

コメント [TS31]: ずばり・・・端的にいえば、とか?

コメント [TS32]: キャピタルゲインについて指摘をするなら、配当等の収入を、インカムゲインということはこの段落で指摘しておくべき。

ここでは、配当という観点から企業価値を評価しましたが、最初に述べたように置かれている立場や観点によってどのように企業価値を測定するかは様々です。第 3 章では、幾つかの企業価値評価モデルを紹介しますが、それぞれのモデルがどのような観点で設計されているかをしっかりと確認する必要があります。

3.1.3 キャッシュフロー (CF) とは

コメント [TS33]: キャッシュ・フロー (中黒)
以下、CFの中黒指摘は省略

「100万円で仕入れた商品を120万円の定価をつけて売り、代金は後日回収することにした」

この時、利益は20万円ですが、売上の回収が完了していないため、現時点での手持ち現金は100万円減ったこととなります。

どれだけ売上・利益をあげていても、キャッシュを伴わなければ、企業は存続していきません。ですから、普段の企業活動によって、どれだけキャッシュを生み出す能力があるかを評価するキャッシュ・フローは企業価値を測るにあたって重要な考え方となるのです。

コメント [TS34]:
「ここでいうキャッシュとは」
現金に準ずるものを、会計学では「現金同等物」と呼ぶことも指摘しても良い。

キャッシュ・フローとは直訳すると「現金の流れ」となるように、企業内におけるキャッシュの流れを言います。キャッシュとは現金または、現金に準ずるものです。企業の本業での稼ぎ、設備投資、資産の売却や取得、銀行からの借入れやその返済、配当の支払いなど、さまざまな形でキャッシュの流れは生まれます。企業活動によって手元のキャッシュがどれだけ増減したかを表すのがキャッシュ・フローです。

コメント [TS35]: インフローとアウトフローの差額となる、「正味キャッシュ・フロー」がキャッシュの増減を示すということを明確にしたほうが良い。

3.1.4 割引

「あなたは宝くじで100万円が当たりました。受け取り方法として、今すぐ100万円をもらうか、10年後に100万円をもらうか、選択することが出来ます」

この場合、あなたはどちらを選ぶでしょうか？おそらく大半の方は今すぐ100万円を受け取ることでしょう。それはおそらく無意識のうちに、今すぐ使えるメリット、逆に言えば10年間もらえない(今すぐもらえない)デメリットを感じ取ったからだと思います。これはファイナンスの世界でも正しい考え方です。

それでは今すぐもらえる100万円と10年後もらえる100万円とでは、どの程度の価値の差が生まれるのでしょうか。この問いに答えを与えるのが「現在価値」という考え方です。

まずあなたは現在100万円持っていて、リスクフリーレートが1%だとします。リスクフリーレートとは、利子の支払いが100%約束される国債などの証券の利率のことです。

もし100万を元手に100万円分の一年物国債(1年後に元本・利子の支払いがある国債)を買ったとすると、一年後あなたは100万円×1.01=101万円もらえることとなります。このときファイナンスでは「今ある100万円と1年後の101万円は等価である(同じ価値をもつ)。」と表現します。あなたは何のリスクを負うことなく100万円を一年間で101万円に増やすことが出来るからです。さらにもし10年物の国債を購入した場合、10年後の受取額は100万円×1.01×1.01×…=110万円となるので、「今持っている100万円=1年後の101万円=10年後の110万円」ということとなります(図1)。

図1

コメント [s36]: 割引価値のほうがよい気が

コメント [TS37]: 表現 (レートが1%という表記があるので、なおさら、ここでこのような%を使う表現を織り交ぜるのはあまり好ましくないのでは?)

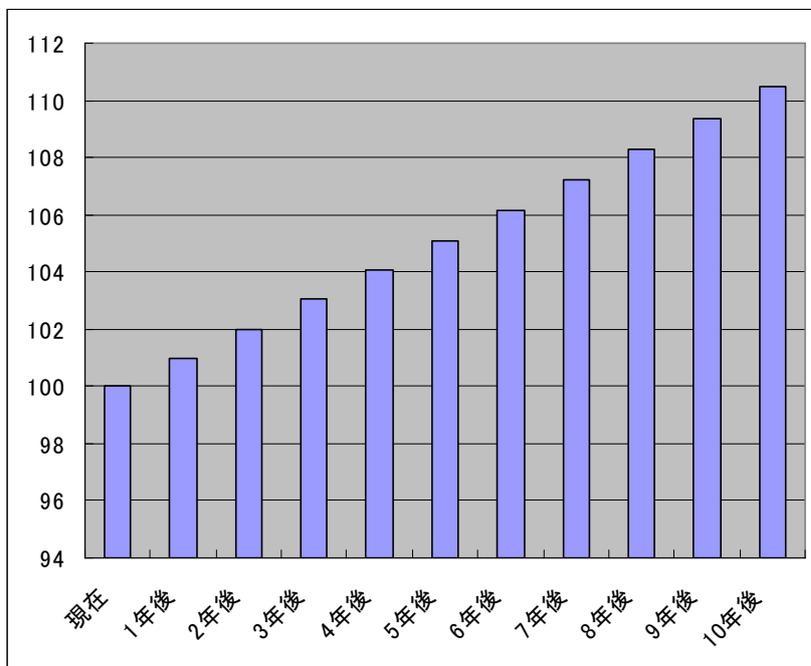
コメント [TS38]: 万円

コメント [TS39]: 100万円×(1.01)^10であると、正確に表現したほうがよいのでは?

コメント [TS40]: 110.4622万円

コメント [TS41]: 110.4622万円

コメント [TS42]: グラフの縦軸の単位



逆に言えばリスクフリーレートが1%の時、2年後の102万円は今すぐもらえる100万円と等価だということになります。計算方法は $102万円 \div 1.01 \div 1.01 = 100万円$ です。このように将来のお金の増減(将来キャッシュフローの増減)を現在の価値に直すことを「割り引く」と言い、将来の価値を現在に割り引くことが「現在価値」の考え方です。

コメント [TS43]: 102.01 万円

コメント [TS44]: 102.01 万円

コメント [TS45]: 不要? 表現

それでは先ほどの問い、今すぐもらえる100万円と10年後もらえる100万円の差を現在価値の考え方を使って算出してみましょう。

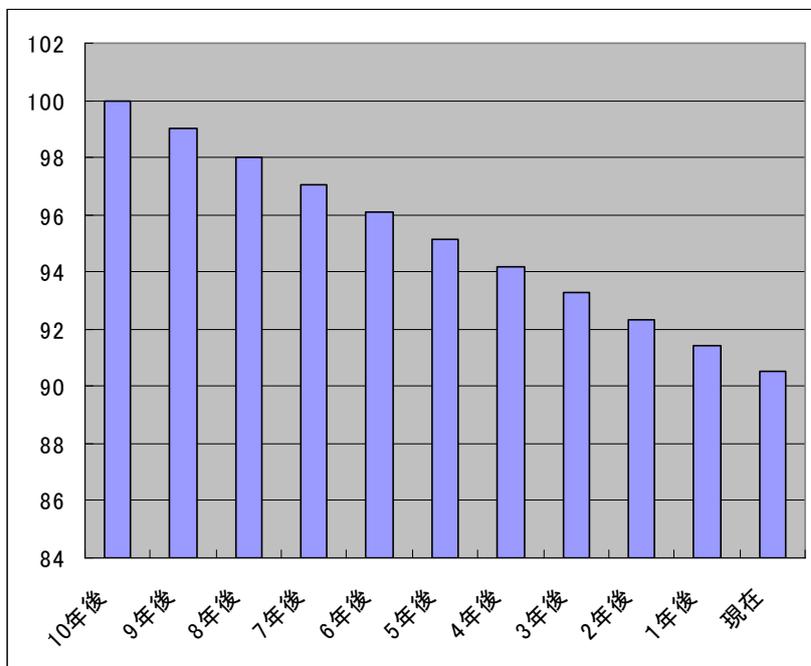
まず今すぐもらえる100万円の価値は現在の価値で言うと100万円ということになります。一方10年後の100万円の現在価値は $100万円 \div (1.01)^{10} = 90.5万円$ ということになります(図2)。約10万円もの差が出てしまいました。このことから現在価値という考え方の重大さが伺えるでしょう。

コメント [TS46]: 90.5287 万円

コメント [TS47]: 表現

図2

コメント [TS48]: グラフの縦軸の単位



3.2 分析モデル

会計学の実証研究を行う上で、企業の価値をどのようにして測定するかということは非常に重要な問題です。

ただし、ここでいう株式の価値とは株式市場で成立している株価とは別のものです。財務諸表などの客観的データから科学的な評価モデルによって算出される価値、すなわち**真実価値**あるいは**適正価値**と呼ばれるものです。

以下では株式の適正価値を測定する為の代表的なモデルとして、①割引キャッシュ・フロー・モデル (DCF)、②**割引配当モデル (DDM)**、③オールソン・モデル (FOM) を紹介します。

3.2.1 DCF 法

DCF (Discounted Cash Flow) 法とは、企業が生み出す価値を CF (キャッシュフロー) であるとし、企業が生み出すであろう将来 CF の割引現在価値の総和を、企業価値とする方法です。

コメント [TS49]: 理論的な価値とも言える。

コメント [TS50]: 配当割引モデルと呼ぶのが一般的では？

桜井のテキスト、ペンマンの訳書では、割引配当モデルと表記されているが、須田、伊藤、大日方、石川、のテキスト・研究著書などでは、配当割引モデルという表記となっている。

ファイナンスのテキストで、割引配当モデルという表記はあまり見ない。(以下、指摘省略)

例えば企業が生み出す将来キャッシュフローが、下図のような場合企業価値 V_e はどのようになるでしょうか。(割引率は 1%とします)

	1 年後	2 年後	3 年後	4 年後	5 年後
CF	100	100	100	100	100

この場合企業価値は、

$$V_e = \frac{100}{(1+0.01)} + \frac{100}{(1+0.01)^2} + \frac{100}{(1+0.01)^3} + \frac{100}{(1+0.01)^4} + \frac{100}{(1+0.01)^5} = 485.3431$$

となります。

ところで、上記のケースでは 5 年後までの CF しか考慮しませんでした、企業は 10 年先、20 年先も存続すると思われま。会計学やファイナンスの世界では、企業は永遠に存続すると仮定しているの、DCF 法を用いるにあたって、CF を無限期間にわたって割り引きます。

しかし、気が遠くなるような将来の CF について完璧な予測を立てることは不可能です、DDM 法のところでも述べたように、実証研究に用いる場合は、適度に予測期間を区切って使用します。

また割引く CF については、FCF(フリーキャッシュフロー)という CF を使用します。FCF は企業の財務情報などの情報より計算して算出するのですが、**難解な概念であるため本書では扱いません。**より詳しく知りたい方は専門書をご覧ください。

コメント [TS51]: FCF は難解？それは言い過ぎではないか。学部学生向けの財務会計、管理会計、財務論等の講義で学習する概念です。別のテキストを指摘するなら、いくつか具体的なテキストを指摘したほうがよいのではないか？

3.2.2 DDM 法 (割引配当モデル)

ある資産の価値を計ろうとする時に有用なのが、その資産が将来どれだけの収益を創出しうるかを計るという方法です。つまり、その資産が将来生み出し得る価値の総和がその資産の価値であるという考え方です。

以上のような考え方に立って、株主にとっての企業 A の価値を考えてみましょう。株主にとっての企業の価値とはすなわち企業が発行している株式の価値と等しいと考えられます。何故なら、株式とは企業の所有権を表わすからです。つまり、企業 A の価値を測定したいならば、企業 A が発行している株式の価値を測れば良いと考えられます。

それでは株式が将来生み出しうる収益とは何でしょうか。もしもあなたが企業 A の株主だとした場合、あなたは毎年度企業 A から配当金を受け取るようになります。この配当金

こそが株式が具体的に生み出しうる収益であり、すなわち株式の価値であると考えることが出来ます。

ただし、ここで測定したいのは株式の現在価値であり、これを求めるには2.4で学んだように将来株式が生み出しうる価値を現在の価値になおさなければなりません。それはあくまで株式が将来生み出す価値であって、株式の現在価値とは異なるからです。

以上の考え方を基に株式の現在価値を式で表わしてみましょう。

企業Aの配当を*i*期の配当を*d_i*と自己資本コスト(株主が株式に対して出資した金額に対して期待するリターンのこと。上記でいう利率に当たる)を*r*とすると、企業Aの株式の価値*V_e*は次の①式のように表わせます。

①式

$$V_e = \frac{d_1}{(1+r)} + \frac{d_2}{(1+r)^2} + \dots + \frac{d_n}{(1+r)^n}$$

このように、現在価値に割引かれた配当の和を求めることによって株式の価値を測定する方法を割引配当モデルと呼びます。このモデルは Discounted Dividends Model の頭文字を取って DDM と呼ばれることもあります。

さて、では上記の割引配当モデルを使って実際に企業Aの価値を算出してみましょう。

上記の①式は理論上の式ということもあり、100年後も1000年後も無限に配当が得られると仮定していますが、このままでは実用には耐えられません。なぜなら、100年後の配当がいくらなのかを予測するのは現実にはほとんど不可能だからです。実際に DDM 法を用いる場合には適当に有限な期間に区切って使用します。

例えば、企業Aの2002年4月時点での株式価値を知りたいと考え、配当データを収集したところ2007年までの配当金のデータが見つかったとします。各年の配当金が毎年100万円、自己資本コストは5パーセントだったとすると、02年時点での企業Aの株式価値は②式のようになります。

図表①

	03年	4年	5年	6年	7年
配当(百万円)	100	100	100	100	100

コメント [TS52]: 式の後に表記したほうが自然

コメント [TS53]: 図表番号の統一は?

コメント [TS54]: 最初が03年なら、以下も04年、05年と表記すべき

②式

$$V_{02} = \frac{100}{(1+0.05)} + \frac{100}{(1+0.05)^2} + \frac{100}{(1+0.05)^3} + \frac{100}{(1+0.05)^4} + \frac{100}{(1+0.05)^5} = 432.9477$$

このように、実証研究においては①式のような理論上の式を、必要に応じて②式のような実用に耐えうる形に改良しなければなりません。以下では①式を理論モデルといい、②式を実証モデルと呼ぶことにします。

3.2.4 オールソンモデル

一般に株主は出資している企業に対して利益を出すことを期待しています。例えば株主が100の資本を出資している企業に対して1%の利益を望んでいる、つまり毎年1の利益を期待しているとします。この企業が6の利益を出したとすると、6-1=5の分は株主の期待を超えて発生した利益です。これを超過利益といいます。

この超過利益を言い換えると、当期純利益から株主資本コスト(株主資本コスト率×期首の資本合計)を控除したものとと言えます。すなわち、企業が株主の期待以上にどれだけの価値を生み出したか、あるいは失ったのかを表しています。

オールソン・モデル(feltham ohlson model)とは、企業の株式価値を期首の資本合計と、将来の超過利益の割引現在価値の合計として評価するモデルです。

例えば、企業Aの期首の資本合計が100、1年後の超過利益が5、割引率が1%だったとします。すると、この時の企業Aの企業価値は、 $100 + \frac{5}{1+0.01} \div 104.95$ となることが分かります。

$$\text{企業価値} = BV_0 + \frac{NI_1 - rBV_0}{1+r} + \frac{NI_2 - rBV_1}{(1+r)^2} + \dots + \frac{NI_n - rBV_{n-1}}{(1+r)^n}$$

BV=資本合計

NI=当期純利益

r=株主資本コスト率

コメント [TS55]: 式の後に表記したほうが自然

コメント [TS56]: ターミナルバリューを無視して、5年分のデータだけで株式価値を求めるというのはいくらなんでも乱暴ではないか？
それならば、無限級数の和として、DDMを運用したほうがより理論的には正しいと考えられます。(V=D/r:ここでの設例では、100/0.05=2000)
また、②式を実証モデルと呼ぶのも少し無理があるのではないのでしょうか？(篠田の個人的見解です。)

コメント [TS57]: 数字に単位はあったほうが良いと思います。十分に分かりやすいですが。

コメント [TS58]: フェルサム=オールソン・モデルについて言及するなら、出典について軽く紹介してはどうか。

コメント [s59]: あとこれをFOモデルとしてよいのかどうか??

割引配当モデルや割引キャッシュフローモデルと違い、貸借対照表から資本合計、損益計算書から当期純利益をそれぞれ使用するモデルであり、財務情報をよく反映したモデルである点が特徴といえます。

それでは、このモデルを実際に用いてみましょう。

実証研究でオールソン・モデルを用いる多くの場合には、予測時点(大多数の企業の決算日が集中する3月末と設定します)を過去にさかのぼり、そこから予測期間分の超過利益の現在価値を足し込んでいきます。また、 r (株主資本コスト率)には、予測時点の10年物長期国債利回りを用いるのが一般的です。

ですから例えば、予測時点を2005年3月31日、予測期間を2年とするモデルは

$$\text{株式価値}_{05} = BV_{05} + \frac{NI_{06} - rBV_{05}}{1+r} + \frac{NI_{07} - rBV_{06}}{(1+r)^2}$$

$BV_n = n$ 年度の資本合計

$NI_n = n$ 年度の当期純利益

$r = 0.0132$ (05年3月10年物長期国債利回り)

と表せるわけです。

(参考)

この超過利益モデルは、割引配当モデルから導き出すことができます。

貸借対照表の期首の資本合計は、当期純利益の獲得によって増加し、配当の支払いによって減少します。これを数式に表すと

$$BV_{t-1} + NI_t - d_t = BV_t$$

となります。つまり

$$d_t = BV_{t-1} + NI_t - BV_t$$

ですから、これを割引配当モデル

$$\text{企業価値} = \frac{d_1}{1+r} + \frac{d_2}{(1+r)^2} + \frac{d_3}{(1+r)^3} + \dots$$

に代入します。

$$\text{企業価値} = \frac{BV_0 + NI_1 - BV_1}{1+r} + \frac{BV_1 + NI_2 - BV_2}{(1+r)^2} + \frac{BV_2 + NI_3 - BV_3}{(1+r)^3} + \dots$$

すると、次のように書き直せます。

コメント [TS60]: 予測期間後の超過利益は、ターミナルバリューとして足し込むものと思いますが。

$$\begin{aligned}
\text{企業価値} &= \frac{BV_0 + NI_1 - BV_1 + rNI_0 - rNI_0}{1+r} + \frac{BV_1 + NI_2 - BV_2 + rNI_1 - rNI_1}{(1+r)^2} \\
&\quad + \frac{BV_2 + NI_3 - BV_3 + rNI_2 - rNI_2}{(1+r)^3} + \dots \\
&= \frac{BV_0(1+r) + NI_1 - rBV_0 - BV_1}{1+r} + \frac{BV_1(1+r) + NI_2 - rBV_1 - BV_2}{(1+r)^2} \\
&\quad + \frac{BV_2(1+r) + NI_3 - rBV_2 - BV_3}{(1+r)^3} + \dots \\
&= BV_0 + \frac{NI_1 - rBV_0}{1+r} - \frac{BV_1}{1+r} + \frac{BV_1}{1+r} + \frac{NI_2 - rBV_1}{(1+r)^2} - \frac{BV_2}{(1+r)^2} + \frac{BV_2}{(1+r)^2} \\
&\quad + \frac{NI_3 - rBV_2}{(1+r)^3} - \frac{BV_3}{(1+r)^3} + \frac{BV_3}{(1+r)^3} + \dots \\
&= BV_0 + \frac{NI_1 - rBV_0}{1+r} + \frac{NI_2 - rBV_1}{(1+r)^2} + \frac{NI_3 - rBV_2}{(1+r)^3} + \dots + \frac{NI_n - rBV_{n-1}}{(1+r)^n}
\end{aligned}$$

3.2.5 CAPM

CAPMとは、ある資産の収益率（注 5.1）と、すべての資産を組み込んだポートフォリオ（注 5.2）の収益率との関係を示すモデルです。

投資できる資産であれば何でも良いのですが、実証研究においては、株式のみに限定して考えることがほとんどです。従って、本章では株式のみに着目して説明します。

株式のみに限定して考えると、CAPM は、ある株式の収益率と、すべての株式を組み込んだポートフォリオ（マーケット、株式市場全体）の収益率の関係を示すモデルとなります。

いくつかの仮定の下で、個別の株式の収益率と、マーケット全体の収益率には、以下のような関係があると言えます。

$$\mu_i - r = \beta (\mu_M - r)$$

ここで、 μ_i はある株式 i の収益率、 μ_M はマーケット全体の収益率、 r はリスクのない資産（注 5. 3）の収益率を表します。 β は株式 i に応じて決まる数値です。

ある資産の収益率から、リスクのない資産の収益率を引いたものをリスクプレミアムといいます。CAPM の式は、ある株式のリスクプレミアムと、マーケットのリスクプレミア

コメント [s61]: これも根拠論文がありません。Sharpe, W.F. [1964] "Capital Asset Prices: A theory of market equilibrium under conditions of risk," *Journal of Finance* No.19, pp.425-442, and Lintner, J., [1965] "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," *Review of Economics and Statistics* No.47, pp.13-37. Treynor unpublished.

ムの関係を表しています。つまり、マーケットのリスクプレミアムが1%変化したとき、ある株式のリスクプレミアムが β %変化する、ということが言えます。

3.1.3 から、 β がちょうど1になるような株式は、マーケットと同じ動きをするような株式といえるでしょう。

同じように、 β が1を越える株式は、マーケットが上昇するときはそれ以上に上昇し、下落するときはそれ以上に下落する、変動の激しい株式といえます。

反対に、 β が1より小さい株式は、マーケットに比べ、変動幅の小さい株式といえます。

このことから分かる通り、 β は株式の変動の相対的な激しさ、つまり相対的なリスクを表す指標として用いることが出来ます。

β は CAPM の式を回帰分析によって推定することで求めることが出来ます。

まず、調べたい株式の株価推移から μ_i を求めます。また、マーケット全体の収益率 μ_M は、(国内株式の場合) TOPIX や日経 225 の収益率で代替するといいいでしょう。また、 r としては(これも国内株式の場合) 10 年物国債の利回りをを用いるのが普通です。このとき、 μ_i 、 μ_M 、 r の適用期間を合わせる必要があります。例えば μ_i に 1 年間の収益率を用いたならば、 μ_M や r も 1 年間の収益率を用いる必要があります。

(ア) β の使い方

3.1.5 のように推定した β は、対象とする株式のリスクについて、投資家たちがどのように評価しているかを表しています。投資家は、株式投資をする際、企業の事業に関わるリスク、財務に関わるリスクを初めとして、企業が抱えるあらゆるリスクをかんがみて投資を行うはずで、したがって、投資家の評価が正しいとすれば、 β を、企業が抱えるすべてのリスクを反映する数値であるとして、実証分析に用いることが出来ます。

注 5.1 収益率

収益率とは、投資から得られる収益を初期投資額(元手)で割って算出した数値です。

注 5.2 ポートフォリオ

ポートフォリオとは、複数の資産を同時に組み合わせて持ったものを言います。

注 5.3 リスクのない資産

リスクとは、将来に対する不確実性のことです。投資した資産の価格が2倍になるかもしれないし、ゼロになるかもしれないという具合に大きく変動する場合、その資産はリスクが高いといえます。逆に、将来の価格の変動が少ない場合、その資産はリスクが低いといえます。そして、まったく将来の変動がない資産のことを、リスクがない資産といえます。たとえば、国債は将来のリターンが保障されていると考えることができ、

コメント [TS62]: たぶん、 β の計算を使用してみれば分かることですが……。日本のデータで、ここで指摘されているように、各指標の期間を統一しようとする、近年のデータにおいて、マーケットの期待収益率が、リスクフリーレートを下回ってしまうことが分かると思います。マーケットの期待収益率については、期間を長めにとるか、リスクフリーレートに一定率を上乗せするかなどの対応をする必要が出てきます。日本のデータでは、CAPMは無効かもしれません。

リスクのない資産とみなすことが多いです。

資産のリスクと収益率には、リスクが高ければ高いほど収益率も高くなるという関係があります。ところが、リスクがゼロであっても、いくらかの収益率は得られます。このリスクのない資産の収益率を、ある基準値として用いることがよくあります。

3.2.6 レリバンス・スタディの基本モデル

レリバンス・スタディ（価値関連性分析）とは、ある会計数値と現実の株価との関係を分析することです。具体的な方法としては、第 5 章で扱う回帰分析を用います。下記回帰式において P に株価、 X に一株あたりの会計数値を代入し、 β が統計的に有意に 0 と異なれば、 X には価値関連性があるということになります。

$$P_i = \alpha + \beta_i X_i + u_i$$

レリバンス・スタディの基本モデルとしてオールソン型回帰モデルがあります。これは株主資本簿価（ BVE ）超過利益（ RI ）を説明変数として用いたモデルで下記のようになります。

$$P_i = \alpha + \beta_i BVE_i + \gamma_i RI_i + u_i$$

また超過利益を当期純利益（ NI ）に置き換えた下記のモデルが使われることもあります。

$$P_i = \alpha + \beta_i BVE_i + \gamma_i NI_i + u_i$$

これらは「資本簿価と利益という、企業にとって最も大切な 2 つの会計情報を用いて株価を説明する」という非常に素朴なモデルであり、感覚的なモデルです。しかし実際に算定した企業価値と実際の株価との相関関係及び株価説明力は高いことが知られています。

レリバンス・スタディの回帰式に会計情報に限らず様々な情報を代入し分析することで、どのような情報が株価に効いているのかを検証してみることが可能です。是非ご自身で研究されてみてください。

コメント [s63]: 大日方先生の本を参照したほうがよい気がします

コメント [s64]: ある会計数値を用いたモデルに基づく企業価値と株価との関係

コメント [TS65]: 素朴かつ感覚的⇨「十分に理論的ではない」ということを指摘すべき

第Ⅱ部 統計学の基礎知識

第4章 統計学を理解するための準備

ここでは、統計学の基礎知識を概観します。

この章の最大の目的は、「検定が行えるようになること」にあります。実証分析など、数値データを用いて議論をする場合には、その主張が「統計的に裏付けられていること（これを、統計的に有意である、といいます）」を示すことが不可欠です。

したがってこの章では「統計的に裏付けられている」とはいかなる場合をいうのということを見て

コメント [TS66]: 表現

ます。
そのために必要な平均と分散、母集団と標本集団、分布、といった知識を1節から3節でカバーします。その後、検定に深くかかわる信頼区間の求め方を提示し、最後に、本章の目的である検定の方法について説明します。

4.1 平均と分散

ここでは、統計学の理解には欠かせない数値である平均と分散について説明します。この2つの数値の役割を説明するために、1つ例を挙げてみましょう。

日本の大学生の英語力を測定するために、ある大学生10人に100点満点のテストを行いました。以下がその結果です。

45、56、72、88、49、67、79、51、90、56

この結果だけを見ても、大学生の英語力について何らかの考察を得ることはできません。そこで、このグループの特徴を表す数値を考えます。この数値のことを**代表値**といいます。平均も分散も、この代表値の仲間なのです。

x_1, x_2, \dots, x_N というデータがあるとき、その平均 μ は、

$$\mu = \frac{1}{N}(x_1 + x_2 + \dots + x_N)$$

で求められます。例えば、前述の大学生のテスト結果の平均は、

$$\mu = \frac{1}{10}(45 + 56 + 72 + 88 + 49 + 67 + 79 + 51 + 90 + 56) = 65.3$$

となります。この平均を見れば、例えば「これらの大学生は100点中65.3点位の英語力を持っている」というような考察を得ることができます。

同様に、 x_1, x_2, \dots, x_N というデータがあるとき、その分散 σ^2 は、

$$\sigma^2 = \frac{1}{N} \left\{ (x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_N - \mu)^2 \right\}$$

で求められます。式は複雑ですが、「各データの平均からの差を二乗して足す」と考えればそれほど難しくありません。例えば、前述の大学生のテスト結果の分散は、

$$\sigma^2 = \frac{1}{10} \left\{ (45 - 65.3)^2 + (56 - 65.3)^2 + \dots + (56 - 65.3)^2 \right\} = 241.61$$

となります。分散は、データの散らばり具合を表す数値です。分散 σ^2 の正の平方根 σ を**標準偏差**といいます。

最後に、**標準化**という作業を説明します。標準化とはデータを、平均が0で、分散が1になるように調整することをいいます。それぞれのデータを、以下のように計算します。

$$z_1 = \frac{x_1 - \mu}{\sigma}, z_2 = \frac{x_2 - \mu}{\sigma}, \dots, z_N = \frac{x_N - \mu}{\sigma}$$

このように計算した z_1, z_2, \dots, z_N というデータは、平均が0で、分散が1になっています。

4.2 母集団と標本集団

「日本の大学生の英語力」を調べるためには、本来全国の大学生全員に英語のテストを行う必要があります。しかし、そんなことは不可能ですよね。そこで実際は、例えば100人の大学生をピックアップしてきてテストを行うのが普通です（注1）。

このように、「全国の大学生全員」のように、本当に調べるべき全体の集団のことを**母集団**といいます。また、「100人の大学生」のように、実際に調査を行う、ピックアップされた集団のことを**標本集団**といいます。

さて、標本集団に対して調査を行って、例えば英語テストの平均が65.3点となったとしましょう。この結果から、「日本の大学生は100点中65.3点の英語力を持っている」という結論を導くことはできると思いますか？…違和感がありますよね。それは、「母集団の平均と標本集団の平均は必ずしも同じとは限らない」からです。以降、この2つを区別するために母集団の平均を**母平均**、標本集団の平均を**標本平均**と呼びます。同様に、母集団の分散を**母分散**、標本集団の分散を**標本分散**と呼びます。

「日本の大学生の英語力」について何か結論を得るために必要なのは母集団の情報です。しかし、手に入るのは標本集団の情報です。このギャップをどう埋めたらいい

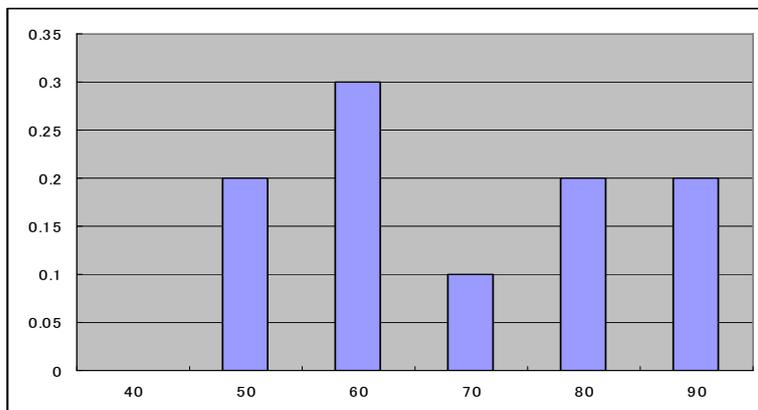
コメント [TS67]: 足したものを標本数で割る

コメント [TS68]: 問いかけるられる表現というのも、可愛らしいですが・・・。

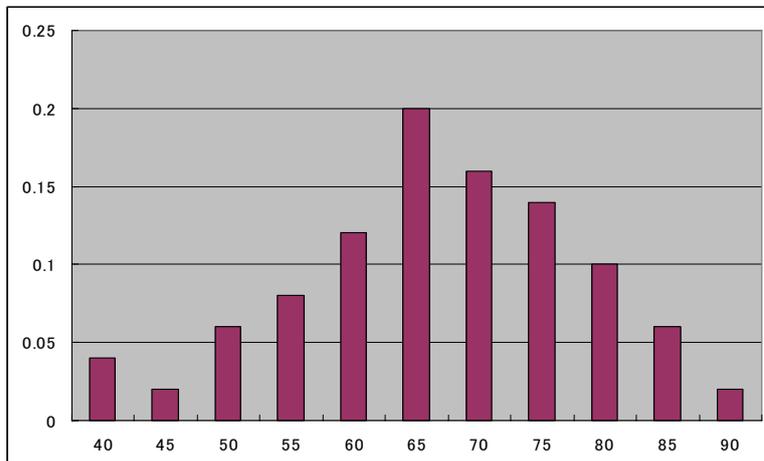
のでしょうか？…ここで、推定・検定という考え方が登場します。実は、標本平均が65.3点と分かっているならば、母平均は大体これぐらいの範囲に入ることができるのです。これが推定・検定の考え方です。

4.3 分布

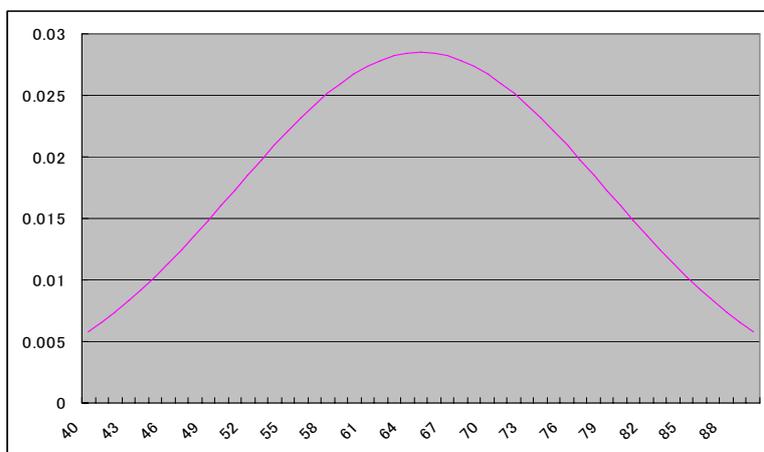
先ほどの英語テストのデータについて、区間ごとにその頻度を示したものをヒストグラムといいます。それをグラフにしたものが下のものです。



また、データをもっと増やして区間を小さくしていくと、下のようなグラフになりました。



このヒストグラムの形は、多くの統計データのヒストグラムに特徴的に現れるもので、平均の頻度をもっとも多く、平均から離れるにつれて次第に頻度が下がっていくようになっています。さらにデータを増やし、区間を小さくすると、下のようなグラフに近づきます。



このグラフは正規分布のグラフと呼ばれます。

あるデータをヒストグラムにするとこのような正規分布のグラフになるとき、そのデータは正規分布に従うといいます。

もちろん、別の形のグラフをもつデータもあります。t分布やカイ2乗分布と呼ばれるグラフをもつものもあり、それぞれt分布に従うデータ、カイ2乗分布に従うデータ

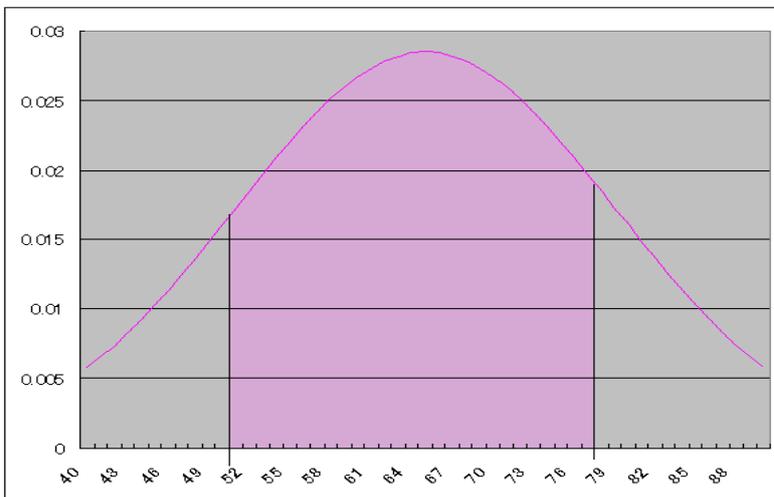
コメント [TS69]: 正規分布のグラフの特徴について言及すべき。

- ・ 平均値を中心に、左右対称な広がりを見せる分布
- ・ 平均からのずれが $\pm\sigma$ 以内の範囲に68.26%の事象が収まり、 $\pm 2\sigma$ 以内の範囲だと95.44%、さらに $\pm 3\sigma$ だと99.74%の事象が収まるような分布。

と呼びます。

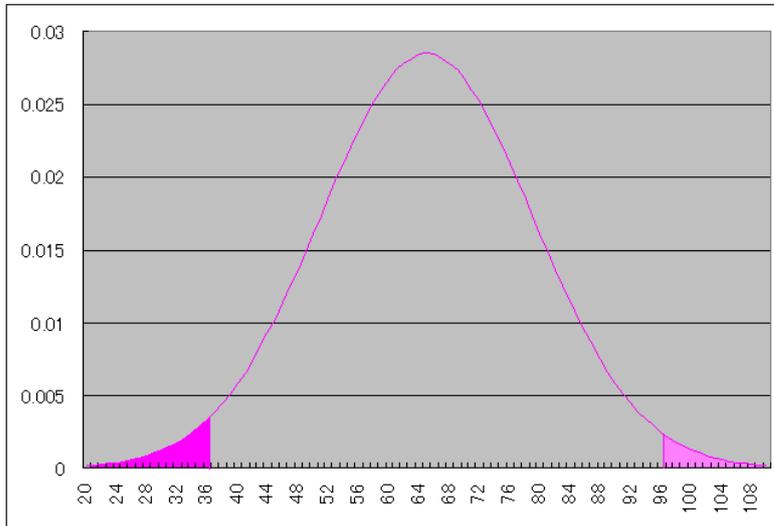
4.4 信頼区間

3で見たような分布のグラフは、非常に便利な性質を持っています。たとえば、52から79までの部分（下のグラフで色を塗った部分）の面積が0.75だったとします。すると、このグラフを持つデータが52から79の間に入る確率が0.75であるといえるのです。

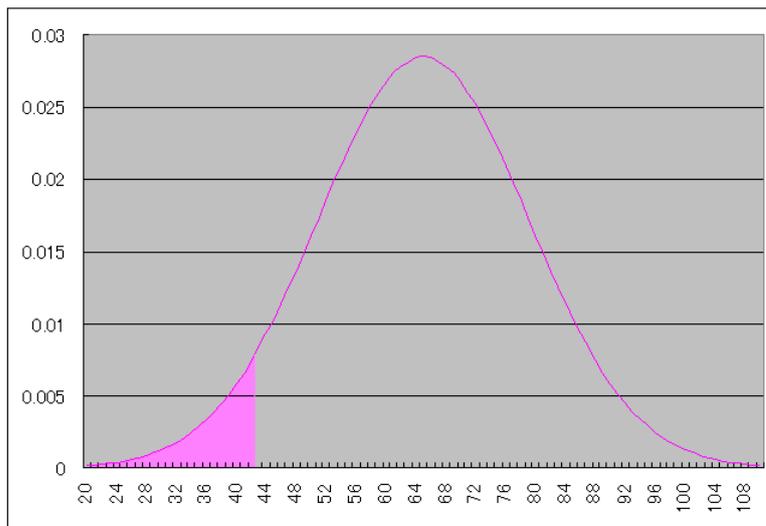


逆に、たとえば95%の確率でデータが入るような区間を求めることもできます。この区間のことを**95%信頼区間**と呼びます。もちろん、99%の確率でデータが入るような区間の場合は**99%信頼区間**と呼びます。また、正規分布だけでなく、t分布やカイ2乗分布でも同様に信頼区間を求めることができます。

95%信頼区間を実際に求める方法を紹介します。一口に95%信頼区間といっても、さまざまなものがあります。ひとつは、両側から2.5%ずつ削ったもので、これを**両側信頼区間**と呼びます。また、左側から5%を削ったものを**左側信頼区間**、右側から5%を削ったものを**右側信頼区間**と呼びます。



(両側 95%信頼区間)



(左側 95%信頼区間)

両側 95%信頼区間の求め方を説明します。

まず、データの 2.5%点を求めます。2.5%点とは、その値より下の値をとる確率が 2.5%であるような点のことです。この点の求め方は後で説明しますが、上の分布の 2.5%点は、37.56 になっています。

同様に、データの 97.5%点を求めます。97.5%点は、その値より上の値をとる確

率が2.5%であるような点のことです。上の分布では、92.44 になっています。

したがって、このデータの両側95%信頼区間は、37.56 から92.44 となります。

左側95%信頼区間については、データの5%点を求めます。上の分布では41.97 になっています。したがって、このデータの左側95%信頼区間は、41.97 から上となります。

○2.5%点や97.5%点の求め方

同じ正規分布でも、平均や分散によって2.5%点や97.5%点は大きく変わりますので、一般にはEXCELを用いる必要があります。しかし、平均が0、分散が1であるような正規分布の2.5%点や97.5%点は巻末の表に載っていることが多いです。

例えば、97.5%点は、巻末の表から1.96 と求めることが出来ます。また、正規分布が左右対称であることを利用すると、2.5%点は、97.5%点にマイナスをつけた、-1.96 となります。

したがって、まず、データを標準化して、平均が0で、分散が1のデータにしてから信頼区間を求めます。例えば、上のグラフに用いたデータですと、平均が65で分散が14の二乗ですので、各データを標準化すると、 $\frac{x_i - 65}{14}$ となります。従って95%信頼区間は、

$$-1.96 < \frac{x_i - 65}{14} < 1.96$$

となります。これを x_i について解くと、

$$37.56 < x_i < 92.44$$

となって、上で求めたような信頼区間を得ることが出来ます。

○標本平均の信頼区間

次の章で説明する検定を行う上で、標本平均の信頼区間を求めることが必要になってきます。ここでは、その方法を先取りして説明します。

平均が μ で分散が σ^2 の正規分布に従っているデータを考えます。このデータから n 個のサンプルを取り出し、標本平均を計算した場合を考えましょう。

サンプルのとり方によって標本平均の値も変わっていきますので、標本平均もなんらかの分布に従っていることがわかります。実は、母集団の分布が正規分布ならば、その標本平均の分布も正規分布に従っているのです。

ここで、標本平均の平均は、母平均と同じく μ になります。標本平均の分散は、母分散をサンプル数で割ったもの、すなわち $\frac{\sigma^2}{n}$ に、標準偏差はその平方根で $\frac{\sigma}{\sqrt{n}}$ になります。

コメント [TS70]: 何の巻末? 統計学のテキストなどと付記すべき

コメント [TS71]: 対照→対称

コメント [TS72]: 標準偏差が14といったほうが自然では?

したがって、標本平均 \bar{x} を標準化すると、 $\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ となり、標本平均の 95% 信頼区間は、

$$-1.96 < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < 1.96$$

となります。

では、実際に数値を代入して考えてみましょう。データは、上のグラフで用いている平均が 65 で、分散が 14 の二乗であるとします。そこから 16 個のサンプルを持ってきたとします。そうすれば、標本平均 \bar{x} の信頼区間は、

$$-1.96 < \frac{\bar{x} - 65}{14/\sqrt{16}} < 1.96$$

なので、

$$58.14 < \bar{x} < 71.86$$

となります。

○母平均の信頼区間

標本平均の信頼区間を求めることが出来た場合、それを利用して母平均の信頼区間を求めることも出来ます。上の例で、母平均がわからず、標本平均は計算できて、64 であると分かったとしましょう。(こちらの状況の方が自然ですが)。すると、母平均 μ の信頼区間は、

$$-1.96 < \frac{64 - \mu}{14/\sqrt{16}} < 1.96$$

なので、

$$57.14 < \mu < 70.86$$

となります。

○t分布

これまでの例では、データは正規分布に従い、母分散がわかっていると仮定していました。しかし、そのようなことは実際はあまり考えられません。「データの分布はわからず、母平均も母分散も分からない」という状況が普通でしょう。このような状況でも、母平均の信頼区間を求めることは可能です。ほとんどの分析はこれ以降の話を利用するので、特に注意して読んでください。

やり方としては簡単です。標本平均を標準化した値 $\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ の σ の代わりに、標本の標準偏差である s を代入すればいいのです。ただし、このようにして出来た $\frac{\bar{x} - \mu}{s/\sqrt{n}}$ という値は、

もはや正規分布には従いません。この値が従う分布は、「自由度が $n-1$ の t 分布」といいます。また、 $\frac{\bar{x}-\mu}{s/\sqrt{n}}$ のことを、t 値と呼びます。

自由度というのは難しい概念なのでここでは説明しませんが、要するにサンプル数によって従う t 分布の種類も違ってくると理解しておいていただければいいと思います。

実際に数値を代入してみましょう。サンプル数が 16、標本平均が 64、標本分散が 16 の二乗であったとします。すると t 値は、 $\frac{64-\mu}{16/\sqrt{16}}$ となります。これは、自由度が 16-1、つまり 15 の t 分布に従います。自由度 15 の t 分布の 97.5% 点は巻末の付表より 2.132 なので、

$$-2.132 < \frac{64-\mu}{16/\sqrt{16}} < 2.132$$

となり、

$$55.472 < \mu < 72.528$$

となることがわかります。

4.5 検定

4.5.1 仮説検定とは

仮説検定とは、ある仮説を立てて、その仮説が間違っていることを統計学的に証明することです。支持したくない仮説を帰無仮説と言い、その帰無仮説を棄却して、採択したい仮説、すなわち対立仮説を支持させるという方法です。

例えば、日本の大学生の平均身長が 170 cm だと一般的に言われているけれども、実際はもっと高いのではないか、という疑問を抱いたとします。この場合、「大学生の平均身長は 170 cm」という仮説を棄却したいわけですから、「大学生の平均身長は 170 cm」が帰無仮説となります。そして「大学生の平均身長は 170 cm よりも高い」という仮説を統計学的に証明したいわけですから、「大学生の平均身長は 170 cm よりも高い」が対立仮説となります。数式を使って整理をすると次のようになります。

$$H_0: \mu = 170$$

$$H_a: \mu > 170$$

H_0 は帰無仮説、 H_a は対立仮説を表しています。

コメント [TS73]: どの巻末の付表か? このテキストにも付表をつけますか?

コメント [s74]: 仮説が妥当ではないことを仮説を棄却することによって証明できるのみであって、必ずしも間違っているかどうかはわからない。

大切なのは帰無仮説や対立仮説をどのように設定するか、ということです。実証研究の場合、帰無仮説は「研究者が否定したい仮説」になります。そして対立仮説は「研究者が考える理論や主張」になります。先ほどの例では、世間では「大学生の平均身長は170cm」と言われているけれど、「大学生の平均身長は170cmよりも高い」に違いないという主張の正しさを証明するために、仮説検定を行っているわけです。仮説検定には研究者の主張が盛り込まれているので、自らの主張と現実世界との整合性を確かめようとする実証研究においては、仮説検定は重要な要素になります。

仮説検定には、母集団の平均を検定するものや分散を検定するものがありますが、本書においては、先ほどの平均身長の場合のように平均の検定を扱います。

4.5.2 仮説検定の方法

検定を行うには、まず、帰無仮説と対立仮説を特定することが必要です。対立仮説の立て方は、片側検定を行うか、両側検定を行うかで2つの方法があります。

片側検定を行うのは、たとえば、「大学生の平均身長は170cmよりも高い」という主張を検証したい場合です。これに対して両側検定を行うのは、たとえば「大学生の平均身長は170cmではない」という主張を検証したい場合です。また、片側検定は右側検定と左側検定の2つに分けることができ、右側検定は母集団の平均が一般に言われているよりも高いことを検証したいときに使用します。一方、左側検定は母集団の平均が一般に言われているよりも低いことを証明したいときに使用します。

ここでは実際に「大学生の平均身長は170cm」を帰無仮説、「大学生の平均身長は170cmよりも高い」を対立仮説として検定を行ってみます。

① サンプリング

実際に確かめたいのは「日本の大学生の平均身長」ですが、日本には莫大な数の大学生がいるので、一人ひとりの身長を調べるには多大な苦勞が伴いますし、物理的・時間的に不可能です。そこで「日本の大学生全体」を母集団として、その母集団から何名かをピックアップして標本平均を求めます。この標本平均を使って「日本の大学生の平均身長」という母平均についての仮説を検定します。

検定や後述する回帰分析等の標本数は多に越したことはありません。可能な限り多くの標本を抽出しましょう。しかし、実証研究においてはデータ上の制約のため限られたサンプルしか無い場合が多く存在します。その際にも標本数は最低でも20を確保しましょう。詳しくはしませんが標本数が少なすぎる場合は検定結果が妥当なものとは言えないからです。

コメント [TS75]: または、低い

コメント [TS76]: 20を確保の真意は？・・・うーん。田中さん、どう思われるでしょうか・・・。

コメント [s77]: まあ、何を検定するかによりますが20はアンケートなどなら…。本当に難しい問題ですね。言及するなら注に落とすか、もっとぼかして20くらいはとか…

コメント [s78]: あと母集団がどの程度になるのかとか、標本は母集団からうまく抽出できているといえるのかなど本当は検討しなくてはならないことも多いですし、ぼかして書くほかないですね。

② 有意水準の決定

仮説検定を行う前に有意水準を決定します。一般的に有意水準は5%や1%を用います。有意水準が5%であるとは、平たく言えば95%以上の確率で帰無仮説は棄却できるということです。1%の場合は99%以上の確率で帰無仮説は棄却できるということになります。従って、より厳密な検定を行いたい場合は1%を採用すると良いでしょう。

残念なことに仮説検定の世界ではこの仮説は「100%否定できる」ということはありません。例えば、ある農場で出来るりんごはどれもだいたい100gだとします。その農場で500gのりんごが生産されることはほとんどありませんが、0%ではありません。0%に限りなく近いかもしれませんが、それでも500gのりんごが生産されるということを「100%否定できない」のです。そのため、仮説検定においても「100%棄却する」のではなく、95%や99%という区切りを設けて仮説を棄却します。

有意水準を決定する際に大切なことは、一度決めた有意水準をむやみに変更してはならないということです。有意水準を恣意的に変更することにより、検定の結果を操作することが可能だからです。例えば、「この仮説を棄却したいんだけど、1%の有意水準では棄却できなかつた。有意水準を5%にすれば棄却できるので変更してしまおう」といった様なことが実証研究においては起こり得ます。誰もが自分の主張の正しさを証明したいのは当然ですが、研究の厳密さを確保するためにも決して歓迎される態度とは言えません。最初に設定した有意水準は守るようにしましょう。

コメント [TS79]: 表現

③ 検定を行う

ここで、前章の、信頼区間を求める場合についての話を思い出してみてください。大学生の身長を例にして、 \bar{X} (標本平均) = 176、s (標本標準偏差) = 13、n (標本数) = 16と仮定すると、t値は、

$$t = \frac{176 - \mu}{13 / \sqrt{16}}$$

となります。この値は、自由度が15のt分布に従います。自由度が15のt分布の5%点は1.645ですから、前章の内容を考えると、95%の確率でt値は1.645よりも小さくならなければなりません。

したがって、帰無仮説で決めた μ の値(170)を代入し、t値が1.645よりも大きくなっ

たとすると、帰無仮説で決めた μ の値は 95%の確率でおかしい、ということになります。このことを、帰無仮説は有意水準 5%で棄却された、といいます。ちなみに、仮説を棄却するかしないかの境界になる 1.645 という数字を、臨界値と呼び、仮説が棄却されるゾーンである「1.645 以上」という範囲を、棄却域と呼びます。

さて、実際に $\mu=170$ を代入してみますと、 t 値は 1.8462 となることがわかります。これは臨界値を上回っているのです、帰無仮説は棄却された、ということになります。

表を用いて臨界値を求めます。その際に、③のステップで決定した検定方法を考慮する必要があります。今回は有意水準を 5%としました。

表より右側確率 5%を与える t の値を見つけます。1.645 となります。したがって棄却域は $t \geq 1.645$ となります。つまり標本平均、標本標準偏差、標本数から t 値を求めて 1.645 を超えれば、めでたく帰無仮説を棄却することが出来ます。よって「大学生の平均身長は 170 cm よりも高い」という仮説は 5%の有意水準によって支持されたわけです。

ちなみに、両側検定をしたい場合についても考えます。自由度が 15 の t 分布の 2.5%点は 2.132 ですから、 t 値は 95%の確率で -2.132 から 2.132 の間に入らなければなりません。したがって棄却域は -2.132 以下、2.132 以上ということになります。先ほど計算した 1.8462 という t 値を思い出しますと、帰無仮説は棄却できない、ということになります。よって「大学生の平均身長は 170 cm ではない」という仮説は 5%の有意水準では支持されませんでした。

さて、今回の例では、片側検定ではうまく帰無仮説が棄却されてくれたのですが、両側検定ではそうはいきませんでした。一般に、片側検定よりも両側検定を使ったほうが帰無仮説は棄却しにくくなっています。ただし、だからといって闇雲に片側検定を用いたらいというわけではありません。片側検定を用いていいのは、たとえば「大学生の平均身長は 170 cm よりも高い」という主張を示したいときに限られるのです。

コメント [TS80]: 表現

4.5.3 研究に際して

実証研究では何回も検定をかけることになると思われますので、逐一、棄却域を求めるのは煩雑極まりません。

そこで検定を行う際の t 値の目安として「2」を常に頭に置いておいて下さい。同じ有意水準であってもサンプル数によって棄却域は左右されますので、極端に少ないサンプルで検定を行う場合を除きますが、ほとんどの場合 t 値が「2」を超えれば右側検定で 5%の有意水準で棄却することができます。左側検定では「-2」が目安となります。是非参考してみてください。

4.6 統計分析の前提

次章より実証研究において有用な統計分析手法である回帰分析を解説していきますが、回帰分析の際にチェックしなければならない事柄が存在します。多重共線性・系列相関・分散不均一です。これらは検定や回帰分析等の統計分析を行う際にクリアしなければならない問題です。

○ 多重共線性

多重共線性とは、後述する重回帰分析の際に、説明変数間の相関が高い場合に起こる現象であり、**回帰分析の結果が歪められてしまうことがあります**。多重共線性が発生すると、決定係数が高い一方で係数の t 値が低く、有意な結果が得られなかったり、推定した回帰係数の符号が理論的に考えられるものと整合しなかったりするなど、望ましくない結果が出てしまいます。多重共線性を排除するには、説明変数間で相関が高い変数のどちらか一方を削るなどします。

○ 系列相関

回帰分析においては、残差が独立して分布していることが前提となりますが、この残差同士に一定の相関が存在することを**系列相関**といいます。系列相関が発生すると回帰分析が有効であるための前提である残差独立性が失われることになります。残差に系列相関が存在するかどうかを確認する方法の一つとして**ダービン・ワトソン比**による方法があります。

○ 分散不均一

回帰分析における誤差項と呼ばれるものについては「独立で、平均が 0 で、分散が σ^2 の正規分布に従う」という前提があるのですが、そのうち分散が σ^2 という一定の値をとらないことを**分散不均一**と呼びます。

なお実際の実証研究の際には、系列相関と分散不均一に関して「母集団に系列相関や分散不均一の問題はない」ということを仮定して実証研究を行うことが一般的です。第〇章の実証研究の実例においてもその仮定を置いています。母集団について何らかの仮定を置く場合は、その前提に立って研究をしているということを忘れないで下さい。

コメント [TS81]: 次の章の中で取り扱ったほうがよいのでは？

コメント [s82]: 統計的分析というより計量的分析（エコノメの意味なので）では？

コメント [TS83]: 表現、この箇所不要。

コメント [TS84]: 脚注などでもよいので、指標の見方だけでも説明するとよいのでは？
エクセルでは、一発で出ないかもしれませんが。

第5章 回帰分析

この章では、身長と体重、営業利益と株価、TOPIX と GDP と鉱工業生産指数など、2つ（もしくはそれ以上）の変数の関係を調べる方法である回帰分析について説明します。

変数の関係を一番簡単に表せるのは、

$$y = ax + b$$

といった一次式です。したがって、この章では統計的な手法を用いてこの式をいかに推定するのかということの説明していきます。

まず、回帰分析について概観した後、2節では簡単のため、2つの変数の関係を考える単回帰に絞って回帰式の求め方を説明します。その後3節で回帰式の推定結果が果たして妥当なのかどうかをどのようにチェックするかを説明します。4節、5節は若干応用の内容で、4節は3変数以上の関係を調べる重回帰分析、5節は「性別」など数字では表せない質的なデータを回帰分析に取り入れる方法であるダミー変数について説明します。

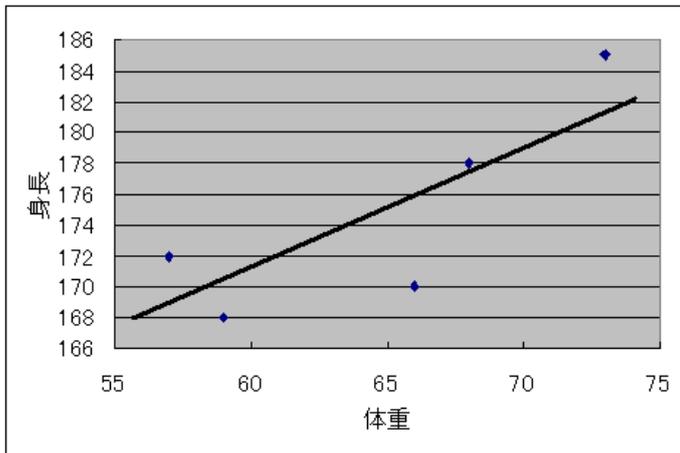
なお、僕たちの研究でもそうでしたが、推定は手計算ではなく EXCEL を用いる場合がほとんどです。むしろ、回帰式の意味を理解した上で、EXCEL の使い方も習得するという姿勢が一番必要だと感じます。したがって本章でも EXCEL の操作を取り混ぜながら回帰分析について説明していきたいと思えます。

5.1 回帰分析とは

回帰分析は、**2つの変数の間の関係を調べる**ことを目的としています。例えば、次に A～E の5人の学生の身長と体重のデータがあります。

	A	B	C	D	E
身長	178	168	172	185	170
体重	68	59	57	73	66

このデータについて、身長と体重の間にどのような関係式が成り立つのかということを知りたいときに、回帰分析を行います。



上の図は、身長と体重を散布図にしたものです。太線で書いたような直線の関係があるように見えてきませんか？この直線がどのような数式で表されるのかを考えるのが回帰分析です。

つまり、身長を y 、体重を x とすると、 y を x の式で表したものを、 $y = \alpha + \beta x$ （これを回帰式とよびます）を求めるのが目的です。

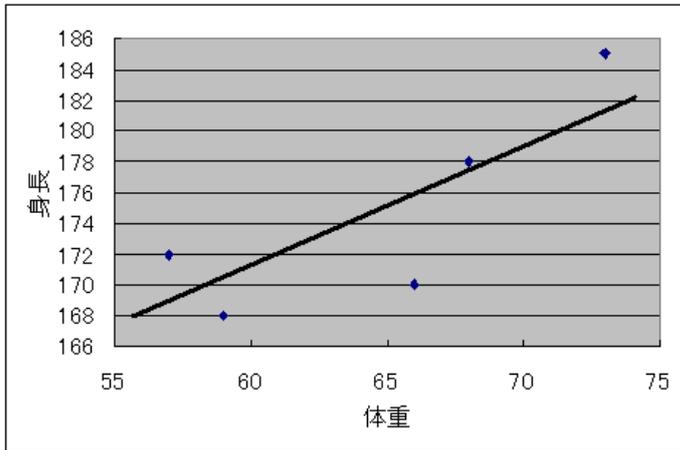
ところで、この式は、「身長を体重の式で表したもの、身長を体重で説明したもの」といえます。このとき、体重を**説明変数**、身長を**非説明変数**と呼びます。これが逆になるとまったく違う分析になるので、注意が必要です。

コメント [TS85]: 非説明→被説明

また、説明変数が1種類であるものを**単回帰分析**、説明変数が2種類以上であるものを**重回帰分析**といいます。例えば、身長を体重と年齢によって説明する分析は、重回帰分析になります。

以降、回帰式の求め方や、その検定の仕方を説明します。まずは、シンプルな単回帰について考えてみましょう。

5.2 回帰式の求め方



では、上のグラフの直線はどのようにして求めればよいでしょうか。

一般的には、最小2乗法という方法を用いて推定します。これは、上の散布図のデータを表す点と直線との距離の二乗和を最小にするように、直線を定める、という方法です。

しかし、最小2乗法を説明するためには、偏微分など少し高度な数学を必要とするので、ここでは、EXCELを用いた推定の仕方と、その結果の解釈の仕方を説明しましょう。

まず、推定の仕方ですが、説明変数、非説明変数をそれぞれセルに入力したあと、ツールの「ツール」→「分析ツール」→「回帰分析」を選択することで、すべて計算することができます。また、「分析ツール」がない場合は、「ツール」→「アドイン」を選択し、「回帰分析」にチェックをしてください。

分析結果は、下のようなのが自動的に計算され、表示されます。

概要									
回帰統計									
重相関 R	0.819881								
重決定 R2	0.672205								
補正 R2	0.56294								
標準誤差	4.570717								
観測数	5								
分散分析表									
	自由度	変動	分散	観測された分散比	有意 F				
回帰	1	128.5256	128.5256	6.152067212	0.089243				
残差	3	62.67436	20.89145						
合計	4	191.2							
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%	
切片	118.9515	22.52879	5.279976	0.013248116	47.25482	190.6482	47.25482	190.6482	
X 値 1	0.861432	0.347304	2.480336	0.089243409	-0.24385	1.96671	-0.24385	1.96671	

さて、この分析結果の中で回帰式は、上の丸の中にある、係数、切片、X 値というところを見れば分かります。この結果だと、回帰式が、

$$y = 118.95 + 0.8614 \times x$$

と推定されたことが分かります。

5.3 回帰分析の妥当性のチェック（注1）

2、で回帰式 $y = \alpha + \beta x$ を求めることができました。しかし、これは（一般的な）身長と体重の関係式を推定した一つの結果にすぎません。なぜなら A～E の 5 組の身長と体重の組合せは、「身長と体重の関係」という母集団の一つのサンプルにすぎないからです。従って、出てきた結果について、それが妥当なのかどうか**検定**を行う必要があります。

コメント [TS86]: 表現：2 節，前節など。

回帰式の検定は色々ありますが、一番代表的なものは、係数 α, β についての t 検定です。先ほども述べたとおり、回帰式は一つの推定結果に過ぎないので、サンプルによって係数 α, β の値は変わってきます。従って、 α, β の分布を考えることができ、それを用いて t 検定を行うことができます。 t 検定のやり方は前章を参照してください。

ただし、係数についての t 検定のやり方はほとんど決まっています。それは、「**帰無仮説を、(係数) = 0 とする**」というものです。係数がゼロになるということは、「説明変数と非説明変数の間にはまったく関係が無い」ということを示します。このような結果になるような分析は無意味です**よね**。従って、このように仮説を立てて、それを棄却することで、回帰式が妥当であることを示そうとするのです。

コメント [TS87]: よね。は、好みの問題ですが、テキストの表現としてはどうでしょうか。

概要									
回帰統計									
重相関 R	0.819881								
重決定 R2	0.672205								
補正 R2	0.56294								
標準誤差	4.570717								
観測数	5								
分散分析表									
	自由度	変動	分散	観測された分散比	有意 F				
回帰	1	128.5256	128.5256	6.152067212	0.089243				
残差	3	62.67436	20.89145						
合計	4	191.2							
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%	
切片	118.9515	22.52879	5.279976	0.013248116	47.25482	190.6482	47.25482	190.6482	
X 値 1	0.861432	0.347304	2.480336	0.089243409	-0.24385	1.96671	-0.24385	1.96671	

帰無仮説の下での t 検定統計量（これを t 値）は、先ほどの分析結果にも表示されています。上の四角の中にあるのがそれです。上段が係数 α についての t 値、下段が係数 β に

についての t 値です。これはいずれも自由度が 3 の t 分布に従っています。自由度は、上の結果の丸の中に示されています。

では、係数 β について実際に検定を行ってみましょう。帰無仮説は、前にも述べたように $\beta = 0$ です。対立仮説は、 $\beta > 0$ とするといいでしょう。なぜなら、体重が増えれば増えるほど身長は一般的には高くなると予測できるからです。有意水準は、ここでは 1% としておきましょう。

自由度が 3 の t 分布の 99% 点は 4.54 です。5.28 > 4.54 ですから、帰無仮説は棄却されます。従って、身長と体重は少なくとも何らかの関係があるということがいえました。

コメント [TS88]: t 分布表をつけたほうが良さそうですね。

もうひとつ、検定結果で注意すべき数値があります。それは R^2 という数値です。これは、説明変数が、非説明変数をどれだけ説明しているか、すなわち回帰式自体の妥当性をチェックするための数値です。

概要								
回帰統計								
重相関 R	0.819881							
重決定 R2	0.672205							
補正 R2	0.56294							
標準誤差	4.570717							
観測数	5							
分散分析表								
	自由度	変動	分散	観測された分散比	有意 F			
回帰	1	128.5256	128.5256	6.152067212	0.089243			
残差	3	62.67436	20.89145					
合計	4	191.2						
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	118.9515	22.52879	5.279976	0.013248116	47.25482	190.6482	47.25482	190.6482
X 値 1	0.861432	0.347304	2.480336	0.089243409	-0.24385	1.96671	-0.24385	1.96671

検定結果では、上の丸の中に示されています（注 2）。

補正 R^2 は 0.56 くらいなので、体重によって、身長の大体 56% くらいが説明できている、ということが出来ます。

補正 R^2 が大体 10% 以上くらいならば、分析として意義のあるもの、といえるようです。逆に補正 R^2 がマイナスになっていれば、それは回帰式自体がおかしいものということの意味しています。

以上をまとめると、「説明変数の妥当性を調べるのが t 検定」、「回帰式全体の妥当性を調べるのが（補正） R^2 」ということになるでしょう。注意すべきなのは、これらのチェックをどちらもクリアして初めて妥当性が示せるということです。

R^2 が十分に高くても t 値が高くなければ説明変数がおかしいことになりすし、 t 値が高くても帰無仮説が棄却されたとしても、（補正） R^2 がマイナスになっていればモデル自体

コメント [TS89]: うーん。これは、少しどうでしょうか。根拠があれば良いのですが。田中さんの意見はどうでしょうか？
会計の分析では、 R^2 はほとんど無視していますし、 R^2 を重視するような予測モデル式の構築を目指す場合などは、通常、0.6 以上は欲しいと考えるものです。
分析目的によって、 R^2 に求められる水準は違いますから、具体的な数字を載せてしまうのは危険だと思います。
統計理論的には R^2 は大切な指標だけれども、会計の実証分析では、 R^2 が非常に低いケースも多く見受けられると言及しておいたほうがよいかもしれません。

がおかしいことになるので、意味のない分析ということになります。

(注1)

ここでは、サンプル数が5という非常に少ない環境で分析をしていますので、 t 値や R^2 の値は高くなっています。普通はサンプル数5という研究はありえず、最低でも20くらいはサンプル数をとっておく必要があります。それに伴って t 値や R^2 の値も低くなってきます。従って、ここで回帰式の妥当性が示されたからといって、身長と体重の関係について何か分かったということではできません。

(注2)

ここで、 R^2 ではなく、補正 R^2 という数値を使っています。実は、説明変数の数を多くしていけばしていくほど、妥当性のいかんにかかわらず R^2 は大きくなっていくのです。そこで、説明変数の数を反映させて R^2 を修正したものが補正 R^2 です。従って、実際の研究では常に補正 R^2 を見てチェックするようにしておくといいでしょう。

コメント [TS90]: プロローグの脚注を参照してもらえばよいのでは？

5.4 重回帰分析

重回帰分析とは、説明変数が2つ以上の回帰分析のことをいいます。その適用例を見るために、次のデータを考えてみましょう。

	A	B	C	D	E	F	G	H	I	J
身長	178	168	172	185	170	165	177	172	166	174
体重	68	59	57	73	66	56	67	58	47	78
胸囲	89	76	78	91	83	75	81	84	74	77
年齢	22	17	28	26	44	33	36	19	15	18

このようなデータについて、身長が体重や胸囲、年齢によってどのように説明できるか調べるのが重回帰分析です。身長を y 、体重を x 、胸囲を z 、年齢を w とすると、推定すべき回帰式は、 $y = \alpha + \beta_1 x + \beta_2 z + \beta_3 w$ となります。では、この回帰式を、EXCELを用いて推定してみましょう。

概要									
回帰統計									
重相関 R	0.910339								
重決定 R2	0.828717								
補正 R2	0.743075								
標準誤差	3.088302								
観測数	10								
分散分析表									
	自由度	変動	分散	観測された分散比	有意 F				
回帰	3	276.8743	92.2914	9.676581535	0.01026				
残差	6	57.22565	9.53761						
合計	9	334.1							
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%	
切片	103.9573	14.19238	7.32487	0.000330745	69.22984	138.6848	69.22984	138.6848	
体重	0.265853	0.134167	1.98151	0.094825788	-0.06244	0.594147	-0.06244	0.594147	
胸囲	0.670782	0.209022	3.20915	0.018386683	0.159325	1.18224	0.159325	1.18224	
年齢	-0.08444	0.110971	-0.761	0.475513803	-0.35598	0.187092	-0.35598	0.187092	

この結果は、回帰式が

$$y = 103.96 + 0.2659x + 0.6708z - 0.0844w$$

と推定できることを示しています。

重回帰分析における回帰式の妥当性も、単回帰のときと同じように検証できます。

例えば、t 検定を考えてみましょう。

体重 x の係数 β_1 について、帰無仮説； $\beta_1 = 0$ 、対立仮説； $\beta_1 > 0$ 、有意水準 5% で検定を行います。

自由度が 6 の t 分布の 5% 点は、2.4469 です。これに対して、t 値は 1.9815 ですので、体重の係数については帰無仮説を棄却することは出来ない、ということになります。

同じように胸囲 z の係数 β_2 について、帰無仮説； $\beta_2 = 0$ 、対立仮説； $\beta_2 > 0$ 、有意水準 5% で検定を行います。

自由度が 6 の t 分布の 5% 点は、前述の通り 2.4469 です。これに対して、t 値は 3.2091 ですので、胸囲の係数については帰無仮説を棄却することが出来、胸囲と身長に少なくとも何らかの関係があることが示されました。

これら単回帰のときと同様な検証に加え、重回帰分析の場合には、F 検定という検定方法があります。これは、変数の係数が、同時に 0 になる場合を帰無仮説とする検定のことです。実際に上の推定結果を用いて考えてみましょう。

体重 x の係数 β_1 、胸囲 z の係数 β_2 、年齢 w の係数 β_3 について、帰無仮説は、「 $\beta_1 = \beta_2 = \beta_3 = 0$ 」であり、対立仮説は、「 β_1 、 β_2 、 β_3 のうち、少なくとも 1 つが 0 でない」となります。

この帰無仮説のもとで、F 値は、結果の「観測された分散比」のところにある 9.6765 という値になります。この F 値は、「回帰」という行の「分散」を、「残差」という行の「分散」で割って求めることができます。また、この F 値は、分子の自由度が、「回帰」という行の「自由度」である 3、分母の自由度が、「残差」という行の「自由度」である 6 の F 分布に従います。

有意水準を 5% としますと、分子の自由度が 3、分母の自由度が 6 である F 分布の 95% 点を巻末付表で調べると、4.76 となります。帰無仮説のもとでの F 値 9.6765 はこれを上回りますので、帰無仮説は棄却され、体重、胸囲、年齢のうち少なくとも 1 つは身長と何らかの関係があることが示されました。

5.5 ダミー変数

ダミー変数は、質的データを扱いたいときに設定する変数のことです。どんなときに使えるのかを考えるために、次のデータセットを見てください。

	A	B	C	D	E	F	G	H	I	J
身長	178	168	172	185	170	145	154	158	150	148
体重	68	59	57	73	66	42	43	51	45	40
性別	男	男	男	男	男	女	女	女	女	女

このデータには、男性と女性それぞれ 5 人ずつの身長と体重があります。ここで、「性別が身長に及ぼす効果」を考えたいとします。しかし、「性別」というのは質的な情報なので、そのまま回帰分析を行うことは出来ません。そこで、仮に男なら 1、女なら 0 という具合に質的な情報を量的な情報に置き換えてみましょう。このように質的データを量的データに変換したものを**ダミー変数**といいます。ダミー変数を用いれば、質的データを含むデータセットの回帰分析をすることが可能になります。

男性を 1、女性を 0 としてダミー変数を設定した上で、回帰分析を行った結果が以下のとおりになります。

概要								
回帰統計								
重相関 R	0.969898							
重決定 R2	0.940702							
補正 R2	0.923759							
標準誤差	3.781005							
観測数	10							
分散分析表								
	自由度	変動	分散	観測された分散比	有意 F			
回帰	2	1587.528	793.764	55.52352242	5.08E-05			
残差	7	100.072	14.296					
合計	9	1687.6						
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	111.4736	10.83159	10.2915	1.76966E-05	85.86098	137.0862	85.86098	137.0862
体重	0.894262	0.242054	3.69447	0.00771096	0.321896	1.466629	0.321896	1.466629
性別	5.357049	5.486463	0.97641	0.361388803	-7.61637	18.33047	-7.61637	18.33047

○で囲んだ「性別」の部分がダミー変数となっています。男性のダミー変数を 1 と設定したので、係数が 5.3570 という正の数であるということは、「男性の方が身長が高い」という意見を支持するものです。しかし、t 値が低いため、検定で帰無仮説を棄却することは出来ず、統計的に支持された結果ということはありませんでした。

コメント [TS91]: たしかにそうですが、ダミー変数の導入に意味があったケースを紹介したほうがよいのではないのでしょうか。データセットを少し変更するだけでできるはずです。

第6章 平均差の検定

この章では平均差の検定について扱います。平均差の検定とは、あるパラメーターについて2つグループ間で差があるかどうかを検定する手法です。会計の実証研究においては頻出するツールですので、確実な理解が必要になります。

6.1 平均差の検定とは

A 大学と B 大学という二つの大学を想定し、それぞれ 51 人選び身長を調査したところ、

A 大学：標本平均 175 c m、標本標準偏差 13

B 大学：標本平均 170 c m、標本標準偏差 5

という結果が得られました。このとき2つの大学の大学生に身長の違いはあるのでしょうか。このように「2つのグループ間の母平均に差があるかどうか」を検証するときに使用するのが**平均差の検定**です。

6.2 仮説の設定

ここでは「A 大学のほうが B 大学よりも平均身長が高い」という仮説を検証してみます。上記の仮説が支持される場合、

A 大学の平均身長 > B 大学の平均身長

このとき、

A 大学の平均身長 - B 大学の平均身長 > 0

となるので、ここでの帰無仮説と対立仮説は

帰無仮説： $\mu_1 - \mu_2 = 0$

対立仮説： $\mu_1 - \mu_2 > 0$

となります。（ここで $\mu_1 = A$ 大学の身長の母平均、 $\mu_2 = B$ 大学の身長の母平均です。）

対立仮説が $\mu_1 - \mu_2 \neq 0$ ではなく、 $\mu_1 - \mu_2 > 0$ であることに注目してください。今回のケースでは、「A 大学と B 大学の平均身長に差がある」ということだけでなく、「A 大学のほうが平均身長は**大きい**」という方向性も指示したいので、不等号を使います。よって今回の検定は片側検定（右側検定）となります。

コメント [TS92]: 高い

6.3 検定の手順

具体的な手順は下記の通りです。

- ① サンプルング及び有意水準の決定
- ② **F** 検定を行う
- ③ データを標準化
- ④ 信頼区間を求める。
- ⑤ 検定統計量を計算し、検定を行う。

① サンプルング及び有意水準の決定

今回標本数はA大学・B大学それぞれ51、有意水準は5%とします。仮説検定の箇所でも述べましたが、標本数は多いほうが望ましいです。しかしデータの制約上標本数が限られる場合は、最低でも両母集団それぞれ標本数20を確保してください。つまり今回の例ではA大学20以上、B大学20以上の標本を用いて検定する必要があります。

② **F** 検定を行う

平均差の検定を行う前に、2つの標本集団の分散について等分散を仮定するのか、あるいは等分散ではないとする仮定をおくのかについて決定しなければなりません。この仮定の決定には**F検定**という手法を用います。

F 検定は分散が等しいかどうかの検定ですので、帰無仮説と対立仮説は以下のようになります。

帰無仮説： $\sigma_A^2 = \sigma_B^2$

対立仮説： $\sigma_A^2 \neq \sigma_B^2$

実際の計算については第10章でEXCELを用いて説明します。EXCELで平均差の検定を行う際には、この検定結果によって用いる分析ツールが異なることに注意してください。

③ データを標準化する

検定統計量を求めるために標本平均の差と標本平均の差の分散を求めます。標本平均の差及び分散は次の式で求めます。

標本平均の差： $\bar{X} = \bar{X}_1 - \bar{X}_2$ 標本平均の差の分散： $V(\bar{X}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

コメント [TS93]: 添字が、A, Bから1, 2となってしまう。

そして帰無仮説における標準化した検定等計量は次の式で与えられます。

$$\frac{\bar{X}}{\sqrt{V(\bar{X})}}$$

コメント [TS94]: 検定統計量Tは・・・

③ 信頼区間を求める

平均差の検定では自由度 $(n_1 + n_2 - 2)$ の t 分布を使用します。よって自由度は $51+51-2=100$ となり、 t 分布表より臨界値は 1.66 となります。よってこの場合の 95% 信頼区間は

コメント [TS95]: 添字は, A, B

$$\frac{\bar{X}}{\sqrt{V(\bar{X})}} < 1.66$$

となります。

④ 検定統計量を計算し、検定を行う。

例のデータを使って、実際に検定統計量を計算してみましょう。

$$\bar{X} = \bar{X}_1 - \bar{X}_2 = 175 - 170 = 5$$

$$V(\bar{X}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{13^2}{51} + \frac{5^2}{51} = 3.803921$$

コメント [TS96]: 添字, 以下, 指摘省略

上記で求めた計算結果を代入すると、

$$\frac{\bar{X}}{\sqrt{V(\bar{X})}} = \frac{5}{\sqrt{3.803921}} = \frac{5}{1.950364} = 2.563624$$

検定等計量 $= 2.563624 > 1.66 =$ 臨界値なので棄却域に入っています。よって「A 大学のほうが B 大学よりも平均身長が高い」という仮説は支持されました。

第 9.10 章の実証研究の実例でもあるように、この平均差の検定は実証分析において強力なツールとなります。平均差の検定を使用する際は検定結果もちろんですが、「何を検証したいのか」という問題意識を明確にすることが大切です。

第7章 χ^2 検定

この章では χ^2 検定（カイ2ジョウケンテイ）を扱います。 χ^2 検定は2つの変数が独立であるかどうかを調べる為に行います。理論上得られるはずの結果と実際に得られた結果を比較することで、検定を行います。ここでいう独立であるというのは、2つの変数の間に確率論的な関連性がないということです。つまり、片方の変数の値が分かっても、もう片方の変数の分布が変化しないということです。

7.1 検定の方法

n 個の観測結果が k 個のカテゴリに分かれているとしましょう。この各カテゴリの観測度数がそれぞれ、 $f_1, f_2, f_3, f_4 \dots f_k$ であり、各カテゴリの理論上の確率が $p_1, p_2, p_3, p_4 \dots p_k$ であるとする、期待度数 $e_1, e_2, e_3, e_4 \dots e_k$ は、 $e_i = n \cdot p_i$ ($i=1, 2, 3, 4 \dots k$) となります。

この時、

$$\chi^2 = \sum (f_i - e_i)^2 / e_i \text{ で、} \chi^2 \text{ を定義します。}$$

χ^2 は、 χ^2 分布という分布に従います。 χ^2 分布の自由度は、「理論が正しくない場合に推定する必要のあるパラメーターの数」と「理論が正しい場合に推定する必要のあるパラメーターの数」の差で定義されます。

この自由度と設定した有意水準を基に分布表から χ^2_a を得て、 $\chi^2 > \chi^2_a$ となった場合、「理論が正しい」という帰無仮説を棄却し、それ以外の場合は帰無仮説を採択します。

7.2 具体的な検証拠

具体的な事例で見てください。

あるコインを投げたときに表が出るという事象と裏が出るという事象が独立であるかどうかを検定したいとします。

H_0 ; 表が出るという事象と裏が出るという事象は互いに独立である

コメント [TS97]: 少し誤解を受けやすい表現では？

a は有意水準であることを明示すべき。

$\chi^2(a)$ は、有意水準が a における χ^2 分布のパーセント点を意味していることをきちんと説明すべきでしょう。

H_1 ; 表が出るという事象と裏が出るという事象には何らかの関係性がある

コインを 100 回投げて、表が出た回数と裏が出た回数を数えてみたしましょう。
その結果が以下の表です。

	表	裏
観測度数	58	42
期待度数	50	50
両度数の差	8	-8

観測度数は表が 58、裏が 42 です。表が出る確率と裏が出る確率が独立だとすれば、表が 50 回・裏が 50 回でなくてはなりませんから、期待度数は 50 です。

さて、この時の χ^2 統計量を求めてみましょう。

$$\chi^2 = (58-50)^2/50 + (42-50)^2/50 = 2.56$$

となります。

次に自由度を求めます。

まず「理論が正しい場合に推定する必要があるパラメータの数」ですが、表が出る確率と裏が出る確率が独立であれば、コインの表が出る確率も裏が出る確率も 1/2 と決まっていますので、推定する必要があるパラメータはありません。従って 0 個です。

次に「理論が正しくない場合に推定する必要があるパラメーターの数」を求めます。表が出る確率と裏が出る確率が独立でなければ、表が出る確率と裏が出る確率という 2 つパラメーターを求める必要があります。しかし、表が出る確率+裏が出る確率=1 です。従って、1 つのパラメーターが推定できれば、自ずともう片方のパラメーターも決まっていきます。従って、「理論が正しくない場合に推定する必要があるパラメーターの数」=1 です。

以上から、 $1 - 0 = 1$ となり、自由度は 1 であると分かります。

自由度が 1・有意水準が 5% の時、 χ^2_a を χ^2 分布表から求めると $\chi^2_a = 3.841459$ となります。

この時、 $\chi^2 < \chi^2_a$ ですから、帰無仮説 H_0 が採択され、コインの表が出る事象と裏が出る事象の独立性が確認されました。

コメント [TS98]: 表記のゆれ: パラメーターか、パラメータか。

第Ⅲ部 Excel の実習ガイダンス

第 8 章 Excel を使った統計分析

この章では、第 4 章から第 7 章で学んだ統計分析手法をコンピュータを使って実践する方法を解説します。Microsoft 社の Microsoft Office Excel という表計算ソフトにデータを入力して処理するのですが、データの処理法を学ぶだけでなく出力された数値データを正しく解釈する為にも、第 4 章から第 7 章と併読して下さい。

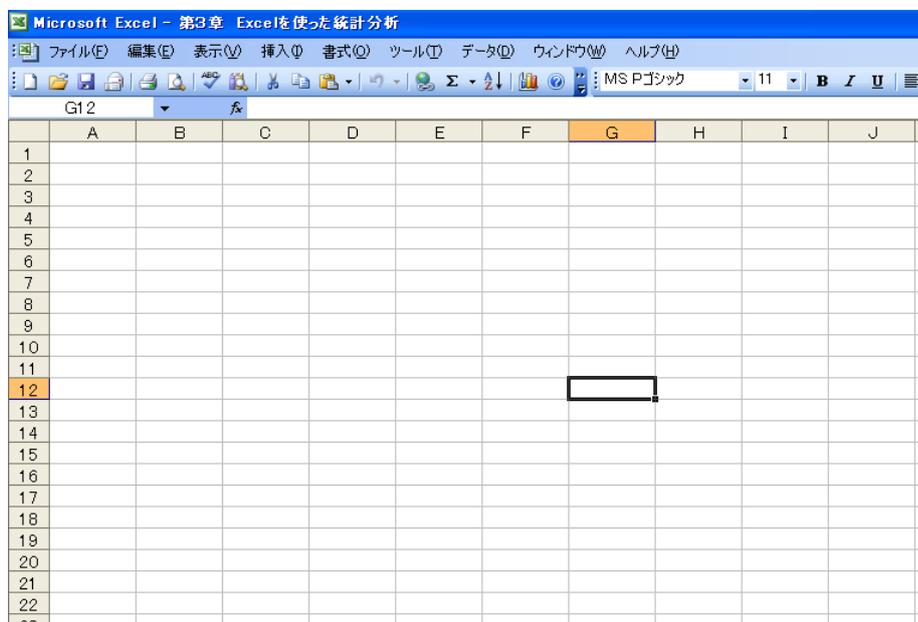
8.1 基本的な計算の方法

本章では、Microsoft 社の Microsoft Office Excel (以下では Excel と表記します) という表計算ソフトを使用します。Microsoft 社のパソコンを始め、多くのパソコンに標準装備されているソフトなので、目にしたことがあるかと思います。



これは windows のスタートコマンドを開いた状態で、楕円で囲まれているのが Excel のアイコンです。まずはこのアイコンをクリックして、Excel を開いてみましょう。

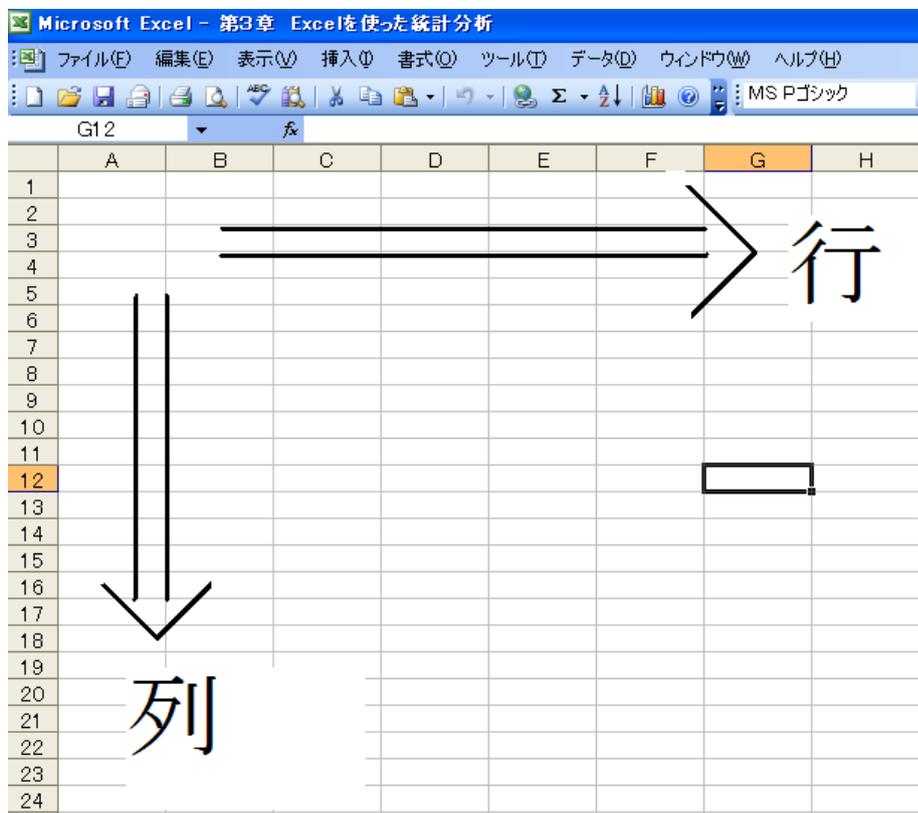
尚、本章では Excel のバージョン 2003 を使用していますが、他のバージョンに関しても同様の操作でフォローできます。



Excel を起動すると、まずこのような画面が開きます。これが Excel の基本画面です。

画面がいくつもの長方形に区切られているのが分かるかと思いますが、この長方形をセルと呼びます。セルは Excel における基本単位で、データはこのセルに入力します。

また、セルの縦の並び（アルファベットから始まるかたまり）のことを列、横の並び（数字から始まるかたまり）のことを行と呼びますので覚えておいてください。



さて、次に Excel を用いて四則計算を行ってみましょう。
 黒い四角で囲まれているセルが現在指定されているセルです。セルが指定されている状態で数値を入力を入力すると、Excel 上で反映されます。

	A	B	C	D
1				
2				
3		100	100	
4				
5				
6				
7				
8				

これは B3 と C4 というセルに 100 という数値を入力した状態です。
 次にこれらを足し算してみましょう。

	A	B	C	D
1				
2				
3		100	100	=
4				
5				
6				
7				

まず D3 のセルに「=」を入力します。

	A	B	C	D
1				
2				
3		100	100	=B3
4				
5				
6				
7				

次に入力した直後に B3 のセルを指定しましょう。すると上のように D3 のセルに、「B3」と表示されます。

	A	B	C	D
1				
2				
3		100	100	=B3+C3
4				
5				
6				
7				

「+ (プラス)」を入力し、C3 のセルを指定します。すると D3 のセルに「B3+C3」という式が表示されます

	A	B	C	D
1				
2				
3		100	100	200
4				
5				
6				
7				

さらに **Enter** キーを押すと「**B3+C3**」が計算され、上のように答えが表示されます。

以上が Excel 上での加法のやり方です。他の四則計算についてもほとんど同様のやり方で計算でき、引き算の時は「**-** (マイナス)」・掛け算の時には「***** (アスタリスク)」・割り算の時には「**/** (スラッシュ)」を使用します。

	A	B	C	D
1				
2				
3		100	100	=B3-C3
4				
5				

	A	B	C	D
1				
2				
3		100	100	=B3*C3
4				
5				

	A	B	C	D
1				
2				
3		100	100	=B3/C3
4				
5				

以上が Excel を使った四則計算の方法です。

Excel にはこのほかにも様々な機能があり、グラフや図表の作成などもできますが、それ

はここでの目的とは異なるので、興味のある方は他の書籍を参照してください。

コメント [TS99]: 具体的にテキストを挙げておいたほうがよいかどうか？

8.2 Excel を用いた回帰分析

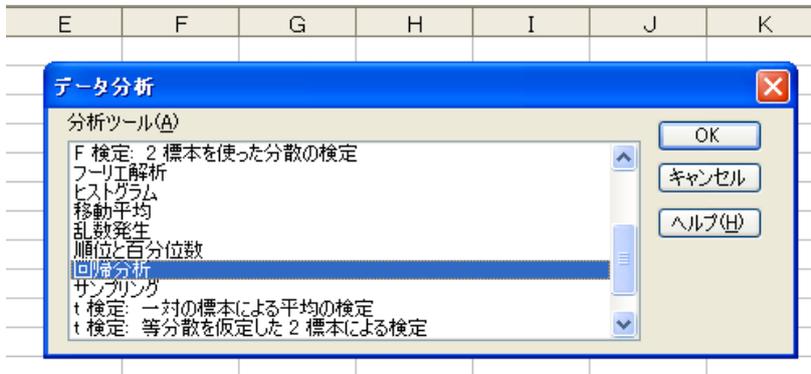
次に回帰分析を行ってみましょう。

第5章でも学びましたが、回帰分析とはある変数 A が別の変数 B をどれだけ説明できるかを測る統計手法です。

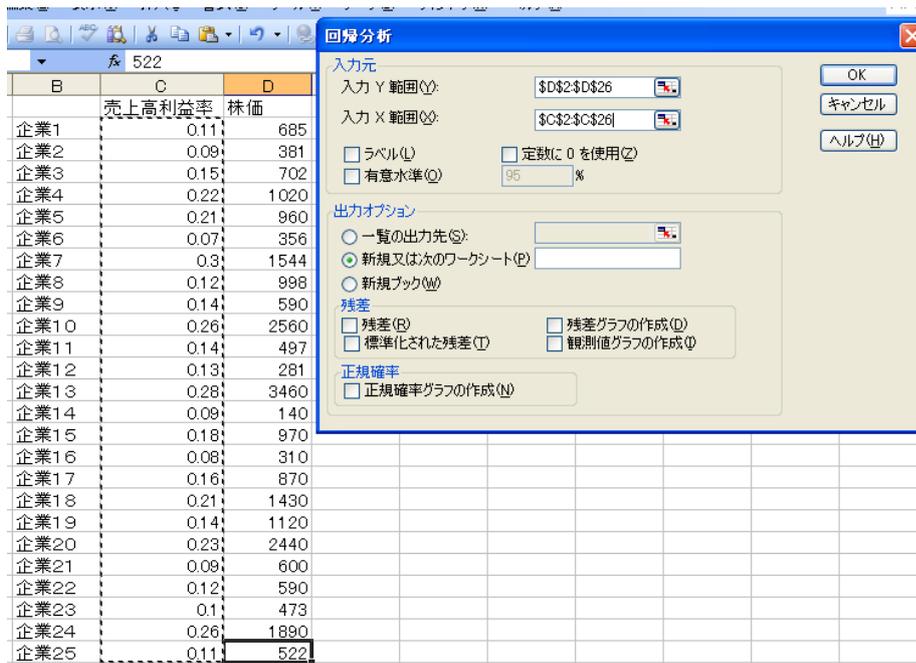
ここでは、ある年の3月31日時点での企業の株価を被説明変数、企業の売上高利益率を説明変数として回帰分析を行ってみましょう。

	A	B	C	D
1			売上高利益率	株価
2		企業1	0.11	685
3		企業2	0.09	381
4		企業3	0.15	702
5		企業4	0.22	1020
6		企業5	0.21	960
7		企業6	0.07	356
8		企業7	0.3	1544
9		企業8	0.12	998
10		企業9	0.14	590
11		企業10	0.26	2560
12		企業11	0.14	497
13		企業12	0.13	281
14		企業13	0.28	3460
15		企業14	0.09	140
16		企業15	0.18	970
17		企業16	0.08	310
18		企業17	0.16	870
19		企業18	0.21	1430
20		企業19	0.14	1120
21		企業20	0.23	2440
22		企業21	0.09	600
23		企業22	0.12	590
24		企業23	0.1	473
25		企業24	0.26	1890
26		企業25	0.11	522
27				

上がサンプルのデータです（このデータは架空のものです）。



ツールバーから「分析ツール」をクリックして、次の画面で回帰分析を指定し OK をクリックします。



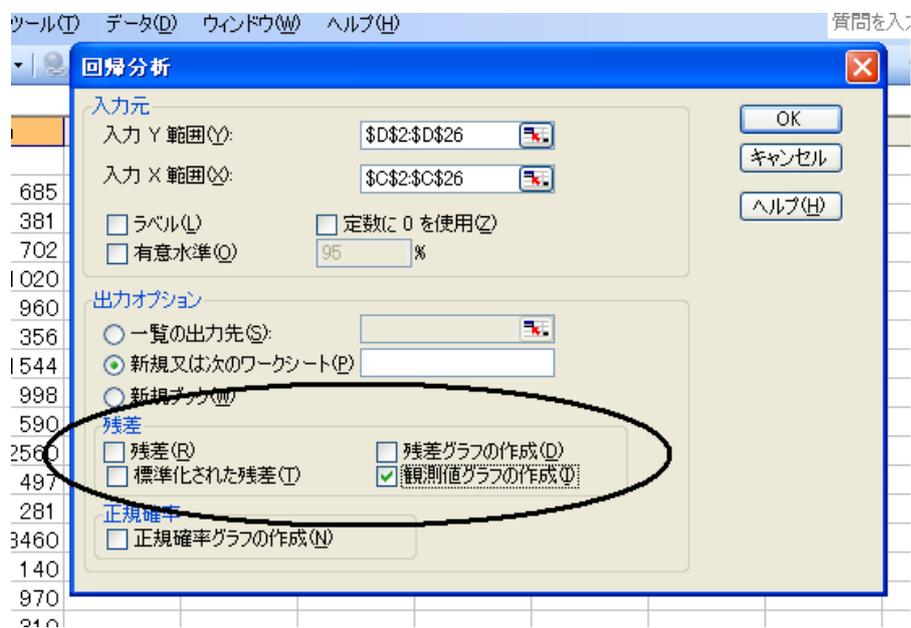
入力 Y 範囲の欄に被説明変数（株価）を、入力 X 範囲の欄に説明変数（売上高利益率）を入力し、OK をクリックすると計算結果が出力されます。

	A	B	C	D	E	F	G	H	I
1	概要								
2									
3	回帰統計								
4	重相関 R	0.837112							
5	重決定 R2	0.700757							
6	補正 R2	0.687747							
7	標準誤差	453.8808							
8	観測数	25							
9									
10	分散分析表								
11		自由度	変動	分散	割された分散	有意 F			
12	回帰	1	11095716	11095716	53.86066	1.82E-07			
13	残差	23	4738180	206007.8					
14	合計	24	15833896						
15									
16		係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
17	切片	-579.442	235.529	-2.46017	0.021824	-1066.67	-92.2136	-1066.67	-92.2136
18	X 値 1	9993.75	1361.735	7.338982	1.82E-07	7176.786	12810.71	7176.786	12810.71
19									

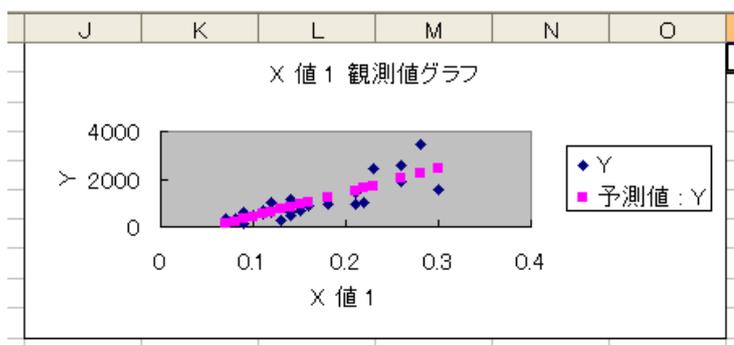
これが出力された計算結果です

まず、見るべきは重決定 R2 と補正 R2 です。これらの値が高いほど、説明変数の説明力が高いといえます。サンプルの補正 R2 は 0.687747 となっていますから、このサンプルの場合、売上高利益率は株価の 68% を説明しているといえます。

次に見るべきは X 値の t 値です。この値が高いほど説明力の安定度が高いことを表わします。t 値が 2 以上あれば十分安定していると言えますが、このサンプルの t 値は 7.338982 となっており、非常に説明力が安定しているといえます。



また、データを入力する際に楕円の「残差」という部分をチェックしておく



上のように、回帰式のグラフが出力されますので、視覚的にも理解の助けになります。

重回帰分析

これまでに扱った回帰分析は、1つの被説明変数に対して1つの説明変数を用いて分析

する単回帰でしたが、以下では1つの被説明変数に対して複数の説明変数を用いる重回帰分析の方法について説明していきたいと思います。

重回帰分析の方法は基本的には単回帰の方法と大きく変わりません。

以上が Excel での回帰分析の方法です。

8.3 Excel を用いた F 検定

それでは、本格的な統計分析を行っていきましょう。

まずは F 検定を行っていききたいと思います。

第 6 章で学んだとおり、F 検定とは 2 つの郡のデータが等分散かどうかを検定する手法です。

ここでは、便宜的にある学校の 3 年 A 組に属する 18 人の身長と 3 年 B 組に属する 20 人の身長というサンプルを仮定して F 検定を行っていきましょう。

コメント [TS100]: ???

説明します→大きく変わりませんか？

それでもいいですが・・・。

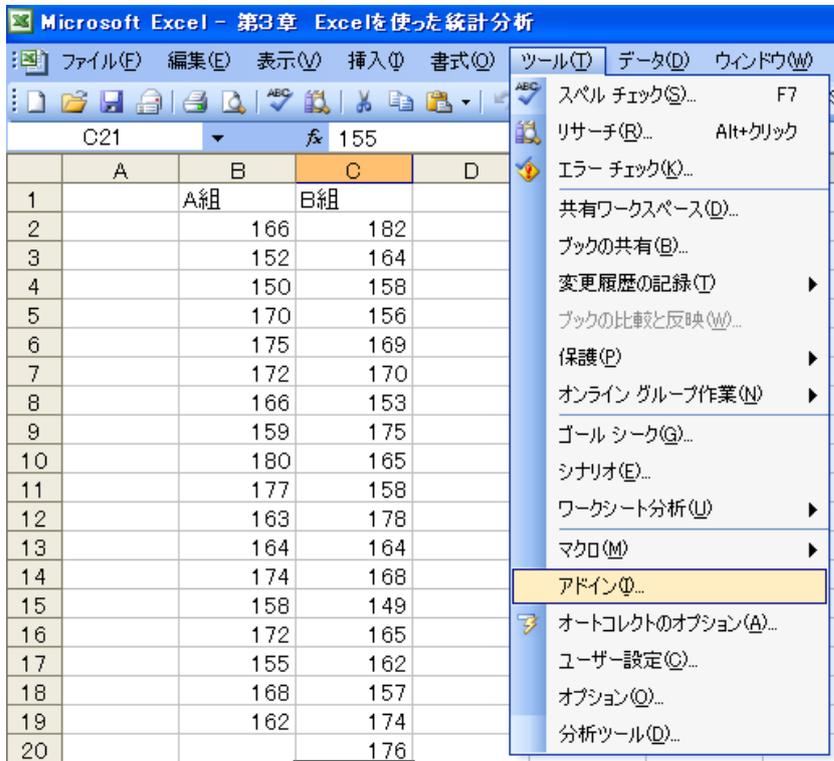
コメント [TS101]: 郡→群

	A	B	C
1		A組	B組
2		166	182
3		152	164
4		150	158
5		170	156
6		175	169
7		172	170
8		166	153
9		159	175
10		180	165
11		177	158
12		163	178
13		164	164
14		174	168
15		158	149
16		172	165
17		155	162
18		168	157
19		162	174
20			176
21			155
22			

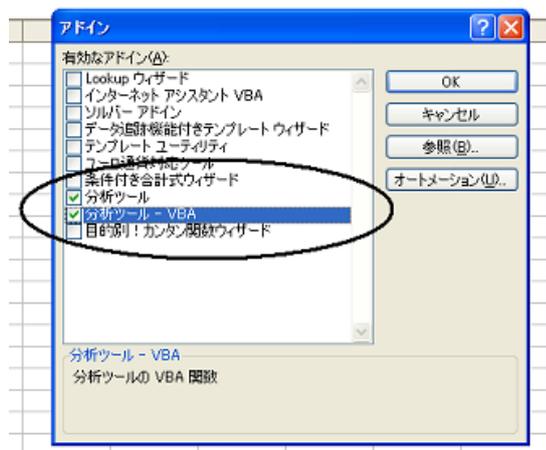
上のようなサンプルについて、以下のような仮説を立て F 検定をかけてみましょう。

H0 : A 組と B 組の身長分散に差がない

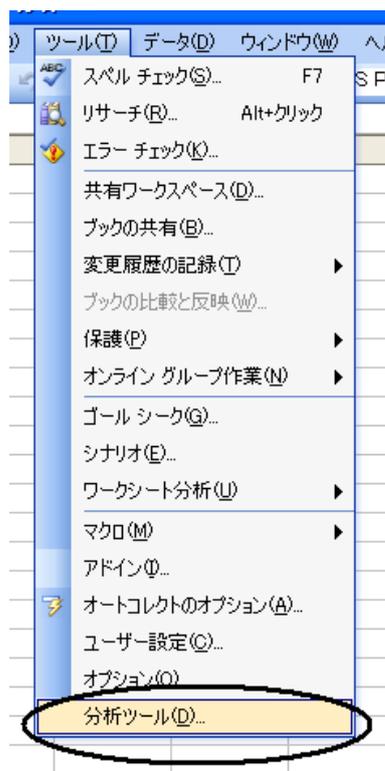
H1 : A 組と B 組の分散には差がある



まず、「ツール」にカーソルを合わせるバーが表示されますので、その中の「アドイン」という項目をクリックしてください。

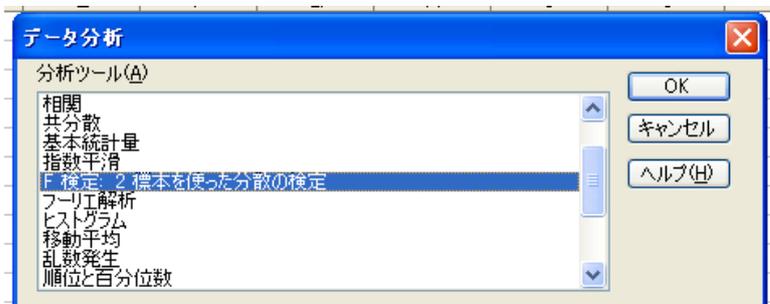


アドインをクリックしますと上のような画面が表示されます。ここに「分析ツール」と「分析ツール VBA」という項目が表示されますので、これらをチェックして OK をクリックしてください。

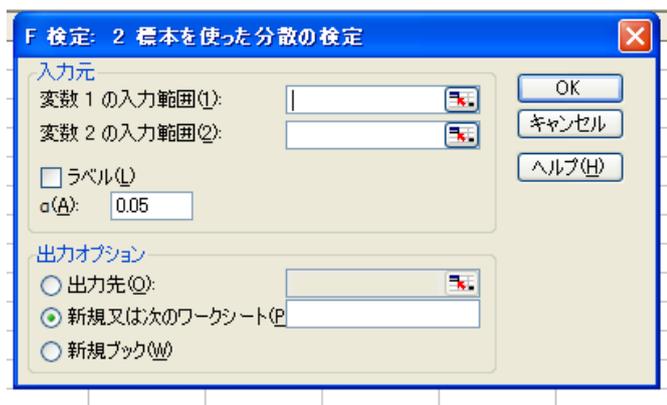


すると、ツールバーに「分析ツール」という項目が追加されます。この分析ツールをクリックしてください。

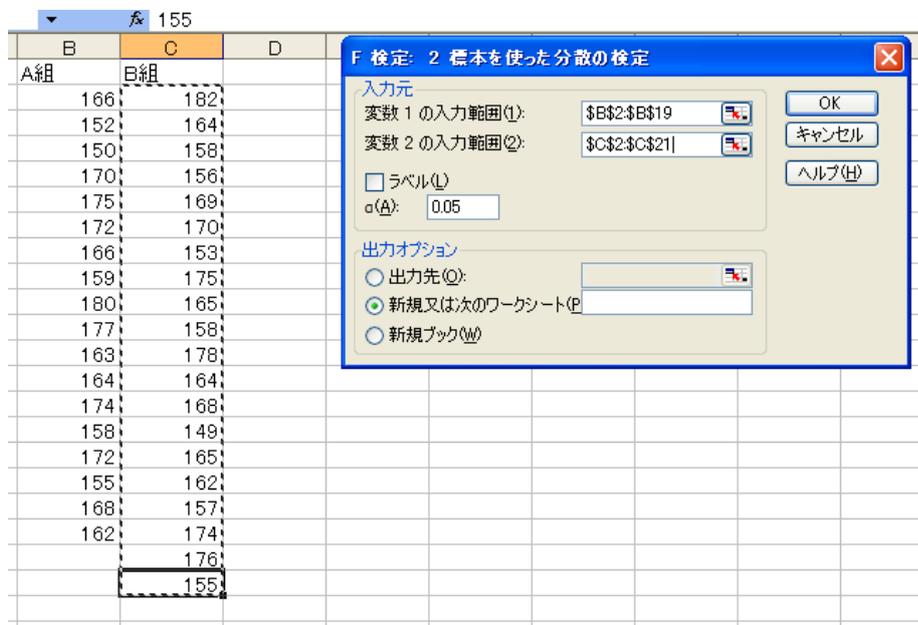
また、F 検定だけでなく、t 検定・回帰分析も分析ツールを使って行いますので、覚えておいてください。



上のような画面が出ますので、「F 検定：2 標本を使った分散の検定」を指定して OK をクリックします



すると上のような表示ができますので、「変数 1 の入力範囲」という項目に 3 年 A 組の身長データを、「変数 2 の入力範囲」という項目に 3 年 B 組の身長データを入力してみましょう。



この状態で OK をクリックすると、Excel 上で F 検定が実行されます。
計算結果はひとまず別のシートに出力されます。

	A	B	C
1	F-検定 : 2 標本を使った分散の検定		
2			
3		変数 1	変数 2
4	平均	165.7222	164.9
5	分散	74.56536	82.30526
6	観測数	18	20
7	自由度	17	19
8	観測された分散比	0.905961	
9	P(F<=f) 片側	0.421816	
10	F 境界値 片側	0.445853	
11			

計算結果はこのように出力されます。

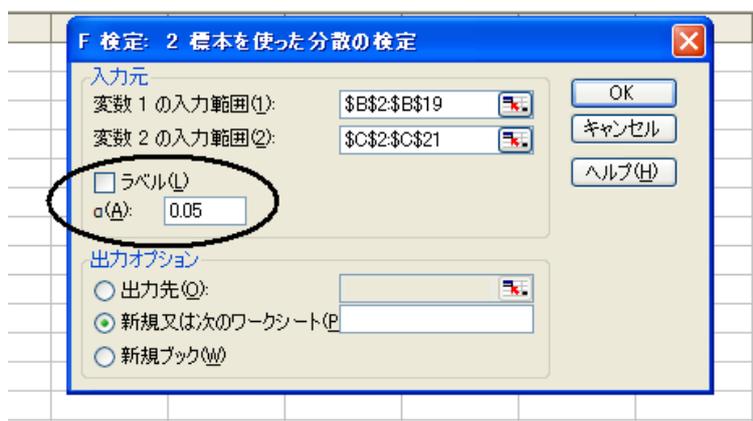
Excel のバージョンによって表示のされ方に多少の差が見られますが、等分散かどうかを判定する上で問題はありません。

F 境界値 片側<P (F<=f) 片側 であれば H0 が棄却されます。しかし、上の計算結果を見てみると P (F<=f) 片側<F 境界値 片側となっていますから、帰無仮説は棄却されず、2郡のデータは等分散であることが分かります。

コメント [TS102]: 郡→群

以上が Excel を用いた F 検定の方法です。

また、ここでは有意水準を 5% に設定していますが、変更することも可能です。



楕円で囲まれた「ラベル」という欄の数値が 0.05 になっていますが、これは「有意水準 5% の場合の境界値を表示する」ということです。従って、有意水準を 1% に設定したい場合は 0.01、10% に設定したい場合は 0.1 と入力してください。

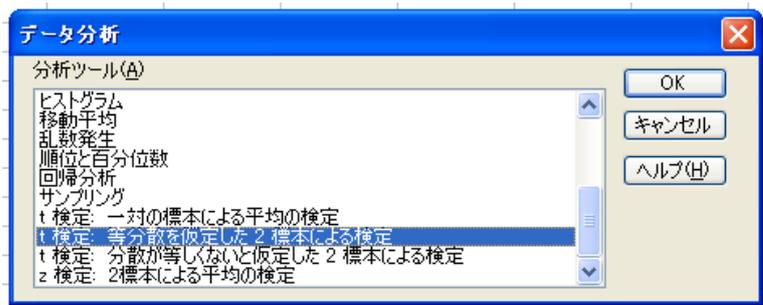
8.4 Excel を用いた t 検定

それでは、次に t 検定を行ってみましょう。

第 6 章でも学びましたが、t 検定とは 2 つの郡の平均値に有意な差があるかどうかを検定するものです。ここでは、F 検定の際に使用したサンプルについて、次の仮説検定を行いたいと思います。

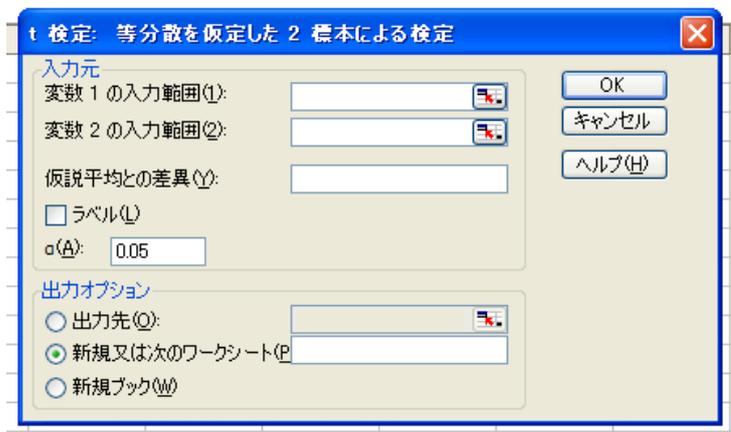
H0 : 2 群の母平均値に差はない

H1 : 2 群の母平均値に差がある

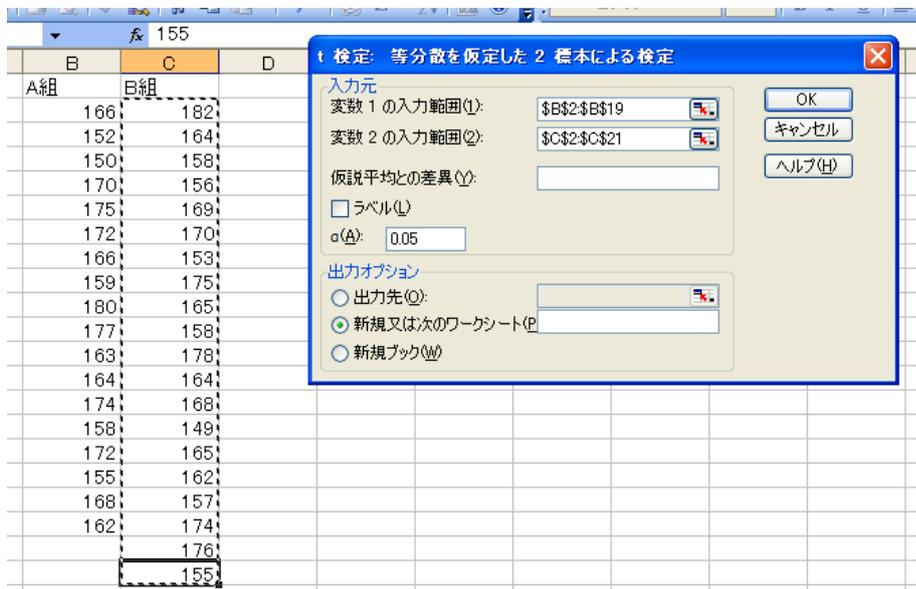


F 検定を行った際と同様に、ツールバーから「分析ツール」を選んでクリックすると上のような画面が表示されます。今回は F 検定の結果から 2 つの郡は等分散であると確認できているため、「t 検定:等分散を仮定した 2 標本による検定」を選びます。

もしも F 検定の結果、2 つの郡が等分散ではなかった場合は、「t 検定:分散が等しくないと仮定した 2 標本による検定」を選び、OK をクリックしてください。



t 検定を選ぶと、上のように F 検定と非常に似た画面が表示されます。



F 検定を行った際と同様に、変数 1 に A 組の身長データ・変数 2 に B 組の身長データを入力し、OK をクリックします。

	A	B	C
1	t-検定：等分散を仮定した2標本による検定		
2			
3		変数 1	変数 2
4	平均	165.7222	164.9
5	分散	74.56536	82.30526
6	観測数	18	20
7	プールされた分散	78.65031	
8	仮説平均との差異	0	
9	自由度	36	
10	t	0.285364	
11	P(T<=t) 片側	0.388501	
12	t 境界値 片側	1.688298	
13	P(T<=t) 両側	0.777001	
14	t 境界値 両側	2.028094	
15			

すると、上のような形で計算結果が表示されます。

計算の結果 $t=0.285364$ と出ました。この t 値を「t 境界値 片側」、或いは「t 境界

値 両側」と比較します。ここでは仮説の特徴からいって両側検定を行った方が適当ですから、「t 境界値 両側」の値と比べましょう。

t 境界値 両側 < t 値となれば H0 は棄却されますが、t 値 < t 境界値 両側 = 2.028094 なので H0 は棄却できません。従って、仮説 H0 が採択され、2 群の母平均値に差はないという結果が得られました。

以上が Excel を用いた t 検定の方法です。

8.5 Excel を用いた独立性の検定

χ^2 乗検定については、基本的に Excel ではフォローされていません。したがって、 χ^2 乗統計量は自分で計算する必要があります。

以下では、第 7 章で χ^2 乗検定を学ぶ際に用いた、コインの表裏についての例を取り上げたいと思います。

	表	裏
観測度数	58	42
期待度数	50	50
両度数の差	8	-8

まず、上のような表を作成します。

G	H	I
	表	裏
観測度数	58	42
期待度数	50	50
両度数の差	8	-8
	=(H5*H5)/H4+(I5*I5)/I4	

次に上のように χ^2 乗統計量を計算します。

$(H5*H5)/H4 + (I5*I5)/I4$ という部分が χ^2 乗統計量の計算式です。8.1 で学んだように、* は乗算を、/ は割算を表わす記号ですから、上記の式は $(8 \times 8) \div 50 + (-8 \times -8) \div 50$ を意味します。

	表	裏
観測度数	58	42
期待度数	50	50
両度数の差	8	-8
χ^2 乗統計量	2.56	

以上を計算した結果、 χ^2 乗統計量は2.56と出ました。

さて、次に $\chi^2 a$ を求めましょう。
 自由度と有意水準を求め、 χ^2 分布表から $\chi^2 a$ を求めることもできますが、Excelには $\chi^2 a$ を計算してくれる機能があるので、利用してみましょう。

コメント [TS103]: ここも、前出のときと同様に、aが有意水準を示していることを明示すべき。
 以下、指摘省略

	表	裏
観測度数	58	42
期待度数	50	50
両度数の差	8	-8
χ^2 乗統計量	2.56	
自由度	1	
有意水準	5%	

まずは自由度と有意水準を入力します。第7章で求めたとおり、自由度は1、有意水準は5%に設定します。

	A	B	C
1		表	裏
2	観測度数	58	42
3	期待度数	50	50
4	両度数の差	8	-8
5			
6	χ^2 乗統計量	2.56	
7			
8	自由度	1	
9	有意水準	5%	
10			
11	$\chi^2 a$	=CHIINV(B9,B8)	
12		CHIINV(確率,自由度)	
13			

次に $\chi^2 a$ を出力したいセルを選択し、chiinv(a,b)と入力します。a には有意水準を入力したセル、b には自由度を入力したセルを入力して下さい（上の表では a=B9、b=B8）。

	A	B	C
1		表	裏
2	観測度数	58	42
3	期待度数	50	50
4	両度数の差	8	-8
5			
6	χ^2 乗統計量	2.56	
7			
8	自由度	1	
9	有意水準	5%	
10			
11	$\chi^2 a$	3.841459	

この状態で Enter Key を押すと、指定した自由度と有意水準の $\chi^2 a$ が計算されます。

あとは、計算した χ^2 統計量と $\chi^2 a$ を比較するだけです。

	A	B	C	D
1		表	裏	
2	観測度数	58	42	
3	期待度数	50	50	
4	両度数の差	8	-8	
5				
6	χ^2 乗統計量	2.56		
7				
8	自由度	1		
9	有意水準	5%		
10				
11	$\chi^2 a$	3.841459		
12				
13		χ^2 乗統計量 2.56 < $\chi^2 a$ 3.84		
14		したがって、帰無仮説は採択された		
15				

第IV部 実証会計の実例

以下の第9・10章では、既にこれまでに学習してきた会計学・ファイナンスの知識や統計手法を用いて、学生が行った実証研究を紹介します。学習してきた内容が実際にどのように研究に用いられているかを実感してください。

研究を開始するにあたって着目したのがオールソンモデルです（第3章参照）。

オールソンモデルの式を再掲すると、以下のようになります。

$$\text{企業価値} = BV_0 + \frac{NI_1 - rBV_0}{1+r} + \frac{NI_2 - rBV_1}{(1+r)^2} + \dots + \frac{NI_n - rBV_{n-1}}{(1+r)^n}$$

BV=資本合計

NI=当期純利益

r =株主資本コスト率

本モデルは資本合計と当期純利益を使用するモデルであり、会計処理を伴う企業の財務情報はモデルにより算出される企業価値に盛り込まれると考えられます。一般的にも、オールソンモデルによる推定株価は実際株価に対する説明力が高いとされています。

しかし、実際にはオールソンモデルによる推定株価と実際株価が一致することは極めて稀です。株価は、財務情報だけでなく、景気・市場の動向・会社の信用度など、会社内外の要因に大きな影響を受けるためです。そこで財務情報以外の要因（以下、非財務情報と呼称）をあぶりだすため、実証研究を行います。

コメント [TS104]: 表現。

ここでは、財務情報以外の要因が株価に与える影響を考察するために、財務情報以外の要因について考察を深めた実証研究の実例を紹介していきます。

くらいでしょうか？

藤井ゼミのゼミ生研究の実例であることを脚注くらいに落としておくといいのではないのでしょうか。

第9章 実証研究の実例 1—浮動株と株価形成に関する考察—

1. サンプルの抽出
2. オールソンモデルによる推定株価を算定、実際株価と比較（回帰分析 第5章）
3. 非財務情報のうち、何に焦点をあてるかを決定
4. 実際株価が推定株価を上回るグループ、実際株価が推定株価を下回るグループに区分し、非財務情報を用いてグループ間の違いを検定（平均差の検定、 χ^2 乗検定、第6・7章）
5. 非財務情報と実際株価の関係性を調査（回帰分析）
6. 実証研究のまとめ

9.1 サンプルの抽出

サンプル企業は、日経新聞株価欄の「商業」に掲載されている企業（以下、「商業全企業」とする。）の中から、三月決算であり、東証一部に上場している企業を抽出しました。決算の時期を指定するのは、検定を行う上での条件を均一にするためです。この結果、サンプル数は以下の62社となりました。

【サンプル企業一覧】

高島	ハウスイ	カナデン	菱電商事
長瀬産業	松田産業	岩谷産業	ニチモウ
豊田通商	トーメンエレクトロニクス	ナイス日榮	伊勢丹
三共生興	ムトウ	兼松エレクトロニクス	丸井
カメイ	第一実業	稲畑産業	上新電機
日本紙パルプ商事	日立ハイテクノロジー	稲畑産業	東日カーライフグループ
東都水	三愛石油	明和産業	Mrmax
山善	サンリオ	ミツウロコ	ヤオコー
椿本興業	ソーダニッカ	シナネン	ソフトバンク
西華産業	AOKI HD	エネクス	伯東
金商又一	日商エレクトロニクス	TOKAI	いなげや
佐藤商事	松屋フーズ	リョーサン	ドトールコーヒー

コメント [TS105]: ここでの分析で、なにを明らかにしたいのか、(論文で言うところの「はじめに」)の部分が、完全に省略されているので、概要が把握できずに、唐突な印象を受けます。

コメント [TS106]: なぜ商業に限定？
こじ付けでも理由は必要。

ユアサ商事	因幡電機産業	新光商事	三城
神鋼商事	ツカモトコーポ レーション	テンアライド	コロワイド
阪和興業	神栄	阪急百貨店	
ゼビオ	ヤマタネ	太平洋興発	

9.2 オールソンモデルによる推定株価を算定、実際株価と比較（回帰分析）

オールソンモデルに従い、2002年度末から2005年度末までの推定株価を算出しました。各年度末の推定株価は、3年分の財務諸表の数値を使用します。

$$\hat{P}_t = \left\{ BV_t + \frac{NI_{t+1} - r BV_t}{1+r} + \frac{NI_{t+2} - r BV_{t+1}}{(1+r)^2} + \frac{NI_{t+3} - r BV_{t+2}}{(1+r)^3} \right\} \div \text{株式数}$$

\hat{P}_t : t年度の推定値
 BV_t : t年度の資本合計
 NI_t : t年度の当期純利益
 r : t年度末の長期国債新発 利率

ただし、2003年度から2005年度までの推定株価を算定するに当たっては、以下の仮定を設けることとしました。これは、入手できる財務情報が限られているためです。

$$NI_{07} = NI_{06} = NI_{05}$$

次に、算出された推定株価と、各年度末における実際株価を比較します。二つの数値の関連性を調べるために、回帰分析を使用しました。その結果、関連性の高さを表す補正R2の値は以下になりました。

【推定株価と実際株価の相関】

	2002年度	2003年度	2004年度	2005年度
補正R2	0.56090	0.54419	0.531558	0.302897

2004年度と2005年度は推定株価の算出において仮定を設けたにも関わらず、高い数値を得ることができました。ここから、オールソンモデルを用いて次の研究を進めていくことが妥当であると判断されます。

9.3 非財務情報のうち、何に焦点をあてるかを決定

コメント [TS107]: なぜ?

これは、かなり論点として問題があるかもしれない。

ターミナルバリュー、まあ、言いようによっては、継続価値ともいえますが、どちらにせよ、予測期間以後の株式価値が欠落しているのは、理論的には問題が大きいと思います。予測期間も3年と大変短いです。いずれにせよ、3年で区切った理屈は、記述が必要だと思います。

他の年数（4年、5年、継続価値を含めたケースなど）の場合と比較して、3年での推定株価が優位にあったのでしょうか？

（学部の学生の研究だから、厳密性は少々このことは無視してもよいとは思いますが。論点も大きいですし、篠田個人の見解なので、他の方のご意見が必要かと思えます。）

・田中さんはどう思われるでしょうか？
 ・また、TAの佐久間さん、板橋君、（今は高橋君も？）あたりは、どのように見ているのでしょうか？

コメント [TS108]: これは、途中で出てくる問題意識というより、当初の分析目的である。

オールソンモデルによる推定株価と、実際株価の差異の原因を「浮動株の割合」に求めることとしました。「浮動株」とは、安定的に保有されている株式ではなく、投機的利益を得ることを目的として、常に市場で売買されている株式のことをいいます。浮動株の割合は、「浮動株比率＝1－特定株比率」として算定しました。「特定株比率」は、会社四季報より入手することができます。

コメント [TS109]: これは、問題意識として、冒頭で明示しておくことではないか？

9.4 二つのグループに区分し、非財務情報を用いてグループ間の違いを検定

サンプルを、実際株価が推定株価を「上回るグループ」と「下回るグループ」の二つのグループに区分し、上回るグループを高評価企業、下回るグループを低評価企業と名付けました。

【グループ化の結果】

	2002 年度	2003 年度	2004 年度	2005 年度
低評価企業	48 社	34 社	32 社	25 社
高評価企業	14 社	28 社	30 社	37 社

次に各年度において、浮動株比率がグループ間の違いを反映しているかどうかを検証しますが、検定を行うためには仮説を立てる必要があります。そこで、2002 年 9 月に発表された東京証券取引所の資料、「浮動株指数の導入について - TOPIX への浮動株比率反映と影響緩和策 -」を参考に仮説を立てることとしました。具体的な浮動株指数導入の経緯を振り返ってみましょう。

TOPIX（東証株価指数）とは、東京証券取引所が日々計算し発表している株価指数で、東証第 1 部の毎日の時価総額（全上場株をある日の終値で評価したものの合計額）を基準日の時価総額で割って算出されるものです。1968(昭和 43)年 1 月 4 日の時価総額を 100 とし計算しており、日経平均株価とならんで、重要な指数の 1 つとなっています。即ち、従来 TOPIX は以下の算定式により計算されています。

$$\text{TOPIX} = \text{上場株式数} \times \text{株価} \div \text{基準時価総額} \times 100$$

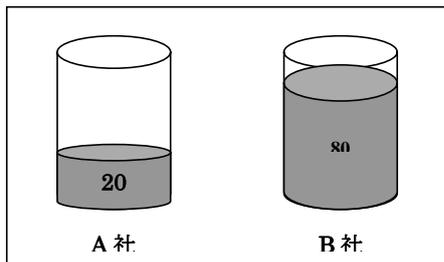
この式から、TOPIX の特徴として、株価が同一であっても、上場株式数が多い銘柄ほど、TOPIX 構成ウェイトが大きくなるということがわかります。

しかしこの方法での TOPIX の算定方法には、問題点が存在しました。インデックス運用では、TOPIX の構成割合に比例して資金配分がなされますが、ここで「浮動株」「固定株」が考慮されていないため、「需給の不一致」が発生する可能性がある、というものです。

具体例で数値を入れて考えてみましょう。図表【インデックス売買 各企業の浮動株の割合】の2銘柄は、上場株式数（浮動株+固定株）ベースで見ると時価総額は100で全く同じとなっていますが、浮動株ベースで見ると、時価総額は20、80と異なります。

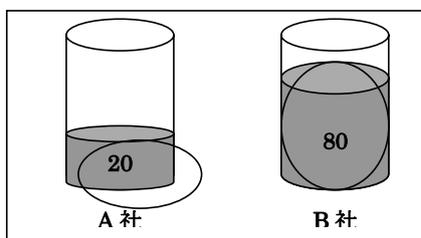
コメント [TS110]: 図表番号:以下指摘省略

【インデックス売買 各企業の浮動株の割合】



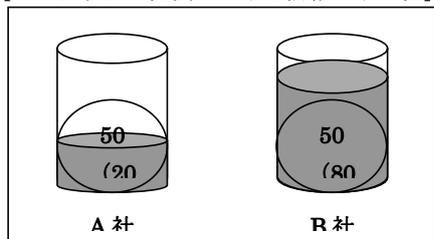
理想的には、インデックス運用は『浮動株ベース』の時価総額に基づくのが望ましいとされています。例えば、上記2銘柄に対し、総額100の運用資産につき、『浮動株ベース』の時価総額に基づき、銘柄Aを20、銘柄Bを80買い付けるとすると、各銘柄の需給関係は図表【インデックス売買における需給関係】のように一致しています。

【インデックス売買における需給関係】



しかし、これまでのTOPIXの下で、図の2銘柄に対し、総額100の運用資産につき、『上場株式数ベース』の時価総額に基づき各銘柄50ずつの買付けを行う場合、図表【インデックス売買における需給の不一致】のとおり、浮動株比率が低い20の銘柄は需要が過剰となり、浮動株比率が高い80の銘柄は理想的な需要よりも需要が多くなっていることが分かります。

【インデックス売買における需給の不一致】



() 内は理想的な需要。

現在、年金を始め、TOPIX との連動を目指す運用（インデックス運用）が増加傾向にある中で、TOPIX の重要性が増しており、海外においても浮動株指数の導入が順次なされています。そこで、新しい TOPIX として、以下の式により TOPIX が算定されることとなりました。

$$\text{TOPIX} = \text{上場株式数} \times \text{浮動株比率} \times \text{株価} \div \text{基準時価総額} \times 100$$

東証の発表を整理すると、浮動株指数導入前では「需給の不一致」が生じているということになります。即ち、次のような仮説を立てることができます。

- ・浮動株比率が低い企業→高評価企業
- ・浮動株比率が高い企業→低評価企業

一方、浮動株指数導入後は「需給の不一致」が解消されるため、導入前の関係は成立しないと考えられます。なお、今回の研究は「効率的市場」（第〇章）を前提として進めることとし、「浮動株指数導入」の時期は、東証がプレスリリースを行った 2004 年 7 月とします。この仮説を、平均差の検定、 χ^2 検定により検証していきます。

(1) 平均差の検定

上記の仮説を、検定を行う上での帰無仮説と対立仮説に転換すると、次のようになります。

帰無仮説：2 グループの浮動株比率に差は無い。

対立仮説：2 グループの浮動株比率は異なる。

(高評価企業の浮動株比率 < 低評価企業の浮動株比率)

これから、片側検定を行い、帰無仮説を棄却できるかどうかを検証します。

なお、t 検定による検証の前に、f 検定（〇章解説）を行う必要があります。f 検定の結果から、等分散を仮定した平均差の検定（2003 年度、2004 年度）、および分散が等しくないと仮定した平均差の検定（2002 年度、2005 年度）を行います。

【t 検定結果】

コメント [TS111]: もう少し丁寧にパラフレーズすることが必要。どうして、浮動株比率が低いと高評価企業といえるのか。

コメント [TS112]: ???
セミストロング

	2002 年度	2003 年度	2004 年度	2005 年度
t 値	3.168267 ***	2.123309 **	-0.208937	0.454055

*** : 1%の水準で有意

** : 5%の水準で有意

* : 10%の水準で有意 (以下共通)

上記の結果から、2002 年度では 1%有意水準で、2003 年度では 5%有意水準で帰無仮説を棄却することができました。2002 年度と 2003 年度の検証結果の詳細を見てみましょう。

【2002 年度の t 検定結果】

	低評価企業	高評価企業
平均	53.3125	40.26429
分散	196.1037	139.4948
観測数	48	14
プールされた分散	183.8384	
仮説平均との差異	0	
t	3.168267	

【2003 年度の t 検定結果】

	低評価企業	高評価企業
平均	52.9	45.50357
分散	158.5606	220.2515
観測数	34	28
プールされた分散	186.3215	
仮説平均との差異	0	
t	2.123309	

いずれの年度においても、「低評価企業の浮動株比率の平均 > 高評価企業の浮動株比率の平均」となっています。ここから、浮動株比率の大小が、低評価企業と高評価企業のグループ間の差異として表れていることがわかります。

(2) 検定

検定のための帰無仮説と対立仮説は次のようになります。

帰無仮説：企業評価と浮動株比率の間に関連が見られない。

対立仮説：企業評価と浮動株比率の間に関連が見られる。

コメント [TS113]: ???

χ² 検定のことか？

ところで、 χ^2 検定を行うためには、高評価企業・低評価企業のグループ分けに加えて、浮動株比率の大小によるグループ分けを行う必要があります。ここで、区分の基準として、恣意性を排除するために、各年度の「商業」全企業の浮動株比率を平均した値（「母集団平均」と呼称）を採用することにしました。

【 χ^2 検定結果】

	2002 年度	2003 年度	2004 年度	2005 年度
母集団平均	48.045	47.028	47.584	47.472
χ^2 値	7.56545 ***	6.53370 **	0.67215	0.62188

上記の結果から、平均差の検定と同様に、2002 年度では 1% 有意水準で、2003 年度では 5% 有意水準で帰無仮説を棄却することができました。2002 年度と 2003 年度の検証結果の詳細は以下のようになります。

【2002 年度の χ^2 検定結果】

浮動株の母集団平均…48.045

実測値	浮動株 > 48.045	浮動株 < 48.045	計
低評価企業	32	16	48
高評価企業	4	10	14
計	36	26	62

期待値	浮動株 > 48.045	浮動株 < 48.045	計
低評価企業	27.87096774	20.12903226	48
高評価企業	8.129032258	5.870967742	14
計	36	26	62

【2003 年度の χ^2 検定結果】

浮動株の母集団平均…47.028

実測値	浮動株 > 47.028	47.028 > 浮動株	計
低評価企業 (社)	26	8	34
高評価企業	13	15	28
計	39	23	62

期待値	浮動株 > 47.028	47.028 > 浮動株	計

低評価企業（社）	21.38709677	12.61290323	34
高評価企業	17.61290323	10.38709677	28
計	39	23	62

いずれの年度においても、期待値に比べ、低評価企業で浮動株比率の高いグループ、高評価企業で浮動株比率の低いグループの構成数が多いことが分ります。

(3) 結果のまとめ

ここまで行ってきた検定結果をまとめてみましょう。2002年度、2003年度は、

- ・浮動株比率が低い企業→高評価企業
- ・浮動株比率が高い企業→低評価企業

という関係が見られましたが、2004年度以降においてはこの関係が見られませんでした。ここから、本項のはじめに立てた仮説は証明されたと考えられます。

9.5 非財務情報と実際株価の関係性を調査（回帰分析）

最後に、これまで行った研究とは違った視点で、『2002年度と2003年度において、浮動株比率が高ければ高いほど株価は低く評価され、浮動株比率が低ければ低いほど株価は高く評価される』という、浮動株比率と株価の関連性を調べます。検証方法は、本章の2.で行った方法に準じ、回帰分析を採用します。

回帰分析では、補正R2による判断ではなく、回帰式 $y = a + b x$ （ y ：実際株価 x ：浮動株比率）の回帰係数 b の有意性を検証することとします。このとき、帰無仮説と対立仮説は次のようになります。

帰無仮説： $b = 0$

対立仮説： $b \neq 0$

回帰係数 b の有意性は、 t 比較検定量を利用して判断することができます。2002年度と2003年度における p 値はそれぞれ次のようになりました。

2002年度：0.000199

2003年度：0.007175

上記の値は、共に5%以下となっているため、算出された傾きは意味のあるものとなることが出来ます。さらに、各年度のデータをプロットしてみると、図表で示されるとおり、負の傾きとなっています。ここから、『2002年度と2003年度において、浮動株比率が高ければ高いほど株価は低く評価され、浮動株比率が低ければ低いほど株価は高く評価される』という線形関係が成り立つことが分りました。

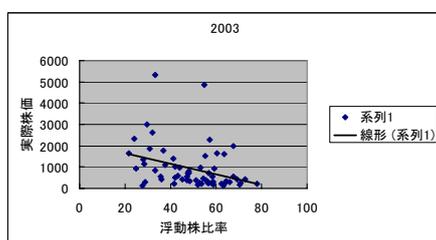
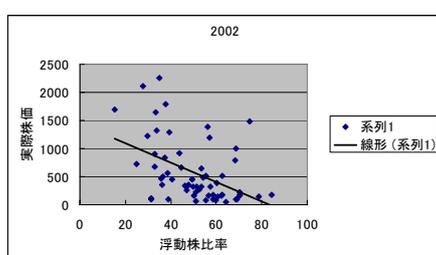
【回帰分析結果】

	2002 年度	2003 年度
補正R 2	0.184067	0.0908123

2002 年度	係数	標準誤差	t	P-値
X 値 1	-17.0736	4.307173	-3.96399	0.000199

2003 年度	係数	標準誤差	t	P-値
X 値 1	-24.7085	8.875756	-2.78382	0.007175

【実際株価と浮動株比率の関係】



9.6 本研究のまとめ

今回の研究では、まず Ohlson モデルを用いて、推定株価の算出を行いました。この推定株価と、実際株価の間には高い相関関係がありましたが、二つの数値の間には乖離があります。この差異を構成する一つの要素として、私たちは浮動株の存在に着目しました。そして、t 検定、 F 検定、回帰分析の結果、次の結論が得られました。

浮動株比率は、2002 年度と 2003 年度において、株価に非財務情報として影響を与え、その結果、浮動株比率が高ければ高いほど、株価が低く評価されていました。これは、インデックス運用の株式においては需給の不一致が発生しており、インデックス運用が投資額全体の 5% とはいえ一つの流れとなっており、他の投資家も同様の動きを見せたためだと推定できます。しかし、2004 年度以降にはこの傾向が見られなくなりました。これは、東証の

コメント [TS114]: 結果から、わかったことと、わかったことから解釈したことは、別であると分かるように記述すべき。

浮動株指数導入発表のインパクトを受けたためと考えられます。現在の浮動株比率と株価の動きについては、因果関係が見られません。浮動株比率の多寡の影響が互いに相殺された状態であると言えます。

ただし、この研究は、「浮動株比率 = 1 - 特定株比率」という仮定をおいて検証を行ったものです。浮動株は、投機的利益を得ることを目的として、常に市場で売買されている株式をいいますが、これを表すものとして会社四季報に掲載されている「浮動株比率」の数値を使うことも可能です。今回は、この数値を用いた検証は省略しますが、実証研究を行う際は、より強固な結論を得るために、ぜひ様々な側面からのアプローチを試みてください。

コメント [TS115]: この章の分析は大変ユニークです。浮動株比率に着目するというのは、素材としてはとても面白いので、感心しました。

ただ、構成が少し気になります。

構成は直せるので、ともかくとしても、根幹部分で気になるのは、①サンプル選択、②仮説設定の背後にあるストーリー（浮動株比率低い→高評価企業）、③オールソン・モデルにおける予測期間後の価値の取り扱いです。

第 10 章 実証研究の実例 2—高齢化の株価形成に与える影響—

1. 第 I 部の第 5 章 5.4 で説明した Ohlson モデルを用いて、各企業の企業価値を算出します。そして、その企業価値を発行済株式総数で割ることによって、推定株価を算出します。
2. 次に、その推定株価と実際株価を比較します。推定株価よりも実際株価の方が高い企業を高評価企業とし、逆に推定株価よりも実際株価の方が低い企業を低評価企業として分類します。
3. 第 II 部第 8 章の平均差の検定を用いて、実際に高齢化が株価にどのように影響するかを研究します

10.1 研究の目的

現在、日本では高齢化が急速に進展し、年金などの社会保障制度に大きな影響を与え、社会問題化しています。そして、高齢化は企業環境にも少なからず影響を与え、それらが企業の株価に影響を与えている可能性は十分にあります。ここでは、**企業の株価に高齢化がどのように影響しているのかを研究します。**

研究対象の業界は、食品業界としました。食品業界を採用した理由は、食品は生活必需品であるため、景気変動の影響を受けにくい業種であるといわれているためです。高齢化の影響を研究する際、長期的な期間にわたって研究する必要があるのですが、対象業種が景気変動の影響を受けやすいと**純粋に高齢化の影響を研究する事が難しい**からです。

10.2 サンプル

サンプルは、東京証券株式取引所に上場する企業を対象としました。サンプリングにおける恣意性を排除する為に東証による分類を用い、**食料品業界から抽出しました。**なお、三月決算企業で、研究の目的から 1998 年三月期から 2006 年三月期まで合併・組織再編などが無い企業としました。この条件に合致した企業数は 56 社でした。

【サンプル企業】

昭和産業	明治乳業	ニチレイ	雪印乳業
協同飼料	森永乳業	東洋水産	プリマハム
中部飼料	ヤクルト本社	オリエンタル酵母工業	日本ハム
日本配合飼料	伊藤ハム	日本食品化工	林兼産業

コメント [TS116]: あまりにも飛躍がある表現のように思います。消費者層の高齢化を見据えた経営戦略をとっている企業を、マーケットがどのように評価しているのかについての分析でだ思うのですが。

コメント [TS117]: 表現

コメント [TS118]: なぜ？
(こじ付けでも、業界選択の理由の明記が必要)

東洋精糖	丸大食品	石井食品	加ト吉
塩水港精糖	キーコーヒー	日清食品	フジッコ
森永製菓	不二製油	ソントン食品工業	理研ビタミン
明治製菓	ポーソー油脂	永谷園	四国ココ・コーラボトリング
中村屋	キッコーマン	日本製粉	焼津水産化学工業
名糖産業	味の素	日本農産工業	福留ハム
井村屋製菓	エスビー食品	日本甜菜製糖	ブルボン
不二家	ハウス食品	三井製糖	亀田製菓
山崎製パン	カゴメ	日新製糖	アリアケジャパン
第一屋製パン	あじかん	江崎グリコ	日本たばこ産業

10.3 Ohlson モデルの適用

今回の研究では、Ohlson モデルの使用にあたって、2 期間モデルを軸にすえました。その理由は、2 期間モデルを使用することによって長い期間の分析が可能になるのと、1 期間モデルでは推定株価が低くなり分析に不向きであると考えられるからです。また、2 期間モデルだけでは実証力に欠ける面もあると思われるますので、補足的に 3 期間モデルも使用しました。

Ohlson モデルを使用して推定株価を算出し、推定株価と実際株価との回帰分析をかけた結果は以下の通りです。なお、この回帰分析に際して、日本たばこ産業の株価が他のサンプル企業に対して非常に高く、より正確な分析を行うためにこの回帰分析からは控除しました。この結果、補正 R^2 はおおむね高いので、Ohlson モデルの株価説明力は高く、Ohlson モデルが実際のサンプル企業においても良く当てはまっている事が分かります。その為、Ohlson モデルによって算出した理論上の株価を各企業の推定株価であると考えても差し支えないという事が出来ます。また、各年度において傾きの t 値が非常に高く、全ての年度において片側 1%水準で有意であるため、強い結果であるといえます。

次に、推定株価と実際株価の比較ですが、各年度においての高評価企業・低評価企業に分類された企業数は以下の通りです。いずれの年度においても、連結情報で一部上手くいかなかった場合もありますが、高評価企業・低評価企業の比率はおおむね 3 : 2 の比率に収まっており、分類は比較的上手くいったといえます。片方により過ぎていると研究が上手くいかない場合があるからです。

【推定株価と実際株価との回帰結果（単独）】
（2 期間モデル）

コメント [TS119]: なぜ？

コメント [TS120]: 理由が弱い……。

コメント [TS121]: 3 期間モデルのほうが統計的には優位である。

（しかし、2 期間モデルと大差はないから、長期間の分析に向いている 2 期間モデルを採用したということか？それならそう明示すべき。）

2 期間でも、3 期間でも、継続価値を足し込んだモデルを採用しないのはなぜか？

	1997	1998	1999	2000
補正R2	0.682225	0.710273	0.55196	0.477811
傾き	1.170644	1.244357	1.260495	1.29414
t値	10.81348	11.549117	8.321360	7.100063

	2001	2002	2003
補正R2	0.638484	0.653908	0.719857
傾き	1.033205	0.906529	0.988303
t値	9.816879	10.150267	11.82197

(3期間モデル)

	1997	1998	1999	2000
補正R2	0.682225	0.71294	0.572511	0.493731
傾き	1.170644	1.180557	1.240451	1.272531
t値	10.75184	11.623859	8.562658	7.325476

	2001	2002
補正R2	0.650602	0.668256
傾き	1.006392	0.868975
t値	10.07728	10.477402

【高評価企業・低評価企業の企業数(単独)】

2期間モデル	1997年	1998年	1999年	2000年	2001年	2002年	2003年
高評価企業数	36	35	32	33	31	25	31
低評価企業数	20	22	24	23	25	31	25

3期間モデル	1997年	1998年	1999年	2000年	2001年	2002年
高評価企業数	36	34	33	32	28	21
低評価企業数	20	22	23	24	28	35

【推定株価と実際株価の回帰結果(連結)】

(2期間モデル)

	1999	2000	2001	2002

補正R2	0.273632	0.176678	0.248873	0.239701
傾き	0.716534	0.643788	0.512168	0.427898
t値	4.619792	3.547946	4.346494	4.245549

	2003
補正R2	0.305653
傾き	0.501013
t値	4.977039

(3期間モデル)

	1999	2000	2001	2002
補正R2	0.288774	0.185540	0.251453	0.245615
傾き	0.710191	0.631797	0.490972	0.412635
t値	4.788035	3.647133	4.374897	4.310622

【高評価企業・低評価企業の企業数（連結）】

2期間モデル	1999年	2000年	2001年	2002年	2003年
高評価企業数	27	29	24	18	29
低評価企業数	29	27	32	38	27

3期間モデル	1999年	2000年	2001年	2002年
高評価企業数	27	27	23	14
低評価企業数	29	29	33	42

10.4 高齢化の影響

10.5

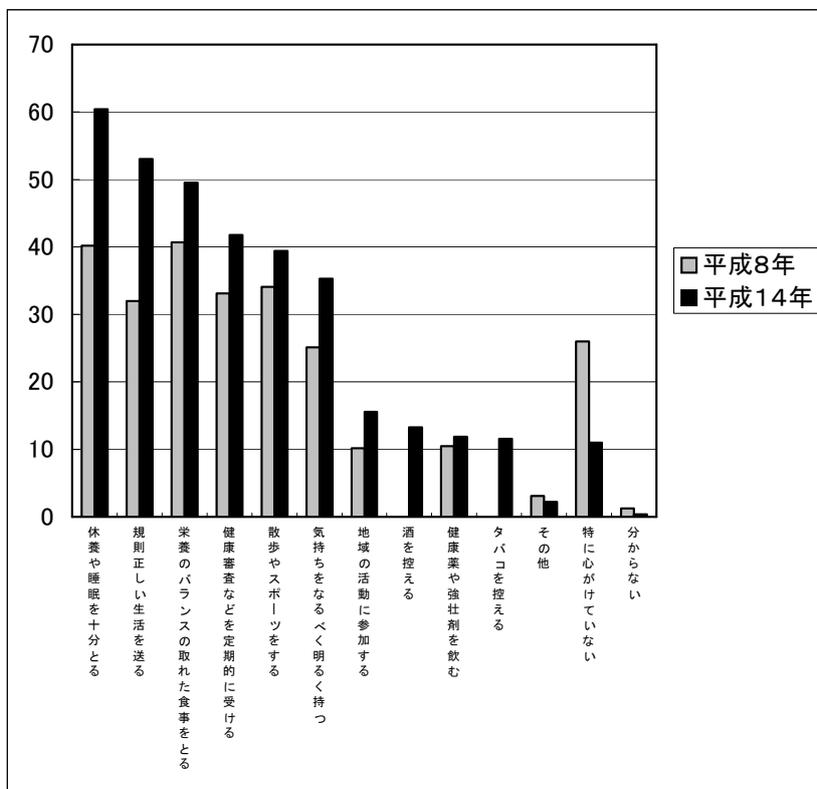
今回は、高齢化の影響が企業の株価にどのように影響を与えるかを研究するに当たって、食品業界の消費者層の高齢化という観点にしばって研究を行いました。食品業界にとっての顧客である消費者は全国民といっても良いほどですが、その国民の高齢化によって各企業の株価に影響を及ぼしている可能性は多分にあります。以降、具体的な影響を研究します。

高齢化といっても、そこにはポジティブな影響とネガティブな影響が想像できます。す

なわち、消費年齢層が高齢化すると考えれば、お年寄りを相手にしていた企業は必然的に将来性が期待され、若者をターゲットにしていた企業は、対応が遅れることとなります。そこで、今回は高齢化を「健康」という切り口から見てみました。高齢化と健康への関心の高まりは全く同じというわけではありませんが、以下の意識調査結果を見ても、高齢者が健康を意識する傾向は強まっていますので、高齢化と少なからず関係があることは明白であると言えます。

コメント [TS122]: 若者層や中年層が健康を意識する傾向は強まっていないのですか？
たとえば、30代の主婦層などは、健康を意識していないのでしょうか？
少し粗さを感じます。

【高齢者の健康に関する意識調査】



平成14年度内閣府共生社会政策統括官『高齢者の健康に関する意識調査結果』より作成

10.5 特定保健用食品について

また、健康といってもさまざまな次元があります。その中で、今回は食品業界と健康を結びつけるものとして、健康食品について分析を行いました。

厚生労働省が身体の生理学的機能などに影響を与える保健機能成分を含んでいると認め
たのが「特定保健用食品」、いわゆるトクホです。「お腹の調子を整える」など、特定の保
健の目的が期待できることを表示できる食品であり、審査を受け、厚生労働大臣の許可を
得なければなりません。

今回の分析では、健康に気を遣うことを示す代理変数として、**各企業の特定保健用食品
の個数を採用しました。**それなりの研究開発費を投じて特定保健用食品を開発していると
するならば、その多寡がほぼ直接的に健康への気の遣い方として現れると思えるからです。

分析手法としては、先ほど分類した高評価企業・低評価企業の間で、その特定保健用食
品の個数の平均値に有意な差があるかどうかを、平均差の検定で検証しました。これは、
Ohlson モデルを用いて算出した推定株価が会計情報を全て含んだ真の株価であると考え
ると、高評価企業では会計情報以外の何かしらの株価に好影響を与えるものが存在し、逆
に低評価企業では悪影響を与えるものが存在していると考えられるからです。つまり、食
品業界では、**高齢化に対応した特定保健用食品を多く有することが株価に好影響を与え、逆
にそれほど有していないことが悪影響を与えるのではないかと予想されます。**検定結果は
以下の通りでした。

【特定保健用食品の個数について平均差の検定（単独）】

H_0 ：高評価企業と低評価企業において個数の平均値に差はない。

H_1 ：高評価企業は低評価企業より個数の平均値が高い。

2 期間モデル	2000 年度	2001 年度	2002 年度	2003 年度
t 値	1.952648**	1.918048**	1.636624*	2.05023**

3 期間モデル	2000 年度	2001 年度	2002 年度
t 値	1.854708**	1.618054*	1.301283

*：10%の水準で有意 **：5%の水準で有意

2 期間モデルにおいてはすべての年度で、また、3 期間モデルにおいても 2000、2001 年
度において有意な差が検出されました。この結果から特定保健用食品の個数に表されるよ
うな、各企業の健康へ気遣いは株価形成に影響を与えているということが出来ます。

10.6 検証の拡張

以上、特定保健用食品に絞って検証を行いました。確かにしっかりした地位をもつ指標
ですが、それでも企業の実態を全て表しているとはいえません。つまり、**企業が消費者の
健康にスポットをあてた製品戦略を行い、それに関して企業間に差がでているとすれば、
トクホだけでその全てを表せているのかという疑問が出てきます。**

コメント [TS123]: トクホ食品の商品数の
ことか?

コメント [TS124]: 疑問があります。
・トクホ＝高齢者を意識かどうか?つまり、
中年層を意識したトクホや、若い女性を意
識したトクホもあるだろうと思われるが、
それらは、分類されているのか? (お腹の
調子を整えるヨーグルトは、高齢者向けの
トクホ?、メタボ対策のダイエット緑茶な
どは、高齢者向けのトクホ?)
これらが明確に分類できていないとすると、
高齢化を見据えた経営戦略を採用している
企業の分析ではなくて、(年齢層に関わら
ず) 消費者の健康志向を見据ええた経営戦
略を採用している企業の分析になりませ
んか?

コメント [TS125]: その通りです。

そこで、追加的な検証として、研究開発費に関して調べてみました。トクホ以外の健康食品や健康志向の食品などの研究開発にどれほど力をいれているのかも、高齢化への対応としてポジティブに評価されるのではないかと考えたからです。具体的な方法としては、有価証券報告書に記載された研究開発費を、企業の規模を排除するために経営資本で割ってグループ間の差を見てみました。ここで経営資本とは、総資産から繰延資産、建設仮勘定、投資その他の資産を減じたものをいいます。検証方法は、先ほどと同じ平均差の検定を用いました。

H_0 : 高評価企業と低評価企業において平均値に差はない。
 H_1 : 高評価企業は低評価企業よりも多額の研究開発費を投じている。

【研究開発費に関する平均差の検定 (単独)】

2 期間モデル	2001 年度	2002 年度	2003 年度
t 値	1.4714595146 *	0.816975911	0.690158346

3 期間モデル	2001 年度	2002 年度
t 値	1.50644937626926 *	0.682202568

* : 10%の水準で有意

2001 年度において 10%水準で有意な結果がでたものの、全体としてはいい結果とは言えませんでした。追加検証においてはトクホの結果を補強することはできませんでした。原因として、やはり健康志向の製品以外の研究開発などが大きく影響したのではないかと考えられます。健康志向の研究開発のみに充てられた研究開発費を検証すればまた違った結果になったかもしれないものの、データの制約から今回はかないませんでした。

10.6 まとめ

まず、第1段階として、Ohlson モデルを利用しました。Ohlson モデルはあくまでも理論的なものですが、実際の企業の事も説明できていたということが分かりました。そして、今回の研究テーマである高齢化の株価形成に与える影響ですが、「健康」というものを媒介にして、高齢化に対応した、つまり健康に気を遣った企業が高評価を受けるといった結果でした。食品業界は、ある意味では消費者に最も密着した産業であり、また健康との関わりも直接的であるため、消費者の健康への対応の度合いはストレートに株価に影響を与えると思われま

高齢化が進む中、消費者の健康に対する意識は高まる一方です。そんな中、健康にスポ

コメント [TS126]: これも難しい分析ですね。後に本文でも指摘されているように、研究開発のどの程度が、トクホのため、あるいは、健康志向の食品のためにあてられているのかについては不明だからです。ただし、(どのような商品開発なのかは特定できなくても,)商品開発に力を入れているということが、市場からの評価につながっているかどうかを分析することはできますね。しかしながら、もうひとつ別の問題点もあります。というのも、研究開発の成果は、数期間遅れて発現すると考えるのが自然です。開発の成果にはタイムラグがあるということです。したがって、同期間の研究開発費と企業の株価評価を、単純に比較することでは、研究開発の効果をきちんと分析できないという問題点も残っています。

コメント [TS127]: そうですね。先に示しておいた方がよいのでは？

コメント [TS128]: 表現:「事」とは？

コメント [TS129]: 食品産業を選択した理由がこれなら、先に明示すべきです。

ットをあてて業績を上げようとする企業が出現してくるはずで、投資家はそのような企業の将来を見込み、高評価を与えるのです。その一方で対応できない企業は評価が下がっていきます。結果、2グループ間の差が現れたのではないのでしょうか。追加で行った研究開発費との関連においては、今回は満足いく結果は得られませんでした。それは、我々がアウトサイダーであるという限界もあると思われまます。しかし、我々はその為に様々な仮定をおいたり、統計を利用するなどして克服しなければなりません。追加的な検証については、今後の課題としたいところです。

コメント [TS130]: ???

企業外部者だから、企業の実態を把握するのに限界があったり、公表データ以外のデータを入手できないことに限界があったりすることをさしているのでしょうか？まず、その点が分かりにくい表現です。

また、この表現では、問題設定は正しいはずだけど、各種のノイズと入手情報の限界のために、想定していたような結果が出なかったという解釈的な結論を示しているように読み取れます。

しかし、問題設定が間違っていたかもしれないという結論もありえるので、見方が公平でなく感じられてしまいます。

書き方の問題ですね。

参考文献

- Ball,R.J. and P.Brown[1968], “An Empirical Evaluation of Accounting Income Numbers,” *Journal of Accounting Research*, Vol.6, pp.159-178.
- Beaver, W.H.[1968], “The Information Content of Annual Earnings Announcements,” *Empirical Research in Accounting: Selected Studies 1968, Supplement to Vol.6 of Journal of Accounting Research*, pp.67-92.
- Freund,J.E. and F.J.Williams[1972], *Elementary Business Statistics: The Modern Approach*, 2nd ed., Prentice-Hall, 福場庸, 大沢豊共訳[1989, 1990]『経済経営系のための統計学入門』上・下, 培風館。
- Watts, R.L. and J.L.Zimmerman[1986], *Positive Accounting Theory*, Prentice-Hall, 須田一幸訳[1991]『実証理論としての会計学』白桃書房。
- 石川博行[2007]『配当政策の実証分析』中央経済社。
- 石塚博司編[2006]薄井彰, 後藤雅俊, 井上達男, 大日方隆, 河榮徳, 須田一幸『実証会計学』中央経済社。
- 石村貞夫+デズモンド・アレン[1997]『すぐわかる統計用語』東京図書。
- 乙政正太[2004]『利害調整メカニズムと会計情報』森山書店。
- 会計情報研究会(代表・石塚博司)[1978]「資本市場における会計情報の有効性(パイロット・テスト): 決算情報は投資家行動に対して影響力をもつか」『企業会計』第 30 巻第 13 号, ■-■頁。
- 國村道雄[1979]「会計情報と株価—実証研究—」『會計』第 115 巻第 3 号, ■-■頁。
- 須田一幸[2000]『財務会計の機能—理論と実証—』白桃書房。
- 編著[2004]『会計制度改革の実証分析』同文館。
- 藤井秀樹[2007]『制度変化の会計学—会計基準のコンバージェンスを見すえて—』中央経済社。

ⁱ 正確にいえば, 母分散が既知の場合と, 母分散は未知で等分散の場合とに分類する必要があります。また, この「2つの母平均の差の検定」は, 「対応のない2つの母平均の差の検定」と「対応のある2つの母平均の差の検定」の2種類に分類され, 統計上の処理方法が異なります。さらに, 母集団が正規分布する場合(パラメトリックな場合)と, しない場合(ノンパラメトリックの場合)によっても分析方法が異なります。ここでは, 以上の点を指摘するにとどめておきます。

ⁱⁱ 統計上の推論手続きに正しくしたとえば, この場合, 「アメリカ人と日本人の平均身長には差がない」という仮説を棄却することによって, 「アメリカ人と日本人の平均身長には差がある」という仮説を支持するという道筋で議論を進めることになります。ここにみるような, 棄却できるかどうかをテストするために設定する仮説を, 統計分析では帰無仮説といいます。そして, それに対立する仮説(この場合は, 「アメリカ人と日本人の平均身長に

は差がある」という仮説)は、対立仮説といえます。つまり、統計分析においては、帰無仮説を棄却できれば対立仮説が支持されるという推論手続きをとるわけです。帰無仮説が棄却できるかどうかは、各種の統計量によって判定されることになります。

iii 本来は、片側検定と両側検定、Z検定とt検定、有意性検定の危険性(第1種の過誤と第2種の過誤)などについても触れるべきですが、ここでは説明の便宜上、すべて割愛します。

iv たとえば、テキストAには、試験の問題と類似した練習問題がたまたまあったのかもしれませんが。このような場合、テキストAで学習した高校生の平均点が相対的に高くなるのは半ば当然で、平均点の差はテキストの良し悪しとはあまり関係ない可能性があります。さらにまた、テキストAは、受験対策に徹し、数学の本質的な考え方の解説はほとんどやっていないというような場合には、テキストAで学習することが「良い」かどうかは大変微妙な問題となり、それについて価値判断を下すのは非常に難しくなるでしょう。

v 上記のケースで χ^2 値を計算すると9.6となります。第9章での解説を参考にして計算してみてください。自由度1の χ^2 値では、6.63を超えると1%水準で有意となります。

vi この場合、「関東人と関西人の間にワインの好みについて差がない」という帰無仮説を棄却することによって、「関東人と関西人の間にワインの好みについて差がある」という対立仮説を支持するという道筋で議論を進めることになります。

vii 重回帰分析を行うときにはふつう、自由度修正済みの R^2 (補正 R^2)を利用します。 R^2 には、説明変数の数を増やしていくと、それだけで値が大きくなっていくという性質があります。とくに問題になるのは、少ないサンプルで重回帰分析を行うときです。このようにときに、説明変数の数をむやみに増やすと、 R^2 の上昇幅は非常に大きなものになります。したがって、重回帰分析を行うときには、たんに説明変数を増やしたことで計算技術的に上昇した R^2 の値を、理論的に適正とみなせる水準にまで下方修正する必要があり、そのために補正 R^2 を利用することになるわけです。

そのさい注意しておきたいのは、公式に従って計算すると、補正 R^2 がマイナスになることがあるということです。これは、サンプル数と説明変数の数にもとづいて実施した下方修正の大きさよりも、修正前の R^2 の値が小さい場合に発生する現象で、数学的には、ありえないことではありません。しかし、本文でふれましたように、 R^2 は0から1の間の値をとる指標で、それがマイナスになることは、統計学的にはありえません。にもかかわらず、公式に従って計算した結果、補正 R^2 がマイナスになったということは、分析モデルが統計学的に意味のあるものではなかったということです。分析モデルが統計学的に意味のあるものではないことを証明するのが研究目的とされるような特殊な場合は別ですが、そうでない場合には、リサーチデザインや変数を変更するなどして、分析モデルそれ自体を改善する必要があります。

viii 実証会計は、しばしば実証研究と呼ばれることもあります。しかし、数量データの統計分析に依拠した社会科学的研究は、会計学以外の領域でもさかんに行われています。とくに経済学ではそうです。というよりもむしろ、歴史的には、経済学の一分野である計量経済学の研究手法が会計学に移植されて、実証会計という新しい研究分野が成立したのです。したがって、会計学の領域で行われる実証研究ということを表現する用語としては、実証研究よりも、実証会計のほうが適切であると考えられます。しかし、実証会計は、会計情報を用いて実証研究をしているだけで、実証研究であることに変わりはないと考えることもできます。このテキストでは実証会計という用語を使いますが、この用語問題はあまり重要なものではありません。重要なのは、実証会計の考え方や手法の基礎を正確に学習することです。

ix Ball and Brown[1968]とBeaver[1968]が、実証会計の嚆矢とされています。

x 会計情報研究会[1978]と國村[1979]が、日本における実証会計の嚆矢とされています。

xi 利益連動型報酬契約とは、企業の報告利益に連動させて経営者の報酬を決めるという条

項を織り込んだ契約をいいます。こうした契約は、企業業績の向上を経営者に動機づけるインセンティブ・システムの代表的な事例です。

xiii このような場合で、厳密に言えば、上記の仮説が支持されたことにはなりません。なぜならば、この場合、利益連動型報酬契約の有無が2つのグループへの分離を生み出す理由になっている可能性だけでなく、それ以外の隠れた要因がその理由になっている可能性もあるからです。この2つの可能性を、 χ^2 検定で識別することはできないのです。入門レベルの実証会計では、このような統計的な論理性をあまり強く意識する必要はありませんが、こうした問題があるということだけは知っておいた方がよいでしょう。