Chapter 3 - Monte Carlo Methods

Monte Carlo Applications to Statistical Inference.
Point Estimators.

Prof. Alex Alvarez, Ali Raisolsadat

School of Mathematical and Computational Sciences
University of Prince Edward Island

Standard Error of an Estimator

In this lecture we will see how to use Monte Carlo methods to estimate the standard error of an estimator.

A reminder that our framework consists on having some observations $x = (x_1, x_2, ..., x_n)$ that are considered a random sample of a random variable $X = (X_1, X_2, ..., X_n) \sim P_\theta$, where $\theta \in \Theta$ is unknown.

Let $\hat{\theta} = \hat{\theta}(X)$ be an estimator of θ. One important quantity that we study on estimators is their standard deviation.

$$s.e._\theta(\hat{\theta}) = stdev\left(\hat{\theta}(X)\right) = \sqrt{Var\left(\hat{\theta}(X)\right)}$$

Standard Error of an Estimator

In some cases we can estimate $s.e._\theta(\hat{\theta})$ easily.

For example, for estimators of the form $\hat{\theta} = \frac{1}{n}\sum_{j=1}^{n} X_j$ we can estimate their standard deviation as

$$\widehat{s.e.}_\theta(\hat{\theta}) = \frac{\hat{s}}{\sqrt{n}}$$

where $\hat{s}$ is the sample standard deviation of the observations $(x_1, x_2, .., x_n)$.

(we used this in previous sections of this chapter to give a confidence interval of a Monte Carlo estimator)

However, for estimators $\hat{\theta}$ that are not given as a sample mean, estimating their standard error may be more complicated.

Estimating the Standard Error

For a given value of θ, if we are able to generate N samples: $\left\{x^{(j)}\right\}_{j=1,2,\ldots,N}$ of the random variable X we could estimate the standard error of the estimator $\hat{\theta}$ as:

$$\widehat{s.e.}_{\theta}(\hat{\theta}) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\hat{\theta}(x^{(j)}) - \overline{\hat{\theta}^{(\cdot)}}\right)^2}$$

where $\overline{\hat{\theta}^{(\cdot)}} = \frac{1}{N}\sum_{j=1}^{N}\hat{\theta}(x^{(j)})$

If necessary, we can do this for a range of values of θ to get a better idea of the standard error of the estimator as a function of θ.

Example

Assume that we would like to estimate an unknown parameter θ from a sample of 20 random numbers that are distributed according to the uniform distribution on $[0, \theta]$. Two unbiased estimators for θ are proposed and we would like to find out which of these two estimators has a smaller variance.

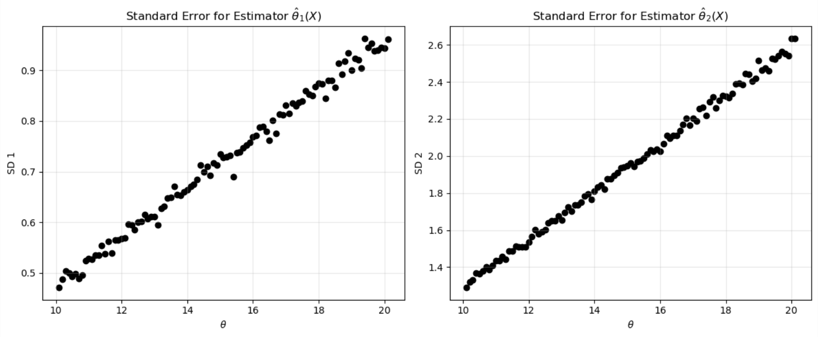
Estimator 1: $\hat{\theta}_1(X) = \frac{21}{20}\max(X_1, X_2, ...X_{20})$

Estimator 2: $\hat{\theta}_2(X) = \frac{2}{N}\sum_{j=1}^{20} X_j$

Use Monte Carlo methods to estimate the standard error for both estimators, for values of θ on $[10, 20]$.

Algorithm 1 Monte Carlo Estimation of Estimator Variability

Input: Number of simulations N, sample size m, parameter grid $\{\theta_j\}_{j=1}^{J}$
for $j = 1$ to J **do**
    Initialize empty vectors for estimator values
    for $i = 1$ to N **do**
        Generate $U_{i1}, \ldots, U_{im} \sim \text{Uniform}(0,\, \theta_j)$
        Compute estimators $\hat{\theta}_{1i} = a_1 \max(U_i) + b_1$ and $\hat{\theta}_{2i} = a_2\, \overline{U_i} + b_2$
    end for
    Compute standard deviations: $\texttt{sd1}_j = \text{sd}(\hat{\theta}_1), \quad \texttt{sd2}_j = \text{sd}(\hat{\theta}_2)$
end for
Output: Standard deviation profiles $\{\texttt{sd1}_j, \texttt{sd2}_j\}_{j=1}^{J}$



Note: Estimator 1 would be preferable as it has a small standard error.

Homework

Assume that we would like to estimate the unknown parameter μ from a sample of 10 random numbers that are distributed according to the normal distribution $N(\mu, 1)$. We don't have the full information about these observations, we only have the min, median and max.

Two unbiased estimators for μ are proposed and we would like to find out which of these two estimators has a smaller variance.

Estimator 1: $\hat{\mu}_1(X) = \dfrac{\min(X) + \max(X)}{2}$

Estimator 2: $\hat{\mu}_2(X) = median(X)$

Use Monte Carlo methods to estimate the standard error for both estimators, for values of μ on $[-5, 5]$