級分け観測値に対するべき正規分布のあてはめ

濱 崎 俊 光[†], 後 藤 昌 司[‡]

Fitting power normal distribution to
grouped observations

Toshimitsu HAMASAKI and Masashi GOTO

In medical science and biology, individuals are divided into the group in advance by some criteria, and statistical processing are often carried out on grouped observations of the individuals. In this case, the midpoint in the interval into which a group is divided is handled as a representative value, and some corrections such as Sheppard may be used, to calculate some statistics or quantities such as mean and variance. However, this procedure may not be sometime appropriate in practice because the numbers of intervals were often small or the lengths of intervals were not uniform. So, the probability that the observed value belongs to the interval is obtained by supposing some underlying distributions, which are appropriate for describing the phenomenon. Thus, the likelihood can be constructed and the inference about parameters of distribution can be carried out based on the likelihood. In practice, it is desirable that the inference can be carried out on the family of the distributions or data-adaptive distribution which include underlying distributions, if it were possible, normal distribution with historical and statistical "property". In this paper, when underlying distributions are unknown, power normal distribution, which includes normal distribution, was adopted. Further, unstructured data were used in this investigation, as it is caught specifically to fit power normal distribution to grouped observations. Through some examples cited from published literature and simulated data, it was shown that interpretation and consideration could be done by using the property of normal distribution intended after the transformation of observation, by performing statistical processing on transformed scale, and inverse-transforming to original scale, that is, returning to original scale of the power normal distribution.

1. はじめに

医学や生物学の研究では、個体について何らかの特性値が観測される際に、個体の背後にある現象の生起過程に関する制約あるいは観測機構の制約により、特性値がある一つの計量値で得られず、ある2個の観測

点の間にあるという情報に限られることがある。また、何らかの基準によって個体が事前に級分けされ、その級分けされた個体の集合に対して統計的な処理が実施されることがある(Cox, 1957)。例えば、このことは、検査値が悉無的な特性をもち、また経済的な制約からその検査が限られた回数しか実施できない場合に生じる。一方で、膨大な量のデータがある場合に、計算のし易さやグラフィカル表示でのとり扱い易さから、事後的に観測値を級分けすることもある(Cox, 1957; Borkowf *et al.*, 1997; Altman, 1998)。このような場合に級分けされた区間の中点を代表値として扱い、可能であれば Sheppard の補正などを施して平均や分散といった統計量を計算することができる。ただし、実地では、区間の個数が少なかったり、区間の長さが均

[†] ファイザー製薬株式会社 バイオメトリクス部 臨床統計グループ 〒163-0461 新宿区西新宿2-1-1 新宿三井ビル内 私書箱226号

[‡] 連絡先氏名 後藤昌司

郵便宛先 〒650-8531 豊中市待兼山町1-3 大阪大学大学院 基礎工学研究科 情報数理系専攻

Department of Informatics and Mathematical Science, Graduate School of Engineering Science, Osaka University

一でないといったことが往々に生じるために、この手法が適切でないことがある (Heitjan, 1989).

そこで、区間内の値をとる確率を、その現象を記述するのに適切な二三の基礎分布 (例えば、生存時間研究であれば、指数分布、Weibull 分布あるいは対数正規分布) を想定して求め、そのうえで尤度を構成し、諸種の推測の問題に進むことができる。このとき、分布の形状を表すパラメータの推定に焦点があてられる (Gjeddebaek, 1949, 1956, 1957, 1959a, 1959b, 1961, 1968: Kulldorff, 1962: Kempthorne, 1966: Barnard, 1967: Kempthorne & Folks, 1971: Copas, 1972: Giesbrecht & Kempthorne, 1976: Stuart & Ord, 1986: Atkinson *et al.*, 1991). そして、これらの基礎分布の選定には、カイ 2 乗統計量による適合度検定か、AIC 規準などの情報量規準が用いられている (仮谷, 1975, 1979: 仮谷・赤坂, 1976: 杉浦, 1980, 1981: 堂前・宮原, 1984). これらの基礎分布のもとで、このような指標を用いることから、理想 (仮説) と現実との乖離を定量的に表現・評価することができる。ただし、これらの基礎分布に固執する限り、データに適応させた柔軟なあてはめを試みることは不可能に近い。実地では、現実のデータに適応的に規定でき、しかも基礎分布を包括する広い分布族、できればこれまでの「財産」の活用を考慮して正規分布を包含する分布族、のもとで推測の行えることが望ましい (松原・後藤, 1979: 後藤, 1986). 本稿では、基礎分布が未知である場合に、このような要件を比較的満たすベキ正規分布を採用し、観測値の分布をパラメトリックに記述することを試みた。このとき、級分け観測値に対するベキ正規分布のあてはめを基本的かつ明示的にとらえるために、データの構造を規定しない場合を主としてとり扱った。2 節では、ベキ正規分布の定義とその計量値の場合のあてはめを簡潔に紹介する。そのうえで、級分け観測値について、ベキ正規分布を基礎分布として想定した場合の尤度を定義し、級分け観測値に対するベキ正規分布のあてはめ方式を示す。3 節では、数件の文献例を通して、級分け観測値に対してベキ正規分布のあてはめの性能を評価する。4 節では、簡単な数値例でもって、ある特性値が計量値として得られているときに、級分けする前の観測値にベキ正規分布をあてはめた場合とそれらの観測値を級分けしてベキ正規分布をあてはめた場合とを対比する。5 節では、本検討で得られた二三の知見を要約し結びに代える。

2 級分け観測値に対するベキ正規分布のあてはめ

2.1 計量値に対するベキ正規分布のあてはめ

ベキ変換は、変換パラメータ λ を指標として

$$Y_i^{(\lambda)} = \begin{cases} \frac{Y_i^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log Y_i, & \lambda = 0 \end{cases} \quad (2.1.1)$$

と定義される (Box & Cox, 1964). ここに、標本サイズを n として、 $Y_i (i=1, \dots, n)$ はベキ変換前の観測値、 $Y_i^{(\lambda)}$ はベキ変換後の観測値である。このとき、ベキ変換前の観測値 Y_i はベキ正規分布 PND (λ, μ, σ^2) に従うといわれる (Goto *et al.*, 1983: Johnson *et al.*, 1994). そのとき、その対数尤度関数は

$$\begin{aligned} \log L(\lambda, \mu, \sigma^2) &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i^{(\lambda)} - \mu)^2 \\ &\quad + (\lambda - 1) \sum_{i=1}^n \log Y_i - n \log A(k) \end{aligned} \quad (2.1.2)$$

で与えられる。ここに、 μ と σ^2 は、 Y_i がベキ変換後に近似的に正規分布に従うときの平均と分散、 $A(k)$ はベキ正規分布の確率比例定数項

$$A(k) = \begin{cases} \Phi(\text{sgn}(\lambda) | k|), & \lambda \neq 0, \\ 1, & \lambda = 0 \end{cases} \quad (2.1.3)$$

および $k = (\lambda\mu + 1)/(\lambda\sigma)$ である。また、 $\Phi(\cdot)$ は標準正規分布の累積密度関数である。ベキ正規分布の定義、およびその性質については Goto *et al.* (1979), Goto & Inoue (1980), Uesaka & Goto (1982), Goto *et al.* (1983), Goto *et al.* (1984), Johnson *et al.* (1994), 上坂・後藤 (1980), 後藤 他 (1991) に詳しい。 λ, μ, σ^2 の最尤推定値はこの対数尤度関数を λ, μ, σ^2 について最大にすることで得られる。ところで、Box & Cox (1964) では $Y_i^{(\lambda)}$ が平均 μ と分散 σ^2 の正規分布に従うことを仮定しているが、それは $A(k)=1$ とおくことと同等である。 λ を与えたときの μ と σ^2 の最尤推定値はそれぞれ

$$\bar{\mu}(\lambda) = \sum_{i=1}^n \frac{Y_i^{(\lambda)}}{n}, \quad \hat{\sigma}^2(\lambda) = \sum_{i=1}^n \frac{\{Y_i^{(\lambda)} - \bar{\mu}(\lambda)\}^2}{n} \quad (2.1.4)$$

となる。これらの推定値を上記の対数尤度関数に代入すると

$$\log L(\lambda) = -\frac{n}{2} \log \hat{\sigma}^2(\lambda) + (\lambda - 1) \sum_{i=1}^n \log Y_i \quad (2.1.5)$$

となり, これを λ について最大にすることでその最尤推定値 $\hat{\lambda}$ が得られる. さらに, $\hat{\lambda}$ を上記の $\hat{\mu}(\lambda)$ と $\hat{\sigma}^2(\lambda)$ に代入してそれらの最尤推定値 $\hat{\mu}(\hat{\lambda})$ と $\hat{\sigma}^2(\hat{\lambda})$ がそれぞれ求められる.

2.2 ベキ正規分布に基づく級分けデータに対する尤度

観測値が $k(k \geq 3)$ 個の区間 $I_1=[Y_0, Y_1), I_2=[Y_1, Y_2), \dots, I_k=[Y_{k-1}, Y_k)$ に級分けされているとする. ここに, $0=Y_0 < Y_1 < \dots < Y_{k-1} < Y_k=\infty$ である. 区間 $I_i=[Y_{i-1}, Y_i)(i=1, 2, \dots, k)$ にある観測値の個数を n_i , $n=\sum_{i=1}^k n_i$ とする. また, 観測値が k 個の排反な区間 $I_1, I_2, \dots, I_k$ の値をとる確率をそれぞれ $p_1, p_2, \dots, p_k$ とする. ここに, $p=\sum_{i=1}^k p_i=1$ である. 区間 I_i にある観測値の頻度を N_i とすれば, これは多項分布に従う.

$$\Pr(N_1=n_1, N_2=n_2, \dots, N_k=n_k) = \frac{n!}{n_1! n_2! \dots n_k!} \prod_{i=1}^k p_i^{n_i} \quad (2.2.1)$$

通常, p_i の最尤推定値 $\hat{p}_i$ は $\hat{p}_i=n_i/n$ で与えられる. ベキ正規分布に基づく級分け観測値に対する尤度の

構成において, 観測値が区間 I_i の値をとる確率 p_i は, 最初に区間 (Y_{i-1}, Y_i) にベキ変換を施し, 次に基礎分布にベキ正規分布を想定することで求めることができる. このとき, $\theta=(\lambda, \mu, \sigma)^T$ とすれば, ベキ正規分布に基づく確率 $p_{\text{PND}i}^*(\theta)$ は

$$p_{\text{PND}i}^*(\theta) = \frac{\Phi(Z_i) - \Phi(Z_{i-1})}{A(k)} \quad (2.2.2)$$

である. ここに, $Z_i = \{Y_i^{(\lambda)} - \mu\}/\sigma$, $Y_i^{(\lambda)}$ は Y_i のベキ変換後の観測値, $p_{\text{PND}1}^*(\theta) = \Phi(Z_1)/A(k)$ および $p_{\text{PND}k}^*(\theta) = \{1 - \Phi(Z_{k-1})\}/A(k)$ である. これらの関係を表2.1と図2.1に示す. 因に, ベキ正規分布の打ち切り確率を無視した場合, すなわち $A(k)=1$ の場合にベキ正規分布に基づく, 観測値が区間 I_i の値をとる確率 $p_{\text{PND}i}(\theta)$ は

$$p_{\text{PND}i}(\theta) = \Phi(Z_i) - \Phi(Z_{i-1}) \quad (2.2.3)$$

で与えられる. ここに, $p_{\text{PND}1}(\theta) = \Phi(Z_1)$ および $p_{\text{PND}k}(\theta) = 1 - \Phi(Z_{k-1})$ である.

したがって, ベキ正規分布を基礎分布とする, 級分けされた観測値に対する対数尤度関数は

表2.1 級分け観測値に対するベキ正規分布のあてはめ

区 間	区間の上限と下限		頻 度	ベキ正規分布に基づく確率
	ベキ正規分布	近似正規分布		
I_1	$[0, Y_1)$	$[-\infty, Y_1^{(\lambda)})$	n_1	$p_{\text{PND}1}^*(\theta) = \Phi(Z_1)/A(k)$
I_2	$[Y_1, Y_2)$	$[Y_1^{(\lambda)}, Y_2^{(\lambda)})$	n_2	$p_{\text{PND}2}^*(\theta) = \{\Phi(Z_2) - \Phi(Z_1)\}/A(k)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I_i	$[Y_{i-1}, Y_i)$	$[Y_{i-1}^{(\lambda)}, Y_i^{(\lambda)})$	n_i	$p_{\text{PND}i}^*(\theta) = \{\Phi(Z_i) - \Phi(Z_{i-1})\}/A(k)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I_{k-1}	$[Y_{k-2}, Y_{k-1})$	$[Y_{k-2}^{(\lambda)}, Y_{k-1}^{(\lambda)})$	n_{k-1}	$p_{\text{PND}k-1}^*(\theta) = \{\Phi(Z_{k-1}) - \Phi(Z_{k-2})\}/A(k)$
I_k	$[Y_{k-1}, \infty)$	$[Y_{k-1}^{(\lambda)}, \infty)$	n_k	$p_{\text{PND}k}^*(\theta) = \{1 - \Phi(Z_{k-1})\}/A(k)$

図2.1 級分け観測値に対するベキ正規分布のあてはめ

$$\begin{aligned}
\log L(\boldsymbol{\theta}) &= \log n! - \sum_{i=1}^k \log n_i! + \sum_{i=1}^k n_i \log p_{\text{FND}_i}^*(\boldsymbol{\theta}) \\
&= \log n! - \sum_{i=1}^k \log n_i! + \sum_{i=1}^k n_i \log p_{\text{FND}_i}(\boldsymbol{\theta}) - n \log A(k)
\end{aligned}
\tag{2.2.4}$$

で与えられる。

上記の対数尤度関数はベキ変換パラメータ λ について非線形であるために、 λ, μ, σ の最尤推定値 $\hat{\lambda}, \hat{\mu}, \hat{\sigma}$ は観測値を用いて明示的に示すことができない。また、 $A(k)=1$ とおき、 λ を所与の値 λ_0 とした場合でも、 μ と σ の最尤推定値 $\hat{\mu}(\lambda_0)$ と $\hat{\sigma}^2(\lambda_0)$ は (2.1.4) のように明示的に示すことはできない。ただし、 $\hat{\sigma}^2(\lambda_0)$ はある状況¹⁾ においては明示的に示すことができる。このとき、 $\hat{\mu}(\lambda_0)$ と $\hat{\sigma}^2(\lambda_0)$ は漸化式

$$\begin{aligned}
\begin{bmatrix} \hat{\mu}_{j+1}(\lambda_0) \\ \hat{\sigma}_{j+1}(\lambda_0) \end{bmatrix} &= \begin{bmatrix} \hat{\mu}_j(\lambda_0) \\ \hat{\sigma}_j(\lambda_0) \end{bmatrix} \\
&+ \left[\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \mu^2} \quad \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \mu \partial \sigma} \right]^{-1} \begin{bmatrix} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \mu} \\ \frac{\partial \log L(\boldsymbol{\theta})}{\partial \sigma} \end{bmatrix}
\end{aligned}
\tag{2.1.5}$$

で与えられる。これらの導関数は補遺に与える。なお、Hernandez & Johnson (1981) の手順に従えば、一定の条件のもとで標本サイズ n が十分に大きければ、これらの推定値が一致推定量になること、そして推定量の分布が漸近的に正規分布に従うことが示される。これらの結果は補遺に与える。次の 2.3 節で、パラメータの推定のためのアルゴリズムを述べる。

2.3 パラメータの推定のためのアルゴリズム

λ, μ, σ の最尤推定値 $\hat{\lambda}, \hat{\mu}, \hat{\sigma}$ は計量値の場合と同様に、2段階で求められる。すなわち、 λ の推定を先行させ、それを用いて μ と σ を推定する。そして、 μ と σ の推定値を用いて、 λ の推定を更新する。以下に、 λ, μ, σ の最尤推定値 $\hat{\lambda}, \hat{\mu}, \hat{\sigma}$ を求めるためのアルゴリズムを示す。

STEP 1: $\hat{\lambda}_m$ の初期値を $\hat{\lambda}_1$ とする。本稿では、 λ を -2 から $+2$ の間で 0.5 の刻みで動かして、その中で対数尤度関数 (2.2.4) を最大にする λ の値を $\hat{\lambda}_1$ として選定した。

STEP 2: $\hat{\mu}_m$ と $\hat{\sigma}_m$ の初期値として、 $\hat{\mu}_1(\hat{\lambda}_1)$ と $\hat{\sigma}_1(\hat{\lambda}_1)$ を選定する。

STEP 3: $\hat{k}_m$ を関係式

$$k_i = \frac{\lambda_m \mu_i + 1}{\lambda_m \sigma_i}$$

を使用して計算する。そして、 $A(k)$ を計算する。

STEP 4: $\hat{\mu}_{i+1}$ と $\hat{\sigma}_{i+1}$ を

$$\begin{aligned}
\begin{bmatrix} \hat{\mu}_{i+1} \\ \hat{\sigma}_{i+1} \end{bmatrix} &= \begin{bmatrix} \hat{\mu}_i \\ \hat{\sigma}_i \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \mu^2} & \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \mu \partial \sigma} & \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \sigma^2} \end{bmatrix}^{-1} \\
&\quad \begin{bmatrix} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \mu} \\ \frac{\partial \log L(\boldsymbol{\theta})}{\partial \sigma} \end{bmatrix} \bigg|_{\lambda=\hat{\lambda}_m, \mu=\hat{\mu}_i, \sigma=\hat{\sigma}_i, k=\hat{k}_i}
\end{aligned}$$

で推定する。このとき、 $|\hat{\mu}_{i+1} - \hat{\mu}_i| + |\hat{\sigma}_{i+1} - \hat{\sigma}_i| > \delta$ であれば、 i を $i+1$ で置きかえ、STEP 3 に戻る。ここに、 δ は収束を判定するための定数であり、本稿では 0.0001 を採用した。

STEP 5: $\hat{\lambda}_{m+1}$ を、 $\hat{\lambda}_m, \hat{\mu}_{i+1}, \hat{\sigma}_{i+1}, \hat{k}_{i+1}$ を使用し

$$\begin{aligned}
\hat{\lambda}_{m+1} &= \hat{\lambda}_m \\
&+ \left[\frac{\partial \log L(\boldsymbol{\theta})}{\partial \lambda} \middle/ \left(-\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \lambda^2} \right) \right] \bigg|_{\lambda=\hat{\lambda}_m, \mu=\hat{\mu}_{i+1}, \sigma=\hat{\sigma}_{i+1}, k=\hat{k}_{i+1}}
\end{aligned}$$

で推定する。

STEP 6: $|\hat{\lambda}_{m+1} - \hat{\lambda}_m| > \varepsilon$ であれば、 m を $m+1$ で置きかえ、STEP 3 に戻る。ここに、 ε は収束を判定するための定数であり、本稿では 0.0001 を採用した。

STEP 7: $\hat{\lambda}_{m+1}, \hat{\mu}_{i+1}, \hat{\sigma}_{i+1}, \hat{k}_{i+1}$ を使用し、対数尤度を計算する。

STEP 8: 反復を終了し、 $\hat{\lambda}_{m+1}, \hat{\mu}_{m,i+1}, \hat{\sigma}_{m,i+1}$ を $\hat{\lambda}, \hat{\mu}, \hat{\sigma}$ で表す。

上記のアルゴリズムでは、 μ と σ の初期値 $\hat{\mu}_1(\hat{\lambda}_1)$ と $\hat{\sigma}_1(\hat{\lambda}_1)$ の選定が重要になる。初期値の選定には区間の中点の指定や重み付き最小 2 乗法といった方法が考えられる (Lindley, 1950 : Swan, 1969 : Benn & Sidebottom, 1976 : Wolynetz, 1979a, 1979b)。ここでは、後者の重み付き最小 2 乗推定を使用した。 μ と σ の重み付き最小 2 重推定値 $\hat{\mu}$ と $\hat{\sigma}$ は以下の手順で得られる。

¹⁾ λ と μ が所与の値 λ_0 と μ_0 であるとし、 $k=3$ とする。 Y_1 と Y_2 が μ_0 について対称であり、また $0 < n_2 < n$ であれば、 σ の最尤推定値 $\hat{\sigma}(\lambda_0, \mu_0)$ が存在し $\hat{\sigma}(\lambda_0, \mu_0) = (Y_2 - \mu_0) / \Phi^{-1}(1/2 + n_2/2n)$ で与えられる。ここに、 $\Phi^{-1}(\cdot)$ は $\Phi(\cdot)$ の逆関数である。

$$\bar{n}_i = \frac{1}{n} \sum_{j=1}^{i-1} n_j, \quad 2 \leq i \leq k, \quad (2.3.1)$$

$$u_i = \Phi^{-1}(\bar{n}_i) \quad (2.3.2)$$

とする。このとき

$$\text{var}(u_i) \cong \frac{\text{var}(\bar{n}_i)}{\{\phi(u_i)\}^2} \cong \frac{\bar{n}_i(1-\bar{n}_i)}{n\{\phi(u_i)\}^2} \quad (2.3.3)$$

であり、この場合に重み w_i は

$$w_i = \begin{cases} \frac{\{\phi(u_i)\}^2}{\bar{n}_i(1-\bar{n}_i)}, & 0 < \bar{n}_i < 1 \\ 0, & \bar{n}_i = 0, 1 \end{cases} \quad (2.3.4)$$

となる。ここに、 $\phi(\cdot)$ は標準正規分布の確率密度関数である。したがって、 μ と σ の重み付き最小2乗推定値 $\hat{\mu}$ と $\hat{\sigma}$ はそれぞれ

$$\begin{aligned} \hat{\mu} &= \frac{\sum_{i=2}^k w_i \{Y_i^{(\lambda)}\}^2 - \bar{\sigma} \sum_{i=2}^k w_i u_i}{\sum_{i=2}^k w_i}, \\ \hat{\sigma} &= \frac{\sum_{i=2}^k w_i \{Y_i^{(\lambda)}\}^2 - \{\sum_{i=2}^k w_i Y_i^{(\lambda)}\}^2 / \sum_{i=2}^k w_i}{\sum_{i=2}^k w_i u_i Y_i^{(\lambda)} - \sum_{i=2}^k w_i u_i \sum_{i=2}^k w_i Y_i^{(\lambda)} / \sum_{i=2}^k w_i} \end{aligned} \quad (2.3.5)$$

で与えられる。これらの推定値を $\hat{\mu}_1(\hat{\lambda}_1) = \hat{\mu}$ および $\hat{\sigma}_1(\hat{\lambda}_1) = \hat{\sigma}$ とする。

2.4 級分け観測値に対するベキ正規分布の適合度

級分け観測値に対するベキ正規分布のあてはめで、その適合度を測る指標として

$$\text{カイ2乗統計量} \quad \sum_{i=1}^k \frac{(n_i - np_{\text{PND}i}^*(\theta))^2}{np_{\text{PND}i}^*(\theta)},$$

$$\text{修正カイ2乗統計量} \quad \sum_{i=1}^k \frac{(n_i - np_{\text{PND}i}^*(\theta))^2}{n_i},$$

$$\text{Kullback-Leibler 情報量} \quad \sum_{i=1}^k n_i \log \frac{n_i}{np_{\text{PND}i}^*(\theta)},$$

$$\text{AIC情報量規準} \quad 2 \log L(\theta) + 2(p+1)$$

といった種々の指標が考えられる。ここに、AIC情報量規準の p はパラメータ数を表す。本稿では、これらの指標の中からカイ2乗統計量を使用して、級分け観測値に対するベキ正規分布の適合度を測る。表2.2にベキ正規分布の適合を測る観測頻度と期待頻度を示す。

表2.2 ベキ正規分布の適合を測る観測頻度と期待頻度

区間	観測頻度	期待頻度
I_1	n_1	$np_{\text{PND}1}^*(\theta)$
I_2	n_2	$np_{\text{PND}2}^*(\theta)$
$\vdots$	$\vdots$	$\vdots$
I_i	n_i	$np_{\text{PND}i}^*(\theta)$
$\vdots$	$\vdots$	$\vdots$
I_{k-1}	n_{k-1}	$np_{\text{PND}k-1}^*(\theta)$
I_k	n_k	$np_{\text{PND}k}^*(\theta)$

3. 文献例による評価

3.1 データの内容

本稿での狙いを具体的かつ実際に把握し、その意図することの心象を明確にするために、文献例で数値的に評価する。ここで、とり扱う文献例の概要を表3.1に示す。

3.2 結果と考察

表3.1に示した20個の文献例に、ベキ正規分布をあてはめた結果を最大対数尤度 $\log L(\hat{\theta}(\hat{\lambda}))$ 、 λ 、 μ 、 σ の最尤推定値 $\hat{\lambda}$ 、 $\mu(\hat{\lambda})$ 、 $\hat{\sigma}(\hat{\lambda})$ 、確率比例定数項 $A(\hat{k})$ 、およびカイ2乗統計量とその p 値に要約して表3.2に示す。また、文献例で示唆された分布とベキ変換パラメータ λ から示唆される分布を表3.3に示す。その結果、文献例①、文献例②、文献例③、文献例⑦、文献例⑩、文献例⑭、文献例⑮、文献例⑱について、カイ2乗統計量の上でベキ正規分布の適合度が顕著に悪かった。また、文献例で示唆されている分布形状とベキ変換パラメータから示唆される分布形状が一致していたのは、文献例④、文献例⑦、文献例⑨、文献例⑮であった。ここでは、とくに文献例①から文献例③および文献例④から文献例⑦について考察する。

文献例①、文献例②、文献例③の三つの例は、杉浦(1980, 1981)から引用した。杉浦(1980, 1981)では、昭和54年度、昭和55年度、昭和56年度に実施された共通1次試験の得点の級分けされた観測値をとりあげ、これらの三つの観測値の分布が正規分布に従わないことから、これらに種々の分布のあてはめを試みている。そして、これらの三つの観測値がそれぞれ打ち切りをもつ Weibull 分布、打ち切りをもつ Johnson 系の S_B 分布、打ち切りをもつ Weibull 分布に従うと結論づけている。これらの三つの級分けられた観測値

表 3.1 文献例の概要

番号	文 献	データの内容
①	杉浦 (1980, 1981)	昭和 54 年度 (1979 年度) に実施された共通 1 次試験の総合得点のデータ。 打ち切りをもつ weibull 分布 が適合することが示されている。
②	杉浦 (1981)	昭和 55 年度 (1980 年度) に実施された共通 1 次試験の総合得点のデータ。 打ち切りをもつ Johnson 系の S_B 分布 が適合することが示されている。
③	杉浦 (1981)	昭和 56 年度 (1981 年度) に実施された共通 1 次試験の総合得点のデータ。 打ち切りをもつ weibull 分布 が適合することが示されている。
④	堂前・宮原 (1984)	昭和 57 年度 (1982 年度) に埼玉大学に入学した 1,084 人の学生の体重 (kg) のデータ。 対数正規分布 に従うことが示されている。
⑤	堂前・宮原 (1984)	昭和 57 年度 (1982 年度) に埼玉大学に入学した 753 人の男子学生の体重 (kg) のデータ。 対数正規分布 に従うことが示されている。
⑥	堂前・宮原 (1984)	昭和 57 年度 (1982 年度) に埼玉大学に入学した 331 人の女子学生の体重 (kg) のデータ。 対数正規分布 に従うことが示されている。
⑦	堂前・宮原 (1984)	昭和 57 年度 (1982 年度) に埼玉大学に入学した学生の中の、昭和 38 年 (1963 年) 生まれの 26 人の男子学生の体重 (kg) のデータ。 対数正規分布 に従うことが示されている。
⑧	堂前・宮原 (1984)	昭和 57 年度 (1982 年度) に埼玉大学に入学した 753 人の男子学生の比体重 [体重 (kg)/身長 (cm)] のデータ。 対数正規分布 に従うことが示されている。
⑨	堂前・宮原 (1984)	昭和 57 年度 (1982 年度) に埼玉大学に入学した学生の中の、昭和 38 年 (1963 年) 生まれの 206 人の女子学生の比体重 [体重 (kg)/身長 (cm)] のデータ。 対数正規分布 に従うことが示されている。
⑩	堂前・宮原 (1984)	昭和 57 年度 (1982 年度) に埼玉大学に入学した 735 人の男子学生の身長 (cm) のデータ。 正規分布 に従うことが示されている。
⑪	Stuart & Ord (1986)	12 回にわたってサイコロを投げたときの成功 (4, 5, 6 のいずれかの目がでれば成功) 回数のデータ。
⑫	Stuart & Ord (1986)	Francis Bacon (英国の随筆家・哲学者) と Geoffrey Chaucer (英国の詩人) の Macaulay 社から出版された随筆における、文の一節における所与の長さのセンテンスの数のデータ。
⑬	Stuart & Ord (1986)	左後頭骨の特定の長さをもつ古代エジプト人の頭蓋骨の数。
⑭	Brown & Hollander (1977)	270 匹の雄のショウジョウバエ (ミバエ) の生存日数のデータ。
⑮	松原・後藤 (1979)	1996 年 1 月から 1974 年 12 月までの 10 年間に厚生省胃癌研究今永班で集積・観察された 1 次研究申の 430 症例のデータ。ベキ正規分布の上でサハザード回帰モデルをあてはめている。このときのベキ変換パラメータの推定値は -0.4 であり、 対数正規分布 が示唆されている。
⑯	Lindsey (1995)	1973 年の最初の四半期に郵便局に就職したスタッフの数と雇用期間のデータ (グレイド 1)。 対数正規分布 をあてはめている。
⑰	Lindsey (1995)	1973 年の最初の四半期に郵便局に就職したスタッフの数と雇用期間のデータ (グレイド 2)。 対数正規分布 をあてはめている。
⑱	Lindsey (1995)	Saxony における 12 歳の子供をもつ 6,115 組の家族の男性の数のデータ。
⑲	Lindsey (1995)	1851 年から 1962 年の英国でおきた炭鉱事故の間の日数。
⑳	Lindsey (1995)	数学あるいは統計学を専攻した女性数学者の月収データ。

表 3.2 文献例に対するベキ正規分布のあてはめの結果

番号	最大対数尤度	パラメータの推定値				分布の適合度	
	$\log L(\hat{\theta}(\hat{\lambda}))$	$\hat{\lambda}$	$\hat{\mu}(\hat{\lambda})$	$\hat{\sigma}(\hat{\lambda})$	$A(\hat{x})$	χ^2 値	p 値
①	-10037.2798	1.7770	57213.3996	19673.6588	0.9982	17575.187	near 0
②	-2268.8531	1.4798	9308.6022	2772.1309	0.9996	4698.313	near 0
③	-5036.6655	1.5567	13954.4045	4581.3445	0.9988	9700.794	near 0
④	-37.5983	0.0616	4.6462	0.1780	1.0000	7.530	0.9413
⑤	-30.7078	-0.7089	1.3347	0.0060	1.0000	10.624	0.7787
⑥	-22.2458	-0.8488	1.1362	0.0042	1.0000	6.441	0.9713
⑦	-26.4005	0.1890	6.2191	0.2307	1.0000	217.450	near 0
⑧	-29.6761	-0.9608	-1.7075	0.2688	1.0000	8.977	0.8787
⑨	-27.3501	0.0032	1.1758	0.1035	1.0000	12.529	0.6386
⑩	-50.3251	1.9904	143.3797	9.2441	1.0000	34.279	0.0031
⑪	-33.4285	1.0504	6.0127	1.8538	0.9999	14.277	0.2180
⑫	-50.0828	0.0336	3.0805	0.6897	1.0000	25.571	0.3215
⑬	-38.4836	0.6496	28.4843	1.1188	1.0000	18.047	0.3859
⑭	6.3435	1.9421	670.9476	513.9325	0.9043	64.300	near 0
⑮	-42.4831	-0.3017	0.9797	0.4596	1.0000	54.182	near 0
⑯	-8.1260	0.1290	0.8857	1.0873	1.0000	8.098	0.7774
⑰	-5.8255	0.1384	0.8505	1.0360	1.0000	10.906	0.5370
⑱	-55.7693	1.0098	5.7867	1.8649	0.9999	31.901	0.0008
⑲	-34.7647	0.2751	9.1202	4.5544	0.9975	17.738	0.1239
⑳	39.4324	0.1437	14.0965	0.8418	1.0000	36.935	0.1790

表 3.3 文献での示唆とベキ変換パラメータから示唆される分布の対比

番号	文献での示唆	ベキ変換パラメータからの示唆
①	打ち切りをもつ Weibull 分布	J 形状分布
②	打ち切りをもつ Johnson 系の S_B 分布	J 形状分布
③	打ち切りをもつ Weibull 分布	J 形状分布
④	対数正規分布	対数正規分布
⑤	対数正規分布	L 形状分布
⑥	対数正規分布	L 形状分布
⑦	対数正規分布	対数正規分布
⑧	対数正規分布	L 形状分布
⑨	対数正規分布	対数正規分布
⑩	正規分布	J 形状分布
⑪	—	対数正規分布
⑫	—	対数正規分布
⑬	—	極大値と極小値をもつ分布
⑭	—	J 形状分布
⑮	対数正規分布	対数正規分布
⑯	対数正規分布	極大値と極小値をもつ分布
⑰	対数正規分布	極大値と極小値をもつ分布
⑱	—	片側打ち切り正規分布
⑲	—	指数形状分布
⑳	—	極大値と極小値をもつ分布

図 3.1 文献例①についてのベキ変換前の分布とベキ変換後の分布

図 3.2 文献例④についてのベキ変換前の分布とベキ変換後の分布

にベキ正規分布をあてはめた結果、ベキ変換パラメータの値はそれぞれ 1.7770, 1.4798, 1.5567 と推定された。これらのベキ変換パラメータの値から、三つの観測値はいずれも右側に鋭く尖る J 形状分布に従うことが示唆される。1 例として、文献例 ① の場合にベキ変換前の観測値の分布とベキ変換後の観測値の分布を図 3.1 に示す。分布の形状に影響を及ぼすような変化(ここでは、試験制度や試験方法の変更といった状況の変化)がなければ、観測値は同一の分布に従うとするのが自然である。また、これらの三つの文献例について、カイ 2 乗検定でみたベキ正規分布の適合度は顕著に悪かった。これは、(2.1.1) で定義されるベキ変換が非負の上限をもたない観測値に適切であるのに対して、これらの三つの文献例において観測値が試験の得点であるために上限 (1000 点) をもっていたことによると考えられる。このような試験の得点や比率といった上限をもつ観測値への対処法を補遺に示す。

文献例 ④ から文献例 ⑦ の四つの文献例は、堂前・宮原 (1984) から引用した。堂前・宮原 (1984) では、昭和 57 年度に埼玉大学に入学した学生の体重の分布

の 3 乗根が正規分布に従うかどうかを検討している。これは、体重が円錐 (あるいは円柱) の体積で近似できること、および円錐の公式の次元的な類推によると考えられる。しかし、ベキ変換パラメータの推定値は文献例 ④ について 0.0616 と推定され、この値から対数正規分布が示唆される。これは堂前・宮原 (1984) の結果と一致する。宮原・堂前 (1984) では、さらにこの観測値を性や生年で層別して、いずれの場合も対数正規分布に従うと結論づけている。文献例 ④ と文献例 ⑦ について、堂前・宮原 (1984) とベキ変換パラメータの推定値から示唆される分布は、対数正規分布で一致した。文献例 ⑤ と文献例 ⑥ については、ベキ変換パラメータの推定値から対数正規分布よりも左側に鋭く尖る L 形状分布が示唆される。以上の結果から、堂前・宮原 (1984) とベキ変換パラメータの値から示唆される結果がほぼ一致することがわかる。ただし、いずれの場合もベキ変換パラメータが $1/3$ に近い値に推定されることはなかった。すなわち、体重の 3 乗根が正規分布に従うとの仮説は認められなかった。1 例として、文献例 ④ の場合にベキ変換前の観測値の分布と

ベキ変換後の観測値の分布を図3.2に示す。ここでは、級分け観測値に対するベキ正規分布のあてはめに焦点をあてたために、これらの観測値を多元分類された観測値としてとり扱わなかった。通常、このように多元分類された観測値に対しては、ベキ変換パラメータの一意性、変換後の観測値の分散の一意性ならびに位置パラメータにおける要因間の交互作用、要因の主効果、要因間の交互作用の非存在の系列仮説を考慮してベキ正規分布をあてはめることが望ましい (Goto *et al.*, 1979: 上坂・後藤, 1979)

3.3 調整ベキ変換による級分け観測値に対するベキ正規分布のあてはめ

本稿でとり扱ったベキ変換は、一般に観測値の変換に伴う尺度の不変性をもたない。そのために、ベキ変換の前後で尺度に依存する量で統計的な処理を行うことは意味をなさない (Box & Cox, 1982: Hinkley & Runger, 1984: Atkinson, 1985: 後藤 他, 1991: 濱崎・後藤, 1996)。このとき、ベキ変換のベキ変換パラ

メータに依存した尺度効果を取り除くために、変換の Jacobian の n 乗根で事後的に変換公式を調整する「調整ベキ変換」が提案されている (Box & Cox, 1964: Hinkley & Runger, 1984: Duan, 1993: 後藤 他, 1991: 濱崎・後藤, 1996)。尺度に関するこの不変化調整により、観測値の尺度がベキ変換の前後で統一的に扱えるだけでなく、ベキ変換の前後の観測値にパラメトリック・モデルをあてはめる場合に、そのモデルに固有のパラメータとベキ変換パラメータの推定値の相関、およびこれらの推定値の分散が調整を施さない通常のベキ変換 (以降では、単に「通常ベキ変換」と呼ぶ) に比べて著しく縮小される (Bickel & Doksum, 1981: Carroll & Ruppert, 1981: Hinkley & Runger, 1984)。また、調整ベキ変換ではベキ変換パラメータとモデル・パラメータの平均と分散の推定値が通常ベキ変換に比べてベキ変換パラメータの変換に対して安定する (Hinkley & Runger, 1984: 濱崎・後藤, 1996)。

上記のことを考慮して、表3.1の20個の文献例に調整ベキ変換を使用してベキ正規分布をあてはめ直し

表3.4 調整ベキ変換による級分け観測値に対するベキ正規分布のあてはめの結果

番号	最大対数尤度	パラメータの推定値				分布の適合度	
	$\log L(\hat{\theta}(\hat{\lambda}))$	$\hat{\lambda}$	$\hat{\mu}(\hat{\lambda})$	$\hat{\sigma}(\hat{\lambda})$	$A(\hat{\lambda})$	χ^2 値	p 値
①	-5433.4851	1.7864	401.2537	139.5067	0.9980	11979.349	near 0
②	-2195.1688	1.4815	448.5139	134.0159	0.9996	4583.097	near 0
③	-4065.8393	1.5607	406.1122	133.6928	0.9988	9336.609	near 0
④	-37.5983	0.0620	209.3325	8.0232	1.0000	7.529	0.9413
⑤	-30.6738	-0.7093	1512.9414	6.8565	1.0000	10.545	0.7841
⑥	-22.2251	-0.8494	1582.7829	5.8715	1.0000	6.364	0.9730
⑦	-26.4005	0.1890	173.7549	6.4467	1.0000	217.260	near 0
⑧	-29.6761	-0.9606	-0.2305	0.0363	1.0000	8.976	0.8787
⑨	-27.3501	0.0028	3.7519	0.3296	1.0000	12.553	0.6368
⑩	-50.3251	1.9904	8.6110	0.5552	1.0000	34.279	0.0031
⑪	-33.5465	1.0502	5.5291	1.7043	0.9999	14.288	0.2175
⑫	-50.0821	0.0340	136.2812	30.5285	1.0000	25.577	0.3212
⑬	-38.4835	0.6501	143.4035	5.6376	1.0000	18.039	0.3864
⑭	26.6373	2.0003	24.3519	19.9633	0.8934	37.665	near 0
⑮	-40.9912	-0.0360	5.5579	3.4230	1.0000	48.866	near 0
⑯	-3.0699	0.2349	2.1823	3.0920	0.9814	7.128	0.8490
⑰	1.0042	0.2499	2.0664	2.8931	0.9820	10.134	0.6042
⑱	-5.8552	1.0080	5.6967	1.8339	0.9999	31.888	0.0008
⑲	-22.4154	0.2592	349.3812	187.0640	0.9705	12.835	0.3811
⑳	-39.4324	0.1440	10962.9335	655.3238	1.0000	36.933	0.1791

表 3.5 通常ベキ変換と調整ベキ変換の Jacobian の n 乗根の値

番号	通常ベキ変換 J_{BC}^n	調整ベキ変換 J_{NP}^n	番号	通常ベキ変換 J_{BC}^n	調整ベキ変換 J_{NP}^n
①	71.5473	1.0000	⑪	1.1036	0.9999
②	12.6762	1.0000	⑫	0.0519	1.0012
③	12.7196	1.0000	⑬	0.2011	1.0005
④	0.0219	1.0014	⑭	25.2507	0.9979
⑤	0.0009	1.0050	⑮	0.7447	1.0007
⑥	0.0007	1.0125	⑯	0.4760	1.0053
⑦	0.0349	1.0042	⑰	0.4934	1.0036
⑧	6.7888	1.0004	⑱	1.0196	1.0000
⑨	3.2779	1.0038	⑲	0.0368	1.0059
⑩	16.7472	0.9991	⑳	0.00142	1.0125

た. 2.1 節で与えた通常ベキ変換 (2.1.1) に対応する調整ベキ変換は

$$Y_{NPi}^{(\lambda)} = \begin{cases} \frac{Y_i^\lambda - 1}{\lambda \bar{Y}^{\lambda-1}}, & \lambda \neq 0, \\ \bar{Y} \log Y_i, & \lambda = 0 \end{cases} \quad (3.3.1)$$

と定義される. ここに, $\bar{Y}$ は観測値 Y_i の幾何平均である. 20 個の文献例にこの調整ベキ変換を使用してベキ正規分布をあてはめ直した結果を表 3.4 に示す. また, 尺度の不変化調整によって, 変換後の観測値の尺度がどの程度まで安定するか, すなわち λ に依存した尺度効果がどの程度までとり除かれるかを変換の Jacobian の n 乗根の値でもって評価した. このとき, (2.1.1) と (3.1.1) の変換に対応する Jacobian はそれぞれ

$$J_{BC} = \begin{cases} \bar{Y}^{n(\lambda-1)}, & \lambda \neq 0, \\ \bar{Y}^n, & \lambda = 0, \end{cases} \quad (3.3.2)$$

$$J_{NP} = \begin{cases} 1 + \frac{1-\lambda}{n\lambda} \sum_{i=1}^{k-1} (1 - Y_i^\lambda), & \lambda \neq 0, \\ 1 + \log \bar{Y}, & \lambda = 0 \end{cases} \quad (3.3.3)$$

である. なお, 変換後の尺度が安定するとは, 変換の Jacobian の n 乗根の値が 1 になることを意味する. 調整ベキ変換の Jacobian の n 乗根の値を通常ベキ変換の場合と併せて表 3.5 に示す. 表 3.4 から, 文献例 ①を除いて, 最大対数尤度 $\log L(\hat{\theta}(\hat{\lambda}))$ とベキ変換パラメータの推定値 $\hat{\lambda}$ について, 通常ベキ変換と調整ベキ変換の間で顕著な違いは認められなかったが, $\hat{\mu}(\hat{\lambda})$ と $\hat{\sigma}(\hat{\lambda})$ については通常ベキ変換と調整ベキ変換の間

で顕著な違いが認められた. これは変換公式の違いによるためである. 表 3.5 から, 20 個のすべての文献例について, 調整ベキ変換後の Jacobian の n 乗根の値は通常ベキ変換の場合に比べて 1 に近かった. このことから, 調整ベキ変換後の観測値の尺度が通常ベキ変換の場合に比べて安定することが示唆される. ただし, いずれの調整ベキ変換の Jacobian の n 乗根の値が正確に 1 にならないことから, 調整ベキ変換後の尺度において, λ に依存した効果が完全にとり除かれるとはいえない. なお, $J_{NP}^n \approx 1$ であるとしても, 調整ベキ変換後の観測値の絶対的な位置は変換前の観測値の位置と同一でなく, λ に依存することに注意したい (Hinkley & Runger, 1984: 濱崎・後藤, 1996).

4. 簡単な数値例

ここでは, 簡単な数値例を通して, 級分けする前の観測値にベキ正規分布をあてはめた場合とそれらの観測値を級分けしてベキ正規分布をあてはめた場合を対比する. そして, それらの結果から, 観測値の級分けによる推定値への影響の程度を把握する.

はじめに, ベキ正規分布の 6 種の典型的な形状, すなわち J 形状分布 ($\lambda^* > 1$), 片側打ち切り正規分布 ($\lambda^* = 1$), 極大値と極小値をもつ分布 ($\lambda_k < \lambda^* < 1$), 指数形状分布 ($\lambda^* = \lambda_k$), 指数形状分布 ($0 < \lambda^* < \lambda_k$), L 形状分布 ($\lambda^* < 0$) を選定し, 打ち切り点 $k=3$ と変換後の変動係数 $\tau=16$ について, 区間 $[0, 1]$ の一様乱数から Box & Müller 法を使用して正規乱数を発生させた. 次に, この正規乱数に (2.1.1) の逆変換を施してベキ

表 4.1 $n=30$ の場合のベキ正規分布のあてはめの結果

	パラメータ	所与の値	最尤推定値		
			UPND	GPND	
				$k=5$	$k=10$
①	λ	2.0	0.9747	-0.4112	0.1471
	μ	0.0106	0.3919	0.2490	0.3018
	σ	0.1702	0.2488	0.3209	0.3739
②	λ	1.0	0.8900	1.3118	1.1823
	μ	0.0213	0.4346	0.5326	0.5626
	σ	0.3404	0.5326	0.7361	0.7703
③	λ	0.60	0.4156	-0.3451	-0.2158
	μ	0.0355	0.0699	-0.0315	0.0026
	σ	0.5674	0.5324	0.5477	0.5658
④	λ	0.3077	0.3083	-0.1841	0.6790
	μ	0.0691	0.1738	0.1012	0.3252
	σ	1.1064	0.8288	0.6764	0.8668
⑤	λ	0.20	0.1921	0.6371	-0.0416
	μ	0.1064	0.3152	0.5741	0.3028
	σ	1.7021	0.9297	0.9602	0.7694
⑥	λ	-1.0	-0.4138	0.3904	-0.1979
	μ	0.0204	0.0210	0.5195	0.2433
	σ	0.3265	0.6200	0.9716	0.7309

表 4.2 $n=50$ の場合のベキ正規分布のあてはめの結果

	パラメータ	所与の値	最尤推定値		
			UPND	GPND	
				$k=5$	$k=10$
①	λ	2.0	0.9558	0.8478	0.6454
	μ	0.0106	0.4763	0.4563	0.4433
	σ	0.1702	0.2383	0.4491	0.4113
②	λ	1.0	0.5693	0.4989	0.4749
	μ	0.0213	0.1551	0.1496	0.1544
	σ	0.3404	0.4189	0.5929	0.6204
③	λ	0.60	0.4316	0.1362	0.4205
	μ	0.0355	0.2212	0.2068	0.2762
	σ	0.5674	0.8624	0.7198	0.8394
④	λ	0.3077	0.1634	0.7928	0.1673
	μ	0.0691	0.2014	0.5404	0.3002
	σ	1.1064	1.0540	1.1076	0.8574
⑤	λ	0.20	0.1310	0.4066	0.5465
	μ	0.1064	-0.1838	0.0551	0.0637
	σ	1.7021	1.0514	0.7577	0.8479
⑥	λ	-1.0	-0.4810	-1.5761	-0.7562
	μ	0.0204	-0.2198	0.2745	0.0861
	σ	0.3265	0.8067	0.1582	0.4821

正規分布に従う標本サイズ 30 と 50 の標本を生成した。ここに、 $\lambda_k = 4/(k^2 + 4)$ である。この生成標本に通常のベキ正規分布 (UPND) をあてはめた。そして、この生成標本を区間 k が 5 と 10 である場合に、級分けした標本の観測値にベキ正規分布 (GPND) をあてはめた。標本サイズが 30 の場合の結果を表 4.1、標本サイズが 50 の場合の結果を表 4.2 に示す。

これらの結果から、標本サイズが 30 と標本サイズが 50 のいずれの場合にも、UPND での $\hat{\lambda}$, $\hat{\mu}(\hat{\lambda})$, $\hat{\sigma}(\hat{\lambda})$ が GPND でそれらの値に比べて所与の値に近い傾向にあった。GPND について、標本サイズが 30 で区間の数が 5 である場合に、とくに $\lambda > 2$ と $\lambda < -1$ の範囲で $\hat{\lambda}$, $\hat{\mu}(\hat{\lambda})$, $\hat{\sigma}(\hat{\lambda})$ の値と所与の値の間に顕著な違いが認められた。標本サイズが 50 で区間の数が 10 である場合に、 $\hat{\lambda}$, $\hat{\mu}(\hat{\lambda})$, $\hat{\sigma}(\hat{\lambda})$ は所与の値に近づく傾向にあり、GPND と UPND でのそれらの値の間に顕著な違いは認められなかった。

5. 結びに代えて

本稿では、データの構造を規定しない観測値の集合の分布を中心に、級分け観測値にベキ正規分布をあてはめ、観測値の分布をパラメトリックに記述することを試みた。数件の文献例と簡単な数値例による検討の結果は、計量値の場合と同様に、ベキ正規分布のあてはめにおける包括的特徴を示している。Cox & Snell (1981) は、統計的データ解析の過程で、「主局面」としての目標に直結する方法と「副次的局面」としての仮定の充足を点検する方法との分離を嫌い、それらの局面を一貫した方法で捉える姿勢を推奨している。本稿で提示した方法もこの姿勢を意識している。すなわち、変換後に意図している正規分布の豊富な財産を利用して、変換尺度上で統計的な処理を行い、それから逆変換を施して、原尺度上のベキ正規分布に戻して解釈・考察することができる。また、本稿では、級分け観測値に対するベキ正規分布のあてはめを基本的かつ明示的に捉えるために、無構造データの場合を主としてとり扱った。観測の機構に何らかの基準が潜在的あるいは顕在的に設けられ、そのことで線形あるいは非線形のモデルが想定され、しかも級分け観測値しか利用できない場合にも本稿で述べたベキ正規分布に基づく方法を援用することが可能である。この有構造データ、とくに回帰への応用については Hamasaki & Goto (1997) と Hamasaki (1998) を参照されたい。こ

のとき、ベキ変換の目標は、① モデルの加法的あるいは線形性、② 誤差分散の均一性、③ 観測値の分布の正規性を充足することにある。実際のデータ解析でこれらの三つの目標を同時に達成することはほぼ不可能である。そこで、ベキ変換を用いるに際して、観測値の応答が変換後も不変で正規分布に従うとする立場、および観測値の分布は正規分布に従い変換後に近似的に正規分布に従う立場に分けて、諸種の接近を試みるのが望ましい (Goto *et al.*, 1983; Goto *et al.*, 1987; Goto, 1992, 1995; 後藤 他, 1991)

謝 辞

本稿の構成に際し、丁寧な査読を通して貴重なご指摘とご意見を頂戴した審査員と編集委員の諸先生方に心から謝意を表します。

補遺

本稿の構成で関係した詳細な計算結果と関連する話題を以下に要約する。

補遺 ①：パラメータについての対数尤度関数の導関数

$$\begin{aligned} & \frac{\partial \log L(\boldsymbol{\theta})}{\partial \lambda} \\ &= \frac{n}{\sigma^2} \sum_{i=1}^k \frac{\hat{p}_i}{p_{\text{PND}_i}(\boldsymbol{\theta})} \left\{ \phi(Z_i) \frac{\partial Y_i^{(\lambda)}}{\partial \lambda} - \phi(Z_{i-1}) \frac{\partial \lambda_{i-1}^{(\lambda)}}{\partial \lambda} \right\} \\ & \quad - \frac{n}{A(k)} \frac{\partial A(k)}{\partial \lambda}, \\ & \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \lambda^2} \\ &= \frac{n}{\sigma^2} \sum_{i=1}^k \hat{p}_i \left[\frac{\phi(Z_i)}{p_{\text{PND}_i}(\boldsymbol{\theta})} \left\{ \sigma^2 \frac{\partial^2 Y_i^{(\lambda)}}{\partial \lambda^2} - Z_i \left(\frac{\partial Y_i^{(\lambda)}}{\partial \lambda} \right)^2 \right\} \right. \\ & \quad - \frac{\phi(Z_{i-1})}{p_{\text{PND}_i}(\boldsymbol{\theta})} \left\{ \sigma^2 \frac{\partial^2 Y_{i-1}^{(\lambda)}}{\partial \lambda^2} - Z_{i-1} \left(\frac{\partial Y_{i-1}^{(\lambda)}}{\partial \lambda} \right)^2 \right\} \\ & \quad - \left. \left\{ \frac{\phi(Z_i)}{p_{\text{PND}_i}(\boldsymbol{\theta})} \frac{\partial Y_i^{(\lambda)}}{\partial \lambda} - \frac{\phi(Z_{i-1})}{p_{\text{PND}_i}(\boldsymbol{\theta})} \frac{\partial Y_{i-1}^{(\lambda)}}{\partial \lambda} \right\}^2 \right] \\ & \quad - \frac{n}{A(k)} \left\{ \frac{1}{A(k)} \left(\frac{\partial A(k)}{\partial \lambda} \right)^2 - \frac{\partial^2 A(k)}{\partial \lambda^2} \right\}, \\ & \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \lambda \partial \mu} \\ &= \frac{n}{\sigma^2} \sum_{i=1}^k \hat{p}_i \left[\frac{\phi(Z_i) - \phi(Z_{i-1})}{p_{\text{PND}_i}^2(\boldsymbol{\theta})} \left\{ \phi(Z_i) \frac{\partial Y_i^{(\lambda)}}{\partial \lambda} \right. \right. \\ & \quad \left. \left. - \phi(Z_{i-1}) \frac{\partial Y_{i-1}^{(\lambda)}}{\partial \lambda} \right\} \right] \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{p_{\text{PND}i}(\boldsymbol{\theta})} \left\{ Z_i \phi(Z_i) \frac{\partial Y_i^{(\lambda)}}{\partial \lambda} - Z_{i-1} \phi(Z_{i-1}) \frac{\partial Y_{i-1}^{(\lambda)}}{\partial \lambda} \right\} \Bigg] \\
& - \frac{n}{A(k)} \left\{ \frac{1}{A(k)} \frac{\partial A(k)}{\partial \lambda} \frac{\partial A(k)}{\partial \mu} - \frac{\partial^2 A(k)}{\partial \lambda \partial \mu} \right\}, \\
\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \lambda \partial \sigma} &= \frac{n}{\sigma^2} \sum_{i=1}^k \hat{p}_i \left[\frac{1}{p_{\text{PND}i}(\boldsymbol{\theta})} \left\{ (Z_i^2 - 1) \phi(Z_i) \frac{\partial Y_i^{(\lambda)}}{\partial \lambda} \right. \right. \\
& \quad \left. \left. - (Z_{i-1}^2 - 1) \phi(Z_{i-1}) \frac{\partial Y_{i-1}^{(\lambda)}}{\partial \lambda} \right\} \right. \\
& \quad \left. + \frac{1}{p_{\text{PND}i}^2(\boldsymbol{\theta})} \{ Z_i \phi(Z_i) - Z_{i-1} \phi(Z_{i-1}) \} \left\{ \phi(Z_i) \frac{\partial Y_i^{(\lambda)}}{\partial \lambda} \right. \right. \\
& \quad \left. \left. - \phi(Z_{i-1}) \frac{\partial Y_{i-1}^{(\lambda)}}{\partial \lambda} \right\} \right] \\
& - \frac{n}{A(x)} \left\{ \frac{1}{A(x)} \frac{\partial A(k)}{\partial \lambda} \frac{\partial A(k)}{\partial \sigma} - \frac{\partial^2 A(k)}{\partial \lambda \partial \sigma} \right\}, \\
\frac{\partial \log L(\boldsymbol{\theta})}{\partial \mu} &= \frac{n}{\sigma} \sum_{i=1}^k \frac{\hat{p}_i}{p_{\text{PND}i}(\boldsymbol{\theta})} \{ \phi(Z_{i-1}) - \phi(Z_i) \} - \frac{n}{A(k)} \frac{\partial A(k)}{\partial \mu}, \\
\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \mu^2} &= \frac{n}{\sigma^2} \sum_{i=1}^k \hat{p}_i \left[\frac{Z_{i-1} \phi(Z_{i-1}) - Z_i \phi(Z_i)}{p_{\text{PND}i}(\boldsymbol{\theta})} \left\{ \frac{\phi(Z_{i-1}) - \phi(Z_i)}{p_{\text{PND}i}(\boldsymbol{\theta})} \right\}^2 \right. \\
& \quad \left. - \frac{n}{A(k)} \left\{ \frac{1}{A(k)} \left(\frac{\partial A(k)}{\partial \mu} \right)^2 - \frac{\partial^2 A(k)}{\partial \mu^2} \right\} \right], \\
\frac{\partial \log L(\boldsymbol{\theta})}{\partial \sigma} &= \frac{n}{\sigma} \sum_{i=1}^k \frac{\hat{p}_i}{p_{\text{PND}i}(\boldsymbol{\theta})} \{ Z_{i-1} \phi(Z_{i-1}) - Z_i \phi(Z_i) \} \\
& \quad - \frac{n}{A(k)} \frac{\partial A(k)}{\partial \sigma}, \\
\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \sigma^2} &= \frac{n}{\sigma^2} \sum_{i=2}^{k-1} \hat{p}_i \left[\frac{(Z_{i-1}^3 - 2Z_{i-1}) \phi(Z_{i-1}) - (Z_i^3 - 2Z_i) \phi(Z_i)}{p_{\text{PND}i}(\boldsymbol{\theta})} \right. \\
& \quad \left. - \left\{ \frac{Z_{i-1} \phi(Z_{i-1}) - Z_i \phi(Z_i)}{p_{\text{PND}i}(\boldsymbol{\theta})} \right\}^2 \right] \\
& \quad - \frac{n}{A(k)} \left\{ \frac{1}{A(k)} \left(\frac{\partial A(k)}{\partial \sigma} \right)^2 - \frac{\partial^2 A(k)}{\partial \sigma^2} \right\}, \\
\frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \mu \partial \sigma} &= \frac{n}{\sigma^2} \sum_{i=1}^k \hat{p}_i \left[\frac{(Z_{i-1}^2 - 1) \phi(Z_{i-1}) - (Z_i^2 - 1) \phi(Z_i)}{p_{\text{PND}i}(\boldsymbol{\theta})} \right. \\
& \quad \left. - \frac{\{ \phi(Z_{i-1}) - \phi(Z_i) \} \{ Z_{i-1} \phi(Z_{i-1}) - Z_i \phi(Z_i) \}}{p_{\text{PND}i}^2(\boldsymbol{\theta})} \right] \\
& \quad - \frac{n}{A(k)} \left\{ \frac{1}{A(k)} \frac{\partial A(k)}{\partial \mu} \frac{\partial A(k)}{\partial \sigma} - \frac{\partial^2 A(k)}{\partial \mu \partial \sigma} \right\}
\end{aligned}$$

ここに

$$\begin{aligned}
\frac{\partial Y_i^{(\lambda)}}{\partial \lambda} &= \begin{cases} \frac{1}{\lambda} \{ Y_i^{(\lambda)} (\log Y_i) - Y_i^{(\lambda)} \}, & \lambda \neq 0, \\ \frac{1}{2} (\log Y_i)^2, & \lambda = 0, \end{cases} \\
\frac{\partial^2 Y_i^{(\lambda)}}{\partial \lambda^2} &= \begin{cases} \frac{1}{\lambda} \left[Y_i^{(\lambda)} (\log Y_i)^2 - 2 \frac{1}{\lambda} \{ Y_i^{(\lambda)} (\log Y_i) - Y_i^{(\lambda)} \} \right], & \lambda \neq 0, \\ \frac{1}{3} (\log Y_i)^3, & \lambda = 0, \end{cases} \\
\frac{\partial A(k)}{\partial \lambda} &= \begin{cases} -\frac{\phi(k)}{\lambda^2 \sigma}, & \lambda > 0, \\ 0, & \lambda = 0, \\ \frac{\phi(k)}{\lambda^2 \sigma}, & \lambda < 0, \end{cases} \\
\frac{\partial^2 A(k)}{\partial \lambda^2} &= \begin{cases} \frac{2\lambda \sigma - x}{\lambda^4 \sigma^2} \phi(x), & \lambda > 0, \\ 0, & \lambda = 0, \\ -\frac{k + 2\lambda \sigma}{\lambda^4 \sigma^2} \phi(x), & \lambda < 0, \end{cases} \\
\frac{\partial^2 A(k)}{\partial \lambda \partial \mu} &= \begin{cases} \frac{x \phi(k)}{\lambda^2 \sigma^2}, & \lambda > 0, \\ 0, & \lambda = 0, \\ -\frac{k \phi(k)}{\lambda^2 \sigma^2}, & \lambda < 0, \end{cases} \\
\frac{\partial^2 A(k)}{\partial \lambda \partial \sigma} &= \begin{cases} \frac{1 - k^2}{\lambda^2 \sigma^2} \phi(k), & \lambda > 0, \\ 0, & \lambda = 0, \\ -\frac{1 - x^2}{\lambda^2 \sigma^2} \phi(k), & \lambda < 0, \end{cases} \\
\frac{\partial A(k)}{\partial \mu} &= \begin{cases} \frac{\phi(k)}{\sigma}, & \lambda > 0, \\ 0, & \lambda = 0, \\ -\frac{\phi(k)}{\sigma}, & \lambda < 0, \end{cases} \\
\frac{\partial^2 A(x)}{\partial \mu^2} &= \begin{cases} -\frac{x \phi(k)}{\sigma^2}, & \lambda > 0, \\ 0, & \lambda = 0, \\ \frac{x \phi(k)}{\sigma^2}, & \lambda < 0, \end{cases} \\
\frac{\partial A(x)}{\partial \sigma} &= \begin{cases} -\frac{k \phi(k)}{\sigma}, & \lambda > 0, \\ 0, & \lambda = 0, \\ \frac{x \phi(k)}{\sigma}, & \lambda < 0, \end{cases} \\
\frac{\partial^2 A(k)}{\partial \sigma^2} &= \begin{cases} \frac{2 - k^2}{\sigma} k \phi(k), & \lambda > 0, \\ 0, & \lambda = 0, \\ -\frac{2 - k^2}{\sigma} k \phi(k), & \lambda < 0, \end{cases}
\end{aligned}$$

$$\frac{\partial^2 A(x)}{\partial \mu \partial \sigma} = \begin{cases} -\frac{1-k^2}{\sigma^2} \phi(k), & \lambda > 0, \\ 0, & \lambda = 0, \\ \frac{1-x^2}{\sigma^2} \phi(k), & \lambda < 0, \end{cases}$$

である。

補遺 ②：推定値の漸近的性質

λ, μ, σ の最尤推定値の漸近的性質を, Hernandez & Johnson (1981) の手順に従い評価する. Hernandez & Johnson (1981) では, Kullback-Leibler 情報量の最小化基準に基づいて, 離散確率および観測確率のベキ変換に基づくパラメータの推定値の性質が検討されている. ここでは, その手順を利用して, 対数尤度関数の挙動を検討することで, 対数尤度関数の最大化基準のもとで得られる λ, μ, σ の最尤推定値の漸近的性質を示す.

$\theta^T = (\theta_1, \theta_2, \theta_3) = (\lambda, \mu, \sigma)$ とし, パラメータ空間 Θ が $\Theta = \{\theta = (\lambda, \mu, \sigma)^T\}$ で与えられる完備集合である, $H(\theta) = \sum_{i=1}^k p_i \log p_{\text{PND}_i}^*(\theta)/p_i$ が $\theta_0 = (\lambda_0, \mu_0, \sigma_0)^T \in \Theta$ で一義的な最小値をもち連続であることを仮定する. また, θ_0 が Θ の内点である, $\nabla^2 H(\theta_0)$ が正則でないとする. なお, 対数尤度関数と推定値が n 個の観測値に基づくことを強調するときに, それらに添え字を付すことにする.

Stirling の近似式を利用すれば

$$\begin{aligned} & \frac{1}{n} \log n! - \frac{1}{n} \sum_{i=1}^k \log n_i! \\ &= \frac{1-k}{2n} \log 2\pi - \sum_{i=1}^k \hat{p}_{i,n} \log \hat{p}_{i,n} \\ &+ \frac{1}{2} \left\{ (1-k) \log n - \sum_{i=1}^k \log \hat{p}_{i,n} \right\} + O(n^{-1}) \end{aligned} \quad (\text{A2.1})$$

と書ける. ここに, $O(n^{-1})$ は θ において一様である. (2.2.4) から

$$\begin{aligned} & \frac{1}{n} \log L_n(\theta) \\ &= \sum_{i=1}^k \hat{p}_{i,n} \log p_{\text{PND}_i}^*(\theta) - \sum_{i=1}^k \hat{p}_{i,n} \log \hat{p}_{i,n} + o(1) \end{aligned} \quad (\text{A2.2})$$

となる. したがって, 不等式

$$\begin{aligned} & \left| \frac{1}{n} \log L_n(\theta) - \sum_{i=1}^k p_i \log \frac{p_{\text{PND}_i}^*(\theta)}{p_i} \right| \\ & \leq \left| \sum_{i=1}^k \hat{p}_{i,n} \log p_{\text{PND}_i}^*(\theta) - \sum_{i=1}^k p_i \log p_{\text{PND}_i}^*(\theta) \right| \\ & + \left| \sum_{i=1}^k \hat{p}_{i,n} \log \hat{p}_{i,n} - \sum_{i=1}^k p_i \log p_i \right| + o(1) \end{aligned} \quad (\text{A2.3})$$

が成りたつ. このとき, $n \rightarrow \infty$ のときに $\theta \in \Theta$ について

$$\left| \sum_{i=1}^k \hat{p}_{i,n} \log p_{\text{PND}_i}^*(\theta) - \sum_{i=1}^k p_i \log p_{\text{PND}_i}^*(\theta) \right| \xrightarrow{a.s.} 0 \quad (\text{A2.4})$$

であること (Hamasaki, 1998), 関数 $x \log x$ が連続であること, および $\hat{p}_{i,n} \rightarrow p_i$ であることから, 不等式 (A2.3) の右辺は確率 1 でゼロになる. すなわち, $n \rightarrow \infty$ のときに $\theta \in \Theta$ について

$$\frac{1}{n} \log L_n(\theta) \xrightarrow{a.s.} H(\theta) \quad (\text{A2.5})$$

である. したがって, 仮定から $n \rightarrow \infty$ のときに $\hat{\theta}_n \xrightarrow{a.s.} \theta_0$ となる.

次に, $\log L_n(\theta)$ の勾配ベクトル $\nabla \log L_n(\theta)$ と Hessian 行列 $\nabla^2 \log L_n(\theta) = (h_{im}(\theta))$ はそれぞれ

$$\begin{aligned} & \nabla \log L_n(\theta) \\ &= \left(\frac{\partial \log L_n(\theta)}{\partial \theta_1}, \frac{\partial \log L_n(\theta)}{\partial \theta_2}, \frac{\partial \log L_n(\theta)}{\partial \theta_3} \right) \end{aligned} \quad (\text{A2.6})$$

と

$$\begin{aligned} & \nabla^2 \log L_n(\theta) \\ &= \begin{pmatrix} \frac{\partial^2 \log L_n(\theta)}{\partial^2 \theta_1^2}, & \frac{\partial^2 \log L_n(\theta)}{\partial \theta_1 \partial \theta_2}, & \frac{\partial^2 \log L_n(\theta)}{\partial \theta_1 \partial \theta_3} \\ \frac{\partial^2 \log L_n(\theta)}{\partial \theta_2 \partial \theta_1}, & \frac{\partial^2 \log L_n(\theta)}{\partial^2 \theta_2^2}, & \frac{\partial^2 \log L_n(\theta)}{\partial \theta_2 \partial \theta_3} \\ \frac{\partial^2 \log L_n(\theta)}{\partial \theta_3 \partial \theta_1}, & \frac{\partial^2 \log L_n(\theta)}{\partial \theta_3 \partial \theta_2}, & \frac{\partial^2 \log L_n(\theta)}{\partial^2 \theta_3^2} \end{pmatrix} \end{aligned} \quad (\text{A2.7})$$

と与えられる. θ_0 のまわりで $n^{-1/2} \nabla \log L_n(\hat{\theta}_n)$ を Taylor 展開すれば

$$\begin{aligned} & \frac{1}{\sqrt{n}} \nabla \log L_n(\hat{\theta}_n) \\ &= \frac{1}{n} \nabla \log L_n(\theta_0) \\ &+ \frac{1}{n} \nabla^2 \log L_n(\theta_{*n}) \{ \sqrt{n}(\hat{\theta}_n - \theta_0) \} \end{aligned} \quad (\text{A2.8})$$

となる. ここに, $\theta_{*n} = \gamma_n \theta_0 + (1 - \gamma_n) \hat{\theta}_n$ ($0 < \gamma_n < 1$) であ

る. θ_0 が Θ の内点であるとの仮定をもって, $n \rightarrow \infty$ のときに $\hat{\theta}_n \xrightarrow{a.s.} \theta_0$ であることから, 十分に大きい n について $\nabla L_n(\hat{\theta}_n) = 0$ となる. したがって, $n \rightarrow \infty$ のときに

$$\frac{1}{\sqrt{n}} \nabla \log L_n(\theta_0) + \frac{1}{n} \nabla^2 \log L_n(\theta_{*n}) \{ \sqrt{n}(\hat{\theta}_n - \theta_0) \} \xrightarrow{a.s.} 0 \quad (A2.9)$$

となり, このことから $n^{-1/2} \nabla \log L_n(\theta_0)$ と $-n^{-1} \nabla^2 \log L_n(\theta_{*n}) \{ \sqrt{n}(\hat{\theta}_n - \theta_0) \}$ が同一の極限分布に従うことがわかる.

ここで, D_i は $Y_i \in I_i$ であれば $D_i(Y_i) = 1$, $Y_i \notin I_i$ であれば $D_i(Y_i) = 0$ をとる指標関数であるとし, $\alpha_i(\theta_0) = (\alpha_{i1}(\theta_0), \alpha_{i2}(\theta_0), \alpha_{i3}(\theta_0))^T$ はその要素が

$$\alpha_{ir}(\theta_0) = \left. \frac{\partial \log p_{\text{NDi}}^*(\theta)}{\partial \theta_r} \right|_{\theta = \theta_0}, \quad r = 1, 2, 3$$

で与えられるとする. また, 確率変数ベクトル $\mathbf{Z}_j(\theta_0)$ が, 互いに独立に同一の分布に従い, $E[\mathbf{Z}_i(\theta_0)] = \nabla H(\theta_0) = 0$ および $E[\mathbf{Z}_i(\theta_0) \mathbf{Z}_i^T(\theta_0)] = \mathbf{W}$ とする. ここに, $\mathbf{W}$ はその要素が

$$W_{kl} = \sum_{i=1}^k p_i \alpha_{il}(\theta_0) \alpha_{im}(\theta_0), \quad l, m = 1, 2, 3$$

で与えられる. このとき, $n^{-1} \nabla \log L_n(\theta_0)$ は

$$\frac{1}{\sqrt{n}} \nabla \log L_n(\theta_0) = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^k D_{ij}(Y_j) \alpha_i(\theta_0) = \frac{1}{n} \sum_{j=1}^k \mathbf{Z}_j(\theta_0) \quad (A2.10)$$

と書ける. したがって, 多変量中心極限定理を適用すれば, $n \rightarrow \infty$ のときに

$$\frac{1}{\sqrt{n}} \nabla \log L_n(\theta_0) \xrightarrow{d} N_3(0, \mathbf{W}) \quad (A2.11)$$

となる. $n \rightarrow \infty$ のときに $\hat{p}_{i,n} \xrightarrow{a.s.} p_i$ であることから

$$\begin{aligned} \frac{1}{n} h_{lm,n}(\theta) &\xrightarrow{a.s.} \frac{\partial^2}{\partial \theta_l \partial \theta_m} \left(\sum_{i=1}^k p_i \log \frac{p_{\text{NDi}}^*(\theta)}{p_i} \right) \\ &= \frac{\partial^2 H(\theta)}{\partial \theta_l \partial \theta_m} \end{aligned} \quad (A2.12)$$

となり, $\theta \in \Theta$ において一様に, $n \rightarrow \infty$ のときに確率1で

$$\frac{1}{n} h_{lm,n}(\theta_{*n}) \xrightarrow{p} \frac{\partial^2 H(\theta)}{\partial \theta_l \partial \theta_m}$$

である. このとき, $V = \{\nabla^2 H(\theta_0)\}^{-1}$ とおき, Slutsky の定理を適用すれば, $n \rightarrow \infty$ のときに

$$\sqrt{n}(\theta_n - \theta_0) \xrightarrow{a.s.} N_3(0, V W V^T) \quad (A2.13)$$

を得る.

これらの結果から, 一定の条件のもとで標本サイズ n が十分に大きければ, λ, μ, σ の最尤推定量が一致推定量になること, そして推定量の分布が漸近的に正規分布に従うことが示される (Hamasaki, 1998).

補遺 ③: 非対称ベキ変換による接近法

文献例①, 文献例②, 文献例③の三つの文献例について, 級分け観測値に対するベキ正規分布の適合度は良好であるとはいえなかった. これは, (2.1.1) で定義されるベキ変換が上限をもたない非負の観測値に対して適切であるのに対して, 文献例①から文献例③では, 観測値が試験の得点であるために上限 (1000 点) をもっていたことによると考えられる. 上限をもつ観測値に対する変換の代表として対称ベキ変換, 非対称ベキ変換, 非対称オッズ比変換が考えられる (Aranda-Ordaz, 1981: Guerrero & Johnson, 1982: Goto *et al.*, 1986: Atkinson, 1985: 河合・後藤, 1997). これら三つの変換の中で, 定義域の広さと解釈の容易さから非対称ベキ変換と非対称オッズ比ベキ変換が推奨されている (Goto *et al.*, 1986: Atkinson, 1985). ここでは, 非対称オッズ比ベキ変換の類推から, 次の変換を使用した.

表 A3.1 非対称ベキ変換による級分け観測値に対するベキ正規分布のあてはめの結果

番号	最大対数尤度	パラメータの推定値				分布の適合度	
	$\log L(\hat{\theta}(\hat{\lambda}))$	$\hat{\lambda}$	$\hat{\mu}(\hat{\lambda})$	$\hat{\sigma}(\hat{\lambda})$	$A(\hat{k})$	χ^2 値	p 値
①	-150.1864	-0.0689	0.5867	0.6044	1.0000	155.888	near 0
②	-255.3088	-0.0125	0.5141	0.5737	1.0000	374.572	near 0
③	-110.5031	-0.1103	0.4488	0.6004	1.0000	74.674	near 0

$$Y_{\text{AOPT}i}^{(\lambda)} = \begin{cases} \frac{1}{\lambda} \left\{ \left(\frac{Y_i}{1000 - Y_i} \right)^\lambda - 1 \right\}, & \lambda \neq 0, \\ \log \left(\frac{Y_i}{1000 - Y_i} \right), & \lambda = 0. \end{cases} \quad (\text{A3.1})$$

上記の変換を使用して文献例①、文献例②、文献例③の三つの文献例にベキ正規分布をあてはめ直した結果を表A3.1に示す。表A3.1から、この変換を使用してベキ正規分布をあてはめることで、通常ベキ変換の場合に比べて分布の適合度が顕著に改善されていることがわかる。

補遺④：離散観測値に対するベキ正規分布のあてはめ

本稿でとり扱った内容は、直接的に離散分布に応用できる。 Y が $i \geq 1$ に対して確率 $p_i = \Pr\{Y=i\}$ で値 i をとる離散確率変数であるとする。このとき、確率 p_i はベキ正規分布に基づく確率 $p_{\text{PND}i}^*(\theta)$ で近似でき、 $p_{\text{PND}i}^*(\theta)$ は

$$p_{\text{PND}i}^*(\theta) = \frac{1}{A(x)} \left\{ \phi \left(\frac{(i+\Delta)^{(\lambda)} - \mu}{\sigma} \right) - \phi \left(\frac{(i-\Delta)^{(\lambda)} - \mu}{\sigma} \right) \right\} \quad (\text{A4.1})$$

で与えられる。ここに、 Δ は連続補正のための定数であり、 $0 < \Delta < 1$ の範囲をとる。 Δ の値として、 $\Delta = 0.5$ を使用することが多い。また、 n 回の試行において、 $i > n$ のときに $p_i = 0$ であれば

$$p_{\text{PND}0}^*(\theta) = \frac{1}{A(x)} \phi \left(\frac{(1-\Delta)^{(\lambda)} - \mu}{\sigma} \right),$$

$$p_{\text{PND}n}^*(\theta) = \frac{1}{A(x)} \left\{ 1 - \phi \left(\frac{(n-1+\Delta)^{(\lambda)} - \mu}{\sigma} \right) \right\} \quad (\text{A4.2})$$

となる。

補遺⑤：関連する話題

観測値の範囲が未知のパラメータに依存するときに、尤度関数が(局所)最大値をもたないことがある。このとき、計量値で与えられた観測値を擬似的に級分けして尤度を構成するか、あるいは観測値の差の積を利用することで、この問題を回避することが試みられている(Kulldorff, 1962; Kempthorne, 1966; Barnard, 1967; Kempthorne & Folks, 1971; Copas, 1972; Giesbrecht & Kempthorne, 1976; Cheng & Amin, 1983; Smith, 1985; Titterton, 1985; Cheng & Iles, 1987; Atkinson *et al.*, 1991; Cheng & Traylor, 1995)。ここでは、前者を擬似的な級分け観測値に基づく接近、後者をMPS (Maximum Product of Spacing) 法と呼ぶ。以下に、これらの二つの方法を簡単に紹介する。MPS法は擬似的な級分け観測値に基づく尤度の形式で解釈することができ、漸近的にこの二つの接近法は等しい(Titterton, 1985; Cheng & Amin, 1983; Atkinson *et al.*, 1991)。

擬似的な級分け観測値に基づく接近： Y_i は確率密度関数 $f(y, \theta)$ をもつ分布から得られるサイズ n の標本であり、 Y_i のすべての観測値が同一の精確さ(幅) Δ をもって観測されるとする。 Y_i を観測することは観測されない計量値が区間 $Y_i \pm \Delta$ にあることを意味する。 Y_i がこの区間の値をとる確率 p_i は

$$p_i = \int_{y_i - \Delta}^{y_i + \Delta} f(y_i, \theta) dy \quad (\text{A5.1})$$

で与えられる。したがって、 n 個の観測値に基づく対数尤度関数は

$$\log L(\theta) = \sum_{i=1}^n \log p_i - n \log(2\Delta) \quad (\text{A5.2})$$

となる。右辺の第2項は級分け区間のサイズについての近似修正である(Atkinson *et al.*, 1991)。したがって、 θ の推定値は上記の対数尤度関数を最大にすることで得られる。

MPS法： Y_i は確率密度関数 $f(y, \theta)$ をもつ分布から得られるサイズ n の標本である。このとき、 θ は分布のパラメータであり、 Y_i は $Y_1 < Y_2 < \dots < Y_i < \dots < Y_n$ に順序付けられているとする。 Y_i が単位区間 $(0, 1)$ に含まれるように、変換 $Z_i = F(y_i, \theta)$ ($i=0, 1, \dots, n+1$)を施す。ここに、 $F(y_i, \theta)$ は Y_i の累積密度関数、 $x_1 = a_1$ および $x_{n+1} = a_2$ である。また、 a_1 と a_2 は未知のパラメータであり、 θ に含まれる。このとき、 θ の推定値は

$$G = \left\{ \prod_{i=1}^{n+1} D_i \right\}^{1/n} \text{ あるいは } \log G = \frac{1}{n} \sum_{i=1}^{n+1} D_i \quad (\text{A5.3})$$

を最大にすることで得られる(Cheng & Amin, 1983; Cheng & Iles, 1987; Cheng & Traylor, 1995)。ここに

$$D_i = y_i - y_{i-1} = \int_{y_{i-1}}^{y_i} f(y, \theta) dy, \quad (i=1, 2, \dots, n+1) \quad (\text{A5.4})$$

である。このとき、明らかに

$$\sum_{i=1}^{n+1} D_i = 1 \quad (\text{A5.5})$$

である。

参考文献

- Altman, D.G. (1998). Categorizing continuous variables. *Encyclopedia of Biostatistics*, ed. by P. Armitage & T. Colton, Vol. 1, 563-567, John Wiley & Sons.
- Aranda-Ordaz, F.J. (1981). On two families of transformations to additivity for binary response data. *Biometrika*, **68**(2), 357-364.
- Atkinson, A.C. (1985). *Plots, Transformations, and Regression*. Oxford University Press [後藤昌司・濱崎俊光 (1998). 変換・回帰・グラフィクス, MPC (印刷中)].
- Atkinson, A.C., Perocchi, L.R. & Smith, R.L. (1991). Grouped likelihood for shifted power transformation. *J. Roy. Statist. Soc.*, **B53**, 473-482.
- Barnard, G.A. (1967). The use of the likelihood function in statistical practice. *5th Berkley Symp. Mathematical Statistics & Probability*, ed. by L.M. LeCan & J. Neyman, **1**, 27-40.
- Benn, R.T. & Sidebottom, S. (1976). Algorithm AS 95: Maximum-likelihood estimation of location and scale parameters from grouped data. *Appl. Statist.*, **25**, 88-93.
- Borkowf, C.B., Gail, M.H., Carroll, R.J. & Gill, R.D. (1997). Analyzing bivariate continuous data grouped into categories defined by empirical quantiles of marginal distributions. *Biometrics*, **53**, 1054-1069.
- Brown, B.W. & Hollander, M. Jr. (1977). *Statistics: A Biomedical Introduction*. John Wiley & Sons [永井武昭・後藤昌司・田崎武信・勘場貢・細野泰彦 (1995). 医学統計解析入門. 改定版, MPC].
- Box, G.P.E. & Cox, D.R. (1964). An analysis of transformations (with discussion). *J. Roy. Statist. Soc.*, **B26**, 211-252.
- Box, G.P.E. & Cox, D.R. (1982). An analysis of transformations revisited, rebutted. *J. Amer. Statist. Assoc.*, **77**, 209-210.
- Carroll, R.J. & Ruppert, D. (1981). On prediction and the power-transformation family. *Biometrika*, **68**, 609-615.
- Cheng, R.C.H. & Amin, N.A.K. (1983). Estimating parameters in continuous univariate distributions with shifted origin. *J. Roy. Statist. Soc.*, **B45**, 394-403.
- Cheng, R.C.H. & Iles, T.C. (1987). Corrected maximum likelihood in non-regular problems. *J. Roy. Statist. Soc.*, **B49**, 95-101.
- Cheng, R.C.H. & Traylor, L. (1995). Non-regular maximum likelihood problems. *J. Ryo. Statist. Soc.*, **B57**, 3-44.
- Copas, J.B. (1972). The likelihood surface in the linear functional relationship. *J. Roy. Statist. Soc.*, **B34**, 274-278.
- Cox, D.R. (1957). Note on grouping. *J. Amer. Statist. Assoc.*, **52**, 543-547.
- Cox, D.R. & Snell, E.J. (1981). *Applied Statistics: Principles and Examples*. Chapman & Hall [後藤昌司・土屋住英 (1985). 応用統計実践教本, MPC].
- Duan, N. (1993). Sensitivity analysis of Box-Cox transformation model: contrast parameters. *Biometrika*, **80**, 885-897.
- Giesbrecht, F. & Kempthorne, O. (1976). Maximum likelihood estimation in three-parameter log-normal distribution. *J. Roy. Statist. Soc.*, **B38**, 257-264.
- Gjeddebaek, N.F. (1949). Contributions to the study of grouped observations: application of the method of maximum likelihood in case of normally distributed observations. *Skandinavisk Aktuarietidskrift*, **32**, 135-159.
- Gjeddebaek, N.F. (1956). Contributions to the study of grouped observations II: loss of information caused by grouping if normally distributed observations. *Skandinavisk Aktuarietidskrift*, **39**, 154-159.
- Gjeddebaek, N.F. (1957). Contributions to the study of grouped observations III: the distribution of estimates of mean. *Skandinavisk Aktuarietidskrift*, **40**, 20-25.
- Gjeddebaek, N.F. (1959a). Contributions to the study of grouped observations IV: some comments on simple estimates. *Biometrics*, **15**, 433-439.
- Gjeddebaek, N.F. (1959b). Contributions to the study of grouped observations V: three-class grouping of normal distributions. *Skandinavisk Aktuarietidskrift*, **42**, 194-207.
- Gjeddebaek, N.F. (1961). Contributions to the study of grouped observations VI. *Skandinavisk Aktuarietidskrift*, **44**, 55-73.
- Gjeddebaek, N.F. (1968). Statistical analysis: grouped observations. *International Encyclopedia of the Social Science*, ed. by D.R. Sills, Macmillan & The Free Press.
- Giesbrecht, F. & Kempthorne, O. (1976). Maximum likelihood estimation in three-parameter log-normal distribution. *J. Roy. Statist. Soc.*, **B38**, 257-264.
- Goto, M. (1992). Extensive views of power trans-

- formations: some recent developments. *Invited paper of Honolulu Conference on Computational Statistics as memorial of 5th anniversary JSCS, JAMS, December, 1-5.*
- Goto, M. (1995). Double power transformations and their applications. *Invited paper of International Conference on Statistical Computing for Quality and Productivity Improvement, Seoul, August, 17-19.*
- Goto, M. & Inoue, T. (1980). Some properties of power-normal distribution. *Bull. Biometric. Soc. Japan*, **1**, 28-54.
- Goto, M., Inoue, T. & Tsuchiya, Y. (1984). On estimation of parameters in power-normal distribution. *Bull. Inform. Cyber.*, **21**(1-2), 41-53.
- Goto, M., Inoue, T. & Tsuchiya, Y. (1986). Power-transformation for ordered categorical data. *Behaviormetrika*, **19**, 17-39.
- Goto, M., Inoue, T. & Tsuchiya, Y. (1987). Double power transformation and its performance: an extensive version of Box-Cox transformation. *J. Japan. Statist. Soc.*, **17**(2), 149-163.
- Goto, M., Matsubara, Y. & Tsuchiya, Y. (1983). Power-normal distribution and its applications. *Rep. Stat. Appl. Res., JUSE.*, **30**(3), 8-28.
- Goto, M., Uesaka, H. Inoue, T. (1979). Some linear models for power transformed data. *Invited Presentation at 10th International Biometric Conference at Brazil, August, 6-10 (Res. Rep. NO.93, Res. Inst. Fond. Inform. Sc., Kyusyu University).*
- Guerrero, V.M. & Johnson, R.A. (1982). Use of the Box-Cox transformation with binary response models. *Biometrika*, **69**(2), 309-314.
- Haitovsky, Y. (1973). *Regression Estimation from Grouped Observations*. Griffin.
- Hamasaki, T. (1998). Inference from grouped observations based on bivariate power-normal distribution. *Submitted*.
- Hamasaki, T. & Goto, M. (1997). Application of power-normal distribution to grouped observations. *Proceedings of the Ninth Korea and Japan Joint Conference of Statistics: Multivariate Analysis and Computing*, 67-72, Cheju, Korea, December 5-6.
- Heitjan, D.F. (1989). Inference from grouped continuous data: a review. *Statistical Science*, **4**(2), 164-183.
- Hernandez, F. & Johnson, R.A. (1981). Transformation of a discrete distribution to near normality. *Statistical Distributions in Scientific Work*, **5**, 259-270.
- Hinkley, D.V. & Runger, G. (1984). The analysis of transformed data. *J. Amer. Statist. Assoc.*, **79**, 302-320.
- Johnson, N.L., Kotz, S. & Balakrishnan, N. (1994). *Continuous Univariate Distributions*. John Wiley & Sons.
- Kempthorne, O. (1966). Some aspects of experimental inference. *J. Amer. Statist. Assoc.*, **61**, 11-34.
- Kempthorne, O. & Folks, L. (1971). *Probability, Statistics, and Data Analysis*. Iowa University Press.
- Kulldorff, G. (1962). *Contributions to the Theory of Estimation Form Grouped and Partially Grouped Samples*. Almqvist & Wiksell.
- Lindley, D.V. (1950). Grouping corrections and maximum likelihood equations. *Proc. Camb. Philos. Soc.*, **46**, 106-110.
- Lindsey, J.K. (1995). *Modelling Frequency and Count Data*. Oxford University Press.
- Smith, R.L. (1985). Maximum likelihood estimation in class of non-regular cases. *Biometrika*, **72**, 67-92.
- Stuart, A. & Ord, J.K. (1986). *Kendall's Advanced Theory of Statistics*. Vol. 1, 6th edition. Edward Arnold.
- Swan, A.V. (1969). Algorithm AS 16: Computing maximum-likelihood estimates for parameters of the normal distribution form grouped and censored data. *Appl. Statist.*, **18**, 110-114.
- Titterton, D.M. (1985). Comment on "Estimating parameters in continuous univariate distributions". *J. Roy. Statist. Soc.*, **B47**, 115-116.
- Uesaka, H. & Goto, R. (1982). Some properties of Box-Cox estimates of transforming parameters in power-normal distributions. *Bull. Biometric. Soc. Japan*, **3**, 1-22.
- Winkelmann, R. (1994). *Count Data Models*. Springer-Verlag.
- Wolynetz, M.S. (1979a). Algorithm AS 138: Maximum likelihood estimation from confined and censored samples. *Appl. Statist.*, **28**, 185-195.
- Wolynetz, M.S. (1979b). Algorithm AS 139: Maximum likelihood estimation in a linear model from confined and censored normal data. *Appl. Statist.*, **28**, 195-206.
- 上坂浩之・後藤昌司 (1980). ベキ正規変換に基づく臨床検査データの解析. 応用統計学, **9**(1), 23-33.
- 仮谷太一 (1975). 初潮年齢分布の最尤推定. 川崎医学会誌, **1**, 75-83.
- 仮谷太一 (1979). 齲蝕(歯牙)年齢分布について. 応用統計学, **9**(2), 73-81.
- 仮谷太一・赤坂守人 (1976). 乳歯萌出年齢分布の母数推定. 応用統計学, **5**(1), 3-18.
- 後藤昌司 (1986). 統計的データ解析の過程. 行動計量学, **13**(2), 48-63.
- 後藤昌司・山本成志・井上俊昭 (1991). ベキ正規分布のパラメータの推定: 推定量の漸近挙動について

- て. 計算機統計学, **4**(1), 45-60.
- 杉浦成昭 (1980). 共通 1 次試験総合得点に対する分布のあてはめ. 応用統計学, **9**(2), 95-107.
- 杉浦成昭 (1981). 共通 1 次試験総合得点に対する分布のあてはめ II. 応用統計学, **10**(1), 39-52.
- 堂前 章・宮原英夫 (1984). 体重の三乗根は正規分布するか?: マイクロ・コンピュータによる保健管理データの処理と分布型へのあてはめ. 行動計量学, **11**(2), 1-11.
- 鍋谷清治 (1983). 区間データに対する最尤推定について. 応用統計学, **12**(2), 59-67.
- 濱崎俊光・後藤昌司 (1996). ベキ変換の変換尺度の不変性調整. 計算機統計学, **9**(1), 37-53.
- 松原義弘・後藤昌司 (1979). ベキ正規分布に基づく生存時間データの回帰分析. 応用統計学, **9**(1), 43-51.