Mark van der Laan & Rachael Phillips
Division of Biostatistics, UC Berkeley

Exercise 6

Super Learning and the Oracle Inequality

1. Consider a candidate estimate ψ and the true parameter value ψ_0. Explain the loss-based dissimilarity between ψ and ψ_0, $d_0 = (\psi, \psi_0)$.

2. Estimators are denoted with _____ which maps _____ into _____.

If we assume the observations are _____, then the data can be identified by

_____.

3. Define the cross-validated selector and the oracle selector.

4. Explain the oracle inequality for a quadratic loss-based dissimilarity or a non-quadratic loss-based dissimilarity in terms of the conditions required for it to hold. *Hint*: These are the conditions required for the remainder term to hold.

5. The cross-validation selector is asymptotically equivalent to the oracle selector. What does this mean in words?