Lista 6

Técnicas Computacionais em Estatística

Tailine J. S. Nonato

June 20, 2025

Lista 6 - MCMC

Exercício 1

Considere a distribuição Rayleigh, com densidade

$$f(x) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}, \quad x \geq 0, \ \sigma > 0.$$

Use o algoritmo Metropolis-Hastings para gerar amostras da Rayleigh quando $\sigma = 2$. Utilize como candidata $q(\cdot \mid x_t)$ o modelo $\chi^2_{x_t}$. Use um burn-in de 2000 e compute a taxa de aceitação. Compare, através de um QQ plot, os quantis amostrais e os quantis teóricos $x_q = F^{-1}(q) = \sigma \left\{-2\log(1-q)\right\}^{1/2}$, $0 < q < 1$.

Solução

Alvo: Rayleigh com $\sigma = 2$.

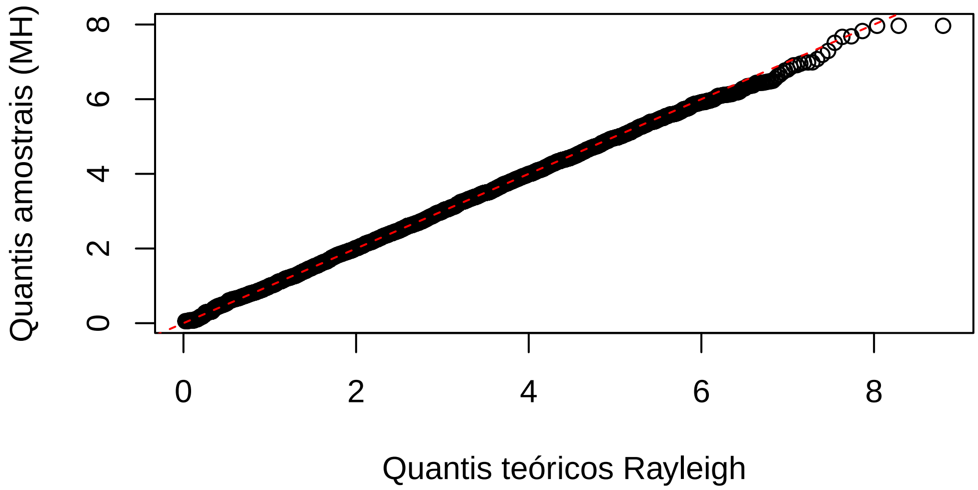
Candidata: $\chi^2_{x_t}$.

```
set.seed(451)

sigma <- 2
f_rayleigh <- function(x) {
  ifelse(x >= 0, (x / sigma^2) * exp(-x^2 / (2 * sigma^2)), 0)}

Nsim <- 10000
burnin <- 2000
```

```r
x <- numeric(Nsim)
x[1] <- 2  # chute inicial x0
aceito <- 0

for (t in 2:Nsim){
  Y <- rchisq(1, df = x[t-1])
  alpha <- min(1, f_rayleigh(Y) * dchisq(x[t-1], df = Y) /
  (f_rayleigh(x[t-1]) * dchisq(Y, df = x[t-1])))
  if (runif(1) < alpha) {
    x[t] <- Y
    aceito <- aceito + 1
  } else {
    x[t] <- x[t-1]
  }}

x_post <- x[-(1:burnin)]

aceitacao <- aceito/(Nsim-1)
cat("Taxa de aceitação:", round(aceitacao, 3), "\n")
```

```
Taxa de aceitação: 0.473
```

```r
rayleigh_q <- function(p, sigma = 2) {
  sigma * sqrt(-2 * log(1 - p))
}

probs <- ppoints(length(x_post))
teo_q <- rayleigh_q(probs, sigma = sigma)
q <- quantile(x_post, probs = probs)

qqplot(teo_q, q,
       xlab = "Quantis teóricos Rayleigh",
       ylab = "Quantis amostrais (MH)")
abline(0, 1, col = "red", lty = 2)
```



Exercício 2

Considere o modelo normal multivariado

$$(X_1, X_2, \dots, X_p) \sim N_p(0, (1-\rho)I + \rho J),$$

em que I é a matriz identidade $p \times p$ e J é uma matriz de uns $p \times p$. Esse é o chamado modelo de equicorrelação, uma vez que $\text{Cor}(X_i, X_j) = \rho$ para todo i e j.

Note que

$$X_i \mid x_{(-i)} \sim N\left(\frac{(p-1)\rho + (p-2)\rho\bar{x}_{(-i)}}{1+(p-2)\rho}, \frac{1+(p-2)\rho-(p-1)\rho^2}{1+(p-2)\rho}\right),$$

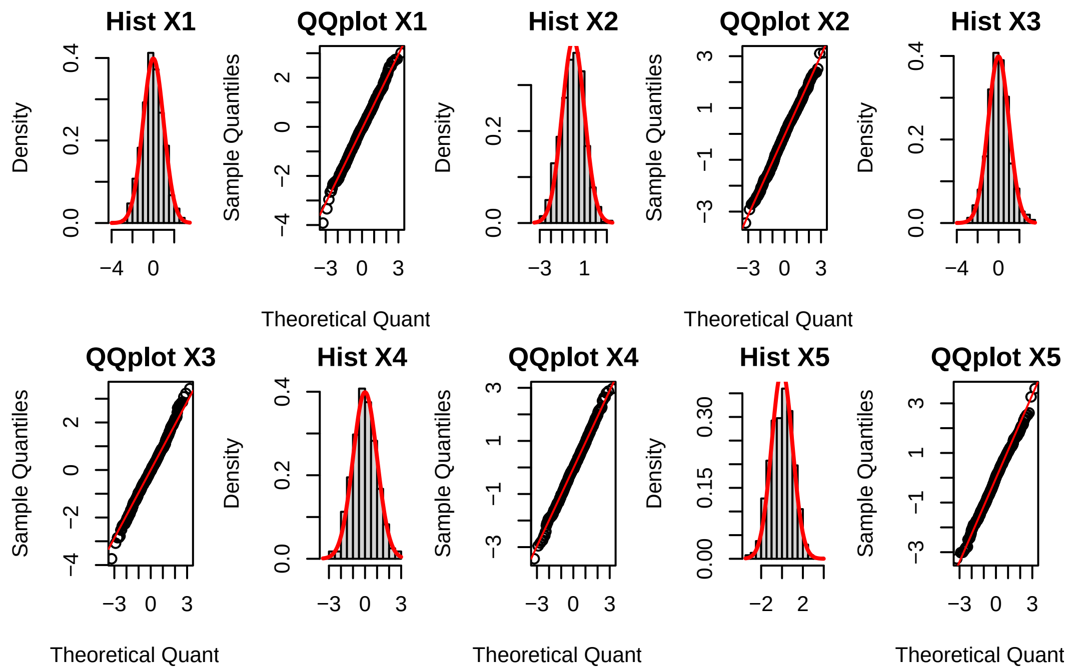
em que $x_{(-i)} = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$ e $\bar{x}_{(-i)}$ é a média desse vetor.

Implemente um amostrador de Gibbs usando a condicional acima. Execute o código em R para $p = 5$ e $\rho = 0{,}25$, e verifique graficamente que as marginais são todas $N(0, 1)$.

Solução

```r
p <- 5           # dimensão
rho <- 0.25      # equicorrelação

Nsim <- 1000
burnin <- 200

X <- matrix(0, nrow = Nsim, ncol = p)

X[1, ] <- rnorm(p)

mu_cond <- function(x_minus_i, i) {
  xbar <- mean(x_minus_i)
  numerador <- (p - 1) * rho * xbar
  denominador <- 1 + (p - 2) * rho
  return(numerador / denominador)
}

var_cond <- function() {
  numerador <- 1 + (p - 2) * rho - (p - 1) * rho^2
  denominador <- 1 + (p - 2) * rho
  return(numerador / denominador)
}

# Gibbs Sampling
for (t in 2:Nsim) {
  x_ant <- X[t - 1, ]
  for (i in 1:p) {
    x_menos_i <- x_ant[-i]
    mu_i <- mu_cond(x_menos_i, i)
    sigma2_i <- var_cond()
    x_ant[i] <- rnorm(1, mean = mu_i, sd = sqrt(sigma2_i))
  }
  X[t, ] <- x_ant
}

#burnin
X_post <- X[-(1:burnin), ]

par(mfrow = c(2, p), mar = c(4, 4, 2, 1))
for (j in 1:p) {
```

```r
  hist(X_post[, j], probability = TRUE,
       main = paste0("Hist X", j),
       xlab = "", breaks = 20)
  curve(dnorm(x), col = "red", lwd = 2, add = TRUE)

  qqnorm(X_post[, j], main = paste0("QQplot X", j))
  qqline(X_post[, j], col = "red")
}
```



Exercício 3

Considere o seguinte modelo:

$$y_i \overset{\text{i.i.d.}}{\sim} N(\mu, \tau^{-1}) \quad (i = 1, \dots, n)$$

em que $\mu \sim N(0, \beta^{-1})$, $\beta \sim \text{Gamma}(2, 1)$ e $\tau \sim \text{Gamma}(2, 1)$. Implemente um amostrador de Gibbs que pode ser usado para estimar μ, β e τ, em que a amostra observada é dada por:

```r
y=c(2.34, 2.55, 1.91, 2.42, 2.26, 1.82, 2.04, 1.88, 2.03, 2.28, 2.16, 2.19,
    1.89, 1.38, 1.95, 2.50, 1.16, 2.09, 1.46, 0.97, 1.68, 1.89, 2.06, 2.07,
    1.72, 1.01, 2.96, 2.84, 2.03, 3.08, 1.80, 2.76, 2.04, 2.05, 1.91, 1.51,
    2.00, 0.77, 1.92, 1.54, 2.29, 2.28, 2.20, 1.60, 2.11, 1.36, 1.46, 2.53,
    1.89, 1.98)
```

Nota: Seja $\bar{y}$ e s_y^2 a média e variância amostrais, então as **condicionais completas** são:

$$\beta \mid \mu, \tau \sim \text{Gamma}\left(5/2,\ 1 + \tfrac{1}{2}\mu^2\right),$$

$$\tau \mid \mu, \beta \sim \text{Gamma}\left(\tfrac{n}{2} + 2,\ 1 + \tfrac{1}{2}(n-1)s_y^2 + \tfrac{1}{2}n(\bar{y} - \mu)^2\right),$$

$$\mu \mid \tau, \beta \sim \mathcal{N}\left(\frac{n\bar{y}\tau}{n\tau + \beta}, \frac{1}{n\tau + \beta}\right).$$

Solução

```r
n <- length(y)
y_bar <- mean(y)
s2_y <- var(y)

B <- 5000

mu <- beta <- tau <- numeric(B)

mu[1] <- mean(y)
beta[1] <- 1
tau[1] <- 1

# Gibbs Sampling
for (i in 2:B) {
  beta[i] <- rgamma(1, shape = 5/2, rate = 1 + 0.5 * mu[i - 1]^2)
  tau[i] <- rgamma(1, shape = n / 2 + 2,
                  rate = 1 + 0.5 * (n - 1) * s2_y + 0.5 * n * (y_bar - mu[i - 1])^2)
  var_mu <- 1 / (n * tau[i] + beta[i])
  mean_mu <- n * y_bar * tau[i] * var_mu
  mu[i] <- rnorm(1, mean = mean_mu, sd = sqrt(var_mu))
}
```

```
#burnin
burnin <- 1000
mu_post <- mu[-(1:burnin)]
beta_post <- beta[-(1:burnin)]
tau_post <- tau[-(1:burnin)]
```