Problem Set 3[1]
Math Camp 2024, UCSB
Instructor: Seonmin (Will) Heo (sheo@ucsb.edu)
Due: 11:59 PM on Monday, Sep 9, 2024

This problem set will help you review the key concepts from the course so far. You are free to use any software but make sure to type the answers and include the code and the results.

1. **Chebychev's Inequality**
   Let $\mathbf{X}$ be a random vector and let $r > 0$ be a positive real number. Assume that $g : \mathbb{R}^n \to \mathbb{R}_+$ takes nonnegative values.

   (a) Check that the following inequality holds for any realization of random vector $\mathbf{X}$.

   $$\frac{g(\mathbf{X})}{r} \geq \mathbb{1}\{g(\mathbf{X}) \geq r\}$$

   (b) Check that

   $$\mathbb{E}\Big[\mathbb{1}\{g(\mathbf{X}) \geq r\}\Big] = P\left(g(\mathbf{X}) \geq r\right)$$

   (c) Check that

   $$\mathbb{E}\left[\frac{g(\mathbf{X})}{r}\right] \geq P\left(g(\mathbf{X}) \geq r\right)$$

   (d) Using (c), show that

   $$\lim_{n\to\infty} \mathbb{E}[\hat{\theta}_n - \theta]^2 = 0 \quad \Rightarrow \quad P\left\{\left|\hat{\theta}_n - \theta\right| < \epsilon\right\} \overset{n\to\infty}{\longrightarrow} 1 \quad \text{for any} \quad \epsilon > 0$$

   Hint: Let $g(\mathbf{X}) := \left(\hat{\theta}_n - \theta\right)^2$ and $r := \epsilon^2$. In addition,

   $$\left(\hat{\theta}_n - \theta\right)^2 < \epsilon^2 \quad \Leftrightarrow \quad \left|\hat{\theta}_n - \theta\right| < \epsilon$$

   Use Sandwich theorem.

2. Let $X_1, \cdots, X_n$ be a random sample from a distribution with PMF

   $$f(x_i|\theta) = \begin{cases} \theta(1-\theta)^{x_i - 1} & \text{if } x_i = 1, 2, \cdots \\ 0 & \text{otherwise} \end{cases}$$

   where $\theta \in (0, 1)$.

   - Find the method of moments estimator for θ.
     Hint: Think about $\mathbb{E}[X] - (1-\theta)\mathbb{E}[X]$.
   - Find the maximum likelihood estimator for θ.

[1] This problem set is created by Woongchan Jeon and modified by Seonmin (Will) Heo.

3. Suppose a coin flip lands on heads with probability θ.

$$f(x_i|\theta) = \begin{cases} \theta & \text{if } x_i = 1 \quad \text{(Head)} \\ 1-\theta & \text{if } x_i = 0 \quad \text{(Tail)} \\ 0 & \text{otherwise} \end{cases}$$

Let $X_1, \cdots, X_n$ be a random sample and let X be a random vector with the same probability distribution as X_i's.

- Apply CLT and derive the probability distribution of $\sqrt{n}(\bar{X}_n - \theta)$.
- (Frequentist Approach) Two mutually exclusive hypotheses on θ are given by

$$H_0 : \theta_0 \in \{.5\} \quad \text{and} \quad H_1 : \theta_A \in \{.8\}$$

  Consider the following test function with level $\alpha = .05$:

$$T_n(\alpha) = \mathbb{I}\left\{ \left| \frac{\sqrt{n}\left(\bar{X}_n - \theta\right)}{\sqrt{\theta(1-\theta)}} \right| > q_{1-\frac{\alpha}{2}} \right\} \quad \text{where} \quad \Phi\left(q_{1-\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2}$$

  Assume that we observed 21 heads out of 32 tosses. What is your p-value and your decision?

- (Optional) (Bayesian Approach) Assume that your prior is given by

$$\pi(H_0) = \pi(H_1) = .5$$

  Assume that we observed 21 heads out of 32 tosses. What is your $\hat{\theta}_n$? What is the posterior probability on H_0? Do you think the evidence is in favor of the null? Hint: Derive the likelihood $P(\hat{\theta}_n|H_0)$ and $P(\hat{\theta}_n|H_1)$. Use Bayes' rule to derive the posterior.

4. You need to build the Monte Carlo of doing a Monte Carlo. We draw 200 independent random variables X_is from the standard normal distribution and calculate the mean $\bar{X}$. We generate m such samples and derive y which is the number of times that satisfies the condition $|\sqrt{200}\bar{X}| > 1.96$. Write a simulation(nSims=100) that uses m=[100, 200, 500, 1000, 2000, 4000, 6000, 8000, 10000] and save y/m for each simulation in each m in a matrix. y/m can be interpreted as the sample mean from the Bernoulli distribution. Make a plot that shows the theoretical and empirical means of y/m together from the simulations(nSims=100) against m. Measure the running time.