比較現代日本論研究演習/現代日本論演習 (田中重人)

第 4 回「サンプルサイズの決定」(2004.11.4)

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

【検定結果の表示】

例 1-1

	平均	標準偏差	(人)
男性	2.62	1.02	(114)
女性	2.24	0.91	(136)
合計	2.41	0.98	(250)

$\eta = 0.198.$　$p < 0.05.$

例 2-1

	平均	標準偏差	(人)
男性	1.77	0.67	(111)
女性	1.89	0.65	(132)
合計	1.84	0.66	(243)

$\eta = 0.086.$　$p > 0.05.$　無回答＝７.

例 1-2

	平均	標準偏差	(人)
男性	2.62	1.02	(114)
女性	2.24	0.91	(136)
合計	2.41	0.98	(250)

$\eta = 0.198^{*}.$　*: 5%水準で有意.

例 2-2

	平均	標準偏差	(人)
男性	1.77	0.67	(111)
女性	1.89	0.65	(132)
合計	1.84	0.66	(243)

$\eta = 0.086^{ns}.$　ns: 5%水準で非有意.

無回答＝７.

【検定力】

power (of a test)

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 → $3.84 / 0.2^2 \fallingdotseq 96$

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 → ϕ 係数とおなじ

連関係数 V →χ^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 $N > \chi^2$ 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 η → 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×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係数とおなじ

順位相関係数類 → 後日