Answer on Question#65384 – Math – Statistics and Probability

Question. Let $X_1, X_2, \dots, X_n$ be a random sample with $E(X_i) = m$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \dots, n$. Show that $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is an unbiased estimator of σ^2.

Proof. Let us compute $E(S^2) = \frac{1}{n-1}\sum_{i=1}^{n} E[(X_i - \bar{X})^2] = \frac{1}{n-1}\sum_{i=1}^{n} E\left[\left((X_i - m) - (\bar{X} - m)\right)^2\right]$

$= \frac{1}{n-1}\sum_{i=1}^{n} E[(X_i - m)^2 - 2(X_i - m)(\bar{X} - m) + (\bar{X} - m)^2] = \frac{1}{n-1}\sum_{i=1}^{n} E[(X_i - m)^2]$

$- \frac{2}{n-1} E \sum_{i=1}^{n}(X_i - m)(\bar{X} - m) + \frac{1}{n-1}\sum_{i=1}^{n} E[(\bar{X} - m)^2] = \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i)$

$- \frac{2}{n-1} E(\bar{X} - m)\sum_{i=1}^{n}(X_i - m) + \frac{n}{n-1} E[(\bar{X} - m)^2] = \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i)$

$- \frac{2}{n-1} E(\bar{X} - m)(\sum_{i=1}^{n} X_i - mn) + \frac{n}{n-1} E[(\bar{X} - m)^2] = \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i)$

$- \frac{2}{n-1} E(\bar{X} - m)(n\bar{X} - nm) + \frac{n}{n-1} E[(\bar{X} - m)^2] = \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i) - \frac{2n}{n-1} E[(\bar{X} - m)^2]$

$+ \frac{n}{n-1} E[(\bar{X} - m)^2] = \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i) - \frac{n}{n-1} E[(\bar{X} - m)^2] = \frac{n\sigma^2}{n-1} - \frac{n}{n-1} Var(\bar{X})$

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

$= \frac{n\sigma^2}{n-1}\left(1 - \frac{1}{n}\right) = \frac{n\sigma^2}{n-1} \cdot \frac{n-1}{n} = \sigma^2$. During these computations we used the following facts:

$Var(X) = E\left[\left(X - E(X)\right)^2\right]$ by definition (see https://en.wikipedia.org/wiki/Variance);

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ by definition (see https://en.wikipedia.org/wiki/Sample_mean_and_covariance), hence $\sum_{i=1}^{n} X_i = n\bar{X}$;

$Var(aX) = a^2 Var(X)$ (see https://en.wikipedia.org/wiki/Variance#Properties);

$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$ when $X_1, X_2, \dots, X_n$ are independent

(see https://en.wikipedia.org/wiki/Variance#Properties); in our case $X_1, X_2, \dots, X_n$ are independent by definition of data sample

(see https://en.wikipedia.org/wiki/Sample_(statistics)).

Since $E(S^2) = \sigma^2$ we conclude that $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is an unbiased estimator of σ^2 by definition of unbiased estimator (see https://en.wikipedia.org/wiki/Estimator#Bias).

Assertion is established.