

5 章 2 群間の量的データの検定

5.1 対応のない検定手順

例えば、男女の成績を比較しようとして試験を実施した場合、男性の集団（群）と女性の集団（群）との比較になりますから、2つの集団に同一人物は1人もいません。しかしその試験で英語と国語の平均点を比較する場合、英語と国語を受験した集団には必ず同じ人がいます。前者のような場合を対応のないデータ、後者の場合を対応のあるデータと呼びます。対応のあるデータについては特別の処理ができるので、章を変えて説明しますが、対応を考えず単純に2つの集団として処理することも可能です。ここでは対応のない2つの集団間の平均値と中央値の比較について見てみましょう。

対応のないデータについての検定手順は以下の図 5.1.1 の四角形の中で表されます。

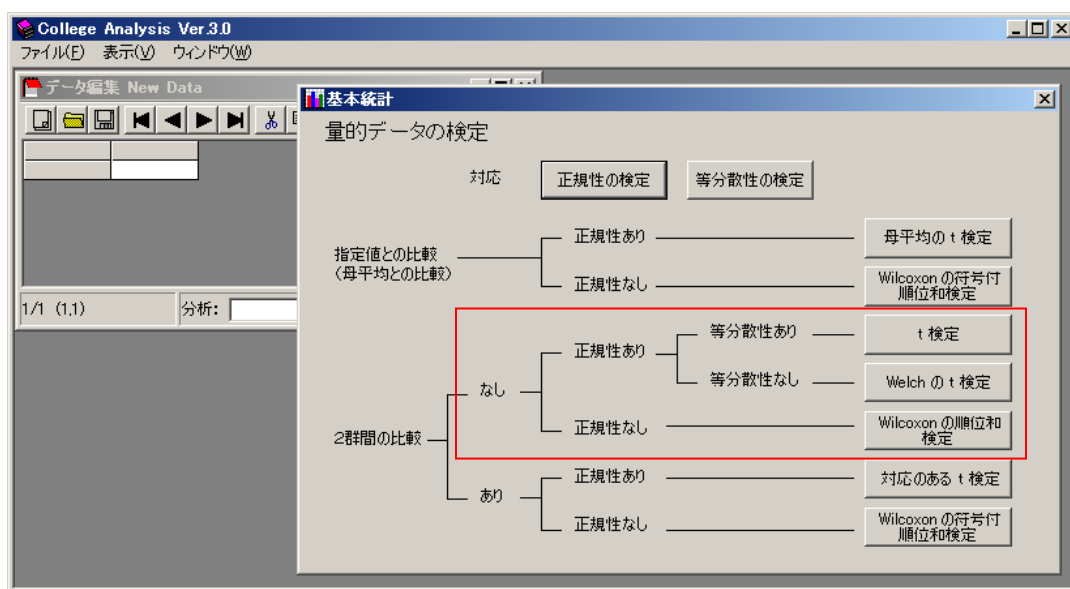


図 5.1.1 対応のない量的データの比較検定手順

この手順はまず2群のデータに正規性があるかどうか調べます。もし2群とも正規性があれば、次の等分散性の検定に進みます。正規性がなければ、検定方法はWilcoxonの順位和検定となります。等分散性の検定ではF検定と呼ばれる検定方法を利用します。これでもし等分散性ありとなれば検定方法はt検定、等分散性なしとなれば検定方法はWelchのt検定になります。それでは各検定を具体的に見て行きましょう。

5.2 対応のない2群間の分散の検定（正規性あり）

データの正規性の判定方法については前章で説明しましたので、ここではこの部分の説明は省略します。データに正規性が認められた場合、次の検定は等分散性の検定です。以下の例を見て下さい。

例

A機を導入した会社 18 社（1 群）と B機を導入した会社 15 社（2 群）について、機械 10 台当り 1 年間の故障発生件数を調べ、不偏分散を求めたら以下の結果を得た。

1 群 10.68 2 群 3.17

分布は正規分布であると仮定して、分散に差があるといえるか有意水準 5%で判定せよ。

まず、図 5.1.1 の量的データ検定メニューの「等分散性の検定」ボタンをクリックして、以下の F 検定メニューを表示します。データはもうすでに集計していますから、ラジオボタンは「集計から」にし、データ数と不偏分散の値を入力します。

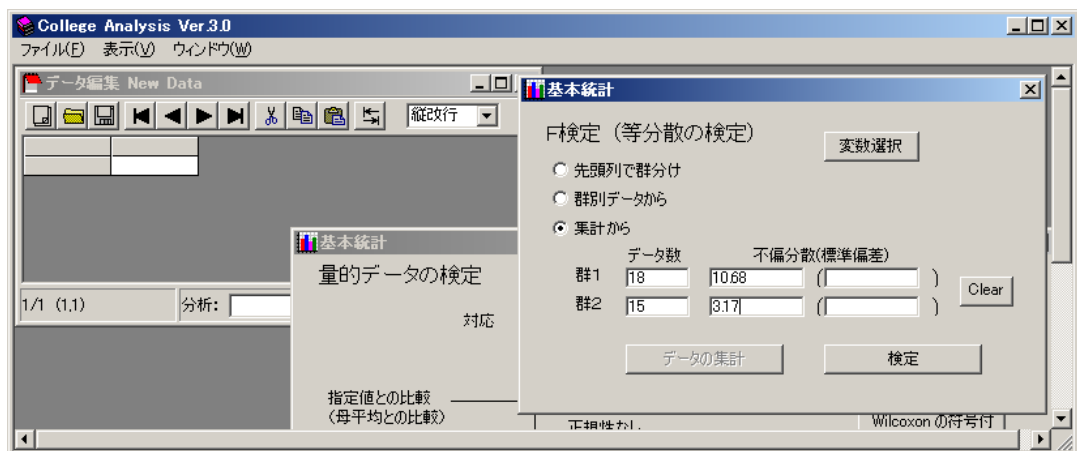


図 5.2.1 等分散性の検定メニュー

その後、「検定」ボタンをクリックすると以下のような結果出力画面が表示されます。

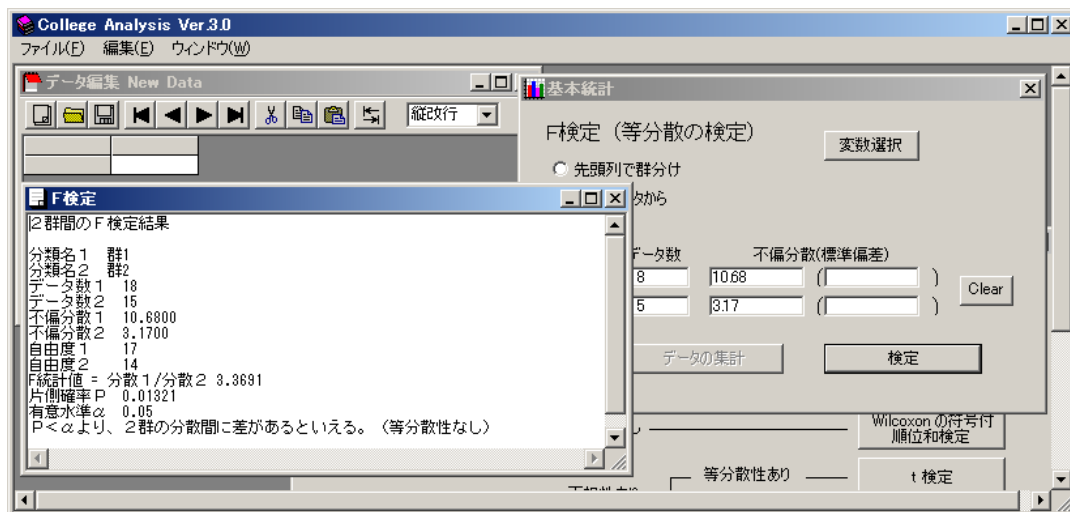


図 5.2.2 等分散性の検定結果画面

これを見ると有意水準 5%で「2 群の分散間に差があるといえる」という結論になっています。言葉を変えて「等分散性なし」というような表現も加えてあります。片側確率が $p > 0.05$ の場合は、「2 群の分散間に差があるといえない」という結論と「等分散性ありとみなす」という表現が併記されます。通常異分散でない場合、等分散とは結論づけられませんが、このような検定の場合は等分散とみなして次の処理に移ることが多いようです。

この検定に使った理論式を書いておきましょう。

理論 F 検定

母分散に差がないとすると

$$F = \frac{u_1^2}{u_2^2} \sim F_{n_1-1, n_2-1} \text{ 分布}$$

5.3 対応のない 2 群間の平均値の検定（正規性あり・等分散）

次の検定は、正規性があつて等分散性もある場合です。検定は t 検定を利用します。以下の例を見て下さい。

例

ある地域の同性・同年齢の児童について、ある要因の有無による 2 つの集団の体重を調べたところ以下のデータを得た。2 つの集団の平均値に差はあるといえるか。正

規性、等分散性を仮定して、有意水準 5%で判定せよ。

	データ数	平均	不偏分散
要因なし	20	40.2	25.5
要因あり	20	36.4	16.0

正規性と等分散性を仮定していますから、図 5.1.1 の分析の選択はt検定に到達します。

「t 検定」ボタンをクリックして、以下の検定メニューを開きます。ラジオボタンは「集計から」にし、メニュー下の集計データ記入欄に上の情報を書き込みます。



図 5.3.1 t 検定メニュー画面

「検定」ボタンをクリックすると以下のような検定結果が表示されます。

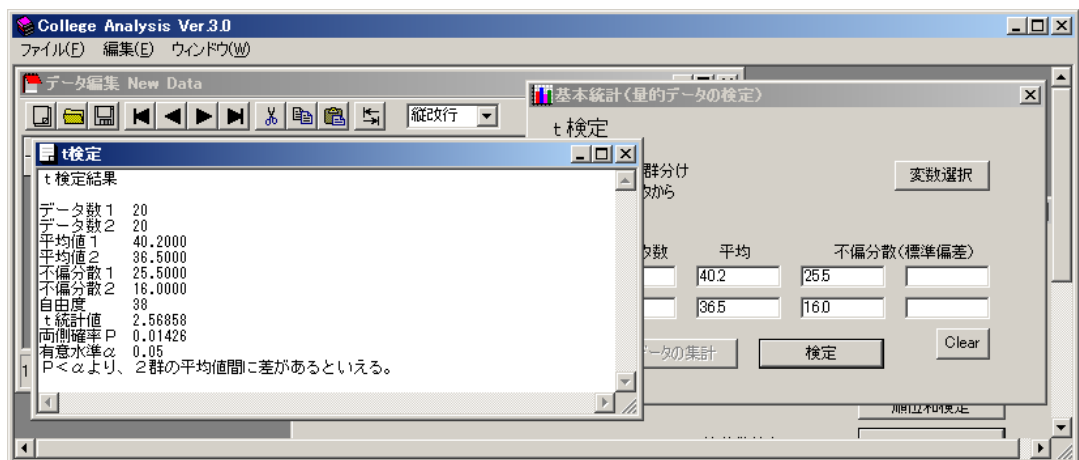


図 5.3.2 t 検定結果表示画面

検定の両側確率が $0.01426 < 0.05$ となり、2 群の平均値間に差があるといえるという結論になります。

この検定に利用した理論は以下の通りです。

理論（student の）t 検定

母平均に差がないとすると

$$t = \frac{\sqrt{\frac{n_1 n_2}{n_1 + n_2}}}{\sqrt{\frac{(n_1 - 1)u_1^2 + (n_2 - 1)u_2^2}{n_1 + n_2 - 2}}} \frac{\bar{x}_1 - \bar{x}_2}{\sim t_{n_1 + n_2 - 2} \text{ 分布}}$$

5.4 対応のない 2 群間の平均値の検定（正規性あり・等分散性なし）

次は正規性があっても、等分散性がない場合の平均値の差の検定です。5.2 節の例の続きと思って下さい。以下の例をご覧ください。

例

A 機を導入した会社 18 社（1 群）と B 機を導入した会社 15 社（2 群）について、機械 10 台当たり 1 年間の故障発生件数を調べ、平均と不偏分散を求めたところ以下の結果を得た。正規性があり、異分散であるとして、2 群間の平均に差があるかどうか有意水準 5% で検定せよ。

	平均	不偏分散
1 群	10.56	10.68
2 群	8.22	3.17

このデータは A 機の会社と B 機の会社で正規性があり、分散が異なると仮定されていますので、図 5.1.1 の分析選択画面をたどるとウェルチ（Welch）の t 検定が選択されます。そこで「Welch の t 検定」ボタンをクリックすると以下のような分析画面が表示されます。ここでもラジオボタンは「集計から」にして、下の集計データの入力部分に上で与えられたデータを入力します。

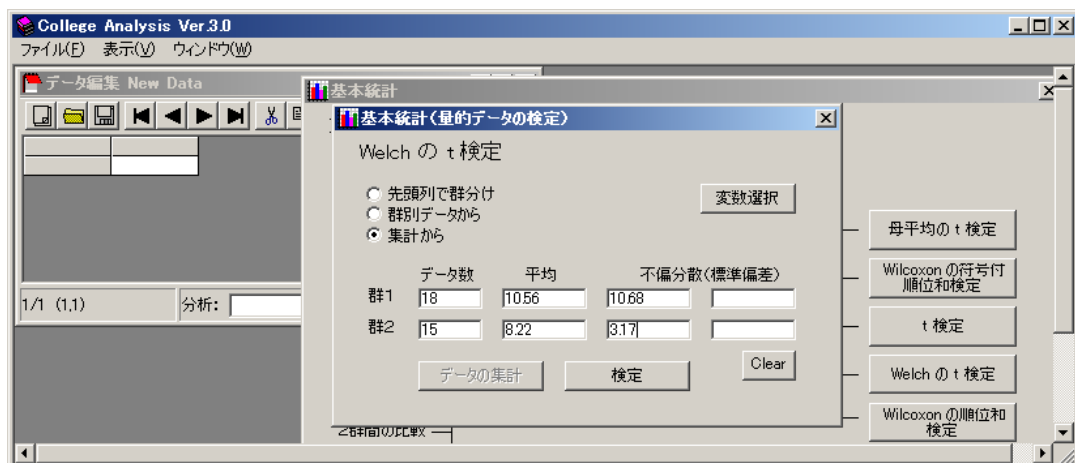


図 5.4.1 Welch の t 検定分析画面

この分析画面で「検定」ボタンをクリックすると、以下の検定結果が表示されます。

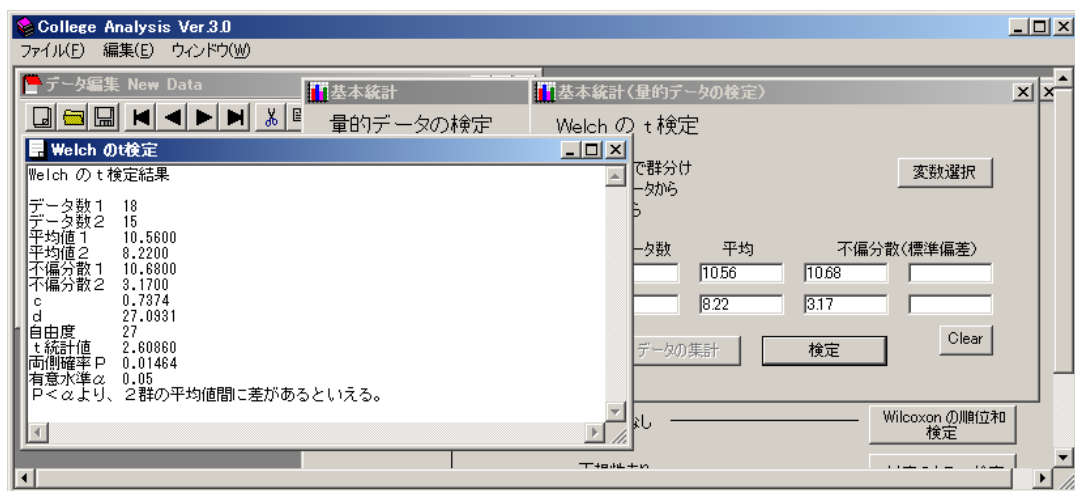


図 5.4.2 Welch の t 検定結果画面

これによると両側検定確率 $p = 0.01464 < 0.05$ で、2 群の平均値間に差があるといえるとなります。

ここで検定に利用した理論は以下の通りです。

理論 Welch(ウェルチ)の t 検定

母平均に差がないとすると

$$c = \frac{u_1^2/n_1}{u_1^2/n_1 + u_2^2/n_2} \quad \text{として、自由度を} \quad d = \frac{1}{\frac{c^2}{n_1 - 1} + \frac{(1-c)^2}{n_2 - 1}} \quad \text{とし、}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{u_1^2/n_1 + u_2^2/n_2}} \sim t_d \text{ 分布}$$

5.5 対応のない2群間の中央値の差の検定（正規性なし）

ここでは2群のデータのどちらともかどちらかのデータに正規性がない場合の処理の方法を学びます。以下の例をご覧ください。

例

あるソフトウェアの販売において、支店の売上伸び率を2つの販売戦略グループで比較したところ、以下の結果が得られた。2群の増加は1群のそれに比べて大きいといえるか。有意水準5%の両側検定で判定せよ。

1群：6, 5, 10 2群：12, 16, 22, 8, 17

データは Samples¥テキスト 5.txt にありますのでそれを開いて下さい。ここでは2つの群のデータが少なすぎて正規性の判定は困難です。このような場合は、分布によらない検定が必要です。そこでデータに正規性がないものと仮定し、図 5.1.1 の検定選択画面をたどって、Wilcoxon（ウィルコクソン）の順位和検定を選択します。「Wilcoxon の順位和検定」をクリックすると以下の分析画面が表示されます。データは群別に分けられているので、ラジオボタンは「群別データから」を選択します。

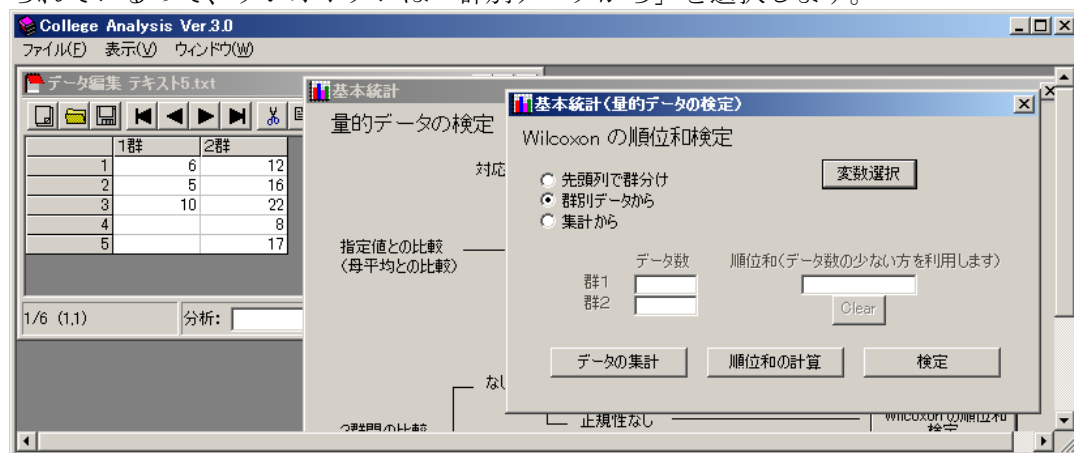


図 5.5.1 Wilcoxon の順位和検定分析画面

「検定ボタン」をクリックすると以下の結果が表示されます。

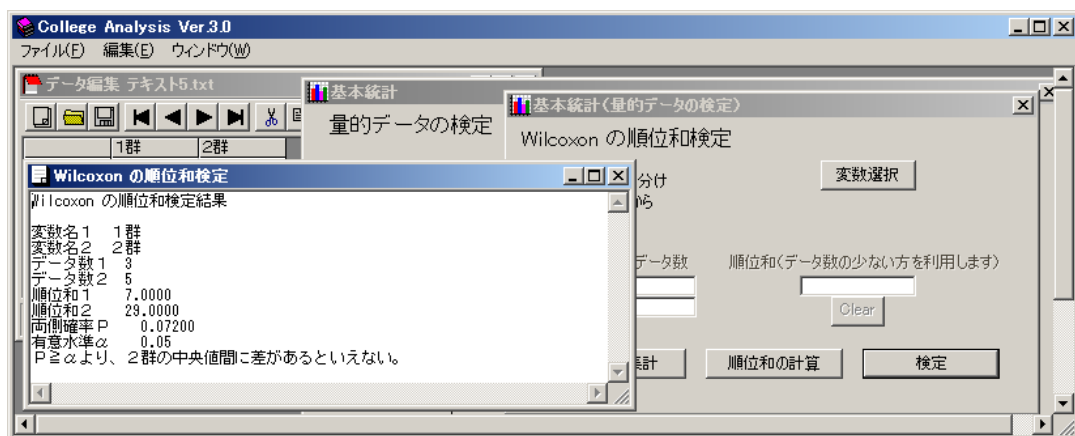


図 5.5.2 Wilcoxon の順位和検定分析結果

Wilcoxon の順位和検定は、後に述べるようにデータの大きさの順位を利用する検定ですので 2 群の平均値を比べる検定ではありません。ここで比べるのは中央値で、この結果によると検定の両側確率は $p = 0.072 > 0.05$ ですので、「2 群の中央値間に差があるといえない」ということになります。何を比べているのかははっきりさせずに「2 群間に差があるといえない」というように表現する場合もあります。

Wilcoxon の順位和検定について簡単に説明しておきます。今データを白丸で表した 1 群と黒丸で表した 2 群の 2 つの群を考えます。図 5.5.3 の上のデータ並びをケース A と下のデータ並びをケース B と呼び、状況の違う 2 つのケースとします。さてどちらのケースの中央値が異なっているように感じるでしょうか。

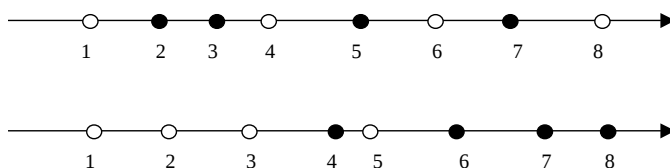


図 5.5.3 Wilcoxon の順位和検定の考え方

上は白と黒が混ざっていますから、下のケースの方が 2 つの群が分離しているように感じます。この感覚をどのように表現するのでしょうか。

今左から順番に番号を付けて行き、白丸と黒丸とでその番号の合計を取り、以下の表を作ります。

表 5.1.1 順位和

	白丸群	黒丸群
ケース A	19	17
ケース B	11	25

この表を見ると、ケース A では白丸群と黒丸群はほぼ同じ順位和ですが、ケース B では 2 つの群で順位和はかなり違います。この違いを利用して 2 群の分布を比較する検定を Wilcoxon の順位和検定といいます。具体的には以下の通りです。

理論

両群のデータの小さい順に順位を付け、データ数の少ない群 ($n_1 \leq n_2$) の順位和を W とする。但し、同じ値のデータにはそれらが異なると考えた場合の順位の平均値を付ける。例えば同順位の 3 位には $(3+4)/2=3.5$ の順位を付ける。

データ数が多い場合 両群の中央値が等しいとすると

$$z = \frac{|W - n_1(n_1 + n_2 + 1)/2| - 1/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}} \sim N(0, 1) \text{ 分布 (正の部分)} \quad (\text{Yates の連続補正})$$

データ数が少ない場合

数表を利用する。

ソフトを利用する限りこの理論を意識することはありませんが、同順位の処理についてはあくまで近似ですので、特に同順位が多い場合には注意が必要です。

問題 1

以下の標本データ (Samples¥テキスト 5.txt) の母平均 (母集団の中央値) には差があるといえるか。検定を選んで有意水準 5% で判定せよ。

1 群 112, 106, 101, 112, 102, 98, 108, 95, 101, 90, 110, 97, 95, 105, 101, 113, 114, 91

2 群 98, 88, 105, 99, 96, 93, 109, 106, 103, 87, 107, 102, 97, 91

検定名 [] 確率 []

判定 母平均 (母集団の中央値) に差があると [いえる・いえない]

問題 2

以下の標本データの母平均 (母集団の中央値) には差があるといえるか。検定を選

んで有意水準 5%で検定せよ。

1 群 358, 469, 397, 350, 329, 446, 393, 379, 443, 348,

455, 332, 311, 424, 420, 354, 353, 390, 434, 430

2 群 335, 387, 385, 343, 394, 351, 404, 391, 330, 363,

319, 334, 348, 396, 408, 403, 415, 353, 377, 399

検定名 [] 確率 []

判定 母平均（母集団の中央値）に差があると [いえる・いえない]

問題 3

ラットの体重増加(g)を、条件を変えた 2 つのグループで測定したところ、以下の結果が得られた。2 群の体重増加に差は認められるか、有意水準 5%で判定せよ。

1 群 : 7.2, 8.3, 5.4, 6.0, 7.3, 11.7, 10.5, 8.0, 9.1

2 群 : 10.1, 13.2, 7.4, 9.1, 16.2, 14.5, 6.3, 11.2, 12.4, 7.4, 12.5, 9.1, 17.0

検定名 [] 確率 []

判定 体重増加に差があると [いえる・いえない]

問題 4

Samples¥テキスト 9.txt のデータを用いて以下の問いに答えよ。

1) 地域別の年収に差があるか、検定を選んで有意水準 5%で判定せよ。

検定名 [] 確率 []

判定 地域別の年収に差があると [いえる・いえない]

2) 地域別の支出に差があるか、検定を選んで有意水準 5%で判定せよ。

検定名 [] 確率 []

判定 地域別の支出に差があると [いえる・いえない]

3) 意見 1 別の年収に差があるか、検定を選んで有意水準 5%で判定せよ。

検定名 [] 確率 []

判定 意見 1 で答え方が違う人で年収に差があると [いえる・いえない]

5.6 対応がある検定手順

2 群間に対応がある場合は、一方の群のデータに対して、他方の群に必ず対応データがあります。検定方法の決定には、この対応するデータ間の差の正規性を調べます。

正規性がある場合は、図 5.6.1 の四角形の中のように、対応がある場合の t 検定、正規性がない場合は Wilcoxon の符号付き順位和検定を利用します。

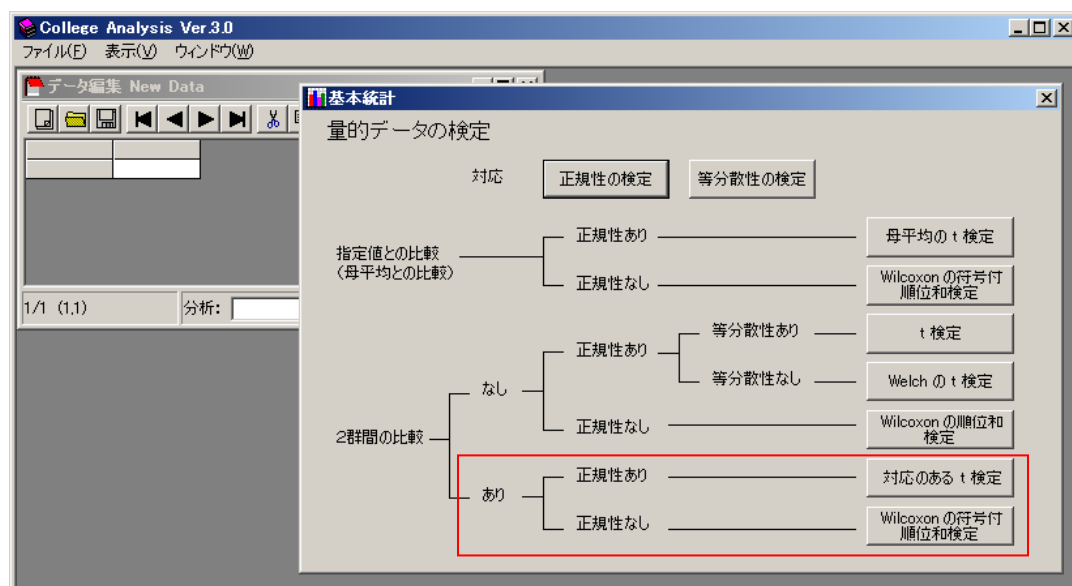


図 5.6.1 対応のある量的データの比較検定手順

対応がある場合の正規性の検定は図 5.6.2 のメニューで対応のあるデータからのラジオボタンを選択して対応するデータを選択して実行します。

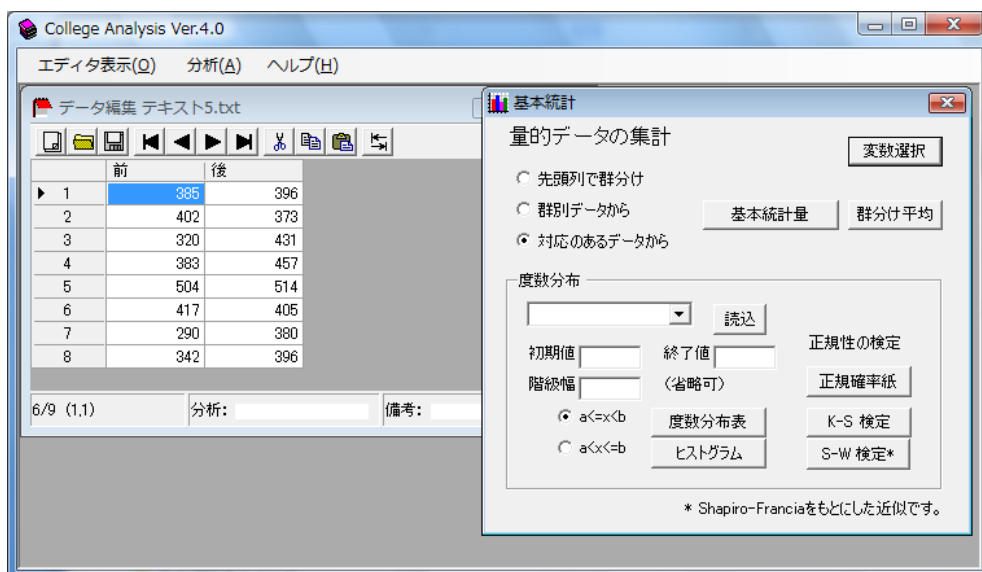


図 5.6.2 対応のあるデータの正規性の検定

以下では正規性のある場合とない場合に分けて差の検定手法を見て行きましょう。

5.7 対応がある 2 群間の平均値の検定（正規性あり）

まず始めは正規性のある例です。

例

ある商品の陳列位置を変える前と後とで売上高（千円）を規模の等しい 8 つの支店で比較したところ、以下の結果を得た。検定を選択して有意水準 5% で差があるかどうか判定せよ。

前	385	402	320	383	504	417	290	342
後	396	373	431	457	514	405	380	396

データは Samples¥テキスト 5.txt にありますので、それを開いてまず、5.6 節で説明した方法で正規性を調べます。その後、量的データの検定メニューから、2 群間の検定で対応ありの中から対応のある t 検定を選んで検定を実行します。実行結果は図 5.6.3 の通りです。

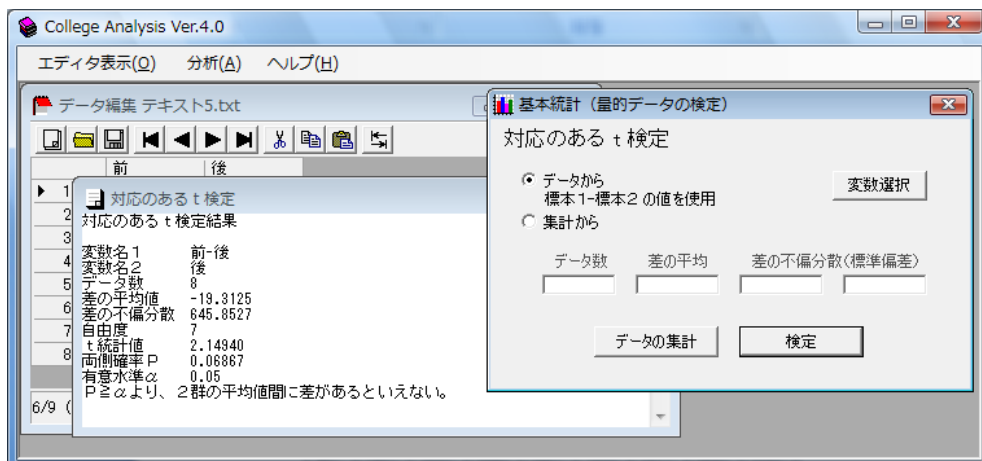


図 5.6.3 対応のある t 検定分析結果

ここで検定に利用した理論は以下の通りです。

理論

対応する各標本の差 ($z_i = \text{標本 1} - \text{標本 2}$) をとる。平均が等しいと仮定すると

$$t = \frac{\sqrt{n} \bar{z}}{u_z} \sim t_{n-1} \text{ 分布}$$

解答

$t = 2.149398$ $p = 0.068675 > 0.05$ より、平均に差があるとはいえない。

5.8 対応がある 2 群間の中央値の検定 (正規性なし)

次は、正規性のない例を見てみましょう。ここでも、中央値の検定と書いていますが、正確には中央値を比較しているわけではありません。ただ表現に困りますので昔からよく使われる中央値を使います。

例

ある商品の陳列位置を変える前と後とで売上高 (千円) を規模の等しい 8 つの支店で比較したところ、以下の結果を得た。検定を選択して有意水準 5% で売上高に差があるかどうか判定せよ。

前	385	402	320	383	504	417	290	342
後	396	310	342	407	514	405	380	365

Samples¥テキスト 5.txt のデータを開いてまず、5.6 節で説明した方法で正規性を調べます。その後、量的データの検定メニューで対応ありの中から、Wilcoxon の符号付き順位和検定を選んで検定を実行します。実行結果は図 5.6.4 の通りです。

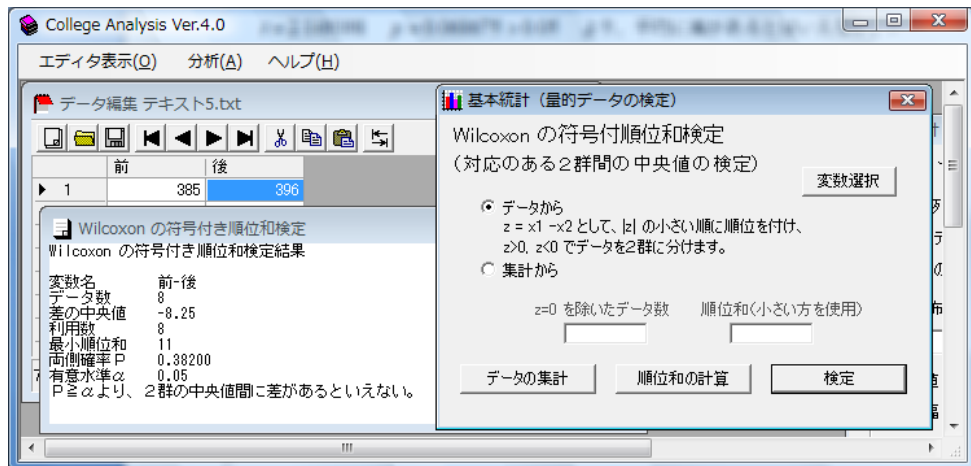


図 5.6.3 Wilcoxon の符号付き順位和検定分析結果

この検定で用いられた手法は以下の通りです。

理論 Wilcoxon の符号付き順位和検定

対応する各標本の差 ($z_i = \text{標本 1} - \text{標本 2}$) について、 z_i の正負で 2 群に分けて順位和を求め、小さい方を R とする。
標本数が多いとき (少ない場合は数表を用いる)

$$z = \frac{|R - n(n+1)/4| - 1/2}{\sqrt{n(n+1)(2n+1)/24}} \sim N(0,1) \text{ 分布 (正の部分)}$$

ここで気が付かれた方もおられると思いますが、Wilcoxon の符号付き順位和検定は 4.4 節で母集団の中央値と指定値との比較、として説明した検定手法です。これがまたここで出てきた理由は、2 つの対応するデータの差を取るという対応するデータ間の差の検定の方法によります。即ち、差を取ったものと 0 (差がない場合) とを比較しているため、0 が指定値と同じ役割になり、4.4 節の方法と同じものが用いられているのです。

問題 5

ある小学生の集団で国語・算数・社会・理科の学力を調べたところ以下のようなデータ（Samples¥テキスト 5.txt）を得た。質問に答えよ。

国語	68	58	60	63	55	69	63	79	62	74	53	75	64	77	66
算数	75	59	58	73	59	69	62	67	68	78	53	67	69	77	70
社会	66	58	50	55	57	66	54	91	57	56	65	55	80	90	63
理科	82	60	61	74	68	74	64	72	70	65	57	79	76	83	74

1) 4科目の平均値と中央値を求める。

	国語	算数	社会	理科
平均値				
中央値				

2) 各科目のデータの正規性を検討する。(みなす・いえない)

	国語	算数	社会	理科
S-W 検定確率				
正規性があると				

3) 各科目間のデータの差の正規性を検討する。

比較科目	S-W 検定確率	比較科目	S-W 検定確率
国語－算数		算数－社会	
国語－社会		算数－理科	
国語－理科		社会－理科	

2群の比較ではデータ間に1対1の対応がある場合、通常対応がある検定手法を利用するが、対応がないとして検定しても間違いではない。以下の問題は両方の方法で検定を行い、結果を比較せよ。

4) 国語と算数の平均値（中央値）に差があるといえるか、有意水準5%で判定する。

	検定名	確率	判定
対応なし			差があると [いえる・いえない]
対応あり			差があると [いえる・いえない]

5) 社会と理科の平均値(中央値)に差があるといえるか、有意水準 5%で判定する。

	検定名	確率	判定
対応なし			差があると [いえる・いえない]
対応あり			差があると [いえる・いえない]

6) 対応がある検定は対応のない検定に比べてどんな場合に有効と思えるか

データのばらつきが対応するデータ同士の差に比べて [大きい・小さい] 場合。