

Un individu ici (ou une observation) est le profil tabagique d'un enfant dans le temps (en  $t_0, t_1, t_2, t_3, t_4$ ) donc  $X_i = (X_i(t_0), X_i(t_1), X_i(t_2), X_i(t_3), X_i(t_4))$

Mieux encore, pour étudier la loi de  $\mathbb{P}(X_i(t_0), X_i(t_1), X_i(t_2), X_i(t_3), X_i(t_4))$   
estimer

$$\hookrightarrow = \mathbb{P}(X_i(t_4), X_i(t_3), X_i(t_2), X_i(t_1) | X_i(t_0)) \mathbb{P}(X_i(t_0))$$

↑  
Inversion  
du temps

On va étudier cette loi conditionnelle dans toute la suite.

On va voir (question A.b) que les 4 matrices  $P(1), P(2), P(3), P(4)$  suffisent pour décrire la loi conditionnelle  $\mathbb{P}(X_i(t_4), X_i(t_3), X_i(t_2), X_i(t_1) | X_i(t_0))$

un élément (une case)  $\tilde{i}, \tilde{j}$  de la matrice  $P(l)$ ,  $l = 1, 2, 3, 4$  n'est rien d'autre que  $P_{\tilde{i}, \tilde{j}}(w_l) = \mathbb{P}(X(t_l) = \tilde{j} | X(t_{l-1}) = \tilde{i})$

- Pour répondre à la question A.a, pensez à

- Que vaut la somme d'une ligne d'une matrice  $P(l)$

- Prendre quelques exemples d'états  $\tilde{i}, \tilde{j}$  dans les états

1: "n'a jamais fumé", 2: "fume actuellement", 3: "a déjà fumé, mais ne fume plus"

et décrire les transitions impossibles.  
↑  
répéter

\* Pour répondre à la question 1.b, procéder comme suit (ce n'est pas l'unique chemin à suivre)

On s'intéresse à :

$$\begin{aligned} & \mathbb{P}(X(t_4)=2, X(t_3)=2, X(t_2)=2, X(t_1)=2 \mid X(t_0)=1) \\ &= \frac{\mathbb{P}(X(t_4)=2, X(t_3)=2, X(t_2)=2, X(t_1)=2, X(t_0)=1)}{\mathbb{P}(X(t_0)=1)} \end{aligned}$$

Ensuite, écrire le numérateur sous la forme. (utiliser les probabilités conditionnelles)

$$\begin{aligned} & \mathbb{P}(X(t_4)=2, X(t_3)=2, X(t_2)=2, X(t_1)=2, X(t_0)=1) \\ &= \mathbb{P}(X(t_4)=2 \mid \dots) \mathbb{P}(X(t_3)=2 \mid \dots) \mathbb{P}(X(t_2)=2 \mid \dots) \mathbb{P}(X(t_1)=2 \mid X(t_0)=1) \\ & \quad \times \mathbb{P}(X(t_0)=1) \end{aligned}$$

- Utiliser la propriété markovienne ! pour conclure que les probabilités de transition (éléments des matrices  $P(1), P(2), P(3), P(4)$ ) suffisent

pour décrire la probabilité d'observer un profil tabagique générique

$$\mathbb{P}(X(t_4)=\text{état}_4, X(t_3)=\text{état}_3, X(t_2)=\text{état}_2, X(t_1)=\text{état}_1 \mid X(t_0)=\text{état}_0)$$

- Pour la log-vraisemblance, utiliser le lien entre les trois éléments d'une ligne d'une matrice  $P(i)$  pour réduire le nombre de paramètres.

1) Les estimateurs par maximum de vraisemblance des paramètres sont les estimateurs naturels (intuitifs).