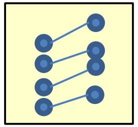
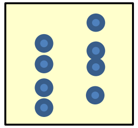
今回とりあげる検定の種類
（標本の正規性仮説）

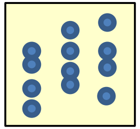
・関連ある2群の検定（1標本t検定）



・独立した2群の検定（2標本t検定）



・多群間の検定（1条件：一元配置分散分析）



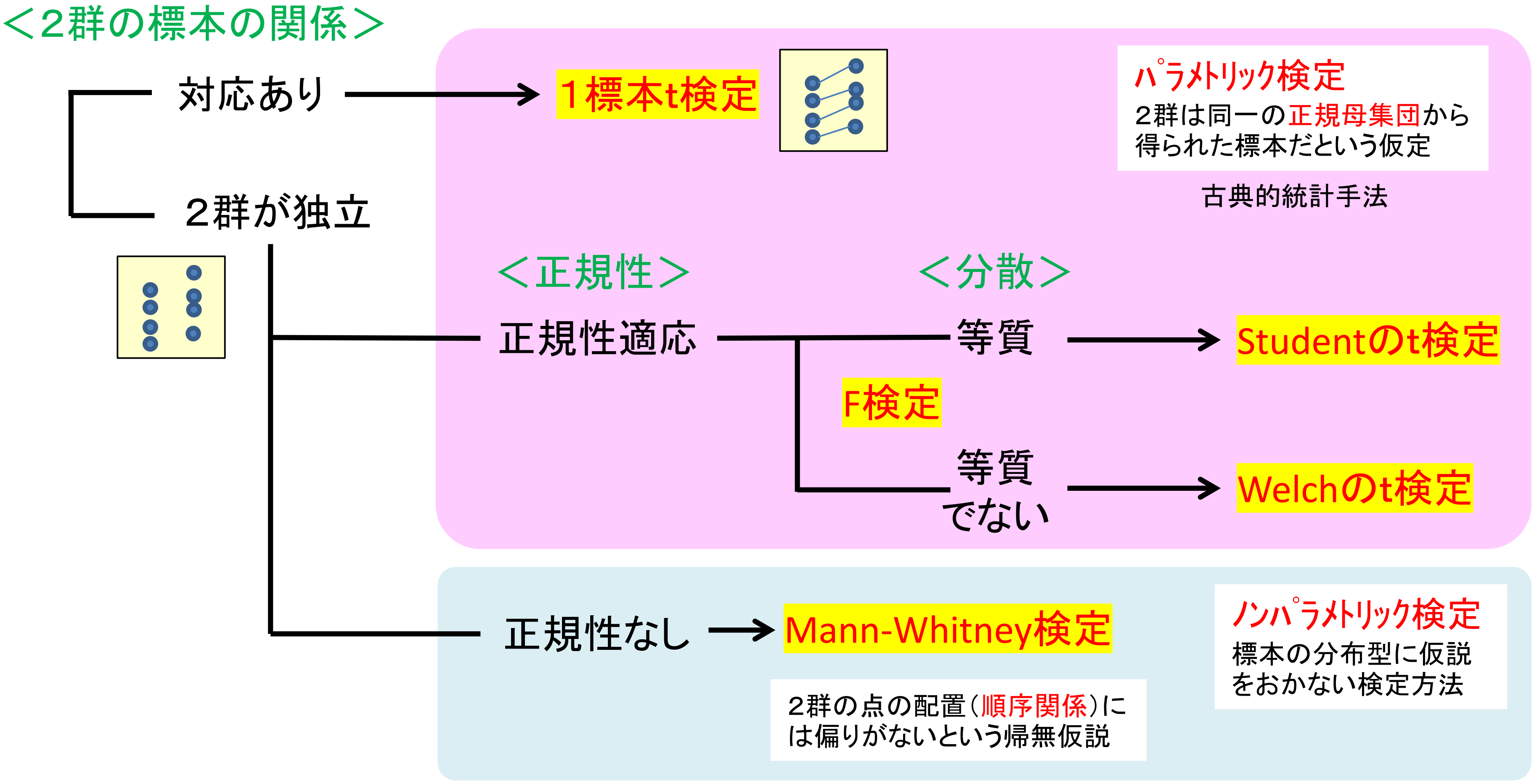
・多群間の検定（2条件：二元配置分散分析）

検定の種類(一例)

パラメトリック検定
データの正規性を仮定

検定の目的	検定対象群の数	検定手法	概要
母平均に差があるかどうかを確認する	2群	Studentのt検定	母平均に差がないという帰無仮説。分散は等質。独立な2群。
		1標本検定	母平均に差がないという帰無仮説。対応のある2群。
		Welchのt検定	母平均に差がないという帰無仮説。分散は等質でなくてよい。独立な2群
		Mann-Whitney検定	**ノンパラメトリック検定** 2群の点の配置（順序関係）には偏りがないという帰無仮説H_0をおく。 サンプル数が少なる得られたデータに正規性を仮定できないときに用いる。
	3群以上	一元配置分散分析	母平均に差がないという帰無仮説。 1つの条件が異なる場合。
		二元配置分散分析	母平均に差がないという帰無仮説。 2つの条件が異なる場合。
母分散に差があるかどうかを確認する	2群	F検定(分散の等質性の検定)	母分散が等しいという帰無仮説。
	3群以上	バートレット検定	母分散が等しいという帰無仮説。
正規性があるかどうかを確認する	2群	コルモゴロフ-スミルノフ検定	正規性があるという帰無仮説。
	2群	シャピロ-ウィルク検定	正規性があるという帰無仮説。

２群の母平均に差があるかの検定方法決定フロー



検定の方法

（1）仮説の設定　（H_0：帰無仮説＝差がない）

（2）検定統計量を求める（p値を求める）

（3）有意水準を設定し棄却域を求める。（一般に有意水準は5％か1％）

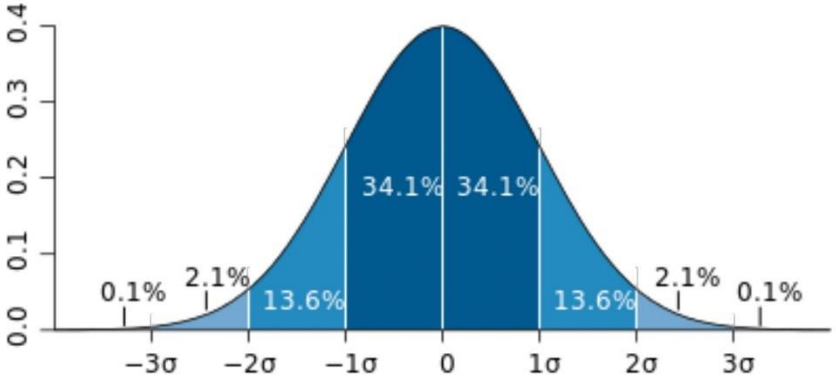
（4）判定（有意水準XX%で差がある・ない）

仮説検定で出てくる主な用語

- 帰無仮説とは、最初に立てる仮説。
- 対立仮説とは、帰無仮説が棄却されたときに採択する仮説。
- 検定統計量とは、仮説検定を行う際に求める値。求めた値のことを検定統計量の実現値という。実現値を算出した後に、この値はよくおこるものか、珍しいものかを判断する。
- 有意水準とは、帰無仮説を棄却する際の基準値。記号にはαを用いられることが多い。
- 棄却域とは、帰無仮説が棄却される検定統計量の範囲。
- p値とは、検定統計量の実現値から求められる確率。

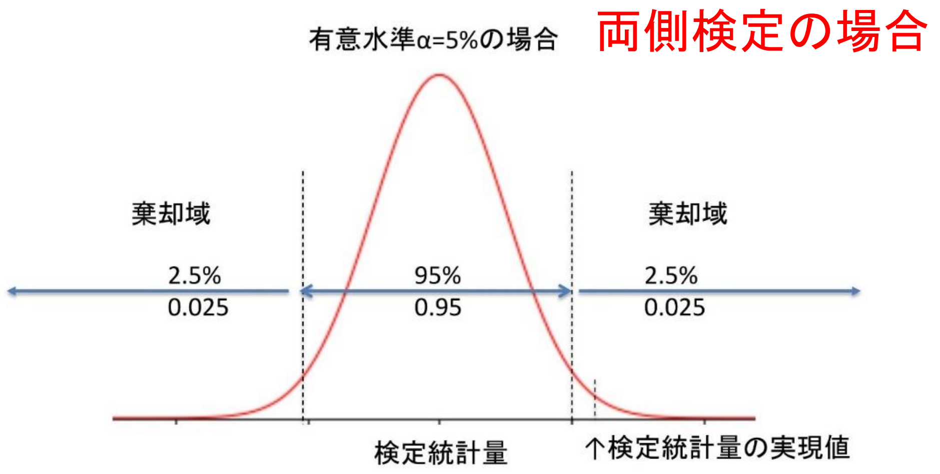
正規分布の性質

- 面積を全て積分すると1になる。(確率分布の性質)
- ある区間の面積の積分値は、その区間のデータが発生する確率として扱える。
- 標準偏差をσとしたとき、±σの範囲には約68%のデータがおさまる。±2σの範囲には約95%のデータがおさまる。
- 発生する確率が低い値を観測すると異常な値であると判定したい時に使える。



「正規分布」『フリー百科事典　ウィキペディア日本語版』。2016年8月19日 (金) 07:11　UTC　URL: http://ja.wikipedia.org

一般的な検定の手順(棄却域で考える場合)

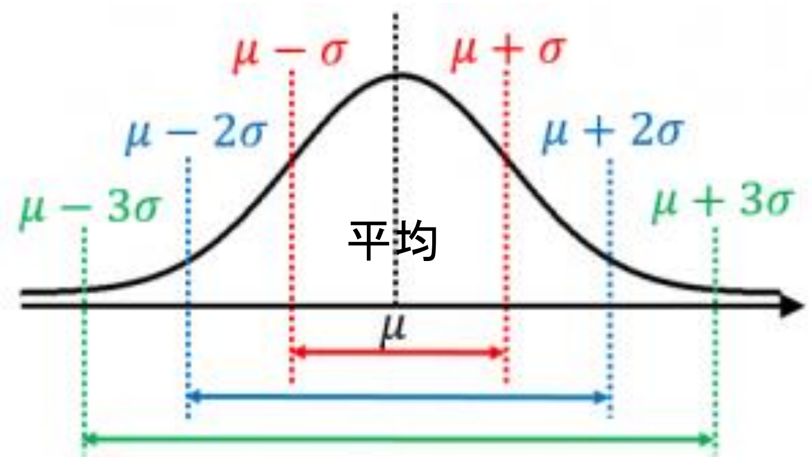


＜両側検定か片側検定のいずれを選ぶべきか＞

片側検定は、あらかじめ変化（差）の向きが理論的に片側にだけ起こると想定される場合（例：降圧剤の効果を調べる実験で、投与後の血圧の上昇を想定しなくてよい場合）に、逆に、処理効果がどちら向きの変化をもたらすか予測できない場合に用いられる。

通常は、変化の向きを予め予測できないとして、両側検定を行う。

正規分布と1σ、2σ、3σ



1σ 区間とは、μ−σ から μ+σ までの区間です。
2σ 区間とは、μ−2σ から μ+2σ までの区間です。
3σ 区間とは、μ−3σ から μ+3σ までの区間です。

1σ 区間におさまる確率→ 約 68.27%
2σ区間におさまる確率→ 約 95.45%
3σ 区間におさまる確率→ 約 99.73%

t検定はどこまで使えるのか？

t検定を実施するには、サンプルの背景にある母集団は、正規分布に従っていることを前提にしています。このため、t検定を行う前に、まずは正規確率プロットなどを使って、正規性の検定を行います。

正規分布に従っていれば、t検定などのパラメトリック分析を行い、正規分布に従っていなければ、1サンプル-ウィルコクソンなどのノンパラメトリック検定を行います。

しかし、厳密にこの手順を守らなければいけないのでしょうか。実は、そうでもありません。サンプルの状況によっては、守らなくてもよい場合があるのです。

例えば、サンプル数が少ない場合、正規性の検定を行っても、サンプル数が少ないために、正規分布に従っているという結果になることが多くあります。このため、極端に分布の形が歪んでいる場合を除いて、ほとんどの場合、サンプル数が少なければt検定が使えるという事実があります。

一方、サンプル数が多い場合、正規性の検定を行うと、わずかなズレでも正規分布に従わないという結果が出ることがあります。本来であれば、サンプル数が多いと情報量が増えるというメリットがあるわけですが、検定に厳密さを求めると、t検定が使えなくなるという結果になってしまいます。しかし、サンプル数が多ければ（以下、2サンプルの場合で説明します）、

＜非正規分布ではあるけれど、等分散となっている場合＞
中心極限定理により、平均値の差は近似的に正規分布に従うことが知られています。また、t値を求める際に利用する標準偏差が正確に計算できるため、t値はほとんど正規分布で近似できます。そのため、t検定が使えます。

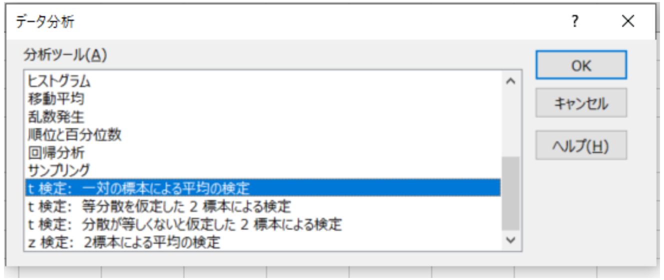
＜非正規分布で、かつ等分散となっていない場合＞
極端なケースではt検定を使えません。ただ、サンプル数が非常に多くあり、その標本分散から母分散をほぼ正確に推定できるのであれば、t検定が使えます。

ということで、t検定は多くの場面で使えることになります。これをt検定の頑健性（ロバストネス）といいます。

結論：t検定でたいていの場合は通用する

https://www4.kke.co.jp/minitab/support/newsletter/mt200912.html

＜エクセルの分析ツールのt検定＞

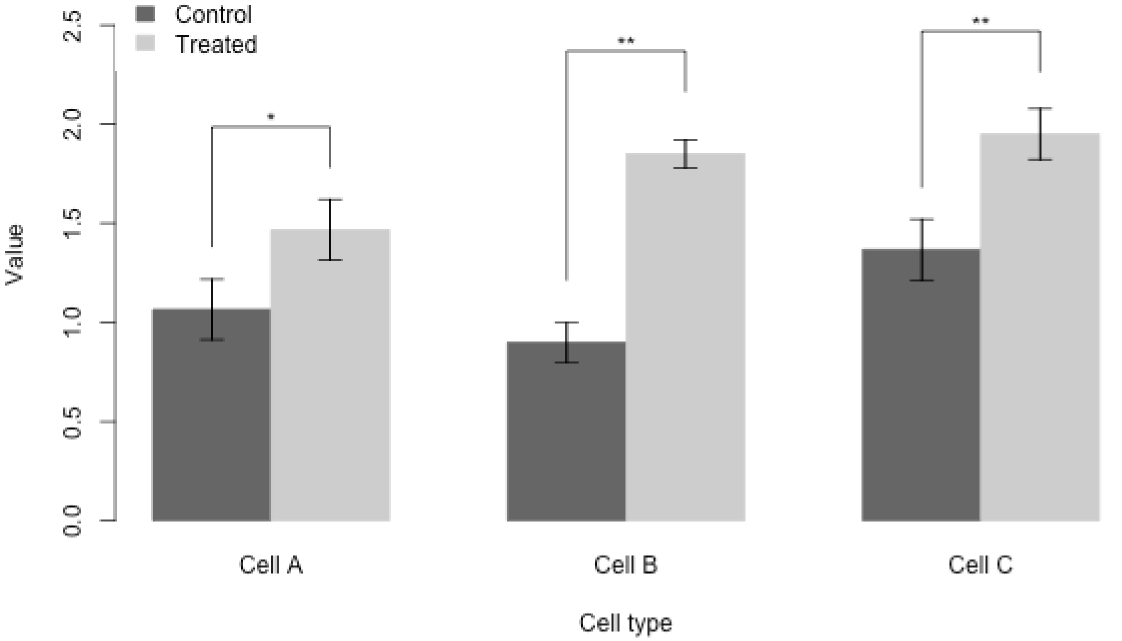


- t検定：一対の標本による平均の検定
- t検定：等分散を仮定した2標本による検定　→　Student(スチューデント)のt検定
- t検定：分散が等しくないと仮定した2標本による検定　→　Welch(ウェルチ)のt検定

＜Student(スチューデント)のt検定＞
https://data-science.gr.jp/theory/tst_student_t_test.html
スチューデントのt検定 (Student t-test) とはパラメトリック検定のひとつである．検定名にあるスチューデントとは，開発者であるゴセット (William Sealy Gosset) が論文執筆時に用いていたペンネーム Student に由来する．スチューデントのt検定に加えて，ウェルチのt検定および対応のあるt検定を含めた種々のt検定はデータXおよびデータYの2つのデータ間の平均値に差があるかどうかを検定する方法であるが，スチューデントのt検定は特に，2つのデータ間に対応がなく，かつ2つのデータの分散に等分散性が仮定できるときに用いる方法である．2つのデータ間の比較を行う場合にはいくつか注意を払うべき点がある．それは以下の3点である．
- データ間の対応の有無
- データの正規性
- データの等分散性

＜ Welch(ウェルチ)のt検定＞
https://data-science.gr.jp/theory/tst_welch_t_test.html
ウェルチのt検定 (Welch t-test) とはスチューデントのt検定と同じく，2つのデータ間の平均値の差に関するパラメトリック検定である．スチューデントのt検定が2つのデータの母分散が等しいと仮定できるときに用いる方法であるのに対し，ウェルチのt検定は2つのデータの母分散が等しいとは限らないときに用いる検定法である．検定法の名前に冠されているウェルチとは，本検定法の開発者である20世紀のイギリスの統計学者 Bernard Lewis Welch に由来する．

統計的有意差の表示例



* 95%確率で有意差あり
(有意水準5%)
** 99%の確率で有意差あり
(有意水準1%)