ON VARIANCE ESTIMATORS

Let \( X_1, \ldots, X_n \) be a random sample from a distribution with mean \( \mu \) and variance \( \sigma^2 \).

Suppose first that \( \mu \) is known. Form \( \hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2 \), an estimator for the distribution variance.

This estimator is unbiased because all \( X_i \) have the same mean \( \mu \) and the same variance \( \sigma^2 \):

\[
E[\hat{\sigma}_0^2] = \frac{1}{n} \sum_{i=1}^{n} E[(X_i - \mu)^2] = \frac{1}{n} \sum_{i=1}^{n} \text{Var}(X_i) = \frac{1}{n} n \sigma^2 = \sigma^2.
\]

If \( \mu \) is not known, it is often replaced by \( \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \). The corresponding variance estimator

\[
\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]

has a bias. Indeed, using that \( \hat{\sigma}_0^2 = \hat{\sigma}^2 + (\bar{X} - \mu)^2 \), we have

\[
E[\hat{\sigma}^2] = \sigma^2 - \text{Var}(\bar{X}).
\]

Thus \( \hat{\sigma}^2 \) tends to underestimate \( \sigma^2 \). The value of \( \text{Var}(\bar{X}) \) depends on the sampling procedure.

When \( X_i \) are uncorrelated, \( \text{Var}(\bar{X}) = \sigma^2/n \). In this case, \( E[\hat{\sigma}^2] = \frac{n-1}{n} \sigma^2 \) and so

\[
S^2 = \frac{n}{n-1} \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.
\]

is an unbiased estimator. The factor \( \frac{n}{n-1} \) is known as Bessel’s correction.

Alternatively, we may see that \( S^2 \) is unbiased by writing

\[
S^2 = \frac{1}{\binom{n}{2}} \sum_{i<j} \frac{(X_i - X_j)^2}{2}
\]

and noting that \( E[(X_i - X_j)^2/2] = \sigma^2 \), for each pair of distinct indices \( i, j \).

Other variance estimators exist. They are designed to meet particular criteria and often depend on the properties of the underlying distribution. For instance, if \( X_i \sim \mathcal{N}(\mu, \sigma^2) \), the variance estimator

\[
\tilde{\sigma}^2 = \frac{n}{n+1} \hat{\sigma}^2 = \frac{1}{n+1} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]

minimizes the mean of the squared relative error \( (1 - \tilde{\sigma}^2/\sigma^2)^2 \). It is, however, biased.

Among the unbiased estimators available, we should probably seek the ones with smaller variance.

The comparison of estimators may be involved. For instance, \( \text{Var}(S^2) = \frac{1}{n} \left( \mu_4 - \frac{n-3}{n-1} \sigma^4 \right) \),

where \( \mu_4 = E[(X - \mu)^4] \) is the fourth central moment (kurtosis). In the normal case, \( \text{Var}(S^2) = \frac{2\sigma^4}{n-1} \).