

「回帰分析 (1): 考え方」

■回帰分析の目的と魅力

今回からは、**回帰分析** (regression analysis) について解説する。回帰分析は、ある 1 つの変数 (従属変数) の値を、他の変数 (独立変数) の値で説明しようとするときに、もっとも頻繁に利用される分析技法である。たとえば、ある大学の先生が、学生の遅刻に頭を悩ませているとする。遅刻の原因として、アルバイトのやりすぎによる疲れがあるのではないかと考え、15 人の学生に調査をしたとする。1 ヶ月のアルバイト時間を X 軸に、遅刻回数を Y 軸にして図 1 のような散布図を描くと、その関係性がわかる。このとき、散布図の上に直線を引いてみたくなることがある。

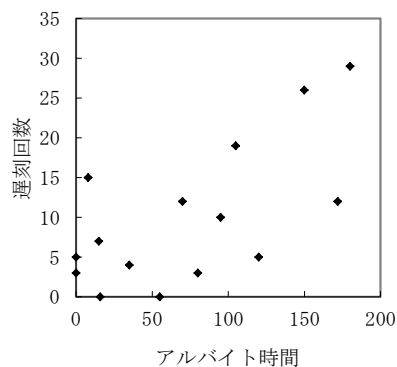


図 1 アルバイト時間と遅刻回数の関係 (仮想データ)

このような直線を引きたくなるのは、次のように考えているからである。「X と Y の関係は、本来この直線のような関数で表せるのではないだろうか。実際のデータがこの直線からいくらかずれているのは、何らかの誤差によるものだろう」と。より定式的に書けば、「本来の Y の値は、X の値から $\hat{Y} = \alpha + \beta X$ という直線の関数で表せる (\hat{Y} は実際の Y の値ではなく、予測値としての Y の値を表す)」と考えている。 α は直線と Y 軸が交わる切片を表し、 β は直線の傾きを表す。 α や β は定数なので、具体的には $\hat{Y} = 4.5 + 0.1X$ といった形で Y の予測式は表される。上のような予測式のことを**回帰式** (regression equation) と呼び、回帰式によって表される線のことを**回帰線** (regression line) と呼ぶ。また、回帰式の α を定数項、 β を**回帰係数** (regression coefficient) と呼ぶ。

回帰分析の目的は、回帰線を最適に調整することを通して、ある変数の値が、その原因と考えられる変数によってどのように説明できるのかを統計的に明らかにすることである。原因と考える X のことを**独立変数** (independent variable)、結果と考える Y のことを**従属変数** (dependent variable) と呼ぶ。X と Y の間に何らかの因果関係を想定して、具体的に X が Y にどの程度の影響を与えるのかを調べることは極めて一般的な関心であり、その疑問に正面から答えを出してくれることが回帰分析の魅力である。

回帰分析の重要な手続きは、次の3点にまとめられる。順に説明しよう。

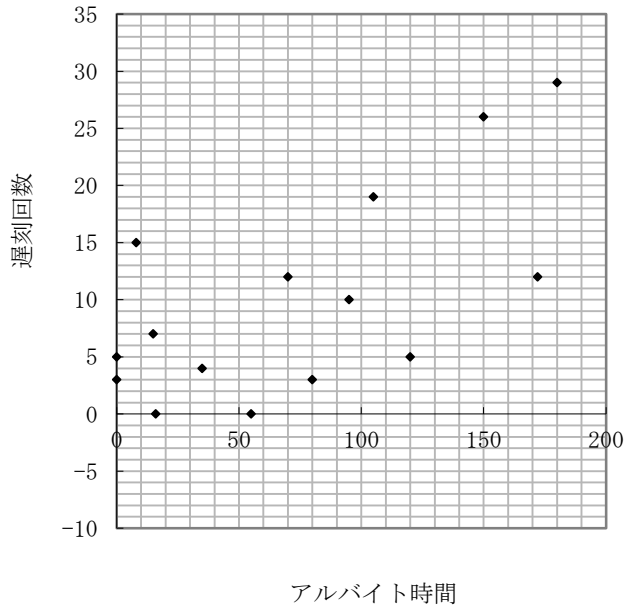
- (1) もっともよい線を引く。
- (2) その線はどのくらいよい線であるかを評価する。
- (3) 母集団についても同様の線を引く価値があるかどうかを判断する。

■最適な回帰式を推定する

回帰分析の最初の手続きとして、散布図の中にもっともよい回帰線を引かなければならない。もっともよい回帰線とは、実際のデータと予測値との差、つまり $Y - \hat{Y}$ (残差と呼ぶ) の合計がもっとも小さくなる線である。ただし、残差がプラス側かマイナス側であるかは、問題ではないので、残差を2乗した値を用いて、その合計値が最も小さくなるようにする。この合計値を**残差平方和** (residual sum of squares) と呼ぶ。残差平方和が小さいほど、その回帰線はよい回帰線と考える。

(練習) ※次ページをみないように！

1. 自分が最適だと思う直線を、散布図の上に定規で引いてみよう。
2. その直線の切片と傾きを読み取って、式に表わしてみよう。



【自分が引いた回帰線】

$$\hat{Y} = \alpha + \beta X$$

↓

$$\hat{Y} = \underline{\hspace{2cm}} + \underline{\hspace{2cm}} X$$

3. 自分が引いた直線について、残差平方和を求め、周りの人と比較してみよう。(残差平方和が小さいほどよい回帰線ということになる)

	アルバイト時間 X	遅刻回数 (観測値) Y	自分が引いた直線		
			予測値 \hat{Y}	残差 $Y - \hat{Y}$	残差平方 $(Y - \hat{Y})^2$
1人目	55	0			
2人目	35	4			
3人目	180	29			
4人目	172	12			
5人目	150	26			
6人目	8	15			
7人目	80	3			
8人目	95	10			
9人目	0	3			
10人目	15	7			
11人目	16	0			
12人目	120	5			
13人目	105	19			
14人目	70	12			
15人目	0	5			

(合計) ↓

残差平方和 = _____

目分量で適当に引いても、そこそよい回帰線が引けると思われるが、数学的には微分方程式を解くことで最適な線を導くことができる。このように数学的に最適な回帰線を求めることを**最小二乗法** (method of least squares) と呼ぶ。

数学的な詳細は省略するが、方程式を解くと、具体的にいまのデータの場合には、 $\beta = 0.095$ 、 $\alpha = 3.01$ が最適である。つまり、 $\hat{Y} = 3.01 + 0.095X$ という回帰式最適である。このとき、残差平方和は 632.13 になり、他にどんな回帰線を考えても、これよりも小さな残差平方和をとることはない。

この回帰線から、次のように具体的な意味を読み取れる。アルバイトをしていない場合 (X が 0 の場合) は遅刻の回数が 3.01 回と予測され、アルバイト時間が 1 時間増えるごとに、0.095 回ずつ遅刻の予測回数が増える。この予測が統計的に最適である。

■ 回帰線の説明力を評価する

最小二乗法によって、最適な回帰線は求まる。しかし、最適な回帰線であったとしても、従属変数の予測に十分な説明力 (予測力) を持つとは限らない。もともと独立変数に従属変数を説明する力がない場合には、最善を尽くしても十分な説明ができるはずはないからである。そこで、2 つ目の手続きとして、その回帰線はどのくらいよい線であるか、説明力の強さを評価する。

回帰線が持つ説明力の評価は、一般に**決定係数** (coefficient of determination) によってなされる。決定係数は 0~1 (0%~100%) の値を取り、従属変数の個人差を独立変数の値でどれだけ説明できるか、その割合を表す。

決定係数は次のような考え方に基づいている。いま、従属変数 (Y) の予測のために独立変数 (X) の情報を用いることができないとしよう。つまり、1 人 1 人のアルバイト時間が分からない中で、遅刻回数をなるべくずれが少ないように予測することを考える。このとき、最適な予測方法は、常に Y の平均値を予測値として用いることである (図 2)。

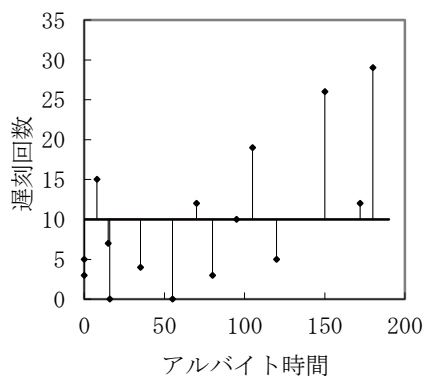


図 2 独立変数を利用しない最善の予測

このときの残差平方和を出発点として、最適な回帰線を用いると残差平方和が何%減少するかを算出したものが決定係数である。今回のデータの場合、遅刻回数をすべて平均値で予測すると、残差平方和が 1124.00 になる。最適な回帰線による残差平方和は 632.13 だったわけなので、この回帰線によって残差平方和は 491.87 だけ改善した。これは、もとの残差平方和の 43.8%にあたる ($491.87/1124.00=0.438$)。つまり、決定係数 $R^2=0.438$

で、遅刻回数の 43.8%がアルバイト時間によるこの回帰式で説明できることを意味する。

ただし、実際のデータ分析では、さらに調整を加えた**調整済み決定係数** (adjusted R²) を記す。標本の回答者数 n が小さいときには、決定係数は母集団における実際の説明力よりも大きくなる偏りを持つ。そこで、その偏りを取り除いて決定係数をやや小さく調整し直すわけである。この辺りの理屈は数学的なことなので、あまり深く理解する必要はない。今回の回帰分析の場合、決定係数は 0.438 だが、調整済み決定係数は 0.395 となる。結局、遅刻回数の個人差の 39.5%がアルバイト時間の違いによって説明できることが分かる。

決定係数がどのくらい大きければ十分なのか、明確な基準はない。学問分野や分析対象、分析目的によって必要な説明力は異なるからである。一般的には、社会調査のデータ分析で求められる説明力（決定係数）の水準は、あまり高くないことが多い。10%を切っても有意義な分析とみなされることも珍しくはない。

■ 回帰分析の結果を統計的に検定する

最後に残された手続きは、この回帰分析の結果が統計的に有意かどうかを検定することである。先ほどの分析では、アルバイト時間が 1 時間増えるごとに遅刻回数が 0.095 回ずつ増えること、そしてこのように考えると遅刻回数の個人差が 39.5%説明できることが示された。しかし、これはたまたま選ばれた 15 人の結果であって、母集団全体では実はアルバイト時間と遅刻回数はまったく関係がない、ということもありえる。そのため、この結果が確率的な偶然の産物ではないと考えてよいのかどうか、確かめなければならない。標本調査の結果は、常にこのような検定の手続きを必要とする。

一般的に、社会科学系の統計分析では、その分析結果が偶然に生じる確率が 5%未満であれば、その結果が偶然ではなく統計的に有意とみなす。このような確率を**有意確率** (probability of significance) と呼び、記号「p」で表す。p<0.05 であれば検定に合格なので、その分析結果を母集団にも当てはめて考えてよい、というルールである。

今回の例の場合は、p=0.007 と算出される。つまり、たまたま偏った人が調査されてアルバイト時間と遅刻回数が強く関係しているような結果が出ただけだ、という確率は 0.7%しかない。こんな偶然が起こったとは考えにくいので、15 人しか調べていないとはいえ、この回帰分析の結果は十分に信用できるということである。

ただし、有意確率を本当に計算するのはコンピューターを用いてもかなり難しい。そこで、実際には検定統計量と呼ばれる簡単に計算できる数値を用いて、およその有意確率を算出している。検定統計量の算出式は、何を検定したいかによって決まっており、回帰分析の結果が母集団を十分に説明できるかどうかを検定するときには、下式のような**F 値**と呼ばれる検定統計量を用いる（今回の分析の場合、F=10.13）。

$$F = \frac{R^2}{(1-R^2)/(n-2)} = \frac{0.438}{(1-0.438)/(15-2)} = 10.13$$

そのため、実際に読み取るのは有意確率だけであるが、レポート等では有意確率を概算するために用いた検定統計量の数値も記しておくことが、一種のマナーになっている。

(練習)

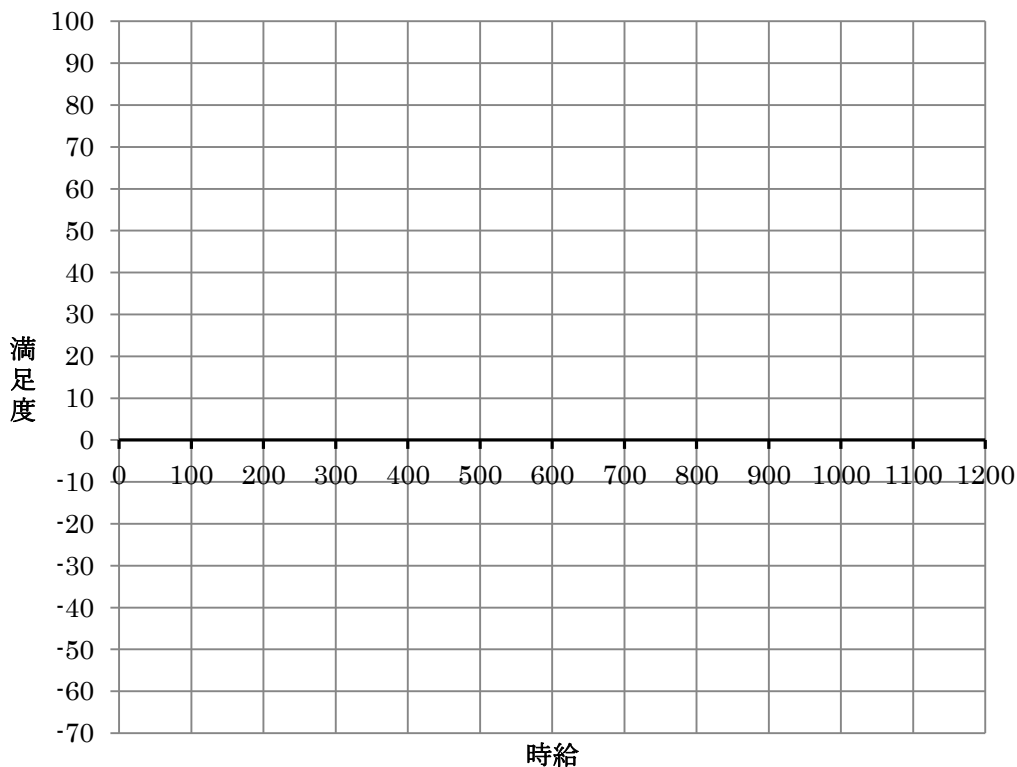
回帰分析の結果が以下のとおりであった場合、具体的にどういう意味が読み取れるか、穴埋めしてみよう。

○飲食店のアルバイト店員 50 名に対するアンケートデータを用いた回帰分析

- ・従属変数は「アルバイトへの満足度 (100 点満点)」
- ・独立変数は「アルバイトの時給」
- ・回帰分析の結果、定数項 $\alpha = -55.8$ 、回帰係数 $\beta = 0.13$
- ・調整済み決定係数 $R^2 = 0.113$
- ・F 値を検定統計量とした検定の結果、有意確率 $p = 0.0098$

↓

回帰分析で求められた最適な回帰式は、 $\hat{Y} =$ _____ で、回帰線をおよそのグラフで図示すると、下のようになる。具体的には、たとえば時給が 800 円ときの満足度は _____ 点と予測されるのに対して、時給が 1000 円ならば、満足度 _____ 点に上昇する。



また、この結果から、アルバイト満足度の個人差は時給の違いでおよそ _____ % 説明できることがわかる。この 50 名の回帰分析の結果を母集団にも一般化してよいかというと、偶然このような結果が得られる確率 (有意確率) が _____ % なので、この分析結果は統計的に有意と {いえる・いえない}。

「回帰分析 (2) : SPSS で実践」

■ SPSS でやってみよう

前回は、回帰分析の考え方について学習した。今回は SPSS を操作して、実際に回帰分析の結果を出力しながら、一通りの手続きを経験しよう。

回帰分析の操作

- ①メニューから、分析 → 回帰 → 線型
- ②説明したい変数 (Y) を [従属変数]、説明に使う変数 (X) を [独立変数] 欄へ移動
(②' 質的変数を独立変数にする場合は、あらかじめダミー変数に変換すること)



- ③ **OK** ボタン

モデル要約 ②

モデル	R	R2 乗	調整済み R2 乗	推定値の標準誤差
1	.662 ^a	.438	.394	6.97320

a. 予測値: (定数)、work アルバイト時間。

分散分析^a ③

モデル		平方和	自由度	平均平方	F 値	有意確率
1	回帰	491.868	1	491.868	10.115	.007 ^b
	残差	632.132	13	48.626		
	合計	1124.000	14			

a. 従属変数 late 遅刻回数
b. 予測値: (定数)、work アルバイト時間。

係数^a

モデル		標準化されていない係数		標準化係数	t 値	有意確率
		B	標準誤差	ベータ		
1	(定数)	3.009	2.841		1.059	.309
	work アルバイト時間	.095	.030	.662	3.180	.007

a. 従属変数 late 遅刻回数

読み取るポイント

- ① 最適な回帰式の α 、 β
- ② 調整済み決定係数
- ③ 全体的な検定結果

(重回帰分析の場合)

- ④ 各独立変数の影響力の検定結果

■独立変数が複数の場合の回帰分析

ここまでは、独立変数が1つの回帰分析を扱ったが、一般的には複数の独立変数を用いた回帰分析がよく行われる。独立変数が複数の場合を**重回帰分析** (multiple regression analysis) と呼ぶこともあるが、回帰分析といえばふつうは重回帰分析のことである。

独立変数が複数ある場合の回帰式は、次のようにどんどん独立変数の効果を足し合わせていく形で表現される。

$$\hat{Y} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots$$

これはつまり、独立変数の値が1増加することは、(他の要素とは関係なく)常に一定の影響力で従属変数の値に作用する、つまり傾きが一定である、という考え方を踏襲している。図形として視覚化することはできないが、多次元空間の散布図の中に1本の最適な線を通して、常に一定の法則が働いていることを主張しようとしていることを意味する。重回帰分析の回帰係数 (β_1 、 β_2 、 β_3 、……) は、とくに偏回帰係数と呼ぶこともある。

具体的には、たとえば遅刻回数 Y を、アルバイト時間 X_1 、通学時間(分) X_2 、睡眠時間 X_3 で説明しようとする重回帰分析では、次のような形で最適な回帰式が析出される。

$$\hat{Y} = 10.21 + 0.22X_1 + 0.04X_2 - 1.31X_3$$

この場合、アルバイトが1時間増えるごとに遅刻が0.22回増え、同様に通学時間が1分長いごとに0.04回遅刻が増える。睡眠時間が1時間長いごとに遅刻は1.31回減る。すべての独立変数が0ならば、遅刻は10.21回と予測される。独立変数が1つの場合と、読み方はまったく同じである。

分析の手続きもほぼ同様であり、以下の4点にまとめられる。

(1) もっともよい回帰式を定める。

(最小二乗法で、 α 、 β_1 、 β_2 、…… β_k の値を定める)

(2) その回帰式は、どのくらいよい式であるかを評価する。

([調整済み] 決定係数によって、説明力を算出する)

(3) 母集団についても、同様の回帰式を定める価値があるかどうか判断する。

(全体的な説明力をF値によって検定する)

(4) 母集団についても、各独立変数を説明に用いる価値があるか、個別に判断する。

(それぞれの独立変数の影響をt値によって検定する)

4つ目の手順だけが重回帰分析に独自のものである。回帰式全体の説明力について検定するだけでなく、1つ1つの独立変数が従属変数を説明するために有効に働いているかどうか、それぞれの影響について検定する。つまり、それぞれの回帰係数 β_1 、 β_2 、……について、母集団でも一定の影響力がある ($\beta \neq 0$) といっただけをどうかを検定する。

この検定は、t値と呼ばれる検定統計量を用い、統計分析ソフトでは対応する有意確率が同時に示される。ここでの有意確率は、つまり、それぞれの独立変数の影響力(回帰係数)がまったくの偶然の産物である確率である。この確率が一定の値(ふつう「.05」つま

り 5%) よりも低ければ、偶然ではなく母集団でもその独立変数による影響力があるとみなしてよいことになる。

(練習)

1. 実際の全国調査 (JGSS-2000) から抽出した 30 代男性のデータを用いて、月給を従属変数、年齢を独立変数とする (月給の違いを年齢で説明する) 回帰分析を実行してみよう。

→読み取るポイント

①最適な回帰式

②調整済み決定係数

③全体的な検定結果

2. 次に、独立変数を、「年齢」「勤続年数」「中 3 の頃の成績」の 3 つに増やして、月給を説明する重回帰分析を実行してみよう。

→読み取るポイント

①最適な回帰式

②調整済み決定係数

③全体的な検定結果

④各独立変数の影響力の検定結果

「回帰分析 (3) : 発展」

■ 質的変数を独立変数にする場合 : ダミー変数

回帰分析の独立変数は量的変数であることが基本である。しかし、質的変数も工夫をすれば独立変数として分析に用いることができる。社会調査データには質的変数が多いので、この応用は重要である。

回帰分析で質的変数を用いる場合には、ダミー変数に変換した上で用いる。ダミー変数とは、0 か 1 のどちらかの値しか取らない変数のことである。たとえば、性別という変数を独立変数に用いたいときには、図 1 のように男性を 1 とするダミー変数 (男性ダミー) か、女性を 1 とするダミー変数 (女性ダミー) のいずれかにリコーディングし、そのダミー変数を回帰分析に用いる。

	元の変数	→	男性ダミー	または	女性ダミー
男性	1	→	1		0
女性	2	→	0		1

図 1 性別のダミー変数

ダミー変数を用いた回帰式の読み取りは簡単である。たとえば、Y が遅刻回数、X₁ が学年、X₂ が男性ダミーの重回帰分析で次のような回帰式が求められたとする。

$$\hat{Y} = 2.0 + 3.9X_1 + 2.2X_2$$

この場合、男子学生は女子学生に比べて 2.2 回多く遅刻することが読み取れる。

性別は 2 つのグループしかない質的変数であったが、3 つ以上のグループ (カテゴリー) がある質的変数の場合はどうすればよいのだろうか。たとえば、学生が所属する学部を独立変数に用いたが、学部は文学部、法学部、工学部、医学部と 4 種類あるとする。この場合、図 2 のように 3 つのダミー変数を作成し、これらすべてを独立変数に用いた回帰分析を行えばよい。

	元の変数	→	文学部 ダミー	法学部 ダミー	工学部 ダミー
文学部	1	→	1	0	0
法学部	2	→	0	1	0
工学部	3	→	0	0	1
医学部	4	→	0	0	0

図 2 学部のダミー変数

もう1つ医学部ダミーが必要ではないかと思うかもしれないが、4つ目のダミー変数は不要である。なぜならば、文学部ダミー、法学部ダミー、工学部ダミーの値がいずれも0である回答者は、自動的に医学部なので、3つのダミー変数さえあれば4つの学部のどれに所属しているか区別できるからである。一般にk個のグループ（カテゴリー）の質的変数の内容は、1つ少ないk-1個のダミー変数で表すことができる。ここでは、医学部ダミーを除いているが、医学部ダミーを分析に加えて他の3つのダミー変数のうち1つを分析から除いてもかまわない。

このようなダミー変数の回帰係数は、省略したカテゴリー（ここでは医学部）と比べて、当該のカテゴリーであることがもたらす影響力を表すことになる。たとえば、文学部ダミーの回帰係数が1.2であれば、それは「医学部と比べて」文学部の方が1.2回だけ遅刻が多いと予測されることを意味する。「文学部以外と比べて」という意味にはならないので注意しよう。

したがって、ダミー変数を省略したカテゴリーは、比較の基準になるという意味で意外と重要な意味を持つ。このようなカテゴリーを**参照カテゴリー [基準カテゴリー]** (reference category) と呼ぶ。いまの例の場合には、医学部が参照カテゴリーである。

参照カテゴリーは、分析者が結果の読み取りやすさを考えて選ぶもので、決まった選び方はない。しかし、次の2点に注意する必要がある。1つは、参照カテゴリーは内容のはっきりとしたグループでなければならない。たとえば、「その他」というグループを参照カテゴリーにすると、何と比べているのか分からなくなるので避ける。もう1つの注意点として、参照カテゴリーのグループに属する回答者は、ある程度人数が多いことが望ましい。あまりに人数が少ないグループを基準にして比較をすると、分析結果が不安定なものになってしまう。

SPSSでは、「他の変数への値の再割り当て」という機能を使って、ダミー変数を作成することができる。ややめんどろであるが、質的変数を回帰分析に活用するためには必要な作業である。

■標準化回帰係数

重回帰分析では、複数の独立変数の中で「いったいどの変数が一番強い影響力をもつか」といったことに関心が向くことがある。単純に回帰係数を比べるだけでは、この疑問に答えることはできない。たとえば、1日の歩行量が1歩増えるごとに、体重が1.5g小さく ($\beta_1 = -1.5$)、1ヶ月にジムに通う回数が1回増えるごとに、体重が500g小さい ($\beta_2 = -500$) としても、ジムに通う回数の方が体重に強く影響するという事にはならない。ジムに通う回数はせいぜい10回程度の話であるが、歩数は人によって数千歩の話だからである。

このような比較をおこなうときに有効なのが、**標準化回帰係数** (standardized regression coefficient) である。標準化回帰係数は、通常回帰係数に独立変数と従属変数の標準偏差の比を掛け合わせたもので、すべての変数を標準得点にしたとき（標準偏差を1に調整したとき）、独立変数が1点増えることが従属変数を何点増やすことになるのかを表す。つまり、すべての変数のばらつき（個人差の程度）をそろえることで、各独立変数の効果を比較できるようにしている。

たとえば、体重の標準偏差が 10,000g (10kg)、歩行量の標準偏差が 2,000 歩、ジムに通う回数の標準偏差が 3 回だったとすると、それぞれの独立変数の標準化回帰係数は、次のようになるので、歩行量の方が影響の規模が大きいことが分かる。

$$\beta_1^* = -1.5 \times \frac{2 \ 0 \ 0 \ 0}{1 \ 0 \ 0 \ 0} = -0.3, \quad \beta_2^* = -500 \times \frac{3}{10000} = -0.15$$

SPSS では「標準化係数ベータ」という列に、自動的に各独立変数の標準化回帰係数が表示されるので、とくに苦勞なくこの値を用いることができる。

■独立変数の出し入れ

重回帰分析では、同じ独立変数でも、他にどのような独立変数を投入したのかによって、回帰係数が変わってくる。たとえば、性別（男性ダミー）と年齢で月給の額を説明しようとしたとき、男性ダミーの回帰係数が 10 万だったとする（男性の方が月給が 10 万円高い）。しかし、これに加えて、正規雇用ダミーを独立変数に加えると、男性ダミーの回帰係数が 5 万に減少したりすることがある。

これは、重回帰分析が「ワンセットの独立変数で」従属変数を説明する回帰線を求めるからである。つまり、「性別と年齢だけで説明しなさい」と言われれば、性別の効果が大きいという説明をせざるをえないが、「正規雇用という原因で説明してもいいよ」と言われれば、性別が男性だからという理由で説明するよりも、正規雇用のおかげで月給が高いと説明する方が適切だ、という解答を回帰分析は示してくれる。

このようなことが起こるのは、そもそも性別と正規雇用の間に強い関連性があるからである（男性の方が正規雇用が多い）。独立変数群の中に関連性の強い変数の組み合わせがあるときには、その回帰係数に注意して、一方の変数を出し入れしてみると、回帰分析の結果がどう変わるかを観察してみよう。扱っている現象に対する理解が深まるはずである（見せかけの関係や媒介関係といった統計的な現象を熟知していれば、理解はより深まる）。

また、このことからわかるように、回帰分析はあくまで分析者が提示したモデル（変数間の因果関係の枠組み）の中で最適な答えを出しているにすぎないことを、忘れないようにしなければならない。回帰分析が「正しい因果関係」を示してくれるわけではない。分析者が想定した因果関係の枠組みの中で、各独立変数の具体的な影響力の大きさ（回帰係数）について最適解を知らせてくれるだけである。したがって、回帰分析はある程度そのメカニズムが理解できている社会現象について、より詳細な情報を得るために用いるべきである。

<参考文献>

岩井紀子・保田時男, 2007, 『調査データ分析の基礎』 有斐閣.

村瀬洋一・高田洋・廣瀬毅士, 2007, 『SPSS による多変量解析』 オーム社.

小田利勝, 2007, 『ウルトラ・ビギナーのための SPSS による統計解析入門』 プレアデス出版.

■（参考）関連する SPSS の操作

リコーディング（値の再割り当て）

リコーディングは、既存の変数の数値（コード）を新しいルールで置き換えて、新しい変数を作る作業である。一般に、調査データの分析を実践するためには、極めて頻繁にリコーディングが必要になる。リコーディングは、主に2つの場面で用いられる。

場面 A：既存のコードを数量として扱いやすいコードに置き換える（とくに大小の逆転）。

例) 1 満足 2 やや満足 3 やや不満 4 不満

→ 4 満足 3 やや満足 2 やや不満 1 不満（この方が結果が読みやすい）

場面 B：既存の細かいコードを、必要十分な大まかな分類でくくり直す。


例) 1フルタイム雇用 2 自営 3.パート・アルバイト 4 派遣 5 内職

↓
1 正規労働

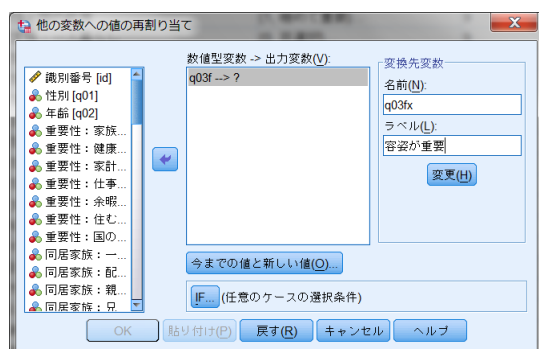
↓
0 非正規労働


いずれの場合も、SPSS の操作は以下のとおりである。

①メニューから、変換 → 他の変数へ値の再割り当て

②リコーディングをしたい変数を左から選択して  ボタン

（※複数の変数を同じルールでリコーディングする場合は、複数選択）

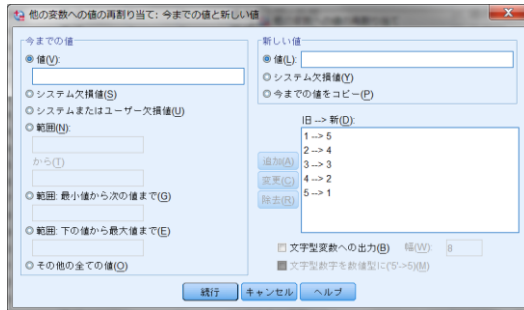


③リコーディング後の新変数の [名前] と [ラベル] を入力して、 ボタン

（※名前はアルファベットの形式変数名、ラベルは変数の内容がわかる日本語）

④ **今までの値と新しい値** ボタンで変換ルールの窓を開く

⑤ 1つ1つの変換ルールについて [今までの値] と [新しい値] を入力して **追加** ボタン
(※ルールの数だけ、これを繰り返す)



⑥ すべてのルールがそろったら **続行** ボタン
(※値を変更しない場合でも、ルールを入れないと空データになるので、必ず全部指定)

⑦ 元の窓で **OK** ボタン

⑧ リコーディングで作成された新しい変数を使って、やりたかった分析を行なう

(連絡)

前の授業で口頭で知らせた通り、「SPSS 基礎」の 10 ページの問題を全部できた人は素点を 10 点加算するので、申告すること。最後の授業まで、いつでも受け付ける。