

第 12 回 続高橋セミナー  
層別因子を含む探索的な回帰分析入門  
2024 年 2 月 11 日

第 8 章 交絡変数と共変量を含む 2 群比較

医薬品あるいは農薬などを含む化合物の安全性の評価の一環として、実験動物を用いた反復投与による各種の毒性試験が行われている。竹内ら (1989), 「統計学辞典, IV 分野別応用, 4.3 毒性評価」に示されているように実験手順は、細かく規定されており、用いる統計的方法についても分野により細かな差異はあるもののおおむね規準化されている。反復投与試験の多くは、実験動物の成長期に行われ、化合物の実験動物に与える影響を多面的に評価するために数多くの変数が測定される。このために、化合物の直接的な作用と推定される反応だけではなく、交絡変数が間接的に作用する反応も拾い出してしまう。間接的な作用であることを識別するために、反応と交絡変数の比などを求めて、交絡変数の影響を取り除く工夫もされている。交絡変数の影響を統計的に取り除く方法として共分散分析が知られており、そのため交絡変数は、解析に際しては、“共変量”として混同されがちである。そこで、多面的な作用が内在する交絡変数を含めた探索的な解析法について例示する。

第 8 章 目 次

8.	交絡変数と共変量を含む 2 群比較	243
8.1.	体重の増加抑制は総摂餌量の影響だけか	243
	体重の増加抑制, 2 群間の $t$ 検定を使った検討, 体重の増加量 および増加率, 前後差か増加率かの判断基準, 交絡変数・総摂餌量 の影響, 共変量・前値の影響, 共変量・胎仔数の影響	
8.2.	Excel による複数の共変量を含む探索的な回帰分析	250
	前後差に対する共変量としての前値, 複数の共変量との交互作用, 複数の共変量を考慮した $2 \times 2$ 格子状の散布図	
	次ページに続く	

8.3. JMP による複数の共変量を含む探索的な回帰分析-----	255
ステップワイズ法による変数選択,	
「標準最小 2 乗法」による予測プロファイルの活用	
8.4. 交絡変数の多彩な影響への対応-----	260
投与終了時の体重は交絡変数, 解析手法の選択のための決定樹,	
1) $X$ と $Y$ に相関があるか, 2) 2 群の回帰直線が同一線上か,	
3) 2 本の直線の傾きは同じか, 4) 回帰直線が原点を通るか,	
5) 2 本の直線が原点を通らず傾きが同じでない場合の場合	
8.5. 体重を交絡変数とする回帰分析-----	268
群により回帰直線の傾きが異なる場合, 体重の総平均での臓器	
重量の比較, 予測プロファイル作成, 回帰直線が同一直線上と	
見なせる場合, 2 本の回帰直線が原点を通る場合,	
交絡変数を常に念頭にした現実的な対応	
文献索引, 索引, 解析用ファイル一覧 -----	(281)

## 第 12 回 続・高橋セミナー

### 層別因子を含む探索的な回帰分析入門

#### 目次 (全章)

はじめに -----	1
1. 層別因子を含む各種の回帰分析の実際 -----	7
2. デザイン行列を活用した 1 因子実験データの解析 -----	55
3. 繰り返しが不揃いの 2 因子実験データの解析 -----	89
4. 欠測値がある直交表の線形モデルによる解析 -----	121
5. デザイン行列を用いた回帰分析の基礎 -----	155
6. 伝統的な共分散分析からの脱却 -----	195
7. 共変量を含む 3 因子実験データの探索的解析 -----	219
<b>8. 交絡変数と共変量を含む 2 群比較 -----</b>	<b>243</b>
9. 前後差の前値は常に共変量 -----	281
10. 層別因子を含むロジスティック曲線のあてはめ -----	327
11. 各種のシグモイド曲線を用いた逆推定 -----	367
12. ミカエリス・メンテン式をめぐる新たな統計解析 -----	405
13. ロジスティック曲線のさらなる活用 -----	455
文献, 文献索引, 索引, 解析用ファイル一覧 -----	489

## 8. 交絡変数と共変量を含む 2 群比較

医薬品あるいは農薬などを含む化合物の安全性の評価の一環として、実験動物を用いた反復投与による各種の毒性試験が行われている。竹内ら (1989), 「統計学辞典, IV 分野別応用, 4.3 毒性評価」に示されているように実験手順は、細かく規定されており、用いる統計的方法についても分野により細かな差異はあるもののおおむね規準化されている。反復投与試験の多くは、実験動物の成長期に行われ、化合物の実験動物に与える影響を多面的に評価するために数多くの変数が測定される。このために、化合物の直接的な作用と推定される反応だけではなく、交絡変数が間接的に作用する反応も拾い出してしまう。間接的な作用であることを識別するために、反応と交絡変数の比などを求めて、交絡変数の影響を取り除く工夫もされている。交絡変数の影響を統計的に取り除く方法として共分散分析が知られており、そのため交絡変数は、解析に際しては、“共変量”として混同されがちである。そこで、多面的な作用が内在する交絡変数を含めた探索的な解析法について例示する。

### 8.1. 体重の増加抑制は総摂餌量の影響だけか

#### 体重の増加抑制

ある化合物に胎児に対する催奇形性が有るか否かの確認のために、動物を対象にした生殖発生毒性試験が行われている。表 8.1 に示すのは、吉村編著 (1987), 「毒性・薬効データの統計解析, 第 3.3 節 体重は増加量でみるべきか」で取り上げられているデータである。これは、ラットを対象に妊娠 7 日から 18 日まで化合物を反復投与し、胎仔に対する各種の影響を検討するために得られたデータの一部である。

このデータは、妊娠 7 日に体重を測定した後に、化合物の反復投与を開始し、妊娠 18 日に体重と総摂餌量を測定した結果である。さらに、出生した仔数が、示されている。化合物の反復投与によって、妊娠ラットの体重増加に何らかの抑制がみられるならば、出生仔数に影響を与えるかもしれない。また、データとして得られていないが、胎仔数が妊娠ラットの体重増加に影響を与えているのかも知れない。化合物の反復投与が体重増加の抑制という影響をもたらしたとすれば、食欲行動抑制のために体重増加の抑制が起きたのか、餌を食べたに

表 8.1 妊娠ラットの体重増加データ [吉村編著 (1987), 表 3-6]

番号	対照群 ( $A_1$ )							処置群 ( $A_2$ )						
	——— 体重g ———				総摂餌量	飼料効率	出生仔数	——— 体重g ———				総摂餌量	飼料効率	出生仔数
	7日	18日	差	率%				7日	18日	差	率			
	$x_1$	$y_1$	$y_1 - x_1$	$\Delta/x_1$	$x_2$	$\Delta/x_2$	$y_2/x_3$	$x_1$	$y_1$	$y_1 - x_1$	$\Delta/x_1$	$x_2$	$\Delta/x_2$	$y_2/x_3$
1	248.7	318.1	69.4	27.9	203.3	0.341	13	257.9	309.1	51.2	19.9	199.2	0.257	13
2	261.3	324.0	62.7	24.0	223.0	0.281	13	260.7	305.4	44.7	17.1	167.9	0.266	13
3	265.6	339.9	74.3	28.0	244.9	0.303	12	277.0	331.5	54.5	19.7	185.9	0.293	14
4	284.0	356.1	72.1	25.4	236.5	0.305	14	283.8	323.2	39.4	13.9	195.6	0.201	13
5	254.9	328.6	73.7	28.9	204.5	0.360	15	281.2	332.1	50.9	18.1	203.6	0.250	12
6	267.4	338.1	70.7	26.4	223.0	0.317	15	292.3	339.6	47.3	16.2	196.7	0.240	13
7	271.0	349.3	78.3	28.9	228.5	0.343	14	270.1	330.2	60.1	22.3	207.7	0.289	13
8	257.4	308.3	50.9	19.8	221.4	0.230	8	265.8	318.7	52.9	19.9	186.3	0.284	11
9	282.2	340.0	57.8	20.5	221.1	0.261	17	269.9	317.1	47.2	17.5	184.1	0.256	11
10	268.3	335.4	67.1	25.0	210.4	0.319	14	269.8	328.7	58.9	21.8	198.8	0.296	16
11	236.8	296.0	59.2	25.0	191.1	0.310	12	267.1	335.4	68.3	25.6	219.3	0.311	15
12	288.0	367.2	79.2	27.5	258.4	0.307	13	259.4	332.0	72.6	28.0	217.0	0.335	14
13	246.9	313.4	66.5	26.9	197.8	0.336	16	246.3	309.9	63.6	25.8	188.6	0.337	14
14	236.6	304.5	67.9	28.7	215.4	0.315	12	230.1	282.5	52.4	22.8	164.5	0.319	12
15	254.8	342.7	87.9	34.5	232.0	0.379	16	223.6	275.4	51.8	23.2	166.7	0.311	13
16	251.5	320.6	69.1	27.5	225.1	0.307	10	277.1	329.5	52.4	18.9	199.3	0.263	14
17	270.8	345.9	75.1	27.7	227.9	0.330	15	269.1	329.1	60.0	22.3	209.9	0.286	14
18	254.4	326.6	72.2	28.4	215.5	0.335	14	231.4	295.7	64.3	27.8	171.9	0.374	15
19	240.2	310.1	69.9	29.1	206.7	0.338	14	239.0	306.5	67.5	28.2	180.8	0.373	15
20	245.8	310.7	64.9	26.4	216.0	0.300	12	257.4	316.0	58.6	22.8	193.0	0.304	14
平均	259.3	328.8	69.4	26.8	220.1	0.316	13.5	261.5	317.4	55.9	21.6	191.8	0.292	13.5
SD	15.2	18.8	8.18	3.2	16.00	0.033	2.1	18.8	17.6	8.57	4.1	16.07	0.043	1.3

もかわらず化合物の投与による何らかの作用によって体重増加が抑制されたのか，を明らかにする必要がある。

図 8.1 に示すように，妊娠 7 日の体重は，化合物の投与前の測定値であり，化合物の影響を受けない前値であるが，総摂餌量は，化合物の投与によって直接的に抑制されるかも知れないし，体重増加の抑制による影響を受けることも考えられる．総摂餌量は，妊娠 18 日の体重に対する交絡変数であり，化合物の投与期間中の直接的な摂餌の抑制，あるいは体重の抑制による間接的に相互の影響を受ける．化合物の投与により，胎仔が発育不全などの原因で死

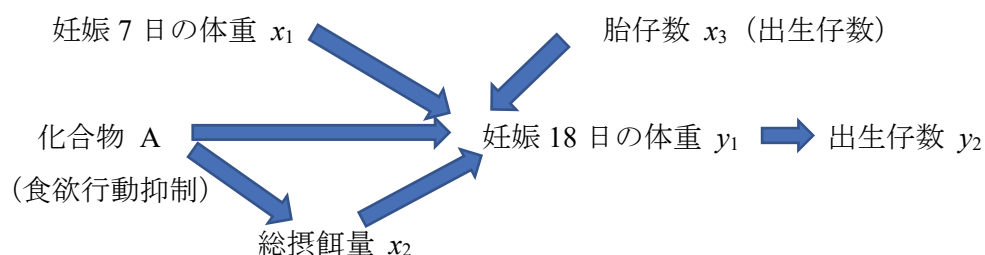


図 8.1 妊娠 18 日の体重に影響を与える交絡変数と共変量

亡していないことを確認するため、出生した仔数も測定されている。なお、出生仔に対しては、さらに、催奇形性が有るか詳細に検討される。

## 2 群間の $t$ 検定を使った検討

今から 30 年以上前の 1980 年代の交絡変数を含むデータに対する統計解析は、2 群間の  $t$  検定を段階的に適用して考察することが一般的であった。計算には、各群の偏差平方和が必要であり、Excel の DevSq() 関数を使うことにより、元のデータから直接求めることができる。対照群  $A_1$  の妊娠 7 日の体重  $x_1$  の偏差平方和  $S_{x_1}^{(1)}$  は、

$$\left. \begin{aligned} S_{x_1}^{(1)} &= \sum_{i=1}^{20} (x_{1,i} - \bar{x}_1)^2 \\ &= \text{DevSq}(248.7, 261.3, \dots, 245.8) = 4380.40 \end{aligned} \right\} \quad (8.1)$$

に示すように、引数に DevSq( $A_1$  群の  $x_1$  の範囲) と設定することにより求められる。表 8.2 に示すように 2 群間の  $t$  検定は、平均値の差に対して差の  $SE$  で割った統計量であり、それぞれの群の偏差平方和を用いて容易に計算できる。なお、Excel には、 $t$  検定の  $p$  値を直接計算する T.test() 関数もあるが、段階的な計算手順を示すために用いなかった。

表 8.2 対象群と処置群の差の  $t$  検定

		対照群 ( $A_1$ )		処置群 ( $A_2$ )		$t$ 検定				
		平均	平方和 $S^{(1)}$	平均	平方和 $S^{(2)}$	群間差	$SE$	$t$	$p$	
$x_1$	7日の体重	259.33	4380.40	261.45	6739.09	2.12	5.4094	0.3919	0.6973	NS
$y_1$	18日の体重	328.78	6715.10	317.38	5897.95	-11.40	5.7613	-1.9779	0.0552	+
$y_1 - x_1$	前後差	69.45	1271.55	55.93	1394.48	-13.52	2.6487	-5.1024	0.0000	***
$(y_1 - x_1)/x_1\%$	率%	26.82	189.78	21.58	317.72	-5.24	1.1557	-4.5369	0.0001	***
$x_2$	総摂餌量	220.13	4866.00	191.84	4908.93	-28.29	5.0718	-5.5769	0.0000	***
$(y_1 - x_1)/x_2$	飼料効率	0.3159	0.0211	0.2923	0.0355	-0.0236	0.0122	-1.9308	0.0610	+
$y_2$	出生仔数	13.45	84.95	13.45	32.95	0.00	0.5570	0.0000	1.0000	NS
偏差平方和 = DevSq ( データの範囲 )						$n = 2 \times 20$		$t = \text{群間差}/SE$		
標準誤差 $SE = \text{Sqrt}((S^{(1)} + S^{(2)})/38) * (1/20 + 1/20)$								$p = \text{T.dist.2T}( t , 38)$		

妊娠 7 日の体重  $x_1$  は、対照群  $A_1$  と処置群  $A_2$  に有意な差 ( $p = 0.6973$ ) ではない。妊娠 18 日の体重  $y_1$  は、対照群  $A_1$  に比べて処置群  $A_2$  は -11.40 と体重の増加が少ないが、 $t$  検定では微妙 ( $p = 0.0552$ ) な結果である。体重の前後差 ( $y_1 - x_1$ ) および増加率  $(y_1 - x_1)/x_1\%$  にすると  $SE$  が減少し、統計的にも明らかな差 ( $p < 0.0001$ ) となっている。体重の前後差 ( $y_1 - x_1$ ) に対して総摂餌量  $x_2$  で割った飼料効率  $(y_1 - x_1)/x_2$  で比較すると、処置群がわずかに小さいが、微妙 ( $p = 0.0610$ ) な結果である。出生仔数  $y_2$  には、差が全くないので化合物の妊娠中の胎仔の生死に関与する作用はないといえる。

体重の増加抑制が、化合物の直接作用なのか、総摂餌量  $x_2$  の抑制による間接的な作用なのか、化合物が摂餌量の食欲行動抑制をしたことによる体重の増加抑制が起きたのかは、これ

らの $t$ 検定の結果からでは、はっきりしたことはわからないが、飼料効率に大きな差がないので、「体重の増加抑制は、主として餌を食べないことが影響しているようである」と吉村編著（1987）では述べられている。

### 体重の増加量および増加率

体重が増加するような場合に、前値 $x_1$ の影響を除くために前後差 $(y_1 - x_1)$ による群間比較が一般的に行われている。この事例は、交絡変数としての総摂餌量 $x_2$ 、共変量としての妊娠7日の体重 $x_1$ を同時に含む場合に、どのような解析を行なったら良いのであろうか。

複数の交絡変数と共変量がある場合には、1変数ごとに散布図を作成し、相互の関連を概観することが欠かせない。図8.2左に示すように、妊娠7日の体重 $x_1$ と妊娠18日の体重 $y_1$ には強い相関関係があり、対照群 $A_1$ に比較し処置群 $A_2$ は、妊娠7日の体重が大きい場合に体重の増加抑制がより大きくなっているように思われる。図8.2右の体重の前後差 $(y_1 - x_1)$ をY軸とした散布図からは、より明確な増加抑制が示されている。表8.2に示した2群間の $t$ 検定による検討では、対照群 $A_1$ の前後差69.45gに対し処置群 $A_2$ は55.93gであり、 $t=5.1024$ と明らかな体重抑制がある。散布図上の回帰直線からは、7日の体重 $x_1$ が大きいほど体重の前後差が拡大しているように思われる。

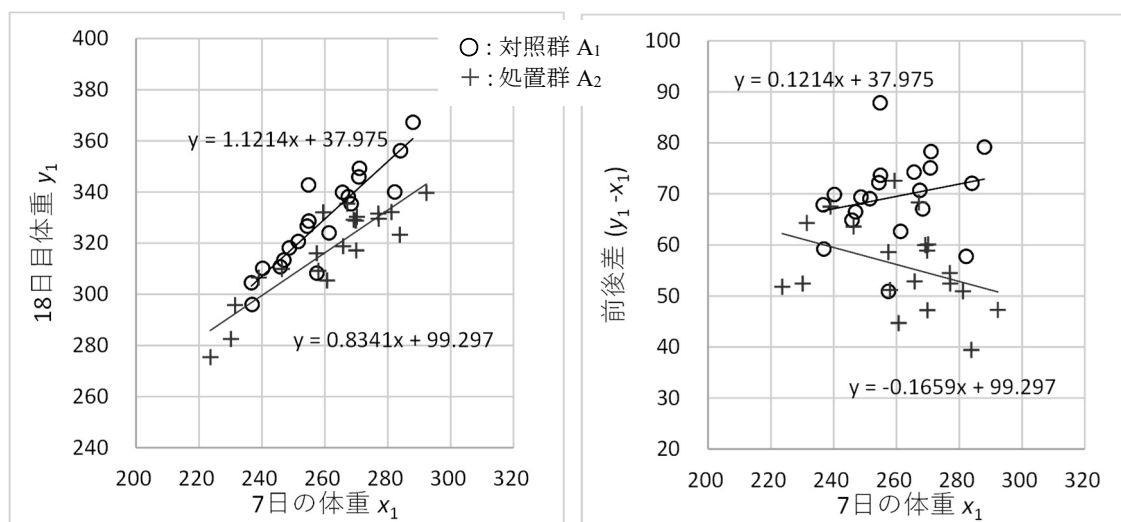


図8.2 妊娠7日の体重 $x_1$ と妊娠18日の体重 $y_1$ および前後差 $(y_1 - x_1)$ の関連

層別散布図の作成は、Excelシート上で対照群のX軸の変数、Y軸の変数を選択し散布図を作りマークを「○」印に変更する。「データの選択」により処置群の変数を「追加」しマークを「+」印に変更する。それぞれの群を選択し、「近似曲線の追加」で「線形近似」を選択し、回帰式を重ね書きし、書式を整えた結果である。

体重の前後差 $(y_1 - x_1)$ に対する $t$ 検定では、図8.2右に示した前値 $x_1$ の体重による前後差の大きさの違いを無視して“平均的”な前後差についての結果であり、そのために、図8.2右

から推測される前値  $x_1$  の体重が大きいほど前後差が拡大、いわゆる交互作用が存在している可能性を無視していいのかとの疑問が生ずる。

前値  $x_1$  と後値  $y_1$  がある場合に、前後差  $(y_1 - x_1)$  での群間比較、あるいは、増加率  $100 \times (y_1 - x_1) / x_1$  での群間比較か、どちらを選ぶか、両方なのか、悩ましい。表 8.2 に示したように群間で  $t$  検定の結果を示すだけであれば、両方の結果を示すことは簡単である。図 8.3 に示す増加率での散布図では、図 8.2 右に示した前後差に比べ、右下がりになるが、全体の位置関係はほとんど同じである。

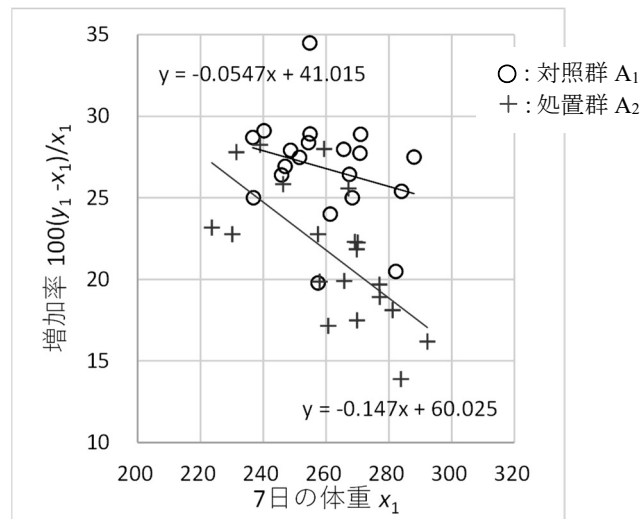


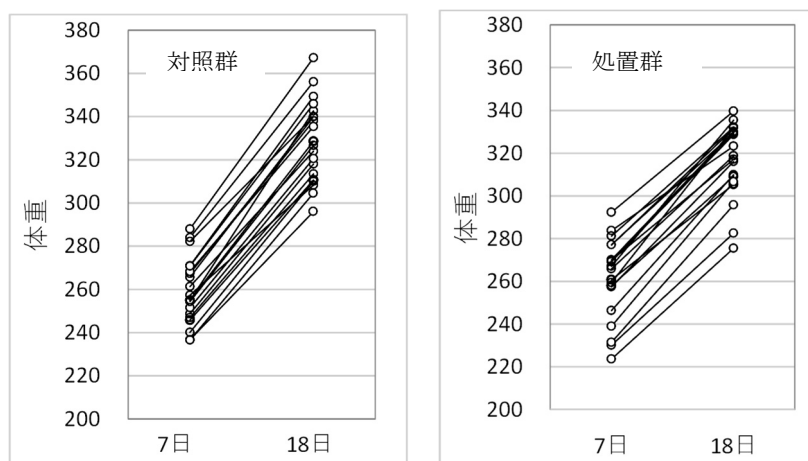
図 8.3 妊娠 7 日の体重  $x_1$  と投与後 18 日の体重  $y_1$  の増加率

### 前後差か増加率かの判断基準

前後差か、増加率か、どちらが望ましいのかについての選択は、図 8.4 に示すように前値  $x_1$  と後値  $y_1$  について個体ごとに線で結んだ線グラフを作成して判断できる。ほぼ平行であれば、前後差  $(y_1 - x_1)$  を選択すればよいのだが、前値  $x_1$  が広い範囲にあり、扇型に広がるような場合には、増加率  $(y_1 - x_1) / x_1$  を用いるべきである。なお、前値が相対的に狭い範囲にあり、扇型に広がるような場合には、増加率  $(y_1 - x_1) / x_1$  を取ることは、前値の誤差の変動が拡大するのでまったく推奨できないので、後値  $y_1$  そのものを使うべきである。

### 交絡変数・総摂餌量の影響

図 8.5 に示すのは、総摂餌量  $x_2$  と妊娠 18 日の体重  $y_1$  の関係、および、前後差  $(y_1 - x_1)$  との関係である。図 8.5 左に示すように、処置群 A2 は、対照群 A1 に比べ総摂餌量  $x_2$  が少ない方にシフトし、妊娠 18 日の体重  $y_1$  も下方にシフトしているにもかかわらず、同じ総摂餌量  $x_2$  で比較すると 18 日の体重  $y_1$  は、逆に大きくなっている。この原因は、図 8.5 右に示すように体重の前後差  $(y_1 - x_1)$  では、同じ総摂餌量  $x_2$  で比較すると処置群 A2 は、前後差  $(y_1 - x_1)$  が小

図 8.4 妊娠 7 日の体重  $x_1$  と投与後 18 日の体重  $y_1$  の線グラフ

Excel の折れ線グラフで、前値と後値の範囲を選択すると 2 本の折れ線が表示される。「データの選択」オプションで「行/列の切り替え」を選択すると 20 本の直線が表示されるので、書式を整えて作成した結果である。

さくなっている。このことから、妊娠 18 日の体重  $y_1$  に対して、体重の前値  $x_1$  と総摂取量  $x_2$  が影響を与えていることになる。

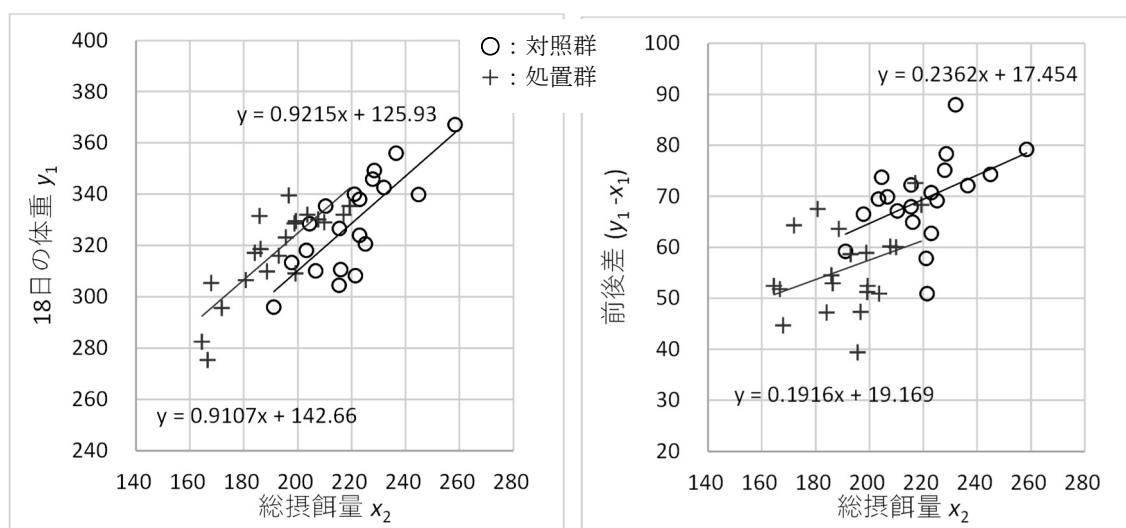


図 8.5 総摂取量と妊娠 18 日の体重および前後差の関連

### 共変量・前値の影響

図 8.6 左に示すように妊娠 7 日の体重  $x_1$  に対する総摂取量  $x_2$  は、処置群で体重に関連なく減少している。図 8.6 右に示す体重の前後差  $(y_1 - x_1)$  を総摂取量で割った摂餌効率  $(y_1 - x_1)/x_2$  では、処置群の体重が重い場合に摂餌効率が悪くなっているように思われる。以上の結果から、化合物の作用は、初期の体重  $x_1$  が大きい場合に餌の食べる量が減少し、その結果として体重の増加抑制が起きていると想定される。この実験データからでは、体重増加を抑制するような別の原因の探求はできないが、他の反復投与による毒性試験の結果から原因が示唆されるかも知れない。



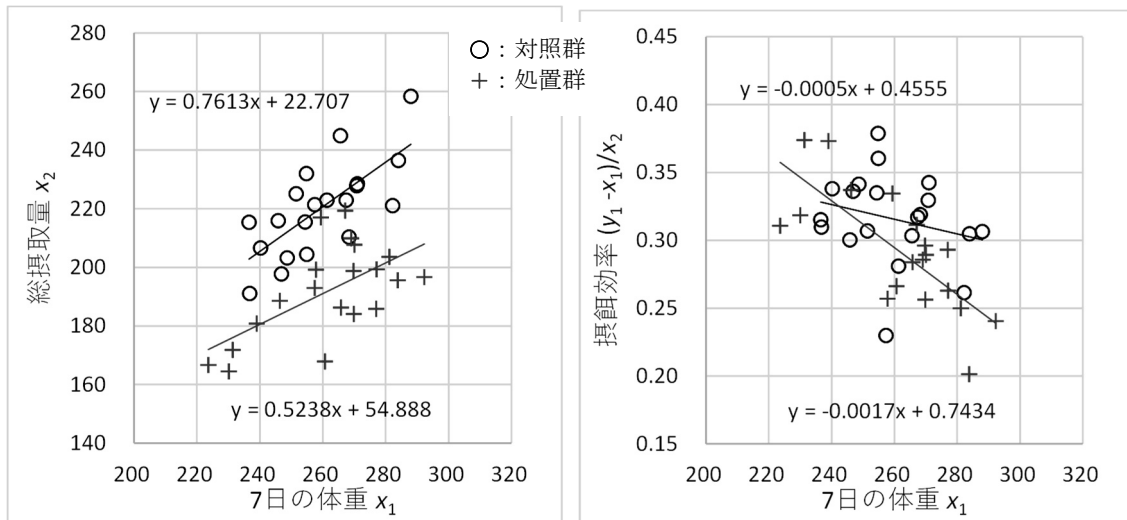
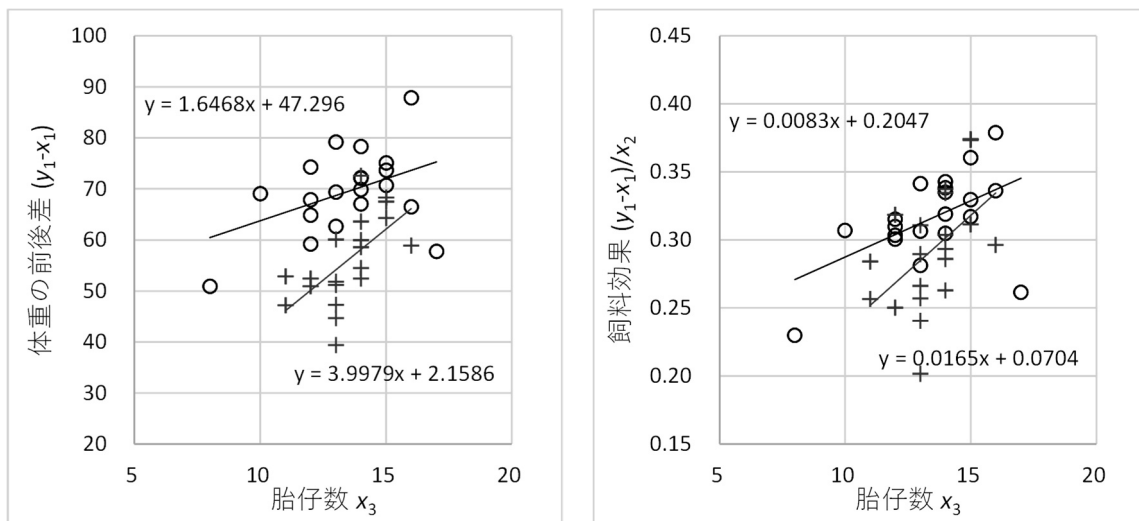


図 8.6 投与前 7 日の体重と総摂餌量および摂取効率

### 共変量・胎仔数の影響

出生仔数  $y_2$  には、表 8.2 で示したように群間で差が無いことから妊娠 7 日の胎仔数  $x_3$  とみなすことにする．図 8.7 に示すように、胎仔数  $x_3$  が体重の前後差  $(y_1 - x_1)$  に対してどのような影響があるか、散布図で確認する．胎仔数  $x_3$  の増加に伴い体重の前後差  $(y_1 - x_1)$  が増加していることが確認される．

図 8.7 胎児数  $x_3$  : ( $y_2$ ) と体重の前後差  $(y_1 - x_1)$  および飼料効率  $(y_1 - x_1)/x_2$  の関係

このように散布図に回帰直線を重ね書きすることにより、2 群間の  $t$  検定の適用だけでは見えなかった関係を浮き彫りにすることができる．毒性試験で得られた膨大な測定変数に対して散布図を描いて考察することは現実的ではないが、何か疑問が生じた場合にその原因を掘り下げる時の最初の一步である．摂餌効率  $(y_1 - x_1)/x_2$  に対する 2 群間の  $t$  検定の結果は、 $p = 0.0610$  と微妙な結果であり、図 8.6 右に示すように、妊娠 7 日の体重  $x_1$  との交互作用が示唆されている．

## 8.2. Excel による複数の共変量を含む探索的な回帰分析

## 前後差に対する共変量としての前値

2 群間の  $t$  検定を使った群間比較では、処置群  $A_2$  は、対照群  $A_1$  に対して体重の前後差  $(y_1 - x_1)$  から明らかな体重の増加抑制がある。しかし、飼料効率  $(y_1 - x_1)/x_2$  には明らかな差が2群間にないことから、「体重の増加抑制は、主として餌を食べないことが影響しているようである」との結論で良いのだろうか。第8.1節で示したように各種の層別散布図により検討を重ねても、なかなかすっきりとした結果が見えてこない。

そこで、対照群  $A_1$  に  $a=1$ 、処置群  $A_2$  に対し  $a=-1$  とする  $(1, -1)$  対比型ダミー変数を与え、前値の体重  $x_1$  (妊娠7日の体重)、総摂餌量  $x_2$ 、胎仔数  $x_3$  (出生仔数  $y_2$  の読み替え) を共変量とする探索的な回帰分析を行なう。回帰分析に際しては、体重の後値  $y_1$  (妊娠18日の体重) についてのみを取り上げ、前後差  $(y_1 - x_1)$  については扱わない。これは、前後差  $(y_1 - x_1)$  に対し、右辺に共変量として前値  $x_1$  を含めることにより、左辺の  $x_1$  を右辺に移し、整理することにより、

$$\left. \begin{aligned} y_1 - x_1 &= \theta_0 + \theta_1 a + \theta_2 x_1 + \theta_3 x_2 + \theta_4 x_3 + \varepsilon \\ y_1 &= \theta_0 + \theta_1 a + (\theta_2 x_1 + x_1) + \theta_3 x_2 + \theta_4 x_3 + \varepsilon \\ &= \theta_0 + \theta_1 a + (\theta_2 + 1)x_1 + \theta_3 x_2 + \theta_4 x_3 + \varepsilon \end{aligned} \right\} \quad (8.2)$$

前値  $x_1$  のパラメータが  $(\theta_2 + 1)$  となるだけで、ダミー変数、他の共変量のパラメータの推定値が、後値  $y_1$  に対するパラメータの推定値と一致し、標準誤差も、 $p$  値も一致する。したがって、後値  $y_1$  に対する回帰分析の結果も、前後差  $(y_1 - x_1)$  による回帰分析の結果も共変量  $x_1$  に対する推定値が互いに  $\pm 1$  異なるだけで、どちらを用いても本質的な違いが無いからである。

## 複数の共変量との交互作用

交絡変数（以後、共変量と扱う）あるいは共変量を含む回帰分析を行なう際に、頭の痛い課題が存在する。それは、質的因子としての群  $A$  と共変量間の交互作用が存在するかの確認である。質的因子と交絡変数および共変量との交互作用に対して、それぞれの平均からの差とする中心化を行ない

$$y_1 = \theta_0 + \theta_1 a + \theta_2 x_1 + \theta_3 x_2 + \theta_4 x_3 + \theta_5 a(x_1 - \bar{x}_1) + \theta_6 a(x_2 - \bar{x}_2) + \theta_7 a(x_3 - \bar{x}_3) + \varepsilon \quad (8.3)$$

による検討を行う。なお、 $(1, -1)$  対比型ダミー変数  $a$  の期待平均は、0 なので中心化を必要としない。交互作用が無視できるのであれば、主効果モデルとして

$$y_1 = \theta_0 + \theta_1 a + \theta_2 x_1 + \theta_3 x_2 + \theta_4 x_3 + \varepsilon \quad (8.4)$$

による解析を行う。幸い質的因子は2水準なので交互作用モデルでは、Excel の回帰分析のパ

ラメータに対する  $t$  検定の  $p$  値が、分散分析の  $F$  検定の  $p$  値（タイプ III の平方和）に一致する。

表 8.3 に交互作用の検討のための Excel シートを示す。群 A については、(1, -1) 対比型ダミー変数とし、前値  $x_1$ 、総摂取量を  $x_2$ 、胎仔数  $x_3$  とし、それらの総平均が (260.390, 205.983, 13.450) として示されている。交互作用についての中心化は、

$$\left. \begin{aligned} ax'_1 &= a(x_1 - \bar{x}_1) = a(x_1 - 260.390) \\ ax'_2 &= a(x_2 - \bar{x}_2) = a(x_2 - 205.983) \\ ax'_3 &= a(x_3 - \bar{x}_3) = a(x_3 - 13.450) \end{aligned} \right\} \quad (8.5)$$

によって計算した結果が示されている。

表 8.3 交互作用の検討のためのデザイン行列

番号	群 A	切片 $x_0$	群 $a$	前値	総摂取量	胎	交互作用			後値	前後差
				7日 $x_1$	餌量 $x_2$	胎仔数 $x_3$	(中心化)			18日 $y_1$	
							$ax'_1$	$ax'_2$	$ax'_3$		
1	A <sub>1</sub>	1	1	248.7	203.3	13	-11.69	-2.68	-0.45	318.1	69.4
2	A <sub>1</sub>	1	1	261.3	223.0	13	0.91	17.02	-0.45	324.0	62.7
:										:	
19	A <sub>1</sub>	1	1	240.2	206.7	14	-20.19	0.72	0.55	310.1	69.9
20	A <sub>1</sub>	1	1	245.8	216.0	12	-14.59	10.02	-1.45	310.7	64.9
21	A <sub>2</sub>	1	-1	257.9	199.2	13	2.49	6.78	0.45	309.1	51.2
22	A <sub>2</sub>	1	-1	260.7	167.9	13	-0.31	38.08	0.45	305.4	44.7
:										:	
39	A <sub>2</sub>	1	-1	239.0	180.8	15	21.39	25.18	-1.55	306.5	67.5
40	A <sub>2</sub>	1	-1	257.4	193.0	14	2.99	12.98	-0.55	316.0	58.6
総平均=				260.390	205.983	13.450					
				$x'_1 = x_1 - 260.390$	$x'_2 = x_2 - 205.983$	$x'_3 = x_3 - 13.450$					

解析モデルに変数の積、あるいは、変数の 2 乗以上の変数を含める場合に用いる変数に対し、それらの平均値を差し引くことにより、多重共線性による行列計算の不具合を防ぐ必要がある。JMP の「モデルのあてはめ」でのデフォルトとなっている。この事例で、中心化を行わないと、ダミー変数の推定値が正しく推定されない。これは、Excel の行列計算でも、回帰分析でも、JMP の「モデルのあてはめ」でも、SAS/GLM でも同様に不可解な結果となる。

表 8.3 に示した (40 行×8 列) のデザイン行列を用い、表 8.4 左に Excel の回帰分析による交互作用モデルの分散分析表およびパラメータの推定結果を示す。交互作用に関する変数 ( $ax'_1$ ,  $ax'_2$ ,  $ax'_3$ ) について推定されたパラメータ (係数) に対する  $p$  値は、(0.7734, 0.3941, 0.5272) であり、交互作用はまったくないとの結果である。表 8.4 右に 3 つの共変量との交互作用を除いた主効果モデルについて分散分析表およびパラメータの推定結果を示す。どの共変量も明らかな統計的な有意差があり、符号がプラスなので、妊娠 18 日の体重  $y$  に対し、右上がりの関係となっている。対照群と処置群のダミー変数  $a$  については、 $\hat{\theta}_1 = -0.0052$  とゼロに近く  $p$  値も 1.0 に近く、妊娠 18 日の体重  $y_1$  に全く差が無いとの結果となっている。

表 8.4 交互作用および主効果モデルに対する分散分析表およびパラメータの推定

分散分析表 後値 $y_1$ 交互作用モデル					
		自由度	変動	分散	分散比
	回帰	7	13104.31	1872.04	74.21
	残差	32	807.20	25.22	
	合計	39	13911.51		
		係数	標準誤差	$t$	$P$ -値
$\theta^{\wedge}_0$	切片	30.6904	14.6666	2.0925	0.0444
$\theta^{\wedge}_1$	$a$	-0.3716	1.3904	-0.2673	0.7910
$\theta^{\wedge}_2$	$x_1$	0.6175	0.0760	8.1277	0.0000
$\theta^{\wedge}_3$	$x_2$	0.4754	0.0764	6.2212	0.0000
$\theta^{\wedge}_4$	$x_3$	2.4319	0.5807	4.1882	0.0002
$\theta^{\wedge}_5$	$ax'_1$	-0.0221	0.0760	-0.2904	0.7734
$\theta^{\wedge}_6$	$ax'_2$	0.0660	0.0764	0.8639	0.3941
$\theta^{\wedge}_7$	$ax'_3$	0.3712	0.5807	0.6393	0.5272

分散分析表 後値 $y_1$ 主効果モデル					
		自由度	変動	分散	分散比
	回帰	4	13064.19	3266.05	134.91
	残差	35	847.32	24.2091	
	合計	39	13911.51		
		係数	標準誤差	$t$	$P$ -値
	切片	29.4769	13.4959	2.1841	0.0357
	$a$	-0.0052	1.2513	-0.0042	0.9967
	$x_1$	0.6437	0.0622	10.3406	0.0000
	$x_2$	0.4515	0.0661	6.8258	0.0000
	$x_3$	2.4529	0.4555	5.3852	0.0000

### 複数の共変量を考慮した $2 \times 2$ 格子状の散布図

共変量が一つの場合には、層別散布図に平行な直線を引くことにより、共分散分析の結果を可視化できる。複数の共変量がある場合には、どうしたら良いのであろうか。第 8.1 節では、様々な層別散布図に、それぞれの層に別々に回帰直線を重ね書きしたのであるが、このような方法では、複数の共変量による調整結果を図示できない。

そこで、3つの共変量の内、総摂餌量  $x_2$  を主たる共変量とし、前値（7日の体重） $x_1$ 、胎仔数  $x_3$  を補助の共変量として、中央値（以下か、以上か）の2区分に分類し、さらに  $x_1$  と  $x_3$  の2区分内の平均値を求めて、後値（18日の体重） $y_1$  の推定値  $\hat{y}_1$  を重ね書きすることにより、可視化する。前値（7日の体重） $x_1$  の全40例の中央値は、261.0なので、それ以下を  $X_1^{(1)}=1$ 、それ以上を  $X_1^{(2)}=2$  と2区分にし、 $X_1^{(1)}=1$  となる区分内の  $x_1$  を平均値は  $\bar{x}_1^{(1)}=246.7$ 、 $X_1^{(2)}=2$  で  $\bar{x}_1^{(2)}=274.1$  に置き換える。胎仔数  $x_3$  の場合の中央値は、14.0なので、それ未満を  $X_3^{(1)}=1$ 、それ以上を  $X_3^{(2)}=2$  とする。 $X_3^{(1)}=1$  となる区分内の  $x_3$  を  $\bar{x}_3^{(1)}=12.1$  に、 $X_3^{(2)}=2$  の場合は  $\bar{x}_3^{(2)}=14.7$  に置き換える。

表 8.3 に示した元データに対し、 $X_1^{(i)}$  および  $X_3^{(i)}$  の2区分変数を追加し、Excel のソート機能を使い、 $(X_1^{(i)}, X_3^{(i)}, A, x_2)$  の順番で並び替える。このような手順により、表 8.4 の結果が得られる。前値（7日の体重） $x_1$  は、 $X_1^{(i)}$  の (1, 2) に対応して、(246.7, 274.1) に置き換えられ、胎仔数  $x_3$  は、 $X_3^{(i)}$  の (1, 2) に対応して、(12.1, 14.7) に置き換えられている。置き換えられた変数を含めて変数名を別名 ( $l_0, l_1, l_2, l_3, l_4$ ) とし与え、表 8.3 で推定されたパラメータ

$$\hat{\theta} = [29.4769 \quad -0.0052 \quad 0.6437 \quad 0.4515 \quad 2.4529]^T$$

を用いて、後値（18日の体重）の推定値  $\hat{y}_1$  を  $\hat{y}_1 = l\hat{\theta}$  により推定している。

表 8.5 前値  $x_1$ ・胎仔数  $x_3$  の区分別の共変量を反映した後値  $y_1$  推定値

前値 群	胎仔 群	投与 群	番号	切片 $l_0$	投与 群 $l_1$	前値 7日の 群平均 $l_2$	総摂 餌量 $l_3$	胎仔 数 群平均 $l_4$	回帰 パラ メータ 推定値 $\theta^\wedge$	後値 18日の 推定値 $\nu_1^\wedge$	後値 18日の 体重 $\nu_1$
$X_1^{(i)}$	$X_3^{(i)}$	A		$x_0$	$a$	$x_1$	$x_2$	$x_3$			
1	1	A <sub>1</sub>	11	1	1	246.7	191.1	12.1	29.4769	304.1	296.0
			1	1	1	246.7	203.3	12.1	-0.0052	309.6	318.1
			14	1	1	246.7	215.4	12.1	0.6437	315.1	304.5
			20	1	1	246.7	216.0	12.1	0.4515	315.3	310.7
			8	1	1	246.7	221.4	12.1	2.4529	317.8	308.3
			16	1	1	246.7	225.1	12.1		319.5	320.6
		A <sub>2</sub>	34	1	-1	246.7	164.5	12.1		292.1	282.5
			35	1	-1	246.7	166.7	12.1		293.1	275.4
			22	1	-1	246.7	167.9	12.1		293.6	305.4
			21	1	-1	246.7	199.2	12.1		307.8	309.1
	2	A <sub>1</sub>	13	1	1	246.7	197.8	14.7		313.6	313.4
			5	1	1	246.7	204.5	14.7		316.7	328.6
			19	1	1	246.7	206.7	14.7		317.6	310.1
			18	1	1	246.7	215.5	14.7		321.6	326.6
			15	1	1	246.7	232.0	14.7		329.1	342.7
		A <sub>2</sub>	38	1	-1	246.7	171.9	14.7		301.9	295.7
			39	1	-1	246.7	180.8	14.7		306.0	306.5
			33	1	-1	246.7	188.6	14.7		309.5	309.9
			40	1	-1	246.7	193.0	14.7		311.5	316.0
			32	1	-1	246.7	217.0	14.7		322.3	332.0
2	1	A <sub>1</sub>	2	1	1	274.1	223.0	12.1		336.1	324.0
			3	1	1	274.1	244.9	12.1		346.0	339.9
			12	1	1	274.1	258.4	12.1		352.1	367.2
		A <sub>2</sub>	29	1	-1	274.1	184.1	12.1		318.6	317.1
			28	1	-1	274.1	186.3	12.1		319.6	318.7
			24	1	-1	274.1	195.6	12.1		323.8	323.2
			26	1	-1	274.1	196.7	12.1		324.3	339.6
			25	1	-1	274.1	203.6	12.1		327.4	332.1
			27	1	-1	274.1	207.7	12.1		329.2	330.2
	2	A <sub>1</sub>	10	1	1	274.1	210.4	14.7		337.0	335.4
			9	1	1	274.1	221.1	14.7		341.8	340.0
			6	1	1	274.1	223.0	14.7		342.7	338.1
			17	1	1	274.1	227.9	14.7		344.9	345.9
			7	1	1	274.1	228.5	14.7		345.2	349.3
			4	1	1	274.1	236.5	14.7		348.8	356.1
		A <sub>2</sub>	23	1	-1	274.1	185.9	14.7		326.0	331.5
			30	1	-1	274.1	198.8	14.7		331.8	328.7
			36	1	-1	274.1	199.3	14.7		332.0	329.5
			37	1	-1	274.1	209.9	14.7		336.8	329.1
			31	1	-1	274.1	219.3	14.7		341.0	335.4

総摂餌量  $x_2$  が,  $(X_1^{(i)}, X_3^{(i)}, A)$  の中で小さい順にソートしてあるのは, Excel の散布図上で推定値  $\hat{y}_1$  を回帰直線として描くためである. 区分  $(X_1^{(1)}=1, X_3^{(1)}=1, A=A_1)$  の最小値  $x_2^{(11)}=191.1$  に対して,  $\mathbf{l}^{(11)}=[1 \ 1 \ 246.7 \ 191.1 \ 12.1]$  なので  $\hat{y}_1^{(11)}=\mathbf{l}^{(11)}\hat{\boldsymbol{\theta}}=304.1$  が得られている. 同じ区分の最大値  $x_2^{(16)}=225.1$  に対しては,  $\hat{y}_1^{(16)}=\mathbf{l}^{(16)}\hat{\boldsymbol{\theta}}=319.5$  が得られている.

表 8.4 の右端には, 後値の体重  $y_1$  が示されている. この区分の 6 個のデータ図 8.8 の左上側の散布上に書き込む. 後値の体重  $y_1$  は, デフォルトのマーカーを ○ 印とし, 推定値  $\hat{y}_1$  は, 回

帰直線としている．この直線の切片  $\hat{\beta}_0$  は， $x_2 = 0$  の場合であり， $\mathbf{l}_{(x_2=0)}^{(1)} = [1 \quad 1 \quad 246.7 \quad 0 \quad 12.1]$  なので  $\hat{\beta}_0 = \mathbf{l}_{(x_2=0)}^{(1)} \hat{\boldsymbol{\theta}} = 217.83$  となる．傾き  $\hat{\beta}_1$  は，表 8.5 から  $\hat{\theta}_3 x_2 = 0.4515 x_2$  なので，

$$(X_1^{(1)}=1, X_3^{(1)}=1, A=A_1) : \hat{y}_1^{\text{対照}} = 217.83 + 0.4515x_2 \quad (8.6)$$

となる．処置群  $A_2$  の場合は， $a = -1$  とした  $\mathbf{l}_{(a=-1)}^{(1)} = [1 \quad -1 \quad 246.7 \quad 191.1 \quad 12.1]$  を用いて

$$(X_1^{(1)}=1, X_3^{(1)}=1, A=A_2) : \hat{y}_1^{\text{処置}} = 217.84 + 0.4515x_2 \quad (8.7)$$

となる．切片が，わずか 0.01 異なるだけで，2本の直線は，完全に重なっているように見える．

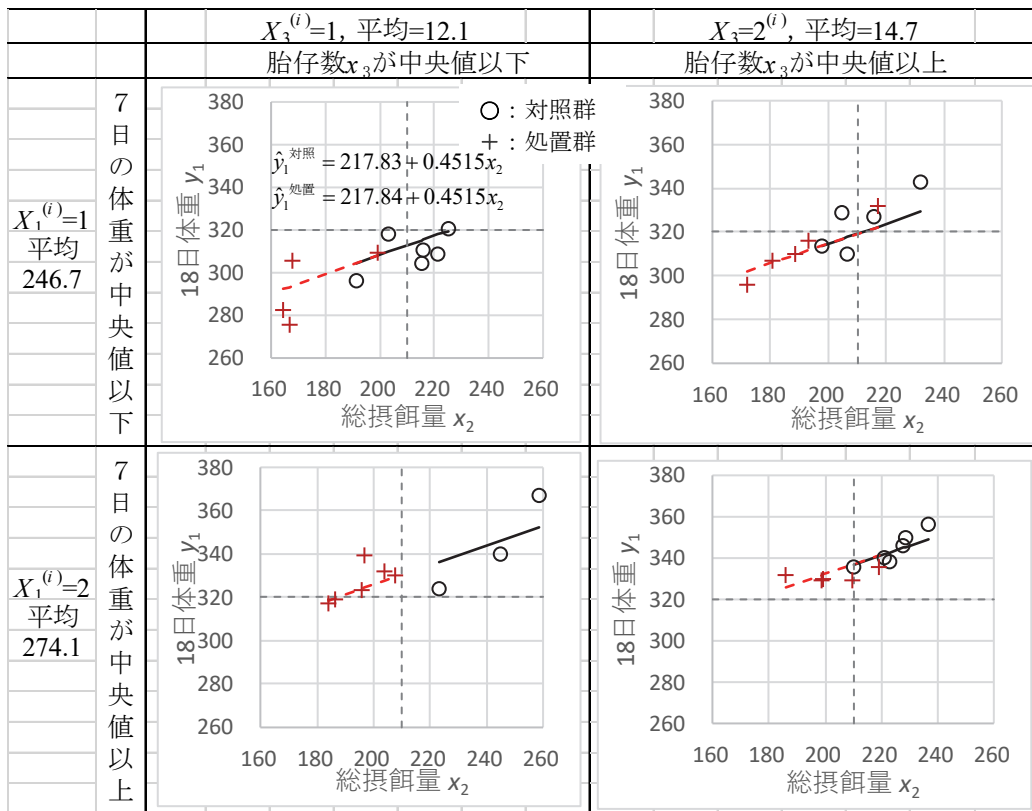


図 8.8 2×2 区分の総摂餌量に対する 18 日体重の散布図

図 8.8 右上側は，胎仔数  $x_3$  が中央値以上の場合であり，図 8.8 左上側に比べて，分布が右上に少々シフトしている．図 8.8 下側は，前値（7 日の体重）が中央値以上の場合であり，後値（18 日の体重）が増加する方向にシフトしていることが観察される．

このような 2×2 の格子状の散布図にすることにより，複数の共変量が絡み合っている場合に，処置群  $A_2$  に対する体重の増加抑制が，体重に対する直接的な作用ではなく，交絡変数である総摂餌量の抑制による間接的な体重減少と判断される．

### 8.3. JMP による複数の共変量を含む探索的な回帰分析

#### ステップワイズ法による変数選択

JMP の「モデルのあてはめ」には、「標準最小 2 乗法」に加えて「ステップワイズ法」も備わっており、(変数増加・変数減少・変数増減)を自動的に行わせる機能だけでなく、 $p$  値を見ながら(モデルに含める変数・除く変数)について手動で取捨選択する機能を備えている。また、変数選択を自動的に行わせる際に、第 7.4 節で示したように主効果と交互作用をセットにして扱う「組み合わせ」のルールがデフォルトで設定されている。

表 8.6 に示すのは、JMP ファイルに Excel のシート上のデータをペーストし、JMP 新規データファイル上に「列名と共に貼り付け」て変数名を整えた結果を示す。下段のそれぞれの変数に対する平均値の計算は、JMP の「表の作成」で計算したもので、Excel のピボットテーブルに対応している。

表 8.6 解析用の JMP データファイル

	番号	群A	前値7日 $x_1$	総摂餌量 $x_2$	胎仔数 $x_3$	後値18日 $y_1$
○	1	1 A1	248.7	203.3	13	318.1
○	2	2 A1	261.3	223.0	13	324.0
○	3	3 A1	265.6	244.9	12	339.9
+	38	38 A2	231.4	171.9	15	295.7
+	39	39 A2	239.0	180.8	15	306.5
+	40	40 A2	257.4	193.0	14	316.0

群A	前値7日 $x_1$	総摂餌量 $x_2$	胎仔数 $x_3$	後値18日 $y_1$
A1 A2	平均	平均	平均	平均
20 20	260.39	205.983	13.45	323.08

群 A と 3 つの共変量の交互作用を含むモデルを JMP の「モデルのあてはめ」で設定し、「手法」を「ステップワイズ法」に切り替える。モデルの設定で、主効果として(「群 A」, 「前値 7 日  $x_1$ 」, 「総摂餌量  $x_2$ 」, 「胎仔数  $x_3$ 」), 交互作用として(「群 A\*前値 7 日  $x_1$ 」, 「群 A\*総摂餌量  $x_2$ 」, 「群 A\*胎仔数  $x_3$ 」)を設定する。ただし、表 8.7 に示すように、ステップワイズ法の設定の画面に現れる変数名は、(「群 A」→「群 A{A2-A1}」)となり、何らかのダミー変数に変換されていることが示唆されている。「群 A」と共変量の交互作用については、それぞれの総平均を差し引いた“中心化”が、内部の変数として使われていることが示されている。

ステップワイズ法には、幾つかの方法があり、「停止ルール」を「閾値  $p$  値」,「方向」を,「変数増減」,「ルール」には,「組み合わせ」を選択している.

表 8.7 に示すのは,「群  $A\{A_2 - A_1\}$ 」を追加し,ロックを掛け,「前値 7 日  $x_1$ 」を手動で追加した結果である. 推定された式は,

$$y_1 = 76.4157 - 6.7016 \times \text{「群 } A\{A_2 - A_1\}\text{」} + 0.9473 \times \text{「前値 7 日 } x_1\text{」} \quad (8.8)$$

であり,「群  $A\{A_2 - A_1\}$ 」の推定値がマイナスなので,内部で生成されたダミー変数は,(-1, 1) 対比型ダミー変数が設定されている. また,「群  $A\{A_2 - A_1\}$ 」の  $p$  値が 1.36E-05 と高度に有意であり,共変量として「前値 7 日  $x_1$ 」を含めた場合に,対照群  $A_1$  と処置群  $A_2$  の「後値 18 日体重  $y_1$ 」に明らかな差があるとの結論となる.

表 8.7 ステップワイズ回帰による「群  $A$ 」と共変量「前値 7 日  $x_1$ 」の選択の結果

後値18日y1のステップワイズ

ステップワイズ回帰の設定

停止ルール: 閾値p値

変数を追加するときのp値0.25

変数を除去するときのp値0.25

方向: 変数増減

ルール: 組み合わせ

→すべて追加モデルの作成

←すべて削除モデルの実行

SSE	DFE	RMSE	R2乗	自由度調整R2乗	Cp	p	AICc	BIC
2635.1242	37	8.4391689	0.8106	0.8003	70.464931	3	290.1702	295.7828

現在の推定値

ロック	追加	パラメータ	推定値	自由度	平方和	"F値"	"p値(F)"
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	切片	76.4157	1	0.0000	0.0000	1
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	群A{A2-A1}	-6.7016	1	1789.2339	25.1228	1.36e-5
<input type="checkbox"/>	<input checked="" type="checkbox"/>	前値7日x1	0.9473	1	9977.9253	140.1009	3.8e-14
<input type="checkbox"/>	<input type="checkbox"/>	総摂餌量x2	0.0000	1	1085.7412	25.2273	1.4e-5
<input type="checkbox"/>	<input type="checkbox"/>	胎仔数x3	0.0000	1	659.8486	12.0259	0.00138
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(前値7日x1-260.39)	0.0000	1	219.0111	3.2633	0.07921
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(総摂餌量x2-205.983)	0.0000	1	2.7498	0.0376	0.84733
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(胎仔数x3-13.45)	0.0000	1	102.3277	1.4544	0.23568

表 8.7 ステップワイズ回帰による「群  $A$ 」と共変量「前値 7 日  $x_1$ 」の選択の結果「総摂餌量  $x_2$ 」および「胎仔数  $x_3$ 」の  $p$  値が, (1.4E-05, 0.00138) であることから,  $p$  値の小さい「総摂餌量  $x_2$ 」をモデルに追加する. 表 8.8 に示すように, 追加後「群  $A\{A_2 - A_1\}$ 」の推定値および  $p$  値が大きく変化し, 有意な差ではなくなっている. さらに, 交互作用を含めて  $p$  値が最も小さいのが,「胎仔数  $x_3$ 」であり,  $p=4.99\text{E-}06$  であることが読み取れる.



表 8.8 ステップワイズ回帰による「総摂餌量 x2」を追加した結果結果

現在の推定値							
ロック	追加	パラメータ	推定値	自由度	平方和	"F値"	"p値(F)"
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	切片	56.3873	1	0.0000	0.0000	1
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	群A{A2-A1}	-0.1493	1	0.3446	0.0080	0.92919
<input type="checkbox"/>	<input checked="" type="checkbox"/>	前値7日x1	0.6739	1	2860.4603	66.4630	1.08e-9
<input type="checkbox"/>	<input checked="" type="checkbox"/>	総摂餌量x2	0.4428	1	1085.7412	25.2273	1.4e-5
<input type="checkbox"/>	<input type="checkbox"/>	胎仔数x3	0.0000	1	702.0632	28.9999	4.99e-6
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(前値7日x1-260.39)	0.0000	1	90.4582	2.1701	0.14965
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(総摂餌量x2-205.983)	0.0000	1	2.7498	0.0622	0.80447
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(胎仔数x3-13.45)	0.0000	1	14.7628	0.3367	0.56546

表 8.9 に示すように「胎仔数  $x_3$ 」を加えると、交互作用の  $p$  値は全て 0.25 以上となり、追加する「変数」ないと判断される。この結果は、表 8.4 で示した Excel での結果に一致する。なお、自由度がすべて 1 なのは、交互作用がモデルに含まれていないためである。

表 8.9 ステップワイズ回帰による「胎仔数 x3」を追加した結果結果

SSE	DFE	RMSE	R2乗	自由度調整R2乗	Cp	p	AICc	BIC
847.31973	35	4.920278	0.9391	0.9321	3.5905223	5	250.1885	257.7763

現在の推定値							
ロック	追加	パラメータ	推定値	自由度	平方和	"F値"	"p値(F)"
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	切片	29.4769	1	0.0000	0.0000	1
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	群A{A2-A1}	0.0052	1	0.0004	0.0000	0.99669
<input type="checkbox"/>	<input checked="" type="checkbox"/>	前値7日x1	0.6437	1	2588.6242	106.9276	3.5e-12
<input type="checkbox"/>	<input checked="" type="checkbox"/>	総摂餌量x2	0.4515	1	1127.9559	46.5922	6.42e-8
<input type="checkbox"/>	<input checked="" type="checkbox"/>	胎仔数x3	2.4529	1	702.0632	28.9999	4.99e-6
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(前値7日x1-260.39)	0.0000	1	9.3823	0.3807	0.54134
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(総摂餌量x2-205.983)	0.0000	1	28.1522	1.1685	0.28732
<input type="checkbox"/>	<input type="checkbox"/>	群A{A2-A1}*(胎仔数x3-13.45)	0.0000	1	14.7628	0.6029	0.44285

JMP の「ステップワイズ法」は、自動的に変数選択を行なう機能もあるが、ここに示したように手作業により変数を取捨選択する機能に優れている。ただし、この結果をもって探索的な解析が終わったわけではない。「ステップワイズ法」での出力は、(推定値, 平方和,  $F$  値,  $p$  値)のみであり、群 A の各水準の最小 2 乗平均などの推定値の計算は行われていないので、第 7 章に示したように、解析結果を考察するためには、種々の推定と結果のグラフ表示が欠かせない。JMP の「ステップワイズ法」は、変数選択に特化しているので、更なる解析のためには、表 8.7 右上の **モデルの作成** または **モデルの実行** を選択することにより、「モデルのあてはめ」の基本である「標準最小 2 乗」が自動的に起動される。

## 「標準最小2乗法」による予測プロファイルの活用

「標準最小2乗」による解析結果を表8.10に示す。分散分析表の「誤差の平方和=847.3197」が、表8.9のSSEに該当し、RMSE=4.9203は、表8.10の「誤差の平均平方=24.2091」の平方根に一致している。「パラメータ推定値」の推定値とp値が、表8.9に一致していることが示されている。もちろん、表8.4に示したExcelの回帰分析による主効果モデルにも一致している。細かいことであるが、群A[A1]のパラメータ推定値は、-0.0052と表8.9の0.0052と異なる。これは、「標準最小2乗法」でのタミー変数が、(-1, 1)対比型ではなく(1, -1)対比型タミー変数に切り替わっていることによる。

表8.10 主効果モデルによる解析結果

分散分析					
要因	自由度	平方和	平均平方	F値	p値(Prob>F)
モデル	4	13064.1900	3266.0475	134.9097	
誤差	35	847.3197	24.2091		
全体(修正済み)	39	13911.5098			<.0001*

パラメータ推定値				
項	推定値	標準誤差	t値	p値(Prob> t )
切片	29.4769	13.4959	2.1841	0.0357*
群A[A1]	-0.0052	1.2513	-0.0042	0.9967
前値7日x1	0.6437	0.0622	10.3406	<.0001*
総摂餌量x2	0.4515	0.0661	6.8258	<.0001*
胎仔数x3	2.4529	0.4555	5.3852	<.0001*

効果の検定					
要因	自由度	平方和	平均平方	F値	p値(F)
群A	1	0.0004	0.0004	0.0000	0.9967
前値7日x1	1	2588.6242	2588.6242	106.9276	<.0001*
総摂餌量x2	1	1127.9559	1127.9559	46.5922	<.0001*
胎仔数x3	1	702.0632	702.0632	28.9999	<.0001*

図8.9に示す予測プロファイルは、表8.10で推定された主効果モデルのパラメータの推定値に対する95%信頼区間、推定された回帰直線とその95%信頼区間が示されている。左端に示されている「323.1 [320.1, 326.1]」は、「群A」のA1水準の $a=1$ 、「前値7日 $x_1$ 」の平均値260.39、「総摂餌量 $x_2$ 」の平均値205.982、「胎仔数 $x_3$ 」の平均値13.45を式(8.9)に代入した推定値（最小2乗平均）と95%信頼区間である。

$$\begin{aligned}
 \hat{y}_1 &= \hat{\theta}_0 x_0 + \hat{\theta}_1 a + \hat{\theta}_2 x_1 + \hat{\theta}_3 x_2 + \hat{\theta}_4 x_3 \\
 &= 29.4769x_0 + (-0.0052)a + 0.6437x_1 + 0.4515x_2 + 2.4529x_3 \\
 &= 323.1
 \end{aligned}
 \quad \left. \vphantom{\begin{aligned} \hat{y}_1 &= \hat{\theta}_0 x_0 + \hat{\theta}_1 a + \hat{\theta}_2 x_1 + \hat{\theta}_3 x_2 + \hat{\theta}_4 x_3 \\ &= 29.4769x_0 + (-0.0052)a + 0.6437x_1 + 0.4515x_2 + 2.4529x_3 \\ &= 323.1 \end{aligned}} \right\} \quad (8.9)$$

群 A を ( $a=1$ ) と固定し、共変量 ( $x_1, x_2, x_3$ ) の平均値に固定し、該当する変数について一定の刻みで変化させたときの回帰直線と 95%信頼区間が示されている。予測プロファイル上に各変数の設定値が示されているが、マウス操作などで自由に変更できるようになっている。群 A については「A1」が選択されているが、マウスで「A2」を選択することもできる。

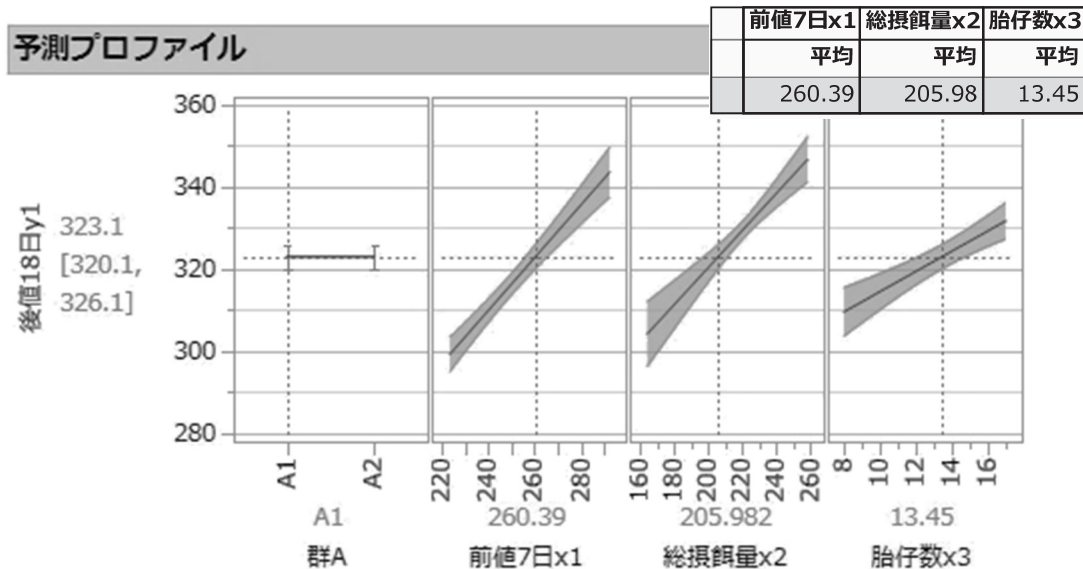


図 8.9 推定値に対する予測プロファイルによる結果の表示

表 8.11 に示すのは、共変量 ( $x_1, x_2, x_3$ ) を平均値に固定したときの、群 A の各水準の最小 2 乗平均、「A1 : 323.072, A2 : 323.083」が推定されている。なお、この「最小 2 乗平均」は、図 8.9 左端に示される推定値にも対応している。表 8.11 の右端の「平均」は、群 A の各水準の算術平均であり、表 8.1 の「体重 18 日」についての「平均」に一致する。これに対し、「最小 2 乗平均」は、式 (8.9) に示したように表 8.10 の回帰分析で得られたパラメータを用いた推定結果である。

表 8.11 群 A についての最小 2 乗平均

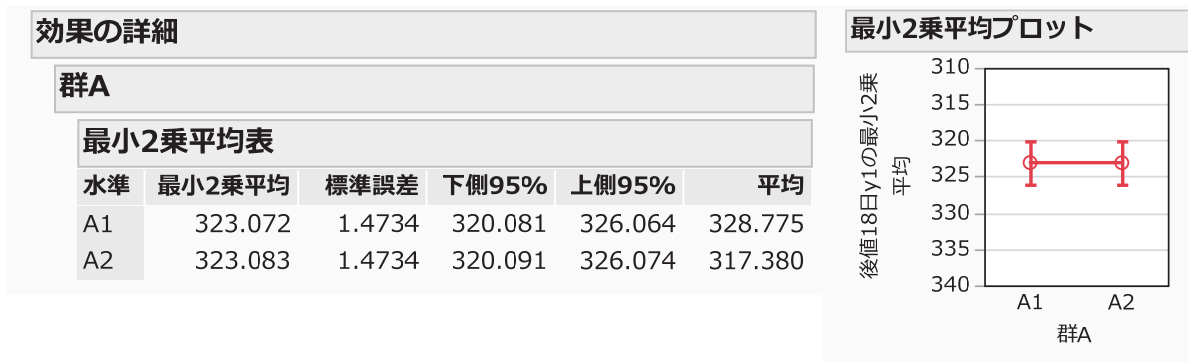


図 8.10 最小 2 乗平均プロット

## 8.4. 交絡変数の多彩な影響への対応

化合物の多面的な毒性を把握するために、反復投与毒性試験が行われている。医薬品の場合は、げっ歯類に対して1カ月、3カ月、6カ月の反復投与が規定されている。本節で取り上げるのはラットを対象にした3カ月の反復投与終了後の測定された体重および臓器重量に対する解析法を取り上げる。

吉村編著（1987）、「毒性・薬効データの統計解析、第3.3節 臓器重量の解析における体重データの利用」には、臓器重量と体重の関係の6通りの典型例が示されており、現実を得られたデータに対し、どのような解析法の選択が望ましいのかを前もって検討するための事例として示されている。さらに、高橋ら（1989）、「SASによる実験データの解析、第7章 共分散分析」では、SASを用いた詳細な解析方法が示されている。これらを踏まえて、ExcelおよびJMPを主体にした解析法を示す。

### 投与終了時の体重は交絡変数

図8.11に示すように反復投与終了時に体重の大きい個体は、一般的に臓器重量も大きくなる。化合物が、体重の増加を抑制すると、それに伴い臓器の重量の増加が抑制される。このときに体重を考慮しないで臓器の重量だけをみると、臓器には化合物の直接の作用がないのに、臓器に直接的な作用があったと誤解してしまう。このような場合に、本当の原因がどれであるか分からなくなる現象を原因の“交絡”confoundingという。

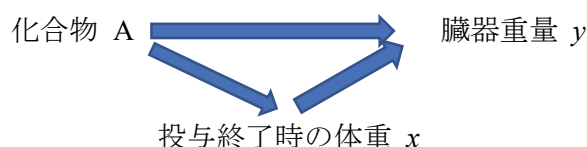


図 8.11 投与終了時の体重が臓器重量におよぼす影響

交絡があると、知りたいことについて正しい判断ができなくなる。図8.12に臓器重量と体重の種々の関連について人工データ（②組は心重量についての実データ）に基づいた6種類の散布図を例示する。

毒性試験における臓器重量の統計解析は、対照群と投与群間での臓器の絶対重量での比較、臓器重量を体重で除した体重比（相対重量）での比較が主体であり、共分散分析をどのような場合に使うかの検討がなされている。交絡変数としての体重 $X$ と臓器重量 $Y$ の關係に、

- i) 相関があるか
- ii) 2群の回帰直線が同一線上か

iii) 直線が原点を通るか

iv) 2 群の直線の傾きは同じか

などを判断基準とし、表 8.12 に示すように ③ 組と ⑤ 組の場合に共分散分析が適用できると示唆している。⑥ 組のように、傾きが異なり、原点を通らない場合は、臓器の絶対重量と

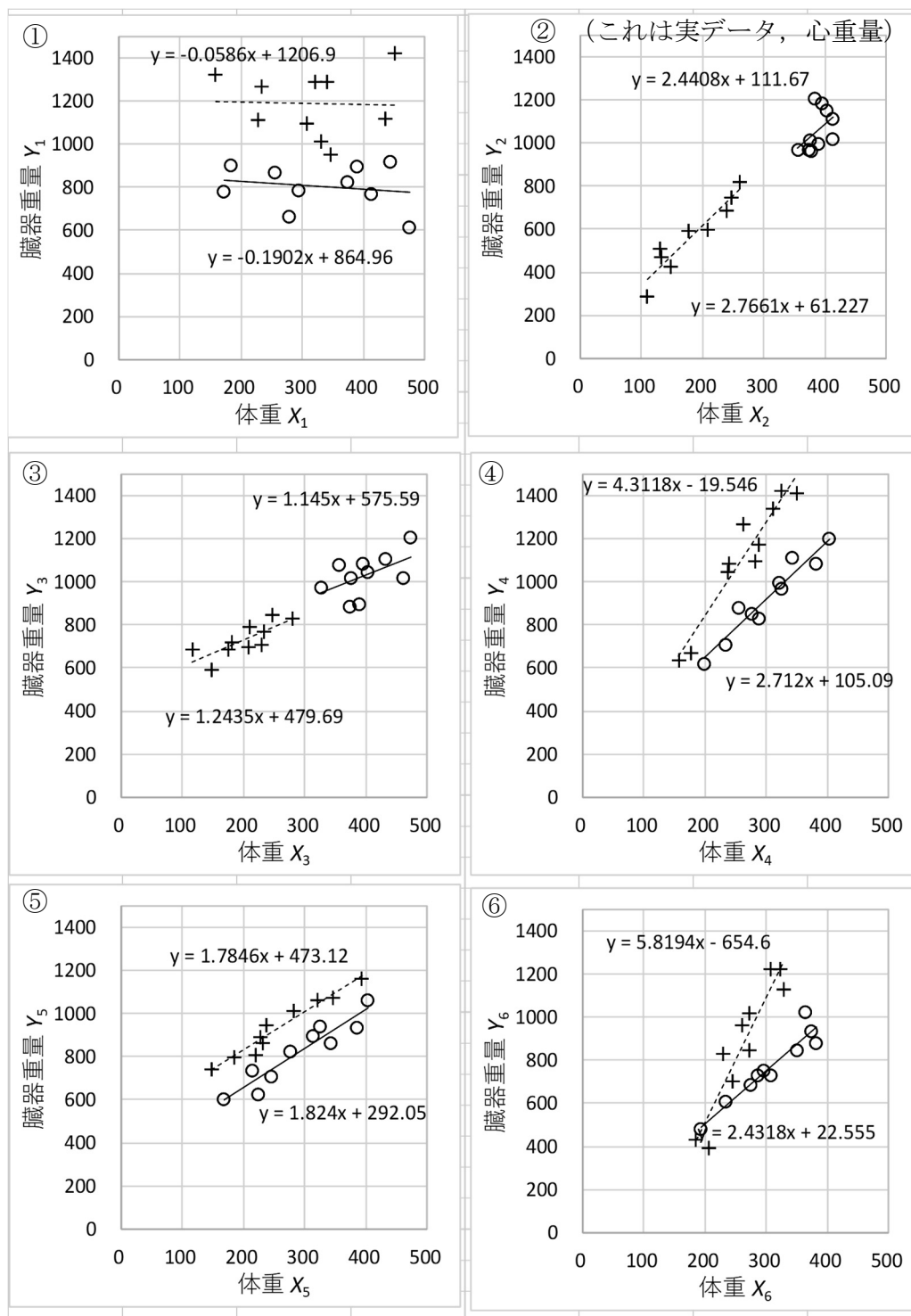


図 8.12 臓器重量と体重の関係の 6 通りの典型例

○印：対照群，＋印：処理群

回帰直線の差の比較を推奨している。このことから、共分散分析は、処置Aと共変量 $X$ の交互作用がないことを前提にしている。②組と③組の場合は、絶対重量と体重比が、①組の場合は、絶対重量と体重を別々に解析することを推奨されている。

表 8.12 体重の有意差が臓器重量の有意差に与える影響 [吉村編著 (1987)]

場合 図8.14に対応	体重データ での有意差	臓器重量 での有意差	臓器への 直接影響	説明
1	無	有	有	臓器重量での解析結果が重要
2	有	有	無	体重比での解析結果が重要
3	有	有	無	共分散分析での解析が必要
4	無	無	有	体重比での解析結果が重要
5	無	無	有	共分散分析での解析が必要
6	無	無	有	回帰直線の差の検定が必要

さて、i) ~ iv) の設問での「相関があるか」については、 $X$ と $Y$ の相関係数を計算して判断すればよいこと分かるが、ii) ~ iv) の設問については、どのような計算をしたらよいのだろうか。図 8.12 から視覚的な判断に基づいて表 8.12 が作成されているが、Excel を用いた計算結果に基づいて検討し直したい。

図 8.12 の元になるデータは、高橋ら (1989)、図表 7.1 に掲載されているので、表 8.13 に引用する。図 8.12 に示したのは、表 8.13 のデータに基づき Excel の散布図によって作成し直し

表 8.13 臓器重量 $X_i$ と体重 $Y_i$ のデータリスト [高橋ら (1989)、図表 7.1]

処置		① 組		② 組		③ 組		④ 組		⑤ 組		⑥ 組	
A	No.	$X_1$	$Y_1$	$X_2$	$Y_2$	$X_3$	$Y_3$	$X_4$	$Y_4$	$X_5$	$Y_5$	$X_6$	$Y_6$
1	1	255	869	355	969	355	1080	255	881	245	710	285	731
1	2	412	771	412	1111	432	1107	382	1082	224	627	382	882
1	3	443	917	413	1017	460	1015	320	995	342	865	350	845
1	4	183	904	383	1204	473	1204	403	1201	403	1064	193	481
1	5	374	826	374	966	374	886	234	707	214	736	234	607
1	6	475	613	375	1013	375	1020	325	970	325	940	275	685
1	7	294	785	394	1185	394	1085	198	621	384	935	364	1021
1	8	278	663	378	963	326	972	276	852	276	822	296	752
1	9	388	897	388	997	388	897	288	831	168	603	308	731
1	10	172	782	403	1150	403	1047	343	1112	313	897	373	932
2	11	347	953	247	744	247	844	287	1174	347	1074	322	1224
2	12	320	1290	110	290	210	790	240	1082	220	806	229	832
2	13	435	1117	260	817	280	827	282	1097	282	1013	272	847
2	14	158	1323	149	423	117	683	157	635	227	893	207	395
2	15	340	1288	240	688	175	688	325	1423	185	798	245	703
2	16	450	1424			181	721	311	1341	231	861	186	431
2	17	233	1269	133	469	233	769	263	1269	393	1162	273	1019
2	18	228	1111	178	591	148	591	178	671	320	1060	308	1221
2	19	330	1012	130	510	230	710	350	1410	148	741	260	960
2	20	308	1097	208	597	208	697	238	1047	238	947	328	1127

たものである。処置 A は、1 が対照群、2 が処置群、 $(X_1, Y_1)$  が ① 組の臓器重量と体重であり、順次 ⑥ 組のデータが示されている。きれいな散布図を短時間で作成することは、統計解析の基礎の基礎でもある。

図 8.12 に示した ① 組の散布図は、対照群  $A_1$  の 10 例について散布図を描き、処置群  $A_2$  の 10 例を「データの選択」で重ね書きし書式を整え、さらに、「近似曲線の追加」で回帰直線を重ね書きする。②～⑥ 組は、① 組全体をコピーし、所定の位置にペーストする。散布図上の対照群  $A_1$  のマークを選択すると該当するデータ範囲がハイライトされるので所定の位置に範囲を移動する。 $A_2$  も同様にする。回帰直線は、自動的に更新される。

### 解析手法の選択のための決定樹

どのような解析手法を選択したら良いのか、図 8.13 に示す「決定樹」に表すと見通しがよくなる。

設問 1:  $X$  と  $Y$  に相関ある?

設問 3: 2 群の直線の傾きは同じか?

設問 2: 2 群が同一直線上か?

設問 4: 原点を通るか?

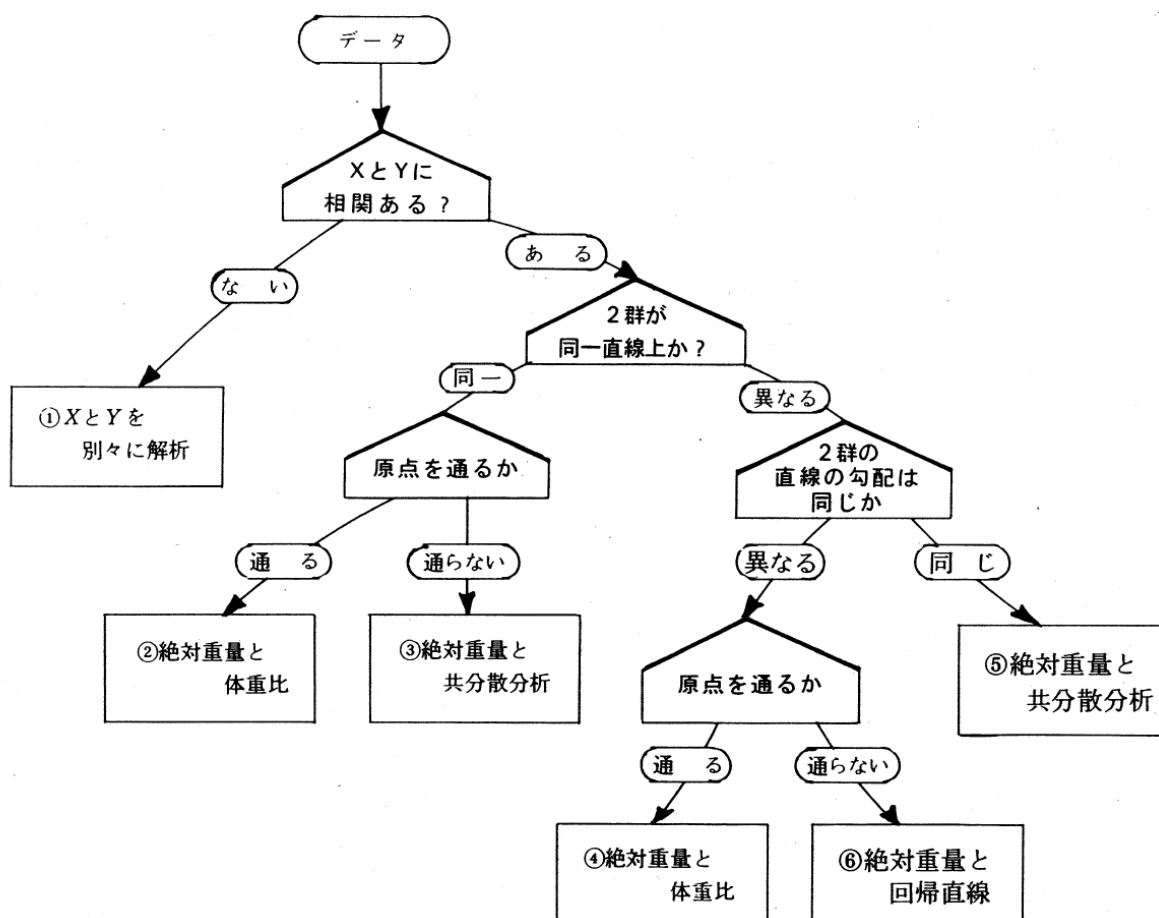


図 8.13 臓器重量の解析のための決定樹 [吉村編著 (1987), 図 3-5 を引用]

この決定樹からも、臓器重量と体重に対する解析は、常に探索的な解析になることが示されている。(相関があるか、原点を通るか)については、一般的な解析法であるが、(2群が同一直線上か、2群の直線の傾きは同じか)は、体重を共変量にした回帰分析が必要になる。しかし、吉村編著(1987)には、図の例示のみでデータおよび解析結果は示されず、別の事例について、偏差平方和をベースにした伝統的な共分散分析の方法が示されているのみである。

### 1) $X$ と $Y$ に相関があるか

視覚的には、「相関があるか」に対し ① 組が該当しないが、② 組の対照群  $A_1$  も該当しないように思われる。表 8.14 に示すように、群ごとに2変数の相関係数を求めるのは、Excel シート上で  $\text{Correl}()$  関数を使うのが効率的である。 $A_1$  の ① 組について、

$$r^{(1,A_1)} = \text{Correl}(X^{(1,A_1)} \text{の範囲}, Y^{(1,A_1)} \text{の範囲}) = -0.1976 \quad (8.10)$$

として計算し、その計算式をコピーし他の組み合わせ位置にペースする。(Xの範囲、Yの範囲)がずれるので、計算式をマウスで選択しハイライトされたデータ範囲を適切な位置にマウスで移動させし、表 8.14 の相関係数を得る。

表 8.14 データセットの組・群別の相関係数  $r$

	データセットの組					
	1	2	3	4	5	6
対処群 $A_1$	-0.1976	0.4711	0.5428 *	0.9628 **	0.9448 **	0.9532 **
処置群 $A_2$	-0.0350	0.9452 **	0.7924 *	0.9583 **	0.9739 **	0.9350 **
	無印: ( $r < 0.5$ )		*: ( $0.5 < r < 0.90$ )		**: ( $r > 0.90$ )	

JMP のグラフビルダーを用いて  $2 \times 6$  の格子状に散布図を描きに確率楕円を重ね書きし、さらに、オプションで相関係数を重ね書きした結果を図 8.14 に示す。

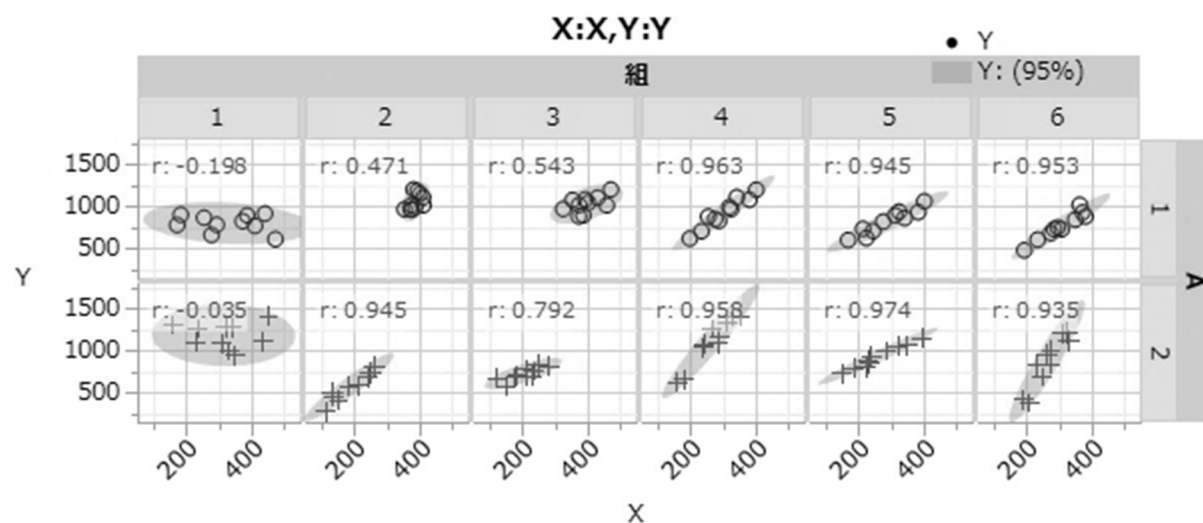


図 8.14 JMP グラフビルダーによる確率楕円と相関係数



## 2) 2 群の回帰直線が同一直線上か

視覚的には、② 組と ③ 組が該当するが、統計的には、臓器重量を  $Y$  とし、体重を共変量  $X$ 、群をダミー変数  $A$  とした回帰分析を行ない、群に関する  $p$  値で判断する。例えば、「 $p > 0.15$  ならば、同一直線」などの基準を前もって設定しておく。② 組の場合に JMP の「モデルのあてはめ」による結果を表 8.15 に示す。JMP は、(1, -1) 対比型ダミー変数なので、対照群  $A_1$  と処置群  $A_2$  の回帰式は、

$$\text{②組 } A_1 : \begin{cases} \hat{y}_i^{(2,A_1)} = 33.2462 - 34.3026 \times 1 + 2.7318 x_i^{(2,A_1)} \\ = -1.0564 + 2.7318 x_i^{(2,A_1)} \end{cases} \quad (8.11)$$

$$\text{②組 } A_2 : \begin{cases} \hat{y}_i^{(2,A_2)} = 33.2462 - 34.3026 \times (-1) + 2.7318 x_i^{(2,A_2)} \\ = 67.5488 + 2.7318 x_i^{(2,A_2)} \end{cases} \quad (8.12)$$

となる。 $A_1$  に対応する  $A[1]$  の検定結果は、 $p = 0.4802$  なので「同一直線上」と判断する。また、切片の推定値が  $p = 0.7945$  と有意ではないので、原点を通るとみなせる。

表 8.15 ② について共分散分析

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	2	1340796	670398	123.01
誤差	16	87199	5450	p値
全体(修正済み)	18	1427995		<.0001*
パラメータ推定値				
項	推定値	標準誤差	t値	p値
切片	33.2462	125.5402	0.2648	0.7945
A[1]	-34.3026	47.4592	-0.7228	0.4802
X2	2.7318	0.4354	6.2742	<.0001*

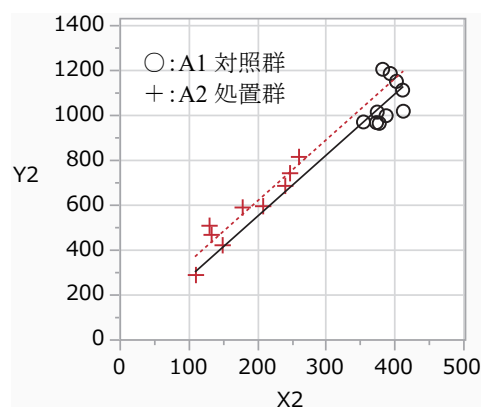


図 8.15 ② についての共分散分析

## 3) 2 本の直線の傾きは同じか

視覚的には、② 組と ⑤ 組が該当するのであるが、「傾きは同じか」について統計的には、臓器重量を  $Y$  とし、体重を共変量  $X$ 、群をダミー変数  $A$ 、さらに、 $X$  と  $A$  の交互作用についての回帰分析を行ない、交互作用  $A \times X$  に関する  $p$  値で判断する。例えば、「 $p > 0.15$  ならば、傾きが同じ」などの基準を前もって設定しておく。表 8.16 に ⑤ 組についての交互作用解析の結果を示す。対照群  $A_1$  と処置群  $A_2$  の回帰式は、

$$A_1 : \begin{cases} \hat{y}_i^{(5,A_1)} = 382.5823 - 90.5357 \times 1 + 1.8043 x_i^{(5,A_1)} + 0.0197 \times 1 x_i^{(5,A_1)} \\ = 292.0466 + 1.824 x_i^{(5,A_1)} \end{cases} \quad (8.13)$$

$$A_2 : \begin{cases} \hat{y}_i^{(5,A_2)} = 382.5823 - 90.5357 \times (-1) + 1.8043 x_i^{(5,A_2)} + 0.0197 \times (-1) \times x_i^{(5,A_2)} \\ = 473.118 + 1.7846 x_i^{(5,A_2)} \end{cases} \quad (8.14)$$

となり、交互作用「A[1]\*X5」に対する検定結果は、 $p=0.8852$  なので傾きは同じと判断できる。

表 8.16 ⑤ 組に対する交互作用の検討

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	3	408190	136063	72.02
誤差	16	30226	1889	p値
全体(修正済み)	19	438416		<.0001*
パラメータ推定値				
項	推定値	標準誤差	t値	p値
切片	382.5823	38.1062	10.04	<.0001*
A[1]	-90.5357	38.1062	-2.38	0.0303*
X5	1.8043	0.1343	13.44	<.0001*
A[1]*X5	0.0197	0.1343	0.15	0.8852

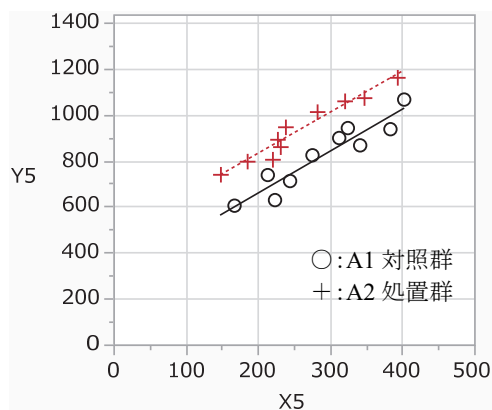


図 8.16 ⑤ 組に対する交互作用の検討

#### 4) 回帰直線が原点を通るか,

視覚的には、② 組と ④ 組が該当するが、それぞれの群別に臓器重量を  $Y$ 、体重を  $X$  とした回帰分析で、切片の  $p$  値で判断する。例えば、「 $p > 0.15$  ならば、原点を通る」などの基準を前もって設定しておく。④ 組の対照群  $A_1$  の場合に JMP の「二変量の関係」での結果を表 8.17 に示す。対照群  $A_1$ 、切片の推定値 83.0290 に対し、95%信頼区間が  $(-86.3777, 296.5526)$  と 0.0 を含んでいるので、原点を通ると判断する。

同様に、表 8.18 に示すように、処置群  $A_2$  についても切片の推定値  $-19.5458$  に対し、95%信頼区間が  $(-301.9799, 262.8839)$  と 0.0 を含んでいるので、原点を通ると判断される。

表 8.17 ④ 組の  $A_1$  についての回帰分析

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	1	276949	276949	101.58
誤差	8	21811	2726	p値
全体(修正済み)	9	298760		<.0001*
パラメータ推定値				
項	推定値	標準誤差	下側95%	上側95%
切片	105.0874	83.0290	-86.3777	296.5526
X4	2.7120	0.2691	2.0915	3.3325

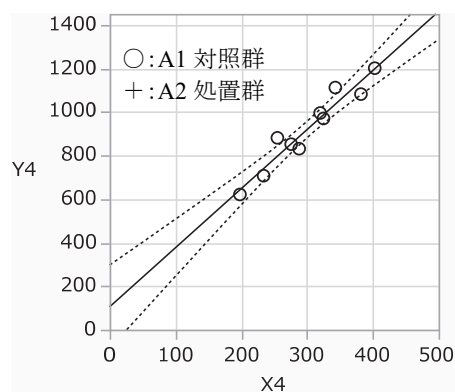
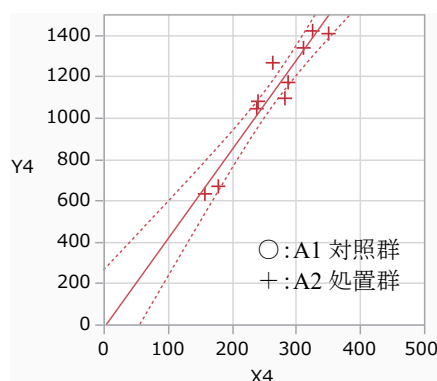
図 8.17 ④  $A_1$  についての回帰直線

表 8.18 ④ 組の  $A_2$  についての回帰分析

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	1	637127	637127	90.04
誤差	8	56608	7076	p値)
全体(修正済み)	9	693735		<.0001*

パラメータ推定値				
項	推定値	標準誤差	下側95%	上側95%
切片	-19.5458	122.4777	-301.980	262.8884
X4	4.3118	0.4544	3.2640	5.3597

図 8.18 ④  $A_2$  についての回帰直線

## 5) 2本の直線が原点を通らず傾きが同じでない場合

視覚的には、⑥ 組が該当するが、JMP の「モデルのあてはめ」による解析結果を表 8.19 に示す。対照群  $A_1$  と処置群  $A_2$  の回帰式は、

$$A_1 : \begin{cases} \hat{y}_i^{(6,A_1)} = -316.0250 + 338.5796 \times 1 + 4.1256 x_i^{(6,A_1)} - 1.6938 \times 1 x_i^{(6,A_1)} \\ = 22.5546 + 2.4318 x_i^{(6,A_1)} \end{cases} \quad (8.15)$$

$$A_2 : \begin{cases} \hat{y}_i^{(6,A_2)} = -316.0250 + 338.5796 \times (-1) + 4.1256 x_i^{(6,A_2)} + (-1.6938) \times (-1) x_i^{(6,A_2)} \\ = -654.605 + 5.8194 x_i^{(6,A_2)} \end{cases} \quad (8.16)$$

となり、交互作用「A[1]\*X6」に対する検定結果は、 $p = 0.0004$  であり、傾きに関して明らかな差がある。表 8.12 によれば、「回帰直線の差の検定が必要」とされているのだが、傾きの差

$$\hat{\beta}_1^{(A_2-A_1)} = 5.8194 - 2.4318 = 3.3876 \quad (8.17)$$

の検定結果は、交互作用の  $p$  値に一致するので、 $p = 0.0004$  と有意な差があるとの結果になるが、これだけで良いのだろうか。Excel の回帰分析を用いて更なる解析方法について、次節で改めて取り上げる。

表 8.19 ⑥ 組に対する交互作用の検討

分散分析				
要因	自由度	平方和	平均平方	F値
モデル	3	959842	319947	42.48
誤差	16	120494	7531	p値
全体(修正済み)	19	1080336		<.0001*

パラメータ推定値				
項	推定値	標準誤差	t値	p値
切片	-316.025	108.5033	-2.91	0.0102*
A[1]	338.5796	108.5033	3.12	0.0066*
X6	4.1256	0.3818	10.81	<.0001*
A[1]*X6	-1.6938	0.3818	-4.44	0.0004*

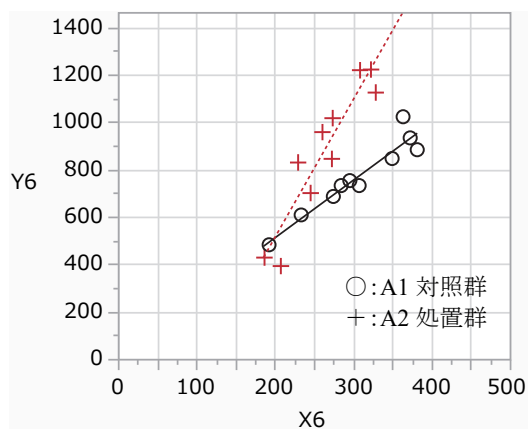


図 8.19 ⑥ に対する交互作用の検討

## 8.5. 体重を交絡変数とする回帰分析

反復毒性試験において化合物が臓器重量に与える影響を検討する場合に、体重が常に交絡変数となる可能性があるが、その影響はさまざまであり一筋縄では対応できない。探索的な解析は、体重を交絡変数（共変量）とする回帰分析が主体となる。この解析方法は、“共分散分析”と認識されており、そのために 2 本の回帰直線の傾きが異なる場合には、使えないとの“迷信”がはびこっている。そこで、“共分散分析”ではなく、“体重を交絡変数とする回帰分析”とすることにする。

前節の図 8.13 で示した決定樹は、探索的な解析を行うための考え方が示されている。最初の分岐である「 $X$  と  $Y$  に相関がある？」について、（ない，ある）の判断基準は、明確に示されていない。その後の分岐でも同様に、判断基準は示されていない。反復毒性試験の統計解析は、あらかじめ定められた方法により自動化されているので、実験研究者に、後から探索的な解析を行うことを期待するのは非現実的である。図 8.13 の決定樹、および表 8.12 の解析方法の説明で、③ 組と ⑤ 組が共分散分析、⑥ 組が回帰直線の傾きの差、について新たな解析の必要性を求めている。これらの解析法については、あらかじめ定式化し、毒性試験終了後に必要に応じて速やかに結果が得られるように、解析システムをあらかじめ準備して置くことが望まれる。

Excel の分析ツールに含まれる回帰分析は、1 変数の単回帰分析のみならず 2 変数以上の“重回帰分析”にも対応している。質的変数としての“群”を量的変数としてのダミー変数とすることにより“共分散分析”にも対応できる。さらに、体重と群の交互作用は、群に関するダミー変数と体重の積を新たな変数として与えることで検定統計量を得ることができる。

### 群により回帰直線の傾きが異なる場合

表 8.13 に示す ⑥ 組のデータのように、傾きが異なると視覚的に判断される場合、傾きの差に関する統計量をどのように求めたら良いのであろうか。群 A について (0, 1) 型のダミー変数と体重  $X$  の交互作用を含めた回帰分析を行なうことにより、交互作用のパラメータが傾きの差の推定値となり、その標準誤差  $SE$  との比による  $t$  値が検定統計量となり、 $p$  値も自動的に得られる。傾きの差に加えて、対照群  $A_1$  と処置群  $A_2$  について体重の総平均  $\bar{x}_{..} = 284.50$  における両群の臓器重量の推定値、いわゆる最小 2 乗平均による比較ができれば、結果の考察がしやすくなる。

表 8.20 左に示すように、切片を  $x_0=1$ ，群 A に対して ( $A_1 : a_2=0$ ， $A_2 : a_2=1$ ) とする (0, 1) 型のダミー変数，体重を  $x$ ，群 A と体重  $x$  の交互作用を ( $a_2x$ ) とする回帰式を

$$y_i = \theta_0 x_{0,i} + \theta_1 a_{2,i} + \theta_2 x_i + \theta_3 (a_{2,i} x_i) + \varepsilon_i, \quad \varepsilon_i \sim \text{正規分布} \quad (8.18)$$

とする．これらの変数をデザイン行列  $X$  としてまとめ，Excel による回帰分析を行うと，表 8.20 に示すように分散分析表および回帰パラメータの推定値を得る．交互作用 ( $a_2x$ ) のパラメータは， $\hat{\theta}_3 = 3.3876$  で  $p = 0.0004$  と有意な傾きの差となる．推定された回帰式は，

$$\left. \begin{aligned} \hat{y}^{(6)} &= \hat{\theta}_0 x_0 + \hat{\theta}_1 a_2 + \hat{\theta}_2 x + \hat{\theta}_3 (a_2 x) \\ &= 22.55 - 677.16 a_2 + 2.4318 x + 3.3876 (a_2 x) \end{aligned} \right\} \quad (8.19)$$

であり，対照群  $A_1$  および処置群  $A_2$  の回帰式，

$$\text{対照群 } A_1 : \begin{cases} \hat{y}^{(6, A_1)} = 22.55 - 677.16 \times 0 + 2.4318 x + 3.3876 \times 0 \\ = 22.55 + 2.4318 x \end{cases} \quad (8.20)$$

$$\text{処置群 } A_2 : \begin{cases} \hat{y}^{(6, A_2)} = 22.55 - 677.16 \times 1 + 2.4318 x + 3.3876 \times (1 \times x) \\ = -654.60 + 5.8194 x \end{cases} \quad (8.21)$$

が導出される．これらは，図 8.12 の ⑥ 組の図に示されている回帰式に一致する．

表 8.20 傾きに差がある場合の交互作用モデルの適用

⑥組	群	— デザイン行列 —					$Y_6$	分散分析表 (定数に 0 を使用) off				
No	A	$x_0$	$a_2$	$x$	$a_2 x_0$	$y$			自由度	変動	分散	分散比
1	1	1	0	285	0	731		回帰	3	959842	319947	42.4848
2	1	1	0	382	0	882		残差	16	120494	<b>7530.87</b>	$=\sigma^2$
3	1	1	0	350	0	845		合計	19	1080336		
4	1	1	0	193	0	481				$t_{0.05} =$	<b>2.1199</b>	
5	1	1	0	234	0	607			係数	標準誤差	$t$	$P$ -値
6	1	1	0	275	0	685	$\theta^{\wedge}_0$	切片 $x_0$	<b>22.55</b>	144.52	0.1561	0.8779
7	1	1	0	364	0	1021	$\theta^{\wedge}_1$	$a_2$	<b>-677.16</b>	217.01	-3.1205	0.0066
8	1	1	0	296	0	752	$\theta^{\wedge}_2$	$x$	<b>2.4318</b>	0.4637	5.2444	0.0001
9	1	1	0	308	0	731	$\theta^{\wedge}_3$	$a_2 x$	<b>3.3876</b>	0.7635	4.4367	0.0004
10	1	1	0	373	0	932						
11	2	1	1	322	322	1224		パラメータの共分散行列				
12	2	1	1	229	229	832		$\Sigma(\theta^{\wedge}) = (X^T X)^{-1} \sigma^2$				
13	2	1	1	272	272	847	$\theta^{\wedge}_0$		<b>20887</b>	-20887	-65.80	65.80
14	2	1	1	207	207	395	$\theta^{\wedge}_1$		-20887	<b>47092</b>	65.80	-162.57
15	2	1	1	245	245	703	$\theta^{\wedge}_2$		-65.80	65.80	<b>0.2150</b>	-0.2150
16	2	1	1	186	186	431	$\theta^{\wedge}_3$		65.80	-162.57	-0.2150	<b>0.5830</b>
17	2	1	1	273	273	1019		切片 $x_0$	$a_2$	$x$	$a_2 x$	
18	2	1	1	308	308	1221						
19	2	1	1	260	260	960						
20	2	1	1	328	328	1127						
総平均		$\bar{x}_{..} = 284.50$										

対照群  $A_1$  の切片の推定値は， $\hat{\theta}_0 = 22.55$  ( $p = 0.8779$ ) と有意ではなので，原点を通るとみなせる．処置群  $A_2$  の切片の推定値は， $\hat{\theta}_0 + \hat{\theta}_1 = -654.60$  であり， $p$  値は示されていないが， $\hat{\theta}_1 = -677.16$  の  $p = 0.0066$  からの類推で原点を通るとは見做せない．したがって，両群の傾き

の差  $\hat{\theta}_3 = 3.3876$  のみによる考察は、不十分である。これは、図 8.12 に例示した 傾きが異なる ⑥ 組の線グラフから、体重が 200 g 近傍で、両群間に臓器重量間に差がないが、体重の総平均  $\bar{x}_{..} = 284.50$  g で両群の回帰直線上の推定値に明らかな差が読み取れる。したがって、臓器重量に対する何らかの要約統計量による比較が必要である。なお、絶対臓器重量での 2 群間の比較では、データの範囲が互いに重なり合い、有意な差とはならないことは明らかであり、不適切である。

### 体重の総平均での臓器重量の比較

体重の総平均における対照群  $A_1$  の臓器重量の推定値と、処置群  $A_2$  の臓器重量の推定値（最小 2 乗平均）を求め、統計的な比較をしたい。そのためには、これまでも繰返し示してきたように、パラメータの共分散行列  $\Sigma(\hat{\theta}) = (X^T X)^{-1} \sigma^2$  が必要となる。これは、表 8.20 右下段にデザイン行列  $X$  と分散分析表の残差の分散  $\hat{\sigma}^2 = 7530.87$  を用いて  $4 \times 4$  の行列として示されている。

表 8.21 には、体重の総平均  $\bar{x}_{..} = 284.50$  g における対照群  $A_1$  および処置群  $A_2$  の要約統計量としての推定値 ( $y_{x=284.5}^{(6,A_1)} = 714.4$ ,  $y_{x=284.5}^{(6,A_2)} = 1001.0$ ) および 95%信頼区間が計算されている。さらに、推定値の差 ( $y_{x=284.5}^{(6,A_2)} - y_{x=284.5}^{(6,A_1)} = 286.6$ ) の 95%信頼区間 (197.3, 375.9) が 0 を含まないので有意な差となっている。

表 8.21 傾きに差がある場合の体重の総平均における臓器重量

	線形和	$l_0$	$l_1$	$l_2$	$l_3$	推定値	分散	幅	95%信頼区間	
⑥ 組	$L^{(i)}$	$x_0$	$a_2$	$x$	$a_2 x$	$y^{\wedge} = l\theta$	$Var(y^{\wedge})$	$t_{0.05} \times SE$	$L_{95\%}$	$U_{95\%}$
$A_1$	$L^{(1)}$	1	0	284.5	0	714.4	852.5	61.9	652.5	776.3
$A_2$	$L^{(2)}$	1	1	284.5	284.5	1001.0	923.2	64.4	936.6	1065.4
$A_1 - A_1$	$L^{(3)}$	0	0	0	0	0.0	0.0	0.0	0.0	0.0
$A_2 - A_1$	$L^{(4)}$	0	1	0	284.5	286.6	1775.7	89.3	197.3	375.9
						$t_{0.05}(20-4) =$		2.1199		

図 8.20 左に示す折れ線グラフは、表 8.21 で計算された推定値と 95%信頼区間を描いた結果である。体重の総平均における対照群  $A_1$  と処置群  $A_2$  の推定値の 95%信頼区間が、互いに重なり合わないので、有意な差があることが視覚的に明らかである。図 8.20 右は、対照群  $A_1$  に対して処置群  $A_2$  の臓器重量の推定値が 286.6 mg 増大し、95%信頼区間が 0 を含まないので、有意な差であることが示されている。

これらの推定値を求めるために、変数 ( $x_0$ ,  $a_2$ ,  $x$ ,  $a_2 x$ ) の別名として ( $l_0$ ,  $l_1$ ,  $l_2$ ,  $l_3$ ) を与え、表 8.20 に示したパラメータの推定値  $\hat{\theta}$  との積和による線形和を  $L^{(i)} = l^{(i)} \hat{\theta}$  として定義する。線形和の係数ベクトル  $l^{(i)}$  を用いて対照群  $A_1$  の推定値  $y_{x=284.5}^{(6,A_1)}$  を得るために  $l^{(1)} = [1, 0, 284.5, 0]$  を与え、 $L^{(1)} = l^{(1)} \hat{\theta}$  によって  $y_{x=284.5}^{(6,A_1)} = 714.4$  が推定される。表 8.21 に示したの

は、係数ベクトル  $\mathbf{l}^{(i)}$  に対する線形和  $L^{(i)}$  の推定値と、その分散、95%信頼区間を計算した結果となっている。

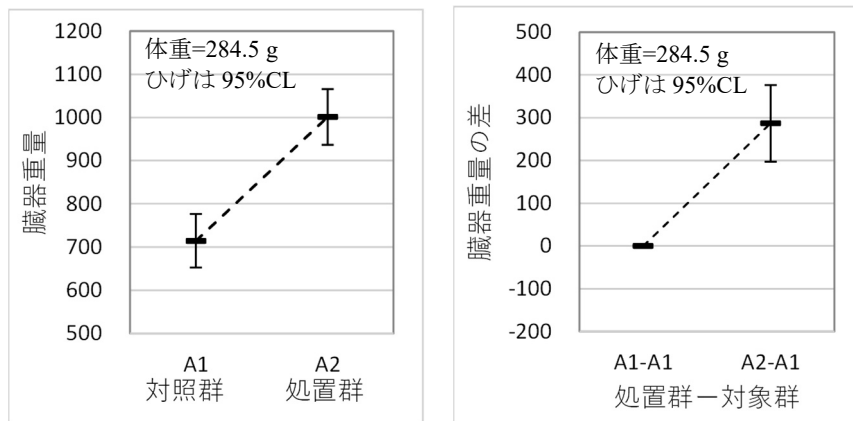


図 8.20 体重が 284.5 g の場合の臓器重量の推定値 (最小 2 乗平均)

実際の計算手順は、次のようである。Excel の行列の積の `Mmult ()` 関数を用いて線形和  $L^{(i)} = \mathbf{l}^{(i)} \hat{\boldsymbol{\theta}}$  は、

$$y_{x=284.5}^{(6,A_1)} = L^{(1)} = \mathbf{l}^{(1)} \hat{\boldsymbol{\theta}} \quad (8.22)$$

=	1	0	284.5	0	22.55	=	714.4
					-677.16		
					2.4318		
					3.3876		

として計算されている。線形和の分散  $\text{Var}(\mathbf{l}^{(1)} \hat{\boldsymbol{\theta}})$  は、パラメータの共分散行列  $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$  に対する  $\mathbf{l}^{(1)}$  の 2 次形式によって、

$$\text{Var}(\mathbf{l}^{(1)} \hat{\boldsymbol{\theta}}) = \mathbf{l}^{(1)} \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}) \mathbf{l}^{(1)T} \quad (8.23)$$

=	1	0	284.5	0	20887	-20887	-65.80	65.80	1	=	852.5
					-20887	47092	65.80	-162.57	0		
					-65.80	65.80	0.2150	-0.2150	284.5		
					65.80	-162.57	-0.2150	0.5830	0		

として求められ、推定値  $y_{x=284.5}^{(6,A_1)} = 714.4$  に対する 95%信頼区間は、

$$\left. \begin{aligned} 95\%CL &= y_{x=284.5}^{(6,A_1)} \pm t(0.05, 16) \sqrt{\text{Var}(\mathbf{l}^{(1)} \hat{\boldsymbol{\theta}})} \\ &= 714.4 \pm 2.1199 \sqrt{852.48} \\ &= (652.5, 776.3) \end{aligned} \right\} \quad (8.24)$$

によって計算されている。ただし、非平行な回帰直線なので、体重を変化させることにより 2 群間の差は、変化することに注意が必要である。

### 予測プロファイル作成

表 8.22 に 2 本の回帰直線の推定値に対する 95%信頼区間が計算されている。計算方法は、表 8.21 に示した体重が  $\bar{x}_{..} = 284.50$  g の場合の対照群  $A_1$  および処置群  $A_2$  の推定値と 95%信

区間の計算と同様であるが、体重を  $x = 0, \dots, 284.5, \dots, 500$  と変化させている．推定値  $y_x^{(6, A_1)}$  および  $y_x^{(6, A_2)}$ ，それらの分散，95%信頼区間の計算法は，表 8.21 と全く同じである．

表 8.22 傾きに差がある場合の回帰直線の 95%信頼区間

	線形和	$l_0$	$l_1$	$l_2$	$l_3$	推定値	分散	幅	95%信頼区間	
⑥ 組	$L^{(i)}$	$x_0$	$a_2$	$x$	$a_2x$	$y^\wedge = \boldsymbol{l\theta}$	$Var(y^\wedge)$	$t_{0.05} \times SE$	$L95\%$	$U95\%$
A <sub>1</sub>	$L^{(5)}$	1	0	0	0	22.6	20887	306.4	-283.8	328.9
回	$L^{(6)}$	1	0	100	0	265.7	9878	210.7	55.0	476.4
帰	$L^{(7)}$	1	0	200	0	508.9	3169	119.3	389.6	628.3
	$L^{(8)}$	1	0	284.5	0	714.4	852	61.9	652.5	776.3
	$L^{(9)}$	1	0	400	0	995.3	2653	109.2	886.1	1104.5
	$L^{(10)}$	1	0	500	0	1238.5	8846	199.4	1039.1	1437.9
A <sub>2</sub>	$L^{(11)}$	1	1	0	0	-654.6	26205	343.2	-997.8	-311.4
回	$L^{(12)}$	1	1	100	100	-72.7	10530	217.5	-290.2	144.9
帰	$L^{(13)}$	1	1	200	200	509.3	2214	99.7	409.5	609.0
	$L^{(14)}$	1	1	284.5	284.5	1001.0	923	64.4	936.6	1065.4
	$L^{(15)}$	1	1	400	400	1673.2	7660	185.5	1487.6	1858.7
	$L^{(16)}$	1	1	500	500	2255.1	21422	310.3	1944.8	2565.4

表 8.23 には，2 本の回帰直線の差の推定値と 95%信頼区間が示されている． $L^{(20)}$  の  $l_3^{(20)}$  が， $a_2x = 233.8$  に対し，95%信頼区間の下限  $L95\%$  が 0 となっているので，体重が 233.8 g において，95%信頼区間の下限  $L95\%$  が 0.0 となり，有意な差となることが示されている．これは，Excel のソルバーで， $L95\%$  が 0 となるように， $l_3^{(20)} = 233.8$  を変化させた結果である．

表 8.23 回帰直線の差についての 95%信頼区間

	線形和	$l_0$	$l_1$	$l_2$	$l_3$	推定値	分散	幅	95%信頼区間	
⑥ 組	$L^{(i)}$	$x_0$	$a_2$	$x$	$a_2x$	$y^\wedge = \boldsymbol{l\theta}$	$Var(y^\wedge)$	$t_{0.05} \times SE$	$L95\%$	$U95\%$
A <sub>2</sub> -A <sub>1</sub>	$L^{(17)}$	0	1	0	0	-677.2	47092	460.0	-1137.2	-217.1
回	$L^{(18)}$	0	1	0	100	-338.4	20407	302.8	-641.2	-35.6
帰	$L^{(19)}$	0	1	0	200	0.4	5383	155.5	-155.2	155.9
	$L^{(20)}$	0	1	0	<b>233.8</b>	114.9	2940	114.9	<b>0.0</b>	229.9
	$L^{(21)}$	0	1	0	400	677.9	10313	215.3	462.6	893.1
	$L^{(22)}$	0	1	0	500	1016.6	30267	368.8	647.8	1385.4

図 8.21 左は，図 8.12 の ⑥ 組のデータで示した体重と臓器重量の散布図に表 8.22 で計算した回帰直線および 95%信頼区間を，Excel の散布図の「データの選択」オプションを使って順次追加し，書式を整えたもので，体重がおおよそ 250 g 以上の場合に，A<sub>2</sub> の  $L95\%$  と A<sub>1</sub> の  $U95\%$  が完全に分離し，統計的な差が明らかになることが読み取れる．

図 8.21 右は，表 8.20 に示されているパラメータの推定値  $\hat{\theta}_1$  と  $\hat{\theta}_3$  についての回帰直線である．図 8.21 左は，2 本の回帰直線の差の推定直線となっている．

$$\begin{aligned}
 \hat{y}^{(6, A_2 - A_1)} &= \hat{\theta}_1 a_2 + \hat{\theta}_3 (a_2 x) \\
 &= -677.16 + 3.3876x
 \end{aligned}
 \tag{8.25}$$



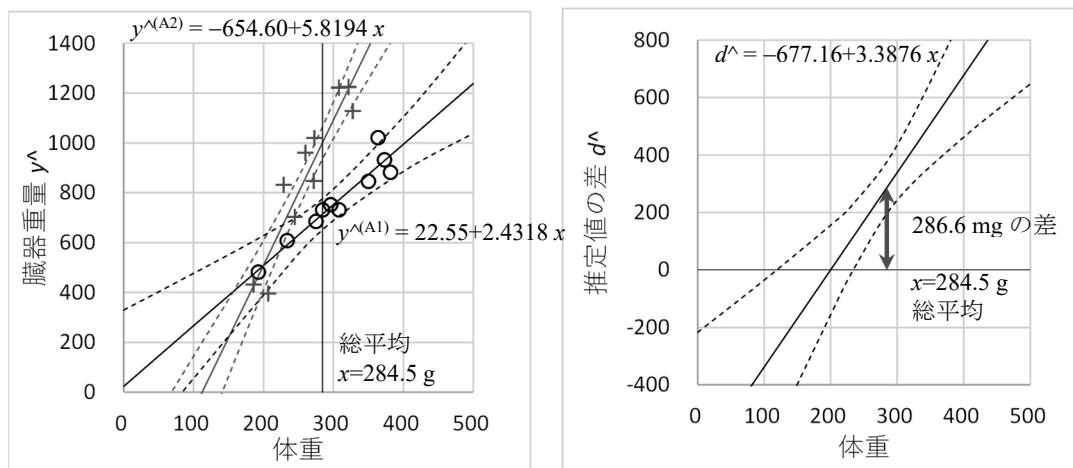


図 8.21 非平行な 2 本の回帰直線および差の回帰直線

また、体重の総平均  $\bar{x}_{..} = 284.50 \text{ g}$  に対する直線の推定値  $286.6 \text{ mg}$  は、図 8.21 左の 2 本の直線間の差となっている。

### 回帰直線が同一直線上と見なせる場合

③ 組のデータは、図 8.12 に示したように、対照群  $A_1$  に対し、処置群  $A_2$  の体重  $x^{(3)}$  および臓器重量  $y^{(3)}$  が共に、明らかに小さくなっているが、それぞれについて回帰直線を重ね書きすると、多少のずれはあるものの同一直線上にあるとみなせ、原点は通らない事例である。

「2 群が同一直線上か」を統計的に判断するためには、群  $A$  と体重  $x$  に交互作用  $A \times x$  がないことを確認し、交互作用  $A \times x$  を除いた（群  $A$ ，体重  $x$ ）を変数とする回帰分析（主効果モデル）で、群間に統計的な差がないことを示す必要がある。③ 組のデータについて、Excel の回帰分析を主体にした解析方法を示す。

表 8.24 左に示すように、群  $A$  を  $(0, 1)$  型ダミー変数とすると処置群  $A_2$  が 1 となるので変数名としては  $a_2$  としている。体重  $X_3$  は  $x$  とし、交互作用を  $a_2x = a_2 \times x$  として計算している。切片に対する変数を  $x_0$  として明示的に与え、 $(x_0, a_2, x, a_2x)$  の 20 行分のデータが“デザイン行列  $\mathbf{X}$ ”となる。臓器重量  $Y_3$  は  $y$  として示す。表 8.24 右上は、交互作用  $a_2x$  を含めたモデルの結果で、表 8.24 右下は、交互作用の変数  $a_2x$  を除いた主効果モデルの結果である。

交互作用モデルで得られた推定式は、

$$\left. \begin{aligned} \hat{y}^{(3)} &= \hat{\theta}_0 x_0 + \hat{\theta}_1 a_2 + \hat{\theta}_2 x + \hat{\theta}_3 (a_2 x) \\ &= 575.59x_0 - 95.90a_2 + 1.1450x + 0.0985(a_2x) \end{aligned} \right\} \quad (8.26)$$

であり、 $A_1$  および  $A_2$  の回帰式は、

表 8.24 回帰直線が同一直線上の場合の交互作用モデル vs. 主効果モデル

③ 組	群	— デザイン行列 $X$ —					$Y_3$	交互作用モデル 推定値, 分散分析表は省略					
No	A	$x_0$	$a_2$	$x$	$a_2x$	$y$			係数	標準誤差	$t$	$P$ -値	
1	1	1	0	355	0	1080	$\theta^{\wedge}_0$	切片 $x_0$	575.59	204.37	2.8164	0.0124	
2	1	1	0	432	0	1107	$\theta^{\wedge}_1$	$a_2$	-95.90	227.60	-0.4214	0.6791	
3	1	1	0	460	0	1015	$\theta^{\wedge}_2$	$x$	1.1450	0.5105	2.2431	0.0394	
4	1	1	0	473	0	1204	$\theta^{\wedge}_3$	$a_2x$	0.0985	0.7016	0.1404	0.8901	
5	1	1	0	374	0	886							
6	1	1	0	375	0	1020		分散分析表 主効果モデル					
7	1	1	0	394	0	1085		自由度	変動	分散	分散比		
8	1	1	0	326	0	972		回帰	2	505458	252729	54.4466	
9	1	1	0	388	0	897		残差	17	78910	4641.78		
10	1	1	0	403	0	1047		合計	19	584369			
11	2	1	1	247	247	844				$t_{0.05} =$	2.1098		
12	2	1	1	210	210	790			係数	標準誤差	$t$	$P$ -値	
13	2	1	1	280	280	827	$\theta^{\wedge}_0$	切片 $x_0$	554.84	137.01	4.0496	0.0008	
14	2	1	1	117	117	683	$\theta^{\wedge}_1$	$a_2$	-65.74	72.99	-0.9006	0.3804	
15	2	1	1	175	175	688	$\theta^{\wedge}_2$	$x$	1.1971	0.3400	3.5213	0.0026	
16	2	1	1	181	181	721		パラメータの共分散行列					
17	2	1	1	233	233	769		$\Sigma(\beta^{\wedge}) = (X^T X)^{-1} \sigma^{\wedge^2}$					
18	2	1	1	148	148	591	$\theta^{\wedge}_0$		18772.4	-9438.9	-46.0007		
19	2	1	1	230	230	710	$\theta^{\wedge}_1$		-9438.9	5327.8	22.5496		
20	2	1	1	208	208	697	$\theta^{\wedge}_2$		-46.0007	22.5496	0.1156		
総平均 $\bar{x}_{..} = 300.5$													

$$\text{対照群 } A_1 : \begin{cases} \hat{y}^{(3,A_1)} = \hat{\theta}_0 \times 1 + \hat{\theta}_1 \times 0 + \hat{\theta}_2 x + \hat{\theta}_3 (0 \times x) \\ = \hat{\theta}_0 + \hat{\theta}_2 x \\ = 575.59 + 1.1450x \end{cases} \quad (8.27)$$

$$\text{処置群 } A_2 : \begin{cases} \hat{y}^{(3,A_2)} = \hat{\theta}_0 x_0 + \hat{\theta}_1 \times 1 + \hat{\theta}_2 x + \hat{\theta}_3 (1 \times x) \\ = (\hat{\theta}_0 + \hat{\theta}_1) + (\hat{\theta}_2 + \hat{\theta}_3) x \\ = 479.69 + 1.2435x \end{cases} \quad (8.28)$$

であり, 傾きの差は  $\hat{\theta}_3 = 0.0985$  ( $p = 0.8901$ ) と有意な差ではない.

表 8.24 右下の交互作用項 ( $a_2x$ ) を除いた主効果モデルによる回帰式は,

$$\left. \begin{aligned} \hat{y}^{(3)} &= \hat{\theta}_0 x_0 + \hat{\theta}_1 a_2 + \hat{\theta}_2 x \\ &= 554.84 \times 1 + (-65.74) \times a_2 + 1.1971x \end{aligned} \right\} \quad (8.29)$$

であり, 群別の回帰式は

$$\text{対照群 } A_1 : \begin{cases} \hat{y}^{(3,A_1)} = \hat{\theta}_0 \times 1 + \hat{\theta}_1 \times 0 + \hat{\theta}_2 x \\ = 554.84 + 1.1971x \end{cases} \quad (8.30)$$

$$\text{処置群 } A_2 : \begin{cases} \hat{y}^{(3,A_2)} = \hat{\theta}_0 \times 1 + \hat{\theta}_1 \times 1 + \hat{\theta}_2 x \\ = 489.10 + 1.1971x \end{cases} \quad (8.31)$$

であり、Y 軸方向の回帰直線の差は  $\hat{\theta}_1 = -65.74$  ( $p = 0.3804$ ) と有意な差ではなく、 $p$  値も 0.4 に近いので、2 本の回帰直線は、「同一直線上」にあると統計的に判断する。

推定された統計量について考察する際には、折れ線グラフ表示が有効である。有意な差がない場合でも、どの程度なのか視覚的に判断することも必要である。表 8.25 に示すように、両群の体重の総平均  $\bar{x}_{..} = 300.5 \text{ g}$  における直線上の推定値を求め、その 95%信頼区間を推定

表 8.25 回帰直線が同一直線上の場合の体重の総平均 300.5 g における推定値

線形和		$l_0$	$l_1$	$l_2$	推定値	分散	幅	95%信頼区間	
③ 組	$L^{(i)}$	$x_0$	$a_2$	$x$	$y^{\wedge}=l\theta$	$Var(y^{\wedge})$	$t_{0.05}\times SE$	$L95\%$	$U95\%$
A <sub>1</sub>	$L^{(1)}$	1	0	300.5	914.5	1564.0	83.4	831.1	998.0
A <sub>2</sub>	$L^{(2)}$	1	1	300.5	848.8	1564.0	83.4	765.3	932.2
A <sub>1</sub> -A <sub>1</sub>	$L^{(3)}$	0	0	0	0.0	0.0	0.0	0.0	0.0
A <sub>2</sub> -A <sub>1</sub>	$L^{(4)}$	0	1	0	-65.7	5327.8	154.0	-219.7	88.3
					$t_{0.05}(20-3)=$		2.1098		

し、図 8.22 左に図示する。平行な直線間の差は、体重  $x$  と関係なく  $-65.7 \text{ mg}$  と推定され、その  $SE$  は、表 8.24 で推定されたダミー変数  $a_2$  の  $SE = 72.99$  に一致する。差 ( $A_2 - A_1$ ) に対する 95%信頼区間を図 8.22 右に図示する。このように、有意な差ではないと結論付けるだけでなく、視覚的に示すことも必要である。

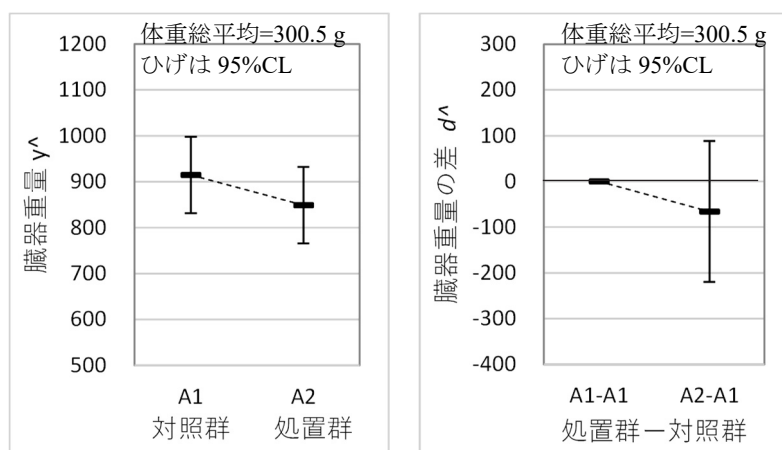


図 8.22 体重の総平均  $\bar{x}_{..} = 300.5 \text{ g}$  における臓器重量および差の推定値

体重の総平均  $\bar{x}_{..} = 300.5 \text{ g}$  における臓器重量および差について、散布図上に平行な直線と 95%信頼区間を重ね書きし、図 8.22 の結果と比較することにより、結果の解釈に際し要約統計量として役立つ。表 8.26 に示すのは、表 8.25 の体重の総平均  $\bar{x}_{..} = 300.5 \text{ g}$  における推定のための計算シートを活用し、0 g から 500 g まで 100 g ごとに引き伸ばした結果である。

表 8.26 回帰直線が同一直線上の場合の 95%信頼区間

		$l_0$	$l_1$	$l_2$	推定値	分散	幅	95%信頼区間	
③ 組	$L^{(i)}$	$x_0$	$a_2$	$x$	$y^{\wedge}=10$	$Var(y^{\wedge})$	$t_{0.05} \times SE$	$L95\%$	$U95\%$
$A_1$	$L^{(5)}$	1	0	0	554.8	18772.4	289.1	265.8	843.9
回	$L^{(6)}$	1	0	100	674.6	10728.1	218.5	456.0	893.1
帰	$L^{(7)}$	1	0	200	794.3	4995.4	149.1	645.2	943.4
	$L^{(8)}$	1	0	300.5	914.6	1562.9	83.4	831.2	998.0
	$L^{(9)}$	1	0	400	1033.7	464.6	45.5	988.2	1079.2
	$L^{(10)}$	1	0	500	1153.4	1666.7	86.1	1067.3	1239.5
$A_2$	$L^{(11)}$	1	1	0	489.1	5222.4	152.5	336.6	641.6
回	$L^{(12)}$	1	1	100	608.8	1688.0	86.7	522.1	695.5
帰	$L^{(13)}$	1	1	200	728.5	465.2	45.5	683.0	774.0
	$L^{(14)}$	1	1	300.5	848.8	1565.2	83.5	765.4	932.3
	$L^{(15)}$	1	1	400	968.0	4954.3	148.5	819.5	1116.5
	$L^{(16)}$	1	1	500	1087.7	10666.2	217.9	869.8	1305.6

図 8.23 に示すのは、表 8.24 の ③ 組のデータについて層別散布図を描き、表 8.26 の推定結果をさらに重ね書きした結果である。体重の総平均  $\bar{x}_{..} = 300.5 \text{ g}$  における回帰直線の推定値と 95%信頼区間が、図 8.22 左に対応する。

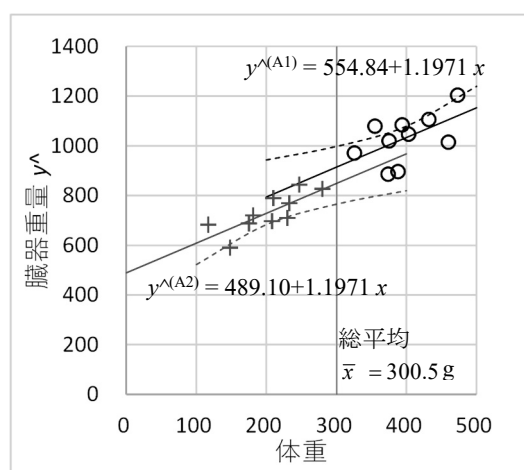


図 8.23 同一直線上の平行な 2 本の回帰直線の 95%信頼区間

回帰直線間の差および 95%信頼区間は、体重  $x$  が変化しても一定であるので、図示するまでもない。このことは、表 8.25 に示したように、処置  $A_2$  と対照  $A_1$  の回帰直線の差の推定に対し、体重  $x$  が 0 となっており、体重に関係なく変数  $a_2$  のパラメータ  $\hat{\theta}_1 = -65.74$  のみによる推定であることから明らかである。さて、ここに示した解析法は、いわゆる“共分散分析”の手順と同じであるが、共変量と見なした交絡変数としての体重の範囲が、処置  $A_2$  と対照  $A_1$  で全く異なり、いわゆる“共分散分析”の適用外である。このように交絡変数の反応に与える影響は、多様であり“明らかな交互作用”がある場合も含め、柔軟な対応が求められる。

## 2本の回帰直線が原点を通る場合

2本の回帰直線が原点を通るとの判定は、別々に回帰直線をあてはめ、切片の95%信頼区間が原点を含むことで対応してきたのであるが、(1, 1) 標示型ダミー変数を用いることにより一括して対応でき、原点を通る2本の回帰直線のあてはめにも応用できる。(1, 1) 標示型ダミー変数は、対照  $A_1$  に対し  $a_1=1$ 、 $A_1$  以外は  $a_1=0$ 、処置  $A_2$  に対し  $a_2=1$ 、 $A_2$  以外は  $a_2=0$  とするダミー変数である。

表 8.27 左に示すように、デザイン行列は対照群  $A_1$  と処置群  $A_2$  それぞれに対する回帰分析の変数を単に並列的に並べた  $(a_1, a_2, a_1x, a_2x)$  形式となっている。ここには、いつもの“切片”のための共通の変数  $x_0$  が含まれていないことに注意してもらいたい。ここでは、 $x_0$  に代えて  $(a_1, a_2)$  が2本の回帰直線のそれぞれの切片を求めるための変数となっている。表 8.27 右上は、Excel の回帰分析で（定数に 0 を使用）オプションを on として求めたパラメータ推定値となっている。推定された回帰式は、

$$\text{対照群 } A_1 \begin{cases} \hat{y}^{(4, A_1)} = \hat{\theta}_{0,1} \times a_1 + \hat{\theta}'_1 \times (a_1x) \\ = 105.09 + 2.7120x \end{cases} \quad (8.32)$$

$$\text{処置群 } A_2 \begin{cases} \hat{y}^{(4, A_2)} = \hat{\theta}_{0,2} \times a_2 + \hat{\theta}'_2 \times (a_2x) \\ = -19.55 + 4.3118x \end{cases} \quad (8.33)$$

表 8.27 2本の回帰直線が原点を通る場合の (1, 1) 標示型ダミー変数による回帰分析

④組	群	— デザイン行列 $X$ —					$Y_4$	(定数に 0 を使用) on, 分散分析表を除く				
No	A	$a_1$	$a_2$	$a_1x$	$a_2x$	$y$			係数	標準誤差	$t$	$P$ -値
1	1	1	0	255	0	881		切片	0.0000	#N/A	#N/A	#N/A
2	1	1	0	382	0	1082	$\theta^{\wedge}_{0,1}$	$a_1$	105.09	111.32	0.9440	0.3592
3	1	1	0	320	0	995	$\theta^{\wedge}_{0,2}$	$a_2$	-19.55	101.93	-0.1918	0.8503
4	1	1	0	403	0	1201	$\theta^{\wedge}_1$	$a_1x$	2.7120	0.3608	7.5171	0.0000
5	1	1	0	234	0	707	$\theta^{\wedge}_2$	$a_2x$	4.3118	0.3782	11.4015	0.0000
6	1	1	0	325	0	970						
7	1	1	0	198	0	621		(定数に 0 を使用) on, ( $a_1$ と $a_2$ )を除くモデル				
8	1	1	0	276	0	852			自由度	変動	分散	分散比
9	1	1	0	288	0	831		回帰	2	21899498	10949749	2172
10	1	1	0	343	0	1112		残差	18	82967	4609	$=\sigma^2$
11	2	0	1	0	287	1174		合計	20	21982465		
12	2	0	1	0	240	1082				$t_{0.05}=$	2.1009	
13	2	0	1	0	282	1097			係数	標準誤差	$t$	$P$ -値
14	2	0	1	0	157	635		切片	0.0000	#N/A	#N/A	#N/A
15	2	0	1	0	325	1423	$\theta^{\wedge}_1$	$a_1x$	3.0458	0.0696	43.7751	0.0000
16	2	0	1	0	311	1341	$\theta^{\wedge}_2$	$a_2x$	4.2411	0.0797	53.2441	0.0000
17	2	0	1	0	263	1269		パラメータの共分散行列				
18	2	0	1	0	178	671		$\Sigma(\beta^{\wedge})=(X^T X)^{-1}\sigma^2$				
19	2	0	1	0	350	1410	$\theta^{\wedge}_1$		0.004841	0.000000		
20	2	0	1	0	238	1047	$\theta^{\wedge}_2$		0.000000	0.006345		
	総平均 $\bar{x}_{..}=$					282.8						

であり、図 8.12 の ④ 組の散布図で、Excel の線形近似により重ね書きした回帰式に一致する。それぞれの切片の推定値に対する  $p$  値から有意な差ではなく、95%信頼区間が原点を含むことが示されている。

原点を通る2本の回帰直線の同時推定は、変数 ( $a_1x$ ,  $a_2x$ ) に対し、(定数に 0 を使用) オプションを on として行うことができる。結果は、表 8.27 右下に示されているように、

$$\text{対照群 } A_1 : \hat{y}^{(4,A_1)} = \hat{\theta}_1 x = 3.0458x \quad (8.34)$$

$$\text{処置群 } A_2 : \hat{y}^{(4,A_2)} = \hat{\theta}_2 x = 4.2411x \quad (8.35)$$

が得られる。共に原点を通る場合には、臓器重量  $y^4$  を体重  $x^4$  で割った“体重比”での比較が正当性を持つのであるが、推定された原点を通る回帰直線を用い、対照群  $A_1$  に対し処置群  $A_2$  の臓器重量の肥大の程度を量的に示したい。

表 8.28 および図 8.24 に示すように、体重の総平均  $\bar{x}_{..} = 282.8$  g における各群の臓器重量は、対照  $A_1$  群で 861.2 mg、処置  $A_2$  群で 1199.2 mg であり、その差は +338.0 mg であり、95% 信頼区間は (275.1~400.8) と明らかな臓器重量を肥大させている。

表 8.28 2本の回帰直線が原点を通る場合の体重の総平均  $\bar{x}_{..} = 282.8$  g における推定値

線形和	$l_1$	$l_2$	推定値	分散	幅	95%信頼区間	
④ 組 $L^{(i)}$	$a_1x$	$a_2x$	$y^4 = l\theta$	$Var(y^4)$	$t_{0.05} \times SE$	$L_{95\%}$	$U_{95\%}$
$A_1$ $L^{(1)}$	282.8	0	861.2	387.0	41.3	819.9	902.5
$A_2$ $L^{(2)}$	0	282.8	1199.2	507.2	47.3	1151.8	1246.5
$A_1 - A_1$ $L^{(3)}$	0	0	0.0	0.0	0.0	0.0	0.0
$A_2 - A_1$ $L^{(4)}$	-282.8	282.8	338.0	894.3	62.8	275.1	400.8
$t_{0.05}(20-2) =$					2.1009		

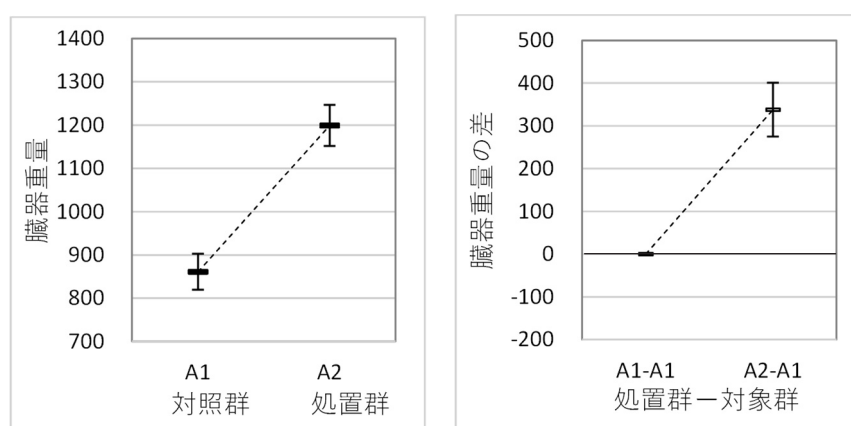


図 8.24 体重の総平均  $\bar{x}_{..} = 282.8$  g の場合の臓器重量および差の推定値

さらに、結果を吟味するために原点を通る回帰直線に 95%信頼区間を重ね書きし、体重の総平均  $\bar{x}_{..} = 282.8$  g における臓器重量の推定値を示したい。表 8.29 に示すのが、体重  $x$  を 0 ~ 500 g まで 100 g ごとに变化させ場合の推定値と 95%信頼区間の計算結果である。一般的な

回帰直線の 95%信頼区間は、データの範囲への外挿に際し広がる曲線なのであるが、原点を通る直線の場合は、まったく異なり 95%信頼区間は、原点を通る直線となっている。奇異な感じがするが、体重が 100 g の場合の対照群の推定値は、 $\hat{y}_{x=100}^{A_1} = 2.7120 \times 100$  であり、分散は、 $Var(a_1x) = x^2 Var(a_1)$ 、標準誤差は、 $SE(a_1x) = x \cdot SE(a_1)$  となるので、95%信頼区間も当然  $x$  に比例し、原点を通る直線となる。

表 8.29 2 本の回帰直線が原点を通る場合の 95%信頼区間

線形和	$l_1$	$l_2$	推定値	分散	幅	95%信頼区間	
④ 組 $L^{(i)}$	$a_1x$	$a_2x$	$y^{\wedge}=l\theta$	$Var(y^{\wedge})$	$t_{0.05} \times SE$	$L95\%$	$U95\%$
A <sub>1</sub> $L^{(5)}$	0	0	0.0	0.0	0.0	0.0	0.0
回 $L^{(6)}$	100	0	304.6	48.4	14.6	290.0	319.2
帰 $L^{(7)}$	200	0	609.2	193.6	29.2	579.9	638.4
$L^{(8)}$	282.8	0	861.2	387.0	41.3	819.9	902.5
$L^{(9)}$	400	0	1218.3	774.6	58.5	1159.8	1276.8
$L^{(10)}$	500	0	1522.9	1210.3	73.1	1449.8	1596.0
A <sub>2</sub> $L^{(11)}$	0	0	0.0	0.0	0.0	0.0	0.0
回 $L^{(12)}$	0	100	424.1	63.4	16.7	407.4	440.8
帰 $L^{(13)}$	0	200	848.2	253.8	33.5	814.7	881.7
$L^{(14)}$	0	282.8	1199.2	507.2	47.3	1151.8	1246.5
$L^{(15)}$	0	400	1696.4	1015.1	66.9	1629.5	1763.4
$L^{(16)}$	0	500	2120.5	1586.2	83.7	2036.9	2204.2

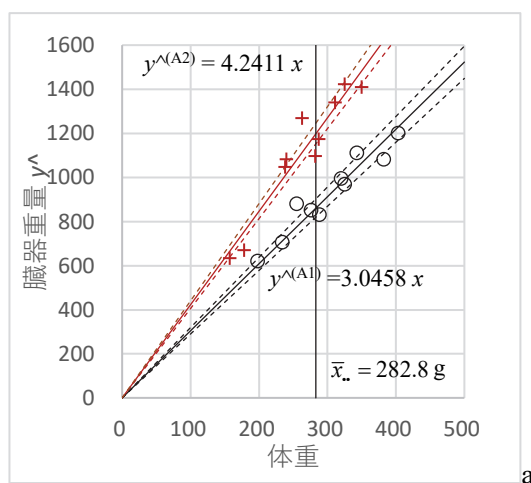


図 8.25 原点を通る 2 本の回帰直線の 95%信頼区間

### 交絡変数を常に念頭にした現実的な対応

反復毒性試験の解剖時の体重と臓器重量を事例とし、交絡変数を常に念頭にした場合の探索的な解析について例示してきた。現実に行われている反復毒性試験では、雄と雌それぞれに対照群に対し最低でも 4 用量の投与群が設定され、対象となる臓器は、約 15 種類あり、前

もって準備されたコンピュータ上の解析システムによって試験終了後速やかに解析報告書が出力され、実験研究者の元に届けられる。このような現実の体制の中で、本章で示した探索的な解析を実験研究者に求めるのは、非現実的である。

図 8.13 に示した「臓器重量の解析のための決定樹」に沿って、解析方法を示してきたのであるが、現実の毒性試験の現場で運用を目指すためには、どうしたら良いのであろうか。通常の解析システムでは、性別・臓器別に対照群と投与群間で絶対重量および体重比での相対重量の平均値に対する多重比較を行っている。着眼点は、対処群に対し有意な差となる投与群がどこからなのかを知ることである。

これまでの検討から、「臓器重量の解析のための決定樹」に沿った解析法は、煩雑で不十分であることを間接的に示してきた。Excel による解析事例として示してきた⑥、③、および、④の事例について共通するのは、体重と臓器重量について層別散布図を描き、交互作用があることを前提とした回帰直線を重ね書きし、体重の総平均におけるそれぞれの回帰直線の推定値および 95%信頼区間、体重の総平均における対処群と処置群の差および 95%信頼区間の計算結果であろう。

SAS の GLM プロシジャを用いることにより、交互作用を含む“共分散分析”であっても群に対する lsmeans ステートメントを用いた最小 2 乗平均が、体重の総平均  $\bar{x}_\cdot$  における推定値と 95%信頼区間を容易に求めることができる。もちろん、JMP の「モデルのあてはめ」によって、SAS と GLM プロシジャと同等の結果が得られる。

いわゆる“共分散分析”の適用に際し、群と共変量の交互作用がないことが、古くから“妄想”のごとく受け継がれている。交互作用がなければ、切片の差が群間の差と見なせるのであるが、交互作用がある場合の群間差と 95%信頼区間の計算は、パラメータの共分散行列  $\Sigma(\hat{\theta})$  を必要とし、第 6 章で示した“手計算”による定式化の範疇外であることに起因している。

これまでに示してきたように、交絡変数の影響は多彩である。交絡変数がある場合には、“共分散分析”を適用すればいいとの短絡的な発想は、“共分散分析”の“交互作用がない場合”に適用できるとの前提条件に絡め取れてしまいがちになる。“交絡変数”は、いわゆる“共変量”ではないことを重ねて注意したい。



## 第8章 文献 索引

高橋ら(1989) - SASによる実験データの解析, 第7章 共分散分析	260
竹内ら(1989) - 統計学辞典, IV分野別応用 4.3 毒性評価	243
吉村編著(1987) - 毒性・薬効データの統計解析, 第3.1節 体重は増加量でみるべきか?	243
- 毒性・薬効データの統計解析, 第3.2節 臓器重量の解析における体重データの利用	260



## 第8章 索引

あ RMSE - 誤差の平均平方	258	か 共変量 - 交絡変数	243, 268
閾値 $p$ 値 - 停止ルール	256	- 前値の影響	248
(1, 1)標示型 - ダミー変数	277	- 胎仔数の影響	249
(1, -1)対比型 - (-1, 1)対比型	258	共変量として前値 - 前後差	250
Excel - Correl () 関数	264	共変量の交互作用 - 中心化	255
- 探索的な回帰分析	250	行/列の切り替え - 折れ線グラフ/ Excel	248
- T.test () 関数	245	行列の積 - Mmult () 関数	271
- DevSq () 関数	245	行列計算の不具合 - 多重共線性	251
- Mmult () 関数	271	組み合わせ - ルール	256
Excel の回帰分析 - 共分散分析	268	グラフビルダー - 格子状に散布図	264
SSE - 誤差の平方和	258	群間の差 - 切片の差	280
$X$ と $Y$ - 相関があるか	264	結果の表示 - 予測プロファイル	259
lsmeansステートメント - 最小2乗平均	280	決定樹 - 解析手法の選択	263
- SASのGLM	280	現実的な対応 - 交絡変数	279
扇型線グラフ - 増加率	247	検定統計量 - 傾きの差	268
折れ線グラフ/ Excel - 行/列の切り替え	248	原点を通る - 回帰直線	266
か 回帰直線 - 傾きが異なる	268	- 95%信頼区間	279
- 原点を通るか	266	- 2本の回帰直線	277
- 同一直線上か	265, 273	交互作用 - 前後差が拡大	247
回帰直線の差 - 95%信頼区間	272	- 中心化	250
回帰分析 - 体重を交絡変数	268	- 複数の共変量	250
- 複数の共変量	250, 255	交互作用の検討 - デザイン行列	251
解析のための決定樹 - 臓器重量	280	交互作用モデル - 傾きの差	269
解析手法の選択 - 決定樹	263	- 主効果モデル	274
確率楕円 - 散布図	264	格子状に散布図 - グラフビルダー	264
化合物 - 毒性	260	交絡変数 - 共変量	243, 268
数多くの変数 - 多重比較	243	- 現実的な対応	279
傾きが異なる - 回帰直線	268	- 総摂取量	244
傾きが同じでない - 2本の直線	267	- 総摂取量の影響	247
傾きに差 - 95%信頼区間	272	- 多彩な影響	260
傾きの差 - 検定統計量	268	- 投与終了時の体重	260
- 交互作用モデル	269	- 2群比較	243
傾きは同じか - 2本の直線	265	誤差の平均平方 - RMSE	258
間接的な作用 - 直接作用	245	誤差の平方和 - SSE	258
95%信頼区間 - 回帰直線の差	272	Correl () 関数 - Excel	264
- 傾きに差	272	- 相関係数	264
- 原点を通る	279	さ 催奇形性 - 生殖発生毒性試験	243
- 切片の推定値	266	最小2乗平均 - lsmeansステートメント	280
- 同一直線上か	276	- 算術平均	259
共分散行列 - 線形和の分散	271	- 線形和の推定値	271
- 2次形式	271	SASのGLM - lsmeansステートメント	280
共分散分析 - Excelの回帰分析	268	差の $t$ 検定 - 対象群と処置群	245
- 高橋ら(1989)	260	算術平均 - 最小2乗平均	259
- 迷信	268	2×2 格子状の散布図 - 複数の共変量	252
- 妄想	280		

さ	散布図 - 確率楕円	264	た	体重を交絡変数 - 回帰分析	268
	- 2×2 区分	254		体重増加データ - 妊娠ラット	244
	JMP - 探索的な回帰分	255		体重比(相対重量) - 臓器の絶対重量	260
	- 二変量の関係	266		対象群と処置群 - 差のt検定	245
	主効果モデル - 交互作用モデル	274		高橋ら(1989) - 共分散分析	260
	種々の関連 - 臓器重量と体重	260		- 臓器重量と体重の関係	260
	主たる共変量 - 補助の共変量	252		竹内ら(1989) - 統計学辞典	243
	飼料効率 - 体重の増加抑制	246		- 毒性評価	243
	ステップワイズ法 - 手作業	257		多彩な影響 - 交絡変数	260
	- 変数選択	255		多重共線性 - 行列計算の不具合	251
	- 変数増加・減少・増減	255		多重比較 - 数多くの変数	243
	- モデルの作成	257		ダミー変数 - (1, 1) 標示型	277
	生殖発生毒性試験 - 催奇形性	243		- (-1, 1) 対比型	256
	成長期 - 実験動物	243		段階的に適用 - 2群間のt検定	245
	切片の差 - 群間の差	280		探索的な回帰分析 - Excel	250
	切片の推定値 - 95%信頼区間	266		探索的な回帰分 - JMP	255
	線グラフ - 前値と後値	247		中心化 - 共変量の交互作用	255
	線形和の推定値 - 最小2乗平均	271		- 交互作用	250
	線形和の分散 - 共分散行列	271		直接作用 - 間接的な作用	245
	前後差 - 共変量として前値	250		直線上の推定値 - 体重の総平均	275
	- 増加率	246		常に探索的 - 臓器重量と体重	264
	- 平行な線グラフ	247		停止ルール - 閾値 $p$ 値	256
	前後差が拡大 - 交互作用	247		T.test () 関数 - Excel	245
	前後差か増加率か - 判断基準	247		デザイン行列 - 交互作用の検討	251
	前値 - 妊娠7日の体重	244		手作業 - ステップワイズ法	257
	前値と後値 - 線グラフ	247		DevSq () 関数 - Excel	245
	前値の影響 - 共変量	248		- 偏差平方和	245
	相関があるか - $X$ と $Y$	264		同一直線上か - 回帰直線	265
	相関係数 - Correl () 関数	264		- 95%信頼区間	276
	増加率 - 扇型線グラフ	247		- 回帰直線	273
	- 前後差	246		統計学辞典 - 竹内ら(1989)	243
	臓器重量の有意差 - 体重の有意差	262		実験動物 - 成長期	243
	臓器の絶対重量 - 体重比(相対重量)	260		投与終了時の体重 - 交絡変数	260
	臓器重量 - 解析のための決定樹	280		毒性 - 化合物	260
	- 体重	260		毒性試験 - 反復投与	243
	臓器重量と体重 - 種々の関連	260		- 反復投与	260
	- 常に探索的	264		毒性評価 - 竹内ら(1989)	243
	- 6 通りの典型例	261	な	2×2区分 - 散布図	254
	臓器重量と体重の関係 - 高橋ら(1989)	260		2区分に分類 - 補助の共変量	252
	- 吉村編著(1987)	260		2群間の $t$ 検定 - 段階的に適用	245
	臓器重量の比較 - 体重の総平均	270		2群比較 - 交絡変数	243
	総摂餌量 - 交絡変数	244		2次形式 - 共分散行列	271
	- 体重の増加抑制	243		二変量の関係 - JMP	266
	- 妊娠 18 日の体重	243		2本の回帰直線 - 原点を通る	277
	総摂餌量の影響 - 交絡変数	247		2本の直線 - 傾きが同じでない	267
た	胎仔数の影響 - 共変量	249		- 傾きは同じか	265
	体重 - 臓器重量	260		2本回帰直線 - 非平行	273
	体重の総平均 - 臓器重量の比較	270		妊娠18日の体重 - 総摂餌量	243
	- 直線上の推定値	275		妊娠7日の体重 - 前値	244
	体重の増加抑制 - 飼料効率	246		妊娠ラット - 体重増加データ	244
	- 総摂餌量	243	は	判断基準 - 前後差か増加率か	247
	- 吉村編著(1987)	246		反復投与 - 毒性試験	243, 260
	体重の有意差 - 臓器重量の有意差	262		非平行 - 2 本回帰直線	273
	体重は増加量でみるべきか - 吉村編著(1987)	243			

は	標準最小 2 乗法 - 予測プロファイル	258	ま	$(-1, 1)$ 対比型 - ダミー変数	256
	複数の共変量 - 回帰分析	250, 255		Mmult () 関数 - Excel	271
	- 交互作用	250		- 行列の積	271
	- $2 \times 2$ 格子状の散布図	252		迷信 - 共分散分析	268
	平行な線グラフ - 前後差	247		妄想 - 共分散分析	280
	偏差平方和 - DevSq () 関数	245		モデルの作成 - ステップワイズ法	257
	変数選択 - ステップワイズ法	255	や	吉村編著(1987) - 臓器重量と体重の関係	260
	変数増加・減少・増減 - ステップワイズ法	255		- 体重の増加抑制	246
	変数増減 - 方向	256		- 体重は増加量でみるべきか	243
	方向 - 変数増減	256		予測プロファイル - 結果の表示	259
	補助の共変量 - 主たる共変量	252		- 標準最小 2 乗法	258
	- 2 区分に分類	252	ら	ルール - 組み合わせ	256
ま	$(-1, 1)$ 対比型 - $(1, -1)$ 対比型	258		6 通りの典型例 - 臓器重量と体重	261

## 第8章 解析用ファイル一覧

サイズ	名前 ^		種類
 93 KB	第08章_01_妊娠ラット		Microsoft Excel ワークシート
 84 KB	第08章_02_妊娠ラット回帰分析		Microsoft Excel ワークシート
 10 KB	第08章_03_ステップワイズ		JMP Data Table
 56 KB	第08章_04_1_臓器重量		Microsoft Excel ワークシート
 10 KB	第08章_04_2_臓器重量_T_相関		JMP Data Table
 16 KB	第08章_04_3_臓器重量_回帰		JMP Data Table
 121 KB	第08章_05_臓器重量_交絡_回帰		Microsoft Excel ワークシート

空白ページ

非売品，無断複製を禁ずる

第 12 回 続高橋セミナー  
層別因子を含む探索的な回帰分析入門

## 第 8 章 交絡変数および共変量を含む 2 群比較

BioStat 研究所 (株)  
〒105-0014 東京都 港区 芝 1-12-3 の1005  
2024 年 2 月 11 日 高橋 行雄  
[takahashi.stat@nifty.com](mailto:takahashi.stat@nifty.com) , FAX : 03-342-8035