

2009 年度 森林統計学 第 12 回資料「平均値の差の検定」での検定方法の選択

1. 検定方法の選択

①「比較する二つの標本の標本平均(\bar{x}_1, \bar{x}_2)が正規分布に従うとみなしてよいかどうか」、
 ②「母集団の標準偏差(σ_1, σ_2)は同じとみなしてよいかどうか」、の二つの基準により、3
 つの方法([1]正規分布による方法・[2] t 検定・[3] Welch の t 検定)のいずれが適切な方法か
 を判断する。①を満たせば[1]、②を満たせば[2]、いずれも満たさなければ[3]が適切な方法
 である。①は中心極限定理(教科書第 6 章 5 節の定理 2)により、標本数が十分大きいかどう
 か、という条件に置きかえられる。②は分散の比が 1 に近いかどうか(1 より大きく異なる
 かどうか)で判断される(このとき母分散は不明なので標本分散[母分散の不偏推定値]の比を
 用いる)。フローチャートを以下に示す。

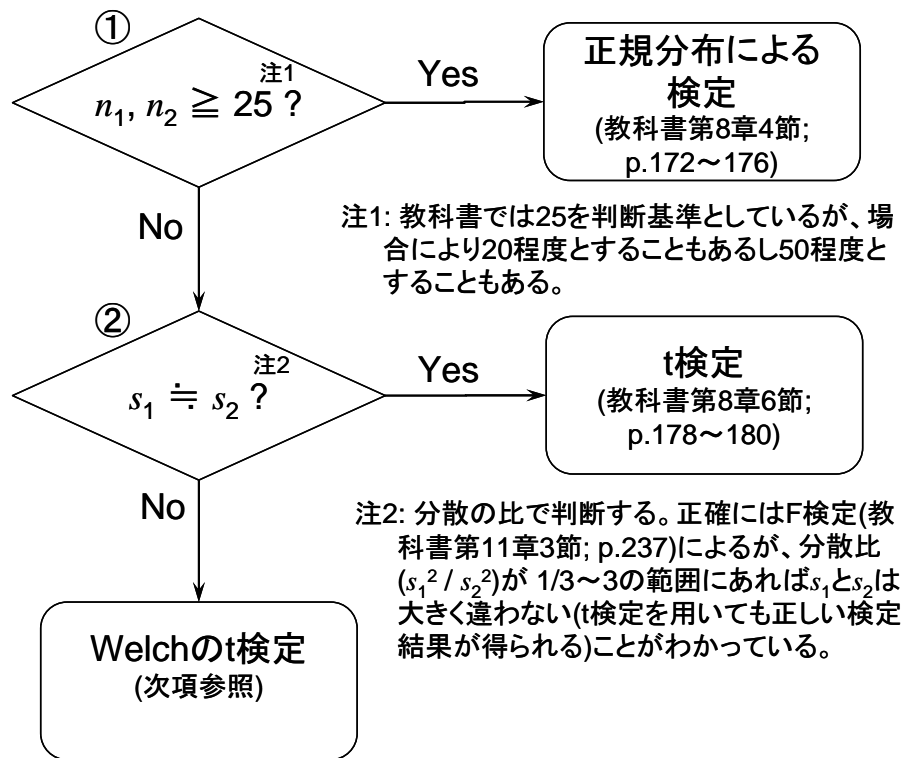


図-1. 「平均値の差の検定」での検定方法の選択

注) n_1, n_2 : 比較するそれぞれの標本(標本 1, 標本 2)のデータ数

s_1, s_2 : それぞれの標本の標準偏差 (s_1^2, s_2^2 : 分散)

2. Welch の t 検定

手順や考え方は[1](正規分布による方法) および[2](t 検定)と同様である。 y には 1,2.と同様に(16)式を用いる。標本平均の標準偏差には 1.と同様に(17)式を用いる。すなわち、統計量 t は次式で求められる。

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (1)$$

注) 通常の両側検定、すなわち「母集団の平均に差がない」という仮説を検定する場合、 $\mu_1 = \mu_2$ から $\mu_1 - \mu_2 = 0$ となるので分子は $\bar{x}_1 - \bar{x}_2$ のみとなる。教科書 p.179 の(6)式 (t 検定の場合の統計量 t の計算式)および教科書 p.173 の定理(4)の式(正規分布を用いる検定の場合の変数「 $\bar{x}_1 - \bar{x}_2$ 」の標準偏差 $\sigma_{\bar{x}_1 - \bar{x}_2}$ の式)も参照。

異なるのは自由度 ν の計算方法で、これは次式で求める。

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} \quad (2)$$

(2)式で計算される自由度は必ずしも整数にならない(しかし $n_1 + n_2 - 2$ と大きくは違わない数値にはなる)が、小数点以下を切り捨てることで安全側の自由度が得られる(α が大きくなる)。)

注) より正確には、t 分布の表から得られる値を補間して用いる(自由度 30 以下では直線補間、より大では逆数補間がよいとされている)。

※ Excel の分析ツールでは、「z 検定: 2 標本による平均の検定」が[1]に、「t 検定: 等分散を仮定した 2 標本による検定」が[2]に、「t 検定: 分散が等しくないと仮定した 2 標本による検定」が[3]に相当する。ただし、分析ツールでは自由度は小数点以下切捨てで検定していると思われるため、3.の場合は補間式を用いて自分で検定を行った方がより正確な結果が得られる。

補足: 分散比の検定(F 検定)

二つの分散の大きさの違いの検定には F 分布を用いる。標本数 n_1 、標本分散 s_1^2 、標本数 n_2 、標本分散 s_2^2 が得られた場合、その分散の比

$$F = \frac{s_1^2}{s_2^2} \quad (3)$$

は分子の自由度 $\nu_1 = n_1 - 1$ 、分母の自由度 $\nu_2 = n_2 - 1$ [自由度(ν_1, ν_2)] の F 分布に従う。F 分布表は分子(Numerator)の自由度と分母(Denominator)の自由度の組合せで棄却限界値が用意されている(教科書「数表」の表 IX; p.300~303)。②の判断の目的のためには、 s_1^2 と s_2^2 の大きな方を分子に、小さな方を分母にしてその比 F を計算し、その値が大きすぎないかどうかをみればよい。すなわち、表 IV の該当する自由度の 5%点($\alpha=0.05$)の数値より F が小さければ、②は満たされる(「分散の違い=標準偏差の違い」は大きくない)と判断する。