

現代日本学演習 V 「実践的統計分析」

第5講 積率相関係数

田中重人 (東北大学文学部准教授)

[テーマ] ピアソンの積率相関係数と相関係数の統計的検定

1 課題

PSPP の「クロス集計表」で、Kendall のタウ b がプラスになる表とマイナスになる表を出力し、クロス表の %を見て解釈する

2 積率相関係数類

2.1 変数の標準化

平均 = 0, 標準偏差 = 1 になるよう変換する。これで単位を気にせずに、変数同士の値を比較できるようになる

具体的には: (その個体の値 - 平均値) / SD

(教科書 pp. 129, 130 を参照)

例題: 前回宿題 (Kendall の順位相関係数) のデータを標準化してみよう。

2.2 Pearson の積率相関係数

標準化済みの変数 X, Y について、それらの積の平均をとったもの:

$$r = \frac{\sum XY}{N} \quad (1)$$

通常、単に「相関係数」といえばこの r をさす

欠点: はずれ値や歪みに弱い

2.3 Spearman の順位相関係数

先に各変数を順位に変換しておく。あとの計算は、Pearson の積率相関係数とおなじ。

r_s または (rho: ロー) であらわす。

2.4 PSPP コマンド

クロス表の「統計量」オプションで「相関係数」を選択。

3 相関係数類の使いわけ

- 順序尺度の場合: Kendall のタウ b または Spearman の
- 間隔尺度の場合
 - 正規分布なら Pearson の r
 - 歪みや外れ値 Spearman の

相関係数が 0 または ± 1 になるのはどのような場合か?

- Goodman-Kruskal の :
- Kendall のタウ b:
- Pearson の r :
- Spearman の :

4 相関係数の検定

Pearson の r の信頼区間は、「Fisher の z 変換」と呼ばれる方法で求められる (森・吉田 1990)。この信頼区間に $r=0$ が含まれるかを判断すれば、統計的検定がおこなえる。

ただし、この方法で正確に信頼区間を求めるのは面倒なので、通常は t 分布を利用した検定だけをおこなう (教科書巻末の数表参照)。Spearman の順位相関係数 についても、おなじ方法が使える。

Kendall の順位相関係数タウ b についての推定・検定は別の方法を使う (Bohrnstedt and Knoke, 1992) が、省略。 r に関する t 検定より検定力が低いことに注意。

文献

池田央 (編) (1989) 『統計ガイドブック』新曜社

森敏明・吉田寿夫 (1990) 『心理学のためのデータ解析テクニカルブック』北大路書房。

Bohrnstedt, G. W. and Knoke, D. (1992) 『社会統計学』(海野道郎・中村隆監訳、学生版) ハーベスト社。