2020年11月19日
Discovery Summit Japan 2020
@WEB

最小2乗平均の謎を予測プロファイルで解く
ー交互作用を考慮した共分散分析を用いてー

高橋 行雄
BioStat 研究所（株）

最小2乗平均とは

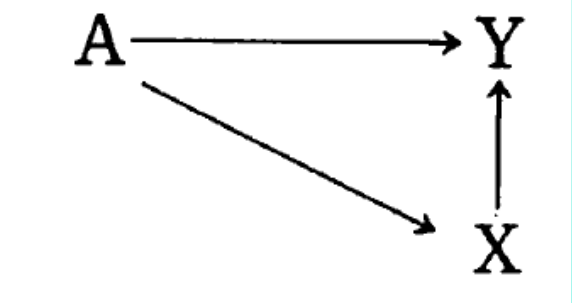
- JMPの「モデルのあてはめ」で得られる「最小2乗平均」および「予測プロファイル」は，探索的な解析の結果の表示になくてはならないものである.
- しかし，共に竹内ら（1989），「統計学辞典」の索引にも載っていない方言みたいなもので，説明するのに苦労することになる.

予測プロファイルとは

- 最小2乗平均は，SASの方言で，SASの一般線形モデル，GLMプロシジャで活躍し始め，LSMEANS として知られている.
- 予測プロファイルは，SASにもないJMPの方言である.
- 最小2乗平均は，JMPでも活躍しており，予測プロファイルを使うと簡単に説明ができる.

理解を深めるための事例

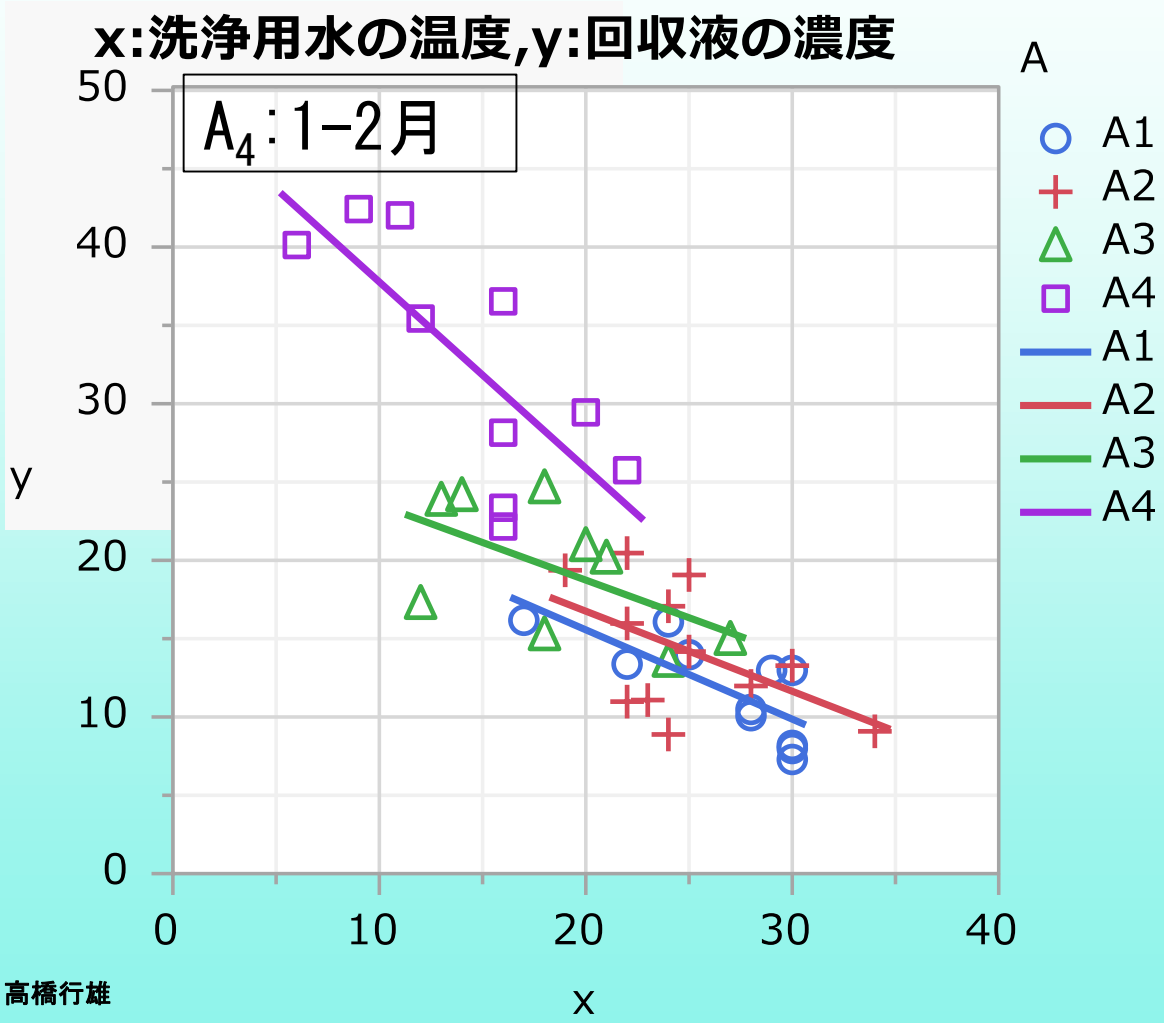
- 奥野ら(1981),「多変量解析法，改訂版」
- 「層別因子を含む回帰分析」の事例
- 亜硫酸ガス回収搭の観察データ
  - 回収液の濃度　　*Y*
  - 洗浄用水の温度　　*X*
  - 4水準の操業時期　　A
  - データ数　　42



データリスト

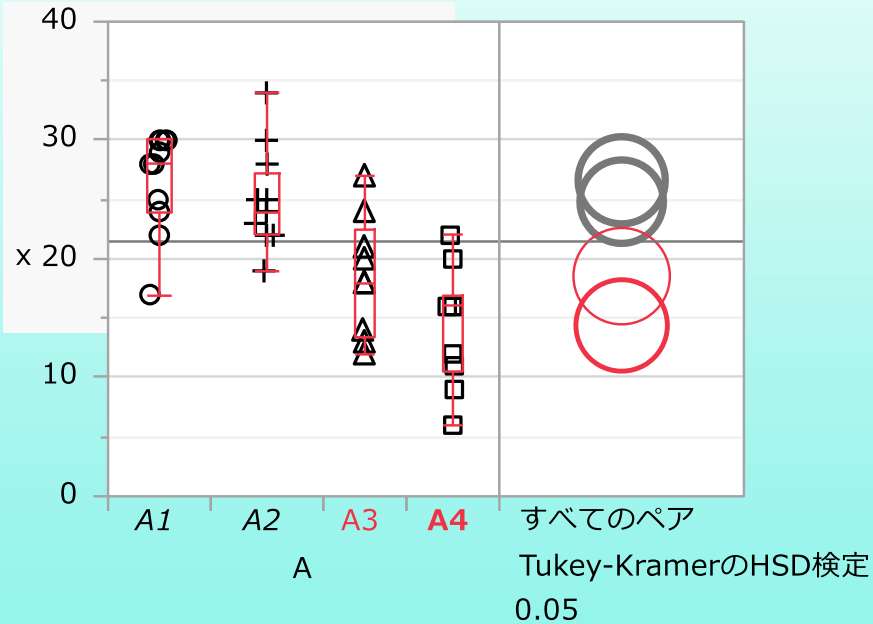
	A_1（7月・8月）		A_2（9月・10月）		A_3（11月・12月）		A_4（1月・2月）	
No.	*x*	*y*	*x*	*y*	*x*	*y*	*x*	*y*
1	30	8.0	34	9.1	27	15.0	16	23.4
2	28	10.5	19	19.4	20	21.0	16	28.2
3	30	8.2	22	20.5	14	24.2	20	29.5
4	29	13.0	25	14.2	18	15.3	16	22.2
5	28	10.1	22	11.0	12	17.3	6	40.2
6	17	16.2	25	19.1	18	24.7	16	36.6
7	24	16.1	22	16.0	13	23.9	12	35.5
8	22	13.4	23	11.1	21	20.2	11	42.1
9	30	13.0	28	12.0	24	13.6	9	42.5
10	30	7.3	24	17.1			22	25.8
11	25	14.0	30	13.3				
12			24	8.9				
平均	26.64	11.80	24.83	14.31	18.56	19.47	14.40	32.60
						総平均	21.4762	19.1119

亜硫酸ガス回収搭の観察結果

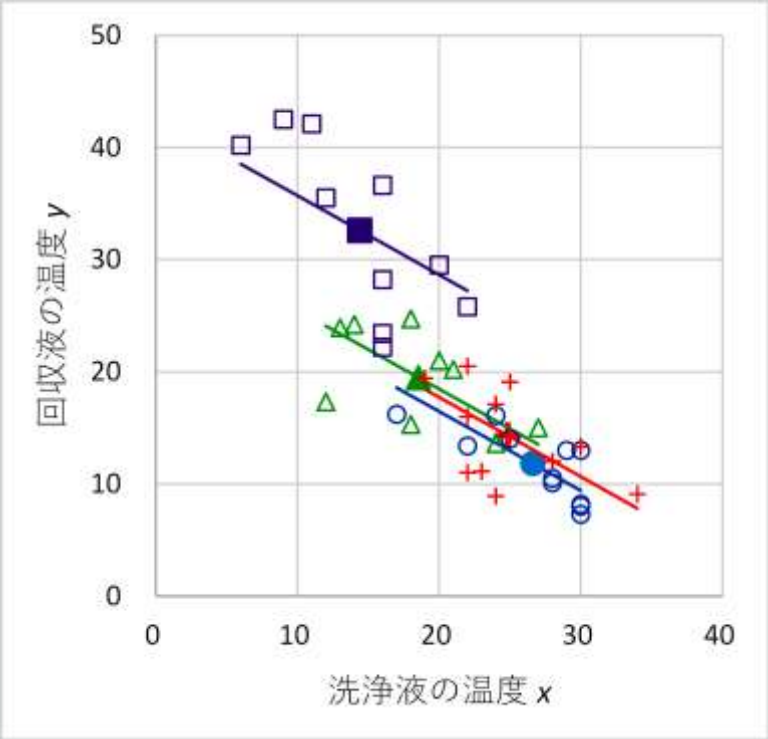


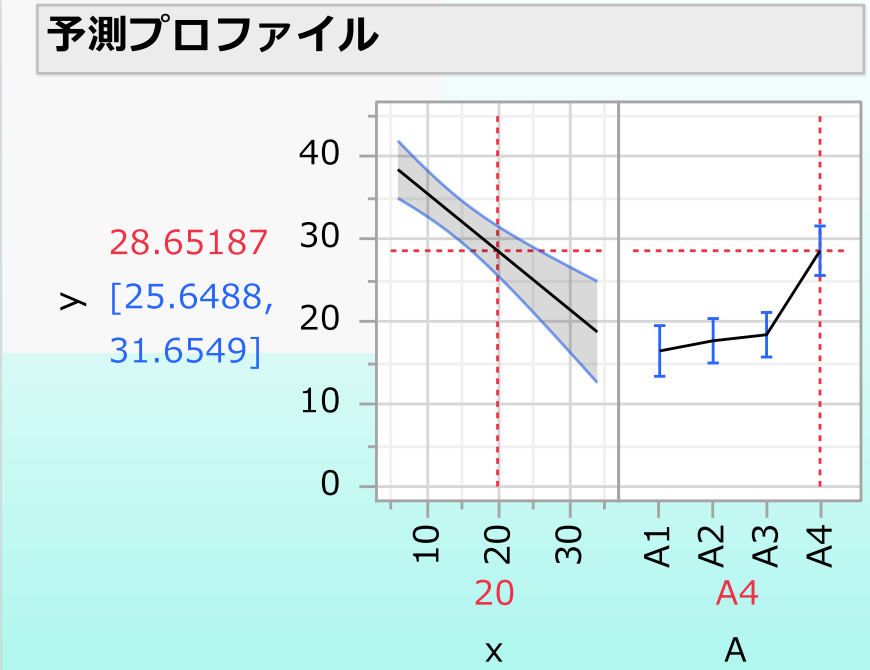
交互作用の扱い

- 奥野ら(1981)は，季節Aと蒸留水の温度*X*との交互作用が有意ではないので，いわゆる共分散分析を行っている．
- 蒸留水の温度*X*は，季節Aにより，明らかな差がある．
- 交互作用を含めた解析も必要と思われる．



平行線のあてはめは妥当か？





A1（夏）からA4（冬）に向け回収液の温度*Y*は上昇しているが，Xの影響だけなのか．

交互作用は有意でない

分散分析

要因	自由度	平方和	平均平方	F値
モデル	7	3145.7754	449.396	30.0158
誤差	34	509.0487	14.972	**p値(Prob>F)**
全体(修正済み)	41	3654.8240		<.0001 *

逐次(タイプ1)検定

要因	自由度	逐次平方和	F値	p値(Prob>F)
x	1	2523.0347	168.5167	<.0001 *
A	3	551.4944	12.2783	<.0001 *
x*A	3	71.2462	1.5862	0.2107

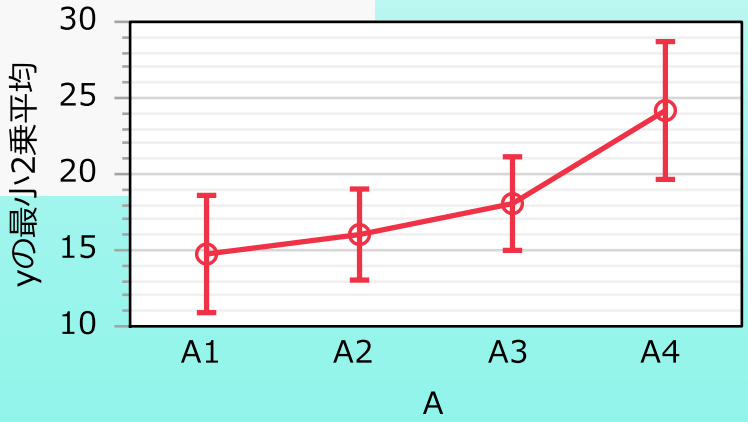
因子Aについての最小2乗平均

A

最小2乗平均表

水準	最小2乗平均	標準誤差	下側95%	上側95%	平均
A1	14.7479	1.8958	10.8953	18.6006	11.8000
A2	16.0280	1.4719	13.0368	19.0192	14.3083
A3	18.0618	1.5129	14.9872	21.1364	19.4667
A4	24.1766	2.2275	19.6499	28.7033	32.6000

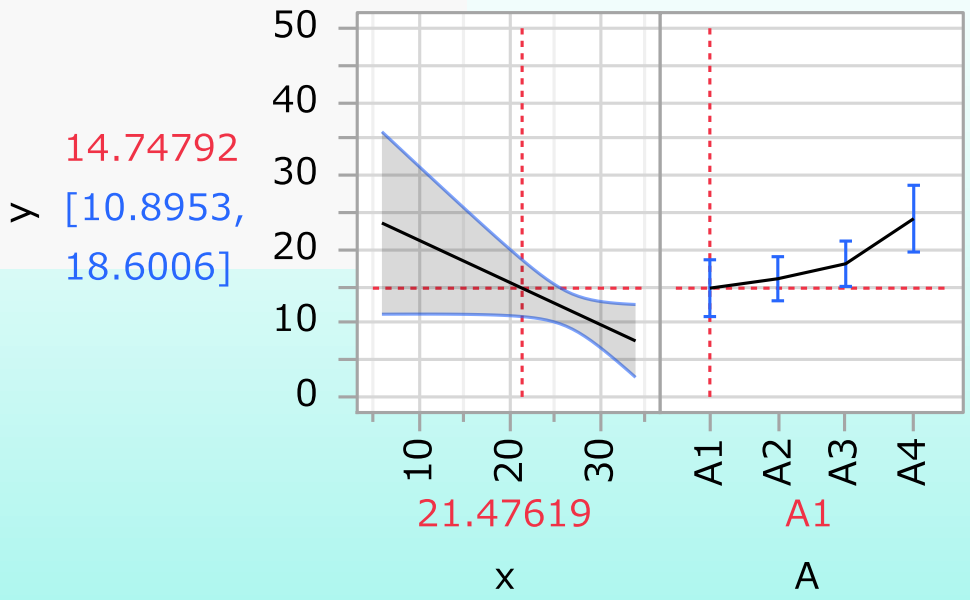
最小2乗平均プロット



◆ 計算式は？

因子A1：予測プロファイル

予測プロファイル

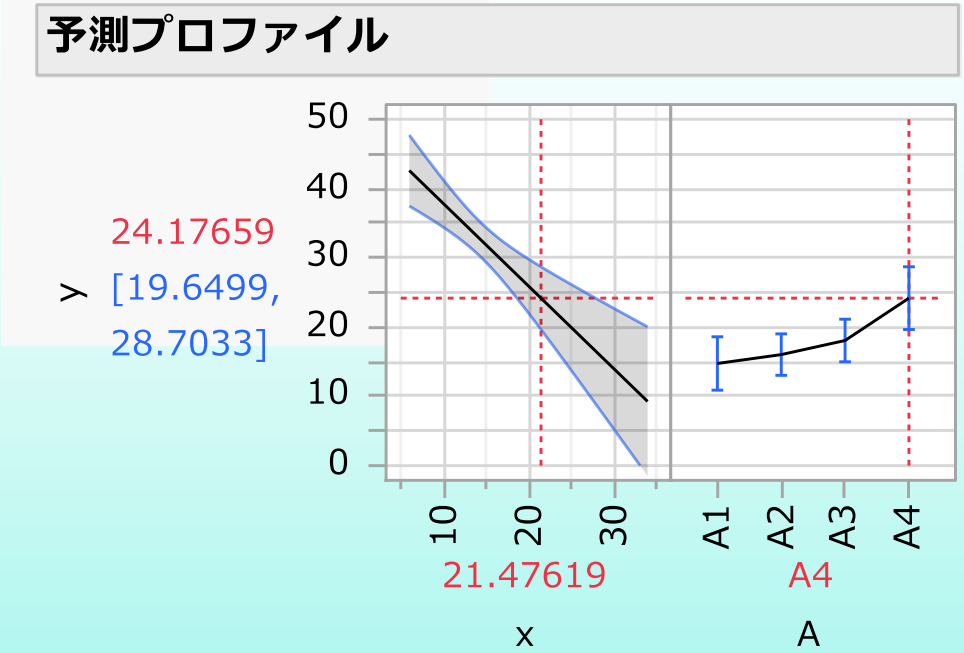


- 総平均 X=21.47619が標示されている
- A=A1を選択
- *Y*=14.7579 [10.8952, 18.6006] 95%信頼区間

最小2乗平均表

水準	最小2乗平均	標準誤差	下側95%	上側95%
A1	14.7479	1.8958	10.8953	18.6006

因子A4：予測プロファイル



- 総平均 X=21.47619 が標示されている
- A=A4を選択
- Y=21.1766 [19.6499, 28.7033] 95%信頼区間

最小2乗平均表

水準	最小2乗平均	標準誤差	下側95%	上側95%
A4	24.1766	2.2275	19.6499	28.7033

回帰直線のパラメータは

- JMPの内部での解析用の変数

対比型デザイン変数			
A	a_1	a_2	a_3
A_1	**1**	0	0
A_2	0	**1**	0
A_3	0	0	**1**
A_4	**-1**	**-1**	**-1**

交互作用のデザイン変数			
A	a_1x	a_2x	a_3x
A_1	x	0	0
A_2	0	x	0
A_3	0	0	x
A_4	$-x$	$-x$	$-x$

- 足してゼロとなる対比型のダミー変数
- SASでは，最後の水準を0とする（1，0）型
- Rでは，最初の水準を0とする（0，1）型

パラメータの推定値

パラメータ推定値

項	推定値
切片	33.0450
x	-0.6887
A[A1]	-6.0281
A[A2]	-6.0157
A[A3]	-4.6528
x*A[A1]	0.1175
x*A[A2]	0.1765
x*A[A3]	0.2077

和
A1切片

和
A1傾き

全水準の推定値

項	推定値
切片	33.0450
x	-0.6887
A[A1]	-6.0281
A[A2]	-6.0157
A[A3]	-4.6528
A[A4]	16.6966
x*A[A1]	0.1175
x*A[A2]	0.1765
x*A[A3]	0.2077
x*A[A4]	-0.5017

和
A4切片

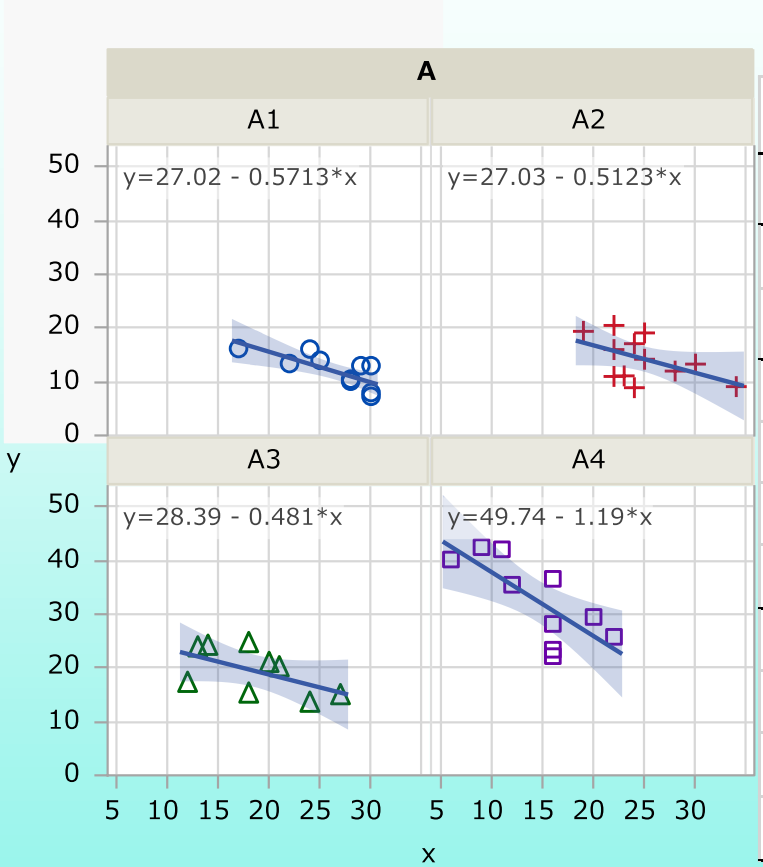
和
A4傾き

A[A4]=−[A1]−[A2]−[A3]=16.6966

x*A[A4] = −0.1175−0.1756−0.2077

= −0.5017

切片と傾きの推定



全水準の推定値						
項	推定値					
切片	33.0450					
x	-0.6887					
A[A1]	-6.0281	+	33.0450	=	27.0169	A1切片
A[A2]	-6.0157	+	33.0450	=	27.0292	A2 "
A[A3]	-4.6528	+	33.0450	=	28.3921	A3 "
A[A4]	16.6966	+	33.0450	=	49.7416	A4 "
x*A[A1]	0.1175	+	-0.6887	=	-0.5713	A1傾き
x*A[A2]	0.1765	+	-0.6887	=	-0.5123	A2 "
x*A[A3]	0.2077	+	-0.6887	=	-0.4810	A3 "
x*A[A4]	-0.5017	+	-0.6887	=	-1.1904	A4 "

A4 : $y\hat{}_{A4}$ = 49.7416 – 1.1904 x

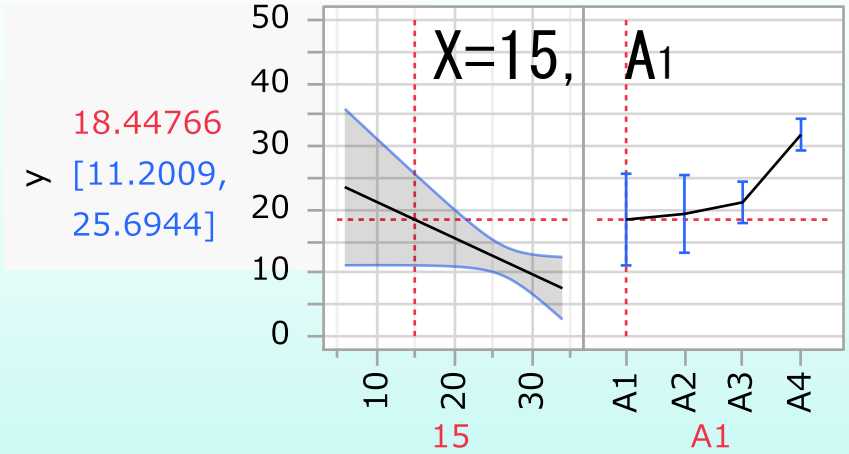
最小2乗平均の計算

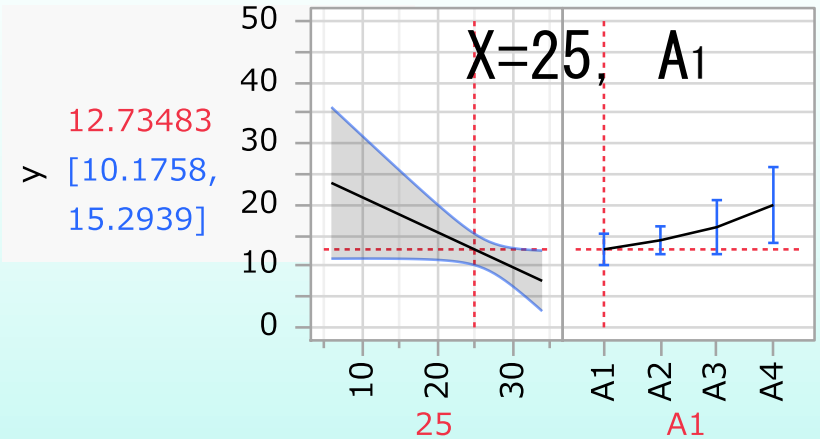
全水準の推定値			最小2乗平均＝					
項	推定値		A_i 切片＋A_i 傾き×x の総平均					
切片	33.0450							
x	-0.6887				x の総平均＝			21.47619
A[A1]	-6.0281	+	33.0450	=	27.0169	A1切片		
A[A2]	-6.0157	+	33.0450	=	27.0292	A2 "		
A[A3]	-4.6528	+	33.0450	=	28.3921	A3 "		
A[A4]	16.6966	+	33.0450	=	49.7416	A4 "	最小2乗平均	
x*A[A1]	0.1175	+	-0.6887	=	-0.5713	A1傾き		**14.7479**
x*A[A2]	0.1765	+	-0.6887	=	-0.5123	A2 "		**16.0280**
x*A[A3]	0.2077	+	-0.6887	=	-0.4810	A3 "		**18.0618**
x*A[A4]	-0.5017	+	-0.6887	=	-1.1904	A4 "		**24.1766**

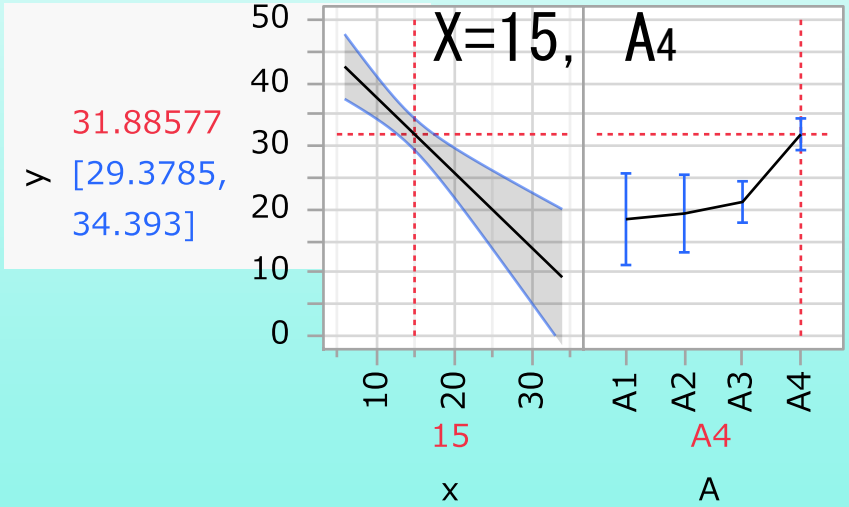
水準	最小2乗平均
A1	14.7479
A2	16.0280
A3	18.0618
A4	24.1766

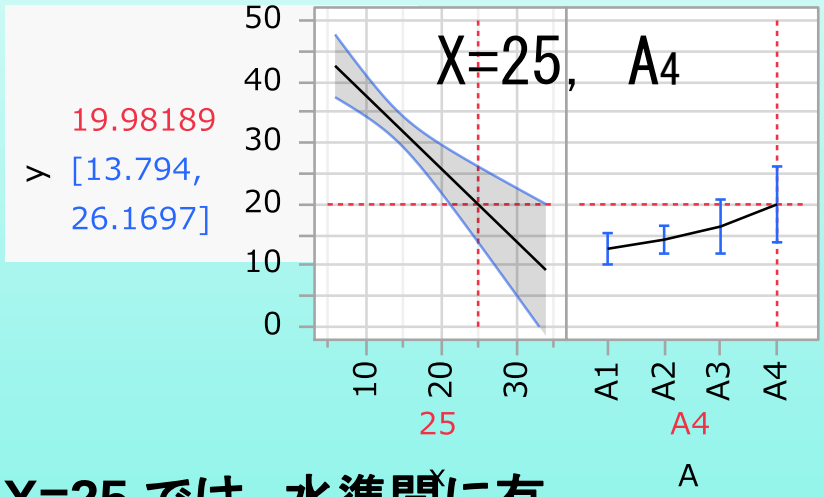
A4：y^A4 = 49.7416 – 1.1904 × 21.47619 = 24.1766

予測プロファイルの活用



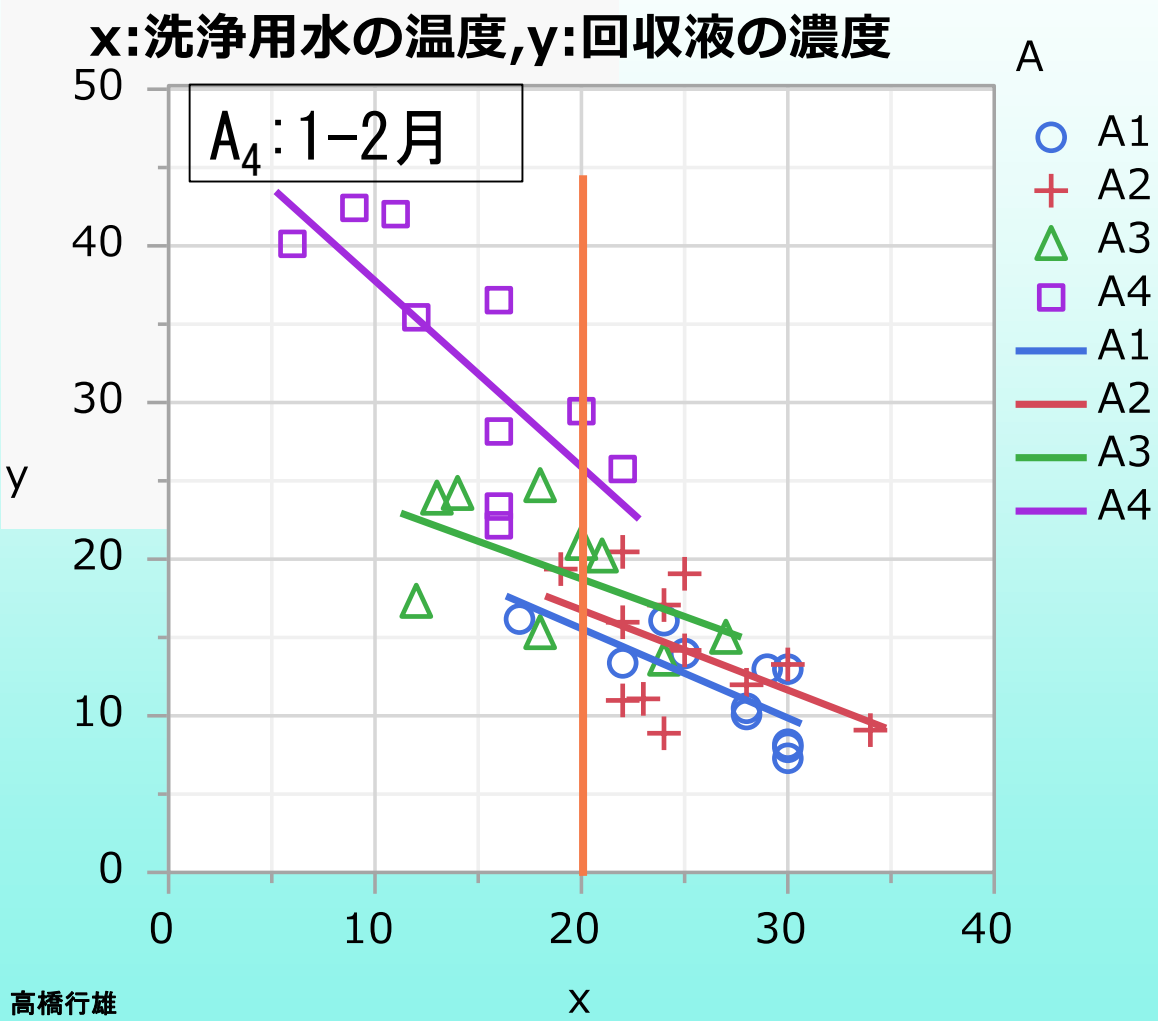




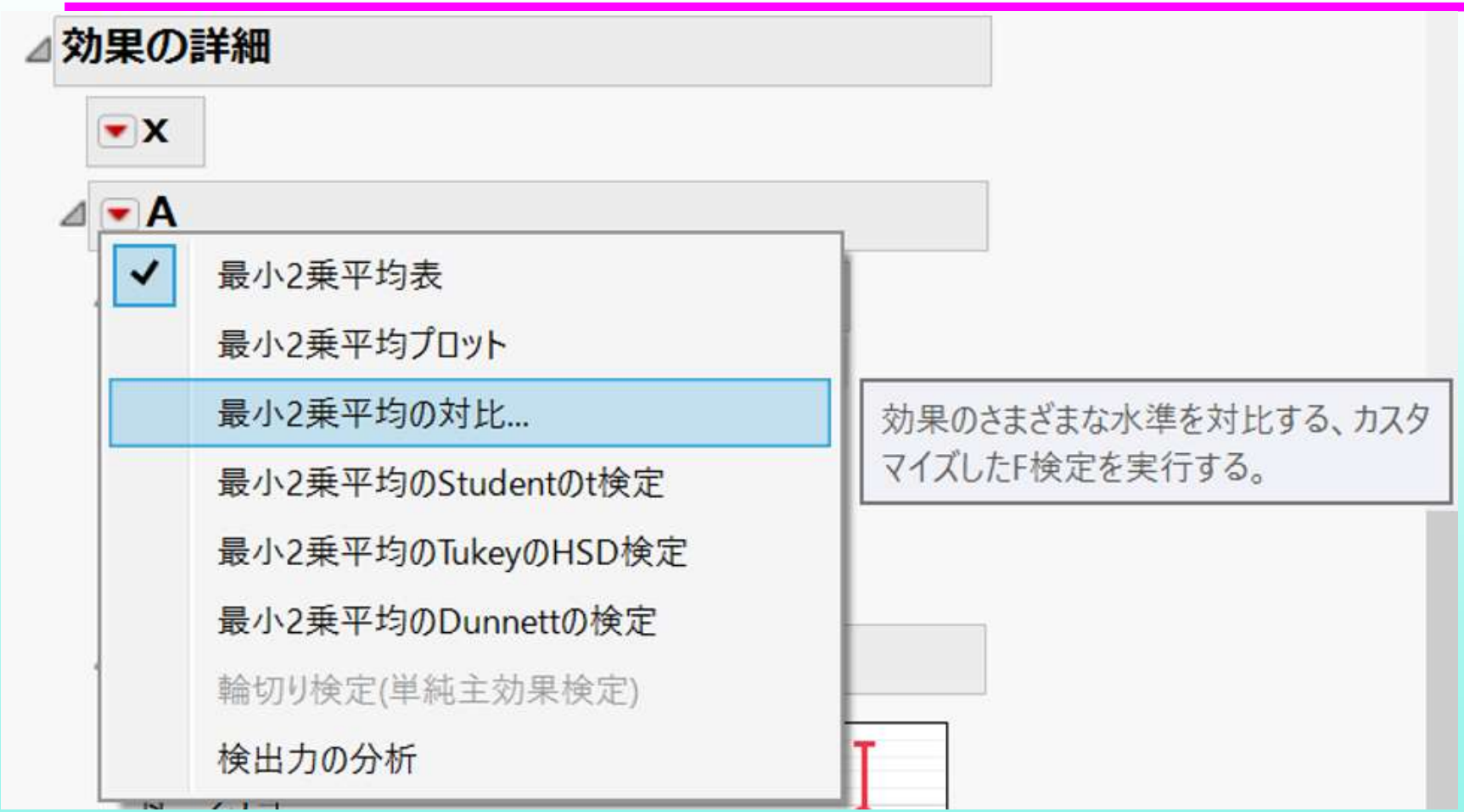


X=25 では，水準間に有意な差はなくなっている．

20 ℃で群間比較をしたい

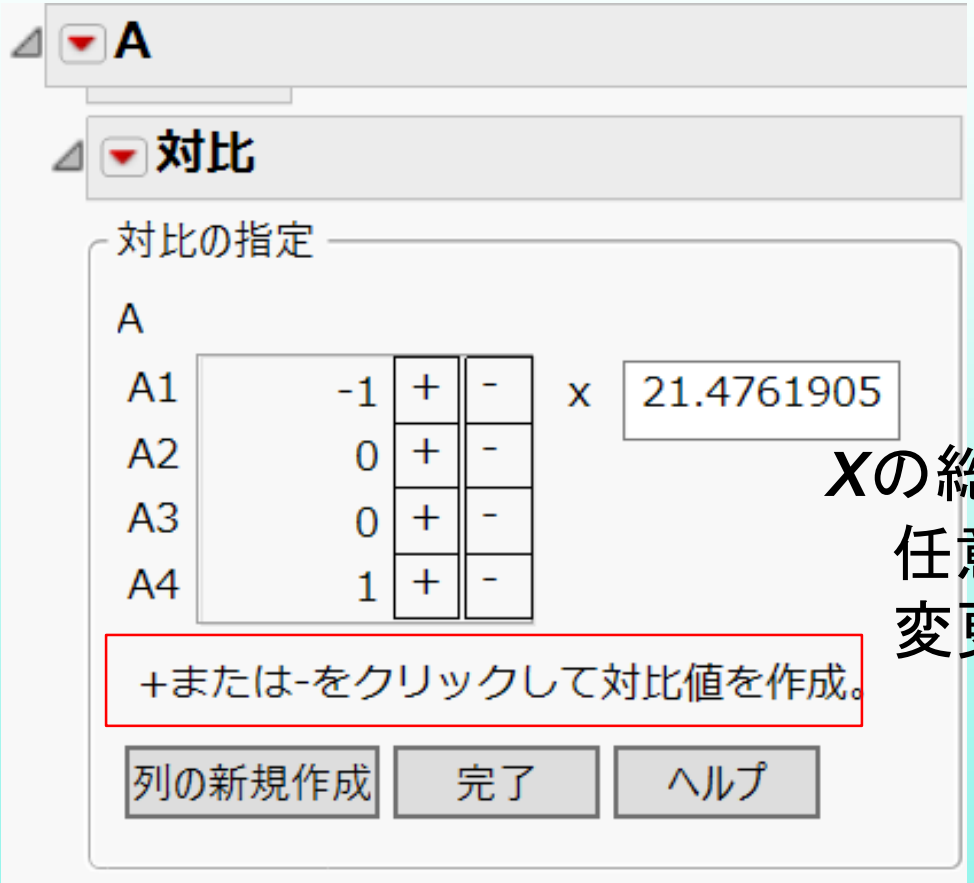


最小2乗平均の対比


効果の詳細
X
A
最小2乗平均表
最小2乗平均プロット
最小2乗平均の対比...
最小2乗平均のStudentのt検定
最小2乗平均のTukeyのHSD検定
最小2乗平均のDunnettの検定
輪切り検定(単純主効果検定)
検出力の分析
効果のさまざまな水準を対比する、カスタマイズしたF検定を実行する。

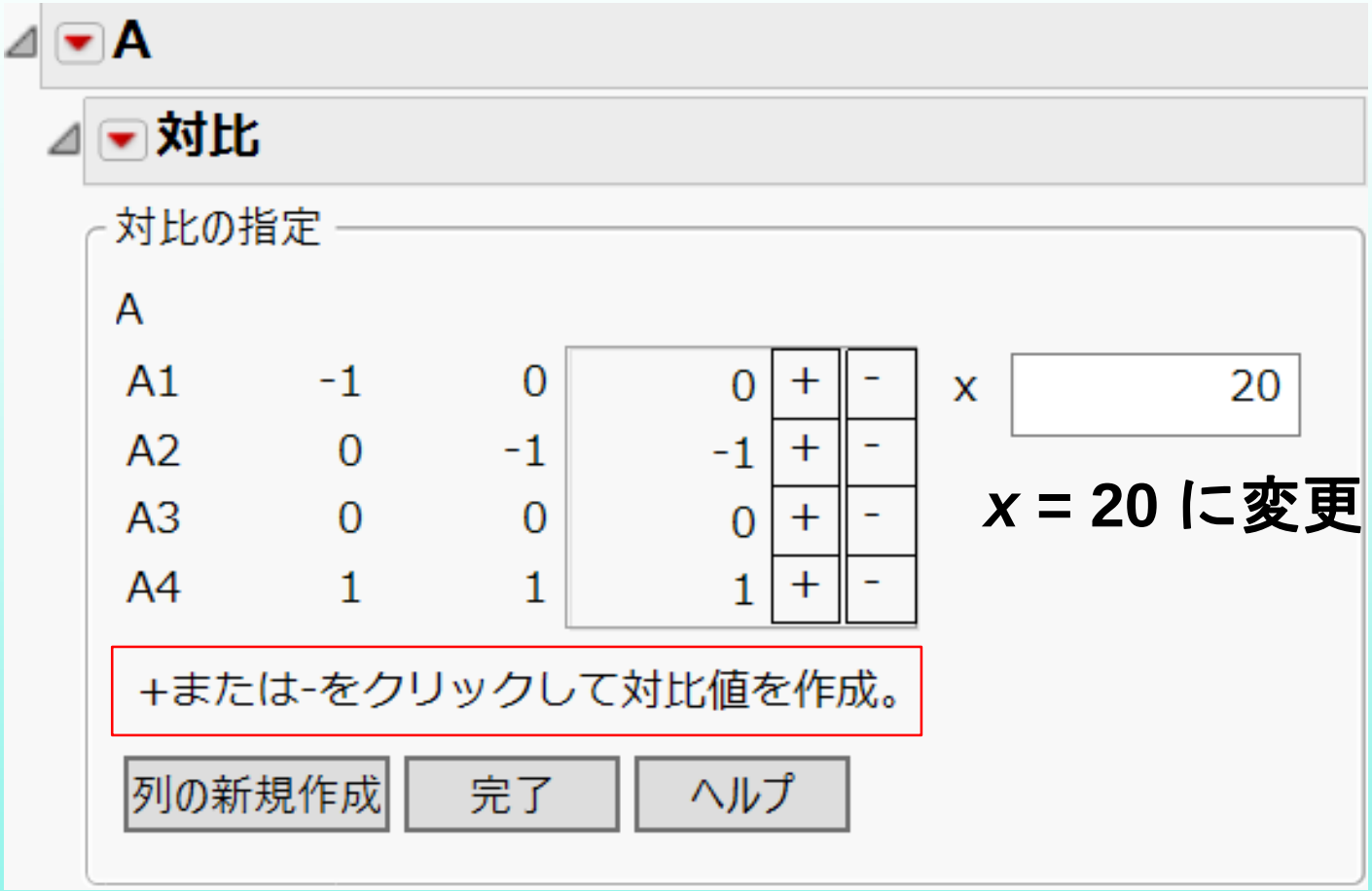
対比による水準間の比較

A4－A1
A4－A2
A4－A3
差の検定
をしたい



*X*の総平均
任意の値に
変更可能

X を 20に変更 対比の設定



***x* = 20 に変更**

$X = 20$ における水準間の差

対比

検定の詳細

A1	-1	0	0
A2	0	-1	0
A3	0	0	-1
A4	1	1	1
推定値	10.343	9.1496	7.162
標準誤差	2.9532	2.6113	2.3417
t値	3.5022	3.5039	3.0585
p値(Prob>|t|)	0.0013	0.0013	0.0043
平方和	183.64	183.81	140.05

すべて有意な差

平方和	分子自由度	分母自由度	F値	p値(Prob>F)
252.8	3	34	5.6281	0.0030 *

x	20

Excel による確認

全水準の推定値					推定値＝			
項	推定値				A_i 切片＋A_i 傾き×x			
切片	33.045							
x	-0.689					x ＝	20	
A[A1]	-6.028	+	33.045	=	27.017	A1切片		
A[A2]	-6.016	+	33.045	=	27.029	A2 "		
A[A3]	-4.653	+	33.045	=	28.392	A3 "		
A[A4]	16.697	+	33.045	=	49.742	A4 "	推定値	差
x*A[A1]	0.117	+	-0.689	=	-0.571	A1傾き	15.591	**10.343**
x*A[A2]	0.176	+	-0.689	=	-0.512	A2 "	16.784	**9.150**
x*A[A3]	0.208	+	-0.689	=	-0.481	A3 "	18.772	**7.162**
x*A[A4]	-0.502	+	-0.689	=	-1.190	A4 "	25.934	規準

推定値	10.343	9.1496	7.162

Excel による95%信頼区間

表 13.12　季節 A_4 を基準とした 2 水準間の差の推定

No.	A	x_0	a_1	a_2	a_3	x	a_1x	a_2x	a_3x	y^	$Var(y$^$)$	SE	$L95\%$	$U95\%$
5	A_1-A_4	0	2	1	1	0	40	20	20	**-10.3426**	8.7211	2.9532	**-16.3441**	**-4.3411**
6	A_2-A_4	0	1	2	1	0	20	40	20	**-9.1496**	6.8189	2.6113	**-14.4564**	**-3.8428**
7	A_3-A_4	0	1	1	2	0	20	20	40	**-7.1620**	5.4834	2.3417	**-11.9208**	**-2.4031**

詳しくは，続・高橋セミナー 第9回 13章

最小2乗平均の謎を予測プロファイルで解く

https://www.yukms.com/biostat/takahasi2/rec/009-13.htm

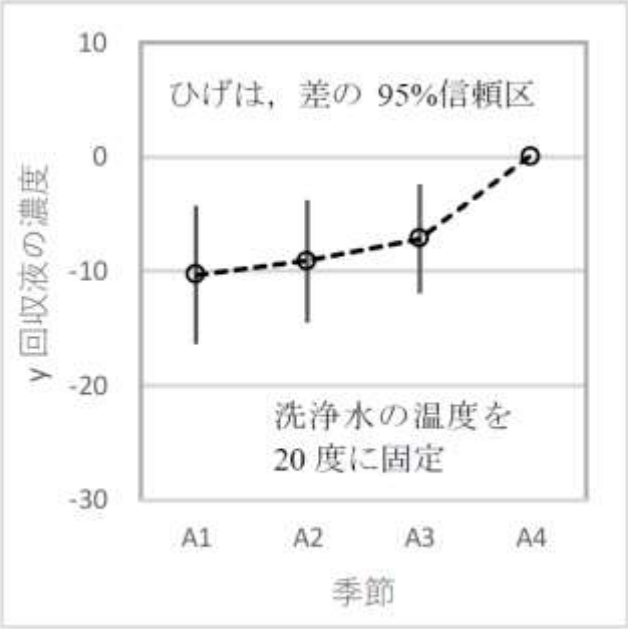


図 13.5　Excel の「折れ線」による差の予測プロファイル

最小2乗平均・予測プロファイル

- 最小2乗平均は，質的因子の水準平均の算出に際して，他の（すべての）変数の平均を考慮した“調整平均”である.
- 予測プロファイルは，（すべての）質的因子の水準平均，（すべての）量的因子の一定範囲に対し，予測値と95%信頼区間を，他の変数の平均などで調整したグラフ表示である.
- ディフォルトで，質的因子の水準については，最小2乗平均と95%信頼区間が標示される.

具体例としては

- 質的因子が1つ，量的変数が1つで，交互作用を含む場合が，最小2乗平均と予測プロファイルを説明する事例として適していることを実感した.
- 複数の量的因子のみの場合では，最小2乗平均は現れない.
- 複数の質的因子のみでは，他の質的因子のダミー変数の平均で調整していることが見えない.

文献

- 竹内啓ら（1989），統計学辞典，1135，東洋経済.
- 奥野忠一，久米均，芳賀敏郎，吉沢正著（1981），多変量解析法 改訂版，112-123，日科技連出版社.
- 高橋行雄・大橋靖雄・芳賀敏郎（1989），SAS による実験データの解析，307-333，東京大学出版会.
- 魚住龍史（2014），LS-Means 再考－GLM とPLM によるモデル推定後のプロセス－，SAS ユーザー総会論文集：449-463，https://www.sas.com/content/dam/SAS/ja_jp/doc/event/sas-user-groups/sugj2014.pdf
- 高橋行雄（2020），最小2乗平均の謎を予測プロファイルで解く，https://www.yukms.com/biostat/takahasi2/rec/009-13.htm

ご清聴ありがとうございました