Bootstrapping for estimating MSE of some estimator

1. When estimating $\theta(F)$, we need to assess the ***quality*** of the estimator $g(\vec{X}) \equiv g(X_1, \ldots, X_n)$; that is, to estimate the *mean square error* (MSE) of $g(\vec{X})$, where $X_i \overset{i.i.d.}{\sim} F, i = 1, \ldots, n$. $\Big($We like to have $\mathrm{E}(g(\vec{X})) = \theta$ and small $\mathrm{Var}(g(\vec{X}))$ or small MSE= $\left[\mathrm{E}(g(\vec{X})) - \theta\right]^2 + \mathrm{Var}(g(\vec{X}))$.$\Big)$

2. For estimator $g(\vec{X})$ to be useful, we need to estimate the MSE of $g(\vec{X})$, denoted by $\mathrm{MSE}(F) \equiv E_F\Big[(g(\vec{X}) - \theta(F))^2\Big]$. $\Big($If $g(\vec{X})$ is unbiased for θ (i.e., $\mathrm{E}(g(\vec{X})) = \theta$), then MSE= $\mathrm{Var}(g(\vec{X}))$. Furthermore, if $g(\vec{X}) \sim$ normal, we have $\mathrm{P}\Big(\big|g(\vec{X}) - \theta\big| \le 2\sqrt{\mathrm{Var}(g(\vec{X}))}\Big) = 95\%$.$\Big)$

3. **When F is known, we can theoretically compute MSE(F).** However, F is often unknown and can be estimated by the empirical distribution function

$$F_e(x) = \frac{\text{number of } i : X_i \le x}{n}, \qquad x \in R$$

$$\iff \quad \mathrm{P}(Y = x_i) = \frac{1}{n}, \qquad i = 1, \ldots, n. \qquad \text{(This can be easily seen in the plot of } F_e(x)\text{)}$$

4. **Another way of thinking about** F_e is that it is the distribution function of a random variable Y which is equally likely to take on any of the n values $x_i,\ i = 1, \ldots, n$. That is, $Y \sim \mathrm{uniform}(x_1, \ldots, x_n)$.

5. By the strong law of large numbers, $F_e \approx F$ as $n \longrightarrow \infty$, implying that

$$\mathrm{MSE}(F_e) \approx \mathrm{MSE}(F)$$

with $\mathrm{MSE}(F_e) \equiv E_{F_e}\Big[(g(\vec{Y}) - \theta(F_e))^2\Big] = \sum_{i_1} \cdots \sum_{i_n} \frac{[g(x_{i_1}, \ldots, x_{i_n}) - \theta(F_e)]^2}{n^n}$, an average of n^n terms.

6. Since F_e is known, we can either **analytically derive** MSE(F_e) or **exactly compute** MSE(F_e) when n^n is not too large.

7. If n^n is large, estimate MSE(F_e) by simulation.

   First compute $\theta(F_e)$, then **sample** $(Y_1^1, \ldots, Y_n^1), \ldots, (Y_1^b, \ldots, Y_n^b)$, and finally evaluate $g(Y_1^1, \ldots, Y_n^1), \ldots, g(Y_1^n, \ldots, Y_n^n)$ to form the estimator

$$\widehat{\mathrm{MSE}}(F_e) = \sum_{j=1}^{b} \frac{[g(Y_1^j, \ldots, Y_n^j) - \theta(F_e)]^2}{b}.$$

   It has been reported that choosing $b = 100$ is usually sufficient.

   Key: $\theta =?$, $g(\vec{X}) =?$, and $\theta(F_e) =?$

Example: When estimating $\theta(F) = \mathrm{E}(X)$ by $\bar{X}$, we can estimate its MSE by S^2/n. Now estimate its MSE by bootstrapping as follows.

Suppose the realized sample is $\{x_1, \ldots, x_n\}$.
Since $Y \sim$ uniform$(x_1, \ldots, x_n)$, we have $\theta(F_e) = \sum_{i=1}^n \frac{1}{n} \cdot x_i = \bar{x}$.

Analytically derive the MSE(F_e) as follows.

$$\begin{aligned}
\mathrm{MSE}(F_e) &= \mathrm{E}_{F_e}\left[\left(\bar{Y} - \underbrace{\bar{x}}_{= \mathrm{E}(Y) = \mathrm{E}(\bar{Y})}\right)^2\right] \\
&= \mathrm{Var}_{F_e}\!\left(\bar{Y}\right) \\
&= \mathrm{Var}_{F_e}(Y)/n.
\end{aligned}$$

Since

$$\mathrm{Var}_{F_e}(Y) = \mathrm{E}_{F_e}\!\left[(Y - \bar{x})^2\right] = \sum_{i=1}^n (x_i - \bar{x})^2/n,$$

we have

$$\mathrm{MSE}(F_e) = \sum_{i=1}^n (x_i - \bar{x})^2/n^2 \;=\; (S^2/n) \cdot \left(\frac{n-1}{n}\right) \approx S^2/n.$$

Key: $\theta = \mu$, $g(\vec{X}) = \bar{X}$, and $\theta(F_e) = \bar{x}$

Example: If $n = 2$ and $X_1 = 1$ and $X_2 = 3$, what is the bootstrap estimate of Var(S^2)?

1. **Bootstrapping technique is used to estimate MSE of some estimator $\widehat{\theta}$.** For unbiased estimator, MSE$(\widehat{\theta}) =$ Var$(\widehat{\theta})$.

2. Since S^2 is an unbiased estimator of σ^2, estimating Var(S^2) is equivalent to estimating MSE(S^2).

   Thus, consider $\theta = \sigma^2$ and $g(\vec{X}) = S^2$. Then

   $$\mathrm{MSE}(F_e) = \mathrm{E}_Y\!\left[(g(\vec{Y}) - \theta(F_e))^2\right]$$

   where $\theta(F_e) = \left[(1-2)^2 + (3-2)^2\right]/2 = 1$ and

   $$g(\vec{Y}) = \begin{cases} 0, & \text{if } (Y_1, Y_2) = (1,1) \\ 0, & \text{if } (Y_1, Y_2) = (3,3) \\ \left[(1-2)^2 + (3-2)^2\right]/(2-1) = 2, & \text{if } (Y_1, Y_2) = (1,3) \\ 2, & \text{if } (Y_1, Y_2) = (3,1), \end{cases}$$

   yielding MSE$(F_e) = \left[(0-1)^2 + (0-1)^2 + (2-1)^2 + (2-1)^2\right]/2^2 = 1$.
   So the bootstrap estimate of Var(S^2) is equal to 1.

   Key: $\theta = \sigma^2$, $g(\vec{X}) = S^2$, and $\theta(F_e) = \sigma^2(F_e) = 1$.