

## 二元配置分散分析（繰り返しのあるデータ）

### 1. 目的

二元配置分散分析は2つの要因に関する平均の差を検定するものです。

ここでは、高齢男女合計12人に記憶試験を施し、年齢（年代）の違い、男女の違いで差があるものを調査するものとします。つまり、

- ・要因Aは 年代（60歳代、70歳代、80歳代）（3択）
- ・要因Bは 男女（2択）

とします。

計測値は、試験結果です。

ここで検討するのは、以下です。

- ・要因Aの年代の違いは、結果に違いをもたらすのか？
- ・要因Bの男女の違いは、結果に違いをもたらすのか？
- ・年代の違い、男女の違いにより、試験結果に相乗効果（交互作用）があるか？

データは以下の通りで、すべての要因A、要因Bの組み合わせにおけるデータ数は同じでなければなりません。この場合は、すべての組み合わせで2となっています。

No	試験 ID	要因A	要因B	値
		年代	性別	計測値
1	男性(1, 1, 1)	60歳代	男性	40
2	男性(1, 1, 2)	60歳代	男性	35
3	男性(2, 1, 1)	70歳代	男性	35
4	男性(2, 1, 2)	70歳代	男性	38
5	男性(3, 1, 1)	80歳代	男性	20
6	男性(3, 1, 2)	80歳代	男性	17
7	女性(1, 2, 1)	60歳代	女性	30
8	女性(1, 2, 2)	60歳代	女性	25
9	女性(2, 2, 1)	70歳代	女性	28
10	女性(2, 2, 2)	70歳代	女性	26
11	女性(3, 2, 1)	80歳代	女性	15
12	女性(3, 2, 2)	80歳代	女性	8

帰無仮説は以下です。

- ・要因Aの年代（60代、70代、80代）の母平均は等しい。（差はない）
- ・要因Bの男女の母平均は等しい。（差はない）
- ・要因Aと要因Bとの間には交互作用がない。

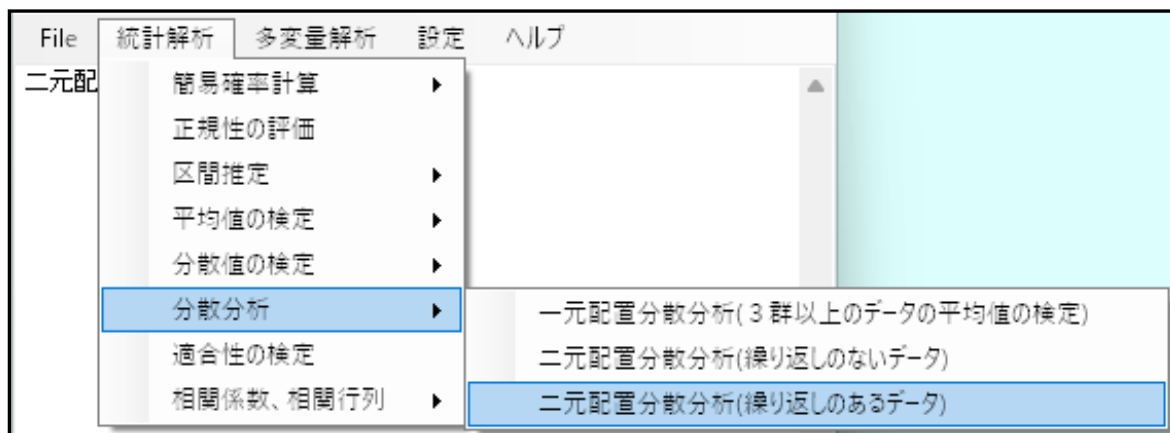
対立仮説は以下です。

- ・要因Aの年代(60, 70, 80代)の母平均は異なる。（年代による差はある）
- ・要因Bの男女の母平均は異なる。（男女による差はある）
- ・要因Aと要因Bとの間には交互作用がある。

## 2. 使用法

### (1) メニューの選択

メニューの「統計解析→分散分析→二元配置分散分析(繰り返しのあるデータ)」を選択します。



### (2) パネルが表示されます。

二元配置分散分析 (繰り返しのある)

有意水準  $\alpha$  (%): 5 ☐ 交互作用の誤差をプーリング 計算実行 使用法 考え方

分散分析表

変動因	平方和	自由度	平均平方	F 値	P 値 (%)	結果	帰無仮説の採択域
要因 A							( , ) 分布関数
要因 B							( , ) 分布関数
要因 AxB							( , ) 分布関数
誤差							
全体							

計算結果が表示される部分

	FactorA	FactorB	個数	合計	平均	(不偏)分散

表データを貼り付け ☐ 先頭行をラベルとして使用 クリア データ分布1 データ分布2

	N0	ID	FactorA	FactorB	Value
▶▶					

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。

### (3) データの入力

パネルのグリッド（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

	NO	ID	FactorA	FactorB	Value
▶*					

データは表計算ソフトのデータをコピーして貼り付けます。  
表データの形式は（ID、要因1、要因2、値）です。  
要因A、要因Bの列には、データがどこに所属するかを示します。

例えば、今回の例では、  
要因Aは年代ですので、“60歳代”～“80歳代”などの文字を入れます。  
要因Bは性別ですので、例えば“男性”、“女性”などの文字が入ります。

		要因A	要因B	値
No	試験 ID	年代	性別	計測値
1	男性 (1, 1, 1)	60歳代	男性	40
2	男性 (1, 1, 2)	60歳代	男性	35
3	男性 (2, 1, 1)	70歳代	男性	35
4	男性 (2, 1, 2)	70歳代	男性	38
5	男性 (3, 1, 1)	80歳代	男性	20
6	男性 (3, 1, 2)	80歳代	男性	17
7	女性 (1, 2, 1)	60歳代	女性	30
8	女性 (1, 2, 2)	60歳代	女性	25
9	女性 (2, 2, 1)	70歳代	女性	28
10	女性 (2, 2, 2)	70歳代	女性	26
11	女性 (3, 2, 1)	80歳代	女性	15
12	女性 (3, 2, 2)	80歳代	女性	8

上記のように表計算ソフトで作成したデータの赤線部分をコピーします。  
表題部分も取り込みますので、「先頭行をラベルとして使用」にチェックを入れます。

表データを貼り付け

をクリックすると、グリッド部分にコピーされます。

同時に、各要因 毎にデータの集計、平均、分散が表形式で表示されます。

	年代	性別	個数	合計	平均	(不偏)分散
▶	60歳代	***	4	130	32.5	41.66667
	70歳代	***	4	127	31.75	32.25
	80歳代	***	4	60	15	26

上記のグリッドは表示が小さい為、スクロールして見ることになりますが、グリッドの左上隅のセルをクリックして、Ctrl-C を押すと、表形式のデータがコピーされます。その際、グリッドが 以下のように青くなるのを確認してください。

ここをクリックして Ctrl-C を押す

	年代	性別	個数	合計	平均	(不偏)分散
▶	60歳代	***	4	130	32.5	41.66667
	70歳代	***	4	127	31.75	32.25
	80歳代	***	4	60	15	26

そのまま、表計算ソフトのシート上で Ctrl-V を押すと、表データがペーストされます。

年代	性別	個数	合計	平均	(不偏)分散
60歳代	***	4	130	32.5	41.66667
70歳代	***	4	127	31.75	32.25
80歳代	***	4	60	15	26
***	男性	6	185	30.83333	95.76667
***	女性	6	132	22	74
60歳代	男性	2	75	37.5	12.5
70歳代	男性	2	73	36.5	4.5
80歳代	男性	2	37	18.5	4.5
60歳代	女性	2	55	27.5	12.5
70歳代	女性	2	54	27	2
80歳代	女性	2	23	11.5	24.5

- 1) 要因Aのカテゴリは年代であり、
  - ・ 60 歳代の件数は 4、平均 32.5、分散 41.7
  - ・ 70 歳代の件数は 4、平均 31.7、分散 32.25
  - ・ 80 歳代の件数は 4、平均 15.0、分散 26
- 2) 要因Bのカテゴリは“男性”、“女性”であり、
  - ・ 男性の件数は 6、平均 30.83、分散 95.76
  - ・ 女性の件数は 6、平均 22、分散 74
- 3) 要因Aと要因Bとの交互部分“60 歳代” x “男性”～“80 歳代” x “女性”のそれぞれについても同様に続きます。

(4) 計算条件の指定

有意水準 $\alpha$ (%) :	5
---------------------	---

“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。

(5) 計算実行

**計算実行** ボタンを押すと計算されます。

分散分析表							
変動因	平方和	自由度	平均平方	F 値	P 値 (%)	結果	帰無仮説の採択域
要因 A	783.16666	2	391.58333	38.834710	0.03688	有意	( 0 , 5.1432528 ) 分布関数
要因 B	234.08333	1	234.08333	23.214876	0.2946	有意	( 0 , 5.9873776 ) 分布関数
要因 A x B	5.1666666	2	2.5833333	0.2561983	78.2	有意でない	( 0 , 5.1432528 ) 分布関数
誤差	60.5	6	10.083333				
全体	1082.9166	11					

平方和、自由度、平方平均、F 値等 が 入力された標本データをもとに計算されます。

要因 A (“60 歳代”、“70 歳代”、“80 歳代”) については：

F 値が 38.83、P 値が 0.036 % となり、“有意”と判断されました。  
つまり 年代別についての有意差が認められました。

要因 B (“男性”、“女性”) については：

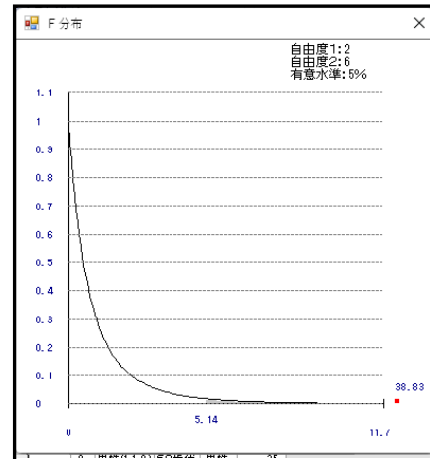
F 値が 23.21、P 値が 0.29 % となり、“有意”と判断されました。  
つまり 男女別についての有意差が認められました。

要因 A x B (交互作用) については：

F 値が 0.25、P 値が 78.2 % となり、“有意でない”と判断されました。  
つまり 交互作用による効果は認められません。

この様子は、**分布関数** ボタンを押すことで確認できます。

例えば要因Aの **分布関数** ボタンを押すと、F分布のグラフが描画されています。  
計算されたF値は38.83なので、グラフのはるか右側に位置することになります。図では、図からはみ出さない位置に 赤い点で描画されています。



## (6) 誤差のプーリング

今回の計算では、“要因A x B”で示される 交互作用による効果が小さいことがわかりました。

変動因	平方和	自由度	平均平方	F 値	P 値 (%)	結果	帰無仮説の採択域	
要因 A	783.16666	2	391.58333	38.834710	0.03688	有意	( 0 , 5.1432528 )	分布関数
要因 B	234.08333	1	234.08333	23.214876	0.2946	有意	( 0 , 5.9873776 )	分布関数
要因 A x B	5.1666666	2	2.5833333	0.2561983	78.2	有意でない	( 0 , 5.1432528 )	分布関数
誤差	60.5	6	10.083333					
全体	1082.9166	11						

そこで、この交互作用を 誤差項に加えて 再度 計算しなおすことができます。

有意水準  $\alpha$  (%) : 5 ☒ 交互作用の誤差をプーリング

「交互作用の誤差をプーリング」にチェックを入れることで、できます。

変動因	平方和	自由度	平均平方	F 値	P 値 (%)	結果	帰無仮説の採択域	
要因 A	783.16666	2	391.58333	47.705583	0.003582	有意	( 0 , 4.4589701 )	分布関数
要因 B	234.08333	1	234.08333	28.517766	0.0694	有意	( 0 , 5.3176550 )	分布関数
要因 A x B	*****	****	*****	*****	*****	*****	( ***** , ***** )	分布関数
誤差	65.666666	8	8.2083333					
全体	1082.9166	11						

## (7) 分布のグラフ描画 (データ分布1)

要因の効果を直感的に理解するためのグラフです。

要因の効果とは、この場合

- ① 年代の違いに差があるか？
- ② 男女の違いに差があるか？
- ③ 年代と男女の組み合わせによる交互作用が認められるか？

となります。

各測定点 (X) から全データの平均 ( $\mu$ ) との差は、以下で表現されます。

$$X - \mu = a + b + (a b) + \varepsilon$$

- ・  $a$  は 要因 A による効果
- ・  $b$  は 要因 B による効果
- ・  $(a b)$  は 要因 A と要因 B との交互作用による効果
- ・  $\varepsilon$  は 誤差

以降の各図は、

要因 A についての図は  $a$  の大きさを表示

要因 B についての図は  $b$  の大きさを表示

要因 A x B についての図は  $(a b)$  の大きさを表示

誤差についての図は 誤差の大きさを表示

となります。

少し正確に書くと、上記については 以下のようになります。

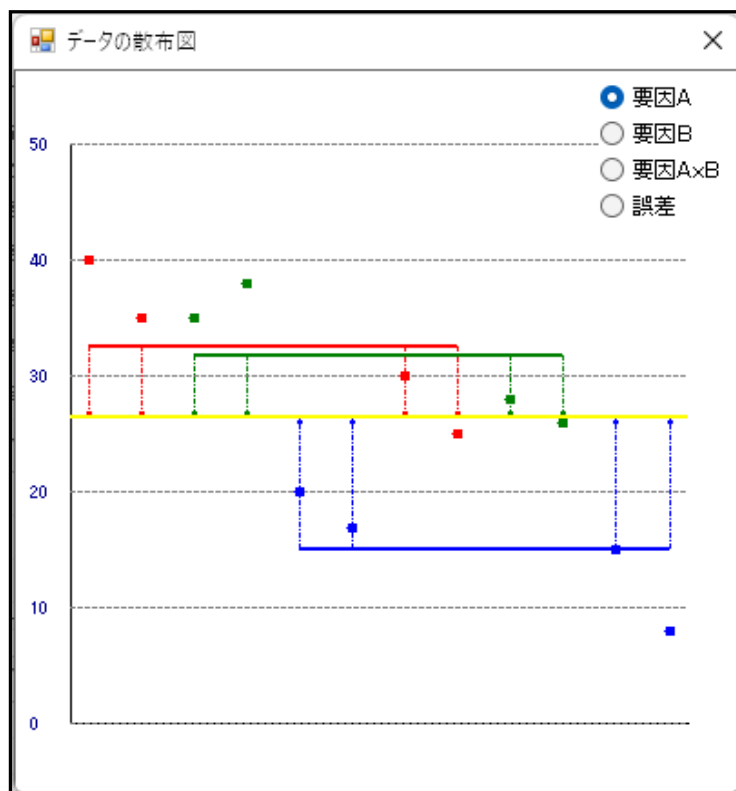
$$\begin{aligned} X_{i,j,k} - \mu &= a_i + b_j + (ab)_{i,j} + \varepsilon_{i,j,k} \\ a_i &\equiv (\mu_i^A - \mu) \\ b_j &\equiv (\mu_j^B - \mu) \\ (ab)_{i,j} &\equiv \mu_{i,j}^{AB} - \mu - (\mu_i^A - \mu) - (\mu_j^B - \mu) \\ \varepsilon_{i,j,k} &\equiv X_{i,j,k} - \mu_{i,j}^{AB} \end{aligned}$$

$$\mu = \frac{\sum_{i=1}^{n_A} \sum_{j=1}^{n_B} \sum_{k=1}^n X_{i,j,k}}{n * n_A * n_B}, \mu_i^A = \frac{\sum_{j=1}^{n_B} \sum_{k=1}^n X_{i,j,k}}{n * n_B}, \mu_j^B = \frac{\sum_{i=1}^{n_A} \sum_{k=1}^n X_{i,j,k}}{n * n_A}, \mu_{i,j}^{AB} = \frac{\sum_{k=1}^n X_{i,j,k}}{n}$$

詳細は「考え方」のドキュメントを参照してください。

データ分布を直観的に理解する為に、 **データ分布1** ボタンを押します。

1) 最初に 要因A についてのグラフを表示します。



図の 赤い点、緑の点、青の点は 入力された測定点です。

この図は 要因Aを選んでますので、

- ・赤点は 要因Aの“60歳代”のデータ
- ・緑点は 要因Aの“70歳代”のデータ
- ・青点は 要因Aの“80歳代”のデータ

を示します。

また 水平の線は、

- ・赤線は 要因Aの“60歳代”の平均を示す線
- ・緑線は 要因Aの“70歳代”の平均を示す線
- ・青線は 要因Aの“80歳代”の平均を示す線

を示し、

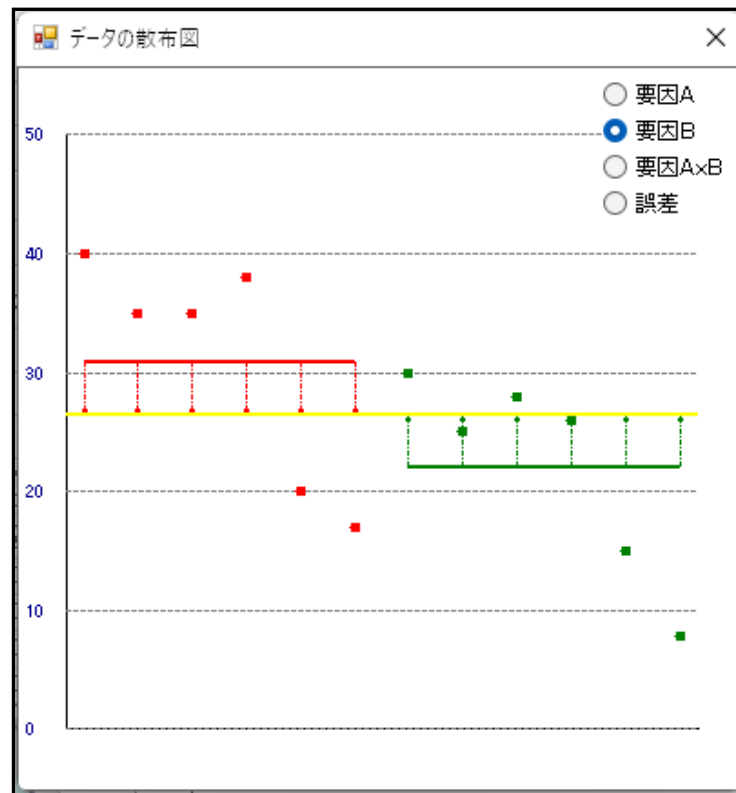
- ・黄色い線は、全体の平均を示す線

です。

赤、緑、青の各線から黄色の線に垂直に矢印線が引かれていますが、この矢印が大きく表示されていると、要因Aについての群間差（年齢による差）が大きいことを示し、逆に小さいと群間差は小さいことを示します。



2) 今度は 要因B を指定してグラフを表示します。



この図は 要因Bを選んでますので、

- ・赤点は 要因Bの“男性”のデータ
- ・緑点は 要因Bの“女性”のデータ

を示します。

先と同様に 水平の線は、

- ・赤線は 要因Bの“男性”の平均を示す線
- ・緑線は 要因Bの“女性”の平均を示す線

を示し、

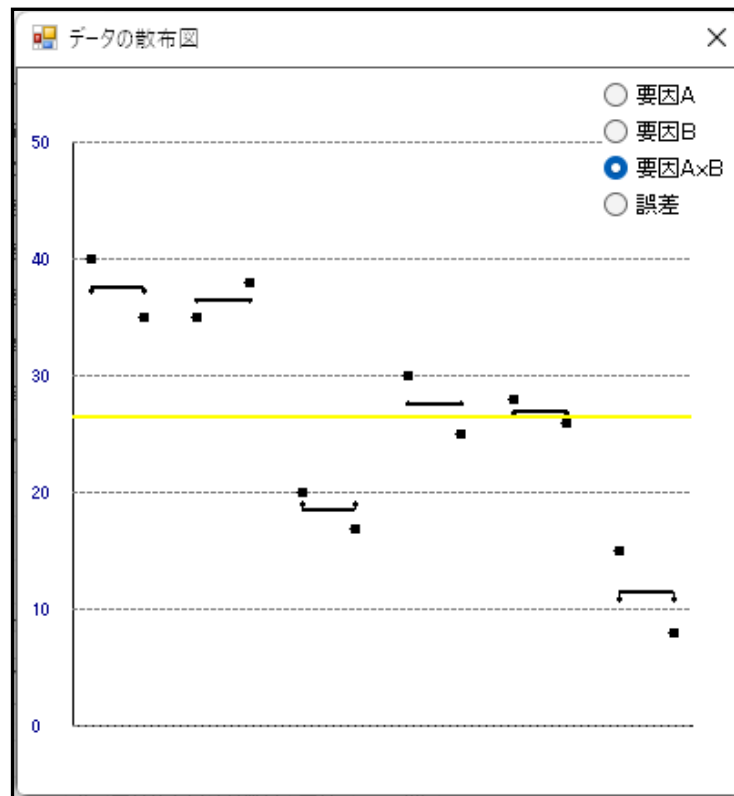
- ・黄色い線は、全体の平均を示す線

です。

赤、緑の各線から黄色の線に垂直に矢印線が引かれていますが、この矢印が大きく表示されていると、要因Bについての群間差（男女別による差）が大きいことを示し、逆に小さいと群間差は小さいことを示します。

3) 続いて 要因A x B を指定してグラフを表示します。

これは 要因Aと要因Bとによる交互作用の有無を評価するものです。



黄色い線は、全体の平均点を示す線です。

この図は要因A x Bを選んでますので、要因A， 要因Bの各条件の組み合わせによる交互作用の大きさを、黒の矢印で表示してます。

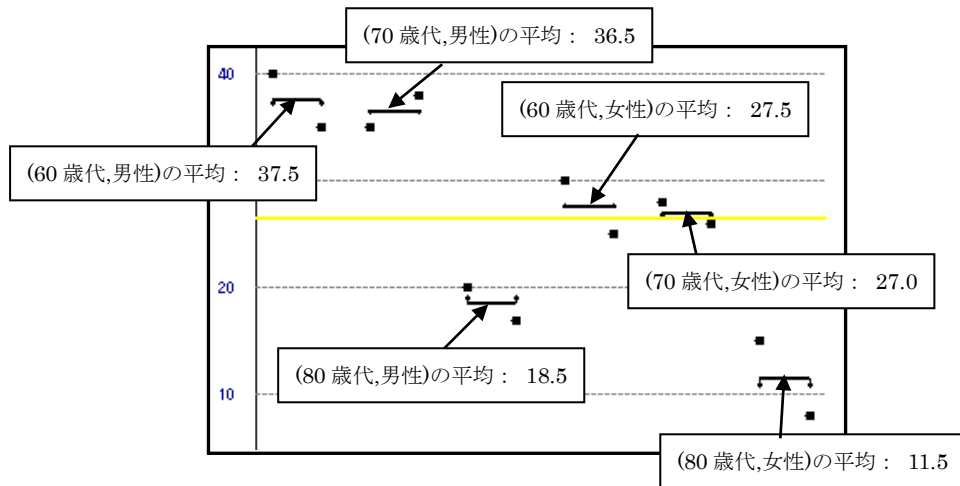
ここでは、すべての測定点を黒で表示します。

各要因の組み合わせは、以下の6通りです。

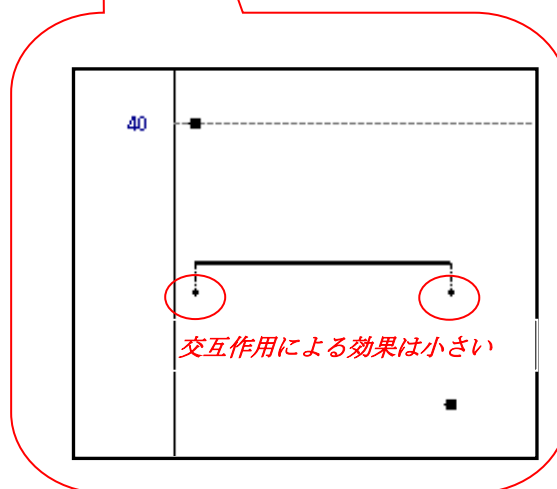
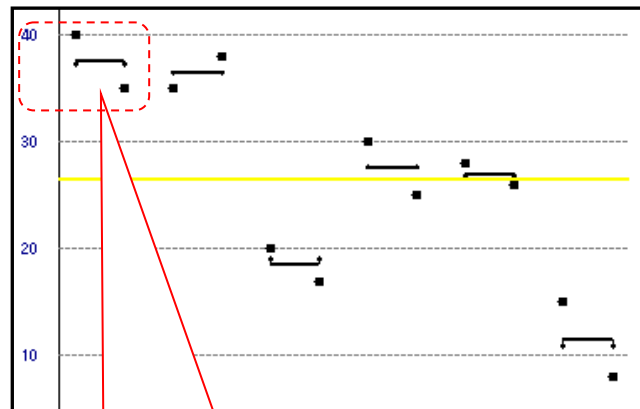
- ・ (60歳代、男性)
- ・ (70歳代、男性)
- ・ (80歳代、男性)
- ・ (60歳代、女性)
- ・ (70歳代、女性)
- ・ (80歳代、女性)

図の黒の水平線は各組合せの平均を示すものです。

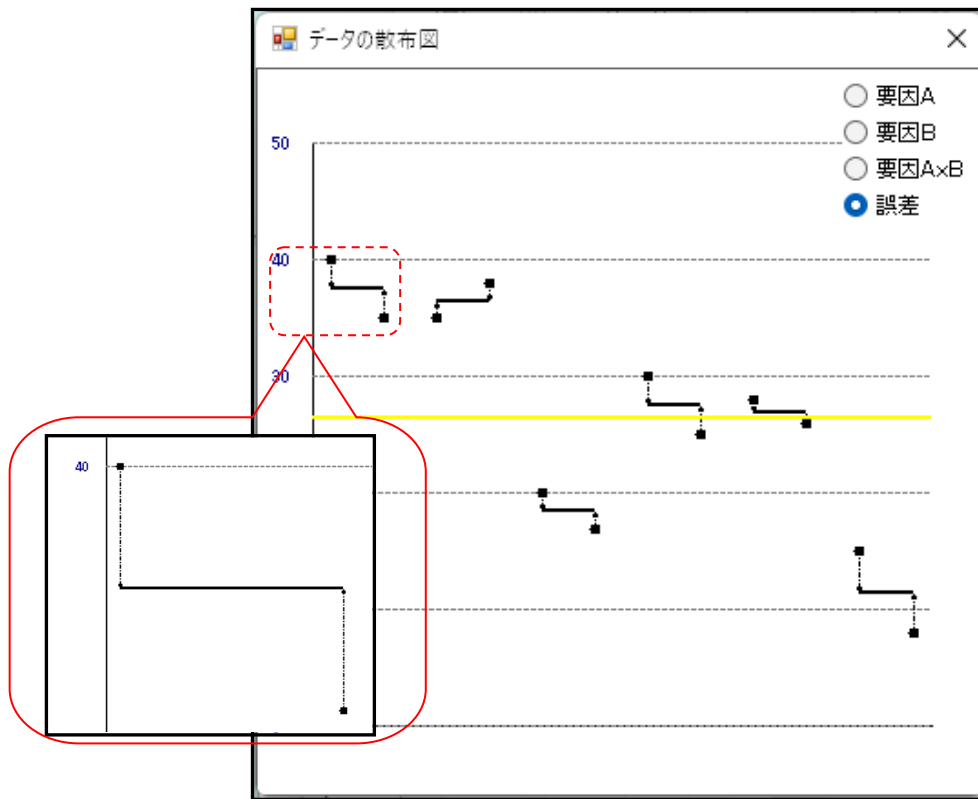
すなわち、



各要因の組み合わせにおける交互作用は、この場合、値が小さくて確認が難しいのですが、最初の部分を拡大します。



4) 続いて 誤差を指定してグラフを表示します。



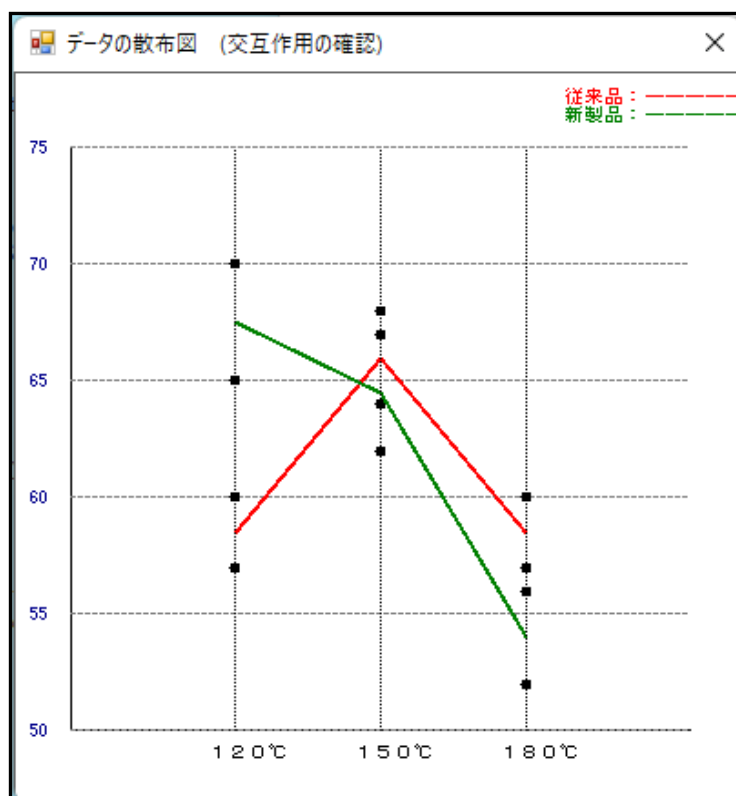
黄色い線は、全体の平均点を示す線です。

測定点から 要因の組み合わせの平均 への誤差を表示しています。

上図は最初の（60歳代、男性）に属する2つのデータから平均への誤差を拡大表示したものです。

(8) 分布のグラフ描画 (データ分布2)

要因Aと要因Bとによる交互作用の有無を評価するものです。



交互作用が有る場合、上図のように線が交絡します。  
なお、交互作用がない場合は、下図のように線が交絡しません。

