Estatística Computacional - Lista 2

Gabriel Lima Novais

October 29, 2019

Exercise 1 (Monte Carlo for Gaussians)

(Q1) Vamos calcular a esperança de $\phi(x)$ em termos de $\pi(x)$:

$$E[\phi(x)] = \int \phi(x)\pi(x)dx = \int \phi(x)\frac{1}{(2\pi)^{d/2}}exp(-\frac{1}{2}x^T x)dx =$$

$$= \int \phi(z+\theta)\frac{1}{(2\pi)^{d/2}}exp(-\frac{1}{2}(z+\theta)^T(z+\theta))dz =$$

$$= \int \phi(z+\theta)\frac{1}{(2\pi)^{d/2}}exp(-\frac{1}{2}(z^T z+2\theta^T z+\theta^T\theta))dz =$$

$$= \int \phi(z+\theta)\frac{1}{(2\pi)^{d/2}}exp(-\frac{1}{2}z^T z)exp(-\frac{1}{2}\theta^T\theta-\theta^T z)dz =$$

$$= \int \phi(z+\theta)\pi(z)exp(-\frac{1}{2}\theta^T\theta-\theta^T z)dz =$$

$$= E[\phi(X+\theta)exp(-\frac{1}{2}\theta^T\theta-\theta^T X)]$$

Como queríamos demonstrar.

(Q2) Pela fórmula da variância e já elevando ao quadrado o termo da primeira esperança encontramos que :

$$\sigma^2(\theta) = E[\phi^2(X+\theta)exp(-\theta^T\theta-2\theta^T X)] - E[\phi(x)]^2$$

Veja que o primeiro termo acima pode ser escrito conforme a integral:

$$E[\phi^2(X+\theta)exp(-\theta^T\theta-2\theta^T X)] = \int \phi^2(x+\theta)exp(-\theta^T\theta-2\theta^T x)\frac{1}{(2\pi)^{d/2}}exp(-\frac{1}{2}x^T x)dx =$$

$$= \frac{1}{(2\pi)^{d/2}}\int \phi^2(x+\theta)exp(-\theta^T\theta-2\theta^T x)exp(-\frac{1}{2}x^T x)dx$$

Como queremos chegar em um resultado que possui o termo $X-\theta$, faremos uma substituição de variável tal que $x = z-\theta$ e trocaremos a derivada de dx para dz:

$$= \frac{1}{(2\pi)^{d/2}}\int \phi^2(z)exp(-\theta^T\theta-2\theta^T(z-\theta))exp(-\frac{1}{2}(z-\theta)^T(z-\theta))dz =$$

$$= \frac{1}{(2\pi)^{d/2}}\int \phi^2(z)exp(-\theta^T\theta-2\theta^T(z-\theta)-\frac{1}{2}(z-\theta)^T(z-\theta))dz =$$

$$= \frac{1}{(2\pi)^{d/2}} \int \phi^2(z) exp(-\theta^T\theta - 2\theta^T(z-\theta) - \frac{1}{2}z^Tz - \frac{1}{2}\theta^T\theta + \theta^T z)dz =$$

$$= \frac{1}{(2\pi)^{d/2}} \int \phi^2(z) exp(\frac{1}{2}\theta^T\theta - \frac{1}{2}z^Tz - \theta^T z)dz =$$

$$= \frac{1}{(2\pi)^{d/2}} \int \phi^2(z) exp(-\frac{1}{2}z^Tz) exp(\frac{1}{2}\theta^T\theta - \theta^T z)dz =$$

$$= E[\phi^2(X) exp(\frac{1}{2}\theta^T\theta - \theta^T X)]$$

E fazendo contas similares as distributivas acima chegamos a conclusão que:

$$E[\phi^2(X) exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))] = E[\phi^2(X) exp(\frac{1}{2}\theta^T\theta - \theta^T X)]$$

Então temos o resultado pedido, pois o segundo termo de $\sigma^2(\theta)$ é idêntico.

(Q3) Vamos efetuar a derivada no primeiro termo do resultado do item anterior uma vez que o segundo termo não está em função do θ. Então:

$$\nabla\sigma^2(\theta) = \nabla E[\phi^2(X) exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))] =$$
$$= E[\phi^2(X)(\theta - X) exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))]$$

Tendo sido feito a primeira derivada vamos calcular a Hessiana, onde denotaremos por $H(\theta)_{i,j}$ cada elemento dessa matrix. Vale ressaltar que como $\theta \in R^d$, temos que o parâmetro em questõa possui dimensão maior que 1, e assim teremos de fato uma matriz. Teremos dois termos principais: (1) elementos da Diagonal e (2) elementos fora da diagonal:

(1) $H(\theta)_{i,i} = E[\phi^2(X) exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))]$ para todo $i \in [1, d]$
(2) $H(\theta)_{i,j} = E[\phi^2(X)(\theta_i - X)(\theta_j - X)^T exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))]$ para todo $i \neq j \in [1, d]$

Portanto somando esses elementos e colocando a Identidade de tamanho $d \times d$ no primeiro elemento, temos o resultado abaixo:

$$H(\theta) = E[\phi^2(X) exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))]I_d + E[\phi^2(X)(\theta-X)(\theta-X)^T exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))]$$

Mas como os termos dentro dos valores esperados são ou multiplicações de termos que geram elementos positivos, ou é um exponencial, então temos que os valores serão sempre positivos. Logo para qualquer vetor $v \in R^d$, não nulo, sabemos que $v^T H(\theta) v > 0$, fazendo com que a matrix seja convexa.

(Q4) Pelos exercícios 2 temos que:

$$-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta)) = \frac{1}{2}\theta^T\theta - \theta^T X)$$

E pelo exercício 3 temos que:

$$\nabla\sigma^2(\theta) = E[\phi^2(X)(\theta - X) exp(-\frac{1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta))]$$

Logo:

$$\nabla\sigma^2(\theta) = E[\phi^2(X)(\theta - X)exp(\frac{1}{2}\theta^T\theta - \theta^T X))] = exp(\frac{1}{2}\theta^T\theta)E[\phi^2(X)(\theta - X)exp(-\theta^T X))]$$

Igualando a zero obteos o resultado pedido:

$$\nabla\sigma^2(\theta) = 0 \implies E[\phi^2(X)(\theta - X)exp(-\theta^T X))] = 0$$

(Q5) Tomemos o $\phi(x) = max(0, \lambda exp(\sigma X) - K)$. Então teremos que:

$$\nabla\sigma^2(\theta) = exp(\frac{1}{2}\theta^T\theta)E[\phi^2(X)(\theta - X)exp(-\theta^T X))] =$$

$$= exp(\frac{1}{2}\theta^T\theta)[\theta E[\phi^2(X)exp(-\theta^T X))] - E[\phi^2(X)Xexp(-\theta^T X))]] =$$

$$= exp(\frac{1}{2}\theta^T\theta)[\theta E[\phi^2(X)exp(-\theta^T X))] - exp(\frac{1}{2}\theta^T\theta)E[\phi^2(X)Xexp(-\theta^T X))]]$$

Tomando que $(\lambda exp(\sigma X) - K)^2 > 0$, temos que para os dois valores esperados acima $\phi^2(x) = (\lambda exp(\sigma X) - K)^2$ e assim o primeiro termo só será menor que o segundo configurando $\nabla\sigma^2(\theta) < 0$ apenas quando $\theta \leq 0$ ou $\theta \in D$, e assim cocnluímos o que foi pedido.

Exercise 2 (Metropolis-Hastings)

(Q1) O algoritmo da simulação pode ser feito tal como se segue:

1 Amostre $X \sim q(X_{t-1})$

2 Amostre $U \sim U[0,1]$

3 Use o condicional:
Se $U \leq \alpha(X_{t-1}, X)$ então $X_t = X$
Caso contrário : $X_{t-1} = X_t$

(Q2) Como temos que $\gamma(x,y) = \gamma(y,x)$:

$$\alpha(x,y) = \frac{\gamma(x,y)}{\pi(x)q(x,y)} \implies \alpha(x,y)\pi(x)q(x,y) = \gamma(x,y)$$

$$\alpha(y,x) = \frac{\gamma(y,x)}{\pi(y)q(y,x)} \implies \alpha(y,x)\pi(y)q(y,x) = \gamma(y,x)$$

$$\alpha(x,y)\pi(x)q(x,y) = \alpha(y,x)\pi(y)q(y,x)$$

O que conclui nossa demonstração.

(Q3) De acordo com a probabilidade de aceitação do MCMC temos que:

$$\alpha(x,y) = min(\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}, 1) \implies$$

$$\implies \frac{\gamma(x,y)}{\pi(x)q(x,y)} = min(\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}, 1) \implies$$

$$\implies \gamma(x,y) = \pi(x)q(x,y)min(\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}, 1) \implies$$

$$\gamma(x,y) = (\pi(y)q(y,x), \pi(x)q(x,y))$$

Veja que o resultado mostra que o MCMC corresponde a uma escolha particular do valor $\gamma(x,y)$.

(Q4) Perceba que conforme foi dado no próprio enunciado da questão temos que $X^{\tau_k} = Y^t$, ou seja, o Y^t representa propostas aceitas em determinado tempo e como isso não necessariamente é feito a todo momento , temos que em alguns t a aceitação não se realiza. Logo como $\tau_{k+1} - \tau_k$ representa o tempo extra até que aquela proposta seja aceita, temos que a soma se traduz em: $\frac{1}{\tau_k - 1} \sum_{i=1}^{k-1} (\tau_{i+1} - \tau_i)\phi(Y^i)$. Agora para achar o Kernel em questão basta observar que a probabilidade de aceitar algum valor na distribuição $q(x, y)$ é dado por $\alpha(x, y)q(x, y)$ e normalizando este valor obtemos o então Kernel citado.

(Q5) Para facilitar as contas suponha que tenhamos $S = \sum_{z \in X} \pi(z)m(z)$. Vamos calcular então $\tilde{\pi}(x)K(x, y)$ e verificar se a reversibilidade de fato ocorre:

$$\tilde{\pi}(x)K(x, y) = \frac{\pi(x)m(x)}{S} \frac{\alpha(x, y)q(x, y)}{m(x)} = \frac{\pi(x)\alpha(x, y)q(x, y)}{S}$$

Pelo item 2 sabemos que $\alpha(x, y)\pi(x)q(x, y) = \alpha(y, x)\pi(y)q(y, x)$, portanto temos que:

$$\tilde{\pi}(x)K(x, y) = \frac{\alpha(y, x)\pi(y)q(y, x)}{S}$$

Agora dividindo por $m(y)$ temos que:

$$\tilde{\pi}(x)K(x, y) = \frac{\pi(y)m(y)}{S} \frac{\alpha(y, x)q(y, x)}{m(y)} = \tilde{\pi}(y)K(y, x)$$

Justamente o que queríamos demonstrar.

(Q6) Dúvida!

Exercise 3 (Metropolis-Hastings)

(Q1) Veja que:

$$\int_X \pi(x)K(x, y)dx = \int_X \pi(x)[\alpha(x)q(y) + (1 - \alpha(x))\delta_x(y)]dx =$$
$$= \int_X \pi(x)\alpha(x)q(y)dx + \int_X (1 - \alpha(x))\delta_x(y)]dx =$$
$$= q(y) \int_X \pi(x)\alpha(x)dx + (1 - \alpha(y))\pi(y) \propto$$
$$\propto q(y) \int_X \frac{q(x)}{\alpha(x)}\alpha(x)dx + (1 - \alpha(y))\pi(y) \propto$$
$$\propto q(y) + (1 - \alpha(y))\pi(y) \propto$$
$$\propto q(y) + (1 - \alpha(y))\frac{q(y)}{\alpha(y)} \propto$$
$$\propto \frac{q(y)}{\alpha(y)} \propto$$
$$\propto \pi(y)$$

Como queríamos demonstrar.

(Q2) Pelo exercício anterior podemos concluir que $\pi(x) \propto \frac{q(x)}{\alpha(x)}$ mas como $\alpha(x) = \alpha$, então $\pi(x) \propto q(x)$. Agora vamos calcular o valor de $COV[X_1, X_k]$ e depois aplicar o limite pois sabemos que $V[x_1] = \sigma^2$, e assim obemos a resposta desejada. Então:

$$COV[X_1, X_k] = E[X_1X_k] - E[X_1]E[X_k] = E[E[X_1X_k|X_1, X_{k-1}]] - E^2[X_1]$$

Com probabilidade $(1-\alpha)$ temos que $X_t = X_{t-1}$ e com probabilidade α teremos que $X_t \sim q(.)$, ou seja teremos que X_t será independente de X_1 e de X_{t-1}. Portanto a esperança condicional será dada por:

$$E[E[X_1X_k|X_1, X_{k-1}]] = (1-\alpha)E[X_1X_{t-1}] - \alpha E[X_1]E[X_t] = (1-\alpha)E[X_1X_{t-1}] - \alpha E^2[X_1]$$

Logo teremos que:

$$COV[X_1, X_k] = (1-\alpha)(E[X_1X_{t-1}] - E^2[X_1]) = (1-\alpha)COV[X_1, X_{k-1}]$$

Fazendo a recursão e lembrando que $COV[X_1, X_1] = V[X_1] = \sigma^2$ chegamos no seguinte resultado: $COV[X_1, X_k] = (1-\alpha)^{k-1}\sigma^2$. Portanto fazendo as contas chegamos que:

$$\sigma_X^2 = \sigma^2 + 2\sum_{k=1}^{\infty}(1-\alpha)^k\sigma^2 = \sigma^2 + 2\frac{1-\alpha}{\alpha}\sigma^2 = \sigma^2(\frac{2-\alpha}{\alpha})$$

Ou seja, quando α vai a 0 teremos que o resultado vai para ∞.E assim concluímos o que desejávamos.

Exercise 4 (Gibbs Sampler)

(Q1) Pelo terorema de Bayes podemos calcular as condicionais da seguinte forma:

$$\pi(x|y) = \frac{\pi(x,y)}{\int \pi(x,y)dx}$$

$$\pi(y|x) = \frac{\pi(x,y)}{\int \pi(x,y)dy}$$

Então, vamos as contas de $\pi(x|y)$ e de forma similar teremos o outro resultado:

$$\pi(x|y) = \frac{\pi(x,y)}{\int \pi(x,y)dx} \propto \frac{\pi(x,y)}{\pi(y)} \propto \pi(x,y) \propto exp(\frac{-1}{2}(x-1)^2(y-2)^2) = exp(\frac{-1}{2(y-2)^{-2}}(x-1)^2)$$

Ou seja, temos que $y|x \sim N(1, 2(y-2)^{-2})$ e de froma similar teremos que $x|y \sim N(2, 2(x-1)^{-2})$

(Q2) Antes de qualquer coisa é preciso verificar se a densidade $\pi(x,y)$ de fato corresponde a uma distribuição de probabilidade, atendendo as propriedades desse tipo de função. Ou seja, integrando em x e y devemos encontrar o valor de 1. Mas veja que tomando $a = \frac{(y-1)^2}{2}$ e $b = 2$ temos que a integral de uma função Gaussiana arbitrária $\int exp(-a(x+b)^2)dx = \sqrt{\frac{\pi}{a}} = \sqrt{2\pi(y-2)^{-2}}$, que é justamente o valor de $\int \pi(x,y)dx$. Agora integrando novamente esse valor vemos que o mesmo diverge, de forma que o valor de $\int \sqrt{2\pi(y-2)^{-2}}dy \neq 1$. Portanto, não faz sentido.

Exercise 5 (Gibbs Sampler)

(Q1) Então temos que:

$$P(X_i = a_i) = \binom{m_i}{a_i}\theta_1^{a_i}(1-\theta_1)^{m_i-a_i}$$

e

$$P(Y_i = a_i) = \binom{n_i}{a_i}\theta_2^{a_i}(1-\theta_2)^{n_i-a_i}$$

Como as variáveis são independentes temos que para $Z_i = X_i + Y_i$ a sua distribuição será tal que:

$$P(Z_i = k_i) = P(X_i + Y_i = k_i) = \sum_{a_i=0}^{k_i} P(X_i = a_i)P(Y_i = k_i - a_i) =$$

$$= \sum_{a_i=0}^{k_i} \binom{m_i}{a_i}\theta_1^{a_i}(1-\theta_1)^{m_i-a_i}\binom{n_i}{k_i-a_i}\theta_2^{k_i-a_i}(1-\theta_2)^{n_i-k_i+a_i}$$

Logo temos que:

$$p(Z_i = z_i|\theta_1,\theta_2) = p(z_i|\theta_1,\theta_2) = \sum_{a_i=0}^{z_i} \binom{m_i}{a_i}\theta_1^{a_i}(1-\theta_1)^{m_i-a_i}\binom{n_i}{z_i-a_i}\theta_2^{z_i-a_i}(1-\theta_2)^{n_i-z_i+a_i}$$

Portanto, como queremos a distribuição conjunta devemos multiplicar todas as marginais de forma que obteremos o seguinte:

$$p(z_1, z_2, ..., z_T|\theta_1,\theta_2) = p(z_1|\theta_1,\theta_2)p(z_2|\theta_1,\theta_2)...p(z_T|\theta_1,\theta_2) =$$

$$= \prod_{i=0}^{T}\left[\sum_{a_i=0}^{z_i} \binom{m_i}{a_i}\theta_1^{a_i}(1-\theta_1)^{m_i-a_i}\binom{n_i}{z_i-a_i}\theta_2^{z_i-a_i}(1-\theta_2)^{n_i-z_i+a_i}\right]$$

E assim chegamos na expressão de verossimilhança da função pedida.

(Q2) Para a aplicação do Gibbs Sampler em $p(\theta_1,\theta_2|z_1,z_2,...z_T)$ precisamos entender que as amostras serão provenientes de $p(\theta_1,\theta_2,x_i,y_i|z_i)$, de forma que precisaremos calcular cada uma das marginais, isto é, precisaremos amostrar das seguintes densidades: $p(\theta_1|\theta_2,x_i,y_i,z_i) = p(\theta_1|x_i)$, $p(\theta_2|\theta_1,x_i,y_i,z_i) = p(\theta_2|y_i)$ e $p(x_i,y_i|\theta_1,\theta_2,z_i)$. Dessa forma, vamos verificar cada de suas formas:

$$p(\theta_1|X_1 = x_1, X_2 = x_2, ..., X_T = x_T) = p(\theta_1|x_1,x_2,...,x_T) \propto p(x_1,x_2,...,x_T|\theta_1)p(\theta_1)$$

Mas como sabemos que $\theta_1 \sim U[0,1]$ e como $p(x_1,x_2,...,x_T|\theta_1) = p(x_1|\theta_1)p(x_2|\theta_1)...p(x_T|\theta_1)$ então temos que :

$$p(\theta_1|x_1,x_2,...,x_T) \propto I_{[0,1]}(x)p(x_1|\theta_1)p(x_2|\theta_1)...p(x_T|\theta_1) \propto$$

$$\propto I_{[0,1]}(x)\binom{m_i}{x_1}\theta_1^{x_1}(1-\theta_1)^{m_i-x_1}\binom{m_i}{x_2}\theta_1^{x_2}(1-\theta_1)^{m_i-x_2}...\binom{m_i}{x_T}\theta_1^{x_T}(1-\theta_1)^{m_i-x_T} \propto$$

$$\propto I_{[0,1]}(x)\theta_1^{x_1}(1-\theta_1)^{m_i-x_1}\theta_1^{x_2}(1-\theta_1)^{m_i-x_2}...\theta_1^{x_T}(1-\theta_1)^{m_i-x_T} \propto$$

$$\propto I_{[0,1]}(x)\prod_{i=0}^{T}\theta_1^{x_1}(1-\theta_1)^{m_i-x_1} \propto I_{[0,1]}(x)\theta_1^{\sum_{i=0}^{T}x_1}(1-\theta_1)^{\sum_{i=0}^{T}m_i-x_1}$$

Com isso obtemos que:

$$(\theta_1|x_1, x_2, ..., x_T) \sim Beta(1 + \sum_{i=0}^{T} x_1, 1 + \sum_{i=0}^{T} m_i - x_1)$$

De forma similar temos que:

$$(\theta_2|y_1, y_2, ..., y_T) \sim Beta(1 + \sum_{i=0}^{T} y_1, 1 + \sum_{i=0}^{T} n_i - y_1)$$

Para a terceira densidade basta observar que que a mesma será igual ao produto das marginais, e como queremos para todos os i's então será o produtório das duas marginais. Isso nos fornecerá o seguinte resultado:

$$p(x_i, y_i|\theta_1, \theta_2, z_i) = \prod_{i=0}^{T} \binom{m_i}{k_i} \theta_1^{k_i}(1-\theta_1)^{m_i-k_i} \binom{n_i}{k_i} \theta_2^{k_i}(1-\theta_2)^{n_i-k_i}$$

Pronto, o único passo agora seria realizar uma amostragem dessa densidade acima, o que pode ser um problema pois ela não é tratável.

Simulation question (Normal mixture model — Gibbs sampling)

(Q1) Queremos mostrar que a densidade de X segue um modelo de mistura normal. Para isso precisamos utilizar a distribuição condicional $X_i|Z_i$ e a distribuição de Z_i. Dessa maneira, olhando para a soma da densidade conjunta de $X_i e Z_i$ chegamos em:

$$f(x_i|p, \mu, \sigma^2) = \sum_{k=1}^{K} f(x_i, z_i|p, \mu, \sigma^2)$$

Mas por Bayes, e levando em consideração que Z é discreto temos que:

$$f(x_i, z_i|p, \mu, \sigma^2) = P(Z_i = k|p, \mu, \sigma^2) f(x_i|Z_i = k, p, \mu, \sigma^2) = p_k \phi(x_i, \mu_i, \sigma_i^2)$$

E assim obtemos que:

$$f(x_i|p, \mu, \sigma^2) = \sum_{k=1}^{K} f(x_i, z_i|p, \mu, \sigma^2) = \sum_{k=1}^{K} p_k \phi(x_i, \mu_i, \sigma_i^2)$$

E obtemos o resultado desejado.

(Q2) De maneira similar ao que foi feito no item anterior obtemos que:

$$P(Z_i = z_i|x, p, \mu, \sigma^2) \propto P(Z_i = z_i|p, \mu, \sigma^2) f(x_i|z_i, p, \mu, \sigma^2) \propto p_k \frac{1}{\sqrt{2\pi\sigma_k^2}} exp(\frac{-1}{2\sigma_k^2}(x_i - \mu_k)^2)$$

E assim temos que:

$$P(Z = z|x, p, \mu, \sigma^2) \propto \prod_{i=0}^{N} p_{z_i} \phi(x_i, \mu_{z_i}, \sigma_{z_i}^2)$$

(Q3) Novamente utilizando Bayes vemos que a posteriori é dada por:

$$\pi(p|x, z, \mu, \sigma^2) \propto f(x, z|p, \mu, \sigma^2)\pi(p)$$

Pelos itens anteriores observamos que:

$$\pi(p|x, z, \mu, \sigma^2) \propto \pi(p)\prod_{i=0}^{N} p_{z_i}\phi(x_i, \mu_{z_i}, \sigma^2_{z_i}) \propto \pi(p)\prod_{i=0}^{N}\prod_{i=0}^{K} p_{z_i} \propto$$

$$\propto \pi(p)\prod_{i=0}^{K}(p_k)^{j_k} \propto \prod_{i=0}^{K}(p_k)^{j_k+\gamma_k-1}$$

Para a realização da amostragem basta fazer a amostra de uma Dirichlet com parâmetros $j_k + \gamma_k$

(Q4) Façamos de forma similar ao que foi feito no item anterior:

$$\pi(\mu|x, z, p, \sigma^2) \propto f(x, z|p, \mu, \sigma^2)\pi(\mu) \propto \prod_{i=0}^{K} exp(\frac{-1}{2\tau^2}(\mu_k - m)^2)\prod_{i=0}^{N} p_{z_i} exp(\frac{-1}{2\sigma^2_{z_i}}(x_i - \mu_{z_i})^2)$$

Agora observando o valor para μ_k temos que o mesmo possui a seguinte fórmula (basta realizar contas simples pelo que foi obtido na equação anterior ou ver que composição de normais gera uma normal):

$$\pi(\mu_k|x, z, p, \sigma^2) \propto exp(\frac{-1}{2}(\tau^{-2} + n_k\sigma_k^{-2})(\mu_k - \frac{(\tau^{-2}m + n_k s_k^{-2})}{(\tau^{-2} + n_k\sigma_k^{-2})})^2))$$

Veja que a mesma deverá possuir um formato de Gaussiana, e assim amostrar dela pode ser feito com técnicas relativamente simples, como por exemplo pelo algoritimo de aceitação e rejeição.

(Q5) Mais uma vez repetindo o procedimento dos itens anteriores chegamos em:

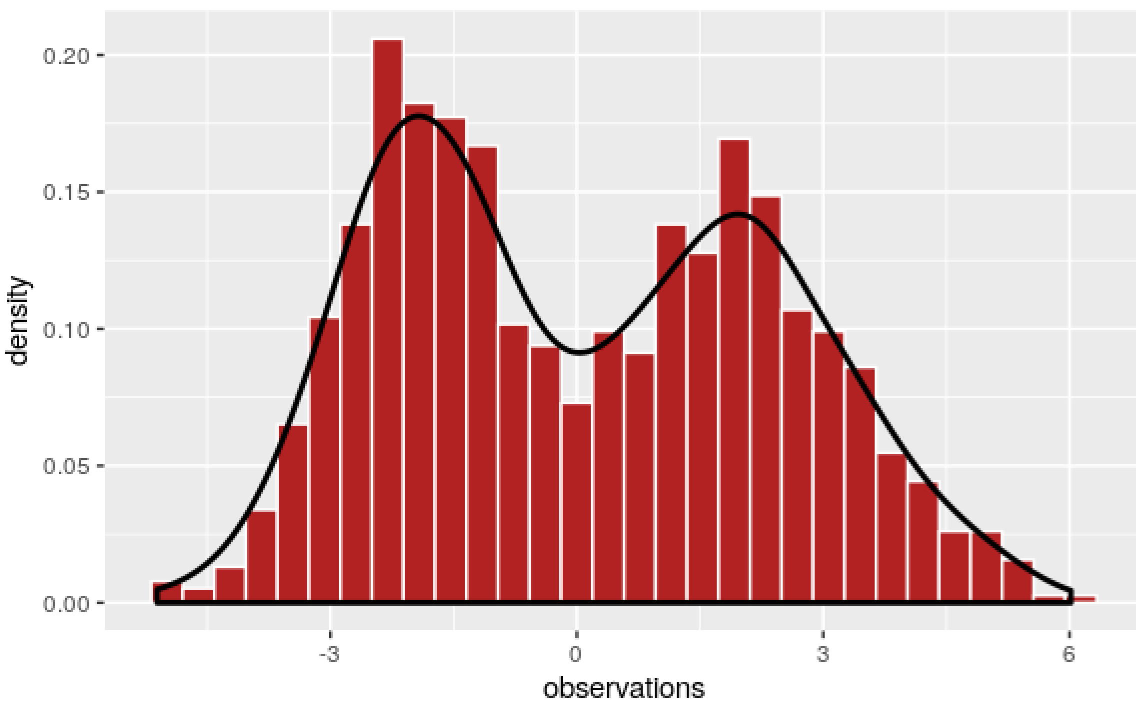
$$\pi(\sigma^2|x, z, p, \mu) \propto f(x, z|p, \mu, \sigma^2)\pi(\sigma^2) \propto \prod_{i=0}^{K}(\sigma_k^2)^{-\alpha-1} exp(-\beta\sigma_k^{-2})\prod_{i=0}^{N}\frac{1}{\sqrt{2\pi\sigma_k^2}} exp(\frac{-1}{2\sigma^2_{z_i}}(x_i - \mu_{z_i})^2)$$

$$\propto \prod_{i=0}^{K}(\sigma_k^2)^{-\alpha-1-\frac{j_k}{2}} exp(-\beta\sigma_k^{-2}) exp(\sum_{i=0}^{N}\frac{-1}{2\sigma^2_{z_i}}(x_i - \mu_{z_i})^2)$$

$$\propto \prod_{i=0}^{K}(\sigma_k^2)^{-\alpha-1-\frac{j_k}{2}} exp(-\beta\sigma_k^{-2} + \sum_{i=0}^{N}\frac{-1}{2\sigma^2_{z_i}}(x_i - \mu_{z_i})^2)$$

Veja que se trata de uma Gamma Inversa e assim amostrar dela seria basicamenrte amostrar de uma Gamma e inverter.

(Q6) Para este exercício foi realizada a construção de uma mistura de duas Gaussianas, e desta maneira, foram selecionadas duas médias (-2 e 2) e duas variâncias (1 e 2) de forma que cada ponto após ser escolhido de maneira aleatória entre uma das duas distribuições (fato decorrente da escolha da variável Z, que possui uma distribuição de índices de acordo com uma probabilidade pré-estabelecida) foi amostrado de uma das Gaussianas com as respectivas médias e variâncias. A curva da densidade encontra-se descrita na figura abaixo:



(Q7) Para a escolha dos hiperparâmetros, foi realizada uma escolha mais subjetiva de forma que os valores fossem os mais comuns. Para isso os valores foram:

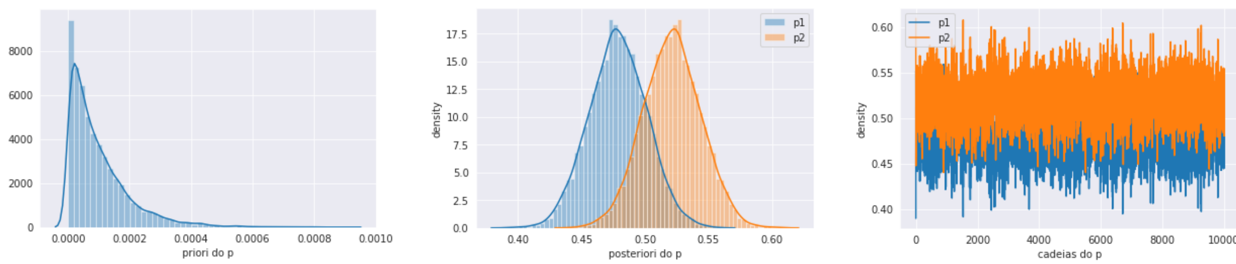
- $\gamma = 1$
- $m = 0$
- $\tau^2 = 1$
- $\alpha = 1$
- $\beta = 1$

Depois de ter sido efetuada a escolha dos hiperparâmetros o próximo passo foi a realização dos cálculos para encontrar as prioris e posterioris.

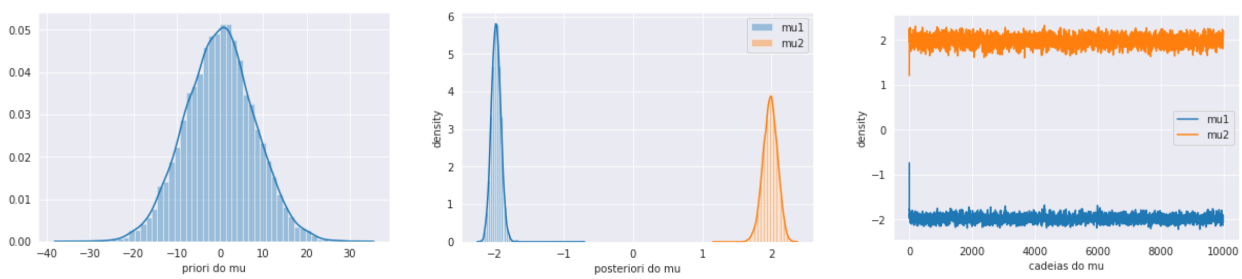
(Q8) Para realizar os itens pedidos no enunciado, foram necessárias a construção de quatro funções básicas para realizar a atualização das densidades condicionais marginais de cada uma das variáveis importantes (Z, μ, p, σ^2). Isso foi importante, pois como estamos lidando com o algoritmo de Gibbs sampling precisamos em cada iteração atualizar tais desnidades com as demais obtidas na iteração anterior, e assim realizar a amostragem. Os resultados obtidos podem ser conferidos nos gráficos abaixo, que representam em ordem por linha a seguinte estrutura: **Priori**, **Posterior** e **Cadeia**.

Vale resssaltar que as prioris seguiram as equações que estão disponíveis no próprio enunciado e dessa forma foram calculadas com base nas densidades de Dirichlet, Normal e Gamma Inversa.

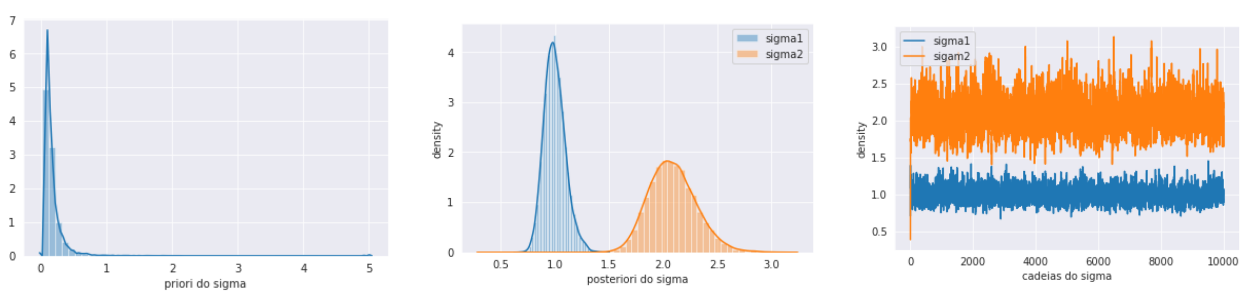
Para o parâmetro de probabilidades p temos que:



Agora para o parâmetro μ obtemos o seguinte resultado:



Finalmente para o parâmetro σ^2 o resultado foi o que se segue:



Com os resultados obtidos nos gráficos acima, podemos concluir que os valores previstos, com base no valor esperado das curvas, para os parâmetros pedidos, foram bem próximos do "real" construido no item 6.