ファジィ観測データに基づく統計的推定[†]

奥田 徹示*・古殿 幸雄*・浅居 喜代治*

Statistical Estimation by Fuzzy Observation Data

Tetsuji OKUDA*, Yukio KODONO* and Kiyoji ASAI*

In practice, there are cases that observations of objects in a population are fuzzy. It is not desirable for the statistical inference to discard the fuzzy data from fuzzy observations or to ignore the effect of fuzziness by reason of the complexity of the treatment of fuzziness.

In this paper, the fuzzy observation with human vagueness on a probability space is considered. The correction formulae eliminating the effect of fuzziness in order to estimate actual moments and actual cumulants of a population are given, when the moments and cumulants are calculated from fuzzy data obtained by fuzzy observations. Furthermore, the validity of the corrections is investigated by the simulation and the experiment on fuzzy observations.

Key Words: fuzzy observations, fuzzy data, corrections for fuzziness, statistical estimates from fuzzy data

1. ま え が き

最近、種々のシステムが複雑・大規模化するに伴って、これらのシステムをモデル化するとき、不確かさやあいまいさが多く含まれるようになってきた。特に、経営システムなど人間を含むシステムでは、取り扱うデータにかなりのあいまいさが含まれる。

従来からの観測では、観測値をできるだけ正確なものにするために、人間の主観的なあいまいさをできるだけ排除する傾向にあり、そのために技術的・時間的・コスト的にかなりの労力を必要とする場合があった。しかしながら、人間の主観的なあいまいさを含むデータを用いて、正確に観測されたデータに基づいて

推定される推定値とそれほど変わらない推定値を得ることのできるデータ処理方法があれば、現実のデータ解析において有効なものになると思われる。そこで本論文では、ファジィ理論に基づいて、人間の主観的なあいまいさを伴う観測データの処理方法とその方法の実用性について考察している。

統計学の分野において、正確な観測がコストなどの関係で困難な場合、ある区間に属する事象が何件発生したかというようなグルーピングデータに関する分析が古くから研究されている。特に、正規分布のような滑らかな分布を念頭においた Sheppard の補正は現在でも有名で、十分実用性をもっていると思われる。これについては、1898年の Sheppard の論文¹⁾以来その研究が始まったと思われる。その後、モーメントやキュムラントの推定に関連した研究がいろいろなされてきている^{2)~7)}。本論文では主にモーメント推定に議論を絞っているが、最近ではグルーピングデータに基づく最尤推定や回帰分析、仮説検定などの研究がなされている^{8)~11)}。

しかしながら、ある区間、たとえば10以上20未満と20以上30未満との区間を考えたとして、19.9および20の値が発生した場合、観測者は19.9は前者の区間に、20は後者の区間に属するということにはっきりと境界を区別して観測していると考えよりも、両者の区間の境界がぼやけた状態で、主観的なあいまいさを伴って観測されたと考えたほうが現実的な場合も多いと考えられる。そこで本論文では、あいまいな境界をもつ区間をファジィ集合として表現して、メンバーシップ関数の形状とあいまいさの幅を考慮して、あいまいな観測データを統計的に処理する方法について検討している。さらに、ここで導いた方法の実用性を、数値例やシミュレーションおよび簡単な観測実験によって検討し、その結果、ある程度の実用可能性のあることが明らかとなっている。

[†] 計測自動制御学会関西支部シンポジウムで発表(1988・9)

* 大阪工業大学 大阪市旭区大宮 5-16-1

* Osaka Institute of Technology, Osaka

(Received June 20, 1989)

(Revised November 6, 1989)

2. ファジィ観測と ファジィデータ

本論文では、1次元実数値空間を X として、 X 上に確率分布が与えられるものとする。通常、観測値あるいは実現値はある1点 x_j として取り扱われるが、ここでは点 x_j の周りのある幅に主観的なあいまいさが伴っている場合を考える。このような観測をファジィ観測と呼び、これを区間 $H_j = [x_j - h/2, x_j + h/2]$ 上で定義されるファジィ事象（ファジィ集合） X_j としてとらえる。このファジィ事象 X_j のメンバシップ関数 $\chi_j: H_j \rightarrow [0, 1]$ は、 $x \in H_j$ のファジィ事象 X_j への帰属度を表わし、 $\chi_j(x_j) = 1$ としておく。そこで、ファジィ観測によって得られたファジィ事象 X_j をファジィ観測値と考え、これをファジィデータと呼ぶことにする。なお、ファジィ事象とメンバシップ関数とは1対1に対応するので、以後は X_j と χ_j とを同一視し、記法の複雑化を避けるため統一的に χ_j を用いる。

さて、ファジィ観測値を統計的に処理するためには、ファジィデータ χ_j の出現確率を計算する必要がある。この確率は、Zadeh の定義¹²⁾を用いて

$$P_j = \int_{x_j - h/2}^{x_j + h/2} \chi_j(x) f(x) dx \quad (2.1)$$

と計算される。ここで、 $f(x)$ は母集団の確率密度関数である。

本論文では、ファジィ観測の基本型として、区間 $[-h/2, h/2]$ 上で定義されるつぎの3種類のメンバシップ関数を定義しておく。

(i) 三角形型：

$$\chi(x) = -\frac{2}{h}|x| + 1, \quad (2.2)$$

(ii) 2次曲線型：

$$\chi(x) = \frac{8}{h^2} \left(|x| - \frac{h}{2} \right)^2 \left[\frac{2}{h} \left(|x| + \frac{h}{4} \right) \right] + \left(1 - \frac{8x^2}{h^2} \right) \left(1 - \left[\frac{2}{h} \left(|x| + \frac{h}{4} \right) \right] \right), \quad (2.3)$$

(iii) 台形型：

$$\chi(x) = \max \left\{ 0, 1 \wedge \left(-\frac{1}{2e}|x| + \frac{h}{8e} + \frac{1}{2} \right) \right\}. \quad (2.4)$$

ここで、 $[A]$ は A の整数部を表わすガウス記号で、 $a \wedge b$ は $\min\{a, b\}$ を表わす。また、 e は台形の形状に対するパラメータでその範囲は $0 < e \leq h/4$ である。

(a) Triangle type (b) Quadratic curve type (c) Trapezoid type

Fig. 1 Three types of fuzzy observations

これらを図示すると Fig. 1 のようになる。

これらの $\chi(x)$ は、 $x=0$ を中心として両側に $h/2$ の幅であいまいが存在していることを示す。すなわち、度合 $\chi(x)$ はファジィ観測に基づいて値 x が発生したと思われる主観的な信頼度を表わしている。これら基本型の $\chi(x)$ を用いて、ファジィデータ χ_j は $\chi_j(x) = \chi(x - x_j)$ と表わされる。そこで、ファジィデータ χ_j を得たとき、区間 H_j の中心値 x_j をファジィデータ χ_j の代表値として選ぶ。この代表値はもはや non-fuzzy な通常の数値であり、これら代表値を用いて通常の統計的な計算方法でデータ処理を行おうとするのがここでの考え方である。ただ、推定問題を取り扱う場合、代表値を用いて処理すると偏りを生じるので、この偏りを除くための補正方法について検討するのが本論文の目的である。

したがって、代表値 x_j を実現値としてとる確率変数を X^R とすると、 $\Pr\{X^R = x_j\}$ の値が(2.1)式で与えられ、観測データは X^R の実現値であると考えてデータ処理を行うことになる。なお、ここでは空間 X を $x_{j+1} = x_j + h/2$ ($j = \dots, -1, 0, 1, \dots$) となるようにファジィ観測の刻みを設定している。したがって、ここで考えるモデルとしては、観測者はファジィ観測によってどのファジィデータ χ_j が発生したかを報告するものとする。ここで述べた三つの基本型に対しては $\sum P_j = 1$ は満たされている。

なお、 $\sum P_j = 1$ の条件は知覚カテゴリーなどのメンバシップ関数に対しては一種の制約となるが、理論上の解析ではこの条件を前提としておく。後の5章で述べるシミュレーションおよび実験の部分では、 $\sum P_j = 1$ の条件を緩和した状況で検討する。

3. モーメントおよびキュムラントの補正

さて、ファジィデータ（代表値）から計算される r 次の生のモーメントは、形式的に

$$\bar{\mu}_r = \sum_{j=-\infty}^{\infty} x_j^r P_j$$

$$= \sum_{j=-\infty}^{\infty} x_j^r \int_{-h/2}^{h/2} \chi(v) f(x_j+v) dv \quad (3.1)$$

と表わされる。オイラー・マクローリンの和分公式⁷⁾を用いて和分を積分近似すると、三角形型の場合、

$$\begin{aligned} \bar{\mu}_r &= \frac{2}{h} \int_{-\infty}^{\infty} x^r \int_{-h/2}^{h/2} \chi(v) f(x+v) dv dx \\ &= \frac{2}{h} \int_{-\infty}^{\infty} \int_{-h/2}^{h/2} (u-v)^r \chi(v) f(u) du dv \\ &= \sum_{j=0}^{[r/2]} \frac{2\mu_{r-2j}}{(2j+2)(2j+1)} \binom{r}{2j} \left(\frac{h}{2}\right)^{2j} \end{aligned} \quad (3.2)$$

となる。ここで、 μ_k は母集団の正確な k 次モーメントを表わしている。

和分から積分への近似の妥当性については本章の数値例の部分で述べるが、ここでは近似が十分意味をなすとして、“ $\doteq$ ”を“ $=$ ”と記すことにすると、具体的な補正公式として次式を得る。

(三角形型)

$$\begin{aligned} \mu_0 &= \bar{\mu}_0 = 1, \quad \mu_1 = \bar{\mu}_1, \quad \mu_2 = \bar{\mu}_2 - \frac{h^2}{24}, \\ \mu_3 &= \bar{\mu}_3 - \frac{h^2}{8} \bar{\mu}_1, \quad \mu_4 = \bar{\mu}_4 - \frac{h^2}{4} \bar{\mu}_2 + \frac{h^4}{160}, \\ &\dots\dots\dots \end{aligned} \quad (3.3)$$

同様に、2次曲線型と台形型の場合についても計算でき、その結果をつぎに記す。

(2次曲線型)

$$\begin{aligned} \bar{\mu}_r &= \sum_{j=0}^{[r/2]} \frac{2\mu_{r-2j}}{(2j+3)(2j+2)(2j+1)} \binom{r}{2j} \\ &\quad \times \left\{ \left(\frac{1}{2}\right)^{2j-2} - \left(\frac{1}{4}\right)^{2j} \right\} h^{2j}, \\ \mu_0 &= \bar{\mu}_0 = 1, \quad \mu_1 = \bar{\mu}_1, \quad \mu_2 = \bar{\mu}_2 - \frac{h^2}{32}, \\ &\dots\dots\dots \end{aligned} \quad (3.4)$$

(台形型)

$$\begin{aligned} \bar{\mu}_r &= \sum_{j=0}^{[r/2]} \frac{2\mu_{r-2j}}{(2j+2)(2j+1)q} \binom{r}{2j} \\ &\quad \times \left\{ \left(\frac{1+q}{4}\right)^{2j+2} - \left(\frac{1-q}{4}\right)^{2j+2} \right\} h^{2j}, \\ \mu_0 &= \bar{\mu}_0 = 1, \quad \mu_1 = \bar{\mu}_1, \\ \mu_2 &= \bar{\mu}_2 - \frac{1}{3} \left(\frac{1}{16} + q^2\right) h^2, \\ &\dots\dots\dots \end{aligned} \quad (3.5)$$

ここで、台形型の場合 $e=qh$ と変換している。これらの(3.2)～(3.5)式は、ファジィデータを用いて計算される生のモーメント $\bar{\mu}_r$ から補正公式によって正確なモーメント μ_r が導き出されることを示している。そして、補正項の大きさはあいまいさの幅 h によって評価でき、 h が比較的小さければ ($h < 1$)、 h^4 以

上の項はほとんど無視できるので、補正は h^2 の項までで十分実用的であろうと考えられる。

モーメントと同じく統計解析においてよく用いられるものにキュムラントがあるが、キュムラントについても同様な補正公式を導くことができる。ファジィデータ(代表値)から計算される積率母関数は

$$\phi(s) = \sum_{j=-\infty}^{\infty} \exp(sx_j) \int_{-h/2}^{h/2} \chi(v) f(x_j+v) dv \quad (3.6)$$

であり、オイラー・マクローリンの和分公式が妥当性を失わないとして、

$$\phi(s) = \frac{2}{h} \int_{-\infty}^{\infty} \int_{-h/2}^{h/2} \exp\{s(u-v)\} \chi(v) f(u) du dv \quad (3.7)$$

と表わせる。そこで、

$$\phi(s) = \frac{2}{h} \int_{-h/2}^{h/2} \exp(-sv) \chi(v) dv \quad (3.8)$$

として、 $\phi(s)$ の対数をとって、

$$\begin{aligned} \log \phi(s) &= \log \phi(s) + \log \int_{-\infty}^{\infty} \exp(su) f(u) du \\ &\equiv \sum_{k=1}^{\infty} c_k s^k / (k!) \end{aligned} \quad (3.9)$$

と表わせば、形式上 c_k はファジィデータから計算される生のキュムラントになる。したがって、 β_k を密度関数 $f(x)$ に対する正確なキュムラントとすると、

$$\sum c_k s^k / (k!) = \log \phi(s) + \sum \beta_k s^k / (k!) \quad (3.10)$$

なる関係が成立する。

ところで、ベルヌーイ多項式に基づく関係

$$\frac{2}{e^z - 1} - \frac{2}{z} + 1 = 2 \sum_{j=2}^{\infty} \frac{B_j z^{j-1}}{j!} \quad (3.11)$$

を利用して、この両辺を 0 から z まで積分すると、

$$\log \frac{2 \cosh z - 2}{z^2} = 2 \sum_{j=2}^{\infty} \frac{B_j z^j}{(j!)j} \quad (3.12)$$

となる。ここで、 B_j はベルヌーイ数と呼ばれ、 $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$, ..., $B_3 = B_5 = \dots = 0$ である。

三角形型の場合、

$$\phi(s) = \frac{2 \cosh \frac{sh}{2} - 2}{\left(\frac{sh}{2}\right)^2} \quad (3.13)$$

となるので、(3.12)式の関係を用いて

$$\log \phi(s) = \sum_{j=2}^{\infty} \frac{B_j h^j}{j^{2j-1}} \frac{s^j}{j!} \quad (3.14)$$

となり、

$$c_r = \frac{B_r h^r}{r^{2r-1}} + \beta_r \quad (r > 1) \quad (3.15)$$

なる関係が得られる。したがって、キュムラントの補正公式として、

$$\begin{aligned}\beta_2 &= c_2 - \frac{h^2}{24}, \quad \beta_4 = c_4 + \frac{h^4}{960}, \\ \beta_6 &= c_6 - \frac{h^6}{8064}, \quad \dots\dots\dots\end{aligned}\quad (3.16)$$

が導かれる。

同様に、2次曲線型の場合は

$$\beta_r = c_r - \frac{B_r(2r+2)h^r}{r2^r} \quad (r>1), \quad (3.17)$$

となり、台形型の場合は

$$\beta_r = c_r - \frac{B_r h^r}{r} \left\{ \frac{1}{2^r} + (2q)^r \right\} \quad (r>1), \quad (3.18)$$

となる。ただし、 $e=qh$ と変換している。

三角形型、2次曲線型、台形型のいずれの場合も、 r 次の補正項のオーダは h^r になっている。したがって、 h がそれほど小さくなく、 h^4 以上の項が無視できるとすれば、キュムラントの補正は2次のキュムラントのみにについて考慮すればよいことになり、解析上の見通しをわかりやすくするといった利点があると考えられる。

なお、ここで導いたモーメントおよびキュムラントの補正公式において、 $e=h/4$ とすれば台形型メンバシップ関数は三角形型と一致し、補正公式も同様に一致する。したがって、台形型は三角形型の一般化とも考えられる。また、2次曲線型はつり鐘型などを代表するタイプとしてここでは考えている。本論文では、あいまいさの幅 h がそれほど大きくない場合を前提としており、メンバシップ関数の型としては、ここで取り上げた3種類ぐらいの設定で十分であろうと思われる。

また、あいまいさの幅を標準偏差 σ との比で考察したい場合、

$$\frac{x}{\sigma} = u, \quad f(\sigma u) \left| \frac{dx}{du} \right| = g(u),$$

$$\chi_j(\sigma u) = \nu_j(u), \quad \frac{h}{\sigma} = \alpha \quad (3.19)$$

などと変換しておけばよい。これらの変換により、 h の役割を α に置き換えれば、ここでの補正公式はそのまま成立する。もちろん、この場合 h が σ をこえないこと、つまり $\alpha < 1$ となることが望ましい。

〔数値例〕ここでは、数値例としてつぎのベータ分布

$$f(x) = \frac{1}{B(10, 2)} x^9 (1-x), \quad 0 \leq x \leq 1 \quad (3.20)$$

を取り上げる。 $h=0.4$ とした三角形型のファジィ観測の場合について、Table 1の頻度を用いてモーメントおよびキュムラントの補正について調べた結果、Table 2, 3のようになった。なお、Table 1の頻度は(3.20)式の $f(x)$ と(2.2)式の $\chi(x)$ とを用いて、(2.1)式に基づき解析的に求めたものである。

Table 2で示されているモーメントの補正では、補正前の誤差が1/100のオーダであるが、補正後の誤差は1/1,000のオーダになっている。Table 3で示されているキュムラントについては、もともと値が小さいため、2次のキュムラントでは補正前の誤差が1/1,000のオーダであったものが補正後は1/10,000の

Table 1 Frequencies of the occurrence of fuzzy event

k	x_k (mid-point)	P_k (probability)
0	0.0	0.0000001
1	0.2	0.0001396
2	0.4	0.0087903
3	0.6	0.1251655
4	0.8	0.5560721
5	1.0	0.3098322

(Total) 0.9999998

Table 2 Corrections of moments

k (order)	μ_k (true)	m_k (raw)	(error)	Corrected h^2	(error)	Corrected h^4	(error)
1	0.8333333	0.8333332	(-0.0000001)				
2	0.7051282	0.7121900	(0.0070618)	0.7055233	(0.0003951)		
3	0.6043956	0.6221406	(0.0177450)	0.6054739	(0.0010783)		
4	0.5238695	0.5540460	(0.0301765)	0.5255584	(0.0016889)	0.5257184	(0.0018488)

Table 3 Corrections of cumulants

k (order)	β_k (true)	c_k (raw)	(error)	Corrected	(error)
1	0.8333333	0.8333332	(-0.0000001)		
2	0.0106838	0.0177458	(0.0070620)	0.0110791	(0.0003953)
3	-0.0010176	-0.0009273	(0.0000903)		
4	-0.0000901	-0.0000014	(0.0000887)	0.0000253	(0.0001154)

オーダとなっている。4 次のモーメントやキュムラントについて注目してみると、厳密に補正をしたからといって必ずしも誤差がより減少するとは限っていない。これは、補正項が h^4 のオーダであり、オイラー・マクローリンの和分公式の近似誤差が影響しているものと思われる。すなわち、 h^4 のオーダの補正項の値がすでに和分から積分へ近似したときの近似誤差の範囲に入っているものと考えられる。しかしながら、誤差そのものの値は微小であり、実用上はそれほど問題にはならないであろう。

なお、ここでの補正が正しく働くかどうかは、オイラー・マクローリンの和分公式の適用の妥当性にかかっているが、数値解析の理論からは、密度関数の尾部が高次の滑らかさを保つことが条件となる。これは、通常の Sheppard の補正に対する議論と同じことである^{6),7)}。

4. 平均補正

前章の数値例では、ファジィ観測の中心値（刻み位置） x_i を固定した場合の例であり、刻み位置が変化すれば補正の効果も変化すると考えられる。そこで、刻み位置が区間 $[X_i - h/4, X_i + h/4]$ でランダムに発生した場合、生のモーメントの平均値がいくらかになるかということについて考察しておく。

いま、 $\{x_j\}$ ($j = \dots, -1, 0, 1, \dots$) を x の刻み位置の集合とする。このとき、 x_i が区間 $[X_i - h/4, X_i + h/4]$ で一様分布に従ってランダムに発生するとする。この状況を Fig. 2 に示す。

そこで、ある刻み位置 x_i が発生したとすると、その場合の生のモーメント $\bar{\mu}_m$ は、 x_i を含む刻み位置の集合 $\{x_p\}$ に対して、

$$\bar{\mu}_m = \sum_{p=-\infty}^{\infty} x_p^m \int_{-h/2}^{h/2} \chi(v) f(x_p + v) dv \quad (4.1)$$

となる。したがって、 $\bar{\mu}_m$ の平均は、 x_i を確率変数の実現値とみなして、

Fig. 2 Location of fuzzy interval mesh

$$\begin{aligned} E(\bar{\mu}_m) &= \frac{2}{h} \int_{X_i - h/4}^{X_i + h/4} \left\{ \sum_{p=-\infty}^{\infty} x_p^m \right. \\ &\quad \cdot \left. \int_{-h/2}^{h/2} \chi(v) f(x_p + v) dv \right\} dx_i \\ &= \frac{2}{h} \sum_{j=-\infty}^{\infty} \int_{X_j - h/4}^{X_j + h/4} \left\{ x_j^m \int_{-h/2}^{h/2} \chi(v) f(x + v) dv \right\} dx \\ &= \frac{2}{h} \int_{-\infty}^{\infty} x_j^m \left\{ \int_{-h/2}^{h/2} \chi(v) f(x + v) dv \right\} dx \end{aligned} \quad (4.2)$$

となる。

この結果は、 $\bar{\mu}_m$ の平均値が $\bar{\mu}_m$ の積分近似と一致していることを示している。すなわち、ここでの補正公式はオイラー・マクローリンの和分公式による近似誤差にかかわらず、多数回の試行においては平均的に補正が正しく働くことになる。

5. シミュレーションおよび観測実験

前章までは、左右対称形のメンバシップ関数を前提として補正公式を検討したが、現実問題としては、左右対称なメンバシップ関数の設定が現実のファジィ観測に即するかどうかは疑問である。そこで本章では、メンバシップ関数が左右非対称となる場合を前提として、簡単なシミュレーションと観測実験とによって本補正公式の実用性について検討する。

ここでのデータの取り扱い方をつぎのように考える。たとえば、Fig. 3 のような状態を考え、現実のデータとして x_i が発生したとき、 $\chi_i^L(x_i)$ あるいは $\chi_i^R(x_i)$ はファジィ観測によってファジィデータ χ_i^L あるいは χ_i^R が発生したと観測者から報告される可能性の度合を表わしているとする。つまり、発生したデータ x_i に対して、ファジィデータ χ_i^L として報告されるかもしれないし、またはファジィデータ χ_i^R として報告されるかもしれない。そこで、それぞれのファジィデータとして報告されるであろう可能性の度合が $\chi_i^L(x_i)$ および $\chi_i^R(x_i)$ で表わされているものとしている。したがって、もし x_i が n 回発生した場合、 χ_i^L が $n \times \chi_i^L(x_i)$ 回、また χ_i^R が $n \times \chi_i^R(x_i)$ 回観測

Fig. 3 Treatment of fuzzy data

者からファジィデータとして報告されると考えておく。もちろん、ファジィデータを実際の統計計算に用いる場合は、Fig. 3 にあるようにその代表値（中心値） v_i^L あるいは v_i^R を用いることになる。

シミュレーションは、台形型のファジィ観測に対して、標準偏差の推定に関して行った。ここで台形型のメンバシップ関数を選んだ理由は、左右非対称性が理論上の補正公式の適用の妥当性に与える影響を調べることをシミュレーションの目的としたため、左右非対称性の効果をわかりやすく分析するためである。ここでは、 i 番目の発生データを x_i として、その両側にある各ファジィデータの代表値を、Fig. 3 のように、 v_i^L 、 v_i^R で表わすことにする。そこで、標本分散を

$$\text{Var}(v) = \frac{1}{n} \sum_{i=1}^n \{ (v_i^L - \bar{v})^2 \chi_i^L(x_i) + (v_i^R - \bar{v})^2 \chi_i^R(x_i) \} \quad (5.1)$$

ただし、

$$\bar{v} = \frac{1}{n} \sum_{i=1}^n \{ v_i^L \chi_i^L(x_i) + v_i^R \chi_i^R(x_i) \} \quad (5.2)$$

として、標準偏差を台形型の補正公式を用いて、

$$S = \sqrt{\text{Var}(v) - \frac{1}{3} \left(\frac{1}{16} + q^2 \right) h^2} \quad (5.3)$$

として推定する。

本シミュレーションにおいては、母集団から発生させる現実（真）のデータに対して、観測者が、観測結果をファジィ観測によりファジィデータとして報告するものと仮定している。そこで、このファジィデータの代表値を用いて計算する生の標準偏差に補正公式を適用して得られる推定値と、発生した現実のデータを用いて標準偏差を推定した場合の推定値との比較検討を行う。

設定した母集団は、平均 50、標準偏差 16 の正規集団で、200 個のデータを発生させるシミュレーションを行った。ファジィ観測の型は台形型を想定し、 $h=20$ 、 $e=2.5$ として、ファジィ観測の代表値の刻みは -10 から 110 までの 10 刻みで設定した。

Fig. 4 はサンプルサイズを 200 とした場合の標準偏差の推定例であるが、台形型メンバシップ関数のパラメータ e は、台形の両側に対してそれぞれ独立に 2.5 の値を中心にして上下 1.0 の幅でランダムに変動させた。そのときの状況を Table 4 に示す。このような状況でのシミュレーション結果が Fig. 4 である。

この Fig. 4 から、現実のデータによる推定値に対して、ファジィデータによる推定値はある一定の幅で誤差が生じており、偏りのあることがわかる。そして補正公式により補正を施すと、ほぼその偏りが解消で

Fig. 4 Simulation result (I) of estimating standard deviation

Table 4 The values e of trapezoid type membership functions

Mid-point of fuzzy interval	Value e (left side)	Value e (right side)
-10	0	1.78676
0	2.33870	3.39567
10	3.20798	2.59636
20	2.37217	3.40250
30	1.83713	1.57777
40	2.38215	2.63584
50	2.51309	1.84037
60	3.49363	2.20505
70	2.33743	2.49354
80	1.80817	3.21170
90	2.41952	1.63883
100	2.43515	2.78079
110	1.86874	0
Standard deviation	0.484557	0.634892
Mean	2.41782	2.46376
Total standard deviation	0.565216	
Total mean	2.44079	

ることがわかる。そこで補正公式の平均的な働きを調べるため、Fig. 4 の場合と同じ条件でシミュレーションを 20 回繰り返し行って、その平均をプロットしたものが Fig. 5 である。この Fig. 5 からわかるように、左右非対称となるような現実的なメンバシップ関数の設定のもとでも、平均的にみて、ファジィデータに基づく推定値を補正することにより、現実のデータに基づく推定値とほとんど変わらない推定値が得られると考えられる。

さらに具体的に、実際に人間が主観的な尺度で観測を行う場合についてのファジィ観測の簡単な実験を行

Fig. 5 Simulation result (II) of estimating standard deviation

った。本実験では、実際のファジィ観測によって得られたファジィデータがどのような形で観測されているかということと、ここでの補正公式の現実のデータへの適用の可能性について検討した。

実験はつぎのような方法で行った。Fig. 6 に示すように、パーソナルコンピュータのディスプレイ上に白色の長方形と、そのなかの一部分に画かれた青色の長方形とを表示させる。観測者には、この青色の長方形の面積が全体の白色の長方形の面積の何%に当たるかを 10% 刻みで答えさせる。青色の長方形は、パーセントの値にして平均 50、標準偏差 16 の正規分布に従って発生させた。したがって、実際には 42% とか 67% とかというような青色の長方形が表示され、観測者はそれを直観的に観測して、だいたい何%の長方形であるかを 10% 刻みで報告する。刻みの設定は 0, 10, ..., 100% とした。

Fig. 6 Indication of the rectangle on the display

データは 1,000 個発生させ、Fig. 7 は 1,000 個のファジィデータから被験者（観測者）のもつメンバシップ関数を推定したものである。たとえば、73% の青色の長方形が何回か発生したとして、そのうちの何回が 70% として報告され、またそのうちの何回が 80% として報告されたかの割合を縦軸にとり、横軸に発生したデータとしてのパーセントの値をとったものが Fig. 7 である。そして、この結果を単純な直線近似で台形型のメンバシップ関数に当てはめたものが Fig. 8 である。この図から、 e の値の平均は約 2.85 となったので、この e の値を用いて標準偏差を推定した結果が Fig. 9 である。

この Fig. 9 より、ファジィデータに基づく推定値が、補正公式の適用により、現実のデータに基づく推定値に近づいていくようすがわかる。なお、実験で用いた刺激は視覚的図形であるので、その知覚にはあいまいさがあって、ファジィデータとしての扱いが当然

Fig. 7 Membership functions estimated from experimental data

Fig. 8 Approximation of membership function

Fig. 9 Result of the experiment

適切なものとなり、non-fuzzy な区間を前提とする従来からの Sheppard の補正を適用するよりは本手法を適用するほうが良い結果を与えている。

6. む す び

本研究では、人間の主観的なあいまいさを伴うファジィデータの統計的な処理方法の提案とその応用の可能性について検討した。

3章での数値例や4章での平均補正の議論から、本論文で導かれた補正公式は理論上は十分妥当であると考えられる。ただし、現実のデータ解析に利用するためには、ファジィ観測が常に左右対称となる理想的なメンバシップ関数で設定できるかどうかはかなり疑問となる点である。そこで、メンバシップ関数の左右対称性を緩和したより現実的な条件のもとでシミュレーションや観測実験を行い、ここでの補正公式の実用性について検討した。その結果、現実のデータ解析に対しても、補正公式がある程度うまく働くことがわかった。特に、平均的な意味ではかなりの妥当性をもってることが明らかにされた。

最後に、本研究を進めるに当たりご助言をいただい

た本学の杉浦寅彦教授に深甚なる謝意を表する。本研究は、科学技術庁の平成元年度科学技術振興調整費による「ファジィシステムとその人間・自然系への適用に関する研究」の一環として行ったものである。

参 考 文 献

- 1) W.H. Sheppard: On the Calculation of the Most Probable Values of Frequency-Constants, for Data Arranged According to Equidistant Divisions of a Scale, Proc. London Math. Soc., **29**, 353/380 (1898)
- 2) K. Pearson, et al.: On Corrections for the Moment-Coefficients of Limited Range Frequency Distributions when There Are Finite or Infinite Ordinates and Any Slopes of the Range, Biometrika, **12**, 231/259 (1919)
- 3) R. A. Fisher: On the Mathematical Foundations of Theoretical Statistics, Phil. Tran. Roy. Soc., A, **222**, 309/368 (1921)
- 4) W.H. Langdon and O. Ore: Semi-Invariants and Sheppard's Correction, Ann. Math., **31**, 230/232 (1930)
- 5) S. Kullback: A Note on Sheppard's Corrections, Ann. Math. Statist., **6**, 158/159 (1935)
- 6) M. G. Kendall: The Conditions under which Sheppard's Corrections Are Valid, J. Roy. Stat. Soc., **101**, 592/605 (1938)
- 7) A. Stuart and J.K. Ord: Kendall's Advanced Theory of Statistics, p. 93, Charles Griffin, 5th ed. (1987)
- 8) A. Indrayan and J.S. Rustagi: Approximate Tests of Hypotheses in Regression Models with Grouped Data, Commun. Statist.-Simula. Computa., B, **10**, 327/341 (1981)
- 9) J. Burridge: A Note on Maximum Likelihood Estimation for Regression Models Using Grouped Data, J.R. Statist. Soc., B, **43**, 41/45 (1981)
- 10) T. Kariya: Analogous t and F Test Statistics Based on Grouped Data, Proceedings of the Pacific Statistical Congress-1985 (eds. by I.S. Francis, B. J. Manly and F.C. Lam), North-Holland, 275/279 (1986)
- 11) M. Schander and F. Schmid: Small Sample Properties of the Maximum Likelihood Estimators μ and σ from a Grouped Sample of a Normal Population, Commun. Statist.-Simula., **17**, 229/239 (1988)
- 12) L. A. Zadeh: Probability Measures of Fuzzy Events, J. Math. Anal. Appl., **28**, 421/427 (1968)