

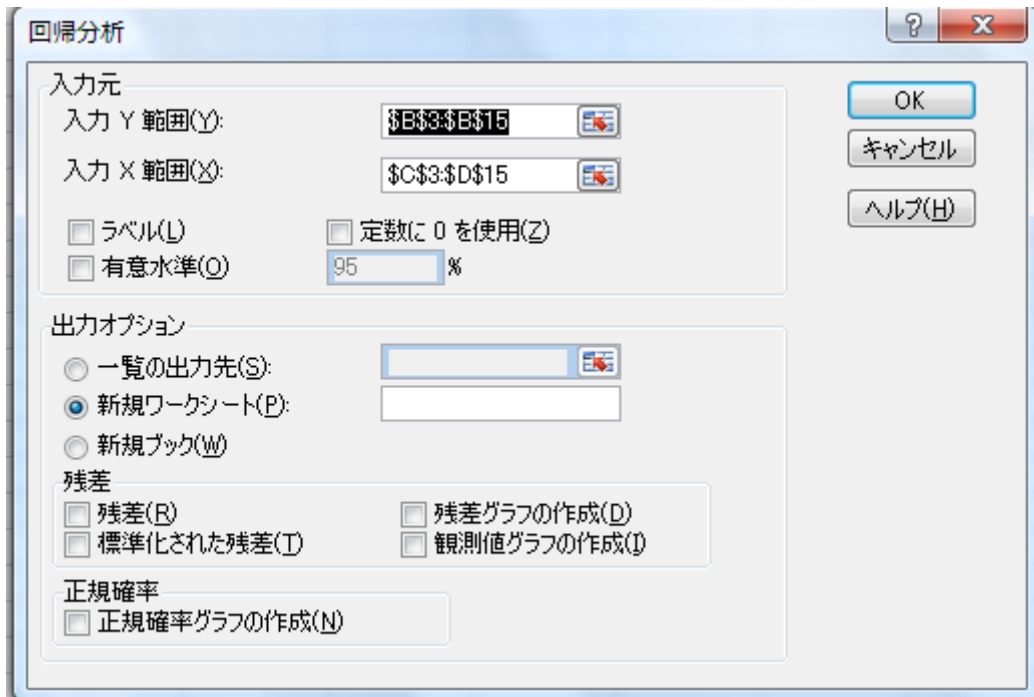
Excel による重回帰分析

教科書、例題3-1のデータに基づき、日本経済の消費関数の推定を行う。ここで、被説明変数は消費(c)、説明変数は、所得(y)と利子率(r)である。

	A	B	C	D	E
1	例題3-1のデータ				
2	year	c	y	r	
3	1994	261,824.60	469,969.10	2.1287	
4	1995	266,649.00	479,181.40	1.1562	
5	1996	273,359.90	492,340.10	0.4117	
6	1997	275,515.20	500,072.30	0.4333	
7	1998	272,523.90	489,824.10	0.3575	
8	1999	274,876.60	489,130.00	0.0375	
9	2000	277,379.50	503,119.80	0.0950	
10	2001	281,913.20	504,047.50	0.0505	
11	2002	284,753.80	505,369.40	0.0010	
12	2003	285,569.30	512,513.00	0.0010	
13	2004	289,964.90	526,577.70	0.0010	
14	2005	294,337.10	536,557.20	0.0010	
15	2006	296,900.00	548,245.90	0.0993	
16					

エクセルにおいて、メニューバーの「データ」→「データ分析」→「回帰分析」を選択すると、次の画面が現れる。データ範囲、出力範囲を設定し、「OK」ボタンをクリックする。

参考: EXCEL2003 の場合は、メニューバーの「ツール」→「分析ツール」→「回帰分析」を選択すると、同様の画面が現れる。



図中の「入力Y範囲」には、被説明変数、「入力X範囲」には説明変数を示す範囲を入力する。入力X範囲は、C3 から D15 の範囲であり、入力Y範囲は、B3 から B15 の範囲である。ラベルには変数の名称をデータ範囲に含めて指定した際にチェックする。データ範囲にラベルを含めていないのにチェックすると、分析結果が違ってくるので注意する必要がある。出力先を指定したら、「OK」ボタンをクリックして、分析を行う。

実行すると、Excel は以下の分析結果を出力する。

「係数」のところには切片と傾きの値が示されている。推定された回帰式は、次のようになる。

$$C = 78203.078 + 0.402Y - 3032.417R, \quad \text{adj}R^2 = 0.967$$

$$(4.581) \quad (12.080) \quad (-2.543)$$

ここで、括弧内の数値はt値である。エクセルのアウトプットにある「補正 R2」は授業中で説明した「自由度修正済み決定係数(adjR²)」に対応し、当てはまりの良さを示す。決定係数は0から1の間の数値となり、1に近いほど当てはまりが良いことを示している。上の例では、adjR²=0.967であり、当てはまりが良いことが分かる。また、「標準誤差」(1909.051)は授業中に説明した回帰の標準誤差(s)を示している。

係数の欄の t 値, p 値は「H₀:係数が0である」という帰無仮説を検定するために用いられる。検定結果から帰無仮説が棄却できない、すなわち統計的に0でないといえない(0かもしれない)となると、説明変数と被説明変数の関係がないことになる。したがって、説明変数を含める意味がなくなってしまう。この検定は「有意性の検定」と呼ばれる。判断の仕方は以下の通りである。

(方法1) p 値 < 分析者が設定する有意水準 → 帰無仮説を棄却

(方法2) 検定統計量(t)の絶対値(|t|) > 分析者が設定した有意水準に対応するt分布の臨界値 → 帰無仮説を棄却

P値が求められている場合には、(方法1)で判断するほうが簡単である。上の例では、所得(X 値 1)の p 値は 0.0000 と得られているので、0.0000 < 0.05 より、有意水準5%で「所得の係数は0である」という帰無仮説は棄却され、所得は有意な説明変数であることがわかる。同様に、利子率(X 値 2)の p 値は 0.0292 と得られているので、0.0292 < 0.05 より、有意水準5%で「利子率の係数は0である」という帰無仮説は棄却され、利子率は有意な説明変数であることがわかる。

参考: 決定係数について

分散分析表において、「変動 合計」の部分(14行C列)は被説明変数の全変動を示している。これに対して、「回帰 変動」の部分(12行C列)は回帰モデルで説明できる変動の大きさ、「残差 変動」の部分(13行C列)は回帰モデルで説明できない変動

の大きさを、それぞれ、示している。決定係数 R^2 は全変動に占める下記モデルで説明できる変動の大きさなので、

$$R^2 = 1271588353.119 / 1308033110.820 = 0.972$$

となり、「重決定 R^2 」(5行B列)の値と一致する。しかし、重回帰分析では、説明変数を増やすにつれて決定係数の値は増加する。そこで、自由度修正済み決定係数 ($\text{adj } R^2$)を用いる。これは、6行B列に示されている「補正 R^2 」の値であり、 $\text{adj } R^2 = 0.967$ である。
