応用計量経済分析 TA セッション
第 6 回：標本分散の分布

TA：北村友宏 *

2015 年 11 月 17 日

1　母平均と母分散

- 母集団分布の平均を**母平均**（population mean）とよぶ。
- 母集団分布の分散を**母分散**（population variance）とよぶ。
  - ⋆ 母平均を μ、母分散を σ^2 と書くことが多い。
  - ⋆ 母集団全体を観測できない場合、どちらも現実には未知。

2　標本分散の期待値

- 無作為標本 $(X_1, X_2, \ldots, X_n)$ の**標本分散**（sample variance）は、

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

例題 1.　母平均 μ、母分散 σ^2 の母集団から抽出した大きさ n の無作為標本 $(X_1, X_2, \ldots, X_n)$ を考える。また、μ（と σ^2）は未知とし、標本分散を $s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ とする。このとき、$E\left(s^2\right) = \sigma^2$ を証明しなさい。ただし、$E\left(\overline{X}\right) = \mu, V\left(\overline{X}\right) = \dfrac{\sigma^2}{n}$ であることを使ってよい。

$\boxed{\text{証明の方針}}$　$(X_i - \mu)^2$ の項と $\left(\overline{X} - \mu\right)^2$ の項を無理やり作る。すると分散の定義から、
$V(X_i) = E(X_i - \mu)^2, V\left(\overline{X}\right) = E\left(\overline{X} - \mu\right)^2$ を適用できる。
（証明）母平均は μ、母分散は σ^2 なので、

$$E(X_1) = E(X_2) = \cdots = E(X_n) = \mu,$$
$$V(X_1) = V(X_2) = \cdots = V(X_n) = \sigma^2$$

* Email: kitamu.tom@gmail.com　　URL: http://tomkitamura.html.xdomain.jp

である。s^2 の総和記号の部分は、

$$
\begin{aligned}
\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 &= \sum_{i=1}^{n}\left(X_i-\mu+\mu-\overline{X}\right)^2 \\
&= \sum_{i=1}^{n}\left[(X_i-\mu)-\left(\overline{X}-\mu\right)\right]^2 \\
&= \sum_{i=1}^{n}\left[(X_i-\mu)^2-2(X_i-\mu)\left(\overline{X}-\mu\right)+\left(\overline{X}-\mu\right)^2\right] \\
&= \sum_{i=1}^{n}(X_i-\mu)^2-\sum_{i=1}^{n}\underbrace{2}_{i\text{ に依存しない}}(X_i-\mu)\underbrace{\left(\overline{X}-\mu\right)}_{i\text{ に依存しない}}+\sum_{i=1}^{n}\underbrace{\left(\overline{X}-\mu\right)^2}_{i\text{ に依存しない}} \\
&= \sum_{i=1}^{n}(X_i-\mu)^2-2\left(\overline{X}-\mu\right)\sum_{i=1}^{n}(X_i-\mu)+n\left(\overline{X}-\mu\right)^2 .
\end{aligned}
$$

ここで、$\overline{X}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$ なので、$n\overline{X}=\displaystyle\sum_{i=1}^{n}X_i$ である。

よって、$\displaystyle\sum_{i=1}^{n}(X_i-\mu)=\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu=n\overline{X}-n\mu=n\left(\overline{X}-\mu\right)$ となる。したがって、

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

である。よって、

$$
\begin{aligned}
E\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right] &= E\left[\sum_{i=1}^{n}(X_i-\mu)^2-n\left(\overline{X}-\mu\right)^2\right] \\
&= E\left[\sum_{i=1}^{n}(X_i-\mu)^2\right]-E\left[n\left(\overline{X}-\mu\right)^2\right] \\
&= E\left[(X_1-\mu)^2+(X_2-\mu)^2+\cdots+(X_n-\mu)^2\right]-nE\left(\overline{X}-\mu\right)^2 \\
&= \underbrace{E(X_1-\mu)^2}_{=V(X_1)}+\underbrace{E(X_2-\mu)^2}_{=V(X_2)}+\cdots+\underbrace{E(X_n-\mu)^2}_{=V(X_n)}-n\underbrace{E\left(\overline{X}-\mu\right)^2}_{=V\left(\overline{X}\right)} \\
&= V(X_1)+V(X_2)+\cdots+V(X_n)-nV\left(\overline{X}\right) \\
&= \underbrace{\sigma^2+\sigma^2+\cdots+\sigma^2}_{n\text{ 個}}-n\cdot\frac{\sigma^2}{n} \\
&= n\sigma^2-\sigma^2 \\
&= (n-1)\sigma^2
\end{aligned}
$$

となる。したがって、

$$E\left(s^2\right) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right] = \frac{1}{n-1}\cdot(n-1)\sigma^2 = \sigma^2$$

である。（証明終）

$\boxed{\text{この式から分かること}}$　母平均 μ が未知のとき、μ を標本平均 $\overline{X}$ で推定して標本分散を $s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ とすれば、$E\left(s^2\right) = \sigma^2$ となる。