

Lista de Exercícios

Inferência Bayesiana

SME 0809
XXXX

Problema 1

$$f(x_1, \dots, x_p \mid \theta_1, \dots, \theta_p) = \binom{n}{n_1, \dots, n_p} \prod_{i=1}^p \theta_i^{x_i}$$

A função de Verossimilhança

A função de verossimilhança é dada por:

$$\begin{aligned} L(\theta_1, \dots, \theta_p; \mathbf{x}) &= \binom{n}{n_1, \dots, n_p} \theta_1^{x_1} \cdots \theta_p^{x_p} \\ &\propto \prod_{i=1}^p \theta_i^{x_i}. \end{aligned}$$

A distribuição a *priori*

A fim de obter posteriori com expressão fechada e manipulação algébrica relativamente mais simples, utilizamos como priori uma distribuição Dirichlet(a_1, \dots, a_p). Assim

$$p(\theta_1, \dots, \theta_p) = \frac{\Gamma(\sum_{i=1}^p a_i)}{\prod_{i=1}^p \Gamma(a_i)} \theta_1^{a_1-1} \cdots \theta_p^{a_p-1} \propto \prod_{i=1}^p \theta_i^{a_i-1}$$

A distribuição a *posteriori* conjunta

A distribuição a posteriori conjunta para $\theta_1, \dots, \theta_p$ é dada por:

$$\begin{aligned} p(\theta_1, \dots, \theta_p \mid \mathbf{x}) &\propto L(\theta_1, \dots, \theta_p; \mathbf{x}) p(\theta_1, \dots, \theta_p) \\ &= \prod_{i=1}^p \theta_i^{x_i} \prod_{i=1}^p \theta_i^{a_i-1} \\ &= \prod_{i=1}^p \theta_i^{x_i+a_i-1}, \end{aligned}$$

portanto, $\theta_1, \dots, \theta_p \mid \mathbf{x} \sim \text{Dirichlet}(x_1 + a_1, \dots, x_p + a_p)$.

A distribuição a *posteriori* marginal

As distribuições marginais são dadas por

$$\begin{aligned}
p(\theta_1 | \mathbf{x}) &= \int_{\Theta^{p-1}} p(\theta_1, \dots, \theta_p | \mathbf{x}) d\boldsymbol{\theta}_{-1} \propto \theta_1^{x_1+a_1-1} \implies \theta_1 | \mathbf{x} \sim \text{Beta}(x_1 + a_1, 1) \\
&\vdots \\
p(\theta_i | \mathbf{x}) &= \int_{\Theta^{p-1}} p(\theta_1, \dots, \theta_p | \mathbf{x}) d\boldsymbol{\theta}_{-i} \propto \theta_i^{x_i+a_i-1} \implies \theta_i | \mathbf{x} \sim \text{Beta}(x_i + a_i, 1) \\
&\vdots \\
p(\theta_p | \mathbf{x}) &= \int_{\Theta^{p-1}} p(\theta_1, \dots, \theta_p | \mathbf{x}) d\boldsymbol{\theta}_{-p} \propto \theta_p^{x_p+a_p-1} \implies \theta_p | \mathbf{x} \sim \text{Beta}(x_p + a_p, 1),
\end{aligned}$$

em que $\boldsymbol{\theta}_{-i} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p)^T$.

Exemplo com dados simulados

Vamos definir, para efeitos de estudo, uma distribuição multinomial com parâmetros $(0.25, 0.25, 0.25, 0.25)$, isto é, classe equiprováveis. Assumindo uma priori Dirichlet com hiperparâmetros $(5, 5, 5, 5)$ temos uma posteriori ainda na classe Dirichlet, com atualização dos parâmetros.

Resumo a posteriori com valores numéricos dos parâmetros do modelo

Na Tabela 1 temos as medidas resumo das distribuições a posteriori marginais de cada parâmetro do modelo Multinomial.

Tabela 1: Medidas resumo da distribuição a posteriori dos parâmetros no modelo Multinomial.

Estatística	θ_1	θ_2	θ_3	θ_4
Mínimo	0.2193	0.2293	0.2015	0.2238
1st Quartil	0.2426	0.2583	0.2260	0.2464
Mediana	0.2494	0.2650	0.2325	0.2530
Média	0.2493	0.2647	0.2328	0.2531
3rd Quartil	0.2558	0.2712	0.2394	0.2594
Máxima	0.2782	0.2962	0.2601	0.2884
desvio-padrão	1.8229	1.9190	1.7822	1.9583

Podemos notar, com base na Tabela 1 que todos os parâmetros do modelo tem média muito próxima ao valor real do parâmetro. Além disso, podemos também perceber uma simetria nas estatísticas, assim sugerindo uma simetria ao redor da média amostral.

Gráfico da distribuição marginal a posteriori

Na Figuras (1) apresentamos os histogramas das distribuições a posteriori marginais.

Podemos notar que em todos os caso o parâmetro tem comportamento similar, ou seja, tende a ser simétrico com relação ao valor real do parâmetro, 0.25.

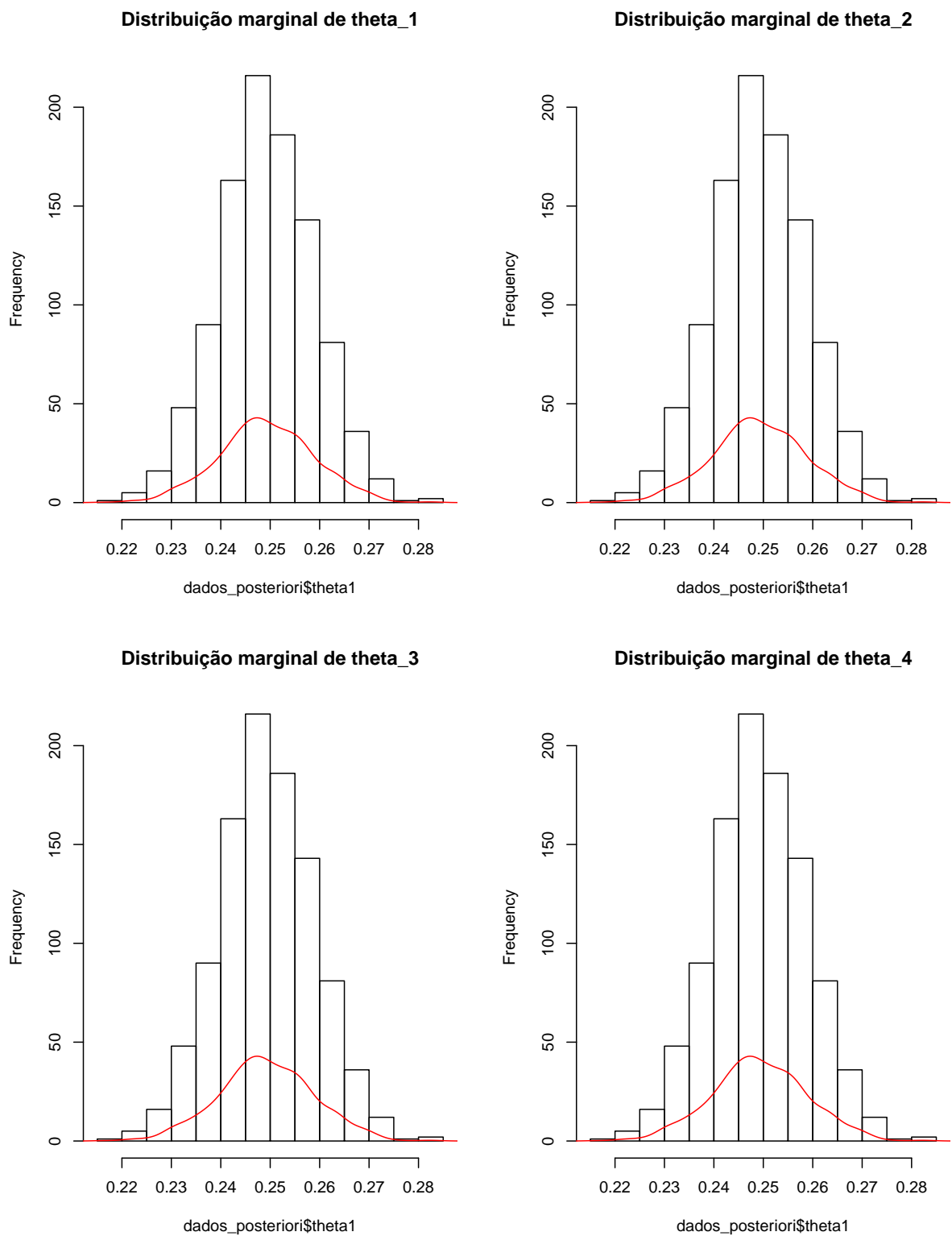


Figura 1: Histograma para amostras da distribuição a posteriori marginal para cada um dos parâmetros para o modelo Multinomial-Dirichlet.

Problema 2

$$Y_i = \alpha + \beta x_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2) \quad (1)$$

A função de Verossimilhança

A função de verossimilhança para o modelo (1) apresentado é dada por:

$$\begin{aligned} L(\alpha, \beta, \sigma^2, \mathbf{x}, \mathbf{y}) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_i - \alpha - \beta x_i)^2}{2\sigma^2} \right\} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n/2} \exp \left\{ -\sum_{i=1}^n \frac{(y_i - \alpha - \beta x_i)^2}{2\sigma^2} \right\} \end{aligned}$$

A distribuição a *priori*

Neste caso temos $\theta = (\alpha, \beta, \sigma^2)$. Vamos assumir priori que sejam conjugas com a verossimilhança, assim

$$\begin{aligned} \pi(\alpha) &= N(a_1, b_1) \\ \pi(\beta) &= N(a_2, b_2) \\ \pi(\sigma^2) &= \text{Gama Inversa}(c, d), \end{aligned}$$

desta forma a priori conjunta é dada por:

$$\begin{aligned} p(\alpha, \beta, \sigma^2) &= \frac{1}{\sqrt{2\pi b_1}} \exp \left\{ -\frac{(\alpha - a_1)^2}{2b_1} \right\} \frac{1}{\sqrt{2\pi b_2}} \exp \left\{ -\frac{(\beta - a_2)^2}{2b_2} \right\} \times \\ &\quad \times \frac{d^c}{\Gamma(c)} (\sigma^2)^{-c-1} \exp \left\{ -\frac{d}{\sigma^2} \right\} \end{aligned}$$

A distribuição a *posteriori* conjunta

$$\begin{aligned} \pi(\alpha, \beta, \sigma^2 \mid y_1, \dots, y_n) &\propto L(\alpha, \beta, \sigma^2, y_1, \dots, y_n) \pi(\alpha) \pi(\beta) \pi(\sigma^2) \\ &\propto (\sigma^2)^{-n/2} \exp \left\{ -\sum_{i=1}^n \frac{(y_i - \alpha - \beta x_i)^2}{2\sigma^2} \right\} \exp \left\{ -\frac{(\alpha - a_1)^2}{2b_1} \right\} \times \\ &\quad \times \exp \left\{ -\frac{(\beta - a_2)^2}{2b_2} \right\} (\sigma^2)^{-c-1} \exp \left\{ -\frac{d}{\sigma^2} \right\} \\ &\propto (\sigma^2)^{-n/2-c-1} \exp \left\{ -\frac{\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 - 2d}{2\sigma^2} - \frac{(\alpha - a_1)^2}{2b_1} - \frac{(\beta - a_2)^2}{2b_2} \right\} \end{aligned}$$

A distribuição a *posteriori* marginal

$$\pi(\sigma^2 \mid \alpha, \beta, \mathbf{x}, \mathbf{y}) \propto (\sigma^2)^{-n/2-c-1} \exp \left\{ -\frac{\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 - 2d}{2\sigma^2} \right\},$$

$$\text{portanto, } \sigma^2 \mid \alpha, \beta, \mathbf{x}, \mathbf{y} \sim \text{Gama Inversa} \left(c + \frac{n}{2}; \frac{\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 + 2d}{2} \right).$$

$$\begin{aligned} \pi(\alpha \mid \beta, \sigma^2, \mathbf{x}, \mathbf{y}) &\propto \exp \left\{ -\frac{\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2}{2\sigma^2} - \frac{(\alpha - a_1)^2}{2b_1} \right\} \\ &\propto \exp \left\{ -\sigma^{-2} \left[\frac{\sum_{i=1}^n (y_i^2 - 2y_i(\alpha + \beta x_i) + (\alpha + \beta x_i)^2)}{2} \right] - b_1^{-1} \frac{(\alpha^2 + 2\alpha a_1 + a_1^2)}{2} \right\} \\ &\propto \exp \left\{ -\sigma^{-2} \left[\frac{\sum_{i=1}^n (-2y_i\alpha + \alpha^2 + 2\alpha\beta x_i)^2}{2} \right] - b_1^{-1} \frac{(\alpha^2 + 2\alpha a_1)}{2} \right\} \\ &\propto \exp \left\{ \frac{-\alpha^2 n\sigma^{-2} - b_1^{-1}\alpha^2 + 2\alpha\sigma^{-2} \sum_{i=1}^n y_i - 2\alpha\sigma^{-2}\beta \sum_{i=1}^n x_i - 2\alpha b_1^{-1}a_1}{2} \right\} \\ &\propto \exp \left\{ \frac{-\alpha^2(n\sigma^{-2} + b_1^{-1}) + 2\alpha \left(\sigma^{-2} \sum_{i=1}^n y_i - \sigma^{-2}\beta \sum_{i=1}^n x_i - b_1^{-1}a_1 \right)}{2} \right\} \\ &\propto \exp \left\{ -(n\sigma^{-2} + b_1^{-1}) \frac{\alpha^2 - 2\alpha \left(\frac{n\sigma^{-2}\bar{y} - n\sigma^{-2}\beta\bar{x} - b_1^{-1}a_1}{n\sigma^{-2} + b_1^{-1}} \right)}{2} \right\} \\ &\propto \exp \left\{ -(n\sigma^{-2} + b_1^{-1}) \frac{\alpha^2 - 2\alpha \left(\frac{n\sigma^{-2}(\bar{y} - \beta\bar{x}) - b_1^{-1}a_1}{n\sigma^{-2} + b_1^{-1}} \right)}{2} \right\} \\ &\propto \exp \left\{ -\frac{\left[\alpha - \left(\frac{n\sigma^{-2}(\bar{y} - \beta\bar{x}) - b_1^{-1}a_1}{n\sigma^{-2} + b_1^{-1}} \right) \right]^2}{2(n\sigma^{-2} + b_1^{-1})} \right\} \end{aligned}$$

$$\text{portanto, } \alpha \mid \beta, \sigma^2, \mathbf{x}, \mathbf{y} \sim N \left[\frac{n\sigma^{-2}(\bar{y} - \beta\bar{x}) - b_1^{-1}a_1}{n\sigma^{-2} + b_1^{-1}}; (n\sigma^{-2} + b_1^{-1})^{-1} \right].$$

$$\begin{aligned}
\pi(\beta \mid \alpha, \sigma^2, \mathbf{x}, \mathbf{y}) &\propto \\
&\exp \left\{ \frac{-\sigma^2 (\sum_{i=1}^n [2y_i(\alpha + \beta x_i) + (\alpha + \beta x_i)^2])}{2} + \frac{-b_2^{-1}(\beta^2 - 2\beta a_2)}{2} \right\} \\
&\propto \exp \left\{ \frac{-\sigma^2 (-2\beta \sum_{i=1}^n y_i x_i + 2\alpha \beta \sum_{i=1}^n x_i + \beta^2 \sum_{i=1}^n x_i^2)}{2} \right\} \times \\
&\times \exp \left\{ \frac{-b_2^{-1}(\beta^2 - 2\beta a_2)}{2} \right\} \\
&\propto \exp \left\{ \frac{2\beta \sigma^{-2} \sum_{i=1}^n x_i y_i - 2\alpha \beta \sigma^{-2} \sum_{i=1}^n x_i - \beta^2 \sigma^{-2} \sum_{i=1}^n x_i^2 - b_2^{-1} \beta^2 - 2\beta b_2^{-1} a_2}{2} \right\} \\
&\propto \exp \left\{ \frac{-\beta^2 (\sigma^{-2} \sum_{i=1}^n x_i^2 + b_2^{-1}) + 2\beta (\sigma^{-2} (\sum_{i=1}^n x_i y_i - \alpha \sum_{i=1}^n x_i) - b_2^{-1} a_2)}{2} \right\} \\
&\propto \exp \left\{ - \left(\sigma^{-2} \sum_{i=1}^n x_i^2 + b_2^{-1} \right) \frac{\left[\beta^2 - 2\beta \left(\frac{\sigma^{-2} (\sum_{i=1}^n x_i y_i - \alpha \sum_{i=1}^n x_i) - b_2^{-1} a_2}{\sigma^{-2} \sum_{i=1}^n x_i^2 + b_2^{-1}} \right) \right]}{2} \right\} \\
&\propto \exp \left\{ - \frac{\left[\beta - \left(\frac{\sigma^{-2} (\sum_{i=1}^n x_i y_i - \alpha \sum_{i=1}^n x_i) - b_2^{-1} a_2}{\sigma^{-2} \sum_{i=1}^n x_i^2 + b_2^{-1}} \right) \right]^2}{2 (\sigma^{-2} \sum_{i=1}^n x_i^2 + b_2^{-1})^{-1}} \right\},
\end{aligned}$$

portanto, $\beta \mid \alpha, \sigma^2, \mathbf{x}, \mathbf{y} \sim N \left[\frac{\sigma^{-2} (\sum_{i=1}^n x_i y_i - \alpha \sum_{i=1}^n x_i) - b_2^{-1} a_2}{\sigma^{-2} \sum_{i=1}^n x_i^2 + b_2^{-1}}; (\sigma^{-2} \sum_{i=1}^n x_i^2 + b_2^{-1})^{-1} \right]$.

Exemplo com dados simulados

Para o exemplo com dados simulados, consideramos uma amostra de tamanho, $n = 100$, e valores para os parâmetros $\alpha = 2$, $\beta = 3$ e $\sigma^2 = 5$.

Resumo a posteriori com valores numéricos dos parâmetros do modelo

Na Tabela 2 apresentamos as medidas descritivas a posteriori para cada um dos do modelo Normal-Gama Inversa.

Pomode notar que a média do parâmetro σ^2 está consideravelmente longe do valor real. Já os parâmetros α e β apresenta médias amostrais muito próximas ao valor real. Além das médias, podemos notar que a distribuição do parâmetro σ^2 é fortemente assimétrica. Já os parâmetros α e β apresentam uma distribuição aproximadamente simétrica. Uma análise cuidadosa do histograma nos dará conclusões mais seguras sobre o comportamento de cada parâmetro.

Gráfico da distribuição marginal a posteriori

Na Figura (2) apresentamos o diagrama de dispersão dos dados simulado, x e y . Além disso, apresentamos o histograma da posteriori para cada

Tabela 2: Medidas resumo da distribuição a posteriori dos parâmetros Normal-Gama Inversa

Estatística	σ^2	α	β
Minimo	2.521	1.395	2.867
1st Quartil	3.382	1.937	2.989
Mediana	3.688	2.089	3.019
Média	3.708	2.090	3.018
3rd Quartil	3.971	2.226	3.047
Máxima	5.724	2.811	3.175
desvio-padrão	0.461	0.214	0.045

parâmetro do modelo Normal-Gama Inversa.

Podemos, notar que o comportamento entre x e y é difuso, porém, podemos ver claramente uma tendência de reta, relacionando x e y . A distribuição de σ^2 é claramente assimétrica, como mencionado na análise da Tabela 2, assim fica evidente o comportamento característico da Gama Inversa. Já os parâmetros α e β apresentam distribuições simétricas ao redor do seu respectivo valor real.

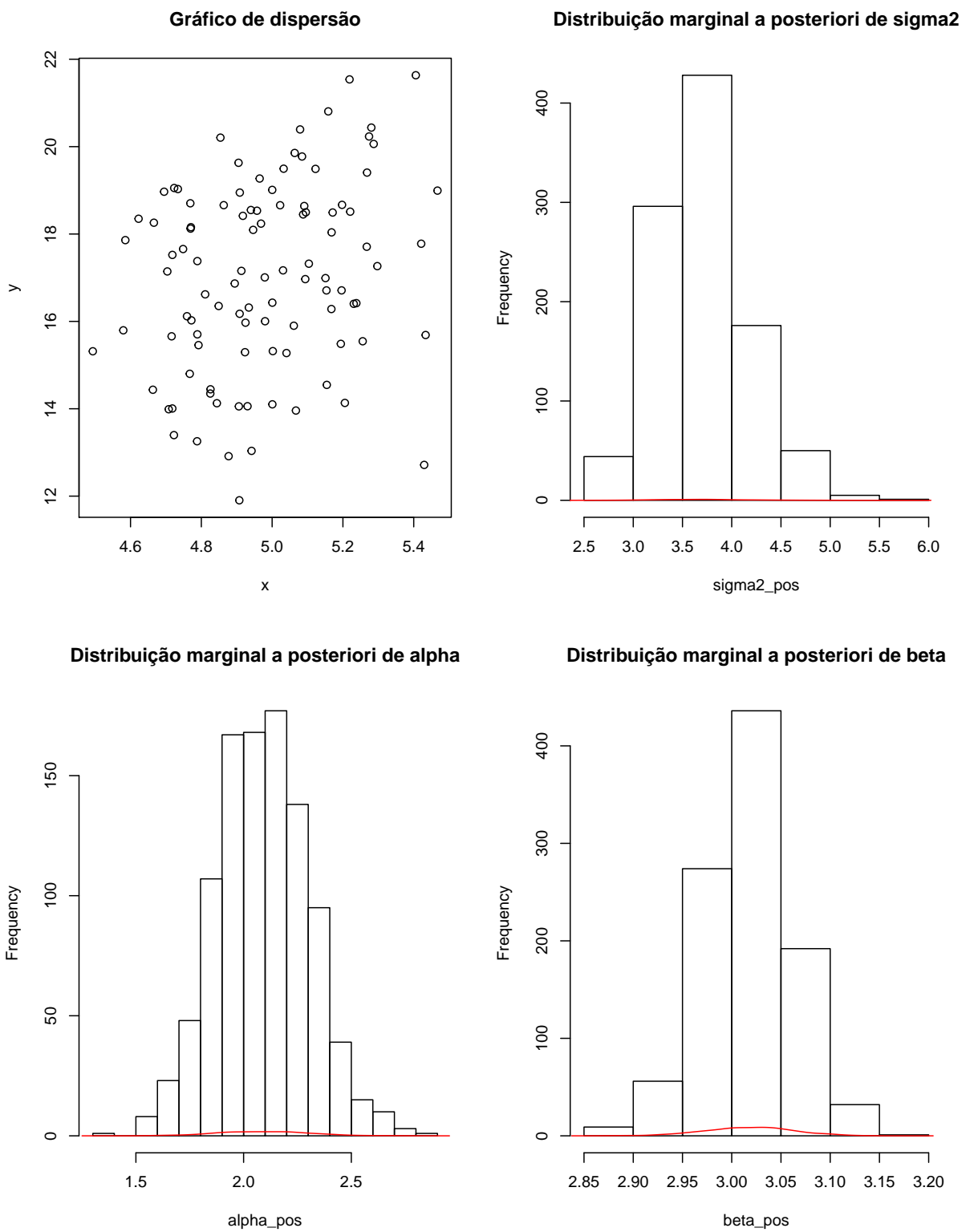


Figura 2: Histograma para amostras da distribuição a posteriori marginal para cada um dos parâmetros no modelo Normal-Gama Inversa.

```
#####
##### ANALISE BAYSIANA COM DADOS SIMULADOS #####
#####

##-----##
## MULTINOMIAL-DIRICHLET ##
##-----##

# vetor de parâmtros
theta <- rep(0.25, 4)

# tamanho da amostra
n0 <- 100

amostra <- rmultinom(n0, 20, theta)

# parametros da priori
hiper_parametros <- rep(5, 4)

# parametros da posteriori
parametros_pos <- hiper_parametros + c(sum(amostra[1,]), sum(amostra[2,]),
                                         sum(amostra[3,]), sum(amostra[4,]))

# amostra a posteriori
library(MCMCpack)

amostra_pos <- rdirichlet(1000, parametros_pos)

# medidas resumo
summary(amostra_pos)
round(sd(amostra[,1]), 4)
round(sd(amostra[,2]), 4)
round(sd(amostra[,3]), 4)
round(sd(amostra[,4]), 4)

# graficos
dados_posteriori <- as.data.frame(amostra_pos)
colnames(dados_posteriori) <- c("theta1", "theta2", "theta3", "theta4")
head(dados_posteriori)
```

```

par(mfrow = c(2, 2))

hist(dados_posteriori$theta1, main = "Distribuição marginal de theta_1")
lines(density(dados_posteriori$theta1), col = "red")

hist(dados_posteriori$theta1, main = "Distribuição marginal de theta_2")
lines(density(dados_posteriori$theta1), col = "red")

hist(dados_posteriori$theta1, main = "Distribuição marginal de theta_3")
lines(density(dados_posteriori$theta1), col = "red")

hist(dados_posteriori$theta1, main = "Distribuição marginal de theta_4")
lines(density(dados_posteriori$theta1), col = "red")

##-----##
## NORMAL-GAMA INVERSA ##
##-----##

# valores dos parâmetros
alpha <- 2
beta <- 3
sigma2 <- 5

x <- rnorm(100, 5, 0.2)
erro <- rnorm(100, 0, sqrt(sigma2))

y <- alpha + beta*x + erro

# parametros priori
a1 <- 2.1; b1 <- 100
a2 <- 2.9; b2 <- 100
c <- 10; d <- 2

n <- length(y)

# parametros poteriori -sigma2
c1 <- c + n/2

```

```

d1 <- d + sum(erro^2)/2

a1_pos <- ((sum(y -beta*x)/sigma2) -a1/b1)/(n/sigma2 + 1/b1)
b1_pos <- 1/(n/sigma2 + 1/b1)

a2_pos <- (((sum(y*x) - alpha*sum(x))/sigma2) -a2/b2)/(sum(x^2)/sigma2 + 1/b2)
b2_pos <- 1/(sum(x^2)/sigma2 + 1/b2)


sigma2_pos <- rinvgamma(1000, c1, d1)
alpha_pos  <- rnorm(1000, a1_pos, sqrt(b1_pos))
beta_pos   <- rnorm(1000, a2_pos, sqrt(b2_pos))


# Medidas resumo

summary(sigma2_pos)
round(sd(sigma2_pos), 3)

summary(alpha_pos)
round(sd(alpha_pos), 3)

summary(beta_pos)
round(sd(beta_pos), 3)


# Figuras
par(mfrow = c(2,2))

plot(x, y, main = "Gráfico de dispersão")

hist(sigma2_pos, main = "Distribuição marginal a posteriori de sigma2")
lines(density(sigma2_pos), col = "red")

hist(alpha_pos, main = "Distribuição marginal a posteriori de alpha")
lines(density(alpha_pos), col = "red")

hist(beta_pos, main = "Distribuição marginal a posteriori de beta")
lines(density(beta_pos), col = "red")

```

