2013 年 07 月 04 日

標本分散、不偏分散が一致推定量であること

新潟工科大学 情報電子工学科 竹野茂治

1 はじめに

「確率・統計」の講義の点推定のところで、教科書などには不偏分散 V_1 と 標本分散 V_2 はどちらも母分散 σ^2 の一致推定量である、と書いてあったが、証明は省略されていたのが気になり自分で計算してみた。多少計算量が必要であったが、ここにそれをまとめておく。

2 一致推定量

この文書では、X_i $(i=1,2,\ldots,n)$ は、ある一つの確率分布 F に従う、互いに独立な確率変数とする。F の (母) 平均 $(=E[X_i])$ を μ とし、F の (母) 分散 $(=V[X_i])$ を σ^2 とする。X_i の **不偏分散** V_1 と **標本分散** V_2 は、次の式で定義される確率変数である。

$$V_1=\frac{S}{n-1},\quad V_2=\frac{S}{n},\quad S=\sum_{i=1}^n(X_i-\overline{X})^2 \tag{1}$$

ここで、$\overline{X}$ は X_i の算術平均 (確率変数としての平均ではない)

$$\overline{X}=\frac{1}{n}\sum_{i=1}^n X_i$$

であり、S は平方和と呼ばれる。S は、容易に次のように変形できる。

$$\begin{aligned} S &= \sum_{i=1}^n(X_i^2-2X_i\overline{X}+\overline{X}^2)=\sum_{i=1}^n X_i^2-2\overline{X}\sum_{i=1}^n X_i+n\overline{X}^2 \\ &= n\overline{X^2}-2\overline{X}\cdot n\overline{X}+n\overline{X}^2=n(\overline{X^2}-\overline{X}^2) \end{aligned} \tag{2}$$

ここで、$\overline{X^k}$ は、X_i^k の算術平均を意味するものとする。

$$\overline{X^k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

$X_1, X_2, \ldots X_n$ によって与えられるある確率変数 $T = T(X_1, X_2, \ldots X_n)$ が、F に関わるあるパラメータ θ の **一致推定量** であるとは、任意の正数 k に対して、

$$\lim_{n\to\infty} P(|T-\theta| > k) = 0 \tag{3}$$

となることを言うようである。これは、n が十分大きければ、T の値は θ の近くに分布して、n を大きくすれば、θ から離れた値を取る確率はいくらでも小さくなる、ということを意味していて、これにより T の値でパラメータ θ の値を推定 (点推定) できることの一つの保証が与えられることになる。

この一致性を示すのに重要なのが、次のチェビシェフの不等式である。

定理 1

確率変数 X、および正数 k に対して、

$$P(|X - E[X]| > k) \leq \frac{V[X]}{k^2} \tag{4}$$

が成り立つ ($E[X]$ は X の平均、$V[X]$ は X の分散)。

証明

分散 $V[X]$ を積分で表現して、(4) の範囲に制限すれば、

$$\begin{aligned} V[X] &= E[|X-E[X]|^2] = \int |X-E[X]|^2 dP \\ &\geq \int_{|X-E[X]|>k} |X-E[X]|^2 dP \geq k^2 P(|X-E[X]| > k) \end{aligned}$$

となるので、k^2 で両辺を割れば (4) が得られる。■

例えば、これを使って、標本平均 $\overline{X}$ が母平均 μ の一致推定量であることが確認してみよう。

平均 E の線形性により、

$$E[\overline{X}] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\cdot n\mu = \mu$$

であり、また、X, Y が独立の場合 $V[X+Y] = V[X] + V[Y]$ であるから、

$$V[\overline{X}] = \frac{1}{n^2}\sum_{i=1}^{n} V[X_i] = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$

となる。よって、$\overline{X}$ にチェビシェフの不等式を適用すると、

$$P(|\overline{X}-\mu| > k) \leq \frac{V[\overline{X}]}{k^2} = \frac{\sigma^2}{nk^2} \to 0 \quad (n \to \infty \text{ のとき})$$

となるので、

$$\lim_{n\to\infty} P(|\overline{X}-\mu| > k) = 0$$

であることがわかり、$\overline{X}$ は μ の一致推定量となる。

3 不偏分散、標本分散の平均

本節では、不偏分散、標本分散の平均 (確率変数としての平均) を計算する。そのために、平方和 S の平均をまず求める。

(2) により、

$$E[S] = nE[\overline{X^2} - \overline{X}^2]$$

となるが、$\overline{X}^2$ を

$$\overline{X}^2 = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} X_i^2 + \sum_{i\neq j} X_i X_j\right)$$

と展開すれば、X_i は互いに独立なので $i \neq j$ のとき $E[X_i X_j] = E[X_i]E[X_j]$ であり、よって、今後 $E[X_i^k] = \xi_k$ と書くことにすれば、

$$\begin{aligned} E[S] &= n \cdot \frac{1}{n} \sum_{i=1}^{n} E[X_i^2] - n \cdot \frac{1}{n^2} \left(\sum_{i=1}^{n} E[X_i^2] + \sum_{i \neq j} E[X_i]E[X_j] \right) \\ &= n\xi_2 - \frac{1}{n} \cdot n\xi_2 - \frac{1}{n} \cdot {}_nP_2\, \xi_1^2 = (n-1)(\xi_2 - \xi_1^2) \end{aligned}$$

となる。ここで、${}_nP_k$ は n 個から k 個を取って並べる順列の数で、

$${}_nP_k = n(n-1)(n-2)\cdots(n-k+1)$$

である。一方、X_i の分散 σ^2 は、

$$\sigma^2 = V[X_i] = E[(X_i - \mu)^2] = E[X_i^2] - E[X_i]^2$$

より、

$$\sigma^2 = \xi_2 - \xi_1^2 \tag{5}$$

となるので、結局、S の平均は、

$$E[S] = (n-1)\sigma^2$$

であることがわかり、よって不偏分散、標本分散の平均は、

$$E[V_1] = \frac{1}{n-1} E[S] = \sigma^2, \quad E[V_1] = \frac{1}{n} E[S] = \frac{n-1}{n} \sigma^2 \tag{6}$$

となる。

4 チェビシェフの不等式の分散への適用

本節では、チェビシェフの不等式を利用して、不偏分散と標本分散の母分散への一致性を、不偏分散の極限を考えることに帰着させる。

まず、V_1 に対してチェビシェフの不等式を適用すると、(6) より、

$$P(|V_1-\sigma^2|>k)\leq\frac{1}{k^2}V[V_1]$$

が言えるので、よって、もし

$$\lim_{n\to\infty}V[V_1]=0 \tag{7}$$

であれば、V_1 の σ^2 に対する一致性が言えることになる。

また、V_2 に対しては、(6) より、チェビシェフの不等式は

$$P\left(\left|V_2-\frac{n-1}{n}\sigma^2\right|>k\right)\leq\frac{1}{k^2}V[V_2] \tag{8}$$

となるが、$|V_2-\sigma^2|>\hat{k}$ ($\hat{k}$ は任意の正数) のとき、

$$\left|V_2-\frac{n-1}{n}\sigma^2\right|\geq|V_2-\sigma^2|-\left|\sigma^2-\frac{n-1}{n}\sigma^2\right|>\hat{k}-\frac{\sigma^2}{n}$$

であり、また、

$$V[V_2]=V\left[\frac{n-1}{n}V_1\right]=\left(\frac{n-1}{n}\right)^2V[V_1]$$

なので、(8) より、

$$P(|V_2-\sigma^2|>\hat{k})\leq P\left(\left|V_2-\frac{n-1}{n}\sigma^2\right|>\hat{k}-\frac{\sigma^2}{n}\right)\leq\frac{1}{\left(\hat{k}-\dfrac{\sigma^2}{n}\right)^2}\left(\frac{n-1}{n}\right)^2V[V_1]$$

となることがわかる。よって、この場合も、(7) が言えれば、

$$\lim_{n\to\infty}P(|V_2-\sigma^2|>\hat{k})=0$$

が言えることになるので、結局 V_1, V_2 の σ^2 に対する一致性は、(7) を示せばよいことになる。

なお、(7) を示すために、今後 $E[X_i^k]=\xi_k$ は、$k=1,2,3,4$ に対して「有限である」と仮定する。

5 不偏分散の自乗の展開

本節では (7) を示すために、不偏分散の分散 (確率変数としての分散) を計算する。

(6) より、

$$V[V_1] = E[(V_1 - \sigma^2)^2] = E[V_1^2] - (\sigma^2)^2 \tag{9}$$

であるが、この $E[V_1^2]$ は (2) より、

$$E[V_1^2] = E\left[\left(\frac{n}{n-1}\right)^2 (\overline{X^2} - \overline{X}^2)^2\right] = \left(\frac{n}{n-1}\right)^2 E[\overline{X^2}^2 - 2\overline{X^2}\,\overline{X}^2 + \overline{X}^4] \tag{10}$$

となる。この (10) の最後の式の中身を順に展開していくが、そのために次のような記号を導入する。$\alpha_1, \ldots, \alpha_k$ を自然数として、

$$SX(\alpha_1, \ldots, \alpha_k) = \sum_{i_1,\ldots,i_k}{}' X_{i_1}^{\alpha_1} \cdots X_{i_k}^{\alpha_k} \tag{11}$$

と定義する。ただし、和 $\sum'_{i_1,\ldots,i_k}$ は、各 i_j が 1 から n まで動き、かつ $i_1, \ldots, i_k$ はすべて互いに異なるものに対する和であるとする。例えば、

$$SX(2) = \sum_{i=1}^n X_i^2, \quad SX(2,1) = \sum_{i \neq j} X_i^2 X_j, \quad SX(2,2) = \sum_{i \neq j} X_i^2 X_j^2 = 2\sum_{i<j} X_i^2 X_j^2$$

などとなる。

命題 2

SX 同士の積について次が成り立つ。

$$\begin{aligned} &SX(\alpha_1, \ldots, \alpha_k) SX(\beta) \\ &\quad = \sum_{j=1}^k SX(\alpha_1, \ldots, \alpha_j + \beta, \ldots, \alpha_k) + SX(\alpha_1, \ldots, \alpha_k, \beta) \end{aligned} \tag{12}$$

証明

(12) の左辺は、

$$SX(\alpha_1,\ldots,\alpha_k)SX(\beta) = \sum_{i_1,\ldots,i_k}{}' X_{i_1}^{\alpha_1}\cdots X_{i_k}^{\alpha_k}\sum_{i=1}^{n} X_j^{\beta}$$

であるが、$SX(\beta)$ の部分を $X_{i_1}^{\beta}$, ..., $X_{i_k}^{\beta}$ と、それ以外に分ければ、

$$\begin{aligned}
&SX(\alpha_1,\ldots,\alpha_k)SX(\beta)\\
&= \sum_{i_1,\ldots,i_k}{}' \sum_{j=1}^{k} X_{i_1}^{\alpha_1}\cdots X_{i_j}^{\alpha_j+\beta}\cdots X_{i_k}^{\alpha_k} + \sum_{i_1,\ldots,i_k,i}{}' X_{i_1}^{\alpha_1}\cdots X_{i_k}^{\alpha_k}X_i^{\beta}\\
&= \sum_{j=1}^{k} SX(\alpha_1,\ldots,\alpha_j+\beta,\ldots,\alpha_k) + SX(\alpha_1,\ldots,\alpha_k,\beta)
\end{aligned}$$

∎

これを使うと、まず $\overline{X^2}^2$ は、

$$\overline{X^2}^2 = \left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right)^2 = \frac{1}{n^2}SX(2)^2 = \frac{1}{n^2}(SX(4)+SX(2,2)) \tag{13}$$

となる。次に、$\overline{X^2}\,\overline{X}^2$ は、

$$\begin{aligned}
\overline{X^2}\,\overline{X}^2 &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 = \frac{1}{n^3}SX(2)SX(1)^2\\
&= \frac{1}{n^3}(SX(3)+SX(2,1))SX(1)\\
&= \frac{1}{n^3}(SX(4)+SX(3,1)+SX(3,1)+SX(2,2)+SX(2,1,1))
\end{aligned}$$

となるので、

$$2\overline{X^2}\,\overline{X}^2 = \frac{2}{n^3}(SX(4)+2SX(3,1)+SX(2,2)+SX(2,1,1)) \tag{14}$$

となる。最後に、$\overline{X}^4$ は、

$$\begin{aligned}
\overline{X}^4 &= \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^4 = \frac{1}{n^4}SX(1)^4 = \frac{1}{n^4}(SX(2)+SX(1,1))SX(1)^2 \\
&= \frac{1}{n^4}(SX(3)+SX(2,1)+SX(2,1)+SX(1,2)+SX(1,1,1))SX(1) \\
&= \frac{1}{n^4}(SX(3)+3SX(2,1)+SX(1,1,1))SX(1) \\
&= \frac{1}{n^4}(SX(4)+SX(3,1)+3SX(3,1)+3SX(2,2)+3SX(2,1,1) \\
&\quad +SX(2,1,1)+SX(1,2,1)+SX(1,1,2)+SX(1,1,1,1))
\end{aligned}$$

となるので、

$$\overline{X}^4 = \frac{1}{n^4}(SX(4)+4SX(3,1)+3SX(2,2)+6SX(2,1,1)+SX(1,1,1,1)) \tag{15}$$

となる。

(13), (14), (15) より、

$$\begin{aligned}
&\overline{X^2}^2 - 2\overline{X^2}\,\overline{X}^2 + \overline{X}^4 \\
&= \frac{1}{n^2}(SX(4)+SX(2,2)) - \frac{2}{n^3}(SX(4)+2SX(3,1)+SX(2,2)+SX(2,1,1)) \\
&\quad + \frac{1}{n^4}(SX(4)+4SX(3,1)+3SX(2,2)+6SX(2,1,1)+SX(1,1,1,1)) \\
&= \frac{(n-1)^2}{n^4}SX(4) - \frac{4(n-1)}{n^4}SX(3,1) + \frac{n^2-2n+3}{n^4}SX(2,2) \\
&\quad - \frac{2(n-3)}{n^4}SX(2,1,1) + \frac{1}{n^4}SX(1,1,1,1)
\end{aligned} \tag{16}$$

となる。

6 不偏分散の分散の極限

次は、(16) の平均の計算である。

$$
\begin{aligned}
E[SX(4)] &= \sum_{i=1}^{n} E[X_i^4] = n\xi_4, \\
E[SX(3,1)] &= \sum_{i\neq j} E[X_i^3 X_j] = \sum_{i\neq j} E[X_i^3]E[X_j] = {}_nP_2\,\xi_3\xi_1, \\
E[SX(2,2)] &= \sum_{i\neq j} E[X_i^2 X_j^2] = {}_nP_2\,\xi_2^2, \\
E[SX(2,1,1)] &= \sideset{}{'}\sum_{i,j,k} E[X_i^2 X_j X_k] = \sideset{}{'}\sum_{i,j,k} E[X_i^2]E[X_j]E[X_k] = {}_nP_3\,\xi_2\xi_1^2, \\
E[SX(1,1,1,1)] &= \sideset{}{'}\sum_{i,j,k,l} E[X_i X_j X_k X_l] = {}_nP_4\,\xi_1^4
\end{aligned}
$$

となるので、(16) より、

$$
\begin{aligned}
&E[\overline{X^2}^2 - 2\overline{X^2}\,\overline{X}^2 + \overline{X}^4] \\
&= \frac{(n-1)^2}{n^4}E[SX(4)] - \frac{4(n-1)}{n^4}E[SX(3,1)] + \frac{n^2-2n+3}{n^4}E[SX(2,2)] \\
&\quad - \frac{2(n-3)}{n^4}E[SX(2,1,1)] + \frac{1}{n^4}E[SX(1,1,1,1)] \\
&= \frac{(n-1)^2}{n^3}\xi_4 - \frac{4(n-1)^2}{n^3}\xi_3\xi_1 + \frac{(n^2-2n+3)(n-1)}{n^3}\xi_2^2 \\
&\quad - \frac{2(n-1)(n-2)(n-3)}{n^3}\xi_2\xi_1^2 + \frac{(n-1)(n-2)(n-3)}{n^3}\xi_1^4 \\
&= \frac{n-1}{n^3}\{(n-1)\xi_4 - 4(n-1)\xi_3\xi_1 \\
&\quad + (n^2-2n+3)\xi_2^2 - 2(n-2)(n-3)\xi_2\xi_1^2 + (n-2)(n-3)\xi_1^4\}
\end{aligned}
$$

となるが、この最後のかっこ内の後半 3 項の和を考えると、(5) より $\xi_1 = \mu, \xi_2 = \sigma^2+\mu^2$ なので、

$$
\begin{aligned}
&(n^2-2n+3)\xi_2^2 - 2(n-2)(n-3)\xi_2\xi_1^2 + (n-2)(n-3)\xi_1^4 \\
&= (n^2-2n+3)(\sigma^2+\mu^2)^2 - 2(n-2)(n-3)(\sigma^2+\mu^2)\mu^2 + (n-2)(n-3)\mu^4 \\
&= (n^2-2n+3)(\sigma^2)^2 + 2(n^2-2n+3-(n-2)(n-3))\sigma^2\mu^2 \\
&\quad + (n^2-2n+3-(n-2)(n-3))\mu^4 \\
&= (n^2-2n+3)(\sigma^2)^2 + 6(n-1)\sigma^2\mu^2 + 3(n-1)\mu^4
\end{aligned}
$$

となることがわかるので、結局

$$\begin{aligned}
&E[\overline{X^2}^2-2\overline{X^2}\,\overline{X}^2+\overline{X}^4]\\
&=\ \frac{n-1}{n^3}\{(n-1)\xi_4-4(n-1)\xi_3\xi_1+(n^2-2n+3)(\sigma^2)^2\\
&\qquad+6(n-1)\sigma^2\mu^2+3(n-1)\mu^4\}\\
&=\ \frac{(n-1)^2}{n^3}(\xi_4-4\xi_3\xi_1+6\sigma^2\mu^2+3\mu^4)+\frac{(n-1)(n^2-2n+3)}{n^3}(\sigma^2)^2
\end{aligned}$$

となる。(9) に戻れば、(10) より

$$\begin{aligned}
V[V_1] &= \left(\frac{n}{n-1}\right)^2 E[\overline{X^2}^2-2\overline{X^2}\,\overline{X}^2+\overline{X}^4]-(\sigma^2)^2\\
&= \frac{1}{n}(\xi_4-4\xi_3\xi_1+6\sigma^2\mu^2+3\mu^4)+\left(\frac{n^2-2n+3}{n(n-1)}-1\right)(\sigma^2)^2\\
&= \frac{1}{n}(\xi_4-4\xi_3\xi_1+6\sigma^2\mu^2+3\mu^4)-\frac{n-3}{n(n-1)}(\sigma^2)^2
\end{aligned}$$

となる。

よって、ξ_k $(1\leq k\leq 4)$ が有限という仮定の元では、

$$\lim_{n\to\infty}V[V_1]=\lim_{n\to\infty}\left(\frac{1}{n}(\xi_4-4\xi_3\xi_1+6\sigma^2\mu^2+3\mu^4)-\frac{n-3}{n(n-1)}(\sigma^2)^2\right)=0$$

が言えることになり、これで V_1, V_2 がともに σ^2 の一致推定量であることが示されたことになる。

7 最後に

この辺りのことがちゃんと書いてある統計の本は読んでいないので、本来はこのような形で証明するものではないかもしれないが、本稿のものでも一応証明にはなっているだろうと思う。

ただ、これはあくまで教科書に書いてあることを個人的な疑問から埋めてみただけのものなので、ちゃんと勉強したい人は、これではなく、ちゃんとした本の証明を読んで勉強した方がいいだろう。