《講　座》

統 計 的 推 論 (3)

——計量データでの平均値を主とした扱い，直線のあてはめを中心とした最小2乗法，非線形モデルでの最小2乗法（コンパートメント説を例にして）

佐 久 間　　昭*

15. 不偏分散

計量データ $x \in \Pi\{\mu, \sigma^2\} \to O_n\{x_i\}$ につき，節5では，σ^2 が既知であるとして，μ について推論し，検定や推定につき，

$$t_0 = \frac{|\bar{x}-\mu_0|}{\sigma}\sqrt{n}$$

$$\mu_L{}^U = \bar{x} \pm t_\alpha[\infty]\frac{\sigma}{\sqrt{n}} \tag{59}$$

を求めた ($y \in \Pi\{\eta, \sigma^2\} \to O_n\{y_i\}$ の表記を用いた).

現実には，σ^2 が未知というのが普通で，標本の値だけで，σ^2 に代わるものを求める必要がある.

$$V = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \tag{60}$$

を求めると，$E\{V\} = \sigma^2$，つまり V を平均的にみれば σ^2 になる性質があるので，**V を不偏分散推定量**といい，具体的に求めた値を不偏分散推定値という．単に不偏分散とか分散ということも多い．$E\{x\} = \mu$, $V\{x\} = \sigma^2$ といった演算子が，しばしば使われ，それぞれ，平均値と分散を示す．

直観的には，

$$V^* = \frac{\sum(x_i-\bar{x})^2}{n}$$

がよさそうであるが，$E\{V^*\} < \sigma^2$ となり，小さく偏っている．

V の分子は展開するとわかるが，$\sum(x_i-\bar{x})^2$ を直接に求めないで，

$$S_{xx} = \sum_{i=1}^{n} x_i{}^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$$

$$= \sum_{i=1}^{n} x_i{}^2 - n\bar{x}^2 \tag{61}$$

として求める方がよい．S_{xx} を (偏差)**平方和**と呼んで，$(\sum x_i)^2/n = n\bar{x}^2$ を**修正項** CF という．S_{xx} の脚符 xx は，x の2乗の感じを示すものである．$\sum x_i{}^2$ では大きすぎるので，平均値に関して $n\bar{x}^2$ の修正を加えているため，修正項の名がある．V の分母の $(n-1)$ は**自由度**と呼ばれ，2乗の項がいくつあるか，つまり，$\sum x_i{}^2$ が n 個，これから $n\bar{x}^2$ を引いて $(n-1)$ 個あることを示すと考えておくとよい．

対になったデータ (x_i, y_i) について，(偏差)**積和**として，

$$S_{xy} = \sum(x-\bar{x})(y-\bar{y})$$

$$= \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}$$

$$= \sum x_i y_i - n\bar{x}\bar{y} \tag{62}$$

が定義されているが，$x_i = y_i$ となった特殊なケースが式(61)の S_{xx} である．S_{xy} は，x_i と y_i との，かかわり，もう少し正確にいえば，直線的なかかわりを示す量で，これを規準化すると，

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \tag{63}$$

で，(単)相関係数という．$x_i = y_i$ では $r=1$ である．S_{xx}, S_{yy} は平方和のとおり，負にはならないが，S_{xy} したがって r は負にもなる．$x_i = -y_i$ では，$r = -1$ となる (→ 節21).

* 東京医科歯科大学難治疾患研究所臨床薬理学部門

【例 12】

$O_5\{4, 7, 9, 11, 14\}$ について，$\bar{x} = 45/5 = 9.0$, $\sum x_i^2 = 463$ である．

$$S_{xx} = 463 - 5\times 81 = 58$$

$x_i^* = x_i - 10$ とすれば，$O_5^*\{-6, -3, -1, 1, 4\}$ で，$\bar{x}^* = -5/5 = -1$, $\sum x_i^{*2} = 63$ で，

$$S_{x^*x^*} = 63 - 5\times(-1)^2 = 58$$

となる．一般に，$x_i^* = ax_i \pm b$ とおくと，

$$S_{x^*x^*} = a^2 S_{xx}$$

となり，b は関係しない．

分散 σ^2 未知の場合には，少し手間がかかるが，$O_n\{x_i\}$ の標本の値だけから V を求めたうえ，式(59) の σ には $\sqrt{V}$ を用いることができる．しかし，S/N 比について，

$$t_0 = \frac{|\bar{x} - \mu_0|}{\sqrt{V}}\sqrt{n} \tag{64}$$

をつくると，$H_0: \mu = \mu_0$ の条件下に，t_0 は $t \in N\{0, 1^2\}$ の実現値にはならない．

$x \in N\{\mu, \sigma^2\} \to O_n\{x_i\}$ の標本に関して，$H_0: \mu = \mu_0$ のもとに，式(64) の t_0 は，自由度 $[n-1]$ の **Student の t 分布**の実現値となる．このため，式(64) の扱いにあたり，つぎのような注意が必要になる．すなわち，平均値の議論にあたっては，

1) $x \in \Pi\{\mu, \sigma^2\}$ と，正規分布にしたがわない場合，分布がほぼ対称であれば，n が 5～10 をこえるあたりから，分布がやや歪んでいるならば，n が 15～30 をこえるあたりからは，$x \in N\{\mu, \sigma^2\}$ の正規分布での議論が利用できる．

2) 分散未知の場合，$O_n\{x_i\}$ で求めた $V = S_{xx}/(n-1)$ を σ^2 の代わりに用いると，正規偏位 $t_\alpha[\infty]$ に関する議論を，Student の t 分布での限界値 $t_\alpha[n-1] (> t_\alpha[\infty])$ に関する議論におきかえることができる．

16. 平均値の推論

$x \in \Pi\{\mu, \sigma^2\} \to O_n\{x_i\}$ につき，n がある程度の大きさであれば，$H_0: \mu = \mu_0$, $H_1: \mu = \mu_1 \neq \mu_0$ の検定にあたり，

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = \sum x_i^2 - n\bar{x}^2$$

$$V = \frac{S_{xx}}{n-1} \tag{65}$$

$$t_0 = \frac{|\bar{x} - \mu_0|}{\sqrt{V}}\sqrt{n}$$

を求め，自由度 $[n-1]$ の t 分布での限界値と比べて $t_0 \geqq t_\alpha[n-1]$ となれば，水準 α で有意とし，H_0 を捨てて，H_1 を採用する．

平均値の $100(1-\alpha)\%$ の信頼限界は，

$$\mu_L{}^U = \bar{x} \pm t_\alpha[n-1]\sqrt{\frac{V}{n}} \tag{66}$$

となる．

いずれの場合も，σ を $\sqrt{V}$ でおきかえ，$t_\alpha[\infty]$ を $t_\alpha[n-1]$ でおきかえて，形式的には分散既知の節 5 の議論が利用できる．

$x \in \Pi\{\mu, \sigma^2\}$ の分布型に問題がある場合，**ノンパラメトリック手法**，ノンパラ手法を用いることが多い．これには，いくつかの方法があるが，代表的なものをあげておく．

まず，$O_n\{x_i\}$ を $x_{(1)} < x_{(2)} < \cdots < x_{(i)} < \cdots < x_{(n)}$ と大きさの順に並べ，順位をつける．ときに同位 $x_{(i)} = x_{(i+1)} = \cdots = x_{(i+t-1)}$ があれば，これらには平均順位，

$$\bar{r} = (i-1) + \frac{t+1}{2} \tag{67}$$

を与える．つまり，第 i から始まる t 個の同位には，すべて $\bar{r}$ を与える．これによって全体の順位和 T_G は $n(n+1)/2$ に保たれる．この辺は節 14の式(56) と同じ要領である．

つぎに，

$$m = \frac{n+1}{2}$$

を求め，$\bar{x}' = x_{(m)}$ とする．つまり，第 m 番目のデータを $\bar{x}'$ とし，これを**中央値** median とするが，母集団でちょうど真中に位置する μ' の推定値に

なる．分布が対称ならば $\mu=\mu'$ であるが，歪んだ分布では，μ' の方が代表性の意味が強いだろう．(i) n が奇数ならば $\bar{x}'=x_{(m)}$ が直ちに求まるが，(ii) n が偶数ならば $\bar{x}'=(x_{(m-0.5)}+x_{(m+0.5)})/2$ と約束する．

中央値の信頼限界を求めるには，

$$r=\frac{n+1}{2}-\frac{t_\alpha[\infty]}{2}\sqrt{n} \tag{68}$$

を計算し，r は整数または整数 +0.5 と半整数に切り捨てた値とする．95%信頼係数では，$r=m-\sqrt{n}$ としてよかろう．そのうえで，小さい方，大きい方から r 番目の値を探し，

$$\mu'_L{}^U=x_{(r)};\ x_{(n+1-r)} \tag{69}$$

を信頼限界とする．

H_0: $\mu'=\mu_0'$, H_1: $\mu'=\mu_1'\neq\mu_0'$ の中心位置のずれの検定に相当するものに，**Wilcoxon の 1 標本検定**，符号つき順位検定がある．まず $O_n\{x_i\}$ につき，$(x_i-\mu_0')$ をつくり，0 のものを除いて，あらためて n を定める．絶対値 $|x_i-\mu_0'|$ に関して順位づけを行ない，同位があれば平均順位を与える．そのうえで，正のものと負のものに分けて順位和 T_+, T_- をとり，$T_++T_-=T_G=n(n+1)/2$ を確かめる．

$$\begin{aligned} A&=\frac{1}{2}T_G,\qquad K=2n(n+1)(2n+1)\\ \kappa&=1-\frac{\sum(t^3-t)}{K}\\ t_0&=\frac{|T-A|-0.5}{\sqrt{\frac{\kappa}{48}K}} \end{aligned} \tag{70}$$

を計算するが，κ の $\sum(t^3-t)$ は，同位のグループが，あちらこちらにあれば，それぞれにつき $(t^3-t)=(t-1)t(t+1)$ を求めて加える．t_0 の T は，T_+, T_- のどちらを用いてもよい．$t_0\geqq t_\alpha[\infty]$ のとき水準 α で有意とし，中心位置のずれを認める．$t_0>2$ のときには，分子の 0.5 を省いて計算しなおして t_0^* とすれば，実質水準をみる場合に，t_0^* の方が近似がよい．κ は，同位によって，ノイズが小さくなるための修正を意味し，同位がなければ $\kappa=1$ となる．

【例 13】

$O_{11}\{1,-1,2,3,0,1,2,1,3,2,1\}$ につき $\bar{x}=15/11=1.3636$, $\sum x_i^2=35$, $S_{xx}=14.5455$, $V=1.4545$, $SD=\sqrt{V}=1.2060$, $SE\{\bar{x}\}=\sqrt{V/n}=0.3636$ などとなる．

H_0: $\mu=\mu_0=0$, H_1: $\mu=\mu_1\neq0$ の検定は，

$$t_0=\frac{1.3636-0}{0.3636}=3.750>t_{0.05}[10]=2.228$$

で有意となる．95% 信頼限界は，

$$\mu_L{}^U=1.3636\pm2.228\times0.3636=0.554;\ 2.174$$

Wilcoxon の 1 標本検定では，$\mu_0=0$ を引いても $O_n\{x_i\}$ はそのままであるが，x=0 を除いた10個につき，絶対値に順位づけをし，

$$R_{10}\{3,3,3,3,3,7,7,7,9.5,9.5\}$$

で $T_-=3$, $T_+=52$, $T_G=55=10\times11/2$. t=5, t=3, t=2 の同位が 1 組ずつあることに注意して，A=27.5, K=4620, $\kappa=0.9675$.

$$\begin{aligned} t_0&=\frac{|3-27.5|-0.5}{\sqrt{\frac{0.9675}{48}\times4620}}\\ &=2.487>t_{0.05}[\infty]=1.960\\ t_0^*&=2.539 \end{aligned}$$

となり，5% 水準で有意となり，母集団の中心位置は $\mu_0=0$ からずれており，大きい．

もとの n=11 につき，m=12/2=6 で，第 6 番目の値を中央値 $\bar{x}'=1$ とする．95% 信頼限界を求めるため，$r=6-\sqrt{11}=2.5$ と切りすてて，$\mu_L'=(0+1)/2=0.5$; $\mu_U'=(2+3)/2=2.5$を得る．

一般に，$x\in N\{\mu,\sigma^2\}$ と正規性がある場合にノンパラ手法を用いるとき効率，感度がおちる．Wilcoxon の 1 標本検定は，しかし，t 分布を用いたいわゆる t 検定に比べて，それほど感度は悪くない．ここで用いた中央値の信頼限界の推定法は，信頼間幅が，t 分布の利用の場合よりも少し広く

なるが，nが小さいときには，それほど感度はおちない．

17. 2つの平均値

対応のない2群データ，$O_m\{x_i\}$, $O_n\{y_i\}$ について，母平均値 μ, η の比較にあたり，

$$V_x=\frac{S_{xx}}{m-1},\quad V_y=\frac{S_{yy}}{n-1}$$

を求め，$V_x \geqq V_y$ とする．$V_x<V_y$ となったなら，x と y を交換して扱う．F分布を利用し，

$$F_0=\frac{V_x}{V_y}<F_{0.025}[m-1,n-1] \tag{71}$$

となれば，消極的に両群の分散は等しいとして，以下の手法が利用できる，F_0 は，等分散のときには，1よりもひどく大きくはならない性質の，2乗の形式のS/N比である．分子と分母につき，それぞれ第1，第2自由度であるが，その水準2.5%の限界値と F_0 を比べる．

不等分散のとき，$\bar{x}$ なり $\bar{y}$ なりの平均値によって分散 V_x, V_y が変化するという場合がある．分散が $k\mu$ の傾向のときには，$\sqrt{x+0.25}$; $k\mu^2$ の傾向のときには $\log x$ ないし $\log(x+1)$ などに変換して扱うとほぼ等分散になるが，この変換値での差の意味を，あらためて考える必要がある．

等分散の場合，共通の分散を求めなおして，平均値の差を吟味するが，検定の S/N 比である t_0 なり，信頼限界の係数 t_α なりは，自由度 $[m+n-2]$ で考える．つまり，共通の分散 V の分子は，$(m-1)+(n-1)$ の2乗項からできている．

$$V=\frac{S_{xx}+S_{yy}}{m+n-2}=\frac{\sum x_i^2-n\bar{x}^2+\sum y_j^2-n\bar{y}^2}{m+n-2}$$

$$t_0=\frac{|\bar{x}-\bar{y}|-\delta_0}{\sqrt{V\left(\dfrac{1}{m}+\dfrac{1}{n}\right)}} \tag{72}$$

$$\delta_L{}^U=(\bar{x}-\bar{y})\pm t_\alpha[m+n-2]\sqrt{V\left(\frac{1}{m}+\frac{1}{n}\right)}$$

t_0, $\delta_L{}^U$ の $\sqrt{V(1/m+1/n)}$ は，差 $(\bar{x}-\bar{y})$ のノイズ $SE\{\bar{x}-\bar{y}\}$ である．t_0 は，$H_0: \mu-\eta=\delta=\delta_0$ について書いたもので，2群に差ありや，の場合には $H_0: \mu=\eta, \delta_0=0$ となるので，t_0 の δ_0 を省略することになる．

なお，2標本t検定で，水準0.05の付近では，節5の式(19)にあたる，検定特性の式は，近似的に，

$$\lambda=\frac{\delta}{\sigma}\div\sqrt{\frac{m+n-2}{(m-1)(n-1)}}(t_\alpha[\infty]+t_\beta[\infty]) \tag{73}$$

となる．

ノンパラ手法で，対応のない2群の**中央値の差**の推定を行なうには，各群の中央値 $\bar{x}', \bar{y}'$ を求めたうえ，点推定を $d'=\bar{x}'-\bar{y}'$ とする．ついで，近似的な区間推定では，

$$w=\frac{t_\alpha[\infty]}{2}\sqrt{\frac{mn}{m+n}},\quad w=\sqrt{\frac{mn}{m+n}}(95\%) \tag{74}$$

を求め，O_m で $(m+1)/2\mp w=r_1; r_2$, O_n で $(n+1)/2\pm w=s_1; s_2$ を求めたうえ，

$$\delta'{}_L{}^U=x_{(r_1)}-y_{(s_1)};\ x_{(r_2)}-y_{(s_2)} \tag{75}$$

とすればよい．

効率のよい2標本検定として，**Wilcoxon の2標本検定**があるが，まず，両群をこみにして順位づけをしたうえ，各群の順位和を求めて S/N をつくるものであるが，節14の後半に記した．この方法は，平均値というよりも，中心位置の比較であり，データの分布についての注文は少ない．

【例 14】

$O_6\{48, 54, 51, 42, 46, 49,\}$, $O_5\{47, 59, 55, 53, 58\}$ につき $\bar{x}=48.3333$, $S_{xx}=85.3333$, $V_x=17.0667$, $\bar{y}=54.4000$, $S_{yy}=91.2000$, $V_y=22.8000$. $F_0=V_y/V_x<F_{0.025}[4,5]=7.388$ で等分散とみて，

$$SE\{\bar{x}-\bar{y}\}=\sqrt{19.6148\left(\frac{11}{30}\right)}=2.682$$

$$t_0=\frac{|48.3333-54.4000|}{2.682}=\frac{6.0667}{2.682}$$

で，$t_{0.05}[9]=2.262$ と同じく，ちょうど5%で有意．引き算を $\bar{y}-\bar{x}$ とおきなおし，

$$\delta_L{}^U = 6.0667 \pm 2.262 \times 2.682 = 0;\ 12.133$$

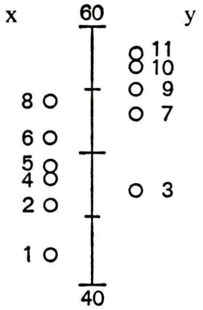
検定と推定が裏腹になっている．

中央値は $\bar{x}'=48.5$, $\bar{y}'=55$ で，差は $d'=-6.5$ となる．$w=\sqrt{30/11}=1.65$ であるから，$r_1=1.5$, $r_2=5$; $s_1=4.5$, $s_2=1$ としたうえ，

$$\begin{aligned}\delta'_L{}^U &= 44-58.5;\ 51-47\\ &= -14.5;\ 4\end{aligned}$$

となるが，$\bar{y}'-\bar{x}'$ の差では 6.5(−4; 14.5) となり，信頼間幅は，かなり広い．

両群のデータを，タテ軸の左右にプロットして順位をつける．



$R_6\{1, 2, 4, 5, 6, 8\}$, $R_5\{3, 7, 9, 10, 11\}$ の順位データとなるが，同位はない．$T_6=26$, $T_5=40$, $T_G=66=11\times12/2$ である．T_6 に着目して，式 (57) から $n_{1\cdot}$, $n_{2\cdot}$ を m, n と読みかえたうえ，$A=66\times6/11=36$ を用い，

$$t_0 = \frac{|26-36|-0.5}{\sqrt{\frac{1}{6}\times5\times36}} = 1.734 < t_{0.05}[\infty] = 1.960$$

で，5% 水準で有意ではなく，両群の中心位置にずれがあるとは言いかねる．$t_0=1.734$ の実質水準は $P=0.083$ である．組み合わせの議論から直接に確率計算をすると，$P=0.0823$ となり，t_0 の近似はよい．

2 標本 t 検定では 5% で有意となったが，Wilcoxon の 2 標本検定では有意にならない．一つには，順位検定の検出力が劣っていたことによるのであろうが，5% ぎりぎりで有意であったので，より強い確信を求めるには，再試験を必要とするであろう．

$\lambda=1.5$ 程度を現実的に意味のある差と考えていたとすれば，式 (73) から，

$$1.5 = \sqrt{0.45}(1.960)+t_{2\beta}[\infty])$$

$$t_{2\beta} = 0.276$$

となり，$\beta \doteqdot 0.4$ で，この例数設計では，検出力 60% となり，そもそも，少なすぎていた．本来 1.5σ 程度の差を 90% で検出するには，各群とも 9〜10 例を用いることが手堅かったといえる．

式 (71) で，**不等分散**と判断されたときには，2 標本 t 検定の手法は利用できなくなる．$m=n$ のときには，形式的に等分散としての式 (72) を用いるが，$t_0 \geqq t_\alpha[n-1]$ のとき有意とする．

$m<n$ と例数が不揃いのときには

$$t_0 = \frac{|\bar{x}-\bar{y}|-\delta_0}{\sqrt{\frac{V_x}{m}+\frac{V_y}{n}}} \tag{76}$$

を求めたうえ，

(1) $t_0<t_2[m+n-2]$ では有意ではない，

(2) $t_0>t_\alpha[m-1]$ では有意である，

(3) 以上で未決定なら，t_0 を重みづけをした

$$t_\alpha{}^* = \frac{\frac{V_x}{m}t_\alpha[m-1]+\frac{V_y}{n}t_\alpha[n-1]}{\frac{V_x}{m}+\frac{V_y}{n}} \tag{77}$$

と比べて判断する．この方法は Cochran によるものであるが，他にも方法がある．

ときに，η/μ の**比**に関心をもつことがあろう．等分散の扱いができるときには，

$$g = \left\{\frac{t_\alpha[m+n-2]}{\bar{x}}\right\}^2 \frac{V}{m}$$

$$\rho_L{}^U = \frac{1}{1-g}\left\{\left(\frac{\bar{y}}{\bar{x}}\right) \pm \frac{\bar{x}}{t_\alpha[m+n-2]} \times \sqrt{V\left[\left(\frac{\bar{y}}{\bar{x}}\right)^2\frac{1}{m}+\frac{1-g}{n}\right]}\right\} \tag{78}$$

によって比の信頼限界が与えられるが，$g<0.1$ では $g=0$ とおいてよいとされる．式 (78) は，Fieller

の式と呼ばれるものの一つである．

【例 15】

例14の $O_6\{x_i\}$, $O_5\{y_j\}$ のデータでは $r=\bar{y}/\bar{x}=1.1255$ となる．$g=0.0072$ であるが，これを省略しないで式 (78) を用いると，

$$\rho_L{}^U=\frac{1}{0.9928}[1.1255\pm0.1327]=1.000;\ 1.267$$

となる．比が1は差が0に対応する．

18. 実験計画法

観測値yが，いくつかの原因系で左右され，また，確率的なゆらぎに影響されるという考えのもとに，節7では，モデルをつくって考察した．モデルを一般的に書けば，

$$y_i=\beta_0+\beta_1x_{i1}+\cdots+\beta_jx_{ij}+\cdots+\beta_px_{ip}+\varepsilon_i \tag{79}$$

といかめしくなるが，**線形モデル**と呼ばれるものになる．

対応のない2群比較で，$O_m\{x_i\}$, $O_n\{y_j\}$ としたが，これらを一括して $O_{m+n}\{y_i\}$ と表現し，y_i が第1群なら x_{i1} は1, x_{i2} は0となり，y_i が第2群なら x_{i1} は0, x_{i2} は1となる約束をおけば，

$$y_i=\beta_0+\beta_1x_{i1}+\beta_2x_{i2}+\varepsilon_i \tag{80}$$

のモデルになり，第1群，第2群の母平均値は $(\beta_0+\beta_1)$, $(\beta_0+\beta_2)$ となる．β_0 を全体の平均値とすれば，β_1, β_2 は，この β_0 からのずれになり，両群の平均値の差が δ なら，β_1, β_2 は $\pm\delta/2$ となる．

一方，ある個体の特性，たとえば年齢 x_i が y_i に直接的に影響するなら，これを共変数，説明変数といい，y_i を目的変数という．

$$y_i=\beta_0+\beta_1x_i+\varepsilon_i \tag{81}$$

が，この際のモデルになり，$P_i\{x_i, y_i\}$ のプロットに直線を通したときの切片が β_0, 勾配が β_1 になる．これも式 (79) の特殊な場合とみられる．

線形モデル式(79)は一貫して，最小2乗法の原理で処理できるが，実際上は，x_{ij} の性質によって区別して扱う習慣がある．

1) 分散分析 ANOVA では，x_{ij} は本質的に1か0をとり，β_j が y_i の変動の原因系になっているか，いないかの指標となる．

2) 回帰分析 REGAN では，x_{ij} は独立変数的であり，目的変数 y_i の変動を説明している性格になる．普通は x_{ij} を留めて，y_i を変量として扱う．

3) 共分散分析 ANOCVA では，x_{ij} の一部は指標的，一部は独立変数的であり，この共変数の影響を除いて y_i の議論をしたり，また共変数そのものの性質を検討したりの場合に利用される．

こうした線形モデルの性質を基盤に，転がりこんでくるデータを待つという態度よりも，積極的に現場に働きかけて，データを取りにいくという態度を重視して組み立てられたものが，**実験計画法**と呼ばれている体系であり，R. A. Fisher の思想の上に発展した．

実質科学的に重要な原因系のいくつかを因子として取りあげ，これを因子水準として小わけし，観測特性との関係において，水準間の比較，因子相互のかかわりなどを，実施可能な範囲で，経済的に，正しく，感度よく捉えることをねらう．

(1) 観測特性，必要なら共変数を指定し，観測条件，方法を明確にする．

(2) 観測の単位を定め，必要に応じて，反応性の類似した単位の集合をブロックとする．

(3) 必要な因子を定め，定量的，半定量的，定性的な水準を設けるが，研究目的の因子，偏り防止の因子，誤差制御の因子，手法上のダミー因子などが区別される．

(4) 因子水準そのもの，異なった因子間の水準の組合わせを処理といい，具体的に取りあげる処理を定め，これと観測単位との対応，割りつけを合理的に決める．

(5) こうして，正当な推論を行なうにふさわしい解析方法，例数と実施可能性を検討し，必要なら，(1)に戻って再吟味する．

19. 分 散 分 析

対応のない2群比較を多群に拡張し，a水準の

処理 $A_1, \cdots, A_i, \cdots, A_a$ につき，それぞれ r 回の同数観測を行なう．ただし，n=ar の観測単位に，$A_1, \cdots, A_a$ を無作為に割りつけるものとする．

A_1	y_{11}	$\cdots$	y_{1j}	$\cdots$	y_{1r}	T_1
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
A_i	y_{i1}	$\cdots$	y_{ij}	$\cdots$	y_{ir}	T_i
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
A_a	y_{a1}	$\cdots$	y_{aj}	$\cdots$	y_{ar}	T_a

$$T_G = \sum_{i=1}^{a} T_i$$

$$y = \frac{T_G}{ar}$$

この形式のデータは，1因子完全無作為化法によるものであるが，A という因子で分類されている観点から**一元配置**のデータと呼んでいることが多い．

y_{ij} は，第 i 水準 A_i のたまたま第 j 番目の観測値である．各水準での和 T_i は，$r\bar{y}_i$ で，A_i の効果の平均値を反映するシグナルである．各 A_i 内でのデータは等分散であるとすれば，これらのばらつきをプールしたものは，同じ処理でのばらつきであるから，ノイズにほかならない．そこで，$T_i=r\bar{y}_i$ 間のばらつきをシグナルとしてまとめ，S/N 比にもっていこうというのが，一元配置の解析方針である．

第 i 処理のデータについては，

$$y_{ij} = (\beta_0+\beta_i)+\varepsilon_{ij}$$

で，$x_{ij}=1$ であるが，他の $\beta_1, \cdots, \beta_a$ に関する指標変数は，すべて 0 となっている．

計算手順から示すと，

$$\begin{aligned} CF &= \frac{T_G^2}{ar} \\ S_{AR} &= \sum_i \sum_j y_{ij}^2 - CF \quad [ar-1] \\ S_A &= \frac{1}{r} \sum T_i^2 - CF \quad [a-1] \\ S_{R(A)} &= S_{AR} - S_A \quad [a(r-1)] \end{aligned} \tag{82}$$

ここに，S_{AR} は全体，S_A は処理，$S_{R(A)}$ は誤差相当の残差に関する平方和であり，

$$\sum_i \sum_j (y_{ij}-\bar{\bar{y}})^2 = r \sum_i (\bar{y}_i-\bar{\bar{y}})^2 + \sum_i \sum_j (y_{ij}-\bar{y}_i)^2$$

のそれぞれに対応しているが，式(82)の計算法が簡単である．ここで重要なことは，平方和と，それに対する自由度の双方に，加法性，

$$\begin{aligned} S_{AR} &= S_A + S_{R(A)} \\ [ar-1] &= [a-1]+[r(a-1)] \end{aligned} \tag{83}$$

が成立する点にある．

ついで，$V_A=S_A/(a-1)$, $V_{R(A)}=S_{R(A)}/(ar-a)$ と，平方和／自由度として，分散を求めたとき，$V_{R(A)}$ は，各処理群内のばらつき ε に関係した誤差分散 σ^2 を推定するが，V_A は，σ^2 と各処理群間のばらつきに関係した σ_A^2 の 2 つの要素から成る．

H_0: $A_1, \cdots, A_a$ の処理効果は一様，H_1: 少なくとも一つは他と異なる，とおいたとき，H_0 が正しいなら $\sigma_A^2=0$ となるから，

$$F_0 = \frac{V_A}{V_{R(A)}} \tag{84}$$

の分散比は，本質的には σ^2/σ^2 で，1 を大きくこえることはあるまい．100 回のうち，たかだか 5 回程度はこえるかもしれないという，F_0 の限界値が，$F_{0.05}[(a-1), (ar-r)]$ である．そこで，F_0 がこれより大になれば，5% 水準で有意とし，H_0: 効果の一様性をすてて，$\sigma_A^2 \neq 0$ と考え，H_1: 少なくとも 1 つの処理は他と異なると考える．

式(82)の値は，分散分析表として整理されるが，Ms は平均平方和で，分散を指す．

要因	SS	Df	Ms	F_0
A: 処理	S_A	$a-1$	V_A	$V_A/V_{R(A)}$
R(A): 残差	$S_{R(A)}$	$ar-a$	$V_{R(A)}$	—
AR: 全体	S_{AR}	$ar-1$	—	—

一元配置で，各処理群の反復数が等しいとき，一般に都合がよいが，$r_i \neq r_{i'}$ と異なるとき，$n=\sum r_i$ を全例数とし，$S_A=\sum T_i^2/r_i - CF$, 残差自由度を $[n-ar]$ と変更すればよい．

観測単位が，互いに類似した b グループに分かれるとき，これらを $B_1, \cdots, B_j, \cdots, B_b$ のブロックとして，n=ab に無作為に $A_1, \cdots, A_i, \cdots, A_a$ を割りつける代わりに，各ブロック B_j ごとに a 水準の処理を無作為化する．これを，**乱塊法**とい

うが，AとBの2因子でデータが分類されているので，**二元配置**と呼んでいることが多い．

$$y_{ij} = \eta+\alpha_i+\beta_j+\varepsilon_{ij}$$

のモデルを考えるわけで，データは，2方向で押えられる．

	$B_1 \cdots B_j \cdots B_b$	
A_1	$y_{11} \cdots\cdots y_{ab}$	T_{A1}
$\vdots$	$\vdots \quad \ddots \quad \vdots$	$\vdots$
A_i	y_{ij}	T_{Ai}
$\vdots$	$\vdots \quad \ddots \quad \vdots$	$\vdots$
A_a	$y_{a1} \cdots\cdots y_{ab}$	T_{Aa}
	$T_{B1} \cdots T_{Bj} \cdots T_{Bb}$	T_G

一元配置の場合と類似の方法で，

$$\begin{aligned}
CF &= \frac{T_G^2}{ab} \\
S_{AB} &= \sum_i \sum_j y_{ij}^2 - CF \quad [ab-1] \\
S_A &= \frac{1}{b}\sum_i T_{Ai}^2 - CF \quad [a-1] \qquad (85) \\
S_B &= \frac{1}{a}\sum_j T_{Bj}^2 - CF \quad [b-1] \\
S_{A\times B} &= S_{AB}-S_A-S_B \quad [(a-1)(b-1)]
\end{aligned}$$

ここに $S_{A\times B}$ は，誤差項であるが，形式上は $A\times B$ を交互作用といい，特別な A_i と B_j の出あいでは，$y_{ij}=\eta+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ij}$ と特別な効果 $(\alpha\beta)_{ij}$ が加わることの表現である．ここの二元配置のモデルでは，$S_{A\times B}$ を流用して誤差相当の残差を示すが，本質的な交互作用は存在しないものとする．

式 (85) の値から，分散分析表は，

要因	SS	Df	Ms	F_0
A	S_A	$a-1$	V_A	$V_A/V_{A\times B}$
B	S_B	$b-1$	V_B	$V_B/V_{A\times B}$
A×B	$S_{A\times B}$	$(a-1)(b-1)$	$V_{A\times B}$	—
AB	S_{AB}	$ab-1$	—	—

Ms は，平方和／自由度の分散である．F_0 はAとBについて求められ，普通は処理効果の比較に関心があり，$F_0(A)=V_A/V_{A\times B} \geqq F_\alpha[(a-1),(a-1)\times(b-1)]$ のとき有意とする．ブロックの一様性をみるには，第1自由度を $(b-1)$ とする．

対応のない2群比較のデータは $a=2$ の一元配置として扱うと，$t_0^2=F_0$ となり，まったく同義になる．対応のある2群比較のデータの場合，差として，1標本t検定を行なうと，これは，$a=2$ の二元配置と同義になる．

【例 16】

二元配置のデータがある．

	B_1	B_2	B_3	B_4	
A_1	5	10	5	6	26
A_2	6	9	7	7	29
A_3	9	13	11	12	45
	20	32	23	25	100

まず，本来は正しくないが，ブロックを無視した一元配置で考えると，

$$\begin{aligned}
CF &= \frac{100^2}{3\times 4} = 833{,}333 \\
S_{AR} &= 916-CF = 82.6667 \\
S_A &= \frac{1}{4}\times 3542-CF = 52.1667 \quad [2] \\
S_{R(A)} &= 82.6667-52.1667 = 30.50000 \quad [9] \\
V_A &= \frac{S_A}{2} = 26.0834 \\
V_{R(A)} &= \frac{S_{R(A)}}{6} = 6.1000 \quad SD = \sqrt{6.1000} \\
F_0 &= \frac{26.0834}{6.1000} = 4.276
\end{aligned}$$

F_0 は $F_{0.05}[2,9]=4.256$ より大きく，5% 水準で有意である．$SD=\sqrt{V_{R(A)}}=2.470$ である．

二元配置としては，$S_{AB}=S_{AR}=82.6667$ であり，$S_B=2574/3-CF=26.000$ を求め，$S_{A\times B}=S_A-S_B=4.5500$ となる．

誤差ないし残差に含まれていたBを取り出したので，一元配置と比べて，$F_0(A)=34.397$ と非常に大きくなり，$SD=\sqrt{V_{A\times B}}=0.871$ となり，いず

要因	SS	Df	Ms	F_0
A	52.1667	2	26.0834	34.397
B	26.0000	3	8.6667	11.429
A×B	4.5500	6	0.7583	—
AB	82.6667	11	—	—

れの処理平均値でも $SE=0.871/\sqrt{4}=0.436$.

一元配置でも二元配置でも，処理Aが有意となった段階では，どれとどれとが異なるかは，はっきりしない．これを明らかにするには，節9で述べた多重比較の手法が必要となる．

乱塊法で，各ブロックごとに順位づけをしたうえ，各処理ごとに和をとり，T_i としたとき，同位の組の修正を行ない，

$$K = 1-\frac{\sum(t^3-t)}{b(a^3-a)}$$
$$\chi_0{}^2 = \frac{6}{\kappa}\left\{\frac{1}{T_G}\sum T_i{}^2-\frac{1}{a}T_G\right\} \geqq \chi_\alpha{}^2[a-1] \tag{86}$$

のとき有意で，処理効果の一様性を否定する方法がある．これは **Freedman の検定**であるが，効率はあまりよくない．上の例では同位なしで，$\kappa=1$, $\chi_0{}^2=6.500>\chi^2_{0.05}[2]=5.991$ で有意となる．

20. 最小2乗法

線形モデルについては節18に紹介したが，$O_n\{x_i, y_i\}$ のデータに直線回帰モデル，

$$y_i = \beta_0+\beta_1x_i+\varepsilon_i \tag{87}$$

をあてはめることを例に，最小2乗法を整理してみる．ベクトルや行列の扱いが盛んに出てくるが，まず，2つのベクトル，

$$\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix}$$

が角度 θ を挾むとき，スカラー積，内積，

$$\mathbf{a}'\mathbf{b} = [\mathbf{a}'\mathbf{b}]' = \mathbf{b}'\mathbf{a} = \sum_{i=1}^{n} a_ib_i$$
$$= ||\mathbf{a}||\cdot||\mathbf{b}||\cos\theta$$

が定義され，要素の積和になる．$\mathbf{a}'$ は転置，$||\mathbf{a}||$ はノルムを示し，

$$\mathbf{a}' = \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix}' = [a_1 \cdots a_i \cdots a_n]$$

$$||\mathbf{a}|| = \sqrt{\mathbf{a}'\mathbf{a}} = \sqrt{\sum_{i=1}^{n} a_i{}^2}$$

である．

このスカラー積の約束で，式(87)は，

$$y_i = [1, x_i]\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}+\varepsilon_i$$

となる．$\mathbf{y}=[y_i \cdots y_n]'$, $\mathbf{j}=[1 \cdots 1]'$, $\mathbf{x}=[x_1 \cdots x_n]'$, $\boldsymbol{\varepsilon}=[\varepsilon_1 \cdots \varepsilon_n]'$ と $(n\times1)$ のタテベクトルと回帰係数ベクトル $\boldsymbol{\beta}=[\beta_0\beta_1]'$，さらに $(n\times2)$ の行列 $\mathbf{X}=[\mathbf{jx}]$ を用意し，全データを，

$$\mathbf{y} = [\mathbf{jx}]\boldsymbol{\beta}+\boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon} \tag{88}$$

と書く．$\mathbf{X}$ の転置はタテの列をヨコの行に並べかえて，$(n\times2)$ を $(2\times n)$ に変えることで，

$$\mathbf{X}' = [\mathbf{jx}]' = \begin{bmatrix} \mathbf{j}' \\ \mathbf{x}' \end{bmatrix}$$

誤差項について，(i) $E\{\varepsilon_i\}=0$ の不偏性，(ii) $V\{\varepsilon_i\}=\sigma^2$ の等分散性，(iii) $C\{\varepsilon_i, \varepsilon_{i'}\}=0$ の独立性のもとに，$Q=\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}=\sum\varepsilon_i{}^2$ を最小にする $\hat{\boldsymbol{\beta}}=[\hat{\beta}_0\hat{\beta}_1]'$ を求めるのが**最小2乗法**でこの解には統計的に好ましい性質 (BLUE) がある．さらに (iv) $\varepsilon_i\in N\{0, \sigma^2\}$ の正規性があれば，検定や区間推定が容易に行なえる．$E\{*\}$, $V\{*\}$, $C\{*\}$ は，平均値，分散，共分散を示す．

$$Q = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = [\mathbf{y}-\mathbf{X}\boldsymbol{\beta}]'[\mathbf{y}-\mathbf{X}\boldsymbol{\beta}]$$
$$= \mathbf{y}'\mathbf{y}-2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}+\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \tag{89}$$

を最小化するが，$[\mathbf{y}'\mathbf{X}\boldsymbol{\beta}]'=\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}$ である．

$$\frac{\partial Q}{\partial\boldsymbol{\beta}} = -2\mathbf{X}'\mathbf{y}+2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \tag{90}$$

と微分し，これを $\mathbf{0}$ とおくが，$[\mathbf{X}'\mathbf{X}]'=\mathbf{X}'\mathbf{X}$ とこ

の (2×2) の行列は対称で，さらにその行列式が $|\mathbf{X}'\mathbf{X}| \neq 0$ であれば，

$$[\mathbf{X}'\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{X}] = \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (91)$$

となる逆行列 $[\mathbf{X}'\mathbf{X}]^{-1}$ が存在するから，式(90)から $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}=\mathbf{X}'\mathbf{y}$ の正規方程式と呼ばれるものをつくり，これに $[\mathbf{X}'\mathbf{X}]^{-1}$ を前から乗じて，

$$\hat{\boldsymbol{\beta}} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{y} \qquad (92)$$

の最小2乗解 $\hat{\boldsymbol{\beta}}=[\hat{\beta}_0\hat{\beta}_1]'$ を得る．$\mathbf{I}$ はベクトルや行列においてはスカラーの1にあたる単位行列である．

$\hat{\boldsymbol{\beta}}$ を式(89)に代入し，最小の Q は，

$$\begin{aligned} Q\{\hat{\boldsymbol{\beta}}\} &= \mathbf{y}'\mathbf{y}-2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}+\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}[\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{y} \\ &= \mathbf{y}'\mathbf{y}-\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \end{aligned} \qquad (93)$$

となる．これは残差平方和と呼ばれる．

誤差自由度は，n 個のデータから2個のパラメターを推定したから $[n-2]$ で誤差相当の残差分散は，

$$V_R = \frac{Q\{\hat{\boldsymbol{\beta}}\}}{n-2}, \quad E\{V_R\} = \sigma^2 \qquad (94)$$

パラメターの分散と共分散は，

$$V\{\hat{\boldsymbol{\beta}}\} = \begin{bmatrix} V\{\hat{\beta}_0\} & C\{\hat{\beta}_0\hat{\beta}_1\} \\ C\{\hat{\beta}_0, \hat{\beta}_1\} & V\{\hat{\beta}_1\} \end{bmatrix} = \sigma^2[\mathbf{X}'\mathbf{X}]^{-1} \qquad (95)$$

となる．

以上の議論は，より一般的な線形モデル，

$$\begin{aligned} y_i = \beta_0+\beta_1 x_{i1}+ \cdots +\beta_j x_{ij}+ \cdots \\ +\beta_p x_{ip}+\varepsilon_i \end{aligned} \qquad (96)$$

にも拡張できるが，$\mathbf{X}=[\mathbf{j}\mathbf{x}_1 \cdots \mathbf{x}_p]$ で $[\mathbf{X}'\mathbf{X}]$ は $\{(p+1)\times(p+1)\}$ と大きくなり，V_R の自由度も $[n-p-1]$ と変わるが，$n>(p+1)$ としておく．

式(96)で，y_i が第 i 例での全般重症度であり，$x_{i1}, \cdots, x_{ip}$ が標的症状ごとの重症度であれば，$O_n\{(x_{i1}, \cdots, x_{ip}), y_i\}$ の多特性データについて求めた $\hat{\boldsymbol{\beta}}=[\hat{\beta}_0\hat{\beta}_1 \cdots \hat{\beta}_p]'$ は，ある意味で，各標的症状の重要性を示している．各症状間に関連があれば，一定の規準によって，重要なものだけを選び出す手法も用意され，こうしたものは**重回帰分析**と呼ばれる．

目的変数 y_i が大きな値では信頼性が低いという事情があれば，$w_i=1/y_i, 1/\hat{y}_i$, とか $w_i=1/y_i^2, 1/\hat{y}_i^2$ などの重みづけをして扱うこともある．この重みを要素とした対角行列，

$$\mathbf{W} = \begin{bmatrix} w_1 & & & & \mathbf{O} \\ & \ddots & & & \\ & & w_i & & \\ & & & \ddots & \\ \mathbf{O} & & & & w_n \end{bmatrix},$$

$$\mathbf{W}^{1/2} = \begin{bmatrix} \sqrt{w_1} & & & & \mathbf{O} \\ & \ddots & & & \\ & & \sqrt{w_i} & & \\ & & & \ddots & \\ \mathbf{O} & & & & \sqrt{w_n} \end{bmatrix}$$

を約束したうえ，式(96)に $\sqrt{w_i}$ を乗じ，

$$\sqrt{w_i}y_i = \sqrt{w_i}(\beta_0+\beta_1 x_{i1}+ \cdots +\beta_p x_{ip})+\sqrt{w_i}\varepsilon_i$$

$$\mathbf{W}^{1/2}\mathbf{y} = [\mathbf{W}^{1/2}\mathbf{X}]\boldsymbol{\beta}+\mathbf{W}^{1/2}\boldsymbol{\varepsilon}$$

とし，式(89)以下に準じて，

$$\left.\begin{aligned} [\mathbf{X}'\mathbf{W}^{1/2}\mathbf{W}^{1/2}\mathbf{X}]\hat{\boldsymbol{\beta}} &= \mathbf{X}'\mathbf{W}^{1/2}\mathbf{W}^{1/2}\mathbf{y} \\ \hat{\boldsymbol{\beta}} &= [\mathbf{X}'\mathbf{W}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{W}\mathbf{y} \\ Q\{\hat{\boldsymbol{\beta}}\} &= \mathbf{y}'\mathbf{W}\mathbf{y}-\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{W}\mathbf{y} \\ V\{\hat{\boldsymbol{\beta}}\} &= \sigma^2[\mathbf{X}'\mathbf{W}\mathbf{X}]^{-1} \end{aligned}\right. \qquad (97)$$

となる．

パラメター $\hat{\boldsymbol{\beta}}$ を用いて，$\mathbf{y}$ の予測値 $\hat{\mathbf{y}}=\mathbf{X}\hat{\boldsymbol{\beta}}$ を得るが，$\mathbf{X}$ の第 i 行 $\mathbf{x}_{i\cdot}'=[1x_{i1} \cdots x_{ip}]$ を用いると，

$$\hat{y}_i = \mathbf{x}_i'\hat{\boldsymbol{\beta}}$$

$$\begin{aligned} V\{\hat{y}_i\} &= \sigma^2\mathbf{x}_{i\cdot}'[\mathbf{X}'\mathbf{X}]^{-1}\mathbf{x}_{i\cdot} \quad \text{ないし，} \\ &= \sigma^2\mathbf{x}_{i\cdot}'[\mathbf{X}'\mathbf{W}\mathbf{X}]^{-1}\mathbf{x}_{i\cdot} \end{aligned} \qquad (98)$$

となる．重みなしでは $\mathbf{I}=\mathbf{W}$ とおいたことに相当する．

21. 直線回帰

$O_n\{x_i, y_i\}$ に直線をあてはめるにあたり，y_i は x_i と関係なく，平均値 η のまわりにばらつくというモデルを並べると，

$$\begin{cases}\omega:\ y_i = \eta + \varepsilon_i \\ \Omega:\ y_i = \beta_0 + \beta_1 x_i + \varepsilon_i\end{cases} \tag{99}$$

となる．ε_i はそれぞれで意味が異なる．Ω では，いわば平均値 $(\beta_0+\beta_1 x_i)$ のまわりに y_i がばらついているとみられる．

モデル ω は単純な線形モデルで，式 (88) の表現では，$\mathbf{X}$ が $\mathbf{j}$ に退化し，

$$\mathbf{y} = \mathbf{j}\eta + \boldsymbol{\varepsilon} = \eta\mathbf{j} + \boldsymbol{\varepsilon} \tag{100}$$

となる．式 (92) 以下を用い，$\mathbf{j}'\mathbf{j} = \sum 1\times 1 = n$, $\mathbf{j}'\mathbf{y} = \sum 1\times y_i$, $\mathbf{y}'\mathbf{y} = \sum y_i^2$ などに注意し，

$$\begin{aligned}\hat{\eta} &= [\mathbf{j}'\mathbf{j}]^{-1}\mathbf{j}'\mathbf{y} = n^{-1}\sum y_i = \bar{y} \\ Q\{\hat{\eta}\} &= \mathbf{y}'\mathbf{y} - \hat{\eta}\mathbf{j}'\mathbf{y} = \sum y_i^2 - n\bar{y}^2 = S_{yy} \\ V_R &= \frac{Q\{\hat{\eta}\}}{n-1} = \frac{S_{yy}}{n-1} = V \\ V\{\hat{\eta}\} &= V\{\bar{y}\} = \sigma^2[\mathbf{j}'\mathbf{j}]^{-1} = \frac{\sigma^2}{n} = SE\{\bar{y}\}\end{aligned} \tag{101}$$

などを得る．いずれも，これまでに何回も利用した性質である．

モデル式 (99) の Ω では少々ややこしいが，

$$\begin{aligned}[\mathbf{X}'\mathbf{X}]^{-1} &= \begin{bmatrix}\mathbf{j}'\mathbf{j} & \mathbf{j}'\mathbf{x} \\ \mathbf{x}'\mathbf{j} & \mathbf{x}'\mathbf{x}\end{bmatrix}^{-1} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2\end{bmatrix}^{-1} \\ &= \begin{bmatrix}\dfrac{1}{n} + \dfrac{\bar{x}}{S_{xx}} & \dfrac{-\bar{x}}{S_{xx}} \\ \dfrac{-\bar{x}}{S_{xx}} & \dfrac{1}{S_{xx}}\end{bmatrix}\end{aligned} \tag{102}$$

$$[\mathbf{X}'\mathbf{y}] = \begin{bmatrix}\mathbf{j}' \\ \mathbf{x}'\end{bmatrix}\mathbf{y} = \begin{bmatrix}\sum y_i \\ \sum x_i y_i\end{bmatrix}$$

としたうえ，式 (62) で，x と y との直線的なかかわり，共変動を，

$$S_{xy} = \sum x_i y_i - n\bar{x}\bar{y} = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}$$

と定義したことを用いると，$\hat{\boldsymbol{\beta}} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{y}$ の具体的な値は，

$$\boldsymbol{\beta} = \begin{bmatrix}\hat{\beta}_0 \\ \hat{\beta}_1\end{bmatrix} = \begin{bmatrix}\bar{y} - \dfrac{S_{xy}}{S_{xx}}\bar{x} \\ \dfrac{S_{xy}}{S_{xx}}\end{bmatrix} \tag{103}$$

となる．切片 $\hat{\beta}_0$ を勾配 $\hat{\beta}_1 = S_{xy}/S_{xx}$ で表現し，

$$\hat{y} = \bar{y} - \hat{\beta}_1\bar{x} + \hat{\beta}_1 x \tag{104}$$

の直線回帰式を得る．

式 (93) の最小の残差平方和は，

$$\begin{aligned}Q\{\hat{\boldsymbol{\beta}}\} &= \sum y_i^2 - (n\bar{y}\hat{\beta}_0 + \hat{\beta}_1\sum x_i y_i) \\ &= \sum y_i^2 - (n\bar{y}^2 - \hat{\beta}_1 n\bar{x}\bar{y} + \hat{\beta}_1\sum x_i y_i) \\ &= S_{yy} - \hat{\beta}_1 S_{xy}\end{aligned} \tag{105}$$

となる．

図5に示すように，$\mathbf{y}, \mathbf{j}, \mathbf{x}$ をとり，$\mathbf{y}$ から $\mathbf{j}$ 上に垂線をおろした点が $\hat{\eta}\mathbf{j}$ であり，$\mathbf{y}$ から $[\mathbf{j}\ \mathbf{x}]$ で定まる平面上に垂線をおろした点が $\mathbf{X}\hat{\boldsymbol{\beta}}$ である．したがって，$\mathbf{j}[\mathbf{j}'\mathbf{j}]^{-1}\mathbf{j}'$ および $\mathbf{X}[\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'$ は，$\mathbf{y}$ を $\mathbf{j}$ 上および $[\mathbf{jx}]$ 上に正射影する働きをもつ $(n\times n)$ の行列である．いずれも対称で，2乗しても同じである（べき等）という性質をもっている．$\|\mathbf{y}\|^2 = \|\hat{\eta}\mathbf{j}\|^2 + Q\{\hat{\eta}\}$ のピタゴラスの定理は，$\sum y_i^2 = n\bar{y}^2 + Q\{\hat{\eta}\}$ で，式 (101) の $Q\{\hat{\eta}\} = S_{yy}$ を意味する．

Ω では，$\{y_i\}$ のばらつきは，$\beta_1 x_i$ と表現される，x_i につられて動く部分と，誤差項 ε_i による部分から成る．ω では $\{y_i\}$ のばらつきは，すべて誤差項 ε_i とした．それぞれの残差平方和は，式 (101)，(105) から，

$$\begin{cases}\omega:\ Q\{\hat{\eta}\} = S_{yy} & [n-1] \\ \Omega:\ Q\{\hat{\boldsymbol{\beta}}\} = S_{yy} - \dfrac{S_{xy}^2}{S_{xx}} & [n-2]\end{cases} \tag{106}$$

であるが，$S_{xy}^2/S_{xx} \geqq 0$ で，ω に $\beta_1 x_i$ を加えた Ω では残差平方和が一般に小さくなり，この減少分 S_{xy}^2/S_{xx} が勾配 β_1 の手柄になり，S_{yy} のうち x の直線的影響で説明できる部分が減少分とみられる．図5の $\hat{\eta}\mathbf{j}$ から $\mathbf{X}\hat{\boldsymbol{\beta}}$ までの長さの2乗がこれにあたる．

$Q\{\hat{\boldsymbol{\beta}}\}$ は $S_{yy} - S_{xy}^2/S_{xx}$ として求めたうえ，以上をまとめると，

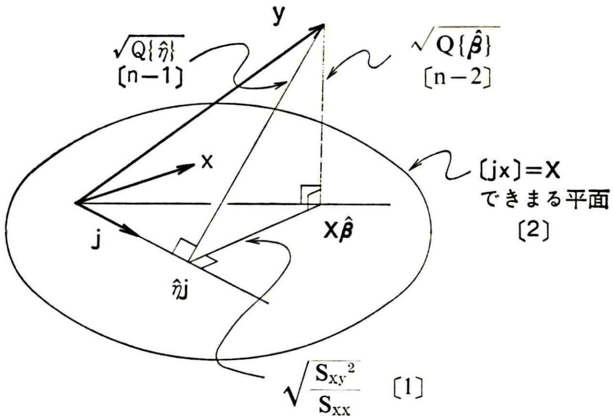


図 5 ω と Ω のモデル

要因	SS	Df	Ms
直線性 L	S_{xy}^2/S_{xx}	1	$V_L=S_{xy}^2/S_{xx}$
残　差 R	$Q\{\hat{\beta}\}$	n−2	$V_R=Q\{\hat{\beta}\}/(n-2)$
全　体 Y	S_{yy}	n−1	—

$H_0: \beta_1=0$, $H_1: \beta_1\neq 0$ の検定は,

$$F_0=\frac{V_L}{V_R}\geqq F_\alpha[1, n-2]=\{t_\alpha[n-2]\}^2 \quad (107)$$

のときに有意で，直線性を認める．もう少し厳密に“直線性”を扱うには，2次，3次，…の項が無視できることを検討する．

式 (95) から $V\{\hat{\beta}_1\}=\sigma^2/S_{xx}$ であるから，$\hat{\beta}_1$ のノイズを $SE\{\hat{\beta}_1\}=\sqrt{V_R/S_{xx}}$ として,

$$t_0=\frac{|\hat{\beta}_1-0|}{\sqrt{V_R}}\sqrt{S_{xx}}$$

$$\beta_{1L}{}^{U}=\hat{\beta}_1\pm t_\alpha[n-2]\sqrt{\frac{V_R}{S_{xx}}} \quad (108)$$

$t_0\geqq t_\alpha[n-2]$ で有意ならば，$[\beta_{1L}, \beta_{1U}]$ はゼロを含まない．t_0^2 は式 (107) の F_0 と完全に一致するはず．

形式的に，x と y と交換して,

$$\hat{x}_i=\beta_0'+\beta_1'y_i+\varepsilon_i'$$

と y に対する x の回帰式が考えられる，これについての F 検定，$\hat{\beta}_1'$ の t 検定も，上述とまったく同じ結果になるが，一般に $\hat{\beta}_1\neq\hat{\beta}_1'$ である．

$\{y_i\}$ のばらつきを変動 S_{yy} で表現すると，その一部は直線性で説明できる．その割合を**関与率**または，x で決定されるという意味で**決定係数**といい,

$$r^2=\frac{S_{xy}^2}{S_{xx}S_{yy}}=\hat{\beta}_1\hat{\beta}_1' \quad (109)$$

とする．形式上は，$\{x_i\}$ のばらつきのうち，y で決定される部分の割合とも考えられる．

回帰式では，独立変数 x で従属変数 y を説明しているが，x と y の直線的な相互関連性を示すのに，(単)**相関係数**ないし Pearson の積率相関係数,

$$r=\frac{S_{xy}^2}{\sqrt{S_{xx}S_{yy}}} \quad (110)$$

が用いられ，$H_0: \rho=0$, $H_1: \rho\neq 0$ を無相関の検定といい,

$$t_0=|r|\sqrt{\frac{n-2}{1-r^2}}\geqq t_\alpha[n-2] \quad (111)$$

のとき有意である．この t_0 は式 (108) の t_0 と同じくなる．式 (111) をみればすぐわかるように，$|r|$ が小さくても，n が大きくなると，容易に有意になるので，直線性あり，とか相関性あり，とかの結論は，現実的にはあまり意味がなく，むしろ r^2 について考えた方がよいと思われる．

対応のない2群のデータをヨコ軸の 0, 1 の座標にそれぞれプロットして，形式的に直線，をあてはめ，式 (108), (111) の t_0 を求めると，平均値

の比較の t 検定の式 (72) の t_0 と正確に一致する.

【例 17】

$O_{11}\{x_i, y_i\}$ のデータがある.

x	1	2	2	3	4	4	5	6	6	7	8	48
y	2	1	4	6	4	5	7	7	9	9	9	63

$\bar{x}=4.3636$, $\bar{y}=5.7273$, $S_{xx}=50.5454$, $S_{yy}=78.1818$, $S_{xy}=57.0909$, $\hat{\beta}_1=1.1295$ などを得る.

これらから,

要因	SS	Df	Ms	F_0
L	64.4840	1	64.4840**	42.368
R	13.6978	9	1.5220	—
Y	78.1818	10	—	—

$t_{0.05}[9]=2.262$ で, 式 (108) からは,

$$\beta_{1L}{}^{U} = 1.1295 \pm 2.262\sqrt{\frac{1.5220}{50.5454}} = 0.737; 1.522$$

また, $t_0=6.509$ で, $t_0{}^2=42.368=F_0$ である.

$$r^2 = \frac{57.0909^2}{50.5454 \times 78.1818} = 0.8248$$

で, y の変動の 82.5% は直線性で説明される. 無相関の検定では,

$$t_0 = \sqrt{0.8248} \cdot \sqrt{\frac{9}{1-0.8248}} = 6.509$$

回帰式は, $\bar{y}-\hat{\beta}_1\bar{x}=0.799$ を計算し,

$$y = 0.799+1.130x$$

ここで, ある値 x^* を指定したとき, そこでの予測値 $\hat{y}^*$ の $100(1-\alpha)\%$ 信頼限界は, 式 (98) の考え方に沿って,

$$\eta_L{}^{*U} = \hat{y}+\hat{\beta}_1(x^*-\bar{x}) \pm t_\alpha[n-2]\sqrt{V_R\left\{\frac{1}{n}+\frac{(x-\bar{x})^2}{S_{xx}}\right\}} \quad (112)$$

となる. $\hat{y}^*$ のまわりの個体に関しては, 根号内が $[V_R\{1+1/n+(x^*-\bar{x})^2/S_{xx}\}]$ となる. いずれにしても $x^*=\bar{x}$ のとき区間幅が最小になる.

散点図と, x^* の断面について 95% の信頼係数を保証したときの 2 種類の信頼域を図 6 に示す.

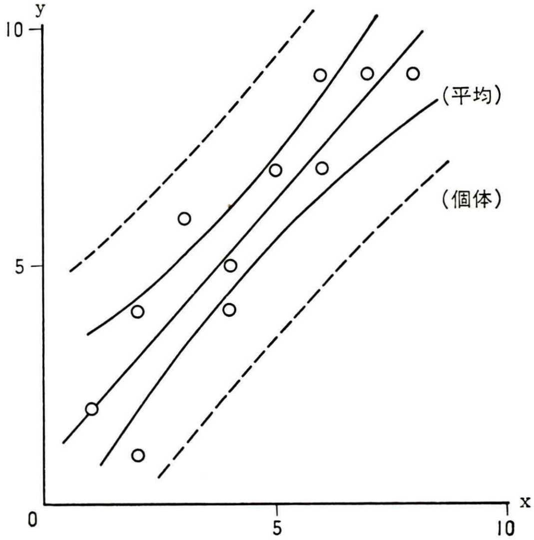


図 6　直 線 回 帰

式 (110) に, なまデータではなしに, x と y それぞれでの順位を用いた場合, Spearman の順位相関係数 (rho) となる. 順位づけをして, 対応するペアーの差 d_i をとると $\sum d_i=0$ となる. 各群内の同位のそれぞれにつき,

$$\tau = \sum (t^3-t)$$

を求めてプールしたものを τ_x, τ_y とする.

$$D = 6\sum d_i{}^2, \qquad A = n^3-n$$

$$r_S = \frac{A-0.5(\tau_x+\tau_y)-D}{\sqrt{(A-\tau_x)(A-\tau_y)}} = 1-\frac{D}{A} \quad (\text{同位なし}) \quad (113)$$

無相関の検定には式 (111) を用いる.

上のデータでは $\sum d_i{}^2=17$, $\tau_x=3\times6=18$, $\tau_y=2\times6+24=36$, $A=1320$ で, Spearman の相関係数は, $r_S=0.9211$ となる.

22. 非線形モデル

ある関数が,

$$f(\mu x+\zeta z) = \mu f(x)+\zeta f(z) \quad (114)$$

と書けるとき線形というが，これまでに扱ったモデルは線形である．

$$y = \exp\{\alpha x + \beta x^2 + \varepsilon\}$$

は，Y=ln y とおけば，

$$Y = \alpha x + \beta x^2 + \varepsilon = \alpha X_1 + \beta X_2 + \varepsilon$$

となり，本質的には線形である．

$$y = \exp\{\alpha x + \beta x^2\} + \varepsilon$$

$$y = Ae^{-\alpha x} + Be^{-\beta t} + \varepsilon$$

$$y = \frac{\alpha}{1+(\beta/x)} + \varepsilon$$

などは，いずれも非線形である．

線形モデル式(96)は，

$$y_i = f(\mathbf{x}_{i\cdot};\boldsymbol{\beta}) + \varepsilon_i = f_i(\boldsymbol{\beta}) + \varepsilon_i$$

$$\mathbf{y} = \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\varepsilon}$$

と書くと，$\mathbf{f}(\boldsymbol{\beta})$ は $f_i(\boldsymbol{\beta})$ を要素としたベクトルで，最小2乗解は，

$$Q\{\boldsymbol{\beta}\} = [\mathbf{y}-\mathbf{f}(\boldsymbol{\beta})]'[\mathbf{y}-\mathbf{f}(\boldsymbol{\beta})]$$

を最小にするパラメター，

$$\hat{\boldsymbol{\beta}} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{y}$$

として与えられる．

ところが，$f_i(\boldsymbol{\beta})$ が**非線形**のときには，このように調子よくはいかない．いま2つのパラメターをもつ非線形式 $f_i(\beta_1, \beta_2)$ について，β_1, β_2 を指定すれば，$f_i(\boldsymbol{\beta})$ が定まり，

$$Q\{\boldsymbol{\beta}\} = \sum_{i=1}^{n}\{y_i - f_i(\boldsymbol{\beta})\}^2 \tag{115}$$

が計算できる．さまざまな $\{\beta_1, \beta_2\}$ の組み合わせにつき式(115)を計算し，最小値を含む領域について，図7のような“地形”ができる．

$\{\beta_1, \beta_2\}$ を，あらゆる組合わせで計算するのも大変であるから，まず，適当な $\boldsymbol{\beta}_0$ を定めて，$Q\{\boldsymbol{\beta}_0\}$ を求め，ここで β_2 を留めて，β_1 を変え $Q\{\boldsymbol{\beta}\} < Q\{\boldsymbol{\beta}_0\}$ かつ最小となる β_1 を探す．ついで，β_1 をそこに留めて，β_2 を変えて $Q\{\boldsymbol{\beta}\}$ がさらに小さく，もっとも小さくなる β_2 を探す．こうして軸ごとに探索して，すり鉢状の盆地の底に達したなら，そこを $\hat{\boldsymbol{\beta}}$ とする．

出発点 $\boldsymbol{\beta}_0$ が図7のように“運のいい”場所であれば右側にみるように，β_1 を動かして真西に探索すれば，1回でほとんど盆地の中心に達することができるが，$\boldsymbol{\beta}_0$ の選び方がまずいと，かなり手間はかかる．

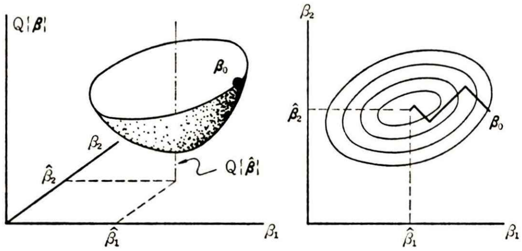


図 7 $Q\{\beta\}$ の探索

もう少し一般的に，$\boldsymbol{\beta}_0$ に立ったとき東西と南北の方向にとらわれず，もっとも急な勾配に沿って下ることが考えられる．$Q\{\boldsymbol{\beta}\}$ を，$\partial Q\{\boldsymbol{\beta}\}/\partial\beta_1$ と微分し，β_{1k} を代入することを $\partial Q_k/\partial\beta_1$ と書けば，

$$\nabla Q_k = \begin{bmatrix} \dfrac{\partial Q}{\partial\beta_1} \\ \dfrac{\partial Q}{\partial\beta_2} \end{bmatrix}_{\boldsymbol{\beta}=\boldsymbol{\beta}_k} = \begin{bmatrix} \dfrac{\partial Q_k}{\partial\beta_1} \\ \dfrac{\partial Q_k}{\partial\beta_2} \end{bmatrix} = \mathbf{g}_k \tag{116}$$

となる．∇Q_k はグラジエントと呼ばれるもので，$\boldsymbol{\beta}_k$ での傾斜であるが，$\mathbf{d}_k = -\mathbf{g}_k$ の方向が最急降下の方向になる．実際は差分を用い，

$$\frac{\partial Q_k}{\partial\beta_1} \doteqdot \frac{Q(\beta_{1k}+\Delta, \beta_{2k}) - Q(\beta_{1k}, \beta_{2k})}{\Delta} \tag{117}$$

とするが，Δ は β_{1k} の 10^{-3}～10^{-6} といった値である．つまり，$\{\beta_{1k}+\Delta, \beta_{2k}\}$ と $\{\beta_{1k}, \beta_{2k}\}$ について $Q\{\boldsymbol{\beta}\}$ を計算し，その差を Δ で割って，これを $\partial Q_k/\partial\beta_1$ として用いる．

最急降下で $\boldsymbol{\beta}_k$ から $\mathbf{d}_k$ 方向に探索した最小値が $Q\{\boldsymbol{\beta}_k + \lambda_k\mathbf{d}_k\}$ であれば，$\boldsymbol{\beta}_{k+1} = \boldsymbol{\beta}_k + \lambda_k\mathbf{d}_k$ をつぎの出発点とし，そこでの最急降下を行なう．$\mathbf{g}_k \doteqdot \mathbf{0}$ とかきざみ $\lambda_k \doteqdot 0$ とかですでに盆地の底に

達したと考える．最急降下法は，局所的には良い方法であるが，$Q\{\boldsymbol{\beta}\}$ の地形が細長い谷状になると，なかなか底まで下れないこともあるので，さらに工夫をこらした方法もある．

他の方法一般についてえるが，出発点 $\boldsymbol{\beta}_0$ の選び方が悪いと，隣りの浅い窪地で終わったり，現実に $\boldsymbol{\beta}$ がとり得ない $\hat{\boldsymbol{\beta}}$ に達したりするほか，かなり大量の計算を必要とするなど非線形モデルでの最小 2 乗法には，厄介な問題が少なくない．いずれにしても，手計算は無理であろう．

23. 線形化の方法

非線形モデルで $f_i(\boldsymbol{\beta})$ を真の解の付近の $\boldsymbol{\beta}_k$ について展開して高次項を省略し，

$$f_i(\boldsymbol{\beta}) \fallingdotseq f_i(\boldsymbol{\beta}_k)+\left[\frac{\partial f_i}{\partial \beta_1}\cdots\frac{\partial f_i}{\partial \beta_p}\right]_{\beta=\beta_k}[\boldsymbol{\beta}-\boldsymbol{\beta}_k]$$

とおけることが多い．パラメターは p 個ある．

$[\nabla f_i(\boldsymbol{\beta}_k)]'$ を行の成分とした $(n\times p)$ の行列，

$$\mathbf{J}(\boldsymbol{\beta}_k)=\mathbf{J}_k=\begin{bmatrix}\frac{\partial f_1}{\partial \beta_1}\cdots\frac{\partial f_1}{\partial \beta_p}\\ \vdots \qquad \vdots\\ \frac{\partial f_n}{\partial \beta_1}\cdots\frac{\partial f_n}{\partial \beta_p}\end{bmatrix}_{\beta=\beta_k} \tag{118}$$

をヤコビの行列というが，これと $f_i(\boldsymbol{\beta}_k)$ を要素としたベクトル $\mathbf{f}_k$ を用いると，

$$\mathbf{y} \fallingdotseq \mathbf{f}_k+\mathbf{J}_k[\boldsymbol{\beta}-\boldsymbol{\beta}_k]+\boldsymbol{\varepsilon} \tag{119}$$

と書ける．これから，

$$\mathbf{y}-\mathbf{f}_k+\mathbf{J}_k\boldsymbol{\beta}_k \fallingdotseq \mathbf{J}_k\boldsymbol{\beta}+\boldsymbol{\varepsilon}$$

と，式 (88) の形式におけば，最小 2 乗解は，

$$\hat{\boldsymbol{\beta}} \fallingdotseq [\mathbf{J}_k'\mathbf{J}_k]^{-1}\mathbf{J}_k'[\mathbf{y}-\mathbf{f}_k]+\boldsymbol{\beta}_k$$

$$V\{\hat{\boldsymbol{\beta}}\} \fallingdotseq \sigma^2[\mathbf{J}_k'\mathbf{J}_k]^{-1}$$

となる．式 (119) で，$\mathbf{d}_k=[\boldsymbol{\beta}-\boldsymbol{\beta}_k]$ とおけば，

$$\hat{\mathbf{d}}_k \fallingdotseq [\mathbf{J}_k'\mathbf{J}_k]^{-1}\mathbf{J}_k'[\mathbf{y}-\mathbf{f}_k] \tag{120}$$

となる．$\hat{\mathbf{d}}_k$ を $\boldsymbol{\beta}_k$ での降下ベクトルとみて，最小の $Q\{\hat{\boldsymbol{\beta}}\}$ をあたえる解を，

$$\hat{\boldsymbol{\beta}} \fallingdotseq \boldsymbol{\beta}_{k+1}=\boldsymbol{\beta}_k+\hat{\mathbf{d}}_k \tag{121}$$

とすることで，反復法が組み立てられるが，これが **Gauss 法** (Gauss-Newton 法) である．重みを用いて，式 (97) の形式で扱ってもよい．

理論的に $\hat{\mathbf{d}}_k$ は下降方向ではあるが，さまざまな条件で式 (121) の反復が収束しなかったり，計算できなかったりすることがある．

$$\boldsymbol{\beta}_{k+1}=\boldsymbol{\beta}_k+\lambda_k\hat{\mathbf{d}}_k$$

として，最小値を与える λ_k を求める方法や，式 (120) に加工して，

$$\hat{\mathbf{d}}_k=[\mathbf{J}_k'\mathbf{J}_k+\mu\mathbf{I}]^{-1}\mathbf{J}_k'[\mathbf{y}-\mathbf{f}_k] \tag{122}$$

とし，$[\mathbf{J}_k'\mathbf{J}_k]$ の対角要素に $\mu>0$ を加えて計算する方法がある (Levenberg-Marquardt の原理)，μ を非常に大きくすると収束は遅くなるが，$\mathbf{d}_k$ は $\boldsymbol{\beta}_k$ での最急降下の方向に近づくという性質がある．どのくらいの μ が良いかは，むしろ経験的に定められるが，$\mu=0$ の本来の Gauss 法で不都合なとき，大きな μ をのせて反復し，しだいに μ を減らし，十分に収束したら $\mu=0$ として成功することがある．

考え方を変えて，$Q\{\boldsymbol{\beta}\}$ の展開を行ない，$(n\times p)$ のヘッセの行列

$$\mathbf{H}(\boldsymbol{\beta}_k)=\mathbf{H}_k=\begin{bmatrix}\frac{\partial^2 f_1}{\partial \beta_1\partial \beta_1}\cdots\frac{\partial^2 f_1}{\partial \beta_1\partial \beta_p}\\ \vdots \qquad \vdots\\ \frac{\partial^2 f}{\partial \beta_p\partial \beta_1}\cdots\frac{\partial^2 f}{\partial \beta_p\partial \beta_p}\end{bmatrix}_{\beta=\beta_k} \tag{123}$$

を用いて，

$$Q\{\boldsymbol{\beta}\} \fallingdotseq Q\{\boldsymbol{\beta}_k\}+\mathbf{d}_k'\nabla Q_k+\frac{1}{2}\mathbf{d}_k'\mathbf{H}_k\mathbf{d}_k$$

$$\frac{\partial Q}{\partial \boldsymbol{\beta}} \fallingdotseq \nabla Q_k+\mathbf{H}_k\mathbf{d}_k$$

$$\hat{\mathbf{d}}_k=-\mathbf{H}_k^{-1}\nabla Q_k=\mathbf{H}_k^{-1}(-\mathbf{g}_k)$$

によって下降ベクトルが求められる．最急勾配は式 (116) で $\mathbf{g}_k$ としたが，最急降下法では $\mathbf{d}_k=-\mathbf{g}_k$ を選んだ．この式 (124) では $\mathbf{H}_k^{-1}$ が，各軸での計量単位を修飾する姿でかかわり，一種の最急降下方の加速になっている．これが **Newton 法**

(Newton-Raphson 法) と呼ばれる．収束が悪いときには，Gauss 法で用いたような変法が利用される．

【例 18】

ラットに Hippuran ^{125}I を静注し，連続的に頭部でモニターし，5分ごとの計測を行なった．

i	1	2	3	4	5	6
分	5	10	15	20	25	30
cpm	1,000	626	440	315	235	173
i	7	8	9	10	11	12
分	35	40	45	50	55	60
cpm	148	124	103	93	83	73

まず，y(cmp) を ln y としてプロットすると図8 (1) の○を得る．

$$y_i = \sum Me^{-\mu t i} + \varepsilon_i$$

の区画モデルを考え，第6 (30分)～第12 (60分) の ln y に直線式

$$\ln y_B = \ln B - \beta t \fallingdotseq \ln 400 - 0.03\ t$$

をあてはめ，いわゆる β 相を，

$$y_B = Be^{-\beta t} \fallingdotseq 400e^{-0.03t}$$

とする．ついで，第1 (5分)～第5 (25分) の，もとの y から y_B の計算値を引き，

$$y_A = y - y_B = Ae^{-\alpha t}$$

とし，あらためて ln y_A をプロットし，●を得るが，ほぼ直線的で，

$$\ln y_A = \ln A - \alpha t \fallingdotseq \ln 1322 - 0.14\ t$$

となる．もしも ln y_A が直線でなければ，第5 (25分) のデータ側から直線をあてはめ，上記の要領で“皮むき”を続け，区画数を増す．このデータでは2区画モデル，

$$y = y_A + y_B + \varepsilon = Ae^{-\alpha t} + Be^{-\beta t} + \varepsilon \qquad (125)$$

でよさそうである．

初期値を $\boldsymbol{\beta}_0 = [1322\ \ 0.14\ \ 400\ \ 0.03]'$ とおいて Gauss 法を用いると，$Q\{\boldsymbol{\beta}_0\} = 996.2734$ で，3回の反復でかなり安定し，$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}_4$ で終了し，$Q\{\hat{\boldsymbol{\beta}}\} = 432.5302$ となる．

あてはめのずれを誤差相当と考え，

$$V_R = \frac{Q\{\hat{\boldsymbol{\beta}}\}}{12-4} = 54.0666$$

$[\mathbf{J}_4{}'\mathbf{J}_4]^{-1}$ の対角要素をひろい，これに V_R を乗じ

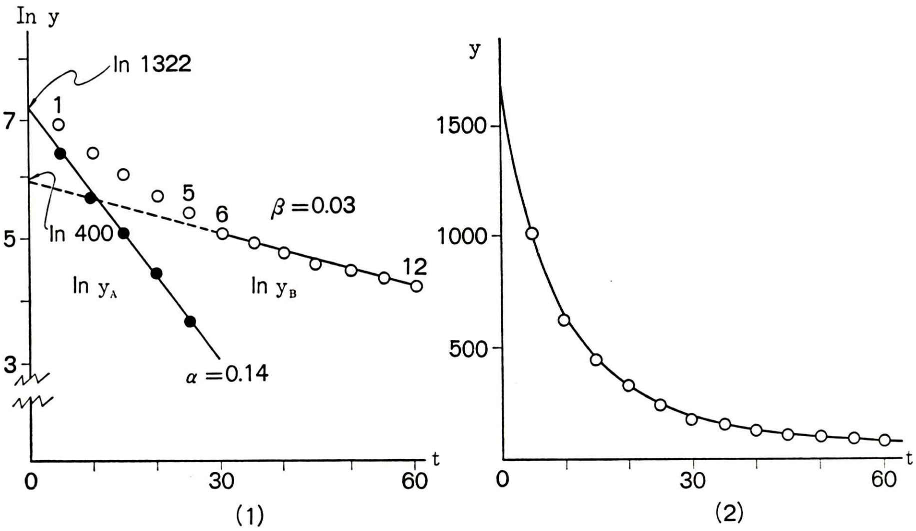


図 8

て各パラメターの分散とし，その平方根をとれば $SE\{\hat{\beta}\}$, $\beta=\hat{A}, \hat{\alpha}, \hat{B}, \hat{\beta}$, となる．

つぎに $w_i=1/y_i$, および $1/y_i^2$ として同じ計算を行なうと，かなり様子がちがってくる．

	A	α	B	β	w_i
$\hat{\beta}$	1315.2	0.1198	310.4	0.02503	1
$SE\{\hat{\beta}\}$	48.22	0.008759	67.9	0.004408	
$\hat{\beta}$	1342.8	0.1099	239.5	0.02025	$\frac{1}{y_i}$
$SE\{\hat{\beta}\}$	34.58	0.006508	44.84	0.003380	
$\hat{\beta}$	1341.6	0.1059	215.0	0.01842	$\frac{1}{y_i^2}$
$SE\{\hat{\beta}\}$	42.9	0.006195	34.91	0.002735	

y の分布にもよるが，$SE\{\beta\}$ からみると，$1/y_i$, $1/y_i^2$ の重みがよさそうである．どのような重みがよいかは，こうした吟味と，経験によって決められるであろう．y_i の誤差分散がわかっており，σ_i^2 であれば，$w_i=1/\sigma_i^2$ ないし，k/σ_i^2 と，分散の逆数に比例した重みを用いると，誤差は等分散になる．また $1/\hat{y}_i$, $1/\hat{y}_i^2$ なども用いられる．

24. パラメターの比較

線形モデルあるいは線形化したモデルでの Gauss 法によるパラメター，たとえば，$\hat{\beta}_1, \hat{\beta}_2$, を得たとき，それぞれの分散を V_1, V_2, 誤差自由度を ν_1, ν_2 とする．

H_0: $\beta_1=\beta_2$, H_1: $\beta_1\neq\beta_2$ の検定にあたり，$\hat{\beta}_1, \hat{\beta}_2$ が同一データから求められたときには，$\nu=\nu_1=\nu_2$ で，$|\hat{\beta}_1-\hat{\beta}_2|$ のノイズは，

$$SE\{\hat{\beta}_1-\hat{\beta}_2\} = \sqrt{V_1+V_2-2C}$$

であるが，C は $C\{\beta_1, \beta_2\}$ で，$[\mathbf{X}'\mathbf{X}]^{-1}$ ないしは $[\mathbf{J}_k'\mathbf{J}_k]^{-1}$ などの，第1行第2列，または第2行第1列の要素に V_R を乗じたものになる．誤差が少なくとも緩い正規性をもつなら，

$$t_0 = \frac{|\hat{\beta}_1-\hat{\beta}_2|}{\sqrt{V_1+V_2-2C}} \geqq t_\alpha[\nu] \qquad (126)$$

のとき有意とし，必要なら反転して区間推定を行なう．

$\hat{\beta}_1, \hat{\beta}_2$ が別個のデータで求められたときは不等分散とした方が安全で，

$$t_0 = \frac{|\hat{\beta}_1-\hat{\beta}_2|}{\sqrt{V_1+V_2}} \qquad (127)$$

を求め，節17の式 (76) 以下の手続きをふめばよいだろう．$\nu_1\leqq\nu_2$ として，

(1) $t_0<t_\alpha[\nu_1+\nu_2]$ では有意ではない．

(2) $t_0>t_\alpha[\nu_1]$ では有意である．

(3) 以上で未決定なら t_0 を

$$t_\alpha^* = \frac{V_1t_\alpha[\nu_1]+V_2t_\alpha[\nu_2]}{V_1+V_2}$$

と比べる．

【例 19】

例18の $w_i=1$ での計算で得た $\hat{\alpha}, \hat{\beta}$ につき

	値	分散	共分散
$\hat{\alpha}$	0.1198	0.0^47672	0.0^43503
$\hat{\beta}$	0.02503	0.0^41943	

ただし，0.0^41 は 0.00001 を示す．式 (126) から

$$t_0 = \frac{|0.1198-0.02503|}{\sqrt{0.0^42609}} = 18.554$$

で，$t_{0.05}[8]=2.306$ より大きく，明らかに $\alpha\neq\beta$ である．

間接的な連続モニターと同時に，8時点で採血し，血漿中の Hippurate ^{125}I を測定したデータがある．

i	1	2	3	4	5	6	7	8
分	5	10	20	30	40	45	50	60
cpm	1,867	1,380	907	577	378	334	289	234

これについての2区画モデルは $w_i=1$ ではどうやら結果は出るが，$w_i=1/y_i$ では，式 (122) の変法で追いつめても，$\mu=0$ で発散してしまう．$w_i=1$ では，

$$y = 1365.9e^{-0.08938t}+1130.6e^{-0.02827t}$$

となる．$\hat{A}, \hat{\alpha}, \hat{B}, \hat{\beta}$ の SE は，$\nu=8-4=4$ で，

804.26, 0.04686, 902.07, 0.01309

と，間接法に比べて，精度が格段に悪い．また $V\{\hat{\alpha}\}=0.0^2196$, $V\{\hat{\beta}\}=0.0^31715$ である．

Hippuran ^{125}I は血球にも入るし，cpm/head, cpm/20 μl plasma と単位も異なるので，A や B は直接法と間接法で異なる．しかし，両者で α, β が等しいならば，都合がよい．

直接法のデータには少々精度に問題があるが，まず α について，前述の間接法でのデータを用いて，

$$t_0=\frac{|0.1198-0.08938|}{\sqrt{0.0^47672+0.0^22196}}=0.638$$

また β について，

$$t_0=\frac{|0.02503-0.02827|}{\sqrt{0.0^41943+0.0^31715}}=0.234$$

いずれも $t_{0.05}[\nu_1=4]=2.776$ より小さく，さしあたりは，α, β のパラメターは間接法，直接法で大きくは変わらないと考えておく．以上の2例のデータは，Am. J. Physiol. **212**, 629,1967 の論文の図から読みとったものである．