MATH562 – Fall 2025

Problem Set: Week 14

1. Suppose that estimator T for parameter $\theta \in \mathbb{R}$ has expectation equal to $\theta(1+\gamma)$, so that the bias is $\theta\gamma$. The bias factor γ can be estimated by $C = \mathrm{E}^*(T^*)/T - 1$, where E^* denotes expectation over bootstrap sampling and T^* is based on the bootstrap samples.

   (a) Argue that the variance estimate $T = n^{-1}\sum(Y_j - \bar{Y})^2$ is such an estimator, and show that in this case C is exactly equal to γ.

   (b) If C were approximated from R resamples by C^*, what would be the simulation variance of C^*?

2. Let T be the median of a random sample of size $n = 2m+1$ with ordered values $y_{(1)} < \cdots < y_{(n)}$; the observed value of T is therefore $t = y_{(m+1)}$.

   (a) Show that $T^* > y_{(l)}$ if and only if fewer than $m+1$ of the Y_j^* are less than or equal to $y_{(l)}$, and deduce that
   $$\Pr{}^*(T^* > y_{(l)}) = \sum_{j=0}^{m} \binom{n}{j} \left(\frac{l}{n}\right)^j \left(1 - \frac{l}{n}\right)^{n-j}.$$
   This specifies the exact resampling distribution of the sample median, and can be used to prove that the bootstrap estimate of $\mathrm{Var}(T)$ is consistent as $n \to \infty$.

   (b) Use the resampling distribution in (a) to show that for $n = 11$,
   $$\Pr{}^*(T^* \leq y_{(3)}) = \Pr{}^*(T^* \geq y_{(9)}) = 0.051,$$
   and deduce that an approximate basic bootstrap 90% confidence interval for the population median is $(2y_{(6)} - y_{(8)}, 2y_{(6)} - y_{(4)})$.