ST 453 Advanced computing for statistical reasoning
Homework problem set 9

November 4, 2024

No R packages are permitted for use in this assignment.

1. Derive a confidence interval for λ based on a random sample $Y_1, \ldots, Y_n \stackrel{\text{iid}}{\sim} \exp(\lambda)$. Design and implement a simulation study to demonstrate that the confidence interval achieves its nominal $1-\alpha$ level coverage for $\alpha \in \{.01, .02, \ldots, .99\}$. Display the nominal versus empirical coverage as a scatter plot.

2. Derive a confidence interval for the intercept β_0 based on the simple linear regression model $Y_i = \beta_0 + \beta_1 x_i + U_i$, for $i \in \{1, \ldots, n\}$, where $U_1, \ldots, U_n \stackrel{\text{iid}}{\sim} \mathrm{N}(0, \sigma^2)$. Assume σ^2 is known. Design and implement a simulation study to demonstrate that the confidence interval achieves its nominal $1-\alpha$ level coverage for $\alpha \in \{.01, .02, \ldots, .99\}$. Display the nominal versus empirical coverage as a scatter plot.

3. Repeat problem 2, but with σ^2 unknown. You will need to first construct an unbiased estimator for σ^2.

4. Derive a confidence interval for β_1 in problem 2, but with σ^2 unknown (recall that we did this in lecture for σ^2 known). Design and implement a simulation study to demonstrate that the confidence interval achieves its nominal $1-\alpha$ level coverage for $\alpha \in \{.01, .02, \ldots, .99\}$. Display the nominal versus empirical coverage as a scatter plot.

5. Derive a prediction interval for $Y_{n+1} \sim \mathrm{N}(x_{n+1}'\beta, \sigma^2)$, assuming σ^2 is unknown, based on the linear regression model $Y \sim \mathrm{N}_n(X\beta, \sigma^2 I_n)$ with $X \in \mathbb{R}^{n\times p}$ and $\mathrm{rank}(X) = p$. Recall from lecture that such an interval was derived in the case of σ^2 known. Design and implement a simulation study to demonstrate that the prediction interval achieves its nominal $1-\alpha$ level coverage for $\alpha \in \{.01, .02, \ldots, .99\}$. Display the nominal versus empirical coverage as a scatter plot.