

第 7 章

2群の平均の差の検定

2群の平均の差の検定：処理の違いによって生じた2群の標本平均を比較することで、それらの母平均にも差があることを検証する。標本間のデータに対応関係がある場合は、個体差を考慮した検定統計量を求めることができる。

7.1 もっともよく使われる検定手法

前章では、統計的検定の基本的な考え方を学ぶため、1つの標本平均（の母平均）と特定の値（定数）との差が統計的に有意であるかどうかを検定しました。しかし、その検定の活躍の場はそう多くはありません。

みなさんが実際に、もっともよく使うのは**2群（標本）の平均の差の検定**（two sample test of means）でしょう。この検定は、例えば肥料を使う（処理）前と使った（処理）後の平均成長速度の差や、農村と都市との平均所得の差など、実験や調査で観測したデータを処理の違いで2群（3群以上については次章）に分け、それら2つの標本の平均を比較することで、抽出元の2つの母集団の平均にも差があるといえるかどうかを判定します。

実は、母平均の検定も、2群の平均の差の検定の1種ではあったのです。しかし、図7.1左のように、片方は特定の値として与えられていたので、2群として考える必要がなかっただけです。なお、2群の平均の差の検定にもいろいろな呼び方があり、例えば「2条件の平均値の差の検定」や「2標本の母平均の差の検定」、「2つの母平均の差の検定」などとも呼ばれています。いずれの名称も、標本やその抽出元の母集団が2つであることや、平均の差であることが強調されていますね。また、「*t*検定を実施しなさい」といわれたら、この2群の平均の差の検定のことを探していると思っていただいて結構です。

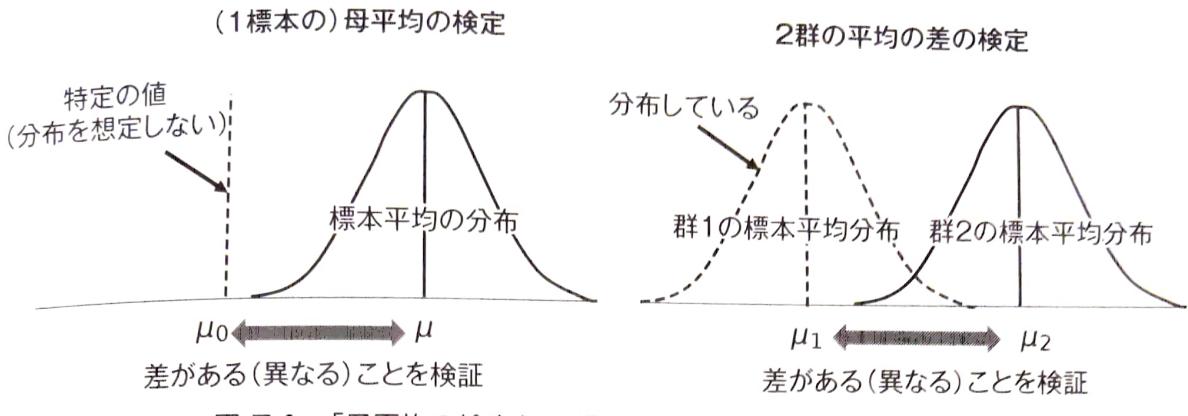


図 7.1 「母平均の検定」と「2群の平均の差の検定」の違い

7.2 対応関係

2群の平均の差の検定といっても、データの取り方によって2種類あります。なぜならば、観測対象となる個人や個体、ケース(case)が、2群(標本)間で**対応のある**関係(paired, dependent)か、**対応のない**独立した関係(unpaired, independent)かで、検定統計量の算出方法が異なるのです。

例えば、新開発のダイエット薬の効果を検証するために、偽薬(プラセボ)と実薬(本物)をそれぞれ4名の被験者に1週間使用してもらって、その間の体重の変化を測定する場面を想像してみてください。図7.2左が「対応のない」場合で、偽薬を試した4名(Aさん, Bさん, Cさん, Dさん)とは異なる4名(Eさん, Fさん, Gさん, Hさん)に実薬を試して体重変化を測定しています。それに対して図7.2右の「対応のある」場合では、同じ4名(Aさん, Bさん, Cさん, Dさん)に、どちらも(偽・実)試して、体重変化を測定しています(つまり、この4名は2回異なる薬を飲んで2回測定することになります)。

どちらも、両群の体重変化の標本平均を計算し、その差が統計的に有意な大きさかどうかを検定するという目的は同じなのですが、対応のある場合には個体差(人ならば個人差)を考慮した検定統計量を推定できるという大きなメ

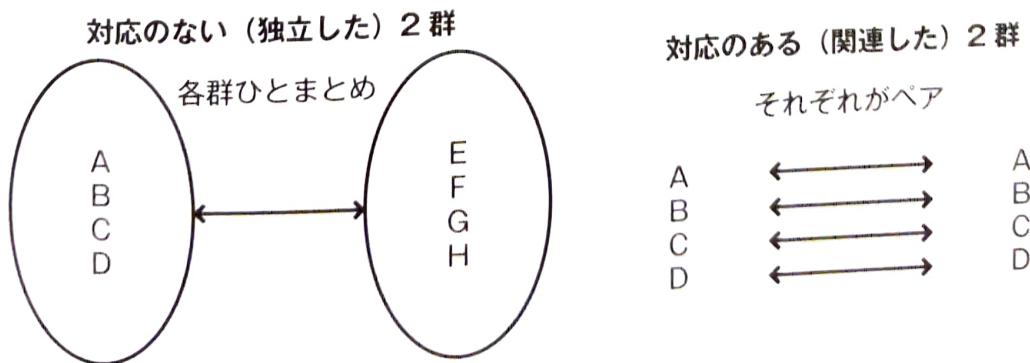


図 7.2 2群の対応関係

リットが生まれます。後ほど詳しく説明しますが、個体差を考慮するとは、分析する側にとって誤差である個体差が大きいときには帰無仮説を棄却しにくくして、保守的な判定を下すようにすることです。ですから、人間や動植物など、個体差が大きいと考えられる分野では、対応のあるデータを観測できるような調査や実験を計画するのが望ましいといえるでしょう。

とはいって、異なる2つの土壤での植物の生育差や、農村と都市での世帯所得差、破壊しなければならない検査など、対応のあるデータを観測するのが物理的に不可能な場合も多いでしょう。その場合には、適用する検定の種類だけ間違わないようにしてください。対応ありデータに対応なし用の検定を使うには問題はないのですが、その逆（対応なしのデータに対応あり用の検定を使うの）は“間違い”なので、くれぐれも気をつけなければなりません。

まずは、どちらの対応関係でも使える、対応のない2群の場合から説明していきます。

7.3 仮説の設定と検定統計量 一対応のない2群

● 仮説の設定

母平均の検定のときと同様、研究目的からすれば棄却したい帰無仮説と、帰無仮説が棄却されたときに採択される対立仮説を設定します。

処理前に観測した標本の平均を \bar{x}_1 、その抽出元である母集団の平均を μ_1 とし、処理後に観測した標本の平均値を \bar{x}_2 、その抽出元である母集団の平均を μ_2 とすると、帰無仮説は図7.3左のように、 $H_0 : \mu_1 = \mu_2$ となります。つまり、どちらの標本も1つの同じ母集団から抽出されていたことになり、処理の効果はなかった（＝母平均に差はなかった）という内容です。

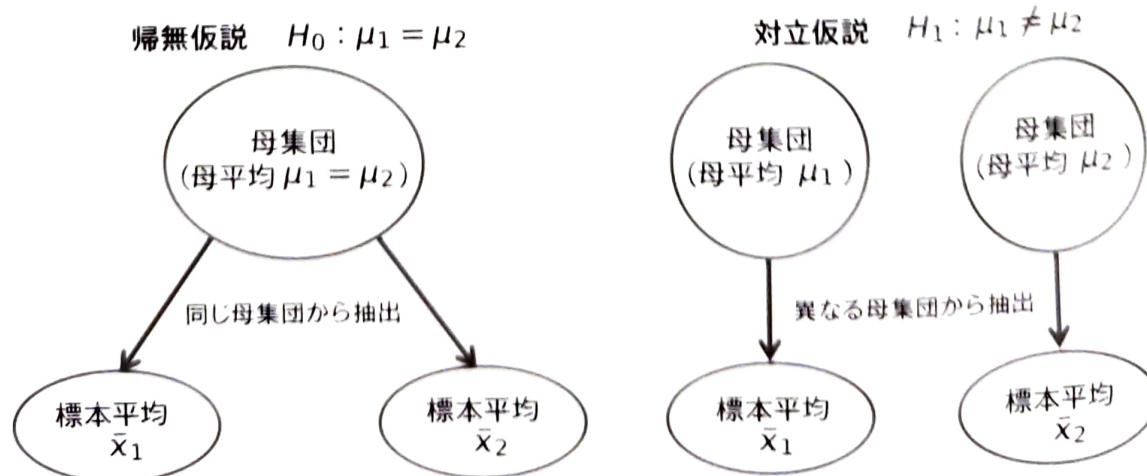


図 7.3 「2群の平均の差の検定」の仮説の考え方

一方、対立仮説は帰無仮説とは逆の内容になるので、 $H_1 : \mu_1 \neq \mu_2$ となり、2つの標本はそれぞれ別な母集団から抽出されたことになり、処理の効果はあった（＝母平均に差があった）という内容です。

● 検定統計量① 一母分散が既知の場合（z検定）――

仮説を設定したら、帰無仮説が正しいと仮定した下での検定統計量を計算します。まずは、母分散が既知の場合にのみ計算できるz検定用の統計量を説明し、次に母分散が未知の場合でも計算できるt検定用の統計量を説明します。

さて、2つの標本平均から統計量の計算をしようとすると、面倒な問題が出てきます。図7.1をもう一度見てください。図左の母平均の検定では、特定の値（定数）が与えられていたため、それと標本平均との離れ具合ははっきりしており、それらの差が誤差の範囲か否かを評価するのは容易でした。しかし、図右の2群の平均の差の検定では、どちらも中心（母平均）のわからない標本分布が2つあるだけなので、その差がはっきりしないのです。つまり、標本平均やその（準）標準化変量のままでは検定統計量として使えません……。

しかし、心配はいりません。ちゃんと、上手な方法が考え出されています。

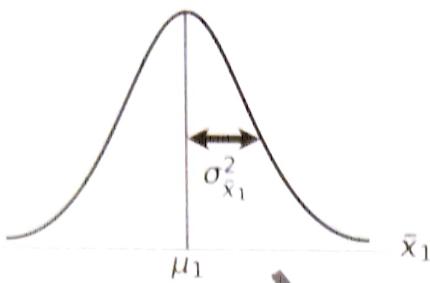
検定統計量とは、帰無仮説が正しい下での標本分布です。本検定の帰無仮説は、 $\mu_1 = \mu_2$ でした。これは、 $\mu_1 - \mu_2 = 0$ のことで、0は定数です。つまり、“（2つの標本の）平均の差の分布”として考えれば、母平均の検定のように、0という特定の値からの離れ具合を簡単に計算できるのです。また、都合が良いことに、正規分布に従う母集団から抽出した標本平均同士の差も正規分布に従います。

以上、2群の平均の差の検定では、標本平均そのままでなく、2つの標本平均の差 ($\bar{x}_1 - \bar{x}_2$)、あるいはそれを標準化した $z_{\bar{x}_1 - \bar{x}_2}$ 、または準標準化した $t_{\bar{x}_1 - \bar{x}_2}$ を検定統計量として用いれば良いことがわかりました。

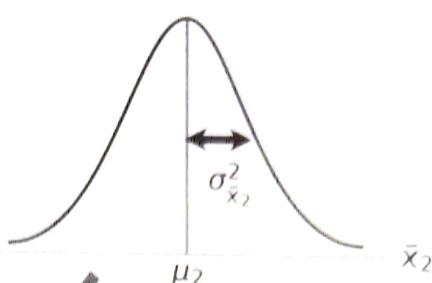
ただし、平均の差から検定統計量を導くにあたって、1つだけ注意することがあります。図7.4のように、平均差の分布の中心（母平均）は、両方の母平均の差 ($\mu_1 - \mu_2$) でOKなのですが、問題はバラツキの代表とではなく“和” ($\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2$) になるのです。

なぜそうなるのか不思議に思われるかもしれません、例えば、ある部品（長さの平均100 mm、分散10 mm²）に毎回異なるドリル（長さの平均が50 mm、分散が5 mm²）で穴を空けて製品を完成させるとします。その製品の穴の残りの厚みの平均と分散を考えてみてください。厚みの平均は差（100 - 50 = 50）

正規分布に従う標本平均 1



正規分布に従う標本平均 2



標本平均の差も正規分布に従う

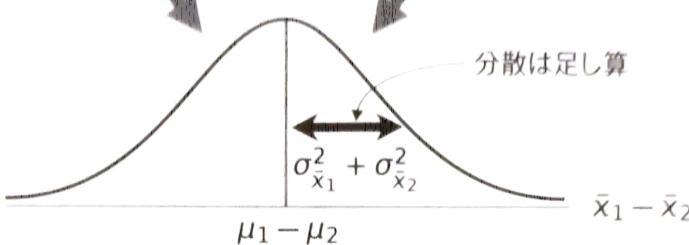


図 7.4 分散の加法性

となるでしょうが、厚みの分散は部品とドリル両方の分散の和 ($10 + 5 = 15$) となることがイメージできるのではないですか？この事例は個別データの分散ですが、もちろん平均の分散（誤差分散）にも同じ加法性が成り立ちます。これを**分散の加法性** (additivity of variance) と呼び、合わせた分散を**合併分散（併合分散）** (weighted average of variance) と呼びます。

以上、2つの標本平均の差 ($\bar{x}_1 - \bar{x}_2$) は、母平均 ($\mu_1 - \mu_2$)、母誤差分散 ($\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2$) の正規分布に従うことがわかりました。

さて、ようやく検定統計量の計算に入りますが、前章のように、(準)標準化していない標本平均差 ($\bar{x}_1 - \bar{x}_2$) から説明しても、対象が異なる度に（分布の中心やバラツキが変わるため）限界値が変化して使いにくいので、本節では平均差の標準化変量 $z_{\bar{x}_1 - \bar{x}_2}$ (z 検定の検定統計量) から始めましょう。

復習ですが、標本平均 \bar{x} から標準化変量 $z_{\bar{x}}$ への変換式は次のようにでした。なお、 $\sigma_{\bar{x}}$ は母標準誤差ですので、母誤差分散の平方根となります。

$$\text{標本平均の標準化変量 } z_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

標本平均の差の標準化変量 $z_{\bar{x}_1 - \bar{x}_2}$ は、この式の \bar{x} を $\bar{x}_1 - \bar{x}_2$ に、 μ を $\mu_1 - \mu_2$ に、 $\sigma_{\bar{x}}$ を $\sigma_{\bar{x}_1 - \bar{x}_2}$ に、それぞれ置き換えた次式になります。

$$\text{標本平均の差の標準化変量 } z_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

ここで、分母の母標準誤差 $\sigma_{\bar{x}_1 - \bar{x}_2}$ は、先ほど説明した合弁分散（母誤差分散）の平方根ですので、次のようにになります。ただし、 σ_1^2 と σ_2^2 はそれぞれの群の母分散、 n_1 と n_2 はそれぞれの群の標本サイズです。

$$\text{標本平均の差の母標準誤差} \quad \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\sigma_{\bar{x}_1 - \bar{x}_2}^2} = \sqrt{\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

いよいよ、2群の平均の差の z 検定の統計量ですが、検定統計量は帰無仮説 ($\mu_1 = \mu_2$) が正しいとした下で計算される統計量です。ですから、この式の分子も分母も帰無仮説でどのようになるのかを考えなければなりません。

まず、 $\mu_1 = \mu_2$ のですから分子の $\mu_1 - \mu_2$ は 0 になります。 $\bar{x}_1 - \bar{x}_2$ だけが残りますね。また分母も、同じ母集団から抽出されたのならば、両群ともバラツキは同じ（等分散；homoscedasticity）はずなので、 σ_1^2 と σ_2^2 を共通の分散である σ^2 に統一して括弧の外に出すことができます（この等分散の条件は重要なので章末で検討します）。よって、 z 検定の統計量は下記のように整理することができます（矢印よりも右が検定統計量）。

z 検定の統計量（対応なし2群）

$$z_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightarrow \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

↑ 帰無仮説の下では0
↑ 帰無仮説の下では等分散

● 検定統計量② 一母分散が未知の場合（ t 検定）一

①で導いた z 検定の統計量は、分母の $\sqrt{\cdot}$ の中にある母分散 σ^2 が既知の場合にしか計算することができません。しかし、母集団のバラツキがわかっている状況は滅多にありません。ですから、実際には、次式のように母分散 σ^2 を不偏分散 $\hat{\sigma}^2$ に置き換えた準標準化変量 t を統計量として、 t 検定を実施します。

$$t \text{ 検定の統計量 (途中)} \quad t_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2}{\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

「これでようやくt検定ができる！」と思ってしまった方もいらっしゃるでしょうが、もう少々お待ちください。このままでは、まだ計算できません。というのも、不偏分散 $\hat{\sigma}^2$ はどのように求めますか？2つの標本から不偏分散 $\hat{\sigma}_1^2$ と $\hat{\sigma}_2^2$ は、それぞれ計算できますが、 $\hat{\sigma}_1^2$ と $\hat{\sigma}_2^2$ の値は当然、異なるでしょう。どちらを使えば良いかなど誰にもわかりません。

そこで、次のように、それぞれの自由度 $(n_1 - 1, n_2 - 1)$ で加重した不偏分散の平均（第1章で学んだ加重平均です）を2群共通の不偏分散 $\hat{\sigma}^2$ とします。

$$\text{2群共通の不偏分散 } \hat{\sigma}^2 = \frac{(n_1 - 1)\hat{\sigma}_1^2 + (n_2 - 1)\hat{\sigma}_2^2}{(n_1 - 1) + (n_2 - 1)}$$

これで「対応のない2群の平均の差のt検定」の統計量の式（標本サイズがアンバランスなとき用）が得られました。

t検定の統計量（対応なし2群、アンバランス）

$$t_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1)\hat{\sigma}_1^2 + (n_2 - 1)\hat{\sigma}_2^2}{(n_1 - 1) + (n_2 - 1)} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

両群の標本がどちらも同じサイズ $(n_1 = n_2 = n)$ でバランスなときは、加重させる必要がないので、次のような単純な形になります。

$$\text{t検定の統計量（対応なし2群、バランス） } t_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}{n}}}$$

以上が、対応のない2群の平均の差のt検定の統計量となります。標本分散 s_1^2 と s_2^2 を使って分母を計算する場合には、 n ではなく自由度 $(n - 1)$ で割ることをお忘れなく。なお、とくにこの検定をスチューデントのt検定と呼ぶことがあります。

● 仮説の検定

検定統計量を解説するだけで時間がかかってしまいましたが、検定の手続き自体は前章で学んだ母平均の検定と変わりません。検定統計量（tやz）の絶対値が、限界値よりも大きいか小さいかを見るだけです。以下、復習がてら確認しておきましょう。

検定統計量の絶対値が限界値よりも大きな値を取った場合（図7.5のように帰無仮説の分布で両端に位置したとき）には、同じ母集団から抽出された（帰無仮説が正しい）にしては標本平均の差が大きすぎる（偶然には起こりにくい）ことになり、「もともとの母集団の平均（母平均）自体にも差があった」→「処理の効果はあった」と判断します。

逆に検定統計量の絶対値が限界値よりも小さな値を取った場合（帰無仮説の分布で中心付近に位置したとき）には、偶然でも十分発生する程度の小さな差なので、同じ母集団から抽出されたという帰無仮説は否定できないため、「母集団の平均には統計的有意差は検出できなかった」→「今回の実験では、処理効果を統計的に確認できなかった」と判断します（帰無仮説を正しいとは断言しないように）。

検定統計量が z でも t でも、この手続きは変わりませんが、限界値の決まり方が異なります。 z 検定では限界値の値は有意水準の大きさだけで決まりますが（例えば有意水準 α が両側5%ならば限界値は ± 1.96 ）、 t 検定では分布形が自由度によって変わるため、限界値の値も有意水準だけでなく自由度の大きさが影響します。自由度の大きさは、アンバランスの場合で $(n_1 - 1) + (n_2 - 1)$ 、バランスな場合で $2(n - 1)$ となります。

なお、ソフトウェアなどで p 値を計算できるならば、有意水準と比較して p 値の方が小さければ帰無仮説を棄却し、大きければ受容するという判定でも構いません。

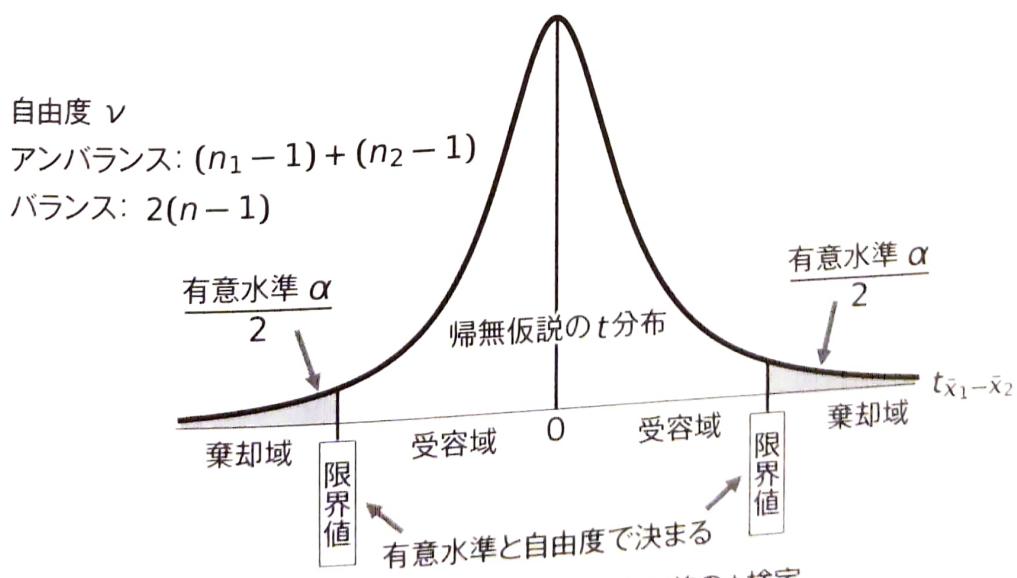


図 7.5 対応のない2群の平均の差の t 検定

例題

第1章のキュウリの収量データは、昼間に養液を与える栽培法Aと夜間に養液を与える栽培法Bとの収量を比較したものでした。栽培法Aの平均収量は2443.7g、栽培法Bは3124.4gで、両栽培法の間には680.7gという収量の差があり、一見、栽培法Bの方が良さそうです。果たして本当に両栽培法の収量に差があるといえるでしょうか？

解：

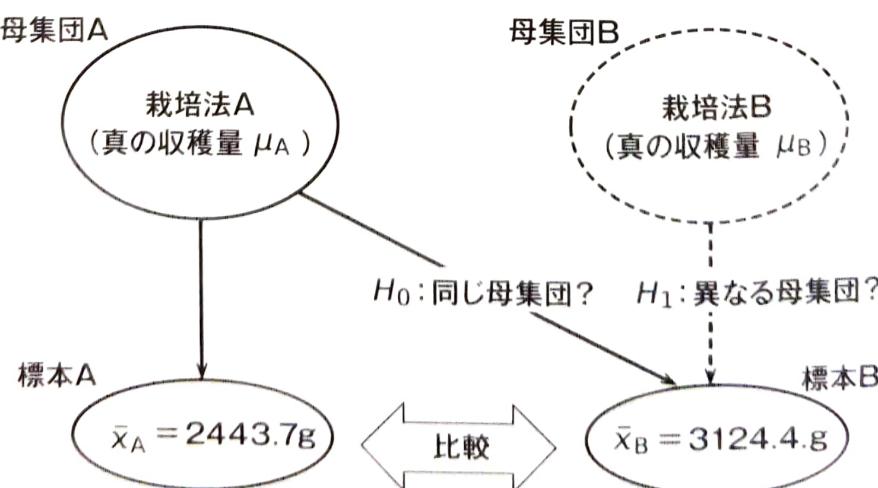
キュウリは2度収穫できないので、2群は対応のない関係です。また、母分散は未知なのでt分布を使ったt検定を実施しましょう。なお、必ず栽培法Bが大きくなるという根拠はないので、両側5%の有意水準を用いましょう。

以下、栽培法Aの収量の標本平均を \bar{x}_A 、母平均を μ_A 、栽培法Bの標本平均を \bar{x}_B 、母平均を μ_B とします。

手順①：仮説の設定

本当に2つの栽培法の収量に差があるということは、実験データの抽出元である母集団の収量（母平均で真の値）に差があることです。よって、帰無仮説は、「どちらの標本も同じ母集団から抽出された」という内容になります。帰無仮説はその逆なので、「異なる母集団から抽出された」という内容です。

$$\left\{ \begin{array}{l} \text{帰無仮説 } H_0 : \mu_A(\bar{x}_A = 2443.7) = \mu_B(\bar{x}_B = 3124.4) \rightarrow \text{どちらの栽培法も同じ収量} \\ \text{対立仮説 } H_1 : \mu_A(\bar{x}_A = 2443.7) \neq \mu_B(\bar{x}_B = 3124.4) \rightarrow \text{栽培法によって収量が異なる} \end{array} \right.$$



例題の仮説の考え方

手順②：検定統計量の計算

帰無仮説が正しい場合の両平均収量の差を標準化したt値を算出します。今回はバランスな（両群の標本サイズが同じ）場合の式が使えます。

両群の不偏分散 ($\hat{\sigma}_A^2$ と $\hat{\sigma}_B^2$) は第3章の例題で計算していますね（栽培法A：183227.0, 栽培法B：133515.4）。それらの値を代入すると、検定統計量tは“-4.685”となります。

$$t_{\bar{x}_A - \bar{x}_B} = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\hat{\sigma}_A^2 + \hat{\sigma}_B^2}{n}}} = \frac{2443.7 - 3124.4}{\sqrt{\frac{183227.0 + 133515.4}{15}}} = -4.685$$

手順③：確率の計算

ソフトウェアを使えば、帰無仮説のt分布における本検定統計量の値（-4.685）より外側の確率（p値）を計算することができます（Excel関数ならば、T.DIST.2T(4.685, 28)です）。ただし、限界値と比較する方法ならば、手作業で仮説を判定できるため、本手順はスキップします。

手順④：仮説の判定（と検出力）

あらかじめ決めておいた両側5%の有意水準 α に対応するt値が限界値となります。付録IIのt分布表から読み取ると、例題の自由度は $2 \times (15 - 1)$ なので $v=28$ の行と、片側確率 $p=0.025$ の列がクロスする“2.048”が限界値となります。もちろん両側で検定しますので、下側（左裾）の“-2.048”も限界値となります。よって、例題の実験で生じた平均の差のt値（-4.685）は、限界値の-2.048よりも小さいため（分布の外側に位置するため）、帰無仮説が正しいには、かなり起きにくいことが起きましたといえます。ですから、むしろ帰無仮説自体が間違っていると考えた方が合理的です。よって、有意水準5%という判定基準が間違っているとされた方が合理的です。

ここで、G*powerの [Post hoc: Compute achieved power] を使って検出力を計算してみましょう。対応のない2群の差の検定は [Means: Difference between two independent means] です。本検定の効果量は、もっとも簡単な推定式である $t \times \sqrt{(n_A + n_B)/(n_A \times n_B)}$ を使って計算すると1.71となります。これに加え、両側有意水準 α に0.05、各標本サイズにどちらも15をパラメータとして入力すると、検出力は0.99と推定されます。今回の検定の検出力がとても強かった理由としては、推定された効果量が大きかったからでしょう。

補足 分析ツールを使った検定

- ① Excelには標準で「分析ツール」という統計解析ソフトが搭載されています。使用するには、最初の1回目だけ、[ファイル]タブから[オプション]を選択して、[アドイン]→[管理から「Excelアドイン」を選択]→[設定]→[「分析ツール」に]で設定する必要があります(Windows版)。すると、[データ]タブに[データ分析]が追加されますので、それをクリックすると分析ツールが現れます。
- ② いろいろな手法がありますが、[t検定：等分散を仮定した2標本による検定]が今回の「対応のない2群の平均の差のt検定」のことですので、これを選んで[OK]をクリックします。ウィンドウが現れたら、変数1の入力範囲に栽培法Aを、変数2に栽培法Bを設定します。データだけでなく栽培法まで選択した場合には[ラベル]をします。有意水準[a(A):]には最初から0.05(5%)が入っていますので、このままで[OK]を押せば、分析完了です。
- ③ 出力結果の見方ですが、tが-4.685となっていますので、この絶対値を限界値である[t境界値両側]の2.048と比較するか、p値である[P(T<=t)両側]の6.56814E-05(これは“ 6.56814×10^{-5} ”のことです)を有意水準(0.05)と比較すれば良いのです(もちろん有意差あります)。

The diagram illustrates the process of performing a t-test using the Analysis ToolPak in Excel, divided into three main steps:

- ① データ分析**: Shows the "Analysis ToolPak (A)" section of the "Data Analysis" dialog box. The "t 検定: 等分散を仮定した 2 標本による検定" option is selected.
- ② t 検定: 等分散を仮定した 2 標本による検定**: Shows the "t Test: Assuming Equal Variances" dialog box. It includes input ranges for data series 1 and 2, a hypothesis difference value, a label check box, and a significance level of 0.05.
- ③ t-検定: 等分散を仮定した 2 標本による検定**: Shows the output table from the t-test. Key values include:
 - t: -4.685
 - t 界界値 両側: 2.048
 - P(T<=t) 片側: 3.28407E-05
 - P(T<=t) 両側: 6.56814E-05
 Annotations with arrows explain the comparison: "検定統計量" points to the t-value (-4.685); "限界値" points to the critical t-value (2.048); and "比較" indicates the comparison between the t-value and the critical value.

7.4 対応のある2群の平均の差の検定

● 個体差とは

肥料を使う前と後での成長速度の差を調べたり（下の事例）、農家が技術講習会に参加する前と後の収量の差を調べたりした場合などは、対応のある（2群間で個体をペアとして関連付けた）データを観測することができます。

こうした対応のあるデータからは、**個体差や個人差**（individual differences）を考慮した検定ができます。個体差は、研究する側からすれば興味のない誤差の一部ですので、その大きさを仮説判定に反映させられるのは、とても大きなメリットといえるでしょう。

ところで、個体差とは何でしょうか？ 理解しやすいように、早速、ここから事例を使って説明しましょう。表7.1は、ある植物を5つ（ポットなどに）用意して、肥料を施す前と後の成長速度を観測した結果です。つまり、この肥料に植物の成長を促進させる効果があるかどうかを検証しようという実験です。また、それを線グラフに描いたのが図7.6です。

表 7.1 ある植物の施肥前後の成長速度比較 (cm/日)

個体番号	施肥前 (群2)	施肥後 (群1)	差 d_i (=群1-群2)
1	0.8	1.1	$d_1 = 0.3$
2	0.7	1.5	$d_2 = 0.8$
3	1.0	1.0	$d_3 = 0.0$
4	0.5	1.6	$d_4 = 1.1$
5	0.6	1.3	$d_5 = 0.7$
平均	$\bar{x}_2 = 0.72$	$\bar{x}_1 = 1.30$	$\bar{d} = 0.58$

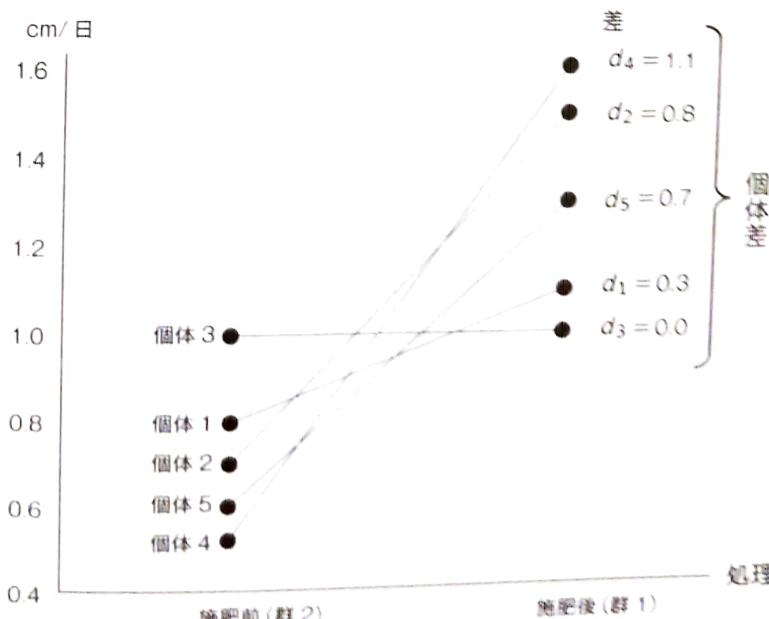


図 7.6 対応のある2群（表7.1の内容）

この図表を眺めて、対応のある場合が持つ“ある特徴”に気がつきましたでしょうか？

施肥前と後の成長速度の変化（2群の差） d_i を個体別に求めることができます。例えば、個体1は施肥前に比べて施肥後は0.3 cm ($d_1 = 0.3$)、個体2は0.8 cm ($d_2 = 0.8$)、……と、それぞれ成長速度が変化した（上がった）ことがわかります。なお、今回は施肥後から施肥前を引いていますが、逆でも問題ございません（統計量の符号が逆になるだけです）。そして、この情報は、同種の植物に同じ肥料を使っても、個体によって成長速度の変化に差がある、つまりバラついていることを意味しています。この個体別の変化（差） d_i のバラツキこそが“個体差”なのです。これは、個体が群間でペアになっていなければ取得できない情報ですね。

● 検定統計量（対応のある2群）

では、この個体差をどのようにして検定統計量に取り入れるのでしょうか。

もう一度、表7.1をご覧ください。対応のある場合には、せっかく個体ごとの2群の変化（差） d_i が求まるのですから、それを使って検定統計量を計算したいですね。でも個別の d_i のままでは誤差（いいかえれば実験結果の精度）を評価できないので、その平均 \bar{d} の分布を考えることにします（そうすれば標本が大きくなれば誤差が小さくなることなどを生かせます）。

さて、この \bar{d} も正規分布に従うので、 \bar{d} のままでも検定できるのですが、どのような対象や単位でも同じ分布表から限界値を読み取るためにには、やはり標準化した $z_{\bar{d}}$ や、準標準化した $t_{\bar{d}}$ で検定した方が便利です。

まず、 z から式を書きますと、次のようにになります。ただし、検定では帰無仮説（対応ありと同じ $H_0 : \mu_1 = \mu_2$ ）が正しい場合の統計量を考えますので、 \bar{d} の真の値 μ （母平均の差 $\mu_1 - \mu_2$ ）は0と仮定しますから、矢印より右の式が検定統計量となります。また、対応があるデータなので、2群ともバランス（同標本サイズ）な場合のみとなります。 n はペアとなっている個体数、つまり d_i の数です。

z 検定の統計量（対応あり2群、バランス）

$$z_{\bar{d}} = \frac{\bar{d} - \mu}{\sigma_{\bar{d}}} = \frac{\bar{d} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \xrightarrow[\text{検定統計量}]{\text{帰無仮説の下}} \frac{\bar{d}}{\sigma_{\bar{d}}} = \frac{\bar{d}}{\sqrt{\frac{\sigma^2}{n}}} \quad \begin{matrix} \leftarrow \text{個体差} \\ (d_i \text{の母分散}) \end{matrix}$$

この式を見ていただければわかるように、対応ありの検定は、差の平均 \bar{d} （対応なしでは「平均の差」でした）と、定数0との離れ具合を統計的に検証しているのです。

この式で注目して欲しいのは、個体差である「 d_i の母分散 σ^2 」が分母（の分子）に入っている点です。これなら、個体差である σ^2 が大きくなるほど z 値は小さくなるため、帰無仮説を棄却するのが難しくなります。つまり、誤差の一種である個体差の大きさを考慮して、念のために保守的な判断を下すのです。なお、対応なしの統計量でも分散を使っていましたが、あくまで2群共通の分散として求めたものなので、個体差を正確には把握できていません（2群がまぜこぜです）。

一応、 d_i の母分散を示しておきます（ μ は d_i の母平均なので普通は未知）。

$$\text{差 } d \text{ の母分散 (個体差)} \quad \sigma^2 = \frac{\sum(d_i - \mu)^2}{n}$$

t 検定のための統計量は、未知の母分散 σ^2 の代わりに不偏分散 $\hat{\sigma}^2$ が入った次式になります。普通は、こちらの式で検定統計量を求めた t 検定を実施することになるでしょう。

t 検定の統計量（対応あり 2 群、バランスのみ）

$$t_{\bar{d}} = \frac{\bar{d} - \mu}{\hat{\sigma}_{\bar{d}}} = \frac{\bar{d} - \mu}{\sqrt{\frac{\hat{\sigma}^2}{n}}} \xrightarrow[\text{検定統計量}]{\text{帰無仮説の下}} \frac{\bar{d}}{\hat{\sigma}_{\bar{d}}} = \frac{\bar{d}}{\sqrt{\frac{\hat{\sigma}^2}{n}}}$$

$$\text{差 } d \text{ の不偏分散 (個体差の不偏推定値)} \quad \hat{\sigma}^2 = \frac{\sum(d_i - \bar{d})^2}{n - 1}$$

この後の検定手順は対応のない場合と同じです。計算した検定統計量の絶対値と、 z や t の分布表から読み取った限界値とを比べて、検定統計量の絶対値の方が大きければ帰無仮説を棄却し、対立仮説を採択します。なお、**対応のある場合の自由度 ν** は、ペアとなっている個体数 n (d の数と同じ) から 1 を引いた $(n - 1)$ 値です。早速、先ほどの事例を例題として検定してみましょう。

例題

表7.1の事例「ある植物の施肥前後の成長速度比較」の結果を見ると、施肥（処理）後の成長速度の方が平均で0.58cm/日ほど上がっているようです。果たして本当に施肥によって成長速度に変化が生じたといえるでしょうか？この肥料が成長速度を変化させる効果を有しているかどうかを、5%の有意水準で検定してみましょう。

解：

このデータは個体ごとに施肥前・後の値が計測され、ペアになっていますので、対応のある2群の平均の差の検定を実施できます。ただし、母分散は未知なのでt検定を用います。また、施肥によって成長速度が下がることも否定はできないので両側検定とします。

手順①：仮説の設定

$$\left\{ \begin{array}{l} \text{帰無仮説 } H_0 : \mu_1 = \mu_2 \rightarrow \text{施肥前後で母集団の平均成長速度に差はない} \rightarrow \text{施肥効果なし} \\ \text{対立仮説 } H_1 : \mu_1 \neq \mu_2 \rightarrow \text{施肥前後で母集団の平均成長速度に差はある} \rightarrow \text{施肥効果あり} \end{array} \right.$$

手順②：検定統計量の計算

t検定の統計量の式に、差の平均 $\bar{d} = 0.58$ （わかりやすく施肥後から前を引いていますが、逆に引いて -0.58 となっても構いません）と、標本サイズ $n = 5$ （差 d の数）、 d の不偏分散 $\hat{\sigma}^2 = 0.187$ を代入すると、 t 値は“2.999”となります。

$$t_{\bar{d}} = \frac{\bar{d}}{\sqrt{\frac{\hat{\sigma}^2}{n}}} = \frac{0.580}{\sqrt{\frac{0.187}{5}}} = 2.999$$

手順③：確率の計算

限界値と比較する方法で仮説を判定するため、本手順はスキップします（Excel関数ならばT.DIST.2T(2.999,4)で両側 p 値が得られます）。

手順④：仮説の判定（と検出力）

あらかじめ決めておいた両側5%の有意水準 α に対応する t 値が限界値となります。付録IIの t 分布表から読み取ると、例題の自由度は $5 - 1 = 4$ となるので、 $\nu = 4$ の行と、片側確率 $p = 0.025$ の列がクロスする“2.776”が限界値となります。よって、例題の実験で観測された差の平均の t 値は、限界値よりも大きい（分布の外側に位置する）ため、帰無仮説が正しいにしては、かなり起きにくいことが起きてしまったといえます。ですから、むしろ帰無仮説自体が間違っていると起きてしまったといえます。

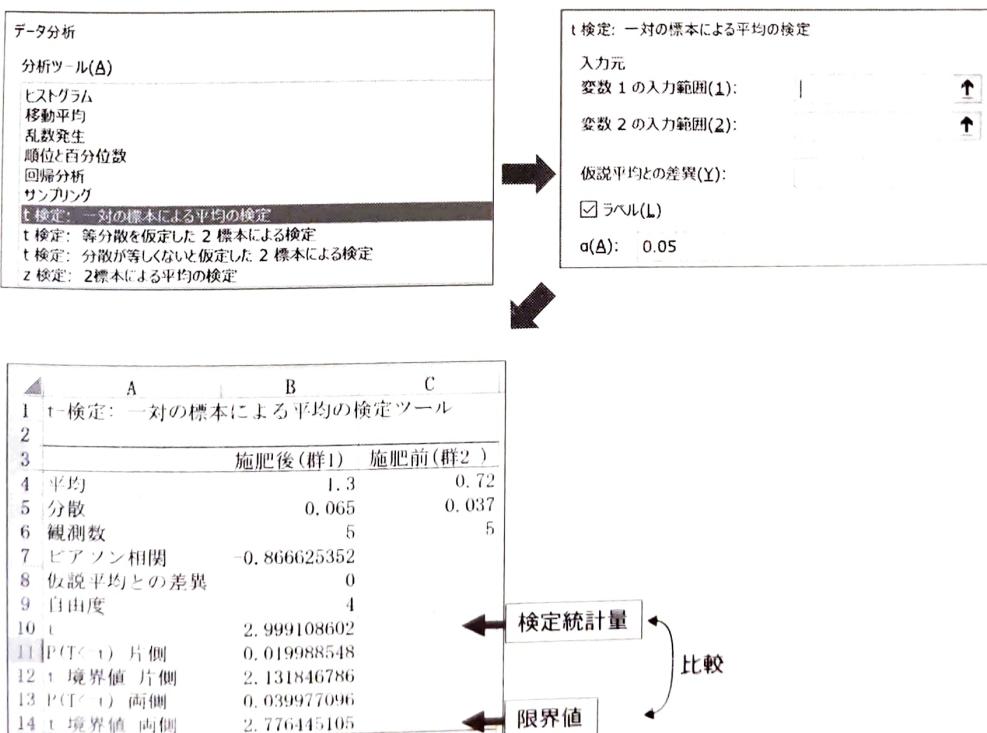
考えた方が合理的です。よって、有意水準5%という判定基準の下で、対立仮説を採択し、「施肥前と施肥後とでは、母集団の平均成長速度に統計的に有意な差がある」→「施肥の効果はあった」ことになります。

ここでも、G*powerの[Post hoc: Compute Achieved Power]を使って例題の検出力を計算してみましょう。対応のある2群の差の検定は[Means: Difference between two dependent means]です。本検定の効果量は、もっとも簡単な推定式である $t \times \sqrt{1/n}$ を使って計算すると1.34と推定できます。これに加え、両側有意水準 α に0.05、標本サイズに5をパラメータとして入力すると、検出力は0.62と推定されます。このように、本検定では検出力が弱かった（第二種の過誤を犯す確率 β が高かった）ことがわかりますが、その理由としては、標本サイズが小さいことが影響しているのでしょうか。

補足 分析ツールを使った検定

[t検定：一対の標本による平均の検定]が今回の「対応のある2群の平均の差のt検定」のことですので、これを選んで[OK]をクリックします。ウィンドウが現れたら、変数1の入力範囲に施肥後を、変数2に施肥前を設定します（適宜[ラベル]に \checkmark ）。有意水準 $[\alpha(A):]$ は0.05(5%)のままで[OK]を押せば、分析完了です。

出力された検定統計量 $t=2.9995$ を、[t境界値両側]の2.776と比較するか、[P(T<=t)両側]の0.03998を有意水準(0.05)と比較します（有意差あり）。



● ウェルチのt検定（等分散が仮定できない場合）

本章では2群の平均の差についての検定を解説してきましたが、本来、この検定を実施するためには、母集団において次の2つの性質が満たされている（と仮定できる）必要があります。

- ① 量的データで、正規分布に従っている（正規性）
- ② 両群の分散が等しい（等分散性）

①の正規性が満たされていなければ、違う形をしている分布同士を比較することになってしまいますし、統計量の確率計算もどのような分布で行えば良いのかわからなくなってしまいます。

②の等分散性については、対応のない2群の平均の差のz検定（アンバランス）で、帰無仮説が正しければ等分散であるとして2群共通の母分散を外に出して式を作り上げたことを思い出してください。母分散が未知の場合に用いるt検定（スチューデントのt検定と呼んで区別します）でも当然、等分散である必要があります（下に式を再掲しておきます）。なお、対応のある2群の場合には、同じ個体から観測しているので、分散は同じと考え（また式自体も、差の平均 \bar{d} の分布から導き出されているので）、等分散性が満たされているかどうかは気にしなくて良いでしょう。

スチューデントのt検定の統計量（対応なし2群、アンバランス） ← p.139の復習

$$t_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

等分散として2群の不偏分散を1つにまとめている

対処法ですが、①の正規性については、主に質的データの場合に問題になります。順序尺度や名義尺度で観測したデータは確率変数ではないからです。そのため、官能試験（味覚、嗅覚など）や意識調査のデータに、本章で学んだz検定やt検定を用いるのはお勧めできません。母集団に正規分布などの確率分布を仮定できなくても問題のない手法（ノンパラメトリック検定）がありますので、そちらを用いるようにしてください（第11章で詳しく解説します）。なお、「量的データで正規分布に従うと思われる対象だが、標本個別の分布が正

規分布とやや異なるようだ」という程度ならば神経質になる必要はありません。なぜならば、第4章で学んだ中心極限定理により、そこそこの標本サイズがあれば、標本平均やその差は正規分布に従うとみなせるからです。

さて、困るのは②の等分散性です。両群が異分散では検定統計量の式自体が導き出せないからです。そこで、Welchという統計学者が、両群の分散をそのまま残して検定する方法を考案しました（**ウェルチのt検定**；Welch's t-test）。とはいっても、ウェルチのt検定の統計量の計算は特別なものではなく、共通の不偏分散を求める前の、2群それぞれの不偏分散が残った状態の次式を用います。

ウェルチのt検定の統計量（対応なし2群、アンバランス）

$$t_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} \quad \text{両群の不偏分散を残している}$$

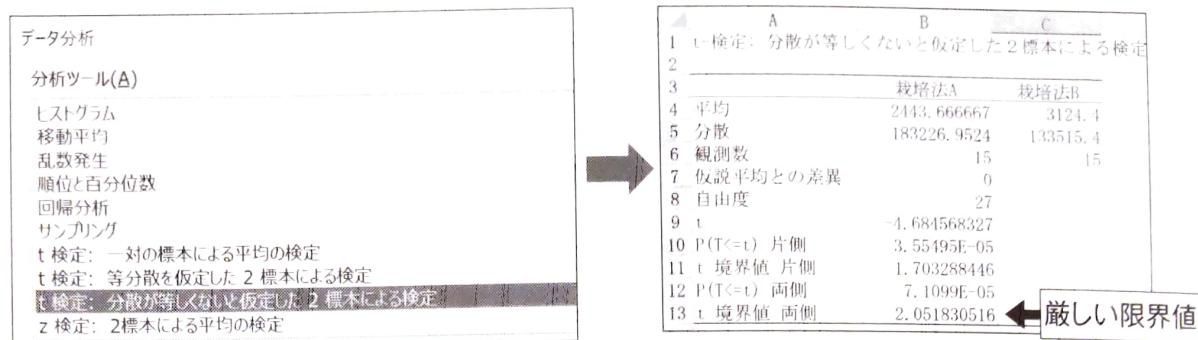
ウェルチの工夫はここからです。この統計量は、共通の不偏分散を用いていないので、自由度 $(n - 1)$ のt分布には従わないのですが、次のような複雑な式で計算した自由度のt分布から限界値を読み取ることを発見したのです。

$$\text{ウェルチのt検定の自由度 } \nu \approx \frac{\left(\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2} \right)^2}{\frac{\hat{\sigma}_1^4}{n_1^2(n_1 - 1)} + \frac{\hat{\sigma}_2^4}{n_2^2(n_2 - 1)}}$$

つまり、自由度さえ、この式で頑張って求めれば、後は普通にt分布表から限界値を読み取って検定を実施できるのです。

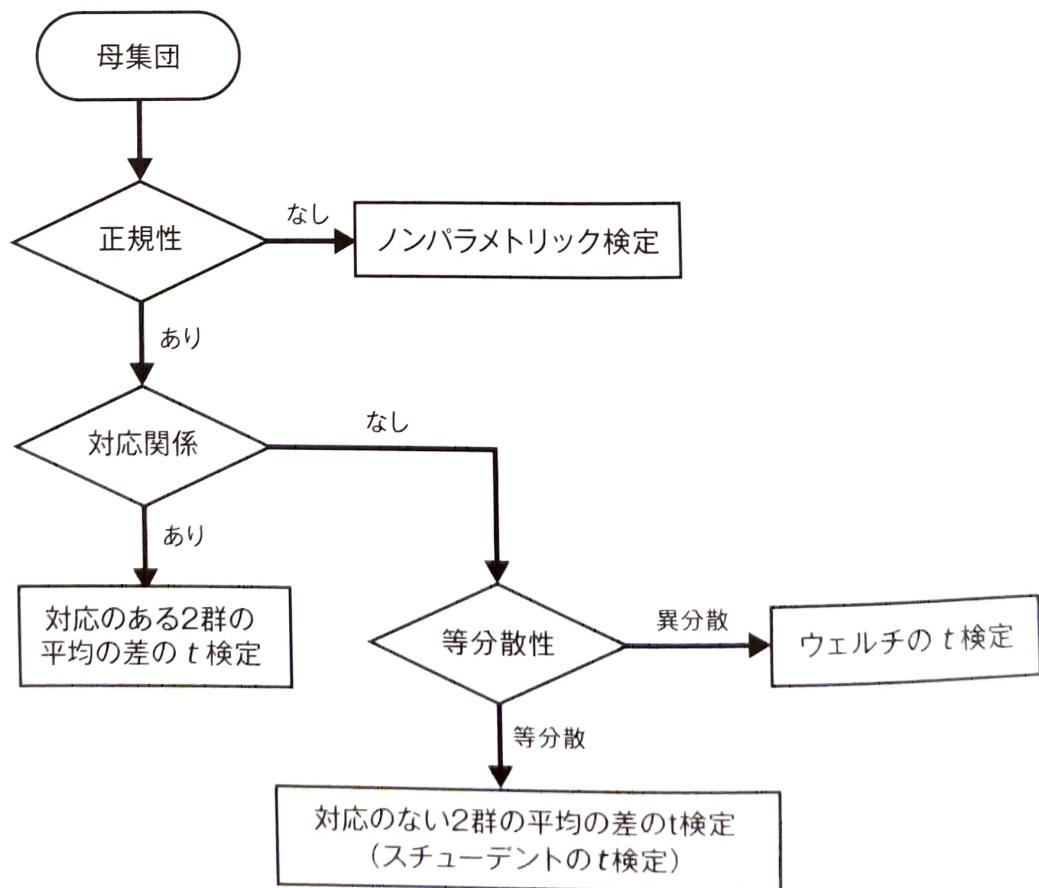
実際には、面倒な計算なのでソフトウェアを使うことになるでしょう。幸いなことに、Excelの分析ツールにもウェルチのt検定が搭載されています。[t検定：分散が等しくないと仮定した2標本による検定]です。スチューデントのt検定を実施した例題（キュウリの収量）に、分析ツールを使ってウェルチのt検定を実施してみたのが図7.7です。スチューデントのt検定と比較してみると、28だった自由度が27（分析ツールでは四捨五入して整数にしてしまってますが、本来は小数点まであります）へと小さくなつたことで、限界値（境界）が2.048から2.052へと厳しくなっていることがわかります（結局、この結果では「有意差あり」という結果は変わりませんでした）。

事例では「有意差あり」という結果は変わりませんでした。
このように、対応のない場合は、2群の母分散が同じか異なるかで使用すべ

図 7.7 分析ツールによるウェルチの t 検定

き検定法が違うのですが、どのようなときに異分散と判定してウェルチの t 検定を使うのでしょうか。3つほど (A~C) 紹介しましょう。

まず、簡単な使い分け方として、A: 2標本のサイズや分散が大きく異なる (1.5~2倍以上) 場合にはウェルチの t 検定を、それ以内ならばスチューデントの t 検定を実施するという目安が示されています。また、B: どうせ母分散はわからないのだから、最初からどのようなときでもウェルチの t 検定を実施するという考え方もあります (等分散の場合にウェルチの t 検定を実施してしまっても、それほど検出力は下がらないといわれています)。

図 7.8 t 検定の使い分け方

しかし、伝統的には、C：予備的に2標本の母分散が同じか異なるかを検定して、その結果をみてから、どちらのt検定を使うかを判断するという方法がとられてきました。次章の予習にもなるので、本章最後の次節ではCの使い分け方に使う検定（等分散の検定）を解説しましょう。

以上、ここまで学んだ2群の平均の差の検定の使い分け方をフローチャートに整理しました（図7.8）。

7.5 等分散の検定

仮説の設定

2標本それぞれの抽出元の母集団の分散が等しい（差はない）か異なる（差はある）かを検証する検定を**等分散の検定**（test for equal variance）と呼びます。まずは、仮説の考え方から説明しましょう。

等分散の検定では、帰無仮説が等分散 ($H_0 : \sigma_1^2 = \sigma_2^2$) となり、対立仮説が異分散 ($H_1 : \sigma_1^2 \neq \sigma_2^2$) となります。もう少し正確にいえば、図7.9のように、帰無仮説は「2つの標本の抽出元は同じ正規（分布に従った）母集団」と考え、対立仮説は「2つの標本の抽出元はそれぞれ異なる正規母集団」と考えます。

よって、この検定で帰無仮説が棄却されたときにはウェルチのt検定を使い、帰無仮説が棄却されなかったときには（等分散と断言はできませんが、その可能性が残されたので）スチューデントのt検定を使うようにします。

本章では、2群の平均の差の検定を扱ってきたため、等分散の検定をt検定の前に実施する予備的な検定として説明していますが、2群の真の分散に差があることを検証するのですから、他にも様々な用途が考えられるでしょう。例

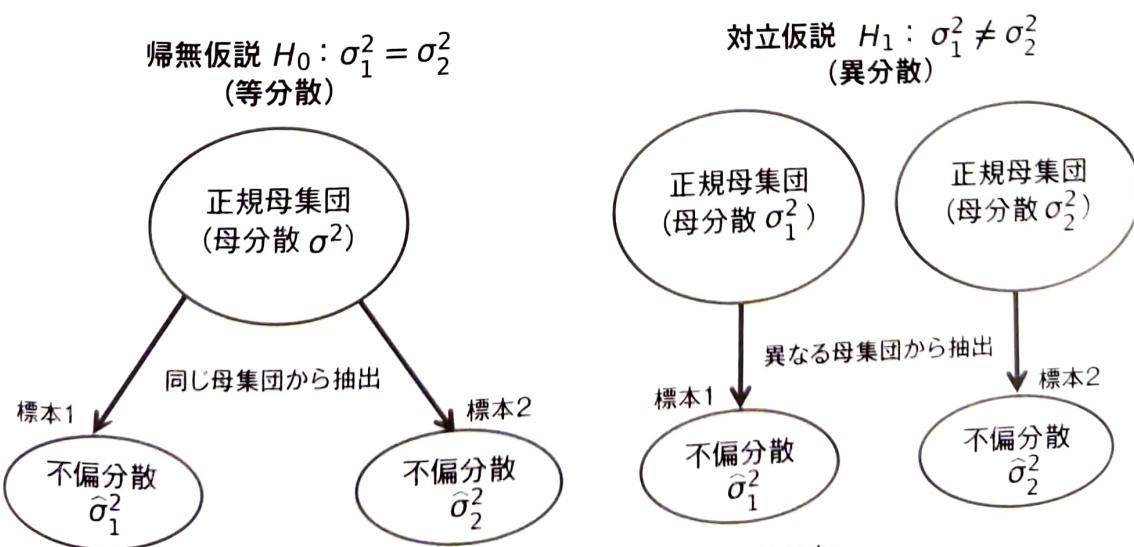


図7.9 「等分散の検定」の仮説の考え方

えば、教員なら自分の担任しているクラスにおいて、男子と女子で成績のバラツキに差があるかどうかを確認したい場合に使えるでしょう。工場長なら、古いラインと新しいラインで、製品の品質のバラツキに差があるかどうかを確認したい場合があるでしょう。

● 検定統計量 F

さて、この等分散の検定ですが、内容は F 分布に従う統計量“ F ”を使った **F 検定 (F-test)** です。第5章後半の復習になりますが、 F とは正規分布に従う2つの母集団から抽出した χ^2/ν の比でした。それを、それぞれの不偏分散 $\hat{\sigma}^2$ と母分散 σ^2 を使った式に変形すると、次のようになりました (103ページ)。

$$F \text{ 統計量 } F_{(\nu_1, \nu_2)} = \frac{\chi^2_{(\nu_1)} / \nu_1}{\chi^2_{(\nu_2)} / \nu_2} = \frac{\frac{\nu_1 \hat{\sigma}_1^2}{\sigma_1^2} / \nu_1}{\frac{\nu_2 \hat{\sigma}_2^2}{\sigma_2^2} / \nu_2} = \frac{\hat{\sigma}_1^2 / \sigma_1^2}{\hat{\sigma}_2^2 / \sigma_2^2}$$

検定統計量は、帰無仮説 ($H_0 : \sigma_1^2 = \sigma_2^2$) が正しい下での分布ですので、分子の母分散 σ_1^2 と分母の母分散 σ_2^2 を同じ σ^2 と考え、最終的には、次式右項のような2群の「不偏分散の比」の形になります。なお、第1自由度 ν_1 も、第2自由度 ν_2 も $n - 1$ です。

$$\text{等分散の検定 (F 検定) の統計量 } F_{(\nu_1, \nu_2)} = \frac{\hat{\sigma}_1^2 / \cancel{\sigma}^2}{\hat{\sigma}_2^2 / \cancel{\sigma}^2} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$$

● 仮説の検定

等分散の検定 (F 検定) の原理はいたって単純です。

抽出元の母集団の分散が同じならば、そこから抽出された2つの標本の分散も似た値になるでしょう。似た値同士を割ると (つまり比は) 1 前後になりますね。 F も分散 (の不偏推定量) の比なのですから、帰無仮説の下の F 分布は、図7.10のようになります (ただし、形状は自由度によって変わります)。

ですから、等分散の検定では、検定統計量である F が限界値よりも 1 に近い値を取ったときに帰無仮説を受容し、1 から離れた大きな値や 0 に近い値 (分散の比ですから負にはなりません) を取ったときに棄却します。

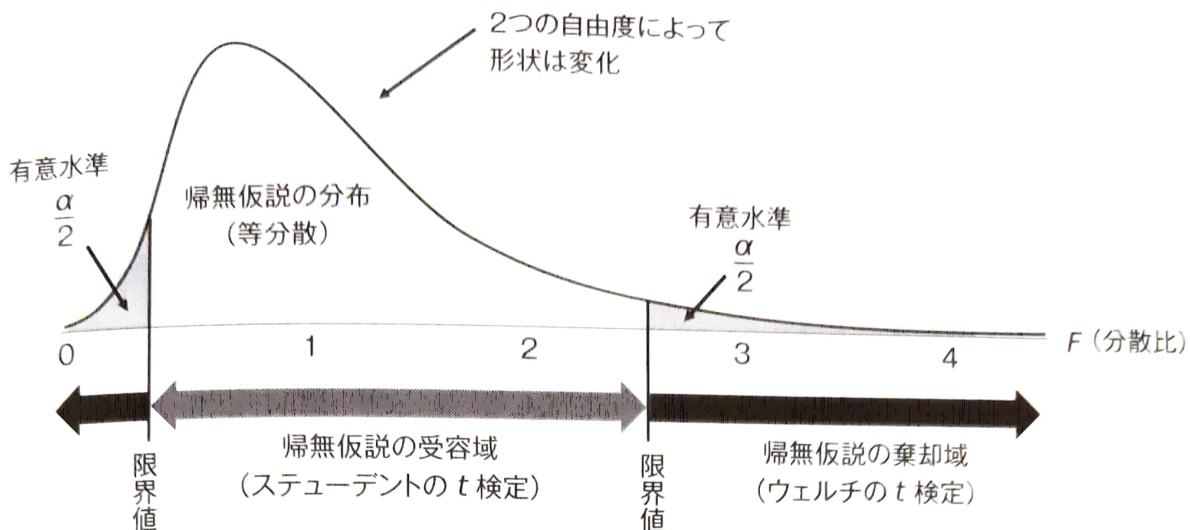


図 7.10 等分散の検定 (F検定) の原理

例題

本章の最初の例題（キュウリの収量）では、栽培法AとBの収量の等分散性について確認せずに「対応のない2群の平均の差のt検定」（スチューデントのt検定）を実施していました。本来、t検定の前に行うべきですが、いまから等分散の検定（両側5%）をこのデータで実施してみましょう。

解：

手順①：仮説の設定

$$\begin{cases} \text{帰無仮説 } H_0 : \sigma_A^2 = \sigma_B^2 \rightarrow \text{栽培法AとBの収量の母分散は等しい} \\ \text{対立仮説 } H_1 : \sigma_A^2 \neq \sigma_B^2 \rightarrow \text{栽培法AとBの収量の母分散は異なる} \end{cases}$$

手順②：検定統計量の計算

2つの不偏分散の比が検定統計量となるF値なので、栽培法Aの不偏分散183227.0を栽培法Bの不偏分散133515.4で割ると1.372となります。よって、検定統計量 $F=1.372$ となります。

$$F_{(14,14)} = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_B^2} = \frac{183227.0}{133515.4} = 1.372$$

大きい値を分子にすると便利

なお、等分散の検定も、両側で検定するのが基本なので（※157ページの補足参考）、どちらをどちらで割っても良いのですが、F分布表は右側（上側）の確率しか掲載していないため、分子に大きな方の不偏分散の値を持って来るようになると簡単です（1よりも大きな統計量にしておく）。

手順③：確率の計算

ソフトウェアを使えば、帰無仮説の F 分布における本検定統計量の値 (1.372) より外側の確率 (p 値) を計算することができます。Excel 関数ならば、 $F.DIST.RT(1.372, 14, 14)$ で上側の確率 (0.281) が求まりますので、両側 p 値はその2倍の 0.562 となります。よって、 $p > 0.05$ で帰無仮説は棄却できないことがわかります。ただし、限界値と比較する方法ならば、次の手順で F 分布表を使って仮説を判定できます。

手順④：仮説の判定（と検出力）

あらかじめ決めておいた両側 5% の有意水準 α に対応する F 値が限界値となります。付録Vの上側確率 2.5% の F 分布表から読み取ると、例題の自由度は第1 (分子)、第2 (分母) ともに 15 - 1 で 14 なので、本来は $\nu_1 = 14$ の行と $\nu_2 = 14$ の列がクロスするところが限界値となります。しかし、 F 分布表の列は粗いため 14 の列がありません。仕方がないので、14 の行と 15 の列がクロスする 2.95 よりもやや大きい 3.0 ぐらいとしておきましょう。

結果、 F 値 (1.372) は、有意水準 5% (両側) に対応する限界値 (3.0) よりも小さいため、帰無仮説は棄却できませんでした。

つまり、栽培法 A と B のキュウリ収量の分散の間には、(今回の実験では) 差は統計的に認められなかったため、等分散性が満たされている可能性は残され、最初の例題で使用したスチューデントの t 検定で問題はなかったろう、ということになります。

さて、本検定も G*power の [Post hoc: Compute Achieved Power] を使って検出力を計算してみましょう。等分散の検定は [F tests] の [Variance: Test of equality (two sample case)] です。本検定の効果量は、 F 値そのもの (Ratio var1/var0) なので 1.372 を入力します (Determine で計算も可)。これに加え、両側有意水準 α に 0.05、標本サイズにどちらも 15 をパラメータとして入力すると、検出力は 0.086 と推定されます。このように、検出力が極めて弱かったのは、標本が 15 ずつしなかったこともありますが、効果量である F が 1 に近かった (小さかった) 影響が大きいでしょう。検出力が弱いということは第二種の過誤を犯す確率 β が大きいということですから、異分散であることを見逃している過誤確率が高いことになり、ウェルチの t 検定を使うべきだった可能性も高いことになります。このようなこともあり、近年では等分散性は確認せず、最初からウェルチの t 検定が用いられるようになってきています。

補足

分析ツールを使った検定

[F検定：2標本を使った分散の検定] が等分散の検定のことです。出力結果を見てみましょう。検定統計量Fは「観測された分散比」なので1.372です。限界値であるF境界値は、片側（上側）の値（2.484）しか示されていません。従いまして片側検定ではありますが、検定統計量F（1.372）<上側限界値（2.484）なので、帰無仮説は受容されました。p値（P）は0.281となっていますが、こちらも片側のみの確率なので、論文に記載するときには2倍して両側の確率とした0.562としてください（これを有意水準 α の0.05と比較すれば両側検定と同じことになります）。

さて、なぜ分析ツールではp値や限界値（F境界値）が片側しか出力されないのでしょうか。実は、本書も改訂前は片側検定で等分散の検定を解説していました。というのは第5章末で導き出した $t^2 = F$ という関係があるため、F検定の片側検定はt検定の両側検定にあたるという考え方があったからです。しかし、これは分散分析（次章）という別の目的のF検定にのみ当てはまることがあるため、近年は、等分散の検定はやはり両側で実施すべきだという流れになってきました。従いまして、SPSSやJMPなど、有名なソフトウェアでも両側のp値が出力されるようになっています。

F 検定: 2 標本を使った分散の検定

入力元
変数 1 の入力範囲(1):
変数 2 の入力範囲(2):
 ラベル(L)
 $\alpha(\Delta): 0.05$

	A	B	C
1	F-検定: 2 標本を使った分散の検定		
2			
3		栽培法A	栽培法B
4	平均	2443.666667	3124.4
5	分散	183226.9524	133515.4
6	観測数	15	15
7	自由度	14	14
8	観測された分散比	1.372328229	← 検定統計量
9	P(F < f) 片側	0.280796311	← p 値(2倍にする)
10	F 界限値 片側	2.483725741	← 上側の限界値

章末問題

問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（サナギでいる）日数を観測した仮想データです。この実験結果から、気温の差がサナギでいる日数に影響を与えたといえるだろうか。両側5%の有意水準で「対応のない2群の平均の差のt検定」（スチューデントのt検定）を実施しなさい。

あるガのマユの形成から羽化までの気温別日数

個体番号	15°C	20°C
1	37	28
2	35	26
3	32	25
4	36	22
5	34	23
6	33	24
7	31	27

問2

- (1) Excelの関数や分析ツールを使って、問1のp値（両側）を計算しなさい。
- (2) G*powerを使って、問1の「効果量」と「検出力」を計算しなさい。
- (3) 本来、スチューデントのt検定は2群が等分散でなければならない。事後ではあるが、問1でスチューデントのt検定を実施したことが妥当だったかどうかを「等分散の検定」で検証しなさい（こちらの事後分析は不要）。