

2002.10.25 比較現代日本論研究演習 I (田中重人)

第4回「サンプルサイズの決定」

1. 検定力
2. ϕ 係数と%の差
3. ϕ 係数と χ^2 臨界値
4. サンプルサイズと検定力

【検定力】

母集団における一定の大きさの関連を
どれくらいの危険率で検出できるか

→ サンプルサイズに依存

【 ϕ 係数と%の差】

2×2 クロス表の%の差

=周辺度数がバランスしていれば、
 ϕ 係数に等しい

【 ϕ 係数と χ^2 臨界値】

2×2 クロス表で独立性の検定が5%有意

$$\chi^2 = N\phi^2 > 3.84$$

【サンプルサイズと検定力】

ある%差を5%水準で検出するのに

必要なサンプルサイズ： $N > 3.84/\phi^2$

20%差 → $3.84 / 0.2^2 \doteq 96$

16%差 →

14%差 →

12%差 →

10%差 →

5%差 →

1%差 →

【サンプルサイズの決定】

- 変数の測定法・分析法をきめる
- どの程度の強さの関連を検出できればよいかを決める
- 必要なサンプルサイズを決める
- 分析のキーとなるカテゴリに均等分配した場合を最低限度とする

※不均等な配分を前提として厳密に求めることも可能

【その他の係数の場合】

Pearson の相関係数 \rightarrow ϕ 係数とほぼおなじ

連関係数 $V \rightarrow \chi^2$ 臨界値が自由度で変わる。

またカテゴリ数(少ない方)を考慮する。

一般に $N > \chi^2 \text{ 臨界値} / (m - 1)V^2$

たとえば 3×3 クロス表なら

$$N > 9.49 / 2V^2$$

相関比 $\eta \rightarrow$ 次の式を使う (k はカテゴリ数) :

$$\frac{\eta^2}{1-\eta^2} \times \frac{N-k}{k} > F_{\text{臨界値}}$$

※ $k \times 2$ クロス表の V 係数とほぼおなじ

※ 2 グループ間の平均比較なら Φ 係数とほぼ同じ

順位相関係数類 \rightarrow 後日