PROJECT SUBMISSION INSTRUCTIONS

Please complete and save your project in pdf file. Before submission, please use the following naming convention:

`firstname-lastname`-P1.pdf

Your pdf file should be submitted to the e-mail address below:

`BIOS276.31r8w8ndztx2a19n@u.box.com`

Please do not use my personal e-mail address. After submission you should receive confirmation of successful uploading. I will let you know if I am missing your project by the day following the due date.

Problems

1. Consider $(X_1, ..., X_n) \sim_{iid} Gamma(\alpha, \beta)$.
(a) Find the method of moments estimators of (α, β).
(b) Find the maximum likelihood estimator of β in closed form.
(c) Consider the following data set:
$\mathbf{x} = (1.33, 1.60, 0.68, 0.28, 1.22, 0.72, 0.16, 0.32, 0.97, 0.46)$;
Plot the likelihood function as a function of α, given $\beta = \hat{\beta}$ and explain how you would find the maximum likelihood estimate of α using the Newton/Raphson procedure.

2. Consider $X_1, ..., X_n \sim_{iid} N(\mu, \sigma^2)$, with joint prior $(\mu, \sigma^2) \sim N(\varepsilon, \sigma^2/\lambda_\mu) \times IG(\lambda_\sigma, \alpha)$.
(a) Derive the parameters of the posterior distribution of (μ, σ^2)
(b) Show that the marginal prior on μ is $T(2\lambda_\alpha, \varepsilon, \alpha/\lambda_\mu\lambda_\sigma)$. [1]
(c) Give the corresponding marginal prior on σ^2.

3. For each of the following situations exhibit a conjugate (Diaconis, Ylvisaker) family for the given distribution and compare to the Jeffrey's prior:
i) $Y \sim Binomial(\theta)$
ii) $Y \sim Poisson(\theta)$

4. Consider the (un-normalized) density

$$f(x) \propto exp(-x^2/2)\{sin(6x)^2 + 3cos(x)^2 sin(4x)^2 + 1\}$$

(a) Plot $f(x)$ and show that it can be bounded by $Mg(x)$, where $g = exp(-x^2/2)/\sqrt{(2\pi)}$. Find the an acceptable, if not optimal value for M. (Hint: use `optimize()`).
(b) Generate 2500 random samples from f using the Accept-Reject Algorithm.
(c) Deduce, from the acceptance rate of this algorithm an approximation to the normalizing constant of f.

5. The Laplace approximation is often useful to obtain fast approximate evaluation of common integrals. In general, a first degree approximation is obtained writing

$$f(x) = \exp n\, h(x),$$

with n is a parameter allowed to go to ∞. Expanding $h(x)$ around its maximum $\hat{x}$

$$h(x) \approx h(\hat{x}) + \frac{(x-\hat{x})^2}{2} h^{''}(\hat{x}) + R(x),$$

where $\lim_{x\to\hat{x}} R(x) = 0$ and representing

$$\int_a^b f(x)dx \approx \exp\{n\, h(\hat{x})\} \int_a^b \exp\left\{n\frac{(x-\hat{x})^2}{2} h^{''}\hat{x}\right\} dx$$

[1] A random variable $x \sim T(\nu, \theta, \sigma^2)$ if $P(x|\nu, \theta, \sigma^2) = \frac{\Gamma((\nu+1)/2)}{\sigma\sqrt{\nu\pi}\Gamma(\nu/2)} \left(1 + \frac{(x-\theta)^2}{\nu\sigma^2}\right)^{(\nu+1)/2}$

you should know where to go from here.

(a) Find the Laplace approximation for the following integrals:

$$(i)\int_a^b \frac{x^{k/2-1}\exp\{-x/2\}}{2^{k/2}\Gamma(k/2)}; \quad \hat{x}=\max(k-2,0).$$

$$(ii)\int_a^b x^{\alpha-1}(1-x)^{\beta-1}; \quad \hat{x}=\frac{\alpha-1}{\alpha+\beta-2}, \quad \text{for } \alpha,\beta>1$$

(b) The integrand in (a.ii) defines the kernel of a $Beta(\alpha,\beta)$ distribution. Use the Laplace approximation to find the appropriate normalizing constant for $(\alpha=2,\beta=2)$ and $(\alpha=1,\beta=2)$, and compare with the exact evaluation of such constant.

6. The two dimensional Banana distribution (Haario, 1999) is defined as:

$$\pi(x_1,x_2)\propto \exp\left\{-\frac{x_1^2}{2}\right\}\exp\left[-\frac{\left\{x_2-2\left(x_1^2-5\right)\right\}^2}{2}\right]; \qquad x_1,x_2\in\mathbb{R}.$$

(a) Describe how to simulate directly from $\pi(x_1,x_2)$.

(b) Describe how to simulate from $\pi(x_1,x_2)$ using the accept-reject method. Implement your algorithm and discuss sampling efficiency in relation to the chosen instrument distribution.

(c) Describe how to simulate from $\pi(x_1,x_2)$ using sampling importance resampling. Implement your algorithm and discuss sampling efficiency in relation to the chosen importance distribution.

(g) Use the algorithms in (a, b, c) to estimate the following quantities: $E(x_1^2), E(x_2)$ and $Pr(x_1+x_2>0)$. Discuss and quantify how different sampling strategies affect Monte Carlo variance.