

相関分析

1 無相関検定

相関係数の大きさにより相関の強さを表現する方法を前節では紹介した。しかし、実際に相関があるということ数を限られた標本から議論することには注意が必要である。実際には知ることのできない母集団に相関があるかどうかを統計的な手法で考えることを無相関検定という。これは、前々回から始まった統計的仮説検定の一つであり、このテキストでは手続き的な話を紹介する。

統計的仮説検定では、まず、実験的に得られた相関係数と母相関係数の間には相関が無い、という帰無仮説を立てる。そして、対立仮説として、それを否定する、相関がある、という仮説も用意する。帰無仮説を否定する際にそれが間違っているという確率（危険率）が通常5%を下回ると、帰無仮説を否定して対立仮説を採用することが可能となる。以下では、実際にそれを試してみよう。

1.1 t 分布表の復習

自由度 ϕ の t 分布について、ある $t(\phi, \alpha) > 0$ の値に対し、 $\pm t(\phi, \alpha)$ の外側に t が現れる確率 α について、 α と $t(\phi, \alpha)$ の関係を与えるものが t 分布表である。 ϕ と α とが重要であるので、そこを間違えないように使用することがポイントとなる。

1.2 無相関検定の事例

無相関検定においては、 t 検定のときと少し異なり、以下の式を用いる。

$$t_0 = \frac{|r|\sqrt{n-2}}{\sqrt{1-r^2}} \quad (1)$$

上式より求めた t_0 の値を t 分布表の危険率ごとの t の値と比較し、判断する。そこで、表1と2に示すデータの組があるとするとき、これらについて、実際に無相関検定を行ってみる。

表 1: サンプルデータ 1

データ番号	x	y
1	2	12
2	5	5
3	6	18
4	9	19
5	12	4
6	14	8
7	15	11
8	19	26
9	25	29
10	27	12
11	28	21
12	31	16
13	34	28
14	36	17

表 2: サンプルデータ 2

データ番号	x	y	データ番号	x	y
1	3	10	12	29	31
2	5	15	13	31	25
3	6	17	14	33	39
4	9	11	15	34	27
5	12	18	16	38	40
6	15	19	17	40	29
7	16	28	18	41	38
8	17	16	19	43	44
9	19	24	20	48	32
10	24	26	21	52	46
11	26	19			

【演習】

表1と2について、相関係数を計算し、式(1)を用いて無相関検定を行うこと。

実際の作業としては、以下の手順で進めることで検定を行う。

1. 相関係数を計算する。

2. 相関係数と自由度から式 (1) を用いて t_0 を計算する.
3. t 分布表より該当する自由度の部分の t の値を見る.
4. $t_0 > t$ となる最大の t の危険率を見る.
5. 危険率が 5%未満であれば, 相関が有意であることが言える. (帰無仮説を否定できる)

【実際の検定の表現の例】

- 帰無仮説 : x と y の間には相関は無い.
- $t_0 > t(\phi, \alpha)$ (有意な場合)
- よって帰無仮説は危険率 α で棄却できる.
- x と y の間の相関は有意である.

- $t_0 < t(\phi, \alpha)$ (有意でない場合)
- よって帰無仮説は棄却できない.
- x と y の間には有意有意な相関は無い.