Μπεϋζιανή Στατιστική

Βασίλης Κατσιάνος

Αύγουστος 2023

Περιεχόμενα

1 Κανονική Πιθανοφάνεια

1.1 Μονοδιάστατη Κανονική Κατανομή

Έστω τυχαίο δείγμα $y_1, \ldots, y_n$ από την κανονική κατανομή $\mathcal{N}\left(\mu, \tau^{-1}\right)$.

α. Θεωρούμε prior ανεξαρτησία με κατανομές $\mu \sim \mathcal{N}\left(a, c^{-1}\right)$ και $\tau \sim \mathrm{Gamma}(p, q)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των παραμέτρων μ και τ.

β. Θεωρούμε τη συζυγή prior κατανομή $\mu \mid \tau \sim \mathcal{N}\left(a, c^{-1}\tau^{-1}\right)$, $\tau \sim \mathrm{Gamma}(p, q)$. Να υπολογιστούν οι δεσμευμένες και οι περιθώριες posterior κατανομές των παραμέτρων μ και τ.

Λύση.

α. Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$\begin{aligned}
\pi(\mu,\tau) &= \pi(\mu)\cdot\pi(\tau)\\
&= \sqrt{\frac{c}{2\pi}}\exp\left\{-\frac{c(\mu-a)^2}{2}\right\}\cdot\frac{q^p}{\Gamma(p)}\tau^{p-1}e^{-q\tau}\\
&\propto \exp\left\{-c\frac{\mu^2-2\mu a+a^2}{2}\right\}\cdot\tau^{p-1}e^{-q\tau}\\
&\propto \exp\left\{-\frac{1}{2}c\mu^2+ca\mu\right\}\cdot\tau^{p-1}e^{-q\tau}.
\end{aligned}$$

Η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$\begin{aligned}
f(y\mid\mu,\tau) &= \prod_{i=1}^{n} f(y_i\mid\mu,\tau)\\
&= \prod_{i=1}^{n}\sqrt{\frac{\tau}{2\pi}}\exp\left\{-\frac{\tau(y_i-\mu)^2}{2}\right\}\\
&\propto \tau^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2\tau\right\}\\
&= \tau^{\frac{n}{2}}\exp\left\{-\sum_{i=1}^{n}\frac{y_i^2-2y_i\mu+\mu^2}{2}\tau\right\}\\
&= \exp\left\{-\frac{1}{2}n\tau\mu^2+n\tau\overline{y}\mu\right\}\cdot\tau^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}y_i^2\tau\right\}.
\end{aligned}$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των μ και τ ως εξής:

$$\begin{aligned}
\pi(\mu\mid\tau,y) &\propto \pi(\mu,\tau\mid y)\\
&\propto \pi(\mu,\tau)\cdot f(y\mid\mu,\tau)\\
&\propto \exp\left\{-\frac{c}{2}\mu^2+ca\mu\right\}\cdot\exp\left\{-\frac{n\tau}{2}\mu^2+n\tau\overline{y}\mu\right\}\\
&= \exp\left\{-\frac{1}{2}\underbrace{(c+n\tau)}_{c_n}\mu^2+(ca+n\tau\overline{y})\,\mu\right\}\\
&= \exp\left\{-\frac{c+n\tau}{2}\mu^2+\underbrace{(c+n\tau)}_{c_n}\underbrace{\frac{ca+n\tau\overline{y}}{c+n\tau}}_{a_n}\mu\right\},
\end{aligned}$$

$$\begin{aligned}
\pi(\tau\mid\mu,y) &\propto \pi(\mu,\tau)\cdot f(y\mid\mu,\tau)\\
&\propto \tau^{p-1}e^{-q\tau}\cdot\tau^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2\tau\right\}\\
&= \tau^{p+\frac{n}{2}-1}\exp\left\{-\left[q+\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^2\right]\tau\right\}.
\end{aligned}$$

Δηλαδή,

$$\mu \mid \tau, y \sim \mathcal{N}\left(\frac{ca + n\tau\bar{y}}{c + n\tau}, \frac{1}{c + n\tau}\right), \quad \tau \mid \mu, y \sim \text{Gamma}\left(p + \frac{n}{2}, q + \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2\right).$$

Μπορούμε να υπολογίσουμε την prior του Jeffreys για τη μονοδιάστατη κανονική κατανομή ως εξής:

$$\log f(y \mid \mu, \tau) = \frac{1}{2}\log\tau - \frac{1}{2}\log(2\pi) - \frac{\tau(y-\mu)^2}{2},$$

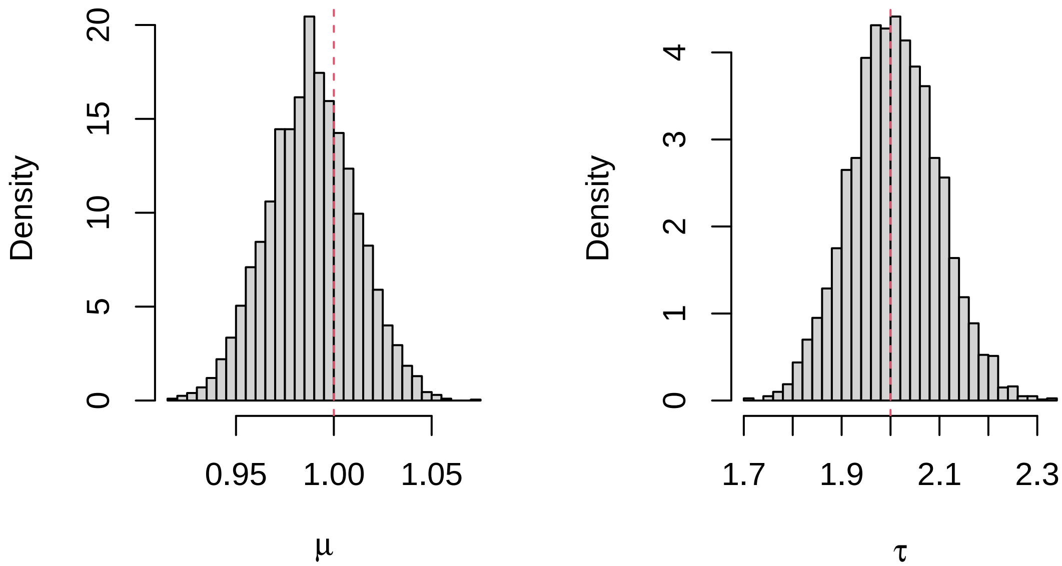
$$\frac{\partial \log f(y \mid \mu, \tau)}{\partial \mu} = \tau(y - \mu), \quad \frac{\partial \log f(y \mid \mu, \tau)}{\partial \tau} = \frac{1}{2\tau} - \frac{(y-\mu)^2}{2},$$

$$\frac{\partial^2 \log f(y \mid \mu, \tau)}{\partial \mu^2} = -\tau, \quad \frac{\partial^2 \log f(y \mid \mu, \tau)}{\partial \tau^2} = -\frac{1}{2\tau^2}, \quad \frac{\partial^2 \log f(y \mid \mu, \tau)}{\partial \mu \partial \tau} = y - \mu,$$

$$\mathcal{I}(\mu, \tau) = -\mathbb{E}\left[\frac{\partial \log f(y \mid \mu, \tau)}{\partial(\mu, \tau)\partial(\mu, \tau)}\right] = \begin{bmatrix} \tau & 0 \\ 0 & \frac{1}{2\tau^2} \end{bmatrix}, \quad J(\mu, \tau) \propto \sqrt{|\mathcal{I}(\mu, \tau)|} = \sqrt{\frac{1}{2\tau}} \propto \tau^{-0.5}.$$

Παρατηρούμε ότι η μη-ορθή prior του Jeffreys προκύπτει για $a = c = q = 0$ και $p = 0.5$

```
MCMCnorm = function(Y, mu0, tau0, a, c, p, q, niter, nburn) {
    n = length(Y)
    S = sum(Y)
    mu = numeric(niter)
    tau = numeric(niter)
    mu[1] = mu0
    tau[1] = tau0
    for (i in 2:niter) {
        mu[i] = rnorm(1, (c * a + tau[i - 1] * S)/(c + n * tau[i - 1]), (c +
            n * tau[i - 1])^(-0.5))
        tau[i] = rgamma(1, p + n/2, q + sum((Y - mu[i])^2)/2)
    }
    return(list(mu = mu[-(1:nburn)], tau = tau[-(1:nburn)]))
}

n = 1000
mu = 1
tau = 2
Y = rnorm(n, mu, tau^(-0.5))
posterior = MCMCnorm(Y, 0, 1, 0, 0, 0.5, 0, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$mu, "FD", freq = FALSE, main = NA, xlab = expression(mu))
abline(v = mu, col = 2, lty = 2)
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



β. Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$
\begin{aligned}
\pi(\mu,\tau) &= \pi(\mu \mid \tau)\cdot\pi(\tau) \\
&= \sqrt{\frac{c\tau}{2\pi}}\exp\left\{-\frac{c\tau(\mu-a)^2}{2}\right\}\cdot\frac{q^p}{\Gamma(p)}\tau^{p-1}e^{-q\tau} \\
&\propto \tau^{p+\frac{1}{2}-1}\exp\left\{-\left[q+\frac{c(\mu-a)^2}{2}\right]\tau\right\} \\
&= \tau^{\frac{1}{2}}\exp\left\{-c\tau\frac{\mu^2-2\mu a+a^2}{2}\right\}\cdot\tau^{p-1}e^{-q\tau} \\
&= \tau^{\frac{1}{2}}\exp\left\{-\frac{1}{2}c\tau\mu^2+c\tau a\mu\right\}\cdot\tau^{p-1}\exp\left\{-\left(q+\frac{ca^2}{2}\right)\tau\right\}.
\end{aligned}
$$

Επομένως, παίρνουμε την από κοινού posterior κατανομή των μ και τ ως εξής:

$$
\begin{aligned}
\pi(\mu,\tau \mid y) &\propto \pi(\mu,\tau)\cdot f(y \mid \mu,\tau) \\
&\propto \tau^{\frac{1}{2}}\exp\left\{-\frac{1}{2}c\tau\mu^2+c\tau a\mu\right\}\cdot\exp\left\{-\frac{1}{2}n\tau\mu^2+n\tau\overline{y}\mu\right\} \\
&\quad\times\tau^{p-1}\exp\left\{-\left(q+\frac{ca^2}{2}\right)\tau\right\}\cdot\tau^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}y_i^2\tau\right\} \\
&= \tau^{\frac{1}{2}}\exp\left\{-\frac{1}{2}\underbrace{(c+n)}_{c_n}\tau\mu^2+\tau\,(ca+n\overline{y})\,\mu\right\}\cdot\tau^{p+\frac{n}{2}-1}\exp\left\{-\left(q+\frac{ca^2}{2}+\frac{1}{2}\sum_{i=1}^{n}y_i^2\right)\tau\right\} \\
&= \tau^{\frac{1}{2}}\exp\left\{-\frac{c+n}{2}\tau\mu^2+\underbrace{(c+n)}_{c_n}\tau\underbrace{\frac{ca+n\overline{y}}{c+n}}_{a_n}\mu\right\}\cdot\tau^{p+\frac{n}{2}-1}\exp\left\{-\left(q+\frac{ca^2}{2}+\frac{1}{2}\sum_{i=1}^{n}y_i^2\right)\tau\right\} \\
&= \tau^{\frac{1}{2}}\exp\left\{-\frac{1}{2}c_n\tau\mu^2+c_n\tau a_n\mu-\frac{1}{2}c_n\tau a_n^2+\frac{1}{2}c_n\tau a_n^2\right\} \\
&\quad\times\tau^{p+\frac{n}{2}-1}\exp\left\{-\left(q+\frac{ca^2}{2}+\frac{1}{2}\sum_{i=1}^{n}y_i^2\right)\tau\right\}
\end{aligned}
$$

$$= \tau^{\frac{1}{2}} \exp\left\{-\frac{c_n \tau(\mu - a_n)^2}{2}\right\} \cdot \tau^{p+\frac{n}{2}-1} \exp\left\{-\left(q + \frac{ca^2}{2} + \frac{1}{2}\sum_{i=1}^n y_i^2 - \frac{c_n a_n^2}{2}\right)\tau\right\}.$$

Υπολογίζουμε ότι:

$$c_n a_n^2 = (c+n)\left(\frac{ca + n\overline{y}}{c+n}\right)^2 = \frac{(ca + n\overline{y})^2}{c+n}.$$

Δηλαδή,

$$\mu \mid \tau, y \sim \mathcal{N}\left(\frac{ca + n\overline{y}}{c+n}, \frac{\tau^{-1}}{c+n}\right), \quad \tau \mid y \sim \text{Gamma}\left(p + \frac{n}{2}, q + \frac{ca^2}{2} + \frac{1}{2}\sum_{i=1}^n y_i^2 - \frac{(ca + n\overline{y})^2}{2(c+n)}\right).$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή του τ ως εξής:

$$\begin{aligned}
\pi(\tau \mid \mu, y) &\propto \pi(\mu, \tau) \cdot f(y \mid \mu, \tau) \\
&\propto \tau^{p+\frac{1}{2}-1} \exp\left\{-\left[q + \frac{c(\mu - a)^2}{2}\right]\tau\right\} \cdot \tau^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^n (y_i - \mu)^2 \tau\right\} \\
&= \tau^{p+\frac{n+1}{2}-1} \exp\left\{-\left[q + \frac{c(\mu - a)^2}{2} + \frac{1}{2}\sum_{i=1}^n (y_i - \mu)^2\right]\tau\right\}.
\end{aligned}$$

Δηλαδή,

$$\tau \mid \mu, y \sim \text{Gamma}\left(p + \frac{n+1}{2}, q + \frac{c(\mu - a)^2}{2} + \frac{1}{2}\sum_{i=1}^n (y_i - \mu)^2\right).$$

Ορισμός 1.1. Λέμε ότι μία τυχαία μεταβλητή X ακολουθεί τη γενικευμένη κατανομή Student's t με μέσο $\mu \in \mathbb{R}$, διασπορά $\sigma^2 > 0$, $\nu > 0$ βαθμούς ελευθερίας και γράφουμε $X \sim t_\nu(\mu, \sigma^2)$, αν έχει συνάρτηση πυκνότητας πιθανότητας:

$$f_X(x \mid \mu, \sigma^2, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi\sigma^2}} \left[1 + \frac{1}{\nu\sigma^2}(x - \mu)^2\right]^{-\frac{\nu+1}{2}}, \quad x \in \mathbb{R}.$$

Τέλος, ορίζουμε:

$$p_n = p + \frac{n}{2}, \quad q_n = q + \frac{ca^2}{2} + \frac{1}{2}\sum_{i=1}^n y_i^2 - \frac{c_n a_n^2}{2}.$$

Τότε, υπολογίζουμε την περιθώρια posterior κατανομή του μ ως εξής:

$$\begin{aligned}
\pi(\mu \mid y) &= \int \pi(\mu, \tau \mid y) d\tau \\
&\propto \int \tau^{p+\frac{n+1}{2}-1} \exp\left\{-\left[q + \frac{c(\mu - a)^2}{2} + \frac{1}{2}\sum_{i=1}^n (y_i - \mu)^2\right]\tau\right\} d\tau \\
&\propto \left[q + \frac{c(\mu - a)^2}{2} + \frac{1}{2}\sum_{i=1}^n (y_i - \mu)^2\right]^{-\left(p+\frac{n+1}{2}\right)} \\
&= \left[q + \frac{ca^2}{2} + \frac{1}{2}\sum_{i=1}^n y_i^2 + \frac{c+n}{2}\mu^2 - (ca + n\overline{y})\mu\right]^{-p_n - \frac{1}{2}}
\end{aligned}$$

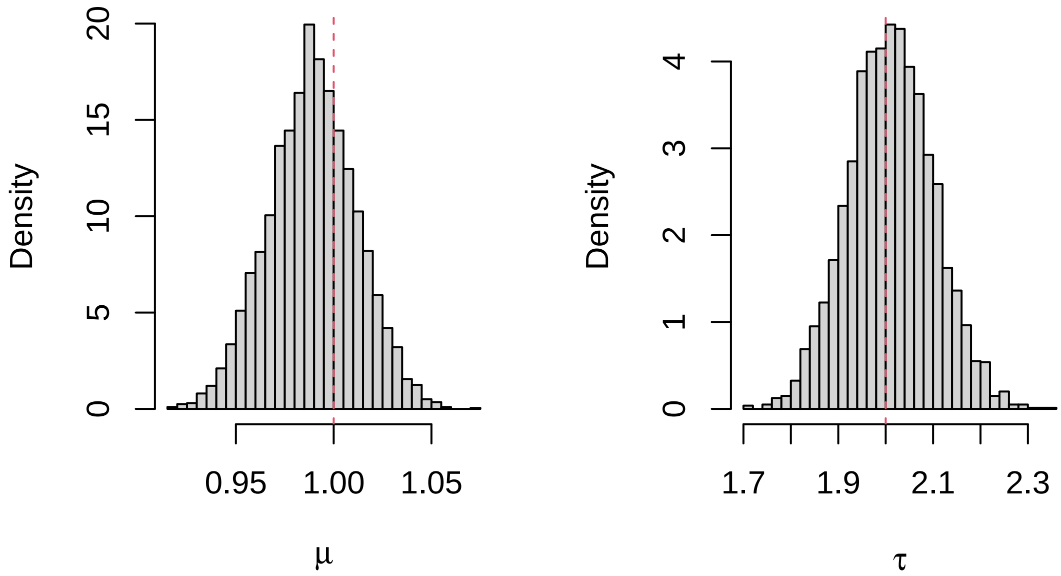
$$
\begin{aligned}
&= \left[q + \frac{ca^2}{2} + \frac{1}{2}\sum_{i=1}^{n} y_i^2 - \frac{c_n a_n^2}{2} + \frac{c_n\left(\mu - a_n\right)^2}{2}\right]^{-\frac{2p_n+1}{2}} \\
&\propto \left[1 + \frac{c_n\left(\mu - a_n\right)^2}{2q_n}\right]^{-\frac{2p_n+1}{2}} \\
&= \left[1 + \frac{1}{2p_n}\frac{\left(\mu - a_n\right)^2}{\frac{q_n}{c_n p_n}}\right]^{-\frac{2p_n+1}{2}}.
\end{aligned}
$$

Δηλαδή,

$$
\mu \mid y \sim t_{2p_n}\left(a_n, \frac{q_n}{c_n p_n}\right).
$$

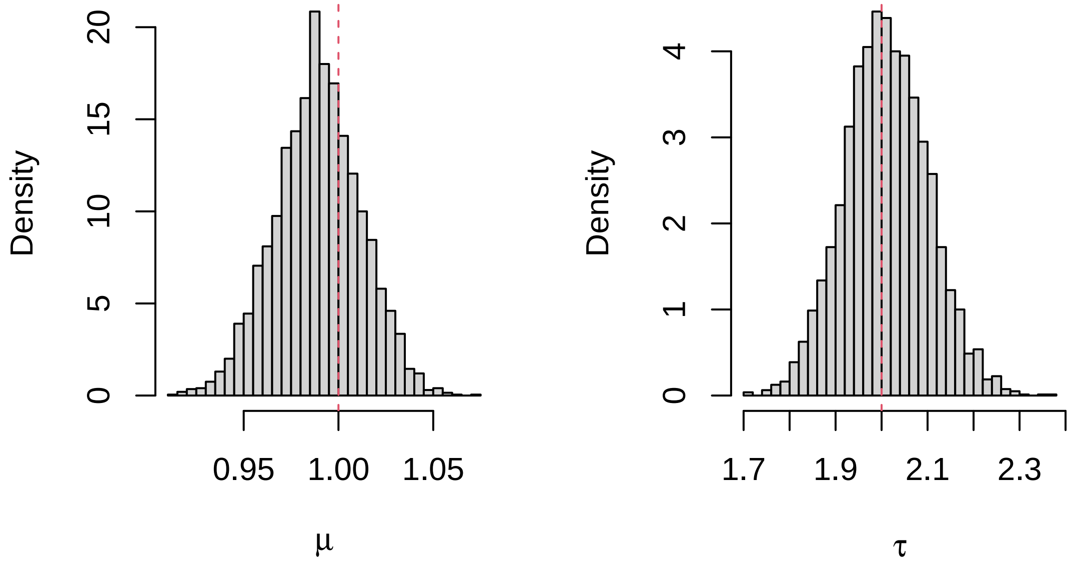
Πρώτα θα υλοποιήσουμε έναν δειγματολήπτη Gibbs που προσομοιώνει εναλλάξ από τις δεσμευμένες posterior κατανομές των παραμέτρων μ και τ.

```
MCMCnorm = function(Y, mu0, tau0, a, c, p, q, niter, nburn) {
    n = length(Y)
    S = sum(Y)
    cn = c + n
    an = (c * a + S)/cn
    mu = numeric(niter)
    tau = numeric(niter)
    mu[1] = mu0
    tau[1] = tau0
    for (i in 2:niter) {
        mu[i] = rnorm(1, an, (cn * tau[i - 1])^(-0.5))
        tau[i] = rgamma(1, p + (n + 1)/2, q + c * (mu[i] - a)^2/2 + sum((Y -
            mu[i])^2)/2)
    }
    return(list(mu = mu[-(1:nburn)], tau = tau[-(1:nburn)]))
}

posterior = MCMCnorm(Y, 0, 1, 0, 0, 0.5, 0, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$mu, "FD", freq = FALSE, main = NA, xlab = expression(mu))
abline(v = mu, col = 2, lty = 2)
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



Έπειτα, θα υλοποιήσουμε τη μέθοδο της σύνθεσης που προσομοιώνει πρώτα από την περιθώρια posterior κατανομή της παραμέτρου τ και στη συνέχεια από τη δεσμευμένη posterior κατανομή της παραμέτρου μ.

```
CMnorm = function(Y, a, c, p, q, niter) {
    n = length(Y)
    S = sum(Y)
    cn = c + n
    an = (c * a + S)/cn
    pn = p + n/2
    qn = q + c * a^2/2 + sum(Y^2)/2 - cn * an^2/2
    mu = numeric(niter)
    tau = numeric(niter)
    for (i in 1:niter) {
        tau[i] = rgamma(1, pn, qn)
        mu[i] = rnorm(1, an, (cn * tau[i])^(-0.5))
    }
    return(list(mu = mu, tau = tau))
}

posterior = CMnorm(Y, 0, 0, 0.5, 0, 4000)
par(mfrow = c(1, 2))
hist(posterior$mu, "FD", freq = FALSE, main = NA, xlab = expression(mu))
abline(v = mu, col = 2, lty = 2)
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



1.2 Γενικευμένη Κατανομή Student's t

Έστω τυχαίο δείγμα $y_1, \ldots, y_n$ από τη γενικευμένη κατανομή Student's t με μέσο $\mu \in \mathbb{R}$, ακρίβεια $\tau > 0$ και $\nu > 0$ βαθμούς ελευθερίας, δηλαδή:

$$f(y_i \mid \mu, \tau, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \sqrt{\frac{\tau}{\nu\pi}} \left[1 + \frac{1}{\nu}\tau(y_i - \mu)^2\right]^{-\frac{\nu+1}{2}}, \quad y_i \in \mathbb{R}.$$

Θεωρούμε τις τυχαίες μεταβλητές $W_i \sim \mathcal{N}\left(0, \tau^{-1}\right)$ και $V_i \sim \chi^2_\nu \equiv \text{Gamma}\left(\frac{\nu}{2}, \frac{1}{2}\right)$. Τότε, παρατηρούμε ότι:

$$Y_i \stackrel{d}{=} \frac{W_i}{\sqrt{\frac{V_i}{\nu}}} + \mu.$$

Θέτουμε $Z_i = \frac{V_i}{\nu}$. Τότε, $Z_i \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$. Παρατηρούμε ότι:

$$Y_i \mid z_i \stackrel{d}{=} \frac{W_i}{\sqrt{z_i}} + \mu \sim \mathcal{N}\left(\mu, \tau^{-1} z_i^{-1}\right).$$

Θεωρούμε γνωστούς τους βαθμούς ελευθερίας ν και prior ανεξαρτησία για τις παραμέτρους μ, τ με κατανομές $\mu \sim \mathcal{N}\left(a, c^{-1}\right)$, $\tau \sim \text{Gamma}(p, q)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των παραμέτρων μ, τ και των λανθανουσών μεταβλητών z_i.

Λύση.

Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$\begin{aligned}
\pi(\mu, \tau) &= \pi(\mu) \cdot \pi(\tau) \\
&= \sqrt{\frac{c}{2\pi}} \exp\left\{-\frac{c(\mu - a)^2}{2}\right\} \cdot \frac{q^p}{\Gamma(p)} \tau^{p-1} e^{-q\tau} \\
&\propto \exp\left\{-c\frac{\mu^2 - 2\mu a + a^2}{2}\right\} \cdot \tau^{p-1} e^{-q\tau}
\end{aligned}$$

$$\propto \exp\left\{-\frac{1}{2}c\mu^2 + ca\mu\right\} \cdot \tau^{p-1}e^{-q\tau}.$$

Ορίζουμε:

$$\overline{zy} = \frac{1}{n}\sum_{i=1}^{n} z_i y_i.$$

Η πλήρης πιθανοφάνεια, δηλαδή η από κοινού πιθανοφάνεια των παρατηρούμενων μεταβλητών y_i και των λανθανουσών μεταβλητών z_i, δίνεται από τη σχέση:

$$\begin{aligned}
f(y, z \mid \mu, \tau) &= \prod_{i=1}^{n} f(y_i, z_i \mid \mu, \tau) \\
&= \prod_{i=1}^{n} f(z_i) f(y_i \mid z_i, \mu, \tau) \\
&= \prod_{i=1}^{n} \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} z_i^{\frac{\nu}{2}-1} e^{-\frac{\nu}{2}z_i} \sqrt{\frac{\tau z_i}{2\pi}} \exp\left\{-\frac{\tau z_i (y_i-\mu)^2}{2}\right\} \\
&\propto \tau^{\frac{n}{2}} \cdot \prod_{i=1}^{n} z_i^{\frac{\nu+1}{2}-1} \exp\left\{-\sum_{i=1}^{n} \frac{\nu + \tau(y_i-\mu)^2}{2} z_i\right\} \\
&= \tau^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} z_i (y_i-\mu)^2 \tau\right\} \cdot \prod_{i=1}^{n} z_i^{\frac{\nu+1}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\} \\
&= \tau^{\frac{n}{2}} \exp\left\{-\sum_{i=1}^{n} \frac{y_i^2 - 2y_i\mu + \mu^2}{2} z_i \tau\right\} \cdot \prod_{i=1}^{n} z_i^{\frac{\nu+1}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\} \\
&= \exp\left\{-\frac{1}{2}n\tau\overline{z}\mu^2 + n\tau\overline{zy}\mu\right\} \cdot \tau^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} z_i y_i^2 \tau\right\} \cdot \prod_{i=1}^{n} z_i^{\frac{\nu+1}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\}.
\end{aligned}$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των παραμέτρων μ και τ ως εξής:

$$\begin{aligned}
\pi(\mu \mid \tau, z, y) &\propto \pi(\mu, \tau, z \mid y) \\
&\propto \pi(\mu, \tau) \cdot f(y, z \mid \mu, \tau) \\
&\propto \exp\left\{-\frac{1}{2}c\mu^2 + ca\mu\right\} \cdot \exp\left\{-\frac{1}{2}n\tau\overline{z}\mu^2 + n\tau\overline{zy}\mu\right\} \\
&= \exp\left\{-\frac{1}{2}\underbrace{(c + n\tau\overline{z})}_{c_n}\mu^2 + (ca + n\tau\overline{zy})\,\mu\right\} \\
&= \exp\left\{-\frac{c + n\tau\overline{z}}{2}\mu^2 + \underbrace{(c + n\tau\overline{z})}_{c_n}\underbrace{\frac{ca + n\tau\overline{zy}}{c + n\tau\overline{z}}}_{a_n}\mu\right\},
\end{aligned}$$

$$\begin{aligned}
\pi(\tau \mid \mu, z, y) &\propto \pi(\mu, \tau) \cdot f(y, z \mid \mu, \tau) \\
&\propto \tau^{p-1}e^{-q\tau} \cdot \tau^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} z_i (y_i-\mu)^2 \tau\right\}
\end{aligned}$$

$$= \tau^{p+\frac{n}{2}-1} \exp\left\{ -\left[q + \frac{1}{2}\sum_{i=1}^{n} z_i (y_i - \mu)^2 \right] \tau \right\}.$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή των λανθανουσών μεταβλητών z_i ως εξής:

$$f(z_i \mid y_i, \mu, \tau) \propto f(y_i, z_i \mid \mu, \tau) \propto z_i^{\frac{\nu+1}{2}-1} \exp\left\{ -\frac{\nu + \tau(y_i-\mu)^2}{2} z_i \right\}.$$

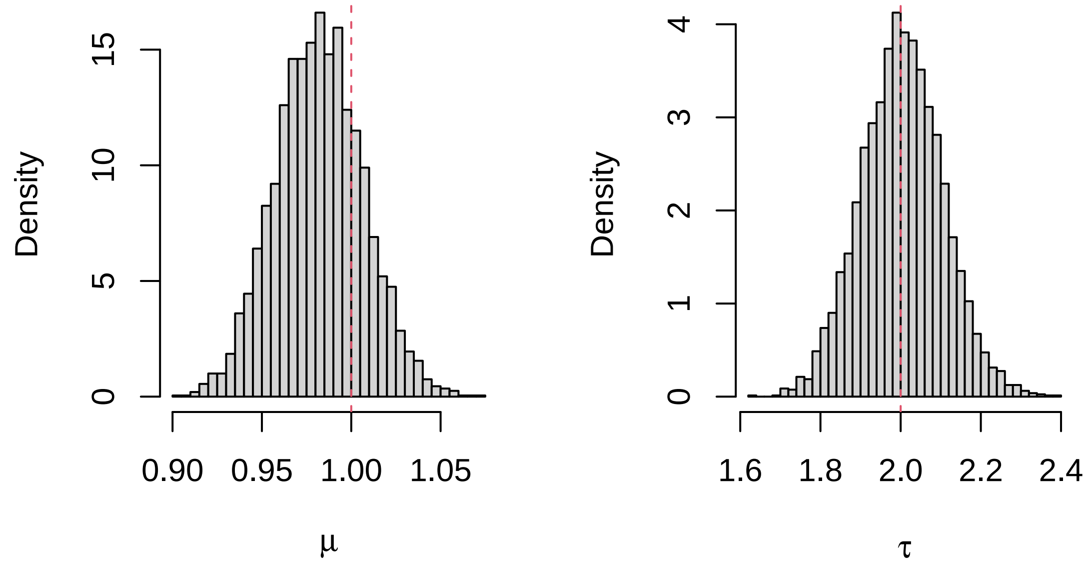
Δηλαδή,

$$\mu \mid \tau, z, y \sim \mathcal{N}\left(\frac{ca + n\tau\overline{zy}}{c + n\tau\overline{z}}, \frac{1}{c + n\tau\overline{z}} \right), \quad \tau \mid \mu, z, y \sim \text{Gamma}\left(p + \frac{n}{2}, q + \frac{1}{2}\sum_{i=1}^{n} z_i (y_i - \mu)^2 \right),$$

$$z_i \mid y_i, \mu, \tau \sim \text{Gamma}\left(\frac{\nu+1}{2}, \frac{\nu + \tau(y_i-\mu)^2}{2} \right).$$

```r
MCMCt = function(Y, mu0, tau0, nu, a, c, p, q, niter, nburn) {
    n = length(Y)
    mu = numeric(niter)
    tau = numeric(niter)
    Z = matrix(0, niter, n)
    mu[1] = mu0
    tau[1] = tau0
    Z[1, ] = rgamma(n, (nu + 1)/2, (nu + tau[1] * (Y - mu[1])^2)/2)
    for (i in 2:niter) {
        mu[i] = rnorm(1, (c * a + tau[i - 1] * sum(Z[i - 1, ] * Y))/(c + tau[i -
            1] * sum(Z[i - 1, ])), (c + n * tau[i - 1])^(-0.5))
        tau[i] = rgamma(1, p + n/2, q + sum(Z[i - 1, ] * (Y - mu[i])^2)/2)
        Z[i, ] = rgamma(n, (nu + 1)/2, (nu + tau[i] * (Y - mu[i])^2)/2)
    }
    return(list(mu = mu[-(1:nburn)], tau = tau[-(1:nburn)], Z = Z[-(1:nburn),
        ]))
}

library(mvtnorm)
n = 1000
mu = 1
tau = 2
nu = 10
Y = rmvt(n, matrix(tau^(-1)), nu, mu)
posterior = MCMCt(Y, 0, 1, nu, 0, 0, 0.5, 0, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$mu, "FD", freq = FALSE, main = NA, xlab = expression(mu))
abline(v = mu, col = 2, lty = 2)
```

```
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



Εναλλακτικά, μπορούμε να σχεδιάσουμε έναν αλγόριθμο Random Walk Metropolis-Hastings. Για την παράμετρο $\mu \in \mathbb{R}$ θεωρούμε την προτείνουσα τυχαία μεταβλητή $\mu^* \mid \mu_{\ell-1} \sim \mathcal{N}\left(\mu_{\ell-1}, \sigma_\mu^2\right)$ με πιθανότητα αποδοχής:

$$A\left(\mu_{\ell-1}, \mu^*\right) = \min\left\{\frac{\pi\left(\mu^*\right)}{\pi\left(\mu_{\ell-1}\right)}\frac{f\left(y \mid \mu^*, \tau_{\ell-1}\right)}{f\left(y \mid \mu_{\ell-1}, \tau_{\ell-1}\right)}, 1\right\}.$$

Για την παράμετρο $\tau > 0$ θεωρούμε την προτείνουσα τυχαία μεταβλητή $\tau^* \mid \tau_{\ell-1} \sim \text{Lognormal}\left(\log \tau_{\ell-1}, \sigma_\tau^2\right)$ με δεσμευμένη συνάρτηση πυκνότητας πιθανότητας:

$$f\left(\tau^\star \mid \tau_{\ell-1}\right) = \frac{1}{\sqrt{2\pi\sigma_\tau^2\tau_{\ell-1}^2}}\exp\left\{-\frac{\left(\log\tau^* - \log\tau_{\ell-1}\right)^2}{2\sigma_\tau^2}\right\}.$$

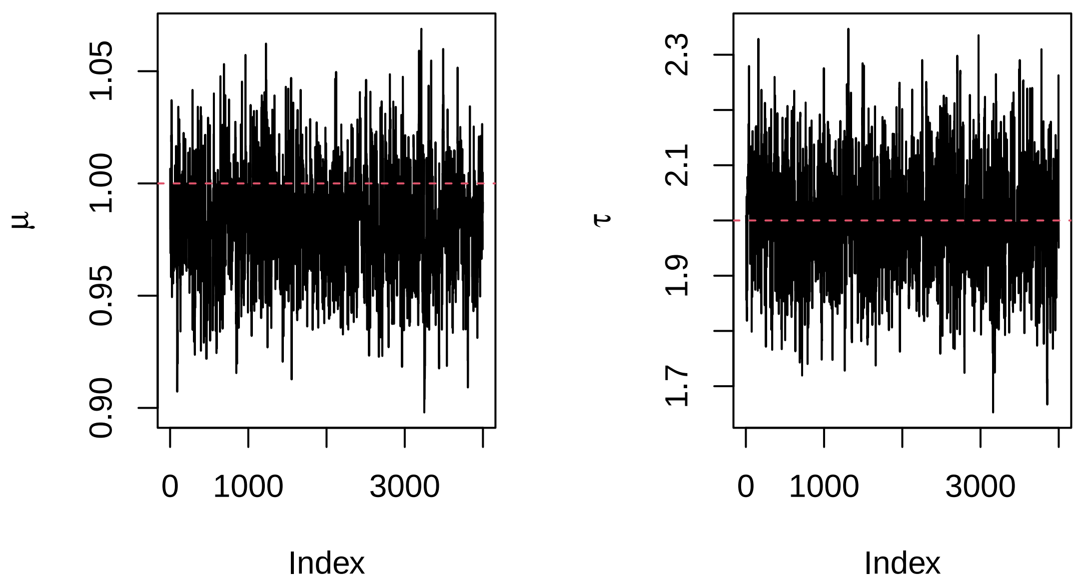
Τότε, η πιθανότητα αποδοχής της τιμής τ^* δίνεται από τη σχέση:

$$A\left(\tau_{\ell-1}, \tau^*\right) = \min\left\{\frac{\pi\left(\tau^*\right)}{\pi\left(\tau_{\ell-1}\right)}\frac{f\left(y \mid \mu_\ell, \tau^*\right)}{f\left(y \mid \mu_\ell, \tau_{\ell-1}\right)}\frac{\tau^*}{\tau_{\ell-1}}, 1\right\}.$$

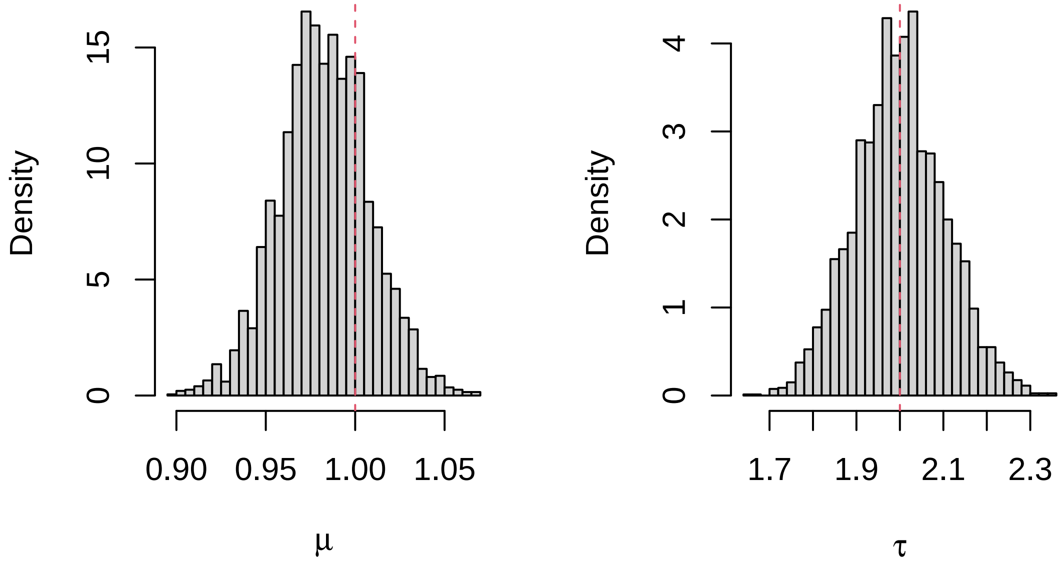
Αν $Z \sim \mathcal{N}\left(0, \sigma_\tau^2\right)$, τότε γνωρίζουμε ότι $\tau^* \mid \tau_{\ell-1} \sim \tau_{\ell-1}e^Z$. Επομένως, η συγκεκριμένη προτείνουσα καλείται πολλαπλασιαστικός τυχαίος περίπατος και μπορεί να γραφτεί ισοδύναμα ως τυχαίος περίπατος σε λογαριθμική κλίμακα με τον ακόλουθο τρόπο: $\log\tau^* \mid \tau_{\ell-1} \sim \mathcal{N}\left(\log\tau_{\ell-1}, \sigma_\tau^2\right)$.

```
RWMHt = function(Y, mu0, tau0, nu, musd, tausd, niter, nburn) {
    library(mvtnorm)
    mu = numeric(niter)
    tau = numeric(niter)
    mu[1] = mu0
    tau[1] = tau0
    for (i in 2:niter) {
        mustar = rnorm(1, mu[i - 1], musd)
```

```r
        logA = sum(dmvt(Y, mustar, matrix(tau[i - 1]^(-1)), nu) - dmvt(Y, mu[i -
            1], matrix(tau[i - 1]^(-1)), nu))
        mu[i] = ifelse(log(runif(1)) < logA, mustar, mu[i - 1])
        taustar = tau[i - 1] * exp(rnorm(1, sd = tausd))
        logA = log(taustar/tau[i - 1])/2 + sum(dmvt(Y, mu[i], matrix(taustar^(-1)),
            nu) - dmvt(Y, mu[i], matrix(tau[i - 1]^(-1)), nu))
        tau[i] = ifelse(log(runif(1)) < logA, taustar, tau[i - 1])
    }
    return(list(mu = mu[-(1:nburn)], tau = tau[-(1:nburn)]))
}

posterior = RWMHt(Y, 0, 1, nu, 0.05, 0.1, 5000, 1000)
par(mfrow = c(1, 2))
plot(posterior$mu, type = "l", ylab = expression(mu))
abline(h = mu, col = 2, lty = 2)
plot(posterior$tau, type = "l", ylab = expression(tau))
abline(h = tau, col = 2, lty = 2)
```




```r
hist(posterior$mu, "FD", freq = FALSE, main = NA, xlab = expression(mu))
abline(v = mu, col = 2, lty = 2)
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



1.3 Μοντέλο Ανάλυσης Διασποράς με Έναν Παράγοντα

Θεωρούμε το μοντέλο ανάλυσης διασποράς $y_{ij} = \mu_i + \varepsilon_{ij}$, όπου $\varepsilon_i \sim \mathcal{N}\left(0, \tau^{-1}\right)$ για $i = 1, 2, \ldots, m$ και $j = 1, 2, \ldots, n_i$. Θεωρούμε prior ανεξαρτησία με κατανομές $\mu_i \sim \mathcal{N}\left(a, c^{-1}\right)$ και $\tau \sim \text{Gamma}(p, q)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των παραμέτρων μ_i και τ.

Λύση.

Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$\begin{aligned}
\pi(\mu, \tau) &= \pi(\tau) \cdot \prod_{i=1}^{m} \pi(\mu_i) \\
&= \frac{q^p}{\Gamma(p)} \tau^{p-1} e^{-q\tau} \cdot \prod_{i=1}^{m} \sqrt{\frac{c}{2\pi}} \exp\left\{-\frac{c(\mu_i - a)^2}{2}\right\} \\
&\propto \exp\left\{-c \sum_{i=1}^{m} \frac{\mu_i^2 - 2a\mu_i + a^2}{2}\right\} \cdot \tau^{p-1} e^{-q\tau} \\
&\propto \exp\left\{-\frac{c}{2} \sum_{i=1}^{m} \mu_i^2 + ca \sum_{i=1}^{m} \mu_i\right\} \cdot \tau^{p-1} e^{-q\tau}.
\end{aligned}$$

Ορίζουμε:

$$n = \sum_{i=1}^{m} N_i, \quad \overline{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}.$$

Τότε, η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$\begin{aligned}
f(y \mid \mu, \tau) &= \prod_{i=1}^{m} \prod_{j=1}^{n_i} f(y_{ij} \mid \mu, \tau) \\
&= \prod_{i=1}^{m} \prod_{j=1}^{n_i} \sqrt{\frac{\tau}{2\pi}} \exp\left\{-\frac{\tau(y_{ij} - \mu_i)^2}{2}\right\}
\end{aligned}$$

$$
\begin{aligned}
&\propto \tau^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\mu_i)^2\tau\right\} \\
&= \tau^{\frac{n}{2}} \exp\left\{-\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{y_{ij}^2-2y_{ij}\mu_i+\mu_i^2}{2}\tau\right\} \\
&= \exp\left\{-\frac{1}{2}\tau\sum_{i=1}^{m}N_i\mu_i^2+\tau\sum_{i=1}^{m}N_i\overline{y}_i\mu_i\right\}\cdot\tau^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}y_{ij}^2\tau\right\}.
\end{aligned}
$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των μ_i και τ ως εξής:

$$
\begin{aligned}
\pi(\mu_i \mid \tau, y) &\propto \pi(\mu, \tau \mid y) \\
&\propto \pi(\mu,\tau)\cdot f(y\mid\mu,\tau) \\
&\propto \exp\left\{-\frac{c}{2}\mu_i^2+ca\mu_i\right\}\cdot\exp\left\{-\frac{1}{2}n_i\tau\mu_i^2+n_i\tau\overline{y}_i\mu_i\right\} \\
&= \exp\left\{-\frac{1}{2}\underbrace{(c+n_i\tau)}_{c_n}\mu_i^2+(ca+n_i\tau\overline{y}_i)\,\mu_i\right\} \\
&= \exp\left\{-\frac{c+n_i\tau}{2}\mu_i^2+\underbrace{(c+n_i\tau)}_{c_n}\underbrace{\frac{ca+n_i\tau\overline{y}_i}{c+n_i\tau}}_{a_n}\mu_i\right\},
\end{aligned}
$$

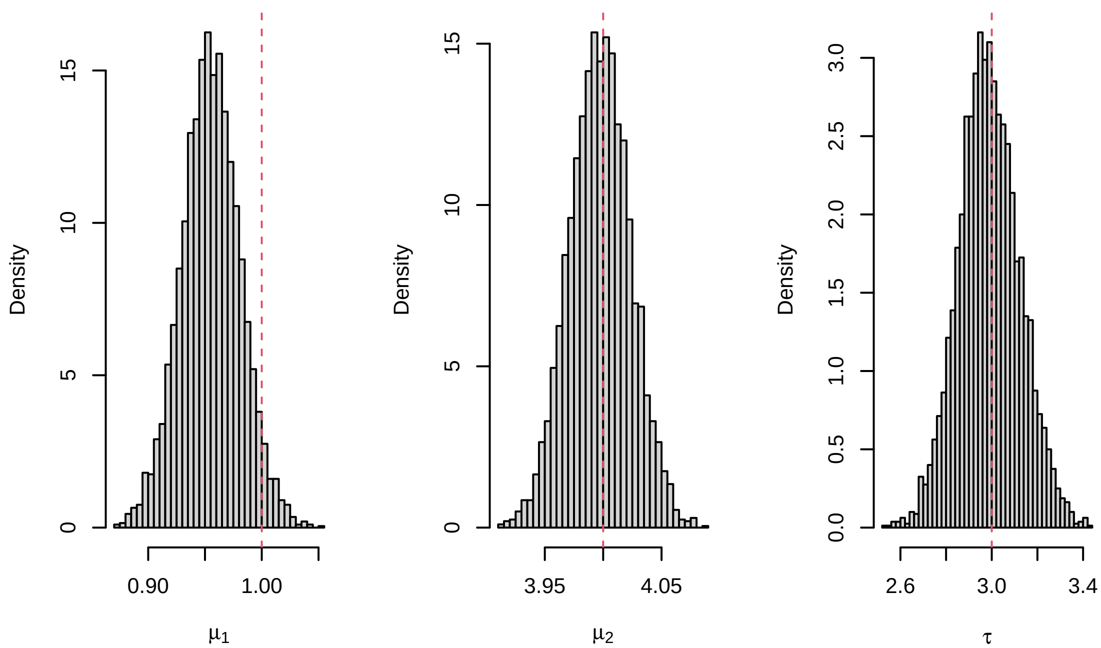
$$
\begin{aligned}
\pi(\tau \mid \mu, y) &\propto \pi(\mu,\tau)\cdot f(y\mid\mu,\tau) \\
&\propto \tau^{p-1}e^{-q\tau}\cdot\tau^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\mu_i)^2\tau\right\} \\
&= \tau^{p+\frac{n}{2}-1}\exp\left\{-\left[q+\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\mu_i)^2\right]\tau\right\}.
\end{aligned}
$$

Δηλαδή,

$$
\mu_i \mid \tau, y \sim \mathcal{N}\left(\frac{ca+n_i\tau\overline{y}_i}{c+n_i\tau}, \frac{1}{c+n_i\tau}\right), \quad \tau\mid\mu,y\sim\text{Gamma}\left(p+\frac{n}{2}, q+\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij}-\mu_i)^2\right).
$$

```
MCMCanova = function(Y, X, mu0, tau0, a, c, p, q, niter, nburn) {
    n = length(Y)
    m = length(levels(X))
    N = table(X)
    S = aggregate(Y ~ X, FUN = sum)[, 2]
    mu = matrix(0, niter, m)
    tau = numeric(niter)
    mu[1, ] = mu0
    tau[1] = tau0
    for (i in 2:niter) {
        mu[i, ] = rnorm(m, (c * a + tau[i - 1] * S)/(c + N * tau[i - 1]), (c +
```

```
            N * tau[i - 1])^(-0.5))
        tau[i] = rgamma(1, p + n/2, q + sum((Y - mu[i, X])^2)/2)
    }
    return(list(tau = tau[-(1:nburn)], mu = mu[-(1:nburn), ]))
}

n = 1000
m = 2
tau = 3
mu = c(1, 4)
X = factor(sample(m, n, replace = TRUE), levels = 1:m)
Y = rnorm(n, mu[X], tau^(-0.5))
posterior = MCMCanova(Y, X, numeric(m), 1, 0, 0, 0.5, 0, 5000, 1000)
par(mfrow = c(1, 3))
hist(posterior$mu[, 1], "FD", freq = FALSE, main = NA, xlab = expression(mu[1]))
abline(v = mu[1], col = 2, lty = 2)
hist(posterior$mu[, 2], "FD", freq = FALSE, main = NA, xlab = expression(mu[2]))
abline(v = mu[2], col = 2, lty = 2)
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



1.4 Γραμμικό Μοντέλο με Student's t Σφάλματα

Θεωρούμε το γραμμικό μοντέλο $y_i = x_i^{\mathrm{T}}\beta + \varepsilon_i$, όπου $\beta \in \mathbb{R}^k$ και τα τυχαία σφάλματα ε_i ακολουθούν τη γενικευμένη κατανομή Student's t με μέσο 0, ακρίβεια $\tau > 0$ και $\nu > 0$ βαθμούς ελευθερίας, δηλαδή:

$$f(y_i \mid \beta, \tau, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\sqrt{\frac{\tau}{\nu\pi}}\left[1 + \frac{1}{\nu}\tau\left(y_i - x_i^{\mathrm{T}}\beta\right)^2\right]^{-\frac{\nu+1}{2}}, \quad y_i \in \mathbb{R}.$$

Θεωρούμε τις τυχαίες μεταβλητές $W_i \sim \mathcal{N}\left(0, \tau^{-1}\right)$ και $V_i \sim \chi^2_\nu \equiv \text{Gamma}\left(\frac{\nu}{2}, \frac{1}{2}\right)$. Τότε, παρατηρούμε ότι:

$$Y_i \stackrel{d}{=} \frac{W_i}{\sqrt{\frac{V_i}{\nu}}} + x_i^{\mathrm{T}}\beta.$$

Θέτουμε $Z_i = \frac{V_i}{\nu}$. Τότε, $Z_i \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$. Παρατηρούμε ότι:

$$Y_i \mid z_i \stackrel{d}{=} \frac{W_i}{\sqrt{z_i}} + x_i^{\mathrm{T}}\beta \sim \mathcal{N}\left(x_i^{\mathrm{T}}\beta, \tau^{-1}z_i^{-1}\right).$$

Θεωρούμε γνωστή την παράμετρο ν και τη δεσμευμένα συζυγή prior κατανομή $\beta \mid \tau \sim \mathcal{N}_k\left(a, \tau^{-1}C^{-1}\right)$, $\tau \sim \text{Gamma}(p, q)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές $\pi(\beta, \tau \mid z, y)$, $\pi(\tau \mid \beta, z, y)$ και $f(z_i \mid y_i, \beta, \tau)$.

Λύση.

Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$\begin{aligned}
\pi(\beta, \tau) &= \pi(\beta \mid \tau) \cdot \pi(\tau) \\
&= (2\pi)^{-\frac{k}{2}} \left|\tau^{-1}C^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{\tau(\beta-a)^{\mathrm{T}}C(\beta-a)}{2}\right\} \cdot \frac{q^p}{\Gamma(p)}\tau^{p-1}e^{-q\tau} \\
&\propto \tau^{p+\frac{k}{2}-1} \exp\left\{-\left[q + \frac{(\beta-a)^{\mathrm{T}}C(\beta-a)}{2}\right]\tau\right\} \\
&= \tau^{\frac{k}{2}} \exp\left\{-\tau\frac{\beta^{\mathrm{T}}C\beta - 2\beta^{\mathrm{T}}Ca + a^{\mathrm{T}}Ca}{2}\right\} \cdot \tau^{p-1}e^{-q\tau} \\
&= \tau^{\frac{k}{2}} \exp\left\{-\frac{\beta^{\mathrm{T}}\tau C\beta}{2} + \beta^{\mathrm{T}}\tau Ca\right\} \cdot \tau^{p-1} \exp\left\{-\left(q + \frac{a^{\mathrm{T}}Ca}{2}\right)\tau\right\}.
\end{aligned}$$

Η πλήρης πιθανοφάνεια, δηλαδή η από κοινού πιθανοφάνεια των παρατηρούμενων μεταβλητών y_i και των λανθανουσών μεταβλητών z_i, δίνεται από τη σχέση:

$$\begin{aligned}
f(y, z \mid \beta, \tau) &= \prod_{i=1}^n f(y_i, z_i \mid \beta, \tau) \\
&= \prod_{i=1}^n f(z_i) f(y_i \mid z_i, \beta, \tau) \\
&= \prod_{i=1}^n \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} z_i^{\frac{\nu}{2}-1} e^{-\frac{\nu}{2}z_i} \sqrt{\frac{\tau z_i}{2\pi}} \exp\left\{-\frac{\tau z_i \left(y_i - x_i^{\mathrm{T}}\beta\right)^2}{2}\right\} \\
&\propto \tau^{\frac{n}{2}} \cdot \prod_{i=1}^n z_i^{\frac{\nu+1}{2}-1} \exp\left\{-\sum_{i=1}^n \frac{\nu + \tau\left(y_i - x_i^{\mathrm{T}}\beta\right)^2}{2} z_i\right\}.
\end{aligned}$$

Ορίζουμε $y = (y_1, \ldots, y_n)^{\mathrm{T}} \in \mathbb{R}^n$, τον πίνακα σχεδιασμού $X = (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^{n \times k}$ και τον διαγώνιο πίνακα βαρών $Z = \text{diag}(z_1, \ldots, z_n) \in \mathbb{R}^{n \times n}$. Τότε, παρατηρούμε ότι $y \mid z \sim N_n\left(X\beta, \tau^{-1}Z^{-1}\right)$. Δηλαδή, η πλήρης πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$\begin{aligned}
f(y,z \mid \beta,\tau) &= f(z)\cdot f(y \mid z,\beta,\tau) \\
&= f(y \mid z,\beta,\tau)\cdot \prod_{i=1}^{n} f(z_i) \\
&= (2\pi)^{-\frac{n}{2}} \left|\tau^{-1}Z^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{\tau(y-X\beta)^{\mathrm{T}}Z(y-X\beta)}{2}\right\}\cdot \prod_{i=1}^{n} \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} z_i^{\frac{\nu}{2}-1} e^{-\frac{\nu}{2}z_i} \\
&\propto \tau^{\frac{n}{2}} \exp\left\{-\frac{(y-X\beta)^{\mathrm{T}}Z(y-X\beta)}{2}\tau\right\}\cdot |Z|^{\frac{1}{2}} \prod_{i=1}^{n} z_i^{\frac{\nu}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\} \\
&= \tau^{\frac{n}{2}} \exp\left\{-\frac{y^{\mathrm{T}}Zy - 2\beta^{\mathrm{T}}X^{\mathrm{T}}Zy + \beta^{\mathrm{T}}X^{\mathrm{T}}ZX\beta}{2}\tau\right\}\cdot \prod_{i=1}^{n} z_i^{\frac{1}{2}} \prod_{i=1}^{n} z_i^{\frac{\nu}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\} \\
&= \exp\left\{-\frac{\beta^{\mathrm{T}}\tau X^{\mathrm{T}}ZX\beta}{2} + \beta^{\mathrm{T}}\tau X^{\mathrm{T}}Zy\right\}\cdot \tau^{\frac{n}{2}} \exp\left\{-\frac{y^{\mathrm{T}}Zy}{2}\tau\right\} \\
&\quad \times \prod_{i=1}^{n} z_i^{\frac{\nu+1}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\}.
\end{aligned}$$

Επομένως, παίρνουμε την από κοινού δεσμευμένη posterior κατανομή των β και τ ως εξής:

$$\begin{aligned}
\pi(\beta,\tau \mid z,y) &\propto \pi(\beta,\tau,z \mid y) \\
&\propto \pi(\beta,\tau)\cdot f(y,z \mid \beta,\tau) \\
&\propto \tau^{\frac{k}{2}} \exp\left\{-\frac{\beta^{\mathrm{T}}\tau C\beta}{2} + \beta^{\mathrm{T}}\tau Ca\right\}\cdot \tau^{p-1} \exp\left\{-\left(q + \frac{a^{\mathrm{T}}Ca}{2}\right)\tau\right\} \\
&\quad \times \exp\left\{-\frac{\beta^{\mathrm{T}}\tau X^{\mathrm{T}}ZX\beta}{2} + \beta^{\mathrm{T}}\tau X^{\mathrm{T}}Zy\right\}\cdot \tau^{\frac{n}{2}} \exp\left\{-\frac{y^{\mathrm{T}}Zy}{2}\tau\right\} \\
&= \tau^{\frac{k}{2}} \exp\left\{-\frac{1}{2}\beta^{\mathrm{T}}\tau \underbrace{\left(C + X^{\mathrm{T}}ZX\right)}_{C_n}\beta + \beta^{\mathrm{T}}\tau\left(Ca + X^{\mathrm{T}}Zy\right)\right\} \\
&\quad \times \tau^{p+\frac{n}{2}-1} \exp\left\{-\left(q + \frac{a^{\mathrm{T}}Ca + y^{\mathrm{T}}Zy}{2}\right)\tau\right\} \\
&= \tau^{\frac{k}{2}} \exp\left\{-\frac{\beta^{\mathrm{T}}\tau C_n\beta}{2} + \beta^{\mathrm{T}}\tau \underbrace{\left(C + X^{\mathrm{T}}ZX\right)}_{C_n} \underbrace{\left(C + X^{\mathrm{T}}ZX\right)^{-1}\left(Ca + X^{\mathrm{T}}Zy\right)}_{a_n}\right\} \\
&\quad \times \tau^{p+\frac{n}{2}-1} \exp\left\{-\left(q + \frac{a^{\mathrm{T}}Ca + y^{\mathrm{T}}Zy}{2}\right)\tau\right\} \\
&= \tau^{\frac{k}{2}} \exp\left\{-\frac{\beta^{\mathrm{T}}\tau C_n\beta}{2} + \beta^{\mathrm{T}}\tau C_n a_n - \frac{a_n^{\mathrm{T}}\tau C_n a_n}{2} + \frac{a_n^{\mathrm{T}}\tau C_n a_n}{2}\right\} \\
&\quad \times \tau^{p+\frac{n}{2}-1} \exp\left\{-\left(q + \frac{a^{\mathrm{T}}Ca + y^{\mathrm{T}}Zy}{2}\right)\tau\right\} \\
&= \tau^{\frac{k}{2}} \exp\left\{-\frac{\tau(\beta-a_n)^{\mathrm{T}}C_n(\beta-a_n)}{2}\right\} \\
&\quad \times \tau^{p+\frac{n}{2}-1} \exp\left\{-\left(q + \frac{a^{\mathrm{T}}Ca + y^{\mathrm{T}}Zy - a_n^{\mathrm{T}}C_n a_n}{2}\right)\tau\right\}.
\end{aligned}$$

Υπολογίζουμε ότι:

$$a_n^{\mathrm{T}} C_n a_n = \left(Ca + X^{\mathrm{T}} Z y\right)^{\mathrm{T}} \left(C + X^{\mathrm{T}} Z X\right)^{-1} \left(Ca + X^{\mathrm{T}} Z y\right).$$

Δηλαδή,

$$\beta \mid \tau, z, y \sim \mathcal{N}_k \left(\left(C + X^{\mathrm{T}} Z X\right)^{-1} \left(Ca + X^{\mathrm{T}} Z y\right), \tau^{-1} \left(C + X^{\mathrm{T}} Z X\right)^{-1} \right),$$

$$\tau \mid z, y \sim \mathrm{Gamma} \left(p + \frac{n}{2}, q + \frac{a^{\mathrm{T}} C a + y^{\mathrm{T}} Z y - a_n^{\mathrm{T}} C_n a_n}{2} \right).$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή του τ ως εξής:

$$\begin{aligned}
\pi(\tau \mid \beta, z, y) &\propto \pi(\beta, \tau) \cdot f(y, z \mid \beta, \tau) \\
&\propto \tau^{p + \frac{k}{2} - 1} \exp\left\{ - \left[q + \frac{(\beta - a)^{\mathrm{T}} C (\beta - a)}{2} \right] \tau \right\} \cdot \tau^{\frac{n}{2}} \exp\left\{ - \frac{(y - X\beta)^{\mathrm{T}} Z (y - X\beta)}{2} \tau \right\} \\
&= \tau^{p + \frac{n+k}{2} - 1} \exp\left\{ - \left[q + \frac{(\beta - a)^{\mathrm{T}} C (\beta - a) + (y - X\beta)^{\mathrm{T}} Z (y - X\beta)}{2} \right] \tau \right\}.
\end{aligned}$$

Δηλαδή,

$$\tau \mid \beta, z, y \sim \mathrm{Gamma} \left(p + \frac{n+k}{2}, q + \frac{(\beta - a)^{\mathrm{T}} C (\beta - a) + (y - X\beta)^{\mathrm{T}} Z (y - X\beta)}{2} \right).$$

Τέλος, παίρνουμε τη δεσμευμένη posterior κατανομή των λανθανουσών μεταβλητών z_i ως εξής:

$$f(z_i \mid y_i, \beta, \tau) \propto f(y_i, z_i \mid \beta, \tau) \propto z_i^{\frac{\nu+1}{2} - 1} \exp\left\{ - \frac{\nu + \tau \left(y_i - x_i^{\mathrm{T}} \beta\right)^2}{2} z_i \right\}.$$

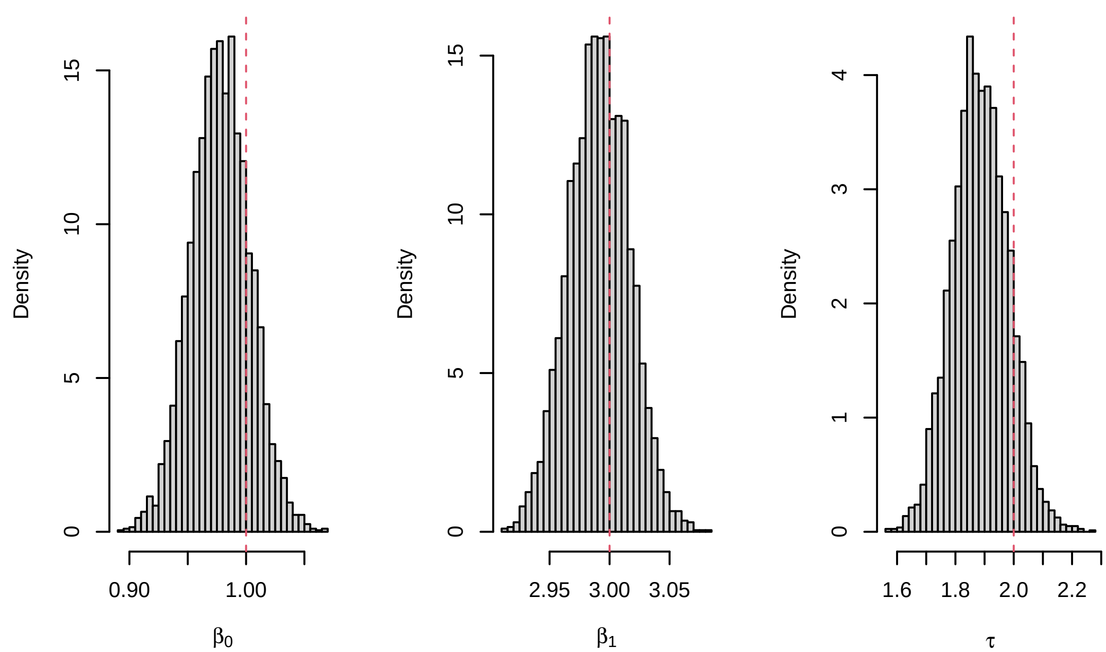
Δηλαδή,

$$z_i \mid y_i, \beta, \tau \sim \mathrm{Gamma} \left(\frac{\nu + 1}{2}, \frac{\nu + \tau \left(y_i - x_i^{\mathrm{T}} \beta\right)^2}{2} \right).$$

Πρώτα θα υλοποιήσουμε έναν δειγματολήπτη Gibbs που προσομοιώνει διαδοχικά από τις δεσμευμένες posterior κατανομές των παραμέτρων β, τ και των λανθανουσών μεταβλητών z_i.

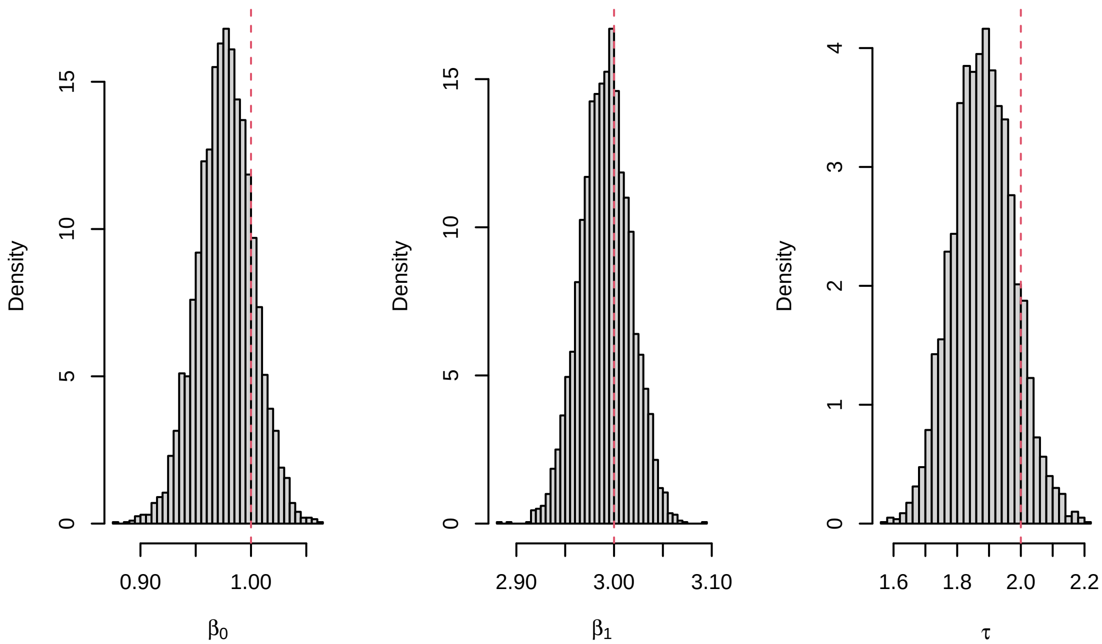
```
MCMCtlm = function(Y, X, beta0, tau0, nu, a, C, p, q, niter, nburn) {
    library(MASS)
    n = length(Y)
    k = dim(X)[2]
    beta = matrix(0, niter, k)
    tau = numeric(niter)
    Z = matrix(0, niter, n)
    beta[1, ] = beta0
    tau[1] = tau0
    Z[1, ] = rgamma(n, (nu + 1)/2, (nu + tau[1] * (Y - X %*% beta[1, ])^2)/2)
    for (i in 2:niter) {
        Cn = C + crossprod(X, Z[i - 1, ] * X)
        an = solve(Cn, C %*% a + crossprod(X, Z[i - 1, ] * Y))
        beta[i, ] = mvrnorm(1, an, solve(Cn)/tau[i - 1])
```

```r
        tau[i] = rgamma(1, p + (n + k)/2, q + (crossprod(beta[i, ] - a, C %*%
            (beta[i, ] - a)) + crossprod(Y - X %*% beta[i, ], Z[i - 1, ] * (Y -
            X %*% beta[i, ])))/2)
        Z[i, ] = rgamma(n, (nu + 1)/2, (nu + tau[i] * (Y - X %*% beta[i, ])^2)/2)
    }
    return(list(beta = beta[-(1:nburn), ], tau = tau[-(1:nburn)], Z = Z[-(1:nburn),
        ]))
}

library(mvtnorm)
n = 1000
k = 2
beta = c(1, 3)
tau = 2
nu = 10
X = cbind(1, rnorm(n))
Y = X %*% beta + rmvt(n, matrix(tau^(-1)), nu)
posterior = MCMCtlm(Y, X, numeric(k), 1, nu, numeric(k), matrix(0, k, k), 0.5,
    0, 5000, 1000)
par(mfrow = c(1, 3))
hist(posterior$beta[, 1], "FD", freq = FALSE, main = NA, xlab = expression(beta[0]))
abline(v = beta[1], col = 2, lty = 2)
hist(posterior$beta[, 2], "FD", freq = FALSE, main = NA, xlab = expression(beta[1]))
abline(v = beta[2], col = 2, lty = 2)
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



Έπειτα, θα υλοποιήσουμε έναν αλγόριθμο Gibbs που προσομοιώνει εναλλάξ από τις δεσμευμένες posterior κατανομές της παραμέτρου (β, τ) και των λανθανουσών μεταβλητών z_i.

```
MCMCtlm = function(Y, X, beta0, tau0, nu, a, C, p, q, niter, nburn) {
    library(MASS)
    n = length(Y)
    k = dim(X)[2]
    beta = matrix(0, niter, k)
    tau = numeric(niter)
    Z = matrix(0, niter, n)
    beta[1, ] = beta0
    tau[1] = tau0
    Z[1, ] = rgamma(n, (nu + 1)/2, (nu + tau[1] * (Y - X %*% beta[1, ])^2)/2)
    for (i in 2:niter) {
        Cn = C + crossprod(X, Z[i - 1, ] * X)
        an = solve(Cn, C %*% a + crossprod(X, Z[i - 1, ] * Y))
        qn = q + (crossprod(a, C %*% a) + crossprod(Y, Z[i - 1, ] * Y) - crossprod(C %*%
            a + crossprod(X, Z[i - 1, ] * Y), an))/2
        tau[i] = rgamma(1, p + n/2, qn)
        beta[i, ] = mvrnorm(1, an, solve(Cn)/tau[i])
        Z[i, ] = rgamma(n, (nu + 1)/2, (nu + tau[i] * (Y - X %*% beta[i, ])^2)/2)
    }
    return(list(beta = beta[-(1:nburn), ], tau = tau[-(1:nburn)], Z = Z[-(1:nburn),
        ]))
}

posterior = MCMCtlm(Y, X, numeric(k), 1, nu, numeric(k), matrix(0, k, k), 0.5,
    0, 5000, 1000)
par(mfrow = c(1, 3))
hist(posterior$beta[, 1], "FD", freq = FALSE, main = NA, xlab = expression(beta[0]))
abline(v = beta[1], col = 2, lty = 2)
hist(posterior$beta[, 2], "FD", freq = FALSE, main = NA, xlab = expression(beta[1]))
abline(v = beta[2], col = 2, lty = 2)
hist(posterior$tau, "FD", freq = FALSE, main = NA, xlab = expression(tau))
abline(v = tau, col = 2, lty = 2)
```



1.5 Πολυδιάστατη Κανονική Κατανομή

Ορισμός 1.2. Ορίζουμε την πολυδιάστατη συνάρτηση Γάμμα ως εξής:

$$\Gamma_k(x) = \pi^{\frac{k(k-1)}{4}} \prod_{j=1}^{k} \Gamma\left(x + \frac{1-j}{2}\right), \quad x > \frac{k-1}{2}, \quad k \in \mathbb{N}.$$

Ορισμός 1.3. Λέμε ότι ένας θετικά ορισμένος τυχαίος πίνακας $X \in \mathbb{R}^{k\times k}$ ακολουθεί την κατανομή Wishart με θετικά ορισμένο πίνακα κλίμακας $A \in \mathbb{R}^{k\times k}$, $\nu > 0$ βαθμούς ελευθερίας και γράφουμε $X \sim \mathcal{W}_k\left(A, \nu\right)$, αν έχει συνάρτηση πυκνότητας πιθανότητας:

$$f_X(x \mid A, \nu) = \frac{1}{2^{\frac{\nu k}{2}} |A|^{\frac{\nu}{2}} \Gamma_k\left(\frac{\nu}{2}\right)} |x|^{\frac{\nu-k-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(A^{-1}x)}, \quad x \in \mathbb{R}^{k\times k}.$$

Έστω τυχαίο δείγμα $y_1, \dots, y_n$ από την πολυδιάστατη κανονική κατανομή $\mathcal{N}_k\left(\mu, \Sigma\right)$.

α. Θεωρούμε prior ανεξαρτησία με κατανομές $\mu \sim \mathcal{N}_k\left(a, C^{-1}\right)$ και $\Omega \sim \mathcal{W}_k\left(A^{-1}, d\right)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των μ και Ω.

β. Θεωρούμε τη συζυγή prior κατανομή $\mu \mid \Omega \sim \mathcal{N}_k\left(a, c^{-1}\Omega^{-1}\right)$, $\Omega \sim \mathcal{W}_k\left(A^{-1}, d\right)$. Να υπολογιστούν οι δεσμευμένες και οι περιθώριες posterior κατανομές των παραμέτρων μ και Ω.

Λύση.

α. Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$
\begin{aligned}
\pi(\mu, \Omega) &= \pi(\mu) \cdot \pi(\Omega) \\
&= (2\pi)^{-\frac{k}{2}} \left|C^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{(\mu - a)^{\mathrm{T}} C (\mu - a)}{2}\right\} \cdot \frac{|A|^{\frac{d}{2}}}{2^{\frac{dk}{2}} \Gamma_k\left(\frac{d}{2}\right)} |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\mathrm{tr}(A\Omega)} \\
&\propto \exp\left\{-\frac{\mu^{\mathrm{T}} C \mu - 2\mu^{\mathrm{T}} C a + a^{\mathrm{T}} C a}{2}\right\} \cdot |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\mathrm{tr}(A\Omega)} \\
&\propto \exp\left\{-\frac{\mu^{\mathrm{T}} C \mu}{2} + \mu^{\mathrm{T}} C a\right\} \cdot |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\mathrm{tr}(A\Omega)}.
\end{aligned}
$$

Λήμμα 1.1. Έστω $x \in \mathbb{R}^n$ και $A \in \mathbb{R}^{n\times n}$. Τότε,

$$
x^{\mathrm{T}} A x = \mathrm{tr}\left(x^{\mathrm{T}} A x\right) = \mathrm{tr}\left(x x^{\mathrm{T}} A\right).
$$

Η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$
\begin{aligned}
f(y \mid \mu, \Omega) &= \prod_{i=1}^{n} f(y_i \mid \mu, \Omega) \\
&= \prod_{i=1}^{n} (2\pi)^{-\frac{k}{2}} \left|\Omega^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{(y_i - \mu)^{\mathrm{T}} \Omega (y_i - \mu)}{2}\right\} \\
&\propto |\Omega|^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\sum_{i=1}^{n} (y_i - \mu)(y_i - \mu)^{\mathrm{T}} \Omega\right]\right\} \\
&= |\Omega|^{\frac{n}{2}} \exp\left\{-\sum_{i=1}^{n} \frac{y_i^{\mathrm{T}} \Omega y_i - 2\mu^{\mathrm{T}} \Omega y_i + \mu^{\mathrm{T}} \Omega \mu}{2}\right\} \\
&= \exp\left\{-\frac{\mu^{\mathrm{T}} n\Omega \mu}{2} + \mu^{\mathrm{T}} n\Omega \overline{y}\right\} \cdot |\Omega|^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\sum_{i=1}^{n} y_i y_i^{\mathrm{T}} \Omega\right)\right\}.
\end{aligned}
$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των μ και Ω ως εξής:

$$
\begin{aligned}
\pi(\mu \mid \Omega, y) &\propto \pi(\mu, \Omega \mid y) \\
&\propto \pi(\mu, \Omega) \cdot f(y \mid \mu, \Omega) \\
&\propto \exp\left\{-\frac{\mu^{\mathrm{T}} C \mu}{2} + \mu^{\mathrm{T}} C a\right\} \cdot \exp\left\{-\frac{\mu^{\mathrm{T}} n\Omega \mu}{2} + \mu^{\mathrm{T}} n\Omega \overline{y}\right\} \\
&= \exp\left\{-\frac{1}{2}\mu^{\mathrm{T}} \underbrace{(C + n\Omega)}_{C_n} \mu + \mu^{\mathrm{T}} \left(Ca + n\Omega \overline{y}\right)\right\} \\
&= \exp\left\{-\frac{\mu^{\mathrm{T}} \left(C + n\Omega\right) \mu}{2} + \mu^{\mathrm{T}} \underbrace{(C + n\Omega)}_{C_n} \underbrace{(C + n\Omega)^{-1} \left(Ca + n\Omega \overline{y}\right)}_{a_n}\right\},
\end{aligned}
$$

$$
\begin{aligned}
\pi(\Omega \mid \mu, y) &\propto \pi(\mu, \Omega) \cdot f(y \mid \mu, \Omega) \\
&\propto |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(A\Omega)} \cdot |\Omega|^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left[\sum_{i=1}^{n}(y_i-\mu)(y_i-\mu)^{\mathrm{T}}\Omega\right]\right\} \\
&= |\Omega|^{\frac{d+n-k-1}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left[\left(A+\sum_{i=1}^{n}(y_i-\mu)(y_i-\mu)^{\mathrm{T}}\right)\Omega\right]\right\}.
\end{aligned}
$$

Δηλαδή,

$$
\mu \mid \Omega, y \sim \mathcal{N}_k\left((C+n\Omega)^{-1}(Ca+n\Omega\overline{y}), (C+n\Omega)^{-1}\right),
$$

$$
\Omega \mid \mu, y \sim \mathcal{W}_k\left(\left(A+\sum_{i=1}^{n}(y_i-\mu)(y_i-\mu)^{\mathrm{T}}\right)^{-1}, d+n\right).
$$

Ορισμός 1.4. Ορίζουμε το γινόμενο Kronecker δύο πινάκων $A \in \mathbb{R}^{n\times m}$ και $B \in \mathbb{R}^{p\times q}$ ως τον πίνακα:

$$
C = A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{bmatrix} \in \mathbb{R}^{np\times mq}.
$$

Λήμμα 1.2. Έστω $A \in \mathbb{R}^{n\times n}$ θετικά ορισμένος πίνακας, $x, a \in \mathbb{R}^n$ και $c \in \mathbb{R}$. Τότε, ισχύει ότι:

$$
\frac{\partial c}{\partial x} = \mathbf{0}_n, \quad \frac{\partial a}{\partial x} = \mathbf{0}_{n\times n}, \quad \frac{\partial Ax}{\partial x} = A, \quad \frac{\partial a^{\mathrm{T}}x}{\partial x} = a, \quad \frac{\partial x^{\mathrm{T}}Ax}{\partial x} = 2Ax,
$$

$$
\frac{\partial aa^{\mathrm{T}}}{\partial x} = \mathbf{0}_{n^2\times n}, \quad \frac{\partial ax^{\mathrm{T}}}{\partial x} = \frac{\partial xa^{\mathrm{T}}}{\partial x} = a \otimes I_n, \quad \frac{\partial xx^{\mathrm{T}}}{\partial x} = x \otimes I_n + I_n \otimes x,
$$

$$
\frac{\partial x^{\mathrm{T}}Ax}{\partial A} = xx^{\mathrm{T}}, \quad \frac{\partial \log|A|}{\partial A} = A^{-1}, \quad \frac{\partial A^{-1}}{\partial A} = -A^{-1} \otimes A^{-1}, \quad \frac{\partial Ax}{\partial A} = x^{\mathrm{T}} \otimes I_n.
$$

Λήμμα 1.3. Έστω $X \in \mathbb{R}^k$ τυχαίο διάνυσμα και $A \in \mathbb{R}^{n\times m}$ σταθερός πίνακας. Τότε, ισχύει ότι:

$$
\mathbb{E}\left(A \otimes X\right) = A \otimes \mathbb{E}(X), \quad \mathbb{E}\left(X \otimes A\right) = \mathbb{E}(X) \otimes A.
$$

Λήμμα 1.4. Έστω $A \in \mathbb{R}^{n\times n}$ και $B \in \mathbb{R}^{m\times m}$. Τότε, ισχύει ότι:

$$
|A \otimes B| = |A|^m |B|^n.
$$

Μπορούμε να υπολογίσουμε την prior του Jeffreys για την πολυδιάστατη κανονική κατανομή ως εξής:

$$
\log f(y \mid \mu, \Omega) = \frac{1}{2}\log|\Omega| - \frac{k}{2}\log(2\pi) - \frac{(y-\mu)^{\mathrm{T}}\Omega(y-\mu)}{2},
$$

$$
\frac{\partial \log f(y \mid \mu, \Omega)}{\partial \mu} = \Omega(y-\mu) \in \mathbb{R}^k, \quad \frac{\partial \log f(y \mid \mu, \Omega)}{\partial \Omega} = \frac{1}{2}\Omega^{-1} - \frac{(y-\mu)(y-\mu)^{\mathrm{T}}}{2} \in \mathbb{R}^{k\times k},
$$

$$
\frac{\partial^2 \log f(y \mid \mu, \Omega)}{\partial \mu \partial \mu} = -\Omega \in \mathbb{R}^{k\times k}, \quad \frac{\partial^2 \log f(y \mid \mu, \Omega)}{\partial \Omega \partial \Omega} = -\frac{1}{2}\Omega^{-1} \otimes \Omega^{-1} \in \mathbb{R}^{k^2\times k^2},
$$

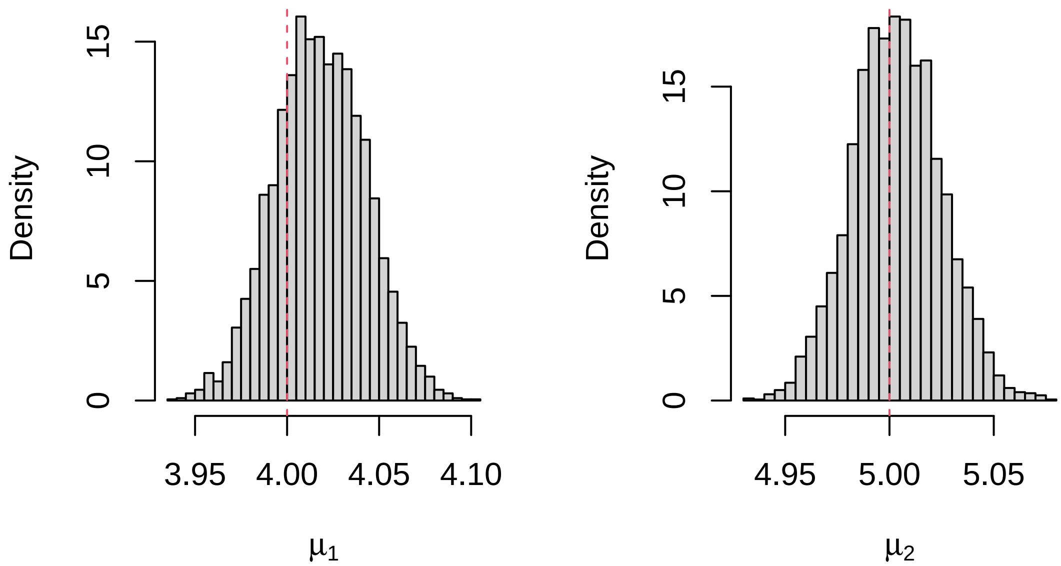
$$\frac{\partial^2 \log f(y \mid \mu, \Omega)}{\partial \Omega \partial \mu} = (y-\mu)^{\mathrm{T}} \otimes I_k \in \mathbb{R}^{k \times k^2}, \quad \frac{\partial^2 \log f(y \mid \mu, \Omega)}{\partial \mu \partial \Omega} = \frac{(y-\mu) \otimes I_k + I_k \otimes (y-\mu)}{2} \in \mathbb{R}^{k^2 \times k},$$

$$\mathcal{I}(\mu, \Omega) = \mathbb{E}\left[-\frac{\partial^2 \log f(y \mid \mu, \Omega)}{\partial(\mu, \Omega)\partial(\mu, \Omega)}\right] = \begin{bmatrix} \Omega & \mathbf{0}_{k \times k^2} \\ \mathbf{0}_{k^2 \times k} & \frac{1}{2}\Omega^{-1} \otimes \Omega^{-1} \end{bmatrix},$$

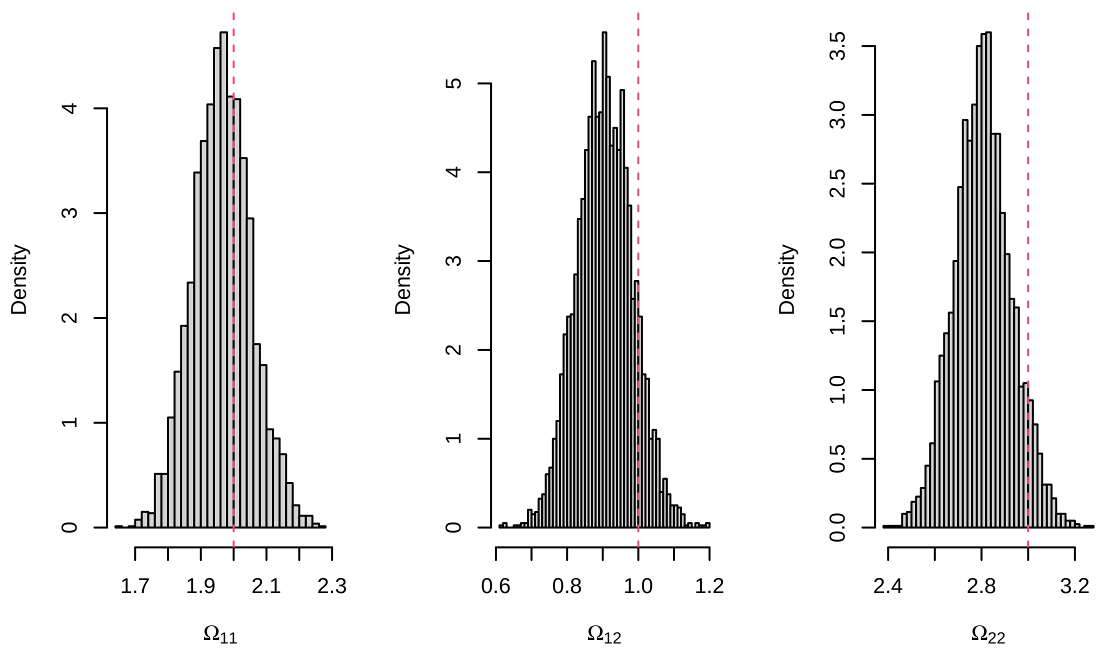
$$J(\mu, \Omega) \propto \sqrt{|\mathcal{I}(\mu, \Omega)|} = \sqrt{|\Omega| \left|\frac{1}{2}\Omega^{-1} \otimes \Omega^{-1}\right|} \propto \sqrt{|\Omega| \, |\Omega|^{-k} \, |\Omega|^{-k}} = |\Omega|^{\frac{1-2k}{2}}.$$

Παρατηρούμε ότι η μη-ορθή prior του Jeffreys προκύπτει για $a = \mathbf{0}_k$, $C = A = \mathbf{0}_{k \times k}$ και $d = 2 - k$

```
MCMCmvnorm = function(Y, mu0, Omega0, a, C, A, d, niter, nburn) {
    library(MASS)
    n = dim(Y)[1]
    k = dim(Y)[2]
    S = colSums(Y)
    mu = matrix(0, niter, k)
    Omega = array(0, c(k, k, niter))
    mu[1, ] = mu0
    Omega[, , 1] = Omega0
    for (i in 2:niter) {
        Cn = C + n * Omega[, , i - 1]
        an = solve(Cn, C %*% a + Omega[, , i - 1] %*% S)
        mu[i, ] = mvrnorm(1, an, solve(Cn))
        Omega[, , i] = rWishart(1, d + n, solve(A + tcrossprod(t(Y) - mu[i,
            ])))
    }
    return(list(mu = mu[-(1:nburn), ], Omega = Omega[, , -(1:nburn)]))
}

library(MASS)
n = 1000
k = 2
mu = c(4, 5)
Omega = matrix(c(2, 1, 1, 3), k)
Y = mvrnorm(n, mu, solve(Omega))
posterior = MCMCmvnorm(Y, numeric(k), diag(k), numeric(k), matrix(0, k, k),
    matrix(0, k, k), 2 - k, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$mu[, 1], "FD", freq = FALSE, main = NA, xlab = expression(mu[1]))
abline(v = mu[1], col = 2, lty = 2)
hist(posterior$mu[, 2], "FD", freq = FALSE, main = NA, xlab = expression(mu[2]))
abline(v = mu[2], col = 2, lty = 2)
```




```
par(mfrow = c(1, 3))
hist(posterior$Omega[1, 1, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[11]))
abline(v = Omega[1, 1], col = 2, lty = 2)
hist(posterior$Omega[1, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[12]))
abline(v = Omega[1, 2], col = 2, lty = 2)
hist(posterior$Omega[2, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[22]))
abline(v = Omega[2, 2], col = 2, lty = 2)
```



β. Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$\begin{aligned}
\pi(\mu, \Omega) &= \pi(\mu \mid \Omega) \cdot \pi(\Omega) \\
&= (2\pi)^{-\frac{k}{2}} \left|c^{-1}\Omega^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{c(\mu-a)^{\mathrm{T}}\Omega(\mu-a)}{2}\right\} \cdot \frac{|A|^{\frac{d}{2}}}{2^{\frac{dk}{2}}\Gamma_k\left(\frac{d}{2}\right)} |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\mathrm{tr}(A\Omega)} \\
&\propto |\Omega|^{\frac{d+1-k-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + c(\mu-a)(\mu-a)^{\mathrm{T}}\right)\Omega\right]\right\} \\
&= |\Omega|^{\frac{1}{2}} \exp\left\{-c\frac{\mu^{\mathrm{T}}\Omega\mu - 2\mu^{\mathrm{T}}\Omega a + a^{\mathrm{T}}\Omega a}{2}\right\} \cdot |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\mathrm{tr}(A\Omega)} \\
&= |\Omega|^{\frac{1}{2}} \exp\left\{-\frac{\mu^{\mathrm{T}}c\Omega\mu}{2} + \mu^{\mathrm{T}}c\Omega a\right\} \cdot |\Omega|^{\frac{d-k-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + caa^{\mathrm{T}}\right)\Omega\right]\right\}.
\end{aligned}$$

Επομένως, παίρνουμε την από κοινού posterior κατανομή των μ και Ω ως εξής:

$$\begin{aligned}
\pi(\mu, \Omega \mid y) &\propto \pi(\mu, \Omega) \cdot f(y \mid \mu, \Omega) \\
&\propto |\Omega|^{\frac{1}{2}} \exp\left\{-\frac{\mu^{\mathrm{T}}c\Omega\mu}{2} + \mu^{\mathrm{T}}c\Omega a\right\} \cdot |\Omega|^{\frac{d-k-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + caa^{\mathrm{T}}\right)\Omega\right]\right\} \\
&\quad \times \exp\left\{-\frac{\mu^{\mathrm{T}}n\Omega\mu}{2} + \mu^{\mathrm{T}}n\Omega\overline{y}\right\} \cdot |\Omega|^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\sum_{i=1}^{n} y_i y_i^{\mathrm{T}}\Omega\right)\right\} \\
&= |\Omega|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\mu^{\mathrm{T}}\underbrace{(c+n)}_{c_n}\Omega\mu + \mu^{\mathrm{T}}\Omega\left(ca + n\overline{y}\right)\right\} \\
&\quad \times |\Omega|^{\frac{d+n-k-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}}\right)\Omega\right]\right\} \\
&= |\Omega|^{\frac{1}{2}} \exp\left\{-\frac{\mu^{\mathrm{T}}(c+n)\Omega\mu}{2} + \mu^{\mathrm{T}}\underbrace{(c+n)}_{c_n}\Omega\underbrace{\frac{ca+n\overline{y}}{c+n}}_{a_n}\right\} \\
&\quad \times |\Omega|^{\frac{d+n-k-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}}\right)\Omega\right]\right\} \\
&= |\Omega|^{\frac{1}{2}} \exp\left\{-\frac{\mu^{\mathrm{T}}c_n\Omega\mu}{2} + \mu^{\mathrm{T}}c_n\Omega a_n - \frac{a_n^{\mathrm{T}}c_n\Omega a_n}{2} + \frac{a_n^{\mathrm{T}}c_n\Omega a_n}{2}\right\} \\
&\quad \times |\Omega|^{\frac{d+n-k-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}}\right)\Omega\right]\right\} \\
&= |\Omega|^{\frac{1}{2}} \exp\left\{-\frac{c_n(\mu-a_n)^{\mathrm{T}}\Omega(\mu-a_n)}{2}\right\} \\
&\quad \times |\Omega|^{\frac{d+n-k-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}} - c_n a_n a_n^{\mathrm{T}}\right)\Omega\right]\right\}.
\end{aligned}$$

Υπολογίζουμε ότι:

$$c_n a_n a_n^{\mathrm{T}} = \frac{\left(ca + n\overline{y}\right)\left(ca + n\overline{y}\right)^{\mathrm{T}}}{c+n}.$$

Δηλαδή,

$$\mu \mid \Omega, y \sim \mathcal{N}_k \left(\frac{ca + n\overline{y}}{c+n}, \frac{1}{c+n}\Omega^{-1} \right),$$

$$\Omega \mid y \sim \mathcal{W}_k \left(\left(A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}} - \frac{(ca + n\overline{y})\,(ca + n\overline{y})^{\mathrm{T}}}{c+n} \right)^{-1}, d+n \right).$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή του Ω ως εξής:

$$\begin{aligned}
\pi(\Omega \mid \mu, y) &\propto \pi(\mu, \Omega) \cdot f(y \mid \mu, \Omega) \\
&\propto |\Omega|^{\frac{d+1-k-1}{2}} \exp\left\{ -\frac{1}{2} \mathrm{tr} \left[\left(A + c(\mu - a)(\mu - a)^{\mathrm{T}} \right) \Omega \right] \right\} \\
&\quad \times |\Omega|^{\frac{n}{2}} \exp\left\{ -\frac{1}{2} \mathrm{tr} \left[\sum_{i=1}^{n} (y_i - \mu)(y_i - \mu)^{\mathrm{T}} \Omega \right] \right\} \\
&= |\Omega|^{\frac{d+n+1-k-1}{2}} \exp\left\{ -\frac{1}{2} \mathrm{tr} \left[\left(A + c(\mu - a)(\mu - a)^{\mathrm{T}} + \sum_{i=1}^{n} (y_i - \mu)(y_i - \mu)^{\mathrm{T}} \right) \Omega \right] \right\}.
\end{aligned}$$

Δηλαδή,

$$\Omega \mid \mu, y \sim \mathcal{W}_k \left(\left(A + c(\mu - a)(\mu - a)^{\mathrm{T}} + \sum_{i=1}^{n} (y_i - \mu)(y_i - \mu)^{\mathrm{T}} \right)^{-1}, d+n+1 \right).$$

Ορισμός 1.5. Λέμε ότι ένα τυχαίο διάνυσμα $X \in \mathbb{R}^k$ ακολουθεί την πολυδιάστατη κατανομή Student's t με μέσο $\mu \in \mathbb{R}^k$, θετικά ορισμένο πίνακα συνδιακύμανσης $\Sigma \in \mathbb{R}^{k \times k}$, $\nu > 0$ βαθμούς ελευθερίας και γράφουμε $X \sim t_\nu(\mu, \Sigma)$, αν έχει συνάρτηση πυκνότητας πιθανότητας:

$$f_X(x \mid \mu, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+k}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{(\nu\pi)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu}(x - \mu)^{\mathrm{T}} \Sigma^{-1} (x - \mu) \right]^{-\frac{\nu+k}{2}}, \quad x \in \mathbb{R}^k.$$

Λήμμα 1.5. Έστω $x, y \in \mathbb{R}^n$ και $A \in \mathbb{R}^{n \times n}$ αντιστρέψιμος πίνακας. Τότε,

$$\left| A + xy^{\mathrm{T}} \right| = |A| \left(1 + y^{\mathrm{T}} A^{-1} x \right).$$

Τέλος, ορίζουμε:

$$d_n = d + n - k + 1, \quad A_n = A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}} - c_n a_n a_n^{\mathrm{T}}.$$

Τότε, υπολογίζουμε την περιθώρια posterior κατανομή του μ ως εξής:

$$\begin{aligned}
\pi(\mu \mid y) &= \int \pi(\mu, \Omega \mid y) d\Omega \\
&\propto \int |\Omega|^{\frac{d+n+1-k-1}{2}} \exp\left\{ -\frac{1}{2} \mathrm{tr} \left[\left(A + c(\mu - a)(\mu - a)^{\mathrm{T}} + \sum_{i=1}^{n} (y_i - \mu)(y_i - \mu)^{\mathrm{T}} \right) \Omega \right] \right\} d\Omega \\
&\propto \left| A + c(\mu - a)(\mu - a)^{\mathrm{T}} + \sum_{i=1}^{n} (y_i - \mu)(y_i - \mu)^{\mathrm{T}} \right|^{-\frac{d+n+1}{2}}
\end{aligned}$$

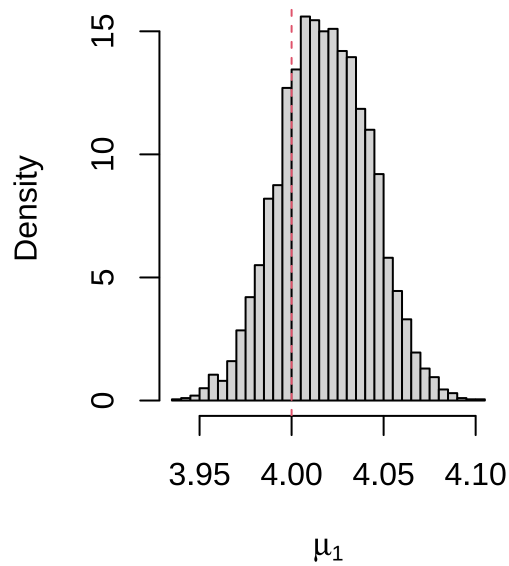
$$
\begin{aligned}
&= \left| A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}} + (c+n)\mu\mu^{\mathrm{T}} - (ca + n\overline{y})\,\mu^{\mathrm{T}} - \mu\,(ca + n\overline{y})^{\mathrm{T}} \right|^{-\frac{d+n-k+1+k}{2}} \\
&= \left| A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}} - c_n a_n a_n^{\mathrm{T}} + c_n\,(\mu - a_n)\,(\mu - a_n)^{\mathrm{T}} \right|^{-\frac{d_n+k}{2}} \\
&\propto \left[1 + c_n\,(\mu - a_n)^{\mathrm{T}} A_n^{-1}\,(\mu - a_n)\right]^{-\frac{d_n+k}{2}} \\
&= \left[1 + \frac{1}{d_n}\,(\mu - a_n)^{\mathrm{T}} c_n d_n A_n^{-1}\,(\mu - a_n)\right]^{-\frac{d_n+k}{2}}.
\end{aligned}
$$

Δηλαδή,

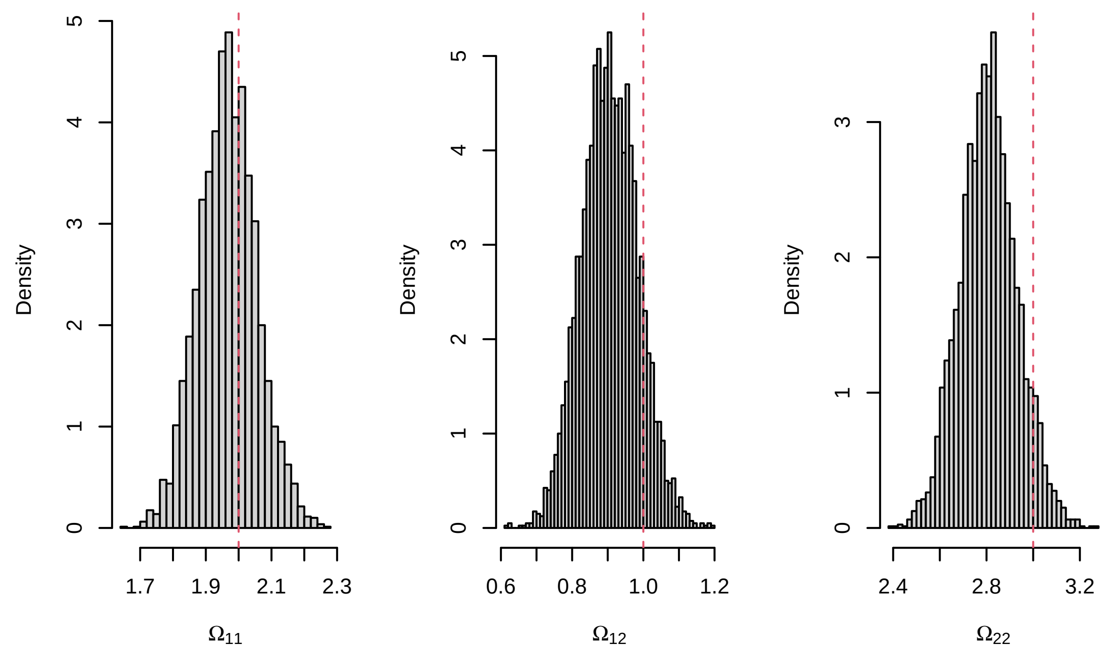
$$
\mu \mid y \sim t_{d_n}\left(a_n, \frac{1}{c_n d_n} A_n\right).
$$

Πρώτα θα υλοποιήσουμε έναν δειγματολήπτη Gibbs που προσομοιώνει εναλλάξ από τις δεσμευμένες posterior κατανομές των παραμέτρων μ και Ω.

```r
MCMCmvnorm = function(Y, mu0, Omega0, a, c, A, d, niter, nburn) {
    library(MASS)
    n = dim(Y)[1]
    k = dim(Y)[2]
    S = colSums(Y)
    cn = c + n
    an = (c * a + S)/cn
    mu = matrix(0, niter, k)
    Omega = array(0, c(k, k, niter))
    mu[1, ] = mu0
    Omega[, , 1] = Omega0
    for (i in 2:niter) {
        mu[i, ] = mvrnorm(1, an, solve(Omega[, , i - 1])/cn)
        Omega[, , i] = rWishart(1, d + n + 1, solve(A + c * tcrossprod(mu[i,
            ] - a) + tcrossprod(t(Y) - mu[i, ])))
    }
    return(list(mu = mu[-(1:nburn), ], Omega = Omega[, , -(1:nburn)]))
}

posterior = MCMCmvnorm(Y, numeric(k), diag(k), numeric(k), 0, matrix(0, k, k),
    2 - k, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$mu[, 1], "FD", freq = FALSE, main = NA, xlab = expression(mu[1]))
abline(v = mu[1], col = 2, lty = 2)
hist(posterior$mu[, 2], "FD", freq = FALSE, main = NA, xlab = expression(mu[2]))
abline(v = mu[2], col = 2, lty = 2)
```



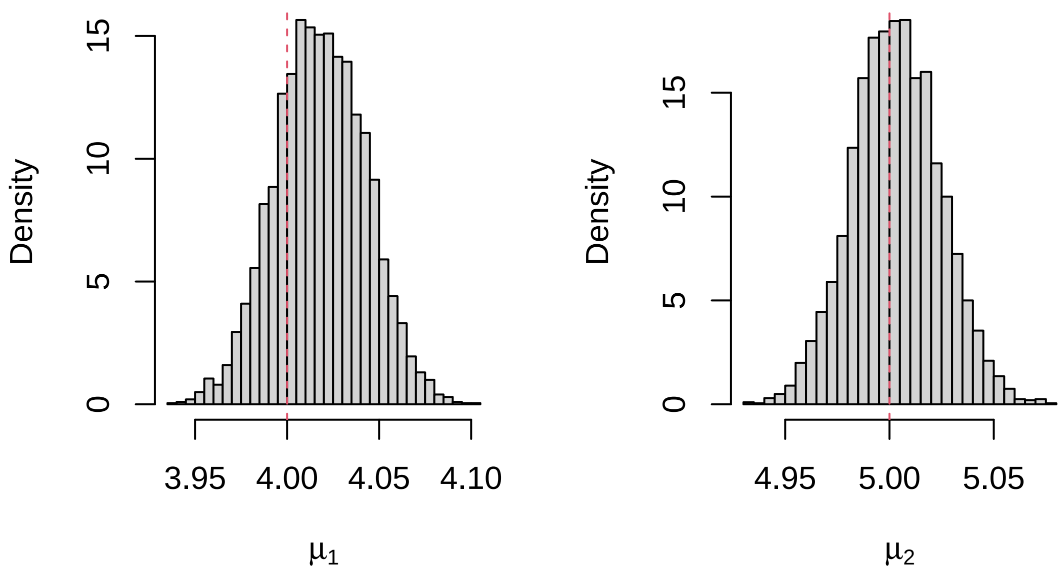

```
par(mfrow = c(1, 3))
hist(posterior$Omega[1, 1, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[11]))
abline(v = Omega[1, 1], col = 2, lty = 2)
hist(posterior$Omega[1, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[12]))
abline(v = Omega[1, 2], col = 2, lty = 2)
hist(posterior$Omega[2, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[22]))
abline(v = Omega[2, 2], col = 2, lty = 2)
```



Έπειτα, θα υλοποιήσουμε τη μέθοδο της σύνθεσης που προσομοιώνει πρώτα από την περιθώρια posterior κατανομή της παραμέτρου Ω και στη συνέχεια από τη δεσμευμένη posterior κατανομή της παραμέτρου μ.

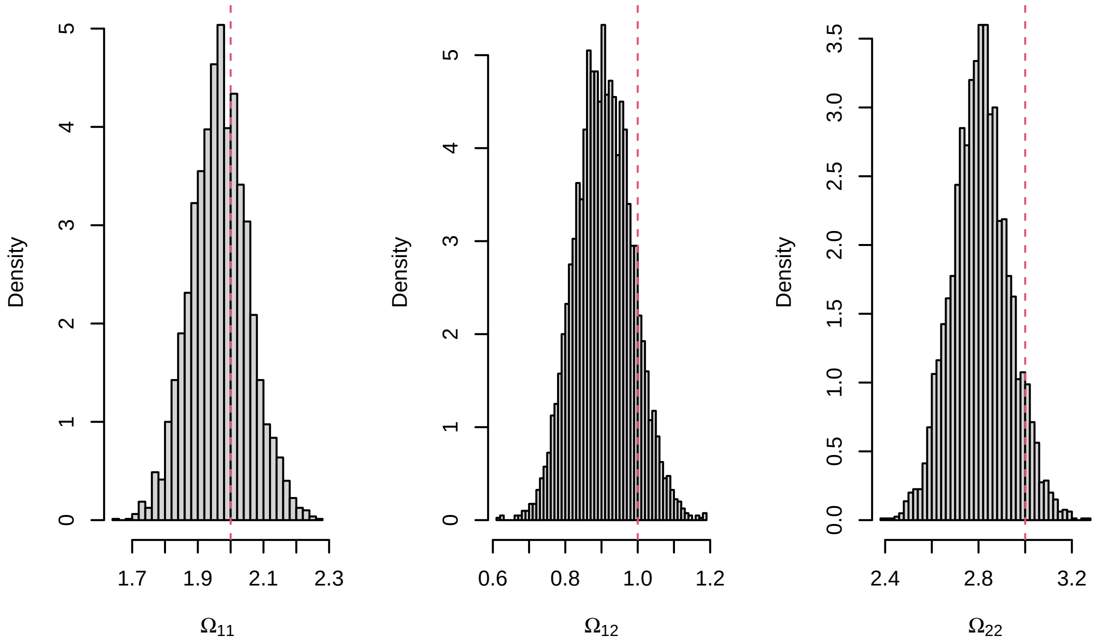
```
CMmvnorm = function(Y, mu0, Omega0, a, c, A, d, niter, nburn) {
    library(MASS)
    n = dim(Y)[1]
```

```
    k = dim(Y)[2]
    S = colSums(Y)
    cn = c + n
    an = (c * a + S)/cn
    An = solve(A + c * tcrossprod(a) + crossprod(Y) - cn * tcrossprod(an))
    mu = matrix(0, niter, k)
    Omega = array(0, c(k, k, niter))
    mu[1, ] = mu0
    Omega[, , 1] = Omega0
    for (i in 2:niter) {
        Omega[, , i] = rWishart(1, d + n, An)
        mu[i, ] = mvrnorm(1, an, solve(Omega[, , i])/cn)
    }
    return(list(mu = mu[-(1:nburn), ], Omega = Omega[, , -(1:nburn)]))
}

posterior = CMmvnorm(Y, numeric(k), diag(k), numeric(k), 0, matrix(0, k, k),
    2 - k, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$mu[, 1], "FD", freq = FALSE, main = NA, xlab = expression(mu[1]))
abline(v = mu[1], col = 2, lty = 2)
hist(posterior$mu[, 2], "FD", freq = FALSE, main = NA, xlab = expression(mu[2]))
abline(v = mu[2], col = 2, lty = 2)
```




```
par(mfrow = c(1, 3))
hist(posterior$Omega[1, 1, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[11]))
abline(v = Omega[1, 1], col = 2, lty = 2)
hist(posterior$Omega[1, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[12]))
abline(v = Omega[1, 2], col = 2, lty = 2)
```

```r
hist(posterior$Omega[2, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[22]))
abline(v = Omega[2, 2], col = 2, lty = 2)
```



1.6 Πολυδιάστατη Κατανομή Student's t

Έστω τυχαίο δείγμα $y_1, \dots, y_n$ από την πολυδιάστατη κατανομή Student's t με μέσο $\mu \in \mathbb{R}^k$, θετικά ορισμένο πίνακα ακρίβειας $\Omega \in \mathbb{R}^{k\times k}$ και $\nu > 0$ βαθμούς ελευθερίας, δηλαδή:

$$f(y_i \mid \mu, \Omega, \nu) = \frac{\Gamma\left(\frac{\nu+k}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} (\nu\pi)^{-\frac{k}{2}} |\Omega|^{\frac{1}{2}} \left[1 + \frac{1}{\nu}(y_i - \mu)^{\mathrm{T}}\Omega(y_i - \mu)\right]^{-\frac{\nu+k}{2}}, \quad y_i \in \mathbb{R}^k.$$

Θεωρούμε τις τυχαίες μεταβλητές $W_i \sim \mathcal{N}_k\left(0, \Omega^{-1}\right)$ και $V_i \sim \chi^2_\nu \equiv \mathrm{Gamma}\left(\frac{\nu}{2}, \frac{1}{2}\right)$. Τότε, παρατηρούμε ότι:

$$Y_i \overset{d}{=} \frac{W_i}{\sqrt{\frac{V_i}{\nu}}} + \mu.$$

Θέτουμε $Z_i = \frac{V_i}{\nu}$. Τότε, $Z_i \sim \mathrm{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$. Παρατηρούμε ότι:

$$Y_i \mid z_i \overset{d}{=} \frac{W_i}{\sqrt{z_i}} + \mu \sim \mathcal{N}_k\left(\mu, z_i^{-1}\Omega^{-1}\right).$$

Θεωρούμε γνωστούς τους βαθμούς ελευθερίας ν και prior ανεξαρτησία για τις παραμέτρους μ, Ω με κατανομές $\mu \sim \mathcal{N}_k\left(a, C^{-1}\right)$ και $\Omega \sim \mathcal{W}_k\left(A^{-1}, d\right)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των παραμέτρων μ, Ω και των λανθανουσών μεταβλητών z_i.

Λύση.

Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$
\begin{aligned}
\pi(\mu, \Omega) &= \pi(\mu) \cdot \pi(\Omega) \\
&= (2\pi)^{-\frac{k}{2}} \left|C^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{(\mu-a)^{\mathrm{T}} C (\mu - a)}{2}\right\} \cdot \frac{|A|^{\frac{d}{2}}}{2^{\frac{dk}{2}} \Gamma_k\left(\frac{d}{2}\right)} |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(A\Omega)} \\
&\propto \exp\left\{-\frac{\mu^{\mathrm{T}} C \mu - 2\mu^{\mathrm{T}} C a + a^{\mathrm{T}} C a}{2}\right\} \cdot |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(A\Omega)} \\
&\propto \exp\left\{-\frac{\mu^{\mathrm{T}} C \mu}{2} + \mu^{\mathrm{T}} C a\right\} \cdot |\Omega|^{\frac{d-k-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(A\Omega)}.
\end{aligned}
$$

Ορίζουμε:

$$
\overline{zy} = \frac{1}{n} \sum_{i=1}^{n} z_i y_i.
$$

Η πλήρης πιθανοφάνεια, δηλαδή η από κοινού πιθανοφάνεια των παρατηρούμενων μεταβλητών y_i και των λανθανουσών μεταβλητών z_i, δίνεται από τη σχέση:

$$
\begin{aligned}
f(y, z \mid \mu, \Omega) &= \prod_{i=1}^{n} f(y_i, z_i \mid \mu, \Omega) \\
&= \prod_{i=1}^{n} f(z_i) f(y_i \mid z_i, \mu, \Omega) \\
&= \prod_{i=1}^{n} \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} z_i^{\frac{\nu}{2}-1} e^{-\frac{\nu}{2} z_i} (2\pi)^{-\frac{k}{2}} |z_i^{-1} \Omega^{-1}|^{-\frac{1}{2}} \exp\left\{-\frac{z_i (y_i - \mu)^{\mathrm{T}} \Omega (y_i - \mu)}{2}\right\} \\
&\propto |\Omega|^{\frac{n}{2}} \cdot \prod_{i=1}^{n} z_i^{\frac{\nu+k}{2}-1} \exp\left\{-\sum_{i=1}^{n} \frac{\nu + (y_i - \mu)^{\mathrm{T}} \Omega (y_i - \mu)}{2} z_i\right\} \\
&= |\Omega|^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left[\sum_{i=1}^{n} z_i (y_i - \mu)(y_i - \mu)^{\mathrm{T}} \Omega\right]\right\} \cdot \prod_{i=1}^{n} z_i^{\frac{\nu+k}{2}-1} \exp\left\{-\frac{\nu}{2} \sum_{i=1}^{n} z_i\right\} \\
&= |\Omega|^{\frac{n}{2}} \exp\left\{-\sum_{i=1}^{n} \frac{y_i^{\mathrm{T}} \Omega y_i - 2\mu^{\mathrm{T}} \Omega y_i + \mu^{\mathrm{T}} \Omega \mu}{2} z_i\right\} \cdot \prod_{i=1}^{n} z_i^{\frac{\nu+k}{2}-1} \exp\left\{-\frac{\nu}{2} \sum_{i=1}^{n} z_i\right\} \\
&= \exp\left\{-\frac{\mu^{\mathrm{T}} n \overline{z} \Omega \mu}{2} + \mu^{\mathrm{T}} n \Omega \overline{zy}\right\} \cdot |\Omega|^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\sum_{i=1}^{n} z_i y_i y_i^{\mathrm{T}} \Omega\right)\right\} \\
&\quad \times \prod_{i=1}^{n} z_i^{\frac{\nu+k}{2}-1} \exp\left\{-\frac{\nu}{2} \sum_{i=1}^{n} z_i\right\}.
\end{aligned}
$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των παραμέτρων μ και Ω ως εξής:

$$
\begin{aligned}
\pi(\mu \mid \Omega, z, y) &\propto \pi(\mu, \Omega, z \mid y) \\
&\propto \pi(\mu, \Omega) \cdot f(y, z \mid \mu, \Omega) \\
&\propto \exp\left\{-\frac{\mu^{\mathrm{T}} C \mu}{2} + \mu^{\mathrm{T}} C a\right\} \cdot \exp\left\{-\frac{\mu^{\mathrm{T}} n \overline{z} \Omega \mu}{2} + \mu^{\mathrm{T}} n \Omega \overline{zy}\right\}
\end{aligned}
$$

$$= \exp\left\{-\frac{1}{2}\mu^{\mathrm{T}}\underbrace{\left(C+n\bar{z}\Omega\right)}_{C_n}\mu+\mu^{\mathrm{T}}\left(Ca+n\Omega\overline{zy}\right)\right\}$$

$$= \exp\left\{-\frac{\mu^{\mathrm{T}}C_n\mu}{2}+\mu^{\mathrm{T}}\underbrace{\left(C+n\bar{z}\Omega\right)}_{C_n}\underbrace{\left(C+n\bar{z}\Omega\right)^{-1}\left(Ca+n\Omega\overline{zy}\right)}_{a_n}\right\},$$

$$\begin{aligned}\pi(\Omega \mid \mu, z, y) &\propto \pi(\mu,\Omega)\cdot f(y,z\mid \mu,\Omega)\\ &\propto |\Omega|^{\frac{d-k-1}{2}}e^{-\frac{1}{2}\mathrm{tr}(A\Omega)}\cdot|\Omega|^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\mathrm{tr}\left[\sum_{i=1}^{n}z_i(y_i-\mu)(y_i-\mu)^{\mathrm{T}}\Omega\right]\right\}\\ &= |\Omega|^{\frac{d+n-k-1}{2}}\exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A+\sum_{i=1}^{n}z_i(y_i-\mu)(y_i-\mu)^{\mathrm{T}}\right)\Omega\right]\right\}.\end{aligned}$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή των λανθανουσών μεταβλητών z_i ως εξής:

$$f(z_i \mid y_i,\mu,\Omega)\propto f(y_i,z_i\mid \mu,\Omega)\propto z_i^{\frac{\nu+k}{2}-1}\exp\left\{-\frac{\nu+(y_i-\mu)^{\mathrm{T}}\Omega(y_i-\mu)}{2}z_i\right\}.$$

Δηλαδή,

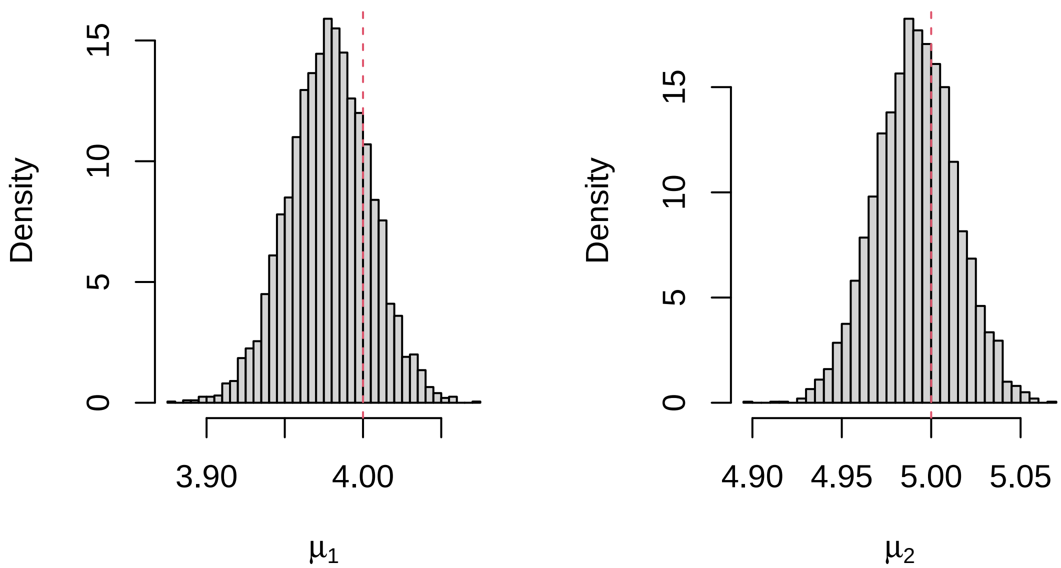
$$\mu \mid \Omega, z, y \sim \mathcal{N}_k\left(\left(C+n\bar{z}\Omega\right)^{-1}\left(Ca+n\Omega\overline{zy}\right),\left(C+n\bar{z}\Omega\right)^{-1}\right),$$

$$\Omega \mid \mu, z, y \sim \mathcal{W}_k\left(\left(A+\sum_{i=1}^{n}z_i(y_i-\mu)(y_i-\mu)^{\mathrm{T}}\right)^{-1},d+n\right),$$

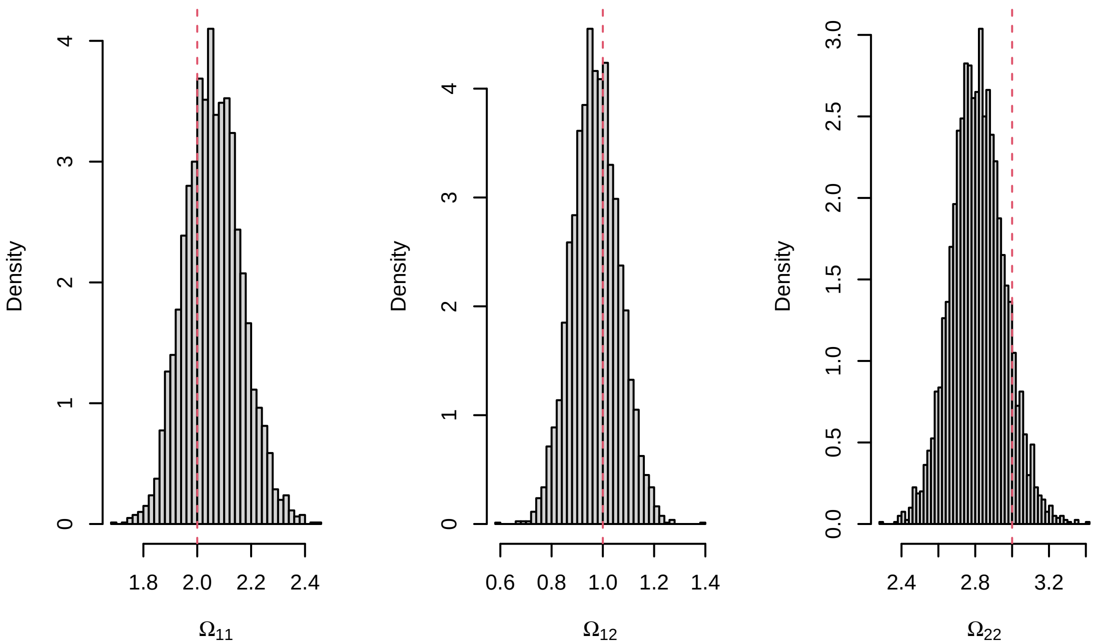
$$z_i \mid y_i,\mu,\Omega \sim \mathrm{Gamma}\left(\frac{\nu+k}{2},\frac{\nu+(y_i-\mu)^{\mathrm{T}}\Omega(y_i-\mu)}{2}\right).$$

```
MCMCmvt = function(Y, mu0, Omega0, nu, a, C, A, d, niter, nburn) {
    library(MASS)
    n = dim(Y)[1]
    k = dim(Y)[2]
    mu = matrix(0, niter, k)
    Omega = array(0, c(k, k, niter))
    Z = matrix(0, niter, n)
    mu[1, ] = mu0
    Omega[, , 1] = Omega0
    Z[1, ] = rgamma(n, (nu + k)/2, (nu + colSums((t(Y) - mu[1, ]) * Omega[,
        , 1] %*% (t(Y) - mu[1, ])))/2)
    for (i in 2:niter) {
        Cn = C + sum(Z[i - 1, ]) * Omega[, , i - 1]
        an = solve(Cn, C %*% a + Omega[, , i - 1] %*% colSums(Z[i - 1, ] * Y))
        mu[i, ] = mvrnorm(1, an, solve(Cn))
        Omega[, , i] = rWishart(1, d + n, solve(A + crossprod(sqrt(Z[i - 1,
            ]) * t(t(Y) - mu[i, ]))))
        Z[i, ] = rgamma(n, (nu + k)/2, (nu + colSums((t(Y) - mu[i, ]) * Omega[,
            , i] %*% (t(Y) - mu[i, ])))/2)
```

```
    }
    return(list(mu = mu[-(1:nburn), ], Omega = Omega[, , -(1:nburn)]))
}

library(mvtnorm)
n = 1000
k = 2
mu = c(4, 5)
Omega = matrix(c(2, 1, 1, 3), k)
nu = 10
Y = rmvt(n, solve(Omega), nu, mu)
posterior = MCMCmvt(Y, numeric(k), diag(k), nu, numeric(k), matrix(0, k, k),
    matrix(0, k, k), 2 - k, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$mu[, 1], "FD", freq = FALSE, main = NA, xlab = expression(mu[1]))
abline(v = mu[1], col = 2, lty = 2)
hist(posterior$mu[, 2], "FD", freq = FALSE, main = NA, xlab = expression(mu[2]))
abline(v = mu[2], col = 2, lty = 2)
```




```
par(mfrow = c(1, 3))
hist(posterior$Omega[1, 1, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[11]))
abline(v = Omega[1, 1], col = 2, lty = 2)
hist(posterior$Omega[1, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[12]))
abline(v = Omega[1, 2], col = 2, lty = 2)
hist(posterior$Omega[2, 2, ], "FD", freq = FALSE, main = NA, xlab = expression(Omega[22]))
abline(v = Omega[2, 2], col = 2, lty = 2)
```

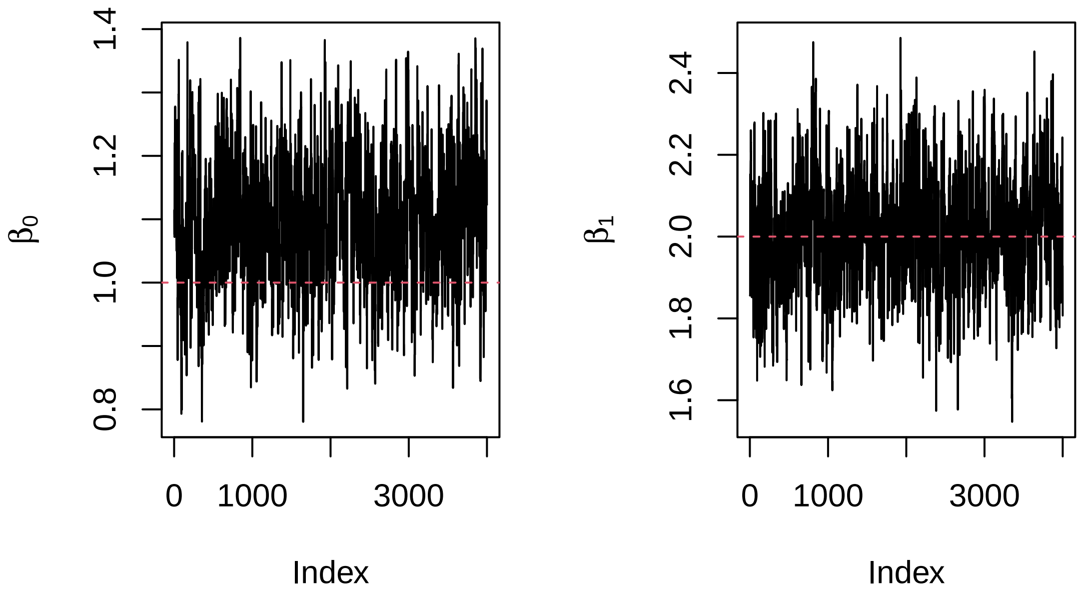

Density
Ω11
Density
Ω12
Density
Ω22

2 Γενικευμένα Γραμμικά Μοντέλα

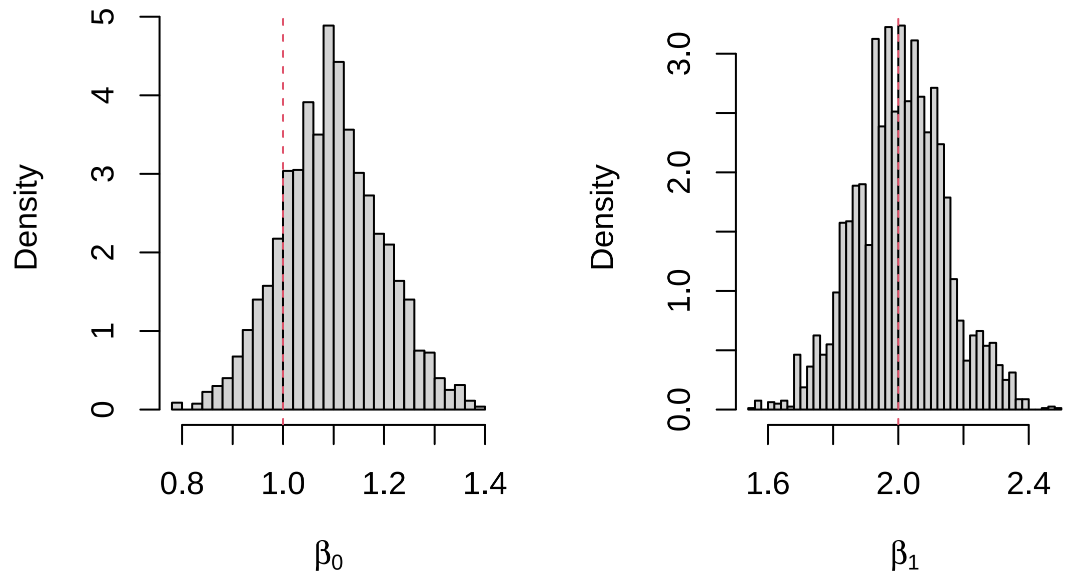
2.1 Λογιστικό Μοντέλο

Θεωρούμε το λογιστικό μοντέλο παλινδρόμησης $y_i \sim \text{Bernoulli}(p_i)$, όπου $\operatorname{logit} p_i = \log \frac{p_i}{1-p_i} = \beta_0 + \beta_1 x_i$, δηλαδή $p_i = \frac{1}{1+e^{-\beta_0-\beta_1 x_i}}$. Θεωρούμε prior ανεξαρτησία με μη-ορθές κατανομές $\pi(\beta_0) \propto 1$ και $\pi(\beta_1) \propto 1$. Μπορούμε να υλοποιήσουμε έναν αλγόριθμο Random Walk Metropolis-Hastings με προτείνουσες τυχαίες μεταβλητές $\beta_0^\star \mid \beta_0^{(\ell-1)} \sim \mathcal{N}\left(\beta_0^{(\ell-1)}, \sigma_0^2\right)$ και $\beta_1^\star \mid \beta_1^{(\ell-1)} \sim \mathcal{N}\left(\beta_1^{(\ell-1)}, \sigma_1^2\right)$.

```r
RWMHlogistic = function(Y, X, beta00, beta10, beta0sd, beta1sd, niter, nburn) {
    beta0 = numeric(niter)
    beta1 = numeric(niter)
    beta0[1] = beta00
    beta1[1] = beta10
    for (i in 2:niter) {
        beta0star = rnorm(1, beta0[i - 1], beta0sd)
        logA = sum(dbinom(Y, 1, (1 + exp(-beta0star - beta1[i - 1] * X))^(-1),
            log = TRUE) - dbinom(Y, 1, (1 + exp(-beta0[i - 1] - beta1[i - 1] *
            X))^(-1), log = TRUE))
        beta0[i] = ifelse(log(runif(1)) < logA, beta0star, beta0[i - 1])
        beta1star = rnorm(1, beta1[i - 1], beta1sd)
        logA = sum(dbinom(Y, 1, (1 + exp(-beta0[i] - beta1star * X))^(-1), log = TRUE) -
            dbinom(Y, 1, (1 + exp(-beta0[i] - beta1[i - 1] * X))^(-1), log = TRUE))
        beta1[i] = ifelse(log(runif(1)) < logA, beta1star, beta1[i - 1])
    }
    return(list(beta0 = beta0[-(1:nburn)], beta1 = beta1[-(1:nburn)]))
}

n = 1000
beta0 = 1
beta1 = 2
X = rnorm(n)
Y = rbinom(n, 1, (1 + exp(-beta0 - beta1 * X))^(-1))
posterior = RWMHlogistic(Y, X, 0, 0, 0.15, 0.25, 5000, 1000)
par(mfrow = c(1, 2))
plot(posterior$beta0, type = "l", ylab = expression(beta[0]))
abline(h = beta0, col = 2, lty = 2)
plot(posterior$beta1, type = "l", ylab = expression(beta[1]))
abline(h = beta1, col = 2, lty = 2)
```




```r
hist(posterior$beta0, "FD", freq = FALSE, main = NA, xlab = expression(beta[0]))
abline(v = beta0, col = 2, lty = 2)
hist(posterior$beta1, "FD", freq = FALSE, main = NA, xlab = expression(beta[1]))
abline(v = beta1, col = 2, lty = 2)
```



2.2 Μοντέλο Πιθανομονάδας

Θεωρούμε το μοντέλο πιθανομονάδας $y_i \sim \text{Bernoulli}(p_i)$, όπου $p_i = \Phi\left(x_i^{\mathrm{T}}\beta\right)$ και $\beta \in \mathbb{R}^k$. Θεωρούμε τις ανεξάρτητες τυχαίες μεταβλητές $z_i = x_i^{\mathrm{T}}\beta + \varepsilon_i$, όπου $\varepsilon_i \sim \mathcal{N}(0,1)$. Τότε, παρατηρούμε ότι:

$$p_i = \mathbb{P}(Y_i = 1) = \Phi\left(x_i^{\mathrm{T}}\beta\right) = \mathbb{P}(\varepsilon_i \leqslant x_i^{\mathrm{T}}\beta) = \mathbb{P}\left(\varepsilon_i \geqslant -x_i^{\mathrm{T}}\beta\right) = \mathbb{P}(Z_i \geqslant 0).$$

Δηλαδή, $y_i \mid z_i \overset{d}{=} \mathbb{1}_{\{z_i \geqslant 0\}}$. Θεωρούμε την prior κατανομή $\beta \sim \mathcal{N}_k\left(a, C^{-1}\right)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές της παραμέτρου β και των λανθανουσών μεταβλητών z_i.

Λύση.

Η prior κατανομή μπορεί να γραφτεί ως:

$$
\begin{aligned}
\pi(\beta) &= (2\pi)^{-\frac{k}{2}} \left|C^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{(\beta-a)^{\mathrm{T}}C(\beta-a)}{2}\right\} \\
&\propto \exp\left\{-\frac{\beta^{\mathrm{T}}C\beta - 2\beta^{\mathrm{T}}Ca + a^{\mathrm{T}}Ca}{2}\right\} \\
&\propto \exp\left\{-\frac{\beta^{\mathrm{T}}C\beta}{2} + \beta^{\mathrm{T}}Ca\right\}.
\end{aligned}
$$

Η πλήρης πιθανοφάνεια, δηλαδή η από κοινού πιθανοφάνεια των παρατηρούμενων μεταβλητών y_i και των λανθανουσών μεταβλητών z_i, δίνεται από τη σχέση:

$$
\begin{aligned}
f(y, z \mid \beta) &= \prod_{i=1}^{n} f(y_i, z_i \mid \beta) \\
&= \prod_{i=1}^{n} f(z_i \mid \beta) f(y_i \mid z_i) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\left(z_i - x_i^{\mathrm{T}}\beta\right)^2}{2}\right\} \mathbb{1}_{\{z_i \geqslant 0\}}^{y_i} \mathbb{1}_{\{z_i < 0\}}^{1-y_i}.
\end{aligned}
$$

Ορίζουμε $z = (z_1, \dots, z_n)^{\mathrm{T}} \in \mathbb{R}^n$ και τον πίνακα σχεδιασμού $X = (x_1, \dots, x_n)^{\mathrm{T}} \in \mathbb{R}^{n\times k}$. Τότε, παρατηρούμε ότι $z \sim N_n\left(X\beta, I_n\right)$. Δηλαδή, η πλήρης πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$
\begin{aligned}
f(y, z \mid \beta) &= f(z \mid \beta) \cdot f(y \mid z) \\
&= f(z \mid \beta) \cdot \prod_{i=1}^{n} f(y_i \mid z_i) \\
&= (2\pi)^{-\frac{n}{2}} \left|I_n\right|^{-\frac{1}{2}} \exp\left\{-\frac{(z-X\beta)^{\mathrm{T}}(z-X\beta)}{2}\right\} \cdot \prod_{i=1}^{n} \mathbb{1}_{\{z_i \geqslant 0\}}^{y_i} \mathbb{1}_{\{z_i < 0\}}^{1-y_i} \\
&\propto \exp\left\{-\frac{z^{\mathrm{T}}z - 2\beta^{\mathrm{T}}X^{\mathrm{T}}z + \beta^{\mathrm{T}}X^{\mathrm{T}}X\beta}{2}\right\} \cdot \prod_{i=1}^{n} \mathbb{1}_{\{z_i \geqslant 0\}}^{y_i} \mathbb{1}_{\{z_i < 0\}}^{1-y_i} \\
&\propto \exp\left\{-\frac{\beta^{\mathrm{T}}X^{\mathrm{T}}X\beta}{2} + \beta^{\mathrm{T}}X^{\mathrm{T}}z\right\} \cdot \prod_{i=1}^{n} \mathbb{1}_{\{z_i \geqslant 0\}}^{y_i} \mathbb{1}_{\{z_i < 0\}}^{1-y_i}.
\end{aligned}
$$

Επομένως, παίρνουμε τη δεσμευμένη posterior κατανομή του β ως εξής:

$$
\begin{aligned}
\pi(\beta \mid z, y) &\propto \pi(\beta, z \mid y) \\
&\propto \pi(\beta) \cdot f(y, z \mid \beta) \\
&\propto \exp\left\{-\frac{\beta^{\mathrm{T}}C\beta}{2} + \beta^{\mathrm{T}}Ca\right\} \cdot \exp\left\{-\frac{\beta^{\mathrm{T}}X^{\mathrm{T}}X\beta}{2} + \beta^{\mathrm{T}}X^{\mathrm{T}}z\right\} \\
&= \exp\left\{-\frac{1}{2}\beta^{\mathrm{T}} \underbrace{\left(C + X^{\mathrm{T}}X\right)}_{C_n} \beta + \beta^{\mathrm{T}}\left(Ca + X^{\mathrm{T}}z\right)\right\}
\end{aligned}
$$

$$= \exp\left\{-\frac{\beta^{\mathrm{T}} C_n \beta}{2} + \beta^{\mathrm{T}} \underbrace{\left(C + X^{\mathrm{T}} X\right)}_{C_n} \underbrace{\left(C + X^{\mathrm{T}} X\right)^{-1} \left(Ca + X^{\mathrm{T}} z\right)}_{a_n}\right\}.$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή των λανθανουσών μεταβλητών z_i ως εξής:

$$f(z_i \mid y_i, \beta) \propto f(y_i, z_i \mid \beta) \propto \exp\left\{-\frac{\left(z_i - x_i^{\mathrm{T}}\beta\right)^2}{2}\right\} \mathbb{1}_{\{z_i \geqslant 0\}}^{y_i} \mathbb{1}_{\{z_i < 0\}}^{1-y_i}.$$

Δηλαδή,

$$\beta \mid z \sim \mathcal{N}_k\left(\left(C + X^{\mathrm{T}} X\right)^{-1} \left(Ca + X^{\mathrm{T}} z\right), \left(C + X^{\mathrm{T}} X\right)^{-1}\right),$$

$$(z_i \mid y_i = 1, \beta) \sim \mathcal{N}\left(x_i^{\mathrm{T}}\beta, 1\right) \mathbb{1}_{\{z_i \geqslant 0\}}, \quad (z_i \mid y_i = 0, \beta) \sim \mathcal{N}\left(x_i^{\mathrm{T}}\beta, 1\right) \mathbb{1}_{\{z_i < 0\}}.$$

Παρατηρούμε ότι:

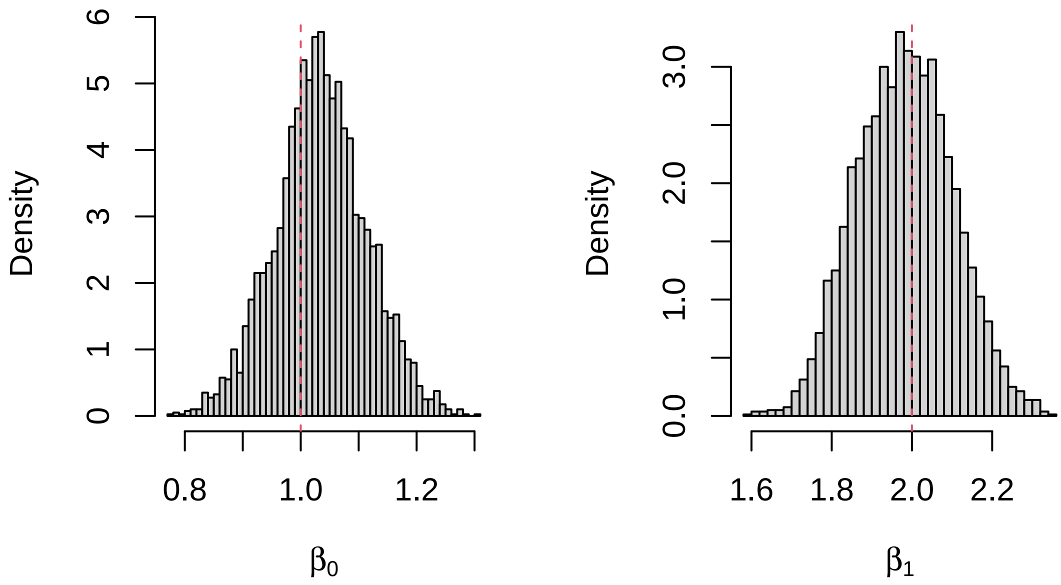
$$F_{Z_i \mid y_i = 1, \beta}(z_i) = \frac{\Phi\left(z_i - x_i^{\mathrm{T}}\beta\right) - \Phi\left(-x_i^{\mathrm{T}}\beta\right)}{1 - \Phi\left(-x_i^{\mathrm{T}}\beta\right)} \mathbb{1}_{\{z_i \geqslant 0\}}, \quad F_{Z_i \mid y_i = 0, \beta}(z_i) = \frac{\Phi\left(z_i - x_i^{\mathrm{T}}\beta\right)}{\Phi\left(-x_i^{\mathrm{T}}\beta\right)} \mathbb{1}_{\{z_i < 0\}}.$$

Αν $U_1, U_2, \ldots, U_n \sim \mathrm{Unif}[0,1]$, τότε παίρνουμε ότι:

$$Z_i = \begin{cases} \Phi^{-1}\left[\Phi\left(x_i^{\mathrm{T}}\beta\right) U_i + 1 - \Phi\left(x_i^{\mathrm{T}}\beta\right)\right] + x_i^{\mathrm{T}}\beta, & y_i = 1 \\ \Phi^{-1}\left[\left(1 - \Phi\left(x_i^{\mathrm{T}}\beta\right)\right) U_i\right] + x_i^{\mathrm{T}}\beta, & y_i = 0 \end{cases}.$$

```
MCMCprobit = function(Y, X, beta0, a, C, niter, nburn) {
    library(MASS)
    n = length(Y)
    k = dim(X)[2]
    beta = matrix(0, niter, k)
    Z = matrix(0, niter, n)
    beta[1, ] = beta0
    prob = pnorm(X %*% beta[1, ])
    U = runif(n)
    Z[1, ] = X %*% beta[1, ] + ifelse(Y == 1, qnorm(prob * U + 1 - prob), qnorm((1 -
        prob) * U))
    for (i in 2:niter) {
        Cninv = solve(C + crossprod(X))
        an = crossprod(Cninv, C %*% a + crossprod(X, Z[i - 1, ]))
        beta[i, ] = mvrnorm(1, an, Cninv)
        prob = pnorm(X %*% beta[i, ])
        U = runif(n)
        Z[i, ] = X %*% beta[i, ] + ifelse(Y == 1, qnorm(prob * U + 1 - prob),
            qnorm((1 - prob) * U))
    }
    return(list(beta = beta[-(1:nburn), ], Z = Z[-(1:nburn), ]))
}
```

```
n = 1000
k = 2
beta = c(1, 2)
X = cbind(1, rnorm(n))
Y = rbinom(n, 1, pnorm(X %*% beta))
posterior = MCMCprobit(Y, X, numeric(k), numeric(k), matrix(0, k, k), 5000,
    1000)
par(mfrow = c(1, 2))
hist(posterior$beta[, 1], "FD", freq = FALSE, main = NA, xlab = expression(beta[0]))
abline(v = beta[1], col = 2, lty = 2)
hist(posterior$beta[, 2], "FD", freq = FALSE, main = NA, xlab = expression(beta[1]))
abline(v = beta[2], col = 2, lty = 2)
```

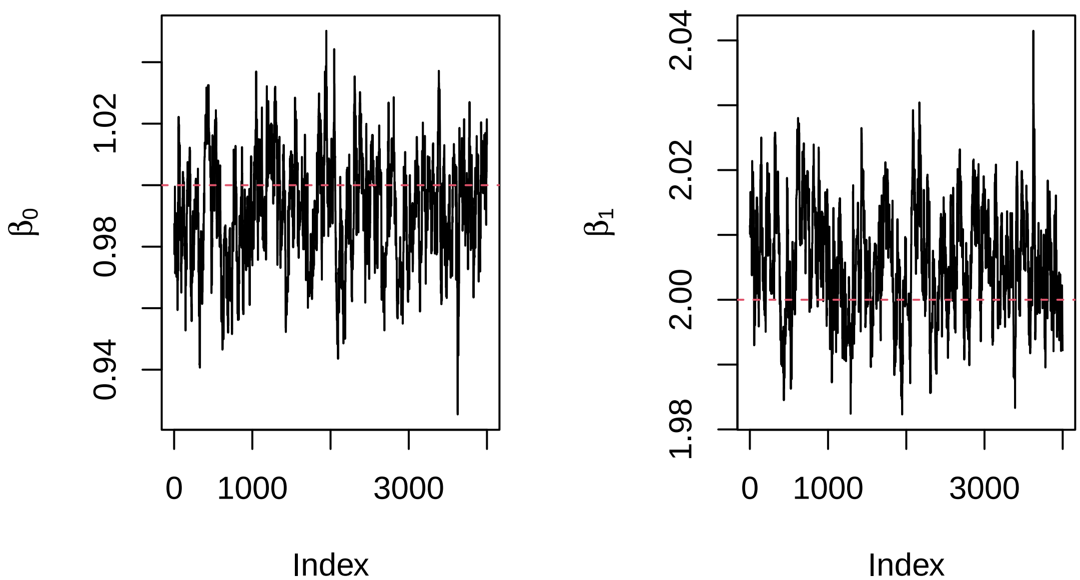


2.3 Λογαριθμικό Μοντέλο Poisson

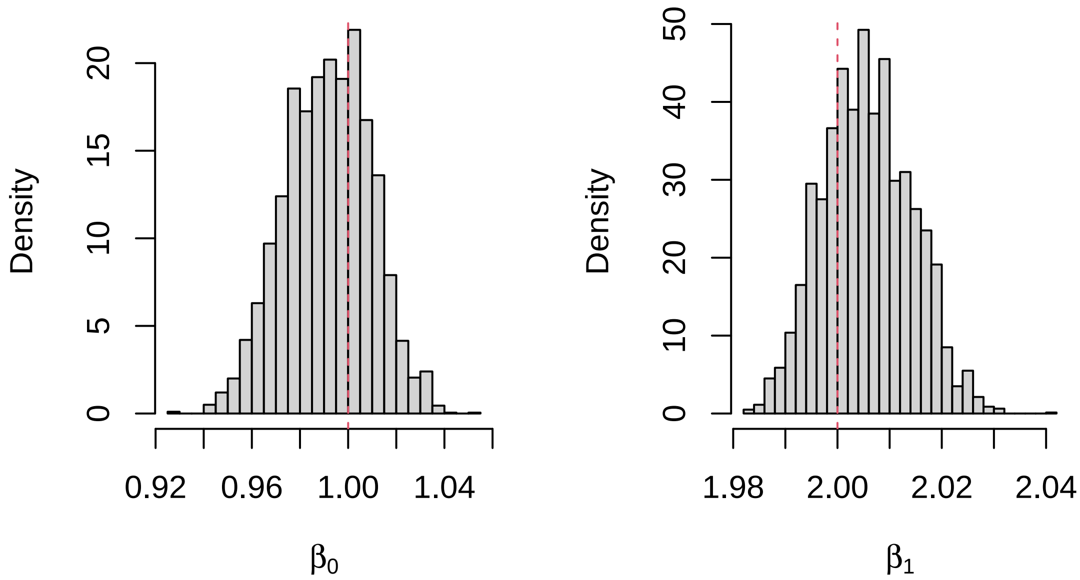
Θεωρούμε το μοντέλο $y_i \sim \text{Poisson}(\lambda_i)$, όπου $\log \lambda_i = \beta_0 + \beta_1 x_i$, δηλαδή $\lambda_i = e^{\beta_0 + \beta_1 x_i}$. Θεωρούμε prior ανεξαρτησία με μη-ορθές κατανομές $\pi(\beta_0) \propto 1$ και $\pi(\beta_1) \propto 1$. Μπορούμε να υλοποιήσουμε έναν αλγόριθμο Random Walk Metropolis-Hastings με προτείνουσες τυχαίες μεταβλητές $\beta_0^\star \mid \beta_0^{(\ell-1)} \sim \mathcal{N}\left(\beta_0^{(\ell-1)}, \sigma_0^2\right)$ και $\beta_1^\star \mid \beta_1^{(\ell-1)} \sim \mathcal{N}\left(\beta_1^{(\ell-1)}, \sigma_1^2\right)$.

```
RWMHpois = function(Y, X, beta00, beta10, beta0sd, beta1sd, niter, nburn) {
    beta0 = numeric(niter)
    beta1 = numeric(niter)
    beta0[1] = beta00
    beta1[1] = beta10
    for (i in 2:niter) {
        beta0star = rnorm(1, beta0[i - 1], beta0sd)
        logA = sum(dpois(Y, exp(beta0star + beta1[i - 1] * X), log = TRUE) -
            dpois(Y, exp(beta0[i - 1] + beta1[i - 1] * X), log = TRUE))
```

```
        beta0[i] = ifelse(log(runif(1)) < logA, beta0star, beta0[i - 1])
        beta1star = rnorm(1, beta1[i - 1], beta1sd)
        logA = sum(dpois(Y, exp(beta0[i] + beta1star * X), log = TRUE) - dpois(Y,
            exp(beta0[i] + beta1[i - 1] * X), log = TRUE))
        beta1[i] = ifelse(log(runif(1)) < logA, beta1star, beta1[i - 1])
    }
    return(list(beta0 = beta0[-(1:nburn)], beta1 = beta1[-(1:nburn)]))
}

n = 1000
beta0 = 1
beta1 = 2
X = rnorm(n)
Y = rpois(n, exp(beta0 + beta1 * X))
posterior = RWMHpois(Y, X, 0, 0, 0.015, 0.0075, 5000, 1000)
par(mfrow = c(1, 2))
plot(posterior$beta0, type = "l", ylab = expression(beta[0]))
abline(h = beta0, col = 2, lty = 2)
plot(posterior$beta1, type = "l", ylab = expression(beta[1]))
abline(h = beta1, col = 2, lty = 2)
```




```
hist(posterior$beta0, "FD", freq = FALSE, main = NA, xlab = expression(beta[0]))
abline(v = beta0, col = 2, lty = 2)
hist(posterior$beta1, "FD", freq = FALSE, main = NA, xlab = expression(beta[1]))
abline(v = beta1, col = 2, lty = 2)
```



2.4 Μηδενικά Διογκωμένο Μοντέλο Poisson

Θεωρούμε το μηδενικά διογκωμένο μοντέλο παλινδρόμησης:

$$\mathbb{P}(Y_i = k) = \begin{cases} 1 - p + pe^{-\lambda}, & k = 0 \\ pe^{-\lambda}\frac{\lambda^k}{k!}, & k = 1, 2, ... \end{cases}$$

Θεωρούμε τις ανεξάρτητες τυχαίες μεταβλητές $z_i \sim \text{Bernoulli}(p)$. Για $k = 0, 1, ...$, παρατηρούμε ότι:

$$\mathbb{P}\left(Y_i = k \mid Z_i = 1\right) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad \mathbb{P}\left(Y_i = 0 \mid Z_i = 0\right) = 1.$$

Δηλαδή, ισχύει ότι $(y_i \mid z_i = 1) \sim \text{Poisson}(\lambda)$ και $(y_i \mid z_i = 0) \overset{d}{=} 0$. Θεωρούμε prior ανεξαρτησία με κατανομές $p \sim \text{Beta}(a, c)$ και $\lambda \sim \text{Gamma}(d, q)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των παραμέτρων p, λ και των λανθανουσών μεταβλητών z_i.

Λύση.

Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

$$\begin{aligned} \pi(p, \lambda) &= \pi(p)\pi(\lambda) \\ &= \frac{\Gamma(a+c)}{\Gamma(a)\Gamma(c)} p^{a-1}(1-p)^{c-1} \cdot \frac{q^d}{\Gamma(d)} \lambda^{d-1} e^{-q\lambda} \\ &\propto p^{a-1}(1-p)^{c-1} \cdot \lambda^{d-1} e^{-q\lambda}. \end{aligned}$$

Η πλήρης πιθανοφάνεια, δηλαδή η από κοινού πιθανοφάνεια των παρατηρούμενων μεταβλητών y_i και των λανθανουσών μεταβλητών z_i, δίνεται από τη σχέση:

$$f(y, z \mid p, \lambda) = \prod_{i=1}^{n} f(y_i, z_i \mid p, \lambda)$$

$$
\begin{aligned}
&= \prod_{i=1}^{n} f(z_i \mid p) f(y_i \mid z_i, \lambda) \\
&= \prod_{i=1}^{n} p^{z_i}(1-p)^{1-z_i} \left(e^{-\lambda} \frac{\lambda^{y_i}}{y_i!}\right)^{z_i} \mathbb{1}_{\{y_i=0\}}^{1-z_i} \\
&\propto \prod_{i=1}^{n} p^{z_i}(1-p)^{1-z_i} e^{-\lambda z_i} \lambda^{y_i} \mathbb{1}_{\{y_i=0\}}^{1-z_i} \\
&\propto p^{n\bar{z}}(1-p)^{n-n\bar{z}} \cdot \lambda^{n\bar{y}} e^{-n\lambda\bar{z}} \cdot \prod_{i=1}^{n} \mathbb{1}_{\{y_i=0\}}^{1-z_i}.
\end{aligned}
$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των p και λ ως εξής:

$$
\begin{aligned}
\pi(p \mid \lambda, z, y) &\propto \pi(p, \lambda, z \mid y) \\
&\propto \pi(p) \cdot f(y, z \mid p, \lambda) \\
&\propto p^{a-1}(1-p)^{c-1} \cdot p^{n\bar{z}}(1-p)^{n-n\bar{z}} \\
&= p^{a+n\bar{z}-1}(1-p)^{c+n-n\bar{z}-1},
\end{aligned}
$$

$$
\begin{aligned}
\pi(\lambda \mid p, z, y) &\propto \pi(\lambda) \cdot f(y, z \mid p, \lambda) \\
&\propto \lambda^{d-1} e^{-q\lambda} \cdot \lambda^{n\bar{y}} e^{-n\lambda\bar{z}} \\
&= \lambda^{d+n\bar{y}-1} e^{-(q+n\bar{z})\lambda}.
\end{aligned}
$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή των λανθανουσών μεταβλητών z_i ως εξής:

$$
f(z_i \mid y_i, p, \lambda) \propto f(y_i, z_i \mid p, \lambda) \propto p^{z_i}(1-p)^{1-z_i} e^{-\lambda z_i} \mathbb{1}_{\{y_i=0\}}^{1-z_i} = \left(pe^{-\lambda}\right)^{z_i} (1-p)^{1-z_i} \mathbb{1}_{\{y_i=0\}}^{1-z_i}.
$$

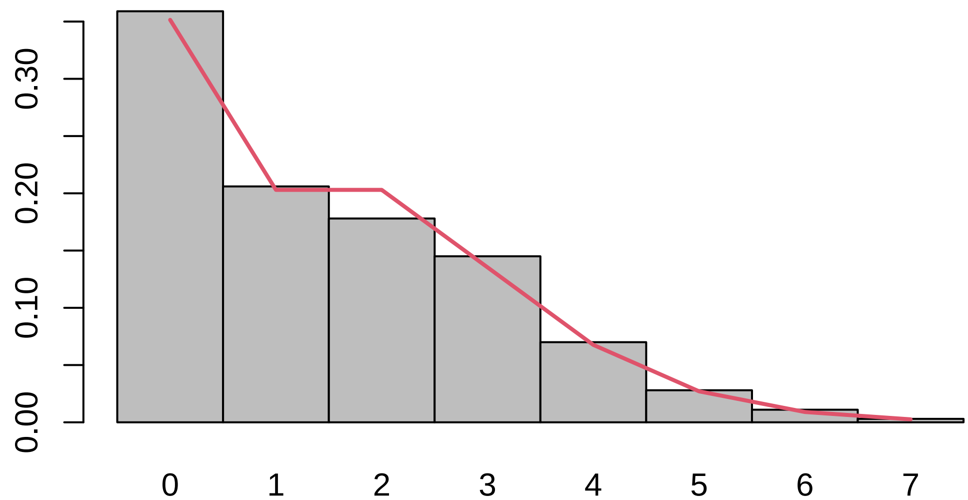
Δηλαδή,

$$
p \mid z, y \sim \text{Beta}\left(a+n\bar{z}, c+n-n\bar{z}\right), \quad \lambda \mid z, y \sim \text{Gamma}\left(d+n\bar{y}, q+n\bar{z}\right),
$$

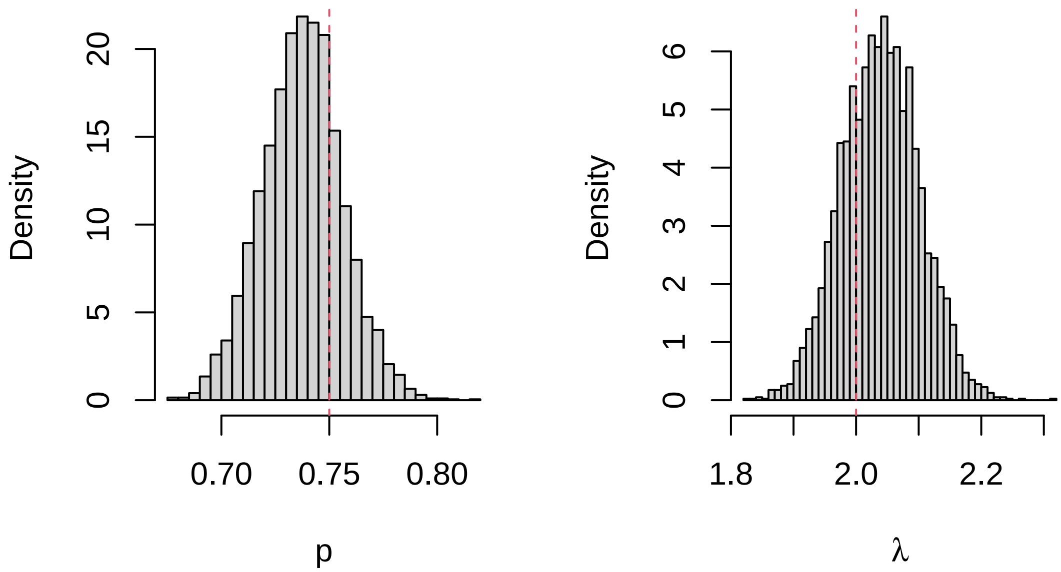
$$
(z_i \mid y_i = 0, p, \lambda) \sim \text{Bernoulli}\left(\frac{pe^{-\lambda}}{pe^{-\lambda}+1-p}\right), \quad (z_i \mid y_i > 0, p, \lambda) \overset{d}{=} 1.
$$

```
MCMCzip = function(Y, p0, lambda0, a, c, p, q, niter, nburn) {
    n = length(Y)
    SY = sum(Y)
    p = numeric(niter)
    lambda = numeric(niter)
    Z = matrix(0, niter, n)
    p[1] = p0
    lambda[1] = lambda0
    Z[1, ] = ifelse(Y == 0, rbinom(n, 1, p[1] * exp(-lambda[1])/(p[1] * exp(-lambda[1]) +
        1 - p[1])), 1)
    for (i in 2:niter) {
        SZ = sum(Z[i - 1, ])
        p[i] = rbeta(1, a + SZ, c + n - SZ)
        lambda[i] = rgamma(1, p + SY, q + SZ)
```

```
        Z[i, ] = ifelse(Y == 0, rbinom(n, 1, p[i] * exp(-lambda[i])/(p[i] *
            exp(-lambda[i]) + 1 - p[i])), 1)
    }
    return(list(p = p[-(1:nburn)], lambda = lambda[-(1:nburn)], Z = Z[-(1:nburn),
        ]))
}

n = 1000
p = 0.75
lambda = 2
Y = ifelse(rbinom(n, 1, p) == 1, rpois(n, lambda), 0)
barplot(table(factor(Y, levels = 0:max(Y)))/n, space = 0)
lines(0:max(Y) + 0.5, c(1 - p, numeric(max(Y))) + p * dpois(0:max(Y), lambda),
    col = 2, lwd = 2)
```




```
posterior = MCMCzip(Y, 0.5, 1, 0.5, 0.5, 0.5, 0, 5000, 1000)
par(mfrow = c(1, 2))
hist(posterior$p, "FD", freq = FALSE, main = NA, xlab = "p")
abline(v = p, col = 2, lty = 2)
hist(posterior$lambda, "FD", freq = FALSE, main = NA, xlab = expression(lambda))
abline(v = lambda, col = 2, lty = 2)
```


Density
0
5
10
15
20
0.70
0.75
0.80
p
Density
0
1
2
3
4
5
6
1.8
2.0
2.2
λ

3 Λύσεις Θεμάτων Εξετάσεων

3.1 Οκτώβριος 2011

Θεωρούμε το ακόλουθο μοντέλο:

- Για $i = 1, 2, \dots, t$, η παρατήρηση x_i είναι μία ανεξάρτητη πραγματοποίηση μίας τυχαίας μεταβλητής Poisson με μέση τιμή θ_1.
- Για $i = t+1, t+2, \dots, n$, η παρατήρηση x_i είναι μία ανεξάρτητη πραγματοποίηση μίας τυχαίας μεταβλητής Poisson με μέση τιμή θ_2.

Υποθέτουμε τις prior $\theta_1 \sim \text{Gamma}(p_1, q_1)$, $\theta_2 \sim \text{Gamma}(p_2, q_2)$ και $t \sim U\{1, 2, \dots, n-1\}$.

α. Γράψτε ως μία σταθερά κανονικοποίησης τις δεσμευμένες posterior κατανομές των παραμέτρων του μοντέλου. Είναι αυτές γνωστές κατανομές; Επίσης βρείτε την περιθώρια posterior κατανομή του t. Είναι αυτή γνωστή κατανομή;

β. Γράψτε έναν αλγόριθμο MCMC που να προσομοιώνει δείγμα από την από κοινού posterior κατανομή των παραμέτρων του μοντέλου. Εξηγείστε πώς θα πραγματοποιηθεί η προσομοίωση του t.

Λύση.

α. Η από κοινού prior κατανομή των θ_1, θ_2 και t μπορεί να γραφτεί ως:

$$\begin{aligned}
\pi(\theta_1, \theta_2, t) &= \pi(\theta_1) \cdot \pi(\theta_2) \cdot \pi(t) \\
&= \frac{q_1^{p_1}}{\Gamma(p_1)} \theta_1^{p_1 - 1} e^{-q_1 \theta_1} \cdot \frac{q_2^{p_2}}{\Gamma(p_2)} \theta_2^{p_2 - 1} e^{-q_2 \theta_2} \cdot \frac{1}{n-1} \\
&\propto \theta_1^{p_1 - 1} e^{-q_1 \theta_1} \cdot \theta_2^{p_2 - 1} e^{-q_2 \theta_2}.
\end{aligned}$$

Ορίζουμε:

$$S_t = \sum_{i=1}^{t} x_i.$$

Τότε, παρατηρούμε ότι:

$$S_n - S_t = \sum_{i=1}^{n} x_i - \sum_{i=1}^{t} x_i = \sum_{i=t+1}^{n} x_i.$$

Η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$\begin{aligned}
f(x \mid \theta_1, \theta_2, t) &= \prod_{i=1}^{t} f(x_i \mid \theta_1) \cdot \prod_{i=t+1}^{n} f(x_i \mid \theta_2) \\
&= \prod_{i=1}^{t} e^{-\theta_1} \frac{\theta_1^{x_i}}{x_i!} \cdot \prod_{i=t+1}^{n} e^{-\theta_2} \frac{\theta_2^{x_i}}{x_i!} \\
&= e^{-t\theta_1} \theta_1^{S_t} \cdot e^{-(n-t)\theta_2} \theta_2^{S_n - S_t} \cdot \prod_{i=1}^{n} \frac{1}{x_i!} \\
&\propto e^{-t\theta_1} \theta_1^{S_t} \cdot e^{-(n-t)\theta_2} \theta_2^{S_n - S_t}.
\end{aligned}$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των θ_1, θ_2 και t ως εξής:

$$\begin{aligned}
\pi(\theta_1 \mid \theta_2, t, x) &\propto \pi(\theta_1, \theta_2, t \mid x) \\
&\propto \pi(\theta_1, \theta_2, t) \cdot f(x \mid \theta_1, \theta_2, t) \\
&\propto \theta_1^{p_1-1} e^{-q_1\theta_1} \cdot e^{-t\theta_1} \theta_1^{S_t} \\
&= \theta_1^{p_1+S_t-1} e^{-(q_1+t)\theta_1},
\end{aligned}$$

$$\begin{aligned}
\pi(\theta_2 \mid \theta_1, t, x) &\propto \pi(\theta_1, \theta_2, t) \cdot f(x \mid \theta_1, \theta_2, t) \\
&\propto \theta_2^{p_2-1} e^{-q_2\theta_2} \cdot e^{-(n-t)\theta_2} \theta_2^{S_n-S_t} \\
&= \theta_2^{p_2+S_n-S_t-1} e^{-(q_2+n-t)\theta_2},
\end{aligned}$$

$$\begin{aligned}
\pi(t \mid \theta_1, \theta_2, x) &\propto \pi(\theta_1, \theta_2, t) \cdot f(x \mid \theta_1, \theta_2, t) \\
&\propto e^{-t\theta_1} \theta_1^{S_t} \cdot e^{-(n-t)\theta_2} \theta_2^{S_n-S_t} \\
&\propto e^{-t\theta_1} \theta_1^{S_t} \cdot e^{t\theta_2} \theta_2^{-S_t} \\
&= e^{-(\theta_1-\theta_2)t} \left(\frac{\theta_1}{\theta_2}\right)^{S_t}.
\end{aligned}$$

Παρατηρούμε ότι οι παράμετροι θ_1 και θ_2 είναι a posteriori ανεξάρτητες δεδομένου του t, δηλαδή:

$$\theta_1 \mid t, x \sim \text{Gamma}\left(p_1 + S_t, q_1 + t\right), \quad \theta_2 \mid t, x \sim \text{Gamma}\left(p_2 + S_n - S_t, q_2 + n - t\right).$$

Η δεσμευμένη posterior κατανομή του t είναι μία διακριτή κατανομή $\pi(t \mid \theta_1, \theta_2, x)$ με πεπερασμένο στήριγμα $\{1, 2, \dots, n-1\}$. Για τον υπολογισμό του διανύσματος πιθανοτήτων $\pi\left(t \mid \theta_1, \theta_2, x\right)$ χρησιμοποιούμε το λεγόμενο κόλπο Log-Sum-Exp. Δηλαδή, ορίζουμε:

$$v_i = -(\theta_1 - \theta_2)i + S_i \log \frac{\theta_1}{\theta_2}, \quad m = \max_{i \in \{1, \dots, n-1\}} v_i.$$

Τότε, παίρνουμε ότι:

$$\pi\left(t \mid \theta_1, \theta_2, x\right) = \frac{e^{v_t - m}}{\sum_{i=1}^{n-1} e^{v_i - m}}.$$

Επιπλέον, παίρνουμε την περιθώρια posterior κατανομή του t ως εξής:

$$\begin{aligned}
\pi(t \mid x) &= \int_0^\infty \int_0^\infty \pi(\theta_1, \theta_2, t \mid x) d\theta_1 d\theta_2 \\
&\propto \int_0^\infty \int_0^\infty \pi(\theta_1, \theta_2, t) f(x \mid \theta_1, \theta_2, t) d\theta_1 d\theta_2 \\
&\propto \int_0^\infty \int_0^\infty \theta_1^{p_1-1} e^{-q_1\theta_1} \cdot \theta_2^{p_2-1} e^{-q_2\theta_2} \cdot e^{-t\theta_1} \theta_1^{S_t} \cdot e^{-(n-t)\theta_2} \theta_2^{S_n-S_t} d\theta_1 d\theta_2 \\
&= \int_0^\infty \theta_1^{p_1+S_t-1} e^{-(q_1+t)\theta_1} d\theta_1 \cdot \int_0^\infty \theta_2^{p_2+S_n-S_t-1} e^{-(q_2+n-t)\theta_2} d\theta_2
\end{aligned}$$

$$= \frac{\Gamma(p_1 + S_t)}{(q_1 + t)^{p_1 + S_t}} \cdot \frac{\Gamma(p_2 + S_n - S_t)}{(q_2 + n - t)^{p_2 + S_n - S_t}}.$$

β. Για να πάρουμε δείγμα από αυτήν την από κοινού posterior κατανομή, χρησιμοποιούμε τον δειγματολήπτη Gibbs που φαίνεται στην επόμενη σελίδα.

Algorithm 3.1 Δειγματολήπτης Gibbs

Ξεκινάμε με αρχικές τιμές $\theta_1^{(0)}$, $\theta_2^{(0)}$, $t^{(0)}$.

Επαναλαμβάνουμε τα παρακάτω βήματα:

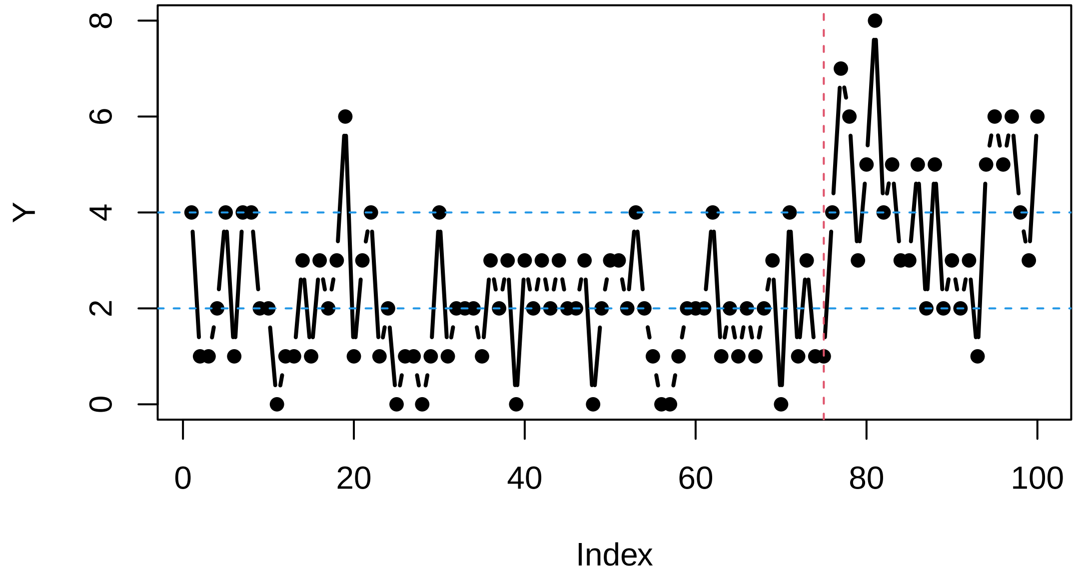
1: Προσομοιώνουμε $\theta_1^{(k)} \sim \text{Gamma}\left(p_1 + S_{t^{(k-1)}}, q_1 + t^{(k-1)}\right)$.

2: Προσομοιώνουμε $\theta_2^{(k)} \sim \text{Gamma}\left(p_2 + S_n - S_{t^{(k-1)}}, q_2 + n - t^{(k-1)}\right)$.

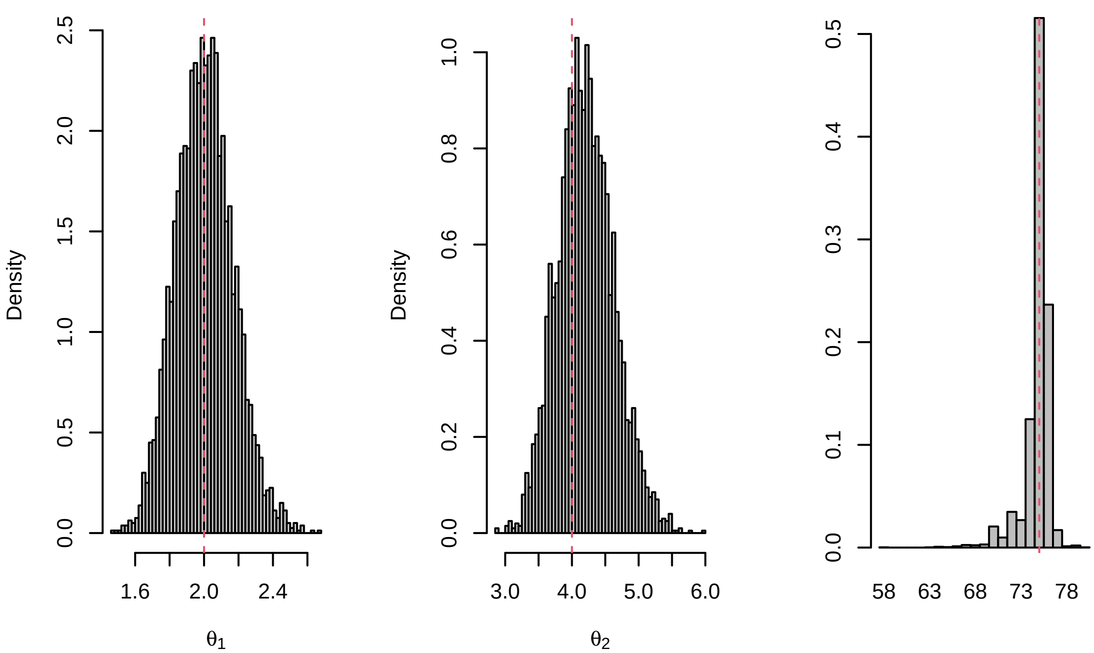
3: Υπολογίζουμε το διάνυσμα πιθανοτήτων $\pi\left(t \mid \theta_1^{(k)}, \theta_2^{(k)}, x\right)$. Προσομοιώνουμε μία τιμή $t^{(k)}$ από το σύνολο $\{1, 2, \ldots, n-1\}$ σύμφωνα με αυτό το διάνυσμα πιθανοτήτων.

```r
MCMCchangepoint = function(Y, theta10, theta20, p1, q1, p2, q2, niter, nburn) {
    n = length(Y)
    S = cumsum(Y)
    theta1 = numeric(niter)
    theta2 = numeric(niter)
    t = numeric(niter)
    theta1[1] = theta10
    theta2[1] = theta20
    logprob = S[-n] * log(theta1[1]/theta2[1]) - (theta1[1] - theta2[1]) * (1:(n -
        1))
    t[1] = sample(n - 1, 1, prob = exp(logprob - max(logprob)))
    for (i in 2:niter) {
        theta1[i] = rgamma(1, p1 + S[t[i - 1]], q1 + t[i - 1])
        theta2[i] = rgamma(1, p2 + S[n] - S[t[i - 1]], q2 + n - t[i - 1])
        logprob = S[-n] * log(theta1[i]/theta2[i]) - (theta1[i] - theta2[i]) *
            (1:(n - 1))
        t[i] = sample(n - 1, 1, prob = exp(logprob - max(logprob)))
    }
    return(list(theta1 = theta1[-(1:nburn)], theta2 = theta2[-(1:nburn)], t = t[-(1:nburn)]))
}

n = 100
theta1 = 2
theta2 = 4
t = 75
Y = c(rpois(t, theta1), rpois(n - t, theta2))
plot(Y, type = "b", pch = 16, lwd = 2)
abline(h = theta1, col = 4, lty = 2)
```

```
abline(h = theta2, col = 4, lty = 2)
abline(v = t, col = 2, lty = 2)
```




```
posterior = MCMCchangepoint(Y, 1, 1, 0.5, 0, 0.5, 0, 5000, 1000)
par(mfrow = c(1, 3))
hist(posterior$theta1, "FD", freq = FALSE, main = NA, xlab = expression(theta[1]))
abline(v = theta1, col = 2, lty = 2)
hist(posterior$theta2, "FD", freq = FALSE, main = NA, xlab = expression(theta[2]))
abline(v = theta2, col = 2, lty = 2)
barplot(table(factor(posterior$t, levels = min(posterior$t):max(posterior$t)))/4000,
    space = 0)
abline(v = t - min(posterior$t) + 0.5, col = 2, lty = 2)
```



3.2 Φεβρουάριος 2009

Έστω τυχαίο δείγμα $y_1, \ldots, y_n$ από την εξής μίξη κατανομών Poisson:

$$f(y_i \mid \alpha, \beta, \gamma) = \gamma f_{\text{Poisson}}(y_i \mid \alpha) + (1-\gamma) f_{\text{Poisson}}\left(y_i \mid \alpha e^{\beta x_i}\right),$$

όπου με $f_{\text{Poisson}}(y_i \mid \theta)$ συμβολίζουμε τη συνάρτηση πιθανότητας της κατανομής Poisson(θ) και x_i είναι μία επεξηγηματική μεταβλητή για την i-οστή παρατήρηση. Θεωρούμε prior ανεξαρτησία ανάμεσα στα α, β, γ με κατανομές $\alpha \sim \text{Gamma}(2,1)$, $\beta \sim \mathcal{N}(0,1)$ και $\gamma \sim U(0,1) \equiv \text{Beta}(1,1)$.

α. Γράψτε ως μία σταθερά κανονικοποίησης την από κοινού posterior κατανομή των παραμέτρων α, β, γ. Για να πάρετε δείγμα από αυτήν την posterior κατανομή, κατασκευάστε έναν αλγόριθμο MCMC με βήματα Metropolis-Hastings για την ανανέωση καθεμιάς από τις παραμέτρους. Για τις παραμέτρους α, γ να χρησιμοποιηθούν ως ανεξάρτητες γεννήτριες προτεινόμενων τιμών οι αντίστοιχες prior κατανομές των παραμέτρων, ενώ για το β να χρησιμοποιηθεί Random Walk Metropolis-Hastings. Για το τελευταίο βήμα πώς θα επιλεγεί η διασπορά της γεννήτριας προτεινόμενων τιμών;

β. Θεωρούμε τώρα την παρακάτω τεχνική αύξησης δεδομένων. Για κάθε y_i εισάγουμε μία δίτιμη τυχαία μεταβλητή z_i τέτοια, ώστε:

$$P(Z_i = 1 \mid \gamma) = 1 - P(Z_i = 0 \mid \gamma) = \gamma.$$

Τότε, η δεσμευμένη συνάρτηση πιθανότητας της y_i δεδομένου του z_i θα είναι:

$$f(y_i \mid z_i, \alpha, \beta) = \begin{cases} f_{\text{Poisson}}(y_i \mid \alpha), & z_i = 1 \\ f_{\text{Poisson}}\left(y_i \mid \alpha e^{\beta x_i}\right), & z_i = 0 \end{cases}.$$

Βρείτε τις δεσμευμένες posterior κατανομές όλων των αγνώστων ποσοτήτων. Είναι αυτές γνωστές κατανομές; Γράψτε έναν αλγόριθμο MCMC που να προσομοιώνει αυτές τις ποσότητες. Πώς θα προσομοιώσετε καθένα από τα z_i;

Λύση.

α. Η από κοινού prior κατανομή των α, β και γ μπορεί να γραφτεί ως:

$$\begin{aligned} \pi(\alpha, \beta, \gamma) &= \pi(\alpha) \cdot \pi(\beta) \cdot \pi(\gamma) \\ &= \alpha e^{-\alpha} \cdot \frac{1}{\sqrt{2\pi}} e^{-\beta^2/2} \cdot 1 \\ &\propto \alpha e^{-\alpha} \cdot e^{-\beta^2/2}. \end{aligned}$$

Η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$\begin{aligned} f(y \mid \alpha, \beta, \gamma) &= \prod_{i=1}^{n} f(y_i \mid \alpha, \beta, \gamma) \\ &= \prod_{i=1}^{n} \left[\gamma f_{\text{Poisson}}(y_i \mid \alpha) + (1-\gamma) f_{\text{Poisson}}\left(y_i \mid \alpha e^{\beta x_i}\right)\right] \end{aligned}$$

$$
\begin{aligned}
&= \prod_{i=1}^{n} \left[\gamma e^{-\alpha} \frac{\alpha^{y_i}}{y_i!} + (1-\gamma) e^{-\alpha e^{\beta x_i}} \frac{\alpha^{y_i} e^{\beta x_i y_i}}{y_i!} \right] \\
&= \prod_{i=1}^{n} \frac{\alpha^{y_i}}{y_i!} \left[\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}} \right] \\
&\propto \alpha^{n\bar{y}} \prod_{i=1}^{n} \left[\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}} \right].
\end{aligned}
$$

Για τον υπολογισμό της πιθανοφάνειας χρησιμοποιούμε πάλι το κόλπο Log-Sum-Exp. Δηλαδή, ορίζουμε:

$$
v_{i1} = \log \gamma + \log f_{\text{Poisson}}\left(y_i \mid \alpha\right), \quad v_{i1} = \log(1-\gamma) + \log f_{\text{Poisson}}\left(y_i \mid \alpha e^{\beta x_i}\right), \quad m_i = \max\left\{v_{i1}, v_{i2}\right\}.
$$

Τότε, παίρνουμε ότι:

$$
\log f(y \mid \alpha, \beta, \gamma) = \sum_{i=1}^{n} \left[m_i + \log\left(e^{v_{i1} - m_i} + e^{v_{i2} - m_i}\right) \right].
$$

Επομένως, παίρνουμε την από κοινού posterior κατανομή των α, β και γ ως εξής:

$$
\begin{aligned}
\pi(\alpha, \beta, \gamma \mid y) &\propto \pi(\alpha, \beta, \gamma) \cdot f(y \mid \alpha, \beta, \gamma) \\
&\propto \alpha e^{-\alpha} \cdot e^{-\beta^2/2} \cdot \alpha^{n\bar{y}} \prod_{i=1}^{n} \left[\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}} \right] \\
&= \alpha^{n\bar{y}+1} e^{-\alpha} \cdot e^{-\beta^2/2} \cdot \prod_{i=1}^{n} \left[\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}} \right].
\end{aligned}
$$

Για να πάρουμε δείγμα από αυτήν την από κοινού posterior κατανομή, χρησιμοποιούμε τον αλγόριθμο Metropolis-Hastings που φαίνεται στην επόμενη σελίδα. Προσαρμόζουμε τη διασπορά σ_β^2 της γεννήτριας προτεινόμενων τιμών για το β έτσι, ώστε το ποσοστό αποδοχής των προτεινόμενων τιμών του αλγορίθμου για το β να είναι περίπου ίσο με 50%.

Algorithm 3.2 Metropolis-Hastings

Ξεκινάμε με αρχικές τιμές $\alpha^{(0)}$, $\gamma^{(0)}$, $\beta^{(0)}$.

Επαναλαμβάνουμε τα παρακάτω βήματα:

1: Προσομοιώνουμε $\alpha^* \sim \text{Gamma}(10, 1)$ και $U_\alpha \sim U(0, 1)$.

2: Υπολογίζουμε τον λόγο:

$$A_\alpha = \frac{f\left(y \mid \alpha^*, \beta^{(\ell-1)}, \gamma^{(\ell-1)}\right)}{f\left(y \mid \alpha^{(\ell-1)}, \beta^{(\ell-1)}, \gamma^{(\ell-1)}\right)}.$$

3: Αν $U_\alpha < A_\alpha$, τότε θέτουμε $\alpha^{(\ell)} = \alpha^*$. Διαφορετικά, θέτουμε $\alpha^{(\ell)} = \alpha^{(\ell-1)}$.

4: Προσομοιώνουμε $\gamma^* \sim U(0, 1)$ και $U_\gamma \sim U(0, 1)$.

5: Υπολογίζουμε τον λόγο:

$$A_\gamma = \frac{f\left(y \mid \alpha^{(\ell)}, \beta^{(\ell-1)}, \gamma^*\right)}{f\left(y \mid \alpha^{(\ell)}, \beta^{(\ell-1)}, \gamma^{(\ell-1)}\right)}.$$

6: Αν $U_\gamma < A_\gamma$, τότε θέτουμε $\gamma^{(\ell)} = \gamma^*$. Διαφορετικά, θέτουμε $\gamma^{(\ell)} = \gamma^{(\ell-1)}$.

7: Προσομοιώνουμε $\beta^* \sim \mathcal{N}\left(\beta^{(\ell-1)}, \sigma_\beta^2\right)$ και $U_\beta \sim U(0, 1)$.

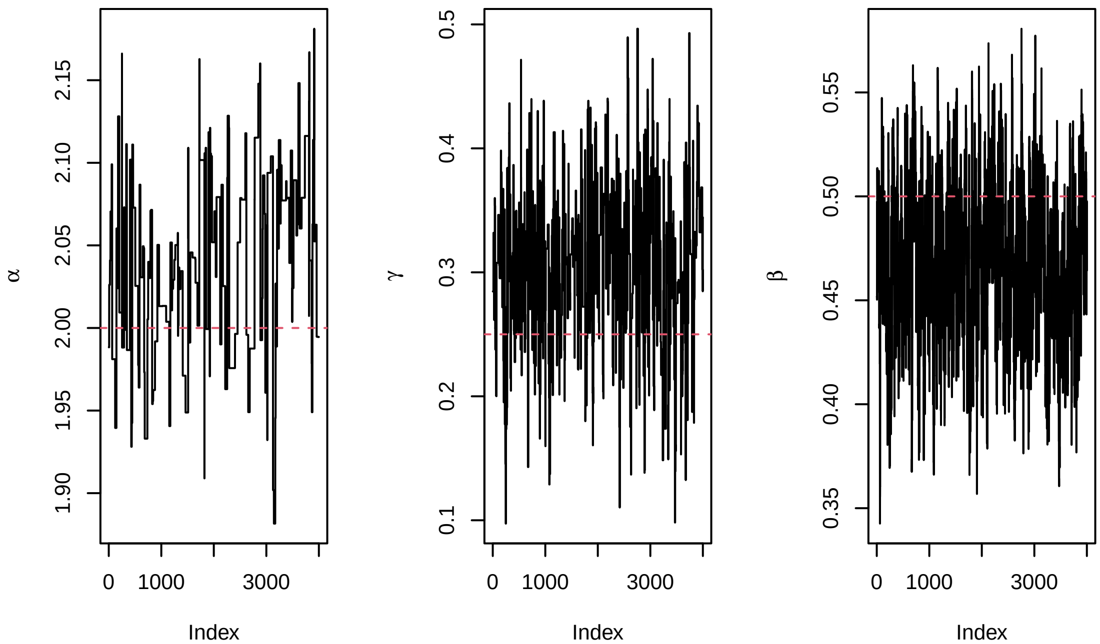
8: Υπολογίζουμε τον λόγο:

$$A_\beta = \frac{\pi\left(\beta^* \mid \alpha^{(\ell)}, \gamma^{(\ell)}, y\right)}{\pi\left(\beta^{(\ell-1)} \mid \alpha^{(\ell)}, \gamma^{(\ell)}, y\right)}.$$

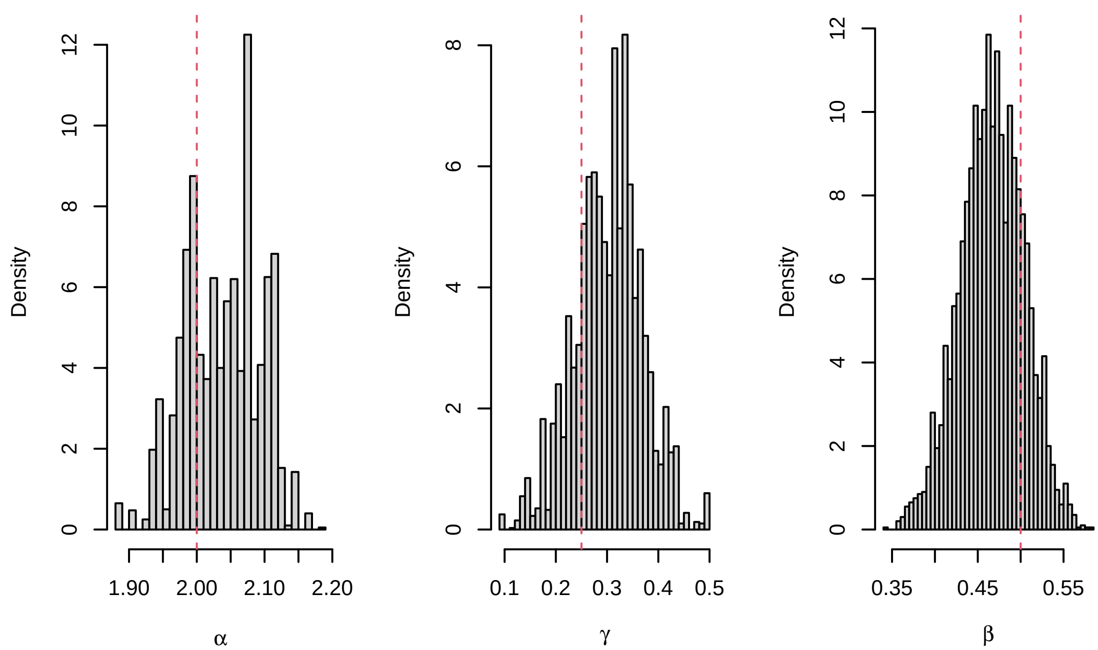
9: Αν $U_\beta < A_\beta$, τότε θέτουμε $\beta^{(\ell)} = \beta^*$. Διαφορετικά, θέτουμε $\beta^{(\ell)} = \beta^{(\ell-1)}$.

```
logdpois = function(Y, X, alpha, gamma, beta) {
    logprob = cbind(log(gamma) + dpois(Y, alpha, log = TRUE), log(1 - gamma) +
        dpois(Y, alpha * exp(beta * X), log = TRUE))
    maximum = apply(logprob, 1, max)
    return(sum(maximum + log(rowSums(exp(logprob - maximum)))))
}

MHpois = function(Y, X, alpha0, gamma0, beta0, betasd, niter, nburn) {
    alpha = numeric(niter)
    gamma = numeric(niter)
    beta = numeric(niter)
    alpha[1] = alpha0
    gamma[1] = gamma0
    beta[1] = beta0
    for (i in 2:niter) {
```

```
        alphastar = rgamma(1, 2)
        logA = logdpois(Y, X, alphastar, gamma[i - 1], beta[i - 1]) - logdpois(Y,
            X, alpha[i - 1], gamma[i - 1], beta[i - 1])
        alpha[i] = ifelse(log(runif(1)) < logA, alphastar, alpha[i - 1])
        gammastar = runif(1)
        logA = logdpois(Y, X, alpha[i], gammastar, beta[i - 1]) - logdpois(Y,
            X, alpha[i], gamma[i - 1], beta[i - 1])
        gamma[i] = ifelse(log(runif(1)) < logA, gammastar, gamma[i - 1])
        betastar = rnorm(1, beta[i - 1], betasd)
        logA = (beta[i - 1]^2 - betastar^2)/2 + logdpois(Y, X, alpha[i], gamma[i],
            betastar) - logdpois(Y, X, alpha[i], gamma[i], beta[i - 1])
        beta[i] = ifelse(log(runif(1)) < logA, betastar, beta[i - 1])
    }
    return(list(alpha = alpha[-(1:nburn)], gamma = gamma[-(1:nburn)], beta = beta[-(1:nburn)]))
}

n = 1000
alpha = 2
gamma = 0.25
beta = 0.5
X = rnorm(n)
Z = rbinom(n, 1, gamma)
Y = ifelse(Z == 1, rpois(n, alpha), rpois(n, alpha * exp(beta * X)))
posterior = MHpois(Y, X, 1, 0.5, 0, 0.05, 5000, 1000)
par(mfrow = c(1, 3))
plot(posterior$alpha, type = "l", ylab = expression(alpha))
abline(h = alpha, col = 2, lty = 2)
plot(posterior$gamma, type = "l", ylab = expression(gamma))
abline(h = gamma, col = 2, lty = 2)
plot(posterior$beta, type = "l", ylab = expression(beta))
abline(h = beta, col = 2, lty = 2)
```




```
hist(posterior$alpha, "FD", freq = FALSE, main = NA, xlab = expression(alpha))
abline(v = alpha, col = 2, lty = 2)
hist(posterior$gamma, "FD", freq = FALSE, main = NA, xlab = expression(gamma))
abline(v = gamma, col = 2, lty = 2)
hist(posterior$beta, "FD", freq = FALSE, main = NA, xlab = expression(beta))
abline(v = beta, col = 2, lty = 2)
```



Παρατηρούμε ότι η prior του α δεν είναι καθόλου αποδοτική ως ανεξάρτητη προτείνουσα συνάρτηση πυκνότητας πιθανότητας.

β. Ορίζουμε:

$$n_1 = \sum_{i=1}^{n} \mathbb{1}_{\{z_i=1\}}, \quad S_{XY} = \sum_{i=1}^{n} \mathbb{1}_{\{z_i=0\}} x_i y_i, \quad S_\beta = \sum_{i=1}^{n} \mathbb{1}_{\{z_i=0\}} e^{\beta x_i}.$$

Η πλήρης πιθανοφάνεια, δηλαδή η από κοινού πιθανοφάνεια των παρατηρούμενων μεταβλητών y_i και των λανθανουσών μεταβλητών z_i, δίνεται από τη σχέση:

$$\begin{aligned}
f(y, z \mid \alpha, \beta, \gamma) &= \prod_{i=1}^{n} f(y_i, z_i \mid \alpha, \beta, \gamma) \\
&= \prod_{i=1}^{n} f(z_i \mid \gamma) f(y_i \mid z_i, \alpha, \beta) \\
&= \prod_{i=1}^{n} \left[P(Z_i = 1 \mid \gamma) f_{\text{Poisson}}(y_i \mid \alpha)\right]^{\mathbb{1}_{\{z_i=1\}}} \left[P(Z_i = 0 \mid \gamma) f_{\text{Poisson}}\left(y_i \mid \alpha e^{\beta x_i}\right)\right]^{\mathbb{1}_{\{z_i=0\}}} \\
&= \prod_{i=1}^{n} \left(\gamma e^{-\alpha} \frac{\alpha^{y_i}}{y_i!}\right)^{\mathbb{1}_{\{z_i=1\}}} \left[(1-\gamma) e^{-\alpha e^{\beta x_i}} \frac{\alpha^{y_i} e^{\beta x_i y_i}}{y_i!}\right]^{\mathbb{1}_{\{z_i=0\}}} \\
&= \prod_{i=1}^{n} \left(\gamma e^{-\alpha}\right)^{\mathbb{1}_{\{z_i=1\}}} \left[(1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}\right]^{\mathbb{1}_{\{z_i=0\}}} \prod_{i=1}^{n} \frac{\alpha^{y_i}}{y_i!} \\
&\propto \gamma^{n_1} e^{-n_1 \alpha} \cdot (1-\gamma)^{n-n_1} e^{\beta S_{XY} - \alpha S_\beta} \cdot \alpha^{n\overline{y}} \\
&= \alpha^{n\overline{y}} e^{-(S_\beta + n_1)\alpha} \cdot \gamma^{n_1} (1-\gamma)^{n-n_1} \cdot e^{S_{XY}\beta}.
\end{aligned}$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των α και γ ως εξής:

$$\begin{aligned}
\pi(\alpha \mid \beta, \gamma, z, y) &\propto \pi(\alpha, \beta, \gamma, z \mid y) \\
&\propto \pi(\alpha, \beta, \gamma) \cdot f(y, z \mid \alpha, \beta, \gamma) \\
&\propto \alpha e^{-\alpha} \cdot \alpha^{n\overline{y}} e^{-(S_\beta + n_1)\alpha} \\
&= \alpha^{n\overline{y}+1} e^{-(S_\beta + n_1 + 1)\alpha},
\end{aligned}$$

$$\begin{aligned}
\pi(\gamma \mid \alpha, \beta, z, y) &\propto \pi(\alpha, \beta, \gamma) \cdot f(y, z \mid \alpha, \beta, \gamma) \\
&\propto \gamma^{n_1} (1-\gamma)^{n-n_1}.
\end{aligned}$$

Δηλαδή,

$$\alpha \mid \beta, z, y \sim \text{Gamma}\left(n\overline{y} + 2, S_\beta + n_1 + 1\right),$$

$$\gamma \mid z \sim \text{Beta}\left(n_1 + 1, n - n_1 + 1\right).$$

Επιπλέον, παίρνουμε τη δεσμευμένη posterior κατανομή του β ως εξής:

$$\begin{aligned}
\pi(\beta \mid \alpha, z, y) &\propto \pi(\alpha, \beta, \gamma) \cdot f(y, z \mid \alpha, \beta, \gamma) \\
&\propto e^{-\beta^2/2 + S_{XY}\beta - \alpha S_\beta},
\end{aligned}$$

η οποία δεν είναι κάποια γνωστή κατανομή.

Τέλος, παίρνουμε τη δεσμευμένη posterior κατανομή των λανθανουσών μεταβλητών z_i ως εξής:

$$f(z_i \mid y_i, \alpha, \beta, \gamma) \propto f(y_i, z_i \mid \alpha, \beta, \gamma) \propto \left(\gamma e^{-\alpha}\right)^{\mathbb{1}_{\{z_i=1\}}} \left[(1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}\right]^{\mathbb{1}_{\{z_i=0\}}}.$$

Δηλαδή,

$$(z_i \mid y_i, \alpha, \beta, \gamma) \sim \text{Bernoulli}\left(\frac{\gamma e^{-\alpha}}{\gamma e^{-\alpha} + (1-\gamma)e^{\beta x_i y_i - \alpha e^{\beta x_i}}}\right).$$

Για να πάρουμε δείγμα από αυτήν την από κοινού posterior κατανομή, χρησιμοποιούμε τον αλγόριθμο MCMC που φαίνεται παρακάτω.

Algorithm 3.3 Markov Chain Monte Carlo

Ξεκινάμε με αρχικές τιμές $\alpha^{(0)}$, $\gamma^{(0)}$, $\beta^{(0)}$, $z^{(0)}$.

Επαναλαμβάνουμε τα παρακάτω βήματα:

1: Προσομοιώνουμε $\alpha^{(\ell)} \sim \text{Gamma}\left(n\overline{y} + 2, S_{\beta^{(\ell-1)}} + n_1 + 1\right)$.

2: Προσομοιώνουμε $\gamma^{(\ell)} \sim \text{Beta}\left(n_1 + 1, n - n_1 + 1\right)$.

3: Προσομοιώνουμε $\beta^* \sim \mathcal{N}\left(\beta^{(\ell-1)}, \sigma_\beta^2\right)$ και $U_\beta \sim U(0,1)$.

4: Υπολογίζουμε τον λόγο:

$$A_\beta = \frac{\pi\left(\beta^* \mid \alpha^{(\ell)}, z^{(\ell-1)}, y\right)}{\pi\left(\beta^{(\ell-1)} \mid \alpha^{(\ell)}, z^{(\ell-1)}, y\right)}.$$

5: Αν $U_\beta < A_\beta$, τότε θέτουμε $\beta^{(\ell)} = \beta^*$. Διαφορετικά, θέτουμε $\beta^{(\ell)} = \beta^{(\ell-1)}$.

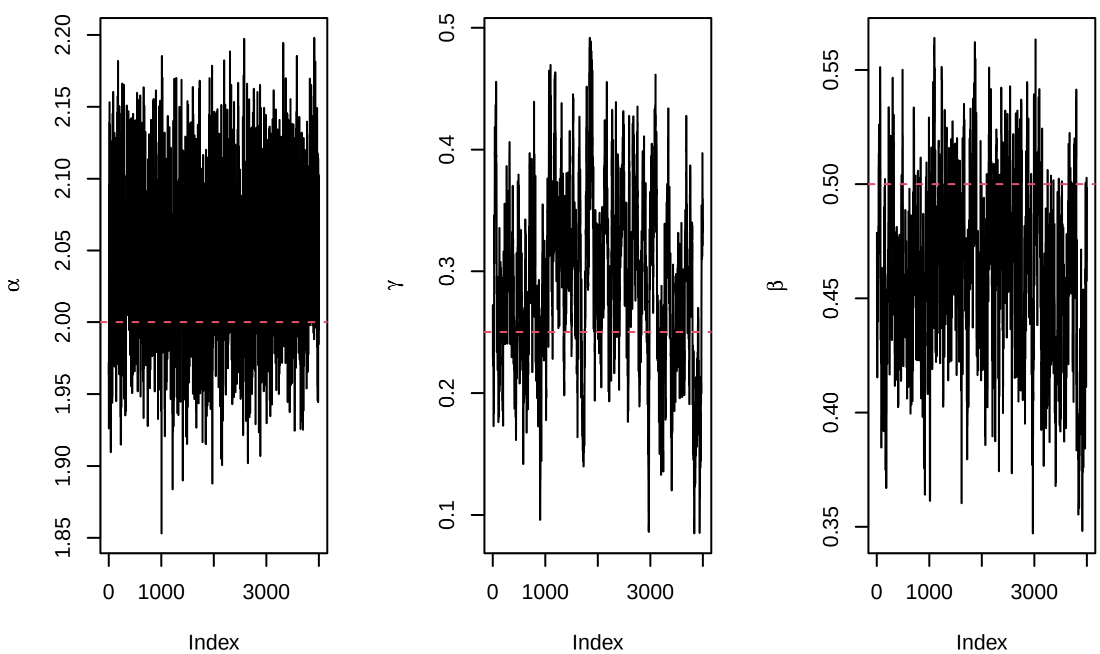
6: Υπολογίζουμε τις πιθανότητες $p_i = P\left(Z_i = 1 \mid y_i, \alpha^{(\ell)}, \gamma^{(\ell)}, \beta^{(\ell)}\right)$.

7: Προσομοιώνουμε $U_i \sim U(0,1)$.

8: Αν $U_i < p_i$, τότε θέτουμε $z_i^{(\ell)} = 1$. Διαφορετικά, θέτουμε $z_i^{(\ell)} = 0$.

```
prob = function(Y, X, alpha, gamma, beta) {
    logprob = cbind(log(gamma) - alpha, log(1 - gamma) + beta * X * Y - alpha *
        exp(beta * X))
    maximum = apply(logprob, 1, max)
    unnormalized = exp(logprob - maximum)
    return(unnormalized[, 1]/rowSums(unnormalized))
}

MCMCpois = function(Y, X, alpha0, gamma0, beta0, betasd, niter, nburn) {
    n = length(Y)
    S = sum(Y)
    alpha = numeric(niter)
    gamma = numeric(niter)
    beta = numeric(niter)
    Z = matrix(0, niter, n)
    alpha[1] = alpha0
    gamma[1] = gamma0
    beta[1] = beta0
    Z[1, ] = rbinom(n, 1, prob(Y, X, alpha[1], gamma[1], beta[1]))
```

```
    for (i in 2:niter) {
        n1 = sum(Z[i - 1, ])
        Sbeta = sum(exp(beta[i - 1] * X[Z[i - 1, ] == 0]))
        alpha[i] = rgamma(1, S + 2, Sbeta + n1 + 1)
        gamma[i] = rbeta(1, n1 + 1, n - n1 + 1)
        betastar = rnorm(1, beta[i - 1], betasd)
        logA = (beta[i - 1]^2 - betastar^2)/2 + sum(X[Z[i - 1, ] == 0] * Y[Z[i -
            1, ] == 0]) * (betastar - beta[i - 1]) + alpha[i] * (Sbeta - sum(exp(betastar *
            X[Z[i - 1, ] == 0])))
        beta[i] = ifelse(log(runif(1)) < logA, betastar, beta[i - 1])
        Z[i, ] = rbinom(n, 1, prob(Y, X, alpha[i], gamma[i], beta[i]))
    }
    return(list(alpha = alpha[-(1:nburn)], gamma = gamma[-(1:nburn)], beta = beta[-(1:nburn)],
        Z = Z[-(1:nburn), ]))
}

posterior = MCMCpois(Y, X, 1, 0.5, 0, 0.05, 5000, 1000)
par(mfrow = c(1, 3))
plot(posterior$alpha, type = "l", ylab = expression(alpha))
abline(h = alpha, col = 2, lty = 2)
plot(posterior$gamma, type = "l", ylab = expression(gamma))
abline(h = gamma, col = 2, lty = 2)
plot(posterior$beta, type = "l", ylab = expression(beta))
abline(h = beta, col = 2, lty = 2)
```




```
hist(posterior$alpha, "FD", freq = FALSE, main = NA, xlab = expression(alpha))
abline(v = alpha, col = 2, lty = 2)
```

```
hist(posterior$gamma, "FD", freq = FALSE, main = NA, xlab = expression(gamma))
abline(v = gamma, col = 2, lty = 2)
hist(posterior$beta, "FD", freq = FALSE, main = NA, xlab = expression(beta))
abline(v = beta, col = 2, lty = 2)
```