

## 効果量について

グループ間の平均値の差を  $t$  検定を利用して評価する時、 $T$  値を計算し、有意確率  $p$  値が有意水準より小さいなら差があると判断する。これは推測統計を利用して、サンプルデータより母集団の平均値の差を推定しているからであるが、サンプルサイズが大きくなると、同じようなデータであっても有意な結果が出やすい。

例えば サンプルサイズ 10 の A グループ、B グループの平均値の差を  $t$  検定で評価した結果は以下であった。

| 計算結果      |                                 |             |                                     |
|-----------|---------------------------------|-------------|-------------------------------------|
| 自由度 :     | 18                              |             |                                     |
| 平均値 1 :   | 39                              | (不偏)分散値 1 : | 11.33333                            |
| 平均値 2 :   | 41                              | (不偏)分散値 2 : | 5.333333                            |
| T 値 :     | -1.543193                       | 帰無仮説の採択域 :  | ( -2.1015 , 2.1015 )                |
| P 値 (%) : | 13.87                           |             | <input type="button" value="分布関数"/> |
| 結 果 :     | 有意でない: 帰無仮説 (2群の平均値は等しい)は棄却できない |             |                                     |
| 効果量 :     | -0.6928203230275                |             |                                     |

検定結果は有意でない、つまり帰無仮説を棄却できないという結果である。

今度は A グループ、B グループを同じデータで 5 倍に膨らませ、サンプルサイズ 50 にして、 $t$  検定を行った結果は以下となった。

| 計算結果      |                            |             |                                     |
|-----------|----------------------------|-------------|-------------------------------------|
| 自由度 :     | 98                         |             |                                     |
| 平均値 1 :   | 39                         | (不偏)分散値 1 : | 10.40816                            |
| 平均値 2 :   | 41                         | (不偏)分散値 2 : | 4.897959                            |
| T 値 :     | -3.614784                  | 帰無仮説の採択域 :  | ( -1.984967 , 1.984967 )            |
| P 値 (%) : | 0.04768                    |             | <input type="button" value="分布関数"/> |
| 結 果 :     | 有意: 帰無仮説 (2群の平均値は等しい)を棄却する |             |                                     |
| 効果量 :     | -0.7229568912920           |             |                                     |

平均値は変わらないが、後者は  $T$  値 (の絶対値) が大きくなり、 $p$  値は 1000 分の 3 程に小さくなり、有意と判定され、帰無仮説は棄却されるという結果となった。

サンプルサイズが大きくなれば精度は高まるので、わずかな差異でも有意と判定できるわけで、これはこれで正しいが、上記の様に意図的に膨らませたデータは見抜けない。

そこで効果量も表示して、実質的な差異も評価することが薦められている。

効果量は、大まかに言うと、(2つのグループの平均の差異) / (標準偏差) であり、具体的な計算式は 各機能の中で説明している。

上記の例では、効果量は -0.69 と -0.72 であり、 $p$  値の大きな差異に比較して、それほど差異でないことを明らかにしている。