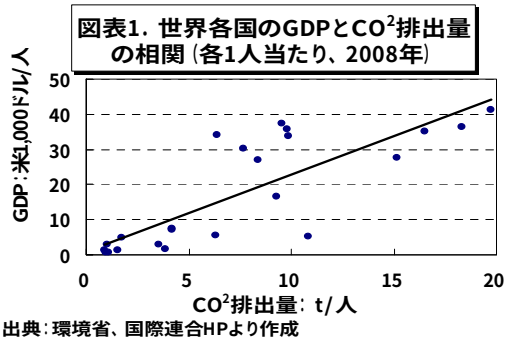
経済統計の基礎知識　＜第 11 回＞

相関と回帰について②　～回帰式の精度と有意性の見方～

寺林　暁良



出典：環境省、国際連合HPより作成

回帰式　$y_i = 974.5 + 2184.6x_i$

②（t＝2.30）　（t＝6.35）

① 自由度修正済決定係数　$R^2 = 0.64$

③ P 値＝$2.7 \times 10^{-6} < 0.05$

今回は、前回例示した図表を用いて、回帰式の精度と有意性の見方を説明する。

回帰式は、x（y）に数値を当てはめることで、もう一方の y（x）が予測するための式であり、$y = ax + b$ となるような一次方程式で表される（図表 1）。ただし、回帰式は、基本的にどんなに相関の低い数値の間でも作ることができてしまうため、精度や有意性をみることが非常に重要となる。①～③は、それを判断するための情報である。

①決定係数で回帰式の精度を確認

①の自由度修正済決定係数（寄与率）は、回帰式の精度を表す指標で、前回紹介した相関係数を 2 乗したものを標本数に考慮して修正した値である。決定係数は、$0 \leqq R^2 \leqq 1$ の範囲で示され、1 に近いほど強い相関がある、すなわち、回帰式の精度が高いということになる。

この決定係数がどれくらいあればよいかという厳密な基準はないが、社会事象の場合には、0.5 以上であれば概ね精度が高く、0.8 以上であれば非常に精度が高いといえる。上記の $R^2 = 0.64$ は、「概ね精度が高い」といえる水準である。

②t 値、③P 値の検定

②の t 値は、個々の係数の有意性（偶然との考えにくさ）を調べるための値である。例えば、回帰式の「傾き」（x の係数、ここでは 2184.6）が 0 と仮定すると、$y = 974.5 + 0 \times x$ となり、x に関係なく y が決まることになる（つまり x と y には相関が無いことになる）。そのため、「仮説検定」を行って相関があることを証明する必要がある。

仮説検定は、「傾き＝0」という「帰無仮説」を立て、この帰無仮説が成り立つ可能性が低いことを示す（棄却する）ことで、x と y には相関があるという「対立仮説」を支持するという順序で行われる。

検定の詳細な理屈は難しいが、覚えて欲しいのは、t 値の場合、だいたい 2 以上（あるいは-2 以下）となった場合に、有意水準 5%（統計的にそうなる確率が 5%以下となる水準で、偶然とは考えにくい）を満たすといえることである。ただし、有意水準をいくつで満たすかは、標本数に応じて 2 から多少前後するため、インターネットなどで「t 分布表」を参照する必要がある。ちなみに②では、どちらの t 値も有意水準を上回っており、有意であるといえる。

また、EXCEL の分析ツールを用いた場合、回帰式全体の有意性を示す F 値を加工した P 値（有意 F と表示される）が求められる。P 値は、0.05 よりも小さい場合に有意水準 5%を満たし、式の係数が 0 であるとする帰無仮説が棄却される。③の P 値は 2.7×10^{-6}（＝0.0000027）と 0.05 より小さいため、有意であるといえる。

以上のように、回帰式は、精度と有意性の条件を満たして、はじめて有効性が確認される。

ここに掲載されているあらゆる内容の無断転載・複製を禁じます