２群の平均の差の検定：資料

芳賀昭弘 [*]

1　２群の平均の差の検定フローチャート

前回は、取り出した標本が母集団の性質（平均や分散）と一致しているかどうかを調べる方法を学んだ（**標本ー母集団**の比較）。しかしながら、医療（臨床研究）では、母集団の性質がわかっているようなケースは稀である。そのため、調べたい標本とともに、それと比較するための対照群を用意して、この二つのグループの平均値の差を求めるということを行う（**標本ー標本**の比較）。図には、その検定の流れを示している。本講義では左側のフローを主に学ぶ。これは「母集団が正規分布をしている」ことを仮定している。そのように母集団の分布がある特定の分布をしていると仮定して検定を行うことを**パラメトリック検定**などと呼ぶことがある。他方、右側のフローでは母集団の分布を特に仮定しておらず、パラメトリック検定と対比的に**ノンパラメトリック検定（ノンパラ）**と呼ばれる（右側のフローで異なっているのは使う検定方法だけであることに注意）。ここでは２群のみの比較を考えるが、場合によっては三つ以上のグループ間の比較が必要な場合もあろう。３群以上の検定に関する内容は次回以降に学ぶ。

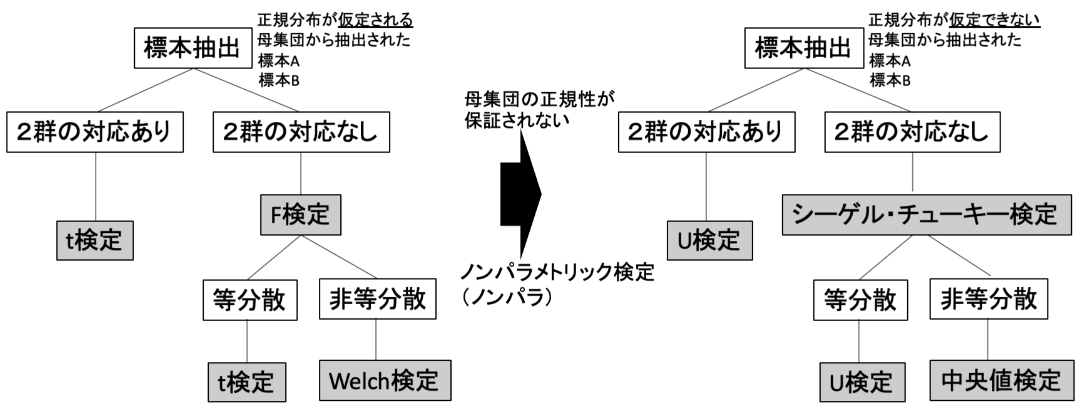

図１　２群の平均の差の検定フローチャート

* Electronic address: haga@tokushima-u.ac.jp

2　二つのグループ（標本）の平均値の差

命題：同じ分散である２つの母集団 $\mathcal{N}(\mu_1, \sigma^2)$、$\mathcal{N}(\mu_2, \sigma^2)$ からの標本 $(x_1, x_2, \cdots, x_m)$、$(y_1, y_2, \cdots, y_n)$ に対して、

$$t = \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{ms_x^2+ns_y^2}} \tag{1}$$

の標本分布は自由度 $m+n-2$ の t-分布に従う。もし母集団の分散が異なる場合には、Welch 法で近似する。この場合、

$$t = \frac{\mathcal{N}(0,1)}{\sqrt{\chi_\nu^2/\nu}} \tag{2}$$

ここで、

$$\nu = \frac{(\sigma_x^2/m+\sigma_y^2/n)^2}{(\sigma_x^2/m)^2/(m-1)+(\sigma_y^2/n)^2/(n-1)} \tag{3}$$

となる。

（対応のある場合）：二つの標本の差の平均を $\bar{d}$、その不変標準偏差と $\hat{\sigma}$ として、

$$t = \frac{\bar{d}}{\hat{\sigma}/\sqrt{n}} \tag{4}$$

は自由度 $n-1$ の t-分布に従う（ここで n は $\bar{d}$ の数）。

母集団が正規分布であるという条件において、２つの標本の平均値の差の検定には、上記のどれかを使うことになります。いずれも t-分布ですね。

3　対応のある t 検定

双子のペアが 10 組あり、そのペアの一方を第１群、もう一方を第２群に分けて取られたデータは、“対応のあるデータ”と考えられます。また、同じ被験者に薬を処方し、服薬してもらう前後の血液データなども “対応のあるデータ”とみなします。その場合、データ群間の分散は等しいと考えられるので、その平均値の差に対しては式 (4) を使った t-検定を行うことになります。

例題１：３名の患者に対して治療前の血液の値が（0.1,0.3,0.3) であった。治療後、血液の値が（1.1,0.7,0.6) となった。治療前後のデータに差があるか検定せよ。
帰無仮説 H0:治療前後のデータに差がない。
検定統計量: 式 (4) で定義される t 値。
有意水準:両側 5% とする。

（解答）治療前後の差の平均：-0.567
治療前後の差の不偏標準偏差：0.379
自由度：2 ($n=3$)

実現値 $t=-2.59$ 自由度 2 に対する t 分布表を見ると、上側確率 0.025（両側で 5% の意味）を与える t 値は 4.303 である（95% 信頼区間は $-4.303 \sim 4.303$ で、今回の実現値-2.59 はその範囲内)。

p 値が 0.05 以上であるので、統計検定量は設定した有意水準 5% 以上であり、帰無仮説は棄却される。

4　対応のない t 検定

前節の対応のあるデータとは対照的に、データをランダムに 2 群に割り振ったようなデータは、独立な 2 群とか、対応のない 2 群と言えます。この場合、データ（の母集団）の等分散性が保証されないため、まず、**分散の等質性の検定**を行います。分散の等質性は、F 分布により検定が可能でした。R では、var.test 関数によって検定を簡単に行うことができます。

例題 2：次のクラス A とクラス B の平均点に有意な差があるでしょうか？有意水準 5% で検定を行ってください。

クラス A	54	55	52	48	50	38	41	40	53	52
クラス B	67	63	50	60	61	69	43	58	36	29

（解答）まず、F 検定を行います。
クラス A の不偏分散：39.79
クラス B の不偏分散：184.49
自由度はともに $9(=n-1)$
F 値の実現値：$F_{9,9}=184.49/39.79=4.64$
上側確率 5% に対する F 分布表を見ると、自由度 (9,9) を与える F 値は 3.18 である (95% 信頼区間は $0 \sim 3.18$ で、今回の実現値 4.64 はその範囲外)。つまり「差がある」ということ。

4.1　分散が等しい場合

例題 2 において分散が等しいと検定された場合には、式 (1) を使って t 値を計算する。
実際には、例題 2 の検定では分散が等しくないと検定されますが、敢えて式 (1) を使って計算をしてみることにしましょう。上で不偏分散を計算しているので、式 (1) を不偏分散で書き直すとともに、$m=n$ なので式 (1) は次のように簡単になる；

$$t=\frac{(\bar{x}-\bar{y})}{\sqrt{(\hat{\sigma_x}^2+\hat{\sigma_y}^2)/n}} \tag{5}$$

ここで $n=10$ であり、クラス A の平均と不偏分散：48.3, 39.79
クラス B の不偏分散：53.6, 184.49
を代入すると、t 値の実現値は 1.12 となる。なお、自由度は $10+10-2=18$ である。t 分布表を見ると自由度 18 の上側確率 2.5% を与える t 値は 2.101（95% 信頼区間が $-2.101 \sim 2.101$）であり、その範囲に実現値が入っていることから、珍しくない結果であると結論できる。
統計検定量は有意水準 5% 以上であり、帰無仮説は棄却されない。

4.2　分散が等しくない場合

実際には、例題 2 の検定では分散が等しくないと検定されます。その場合、自由度は式 (3) を使って算出します。Welch の t-検定などと呼んでいます。t 値の計算は式 (1)（$n = m$ の時は式 (5)）と同じ。式 (3) を使って自由度を計算してみると（大変ですが）12.71 となります。自由度 12 と自由度 13 の t 分布表より、自由度 12.71 の t 分布の 95% 信頼区間は $(-2.17 \sim 2.17)$ くらいであるとわかるので、t 値の実現値 1.12 はその範囲内です。統計検定量は定めた有意水準 5% 以上であるので、帰無仮説は棄却されない。

5　演習

1. 次の表は、２種類の気温に保った温室で、あるガの幼虫を７匹ずつ使ってマユを形成してから成虫へ羽化するまでの生育日数を観測したデータである。この実験から 15 度と 20 度の気温の差が、ガの生育日数に影響を与えたと言って良いだろうか？

No.	15deg.	20deg.
1	37	28
2	35	26
3	32	25
4	36	22
5	34	23
6	33	24
7	31	27