Homework 4

DSO603, due by the end of Nov 11th, 2025

1. The bootstrap, jackknife, jackknife-after-bootstrap (JAB), and infinitesimal jackknife (IJ) all aim to approximate sampling variability, but they differ in philosophy and computational cost.

   (a) Briefly explain the main idea and goal of each method: Bootstrap, Jackknife, JAB, and IJ. In one or two sentences per method, describe what quantity each estimates (e.g., sampling variance, Monte Carlo variance, or analytic approximation).

   (b) Suppose we use a causal forest with $n = 2000$ and $B = 1000$ trees. Which resampling method(s) are practical for variance estimation, and why? Comment on the computational feasibility and stability of each.

2. **Simulation Exercise**

   Generate data as follows:

$$X_i \sim \text{Uniform}[0,1], W_i \sim \text{Bernoulli}(0.5), \varepsilon_i \sim N(0,1),$$
$$Y_i = 1 + 2X_i + W_i(1 + X_i) + \varepsilon_i.$$

   The true conditional average treatment effect (CATE) is:

$$\tau(x) = 1 + x.$$

   Use $n = 1000$ observations as sample size. Split the data into 60% training, and 40% testing sets. Set the number of trees in forest 500.

   - Use the `causal_forest()` function from the `grf` package in R (https://github.com/grf-labs/grf).

- Fit two models:
  (a) An adaptive causal forest with `honesty = FALSE`.
  (b) An honest causal forest with `honesty = TRUE` (default).
- Keep all other tuning parameters identical across the two models (e.g., `num.trees`, `min.node.size`, `sample.fraction`).

Repeat the Monte Carlo simulation 50 times. Produce the following outputs:

(1) Predicted Effect Plot: Plot the true treatment effect $\tau(x) = 1 + x$ (solid black line) and the estimated effects $\hat{\tau}_{RF}(x)$ and $\hat{\tau}_{CF}(x)$ (colored lines) over the test set. $\hat{\tau}_{RF}(x)$ and $\hat{\tau}_{CF}(x)$ should be average across 50 repetitions.

(2) Performance Summary Table: Compute the following test-set metrics for both estimators:

$$\text{MSE} = \frac{1}{n_{\text{test}}} \sum_i (\hat{\tau}(X_i) - \tau(X_i))^2, \quad \text{Bias} = \frac{1}{n_{\text{test}}} \sum_i (\hat{\tau}(X_i) - \tau(X_i)).$$

Present a 2 by 2 table comparing MSE and Bias for RF and CF. Your MSE and Bias should be average across 50 repetitions.

(3) Uncertainty Output for honest CF: For each Monte Carlo repetition, using the variance estimates from honest CF (extracted from the package), compute the mean standard error

$$\overline{SE}_{IJ} = \frac{1}{n_{\text{test}}} \sum_i \widehat{SE}_{IJ}(X_i),$$

and report the 95% coverage rate:

$$\text{Coverage} = \frac{1}{n_{\text{test}}} \sum_i \mathbf{1}\{|\hat{\tau}(X_i) - \tau(X_i)| \le 1.96\, \widehat{SE}_{IJ}(X_i)\}.$$

Summarize these two statistics in a short table. Your reported values should be average across 50 repetitions.

Briefly discuss your simulation outputs.