

薬物血中濃度モニタリングのための

Population

Pharmacokinetics 入門

京都大学 教授 堀了平 監修  
医学部附属病院薬学部門

薬業時報社

## 本書に用いられている主な記号

ただし、個々によって異なる場合は、記載された箇所における説明が優先される

$C_p$	血中（血漿中）薬物濃度	$\bar{\theta}$	パラメータ $\theta$ の算術平均値
$k_e$	消失速度定数	$\bar{\theta}$	パラメータ $\theta$ の母集団（平均）値（真値）
$k_a$	吸収速度定数	$\hat{\theta}$	パラメータ $\theta$ の母集団（平均）値の推定値
$V_d$	分布容積	$\theta_i$	パラメータ $\theta$ の個人 (i) の値（真値）
$CL$	(総)クリアランス	$\hat{\theta}_i$	パラメータ $\theta$ の個人 (i) の値の推定値
$D$	投与量	$t_{ij}$	個人 (i) の j 番目の測定時間
$F$	吸収分率，バイオアベイラビリティの分率	$C_{p_{ij}}$	個人 (i) の j 番目の血中濃度（実測値）
$t_{1/2}$	半減期	$\hat{C}_{p_{ij}}$	個人 (i) の j 番目の血中濃度の推定値
$CL_{cr}$	クレアチニンクリアランス	$\eta_\theta$	パラメータ $\theta$ の固体間変動
$S_{cr}$	血清クレアチニン値	$\omega_\theta^2$	パラメータ $\theta$ の固体間変動の分散
$WT$	体重	$\varepsilon_\theta$	パラメータ $\theta$ の固体内変動
$GFR$	糸球体濾過速度	$\sigma_\theta^2$	パラメータ $\theta$ の固体内変動の分散
$V_m$	最大代謝速度	$SS$	残差平方和
$K_m$	ミカエリス定数	$Ob$	目的関数
$R_{inf}$	点滴速度	$W$	データ（パラメータ）の重み
$T$	点滴時間		
$\tau$	投与間隔		
$C_{p_{ss}}$	定常状態時血中濃度		
$C_{p_{ss}}^{ave}$	定常状態時平均血中濃度		
$D_L$	負荷投与量		
$D_M$	維持投与量		

第 1 章

# 理論的背景

## 第 1 節

# Population pharmacokinetics とベイジアン法

### 1. 臨床における pharmacokinetics の適用

Pharmacokinetics の概念、手法を臨床に適用する目的は、(1) 投与された医薬品の血中濃度の推移や病巣中濃度の推移の法則性を知り、その薬物の有効性の発現・維持や副作用発現との関連性を明らかにすること、および (2) 明らかとなった薬物の体内動態の法則性をもとに、必要とする（あるいは望ましい）血中濃度あるいは体内薬物量を得るための投与設計を個々の患者に対し行うことの 2 点に置かれる。

(1) はそれぞれの薬物の体内動態の法則性を明らかにする方向であり、たとえば、血中濃度推移をコンパートメントモデルにあてはめ、平均的なパラメータ値を求めるとともに、それらパラメータ値の変動要因を明らかにすることなどがあげられる。ここでは、一人一人の患者における薬物動態ではなく、患者の集団としての平均的な関係、population pharmacokinetic parameter (母集団パラメータ) を得ることを目的としている。(2) は個別化を行う方向であり、一般化されたその薬物に関する情報を個々の患者の状態に即して還元し、有効で合理的な薬物治療の遂行に参画することを目的としている。

### 2. Population pharmacokinetics

薬物の体内動態の平均的なパラメータは、ある共通の要因を持つ多数の被験者（たとえば、健康被験者群、ある嗜好・食習慣を持つ被験者群、患者群、ある特別の疾患あるいはある範囲の臨床検査値を示す患者群など）を対象とする薬物の投与実験によって得られてきた。まず、一人一人の被験者から、多数回の血中濃度値を得ることにより、個々の被験者のパラメータ値を算出し（このとき、パラメータ値と個人内誤差が得られる）、それぞれのパラメータの平均値とその分散（個人間変動）を求める（2 段階を踏むので STS 法という）。この方法は、実験プロトコルに従い、厳密にコントロールされた条件下で実験が行われる場合が多く、得られたパラメータ値の信頼性は高い。しかし、1 被験者に対し多くの採血点を必要とし、しかも、同一要因を持つ被験者を多く必要とするなど、被験者群を幅広くとるには困難性があり、母集団を広く想定することが困難であるという点に問題がある。

母集団パラメータの算出に、先の STS 法とは異なった方法が Sheiner & Beal によって

開発された。これは、ある患者の血中濃度の推移を具体的な体内動態パラメータによって説明される部分（固定効果という）とその他の未知の部分（個体内変動、個体間変動として表現され、変量効果という）とにわけ、それぞれの値を多数の患者の比較的ランダムな血中濃度値を用いて同時に解析しようとするものである。そのためのコンピュータプログラムとしてNONMEM (Nonlinear Mixed Effect Model) が開発された。一般に population pharmacokinetics と表現した場合、この同時的解法をさす。

この同時的解法によれば、被験者をあらかじめ同一の要因を持つ群に層別化する必要がないこと（ただし、できるだけ患者の臨床検査値などの情報は集め、解析の段階で体内動態パラメータとの相関性を検討する）、各被験者からはほぼランダムな時点で、最低1ポイントの採血点 が得られればよいことなど、先のSTS法の欠点を補うことのできる構成となっている。そのため、臨床における血中濃度モニタリング業務などによって得られたルーチンの測定値を解析に利用できること、また、ランダムな条件下のランダムな時点の血中濃度値を収集すればよいことから被験者の対象を広く取ることができることなど、薬物の母集団パラメータを得るための情報源を、従来の研究室レベルの厳密な実験の枠から臨床の日常活動のレベルにまで拡大できることがこの方法の特徴であり、注目される所以でもある。

### 3. 投与設計

患者個々の状態に応じた医薬品の投与量、投与間隔などの設計は、薬物治療の精密化と合理化のためのひとつのゴールである。そのためには、いかに迅速に精度よく個々の患者の体内動態パラメータを得るかが前提となる。

しかし、個々の患者のその時点でのパラメータ値を事前に得ることは現実にはできない。そのため、その時々患者のパラメータ値を知るために、すでに報告されているそのパラメータの平均値（母集団パラメータ）をそのまま用いるか、または患者の状態を結びつける付加的情報を組み合わせて利用する方法が従来からとられている。付加的情報としては、体重、年齢、クレアチニンリアランス、血清クレアチニン値、あるいは薬物の試験的投与で得られる1ないし2点の血中濃度値などである。母集団パラメータとこれら付加的情報とを組み合わせることによる個々の患者のパラメータの推定は、経験式、ノモグラムなどによって、個々の薬物ごとに多くの研究者によって提示されてきている。しかし、これらの関係式やノモグラムは、平均的關係として得られたものであり、とくに変量効果の因子を考慮できない欠点があり、予測の精度に問題がある場合も多い。

これらの欠点に対し、近年、統計学のベイズ理論にもとづく予測法が提示された。これは、事前情報としての母集団パラメータと新たに得られた患者に関する情報（1点以上の血中濃度、あるいは体内動態の変動を説明する臨床検査値など）とのバランスの上に（パラメータの分散、測定値の誤差なども考慮し）、最もあてにできそうな患者のパラメータを予測するというものである。この方法を適用するためには、事前情報として、平均的体内動態パラメータ（母集団パラメータ）およびその分散、測定値の誤差などが必要であり、これは、先に述べたNONMEMに代表される同時的解法により得られる情報と一致する。すなわち、パラメータ

の解析は同時的解法で行い，得られた情報を事前情報としてベイジアン法により各患者の個別のパラメータ推定を行うことが，現在，最も被験者に対する負担が少なく，予測性において優れた方法とされるにいたった。

本章の第2，3節では population pharmacokinetics とベイジアン法の理解に必要な概念の解説を試みた。そのうち，第2節ではできるだけ数式を用いず，平易に概念をまとめた。この節だけでも第2章以降に述べられている内容は充分フォローできるように心掛けた。第3節はさらに詳しく論理的内容に踏み込むための手引として利用されたい。

〔緒方宏泰，魚井 徹〕

## 第 2 節

# 概念的枠組み

### 1. モデル式のあてはめ

一人の被験者に薬物を静脈投与し、経時的に血中濃度の測定を行ったときの測定時刻を  $t_i$ 、濃度を  $C_{pi}$  とする。薬物の血中からの消失速度が、その時点での濃度に比例すると考えると、血中濃度曲線は、 $C_{pi} = ae^{-\alpha t_i}$  という静注 1-コンパートメントのモデル式によって表される。

$a$ 、 $\alpha$  が種々の値を取るとき、曲線は図 1-1 の①のように、データにうまく適合したり、②のように、データとの適合性が悪かったりする。データによく適合する曲線を求める（すなわち、 $a$ 、 $\alpha$  の値を決める）方法に、最小二乗法と呼ばれる方法がある。最小二乗法では、図 1-1 に示した  $d_i$ （実測値と曲線間の距離）の二乗和  $S = \sum_i d_i^2$  を目安にして、 $S$  を最小にする  $a$ 、 $\alpha$  の値を定める。

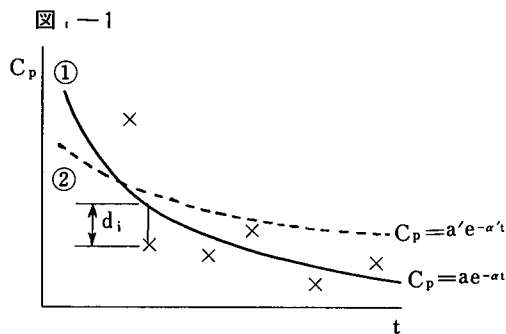


図 1-2 パラメータ推定の手順

モデル式;  $C_p = ae^{-\alpha t}$

データ;  $\{C_{pi}, t_i\} = 1 \sim k$

↓

モデル式のあてはめ計算; 最小二乗法

↓

パラメータの推定値;  $\hat{a}, \hat{\alpha}$

この間の手続きを図式化すると図1-2のようになる。データに対してモデル式を想定し、モデル式のあてはめ計算を行うことによって、モデル式中の未知パラメータの値が推定される。

データ中には誤差が含まれているため、求められたパラメータの値は個体の真の値ではなく、真の値を推定した値となっている。パラメータの真の値はこのくらいであろうというように、ヤマをかける気持ちをこめて、推定値にヤマ記号を付して  $\hat{a}$ ,  $\hat{a}$  と記す。

$d_j$  は残差と呼ばれ、測定された個体での血中濃度値の短時間内のゆらぎ、血中濃度の測定誤差、および想定したモデル式が現実には適合しないことによる実測値との開きなどから成り立っている。

## 2. 集団に対するモデル式のあてはめ

血中濃度が複数の被験者について測定されているとき、各個体をそれぞれ問題にするのであれば、前記のあてはめ計算を各人について行えばよい。しかし、測定された集団についての平均的な薬物動態を知ろうとするときには種々の方法が考えられる。

同一の個体のデータであるか否かを無視して、全データをプールし、あたかも一個体のデータであるかのように考えてあてはめ計算を行う方法は NPD 法 (naive-pooled data method) と呼ばれる。

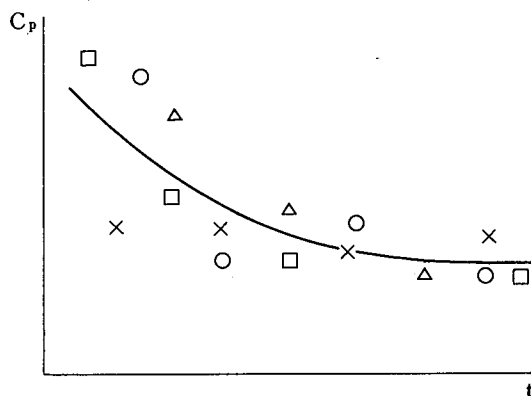
図1-3に NPD 法の概略を示す。図の○, ×, △, □は異なった4名のデータに対応する。曲線のあてはめは、個体を無視して行われるため、個体差に関する情報は得られない。

すべての個体についての測定の時点が揃っているときには、各時間ごとに測定値の平均値を算出し、平均値に対してあてはめ計算を行うことができ、NAD 法 (naive-averaging-of-data method) と呼ばれる。

各個体について十分な測定点の数があり、個体ごとにあてはめが可能となるときには、各人から得られるパラメータの値の適当な代表値を求めることにより、集団の平均的な動態を知ることができるが、この方法は、STS 法 (standard two-stage method) と呼ばれる。

モデルが  $C_p = ae^{-\alpha t}$  のときの例を図1-4に示す。図の左にみられるように、各個体の  $\hat{a}_i$ ,

図1-3 NPD法





$\hat{a}_i$  が求められ、図の右のように、 $\hat{a}_i$ ,  $\hat{\alpha}_i$  から平均的な  $\bar{a}$ ,  $\bar{\alpha}$  および、個体間のばらつき  $\sigma_a$ ,  $\sigma_\alpha$  が求められる。

一個体ごとにあてはめ計算を行わずに、モデル式を  $C_{pij} = a_i e^{-\alpha t_j}$  ( $i=1 \sim n$ ) として、各人に個有のパラメータ  $a_i$  を導入し、個体ごとの残差二乗和を全個体について加えた値を最小にするように  $a_i$  ( $i=1 \sim n$ ) と  $\alpha$  を決める方法もある。

$$d_{ij} = a_i e^{-\alpha t_j} - \hat{C}_{pij}$$

$$S_1 = d_{11}^2 + d_{12}^2 + \dots + d_{1k}^2 \quad (1人目)$$

$$S_2 = d_{21}^2 + d_{22}^2 + \dots + d_{2k}^2 \quad (2人目)$$
  

$$S_n = d_{n1}^2 + d_{n2}^2 + \dots + d_{nk}^2 \quad (n人目)$$

$$S_T = S_1 + S_2 + \dots + S_n \Rightarrow \text{最小にする。}$$

ここで各人について  $a_i$  を別々に決めることは、分布容積などの Size Factor を個人ごとに変えることを意味する。必要に応じ、 $\alpha$  の個体差も考えて  $\alpha_i$  とすることもできる。この方法は、複数の個体を同時に取り扱うことから同時的解法と呼ぶことができ、個体差  $\sigma_a$  ( $\sigma_\alpha$ ) についての情報が得られる (図1-5)。

図1-4 STS法

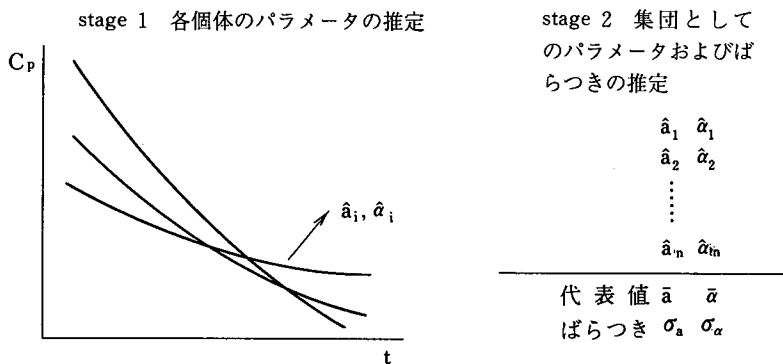


図1-5 同時的解法

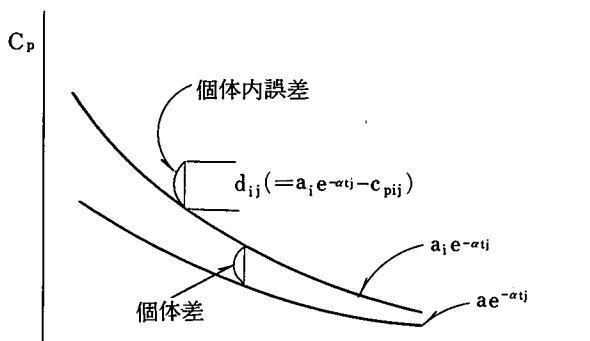
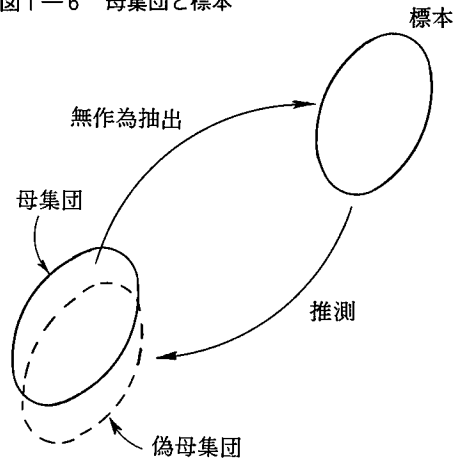


図1-6 母集団と標本



上記のいずれの方法を用いるにせよ、集団についての平均的な動態を調べたことを、population pharmacokinetics と呼ぶ。Population という語は、統計学では母集団を指している。統計学では、特定の集団から無作為に見本（標本、サンプル）を抽出し、その標本についての測定結果から、もとの母集団に関する推測をするという手順をとる。実際の臨床の場では、母集団から無作為に被験者を抽出することは稀であり、たまたま集まった個体について試験・測定を行う。この場合には、試験によって得られた知見を適用すべき集団として、逆に母集団を想定しているので偽母集団とも呼ばれる（図1-6）。

### 3. 誤差のモデル化

実測値が、モデルで与えられる部分  $C_p = ae^{-\alpha t}$  に誤差  $\varepsilon$  を付加したものであると考えたとき、誤差は種々の方法で付加される。

$C_p = ae^{-\alpha t} + \varepsilon$  の形で付加したとき、誤差に関して加法モデルとなるが、 $\varepsilon$  は 0 を中心に正・負両方の値をとる。しかし、 $C_p$  が負となる（誤差が  $-ae^{-\alpha t}$  よりさらに小さい）場合は意味をなさない。このモデルでは、 $C_p$  の大小にかかわらず一定の  $\varepsilon$  を想定しているが、一般のデータでは濃度  $C_p$  が大きくなるとともに誤差も大きくなる場合が多い点が不都合となる。 $\varepsilon$  の散らばり具合（分布）としては、正負等しく出現すると考えられる場合には、正規分布が用いられる。

$C_p = ae^{-\alpha t} \varepsilon$  として、乗法モデルにより誤差を付加することも可能である。この場合、 $\varepsilon$  の値は正であり、誤差がない場合は  $\varepsilon = 1$  に対応する。このモデルは、 $\varepsilon$  が  $0 \sim 1$  では稠密であり、1 より大きくなるに従って粗くなることが期待されるときにつづがうがよく、 $\varepsilon$  の分布としては対数正規分布（ $\varepsilon$  の対数値が正規分布する）を想定するのがよい。乗法モデルは、誤差が測定値に比例する場合に適している。

$C_p = ae^{-\alpha t} \varepsilon$  の両辺の対数をとると、 $\ln C_p = \ln a - \alpha t + \ln \varepsilon$  となり、一次式（線形式）に帰着できる。ところが、モデル式が  $C_p = ae^{-\alpha t} + be^{-\beta t}$ （静注 2-コンパートメントモデル）の場

合には、誤差  $\varepsilon$  を乗法的に付加したとき、誤差  $\varepsilon$  については対数変換により加え算の形となるが、パラメータ  $a$ ,  $\alpha$ ,  $b$ ,  $\beta$  に関しては一次式の形とすることができない。適当な変換により一次式とすることができる場合を本質的には線形であるといい、一次式とすることができない場合を本質的に非線形であるという。

前節の個体差  $a_i$  について考えたとき、 $a_i = \bar{a}e^{\eta_i}$  の形で個体差をモデル化し、個体内の誤差を  $e^\varepsilon$  の形で乗法的にモデル化すると、 $C_p = \bar{a}e^{\eta_i}e^{-\alpha t}e^\varepsilon$  となり、両辺の対数をとると、 $\ln C_p = \ln \bar{a} + \eta_i - \alpha t + \varepsilon$  の形に線形化される。

#### 4. 回帰関係の扱い

薬物の動態に対するサイズ要因としての体重や、腎機能による影響を調べるとき、集団の平均的動態の検討法の差異に応じて種々の扱いが可能である。例として、静注1-コンパートメントモデルにつきクレアチニンクリアランス  $CL_{cr}$  による影響を検討することを考える。

NPD法およびNAD法では、被験者を  $CL_{cr}$  の値によっていくつかのサブグループに分割し、それぞれのグループごとにパラメータ値を求める。次にサブグループごとの  $CL_{cr}$  の代表値を横軸にとり、縦軸にパラメータ値をプロットする(図1-7)。

TS法では、各人についてパラメータ値が得られるので、この値を各人の  $CL_{cr}$  の値に対してプロットすればよい(図1-8)。

同時解法では、 $\alpha = \alpha_0 + sCL_{cr}$  と回帰関係をモデル化し、 $C_p = ae^{-(\alpha_0 + sCL_{cr})t}$  のモデル式を用いて、前述の(同時的)最小二乗法を行うことにより、 $a$ ,  $\alpha_0$ ,  $s$  の推定値を得る。モデル  $C_p = ae^{-(\alpha_0 + sCL_{cr})t}$  において、サイズ要因が体重  $WT$  のべき乗に指数的に比例すると仮定すれば、 $a = a_0e^{WT^b}$ ,  $\alpha = \alpha_0 + sCL_{cr}$  として、個体内の誤差を  $e^\varepsilon$  の形で乗法的にモデル化すると、

図1-7 NPD法, NAD法による回帰関係の推定

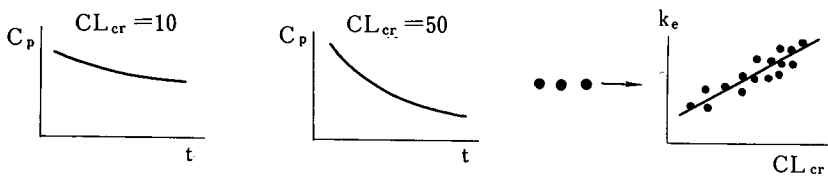
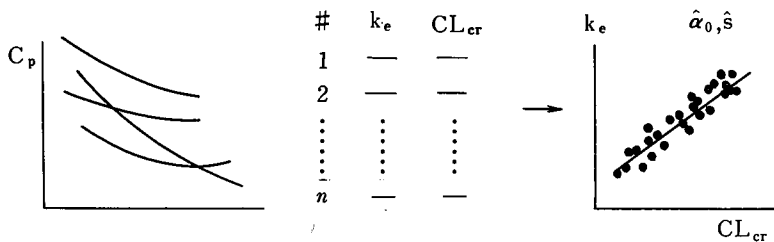


図1-8 TS法による回帰関係の推定



$$C_p = a_0 e^{WT^\theta} e^{-(\alpha_0 + SCL_{cr})t} e^\varepsilon \quad (1.1)$$

となり、両辺の対数をとると、

$$\ln C_p = \ln a_0 + WT^\theta - (\alpha_0 + SCL_{cr})t + \varepsilon \quad (1.2)$$

となる。ここで、 $WT^\theta$  を  $WT$  の二次式で近似してみる。改めて定数  $p$ ,  $q$ ,  $r$  を定めると、

$$\ln C_p = p + qWT + rWT^2 - (\alpha_0 + SCL_{cr})t + \varepsilon \quad (1.3)$$

の形になる。

このモデルの場合、体重  $WT$ 、クレアチニンクリアランス  $CL_{cr}$  および血中濃度がわかれば、通常の最小二乗法により推定値  $\hat{p}$ ,  $\hat{q}$ ,  $\hat{r}$ ,  $\hat{s}$  を求めることができる。

サイズ要因および  $CL_{cr}$  の影響をとり込むための前記の回帰式において個体差を考える。

$$a = a_0 e^{(WT^\theta + \eta_a)} \quad (1.4)$$

$$\alpha = \alpha_0 + SCL_{cr} + \eta_\alpha \quad (1.5)$$

そうすると、

$$C_p = a_0 e^{(WT^\theta + \eta_a)} e^{-(\alpha_0 + SCL_{cr} + \eta_\alpha)t} e^\varepsilon \quad (1.6)$$

というモデルが得られる。

## 5. モデルの実例

ここでは、モデルを簡略に示すために、個体と時点を示す添え字は省いた。

### 経口1-コンパートメントモデル

$$C_p = \frac{k_a F D}{Vd(k_a - k_e)} (e^{-k_e t} - e^{-k_a t}) e^\varepsilon \quad (1.7)$$

$$k_a = \tilde{k}_a e^{\eta_{k_a}} \quad (1.8)$$

$$Vd = \tilde{V}d e^{\eta_{Vd}} \quad (1.9)$$

$$k_e = CL/Vd \quad (1.10)$$

$$CL = \tilde{C}L e^{\eta_{CL}} \quad (1.11)$$

このモデルでは、個体内誤差は  $e^\varepsilon$  の形で乗法的に扱っている。また、 $k_a$ ,  $Vd$ ,  $CL$  について個体差を考慮し、それぞれ  $e^\eta$  の形で乗法的に付している。

### 静注2-コンパートメントモデル

$$C_p = (Ae^{-\alpha t} + Be^{-\beta t}) e^\varepsilon \quad (1.12)$$

$$A = \tilde{A} e^{\eta_A} \quad (1.13)$$

$$\alpha = \bar{\alpha} e^{\eta_\alpha} \quad (1.14)$$

$$B = \bar{B} e^{\eta_B} \quad (1.15)$$

$$\beta = \bar{\beta} e^{\eta_\beta} \quad (1.16)$$

このモデルでも、個体内誤差は  $e^\varepsilon$  の形で乗法的に付し、 $A$ ,  $\alpha$ ,  $B$ ,  $\beta$  にそれぞれ個体差を乗法的に付している。

**Michaelis-Menten 型の薬物動態における投与量と定常状態の濃度の関係式**

$$R = V_m C_{PSS} / (K_m + C_{PSS}) + \varepsilon \quad (1.17)$$

$$V_m = \bar{V}_m e^{\eta_{V_m}} \quad (1.18)$$

$$K_m = \bar{K}_m e^{\eta_{K_m}} \quad (1.19)$$

ここでは、 $\varepsilon$  は加法的に、 $\eta$  は乗法的に扱っている。

## 6. モデル式の線形化

モデル式  $C_p = ae^{-\alpha t} + be^{-\beta t}$  において、 $\beta = \beta_0 + sCL_{cr} + \eta_\beta$  として回帰関係および個体差をとり入れたとする。このモデルでは、今までの例のように両辺の対数をとって線形化することはできない。そこで、マクローリン展開を利用して線形化を行う。

マクローリン展開は次の式により与えられる。

$$f(x) = f(0) + f'(0)x + H \cdot O \cdot T \quad (1.20)$$

ここで、 $H \cdot O \cdot T$  は、二次以上の項を示す。たとえば、

$$e^x = e^0 + d(e^x)/dx|_{x=0} x + H \cdot O \cdot T = 1 + x + H \cdot O \cdot T \quad (1.21)$$

マクローリン展開を用いて  $C_p = ae^{-\alpha t} + be^{-(\beta_0 + sCL_{cr} + \eta_\beta)t}$  を  $\eta_\beta$  に関して線形化しよう。

$$e^{-(\beta_0 + sCL_{cr} + \eta_\beta)t} \Big|_{\eta_\beta=0} = e^{-(\beta_0 + sCL_{cr})t} \quad (1.22)$$

$$d(e^{-(\beta_0 + sCL_{cr} + \eta_\beta)t})/d\eta_\beta \Big|_{\eta_\beta=0} = -e^{-(\beta_0 + sCL_{cr})t} t \quad (1.23)$$

であるから、

$$e^{-(\beta_0 + sCL_{cr} + \eta_\beta)t} = e^{-(\beta_0 + sCL_{cr})t} - e^{-(\beta_0 + sCL_{cr})t} t \eta_\beta + H \cdot O \cdot T \quad (1.24)$$

となる。

高次の項を無視すると、

$$C_p = ae^{-\alpha t} + b(e^{-(\beta_0 + sCL_{cr})t} - e^{-(\beta_0 + sCL_{cr})t} t \eta_\beta) \quad (1.25)$$

となり、 $\eta_\beta$  については線形の式が得られる。

線形化する前の式を使用して未知パラメータの推定を行う直接法も可能ではあるが、非常に困難であるため、モデル式を個体差について線形化し、その上でパラメータの値を推定しよう

とするのが後述の NONMEM の方法である。

モデル式が、個体内誤差  $\epsilon$  に関しても非線形であるときには、 $\epsilon$  についても同様に線形化を行う。

## 7. 投与設計法

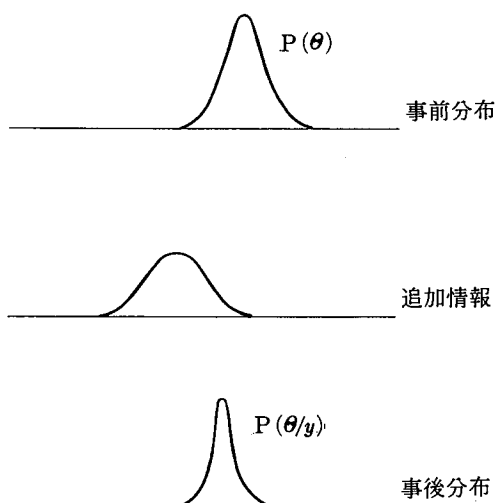
新しい個体についての動態を予測するとき、その個体についての情報がなければ、すでに知られている平均的な動態により予測を行う。もし、その個体について、体重、 $CL_{cr}$  等の回帰変数の値が得られていれば、この値を回帰関係にとり入れたモデルに適用することにより、予測の精度を上げることができる。ノモグラムを利用して動態を予測し、投与量を決定する方法がこれに該当する。

予測の精度を高めるために、問題とする個体に試験的な用量を投与して、その個体についてパラメータを算出し、以後の予測に用いるという方法がある。この方法では、精度の良い推定をすることが可能であるが、問題としている個体の動態が変化するともう一度パラメータを求め直すことが必要になる。さらに、動態を解析できるスタッフが常駐している必要があること、パラメータを求めるために多くのサンプル点数が必要であり、費用が大となる一方、結果が出るまでに時間がかかる等の欠点がある。

これと似た方法として、投与後一定時間経過後の血中濃度と平均的な動態から、個体での動態を予測する方法や最小の点数のデータからパラメータの値を方程式によって解く方法等がある。

集団としての平均的な動態がわかっているとき、新しい個体について得られた動態の情報によりもとの平均的な動態を修飾し、個体に合う予測値を求める方法にベイズ (Bayes) の方法がある。

図1-9 事前分布, 追加情報, 事後分布の関係



この方法の概略を図1—9に示す。図の上段には新しい個体についての情報が得られる前のパラメータ  $\theta$  の分布（事前分布と呼ばれる）を示す。中段には新しいデータが得られたとき、パラメータ  $\theta$  からその値が得られる確率に相当する値（尤度）のグラフを示している。下段には追加された中段の情報を利用して改良された  $\theta$  の分布（事後分布と呼ばれる）を示す。

図にみられるように、事前分布が追加情報によって修飾され、追加データの方向に寄った事後分布が得られる。

ベイズの方法によると、追加情報が少ない間は平均的動態による推定の性格が強く、追加情報が増すに従って、すなわち、新しい個体についての動態がよりよくわかるに従って、その個体固有の推定値が得られるという形で推定が行われる。経験的ベイズの方法によると、実際のパラメータの推定は次のように行われる。

$\theta_k$  を集団のパラメータ、 $\hat{\theta}_k$  を個体のパラメータ、 $\omega_k^2$  を  $\theta_k$  の分散とし、 $C_{pj}$  を新たに得られた測定値、 $\hat{C}_{pj}$  をパラメータ  $\hat{\theta}_k$  を用いた予測値、 $\sigma^2$  を  $C_{pj}$  の分散としたとき、

$$Ob = \sum_j (C_{pj} - \hat{C}_{pj})^2 / \sigma^2 + \sum_k (\theta_k - \hat{\theta}_k)^2 / \omega_k^2 \quad (1.26)$$

の値を最小にするように  $\hat{\theta}_k$  を定める。この式は、測定値による項と、パラメータによる項の重みつき最小二乗法の形となっている。

ベイズの方法では、個体について新しい情報が得られるたびに、逐次的にパラメータの推定値を改良することが可能である。

## 8. NONMEM の概要

個体差、さらには回帰関係を考慮に入れて、同時的解法により薬物動態パラメータを推定するコンピュータプログラムに NONMEM がある。NONMEM という名前は、Nonlinear Mixed Effect Model に由来している。

一般的なモデル式を、

$$y = f(t, \theta, x, \eta, \epsilon)$$

$y$  : 血中濃度など、解析を目的とする変数、 $t$  : 時刻、 $\theta$  : 薬物動態パラメータ、 $x$  : 回帰変数、 $\eta$  : 個体差、 $\epsilon$  : 個体内誤差

としたとき、NONMEM を使用するために、

$$y = f(t, \theta, x, 0, 0) + g(t, \theta, x)\eta + h(t, \theta, x)\epsilon$$

の形に線形化する。NONMEM では、この  $f, g, h$  をユーザーが作成するプログラム（サブルーチン）により計算し、NONMEM 側に戻すようになっており、広汎なモデルに対応できる。

NONMEM で解析すべきデータおよび解析条件の指示は、カード（または相等の媒体）から読み込ませる。入力構成はおおよそ、

- ① 変数の数、名前等データの記述
- ② データ

- ③ パラメータの個数等モデルの記述と、解析のためのパラメータの初期値
- ④ 出力オプション等の指定

より成り立っている。

NONMEM は、指定された条件のもとに、最小二乗法と類似の方法により、パラメータの推定を行う。この方法の考え方の概略を示すと、次のようになる。

通常の最小二乗法では、

$$Ob = \sum (y_i - \hat{y}_i)^2$$

として、目的関数  $Ob$  の値を最小とするパラメータの値を推定値とする。次に、重みづき最小二乗法では、目的関数  $Ob$  を、

$$Ob = \sum (y_i - \hat{y}_i) W_i^{-1} (y_i - \hat{y}_i)$$

として、 $Ob$  を最小にするように推定値を求める。ところが、重み  $W_i$  の中に未知パラメータが含まれる場合には、本質的に  $Ob$  の値が改良されていないにもかかわらず、 $W_i$  が大きくなることによって  $Ob$  の値が小さくなり得る。このことに対するペナルティを付加した形の

$$Ob = \sum (\ln W_i + (y_i - \hat{y}_i) W_i^{-1} (y_i - \hat{y}_i))$$

を目的関数としてパラメータの推定を行う。この目的関数の形は、誤差に正規分布を仮定したときの最尤法とよばれる方法によって導くことができる（詳細は次節参照）。

実際のパラメータ推定の計算は次のようにして行われる。まず、入力されたパラメータの初期値  $\theta_0$  より、 $Ob$  の値を減少させる方向に  $\Delta\theta_1$  だけ  $\theta$  の値を変化させ、 $\theta_1 = \theta_0 + \Delta\theta_1$  とする。次に、 $Ob$  の値をさらに減少させる方向に  $\Delta\theta_2$  だけずらし、 $\theta_2 = \theta_1 + \Delta\theta_2$  とする。以下、次々に  $\theta_i$  の値を更新し、 $Ob$  の減少があらかじめ指定された大きさ以下になったときに計算を終了し、そのときの  $\theta_i$  の値を、推定値  $\hat{\theta}$  とする。

〔緒方宏泰，魚井 徹〕



## 第 3 節

# 統計的背景

### 1. ベクトルと行列

ここでは、以後の解説や、関連文献を読む上で必要なベクトル・行列に関する用語を説明する。

まず、ベクトルのスカラー倍、ベクトルとベクトルの和、ベクトルとベクトルの内積、行列、行列の和、行列の積などの基本的な演算規則を記す。次に、線形および非線形最小二乗法を解く際に必要となる連立一次方程式の解法に関係する逆行列を、また、NONMEMで解析するときに、目的関数の中に現れる行列式について記す。最後に、ベクトル記法に慣れるための練習として、統計量やモデルがベクトル記法によってどのように表されるのかをみる。

複数の測定値の集まりをベクトルと呼ぶ。測定値が  $y_i (i=1 \sim n)$  であるとき、 $y_i$  を縦方向に並べたものを一つの記号  $y$  で表し、列ベクトルと呼ぶ。また、横方向に並べたものを記号  $y'$  で表し、行ベクトルと呼ぶ。

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad y' = (y_1, y_2, \dots, y_n) \quad (1.27)$$

ベクトルに対して、単一の測定値をスカラーと呼ぶ。

ベクトルのスカラー  $c$  倍は、ベクトルの各要素  $y_1 \sim y_n$  を  $c$  倍したものとする。

$$cy = \begin{pmatrix} cy_1 \\ cy_2 \\ \vdots \\ cy_n \end{pmatrix} \quad (1.28)$$

また、二つのベクトル  $x$  と  $y$  の和  $x+y$  は、各々の対応する要素を加え合わせたものとする。

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad (1.29)$$

そうすると、 $a\mathbf{x} + b\mathbf{y}$  は次のようになる。

$$a\mathbf{x} + b\mathbf{y} = \begin{pmatrix} ax_1 + by_1 \\ ax_2 + by_2 \\ \vdots \\ ax_n + by_n \end{pmatrix} \quad (1.30)$$

すべての要素が0であるベクトルをゼロベクトルと呼び、 $\mathbf{0}$  で表す。 $-\mathbf{x}$  は、各要素を負にしたものとする。

行ベクトル  $\mathbf{x}'$  と列ベクトル  $\mathbf{y}$  の対応する要素  $x_i, y_i$  をかけ合わせたものの和 (積和) を  $\mathbf{x}'\mathbf{y}$  で表し、 $\mathbf{x}$  と  $\mathbf{y}$  の内積と呼ぶ。

$$\mathbf{x}'\mathbf{y} = \sum_i x_i y_i \quad (1.31)$$

ベクトル  $\mathbf{x}$  と  $\mathbf{x}$  自身の内積は、 $\mathbf{x}'\mathbf{x} = \sum_i x_i^2$  であり、 $\mathbf{x}$  の長さの二乗となっている。 $\mathbf{x}$  の長さを  $\|\mathbf{x}\|$  で表すと、

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}'\mathbf{x}} \quad (1.32)$$

となる。

二つのベクトル  $\mathbf{x}, \mathbf{y}$  の長さ  $\|\mathbf{x}\|, \|\mathbf{y}\|$  と、 $\mathbf{x}$  と  $\mathbf{y}$  の内積の間には次の関係がある。

$$\mathbf{x}'\mathbf{y} = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos \theta \quad (1.33)$$

ここで、 $\theta$  は  $\mathbf{x}$  と  $\mathbf{y}$  の間の角度であり、二つのベクトルの向いている方向の近さを表す。

$n$  人について時刻  $t_1, t_2, \dots, t_k$  で測定が行われ、値  $y_{ij}$  が得られているとき、 $y_{ij}$  の集まりを一つの記号  $\mathbf{Y}$  で表し、 $n$  行  $k$  列の行列と呼ぶ。

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1k} \\ y_{21} & & & \\ \vdots & & & \\ y_{n1} & \dots & \dots & y_{nk} \end{pmatrix} \quad (1.34)$$

行列  $\mathbf{Y}$  のスカラー倍  $c\mathbf{Y}$  は、行列  $\mathbf{Y}$  の各要素を  $c$  倍したものとし、二つの  $n$  行  $k$  列の行列の和は、各々の対応する要素を加え合わせた値を要素とする行列によって定義する。

$$CY = \begin{pmatrix} cY_{11} & cY_{12} \cdots cY_{1k} \\ cY_{21} \\ \vdots \\ cY_{n1} \cdots \cdots cY_{nk} \end{pmatrix} \quad (1.35)$$

$$X+Y = \begin{pmatrix} x_{11}+y_{11} & x_{12}+y_{12} \cdots x_{1k}+y_{1k} \\ x_{21}+y_{21} \\ \vdots \\ x_{n1}+y_{n1} \cdots \cdots x_{nk}+y_{nk} \end{pmatrix} \quad (1.36)$$

行列は、縦ベクトルを横方向に並べたものとも、横ベクトルを縦方向に並べたものとも考えられる。

$$X = \begin{pmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_n \end{pmatrix} = (b_1 \quad b_2 \cdots b_k) \quad (1.37)$$

$$a'_i = (x_{i1} \quad x_{i2} \cdots x_{ik}) \quad (1.38)$$

$$b_j = \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} \quad (1.39)$$

二つの行列 X, Y の積は次のように定義される。

$$X = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} \quad (1.40)$$

$$Y = (y_1 \quad y_2 \cdots y_n) \quad (1.41)$$

$$Z = XY \quad (1.42)$$

ここで、 $z_{ij} = x'_i y_j$  ( $x_i$  と  $y_j$  の内積) である。この積が定義されるためには、 $x'_i$  と  $y_j$  の長さ、すなわち、X の列の長さと Y の行の長さが等しくなければならない。行列 Z の行と列を入れかえたものを転置行列と呼び Z' で示すが、 $(XY)' = Y'X'$  の関係がある。

縦、横の長さが等しい行列を正方行列と呼ぶ。正方行列で、すべての対角要素  $x_{ii}$  が 1 で、残りの要素  $x_{ij}$  ( $i \neq j$ ) がゼロである行列を単位行列と呼び、I で表す。

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \quad (1.43)$$

スカラー  $c$  に何かを乗じて 1 になるようなものを逆数と呼び、 $c^{-1}$  で表していた。

$$c \cdot c^{-1} = 1 \quad (1.44)$$

これと同様に、正方行列  $\mathbf{X}$  に別の行列を乗じて単位行列となるとき、その乗じた行列を逆行列と呼び、 $\mathbf{X}^{-1}$  で表す。

$$\mathbf{X}\mathbf{X}^{-1} = \mathbf{I} \quad (1.45)$$

連立一次方程式  $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ik}x_k = b_i$  ( $i=1 \sim k$ ) は、ベクトルと行列を用いて次のように表される。

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & & & \\ \vdots & & & \\ a_{k1} & \cdots & \cdots & a_{kk} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix} \quad (1.46)$$

すなわち、

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (1.47)$$

となる。

ここで、行列  $\mathbf{A}$  とベクトル  $\mathbf{x}$  のかけ算は、ベクトル  $\mathbf{x}$  を  $k$  行 1 列のマトリックスと考えて、マトリックスのかけ算の規則を使用している。すなわち、

$$\begin{aligned} \begin{pmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_k \end{pmatrix} \mathbf{x} &= \begin{pmatrix} a'_1 \mathbf{x} \\ a'_2 \mathbf{x} \\ \vdots \\ a'_k \mathbf{x} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kk}x_k \end{pmatrix} \end{aligned} \quad (1.48)$$

となっている。

行列・ベクトルに単位行列  $\mathbf{I}$  を乗じても値は変化しない。

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & & & \\ \vdots & & & \\ a_{k1} & \cdots & \cdots & a_{kk} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & & & \\ \vdots & & & \\ a_{k1} & \cdots & \cdots & a_{kk} \end{pmatrix} \quad (1.49)$$

これから、連立方程式  $Ax=b$  を解くには、

$$Ax = b \quad (1.50)$$

$$A^{-1}Ax = A^{-1}b \quad (1.51)$$

$$Ix = A^{-1}b \quad (1.52)$$

$$x = A^{-1}b \quad (1.53)$$

となり、 $A^{-1}$  が求めれば  $x$  を求めることができる ( $A^{-1}$  を求めるプログラムはコンピュータプログラムパッケージ中に通常含まれている)。

連立方程式中に同じ式が2度以上現れるときなどでは、方程式を解くことができず、不定となる。このときには、 $A^{-1}$  を求めることができないわけであるが、 $A^{-1}$  が求められる (すなわち、方程式が解ける) ための条件とは、 $|A| \neq 0$  であることが示される。

ここで、 $|A|$  は行列式と呼ばれ、次のように定義される。

数の並び、1, 2, 3の順序を入れかえたものを列挙するとともに、もとの並びを併記する。

1 2 3	1 2 3
1 3 2	1 2 3
2 1 3	1 2 3
3 1 2	1 2 3
2 3 1	1 2 3
3 2 1	1 2 3

次に、二つの並びの対応したものを組にして、下記の並びを作成する。

符合(sign)				
1	1	2	3	+
1	1	2	3	-
1	2	2	3	+
1	3	2	3	-
1	2	3	1	+
1	3	2	1	-

ここで、枠で囲った数字はもとの並び、囲わないものは並べかえた方の並び (並べかえなしも含む) である。また、符号 (sign) はもとの並びから偶数回の入れかえによって得られた場合に+、奇数回の場合に-とする。この数字の並びと符号を使用して、行列式は次式で与えられる。

$$|A| = \sum \text{sign } a_{1i} a_{2j} a_{3k} \quad (1.54)$$

( $\sum$  は並びすべてについての和を表す)

上記は三次の場合について記したが、 $n$  次の場合でも同様である（行列式の計算プログラムも、通常のパッケージ中に含まれている）。

## 2. モデルのベクトル表現

測定値  $y_i$  は、平均値  $\mu$  に誤差  $\varepsilon_i$  が加わったものであると考え、 $y_i = \mu + \varepsilon_i$ 、ベクトル表現では  $\mu' = (\mu, \mu, \mu, \dots)$  として、

$$y = \mu + \varepsilon \quad (1.55)$$

で表される。

平均値は、単位ベクトル  $j' = (1 \ 1 \ 1 \ \dots)$  として、

$$\frac{\sum x_i}{n} = \frac{1}{n} x' j \quad (1.56)$$

また、分散は、平均値ベクトル  $\bar{x} = (\bar{x} \ \bar{x} \ \dots)$  を用いて、

$$\begin{aligned} S_{xx} &= \frac{\sum (x_i - \bar{x})^2}{n} = \frac{1}{n} (x - \bar{x})'(x - \bar{x}) \\ &= \frac{1}{n} \|x - \bar{x}\|^2 \end{aligned} \quad (1.57)$$

同様に、 $x$  と  $y$  の共分散は、

$$\begin{aligned} S_{xy} &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n} \\ &= \frac{1}{n} (x - \bar{x})'(y - \bar{y}) \end{aligned} \quad (1.58)$$

等とベクトルによって表される。

さらに、相関係数は

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{(x - \bar{x})'(y - \bar{y})}{\|x - \bar{x}\| \cdot \|y - \bar{y}\|} \quad (1.59)$$

で与えられる。

ところで、 $(x - \bar{x})'(y - \bar{y}) = \|x - \bar{x}\| \cdot \|y - \bar{y}\| \cos \theta$  であったから、これを上式に代入すると、 $r = \cos \theta$  となり、相関係数は二つのベクトル  $x$ 、 $y$  の間の角度、すなわち、 $x$ 、 $y$  の方向の近さを測っている。

分散と共分散を次のように並べた行列を分散共分散行列という。

$$\Sigma = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{pmatrix} \quad (1.60)$$

分散共分散行列は、多変量の場合でも同様に定義され、その行列式 $|\Sigma|$ は一般化分散と呼ばれ、多次元空間での広がりを表す。

説明変数が複数個あるときの重回帰式は下記で表される。

$$y_1 = \alpha_1 x_{11} + \alpha_2 x_{12} + \cdots + \varepsilon_1 \quad (1.61)$$

$$y_2 = \alpha_1 x_{21} + \alpha_2 x_{22} + \cdots + \varepsilon_2$$

⋮

$$y_n = \alpha_1 x_{n1} + \alpha_2 x_{n2} + \cdots + \varepsilon_n$$

これをベクトル表現で表すと、

$$y = X\alpha + \varepsilon \quad (1.62)$$

となり、

最小二乗法によるパラメータベクトル $\alpha$ の推定値 $\hat{\alpha}$ は後に示すが、

$$\hat{\alpha} = (X'X)^{-1}X'y \quad (1.63)$$

によって与えられる。

### 3. 正規分布

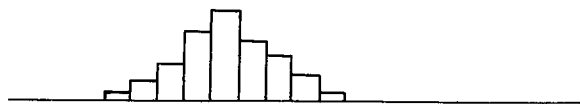
統計的推定、検定を行ううえで、誤差の分布として正規分布を仮定することが多い。ここでは、正規分布、対数正規分布について記すとともに、期待値、母平均、母分散について記す。後に、ELS関数の形を導く上で、正規分布の仮定が使用される。

連続的な値をとる測定値について、ヒストグラムを作成したとき、図1-10のように単峰性で左右対称な形になることが多い。このヒストグラムに当てはまる関数として、

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1.64)$$

を想定する。式(1.64)は正規分布の密度関数と呼ばれる。

図1-10 正規分布に従う変数のヒストグラム



$f(x)$ は有限の $x$ について常に正であり、 $(x-\mu)$ を $-(x-\mu)$ とおきかえても式の値が不変で

あるので、 $x=\mu$  に関して対称である。また、 $\lim_{x \rightarrow \infty} f(x)=0$  であるので、 $x$  が大きくなるにつれて  $f(x)$  は  $x$  軸に漸近する。

式(1.64)を  $x$  について微分すると、

$$f'(x) = -\frac{(x-\mu)}{\sigma^2} f(x) \quad (1.65)$$

$$f''(x) = -\frac{1}{\sigma^2} \left(1 + \frac{x-\mu}{\sigma}\right) \left(1 - \frac{x-\mu}{\sigma}\right) f(x) \quad (1.66)$$

となり、 $f'(\mu)=0$ 、 $f''(\mu)<0$  であることから、 $f(x)$  は  $\mu$  で最大値をとり、 $x=\mu \pm \sigma$  で  $f''(x)$  の符号が変化するので、変曲点となっている注1)。

また、

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad (1.67)$$

となっており、これは、“ $x$  が  $-\infty$  から  $\infty$  の間のいずれかの値をとる確率は全体で1”であることに対応する。

関数  $g(x)$  に  $x$  の値が起こり得る確率(密度)  $f(x)$  をかけて、 $x$  の全域にわたって積分した値 ( $g(x)$  の加重平均) は、 $g(x)$  の期待値と呼ばれる。 $g(x)=x$  の場合、期待値は母平均と呼ばれ、 $g(x)=(x-\mu)^2$  の場合は母分散と呼ばれる。

$$E(x) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \mu \text{ (母平均)} \quad (1.68)$$

$$V(x) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x) dx = \sigma^2 \text{ (母分散)} \quad (1.69)$$

変数を  $x$ 、 $y$  とし、定数を  $a$ 、 $b$  としたとき、

$$E(ax \pm b) = aE(x) \pm b \quad (1.70)$$

$$V(ax \pm b) = a^2 V(x) \quad (1.71)$$

$$E(ax \pm by) = aE(x) + bE(y) \quad (1.72)$$

$$V(ax \pm by) = a^2 V(x) \pm 2abC(x, y) + b^2 V(y) \quad (1.73)$$

注1) 微分の公式

$$(x^n)' = nx^{n-1}$$

$$(cf(x))' = cf'(x)$$

$$(f(x) \cdot g(x))' = f'(x)g(x) + f(x)g'(x)$$

$$(e^x)' = e^x$$

$$(\ln x)' = 1/x$$



となる。ここで、 $C(x, y)$  は  $x$  と  $y$  の共分散であり、

$$C(x, y) = E[(x - \mu_x)(y - \mu_y)] \quad (1.74)$$

で与えられる。母相関係数は、

$$\rho = \frac{C(x, y)}{\sqrt{V(x)}\sqrt{V(y)}} \quad (1.75)$$

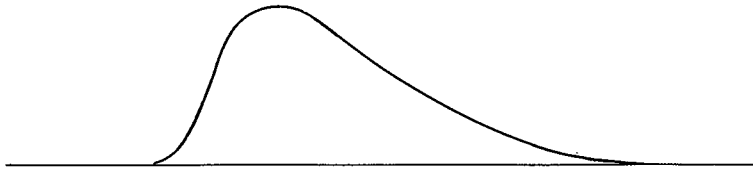
によって定義される。

式(1.64)に示した正規分布については、 $E(x) = \mu$ 、 $V(x) = \sigma^2$ となっている。

統計学では、確率的に変動する変数を変量または確率変数と呼ぶが、確率変数が母平均  $\mu$ 、母分散  $\sigma^2$  の正規分布に従うとき、 $x \sim N(\mu, \sigma^2)$  と記す。

薬物動態のデータでは、測定値が正規分布ではなく、0に近い値から無限大まで図1-11のような分布を示すことが多い。横軸  $x$  の代わりに  $\log(x)$  を軸として分布を描いて正規分布となるとき、 $x$  は対数正規分布に従うといわれる。

図1-11 対数正規分布



薬物動態のモデル  $C_p = C_{p0} e^{-k_e t}$  で、 $k_e$  が誤差を伴い、 $k_e = k_{e0} + \epsilon$ 、 $\epsilon \sim N(0, \sigma^2)$  であるとすると、

$$C_p = C_{p0} e^{-(k_{e0} + \epsilon)t} \quad (1.76)$$

両辺の対数をとると、

$$\ln C_p = \ln C_{p0} - (k_{e0} + \epsilon)t \quad (1.77)$$

となる。 $t$  が特定の時刻  $t_0$  では、 $\ln C_p$  は正規分布  $N(\ln C_{p0} - k_{e0} t_0, t_0^2 \sigma^2)$  に従うので、対数正規分布となっている。

#### 4. 最小二乗法

パラメータの推定において、最小二乗法は重要な役割を果たす。以下には、線形の最小二乗法を例にとり、正規方程式、パラメータの信頼区間の求め方を記し、次に、非線形最小二乗法ならびに重みつき非線形最小二乗法の考え方と、パラメータの区間推定のための推定値の分散・共分散行列を示す。

線形のモデル  $y = X'\beta + \varepsilon$  を考える。パラメータ  $\beta$  の推定値  $\hat{\beta}$  が求まったとき、その  $\hat{\beta}$  を使って  $y$  の推定値は  $\hat{y} = X'\hat{\beta}$  によって与えられる。残差  $d$  を  $d = y - \hat{y} (= y - X'\hat{\beta})$  として、実測値  $y$  と推定値  $\hat{y}$  の差とする。残差二乗和  $S = d'd$  を最小にするように  $\hat{\beta}$  を定める方法が最小二乗法である。そのためには、 $\partial S / \partial \hat{\beta} = 0$  を解けばよい。

$$S = d'd \quad (1.78)$$

$$\begin{aligned} &= (y - X'\hat{\beta})'(y - X'\hat{\beta}) \\ &= y'y - y'X'\hat{\beta} - \hat{\beta}'Xy + \hat{\beta}'XX'\hat{\beta} \\ &= yy' - 2\hat{\beta}'Xy + \hat{\beta}'XX'\hat{\beta} \end{aligned}$$

微分の公式  $\partial \beta' a / \partial \beta = a$ ,  $\partial \hat{\beta}' XX' \hat{\beta} / \partial \hat{\beta} = 2XX'\hat{\beta}$  を利用すると

$$\partial S / \partial \hat{\beta} = -2Xy + 2XX'\hat{\beta} = 0 \quad (1.79)$$

$$XX'\hat{\beta} = Xy \quad (1.80)$$

$$\hat{\beta} = (XX')^{-1}Xy \quad (1.81)$$

として、 $\beta$  の推定値を求めることができる。式(1.80)は正規方程式と呼ばれる。

正規方程式の形は、形式的に次のように覚えるとよい。

$$y = X'\beta + \varepsilon \quad (1.82)$$

$$y = X'\hat{\beta} \quad \text{形式的に } \varepsilon \text{ を落とす} \quad (1.83)$$

$$Xy = XX'\hat{\beta} \quad \text{両辺に } X \text{ をかける} \quad (1.84)$$

$$\hat{\beta} = (XX')^{-1}Xy \quad \text{両辺に } (XX')^{-1} \text{ をかける} \quad (1.85)$$

単回帰モデル  $y_i = a + bx_i + \varepsilon_i$  における  $a$ ,  $b$  の推定を上記の正規方程式を利用して行ってみよう。モデルは、ベクトル表現では、

$$y = X'\beta + \varepsilon \quad (1.86)$$

ここで、

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad (1.87) \quad X' = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \quad (1.88) \quad \beta = \begin{pmatrix} a \\ b \end{pmatrix} \quad (1.89)$$

したがって、

$$XX' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \quad (1.90)$$

$$XY = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \quad (1.91)$$

正規方程式は、

$$\begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \quad (1.92)$$

クラームルの公式により、

$$\begin{aligned} b &= \frac{\begin{vmatrix} n & \sum y_i \\ \sum x_i & \sum x_i y_i \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}} \\ &= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - \sum x_i \sum x_i} \\ &= \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2} \left( = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})/n}{\sum (x_i - \bar{x})^2/n} \right) \\ &= \frac{S_{xy}}{S_{xx}} \end{aligned} \quad (1.93)$$

また、式(1.92)より

$$na + b \sum x_i = \sum y_i \quad (1.94)$$

したがって、

$$a + b \bar{x} = \bar{y} \quad (1.95)$$

$$a = \bar{y} - b \bar{x} \quad (1.96)$$

これに、上記で求めた  $b$  を代入すれば  $a$  の値が求まる。

最小二乗法によって求まる  $\hat{\beta}$  の分散は、モデル  $y = X\beta + \varepsilon$  において、 $E(\varepsilon) = \mathbf{0}$ 、 $V(\varepsilon) = \sigma^2 \mathbf{I}$  とすると、後に式(1.113)に示すように、

$$V(\hat{\beta}) = \sigma^2 (XX')^{-1} \quad (1.97)$$

によって与えられる。これにより、パラメータの数を  $p$ 、 $(XX')^{-1}$  の第  $i$  対角要素を  $V_{ii}$  とし、 $\varepsilon$  が正規分布に従うとすると、

$$t = \frac{\hat{\beta}_i - \beta_i}{\sqrt{V_{ii} \hat{\sigma}^2}} \quad (1.98)$$

$$\hat{\sigma}^2 = (y - \hat{y})'(y - \hat{y}) / (n - p - 1) \quad (1.99)$$

となり, t 検定により  $\beta_i$  の有意性を検定することができる。また,  $\beta_i$  の信頼区間は,

$$\beta_{iL}^u = \hat{\beta}_i \pm t_{\alpha(n-p-1)} \sqrt{V_{ii} \hat{\sigma}^2} \quad (1.100)$$

によって与えられる。

モデルが非線形で,  $y = h(x, \theta) + \varepsilon$  で表されるとき,  $f = y - h(x, \hat{\theta})$  とすると, 最小二乗法は  $S = f'f$  を最小とする  $\hat{\theta}$  を求めることになる。このための計算方法は後に記すが, 推定値  $\hat{\theta}$  が求まったとき,

$$\hat{\sigma}^2 = f'f / (n - p) \quad (1.101)$$

$$\hat{V}(\hat{\theta}) = \hat{\sigma}^2 [J(\hat{\theta})J'(\hat{\theta})]^{-1} \quad (1.102)$$

と, 線形のモデルの場合と同様な形式で  $\hat{\theta}$  の分散が求められる。ここで,  $J'(\hat{\theta})$  は次のヤコビ行列である。

$$J(\hat{\theta}) = \begin{pmatrix} \frac{\partial f^1}{\partial \hat{\theta}_1} & \frac{\partial f^1}{\partial \hat{\theta}_2} & \dots & \frac{\partial f^1}{\partial \hat{\theta}_p} \\ \frac{\partial f^2}{\partial \hat{\theta}_1} & & & \\ \vdots & & & \\ \frac{\partial f^n}{\partial \hat{\theta}_1} & \dots & \dots & \frac{\partial f^n}{\partial \hat{\theta}_p} \end{pmatrix} \quad (1.103)$$

$$f = \begin{pmatrix} f^1 \\ f^2 \\ \vdots \\ f^n \end{pmatrix} \quad (1.104)$$

線形モデル  $y = X'\beta + \varepsilon$  の場合について,  $J'(\theta)$  を求めよう。

$$f^i = y_i - x_i' \hat{\beta} \quad (1.105)$$

$$\frac{\partial f^i}{\partial \hat{\beta}_j} = -x_{ij} \quad (1.106)$$

ここで,  $x_i'$  は  $X'$  のヨコベクトルを表す。

上記より,  $J'(\theta) = -X'$  となっており, 式(1.102)の表現は線形のモデルについても適用できることがわかる。

最小二乗法において,  $S = \sum (y_i - \hat{y}_i)^2$  の代わりに, 各データに重み  $W_i$  を設けて,  $S = \sum W_i (y_i - \hat{y}_i)^2$  を最小とすると, 重みづき最小二乗法と呼ばれる。一般に, 重みを  $W$  と

したとき、正規方程式は、

$$\mathbf{XWX}'\hat{\beta}=\mathbf{XW}\mathbf{y} \quad (1.107)$$

で与えられるが、

$$\hat{\beta}=(\mathbf{XWX}')^{-1}\mathbf{XW}\mathbf{y} \quad (1.108)$$

$$\begin{aligned} E(\hat{\beta}) &= (\mathbf{XWX}')^{-1}\mathbf{XWE}(\mathbf{y}) \\ &= (\mathbf{XWX}')^{-1}\mathbf{XWX}'\beta \\ &= \beta \end{aligned} \quad (1.109)$$

となり、 $\hat{\beta}$  は推定値の期待値が母数と一致する不偏推定量となっている。また、

$$\begin{aligned} \hat{\beta}-\beta &= (\mathbf{XWX}')^{-1}\mathbf{XW}\mathbf{y}-(\mathbf{XWX}')^{-1}\mathbf{XWX}'\beta \\ &= (\mathbf{XWX}')^{-1}\mathbf{XW}(\mathbf{y}-\mathbf{X}'\beta) \end{aligned} \quad (1.110)$$

したがって、 $\hat{\beta}$  の分散・共分散行列は、

$$\mathbf{V}_{\hat{\beta}}=E[(\mathbf{XWX}')^{-1}\mathbf{XW}(\mathbf{y}-\mathbf{X}'\hat{\beta})(\mathbf{y}-\mathbf{X}'\hat{\beta})'\mathbf{W}'\mathbf{X}'((\mathbf{XWX}')^{-1})] \quad (1.111)$$

ここで、 $\mathbf{W}=\mathbf{V}_\varepsilon^{-1}$  として、残差の分散・共分散行列の逆行列を採ると、

$$\begin{aligned} \mathbf{V}_{\hat{\beta}} &= (\mathbf{XV}_\varepsilon^{-1}\mathbf{X}')^{-1}\mathbf{XV}_\varepsilon^{-1}\mathbf{V}_\varepsilon\mathbf{V}_\varepsilon^{-1}\mathbf{X}'(\mathbf{X}'\mathbf{V}_\varepsilon^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{V}_\varepsilon^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{XV}_\varepsilon^{-1}\mathbf{X}')^{-1} \end{aligned} \quad (1.112)$$

さらに、 $\mathbf{V}_\varepsilon=\sigma^2\mathbf{I}$  とすると、

$$\mathbf{V}_{\hat{\beta}}=(\mathbf{XX}')^{-1}\sigma^2 \quad (1.113)$$

となり、重みなしの場合に帰着する。

## 5. 尤 度

NONMEM による解法は、最尤法に相等している。また、モデル式の良さの評価のために尤度比検定あるいは AIC (赤池情報量基準) が使用される。さらに、ベイズ推定においても尤度が重要な役割を演ずる。

ここでは、尤度、対数尤度、最尤法と最大対数尤度、尤度比検定について記す。また、ELS 関数の形が最尤法により導かれることを示し、AIC と最大対数尤度の関係式を記す。

測定値  $x_0$  が得られたとき、 $x_0 \sim N(10, 2^2)$  であるとする、値  $x_0$  が得られる確率は、

$$f(x_0|\mu=10, \sigma^2=2^2)=\frac{1}{\sqrt{2\pi}\cdot 2}\exp(-(x_0-10)^2/(2\cdot 2^2)) \quad (1.114)$$

に比例する。

$x_0$  が別の正規分布  $N(20, 3^2)$  から得られたとすると、

$$f(x_0 | \mu=20, \sigma^2=3^2) = \frac{1}{\sqrt{2\pi} \cdot 3} \exp(-(x_0-20)^2/(2 \cdot 3^2)) \quad (1.115)$$

に比例する。

一般に、測定値  $x_0$  が得られたとき、それが正規分布  $N(\mu_0, \sigma_0^2)$  からの標本であるとする、 $\mu_0, \sigma_0^2$  が与えられたという条件での  $x_0$  の出現確率は、

$$f(x_0 | \mu=\mu_0, \sigma^2=\sigma_0^2) = \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \exp(-(x_0-\mu_0)^2/(2 \cdot \sigma_0^2)) \quad (1.116)$$

に比例し、この値を尤度と呼ぶ（以下では、尤度の式を  $\mu_0, \sigma_0^2$  の関数とみなすので、母数の添字 0 は省略する。）

$x_i (i=1 \sim n)$  が得られたとき、 $x_i$  が互いに独立に  $N(\mu, \sigma^2)$  から得られたとすると、独立な事象が同時に起こる確率は各々の確率の積で与えられるので、尤度  $L$  は、

$$L = \prod_i \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp(-(x_i-\mu)^2/(2 \cdot \sigma^2)) \quad (1.117)$$

で表される。ここで記号  $\prod_i$  は、 $\prod_i$  以下の式中で  $i$  を 1 から  $N$  まで変化させたものをすべてかけ合せることを意味する。

式 (1.117) の両辺の対数をとると、

$$\begin{aligned} l = \ln L &= \sum_i \left( \ln \frac{1}{\sqrt{2\pi} \cdot \sigma} - \frac{(x_i-\mu)^2}{2 \cdot \sigma^2} \right) \\ &= -\frac{n}{2} \ln 2\pi\sigma^2 - \sum_i \frac{(x_i-\mu)^2}{2 \cdot \sigma^2} \end{aligned} \quad (1.118)$$

となり、 $l (= \ln L)$  は対数尤度と呼ばれる。

ここで、 $\mu, \sigma$  を種々に変化させたとき、 $L$  を最大にする（したがって、 $l$  を最大にする）値  $\hat{\mu}, \hat{\sigma}$  を  $\mu, \sigma$  の推定値としてとる方法を最尤法と呼び、このようにして求められた推定値  $\hat{\mu}, \hat{\sigma}$  を最尤推定値と呼ぶ。

式 (1.118) で  $l$  を最大にするためには、 $\partial l / \partial \mu = 0, \partial l / \partial \sigma^2 = 0$  を満たさなければならないので、

$$\frac{\partial l}{\partial \mu} = \frac{1}{\sigma^2} \sum (x_i - \mu) = 0 \quad (1.119)$$

$$\frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum (x_i - \mu)^2 = 0 \quad (1.120)$$

式(1.119)より  $\sum x_i - n\mu = 0$ , したがって,  $\hat{\mu} = \sum x_i / n$ , また, 式(1.120)より  $-n\sigma^2 + \sum (x_i - \mu)^2 = 0$ , したがって,  $\hat{\sigma}^2 = \sum (x_i - \mu)^2 / n$  となり, 通常使用している標本平均, 標本分散は最尤推定値となっていることが導かれる (推定値  $\hat{\mu}$ ,  $\hat{\sigma}^2$  を求めるには, 最尤法以外の方法があることに注意)。

k 群についてそれぞれ r 個体のデータが得られているとき,

		個体			
		1	2	⋯	r
群	1	$x_{11}$	$x_{12}$	$\dots x_{1r}$	
	2	$x_{21}$			
	⋮				
	k	$x_{k1}$	$\dots$		$x_{kr}$

(例—1)

各群の平均値を  $\mu_i (i=1 \sim k)$ , 誤差は各群に共通で  $\sigma^2$ ,  $x_{ij}$  は互いに独立であるとする, 第 i 群の第 j 個体の対数尤度  $l_{ij}$  は,

$$l_{ij} = -\frac{1}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} (x_{ij} - \mu_i)^2 \tag{1.121}$$

したがって,  $l (= \sum_i \sum_j l_{ij})$  は,

$$l = -\frac{kr}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_i \sum_j (x_{ij} - \mu_i)^2 \tag{1.122}$$

となる。ここで, 推定すべき母数は  $\mu_i (i=1 \sim k)$  および  $\sigma^2$  である。l を  $\mu_i$  および  $\sigma^2$  について偏微分すると,

$$\frac{\partial l}{\partial \mu_i} = \frac{1}{\sigma^2} \sum_j (x_{ij} - \mu_i) = 0 \quad (i=1 \sim k) \tag{1.123}$$

$$\frac{\partial l}{\partial \sigma^2} = -\frac{kr}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_i \sum_j (x_{ij} - \mu_i)^2 = 0 \tag{1.124}$$

となる。これより,

$$\hat{\mu}_i = \sum_j x_{ij} / r \tag{1.125}$$

$$\hat{\sigma}^2 = \sum_i \sum_j (x_{ij} - \hat{\mu}_i)^2 / kr \tag{1.126}$$

$$\begin{aligned} &= \sum_i \sum_j \left( x_{ij} - \frac{\sum_j x_{ij}}{r} \right) / kr \\ &= \sum_i \sum_j (x_{ij} - \bar{x}_i) / kr \\ & \quad (\bar{x}_i = \sum_j x_{ij} / r) \end{aligned}$$

が得られる。

最尤推定値  $\hat{\mu}_i = \sum_j x_{ij}/r$ ,  $\hat{\sigma}^2 = \sum_i \sum_j (x_{ij} - \bar{x}_i)/kr$  を対数尤度の式 (1.122) に代入すると最大対数尤度が得られる。

$$l_1 = -\frac{kr}{2} \ln 2\pi \frac{\sum_i \sum_j (x_{ij} - \bar{x}_i)^2}{kr} - \frac{1}{2 \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 / kr} \cdot \sum_i \sum_j \left( x_{ij} - \frac{\sum_j x_{ij}}{r} \right)^2 \quad (1.127)$$

$$= -\frac{kr}{2} \ln(2\pi \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 / kr) - \frac{kr}{2}$$

ここで、群を区別せずに  $kr$  箇のデータがすべて  $N(\mu, \sigma^2)$  から得られたとすると、最大対数尤度は、

$$l_0 = -\frac{kr}{2} \ln(2\pi \sum_i \sum_j (x_{ij} - \bar{x})^2 / kr) - \frac{kr}{2} \quad (1.128)$$

ここで、 $\bar{x} = \sum_i \sum_j x_{ij} / kr$

$l_1, l_0$  を逆対数変換して尤度  $L_1, L_0$  に戻すと次のようになる。

$$L_1 = e^{l_1} \quad (1.129)$$

$$L_0 = e^{l_0} \quad (1.130)$$

尤度は、字が示す通り、もっともらしさを表しており、 $L_1$  は平均値  $\mu_i$  が群ごとに異なる ( $\mu_i \neq \mu_j$ ) と考えた場合 (対立仮説と呼び、 $H_1$  で表す) のもっともらしき、 $L_0$  は群間に差がない ( $\mu_i \equiv \mu_0$ ) と考えた場合 (帰無仮説と呼び、 $H_0$  で表す) のもっともらしさを表している。群間の平均値が異なると考えるか、同一と考えるかを統計的に定める手続きが統計的検定であるが、ここでは、尤度を手がかりに決定することを考える。

対立仮説  $H_1$  の仮定のもとでの最大尤度  $L_1$  と、帰無仮説  $H_0$  のもとでの最大尤度  $L_0$  を比べ、 $L_1$  の方が  $L_0$  に比して大きければ、 $H_1$  をよりもっともらしいとして採用し、 $L_1$  と  $L_0$  の大きさが似ている場合には  $H_0$  (平均値に差がない) を採用することにする。そのために、尤度比  $\lambda = L_0/L_1$  を求め、 $\lambda$  の値が小さい場合には  $H_0$  を棄てて  $H_1$  をとることにする。このような検定の方式は尤度比検定と呼ばれる。

今の例では、 $\lambda = L_0/L_1$  を計算すると、

$$\lambda = C \frac{\sum_i \sum_j (x_{ij} - \bar{x}_i)^2}{\sum_i \sum_j (x_{ij} - \bar{x})^2} = \frac{S_E}{S_T} \quad (C \text{ は定数}) \quad (1.131)$$

となっている。ここで、 $S_E$  は各群内での群平均からの隔たりを集めたものであり、誤差の偏差平方和になっており、 $S_T$  は全体の偏差平方和になっている。

$\lambda$  の分布が既知の場合 (上の例では  $F$  分布に帰着する) には、その分布を用いて検定を行



うことができるが、より複雑な問題では、 $\lambda$ の分布がどのようになるかわからないことが多い。この場合には、 $\lambda$ の代りに  $-2 \ln \lambda = -2(l_0 - l_1)$  が、データが多い場合には近似的に  $\chi^2$  分布に従うことを利用して検定が行われる。その際、自由度は  $H_1$  から  $H_0$  の間で落したパラメータの個数となる。

(例-1) と形の上では似ているが、別の (例-2) を考える。ここでは、 $k$  人の個体について、時刻  $t_i (i=1-r)$  で測定が行われているものとする。

		時刻				
		$t_1$	$t_2$	.....	$t_r$	
個体	1	$x_{11}$	$x_{12}$	.....	$x_{1r}$	(例-2)
	2	$x_{21}$	.....	.....	$\vdots$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	k	$x_{k1}$	.....	.....	$x_{kr}$	

各個体の平均的な水準を  $\eta_i$  とし、測定値  $x_{ij}$  は  $\eta_i$  に誤差  $\epsilon_{ij}$  が付加されたものであると考える。 $\eta_i \sim N(\mu, \omega_\eta^2)$ ,  $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$  と仮定し、 $\eta_i, \epsilon_{ij}$  はすべて独立であると考え、第  $i$  例の第  $t_j$  時点での測定値  $x_{ij}$  の尤度は、

$$L_{ij} = \frac{1}{\sqrt{2\pi} \omega_\eta} \exp\left(-\frac{(\eta_i - \mu)^2}{2\omega_\eta^2}\right) \cdot \frac{1}{\sqrt{2\pi} \sigma_\epsilon} \exp\left(-\frac{(x_{ij} - \eta_i)^2}{2\sigma_\epsilon^2}\right) \tag{1.132}$$

となり、 $x_{ij}$  と  $x_{ij'}$  の間が独立でないために第  $i$  例についての尤度を  $\prod_j L_{ij}$  とすることができない。

これを、(例-1) の場合と比べると、形式的に似たデータであっても、尤度の形が異なり、以降の扱いも異なったものとなる。これは、(例-1) では群平均  $\mu_i$  は固定のもので、確率変数ではないが、(例-2) では、個体の水準  $\eta_i$  は  $N(\mu, \omega_\eta^2)$  に従う確率変数として考えていることによる。(例-1) の場合を母数型 (fixed effect) のモデルと呼び、(例-2) の場合を変量型 (random effect) のモデルと呼ぶ。また、(例-2) のように、同一個体について繰り返し測定されているときは反復測定 (repeated measurements) と呼ばれる。

最尤法でパラメータの推定を行うとき、必ずしも (例-1) のように解析的に推定式を求めることができない。また、解析的に解いて推定する場合には、別の問題では改めて解を求めることが必要になる。このようなことから、対数尤度を直接最大にする方法が考えられる。

対数尤度を最大にするパラメータの値を求めるには、パラメータ  $\theta$  ( $\mu, \sigma^2$  など) の適当な値  $\theta_0$  を初期値として対数尤度を求め、次に、改良されたパラメータ値  $\theta_1$  を対数尤度が増加するように定める。この手続きを次々に繰り返し、 $\theta_0, \theta_1, \theta_2, \dots, \theta_n$  を求め、対数尤度がそれ以上増加しなくなったときの値  $\hat{\theta}$  を推定値とする (詳細は、非線形最適化の項参照)。

モデルが  $y = f(x, \theta) + \epsilon$  で表され、 $\epsilon$  が平均 0、分散・共分散行列  $\Sigma$  の多変量正規分布に従うとき、尤度は、

$$L = \prod_i \frac{1}{(\sqrt{2\pi})^p \sqrt{|\Sigma|}} \exp\left(-\frac{1}{2} (y_i - \bar{y}_i)' \Sigma^{-1} (y_i - \bar{y}_i)\right) \quad (1.133)$$

で表される。両辺の対数をとると、

$$l = C - \sum \left( \frac{1}{2} \ln |\Sigma| - \frac{1}{2} (y_i - \bar{y}_i)' \Sigma^{-1} (y_i - \bar{y}_i) \right) \quad (1.134)$$

C: 定数

となる。最尤法ではこの  $l$  を最大とする  $\theta$  を求めるが、

$$S = \sum (\ln |\Sigma| + (y_i - \bar{y}_i)' \Sigma^{-1} (y_i - \bar{y}_i)) \quad (1.135)$$

として、 $S$  の最小値を求めても同じである。式(1.135)は ELS 関数と呼ばれ、最小二乗法を拡張した形となっている。右辺第2項は重みづき最小二乗法で、重み  $W = \Sigma^{-1}$  とした形となっている。第1項は、 $|\Sigma|$  が大きくなることによって本質的ではないのに残差が小さくなることに対するペナルティの役割を果たしており、 $|\Sigma|$  が大きくなることのみによっては、 $S$  の最小値が求まらなくなっている。

最尤推定量  $\hat{\theta}$  の分布は、 $y$  の密度関数を  $f(y, \theta)$  として、 $y$  が独立であるとするとき  $n$  が大きいとき、

$$\hat{\theta} \sim N\left(\theta, \frac{1}{n} F\right) \quad (1.136)$$

となり、近似的に正規分布に従う。ここで、 $F$  は  $(i, j)$  要素が下記で示される Fisher 情報行列である。

$$F_{ij} = E \left[ \frac{\partial l}{\partial \theta_i} \frac{\partial l}{\partial \theta_j} \right]_{\hat{\theta} = \theta} \quad (1.137)$$

$$l = \ln(f(y, \theta)) \quad (1.138)$$

モデルの良さを量として AIC (赤池情報量規準) がある。

$$AIC = -2 \times (\text{モデルの最大対数尤度}) + 2 \times (\text{モデルの自由パラメータ数}) \quad (1.139)$$

AIC を最小とするモデルを最適のモデルとすることができるが、第2項は、パラメータを増すことにより本質的に良いモデルではないにもかかわらず、第1項が減少することを補っている。モデルの自由パラメータの数は、次のモデルの場合、 $\bar{k}_a$ ,  $\omega_{ka}$ ,  $\bar{V}d$ ,  $\omega_{vd}$ ,  $\bar{C}L$ ,  $\omega_{cl}$ ,  $\sigma_e$  の七つである。

$$C_p = \frac{k_a \cdot D}{Vd(k_a - k_e)} (e^{-k_e t} - e^{-k_a t}) e^e \quad (1.140)$$

$$k_a = \bar{k}_a e^{\eta_{ka}} \quad (1.141)$$

$$Vd = \bar{V}d e^{\eta_{va}} \quad (1.142)$$

$$k_e = CL/Vd \quad (1.143)$$

$$CL = \bar{C}L e^{\eta_{cl}} \quad (1.144)$$

## 6. 非線形最適化

非線形最小二乗法によりパラメータを推定する場合、推定値を解析的な式の形として示すことができず、数値計算により求める。数値計算には種々の方法があるが、傾斜法の考え方と傾斜法のバリエーションとしての最急降下法、ニュートン法、ガウス-ニュートン法、マルカート変法について記す。

薬物動態定数  $\theta$  を、最小二乗法または最尤法によって推定するには、未知母数  $\theta$  の関数  $Ob = F(\theta)$  を最小または最大とする  $\theta$  を推定値としてとればよい。 $Ob = F(\theta)$  は、目的関数と呼ばれ、 $Ob$  を最大とする問題は、 $-Ob$  を最小にする問題と同じである。

最適解  $\hat{\theta}$  が観測値  $x$  の関数として解析的に示されない場合の一般的手順は、仮の解  $\theta_0$  から出発して、目的関数  $Ob$  の値が減少するように  $\theta_1, \theta_2 \dots$  と次々に解を改良する方法である。

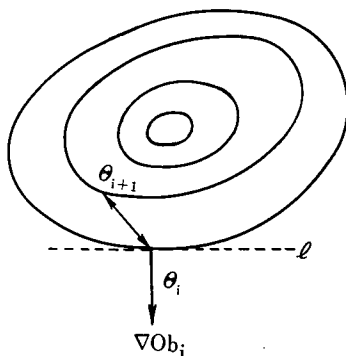
このような逐次的方法には、大別して、直接探索法と  $\theta_i$  の近辺の関数の傾斜を利用して  $\theta_i$  を改良する傾斜法がある。薬物動態定数の推定では、主として傾斜法が用いられる。

$\theta$  が二次元の場合の傾斜法を図1—12に示す。最適解の近くでは、目的関数が二次曲面で近似できると考えられる。また、最適値が存在すれば図1—12のような形状となっている。第  $i$  番目の解が求まったとき、その点を通る等高線の接線を  $l$  とすると、傾きベクトル  $\nabla Ob_i (= \partial Ob / \partial \theta |_{\theta = \theta_i})$  は図に示した方向を向いている。この図から、改良された  $\theta_{i+1}$  は  $\nabla Ob_i$  と反対の方向を向いていると考えられるが、その条件は、

$$\Delta \theta_i \nabla Ob_i < 0 \quad (1.145)$$

で表される。式(1.145)を満たす  $\Delta \theta_i$  は谷向きであるといわれる。

図1—12 傾斜法



谷向きの（目的関数が減少する） $\Delta\theta_i$ の決め方は種々あるが、傾斜法は、この $\Delta\theta_i$ の決め方により区別される。 $\theta$ の改良方向が定めれば、その方向にどれだけ進むかという大きさ $\lambda_i$ を定めて、

$$\theta_{i+1} = \theta_i + \lambda_i \Delta\theta_i \quad (1.146)$$

により、改良値が得られる。

ステップ幅 $\lambda_i$ の求め方としては、 $\Delta\theta_i$ の方向で目的関数 $Ob$ を最小とする $\lambda_i$ を求める場合と、 $Ob_{i+1} < Ob_i$ となる $\lambda_i$ であればよいとする擬最適法があるが、計算量がより少なく、収束性もよいことから、擬最適法が多く用いられる。

$\Delta\theta$ の方向を決めるにあたり、 $\theta$ の空間中で、

$$d(\theta_1, \theta_2) = ((\theta_1 - \theta_2)' A (\theta_1 - \theta_2))^{1/2} \quad (1.147)$$

として、距離 $d$ を定める。ここで、行列 $A$ は $|A| > 0$ で対称な行列とするが、測定の尺度を導入したことに相等し、計量行列と呼ばれる。計量行列 $A$ を導入した場合、 $Ob$ が最も減少する方向（最急降下の方向）は $\nabla Ob$ を $Ob$ の勾配ベクトルとして、

$$\Delta\theta = -A_k^{-1} \nabla Ob_k \quad (1.148)$$

で与えられる。

式(1.148)で $A=I$ （単位行列）とすると、 $\Delta\theta_k$ は接線ベクトル $\nabla Ob_k$ と反対の方向を向いており、通常の意味での最急降下の方向であるので最急降下法と呼ばれる。最急降下法は一般に最も能率のよい方法のように思われるが、探索の方向が制限される等の理由で、局所的にはよい方法であるが、一般的にはよい方法とはされていない。

次に、 $A_k = \text{Hess}_k$ とした場合はニュートンの方法と呼ばれる。ここで、 $\text{Hess}$ はヘッセ行列と呼ばれ、次式で与えられる。

$$\text{Hess} = \left[ \frac{\partial^2 Ob}{\partial \theta_i \partial \theta_j} \right] \quad (1.149)$$

$A_k$ として $\text{Hess}_k$ を用いる考え方は、 $Ob$ を二次の項までテーラー展開し、 $Ob$ を最小にするための条件として $\partial Ob / \partial \theta = 0$ とすることにより導かれる。

$A_k$ としてヘッセ行列と単位行列を適当な比率で混ぜ合せた形 $A_k = \text{Hess}_k + \alpha I$ とした場合は、マルカート(Marquardt)法と呼ばれる。本法は、式の形からわかるように、 $\alpha=0$ の場合はニュートン法となり、 $\alpha$ が大きくなり、ヘッセ行列が事実上無視できるときには最急降下法となる。

上記以外にも種々の最適化の方法がある。

次に、非線形最小二乗法の場合について、ニュートン法がどんな形になるかをみてみよう。

ニュートン法は、前に記したごとく、

$$\Delta\theta = -\mathbf{H}_k^{-1} \nabla \text{Ob}_k \quad (1.150)$$

により改良方向が求められる。最小二乗法は、

$$\mathbf{f} = \mathbf{y} - \mathbf{h}(\mathbf{x}, \theta) \quad (1.151)$$

$$\text{Ob} = \mathbf{f}'\mathbf{f} \quad (1.152)$$

として、Ob を最小とすることであった。ニュートン法のために、傾きベクトル  $\nabla \text{Ob}_k$  を添え字の k を省略して求めると、

$$\nabla \text{Ob} = \begin{pmatrix} \frac{\partial \text{Ob}}{\partial \theta_1} \\ \frac{\partial \text{Ob}}{\partial \theta_2} \\ \vdots \\ \frac{\partial \text{Ob}}{\partial \theta_p} \end{pmatrix} \quad (1.153)$$

の形をしている。

ところで、第 i 例の残差を  $f^i$  とすると、

$$\frac{\partial f^{i2}}{\partial \theta_j} = 2f^i \frac{\partial f^i}{\partial \theta_j} \quad (1.154)$$

であるから、

$$\frac{\partial \text{Ob}}{\partial \theta_j} = 2 \sum_{i=1}^n f^i \frac{\partial f^i}{\partial \theta_j} \quad (1.155)$$

$$= 2 \left( \frac{\partial f^1}{\partial \theta_j} \frac{\partial f^2}{\partial \theta_j} \cdots \frac{\partial f^n}{\partial \theta_j} \right) \begin{pmatrix} f^1 \\ f^2 \\ \vdots \\ f^n \end{pmatrix}$$

したがって、前に記したヤコビ行列  $\mathbf{J}'$  を使って、

$$\nabla \text{Ob} = 2\mathbf{J}\mathbf{f} \quad (1.156)$$

となる。

次に、ヘッセ行列  $\mathbf{H}_k$  を、添え字 k を省略して求めると、第 m, n 要素は、

$$\begin{aligned}\frac{\partial^2 f^i{}^2}{\partial \theta_m \partial \theta_n} &= \frac{\partial}{\partial \theta_m} \left( 2f^i \frac{\partial f^i}{\partial \theta_n} \right) \\ &= 2 \left[ f^i \frac{\partial^2 f^i}{\partial \theta_m \partial \theta_n} + \frac{\partial f^i}{\partial \theta_m} \cdot \frac{\partial f^i}{\partial \theta_n} \right]\end{aligned}\quad (1\cdot157)$$

であるから,

$$H_{ij} = 2 \sum_{i=1}^n \left[ f^i \frac{\partial^2 f^i}{\partial \theta_m \partial \theta_n} + \frac{\partial f^i}{\partial \theta_m} \cdot \frac{\partial f^i}{\partial \theta_n} \right] \quad (1\cdot158)$$

これより,  $f^i$  のヘッセ行列を  $G^i$  とすると,

$$H = 2 \left[ JJ' + \sum_{i=1}^n f^i G^i \right] \quad (1\cdot159)$$

となる。したがってニュートン法は, これらの結果を式(1.148)に代入して,

$$\Delta \theta_k = -2H_k^{-1} J_k f_k \quad (1\cdot160)$$

となり,  $\Delta \theta$  の方向のみを考えるなら,

$$\Delta \theta_k = -H_k^{-1} J_k f_k \quad (1\cdot161)$$

となる。

式(1.159)のヘッセ行列で第一項のみをとって, 方向のみを問題とするために定数2を省くと,

$$H = JJ' \quad (1\cdot162)$$

となる。

この  $H$  を用いた場合をガウス-ニュートン法と呼ぶ。ガウス-ニュートン法では,

$$\Delta \theta_k = -[J_k J_k']^{-1} J_k f_k \quad (1\cdot163)$$

となり, 正規方程式

$$J_k J_k' \Delta \theta_k = -J_k f_k \quad (1\cdot164)$$

の解の形となっている。

ガウス-ニュートン法で,  $J_k J_k'$  の代わりに計量行列として,

$$A = J_k J_k' + \alpha I \quad (1\cdot165)$$

とすると, ガウス-ニュートン法のマルカート変法となる。

ガウス-ニュートン法中出现する項  $J_k J_k'$  は, 最適値が得られたときには前に記したごとく,  $\theta$  の分散・共分散行列の推定に用いられる。

## 7. ベイズ推定

個体ごとの薬物動態パラメータを推定する方法にベイズの方法がある。以下には、ベイズの定理を離散量を例にとって導き、類似の形の連続量での定理の形を記す。次に、ベイズの定理から、経験的ベイズの方法により重みづき最小二乗法の形で個体ごとのパラメータ推定が行われることを示す。

正12面体のサイコロを振って目が偶数となることをA、目が8以下となることをBで表す。Aの起こる（偶数の目が出る）確率を $P(A)$ で、Bの起こる（8以下の目が出る）確率を $P(B)$ で表す。

$$P(A)=1/2 \quad (1\cdot166)$$

$$P(B)=2/3 \quad (1\cdot167)$$

AかつB（偶数で8以下）となる確率を $P(A\cap B)$ で表す。

$$\begin{aligned} P(A\cap B) &= 4/12 & (1\cdot168) \\ &= 1/3 (=P(B\cap A)) \end{aligned}$$

Bが起ったという条件のもとで、Aの起こる確率を $P(A|B)$ で表し、条件付き確率と呼ぶ。Bの場合の数は8通り、その中でA（偶数）となる場合の数は4通りであるから、

$$P(A|B)=4/8=1/2 \quad (1\cdot169)$$

となる。

ところで、 $P(A\cap B)$ は、場合の数をかぞえることなく、Bが起こる確率にBが起ったときにAが起こる確率をかけて求めることができる。

$$\begin{aligned} P(A\cap B) &= P(B) \cdot P(A|B) & (1\cdot170) \\ &= 2/3 \cdot 1/2 = 1/3 \end{aligned}$$

式(1.170)を書きかえると、

$$P(A|B)=P(A\cap B)/P(B) \quad (1\cdot171)$$

となる。

いま、2種類の箱 $A_1$ 、 $A_2$ があり、それぞれの箱の中には異なる比率で赤玉と白玉が入っているとす。2種類の箱を比率 $r$ で準備しておき、まずランダムに箱の一つ引き、次に、その箱からランダムに玉を取り出すとしたとき、赤玉を取り出す確率はどのようになるだろうか。

一方の種類箱を引く確率を $P(A_1)$ 、他方種類箱を引く確率を $P(A_2)$ とする。いずれか種類箱は必ず引くことになるので、 $P(A_1)+P(A_2)=1$ となっている。

$A_1$ の箱を引いたときの赤玉を引く条件付き確率を $P(B|A_1)$ とし、 $A_2$ の箱を引いたときに赤

玉を引く確率を  $P(B|A_2)$  とする。

箱  $A_1$  を引いて、かつ、赤玉となる確率は  $P(A_1) \cdot P(B|A_1)$ 、同様に、箱  $A_2$  では  $P(A_2) \cdot P(B|A_2)$  であるから、いずれの箱を引いたにしても結果的に赤玉が出る確率は、

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) \quad (1.172)$$

となる。一般に、箱の種類が  $k$  種あるときには、

$$P(B) = \sum_i P(A_i) \cdot P(B|A_i) \quad (1.173)$$

で与えられる。

次に、赤玉が得られたときに箱  $A_i$  を引いていた確率を求める。これは  $P(A_i|B)$  で表されるが、式 (1.171) を用いると、

$$\begin{aligned} P(A_i|B) &= P(A_i \cap B) / P(B) \\ &= P(B \cap A_i) / P(B) \\ &= P(B|A_i) \cdot P(A_i) / P(B) \end{aligned} \quad (1.174)$$

ここで式 (1.173) を用いると、

$$P(A_i|B) = P(B|A_i) \cdot P(A_i) / (\sum_j P(A_j) \cdot P(B|A_j)) \quad (1.175)$$

となり、 $P(A_i|B)$  が、 $P(A_i)$  と  $P(B|A_i)$  によって表される。式 (1.175) をベイズの定理と呼ぶ。

連続分布についてのベイズの定理は、 $x \sim N(\mu, \sigma_x^2)$  とし、 $\mu \sim N(\mu, \sigma_\mu^2)$  とするとき、離散量の式 (1.175) と同様な式

$$h(\mu|x) = \frac{f(x|\mu) \cdot g(\mu)}{\int f(x|\mu) \cdot g(\mu) d\mu} \quad (1.176)$$

で表される。ここで、 $f(x|\mu)$  は母数を固定したときの確率密度であるから、すでに出てきた尤度である。 $\mu$  の分布  $g(\mu)$  は事前分布と呼ばれ、 $h(\mu|x)$  は事後分布と呼ばれる。

式 (1.176) は、 $\mu$  に関する情報  $g(\mu)$  が事前に知らされているとき、新しいデータが得られたら、それに対する尤度  $f(x|\mu)$  を用いて  $\mu$  の分布を新たに推定し直すことができることを意味する。

事前分布  $g(\mu)$  としては種々のものがとられるが、経験的ベイズの方法では  $g(\mu)$  をデータより推定する。

いま、母数  $\theta$  の分布が以前の情報により既知であるとし、 $\theta$  の要素  $\theta_i$  が互いに独立に正規分布  $N(\bar{\theta}_i, \omega_{\theta_i}^2)$  に従うとする。また、 $x_i$  はこの母集団から独立にとられたランダムな標本とし、 $x - \hat{x}$  が正規分布  $N(0, \sigma_x^2)$  に従うとすると、

$$\text{尤度} = \prod_i \frac{1}{\sqrt{2\pi} \sigma_x} \exp\left(-\frac{(x_i - \hat{x})^2}{2\sigma_x^2}\right) \quad (1.177)$$



$$\text{事前分布} = \prod_i \frac{1}{\sqrt{2\pi} \sigma_{\theta_i}} \exp\left(-\frac{(\tilde{\theta}_i - \theta_i)^2}{2\omega_{\theta_i}^2}\right) \quad (1.178)$$

となる。ここで、 $\theta_i$  の推定を行うために、一般化最尤法として事後分布密度を最大にする  $\theta_i$  をとることとする。

ベイズの定理の式(1.176)中で、分母は母数  $\theta$  について積分しているので定数であるから、式(1.177)、(1.178)の積を最大にする  $\theta_i$  を求めればよい。この積を  $S$  で表すと、

$$S = \prod_i \frac{1}{\sqrt{2\pi} \sigma_x} \exp\left(-\frac{(x_j - \hat{x}_j)^2}{2\sigma_x^2}\right) \cdot \prod_i \frac{1}{\sqrt{2\pi} \sigma_{\theta_i}} \exp\left(-\frac{(\tilde{\theta}_i - \theta_i)^2}{2\omega_{\theta_i}^2}\right) \quad (1.179)$$

となり、 $s = \ln S$  とし、 $C_0$  を定数とすると、

$$s = C_0 - \left( \sum_j \frac{(x_j - \hat{x}_j)^2}{2\sigma_x^2} + \sum_i \frac{(\tilde{\theta}_i - \theta_i)^2}{2\omega_{\theta_i}^2} \right) \quad (1.180)$$

となり、 $s$  を最大にすればよいが、これは、

$$s' = \sum_j \frac{(x_j - \hat{x}_j)^2}{2\sigma_x^2} + \sum_i \frac{(\tilde{\theta}_i - \theta_i)^2}{2\omega_{\theta_i}^2} \quad (1.181)$$

を最小にすることと同じである。

式(1.181)において、 $x_j$  は測定値、 $\sigma_x^2$  は  $x_j - \hat{x}_j$  の母分散、 $\hat{x}_j$  は  $\theta_i$  の推定値  $\hat{\theta}_i$  での  $x_j$  の推定値、 $\tilde{\theta}_i$  は事前に知られている母集団での  $\theta_i$  の値、 $\omega_{\theta_i}^2$  は  $\theta_i$  の母分散である。

## 8. ELS 法による推定

ここでは、ELS 法による推定、とくに NONMEM を利用してパラメータの推定を行うための一次近似による ELS 法について記す。NONMEM を使用する上で必要な個体間分散、個体内分散、ユーザープログラム (サブルーチン) を記す上で必要な一次近似式の作成法を記す。

$n$  人の個体について、それぞれ  $r_i (i=1 \sim n)$  時点で測定が行われているものとする。測定値はスカラーのこともあるが、一般にはベクトルであるとして、 $y_{ij}$  で表す。

	時点	個体当り 測定回数
個体	1 2 .....	
1		$r_1$
2	$y_{ij}$	$\vdots$
$\vdots$		$\vdots$
$n$		$r_n$

$y_{ij}$  について、一般的な次のモデルを仮定する。

$$y_{ij} = \mathbf{M}(\bar{\theta}, t_{ij}, \eta_i, \varepsilon_{ij}) \quad (1.182)$$

ここで、 $\bar{\theta}$  は全個体共通の母数、 $t_{ij}$  は測定時刻、 $\eta_i$  は個体の効果 ( $\theta$  のうちで、個体差を想定するもの)、 $\varepsilon_{ij}$  は個体内変動である。また、 $y_{ij}$  が  $p$  次元であるとしたとき、 $\mathbf{M}$  は  $p$  次元のベクトル関数、 $\bar{\theta}$  は  $q$  次元 (モデルの必要性に応じる)、 $\eta_i$  は  $q$  次元以下 ( $s$  次元)、 $\varepsilon_{ij}$  は  $p$  次元。

さらに、誤差の分布に関して次の仮定をする。

$$\begin{aligned} \bar{\theta} \text{ と } \eta_i \text{ は独立} & \quad (1.183) \\ \bar{\theta} \text{ と } \varepsilon_{ij} \text{ は独立} & \\ \eta_i \text{ と } \varepsilon_{ij} \text{ は独立} & \\ \varepsilon_{ij} (i \neq j) \text{ は独立} & \\ \mathbf{E}(\eta_i) = \mathbf{0}, \mathbf{V}(\eta_i) = \mathbf{\Omega} & \\ \mathbf{E}(\varepsilon_{ij}) = \mathbf{0}, \mathbf{V}(\varepsilon_{ij}) = \mathbf{\Sigma} & \end{aligned}$$

したがって、

$$\begin{aligned} \mathbf{\Omega} \text{ は } s \times s \text{ の行列} \\ \mathbf{\Sigma} \text{ は } p \times p \text{ の行列} \end{aligned}$$

$\mathbf{M}$  が誤差  $\eta_i$ ,  $\varepsilon_{ij}$  に関して線形なモデルは次式で表される。

$$y_{ij} = \mathbf{F}_{ij}(\bar{\theta}, t_{ij}) + \mathbf{G}_{ij}(\bar{\theta}, t_{ij})\eta_i + \mathbf{H}_{ij}(\bar{\theta}, t_{ij})\varepsilon_{ij} \quad (1.184)$$

ここで、 $\mathbf{F}_{ij}$  は  $p$  次元ベクトル、 $\mathbf{G}_{ij}$  は  $p \times s$  の行列、 $\mathbf{H}_{ij}$  は  $p \times p$  の行列。

$\mathbf{M}$  が  $\eta_i$ ,  $\varepsilon_{ij}$  に関して線形でないとき、簡単なモデルではそのまま ELS 法が適用できるが、複雑なモデルでは適用困難となる。その場合には、 $\mathbf{M}$  を  $\eta_i = \mathbf{0}$ ,  $\varepsilon_{ij} = \mathbf{0}$  の周りでテラー展開して、

$$y_{ij} = \mathbf{M}_{ij}(\bar{\theta}, t_{ij}, \mathbf{0}, \mathbf{0}) + \frac{\partial \mathbf{M}_{ij}(\bar{\theta}, t_{ij}, \mathbf{0}, \mathbf{0})}{\partial \eta_i'} \eta_i + \frac{\partial \mathbf{M}_{ij}(\bar{\theta}, t_{ij}, \mathbf{0}, \mathbf{0})}{\partial \varepsilon_{ij}'} \varepsilon_{ij} \quad (1.185)$$

として、一次近似により推定を行う。ここで、次元は式 (1.184) の場合と同様である。以下では、モデルが  $\eta_i$ ,  $\varepsilon_{ij}$  に関して、線形の場合について考える。

ベクトル  $y_{ij}$  を縦方向に並べたベクトルを  $y_i$  で表す。

$$y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iri} \end{pmatrix} \begin{matrix} ) p \\ ) p \\ ) p \\ ) p \end{matrix} \quad y_i \text{ は } p \cdot r_i \text{ 次元} \quad (1.186)$$

$y_i$  の分散・共分散行列を  $\mathbf{C}_i(p \cdot r_i \times p \cdot r_i)$  とすると、ELS 関数は次式で与えられる。

$$Ob = \sum_i [\log |C_i| + (y_i - \bar{y}_i)' C_i^{-1} (y_i - \bar{y}_i)] \quad (1.187)$$

ここで

$$\bar{y}_i = E(y_i) \quad (1.188)$$

$$C_i = E[(y_i - \bar{y}_i)(y_i - \bar{y}_i)'] \quad (1.189)$$

モデル式(1.184)を式(1.188), (1.189)に代入して  $\bar{y}_i$  と  $C_i$  の形を求めるために次の記号を定める。

$$y'_i = (y'_{i1}, y'_{i2}, \dots, y'_{iri}) \quad (1.190)$$

であったのに対応して,

$$F'_i = (F'_{i1}, F'_{i2}, \dots, F'_{iri}) \quad (p \cdot r_i) \quad (1.191)$$

$$G'_i = (G'_{i1}, G'_{i2}, \dots, G'_{iri}) \quad (s \times p \cdot r_i) \quad (1.192)$$

$$H'_i = \begin{pmatrix} H_{i1} & 0 \\ & H_{i2} \\ 0 & \dots \\ & H_{iri} \end{pmatrix} \quad (p \cdot r_i \times p \cdot r_i) \quad (1.193)$$

$$\epsilon'_i = (\epsilon'_{i1}, \epsilon'_{i2}, \dots, \epsilon'_{iri}) \quad (p \cdot r_i) \quad (1.194)$$

$$\Sigma = E(\epsilon_i \epsilon_i') \quad (p \cdot r_i \times p \cdot r_i) \quad (1.195)$$

とする。そうすると,

$$y_i = F_i + G_i \eta_i + H_i \epsilon_i \quad (1.196)$$

$$\bar{y}_i = E(y_i) = F_i + G_i E(\eta_i) + H_i E(\epsilon_i) = F_i \quad (1.197)$$

$$C_i = E[(y_i - \bar{y}_i)(y_i - \bar{y}_i)'] \quad (1.198)$$

$$\begin{aligned} &= E[(G_i \eta_i + H_i \epsilon_i)(G_i \eta_i + H_i \epsilon_i)'] \\ &= E(G_i \eta_i \eta_i' G_i') + E(G_i \eta_i \epsilon_i' H_i') + E(H_i \epsilon_i \eta_i' G_i') + E(H_i \epsilon_i \epsilon_i' H_i') \\ &= G_i \Omega G_i' + H_i \Sigma H_i' \end{aligned}$$

したがって,  $F_i, G_i, H_i, \Omega, \Sigma, \theta$  の値により  $Ob$  の値が計算でき, ELS の計算が行われることによって  $\hat{\Omega}, \hat{\Sigma}, \hat{\theta}$  が求められる。

例として,  $y_{ij} = \mu + \eta_i + \epsilon_{ij}$  の場合について計算の様子をみると,

$$y_i = \begin{pmatrix} \mu + \eta_i + \epsilon_{i1} \\ \mu + \eta_i + \epsilon_{i2} \\ \vdots \\ \mu + \eta_i + \epsilon_{iri} \end{pmatrix} \quad (1.199)$$

$$y_i = E(y_i) = \mu \quad (1.200)$$

$$\mathbf{G}_i = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \mathbf{r}_i \quad (1 \cdot 201)$$

$$\mathbf{H}_i = \begin{pmatrix} 1 & 0 \\ & 1 \\ & \dots \\ 0 & \dots & 1 \end{pmatrix} \mathbf{r}_i \quad (1 \cdot 202)$$

$$\mathbf{\Omega} = \sigma_\eta^2 \quad (1 \cdot 203)$$

$$\mathbf{\Sigma} = \sigma_\varepsilon^2 \quad (1 \cdot 204)$$

$$\mathbf{C}_i = \mathbf{I} \sigma_\eta^2 \mathbf{I}' + \mathbf{I} \sigma_\varepsilon^2 \mathbf{I}' \quad (1 \cdot 205)$$

$$\begin{aligned} &= \begin{pmatrix} \sigma_\eta^2 & \sigma_\eta^2 \dots \sigma_\eta^2 \\ \vdots & \dots \\ \sigma_\eta^2 & \dots \sigma_\eta^2 \end{pmatrix} + \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \dots \sigma_\varepsilon^2 \end{pmatrix} \\ &= \begin{pmatrix} \sigma_\eta^2 + \sigma_\varepsilon^2 & \sigma_\eta^2 \\ \sigma_\eta^2 & \dots \sigma_\eta^2 + \sigma_\varepsilon^2 \end{pmatrix} \end{aligned}$$

最後に、具体的なモデルについて、線形化の手順をみておく。

**Michaelis-Menten** 型の薬物動態についての投与量と定常状態の濃度の関係式

$$\mathbf{R}_{ij} = V_{mj} C_{PSSij} / (K_{mj} + C_{PSSij}) \quad (1 \cdot 206)$$

について、

$$\mathbf{R}_{ij}^0 = \mathbf{R}_{ij} + \varepsilon_{ij} \quad (1 \cdot 207)$$

$$V_{mj} = \tilde{V}_m \exp \eta_{vm}^i \quad (1 \cdot 208)$$

$$K_{mj} = \tilde{K}_m \exp \eta_{km}^i \quad (1 \cdot 209)$$

$$\mathbf{E}(\varepsilon_{ij}) = 0 \quad (1 \cdot 210)$$

$$\mathbf{V}(\varepsilon_{ij}) = \sigma_\varepsilon^2 \quad (1 \cdot 211)$$

$$\mathbf{E}(\eta_{vm}^i) = 0 \quad (1 \cdot 212)$$

$$\mathbf{V}(\eta_{vm}^i) = \omega_{vm}^2 \quad (1 \cdot 213)$$

$$\mathbf{E}(\eta_{km}^i) = 0 \quad (1 \cdot 214)$$

$$\mathbf{V}(\eta_{km}^i) = \omega_{km}^2 \quad (1 \cdot 215)$$

とすると、

$$R_{ij}^0 = \frac{\tilde{V}_m \exp(\eta_{vm}^i) C_{pssij}}{\tilde{K}_m \exp(\eta_{km}^i) + C_{pssij}} + \varepsilon_{ij} \quad (1 \cdot 216)$$

これを  $\eta_{vm}^i=0$ ,  $\eta_{km}^i=0$  の周りでテーラー展開する。

$$\frac{\partial R_{ij}^0}{\partial \eta_{vm}^i} = \frac{\tilde{V}_m \exp(\eta_{vm}^i) C_{pssij}}{\tilde{K}_m \exp(\eta_{km}^i) + C_{pssij}} \quad (1 \cdot 217)$$

$$\left. \frac{\partial R_{ij}^0}{\partial \eta_{vm}^i} \right|_{\eta_{vm}^i=0, \eta_{km}^i=0} = \frac{\tilde{V}_m C_{pssij}}{\tilde{K}_m + C_{pssij}} \quad (1 \cdot 218)$$

$$\frac{\partial R_{ij}^0}{\partial \eta_{km}^i} = \frac{-\tilde{V}_m \exp(\eta_{vm}^i) C_{pssij}}{(\tilde{K}_m \exp(\eta_{km}^i) + C_{pssij})^2} \quad (1 \cdot 219)$$

$$\left. \frac{\partial R_{ij}^0}{\partial \eta_{km}^i} \right|_{\eta_{vm}^i=0, \eta_{km}^i=0} = \frac{-\tilde{V}_m C_{pssij}}{(\tilde{K}_m + C_{pssij})^2} \quad (1 \cdot 220)$$

したがって、

$$\begin{aligned} R_{ij}^0 &\doteq \tilde{V}_m C_{pssij} / (\tilde{K}_m + C_{pssij}) + \tilde{V}_m C_{pssij} / (\tilde{K}_m + C_{pssij}) \cdot \eta_{vm}^i \\ &\quad - \tilde{V}_m C_{pssij} / (\tilde{K}_m + C_{pssij})^2 \cdot \eta_{km}^i + \varepsilon_{ij} \\ &= \frac{\tilde{V}_m C_{pssij}}{(\tilde{K}_m + C_{pssij})} \left( 1 + \eta_{vm}^i + \frac{\eta_{km}^i}{(\tilde{K}_m + C_{pssij})} \right) + \varepsilon_{ij} \end{aligned} \quad (1 \cdot 221)$$

となり、 $\eta$ ,  $\varepsilon$  に関して線形化される。

#### 参 考 文 献

- 1) L. B. Sheiner, B. Rosenberg, V. V. Marathe : Estimation of Population Characteristics of Pharmacokinetic Parameters from Routine Clinical Data, *J. Pharmacokin. Biopharm.*, 5, 445~479, (1977).
- 2) L. B. Sheiner, S. L. Beal : Evaluation of Methods for Estimating Population Pharmacokinetic Parameters. I. Michaelis-Menten Model ; Routine Clinical Pharmacokinetic Data, *J. Pharmacokin. Biopharm.*, 8, 553~571, (1980).
- 3) L. B. Sheiner, S. L. Beal : Evaluation of Methods for Estimating Population Pharmacokinetic Parameters. II . Biexponential Model and Experimental Pharmacokinetic Data, *J. Pharmacokin. Biopharm.*, 9, 635~651, (1981).
- 4) L. B. Sheiner, S. L. Beal : Evaluation of Methods for Estimating Population Pharmacokinetic Parameters. III. Mono exponential Model ; Routine Clinical Pharmacokinetic Data, *J. Pharmacokin. Biopharm.*, 11, 303~319, (1983)
- 5) L. B. Sheiner : The Population Approach to Pharmacokinetic Data Analysis ; Rationale and Standard Data Analysis Methods, *Drug Metabolism Reviews*, 15, 153~171, (1984).
- 6) S. L. Beal : Population Pharmacokinetic Data and Parameter Estimation Based on Their First Two Statistical Moments, *Drug Metabolism Reviews*, 15, 173~193, (1984).
- 7) M. Rowland, L. B. Sheiner, J. L. Steimer : *Valiability in Drug Therapy*, Raven, New York, 1985.
- 8) M. E. Burton, M. R. Vasko, D. C. Brater : Comparison of Drug Dosing Methods, *Clin. Pharmacol.*

- kin., 10, 1~37, (1985).
- 9) L. B. Sheiner, S. L. Beal : Bayesian Individualization of Pharmacokinetics ; Simple Implementation and Comparison with Non-Bayesian Methods, J. Pharm. Sci., 71, 1344~1348, (1982).
  - 10) C. R. Rao : Linear Statistical Inference and Its Applications, 2nd edition, Wiley, New York, 1973.
  - 11) 佐久間昭 : 薬効評価—計画と解析 II, 東京大学出版会, 1981.
  - 12) Y. Bard : Nonlinear Parameter Estimation, Academic Press, New York, 1974.
  - 13) W. J. Kennedy, Jr., J. E. Gentle : Statistical Computing, Dekker, New York, 1980.
  - 14) S. L. S. Jacoby, J. S. Kowalik, J. T. Pizzo, 関根智明 (訳) : 非線形最適化問題の反復解法, 培風館, 1976.
  - 15) J. O. Berger : Statistical Decision Theory, Springer, New York, 1980.

[魚井 徹, 緒方宏泰]