2009年5月20日作成
2011年4月18日修正

標本分散の誤差の評価

Estimation of errors in sample variance

井田　隆
名古屋工業大学セラミックス基盤工学研究センター

はじめに

一般的に，測定により得られる強度データは，系統的な誤差 systematic error と，統計的な誤差 statistical error とを含んでいます。さらに，粉末回折強度が計数法により測定される場合には，そのデータが含む統計的な誤差としては，「計数統計誤差」と「粒子統計誤差」が主なものと考えられています。

- 粉末回折強度の誤差
  - 系統的な誤差
  - 統計的な誤差
    - 計数統計誤差
    - 粒子統計誤差

このノートでは，これら統計的な誤差について実験的な評価を行う研究のために必要であった数学的な形式の導出のしかたを示します。導かれる形式は，粉末回折測定に限らず広い範囲で適用可能と思われます。

このような方法で導かれた数式を，はじめにX線検出器の数え落としの影響を受けた計数統計誤差を実験的に評価する研究 [T. Ida. *J. Appl. Cryst*. **41**(6), 1019-1023 (2008)] で用い，さらに粒子統計誤差を実験的に評価する研究 [T. Ida, T. Goto & H. Hibino. *J. Appl. Cryst*. **42**, 597-606 (2009)] で用いました。

任意の n 個の標本値 $\{X_j\}$ ($j = 0, ..., n - 1$) に対して，標本平均が

$$\bar{X} \equiv \frac{1}{n}\sum_{j=0}^{n-1} X_j$$

で計算され，標本分散は

$$V = \frac{1}{n-1}\sum_{j=0}^{n-1}\left(X_j - \bar{X}\right)^2$$

で計算されること，標本平均の分散は

$$\left\langle \left(\bar{X} - m\right)^2 \right\rangle = \frac{V}{n}$$

標本分散の分散は，

$$\left\langle (V-\sigma)^2 \right\rangle = \frac{1}{(n-1)(n^2-3n+3)}\left[n\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^4 - \left(n^2-3\right)V^2 \right]$$

$$\sim \frac{1}{n^2}\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^4 - \frac{V^2}{n}$$

という式で，実験データから推定されることを示します。分散の平方根を統計的な誤差とみなすことができます。

標本数によって標本分散の誤差がどのように変わるかを別のノート「典型的な統計モデルにおける標本分散の誤差」に示します。

標本値

平均 m，分散 σ^2（標準偏差 σ）の共通の確率分布に従う n 個の標本値 $\{X_j\}$ ($j = 0, ..., n - 1$) があるとします。また，異なる標本値は互いに独立であるとします。まったく同じ条件で強度測定を繰り返して得られるデータは，測定機器の反応時間が十分に短ければ，このような性格を持つと仮定することができます。

このときに，個々の標本値の積の期待値について，

$$\left\langle X_j X_k \right\rangle = \begin{cases} m^2+\sigma^2 & [j=k] \\ m^2 & [j\neq k] \end{cases} \tag{1}$$

あるいは

$$\left\langle (X_j-m)(X_k-m) \right\rangle = \begin{cases} \sigma^2 & [j=k] \\ 0 & [j\neq k] \end{cases} \tag{2}$$

の関係が成り立ちます。ただし，$\langle X \rangle$ は X の期待値を表すとします。

クロネッカーのデルタ記号

$$\delta_{jk} = \begin{cases} 1 & [j=k] \\ 0 & [j\neq k] \end{cases} \tag{3}$$

を用いれば，上の関係を

$$\left\langle X_j X_k \right\rangle = m^2 + \sigma^2\delta_{jk} \tag{4}$$

とも書けます。

標本平均

標本平均の期待値

標本平均 $\bar{X}$ を

$$\bar{X} \equiv \frac{1}{n}\sum_{j=0}^{n-1} X_j \tag{5}$$

で定義します。標本平均 $\bar{X}$ は実測のデータから計算される値であり，真の平均 m とは異なる値であることに注意してください。

しかし，標本平均の期待値は

$$\langle \bar{X} \rangle = \left\langle \frac{1}{n}\sum_{j=0}^{n-1} X_j \right\rangle = \frac{1}{n}\sum_{j=0}^{n-1}\langle X_j \rangle = \frac{1}{n}\sum_{j=0}^{n-1} m = m \tag{6}$$

となり，真の平均 m の値と一致します。この関係は，標本の平均が真の平均 m から大小どちらかに偏るようなことがないことは期待できることを意味します。

標本平均の分散の期待値

標本平均の分散の期待値は，

$$\begin{aligned}
\left\langle \left(\bar{X}-m\right)^2 \right\rangle &= \left\langle \left(\frac{1}{n}\sum_{j=0}^{n-1} X_j - m\right)^2 \right\rangle = \left\langle \left(\frac{1}{n}\sum_{j=0}^{n-1} X_j - m\right)\left(\frac{1}{n}\sum_{k=0}^{n-1} X_k - m\right) \right\rangle \\
&= \frac{1}{n^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\left\langle \left(X_j - m\right)\left(X_k - m\right)\right\rangle = \frac{1}{n^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\left(\langle X_j X_k\rangle - \langle X_j\rangle m - m\langle X_k\rangle + m^2\right) \\
&= \frac{1}{n^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\left(\langle X_j X_k\rangle - m^2\right) = \frac{1}{n^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sigma^2\delta_{jk} = \frac{1}{n^2}\sum_{j=0}^{n-1}\sigma^2 \\
&= \frac{\sigma^2}{n}
\end{aligned} \tag{7}$$

となります。分散の平方根を誤差と見なすことにすれば，この関係は「標本平均の誤差は，標本数の平方根に反比例して減少する」ことを意味します。

標本分散

標本分散の期待値

標本分散を

$$V = \frac{1}{n-1}\sum_{j=0}^{n-1}\left(X_j - \bar{X}\right)^2 \tag{8}$$

で定義します。 式 (5) の標本平均 $\bar{X}$ の計算では和を n で割っていましたが，式 (8) では和を (n - 1) で割っていることに注意してください。式 (8) で定義される標本分散の期待値 $\langle V \rangle$ は真の分散の値 σ^2 と一致します。この関係は以下のように導出されます。

$$\begin{aligned}
\langle V \rangle &= \frac{1}{n-1}\sum_{j=0}^{n-1}\left\langle \left(X_j - \bar{X}\right)^2 \right\rangle = \frac{1}{n-1}\sum_{j=0}^{n-1}\left\langle \left(X_j - \frac{1}{n}\sum_{k=0}^{n-1}\bar{X}\right)^2 \right\rangle \\
&= \frac{1}{n-1}\sum_{j=0}^{n-1}\left\langle \left(X_j - \frac{1}{n}\sum_{k=0}^{n-1}X_k\right)\left(X_j - \frac{1}{n}\sum_{l=0}^{n-1}X_l\right) \right\rangle
\end{aligned}$$

$$= \frac{1}{n^2(n-1)} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} \left\langle \left(X_j - X_k\right)\left(X_j - X_l\right) \right\rangle$$

$$= \frac{1}{n^2(n-1)} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} \left(\left\langle X_j^2 \right\rangle - \left\langle X_j X_l \right\rangle - \left\langle X_k X_j \right\rangle + \left\langle X_k X_l \right\rangle \right)$$

$$= \frac{1}{n^2(n-1)} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} \left[\left(m^2 + \sigma^2\right) - \left(m^2 + \sigma^2 \delta_{jk}\right) \right]$$

$$= \frac{1}{n^2(n-1)} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} \sigma^2 \left(1 - \delta_{jk}\right) = \frac{\sigma^2}{n^2(n-1)} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} \left(1 - \delta_{jk}\right)$$

$$= \frac{\sigma^2}{n^2(n-1)} \left(n^3 - n^2\right) = \sigma^2 \tag{9}$$

式 (8) の定義を用いれば，標本分散の期待値 $\langle V \rangle$ が真の分散の値 σ^2 から大小どちらかに偏ることがありません。このことから，このように定義される標本分散を「不偏分散」と呼びます。

一方で，かりに標本分散を定義するときに，和を n で割った値

$$V' = \frac{1}{n} \sum_{j=0}^{n-1} \left(X_j - \bar{X}\right)^2 \tag{10}$$

を使うとすると，この期待値は

$$\langle V' \rangle = \frac{n-1}{n} \sigma^2 \tag{11}$$

となり，真の分散の値 σ^2 よりも，平均的には常に少し小さめの値になってしまいます。

そこで，分散を実験的に評価する場合には，式 (8) で定義される不偏分散を使うのが普通です。ただし絶対そうしなければいけないかと言うと，「それほどのことではない」とも言えます。たとえば，「実現確率が最大となるような推定」のしかたは最尤推定と呼ばれてよく使われる方法ですが，統計分布が対称でなければ不偏推定（平均が一致するような推定）と最尤推定は一致しませんし，どちらが良い推定値かは明確ではありません。

標本分散の分散の期待値

不偏分散により「真の分散の値」を推定するとしても，有限の数の標本の値から実験的に評価される分散の値は，当然誤差をともなうものになります。標本分散の分散（統計的なばらつき）を推定する式は

$$\left\langle (V-\sigma)^2 \right\rangle = \frac{1}{(n-1)\left(n^2-3n+3\right)} \left[n \sum_{j=0}^{n-1} \left(X_j - \bar{X}\right)^4 - \left(n^2-3\right) \left(\frac{1}{n-1} \sum_{j=0}^{n-1} \left(X_j - \bar{X}\right)^2 \right)^2 \right] \tag{12}$$

となります。実際には，以下の近似式を使うのでも充分だと思われます。

$$\left\langle (V-\sigma)^2 \right\rangle \sim \frac{1}{n^2} \sum_{j=0}^{n-1} \left(X_j - \bar{X}\right)^4 - \frac{1}{n} \left(\frac{1}{n-1} \sum_{j=0}^{n-1} \left(X_j - \bar{X}\right)^2 \right)^2 \tag{13}$$

以下に式 (12) の導出の考え方を示します。

標本分散の分散の期待値は，以下の計算で導かれます。

$$\left\langle (V-\sigma)^2 \right\rangle = \left\langle \left[\frac{1}{n-1}\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^2-\sigma^2 \right]^2 \right\rangle = \frac{1}{(n-1)^2}\left\langle \left[\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^2-(n-1)\sigma^2 \right]^2 \right\rangle$$

$$=\frac{1}{(n-1)^2}\left\langle \left[\sum_{j=0}^{n-1}\left(X_j-\frac{1}{n}\sum_{k=0}^{n-1}X_k \right)^2-(n-1)\sigma^2 \right]^2 \right\rangle$$

$$=\frac{1}{(n-1)^2}\left\langle \left[\sum_{j=0}^{n-1}\left(X_j-\frac{1}{n}\sum_{k=0}^{n-1}X_k \right)\left(X_j-\frac{1}{n}\sum_{l=0}^{n-1}X_l \right)-(n-1)\sigma^2 \right]^2 \right\rangle$$

$$=\frac{1}{n^4(n-1)^2}\left\langle \left[\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left(X_j-X_k\right)\left(X_j-X_l\right)-n^2(n-1)\sigma^2 \right]^2 \right\rangle$$

$$=\frac{1}{n^4(n-1)^2}\left\langle \left[\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left(X_j-X_k\right)\left(X_j-X_l\right) \right]^2 \right\rangle$$

$$-\frac{2}{n^4(n-1)^2}n^2(n-1)\sigma^2\left\langle \sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left(X_j-X_k\right)\left(X_j-X_l\right) \right\rangle$$

$$+\frac{1}{n^4(n-1)^2}\left\langle \left[n^2(n-1)\sigma^2 \right]^2 \right\rangle$$

$$=\frac{1}{n^4(n-1)^2}\left\langle \left[\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left(X_j-X_k\right)\left(X_j-X_l\right) \right]^2 \right\rangle-\frac{2\sigma^2}{n^2(n-1)}\left\langle \sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left(X_j-X_k\right)\left(X_j-X_l\right) \right\rangle$$

$$+\sigma^2 \tag{14}$$

上の式の第 1 項と第 2 項を別々に計算します。

第 1 項：

$$(\text{第 1 項}) = \frac{1}{n^4(n-1)^2}\left\langle \left[\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left(X_j-X_k\right)\left(X_j-X_l\right) \right]^2 \right\rangle$$

$$=\frac{1}{n^4(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\ \sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\sum_{l'=0}^{n-1}\ \left\langle \left(X_j-X_k\right)\left(X_j-X_l\right)\left(X_{j'}-X_{k'}\right)\left(X_j-X_{l'}\right) \right\rangle$$

$$=\frac{1}{n^4(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\sum_{l'=0}^{n-1}\left\langle \left(X_j^2-X_jX_k-X_jX_l+X_kX_l\right)\left(X_{j'}^2-X_{j'}X_{k'}-X_{j'}X_{l'}+X_{k'}X_{l'}\right) \right\rangle$$

$$=\frac{1}{n^4(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\ \sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\sum_{l'=0}^{n-1}\ \left\langle \left(X_j^2-2X_jX_k+X_kX_l\right)\left(X_{j'}^2-2X_{j'}X_{k'}+X_{k'}X_{l'}\right) \right\rangle$$

$$=\frac{1}{n^4(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\ \sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\sum_{l'=0}^{n-1}$$

$$\left\langle X_j^2\left(X_{j'}^2-2X_{j'}X_{k'}+X_{k'}X_{l'}\right)-2X_jX_k\left(X_{j'}^2-2X_{j'}X_{k'}+X_{k'}X_{l'}\right)\right.$$

$$\left.+X_kX_l\left(X_{j'}^2-2X_{j'}X_{k'}+X_{k'}X_{l'}\right)\right\rangle$$

$$
\begin{aligned}
&= \frac{1}{n^4(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\ \sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\sum_{l'=0}^{n-1} \\
&\quad \Big\langle X_j^2X_{j'}^2 - 2X_j^2X_{j'}X_{k'} + X_j^2X_{k'}X_{l'} - 2X_jX_kX_{j'}^2 + 4X_jX_kX_{j'}X_{k'} - 2X_jX_kX_{k'}X_{l'} \\
&\quad + X_kX_lX_{j'}^2 - 2X_kX_lX_{j'}X_{k'} + X_kX_lX_{k'}X_{l'}\Big\rangle \\
&= \frac{1}{n^4(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\ \sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\sum_{l'=0}^{n-1}\ \Big\langle X_j^2X_k^2 - X_j^2X_kX_l - 2X_j^2X_kX_l + 2X_jX_kX_lX_{j'} \\
&\quad + X_j^2X_kX_l - X_jX_kX_lX_{j'}\Big\rangle \\
&= \frac{1}{n^4(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\ \sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\sum_{l'=0}^{n-1}\ \Big\langle X_j^2X_k^2 - 2X_j^2X_kX_l + X_jX_kX_lX_{j'}\Big\rangle \\
&= \frac{1}{n^2(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\ \Big\langle X_j^2X_k^2 - 2X_j^2X_kX_l + X_jX_kX_lX_{j'}\Big\rangle \\
&= \frac{1}{n^2(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\ \Big\langle X_j^2X_k^2\Big\rangle - \frac{2}{n^2(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\ \Big\langle X_j^2X_kX_l\Big\rangle \\
&\quad + \frac{1}{n^2(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\ \Big\langle X_jX_kX_lX_{j'}\Big\rangle \\
&= \frac{1}{(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\Big\langle X_j^2X_k^2\Big\rangle - \frac{2}{n(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\Big\langle X_j^2X_kX_l\Big\rangle + \frac{1}{n^2(n-1)^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\Big\langle X_jX_kX_lX_{j'}\Big\rangle \\
&= \frac{1}{(n-1)^2}\sum_{j=0}^{n-1}\left[\Big\langle X_j^4\Big\rangle + \sum_{k\neq j}\ \Big\langle X_j^2X_k^2\Big\rangle\right] \\
&\quad - \frac{2}{n(n-1)^2}\sum_{j=0}^{n-1}\left[\Big\langle X_j^4\Big\rangle + \sum_{k\neq j}\Big\langle X_j^2X_k^2\Big\rangle + 2\sum_{k\neq j}\Big\langle X_j^3X_k\Big\rangle + \sum_{k\neq j}\sum_{\substack{l\neq j\\ l\neq k}}\Big\langle X_j^2X_kX_l\Big\rangle\right] \\
&\quad + \frac{1}{n^2(n-1)^2}\sum_{j=0}^{n-1}\Bigg[\ \Big\langle X_j^4\Big\rangle + 3\sum_{k\neq j}\Big\langle X_j^2X_k^2\Big\rangle + 3\sum_{k\neq j}\Big\langle X_j^3X_k\Big\rangle + \sum_{k\neq j}\Big\langle X_jX_k^3\Big\rangle \\
&\qquad + 3\sum_{k\neq j}\sum_{\substack{l\neq j\\ l\neq k}}\Big\langle X_j^2X_kX_l\Big\rangle + 3\sum_{k\neq j}\sum_{\substack{l\neq j\\ l\neq k}}\Big\langle X_jX_k^2X_l\Big\rangle + \sum_{k\neq j}\ \sum_{\substack{l\neq j\\ l\neq k}}\ \sum_{\substack{j'\neq j\\ j'\neq k\\ j'\neq l}}\ \Big\langle X_jX_kX_lX_{j'}\Big\rangle\ \Bigg]
\end{aligned}
\tag{15}
$$

ここで，$\left\langle X_j^k\right\rangle \equiv \left\langle x^k\right\rangle$ とすると，

$$
\begin{aligned}
(\text{第 1 項}) &= \frac{n}{(n-1)^2}\left[\left\langle x^4\right\rangle + (n-1)\left\langle x^2\right\rangle^2\right] \\
&\quad - \frac{2}{(n-1)^2}\left[\left\langle x^4\right\rangle + (n-1)\left\langle x^2\right\rangle^2 + 2(n-1)\left\langle x^3\right\rangle\left\langle x\right\rangle + (n-1)(n-2)\left\langle x^2\right\rangle\left\langle x\right\rangle^2\right] \\
&\quad + \frac{1}{n(n-1)^2}\Big[\ \left\langle x^4\right\rangle + 3(n-1)\left\langle x^2\right\rangle^2 + 4(n-1)\left\langle x^3\right\rangle\left\langle x\right\rangle + 6(n-1)(n-2)\left\langle x^2\right\rangle\left\langle x\right\rangle^2
\end{aligned}
$$

$$+(n-1)(n-2)(n-3)\langle x\rangle^4 \Big]$$

$$=\frac{n}{(n-1)^2}\left[\left\langle x^4\right\rangle+(n-1)\left\langle x^2\right\rangle^2\right]$$

$$-\frac{2}{(n-1)^2}\left[\left\langle x^4\right\rangle+(n-1)\left\langle x^2\right\rangle^2+2(n-1)\left\langle x^3\right\rangle\langle x\rangle+(n-1)(n-2)\left\langle x^2\right\rangle\langle x\rangle^2\right]$$

$$+\frac{1}{n(n-1)^2}\Big[\quad\left\langle x^4\right\rangle+3(n-1)\left\langle x^2\right\rangle^2+4(n-1)\left\langle x^3\right\rangle\langle x\rangle+6(n-1)(n-2)\left\langle x^2\right\rangle\langle x\rangle^2$$

$$+(n-1)(n-2)(n-3)\langle x\rangle^4 \Big]$$

$$=\frac{n^2-2n+1}{n(n-1)^2}\left\langle x^4\right\rangle+\frac{(n-1)\left(n^2-2n+3\right)}{n(n-1)^2}\left\langle x^2\right\rangle^2-\frac{(n-1)(4n-4)}{n(n-1)^2}\left\langle x^3\right\rangle\langle x\rangle$$

$$-\frac{(n-1)(n-2)(2n-6)}{n(n-1)^2}\left\langle x^2\right\rangle\langle x\rangle^2+\frac{(n-1)(n-2)(n-3)}{n(n-1)^2}\langle x\rangle^4$$

$$=\frac{1}{n}\left\langle x^4\right\rangle+\frac{\left(n^2-2n+3\right)}{n(n-1)}\left\langle x^2\right\rangle^2-\frac{4}{n}\left\langle x^3\right\rangle\langle x\rangle-\frac{2(n-2)(n-3)}{n(n-1)}\left\langle x^2\right\rangle\langle x\rangle^2+\frac{(n-2)(n-3)}{n(n-1)}\langle x\rangle^4 \tag{16}$$

第２項：

$$（第２項）=-\frac{2\sigma^2}{n^2(n-1)}\left\langle\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left(X_j-X_k\right)\left(X_j-X_l\right)\right\rangle$$

$$=-\frac{2\sigma^2}{n^2(n-1)}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\left\langle X_j^2-X_jX_k-X_jX_l+X_kX_l\right\rangle$$

$$=-\frac{2\sigma^2}{n(n-1)}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\left\langle X_j^2-X_jX_k\right\rangle=-\frac{2\sigma^2}{n(n-1)}\left(n\sum_{j=0}^{n-1}\left\langle X_j^2\right\rangle-\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\left\langle X_jX_k\right\rangle\right)$$

$$=-\frac{2\sigma^2}{n(n-1)}\left\{n^2\left\langle x^2\right\rangle-n\left[\left\langle x^2\right\rangle+(n-1)\langle x\rangle^2\right]\right\}$$

$$=-2\sigma^2\left(\left\langle x^2\right\rangle-\langle x\rangle^2\right)=-2\sigma^4 \tag{17}$$

式 (16), (17) から，標本分散の分散（第１項＋第２項＋第３項）は

$$\left\langle\left(V-\sigma^2\right)^2\right\rangle=\frac{1}{n}\left\langle x^4\right\rangle+\frac{\left(n^2-2n+3\right)}{n(n-1)}\left\langle x^2\right\rangle^2-\frac{4}{n}\left\langle x^3\right\rangle\langle x\rangle$$

$$-\frac{2(n-2)(n-3)}{n(n-1)}\left\langle x^2\right\rangle\langle x\rangle^2+\frac{(n-2)(n-3)}{n(n-1)}\langle x\rangle^4$$

$$-\sigma^4$$

$$
\begin{aligned}
&= \frac{1}{n}\left\langle x^4 \right\rangle + \frac{\left(n^2-2n+3\right)}{n(n-1)}\left\langle x^2 \right\rangle^2 - \frac{4}{n}\left\langle x^3 \right\rangle\langle x\rangle - \frac{2(n-2)(n-3)}{n(n-1)}\left\langle x^2 \right\rangle\langle x\rangle^2 + \frac{(n-2)(n-3)}{n(n-1)}\langle x\rangle^4 \\
&\qquad -\left(\left\langle x^2 \right\rangle - \langle x\rangle^2\right)^2 \\
&= \frac{1}{n}\left\langle x^4 \right\rangle + \frac{\left(n^2-2n+3\right)}{n(n-1)}\left\langle x^2 \right\rangle^2 - \frac{4}{n}\left\langle x^3 \right\rangle\langle x\rangle - \frac{2(n-2)(n-3)}{n(n-1)}\left\langle x^2 \right\rangle\langle x\rangle^2 + \frac{(n-2)(n-3)}{n(n-1)}\langle x\rangle^4 \\
&\qquad -\left\langle x^2 \right\rangle^2 + 2\left\langle x^2 \right\rangle\langle x\rangle^2 - \langle x\rangle^4 \\
&= \frac{1}{n}\left\langle x^4 \right\rangle + \frac{\left(n^2-2n+3\right)-\left(n^2-n\right)}{n(n-1)}\left\langle x^2 \right\rangle^2 - \frac{4}{n}\left\langle x^3 \right\rangle\langle x\rangle - \frac{2\left[(n-2)(n-3)-n(n-1)\right]}{n(n-1)}\left\langle x^2 \right\rangle\langle x\rangle^2 \\
&\qquad + \frac{(n-2)(n-3)-n(n-1)}{n(n-1)}\langle x\rangle^4 \\
&= \frac{1}{n}\left\langle x^4 \right\rangle - \frac{n-3}{n(n-1)}\left\langle x^2 \right\rangle^2 - \frac{4}{n}\left\langle x^3 \right\rangle\langle x\rangle + \frac{4(2n-3)}{n(n-1)}\left\langle x^2 \right\rangle\langle x\rangle^2 - \frac{2(2n-3)}{n(n-1)}\langle x\rangle^4 \qquad (18)
\end{aligned}
$$

と書ける。一方，４次の中心積率（モーメント）は

$$
\begin{aligned}
\left\langle (x-\langle x\rangle)^4 \right\rangle &= \left\langle x^4 - 4x^3\langle x\rangle + 6x^2\langle x\rangle^2 - 4x\langle x\rangle^3 + \langle x\rangle^4 \right\rangle \\
&= \left\langle x^4 \right\rangle - 4\left\langle x^3 \right\rangle\langle x\rangle + 6\left\langle x^2 \right\rangle\langle x\rangle^2 - 3\langle x\rangle^4 \qquad (19)
\end{aligned}
$$

と表される。式 (18), (19) から，標本分散と「４次中心積率を n で割った値」との差は

$$
\begin{aligned}
&\left\langle \left(V-\sigma^2\right)^2 \right\rangle - \frac{1}{n}\left\langle (x-\langle x\rangle)^4 \right\rangle = \frac{1}{n}\left\langle x^4 \right\rangle - \frac{n-3}{n(n-1)}\left\langle x^2 \right\rangle^2 - \frac{4}{n}\left\langle x^3 \right\rangle\langle x\rangle + \frac{4(2n-3)}{n(n-1)}\left\langle x^2 \right\rangle\langle x\rangle^2 \\
&\qquad - \frac{2(2n-3)}{n(n-1)}\langle x\rangle^4 - \frac{\left\langle x^4 \right\rangle - 4\left\langle x^3 \right\rangle\langle x\rangle + 6\left\langle x^2 \right\rangle\langle x\rangle^2 - 3\langle x\rangle^4}{n} \\
&= -\frac{n-3}{n(n-1)}\left\langle x^2 \right\rangle^2 + \frac{(8n-12)-(6n-6)}{n(n-1)}\left\langle x^2 \right\rangle\langle x\rangle^2 - \frac{(4n-6)-(3n-3)}{n(n-1)}\langle x\rangle^4 \\
&= -\frac{n-3}{n(n-1)}\left\langle x^2 \right\rangle^2 + \frac{2(n-3)}{n(n-1)}\left\langle x^2 \right\rangle\langle x\rangle^2 - \frac{n-3}{n(n-1)}\langle x\rangle^4 = -\frac{n-3}{n(n-1)}\left(\left\langle x^2 \right\rangle - \langle x\rangle^2\right)^2 \\
&= -\frac{n-3}{n(n-1)}\sigma^4 \qquad (20)
\end{aligned}
$$

となります。

したがって，標本分散の分散 $\left\langle \left(V-\sigma^2\right)^2 \right\rangle$ は，４次中心積率 $\left\langle (x-\langle x\rangle)^4 \right\rangle$ を用いて

$$
\left\langle \left(V-\sigma^2\right)^2 \right\rangle = \frac{1}{n}\left\langle (x-\langle x\rangle)^4 \right\rangle - \frac{n-3}{n(n-1)}\sigma^4 \qquad (21)
$$

と表されます。標本数 n が十分に多いとみなせる場合には,

$$
\left\langle \left(V-\sigma^2\right)^2 \right\rangle \sim \frac{1}{n}\left[\left\langle (x-\langle x\rangle)^4 \right\rangle - \sigma^4\right] \qquad (22)
$$

と近似されます。

次に，４次中心積率を，測定データから推定することを考えてみます。まず

$$K' \equiv \frac{1}{n}\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^4 \tag{23}$$

を定義すると，この値の期待値は

$$\begin{aligned}
\langle K'\rangle &= \frac{1}{n}\sum_{j=0}^{n-1}\left\langle\left(X_j-\bar{X}\right)^4\right\rangle = \frac{1}{n}\sum_{j=0}^{n-1}\left\langle\left(X_j-\frac{1}{n}\sum_{k=0}^{n-1}X_k\right)^4\right\rangle \\
&= \frac{1}{n^5}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\left\langle\left(X_j-X_k\right)\left(X_j-X_l\right)\left(X_j-X_{j'}\right)\left(X_j-X_{k'}\right)\right\rangle \\
&= \frac{1}{n^5}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\left\langle\left(X_j-X_k\right)\left(X_j-X_l\right)\left(X_j^2-X_jX_{j'}-X_jX_{k'}+X_{j'}X_{k'}\right)\right\rangle \\
&= \frac{1}{n^5}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\Big\langle\left(X_j-X_k\right) \\
&\qquad \left(X_j^3-X_j^2X_{j'}-X_j^2X_{k'}+X_jX_{j'}X_{k'}-X_j^2X_l+X_jX_lX_{j'}+X_jX_lX_{k'}-X_lX_{j'}X_{k'}\right)\Big\rangle \\
&= \frac{1}{n^5}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\Big\langle X_j\Big(X_j^3-X_j^2X_{j'}-X_j^2X_{k'}+X_jX_{j'}X_{k'} \\
&\qquad -X_j^2X_l+X_jX_lX_{j'}+X_jX_lX_{k'}-X_lX_{j'}X_{k'}\Big) \\
&\qquad -X_k\Big(X_j^3-X_j^2X_{j'}-X_j^2X_{k'}+X_jX_{j'}X_{k'} \\
&\qquad -X_j^2X_l+X_jX_lX_{j'}+X_jX_lX_{k'}-X_lX_{j'}X_{k'}\Big)\Big\rangle \\
&= \frac{1}{n^5}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\Big\langle X_j^4-X_j^3X_{j'}-X_j^3X_{k'}+X_j^2X_{j'}X_{k'} \\
&\qquad -X_j^3X_l+X_j^2X_lX_{j'}+X_j^2X_lX_{k'}-X_jX_lX_{j'}X_{k'} \\
&\qquad -X_j^3X_k+X_j^2X_kX_{j'}+X_j^2X_kX_{k'}-X_jX_kX_{j'}X_{k'} \\
&\qquad +X_j^2X_kX_l-X_jX_kX_lX_{j'}-X_jX_kX_lX_{k'}+X_kX_lX_{j'}X_{k'}\Big\rangle \\
&= \frac{1}{n^5}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\sum_{k'=0}^{n-1}\left\langle X_j^4-4X_j^3X_k+6X_j^2X_kX_l-3X_jX_kX_lX_{j'}\right\rangle \\
&= \frac{1}{n^4}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\left\langle X_j^4-4X_j^3X_k+6X_j^2X_kX_l-3X_jX_kX_lX_{j'}\right\rangle \\
&= \frac{1}{n}\sum_{j=0}^{n-1}\left\langle X_j^4\right\rangle-\frac{4}{n^2}\sum_{j=0}^{n-1}\left(\left\langle X_j^4\right\rangle+\sum_{k\neq j}\left\langle X_j^3X_k\right\rangle\right) \\
&\qquad +\frac{6}{n^3}\sum_{j=0}^{n-1}\left[\left\langle X_j^4\right\rangle+\sum_{k\neq j}\left\langle X_j^2X_k^2\right\rangle+2\sum_{k\neq j}\left\langle X_j^3X_k\right\rangle+\sum_{k\neq j}\sum_{\substack{l\neq j\\ l\neq k}}\left\langle X_j^2X_kX_l\right\rangle\right]
\end{aligned}$$

$$-\frac{3}{n^4}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\sum_{j'=0}^{n-1}\Bigg[\left\langle X_j^4\right\rangle+3\sum_{k\neq j}\left\langle X_j^2X_k^2\right\rangle+3\sum_{k\neq j}\left\langle X_j^3X_k\right\rangle+\sum_{k\neq j}\left\langle X_jX_k^3\right\rangle$$

$$+3\sum_{k\neq j}\sum_{\substack{l\neq j\\ l\neq k}}\left\langle X_j^2X_kX_l\right\rangle+3\sum_{k\neq j}\sum_{\substack{l\neq j\\ l\neq k}}\left\langle X_jX_k^2X_l\right\rangle+\sum_{k\neq j}\sum_{\substack{l\neq j\\ l\neq k}}\sum_{\substack{j'\neq j\\ j'\neq k\\ j'\neq l}}\left\langle X_jX_kX_lX_{j'}\right\rangle\Bigg]$$

$$=\left\langle x^4\right\rangle-\frac{4}{n}\left[\left\langle x^4\right\rangle+(n-1)\left\langle x^3\right\rangle\langle x\rangle\right]$$

$$+\frac{6}{n^2}\left[\left\langle x^4\right\rangle+(n-1)\left\langle x^2\right\rangle^2+2(n-1)\left\langle x^3\right\rangle\langle x\rangle+(n-1)(n-2)\left\langle x^2\right\rangle\langle x\rangle^2\right]$$

$$-\frac{3}{n^3}\Bigg[\left\langle x^4\right\rangle+3(n-1)\left\langle x^2\right\rangle^2+4(n-1)\left\langle x^3\right\rangle\langle x\rangle+6(n-1)(n-2)\left\langle x^2\right\rangle\langle x\rangle^2$$

$$+(n-1)(n-2)(n-3)\langle x\rangle^4\Bigg]$$

$$=\left\langle x^4\right\rangle-\frac{4}{n}\left\langle x^4\right\rangle-\frac{4(n-1)}{n}\left\langle x^3\right\rangle\langle x\rangle$$

$$+\frac{6}{n^2}\left\langle x^4\right\rangle+\frac{6(n-1)}{n^2}\left\langle x^2\right\rangle^2+\frac{12(n-1)}{n^2}\left\langle x^3\right\rangle\langle x\rangle+\frac{6(n-1)(n-2)}{n^2}\left\langle x^2\right\rangle\langle x\rangle^2$$

$$-\frac{3}{n^3}\left\langle x^4\right\rangle-\frac{9(n-1)}{n^3}\left\langle x^2\right\rangle^2-\frac{12(n-1)}{n^3}\left\langle x^3\right\rangle\langle x\rangle-\frac{18(n-1)(n-2)}{n^3}\left\langle x^2\right\rangle\langle x\rangle^2$$

$$-\frac{3(n-1)(n-2)(n-3)}{n^3}\langle x\rangle^4$$

$$=\left\langle x^4\right\rangle-\frac{4}{n}\left\langle x^4\right\rangle+\frac{6}{n^2}\left\langle x^4\right\rangle-\frac{3}{n^3}\left\langle x^4\right\rangle$$

$$-\frac{4(n-1)}{n}\left\langle x^3\right\rangle\langle x\rangle+\frac{12(n-1)}{n^2}\left\langle x^3\right\rangle\langle x\rangle-\frac{12(n-1)}{n^3}\left\langle x^3\right\rangle\langle x\rangle$$

$$+\frac{6(n-1)}{n^2}\left\langle x^2\right\rangle^2-\frac{9(n-1)}{n^3}\left\langle x^2\right\rangle^2$$

$$+\frac{6(n-1)(n-2)}{n^2}\left\langle x^2\right\rangle\langle x\rangle^2-\frac{18(n-1)(n-2)}{n^3}\left\langle x^2\right\rangle\langle x\rangle^2$$

$$-\frac{3(n-1)(n-2)(n-3)}{n^3}\langle x\rangle^4$$

$$=\frac{n^3-4n^2+6n-3}{n^3}\left\langle x^4\right\rangle$$

$$-\frac{4(n-1)\left(n^2-3n+3\right)}{n^3}\left\langle x^3\right\rangle\langle x\rangle$$

$$+\frac{(n-1)(6n-9)}{n^3}\left\langle x^2\right\rangle^2$$

$$+\frac{(n-1)(n-2)(6n-18)}{n^3}\left\langle x^2\right\rangle\langle x\rangle^2$$

$$-\frac{3(n-1)(n-2)(n-3)}{n^3}\langle x\rangle^4$$

$$\langle K'\rangle = \frac{n^3-4n^2+6n-3}{n^3}\langle x^4\rangle - \frac{4(n-1)(n^2-3n+3)}{n^3}\langle x^3\rangle\langle x\rangle + \frac{3(n-1)(2n-3)}{n^3}\langle x^2\rangle^2$$
$$+\frac{6(n-1)(n-2)(n-3)}{n^3}\langle x^2\rangle\langle x\rangle^2 - \frac{3(n-1)(n-2)(n-3)}{n^3}\langle x\rangle^4$$

$$\begin{array}{r|llll}
 & n^2 & -3n & +3 & \\
\hline
n-1\ ) & n^3 & -4n^2 & +6n & -3 \\
 & n^3 & -n^2 & & \\
\hline
 & & -3n^2 & +6n & \\
 & & -3n^2 & +3n & \\
\hline
 & & & 3n & -3 \\
 & & & 3n & -3 \\
\hline
 & & & & 0
\end{array}$$

$$\langle K'\rangle = \frac{(n-1)(n^2-3n+3)}{n^3}\langle x^4\rangle - \frac{4(n-1)(n^2-3n+3)}{n^3}\langle x^3\rangle\langle x\rangle + \frac{(n-1)(6n-9)}{n^3}\langle x^2\rangle^2$$
$$+\frac{6(n-1)(n-2)(n-3)}{n^3}\langle x^2\rangle\langle x\rangle^2 - \frac{3(n-1)(n-2)(n-3)}{n^3}\langle x\rangle^4 \qquad (24)$$

ここで,

$$\frac{(n-1)(n^2-3n+3)}{n^3}\langle (x-m)^4\rangle = \frac{(n-1)(n^2-3n+3)}{n^3}\langle x^4\rangle - \frac{4(n-1)(n^2-3n+3)}{n^3}\langle x^3\rangle\langle x\rangle$$
$$+\frac{6(n-1)(n^2-3n+3)}{n^3}\langle x^2\rangle\langle x\rangle^2 - \frac{3(n-1)(n^2-3n+3)}{n^3}\langle x\rangle^4 \qquad (25)$$

なので,

$$\langle K'\rangle - \frac{(n-1)(n^2-3n+3)}{n^3}\langle (x-m)^4\rangle$$
$$= \frac{3(n-1)(2n-3)}{n^3}\langle x^2\rangle^2$$
$$+\frac{6(n-1)(n-2)(n-3)}{n^3}\langle x^2\rangle\langle x\rangle^2 - \frac{6(n-1)(n^2-3n+3)}{n^3}\langle x^2\rangle\langle x\rangle^2$$
$$-\frac{3(n-1)(n-2)(n-3)}{n^3}\langle x\rangle^4 + \frac{3(n-1)(n^2-3n+3)}{n^3}\langle x\rangle^4$$
$$= \frac{3(n-1)(2n-3)}{n^3}\langle x^2\rangle^2$$
$$+\frac{6(n-1)\left[(n^2-5n+6)-(n^2-3n+3)\right]}{n^3}\langle x^2\rangle\langle x\rangle^2$$
$$-\frac{3(n-1)\left[(n^2-5n+6)-(n^2-3n+3)\right]}{n^3}\langle x\rangle^4$$

$$= \frac{3(n-1)(2n-3)}{n^3}\langle x^2\rangle^2 - \frac{6(n-1)(2n-3)}{n^3}\langle x^2\rangle\langle x\rangle^2 + \frac{3(n-1)(2n-3)}{n^3}\langle x\rangle^4$$
$$= \frac{3(n-1)(2n-3)}{n^3}\left(\langle x^2\rangle^2 - 2\langle x^2\rangle\langle x\rangle^2 + \langle x\rangle^4\right)$$
$$= \frac{3(n-1)(2n-3)}{n^3}\left(\langle x^2\rangle - \langle x\rangle^2\right)^2$$
$$= \frac{3(n-1)(2n-3)}{n^3}\sigma^4 \tag{26}$$

の関係が導かれます。まとめると,

$$\langle K'\rangle = \frac{(n-1)(n^2-3n+3)}{n^3}\left\langle (x-m)^4\right\rangle + \frac{3(n-1)(2n-3)}{n^3}\sigma^4 \tag{27}$$

となり，この式を変形して

$$\left\langle (x-m)^4\right\rangle = \frac{n^3}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{3(2n-3)}{(n^2-3n+3)}\sigma^4 \tag{28}$$

だから,

$$\left\langle \left(V-\sigma^2\right)^2\right\rangle = \frac{1}{n}\left\langle (x-m)^4\right\rangle - \frac{n-3}{n(n-1)}\sigma^4$$
$$= \frac{1}{n}\left[\frac{n^3}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{3(2n-3)}{(n^2-3n+3)}\sigma^4\right] - \frac{n-3}{n(n-1)}\sigma^4$$
$$= \frac{n^2}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{3(2n-3)}{n(n^2-3n+3)}\sigma^4 - \frac{n-3}{n(n-1)}\sigma^4$$
$$= \frac{n^2}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{3(n-1)(2n-3)}{n(n-1)(n^2-3n+3)}\sigma^4 - \frac{(n-3)(n^2-3n+3)}{n(n-1)(n^2-3n+3)}\sigma^4$$
$$= \frac{n^2}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{3(2n^2-5n+3)+(n^3-3n^2+3n-3n^2+9n-9)}{n(n-1)(n^2-3n+3)}\sigma^4$$
$$= \frac{n^2}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{(6n^2-15n+9)+(n^3-6n^2+12n-9)}{n(n-1)(n^2-3n+3)}\sigma^4$$
$$= \frac{n^2}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{n^3-3n}{n(n-1)(n^2-3n+3)}\sigma^4$$
$$= \frac{n^2}{(n-1)(n^2-3n+3)}\langle K'\rangle - \frac{n^2-3}{(n-1)(n^2-3n+3)}\sigma^4$$
$$= \frac{1}{(n-1)(n^2-3n+3)}\left[n^2\langle K'\rangle - (n^2-3)\sigma^4\right] \tag{29}$$

結局，標本分散の分散を推定する式として，式 (12) :

$$\left\langle \left(V-\sigma^2\right)^2\right\rangle = \frac{1}{(n-1)(n^2-3n+3)}\left[n\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^4 - (n^2-3)\left(\frac{1}{n-1}\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^2\right)^2\right]$$

あるいはこれを近似する式 (13) :

$$\left\langle \left(V-\sigma^2\right)^2 \right\rangle \sim \frac{1}{n^2}\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^4 - \frac{1}{n}\left(\frac{1}{n-1}\sum_{j=0}^{n-1}\left(X_j-\bar{X}\right)^2\right)^2$$

を用いれば良いと思われるわけです。