HBC HW2 - Math 425

oqueen

March 2021

1 Problem

In this document, I will show three examples of estimators that violate 1) sufficiency, 2) efficiency, and 3) consistency. Specifically, I will focus on estimators that do not consider the full amount of the given sample. For all problems assume that we have a random sample $Y_1, ..., Y_n$ with observed values $y_1, ..., y_n$ from some distribution Y. We will consider the case where we take $Y_1, ..., Y_m$ where $m < n$, i.e. a smaller sample than the one we have overall.

1.1 Sufficiency

For an estimator to be sufficient, we know that the following function:

$$P(X_1 = k_1, ..., X_n = k_n | \hat{\theta} = \theta_e)$$

is not dependent on θ. If this probability is dependent on θ, then $\hat{\theta}$ is not a sufficient estimator. For example, suppose we have a random variable X, where our sample is $X_1, ..., X_n$ with observed values $x_1, ..., x_n$, and X had a binomial distribution. If we take $\hat{\theta^*}$ to be only dependent on a subset of the total sample, call this set X^*, then we would get:

$$P(X_1 = k_1, ..., X_n = k_n | \hat{\theta^*} = h(X^*)) = \frac{P(X_1 = k_1, ..., X_n = k_n)}{P(\hat{\theta^*} = h(X^*))}$$

This will expand to:

$$P(X_1 = k_1, ..., X_n = k_n | \hat{\theta^x} = h(X^*)) = \frac{\theta^{\sum_{i=1}^{n} k_i} (1-\theta)^{n - \sum_{i=1}^{n} k_i}}{\theta^{h(X^*)(1-\theta)^{1-h(X^*)}}}$$

and there is no way that this will simplify to an equation that does not depend on θ, the original parameter in the random variable, unless $h(X^*) = \sum_{i=1}^{n} k_i$. Therefore, we must consider the entire sample $X_1, ..., X_n$ to calculate θ in order for our estimator to be sufficient.

1.2 Efficiency

Consider an estimator $\hat{\theta}$ for θ such that $\hat{\theta}$ only considers a fixed number of m samples. By the Cramer-Rao inequality we have:

$$Var(\hat{\theta}) \geq \left[nE\left[\left(\frac{\partial ln f_Y(Y;\theta)}{\partial \theta} \right)^2 \right] \right]^{-1}$$

Let's assume that for some fixed n^* sample size, our estimator $\hat{\theta}$ is efficient. However, if we increase our total sample size to $n^* + 1$, this causes the CR lower bound to decrease, thereby invalidating the efficiency for our estimator. Therefore, we must consider all samples in any particular random sample in order to create an efficient estimator.

1.3 Consistency

To satisfy consistency, we must show:

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| < \epsilon) = 1$$

This is usually shown by Chebyshev's inequality:

$$P(|Y - E(Y)| < \epsilon) \geq 1 - \frac{\sigma^2}{\epsilon^2}$$

For all $\epsilon > 0$. Let us consider our $\hat{\theta}$ estimator to only depend on a fixed number of samples, (i.e. the $Y_1, ..., Y_m$ discussed in the introduction). However, if this is the case, then σ^2 would be constant, as would $E(\hat{\theta})$. Therefore, as we took the limit to ∞, the $\frac{\sigma^2}{\epsilon^2}$ term in Chebyshev's inequality would not change, causing $1 - \frac{\sigma^2}{\epsilon^2}$ to never be equal to 1 (assuming variance of our estimator is not zero, meaning it would be a fixed value). Therefore, our estimator would not be consistent; we must also consider every sample to obtain a proper consistent estimator.