An Unbiased Estimator of the Variance of iid RV's

Let $X_1, \ldots, X_n$ be iid rv's with mean μ and variance σ^2. This example shows that $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is an unbiased estimator of σ^2; i.e., $E[S^2] = \sigma^2$.

EXAMPLE 4a

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having expected value μ and variance σ^2, and as in Example 2c, let $\bar{X} = \sum_{i=1}^{n} X_i/n$ be the sample mean. The quantities $X_i - \bar{X}, i = 1, \ldots, n$, are called *deviations*, as they equal the differences between the individual data and the sample mean. The random variable

$$S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$$

is called the *sample variance*. Find (a) $\text{Var}(\bar{X})$ and (b) $E[S^2]$.

Solution.

$$\begin{aligned}
\text{(a)} \quad \text{Var}(\bar{X}) &= \left(\frac{1}{n}\right)^2 \text{Var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \text{Var}(X_i) \quad \text{by independence} \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

(b) We start with the following algebraic identity:

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

Taking expectations of the preceding yields that

$$\begin{aligned}
(n-1)E[S^2] &= \sum_{i=1}^{n} E[(X_i - \mu)^2] - nE[(\bar{X} - \mu)^2] \\
&= n\sigma^2 - n\text{Var}(\bar{X}) \\
&= (n-1)\sigma^2.
\end{aligned}$$

Thus $E[S^2] = \sigma^2$.