U.D.C. 519.2

移動範囲による標準偏差の推定について

A Note on Estimating the Standard Deviation by Moving Range

島　田　正　三*
Shozo Simada

内　容　梗　概

本報は，従来標準偏差の推定にしばしば用いられてきた移動範囲による標準偏差の推定量精度について論じたものである。移動範囲による推定とは測定値 x_1, x_2, x_3, ……, x_n に対し，まず

x_1, x_2, ……, x_k

についての範囲 R_1 を求め，ついで，測定値を1個ずらせて

x_2, x_3, ……, x_{k+1}

についての範囲 R_2 を求める。同様にして R_1, R_2, R_3……, R_{n-k+1} を求めて，その平均 $\overline{R}$ をとり，これに適当な修正係数を乗じたものをもつて σ の推定値としようという方法である。

このような場合，R_1 と R_2 との算出には，測定値 x_2, x_3, ……x_k が共通に使われる。そのため R_1 と R_2 とは，統計的に独立でない。同様に R_1 と $R_i(i=2,\ 3,……k)$ が独立でなくなる。本報はまず R_1 と R_i との相関を求め，さらにこれを用いて推定量の分散を求めたものである。なお k としては 2～5 の4通りについて行つた。精度は k を2から3にすると大分向上するがさらに4，5としてもあまり増さない。$k=3$ 程度がよいのではないかと思われる。

〔I〕緒　　言

一定の間隔をおいて測定したデータ x_1, x_2, x_3, ……, x_t, …… の標準偏差を求める場合に，よく移動範囲が用いられる。本報は，移動範囲から標準偏差を推定した場合の分散を求めたものである。

同一母集団から抽出した大きさ n の標本について，最小値および最大値を求め，それをそれぞれ $x_{(1)}$, $x_{(n)}$ にて表わすものとする。範囲

$$R\equiv x_{(n)}-x_{(1)} \qquad (1)$$

を用いて，母集団の標準偏差を推定する場合には，R を母集団の分布，および標本の大きさ n に関係するある常数 d_n で割つてやればよい。つまり σ の推定値 $\hat{\sigma}$ は次式によつて与えられる。

$$\hat{\sigma}=R/d_n \qquad (2)$$

母集団の分布が正規分布にしたがう場合の d_n の値は，つとに計算され広く実用に供せられている。いうまでもなく，d_n の値は，(2) 式の $\hat{\sigma}$ が不偏推定量となるように定められた修正係数である。

$$E(R)=d_n\sigma \qquad (3)$$

さらに，標本の大きさが10ないしそれ以下の場合は，(2) 式から求めた推定量の効率は

$$s\equiv\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2/(n-1)}\text{，ここに}\ \overline{x}\equiv\Big(\sum_{i=1}^{n}x_i\Big)/n$$

にある修正係数をかけて求めた推定量に比して，あまり劣つていないことがわかつている。しかし，大きさがかなり大きな場合には，その効率は非常に悪いものとなる。その一つの対策としては，全測定値をいくつかの組に分け，各組についての範囲を求め，さらにそれを平均したもの（平均範囲）を用いて σ を推定するという方法が考えられている。この際推定の効率をよくするには，組の大きさを7ないし8にとるのがもつともよいことが証明されている。

あと一つの対策としては，移動範囲を用いる方法が考えられる。移動範囲による方法とはまず最初の k 個の測定値

x_1, x_2, ……, x_k

について，範囲 R_1 を求め，ついで

x_2, x_3, ……, x_{k+1}

の範囲 R_2 を求める。同様に

x_i, x_{i+1}, ……, x_{i+k-1}

の範囲 $R_i(i=1,\ 2,\ 3,……,\ n-k+1)$ を求め，これらの平均

$$\overline{R}=\Big(\sum_{i=1}^{n-k+1}R_i\Big)/(n-k+1) \qquad (4)$$

を d_k で割つたものをもつて，σ の推定量としようという方法である。容易にわかるように，R_i と R_{i+1} との算出には，$k-1$ 個の測定値 x_{i+1}, x_{i+2}, ……, x_{i+k-1} が共通に入り込んでくるため，両者は独立ではない。同様に R_i と $R_{i+j}(j=1,\ 2,……,\ k-1)$ とは独立とならない。本報は，x が正規分布にしたがう場合 $k=2,\ 3,\ 4,\ 5$ の4通りについて R_i と R_{i+j} との相関を求め，ついで移動範囲による σ の推定量の分散を求めたものである。なお k のことを範囲の大きさと呼ぶこととする。

〔II〕二つの範囲の間の相関係数

前述のごとく，範囲の大きさが k の場合，R_1 と R_{1+s}

* 日立製作所中央研究所

とは，s が $k-1$ ないし，それより小さい場合には，$k-s$ 個の測定値を共有しているため，たがいに独立ではない。そこでまず最初に $R_1 \cdot R_{1+s}$ の期待値を計算する。

なお本節では，母集団分布は母平均0，分散1の正規分布にしたがうものと仮定する。

R_1 と R_{1+s} に共通に使われる測定値

$$x_{s+1},\ x_{s+2}, \ldots\ldots,\ x_k\ (s<k-1 \text{ とする})$$

の内の最大値を y，最小値を z にて表わす，また R_1 の算出に用いられた k 個の測定値のうち

$$x_1,\ x_2 \ldots\ldots, x_s$$

の s 個は R_{1+s} の算出には使われない。この s 個の測定値のうちの最大値，および最小値をそれぞれ y_1, z_1 とすれば，y, z と y_1, z_1 との大小関係には，次の四つのケースが考えられる。

Case 1：　s 個全部が z より大きく，さらにそのうちのいくつかは y より大きい場合 $(y_1>y,\ z_1>z)$

Case 2：　s 個全部が y と z との間にある場合 $(y_1<y,\ z_1>z)$

Case 3：　s 個全部が y よりは小さく，さらにそのうちのいくつかは z より小さい場合 $(y_1<y,\ z_1<z)$

Case 4：　s 個のうちいくつかは y より大きく，さらにいくつかは z より小さい場合 $(y_1>y,\ z_1<z)$

Case 1 の場合には，範囲は y_1-z によつて与えられ，さらに最大値が y_1 に等しく，最小値が z よりも小でない確率は

$$\frac{s!}{(s-1)!}\phi(y_1)\Big\{\Phi(y_1)-\Phi(z)\Big\}^{s-1}$$

Case 2 の場合には，範囲は $y-z$ によつて与えられ，さらに s 個全部が y, z の間にある確率は

$$\Big\{\Phi(y)-\Phi(z)\Big\}^{s}$$

Case 3 の場合には，範囲は $y-z_1$，によつて与えられ，さらに最小値 z_1 に等しく，最大値が y より小さい確率は

$$\frac{s!}{(s-1)!}\phi(z_1)\Big\{\Phi(y)-\Phi(z_1)\Big\}^{s-1}$$

Case 4 の場合には，範囲は y_1-z_1 によつて与えられ，さらに最大値が $y_1(>y)$ に等しく，最小値が $z_1(<z)$ に等しい確率は

$$\frac{s!}{(s-2)!}\phi(y_1)\phi(z_1)\Big\{\Phi(y_1)-\Phi(z_1)\Big\}^{s-2}$$

ここに

$$\phi(u)\equiv\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$$

$$\Phi(u)\equiv\int_{-\infty}^{u}\phi(t)\,dt$$

まつたく同様のことが，R_{1+s} の算出に用いた k 個のうち，R_1 と共通でない s 個の測定値

$$x_{k+1},\ x_{k+2}, \ldots\ldots,\ x_{k+s}$$

についても成立する。両者を区別するために，さきに Case 1～4 で定義した y_1, z_1 と同様の量を，プライムを附して，それぞれ y_1', z_1' にて表わすものと約束する。さらに R_1 の Case 1～4 に相当するものを，それぞれ Case 1′, 2′, 3′, 4′ を表わすこととする。

このようにすると，全部で $4^2=16$ の組合せが存在する。これらを Case $(i,\ j')$ のように表わすこととしよう。たとえば Case (4, 3′) とは x_1, $x_2, \ldots\ldots, x_s$ の最大値 y_1 は y より大，また最小値 z_1 は z より小であり，また x_{k+1}, $x_{k+2}, \ldots\ldots, x_{k+s}$ は全部 y より小さく，さらに最小値 z_1' は z より小さい場合に相当する。この場合の $R_1 \cdot R_{1+s}$ は

$$R_1 \cdot R_{1+s}=(y_1-z_1)(y-z_1')$$

に等しい。

なお，R_1 と R_{1+s} とに共通の $k-s$ 個の測定値の最大値，最小値がそれぞれ y, z に等しい確率密度は

$$\frac{(k-s)!}{(k-s-2)!}\phi(y)\phi(z)\Big\{\Phi(y)-\Phi(z)\Big\}^{k-s-2}$$

に等しいから，Case (4, 3′) の $E(R_1,\ R_{1+s})$ に対する寄与は，次のごとく表わされる

$$\int_{-\infty}^{\infty}dy\int_{-\infty}^{y}dz\,\frac{(k-s)!}{(k-s-2)!}\phi(y)\phi(z)\Big\{\Phi(y)-\Phi(z)\Big\}^{k-s-2}\cdot$$

$$\times\int_{-\infty}^{z}dz_1'\int_{y}^{\infty}dy_1\int_{-\infty}^{z}dz_1(y_1-z_1)(y-z_1')$$

$$\times\frac{s!}{(s-2)!}\phi(y_1)\phi(z_1)\Big\{\Phi(y_1)-\Phi(z_1)\Big\}^{s-2}\cdot$$

$$\times\frac{s!}{(s-1)!}\phi(z_1')\Big\{\Phi(y)-\Phi(z_1')\Big\}^{s-1}$$

$$=\int_{-\infty}^{\infty}dy\int_{-\infty}^{y}dz\,\frac{(k-s)!}{(k-s-2)!}\phi(y)\phi(z)\Big\{\Phi(y)-\Phi(z)\Big\}^{k-s-2}$$

$$\times\Bigg[\int_{y}^{\infty}dy_1\int_{-\infty}^{z}dz_1\frac{s!}{(s-2)!}(y_1-z_1)\phi(y_1)\phi(z_1)$$

$$\times\Big\{\Phi(y_1)-\Phi(z_1)\Big\}^{s-2}\Bigg]$$

$$\times\Bigg[\int_{-\infty}^{z}dz_1'\frac{s!}{(s-1)!}(y-z_1')\phi(z_1')$$

$$\times\Big\{\Phi(y)-\Phi(z_1')\Big\}^{s-1}\Bigg]\qquad\cdots\cdots(5)$$

つまり，上式の最初の〔　〕は x_1, $x_2, \ldots\ldots, x_s$ に関するものであり，第2の〔　〕は x_{k+1}, $x_{k+2}, \ldots\ldots, x_{k+s}$ に関するものである。上のように，x_1, $x_2, \ldots\ldots, x_s$ と x_{k+1}, $x_{k+2}, \ldots\ldots, x_{k+s}$ との部分に因数分解される。した

がつて16個の総ての組合せについてこれを加え合せると，$R_1 \cdot R_{1+s}$ の期待値は次式で与えられることがわかる。

$$E\,(R_1 \cdot R_{1+s}) = \frac{(k-s)!}{(k-s-2)!} \int_{-\infty}^{\infty} dy \int_{-\infty}^{y} dz \phi(y) \phi(z) \times \left\{ \Phi(y) - \Phi(z) \right\}^{k-s-2} I^2 \cdots\cdots\cdots\cdots (6)$$

ここに

$$\begin{aligned} I = & \int_{y}^{\infty} \frac{s!}{(s-1)!} (y_1 - z) \phi(y_1) \left\{ \Phi(y_1) - \Phi(z) \right\}^{s-1} dy_1 \\ & + (y-z) \left\{ \Phi(y) - \Phi(z) \right\}^{s} \\ & + \frac{s!}{(s-1)!} \int_{-\infty}^{z} (y - z_1) \phi(z_1) \left\{ \Phi(y) - \Phi(z_1) \right\}^{s-1} \\ & + \frac{s!}{(s-2)!} \int_{y}^{\infty} dy_1 \int_{-\infty}^{z} dz_1 (y_1 - z_1) \phi(y_1) \phi(z_1) \\ & \times \left\{ \Phi(y_1) - \Phi(z_1) \right\}^{s-2} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (7) \end{aligned}$$

なお，$s=k-1$ の場合，つまり R_1 と R_{1+s} とに共通の測定値が1個の場合には，(7) 式は

$$E(R_1 \cdot R_k) = \int_{-\infty}^{\infty} dx_{(k)} \phi(x_{(k)}) I^2$$

によつて与えられる。計算の結果は次の通りである。ただし簡単のために

$$\int_{-\infty}^{\infty} \phi^i(x) \Phi^i(x) dx \equiv (i,\ j)$$

$$\phi(x) \equiv \phi_x, \quad \Phi(x) \equiv \Phi_x$$

とかくこととする。

$k=2$:

$$E(R_1 \cdot R_2) = \frac{1}{3} + \frac{2\sqrt{3}}{\pi}$$

$k=3$:

$$E(R_1 R_2) = 1 + \left(2 + \frac{9}{\sqrt{3}}\right) \frac{1}{\pi}$$

$$E(R_1 R_3) = \frac{1}{3} + \left(5 + \frac{6}{\sqrt{3}}\right) \frac{1}{\pi}$$

$k=4$:

$$\begin{aligned} E(R_1 R_2) = & \frac{6}{5} + (\sqrt{3} - 6) \frac{1}{\pi} + 48(3,2) \\ & + 144 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x dx \int_{-\infty}^{x} \phi_t{}^2 d_t \end{aligned}$$

$$\begin{aligned} E(R_1 R_3) = & \frac{2}{3} - \left(\frac{8}{\sqrt{3}} + \frac{39}{2} \right) \frac{1}{\pi} + 112(2,2)/\sqrt{\pi} \\ & + 120(3,2) - 16(3,3) \\ & + 304 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x dx \int^{x} \phi_t{}^2 dt - 256 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 dt \end{aligned}$$

$$\begin{aligned} E(R_1 R_4) = & \frac{19}{70} - \left(\frac{21}{\sqrt{3}} + \frac{9}{4} \right) \frac{1}{\pi} \\ & + \frac{180(2,2)}{\sqrt{\pi}} - \frac{90(2,3)}{\sqrt{\pi}} \\ & + 306(3,2) - 276(3,3) + 108(3,4) \\ & - 180 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x dx \int^{x} \phi_t{}^2 dt \\ & + 396 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 dt \\ & - 576 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 \Phi_t dt \end{aligned}$$

$k=5$:

$$\begin{aligned} E(R_1 R_2) = & \frac{4}{3} + 60(3,0) - 180(3,1) + 180(3,2) \\ & - 216(2,1)^2 + 192(2,0)(2,2) \\ & - 96(2,0)(2,1) + 240 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 \Phi_t dt \end{aligned}$$

$$\begin{aligned} E(R_1 R_3) = & \frac{6}{7} + 75(3,0) - 300(3,1) + 462(3,2) \\ & - 324(3,3) + 162(3,4) \\ & - 216(2,1)^2 + 432(2,1)(2,2) + 40(2,0)^2 \\ & - 24(2,0)(2,1) + 96(2,0)(2,2) \\ & - 144 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x dx \int^{x} \phi_t{}^2 dt \\ & + 432 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 dt \\ & - 864 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 \Phi_t dt \end{aligned}$$

$$\begin{aligned} E(R_1 R_4) = & \frac{1}{2} + 72(3,0) - 360(3,1) \\ & + 738(3,2) - 756(3,3) \\ & - 378(3,4) + 120(2,0)^2 - 240(2,0)(2,1) \\ & + 252(2,1)^2 \\ & + 432(2,1)(2,2) + 288(2,2)^2 \\ & - 456 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x dx \int^{x} \phi_t{}^2 dt \\ & + 1008 \int_{-\infty}^{\infty} \phi_x{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 dt \\ & - 2016 \int_{-\infty}^{\infty} \phi_t{}^2 \Phi_x{}^2 dx \int^{x} \phi_t{}^2 \Phi_t dt \end{aligned}$$

$$\begin{aligned} E(R_1 R_5) = & \frac{23}{105} + 48(3,0) - 288(3,1) \\ & + 720(3,2) - 864(3,3) \\ & + 432(3,4) + 240(2,0)^2 + 1440(2,1)^2 \end{aligned}$$

$$-960(2,0)(2,1)$$
$$+1296(2,2)^2+480(2,0)(2,2)$$
$$-1440(2,1)(2,2)$$
$$-480\int_{-\infty}^{\infty}\phi_x{}^2\Phi_x dx\int^{x}\phi_t dt$$
$$+1152\int_{-\infty}^{\infty}\phi_x{}^2\Phi_x{}^2 dx\int^{x}\phi_t{}^2 dt$$
$$-2304\int_{-\infty}^{\infty}\phi_x{}^2\Phi_x{}^2 dx\int^{x}\phi_t{}^2\Phi_t dt$$

$(i,\ j)$ および

$$\int_{-\infty}^{\infty}\phi_x{}^2\Phi_x dx\int^{x}\phi_t{}^2 dt,\qquad \int_{-\infty}^{\infty}\phi_x{}^2\Phi_x{}^2 dx\int^{x}\phi_t{}^2 dt,$$
$$\int_{-\infty}^{\infty}\phi_x{}^2\Phi_x{}^2 dx\int^{x}\phi_t{}^2\Phi_t dt$$

などの数値計算は，後の機会にゆずり，本報では述べないこととする。その結果を用いて，R_1 と R_i $(i=1, 2\ldots k)$ の共分散を求めると第1表のようになる。

第1表　$k=2, 3, 4, 5$ の場合の R_1 と R_i $(i=1, 2, \ldots k)$ との共分散 Cov $(R_1,\ R_i)$

i	$k=2$	$k=3$	$k=4$	$k=5$
1	0.72676	0.78920	0.77406	0.74662
2	0.16275	0.42582	0.51479	0.54899
3		0.16275	0.30517	0.41832
4			0.16425	0.23512
5				0.10875

〔III〕 移動範囲を用いる σ の推定

本節では，移動範囲を用いた σ の推定量と，平均範囲を用いた推定量との分散の比較を行う。

両者とも，範囲の大きさを k，標本の大きさを n とする。さらに，簡単のために，$k \ll n$ としよう。移動範囲による推定量を σ_1 平均範囲によるものを $\hat{\sigma}_2$ として区別する。

$$\mathrm{var}(\hat{\sigma}_1) \fallingdotseq \frac{n\,\mathrm{var}(R)+2n\sum_{i=2}^{k}\mathrm{cov}(R_1, R_i)}{\{nE(R)\}^2}$$

$$\mathrm{var}(\hat{\sigma}_2) \fallingdotseq \frac{\mathrm{var}(R)}{\dfrac{n}{k}\left\{E(R)\right\}^2}$$

$k=2,\ 3,\ 4,\ 5$ の場合について，両者を求めてみると

	$k=2$	$k=3$	$k=4$	$k=5$
$n\mathrm{var}(\hat{\sigma}_1)$	0.8265	0.6864	0.6328	0.6227
$n\mathrm{var}(\hat{\sigma}_2)$	1.1415	0.8265	0.7153	0.6900
$\mathrm{var}(\hat{\sigma}_1)/\mathrm{var}(\hat{\sigma}_2)$	0.724	0.830	0.885	0.902

つまり，$k=2, 3$ くらいまでのところは，$\hat{\sigma}_1$ の方が分散が小さいが，$k=5$ 程度となるが，それほど大きな違いは生じない。k が大きくなると，平均範囲にくらべて移動範囲の計算は，かなり繁雑なものとなろう。したがつて，一般には k の大きな場合，移動範囲によつて σ を推定することは，あまり好ましないということができよう。

〔IV〕 結　　言

従来，標準偏差 σ の推定法としては，いわゆる不偏分散 s^2 によることが多く，範囲による推定が推奨されながらも，なにか効率の点で損をしたような気分が抜け切れぬためか，管理図など一部を除いては，あまり用いられなかつた。しかし，標本の大きさが大きな場合でも，それを適当な組に分け，各組の範囲から平均範囲 $\overline{R}$ を求めるとか，または本報の移動範囲を用いることによつて，精度の損失は大巾に縮少させることができる。工場の現場においては，拙速を尊ぶばかりでなく，現場技術者は計算にあまりなじんでいない。そのため，s による推定は計算違いを起しやすい。このような点を考慮に入れると，s による推定は，現場ではあまり適したものといえないと思う。

なお，移動範囲をとるべき組の大きさ k について考えると，本報では $k=2,\ 3,\ 4,\ 5$ までしか計算していないから，これ以上の場合についてはあきらかでないが，k を2から3に変えると推定の精度が大分向上するに対しそれ以上 4，5 と増していつても，その割には向上しない。この点から見て，おそらく k を6ないしそれ以上にしても，大巾な精度の向上は望めまい。一方 k が大きくなると，移動範囲の計算はかなり厄介となる。それ故 k としては，3くらいをとるのがよいのではないかと思われる。

なお，測定値 $x_1,\ x_2, \ldots\ldots$ が時間の経過につれて求められたものであり，個々の測定値がとり出された母集団の母平均が，必ずしも一定でない場合には σ の推定量の中には偏りが入り込んでくる。これについては，本報では述べないが，結論は k が小さい程偏りは小さくなるのである。たとえば（$k=2$ の場合を，s に含まれる偏りの大きさと比較すると，範囲による推定量の方が桁ちがいに小さい）したがつて，このような場合には，ことさら k は小さくすることが望まれる。

終りに，本研究を行うに当り，種々御激励を賜つた日立製作所中央研究所菊田所長，浜田副所長，宮城部長，ならびに種々御指導御鞭撻を賜つた高田主任研究員に衷心より感謝の意を表するものである。