TP 3 Autour de l'exercice 1

* Trajectoire moyenne : $d(t) = p_0 + v_0(t - t_0)$

avec $t_0 \sim \mathcal{N}(\bar{t}_0, \sigma_{t_0})$ et $v_0 \sim \mathcal{N}(\bar{v}_0, \sigma_{v_0})$

et où on suppose que p_0, σ_{t_0} et σ_{v_0} sont fixés.

* Trajectoire individu i : $d_i(t) = d(\alpha_i(t - t_0 - \tau_i) + t_0)$

α_i = paramètre d'accélération de l'individu i p/à d
τ_i = ———— décalage temporel.

Idée : On observe un processus dont l'évolution moyenne (théorique) serait d.

On note "t_0" le début du processus moyen. Celui-ci évolue à vitesse 1.

$\leadsto$ avec α_i, τ_i, pour l'individu i : accélération en α_i
et initialisation du temps $t_0 + \tau_i$

* Observations : Pour l'individu i, on se donne k_i observations.

i.e 2 vecteurs $\begin{cases} (t_{ij})_{1 \leq j \leq k_i} & \text{tps d'observation} \\ (y_{ij})_{1 \leq j \leq k_i} & \text{données observées} \end{cases}$

et on suppose que $y_{ij} = d_i(t_{ij}) + \varepsilon_{ij}$ avec $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

* Variables cachées : $z_{pop} = (t_0, v_0)$ effets fixes, commun à l'ensemble de la population

et $z_i = (\xi_i, \tau_i)$ avec $\alpha_i = e^{\xi_i}$ effets aléatoires, modélisant la variabilité inter-individus

Rque : Pourquoi poser $\alpha_i = e^{\xi_i}$ et ne pas travailler directement avec α_i ?
(i) On est assuré de la positivité de α_i sans la contraindre (plus)
(ii) On a envie qu'en moyenne l'accélération individuelle vale 1 (sinon c'est que notre individu statistiquement moyen est mal choisi). Plus simple d'imposer $\xi_i \sim \mathcal{N}(0, *)$ que $\alpha_i \sim \mathcal{N}(1, *)$ et équivalent

On suppose donc $\xi_i \sim \mathcal{N}(0, \sigma_\xi^2)$ et $\tau_i \sim \mathcal{N}(0, \sigma_\tau)$ pour tout individu,

i.e que l'accélération et le décalage temporel sont en moyenne nuls

* Paramètres (que l'on cherche à estimer) $\begin{cases} \theta_{pop} = (\bar{t}_0, \bar{v}_0) \\ \theta_i = (\sigma_\xi, \sigma_\tau, \sigma) \end{cases}$ $\theta = (\theta_{pop}, \theta_i)$

* Cadre bayésien $\theta = (\overline{t_0}, \overline{v_0}, \sigma_\xi, \sigma_\tau, \sigma)$

On impose les a priori suivants : $\mathcal{N}$ sur les moyennes et $\mathcal{W}^{-1}$ sur les variances

i.e $\begin{cases} \overline{t_0} \sim \mathcal{N}(\overline{\overline{t_0}}, s^2_{t_0}) \; ; \; \overline{v_0} \sim \mathcal{N}(\overline{\overline{v_0}}, s^2_{v_0}) \\ \sigma_\xi \sim \mathcal{W}^{-1}(v_\xi, m_\xi) \; ; \; \sigma_\tau \sim \mathcal{W}^{-1}(v_\tau, m_\tau) \; , \; \sigma \sim \mathcal{W}^{-1}(v, m) \end{cases}$

où $\mathcal{W}^{-1}(v, m)$ est la distribution de densité (p/à la mesure de Lebesgue)

$$\forall \sigma \in \mathbb{R}_+, \quad q_{\mathcal{W}^{-1}(v,m)}(\sigma) = \left(\frac{v}{\sigma\sqrt{2}} \exp\left(-\frac{v}{2\sigma^2}\right)\right)^m \quad \text{[Wishart - Inverse]}$$

$\boxed{\text{I}}$ La log-vraissemblance du modèle

1. $q(y, z, \theta) = q(y \mid z, \theta)\, q(z_{pop} \mid \theta)\, q((z_i)_{1 \leq i \leq N} \mid \theta)\, q_{prior}(\theta)$

$$\overset{(i)}{=} \left[\prod_{i=1}^{N} q(y_i \mid z_{pop}, z_i, \theta)\, q(z_i \mid \theta) \right] q(z_{pop} \mid \theta)\, q_{prior}(\theta)$$

$$\overset{(ii)}{=} \left[\prod_{i=1}^{N} \left(\prod_{j=1}^{k_i} q(y_{ij} \mid z_{pop}, z_i, \theta) \right) q(z_i \mid \theta) \right] q(z_{pop} \mid \theta)\, q_{prior}(\theta)$$

(i) Indépendance des individus les uns p/ aux autres OK.

(ii) Indépendance des mesures de l'individu i les unes p/ aux autres
→ hypothèse simplificatrice ; faux en toute rigueur.

$$\log q(y, z, \theta) = \sum_{i=1}^{N} \sum_{j=1}^{k_i} \log q(y_{ij} \mid z_{pop}, z_i, \theta) + \sum_{i=1}^{N} \log q(z_i \mid \theta) + \log q(z_{pop} \mid \theta) + \log q_{prior}(\theta)$$

2. $y_{ij} = \underbrace{d_i(t_{ij})}_{z_{pop}, z_i, \theta} + \varepsilon_{ij}$ avec $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

donc pour tout $i \in [\![1, N]\!]$, pour tout $j \in [\![1, k_i]\!]$

$y_{ij} \mid z_{pop}, z_i, \theta \sim \mathcal{N}(d_i(t_{ij}), \sigma^2)$ [transformation affine]

$$\log q(y_{ij} \mid z_{pop}, z_i, \theta) = -\frac{1}{2}\left(\frac{y_{ij} - d_i(t_{ij})}{\sigma}\right)^2 - \log(\sigma) + \dots$$

3. $z_i = (\xi_i, \tau_i)$ avec $\xi_i \perp\!\!\!\perp \tau_i$, $\xi_i \sim \mathcal{N}(0, \sigma_\xi^2)$, $\tau_i \sim \mathcal{N}(0, \sigma_\tau^2)$

donc, pour tout $i \in [\![1, N]\!]$, $\log q(z_i \mid \theta) = -\frac{1}{2}\left(\left(\frac{\xi_i}{\sigma_\xi}\right)^2 + \left(\frac{\tau_i}{\sigma_\tau}\right)^2\right) + \dots$

4. $z_{pop} = (t_0, v_0)$ avec $t_0 \perp\!\!\!\perp v_0$, $t_0 \sim \mathcal{N}(\overline{t_0}, \sigma_{t_0}^2)$, $v_0 \sim \mathcal{N}(\overline{v_0}, \sigma_{v_0}^2)$

donc, $\log q(z_{pop} \mid \theta) = -\frac{1}{2}\left[\left(\frac{t_0 - \overline{t_0}}{\sigma_{t_0}}\right)^2 + \left(\frac{v_0 - \overline{v_0}}{\sigma_{v_0}}\right)^2\right] + \dots$

5. $\theta = (\overline{t_0}, \overline{v_0}, \sigma_\xi, \sigma_\tau, \sigma)$ $\quad \overline{t_0} \sim \mathcal{N}(\overline{\overline{t_0}}, s^2_{t_0})$ $\quad \overline{v_0} \sim \mathcal{N}(\overline{\overline{v_0}}, s^2_{v_0})$

$$\sigma_\xi \sim W^{-1}(v_\xi, m_\xi) \quad , \quad \sigma_\tau \sim W^{-1}(v_\tau, m_\tau) \quad , \quad \sigma \sim W^{-1}(v, m)$$

$$\Rightarrow \log q_{prior}(\theta) = -\frac{1}{2}\left[\left(\frac{\overline{t_0}-\overline{\overline{t_0}}}{s^2_{t_0}}\right)^2 + \left(\frac{\overline{v_0}-\overline{\overline{v_0}}}{s^2_{v_0}}\right)^2\right] + \frac{m_\xi}{2}\left[2\log\left(\frac{v_\xi}{\sigma_\xi}\right) - \left(\frac{v_\xi}{\sigma_\xi}\right)^2\right] + \frac{m_\tau}{2}\left[2\log\left(\frac{v_\tau}{\sigma_\tau}\right) - \left(\frac{v_\tau}{\sigma_\tau}\right)^2\right] + \frac{m}{2}\left[2\log\left(\frac{v}{\sigma}\right) - \left(\frac{v}{\sigma}\right)^2\right] + \dots$$

6. Finalement, en posant $k = \sum_{i=1}^{N} k_i$,

$$\begin{aligned}\log q(y, z, \theta) = & -\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{k_i}\left(y_{ij} - d_i(t_{ij})\right)^2 - kN\log(\sigma) \\ & -\frac{1}{2}\sum_{i=1}^{N}\left[\left(\frac{\xi_i}{\sigma_\xi}\right)^2 + \left(\frac{\tau_i}{\sigma_\tau}\right)^2\right] - N\log(\sigma_\xi\sigma_\tau) \\ & -\frac{1}{2}\left[\left(\frac{t_0-\overline{t_0}}{\sigma_{t_0}}\right)^2 + \left(\frac{v_0-\overline{v_0}}{\sigma_{v_0}}\right)^2\right] \\ & -\frac{1}{2}\left[\left(\frac{\overline{t_0}-\overline{\overline{t_0}}}{s_{t_0}}\right)^2 + \left(\frac{\overline{v_0}-\overline{\overline{v_0}}}{s_{v_0}}\right)^2\right] \\ & +\frac{m_\xi}{2}\left[2\log\left(\frac{v_\xi}{\sigma_\xi}\right) - \left(\frac{v_\xi}{\sigma_\xi}\right)^2\right] \\ & +\frac{m_\tau}{2}\left[2\log\left(\frac{v_\tau}{\sigma_\tau}\right) - \left(\frac{v_\tau}{\sigma_\tau}\right)^2\right] \\ & +\frac{m}{2}\left[2\log\left(\frac{v}{\sigma}\right) - \left(\frac{v}{\sigma}\right)^2\right] + \dots\end{aligned}$$

cstes qui ne changent pas lors updates.

□ θ

□ $y = (y_{ij})_{ij}$ ou $z = (z_{pop}, (z_i)_i)$

Par la suite, on va mettre en oeuvre un SAEM aussi on va exhiber les statistiques exhaustives (ou suffisantes)

Pour faire ça, on cherche les termes couplés (y, z) et θ et on regarde si on peut les mettre sous forme de produits (éventuellement scalaires) $(y, z) \times \theta$.

* Statistiques exhaustives On pose donc

$$S_1(y, z) = \frac{1}{kN}\sum_{i=1}^{N}\sum_{j=1}^{k_i}\left(y_{ij} - d_i(t_{ij})\right)^2$$

$$S_2(y, z) = \frac{1}{N}\sum_{i=1}^{N}\xi_i^2 \qquad S_3(y, z) = \frac{1}{N}\sum_{i=1}^{N}\tau_i^2$$

$$S_4(y, z) = t_0 \qquad S_5(y, z) = v_0$$

Ainsi, $$\langle S(y, z), \varphi(\theta)\rangle = S_1(y, z) \times \left(\frac{-kN}{2\sigma^2}\right) + S_2(y, z) \times \left(\frac{-N}{2\sigma_\xi^2}\right) + S_3(y, z) \times \left(\frac{-N}{2\sigma_\tau^2}\right) + S_4(y, z) \times \left(\frac{\overline{t_0}}{\sigma_{t_0}^2}\right) + S_5(y, z) \times \left(\frac{\overline{v_0}}{\sigma_{v_0}^2}\right)$$

et "on met le reste dans $\Phi(\theta)$"

II L'algorithme SAEM - Stochastic Approximation EM

On peut réécrire notre modèle sous la forme hiérarchique suivante

$$\begin{cases} y = (y_{ij})_{ij} \mid z_{pop}, (z_i)_i, \theta \sim \mathcal{N}(d_i(t_{ij}), \sigma^2) & [\text{Observé}] \\ z_{pop} \mid \theta \sim \mathcal{N}(\bar{t}_0, \sigma_{t_0}^2) \otimes \mathcal{N}(\bar{v}_0, \sigma_{v_0}^2) & [\text{Inconnu}] \\ z_i \mid \theta \sim \mathcal{N}(0, \sigma_\xi^2) \otimes \mathcal{N}(0, \sigma_\tau^2), \text{ pour tout } i & [\text{Inconnu}] \\ \oplus \text{ A priori sur } \theta : \theta \sim \mathcal{N}(\bar{\bar{t}}_0, s_{t_0}^2) \otimes \mathcal{N}(\bar{\bar{v}}_0, s_{v_0}^2) \otimes W^{-1}(m_\xi, v_\xi) \otimes \dots \end{cases}$$

On cherche à résoudre le pbm du maximum de vraissemblance pour la log-vraissemblance complète (ie avec les a priori), ou encore le pbm du maximum a posteriori (MAP)

On suppose que ce MAP existe (ce n'est pas difficile à montrer mais pas "gratuit" non plus)

* L'algorithme EM : (E) Calculer $Q_k(\theta) = \mathbb{E}_{z}\left[\log q(y, \boxed{z}, \theta) \mid y, \theta^{(k-1)}\right]$

(M) Maj $\theta^{(k)} = \underset{\theta \in \Theta}{\operatorname{argmax}}\ Q_k(\theta)$

Problème, dans le modèle proposé, on ne connais PAS z et on ne peut donc pas calculer Q_k $\rightsquigarrow$ Approximation Stochastic.

* L'algorithme SAEM : On découpe l'étape (E) en 2 sous-étapes

(E) $\begin{cases} \text{(Sim) Simuler } z^{(k)} = (z_{pop}^{(k)}, (z_i^{(k)})_i) \text{ à partir de } q(z \mid y, \theta) \\ \text{(SA) (Stochastic Approximation) } Q_k(\theta) = Q_{k-1}(\theta) + \varepsilon_k \left[\log q(y, z^{(k)}, \theta) - Q_{k-1}(\theta)\right] \\ \quad \text{où } (\varepsilon_k)_k \text{ suite} \searrow \text{ tq } \sum \varepsilon_k = +\infty, \ \sum \varepsilon_k^2 < +\infty \\ \quad [\text{ i.e qui décroît mais pas trop vite }] \end{cases}$

(M) Maj $\theta^{(k)} = \underset{\theta \in \Theta}{\operatorname{argmax}}\ Q_k(\theta)$

En pratique, pour le choix de (ε_k), on se donne un paramètre N_b et on pose : $\forall k \in \mathbb{N}^*$ $\varepsilon_k = \begin{cases} 1 & \text{si } k \in [\![1, N_b]\!] \\ (k - N_b)^{-\alpha} & \text{sinon} \end{cases}$

où $\alpha \in \left[\frac{1}{2}, 1\right[$ pour assurer la cvgce de l'algo. (Par expl $\alpha = 0{,}65$)

Puisque le modèle est exponentiel [cf page 3] i.e $\exists \Phi_1, \Phi_2, \Psi$

tq $\forall \theta \in \Theta$, $\log q(y, z, \theta) = \Phi_1(y, z) - \Phi_2(\theta) + \langle S(y, z) ; \Psi(\theta) \rangle$

on peut en fait faire porter l'Approximation Stochastique (SA) sur les statistiques exhaustives. Plus précisément, on remplace (SA) par (SA') pour chacune des statistiques exhaustives :

$$(SA') \quad S_k = S_{k-1} + \varepsilon_k \left(S(y, z^{(k)}) - S_{k-1} \right) \quad \text{avec} \quad S_0 = 0$$

Dans ce cas, $\forall \theta \in \Theta$, $Q_k(\theta) = -\Phi(\theta) + \langle S_k, \Psi(\theta) \rangle$

* Paramètres θ optimaux [cf page 3] $\theta = (\overline{t_0}, \overline{v_0}, \sigma_\xi, \sigma_\tau, \sigma)$

On a $\Phi_2(\theta) = -kN \log(\sigma) - N \log(\sigma_\xi \sigma_\tau) - \frac{1}{2}\left[\left(\frac{\overline{t_0}}{\sigma_{t_0}}\right)^2 + \left(\frac{\overline{v_0}}{\sigma_{v_0}}\right)^2 \right] + \log q_{prior}(\theta)$

$\Phi_1(y, z) = -\frac{1}{2}\left[\left(\frac{S_4}{\sigma_{t_0}}\right)^2 + \left(\frac{S_5}{\sigma_{t_0}}\right)^2 \right]$

Rque : On note $S_k = [S_1^{(k)}, S_2^{(k)} \dots]$ pour les updates

(i) $\boxed{t_0}$ $\frac{\partial Q_k}{\partial t_0}(\theta) = \frac{1}{\sigma_{t_0}^2}(S_4^{(k)} - t_0) - \frac{1}{s_{t_0}^2}(t_0 - \overline{t_0}) = 0$

$$\Rightarrow \quad t_0^{(k)} = \left(\frac{1}{\sigma_{t_0}^2} + \frac{1}{s_{t_0}^2} \right)^{-1} \left(\frac{1}{\sigma_{t_0}^2} S_4^{(k)} + \frac{1}{s_{t_0}^2} \overline{t_0} \right)$$

$\rightsquigarrow$ $t_0^{(k)}$ est le Barycentre entre $\underset{\text{model}}{\left(S_4^{(k)}, \frac{1}{\sigma_{t_0}^2}\right)}$ et $\underset{\text{a priori}}{\left(\overline{t_0}, \frac{1}{s_{t_0}^2}\right)}$

(ii) $\boxed{\overline{v_0}}$ De m, $\overline{v_0}^{(k)} = \left(\frac{1}{\sigma_{v_0}^2} + \frac{1}{s_{v_0}^2} \right)^{-1} \left(\frac{1}{\sigma_{v_0}^2} S_5^{(k)} + \frac{1}{s_{v_0}^2} \overline{v_0} \right)$ $\oplus$ Barycentre

(iii) $\boxed{\sigma_\xi}$ $\frac{\partial Q_k}{\partial \sigma_\xi}(\theta) = \frac{-N}{\sigma_\xi} + \frac{N}{\sigma_\xi} S_2 - \frac{m_\xi}{\sigma_\xi} + \frac{m_\xi v_\xi^2}{\sigma_\xi^3} = 0$

$$\Rightarrow \quad \sigma_\xi^{2\,(k)} = \frac{N S_2^{(k)} + m_\xi v_\xi^2}{N + m}$$

$\rightsquigarrow$ Idem, Barycentre entre $(S_2^{(k)}, N)$ et (v_ξ^2, m_ξ)

(iv) $\boxed{\sigma_\tau}$ De m, $\sigma_\tau^{2\,(k)} = \frac{N S_3^{(k)} + m_\tau v_\tau^2}{N + m_\tau}$ $\oplus$ Barycentre $(S_3^{(k)}, N)$ et (v_τ^2, m_τ)

(v) $\boxed{\sigma}$ $\frac{\partial Q_k}{\partial \sigma}(\theta) = -\frac{kN}{\sigma} + \frac{m v^2}{\sigma^3} - \frac{m}{\sigma} + \frac{kN}{\sigma^3} S_1^{(k)} = 0$

$$\Rightarrow \quad \sigma^{2\,(k)} = \frac{kN S_1^{(k)} + m v^2}{kN + m}$$ $\oplus$ Barycentre $(S_1^{(k)}, Nk)$ et (v^2, m)

Pour implémenter le SAEM, il ne reste plus qu'à réaliser (Sim), i.e la simulation de $z = (z_{pop}, (z_i)_i)$

III - L'algorithme de Metropolis-Hasting

On cherche à simuler $z = (z_{pop}, (z_i)_i)$. Plusieurs possibilités :

a) On simule les $2N+2$ variables en même temps ;

b) On simule les variables une par une (Gibbs "simple") ;

c) On regroupe les variables par blocs (intelligemment) et on simule les variables bloc par bloc (Gibbs bloc)

Dans tous les cas, on est obligé de recourir à Metropolis-Hasting car on n'est pas capable de simuler directement $q(z | y, \theta)$.

a) Pbm Calculs (potentiellement) compliqués du fait du "grand" nombre de variables à simuler en même temps.

b) Pbm : Chaque étape de simulation reste coûteuse, en réaliser $2N+2$ n'est donc pas forcément une bonne idée.

c) On trouve un compromis entre les 2 configurations précédentes.
Une façon naturelle de découper en blocs (de taille équitable) :
1 bloc pour la population + 1 bloc par individu, soit N blocs.

Rque : On considère une version simplifiée du modèle proposé par Jean-Baptiste Schiratti dans sa thèse aussi (b) et (c) se valent plus ou moins dans ce cas. La distinction est plus claire sur un modèle complexe.

⇝ On choisit donc comme échantillonneur "(Block) Gibbs within Metropolis-Hasting" *

et on cherche à simuler $q(z | y, \theta) = q(z_{pop}, (z_i)_i | y, \theta)$

(*) J'écris un Gibbs simple après mais pas beaucoup plus compliqué d'écrire un "par bloc"

$*$ (Bloc) "Population" : $z_{pop} = (t_0, v_0)$ et $t_0 \perp\!\!\!\perp v_0$

On introduit 2 nouveaux paramètres $\xi_{t_0}^2$ et $\xi_{p_0}^2$ qui vont être les variances pour les "proposals" dans HM.

. $\boxed{t_0}$ On choisit comme probabilité de transition $t_0^* | t_0^{(k-1)} \sim \mathcal{N}(t_0^{(k-1)}, \xi_{t_0}^2)$

Ainsi, $P(t_0^* | t_0^{(k-1)}) = \dfrac{1}{\xi_{t_0}^2 \sqrt{2\pi}} \exp\left(\dfrac{-(t_0^* - t_0^{(k-1)})^2}{2\xi_{t_0}^2}\right) = P(t_0^{(k-1)} | t_0^*)$

et on peut écrire le ratio d'acceptation plus facilement :

$$\alpha(t_0^*, t_0^{(k-1)}) = \frac{\pi(t_0^*) P(t_0^{(k-1)} | t_0^*)}{\pi(t_0^{(k-1)}) P(t_0^* | t_0^{(k-1)})} \wedge 1 = \frac{\pi(t_0^*)}{\pi(t_0^{(k-1)})} \wedge 1$$

où π désigne la loi cible : $\pi(.) = q(. | y, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i, \theta^{(k-1)})$

Or, $$\pi(t_0^*) = \frac{q(y, t_0^*, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)})}{q(y, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)})}$$

$$= \frac{q(y | t_0^*, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i, \theta^{(k-1)}) \times q(t_0^*, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)})}{q(y, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)})}$$

et de même pour $\pi(t_0^{(k-1)})$ d'où :

$$\alpha(t_0^*, t_0^{(k-1)}) = \frac{q(y | t_0^*, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)}) \times q(t_0^*, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)})}{q(y | t_0^{(k-1)}, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)}) \times q(t_0^{(k-1)}, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)})} \wedge 1$$

$$= \frac{q(y | t_0^*, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)}) \times q(t_0^* | \theta^{(k-1)})}{q(y | t_0^{(k-1)}, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)}) \times q(t_0^{(k-1)} | \theta^{(k-1)})} \wedge 1$$

. $\boxed{v_0}$ De même, $v_0^* | v_0^{(k-1)} \sim \mathcal{N}(v_0^{(k-1)}, \xi_{v_0}^2)$

$$\alpha(v_0^*, v_0^{(k-1)}) = \frac{q(y | t_0^{(k)}, v_0^*, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)}) \times q(v_0^* | \theta^{(k-1)})}{q(y | t_0^{(k)}, v_0^{(k-1)}, (\xi_i^{(k-1)})_i, (\tau_i^{(k-1)})_i | \theta^{(k-1)}) \times q(v_0^{(k-1)} | \theta^{(k-1)})} \wedge 1$$

En pratique, on considère $\log(\alpha)$ et dans MH $\Rightarrow \log(\underset{\text{uniforme}}{u}) < \log(\alpha)$

$\rightarrow \log q(y | t_0, v_0, (\xi_i)_i, (\tau_i)_i, \theta) = \dfrac{-1}{2\sigma^2} \displaystyle\sum_{i=1}^{N} \sum_{j=1}^{k_i} (y_{ij} - d_i(t_{ij}))^2$

$\rightarrow \log(t_0 | \theta) = \dfrac{-1}{2\sigma_{t_0}^2}(t_0 - \overline{t_0})^2$ où $d_i(t_{ij})$ dépend de z_{pop}...

$\rightarrow \log(v_0 | \theta) = \dfrac{-1}{2\sigma_{v_0}^2}(v_0 - \overline{v_0})^2$

* (Bloc) individu i $\forall i \in [\![1, N]\!]$, $z_i = (\xi_i, \tau_i)$

On procède comme pour z_{pop}. Soit γ_ξ^2 et γ_τ^2. Alors : [calcul similaire à z_{pop}]

- $\boxed{\xi_i}$: $\xi_i^* \mid \xi_i^{(k-1)} \sim \mathcal{N}(\xi_i^{(k-1)}, \gamma_\xi^2)$

$$\alpha(\xi_i^*, \xi_i^{(k-1)}) = \frac{q(y_i \mid t_0^{(k)}, v_0^{(k)}, \xi_i^*, \tau_i^{(k-1)}, \theta^{(k-1)}) \times q(\xi_i^{(k-1)} \mid \theta^{(k-1)})}{q(y_i \mid t_0^{(k)}, v_0^{(k)}, \xi_i^{(k-1)}, \tau_i^{(k-1)}, \theta^{(k-1)}) \times q(\xi_i^* \mid \theta^{(k-1)})} \wedge 1$$

- $\boxed{\tau_i}$ $\tau_i^* \mid \tau_i^{(k-1)} \sim \mathcal{N}(\tau_i^{(k-1)}, \gamma_\tau^2)$

$$\alpha(\tau_i^*, \tau_i^{(k-1)}) = \frac{q(y_i \mid t_0^{(k)}, v_0^{(k)}, \xi_i^{(k)}, \tau_i^*, \theta^{(k-1)}) \times q(\tau_i^{(k-1)} \mid \theta^{(k-1)})}{q(y_i \mid t_0^{(k)}, v_0^{(k)}, \xi_i^{(k)}, \tau_i^{(k-1)}, \theta^{(k-1)}) \times q(\tau_i^* \mid \theta^{(k-1)})} \wedge 1$$

Avec :
$$\to \log q(y_i \mid t_0, v_0, \xi_i, \tau_i, \theta) = \sum_{j=1}^{k_i} (y_{ij} - d_i(t_{ij}))^2$$
$$\to \log q(\xi_i \mid \theta) = \frac{-1}{2}\left(\frac{\xi_i}{\sigma_\xi}\right)^2$$
$$\to \log q(\tau_i \mid \theta) = \frac{-1}{2}\left(\frac{\tau_i}{\sigma_\tau}\right)^2$$

* Finalement l'algo s'écrit : (i) Simuler $z_{pop} \to t_0, v_0$
(ii) For $i = 1 \to N$, Simuler $z_i \to \xi_i, \tau_i$

Rque : Le choix retenu pour les proposal dans HM est particulier.
On parle de "Metropolis-Hastings Symetric Random Walk"
La symétrie de la distribution selon laquelle on propose permet la simplification dans α et ainsi α devient indep du proposal.

Bref, en recollant tous les morceaux, on est capable d'implémenter "Metropolis-Hastings within Gibbs - Stochastic Approximation EM" !