

現代日本学演習 V 「実践的統計分析」

第4講 順位相関係数

田中重人 (東北大学文学部准教授)

[テーマ] 順序尺度の相関を測る方法

1 前回課題について

比率の差については、第2講資料の正規分布を利用した信頼区間の幅が x より狭くなる条件を求めればよい。

$$x > 2 \times 1.96 \sqrt{\frac{m(1-m)}{n}} \quad (1)$$

$$n > 4 \times 3.84 \times \frac{m(1-m)}{x^2} \quad (2)$$

平均値の差については、前期第11講の平均値の差の信頼区間の公式をあてはめればよい。 n 人を半分ずつの人数 ($=n/2$ 人ずつ) にわけるとすると、95%信頼区間の幅は

$$x > 2 \times 1.96 \times SD \times \sqrt{\frac{1}{n/2} + \frac{1}{n/2}} \quad (3)$$

$$x > 4 \times 1.96 \times \frac{SD}{\sqrt{n}} \quad (4)$$

$$n > 16 \times 3.84 \times \frac{SD^2}{x^2} \quad (5)$$

なお、標本における平均値の差 d がこの信頼区間の幅の半分より大きいと、検定結果は5%水準で有意になる。このための条件は、式(4)より

$$d > 2 \times 1.96 \times \frac{SD}{\sqrt{n}} \quad (6)$$

$$n > \left(2 \times 1.96 \times \frac{SD}{d}\right)^2 = 15.37 \frac{1}{ES^2} \quad (7)$$

第2講資料の「平均値の信頼区間」の幅が x より小さい、と考えて求めても同じ値になる。

これは簡略な求め方で、実際には、人数が均等に分かれていないことが多いし、厳密には自由度や併合SDなども考慮する必要がある。さらに、本来は、母集団における一定以上の大きさの平均値の差を検出するための条件を求めなければならないので、正確に計算するのは相当面倒である。永田(2003)を参照。

調査規模の目安として、つぎのことをおぼえておくとよい(5%水準の場合)：

- 100ケースで有意差がでるのは：比率の20%以上の差、エフェクト・サイズが0.4以上になるような平均値の差
- 400ケースで有意差がでるのは：比率の10%以上の差、エフェクト・サイズが0.2以上になるような平均値の差

2 尺度水準と分析法

- 名義 × 名義 クロス表
- 名義 × 間隔 分散分析・平均値の比較
- 順序 × 順序 順位相関係数 (rank correlation coefficient)
 - Goodman-Kruskal の
 - Kendall の τ_b
 - Spearman の r_s (と書くこともある)
- 間隔 × 間隔 積率相関係数 (product-moment correlation coefficient)
 - Pearson の r

3 相関係数とは

ふたつの変数どうしが正 (+) の関係にあるか、負 (-) の関係にあるかを、 $-1 \sim +1$ の範囲の値であらわす。

- 無関連のときゼロ
- 完全な関連のとき ± 1

「相関図」(または「散布図」(scattergram) ともいう) を描いて考えるとよい (教科書 p. 75)。

4 順位相関係数

4.1 Pair

相関図上の任意の2点を直線で結んだとき

- 右上がり Concordant
- 左上がり Discordant

それぞれのペアの個数を C, D とする。

4.2 グッドマンとクラスカルの「ガンマ」係数

$$\text{Goodman-Kruskal's } \gamma = \frac{C - D}{C + D} \quad (8)$$

同順位ペアをうまく扱えないので、あまり使われない

4.3 ケンドールの順位相関係数 (タウ b)

- K : x について同順位でないペア数
- L : y について同順位でないペア数

$$\text{Kendall's } \tau_b = \frac{C - D}{\sqrt{KL}} \quad (9)$$

同順位ペアがなければ、Goodman-Kruskal の と同じ値になる。

4.4 PSPP コマンド

- 「分析」 「記述統計量」 「クロス集計表」
- 変数を指定
- 「統計量」オプション 「Kendall のタウ b」を選択

5 課題

(x, y) の値がつぎの組み合わせであるような 6 人の標本があるとする：

$(1, 2) (2, 4) (2, 4) (4, 3) (4, 5) (5, 5)$

この標本について、Kendall の順位相関係数タウ b を求めよ。

6 次回予習

教科書の第 3 章、第 8 章 7 節を読んでおくこと。

文献

永田 靖 (2003) 『サンプルサイズの決め方』朝倉書店.