样本方差的无偏估计

Guangyao Zhao

2022-08-26

Contents

样本方差的无偏估计公式如下：

$$
\begin{aligned}
&\mathrm{E}\left[S^2\right]=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right]=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left((X_i-\mu)-(\bar{X}-\mu)\right)^2\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left((X_i-\mu)^2-2(\bar{X}-\mu)(X_i-\mu)+(\bar{X}-\mu)^2\right)\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-\frac{2}{n}(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)+\frac{1}{n}(\bar{X}-\mu)^2\sum_{i=1}^{n}1\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-\frac{2}{n}(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)+\frac{1}{n}(\bar{X}-\mu)^2\cdot n\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-\frac{2}{n}(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)+(\bar{X}-\mu)^2\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-\frac{2}{n}(\bar{X}-\mu)\cdot n\cdot(\bar{X}-\mu)+(\bar{X}-\mu)^2\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-2(\bar{X}-\mu)^2+(\bar{X}-\mu)^2\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-(\bar{X}-\mu)^2\right]\\
&=\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right]-\mathrm{E}\left[(\bar{X}-\mu)^2\right]\\
&=\sigma^2-\mathrm{E}\left[(\bar{X}-\mu)^2\right]
\end{aligned}
$$

所以只有样本均值和真值均值相等的时候，样本方差的均值才和真值方差相等。由于样本的随机性，样本均值取值不一定，所以上式的 $\mathrm{E}\left[(\bar{X}-\mu)^2\right]$ (即样本期望的方差) 可能并不为 0：

$$
\begin{aligned}
E(\bar{X}-\mu)^2 = E(\bar{X}-E[\bar{X}])^2 &= \operatorname{var}(\bar{X}) \\
&= \operatorname{var}\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) \\
&= \frac{1}{n^2} \operatorname{var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \operatorname{var}\left(X_i\right) \\
&= \frac{n\sigma^2}{n^2} \\
&= \frac{\sigma^2}{n}
\end{aligned}
$$

综上，方差的无偏估计为：

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

但是当 n 大到一定程度的时候 $E\left[S^2\right]$ 是整体方差的相合估计。