Statistique Bayésienne
Devoir Maison 1

Exercise 1

1) La loi a posteriori de τ^2 est :

$$\begin{aligned}
\pi\left(\tau^2 \mid y^{(T)}, \sigma^{(T)}, \sigma_0, \beta_0, \beta_1\right) &\propto f(y^{(T)}, \sigma^{(T)}, \sigma_0, \beta_0, \beta_1 \mid \tau^2)\pi(\tau^2) \\
&\propto \underbrace{f(y^{(T)} \mid \sigma^{(T)}, \sigma_0, \beta_0, \beta_1, \tau^2)}_{\perp\!\!\!\perp \tau^2} f(\sigma^{(T)} \mid \sigma_0, \beta_0, \beta_1, \tau^2) \underbrace{f(\sigma_0, \beta_0, \beta_1 \mid \tau^2)}_{\perp\!\!\!\perp \tau^2} \pi(\tau^2) \\
&\propto f(\sigma^{(T)} \mid \sigma_0, \beta_0, \beta_1, \tau^2)\, \pi(\tau^2) \\
&\propto \prod_{t=1}^{T} f(\underbrace{\sigma_t \mid \sigma^{(t-1)}, \sigma_0, \beta_0, \beta_1, \tau^2}_{\sim \mathcal{N}(\beta_0 + \beta_1 \sigma_{t-1}, \tau^2)})\, \pi(\tau^2) \\
&\propto \prod_{t=1}^{T} \frac{1}{|\tau|} e^{-\frac{1}{2}\frac{(\sigma_t - \beta_0 - \beta_1 \sigma_{t-1})^2}{\tau^2}} (\tau^2)^{-c_0 - 1} e^{-\frac{d_0}{\tau^2}} \\
&\propto \frac{1}{|\tau|^T} (\tau^2)^{-c_0-1} e^{-\frac{\left(d_0 + \frac{1}{2}\sum_{t=1}^{T}(\sigma_t - \beta_0 - \beta_1 \sigma_{t-1})^2\right)}{\tau^2}}
\end{aligned}$$

Comme $|\tau|^T = \sqrt{\tau^2}^T = (\tau^2)^{\frac{T}{2}}$, on a alors :

$$\boxed{\tau^2 \mid y^{(T)}, \sigma^{(T)}, \sigma_0, \beta_0, \beta_1 \sim I\Gamma\left(c_0 + \frac{T}{2},\, d_0 + \frac{1}{2}\sum_{t=1}^{T}(\sigma_t - \beta_0 - \beta_1 \sigma_{t-1})^2\right)}$$

2) La loi a posteriori de β_0 est :

$$\begin{aligned}
\pi\left(\beta_0 \mid y^{(T)}, \sigma^{(T)}, \sigma_0, \tau^2, \beta_1\right) &\propto f(y^{(T)}, \sigma^{(T)}, \sigma_0, \beta_1, \tau^2 \mid \beta_0)\pi(\beta_0) \\
&\propto \underbrace{f(y^{(T)} \mid \sigma^{(T)}, \sigma_0, \beta_0, \beta_1, \tau^2)}_{\perp\!\!\!\perp \beta_0} f(\sigma^{(T)} \mid \sigma_0, \beta_0, \beta_1, \tau^2) \underbrace{f(\sigma_0, \beta_1, \tau^2 \mid \beta_0)}_{\perp\!\!\!\perp \beta_0} \pi(\beta_0) \\
&\propto \prod_{t=1}^{T} f(\underbrace{\sigma_t \mid \sigma^{(t-1)}, \sigma_0, \beta_0, \beta_1, \tau^2}_{\sim \mathcal{N}(\beta_0 + \beta_1 \sigma_{t-1}, \tau^2)})\, \pi(\beta_0) \\
&\propto \prod_{t=1}^{T} e^{-\frac{1}{2}\frac{(\sigma_t - \beta_0 - \beta_1 \sigma_{t-1})^2}{\tau^2}} e^{-\frac{1}{2}\frac{(\beta_0 - \alpha_0)^2}{\gamma_0}} \\
&\propto e^{-\frac{1}{2}\left[\frac{1}{\tau^2}\sum_{t=1}^{T}\left(\beta_0^2 - 2\beta_0(\sigma_t - \beta_1\sigma_{t-1})\right) + \frac{\beta_0^2}{\gamma_0} - \frac{2\beta_0\alpha_0}{\gamma_0}\right]} \\
&\propto e^{-\frac{1}{2}\left[\left(\frac{T}{\tau^2} + \frac{1}{\gamma_0}\right)\beta_0^2 - 2\beta_0\left(\frac{1}{\tau^2}\sum_{t=1}^{T}(\sigma_t - \beta_1\sigma_{t-1}) + \frac{\alpha_0}{\gamma_0}\right)\right]} \\
&\propto e^{-\frac{1}{2}\left(\frac{T}{\tau^2} + \frac{1}{\gamma_0}\right)\left(\beta_0 - \frac{\frac{1}{\tau^2}\sum_{t=1}^{T}(\sigma_t - \beta_1\sigma_{t-1}) + \frac{\alpha_0}{\gamma_0}}{\frac{T}{\tau^2} + \frac{1}{\gamma_0}}\right)^2} \\
&\propto e^{-\frac{1}{2}\left(\frac{T}{\tau^2} + \frac{1}{\gamma_0}\right)\left(\beta_0 - \frac{1}{\gamma_0 T + \tau^2}\left(\gamma_0 \sum_{t=1}^{T}(\sigma_t - \beta_1\sigma_{t-1}) + \alpha_0\tau^2\right)\right)^2}
\end{aligned}$$

Ainsi,

$$\boxed{\beta_0 \mid y^{(T)}, \sigma^{(T)}, \sigma_0, \tau^2, \beta_1 \sim \mathcal{N}\left(\frac{\tau^2}{\gamma_0 T + \tau^2}\alpha_0 + \frac{\gamma_0}{\gamma_0 T + \tau^2}\sum_{t=1}^{T}\left(\sigma_t - \beta_1\sigma_{t-1}\right), \frac{\gamma_0\tau^2}{\gamma_0 T + \tau^2}\right)}$$

3) La loi a posteriori de β_1 est :

$$\begin{aligned}
\pi\left(\beta_1 \mid y^{(T)}, \sigma^{(T)}, \sigma_0, \tau^2, \beta_0\right) &\propto f(y^{(T)}, \sigma^{(T)}, \sigma_0, \beta_0, \tau^2 \mid \beta_1)\pi(\beta_1) \\
&\propto \underbrace{f(y^{(T)} \mid \sigma^{(T)}, \sigma_0, \beta_1, \beta_0, \tau^2)}_{\perp\!\!\!\perp \beta_1} f(\sigma^{(T)} \mid \sigma_0, \beta_0, \beta_1, \tau^2) \underbrace{f(\sigma_0, \beta_0, \tau^2 \mid \beta_1)}_{\perp\!\!\!\perp \beta_1} \pi(\beta_1) \\
&\propto \prod_{t=1}^{T} f(\underbrace{\sigma_t \mid \sigma^{(t-1)}, \sigma_0, \beta_0, \beta_1, \tau^2}_{\sim\mathcal{N}(\beta_0+\beta_1\sigma_{t-1}, \tau^2)})\ \pi(\beta_1) \\
&\propto \prod_{t=1}^{T} e^{-\frac{1}{2}\frac{(\sigma_t-\beta_0-\beta_1\sigma_{t-1})^2}{\tau^2}} e^{-\frac{1}{2}\frac{(\beta_1-\alpha_1)^2}{\gamma_1}} \mathbb{1}_{(-1,1)}(\beta_1) \\
&\propto e^{-\frac{1}{2}\left[\frac{1}{\tau^2}\sum_{t=1}^{T}\left(\beta_1^2\sigma_{t-1}^2 - 2\beta_1\sigma_{t-1}(\sigma_t-\beta_0)\right) + \frac{\beta_1^2}{\gamma_1} - \frac{2\beta_1\alpha_1}{\gamma_1}\right]} \mathbb{1}_{(-1,1)}(\beta_1) \\
&\propto e^{-\frac{1}{2}\left[\left(\frac{1}{\tau^2}\sum_{t=1}^{T}\sigma_{t-1}^2 + \frac{1}{\gamma_1}\right)\beta_1^2 - 2\beta_1\left(\frac{1}{\tau^2}\sum_{t=1}^{T}\sigma_{t-1}(\sigma_t-\beta_0) + \frac{\alpha_1}{\gamma_1}\right)\right]} \mathbb{1}_{(-1,1)}(\beta_1) \\
&\propto e^{-\frac{1}{2}\left(\frac{1}{\tau^2}\sum_{t=1}^{T}\sigma_{t-1}^2 + \frac{1}{\gamma_1}\right)\left(\beta_1 - \frac{\frac{1}{\tau^2}\sum_{t=1}^{T}\sigma_{t-1}(\sigma_t-\beta_0) + \frac{\alpha_1}{\gamma_1}}{\frac{1}{\tau^2}\sum_{t=1}^{T}\sigma_{t-1}^2 + \frac{1}{\gamma_1}}\right)^2} \mathbb{1}_{(-1,1)}(\beta_1) \\
&\propto e^{-\frac{1}{2}\left(\frac{1}{\tau^2}\sum_{t=1}^{T}\sigma_{t-1}^2 + \frac{1}{\gamma_1}\right)\left(\beta_1 - \frac{1}{\gamma_1\sum_{t=1}^{T}\sigma_{t-1}^2 + \tau^2}\left(\gamma_1\sum_{t=1}^{T}\sigma_{t-1}(\sigma_t-\beta_0) + \alpha_1\tau^2\right)\right)^2} \mathbb{1}_{(-1,1)}(\beta_1)
\end{aligned}$$

On obtient donc :

$$\boxed{\beta_1 \mid y^{(T)}, \sigma^{(T)}, \sigma_0, \tau^2, \beta_0 \sim \mathcal{N}\left(\frac{\gamma_1\sum_{t=1}^{T}\sigma_{t-1}\left(\sigma_t - \beta_0\right) + \alpha_1\tau^2}{\gamma_1\sum_{t=1}^{T}\sigma_{t-1}^2 + \tau^2}, \frac{\gamma_1\tau^2}{\gamma_1\sum_{t=1}^{T}\sigma_{t-1}^2 + \tau^2}\right)\mathbb{1}_{(-1,1)}(\beta_1)}$$

Exercise 2

1) La fonction de vraisemblance du modèle est :

$$\begin{aligned}\mathcal{L}(y^{(T)} \mid \gamma, \delta, \lambda) &= \prod_{t=1}^{T} f(y_t \mid y^{(t-1)}, \gamma, \delta, \lambda) \\ &= \prod_{t=1}^{\lambda} \frac{\gamma^{y_t} e^{-\gamma}}{y_t!} \prod_{t=\lambda+1}^{T} \frac{\delta^{y_t} e^{-\delta}}{y_t!}\end{aligned}$$

avec $y^{(T)}$ la même notation que dqns l'exercice 1.

Ainsi,

$$\boxed{\mathcal{L}(y^{(T)} \mid \gamma, \delta, \lambda) = \left(\prod_{t=1}^{T} \frac{1}{y_t!}\right) \gamma^{\left(\sum_{t=1}^{\lambda} y_t\right)} \delta^{\left(\sum_{t=\lambda+1}^{T} y_t\right)} e^{-\lambda\gamma-(T-\lambda)\delta}}$$

2) **L'algorithme de Gibbs**

Pour utiliser l'algorithme de Gibbs, il est nécessaire de calculer les lois conditionnelles suivantes $(\gamma \mid y^{(T)}, \delta, \lambda)$, $(\delta \mid y^{(T)}, \gamma, \lambda)$ et $(\lambda \mid y^{(T)}, \gamma, \delta)$.

D'après le théorème de Bayes, on a

$$\begin{aligned}\pi(\gamma \mid y^{(T)}, \delta, \lambda) &\propto f(y^{(T)} \mid \gamma, \delta, \lambda) \underbrace{f(\delta, \lambda \mid \gamma)}_{\perp\!\!\!\perp \gamma} \pi(\gamma) \\ &\propto \mathcal{L}(y^{(T)} \mid \gamma, \delta, \lambda) \quad \underbrace{\pi(\gamma)}_{\sim\Gamma(a1,a2)} \\ &\propto \gamma^{\left(\sum_{t=1}^{\lambda} y_t\right)} \delta^{\left(\sum_{t=\lambda+1}^{T} y_t\right)} e^{-\lambda\gamma-(T-\lambda)\delta} \gamma^{a_1-1} e^{-\frac{\gamma}{a_2}} \mathbb{1}_{\gamma\geq 0} \\ &\propto \gamma^{a_1+\left(\sum_{t=1}^{\lambda} y_t\right)-1} e^{-\gamma\left(\lambda+\frac{1}{a_2}\right)} \mathbb{1}_{\gamma\geq 0}\end{aligned}$$

On reconnait ici une loi Γ, on a donc

$$\boxed{\gamma \mid y^{(T)}, \delta, \lambda \sim \Gamma\left(a_1 + \sum_{t=1}^{\lambda} y_t, \frac{a_2}{1+\lambda a_2}\right)}$$

De même, on obtient

$$\boxed{\delta \mid y^{(T)}, \gamma, \lambda \sim \Gamma\left(d_1 + \sum_{t=\lambda+1}^{T} y_t, \frac{d_2}{1+(T-\lambda)\, d_2}\right)}$$

Quant à λ, on a

$$\begin{aligned}\pi(\lambda \mid y^{(T)}, \gamma, \delta) &\propto f(y^{(T)} \mid \gamma, \delta, \lambda) \underbrace{f(\gamma, \delta \mid \lambda)}_{\perp\!\!\!\perp \lambda} \pi(\lambda) \\ &\propto \gamma^{\left(\sum_{t=1}^{\lambda} y_t\right)} \delta^{\left(\sum_{t=\lambda+1}^{T} y_t\right)} e^{-\lambda(\gamma-\delta)} \frac{1}{T-1} \mathbb{1}_{\lambda \in [\![1\,;T-1]\!]}\end{aligned}$$

Une fois ces trois distributions définies, on peut appliquer l'algorithme de Gibbs.

On commence par initialiser les paramètres $\theta^{(0)} = (\gamma^{(0)}, \delta^{(0)}, \lambda^{(0)})$. On peut les tirer des distributions a priori.

Ensuite, on tire :

— $\gamma^{(1)}$ depuis $\pi(\gamma \mid y^{(T)}, \delta^{(0)}, \lambda^{(0)}) \sim \Gamma\left(a_1 + \sum_{t=1}^{\lambda^{(0)}} y_t, \frac{a_2}{1+\lambda^{(0)} a_2}\right)$

— $\delta^{(1)}$ depuis $\pi(\delta \mid y^{(T)}, \gamma^{(1)}, \lambda^{(0)}) \sim \Gamma\left(d_1 + \sum_{t=\lambda^{(0)}+1}^{T} y_t, \frac{d_2}{1+\left(T-\lambda^{(0)}\right) d_2}\right)$

— Et $\lambda^{(1)}$ depuis $\pi(\lambda \mid y^{(T)}, \gamma^{(1)}, \delta^{(1)}) \propto \gamma^{(1)\left(\sum_{t=1}^{\lambda} y_t\right)} \delta^{(1)\left(\sum_{t=\lambda+1}^{T} y_t\right)} e^{-\lambda(\gamma^{(1)}-\delta^{(1)})} \mathbb{1}_{\lambda \in [\![1\,;T-1]\!]}$

Finalement, on réitère cette étape avec $(\gamma^{(1)}, \delta^{(1)}, \lambda^{(1)})$, et ainsi de suite.

3) **Implémentation sur des données**

Pour cette question, l'implémentation de l'algorithme de Gibbs a été réalisée avec Python sur $n_{iter} = 5000$ itérations.

Dans un premier temps, on a pris les paramètres suivants pour les a priori :

— $a_1 = 1$

— $a_2 = 0.5$

— $d_1 = 0.5$

— $d_2 = 0.5$

On obtient les résultats suivants

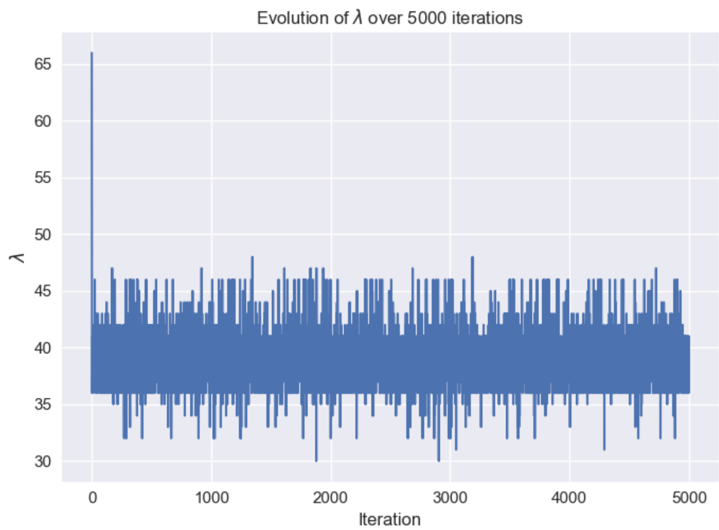


FIGURE 1 – Évolution de λ au cours des 5000 itérations, pour $a_1 = 1, a_2 = 0.5, d_1 = 0.5$ et $d_2 = 0.5$

On voit λ oscille autour de sa moyenne 39.25 avec un écart-type de 2.48.

Ensuite, on peut regarder en détails la distribution des λ :

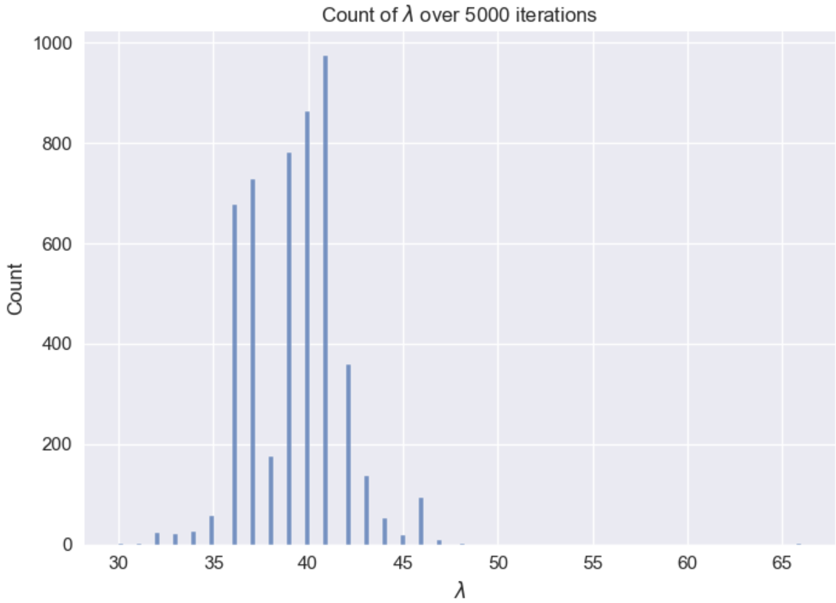
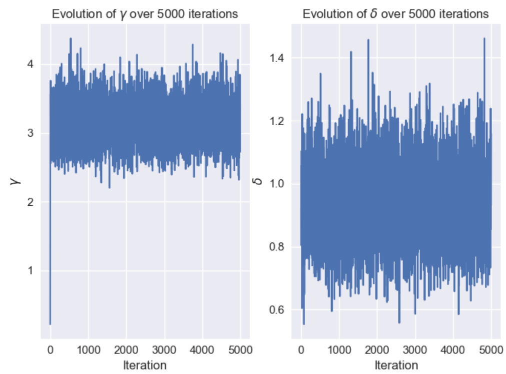
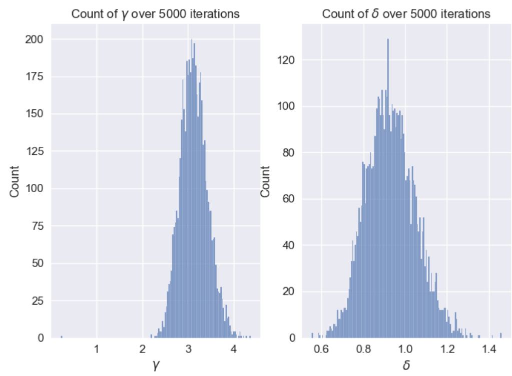


FIGURE 2 – Distribution des λ au cours des 5000 itérations, pour $a_1 = 1, a_2 = 0.5, d_1 = 0.5$ et $d_2 = 0.5$

De même, on peut regarder les distributions pour γ et δ.



(a) Évolution de γ et δ



(b) Histogrammes de γ et δ

FIGURE 3 – Graphes pour γ et δ pour 5000 itérations, $a_1 = 1, a_2 = 0.5, d_1 = 0.5$ et $d_2 = 0.5$

γ a une moyenne de 3.13 et un écart-type de 0.30. Alors que δ a une moyenne de 0.93 et un écart-type de 0.12.

Ainsi, on a donc défini le point de changement en $\lambda = 39$ si on utilise la moyenne (il vaut 40 si on prend la médiane) et on observe effectivement un changement de comportement dans les données.

De plus, on constate que le choix des a priori ne changent pas le résultat.

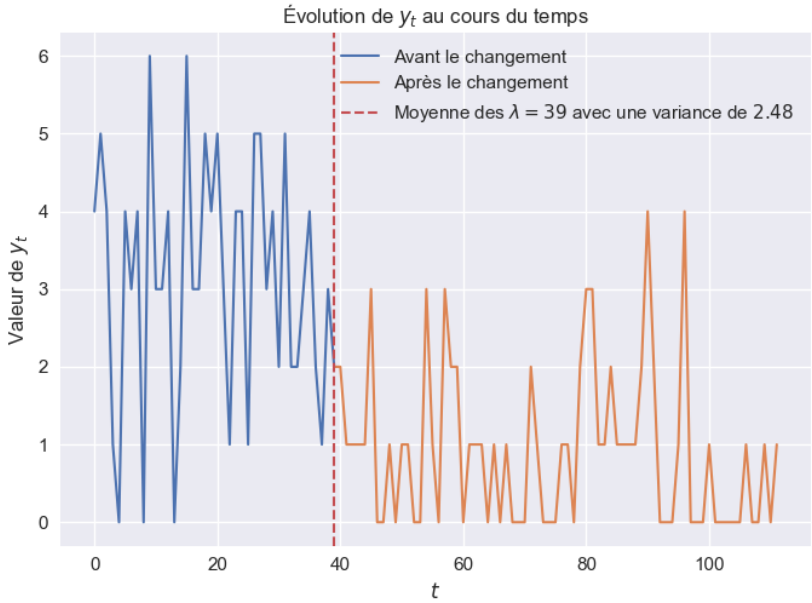


FIGURE 4 – Évolution du nombre d'accidents dans le temps séparé par le point de changement.

— Pour $a_1 = 1, a_2 = 1, d_1 = 2, d_2 = 0.5$, on a $\text{moy}(\lambda) = 39.44, \sigma(\lambda) = 2.52$ et $\text{median}(\lambda) = 40$.

— Pour $a_1 = 0.1, a_2 = 0.1, d_1 = 1, d_2 = 3$, on a $\text{moy}(\lambda) = 39.11, \sigma(\lambda) = 2.46$ et $\text{median}(\lambda) = 39$.

— Pour $a_1 = 1, a_2 = 1, d_1 = 1, d_2 = 1$, on a $\text{moy}(\lambda) = 39.33, \sigma(\lambda) = 2.49$ et $\text{median}(\lambda) = 40$.

On obtient également des résultats similaires quand on change le *seed*.

Finalement, quand on remarque que

— La moyenne des accidents avant le changement est de 3.15, avec un écart-type de 1.59.

— Puis, qu'il passe à 0.93 après le changement, avec un écart-type de 1.02.

Il y a donc bien un changement de comportement majeur après $t = 39$.

Le code est disponible sur GitHub à l'adresse suivante : https://github.com/gwatkinson/DMStatsBayesiennes