第 IX 部　母平均の差の検定

目次

8　母分散が未知の場合

8.1　はじめに

- 母分散が未知の場合の平均値の差の検定は、σ_x と σ_y が等しいかどうかで手法が異なります。
- ここでは、σ_x と σ_y が等しい場合を説明します。
- ここでは、σ_x と σ_y が等しくない場合も紹介します。

8.1.1　ポイント

- 等分散の検定
- 等分散の場合の母平均の差の検定
- ウェルチの t 検定

8.1.2　標準正規分布

平均 0, 分散 1 の正規分布を標準正規分布といい $N(0,1)$ とあらわす。

8.1.3　正規分布

$X=x$ が標準正規分布に従う $Z=z$ を用いて

$$x=\mu+\sigma z \quad ;\sigma\neq 0 \tag{1}$$

とあらわせるとき、X が従う分布を『**平均 $\boldsymbol{\mu}$, 分散 $\boldsymbol{\sigma^2}$ の正規分布**』といい、$N\left(\mu,\ \sigma^2\right)$ であらわす。

x が $N\left(\mu,\ \sigma^2\right)$ に従うとき、

$$z=\frac{x-\mu}{\sigma} \quad ;\sigma\neq 0 \tag{2}$$

とすると、z は、標準正規分布 $N(0,\ 1)$ に従う。

8.1.4 χ^2 分布

標準正規分布 $N(0,\ 1)$ に従う n 個の独立な確率変数 $z_1,\ z_2,\ \cdots,\ z_n$ の平方和を作り $\chi^2_{(z)}$ とあらわす。

$$\chi^2_{(z)} = \sum_{i=1}^{n} {z_i}^2 \tag{3}$$

このとき n を『**自由度**』といい ν であらわす。$\chi^2_{(z)}$ が従う分布を『**自由度 $\boldsymbol{\nu = n}$ の $\boldsymbol{\chi^2}$ 分布**』という。

8.1.5 χ^2 の和

2 つの確率変数 U_1, U_2 がそれぞれ独立に χ^2 分布に従うとする。U_1 の自由度を $\nu = n$,U_2 の自由度を $\nu = m$ とする。

$$U_1 = \sum_{i=1}^{n} {z_i}^2 \tag{4}$$

$$U_2 = \sum_{j=1}^{m} {z_j}^2 \tag{5}$$

ここで

$$U_3 = U_1 + U_2 \tag{6}$$

を考える。(4)(5) を代入すると

$$U_3 = \sum_{i=1}^{n} {z_i}^2 + \sum_{j=1}^{m} {z_j}^2 \tag{7}$$

$$= \sum_{i=1}^{n} {z_i}^2 + \sum_{i=n+1}^{n+m} {z_i}^2 \tag{8}$$

$$= \sum_{i=1}^{n+m} {z_i}^2 \tag{9}$$

よって、U_3 は自由度 $\nu = n + m$ の χ^2 分布に従う。

8.1.6　標本平均

サンプル・サイズが n である標本 A に属する要素 x_i の平均を『**標本平均**』といい、$\bar{X}$ であらわす。標本に属する要素の数を n とし i 番目の要素を x_i とあらわすと、

$$\bar{X} = \frac{1}{n} \sum_{x_i \in A}^{n} x_i \tag{10}$$

である。

8.2 $\boldsymbol{x \sim N(\mu, \sigma^2)}$ が従う $\boldsymbol{\chi^2}$ 分布

$N\left(\mu,\ \sigma^2\right)$ に従う確率変数 $X = x$ を標準化

$$z_i = \frac{x_i - \mu}{\sigma} \tag{11}$$

すると、$Z = z$ は $N(0,\ 1)$ に従う。

ここで、$N\left(\mu,\ \sigma^2\right)$ からサンプル・サイズ n の標本 A を抽出し、

$$\chi^2_\mu = \sum_{x_i \in A}^{n} \left(\frac{x_i - \mu}{\sigma}\right)^2 \tag{12}$$

をつくると、χ^2_μ は自由度 $\nu = n$ の χ^2 分布に従う。

また、μ の代わりに $\bar{X}$ をもちいて

$$\chi^2_{\bar{X}} = \sum_{x_i \in A}^{n} \left(\frac{x_i - \bar{X}}{\sigma}\right)^2 \tag{13}$$

をつくると、$\chi^2_{\bar{X}}$ は自由度 $\nu = n - 1$ の χ^2 分布に従う。

8.2.1 $\boldsymbol{\chi^2_{\bar{X}} + \chi^2_{\bar{Y}}}$ の自由度

$N\left(\mu_{(X)},\ \sigma^2\right)$ から抽出されたサンプル・サイズ n の標本 A において、$\chi^2_{\bar{X}}$ は自由度 $\nu = n - 1$ の χ^2 分布に従う。

$N\left(\mu_{(Y)},\ \sigma^2\right)$ から抽出されたサンプル・サイズ m の標本 B において、$\chi^2_{\bar{Y}}$ は自由度 $\nu = m - 1$ の χ^2 分布に従う。

よって $\chi^2_{\bar{X}} + \chi^2_{\bar{Y}}$ は自由度

$$\nu = (n-1) + (m-1) \tag{14}$$

$$= n - 1 + m - 1 \tag{15}$$

$$= n + m - 2 \tag{16}$$

の χ^2 分布に従う。

問題 IX − 8 − 1

$$\chi^2_{\bar{X}} = \sum_{x_i \in A}^{n} \left(\frac{x_i - \bar{X}}{\sigma} \right)^2$$
$$\chi^2_{\bar{Y}} = \sum_{y_j \in B}^{m} \left(\frac{y_j - \bar{Y}}{\sigma} \right)^2$$

とする。

$$\chi^2_{\bar{X}} + \chi^2_{\bar{Y}}$$

を計算しなさい。

解例 IX − 8 − 1

$$
\begin{aligned}
\chi^2_{\bar{X}} + \chi^2_{\bar{Y}} &= \sum_{x_i \in A}^{n} \left(\frac{x_i - \bar{X}}{\sigma} \right)^2 + \sum_{y_j \in B}^{m} \left(\frac{y_j - \bar{Y}}{\sigma} \right)^2 \\
&= \frac{1}{\sigma^2} \left(\sum_{x_i \in A}^{n} \left(x_i - \bar{X} \right)^2 + \sum_{y_j \in B}^{m} \left(y_j - \bar{Y} \right)^2 \right)
\end{aligned}
$$

8.2.2 $\mathsf{V}\left(\bar{X}\right) + \mathsf{V}\left(\bar{Y}\right)$

$N\left(\mu_{(X)},\ \sigma^2\right)$ から抽出された n 個の要素を持つ標本を X とし、

$N\left(\mu_{(y)},\ \sigma^2\right)$ から抽出された m 個の要素を持つ標本を Y とする。

$$X = \{x_1,\ x_2,\ \cdots,\ x_n\} \tag{17}$$

$$Y = \{y_1,\ y_2,\ \cdots,\ y_m\} \tag{18}$$

標本平均の分散はそれぞれ

$$\mathsf{V}\left(\bar{X}\right) = \mathsf{V}\left(\frac{1}{n}\sum_{x_i \in A}^{n} x_i\right) \tag{19}$$

$$\mathsf{V}\left(\bar{Y}\right) = \mathsf{V}\left(\frac{1}{m}\sum_{y_j \in B}^{m} y_j\right) \tag{20}$$

である。

ここで、

$$\mathsf{V}\left(\bar{X}+\bar{Y}\right)=\mathsf{V}\left(\bar{X}\right)+\mathsf{V}\left(\bar{Y}\right) \tag{21}$$

$$=V\left(\frac{1}{n}\sum_{x_i\in A}^{n} x_i\right)+\mathsf{V}\left(\frac{1}{m}\sum_{y_j\in B}^{m} y_j\right) \tag{22}$$

$$=\frac{1}{n^2}\mathsf{V}\left(\sum_{x_i\in A}^{n} x_i\right)+\frac{1}{m^2}\mathsf{V}\left(\sum_{y_j\in B}^{m} y_j\right) \tag{23}$$

x_i、y_j はそれぞれ独立なので i および j ごとの分散であらし

$$\begin{aligned}&=\frac{1}{n^2}\left(\mathsf{V}\left(x_1\right)+\mathsf{V}\left(x_2\right)+\cdots+\mathsf{V}\left(x_n\right)\right)\\&+\frac{1}{m^2}\left(\mathsf{V}\left(y_1\right)+\mathsf{V}\left(y_2\right)+\cdots+\mathsf{V}\left(y_m\right)\right)\end{aligned} \tag{24}$$

$$=\frac{1}{n^2}n\sigma^2+\frac{1}{m^2}m\sigma^2 \tag{25}$$

$$=\frac{\sigma^2}{n}+\frac{\sigma^2}{m} \tag{26}$$

$$=\sigma^2\left(\frac{1}{n}+\frac{1}{m}\right) \tag{27}$$

8.2.3 標本平均からの偏差の二乗和

母平均を μ, 母分散を σ^2 とする。ここで

$$
\begin{aligned}
\mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2\right) &= \mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \mu + \mu - \bar{X}\right)^2\right) && (28) \\
&= \mathsf{E}\left(\sum_{x_i \in A}^{n} \left(\left(x_i - \mu\right) + \left(\mu - \bar{X}\right)\right)^2\right) && (29) \\
&= \mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \mu\right)^2 + \sum_{x_i \in A}^{n} 2\left(x_i - \mu\right)\left(\mu - \bar{X}\right) + \sum_{x_i \in A}^{n} \left(\mu - \bar{X}\right)^2\right) && (30) \\
&= \mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \mu\right)^2 + 2\left(\mu - \bar{X}\right)\sum_{x_i \in A}^{n} \left(x_i - \mu\right) + n\left(\mu - \bar{X}\right)^2\right) && (31)
\end{aligned}
$$

和の期待値は期待値の和なので

$$= \mathsf{E}\left(\sum_{x_i \in A}^{n} (x_i - \mu)^2\right) + \mathsf{E}\left(2\left(\mu - \bar{X}\right)\sum_{x_i \in A}^{n} (x_i - \mu)\right) + \mathsf{E}\left(n\left(\mu - \bar{X}\right)^2\right) \tag{32}$$

ここで、第 1 項は i ごとに母分散の定義 $\mathsf{E}\left((x_i - \mu)^2\right) = \sigma^2$ より、$n\sigma^2$。

第 2 項総和部分は

$$\sum_{x_i \in A}^{n} (x_i - \mu) = \sum_{x_i \in A}^{n} x_i - n\mu \tag{33}$$

$$= n\bar{X} - n\mu \tag{34}$$

$$= -n\left(\mu - \bar{X}\right) \tag{35}$$

よって第 2 項は

$$\mathsf{E}\left(2\big(\mu - \bar{X}\big)\left(-n\left(\mu - \bar{X}\right)\right)\right) = -2n\mathsf{E}\left(\left(\mu - \bar{X}\right)^2\right) \tag{36}$$

従って (32) は

$$= n\sigma^2 - 2n\mathsf{E}\left(\left(\mu - \bar{X}\right)^2\right) + n\mathsf{E}\left(\left(\mu - \bar{X}\right)^2\right) \tag{37}$$

$$= n\sigma^2 - n\mathsf{E}\left(\left(\mu - \bar{X}\right)^2\right) \tag{38}$$

(38) 第 2 項期待値部分は標本平均 $\bar{X}$ の分散なので $\mathsf{V}\left(\bar{X}\right)$ とあらわすと

$$\mathsf{V}\left(\bar{X}\right) = \mathsf{V}\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \tag{39}$$

$$= \frac{1}{n^2}\mathsf{V}\left(x_1 + x_2 + \cdots + x_n\right) \tag{40}$$

各 x_i は独立なので分散の和に分解すると

$$= \frac{1}{n^2}\left(\mathsf{V}\left(x_1\right) + \mathsf{V}\left(x_2\right) + \cdots + \mathsf{V}\left(x_n\right)\right) \tag{41}$$

定義により $\mathsf{V}\left(x_i\right) = \sigma^2$ なので

$$= \frac{1}{n^2}n\sigma^2 \tag{42}$$

$$= \frac{\sigma^2}{n} \tag{43}$$

よって (38) は

$$n\sigma^2 - n \times \frac{\sigma^2}{n} = n\sigma^2 - \sigma^2 \tag{44}$$

$$= (n-1)\,\sigma^2 \tag{45}$$

(28) 左辺と (45) を並べると

$$\mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2\right) = (n-1)\,\sigma^2 \tag{46}$$

ここで、$\mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2 + \sum_{y_j \in B}^{m} \left(y_j - \bar{Y}\right)^2\right)$ を求める。

$$\mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2 + \sum_{y_j \in B}^{m} \left(y_j - \bar{Y}\right)^2\right) = \mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2\right) + \mathsf{E}\left(\sum_{y_j \in B}^{m} \left(y_j - \bar{Y}\right)^2\right) \tag{47}$$

$$= (n-1)\,\sigma^2 + (m-1)\,\sigma^2 \tag{48}$$

$$= \sigma^2\left((n-1) + (m-1)\right) \tag{49}$$

$$= \sigma^2\left(n + m - 2\right) \tag{50}$$

8.3 $\boldsymbol{F}$ 分布

自由度 $\nu = n$ の χ^2 分布に従う確率変数を χ_x^2，自由度 $\nu = m$ の χ^2 分布に従う確率変数を χ_y^2 とする。この二つの χ^2 分布に従う変数 χ_x^2，χ_y^2，を用いて統計量

$$F_{(n,m)} = \frac{\left(\frac{\chi_x^2}{n}\right)}{\left(\frac{\chi_y^2}{m}\right)} \tag{51}$$

を作る。このとき統計量 $F_{(n,m)}$ が従う分布を『**自由度 $(\boldsymbol{n},\ \boldsymbol{m})$ の $\boldsymbol{F}$ 分布**』という。

8.4 t 分布

標準正規分布 $N(0,\ 1)$ に従う確率変数 $Z = z$ と、これとは独立に自由度 $\nu = n$ の χ^2 分布に従う確率変数 χ_z^2 を用いて

$$t_0 = \frac{z}{\sqrt{\frac{\chi_z^2}{n}}} \tag{52}$$

をつくる。このとき t_0 が従う分布を『**t 分布**』といい、$\nu = n$ を『**t 分布の自由度**』という。

8.5　母分散が未知の場合の母平均の差の検定

母分散 σ^2 が未知の場合は、母分散が等しいかどうかで推定方法が異なる。ここでは母分散が等しい場合を説明する。

$$x_i \sim N\left(\mu_x, \sigma_x^2\right) \quad ; i = 1, 2, \cdots, n \tag{53}$$

$$y_j \sim N\left(\mu_y, \sigma_y^2\right) \quad ; j = 1, 2, \cdots, m \tag{54}$$

とする。

8.5.1　等分散の検定

母分散 σ_x^2, σ_y^2 が未知なので、分散が等しいかどうかを検定する。

仮説を

$$H_0 : \quad \sigma_x^2 = \sigma_y^2 \tag{55}$$

$$H_1 : \quad \sigma_x^2 \neq \sigma_y^2 \tag{56}$$

とする。

それぞれの不偏分散

$$ {s_x}^2 = \frac{1}{n-1} \sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2 \tag{57} $$

$$ {s_y}^2 = \frac{1}{m-1} \sum_{y_j \in B}^{m} \left(y_j - \bar{Y}\right)^2 \tag{58} $$

のうち、値の大きい方を**分子**にして統計量 F_0 をつくる。ここで、

$$ {s_x}^2 > {s_y}^2 \tag{59} $$

とすると、F_0 は

$$ F_0 = \frac{{s_x}^2}{{s_y}^2} \tag{60} $$

である。

ここで、${s_x}^2$ は自由度 $(n-1)$ の χ^2 分布に従い、${s_y}^2$ は自由度 $(m-1)$ の χ^2 分布に従うから、F_0 は自由度 $(n-1,\ m-1)$ の F 分布に従う。

　そこで有意水準両側 5 % で検定すると

$$F_0 \geqq F(n-1,\ m-1, 0.025) \tag{61}$$

のとき、H_0 が棄却される。

　等分散の検定で H_0 が棄却されず、X と Y の分散が異なるとは言えない場合、

$$\sigma_x^2 = \sigma_y^2 = \sigma^2 \tag{62}$$

として次の『**母平均の差の検定**』を行う。

8.5.2　母分散が等しい場合の母平均の差の検定

仮定を

$$H_0 : \quad \mu_x - \mu_y = 0 \tag{63}$$

$$H_1 : \quad \mu_x - \mu_y \neq 0 \tag{64}$$

である。ここで

$$\mathsf{E}\left(\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2 + \sum_{y_j \in B}^{m} \left(y_j - \bar{Y}\right)^2\right) = \sigma^2 \left(n + m - 2\right) \tag{65}$$

なので、X, Y の共通の分散を

$$\hat{\sigma}^2 = \frac{\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2 + \sum_{y_j \in B}^{m} \left(y_j - \bar{Y}\right)^2}{n + m - 2} \tag{66}$$

で推定する。

$$\bar{X} - \bar{Y} \sim N\left(\mu_x - \mu_y,\ \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}\right) \tag{67}$$

において、

$$\sigma_x^2 = \sigma_y^2 = \frac{\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2 + \sum_{y_j \in B}^{m} \left(y_j - \bar{Y}\right)^2}{n + m - 2} \tag{68}$$

とすると、標本平均の差の分散 $\mathsf{V}\left(\bar{X}-\bar{Y}\right)$ は

$$\mathsf{V}\left(\bar{X}-\bar{Y}\right)=\frac{\left(\frac{\sum_{x_i\in A}^{n}\left(x_i-\bar{X}\right)^2+\sum_{y_j\in B}^{m}\left(y_j-\bar{Y}\right)^2}{n+m-2}\right)}{n}+\frac{\left(\frac{\sum_{x_i\in A}^{n}\left(x_i-\bar{X}\right)^2+\sum_{y_j\in B}^{m}\left(y_j-\bar{Y}\right)^2}{n+m-2}\right)}{m} \tag{69}$$

$$=\left(\frac{\sum_{x_i\in A}^{n}\left(x_i-\bar{X}\right)^2+\sum_{y_j\in B}^{m}\left(y_j-\bar{Y}\right)^2}{n+m-2}\right)\left(\frac{1}{n}+\frac{1}{m}\right) \tag{70}$$

である。

このとき、

$$\chi^2_{\bar{X}+\bar{Y}} = \chi^2_{\bar{X}} + \chi^2_{\bar{Y}} \tag{71}$$

は自由度 $\nu = n + m - 2$ の χ^2 分布に従うので、

$$t_0 = \frac{\left(\bar{X} - \bar{Y}\right) - (\mu_x - \mu_y)}{\sqrt{\left(\frac{\sum_{x_i \in A}^{n}\left(x_i - \bar{X}\right)^2 + \sum_{y_j \in B}^{m}\left(y_j - \bar{Y}\right)^2}{n+m-2}\right)\left(\frac{1}{n} + \frac{1}{m}\right)}} \tag{72}$$

は自由度 $n + m - 2$ の t 分布に従う。

この t_0 を

$$H_0: \quad \mu_x - \mu_y = 0 \tag{73}$$

の下で、両側 5% で検定すると、

$$t_0 \leqq -t(n+m-2,\ 0.05) \tag{74}$$

または、

$$t_0 \geqq t(n+m-2,\ 0.05) \tag{75}$$

のとき H_0 が棄却される。

例題

表 1　数値例

A 群	10.5	9.8	10.9	9.7	10.6
B 群	11.1	11.7	10.8	11.5	10.9

A 群と B 群の母平均に差があるかを有意水準 5% で検定する。

等分散の検定

母分散に対する仮説を

$$H_0 \quad : \sigma_a^2 = \sigma_b^2 \tag{76}$$

$$H_1 \quad : \sigma_a^2 \neq \sigma_b^2 \tag{77}$$

とする。

$$\bar{A} = \frac{10.5 + 9.8 + 10.9 + 9.7 + 10.6}{5} = 10.3 \tag{78}$$

$$\bar{B} = \frac{11.1 + 11.7 + 10.8 + 11.5 + 10.9}{5} = 11.2 \tag{79}$$

$$\begin{aligned}\left(x_i - \bar{X}\right)^2 &= (10.5 - 10.3)^2 + (9.8 - 10.3)^2 + (10.9 - 10.3)^2 \\ &\quad + (9.7 - 10.3)^2 + (10.6 - 10.3)^2 = 1.10 \end{aligned} \tag{80}$$

$$\begin{aligned}\left(y_i - \bar{Y}\right)^2 &= (-0.1)^2 + (0.5)^2 + (-0.4)^2 + (0.3)^2 + (-0.3)^2 \\ &= 0.60 \end{aligned} \tag{81}$$

$$s_a{}^2 = \frac{1.10}{5-1} = 0.275 \tag{82}$$

$$s_b{}^2 = \frac{0.60}{4} = 0.15 \tag{83}$$

$$F_0 = \frac{s_a{}^2}{s_b{}^2} = \frac{0.275}{0.15} = 1.83 \tag{84}$$

有意水準 α の F 分布の上側の臨界値を $F(\nu_1,\ \nu_2,\ \alpha/2)$ であらわすと、臨界値は

$$F(4, 4, 0.025) = 9.604\,529\,885 \tag{85}$$

なので[1]

$$F_0 \leqq F(4,\ 4,\ 0.025) \tag{86}$$

なので、H_0 は棄却されない。よって『分散が異なる』とは言えないので等分散の場合の平均値の差の検定を行う。

1） Excel を用いて F.INV.RT(0.025,4,4) で算出。F 分布表であれば 9.60。

母平均の差の検定

母平均の差に対する仮説を

$$H_0: \quad \mu_a - \mu_b = 0 \tag{87}$$

$$H_1: \quad \mu_a - \mu_b \neq 0 \tag{88}$$

とする。$\mathrm{V}\left(\bar{X} - \bar{Y}\right)$ の推定値 $\hat{\sigma^2}$ は

$$\hat{\sigma^2} = \frac{\sum_{x_i \in A}^{n} \left(x_i - \bar{X}\right)^2 + \sum_{y_j \in B}^{m} \left(y_i - \bar{Y}\right)^2}{n + m - 2} \tag{89}$$

$$= \frac{1.1 + 0.6}{5 + 5 - 2} = 0.2125 \tag{90}$$

よって

$$t_0 = \frac{\bar{A} - \bar{B}}{\sqrt{\frac{\sum\left(x_i - \bar{X}\right)^2 + \sum\left(y_i - \bar{Y}\right)^2}{n+m-2}}\sqrt{\frac{1}{n} + \frac{1}{m}}} \tag{91}$$

$$= \frac{10.3 - 11.2}{\sqrt{\frac{1.1+0.6}{5+5-2}}\sqrt{\frac{1}{5} + \frac{1}{5}}} \tag{92}$$

$$= \frac{-0.9}{\sqrt{\frac{1.7+0.6}{8}}\sqrt{\frac{2}{5}}} \tag{93}$$

$$= -3.09 \tag{94}$$

有意水準 α の自由度 ν の t 分布の臨界値を $t(\nu,\ \alpha)$ とあらわすと、 臨界値は[2)]

$$t(8,\ 0.05) = 2.306\,004\,135 \tag{95}$$

なので

$$|t_0| > t(8,\ 0.05) \tag{96}$$

H_0 は棄却され、A 群と B 群には『差がない』とは言えない。

2） Excel で T.INV.2T(0.05,8) で算出。t 分布表であれば 2.30。

8.6　母分散が異なる場合の平均値の差の検定【参考】

等分散の検定で

$$H_0: \quad \sigma_a^2 = \sigma_b^2 \tag{97}$$

が棄却された場合は、

$$u_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} \tag{98}$$

の σ_x^2, σ_y^2 の代わりに不偏分散 s_x^2, s_y^2 を用いて、

$$t_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} \tag{99}$$

で検定を行う。この場合 t 分布の自由度は

$$\nu_0 = \frac{\left(\frac{s_x^2}{n} + \frac{s_y^2}{m}\right)^2}{\frac{\left(\frac{s_x^2}{n}\right)^2}{n-1} + \frac{\left(\frac{s_y^2}{m}\right)^2}{m-1}} \tag{100}$$

この検定方法を Welch の t 検定（ウェルチの t 検定) という。

8.7 まとめ

- 母分散は通常未知である。
- 未知の場合は等分散の検定を行い等分散であることを確認する。
- 等分散であることが確認出来たら F 検定を行う。
- ここでは不偏分散が小さい方を分母にしたので上側で検定しているが、分母が小さい場合は下側検定になる。
- 参考レベルで母分散が等しくない場合も紹介をした。

参考文献

1. 『経済数学教室 1 巻』小山昭雄著「岩波書店」1994 年 5 月 30 日
2. 『新体系 大学数学 入門の教科書 上下』吉沢光雄著「講談社」2022 年 6 月 16 日
3. 『技術者のための高等数学 7 確率と統計（原書代 8 版）』E・クライツィグ著　田栗正章訳「培風館」2004 年 11 月 30 日第 8 版
4. 『データ解析のための数理統計入門』久保川達也著「共立出版」2023 年 10 月 15 日初版
5. 『統計学入門』東京大学教養学部統計学教室篇「東京大学出版会」1991 年 7 月 9 日初版
6. 『行動科学における統計解析法」芝祐順・南風原朝和著　「東京大学出版会」1990 年 3 月 20 日　初版
7. 『計量経済学の理論』A.S. ゴールドバーガー著, 福地崇生・森口親司訳「東洋経済新報社」

1970 年 9 月 10 日 第 1 刷

8.『計量経済学序説』R.J. ウォナコット・T.H. ウォナコット著, 国府田恒夫・田中一盛訳「培風館」 1975 年 6 月 10 日初版
9.『マーケティング・サイエンス』片平秀貴著　「東京大学出版会」1987 年 4 月 20 日初版
10.『回帰分析』佐和光男著 「朝倉書店」1979 年 4 月 20 日初版
11.『完全独習　統計学入門』小島寛之著「ダイヤモンド社」2006 年 9 月 28 日　初版
12.『分散分析の基礎』高橋敬子著「プレアデス出版」2009 年 10 月 1 日