



慶應義塾大学ビジネス・スクール 計量分析の見方(基礎編)¹

計量モデルの基本的な考え方

このノートでは、計量分析の基礎的な見方をいくつかの分析例を使って説明します。まず最初の例として、ある製品の価格がその製品の需要にどのような影響を与えるかを考えてみましょう。ここで簡単化のために、しばらくの間、その需要に影響を与える他の要因（ライバル企業の製品の価格等）は考えないことにします。需要と価格に関して、図1のようなデータがプロットされていると仮定します。計量分析の目的は、その需要と価格の関係を最もうまく表す式を求めることです。

5

10

具体的に、需要と価格の関係を次の式で表します。

$$Q = \alpha - \beta P + u$$

15

ここで、 Q は需要量、 P は価格水準、 u は誤差項を表します。 α (アルファ) と β (ベータ) はパラメーターといいます。ここで、 α と β は正の値であると仮定します。すなわち、価格の上昇は需要を低下させると考えています (β の前のマイナスの符号に注目して下さい)。需要の変化を価格の変化で説明しようとするところから、需要量を被説明変数 (結果となる変数)、価格を説明変数 (原因となる変数) といいます。また、

20

$$Q = \alpha - \beta P$$

と表される式は回帰式とよばれ、 α と β はそれぞれ切片と直線の傾きを表しています (図1 参照)。計量分析の具体的な目的は、その需要と価格の関係を最もうまく表すパラメーターを求めることです。

25

¹ このノートは、これまで行われた計量分析の結果を理解するためのものです。実際にデータを使い、これから計量分析をしようという人は、MEノート(7、8)、和合・伴(1995)他を参照して下さい。

本ケースは慶應義塾大学ビジネススクール専任講師中村洋によって作成された。
本ケースの記述は経営管理の巧拙を示すものではなく、クラス討議の資料として作成されたものである。(1996年12月)

30

最小二乗法

ここで、回帰式の最適なパラメーター値を求める最も一般的な方法である最小二乗法を説明します。最も簡単な例として上記の需要と価格の関係を考えてみましょう。最小二乗法は、図1において各点から回帰線（回帰式で示される線）までの長さ（ u ）（=誤差項）の二乗の和を最小にするようにパラメーター値を導き出すというものです。² 具体的な手順は以下の通りです。

（1）ある一定の水準のパラメーター値（ここでは、 α と β の値）を仮定し、各期の説明変数（ここでは価格）のデータを使って、各期の需要量の推計値（ \hat{Q} ）を得ます。³ 次に、各期の現実の需要量（ Q ）と各期の推計値 \hat{Q} との差（ $Q - \hat{Q}$ ）を二乗します。⁴ そして、各期の二乗した結果を合計します。

（2）他にも、無数に数多くの α と β の値を仮定して、それぞれ同様に、 $(Q - \hat{Q})$ の二乗の和を計算します。

（3）その中で、 $(Q - \hat{Q})$ の二乗の和が最小になるような α と β のパラメーター値を「最適値」として選び出します。

複数の説明変数

この例においては、説明変数が価格の1つだけでした。説明変数が2個以上の場合のパラメーターの推定も、概念的には同じです。例えば、ある製品への需要は、その製品の価格のみならず、ライバル企業の製品の価格、広告費、販売地域の平均所得、販売地域の人口にも影響されると考えられます。その時、需要関数は次のように表すことができます。

$$Q = \alpha - \beta P + \gamma P_r + \delta A + \zeta I + \eta W + u$$

ここで、 P_r はライバル企業の製品の価格、 A は広告費、 I は販売地域の平均所得、 W は

² (数学が得意な人のみ参照)

この回帰式の最適なパラメーター値は数学的に次のように表されます。

$$(\alpha, \beta) = \min_{(\alpha, \beta)} \sum_{i=1}^N (Q_i - \alpha + \beta P_i)^2$$

ここで、各データに1からNまで番号が付けられています。Nはデータ数を示します。

³ この推計値が、回帰線上の値となります。

⁴ 例えば、 $\alpha=2$ 、 $\beta=1$ と仮定したとしましょう。ある期の価格が1、需要量が2というデータが与えられた場合、需要量の推計値（ \hat{Q} ）と、

現実の需要量と推計値の差（ $Q - \hat{Q}$ ）の二乗は次のように表されます。

$$\hat{Q} = 2 - 1 \times 1 = 1$$

$$(Q - \hat{Q})^2 = (2 - 1)^2 = 1$$

販売地域の人口を表します。 γ （ガンマ）はライバル企業の製品の価格の、 δ （デルタ）は広告費の、 ζ （ゼータ）は販売地域の平均所得の、 η （イータ）は販売地域の人口のパラメーターをそれぞれ表します。同様に、最適二乗法を使い、これらの説明変数のパラメーターの「最適値」を求めることができます。⁵

5

統計量値

計量分析には、こうして求められた最適なパラメーター値を評価するための統計量値が与えられています。⁶

(1) t 値 (t-value)

t 値は各説明変数の推定パラメータについて評価する統計量値で、その絶対値がある水準よりも高ければ、「有意」といい、「その説明変数のパラメーターが統計的に見てゼロではない」ことを意味します。上記の需要と価格の例では、「 $\beta = 0$ すなわち価格が需要に全く影響を与えない」という仮説が統計的に棄却されること（≒価格が需要に影響を与えている）を意味します。

通常、計量分析には、何%の水準で有意ということが明示されています。例えば、5%の水準で有意ということは、「影響を与えていないのに、間違って影響を与えているとする確率が5%」ということを意味します。何%の水準で有意とするかは、計量分析者の判断です。通常は1、5、10%の水準で判断します。⁷

(2) 決定係数 (R^2 : R-squared)

0以上1以下の数字をとり、説明変数全ての変化で、どれだけ被説明変数の変化を説明できるかを示します。決定係数の値が低ければ、被説明変数の変化を説明できる他の要因が存在すると考えられます。経験的に、変数相互の時間に依存する関係を分析する時系列

⁵ (数学が得意な人のみ参照)

この回帰式の最適なパラメーター値は数学的に次のように表されます。

$$(\alpha, \beta, \gamma, \delta, \zeta, \eta) = \min_{(\alpha, \beta, \gamma, \delta, \zeta, \eta)} \sum_{i=1}^n (Q_i - \alpha + \beta P_i - \gamma Pr_i - \delta A_i - \zeta I_i - \eta W_i)^2$$

⁶ このノートの目的は、これまで行われた計量分析の結果を理解することです。したがって、以下の統計量値をどのようにして求めるかはこのノートでは取り扱いません。

⁷ 何%の水準で有意ということが明示されていない場合は、目安として絶対値で2以上であれば、有意と考えてもかまいません。 厳密には、説明変数の数、データの数によって、判断基準となる値は変わります。

⁸ 例えば、ある企業の製品に対する需要にどのような要因が影響しているのかを十数年のデータを使って分析するのが時系列分析と呼ばれます。一方、ある年の上場企業各社の売上げ総利益率にどのような要因が影響しているのかを分析するのがクロス・セクション（横断面）分析と呼ばれます。計量分析の例1と2を参照してください。

10

15

20

25

30

分析では高く、変数相互の時間に依存しない関係を分析するクロス・セクション分析ではかなり低くです。⁸

(3) ダービン・ワトソン比 (DW : Durbin-Watson ratio)

誤差項 (u) の系列相関を示します。この比率が 2 に近ければ「誤差項は互いに無相関である」と言えます。ここで、「誤差項は互いに無相関である」とは、誤差が正である期間がしばらく続き次は負である期間がしばらく続いたり（正の相関関係）、正と負の誤差がほぼかわるがわる表れたり（負の相関関係）しないということです（図 2 参照）もし、誤差項に相関関係があれば、分析結果は統計的に正しいとは言えません。^{9 10}

10

計量分析の例 1 (時系列分析)

ある食品メーカ製造のフルーツパイへの需要に影響すると考えられる要因を使って、計量分析をした例が次のように示されています。

15

被説明変数： フルーツパイへの需要

説明変数	パラメーター値	t 値
定数	-4516.291	-0.91
自己の価格	-35.985	5.13 (s5)
広告費	203.713	0.64
競争者の価格	37.96	5.37 (s5)
販売地域の平均所得	777.051	11.7 (s5)
販売地域の人口	0.256	1.04

25

注 1：決定係数 = 93%

注 2：(s5) は、5% の水準で有意であることを示します。

この分析から、次のことが言えます。

30 (1) 自己の価格に関する t 値から「5% で有意」であることがわかるので、自己の価

⁹ 最小二乗法が適用可能な条件として、「誤差項は互いに無相関である」ことが必要条件です。

¹⁰ 以下の例でもわかるように、このダービン・ワトソン比はしばしば省略されます。

¹¹ 厳密には、「影響を与えていないとはいえない」と言うのが正しい。

格の上昇は需要にマイナスの影響を与えていていると言える。¹¹

(2) 広告費に関する t 値から「有意でない」ことがわかるので、広告費の増加は需要に影響を与えていないと言える。

(3) 競争者の価格に関する t 値から「5%で有意」であることがわかるので、競争者の価格の上昇は需要にプラスの影響を与えていていると言える。 5

(4) 販売地域の平均所得に関する t 値から「5%で有意」であることがわかるので、販売地域の平均所得の上昇は需要にプラスの影響を与えていていると言える。

(5) 販売地域の人口に関する t 値から「有意でない」ことがわかるので、販売地域の人口の上昇は需要に影響を与えていないと言える。 10

確認のための練習問題

次の文章は正しいか間違っているかを、簡単な理由を付けて示してください。

(1) 自己の価格の t 値よりも販売地域の平均所得の t 値が大きいので、販売地域の平均所得の影響の方が大きい。したがって、これからは平均所得の高い地域に出店するべきだ。 15

(2) 決定係数が高いので、これらの説明変数は非説明変数に影響を与えていると言える。

(3) 販売地域の人口の代わりに販売地域の人口密度を説明変数として使用した場合の、決定係数が 94% としました。販売地域の人口を説明変数として使った場合より決定係数が高いので、人口密度を説明変数として使用した方がよい。 20

(答え)

(1) 間違い。 t 値は影響の有無を判断するもので影響の大小を見るものではありません。 25

(2) 間違い。決定係数は「説明力」を見るもので、影響の有無や大小を見るものではありません。

(3) 間違い。決定係数のわずかな差が、統計的に意味があるかどうかは判断できません。「販売地域の人口」と「販売地域の人口密度」のどちらの説明変数を使うのが望ましいのかは統計的に判断できません。 30

計量分析の例2（クロス・セクション分析）

竹中（1995）は、1989/90年（度）について日本の製造企業85社の売上総利益率に影響を与えると考えられる要因を使って、計量分析をしました。

被説明変数： 売上総利益率

説明変数	パラメーター値	t値	
定数	0.16	3.07	(s1)
広告宣伝	2.30	5.82	(s1)
企業輸出比率	-0.07	-1.31	
成長率	-0.05	-0.28	
輸入シェア	-0.02	-0.60	
固定資産比率	0.08	1.73	(s10)
マーケットシェア	0.21	2.76	(s1)
生産4社集中度	0.03	0.51	
稼働率	-0.04	-0.75	

注1：決定係数=37%¹²

注2：(s1)は、1%の水準で有意であることを示します。

(s10)は、10%の水準で有意であることを示します。

この計量分析から、広告宣伝、固定資産比率、マーケットシェアに関するt値から「有意」であることがわかるので、それらは売上総利益率に影響を与えていると言える。一方、企業輸出比率、成長率、輸入シェア、生産4社集中度、稼働率に関するt値から「有意でない」ことがわかるので、それらの変数は売上総利益率に影響を与えていないと言える。ここで、10%の水準で有意かどうかを判断基準としています。

30

¹² クロス・セクション分析であるので決定係数は低くでています。

¹³ 離散的選択モデルにおいて、最適なパラメーター値の導出には最小二乗法とは異なった手法が使われますが、t値や決定係数の見方は同じです。

離散的選択モデル

最後に、最近マーケティングの分野などでよく使われている離散的選択モデルを簡単に説明します。このモデルは、二者択一の選択あるいは多数の選択肢からの選択を確率の問題としてとらえ、その確率がどのような要因によって影響を受けるかを分析する手法です。¹³ 5 二者択一の選択の例として、自動車を買うか買わないかという選択が考えられます。

程勲（1996）は、アメリカ企業がどこの国に工場を建てるかという多数の選択肢からの選択が、立地経験、人口密度、労働力等にどのような影響を受けるかを多くの産業を対象に分析しました。ここでは、繊維産業に関する分析結果を引用します。

計量分析の例3（離散的選択モデル）

説明変数	パラメーター値	t値
GDP定数	0.145	0.257
労働コスト	-0.129	-0.559
安全度	12.750	2.594 (s1)
人的資本	-0.774	-0.619
一人当たりの面積	-0.794	-2.048 (s5)
立地累積値	1.002	2.122 (s5)
経済集積	-2.460	-1.198
経済集積の二乗値	0.056	0.461
製品輸入比率	-5.086	-1.222
日本ダミー	-2.257	-1.765 (s10)

注1：（疑似）決定係数=24.5%

注2：(s1)は、1%の水準で有意であることを示します。

(s5)は、5%の水準で有意であることを示します。

(s10)は、10%の水準で有意であることを示します。

この結果から、彼は次のような主張を行いました。

(1) 安全性が高い（債務不履行の可能性が少ない）国ほど、直接投資が行われる確率が増える。

(2) 一人当たりの面積が少ない国ほど、投資が行われる確率が増す。つまり、面積と

比べて相対的労働人口の豊富な国を、投資対象国として選択する傾向がある。

(3) 立地累積値（立地経験）が多い国ほど、投資が行われる確率が増す。

(4) 投資対象国が日本であると投資が行われる確率が低下する。このことは日本市場の閉鎖性を暗示している。

5

参考文献

刈屋武昭監修（1984）「計量経済分析の基礎と応用」東洋経済新報社

竹中康治（1995）「寡占における過剰設備」植草益 編「日本の産業組織」第8章

程勲（1996）「米国対外直接投資の決定要因と日本市場の閉鎖性について」日本経済研究

10 和合・伴（1995）「T S Pによる経済データの分析」東京大学出版会、

15

20

25

30

図1

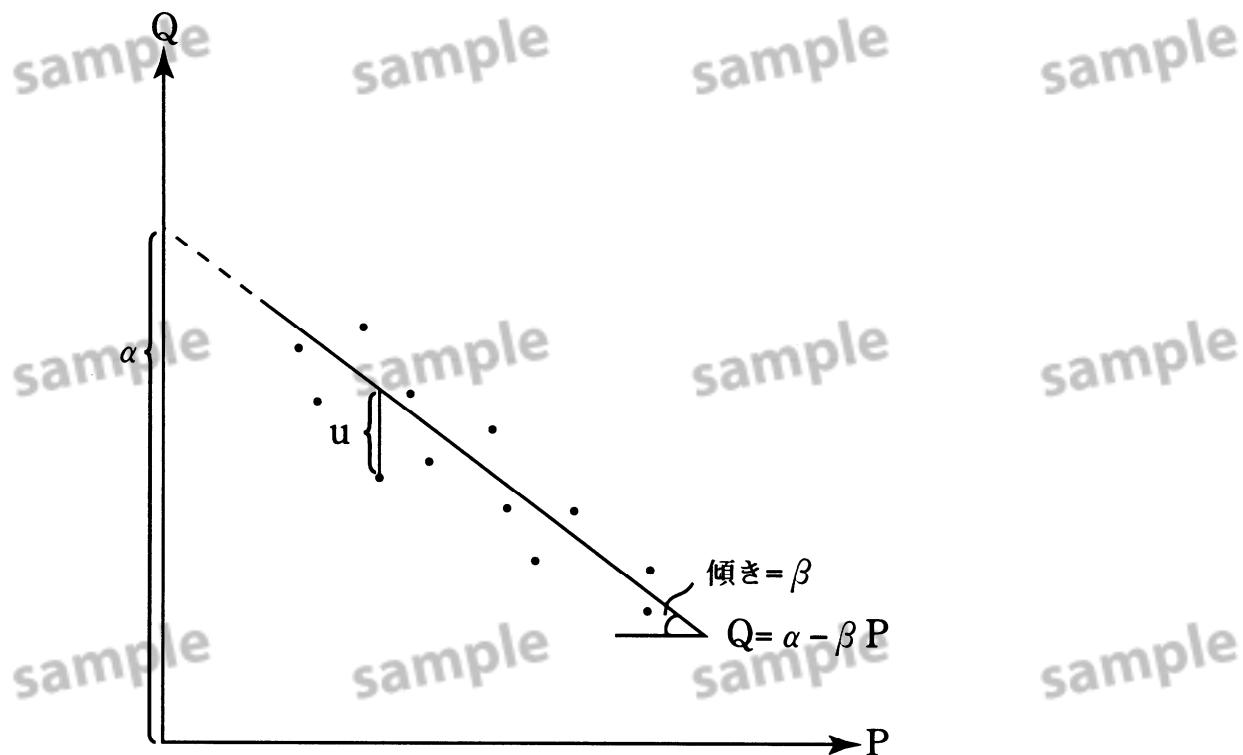
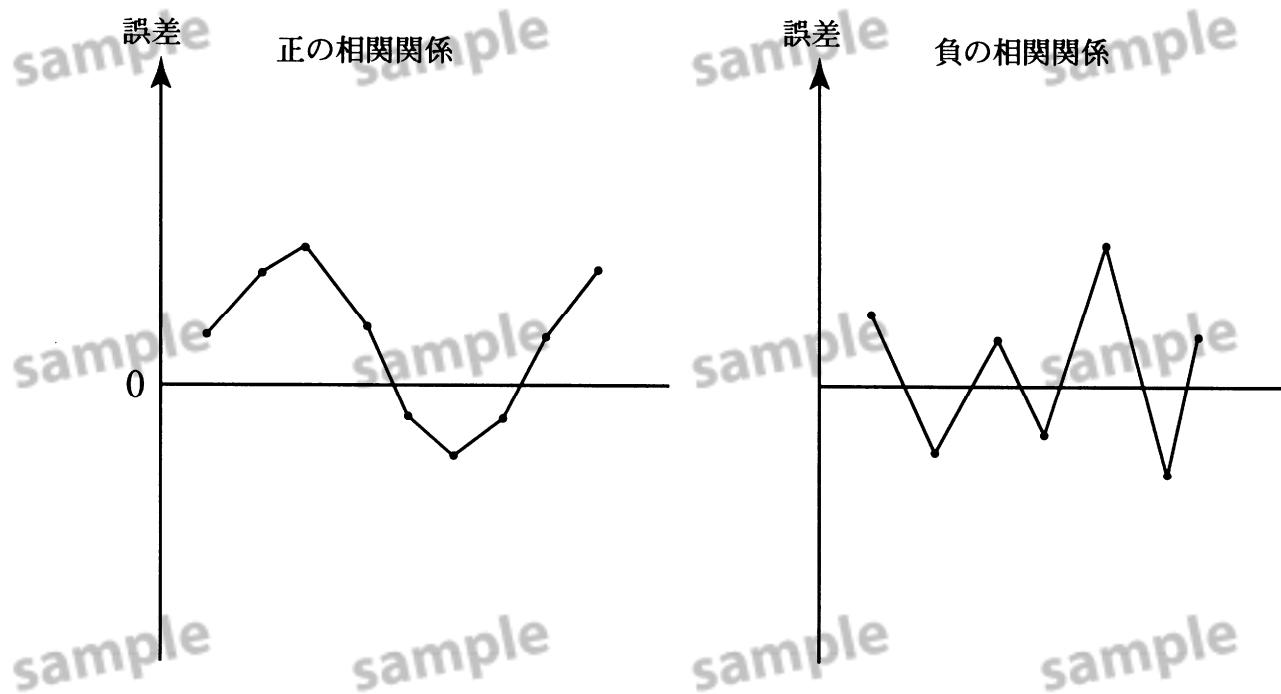


図2



不許複製

慶應義塾大学ビジネス・スクール

Contents Works Inc.