

2つの群の平均値の検定（等分散を前提としない）

1. 目的

2つの群の平均値が等しいかどうかを検定します。3パターンの検査ができます。
ただし ここでは 2つの群の等分散を仮定しません。

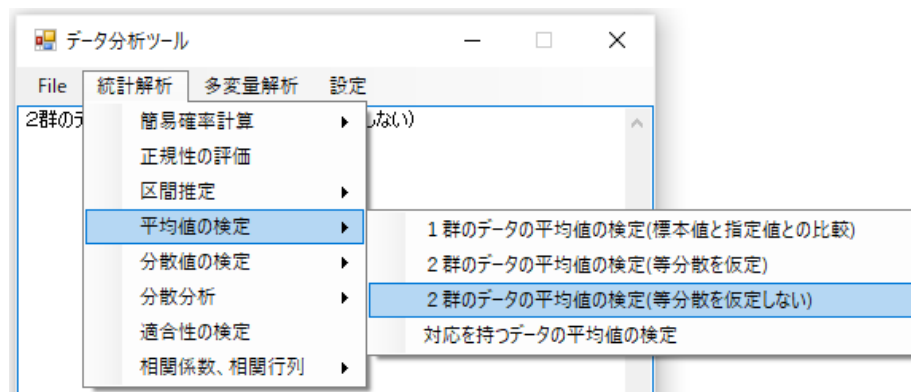
μ_1 を第1群の母平均、 μ_2 を第2群の母平均とした時

- (a) 2群間の平均値には差がない。つまり $\mu_1 = \mu_2$ といえるのかの検査
- (b) $\mu_1 < \mu_2$ といえるかの検査
- (c) $\mu_1 > \mu_2$ といえるかの検査

2. 使用法

(1) メニューの選択

メニューの「統計解析→平均値の検定→2群のデータの平均値の検定(等分散を仮定しない)」を選択します。



(2) パネルが表示されます。

棄却域の確率を示します。
通常 5%を利用するので、
デフォルトで5が指定されて
いる。変更可能。

計算結果が表示される部分

計算結果

自由度:

平均値 1: (不偏)分散値 1:

平均値 2: (不偏)分散値 2:

T 値: 帰無仮説の採択域: (,)

P 値 (%): 分布関数

結果:

効果量:

☐ 先頭行をラベルとして使用 直接入力可能

第1群のデータ 第2群のデータ

No.	ID	Value
*		

No.	ID	Value
*		

(3) データの入力

パネルの下の部分にデータを入力します。

☐ 先頭行をラベルとして使用

データ分布1

データ分布2

直接入力可能

第1群のデータ

表データを貼り付け

クリア

第2群のデータ

表データを貼り付け

クリア

	N0	ID	Value
*			

	N0	ID	Value
*			

データの入力方法は以下の2つあります。

- ① 表計算ソフトのデータをコピーして貼り付ける方法
- ② 直接数値を入力する方法

2つの車の燃費データの比較を例に説明します。

2つのメーカーの2車種のそれぞれ40台の計測された燃費データがあるとします。

いずれも 表計算ソフトのデータをコピーして、上記の表にコピーします。

☒ 先頭行をラベルとして使用

データ分布1

データ分布2

直接入力可能

第1群のデータ

表データを貼り付け

クリア

第2群のデータ

表データを貼り付け

クリア

	N0	試験車ID	燃費 (Km/l)
▶	1	A-1	22.79
	2	A-2	27.03
	3	A-3	23.49
	4	A-4	25.44
	5	A-5	20.45
	6	A-6	24.57
	7	A-7	24.29

	N0	試験車ID	燃費 (Km/l)
▶	1	B-1	25.67
	2	B-2	23.34
	3	B-3	25.85
	4	B-4	22.18
	5	B-5	23.41
	6	B-6	25.34
	7	B-7	20.36

ここでは第1群のデータ数と、第2群のデータ数はたまたま同じですが、等しくなくても構いません。

(4) 計算条件の指定

2群のデータの平均値の検定(等分散を仮定しない) (WelchのT検定)

有意水準 α (%): 5

☒ 両側 ☐ 片側検定(左) ☐ 片側検定(右)

“有意水準”には デフォルトで 5 が指定されています。変更できます。

その右側の ☒ 両側 ☐ 片側検定(左) ☐ 片側検定(右) は、

以下の3つの検査目的に応じて選択します。つまり、

「

(a) 第1群の母平均 μ_1 と 第2群の母平均 μ_2 とが等しいとみなせるかの検査

(b) 第1群の母平均 μ_1 < 第2群の母平均 μ_2 とみなせるかの検査

(c) 第1群の母平均 μ_1 > 第2群の母平均 μ_2 とみなせるかの検査

」

・「(a) 第1群の母平均 μ_1 と第2群の母平均 μ_2 とが等しいとみなせるかの検査」を目的とする場合は、☒ **両側** を選択します。

$\mu_1 = \mu_2$ の判断に利用されます。

棄却された場合（対立仮説が採択される場合）、 $\mu_1 \neq \mu_2$ ということになりますが、 $\mu_1 < \mu_2$ なのか $\mu_1 > \mu_2$ なのかは 問題にしません。

“ $\mu_1 = \mu_2$ ”であるかどうかを問題にするという立場です。

棄却されなければ（帰無仮説が採択される場合）、“ $\mu_1 = \mu_2$ である”という主張を覆せません。

・「(b) 第1群の母平均 $\mu_1 < \mu_2$ と第2群の母平均 μ_2 とみなせるかの検査」を目的とする場合は、☒ **片側検定(左)** を選択します。

$\mu_1 < \mu_2$ の判断に利用されます。

棄却された場合（対立仮説が採択される場合）、 $\mu_1 < \mu_2$ と判断できます。

棄却されなければ（帰無仮説が採択される場合）、“ $\mu_1 = \mu_2$ である”という主張を覆せません。

・「(c) 第1群の母平均 $\mu_1 > \mu_2$ と第2群の母平均 μ_2 とみなせるかの検査」を目的とする場合は、☒ **片側検定(右)** を選択します。

$\mu_1 > \mu_2$ の判断に利用されます。

棄却された場合（対立仮説が採択される場合）、 $\mu_1 > \mu_2$ と判断できます。

棄却されなければ（帰無仮説が採択される場合）、“ $\mu_1 = \mu_2$ である”という主張を覆せません。

(5) 計算実行

ボタンを押すと計算されます。

(6) 計算結果

☒ **両側** を選択しての計算結果を表示しています。

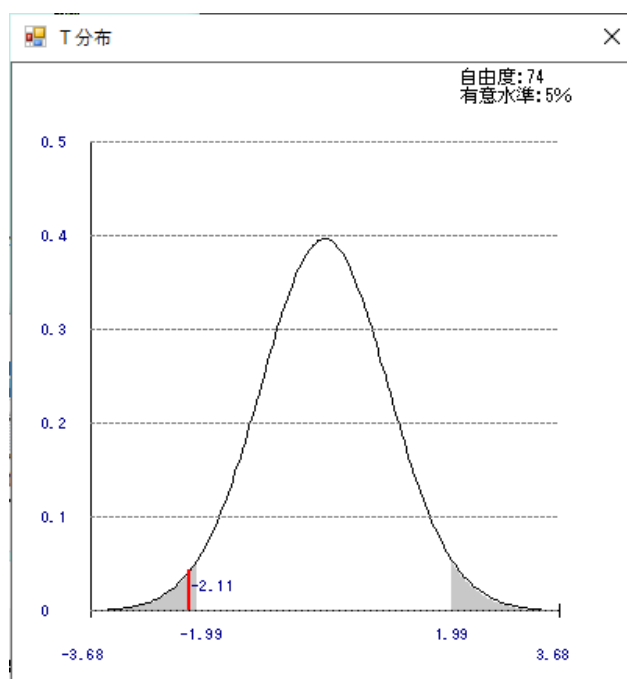
計算結果			
自由度:	74		
平均値 1:	23.84575	(不偏)分散値 1:	2.968451
平均値 2:	24.77225	(不偏)分散値 2:	4.718582
T 値:	-2.11347	帰無仮説の採択域:(-1.993049 , 1.993049)
P値(%):	3.79319		<input type="button" value="分布関数"/>
結 果:	有意: 帰無仮説(2群の平均値は等しい)を棄却する		
効果量:	-0.4725862237247		

自由度、平均値、不偏分散値は、入力された標本データをもとに計算されます。それぞれの平均値は標本データの平均ですので、母集団平均ではありませんが、母集団平均はそれぞれの標本平均値の近傍にあることになります。

“T 値”は“ $\mu_1 = \mu_2$ である”と仮定し標本データから計算された統計量で、帰無仮説の採択域（-1.99, 1.99）の中にあれば帰無仮説を採択し、外れていれば帰無仮説を棄却、つまり対立仮説を採択することになります。

この場合 T=-2.11 で、帰無仮説の採択域（-1.99, 1.99）から外れており、また有意確率（P 値）が 3.8（%）と有意水準（5%）より小さいので、対立仮説を採択することになります。

この様子は **分布関数** を選択することで、直観的に判断できます。
次のグラフが表示されます。



この T 分布を見ると、“ $\mu_1 = \mu_2$ である”と仮定して計算された T 値が-2.11 ですから、“ $\mu_1 = \mu_2$ である”としたことが棄却されるわけです。

以上をまとめた結論が、下の 1 行に記述されます。

結 果： **有意：帰無仮説(2群の平均値は等しい)を棄却する**

ここでいう“有意:”とは、帰無仮説が棄却される(対立仮説が採択される)ことを意味します。
逆に“有意でない:”とは、帰無仮説が棄却されないことを意味します。

なお、サンプルサイズ N_1 , N_2 が大きいと、統計量 T の式の一部の $\frac{\sqrt{N_1 N_2}}{\sqrt{N_1 U_y^2 + N_2 U_x^2}}$ が大きくなり、 T の絶対値は大きく、p 値は小さくなり、帰無仮説が棄却されることが多くなります。

その為、サンプルサイズの大小に影響されにくい 効果量 d を同時に表示します。

効果量 $d = \frac{\bar{X} - \bar{Y}}{U_{xy}}$ ただし $U_{xy}^2 = \frac{(N_1 - 1)U_x^2 + (N_2 - 1)U_y^2}{N_1 + N_2 - 2}$