BIOS 7717
Homework Assignment 2

1. In the notes we considered a Monte Carlo method that uses the unit circle to estimate π. Devise and implement a Monte Carlo method that uses the unit sphere to estimate π. How does your estimator perform relative to the circle-based estimator? Why? (Note that a large Monte Carlo sample size will be required to get good estimates of the biases.)

2. The Pearson type VII distribution with free parameters μ and σ, and shape parameter $3/2$, has density

$$f(x) = \frac{1}{2\sigma}\left\{1 + \left(\frac{x-\mu}{\sigma}\right)^2\right\}^{-3/2}.$$

   This distribution has mean μ but no higher moments (σ is not a variance parameter; it is merely a scale parameter). Denote the distribution as $\mathcal{P}(\mu, \sigma)$.

   Design and implement a Metropolis–Hastings random walk algorithm to sample from the posterior $\pi(\mu, \sigma \mid \boldsymbol{y})$, where $\boldsymbol{y}$ is an iid sample from the $\mathcal{P}(\mu, \sigma)$ distribution. You will of course have to choose prior distributions; justify your choices. Include your Metropolis–Hastings ratio in your write-up. Use R package `pearson7` to simulate data for testing your implementation.

   Finally, use your software to analyze the dataset provided on our course page (filename `pearson7.rData`). Present the results of your analysis, concisely. Include appropriate plots. Report Monte Carlo standard errors. How much tuning was required to get good acceptance rates?

3. Carry out a simulation study to determine how well your Bayesian approach from Problem 2 performs relative to maximum likelihood inference for the $\mathcal{P}(\mu, \sigma)$ distribution. Consider bias, variance, mean squared error, and running time.

4. Note that $(\text{med}_n, \sqrt{3}\,\text{mad}_n)'$ is $\sqrt{n}$-consistent for $(\mu, \sigma)'$, where med_n denotes the sample median and mad_n denotes the sample median absolute deviation from the median. How well does this estimator perform relative to those considered in Problems 2 and 3?

We will solve the following problem together during class. I have included the description here so that you will have a chance to ponder the setup beforehand.

Bonus: Here we consider a Gaussian copula model having Kumaraswamy marginals. The stochastic form of the model is given by

$$\begin{aligned} \boldsymbol{Z} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \\ \boldsymbol{U} &= \Phi(\boldsymbol{Z}) \\ \boldsymbol{Y} &= F^{-1}(\boldsymbol{U}), \end{aligned}$$

where $\boldsymbol{\Omega}$ is a correlation matrix such that all off-diagonal elements equal $\rho \in (-1, 1)$, and F^{-1} is the quantile function for the Kumaraswamy distribution. The pdf, cdf, and quantile function for this distribution are given by

$$\begin{aligned} f(x) &= abx^{a-1}(1-x^a)^{b-1} \\ F(x) &= 1 - (1-x^a)^b \\ F^{-1}(p) &= \{1 - (1-p)^{1/b}\}^{1/a}, \end{aligned}$$

where $x \in (0, 1)$ and $a, b > 0$.

Design and implement a Metropolis–Hastings algorithm to sample from the posterior $\pi(a, b, \rho \mid \boldsymbol{y})$, where $\boldsymbol{y}$ is a sample from the model. Note that the log likelihood is given by

$$\ell(a, b, \rho \mid \boldsymbol{y}) = -\frac{1}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2}\boldsymbol{z}'(\boldsymbol{\Omega}^{-1} - \mathbf{I})\boldsymbol{z} + \sum_{i=1}^{n} \log f(y_i),$$

where $z_i = \Phi^{-1}\{F(y_i)\}$. Apply your method to the data stored in file `Kumaraswamy.rData`.