計数法による各種標本平均粒径の統計的偏りの理論的証明

森 芳 郎・菅 沼 彰

東京大学工学部 化学工学科

緒 言

液滴、気泡、粉体などの分散系を対象とする工学では、現象を適確に把握するために個数分布や個数平均径以外の量、たとえば重量分布や各種の平均粒径を使用することが多い。これらの量を測定するためには、たとえば液滴の写真観測や粉体粒子の顕微鏡計測のような計数法粒度測定が行われる場合が少なくない。計数法粒度測定の結果を有効に利用するためには、データの推計学的取り扱いが重要であると思われる。個数平均径や個数分布を取り扱うための推計学的手法はよく開発されているが、それ以外の場合については未解決の問題が多い。本報では、よく使用される平均径の中で、標本平均径が母平均径の不偏推定量になるのは個数平均径だけであり、それ以外はいずれも母分布のいかんにかかわらず統計的に偏った推定値になることを証明する。

1. 問題の輪郭

粒度分布を有する粉体の代表粒度としては、Table 1 に列挙した平均粒径と呼ばれる各種の代表値が用いられることが多いようである¹。これらの平均粒径にはそれぞれ名前がつけられているが、現在の慣用名には不統一不適切なものもあり、分野によって名称が違う例もあるのでここでは名前を記さない。ただし区別するための添字は著者らの1人が提案中の命名法⁹⁾を念頭において選んである。ここに x は粒径、 $f(x)$ は x の母個数頻度関数、 μ は母平均粒径である¹¹。粒子の形状、粒子の密度は x によらず一定であると考ええる。

計数法粒度測定によって得られるデータは、計測粒子数を n とするとき、

$$x_1, x_2, x_3, \dots, x_i, \dots, x_n \quad (10)$$

すなわち粒径 x の数値リストである。測定法によっては粒径 x 以外の量、たとえば個々の粒子の体積 v のリストが得られる場合もあるが、本報では $v_i \rightarrow x_i$ の換算に伴う誤差の類を検討の対象から除外して議論するので、測定データは粒径 x のリストであるとしても一般性を失わ

ない。以下、測定データは無限母集団からランダムに採取されたサイズ n の標本集団であると考ええる。サイズ n の標本集団における x の個数頻度関数 $f_n^*(x)$ は、

$$f_n^*(x) = (\nu_x/n)\Delta \quad (11)$$

で与えられる。ここに Δ はデルタ関数である。たとえば標本集団に $x=a$ なる粒子が ν_a 個含まれているならば、 $f_n^*(a) = (\nu_a/n)\Delta$ というのがEq. (11)の意味である¹¹⁾。 $f(x)$ は母数であるが、 $f_n^*(x)$ は標本値であって標本集団ごとに値が異なる。母平均粒径 μ の定義式Eqs. (1)~(9)の $f(x)$ を機械的に $f_n^*(x)$ で置き換えると標本平均粒径 X_n の定義式を得る。これを通常の計算公式と共に一括してTable 2に示す。 X_n は標本平均粒径、すなわちサイズ n の標本集団の平均粒径であって、標本集団ごとにその値は異なる。粒度測定を行って平均粒径を求めようとする場合、その目的は標本平均粒径 X_n の値を確定することではなく、母平均粒径 μ を推定することである。

母分布、すなわち測定対象となる粉体の真の粒度分布に関する仮定なしでの μ の推定値としては、現在までのところTable 2の X_n 以外には知られていないので、 X_n を計算して μ の推定値とすることが一般に行われている。粒径 x の母分布が対数正規分布であるという条件の下では標本集団にHatchの公式を適用して計算した標本値 X_n^* を μ の推定値にすることも可能であり、 X_n^* の推測統計学的な検討には若干の進展がみられる^{3,7,8)}。しかしながら母集団が厳密に対数正規分布に従う場合であっても、同じ標本集団について計算した X_n^* と X_n とが一般には違う値になることから明らかなように、 X_n^* と X_n とは別の統計量である。一般にある母数 γ を推定する手段は幾とおりもありうる。そして推定の方法が異なれば母数 γ とその推定値 g との関係も当然異なる^{1,2,10,12)}。本報で論ずるのは μ と X_n との関係である。Table 2に列挙した X_n の中で、推測統計学的な意味での μ との関係がよくわかっているのは x の標本算術平均値である ${}_nX_n$ だけであるように思われる。推測統計学で平均値といえば算術平均値を意味するのが普通であり、標本平均の分散、母平均の推定、母平均の有意差の検定、などというのはいずれも算術平均の分散、推定、検定を意味している。粉体や液滴などの分散系を対象とする工学では目的に応じて個数平均径以外の平均径も利用しなければならないが、そのような場合には名前は平均値であっても算術平均値に関する推計学の公式や手法

¹ モードやメディアンは本報の対象にしないが、平均粒径以外の平均粒度、たとえば平均沈降速度、比表面積などに対しては本報の内容を適用できる。

¹¹ $f(x)$ 、 μ は測定サンプルの値ではなくて母集団の値であることを明確にしておく必要があるので、それぞれ母個数頻度関数、母平均粒径という。

¹¹¹ デルタ関数を導入したのは、 $f_n^*(x)$ についても $f(x)$ と同様にたとえば $\int f_n^*(x)dx=1$ というような積分記号による表現を可能にするためである。

Table 1 Definitions of mean particle diameters (population values): $\int$ denotes $\int_0^{+\infty}$

$$N\mu = \int x f(x) dx \quad (1)$$

$$G\mu = \exp \left\{ \int (\ln x) f(x) dx \right\} \quad (2)$$

$$H\mu = \left\{ \int x^{-1} f(x) dx \right\}^{-1} \quad (3)$$

$$L\mu = \int x^2 f(x) dx / \int x f(x) dx \quad (4)$$

$$A\mu = \int x^3 f(x) dx / \int x^2 f(x) dx \quad (5)$$

$$W\mu = \int x^4 f(x) dx / \int x^3 f(x) dx \quad (6)$$

$$NA\mu = \left\{ \int x^2 f(x) dx \right\}^{1/2} \quad (7)$$

$$NV\mu = \left\{ \int x^3 f(x) dx \right\}^{1/3} \quad (8)$$

$$LA\mu = \left\{ \int x^3 f(x) dx / \int x f(x) dx \right\}^{1/2} \quad (9)$$

をそのまま適用することはできない。粒度測定それ自体が粒度測定の目的であるならば、推計学的根拠の明確なデータ処理手法が完備している個数平均径だけを考えることにしてもよいであろうが、分散系のプロセスや現象を説明、記述、解析、さらには予測する手段としての粒度測定であるから、推計学の立場からみれば扱いにくい量であってもそれを測定しなければならない。すなわち Table 1 に列挙したような母平均径をできるだけ正確に推定する推計学的手法の開発は、分散系を扱う工学にとっては大変重要な課題である。

著者らは先に、母重量平均径 $w\mu$ の推定値としての標本重量平均径 wX_n は負の偏りを伴う推定量であることを指摘した⁴⁻⁶⁾。

$$E(wX_n) - w\mu < 0 \quad (12)^{\dagger}$$

本報では、Table 2 の標本平均径で個数平均径以外のものはいずれも母平均径に対して偏りのある推定量であること、すなわち、

$$E(jX_n) - j\mu \begin{cases} > 0 & j=H, G \\ = 0 & j=N \\ < 0 & j=L, A, W, NA, NV, LA \end{cases} \quad (13)$$

を理論的に証明する。

2. 用語の定義と補助定理

2・1 用語の定義

一般にある母数 γ の推定量として、サイズ n の標本集

[†] ここに $E(g)$ は g の期待値、すなわち統計量 g を求める操作を無限回反復して得られる g の算術平均値である。

^{**} γ の標本値という表現はここに言う狭義の標本値以外の統計量をも含む場合があるので、区別を明確にした。

^{***} 一般に統計量 g_n が γ に確率収束することを、

$$p \lim_{n \rightarrow \infty} g_n = \gamma$$

と書く。その意味は、 $|g_n - \gamma| \geq \epsilon$ となる確率を $P(\epsilon)$ で表したとき、任意の $\epsilon > 0$ について、

$$\lim_{n \rightarrow \infty} P(\epsilon) = 0$$

が成立することである。

Table 2 Definitions and calculating formulas of sample mean particle diameters: $\int$ and $\sum$ respectively denote $\int_0^{+\infty}$ and $\sum_{i=1}^n$

$$NX_n = \int x f_n^*(x) dx = \sum x_i / n \quad (1S)$$

$$GX_n = \exp \left\{ \int (\ln x) f_n^*(x) dx \right\} = \exp \left\{ \sum \ln x_i / n \right\} \quad (2S)$$

$$HX_n = \left\{ \int x^{-1} f_n^*(x) dx \right\}^{-1} = \left\{ \sum x_i^{-1} / n \right\}^{-1} \quad (3S)$$

$$LX_n = \int x^2 f_n^*(x) dx / \int x f_n^*(x) dx = \sum x_i^2 / \sum x_i \quad (4S)$$

$$AX_n = \int x^3 f_n^*(x) dx / \int x^2 f_n^*(x) dx = \sum x_i^3 / \sum x_i^2 \quad (5S)$$

$$WX_n = \int x^4 f_n^*(x) dx / \int x^3 f_n^*(x) dx = \sum x_i^4 / \sum x_i^3 \quad (6S)$$

$$NAX_n = \left\{ \int x^2 f_n^*(x) dx \right\}^{1/2} = \left\{ \sum x_i^2 / n \right\}^{1/2} \quad (7S)$$

$$NVX_n = \left\{ \int x^3 f_n^*(x) dx \right\}^{1/3} = \left\{ \sum x_i^3 / n \right\}^{1/3} \quad (8S)$$

$$LAX_n = \left\{ \int x^3 f_n^*(x) dx / \int x f_n^*(x) dx \right\}^{1/2} = \left\{ \sum x_i^3 / \sum x_i \right\}^{1/2} \quad (9S)$$

団によって値が定まる統計量 g_n を選んだ場合に、Eq. (14) が成立すれば g_n は γ の一致推定量、Eq. (15) が成立すれば g_n は γ の不偏推定量といわれる。

$$\lim_{n \rightarrow \infty} E(g_n) - \gamma = 0 \quad (14)$$

$$E(g_n) - \gamma = 0 \quad (15)$$

母数 γ の定義式中の $f(x)$ を標本頻度関数 $f_n^*(x)$ で置き換えることによって定義される統計量を、本報では γ の狭義の標本値ということにする^{**}。 $f_n^*(x)$ は $f(x)$ に確率収束する¹⁾。

$$p \lim_{n \rightarrow \infty} f_n^*(x) = f(x) \quad (16)^{***}$$

したがって一般に γ の推定量 g_n として γ の狭義の標本値を選ぶならば、

$$p \lim_{n \rightarrow \infty} g_n = \gamma \quad (17)$$

となり、この g_n は明らかに γ の一致推定量になっている。Table 2 の X_n はそれぞれ Table 1 の μ の狭義の標本値であるから μ の一致推定量である。

$$\lim_{n \rightarrow \infty} E(X_n) - \mu = 0 \quad (18)$$

g_n が母数 γ の狭義の標本値であっても、 γ の不偏推定量になるとはかぎらない。母分散、母相関係数、母歪度、母尖度などの推定量としてそれらの狭義の標本値を用いると偏った推定になることは初等統計学でもよく知られている。

2・2 偏りを証明するための補助定理

標本平均径の偏り (Eq. (13)) の証明に使用する補助定理を一括して本節に示す。

補助定理 1 統計量 g_n が母数 γ の一致推定量であって、

$$\delta = E(g_n) - E(g_{n+1}) \quad (19)$$

が n によらず常に正であるならば g_n は正の偏り、常に負であるならば負の偏りを伴う推定量であり、常にゼロであるならば不偏推定量である。

補助定理 2 $m \geq 2$, $1 \leq k \leq m$ とし, m 個の実数値 y の集合を ${}_m\Phi$, ${}_m\Phi$ から k 番目の成分を除去して作った集合を ${}_m\varphi_k$, m 個の実数値中の最大値を $y_{\max}$, 最小値を $y_{\min}$ で表すことにする。

$${}_m\Phi = (y_1, y_2, \dots, y_{k-1}, y_k, y_{k+1}, \dots, y_m) \quad (20)$$

$${}_m\varphi_k = (y_1, y_2, \dots, y_{k-1}, y_{k+1}, \dots, y_m) \quad (21)$$

l 個の実数値 ($l \geq 1$) が与えられたときに l 個の実数値全部が関与してその値が定まるような関数を一般に F , ${}_m\Phi$, ${}_m\varphi_k$ によって定まる F の値をそれぞれ $F({}_m\Phi)$, $F({}_m\varphi_k)$ で表すことにして, Eq. (22) で D を定義する。

$$D = \frac{1}{m} \sum_{k=1}^m F({}_m\varphi_k) - F({}_m\Phi) \quad (22)$$

$w(z)$ は $y_{\min} \leq z \leq y_{\max}$ の範囲内では常に正の値をとる単調関数であって定数ではないものとする。 $h(z)$ は $y_{\min} \leq z \leq y_{\max}$ の範囲内で微分可能な非線型単調関数であり, 変曲点をもたないものとする。 h の逆関数を h^{-1} , 導関数を h' で表すものとする。

以上の条件と表記法の下では補助定理 2・1~2・5 が成立する。証明は一括して Appendix に示す。

補助定理 2・1 a, b を定数, ただし $a \neq 0$ として,

$$F = \sum_{i=1}^l (ay_i + b) / l \quad (23)^*$$

の形に書けるならば,

$$D = 0 \quad (24)$$

補助定理 2・2 Eqs. (25), (28) または Eq. (31) の形に書けて $y_{\max} - y_{\min} = 0$ ならば $D = 0$ である。

補助定理 2・3

$$F = \sum_{i=1}^l y_i w(y_i) / \sum_{i=1}^l w(y_i) \quad (25)$$

の形に書けるならば,

$$D > 0, \quad w(y_{\max}) - w(y_{\min}) < 0 \quad (26)$$

$$D < 0, \quad w(y_{\max}) - w(y_{\min}) > 0 \quad (27)$$

補助定理 2・4

$$F = h^{-1} \left\{ \sum_{i=1}^l h(y_i) / l \right\} \quad (28)$$

の形に書けるならば,

$$D > 0, \quad \{h(y_{\max}) - h(y_{\min})\} \{h'(y_{\max}) - h'(y_{\min})\} < 0 \quad (29)$$

$$D < 0, \quad \{h(y_{\max}) - h(y_{\min})\} \{h'(y_{\max}) - h'(y_{\min})\} > 0 \quad (30)$$

補助定理 2・5

$$F = h^{-1} \left\{ \sum_{i=1}^l h(y_i) w(y_i) / \sum_{i=1}^l w(y_i) \right\} \quad (31)$$

の形に書けるならば,

$$D > 0, \quad \{h(y_{\max}) - h(y_{\min})\} \{h'(y_{\max}) - h'(y_{\min})\} < 0$$

* $F({}_m\Phi) = \sum_{i=1}^m (ay_i + b) / m$

$F({}_m\varphi_k) = \sum_{i \neq k}^{m-1} (ay_i + b) / (m-1)$

$i < k$ で $i^* = i$, $k < i$ で $i^* = i-1$ とする。補助定理 2・3~2・5 でも同様に考えるものとする。

$$\& \quad w(y_{\max}) - w(y_{\min}) < 0 \quad (32)$$

$$D < 0, \quad \{h(y_{\max}) - h(y_{\min})\} \{h'(y_{\max}) - h'(y_{\min})\} > 0$$

$$\& \quad w(y_{\max}) - w(y_{\min}) > 0 \quad (33)$$

3. 標本平均径の偏りの証明

無限母集団から n 個の標本をランダムに採取する操作は, 第 1 段階で $(n+1)$ 個の標本を採取し, 第 2 段階で $(n+1)$ 個の標本からランダムに 1 個の標本を除去する操作と等価である。第 1 段階でサイズ $(n+1)$ なる標本集団,

$$\Theta = (x_1, x_2, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_{n+1}) \quad (34)$$

を得る確率密度を $p(\Theta)$, この標本集団について計算した標本平均径を ${}_jX(\Theta)$ で表すことにすると,

$$E({}_jX_{n+1}) = \int {}_jX(\Theta) p(\Theta) d\Theta \quad (35)$$

第 1 段階で Θ を得たとき, 第 2 段階で x_k が除去されてサイズ n なる標本集団,

$$\theta_k = (x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_{n+1}) \quad (36)$$

を得る確率は k によらず $1/(n+1)$ である。標本集団 θ_k について計算した標本平均径を ${}_jX(\theta_k)$, 第 1 段階で Θ を得たという条件の下での標本平均径 ${}_jX_n$ の期待値を $E({}_jX_n|\Theta)$ で表すことにすると,

$$E({}_jX_n|\Theta) = \frac{1}{n+1} \sum_{k=1}^{n+1} {}_jX(\theta_k) \quad (37)$$

よって ${}_jX_n$ の期待値 $E({}_jX_n)$ は Eq. (38) で与えられる。

$$\begin{aligned} E({}_jX_n) &= \int E({}_jX_n|\Theta) p(\Theta) d\Theta \\ &= \int \left\{ \frac{1}{n+1} \sum_{k=1}^{n+1} {}_jX(\theta_k) \right\} p(\Theta) d\Theta \end{aligned} \quad (38)$$

Eqs. (19), (35), (38) により,

$$\begin{aligned} {}_j\delta &= E({}_jX_n) - E({}_jX_{n+1}) \\ &= \int \left\{ \frac{1}{n+1} \sum_{k=1}^{n+1} {}_jX(\theta_k) - {}_jX(\Theta) \right\} p(\Theta) d\Theta \end{aligned} \quad (39)$$

それぞれの定義から明らかなように,

$$\Theta = {}_{n+1}\Phi \quad (40)$$

$$\theta_k = {}_{n+1}\varphi_k \quad (41)$$

$${}_jX(\Theta) = F({}_{n+1}\Phi) \quad (42)$$

$${}_jX(\theta_k) = F({}_{n+1}\varphi_k) \quad (43)$$

とおくことができるから,

$${}_jD = \frac{1}{n+1} \sum_{k=1}^{n+1} {}_jX(\theta_k) - {}_jX(\Theta) \quad (44)$$

Eqs. (39) と (44) により,

$${}_j\delta = \int {}_jD p(\Theta) d\Theta \quad (45)$$

$j=N$ の場合 Eq. (23) で $y_i = x_i$, $a=1$, $b=0$ とおいた場合に相当するから, 補助定理 2・1 によりすべての Θ について ${}_ND=0$ である。ゆえに,

$${}_N\delta = 0 \quad (46)$$

Eq. (46) と補助定理 1 により,

$$E(NX_n) - N\mu = 0 \quad (47)$$

$j=L, A, W$ の場合 Eq. (25) で $y_i = x_i$, $w(y_i) = x_i^{a(j)}$ といった場合に相当し, $a(L)=1$, $a(A)=2$, $a(W)=3$ である。 $x_{\max} \neq x_{\min}$ ならば,

$$w(y_{\max}) - w(y_{\min}) = x_{\max}^{a(j)} - x_{\min}^{a(j)} > 0 \quad (48)$$

Eq. (48) と補助定理 2・3 により,

$$jD < 0, \quad x_{\max} \neq x_{\min} \quad (49)$$

ただしここに $x_{\max}$, $x_{\min}$ はそれぞれ θ における x の最大値と最小値であり, θ を構成する $(n+1)$ 個の粒子の粒径がすべて等しい場合にのみ $x_{\max} = x_{\min}$ となる。その場合には補助定理 2・2 により,

$$jD = 0, \quad x_{\max} = x_{\min} \quad (50)$$

Eq. (49) と Eq. (50) を Eq. (45) に代入すると $j\delta < 0$ を得る[†]。補助定理 1 により,

$$E(jX_n) - j\mu < 0, \quad j=L, A, W \quad (51)$$

$j=H$ の場合 Eq. (28) で $y_i = x_i$ とおき,

$$h(y_i) = h(x_i) = x_i^{-1} \quad (52)$$

とした場合に相当する。

$$h'(z) = -z^{-2} \quad (53)$$

よって $x_{\max} - x_{\min} \neq 0$ の場合には,

$$\{h(x_{\max}) - h(x_{\min})\} \{h'(x_{\max}) - h'(x_{\min})\} \\ = (x_{\max}^{-1} - x_{\min}^{-1})(x_{\min}^{-2} - x_{\max}^{-2}) < 0 \quad (54)$$

よって補助定理 2・4 により $H D > 0$ を得る。以下 $j=L, A, W$ の場合と同様の手順で,

$$H\delta > 0 \quad (55)$$

補助定理 1 により,

$$E(HX_n) - H\mu > 0 \quad (56)$$

$j=G$ の場合 Eq. (28) で $y_i = x_i$ とおき,

$$h(y_i) = h(x_i) = \ln x_i \quad (57)$$

とした場合に相当する。

$$h'(z) = z^{-1} \quad (58)$$

よって $x_{\max} \neq x_{\min}$ の場合には

$$\{h(x_{\max}) - h(x_{\min})\} \{h'(x_{\max}) - h'(x_{\min})\} \\ = (\ln x_{\max} - \ln x_{\min})(x_{\max}^{-1} - x_{\min}^{-1}) < 0 \quad (59)$$

以下, $j=H$ の場合と同様の手順で $G\delta > 0$, よって,

$$E(GX_n) - G\mu > 0 \quad (60)$$

$j=NA, NV$ の場合 Eq. (28) で $y_i = x_i$ とおき,

$$h(y_i) = h(x_i) = x_i^{a(j)} \quad (61)$$

とした場合に相当し, $a(NA)=2$, $a(NV)=3$ である。

$$h'(z) = a(j)z^{a(j)-1} \quad (62)$$

$$\{h(x_{\max}) - h(x_{\min})\} \{h'(x_{\max}) - h'(x_{\min})\} \\ = a(j)(x_{\max}^{a(j)} - x_{\min}^{a(j)})(x_{\max}^{a(j)-1} - x_{\min}^{a(j)-1}) > 0 \quad (63)$$

よって $j\delta < 0$ となる。補助定理 1 により,

$$E(jX_n) - j\mu < 0, \quad j=NA, NV \quad (64)$$

[†] 無限母集団におけるすべての x の値が等しい, すなわち均一粒径粒子からなる粉体という事実上無意味な場合を除くすれば, すべての θ について $x_{\max} = x_{\min}$ となるようなことは起こりえない。したがって $j\delta = 0$ となることはない。

$j=LA$ の場合 Eq. (31) で $y_i = x_i$ とおき,

$$w(y_i) = w(x_i) = x_i \quad (65)$$

$$h(y_i) = h(x_i) = x_i^2 \quad (66)$$

とした場合に相当し, Eq. (65) は Eq. (48) に, Eq. (66) は Eq. (61) に含まれている。ゆえに補助定理 2・5 の Eq. (33) が成立するので他の場合と同様の手順により,

$$E(LAX_n) - LA\mu < 0 \quad (67)$$

結 言

以上, よく用いられる標本平均径で母平均径の不偏推定量になっているのは個数平均径だけであり, それ以外には偏りのある推定量にすぎないことを証明した。この証明は母分布の形に関する仮定をおいていないという意味で一般的な証明であり, 証明の途中で近似を用いていないという意味で厳密な証明である。なお, 証明の手順から明らかなように,

$$E(jX_n) - E(jX_{n+1}) \begin{cases} > 0 & j=H, G \\ < 0 & j=L, A, W, NA, NV, LA \end{cases} \quad (68)$$

である。この式と Eq. (13) とから, 偏りの符号は計測粒子数 n によらないが偏りの絶対値は n の単調減少関数であることがわかる。したがって, 自動化されていない計数法粒度測定などで n が小さい場合には, ばらつきだけではなく偏りにも充分な注意を払うべきであろう。

Appendix

補助定理 2・1 の証明 Eqs. (20), (21), (23) を Eq. (22) に代入する。

$$D = \frac{1}{m(m-1)} \sum_{k=1}^m \sum_{i^*=1}^{m-1} (ay_{i^*} + b) - \frac{1}{m} \sum_{i=1}^m (ay_i + b) \quad (A-1)$$

$$\begin{aligned} & \sum_{k=1}^m \sum_{i^*=1}^{m-1} (ay_{i^*} + b) \\ &= \sum_{k=1}^m \left[\left\{ \sum_{i=1}^m (ay_i + b) \right\} - (ay_k + b) \right] \\ &= m \sum_{i=1}^m (ay_i + b) - \sum_{k=1}^m (ay_k + b) \\ &= (m-1) \sum_{i=1}^m (ay_i + b) \end{aligned} \quad (A-2)$$

これを Eq. (A-1) の右辺第 1 項に代入すれば,

$$\begin{aligned} D &= \frac{1}{m} \sum_{i=1}^m (ay_i + b) - \frac{1}{m} \sum_{i=1}^m (ay_i + b) \\ &= 0 \end{aligned} \quad (\text{証明終わり})$$

補助定理 2・2 の証明 $y_{\max} - y_{\min} = 0$ であるからすべての y の値が等しい。

$$y_1 = y_2 = \cdots = y_n = a \quad (A-3)$$

Eq. (A-3) を Eqs. (25), (28), または Eq. (31) に代入するといずれの場合にも $F=a$ となるから,

$$D = a - a = 0 \quad (\text{証明終わり})$$

補助定理 2・3 の証明 $y_{\max}-y_{\min} \neq 0$, すなわち $y_{\max} > y_{\min}$ の場合だけを考えればよい。 $F(m\phi) = \sum y_k w(y_k) / \sum w(y_k)$ であるから, $\sum y_k w(y_k) = \sum \{F(m\phi) w(y_k)\}$, よって,

$$\sum_{k=1}^m \left[\left\{ F(m\phi) - y_k \right\} w(y_k) \right] = 0 \quad (\text{A-4})$$

定義により $w(y) > 0$ であるから Eq. (A-4) が成立するには,

$$y_{\max} > F(m\phi) > y_{\min} \quad (\text{A-5})$$

でなければならない。次に,

$$W = \sum_{k=1}^m w(y_k) \quad (\text{A-6})$$

とおくことにすると,

$$F(m\phi) = \sum_{k=1}^m y_k w(y_k) / W \quad (\text{A-7})$$

$$F(m\phi_k) = \frac{F(m\phi) W - y_k w(y_k)}{W - w(y_k)} \quad (\text{A-8})$$

Eqs. (A-7), (A-8) を Eq. (22) に代入して整理すれば,

$$D = \sum_{k=1}^m \left[\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w(y_k)} \right] \quad (\text{A-9})$$

を得る。定義により $W > w(y_k) > 0$ であるから,

$$\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w(y_k)} \leq 0, \quad F(m\phi) \geq y_k \quad (\text{A-10})$$

よって D_+ を Eq. (A-11), D_- を Eq. (A-12) で定義すれば D_+ の右辺各項は 0 または正, D_- の右辺各項はすべて負である。

$$D_+ = \sum_{F(m\phi) \geq y_k} \left[\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w(y_k)} \right] \quad (\text{A-11})$$

$$D_- = \sum_{y_k > F(m\phi)} \left[\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w(y_k)} \right] \quad (\text{A-12})$$

明らかに,

$$D = D_+ + D_- \quad (\text{A-13})$$

以下まず $w(y_{\max}) > w(y_{\min})$, すなわち $w(y)$ が単調増加関数である場合を考える。

$$w(y_k) \geq w\{F(m\phi)\}, \quad y_k \leq F(m\phi)$$

$$\therefore \frac{1}{W - w(y_k)} \leq \frac{1}{W - w\{F(m\phi)\}}, \quad y_k \leq F(m\phi) \quad (\text{A-14})$$

Eq. (A-11) と Eq. (A-14) により Eq. (A-15), Eq. (A-12) と Eq. (A-14) により Eq. (A-16) を得る。

$$D_+ < \sum_{F(m\phi) \geq y_k} \left[\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w\{F(m\phi)\}} \right] \quad (\text{A-15})$$

$$D_- < \sum_{y_k > F(m\phi)} \left[\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w\{F(m\phi)\}} \right] \quad (\text{A-16})$$

Eq. (A-13) に Eq. (A-15) と Eq. (A-16) を代入して Eq. (A-4) を援用すれば,

$$D < \sum_{k=1}^m \left[\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w\{F(m\phi)\}} \right] = 0 \quad (\text{A-17})$$

次に $w(y_{\max}) < w(y_{\min})$ の場合を考えると, Eq. (A-14) の不等号, したがって Eq. (A-15), Eq. (A-16) の不等号も逆転し,

$$D > \sum_{k=1}^m \left[\frac{w(y_k) \{F(m\phi) - y_k\}}{W - w\{F(m\phi)\}} \right] = 0 \quad (\text{A-18})$$

補助定理 2・4 の証明 $y_{\max} - y_{\min} > 0$ である。 h は定義により単調増加関数であるが, 以下まず h が単調増加関数である場合について考える。

$$h(y_{\max}) - h(y_{\min}) > 0 \quad (\text{A-19})$$

$$h\{F(m\phi)\} = h \left[h^{-1} \left\{ \sum_{i=1}^m h(y_i) / m \right\} \right] = \sum_{i=1}^m h(y_i) / m \quad (\text{A-20})$$

$$\begin{aligned} h\{F(m\phi_k)\} &= h \left[h^{-1} \left\{ \sum_{i \neq k}^m h(y_i) / (m-1) \right\} \right] \\ &= \sum_{i \neq k}^m h(y_i) / (m-1) \\ &= h\{F(m\phi)\} + \delta_k^* \end{aligned} \quad (\text{A-21})$$

ただしここに,

$$\delta_k^* = [h\{F(m\phi)\} - h(y_k)] / (m-1) \quad (\text{A-22})$$

であって,

$$\sum_{k=1}^m \delta_k^* = \frac{1}{(m-1)} \left[m h\{F(m\phi)\} - \sum_{k=1}^m h(y_k) \right] = 0 \quad (\text{A-23})$$

であるから, δ_k^* には正のものと負のものとが必ず存在する。そうでなければ δ_k^* はすべて 0 でなければならず, $y_{\max} - y_{\min} > 0$ の下ではそのようなことは起こりえない。 Eq. (A-21) により,

$$F(m\phi_k) = h^{-1} [h\{F(m\phi)\} + \delta_k^*] \quad (\text{A-24})$$

$$d_k = F(m\phi_k) - F(m\phi) \quad (\text{A-25})$$

で d_k を定義すれば,

$$D = \sum_{k=1}^m d_k / m \quad (\text{A-26})$$

$$d_k = h^{-1} [h\{F(m\phi)\} + \delta_k^*] - h^{-1} [h\{F(m\phi)\}] \quad (\text{A-27})$$

$$d_k = a_k \delta_k^* \quad (\text{A-28})$$

とおけば, $0 \leq b_k \leq 1$ の範囲で b_k を適当に選ぶことによって,

$$a_k = dh^{-1}(z) / dz, \quad z = h\{F(m\phi) + b_k \delta_k^*\} \quad (\text{A-29})$$

が成立する。次に Eq. (A-30) で A を定義する。

$$A = \left\{ \frac{dh^{-1}(z)}{dz} \right\}_{z=h\{F(m\phi)\}} \quad (\text{A-30})$$

h が単調増加関数である場合を考えているのであるから, h^{-1} も単調増加関数である。よって,

$$a_k > 0, \quad A > 0 \quad (\text{A-31})$$

h は定義により変曲点の無い非線型関数であるから上に凸か下に凸かのいずれかであり, h が下に凸ならば h^{-1} は上に凸, h が上に凸ならば h^{-1} は下に凸である。まず h が下に凸,

$$h'(y_{\max}) - h'(y_{\min}) > 0 \quad (\text{A-32})$$

の場合を考えると, $dh^{-1}(z)/dz$ が z の単調減少関数になるので,

$$a_k (\delta_k^* < 0) > A \geq a_k (\delta_k^* \geq 0) \quad (\text{A-33})$$

Eq. (A-34) で D_+ と D_- を定義する。

$$D_+ = \sum_{\delta_k^* \geq 0} d_k, \quad D_- = \sum_{\delta_k^* < 0} d_k \quad (\text{A-34})$$

$$D = (D_+ + D_-)/m \quad (\text{A-35})$$

$$D_+ = \sum_{\delta_k^* \geq 0} a_k \delta_k^* < \sum_{\delta_k^* \geq 0} A \delta_k^* = A \sum_{\delta_k^* \geq 0} \delta_k^* \quad (\text{A-36})$$

$$D_- = \sum_{\delta_k^* < 0} a_k \delta_k^* < \sum_{\delta_k^* < 0} A \delta_k^* = A \sum_{\delta_k^* < 0} \delta_k^* \quad (\text{A-37})$$

Eq. (A-35) に Eq. (A-36) と Eq. (A-37) を代入し Eq. (A-23) を援用すれば,

$$D = (D_+ + D_-)/m < \frac{A}{m} \sum_{k=1}^m \delta_k^* = 0 \quad (\text{A-38})$$

次に h が上に凸,

$$h'(y_{\max}) - h'(y_{\min}) < 0 \quad (\text{A-39})$$

の場合を考えると, $dh^{-1}(z)/dz$ が z の単調増加関数になるので Eq. (A-33) の不等号が逆転する。その結果として Eqs. (A-36), (A-37) の不等号も逆転するから,

$$D = (D_+ + D_-)/m > \frac{A}{m} \sum_{k=1}^m \delta_k^* = 0 \quad (\text{A-40})$$

以上で h が単調増加関数である場合についての証明が終わった。

h が単調減少関数である場合についても同様の手順で証明することができるが, 簡単に証明するには Eq. (A-41) で定義される h_* を考えればよい。

$$h_*(z) = -h(z) \quad (\text{A-41})$$

h が定義を満足するならば h_* も h の定義を満足することは明らかであり, Eq. (A-25) が成立するならば,

$$F = h_*^{-1} \left\{ \sum_{i=1}^l h_*(y_i) / l \right\} \quad (\text{A-42})$$

が成立することも明らかである。そして, h が単調減少関数であるときは h_* は単調増加関数である。よって,

$$D > 0, \{h_*(y_{\max}) - h_*(y_{\min})\} \{h_*'(y_{\max}) - h_*'(y_{\min})\} < 0 \quad (\text{A-43})$$

$$D < 0, \{h_*(y_{\max}) - h_*(y_{\min})\} \{h_*'(y_{\max}) - h_*'(y_{\min})\} > 0 \quad (\text{A-44})$$

はすでに証明されている。 h_* の定義により明らかに,

$$\{h_*(y_{\max}) - h_*(y_{\min})\} \{h_*(y_{\max}) - h_*(y_{\min})\} < 0 \quad (\text{A-45})$$

$$\{h_*'(y_{\max}) - h_*'(y_{\min})\} \{h_*'(y_{\max}) - h_*'(y_{\min})\} < 0 \quad (\text{A-46})$$

であるから, $\{h_*(y_{\max}) - h_*(y_{\min})\} \{h_*'(y_{\max}) - h_*'(y_{\min})\}$ の符号と, $\{h(y_{\max}) - h(y_{\min})\} \{h'(y_{\max}) - h'(y_{\min})\}$ の符号は互いに等しい。以上で h が単調減少関数である場合についての証明が終わった。

補助定理 2・5 の証明 $y_{\max} - y_{\min} > 0$ である。 h は定義により単調関数であるが, 以下まず h が単調増加関数である場合について考える。

$$h(y_{\max}) - h(y_{\min}) > 0 \quad (\text{A-47})$$

B, δ_k^*, d_k, D^* を次のように定義する。

$$B = \frac{1}{m} \sum_{k=1}^m h\{F(m\varphi_k)\} \quad (\text{A-48})$$

$$\delta_k^* = h\{F(m\varphi_k)\} - B \quad (\text{A-49})$$

$$d_k = F(m\varphi_k) - h^{-1}(B) \quad (\text{A-50})$$

$$D^* = \frac{1}{m} \sum_{k=1}^m d_k \quad (\text{A-51})$$

明らかに,

$$\sum_{k=1}^m \delta_k^* = 0 \quad (\text{A-52})$$

$$D = D^* + h^{-1}(B) - F(m\Phi) \quad (\text{A-53})$$

$$F(m\varphi_k) = h^{-1}[h\{F(m\varphi_k)\}] = h^{-1}(B + \delta_k^*) \quad (\text{A-54})$$

$$d_k = h^{-1}(B + \delta_k^*) - h^{-1}(B) \quad (\text{A-55})$$

であるから, 以下, 補助定理 2・4 の証明における Eq.

(A-28) から Eq. (A-40) に至る手順と同様にして,

$$D^* \leq 0, \quad h'(y_{\max}) - h'(y_{\min}) \leq 0 \quad (\text{A-56})$$

を得る。Eqs. (A-47), (A-53), (A-56) より,

$$D \leq h^{-1}(B) - F(m\Phi), \quad \{h(y_{\max}) - h(y_{\min})\} \{h'(y_{\max}) - h'(y_{\min})\} \leq 0 \quad (\text{A-57})$$

一方,

$$y^* = h(y), \quad w^*(y^*) = w(y) \quad (\text{A-58})$$

において Eq. (31) に代入してみると,

$$F = h^{-1} \left\{ \sum_{i=1}^l y_i^* w^*(y_i^*) / \sum_{i=1}^l w^*(y_i^*) \right\} \quad (\text{A-59})$$

を得る。

$$F(m\Phi) = h^{-1} \left\{ \sum_{k=1}^m y_k^* w^*(y_k^*) / \sum_{k=1}^m w^*(y_k^*) \right\} \quad (\text{A-60})$$

$$F(m\varphi_k) = h^{-1} \left\{ \sum_{i \neq k}^m y_i^* w^*(y_i^*) / \sum_{i \neq k}^m w^*(y_i^*) \right\} \quad (\text{A-61})$$

Eq. (A-48) に Eq. (A-61) を代入すれば,

$$B = \frac{1}{m} \sum_{k=1}^m \left\{ \sum_{i \neq k}^m y_i^* w^*(y_i^*) / \sum_{i \neq k}^m w^*(y_i^*) \right\} \quad (\text{A-62})$$

すでに証明された補助定理 2・3 により,

$$B - h\{F(m\Phi)\} \leq 0, \quad w^*(y_{\max}^*) - w^*(y_{\min}^*) \leq 0 \quad (\text{A-63})$$

h が単調増加関数であるから h^{-1} も単調増加関数である。

$$h^{-1}(B) - F(m\Phi) \geq 0, \quad w^*(y_{\max}^*) - w^*(y_{\min}^*) \leq 0 \quad (\text{A-64})$$

h が単調増加関数であることと, Eq. (A-58) とにより,

$$\{w(y_{\max}) - w(y_{\min})\} \{w^*(y_{\max}^*) - w^*(y_{\min}^*)\} > 0 \quad (\text{A-65})$$

Eq. (A-64) と Eq. (A-65) により,

$$h^{-1}(B) - F(m\Phi) \geq 0, \quad w(y_{\max}) - w(y_{\min}) \leq 0 \quad (\text{A-66})$$

Eq. (A-57) と Eq. (A-66) とにより,

$$D \geq 0, \quad \{h(y_{\max}) - h(y_{\min})\} \{h'(y_{\max}) - h'(y_{\min})\} \geq 0 \quad \& \quad w(y_{\max}) - w(y_{\min}) \geq 0$$

以上で h が単調増加関数である場合についての証明が終わった。

次は h が単調減少関数である場合の証明である。

$$h_*(y) = -h(y) \quad (\text{A-67})$$

とおけば $h_*(y)$ は単調増加関数であり,

$$\begin{aligned} F &= h^{-1} \left\{ \frac{\sum_{i=1}^l h(y_i) w(y_i)}{\sum_{i=1}^l w(y_i)} \right\} \\ &= h_*^{-1} \left\{ \frac{\sum_{i=1}^l h_*(y_i) w(y_i)}{\sum_{i=1}^l w(y_i)} \right\} \quad (\text{A-68}) \end{aligned}$$

以下補助定理 2・4 の証明と同様の論法で証明を完結することができる。

Nomenclature

$E(g)$	= expectation of statistic, g	
(x)	= number frequency distribution function of x	
	(population value)	$[\mu^{-1}]$
$f_n^*(x)$	= number frequency distribution function of x	
	(sample value)	$[\mu^{-1}]$
n	= sample size	$[-]$
X_n	= mean particle diameter (sample value)	$[\mu]$
x	= particle diameter	$[\mu]$
μ	= mean particle diameter (population value)	$[\mu]$

Literature cited

- 1) Hollander, M. and D. A. Wolfe: "Nonparametric Statistical Methods", p.1, John Wiley & Sons, Inc. (1973)
- 2) Kondo, Y. and W. Funasaka: "Tōkei-teki Hōhō", p. 79, Kyōritsu Syuppan (1967)
- 3) Masuda, H. and K. Inoya: *J. Chem. Eng. Japan*, **4**, 60 (1971)
- 4) Mori, Y. and A. Suganuma: *Kagaku Kōgaku*, **32**, 1099 (1968)
- 5) Mori, Y. and A. Suganuma: *ibid.*, **34**, 445 (1970)
- 6) Yamamoto, H., A. Suganuma and Y. Mori: *Kagaku Kōgaku Ronbunshu*, **1**, 226 (1975)
- 7) Mori, Y. and A. Suganuma: *ibid.*, **1**, 197 (1975)
- 8) Mori, Y. and A. Suganuma: Preprints of the 7th Autumn Meeting of The Soc. of Chem. Engrs., Japan, Nagoya (1973)
- 9) Suganuma, A.: *Funsai (The Micromeritics, Japan)*, **18**, 162 (1973)
- 10) Van der Waerden, B. L.: "Mathematische Statistik", p. 148, Springer-Verlag (1965)
- 11) Wilks, S. S.: "Mathematical Statistics", p. 103, John Wiley & Sons, Inc. (1962)
- 12) Wilks, S. S.: *ibid.*, p. 339 and p. 473 (1962)

(1974年2月8日受理)

A Theoretical Proof of Statistical Bias of Miscellaneous Sample Mean Particle Diameters by Counting Methods

Yoshiro Mori and Akira Suganuma

Dept. of Chem. Eng., Faculty of Eng., Univ. of Tokyo, Tokyo

Without any assumption of particle size distribution function, the following arguments are generally and rigorously proven.

1) Although a sample value of number mean diameter, ${}_nX_n$ is an unbiased estimate of its population value, ${}_n\mu$, sample values of many other mean diameters, ${}_jX_n$ ($j \neq N$) are, respectively, biased estimates of their population values, ${}_j\mu$.

2) Sign of bias depends neither on particle size distribution of population nor on sample size, n . Bias of sample geometric mean diameter and that of sample harmonic mean diameter are always positive. On the other hand, many other sample mean diameters, such as $\sum x^4 / \sum x^3$, $\sum x^3 / \sum x^2$, $\sum x^2 / \sum x$, $(\sum x^3 / n)^{1/3}$, $(\sum x^2 / n)^{1/2}$ and $(\sum x^3 / \sum x)^{1/2}$ are always biased negatively.

3) The smaller the sample size, n , the larger becomes the absolute value of statistical bias.

~~~~~