Second homework assignment

PhD in Business Economics | Advanced Bayesian Econometrics
Hedibert Freitas Lopes | Due date: 9am, September 10th, 2024.

Please submit either your file (handwritten or typed) in PDF or HTML. The file must be a single PDF/HTML document for submission to me at `hedibertfl@insper.edu.br`. Students should follow the deadlines for submissions. This homework assignment should be done individually.

PART I: Monte Carlo integration

Let θ follow a Gamma distribution with parameters a and b, denoted by $\theta|a,b \sim G(a,b)$, such that its probability density function is given by

$$p(\theta|a,b) = \frac{b^a}{\Gamma(a)}\theta^{a-1}\exp\{-b\theta\},$$

for $\theta, a, b > 0$.

i) Show that $\mu = E(\theta|a,b) = a/b$ and $\sigma^2 = V(\theta|a,b) = a/b^2$;

ii) For $(a,b) \in \{(2,1),(2,2),(3,2),(2,3)\}$, plot $p(\theta|a,b)$ and compute μ and σ^2;

iii) Sample M draws from $p(\theta|a,b)$, for the four pairs (a,b) given in ii), and for increasing Monte Carlo size:

$$M \in \{1000, 10000, 100000\}.$$

Let the samples be summarized as $\{\theta^{(1)},\ldots,\theta^{M}\}$. Approximate μ and σ^2 by their Monte Carlo integration approximation:

$$\mu_{MC} = \frac{1}{M}\sum_{i=1}^{M}\theta^{(i)} \quad \text{and} \quad \sigma^2_{MC} = \frac{1}{M}\sum_{i=1}^{M}(\theta^{(i)} - \mu_{MC})^2,$$

respectively. Compare the MC-based approximated integrals, μ_{MC} and σ^2_{MC}, to the exact, closed form values, μ and σ^2;

iv). To verify empirically that the MC approximations become increasingly closer to the integrals (μ_{MC} and σ^2_{MC}), let us repeat iii) $R = 100$ times (a replication study) and store μ_{MC} and σ^2_{MC}:

$$\{\mu^{(1)}_{MC},\ldots,\mu^{(R)}_{MC}\} \quad \text{and} \quad \{\sigma^{2(1)}_{MC},\ldots,\sigma^{2(R)}_{MC}\},$$

for each value of M and pair (a,b). *Hint:* Comparing the boxplots of the replications as M increases will highlight the decreasing variability of the MC approximation.

Note: You can easily find textbooks and wikipedia-like pages deriving the law of large numbers and central limit theorems for various MC-based integration approximation, at various levels of formality.

PART II: Monte Carlo Sampling

Here we introduce algorithmically the MC methods commonly known as SIR. The idea is to obtain draws from a *target* density, say $\pi(\theta)$, based on resampled draws from a *candidate* density, say $q(\theta)$:

STEP 1: Sampling from the candidate density $q(\theta)$

$$\text{Sample } \{\widetilde{\theta}^{(1)}, \ldots, \widetilde{\theta}^{(M)}\} \text{ from } q(\theta);$$

STEP 2: Computing resampling weights

$$w_i \equiv w(\widetilde{\theta}^{(i)}) = \frac{\pi(\widetilde{\theta}^{(i)})}{q(\widetilde{\theta}^{(i)})} \qquad i = 1, \ldots, M.$$

STEP 3: Resampling
Sample N draws from the discrete set $\{\widetilde{\theta}^{(1)}, \ldots, \widetilde{\theta}^{(M)}\}$ with weights $\{w_1, \ldots, w_M\}$.
Call the resampled draws $\{\theta^{(1)}, \ldots, \theta^{(M)}\}$.

A few rule of thumbs:

1. The performance of the MC approximation depends on $q(\theta)$ being able to envelop $\pi(\theta)$;
2. The closer $q(\theta)$ is to $\pi(\theta)$, the better;
3. One need to be able to sample from $q(\theta)$ and point-wise evaluate both $q(\theta)$ and $\pi(\theta)$, up to normalizing constants.

Let us practice?

Let $\pi_A(\theta) \propto \exp\{-0.5\theta^2\}$ and $\pi_B(\theta) \propto \exp\{-|\theta|\}$, for $\theta \in \Re$. It is easy to see that these are the kernels of the standards normal and Laplace (double-exponential) densities, respectively. The normalizing constants are $(2\pi)^{-1/2}$ and $1/2$ respectively, but unnecessary for the implementation of the SIR scheme. As proposal density, let us use a Student's t with $\nu = 3$ degrees of freedom, i.e. $q(\theta) \propto (1 + \theta^2/\nu)^{-(\nu+1)/2}$.

i) Let $M = 5.000$ and $N = 1.000$ and obtain draws from π_A and π_B.

ii) Compute the SIR-based approximations to $(E_{\pi_A}(\theta), V_{\pi_A}(\theta))$ and $(E_{\pi_B}(\theta), V_{\pi_B}(\theta))$. It is easy to see that the exact values are $(0, 1)$ and $(0, 2)$, respectively.

iii) What happens when instead $q(\theta)$ is a uniform distribution in $(-10, 10)$? Probably, M will need to be much bigger, say $M = 50.000$, because this new proposal might be less efficient than the Student's t one.

iv) What if instead we want to approximate via SIR tail areas $P_A(\theta > 2)$ and $P_B(\theta > 2)$? M and N might need to be bigger. Check it out. By the way, the exact value of the tail probabilities are 0.02275013 and 0.06766196, respectively.