2003.7.15 現代日本論演習 II (田中重人)

第 13 回「統計的検定」

1. 平均値の差の推定
2. 区間推定と統計的検定
3. 分散分析と F 検定
4. クロス表の独立性の検定
5. 検定結果の書きかた

【平均値の差の推定】

2 層間の **平均値の差** についても

平均値そのものと同様の区間推定ができる：

このとき 95%信頼区間はおよそ

$$\underbrace{d}_{\text{平均値の差}} \pm 1.96 \times \overbrace{\text{併合SD} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}^{\text{標準誤差}}$$

ただし n_1, n_2 はそれぞれの層の人数

各層の人数が多いほど
平均値の差の信頼区間が狭くなる

➡標本を均等にわけるほうが
信頼性が高い

【SPSSのコマンド】

「平均値の比較」→「独立したサンプルのＴ検定」

◎「グループ化変数」は、数値を指定しないといけない。

連続量を一定の値で切ることもできる

出力は「独立サンプルの検定」の１行目
「等分散を仮定する」を見る

【区間推定と統計的検定】

Statistical test

統計的検定＝特定の値を設定して、その値が信頼区間に含まれているかどうかを判定する

0に設定するのがふつう

95%信頼区間が 0 をふくまない

⇔「5%水準で有意」

※　統計的検定の論理は本当はもっと複雑である。

【有意確率とは】

信頼区間をひろげていくと、
どこかでゼロをふくむようになる

→このときの危険率のことを「有意確率」
または「有意水準」**(level of significance)**
という。

分析の際は、

- 前もって危険率を設定しておく
  (通常は 5%または 1%)
- 有意確率がその値を
  下回っているかどうか判別する

例:

有意確率が 0.007 → 1%水準で有意（5%水準でも有意）

有意確率が 0.023 → 1%水準で非有意（5%水準では有意）

有意確率が 0.088 → 1%水準で非有意（5%水準でも非有意）

【統計的検定のいろいろ】

★　平均値の差の t 検定

コマンドの指定は区間推定とおなじ。

出力の「有意確率　(両側)」を見る

※　2 層の間の差の検定にしか使えない

※　「母集団では正規分布」を前提とする

※　2 層の間で分散が等しいことが前提

★　分散分析と F 検定

「平均値の比較」→「グループの平均」

オプション「分散分析表とイータ」を指定

出力「分散分析表」の右端「有意確率」

※　3 層以上の場合に使う。

η の信頼区間を使って判断するのと同じである。

※　2 層の場合にも使えるが、t 検定と同じ結果になる

※　必要とする前提も t 検定と同様

★　クロス表の独立性の検定

「クロス集計表」の「統計」で
「カイ２乗」を指定。
出力の「Pearson」の列の右端が有意確率

※　Vの信頼区間を使って判断するのとおなじ
※　各セルの期待度数が 5 以上であることを前提とする

【検定結果の表示】

表の下に書く

(1) $p < 0.xx$ 形式：

平均値の差＝0.04 ($p > 0.05$)

相関比 η ＝0.198 ($p < 0.05$)

Cramer's V＝0.258 ($p < 0.01$)

(2) 「xx%水準で有意」 形式：

平均値の差＝0.04 (5%水準で非有意)

相関比 η ＝0.198 (5%水準で有意)

Cramer’s V＝0.258 (1%水準で有意)

(3) 星印 (*) 形式：

平均値の差＝0.04^{ns}

相関比 η ＝0.198*

Cramer's V＝0.258**

→ 注釈をどこかに書いておく

(ns: $p > 0.05$; *: $p < 0.05$; **: $p < 0.01$)