Estimators

Introduction to Machine Learning (CSCI 1950-F), Summer 2011

Solve the following problems. Provide mathematical justification for your answers.

Let $X_1, \ldots, X_n \sim$Normal(μ, σ^2) i.i.d., with $n > 1$. We will assume that μ is known, and consider various estimators of the variance σ^2. For these exercises, use the square loss: $L(\sigma_1^2, \sigma_2^2) = (\sigma_1^2 - \sigma_2^2)^2$. Recall that the "unbiased" sample variance is

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

where $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ is the sample mean. Consider the family of estimators

$$v_b = bs^2$$

parametrized by the real numbers $b \geq 0$. Note that the "biased" sample variance, defined as

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

is the member of this family obtained by taking $b = (n-1)/n$. (Of course, the "unbiased" sample variance is also a member of the family.)

(1) Unbiasedness

Compute the expected value of v_b. Use this to show that $\hat{\sigma}^2$ is indeed biased and that s^2 is indeed unbiased.

(2) Inadmissibility

(A) The variance of these estimators is very tedious to compute directly. Assuming (without proof) that the variance of s^2 is $2\sigma^4/(n-1)$, compute the variance of v_b.

(B) Use the bias-variance decomposition to obtain an expression for the (frequentist) risk for v_b,

$$R(\sigma^2, v_b) = EL(\sigma^2, v_b),$$

where σ^2 is the true value of the variance.

(C) Does $\hat{\sigma}^2$ dominate s^2? Does s^2 dominate $\hat{\sigma}^2$? Draw a sketch of the risk functions $R(\sigma^2, s^2)$ and $R(\sigma^2, \hat{\sigma}^2)$, with σ^2 on the "x"-axis and the risk on the "y"-axis. Among all estimators v_b in the family, is there one that is "best" in some sense? Why or why not?