Module 9: Parallel computing and simulations

July 23, 2024

Simulations on Cauchy

Suppose $\boldsymbol{X} = (X_1, \ldots, X_n)$ is an i.i.d. sample from the shifted Cauchy distribution with density

$$f(x \mid \theta) = \frac{1}{\pi\left(1 + (x - \theta)^2\right)}, \quad x \in \mathbb{R}$$

Our goal is to compare the following 4 estimators of the parameter θ.

- Sample mean
$$\hat{\theta}_n^{(1)} = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$
- Sample median
$$\hat{\theta}_{\mathrm{n}}^{(2)} = M_{\mathrm{n}} = \frac{1}{2}\left(X_{(k)} + X_{(k+1)}\right)$$
- Modified sample mean
$$\hat{\theta}_n^{(3)} = M_n + \frac{2}{n} \cdot \left.\frac{\partial \ell}{\partial \theta}\right|_{\theta = M_n}$$
where ℓ is the log-likelihood function.
- Maximum likelihood estimator (MLE) $\hat{\theta}_{\mathrm{n}}^{(4)}$ defined by
$$\ell\left(\hat{\theta}_n^{(4)} \mid X\right) = \max_{\theta \in \mathbb{R}} \ell(\theta \mid \boldsymbol{X})$$
where ℓ is the log-likelihood function.

1. Derive the likelihood function and log-likelihood function.
2. Simulate data from Cauchy distribution with location 5, and scale 1.
3. Choose your number of simulations.
4. Verify consistency of the estimators. There are different approaches. You can samples the data sequentially and plot the sequence of the results as a function of n. What do you observe if it is a consistent estimator? The second approach is to use only the representative increasing values of the sample size. e.g. use $n = 10, 50, 100, 200, \ldots, 1000$ and what do you observe?
5. Calculate the mean square error of the estimators.
6. Calculate the coverage probability of the estimators. Calculate $\mathbf{P}_\theta\left(\left|\hat{\theta}_n - \theta\right| \leq \varepsilon\right)$, for $\varepsilon = 0.1$, and $\theta = 5$.