

2025 年度 データサイエンス レポート課題

25G1XXX 情報 工太

選択した分析手法

仮説検定・回帰分析・判別分析・ニューラルネットワーク（選択したもの以外を削除すること）

例題

（記載例）ある運動の前後に脈拍を記録したところ、以下のようになった。下記データについて有意水準 5% の検定を行え。なお、差 d は正規分布に従うと考えて良いが、母標準偏差 σ_d は未知であるとする。

表 1 運動前後での脈拍の測定結果

	運動前	運動後
被験者 A	78	82
被験者 B	70	71
被験者 C	73	74
被験者 D	72	76
被験者 E	76	74
被験者 F	71	73
被験者 G	77	82

解答

（記載例）帰無仮説 H_0 は「運動前後の脈拍は等しい」である。まず、運動前後の脈拍の差 d を計算する（ここでは運動後から運動前を引く）。差 d の標本平均は $\bar{d} = 2.14$ 、標本標準偏差は $s_d = 2.41 \dots$ 、標本サイズは $n = 7$ である。標本標準偏差 s_d の自由度は $df = n - 1 = 6$ である。そこで、統計検定量の Student の t は

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{2.14}{2.41 \dots / \sqrt{7}} = 2.35 \dots$$

となる。 t 分布表の $df = 6$ と $\alpha = 0.05$ から、臨界値 $t_{0.025}(6) = 2.447$ を得る。そこで、以下の不等式が成立する。

$$2.35 \dots = |t| < t_{0.025}(6) = 2.447$$

この結果、検定統計量 t は棄却域に入らないことがわかる。そこで、「運動前後の脈拍の間に統計的に有意な差は認められなかった」と結論する。

ソースコード 1 プログラム例

```
1  before = np.array([78, 70, 73, 72, 76, 71, 77]) # 運動前
2  after  = np.array([82, 71, 74, 76, 74, 73, 82]) # 運動後
3
4  # 差を計算
5  d = after - before
6  mean_d = np.mean(d) # 平均差
7  std_d = np.std(d, ddof=1) # 不偏標準偏差
8  n = len(d) # サンプルサイズ
9
10 # t 検定の計算
11 t_statistic = mean_d / (std_d / np.sqrt(n)) # t 値
12
13 # p 値の計算 (両側検定)
14 p_value = 2 * (1 - stats.t.cdf(abs(t_statistic), df=n-1))
15
16 # 有意水準
17 alpha = 0.05
18
19 # 結果の表示
20 print(f"平均差: mean_d:.2f")
21 print(f"不偏標準偏差: std_d:.2f")
22 print(f"計算された t 値: t_statistic:.2f")
23 print(f"p 値: p_value:.4f")
24
25 if p_value < alpha:
26     print("結果: 統計的に有意です (P<0.05) (帰無仮説を棄却します). ")
27 else:
28     print("結果: 統計的に有意ではありません (帰無仮説を棄却できません). ")
```
