MI 602 A - Métodos Computacionais em Estatística
Primeiro semestre de 2012
Lista de Exercícios II
Data da entrega: 23/04/2012 (todas as questões, exceto a primeira)

1. Considere que X é uma v.a. que corresponde à uma mistura finita de duas distribuições normais independentes, ou seja, tal que

$$f(x;\boldsymbol{\theta}) = p f_{X_1}(x;\mu_1,\sigma_1^2) + (1-p) f_{X_2}(x;\mu_2,\sigma_2^2),$$

$$\boldsymbol{\theta} = (p,\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)), p \in (0,1), \mu_i \in \mathcal{R}, \sigma_i^2 \in \mathcal{R}^2, i = 1,2$$

em que $f_{X_i}(.;.)$ é a fdp de uma distribuição $N(\mu_i,\sigma_i^2), i = 1,2$.

Proponha um algoritmo para simular uma amostra aleatória de tamanho n da distribuição acima e o implemente em R.

2. Em relação à distribuição da questão 1, proponha um algoritmo para calcular $h = \displaystyle\int_a^b f_{(x;\boldsymbol{\theta})}(x)dx$, usando integração por quadratura (não-adaptativa), e o implemente em R. Usando o programa implementado em R, calcule a integral em questão para os seguintes pares (a,b): $(-\infty,\infty)$, $(-1,0)$, $(0,1)$ e $(5,10)$, para $(\mu_1 = 2, \mu_2 = -1, \sigma_1^2 = 2, \sigma_2^2 = 8)$. Considere que $p \in \{0,3;0,5;0,7\}$. Compare com os valores verdadeiros.

3. Em relação à distribuição da questão 1, proponha um algoritmo para calcular $h = \displaystyle\int_{-\infty}^{\infty} w(x) f_{(x;\boldsymbol{\theta})}(x)dx$, usando integração por quadratura adaptativa, e o implemente em R. Usando o programa implementado em R, calcule a integral em questão para $w(x) = 1, w(x) = x, w(x) = x^2$ e considere que $p \in \{0,3;0,5;0,7\}$. Compare com os valores verdadeiros.

4. Considere a seguinte integral $h = \int_{1/2}^{2/3} x^2(1-x^2)dx$. Descreva um algoritmo para calculá-la via Monte Carlo e outro para calculá-la via amostragem por importância. Procure usar distribuições próximas nos dois métodos. Compare os valores entre si e com o verdadeiro valor para $m \in \{1000, 2000, 3000, ..., 10000\}$.

5. Considere a seguinte integral $h = \int_0^{1/3} \left(\int_0^{2/3} x^3 y^2 (1-x-y)^6 dx \right) dy$. Descreva um algoritmo para calculá-la via Monte Carlo e outro para calculá-la via amostragem por importância. Procure usar distribuições próximas nos dois métodos. Compare os valores entre si e com o verdadeiro valor para $m \in \{1000, 2000, 3000, ..., 10000\}$.

6. Seja $X \sim$ Poisson(λ), truncada à direita do zero, ou seja $f_X(x;\lambda) = \dfrac{e^{-\lambda}\lambda^x}{(1-e^{-\lambda})x!} \mathbb{1}_{\{1,2,\ldots\}}(x)$. Considere uma amostra aleatória de tamanho n desta distribuição. Responda os itens:

   a) Obtenha a log-verossimilhança, a função escore, a função Hessiana e a informação de Fisher.

b) Descreva o algoritmo Escore de Fisher, para a obtenção da estimativa de máxima verossimilhança (emv) e o implemente em R.

c) Considere que $n \in \{10, 30, 50, 200\}$ e $\lambda \in \{2, 7\}$. Para cada um dos tamanhos amostrais e valores de λ, simule R = 500 conjuntos de respostas com base no algoritmo construído na Lista I, obtenha a estimativa de máxima verossimilhança e o respectivo erro-padrão assintótico. Faça um histograma para cada um dos quatro conjuntos de estimativas. O que ocorre com a distribuição do emv, sua média e seu erro-padrão?

7. Seja $X \sim$ beta(a, b), ou seja $f_X(x; a, b) = \dfrac{1}{\beta(a, b)} x^{a-1}(1 - x)^{b-1}\mathbb{1}_{(0,1)}(x), a, b > 0$. Considere uma amostra aleatória de tamanho n desta distribuição e que (a, b) são desconhecidos. Responda os itens:

a) Obtenha a log-verossimilhança, o vetor escore, a matriz Hessiana e a informação de Fisher.

b) Descreva o algoritmo Escore de Fisher, para a obtenção da estimativa de máxima verossimilhança (emv) e o implemente em R.

b) Implemente o algoritmo BFGS usando a função optim no R.

c) Considere que $n \in \{10, 30, 50, 200\}$ e $(a, b) \in \{(1, 2), (2, 1)\}$. Para cada um dos tamanhos amostrais e valores de (a,b), simule R = 500 conjuntos de respostas (você pode usar a função *rbeta*), obtenha as estimativas de máxima verossimilhança (sando os algoritmos Escore de Fisher e BFGS) e os respectivos erros-padrão assintóticos. Faça um histograma para cada um dos quatro conjuntos de estimativas de (a,b). O que ocorre com as distribuições marginais doe emv, suas médias e seus erros-padrão?