Μπεϋζιανή Στατιστική
Λύσεις Ασκήσεων και Θεμάτων Εξετάσεων

Βασίλης Κατσιάνος

1 Ανάλυση Κανονικής Πιθανοφάνειας

1.1 Μονοδιάστατη Κανονική

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

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

β. Θεωρούμε τη συζυγή prior κατανομή $\mu \mid \tau \sim N\left(a, c^{-1}\tau^{-1}\right)$, $\tau \sim \text{Gamma}(p, q)$. Να υπολογιστούν η από κοινού posterior κατανομή των παραμέτρων μ και τ και η δεσμευμένη posterior κατανομή του τ.

Λύση.

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

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

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

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:

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

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

Δηλαδή,

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

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:

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

Υπολογίζουμε ότι:

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

Δηλαδή,

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

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

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

Δηλαδή,

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

1.2 Γενικευμένη Student's t

Έστω τυχαίο δείγμα $y_1, \ldots, y_n$ από τη γενικευμένη κατανομή t με μέσο $\mu \in \mathbb{R}$, ακρίβεια $\tau > 0$ και $\nu > 0$ βαθμούς ελευθερίας, δηλαδή:

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

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

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

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

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

Θεωρούμε γνωστή την παράμετρο ν και prior ανεξαρτησία ανάμεσα στις παραμέτρους μ, τ, με κατανομές $\mu \sim N\left(a, c^{-1}\right)$ και $\tau \sim \text{Gamma}(p, q)$. 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:

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

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

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

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

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

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

$$= \exp\left\{ -\frac{c+n\tau\overline{z}}{2}\mu^2 + \underbrace{(c+n\tau\overline{z})}_{c_n} \underbrace{\frac{ca+n\tau\overline{zy}}{c+n\tau\overline{z}}}_{a_n} \mu \right\},$$

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

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,

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

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

1.3 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 t με μέσο 0, ακρίβεια $\tau > 0$ και $\nu > 0$ βαθμούς ελευθερίας, δηλαδή:

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

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

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

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

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

Θεωρούμε γνωστούς τους βαθμούς ελευθερίας ν και ορίζουμε τη δεσμευμένα συζυγή prior κατανομή $\beta \mid \tau \sim N_k\left(a, \tau^{-1}C\right)$, $\tau \sim \text{Gamma}(p,q)$. 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:

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

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:

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

$$
\begin{aligned}
&= \tau^{\frac{n}{2}} \exp\left\{-\frac{y^{\mathrm{T}}Zy - 2\beta^{\mathrm{T}}X^{\mathrm{T}}Zy + \beta^{\mathrm{T}}X^{\mathrm{T}}ZX\beta}{2}\tau\right\} \cdot \prod_{i=1}^{n} z_i^{\frac{1}{2}} \prod_{i=1}^{n} z_i^{\frac{\nu}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\} \\
&= \exp\left\{-\frac{\beta^{\mathrm{T}}\tau X^{\mathrm{T}}ZX\beta}{2} + \beta^{\mathrm{T}}\tau X^{\mathrm{T}}Zy\right\} \cdot \tau^{\frac{n}{2}} \exp\left\{-\frac{y^{\mathrm{T}}Zy}{2}\tau\right\} \\
&\quad \times \prod_{i=1}^{n} z_i^{\frac{\nu+1}{2}-1} \exp\left\{-\frac{\nu}{2}\sum_{i=1}^{n} z_i\right\}.
\end{aligned}
$$

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

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

Υπολογίζουμε ότι:

$$
a_n^{\mathrm{T}}C_n^{-1}a_n = \left(C^{-1}a + X^{\mathrm{T}}Zy\right)^{\mathrm{T}}\left(C^{-1} + X^{\mathrm{T}}ZX\right)^{-1}\left(C^{-1}a + X^{\mathrm{T}}Zy\right).
$$

Δηλαδή,

$$
\beta \mid \tau, z, y \sim N_k\left(\left(C^{-1} + X^{\mathrm{T}}ZX\right)^{-1}\left(C^{-1}a + X^{\mathrm{T}}Zy\right), \left(C^{-1} + X^{\mathrm{T}}ZX\right)^{-1}\right),
$$

$$\tau \mid z, y \sim \text{Gamma}\left(p + \frac{n}{2}, q + \frac{a^{\mathrm{T}} C^{-1} a + y^{\mathrm{T}} Z y - a_n^{\mathrm{T}} C_n^{-1} a_n}{2}\right).$$

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

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

Δηλαδή,

$$\tau \mid \beta, z, y \sim \text{Gamma}\left(p + \frac{n+k}{2}, q + \frac{(\beta - a)^{\mathrm{T}} C^{-1} (\beta - a) + (y - X\beta)^{\mathrm{T}} Z (y - X\beta)}{2}\right).$$

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,

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

1.4 Πολυδιάστατη Κανονική

Έστω τυχαίο δείγμα $y_1, \ldots, y_n$ από την πολυδιάστατη κανονική κατανομή $N_k(\mu, \Sigma)$.

α. Θεωρούμε prior ανεξαρτησία ανάμεσα στις παραμέτρους, με κατανομές $\mu \sim N_k(a, C)$ και $\Sigma \sim \text{IW}(A, d)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των μ και Σ.

β. Θεωρούμε τη συζυγή prior κατανομή $\mu \mid \Sigma \sim N\left(a, c^{-1}\Sigma\right)$, $\Sigma \sim \text{IW}(A, d)$. Να υπολογιστεί η από κοινού posterior κατανομή των μ και Σ και η δεσμευμένη posterior κατανομή του Σ.

Λύση.

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

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

$$\propto \exp\left\{-\frac{\mu^{\mathrm{T}}C^{-1}\mu}{2}+\mu^{\mathrm{T}}C^{-1}a\right\}\cdot|\Sigma|^{-\frac{d+k+1}{2}}e^{-\frac{1}{2}\mathrm{tr}\left(A\Sigma^{-1}\right)}.$$

Λήμμα 1. Έστω $x \in \mathbb{R}^n$ και $A \in \mathbb{R}^{n\times n}$. Τότε,

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

Η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

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

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

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

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

Δηλαδή,

$$\mu \mid \Sigma, y \sim N_k\left(\left(C^{-1}+n\Sigma^{-1}\right)^{-1}\left(C^{-1}a+n\Sigma^{-1}\overline{y}\right), \left(C^{-1}+n\Sigma^{-1}\right)^{-1}\right),$$

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

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

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

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

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

$$\times |\Sigma|^{-\frac{d+n+k+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\left(A + caa^{\mathrm{T}} + \sum_{i=1}^{n} y_i y_i^{\mathrm{T}} - c_n a_n a_n^{\mathrm{T}}\right)\Sigma^{-1}\right]\right\}.$$

Υπολογίζουμε ότι:

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

Δηλαδή,

$$\mu \mid \Sigma, y \sim N_k\left(\frac{ca + n\overline{y}}{c+n}, \frac{\Sigma}{c+n}\right),$$

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

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

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

Δηλαδή,

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

1.5 Πολυδιάστατη Student's t

Έστω τυχαίο δείγμα $y_1, \ldots, y_n$ από την πολυδιάστατη κατανομή t με μέσο $\mu \in \mathbb{R}^k$, θετικά ορισμένο πίνακα συνδιακύμανσης $\Sigma \in \mathbb{R}^{k \times k}$ και $\nu > 0$ βαθμούς ελευθερίας, δηλαδή:

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

Θεωρούμε τις τυχαίες μεταβλητές $W_i \sim N_k(0, \Sigma)$ 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, $Z_i \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$. Παρατηρούμε ότι:

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

Θεωρούμε γνωστή την παράμετρο ν και prior ανεξαρτησία ανάμεσα στις παραμέτρους μ, Σ, με κατανομές $\mu \sim N_k(a, C)$ και $\Sigma \sim \text{IW}(A, d)$. Να υπολογιστούν οι δεσμευμένες posterior κατανομές των παραμέτρων μ, Σ και των λανθανουσών μεταβλητών z_i.

Λύση.

Η από κοινού prior κατανομή μπορεί να γραφτεί ως:

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

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

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

όπου $\overline{zy} = \frac{1}{n}\sum\limits_{i=1}^{n} z_i y_i$.

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

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

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

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

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

Δηλαδή,

$$\mu \mid \Sigma, z, y \sim N_k\left(\left(C^{-1} + n\bar{z}\Sigma^{-1}\right)^{-1}\left(C^{-1}a + n\Sigma^{-1}\overline{zy}\right), \left(C^{-1} + n\bar{z}\Sigma^{-1}\right)^{-1}\right),$$

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

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

2 Markov Chain Monte Carlo

2.1 Οκτώβριος 2011

Θέμα 1. Θεωρούμε το μοντέλο:

M_1 : Για $i = 1, 2, \ldots, t_0$, η παρατήρηση x_i είναι μία ανεξάρτητη πραγματοποίηση μίας τυχαίας μεταβλητής Poisson με μέση τιμή θ_1.

Για $i = t_0+1, t_0+2, \ldots, T$, 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 $t_0 \sim U\{1, 2, \ldots, T-1\}$.

α. Γράψτε ως μία σταθερά κανονικοποίησης τις δεσμευμένες posterior κατανομές των παραμέτρων του μοντέλου M_1. Είναι αυτές γνωστές κατανομές; Επίσης βρείτε την περιθώρια posterior κατανομή του t_0. Είναι αυτή γνωστή κατανομή;

β. Γράψτε έναν αλγόριθμο MCMC που να προσομοιώνει δείγμα από την από κοινού posterior κατανομή των παραμέτρων του μοντέλου M_1. Εξηγείστε πώς θα πραγματοποιηθεί η προσομοίωση του t_0.

Λύση.

α. Η από κοινού prior κατανομή των θ_1, θ_2 και t_0 μπορεί να γραφτεί ως:

$$\begin{aligned}
\pi(\theta_1, \theta_2, t_0) &= \pi(\theta_1) \cdot \pi(\theta_2) \cdot \pi(t_0) \\
&= \frac{q_1^{p_1}}{\Gamma(p_1)} \theta_1^{p_1-1} e^{-q_1\theta_1} \cdot \frac{q_2^{p_2}}{\Gamma(p_2)} \theta_2^{p_2-1} e^{-q_2\theta_2} \cdot \frac{1}{T-1} \\
&\propto \theta_1^{p_1-1} e^{-q_1\theta_1} \cdot \theta_2^{p_2-1} e^{-q_2\theta_2}.
\end{aligned}$$

Ορίζουμε:

$$s(t_0) = \sum_{i=1}^{t_0} x_i.$$

Τότε,

$$s(T) - s(t_0) = \sum_{i=1}^{T} x_i - \sum_{i=1}^{t_0} x_i = \sum_{i=t_0+1}^{T} x_i.$$

Η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

$$\begin{aligned}
f(x \mid \theta_1, \theta_2, t_0) &= \prod_{i=1}^{t_0} f(x_i \mid \theta_1) \cdot \prod_{i=t_0+1}^{T} f(x_i \mid \theta_2) \\
&= \prod_{i=1}^{t_0} e^{-\theta_1} \frac{\theta_1^{x_i}}{x_i!} \cdot \prod_{i=t_0+1}^{T} e^{-\theta_2} \frac{\theta_2^{x_i}}{x_i!} \\
&= e^{-t_0\theta_1} \theta_1^{s(t_0)} \cdot e^{-(T-t_0)\theta_2} \theta_2^{s(T)-s(t_0)} \cdot \prod_{i=1}^{T} \frac{1}{x_i!} \\
&\propto e^{-t_0\theta_1} \theta_1^{s(t_0)} \cdot e^{-(T-t_0)\theta_2} \theta_2^{s(T)-s(t_0)}.
\end{aligned}$$

Επομένως, παίρνουμε τις δεσμευμένες posterior κατανομές των θ_1, θ_2 και t_0 ως εξής:

$$\pi(\theta_1 \mid \theta_2, t_0, x) \propto \pi(\theta_1, \theta_2, t_0 \mid x)$$

$$
\begin{aligned}
&\propto \pi(\theta_1, \theta_2, t_0) \cdot f(x \mid \theta_1, \theta_2, t_0) \\
&\propto \theta_1^{p_1-1} e^{-q_1\theta_1} \cdot e^{-t_0\theta_1} \theta_1^{s(t_0)} \\
&= \theta_1^{p_1+s(t_0)-1} e^{-(q_1+t_0)\theta_1},
\end{aligned}
$$

$$
\begin{aligned}
\pi(\theta_2 \mid \theta_1, t_0, x) &\propto \pi(\theta_1, \theta_2, t_0 \mid x) \\
&\propto \pi(\theta_1, \theta_2, t_0) \cdot f(x \mid \theta_1, \theta_2, t_0) \\
&\propto \theta_2^{p_2-1} e^{-q_2\theta_2} \cdot e^{-(T-t_0)\theta_2} \theta_2^{s(T)-s(t_0)} \\
&= \theta_2^{p_2+s(T)-s(t_0)-1} e^{-(q_2+T-t_0)\theta_2},
\end{aligned}
$$

$$
\begin{aligned}
\pi(t_0 \mid \theta_1, \theta_2, x) &\propto \pi(\theta_1, \theta_2, t_0 \mid x) \\
&\propto \pi(\theta_1, \theta_2, t_0) \cdot f(x \mid \theta_1, \theta_2, t_0) \\
&\propto e^{-t_0\theta_1} \theta_1^{s(t_0)} \cdot e^{-(T-t_0)\theta_2} \theta_2^{s(T)-s(t_0)} \\
&\propto e^{-t_0\theta_1} \theta_1^{s(t_0)} \cdot e^{t_0\theta_2} \theta_2^{-s(t_0)} \\
&= e^{-(\theta_1-\theta_2)t_0} \left(\frac{\theta_1}{\theta_2}\right)^{s(t_0)} = p(t_0).
\end{aligned}
$$

Παρατηρούμε ότι οι παράμετροι θ_1 και θ_2 είναι a posteriori ανεξάρτητες δεδομένου του t_0, δηλαδή:

$$
\theta_1 \mid t_0, x \sim \mathrm{Gamma}\left(p_1 + s(t_0), q_1 + t_0\right), \quad \theta_2 \mid t_0, x \sim \mathrm{Gamma}\left(p_2 + s(T) - s(t_0), q_2 + T - t_0\right).
$$

Η δεσμευμένη posterior κατανομή του t_0 είναι μία διακριτή κατανομή με πεπερασμένο στήριγμα $\{1, 2, \ldots, T-1\}$, δηλαδή:

$$
\pi(t_0 \mid \theta_1, \theta_2, x) = \frac{p(t_0)}{\sum_{i=1}^{T-1} p(i)}.
$$

Επιπλέον, παίρνουμε την περιθώρια posterior κατανομή του t_0 ως εξής:

$$
\begin{aligned}
\pi(t_0 \mid x) &= \int_0^\infty \int_0^\infty \pi(\theta_1, \theta_2, t_0 \mid x) d\theta_1 d\theta_2 \\
&\propto \int_0^\infty \int_0^\infty \pi(\theta_1, \theta_2, t_0) f(x \mid \theta_1, \theta_2, t_0) d\theta_1 d\theta_2 \\
&\propto \int_0^\infty \int_0^\infty \theta_1^{p_1-1} e^{-q_1\theta_1} \cdot \theta_2^{p_2-1} e^{-q_2\theta_2} \cdot e^{-t_0\theta_1} \theta_1^{s(t_0)} \cdot e^{-(T-t_0)\theta_2} \theta_2^{s(T)-s(t_0)} d\theta_1 d\theta_2 \\
&= \int_0^\infty \theta_1^{p_1+s(t_0)-1} e^{-(q_1+t_0)\theta_1} d\theta_1 \cdot \int_0^\infty \theta_2^{p_2+s(T)-s(t_0)-1} e^{-(q_2+T-t_0)\theta_2} d\theta_2 \\
&= \frac{\Gamma(p_1 + s(t_0))}{(q_1 + t_0)^{p_1+s(t_0)}} \cdot \frac{\Gamma(p_2 + s(T) - s(t_0))}{(q_2 + T - t_0)^{p_2+s(T)-s(t_0)}} = q(t_0)
\end{aligned}
$$

Η περιθώρια posterior κατανομή του t_0 είναι μία διακριτή κατανομή με πεπερασμένο στήριγμα

$\{1, 2, \ldots, T-1\}$, δηλαδή:

$$\pi(t_0 \mid x) = \frac{q(t_0)}{\sum_{i=1}^{T-1} q(i)}.$$

β. Χρησιμοποιούμε τον ακόλουθο δειγματολήπτη Gibbs:

Algorithm 1 Δειγματολήπτης Gibbs

Ξεκινάμε με μία αρχική τιμή $t_0^{(0)}$.

Επαναλαμβάνουμε τα παρακάτω βήματα:

1: Προσομοιώνουμε $\theta_1^{(k)} \sim \text{Gamma}\left(p_1 + s\left(t_0^{(k-1)}\right), q_1 + t_0^{(k-1)}\right)$.

2: Προσομοιώνουμε $\theta_2^{(k)} \sim \text{Gamma}\left(p_2 + s(T) - s\left(t_0^{(k-1)}\right), q_2 + T - t_0^{(k-1)}\right)$.

3: Υπολογίζουμε το διάνυσμα πιθανοτήτων $\pi\left(t_0 \mid \theta_1^{(k)}, \theta_2^{(k)}, x\right)$. Προσομοιώνουμε μία τιμή $t_0^{(k)}$ από το σύνολο $\{1, 2, \ldots, T-1\}$ σύμφωνα με αυτό το διάνυσμα πιθανοτήτων.

2.2 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 $\alpha \sim \text{Gamma}(10, 1)$, $\beta \sim N(0, 4)$ 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 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 άγνωστων 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:

$$\begin{aligned}
\pi(\alpha, \beta, \gamma) &= \pi(\alpha) \cdot \pi(\beta) \cdot \pi(\gamma) \\
&= \frac{1}{9!}\alpha^9 e^{-\alpha} \cdot \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}\beta^2} \cdot 1 \\
&\propto \alpha^9 e^{-\alpha} \cdot e^{-\frac{1}{8}\beta^2}.
\end{aligned}$$

Η πιθανοφάνεια ολόκληρου του δείγματος δίνεται από τη σχέση:

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

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

$$\begin{aligned}
\pi(\alpha, \beta, \gamma \mid y) &\propto \pi(\alpha, \beta, \gamma) \cdot f(y \mid \alpha, \beta, \gamma) \\
&\propto \alpha^9 e^{-\alpha} \cdot e^{-\frac{1}{8}\beta^2} \cdot \alpha^{n\bar{y}} \prod_{i=1}^{n} \left[\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}\right] \\
&= \alpha^{n\bar{y}+9} e^{-\alpha} \cdot e^{-\frac{1}{8}\beta^2} \cdot \prod_{i=1}^{n} \left[\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}\right].
\end{aligned}$$

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

$$\begin{aligned}\pi(\beta \mid \alpha, \gamma, y) &\propto \pi(\alpha, \beta, \gamma \mid y) \\ &\propto e^{-\frac{1}{8}\beta^2} \cdot \prod_{i=1}^{n} \left[\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}\right].\end{aligned}$$

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 $30-40\%$.

Algorithm 2 Metropolis-Hastings

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

Επαναλαμβάνουμε τα παρακάτω βήματα:

1: Προσομοιώνουμε $\beta^{\text{can}} \sim N\left(\beta^{(k-1)}, \sigma_\beta^2\right)$ και $U_\beta \sim U(0,1)$.

2: Υπολογίζουμε τον λόγο:

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

3: Αν $U_\beta < A_\beta$, τότε θέτουμε $\beta^{(k)} = \beta^{\text{can}}$. Διαφορετικά, θέτουμε $\beta^{(k)} = \beta^{(k-1)}$.

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

5: Υπολογίζουμε τον λόγο:

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

6: Αν $U_\alpha < A_\alpha$, τότε θέτουμε $\alpha^{(k)} = \alpha^{\text{can}}$. Διαφορετικά, θέτουμε $\alpha^{(k)} = \alpha^{(k-1)}$.

7: Προσομοιώνουμε $\gamma^{\text{can}} \sim U(0,1)$ και $U_\gamma \sim U(0,1)$.

8: Υπολογίζουμε τον λόγο:

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

9: Αν $U_\gamma < A_\gamma$, τότε θέτουμε $\gamma^{(k)} = \gamma^{\text{can}}$. Διαφορετικά, θέτουμε $\gamma^{(k)} = \gamma^{(k-1)}$.

β. Έστω $n_1(z)$ το πλήθος των z_i που είναι ίσα με μονάδα. Η πλήρης πιθανοφάνεια, δηλαδή η από κοινού πιθανοφάνεια των παρατηρούμενων μεταβλητών y_i και των λανθανουσών μεταβλητών z_i, δίνεται από τη σχέση:

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

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

$$
\begin{aligned}
\pi(\beta \mid \alpha, \gamma, z, y) &\propto \pi(\alpha, \beta, \gamma, z \mid y) \\
&\propto \pi(\alpha, \beta, \gamma) \cdot f(y, z \mid \alpha, \beta, \gamma) \\
&\propto \exp\left\{-\frac{1}{8}\beta^2 + \beta \sum_{i:z_i=0} x_i y_i - \alpha \sum_{i:z_i=0} e^{\beta x_i}\right\},
\end{aligned}
$$

η οποία δεν είναι κάποια γνωστή κατανομή. Παρατηρούμε ότι $\pi(\beta \mid \alpha, \gamma, z, y) = \pi(\beta \mid \alpha, z, y)$.

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

$$
\begin{aligned}
\pi(\alpha \mid \beta, \gamma, z, y) &\propto \pi(\alpha, \beta, \gamma, z \mid y) \\
&\propto \pi(\alpha, \beta, \gamma) \cdot f(y, z \mid \alpha, \beta, \gamma) \\
&\propto \alpha^9 e^{-\alpha} \cdot \alpha^{n\bar{y}} \exp\left\{-\left[\sum_{i:z_i=0} e^{\beta x_i} + n_1(z)\right] \alpha\right\} \\
&= \alpha^{n\bar{y}+9} \exp\left\{-\left[\sum_{i:z_i=0} e^{\beta x_i} + n_1(z) + 1\right] \alpha\right\},
\end{aligned}
$$

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

Δηλαδή,

$$\alpha \mid \beta, z, y \sim \text{Gamma}\left(n\overline{y} + 10, \sum_{i:z_i=0} e^{\beta x_i} + n_1(z) + 1\right),$$
$$\gamma \mid z \sim \text{Beta}\left(n_1(z) + 1, n - n_1(z) + 1\right).$$

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

$$\begin{aligned}
P(Z_i = 1 \mid y_i, \alpha, \beta, \gamma) &\propto f(y_i, z_i = 1 \mid \alpha, \beta, \gamma) \\
&= P(Z_i = 1 \mid \gamma) f(y_i \mid z_i = 1, \alpha, \beta) \\
&= \gamma f_{\text{Poisson}}(y_i \mid \alpha) \\
&= \gamma e^{-\alpha} \frac{\alpha^{y_i}}{y_i!} \\
&\propto \gamma e^{-\alpha} \alpha^{y_i},
\end{aligned}$$

$$\begin{aligned}
P(Z_i = 0 \mid y_i, \alpha, \beta, \gamma) &\propto f(y_i, z_i = 0 \mid \alpha, \beta, \gamma) \\
&= P(Z_i = 0 \mid \gamma) f(y_i \mid z_i = 0, \alpha, \beta) \\
&= (1-\gamma) f_{\text{Poisson}}\left(y_i \mid \alpha e^{\beta x_i}\right) \\
&= (1-\gamma) e^{-\alpha e^{\beta x_i}} \frac{\alpha^{y_i} e^{\beta x_i y_i}}{y_i!} \\
&\propto (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}} \alpha^{y_i}.
\end{aligned}$$

Δηλαδή,

$$P(Z_i = 1 \mid y_i, \alpha, \beta, \gamma) = \frac{\gamma e^{-\alpha} \alpha^{y_i}}{\gamma e^{-\alpha} \alpha^{y_i} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}} \alpha^{y_i}} = \frac{\gamma e^{-\alpha}}{\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}},$$

$$P(Z_i = 0 \mid y_i, \alpha, \beta, \gamma) = 1 - P(Z_i = 1 \mid y_i, \alpha, \beta, \gamma) = \frac{(1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}}{\gamma e^{-\alpha} + (1-\gamma) e^{\beta x_i y_i - \alpha e^{\beta x_i}}}.$$

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

Algorithm 3 Markov Chain Monte Carlo

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

Επαναλαμβάνουμε τα παρακάτω βήματα:

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

2: Προσομοιώνουμε $U_i \sim U(0,1)$.

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

4: Προσομοιώνουμε $\alpha^{(k)} \sim \text{Gamma}\left(n\overline{y} + 10, \sum_{i:z_i^{(k)}=0} \exp\left\{\beta^{(k-1)} x_i\right\} + n_1\left(z^{(k)}\right) + 1\right)$.

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

6: Προσομοιώνουμε $\beta^{\text{can}} \sim N\left(\beta^{(k-1)}, \sigma_\beta^2\right)$ και $U_\beta \sim U(0,1)$.

7: Υπολογίζουμε τον λόγο:

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

8: Αν $U_\beta < A_\beta$, τότε θέτουμε $\beta^{(k)} = \beta^{\text{can}}$. Διαφορετικά, θέτουμε $\beta^{(k)} = \beta^{(k-1)}$.