

Processus de branchement avec mutations à la naissance.

Mathieu Richard

Laboratoire de Probabilités et Modèles Aléatoires, Paris 6



JPS 2012

- 1 Le modèle
- 2 Propriétés sans mutation
- 3 Spectre des fréquences

- 1 Le modèle
- 2 Propriétés sans mutation
- 3 Spectre des fréquences

Modèle

On considère le modèle de population suivant :

- tous les individus **évoluent indépendamment** les uns des autres et ont des **durées de vie i.i.d.**,
- au cours de sa vie, chaque individu donne **naissance à taux constant b** à des individus qui évoluent suivant la même dynamique.

Modèle

On considère le modèle de population suivant :

- tous les individus **évoluent indépendamment** les uns des autres et ont des **durées de vie i.i.d.**,
- au cours de sa vie, chaque individu donne **naissance à taux constant b** à des individus qui évoluent suivant la même dynamique.

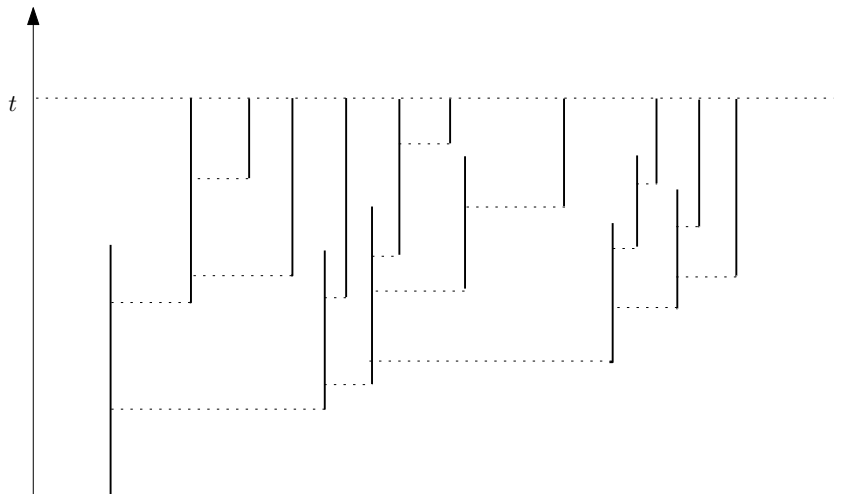


FIGURE: Un exemple d'arbre de ramification. L'axe vertical représente le temps et l'axe horizontal les relations de filiation.

Types/Allèles

- On suppose de plus que les individus portent des **types** (ou allèles).
- A chaque événement de naissance,
 - ▶ avec probabilité $(1 - p)$, un enfant est un **clone** de sa mère (*i.e.* hérite de son type).
 - ▶ avec probabilité p , un enfant est un **mutant** dont le type n'a jamais été porté auparavant (modèle à **infinité d'allèles**).
- Le modèle est **neutre** puisque aucun allèle n'est avantageé ou désavantageé.
- On appelle **famille** tous les individus partageant le même allèle.

Types/Allèles

- On suppose de plus que les individus portent des **types** (ou allèles).
- A chaque événement de naissance,
 - ▶ avec probabilité $(1 - p)$, un enfant est un **clone** de sa mère (*i.e.* hérite de son type).
 - ▶ avec probabilité p , un enfant est un **mutant** dont le type n'a jamais été porté auparavant (modèle à **infinité d'allèles**).
- Le modèle est **neutre** puisque aucun allèle n'est avantageé ou désavantageé.
- On appelle **famille** tous les individus partageant le même allèle.

Types/Allèles

- On suppose de plus que les individus portent des **types** (ou allèles).
- A chaque événement de naissance,
 - ▶ avec probabilité $(1 - p)$, un enfant est un **clone** de sa mère (*i.e.* hérite de son type).
 - ▶ avec probabilité p , un enfant est un **mutant** dont le type n'a jamais été porté auparavant (modèle à **infinité d'allèles**).
- Le modèle est **neutre** puisque aucun allèle n'est avantageé ou désavantageé.
- On appelle **famille** tous les individus partageant le même allèle.

Types/Allèles

- On suppose de plus que les individus portent des **types** (ou allèles).
- A chaque événement de naissance,
 - ▶ avec probabilité $(1 - p)$, un enfant est un **clone** de sa mère (*i.e.* hérite de son type).
 - ▶ avec probabilité p , un enfant est un **mutant** dont le type n'a jamais été porté auparavant (modèle à **infinité d'allèles**).
- Le modèle est **neutre** puisque aucun allèle n'est avantageé ou désavantageé.
- On appelle **famille** tous les individus partageant le même allèle.

Types/Allèles

- On suppose de plus que les individus portent des **types** (ou allèles).
- A chaque événement de naissance,
 - ▶ avec probabilité $(1 - p)$, un enfant est un **clone** de sa mère (*i.e.* hérite de son type).
 - ▶ avec probabilité p , un enfant est un **mutant** dont le type n'a jamais été porté auparavant (modèle à **infinité d'allèles**).
- Le modèle est **neutre** puisque aucun allèle n'est avantageé ou désavantageé.
- On appelle **famille** tous les individus partageant le même allèle.

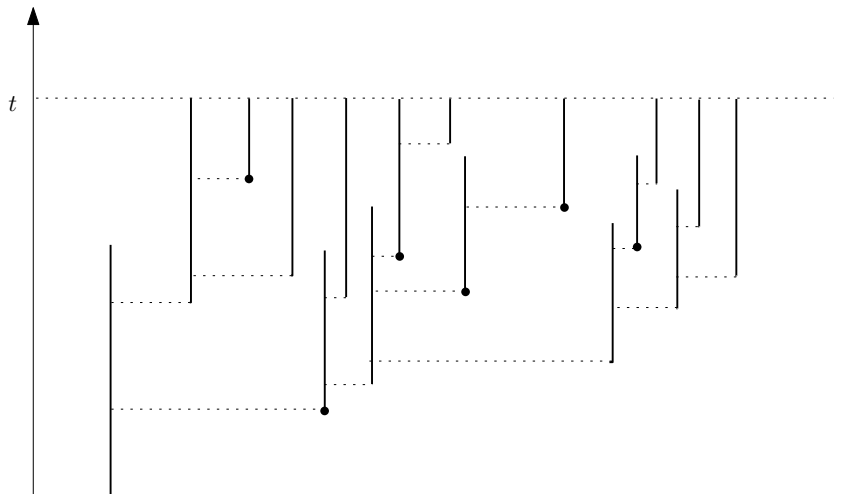


FIGURE: Un exemple d'arbre de ramification avec mutations à la naissance (représentées par des points noirs).

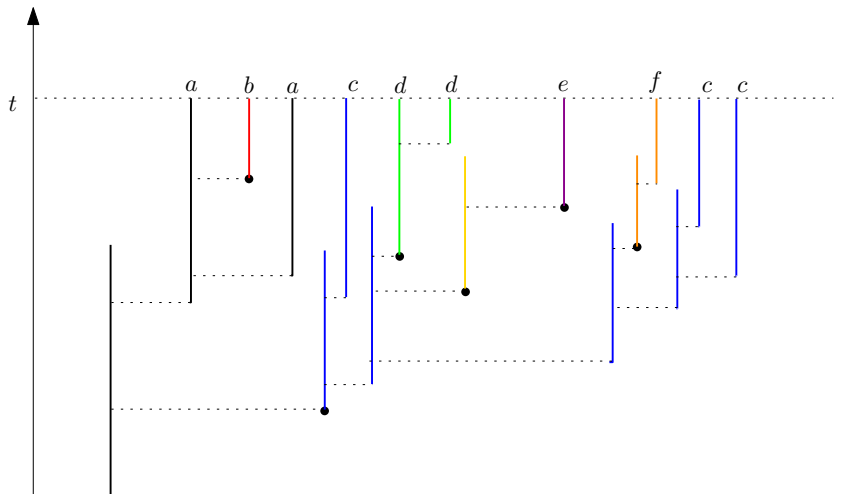


FIGURE: Les différents allèles présents au temps t . Les familles sont représentées par différentes couleurs.

Spectre des fréquences

On s'intéresse à la **partition allélique** de la population au cours du temps. Pour cela, on étudie le **spectre des fréquences** : on appelle

- M_t^i le nombre de familles ayant i représentants vivants au temps t
- $M_t^{i,a}$ le nombre de familles "d'âges" inférieurs à a et ayant i représentants au temps t .
- M_t le nombre de types présents au temps t .

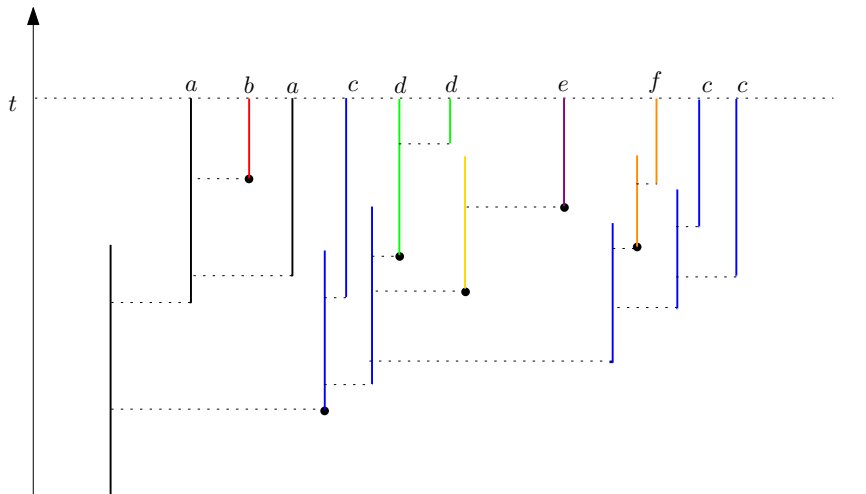


FIGURE: Le spectre des fréquences ($M_t^i, i \geq 1$) vaut $(3, 2, 1, 0, \dots)$ car trois allèles (b, e, f) sont portés par seulement un individu, a et d sont portés par deux individus et c est le seul porté par trois.

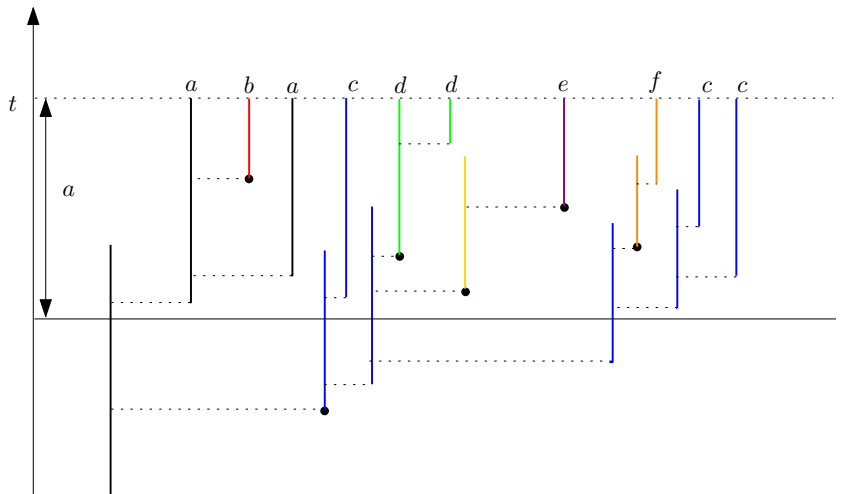


FIGURE: Le spectre des fréquences (avec âge a) ($M_t^{i,a}, i \geq 1$) vaut $(3, 1, 0, \dots)$.

- 1 Le modèle
- 2 Propriétés sans mutation**
- 3 Spectre des fréquences

- On notera $\Lambda(\cdot)/b$ la loi commune des durées de vie où Λ est une mesure positive sur $(0, \infty)$ de masse finie b .
- On se restreint à des arbres **surcritiques** :

$$m := \int_{(0, \infty)} r \Lambda(dr) > 1.$$

- Si $X(t)$ représente le nombre d'individus vivants au temps t , le processus $(X(t), t \geq 0)$ est un processus de **Crump-Mode-Jagers binaire** et *homogène*.
- Ce processus est **non markovien** sauf si $\Lambda(\cdot)/b \sim \mathcal{E}(d)$ (processus de naissance et mort "classique" avec taux de naissance b et taux de mort d).

- On notera $\Lambda(\cdot)/b$ la loi commune des durées de vie où Λ est une mesure positive sur $(0, \infty)$ de masse finie b .
- On se restreint à des arbres **surcritiques** :

$$m := \int_{(0, \infty)} r \Lambda(dr) > 1.$$

- Si $X(t)$ représente le nombre d'individus vivants au temps t , le processus $(X(t), t \geq 0)$ est un processus de **Crump-Mode-Jagers** *binaire* et *homogène*.
- Ce processus est **non markovien** sauf si $\Lambda(\cdot)/b \sim \mathcal{E}(d)$ (processus de naissance et mort "classique" avec taux de naissance b et taux de mort d).

- On notera $\Lambda(\cdot)/b$ la loi commune des durées de vie où Λ est une mesure positive sur $(0, \infty)$ de masse finie b .
- On se restreint à des arbres **surcritiques** :

$$m := \int_{(0, \infty)} r \Lambda(dr) > 1.$$

- Si $X(t)$ représente le nombre d'individus vivants au temps t , le processus $(X(t), t \geq 0)$ est un processus de **Crump-Mode-Jagers** *binaire* et *homogène*.
- Ce processus est **non markovien** sauf si $\Lambda(\cdot)/b \sim \mathcal{E}(d)$ (processus de naissance et mort "classique" avec taux de naissance b et taux de mort d).

- On notera $\Lambda(\cdot)/b$ la loi commune des durées de vie où Λ est une mesure positive sur $(0, \infty)$ de masse finie b .
- On se restreint à des arbres **surcritiques** :

$$m := \int_{(0, \infty)} r \Lambda(dr) > 1.$$

- Si $X(t)$ représente le nombre d'individus vivants au temps t , le processus $(X(t), t \geq 0)$ est un processus de **Crump-Mode-Jagers** *binaire* et *homogène*.
- Ce processus est **non markovien** sauf si $\Lambda(\cdot)/b \sim \mathcal{E}(d)$ (processus de naissance et mort "classique" avec taux de naissance b et taux de mort d).

Notations

- On note

$$\psi(\lambda) = \lambda - \int_0^\infty (1 - e^{-\lambda r}) \Lambda(dr), \quad \lambda > 0.$$

- Cette fonction est convexe et vérifie $\psi(0) = 0$: on note η sa plus grande racine.
- La fonction d'échelle W associée à ψ est telle que

$$\int_0^\infty e^{-\lambda x} W(x) dx = \frac{1}{\psi(\lambda)}, \quad \lambda > \eta.$$

- Rq : En général, on ne sait pas calculer explicitement W .

Notations

- On note

$$\psi(\lambda) = \lambda - \int_0^\infty (1 - e^{-\lambda r}) \Lambda(dr), \quad \lambda > 0.$$

- Cette fonction est convexe et vérifie $\psi(0) = 0$: on note η sa plus grande racine.
- La fonction d'échelle W associée à ψ est telle que

$$\int_0^\infty e^{-\lambda x} W(x) dx = \frac{1}{\psi(\lambda)}, \quad \lambda > \eta.$$

- Rq : En général, on ne sait pas calculer explicitement W .

Notations

- On note

$$\psi(\lambda) = \lambda - \int_0^\infty (1 - e^{-\lambda r}) \Lambda(dr), \quad \lambda > 0.$$

- Cette fonction est convexe et vérifie $\psi(0) = 0$: on note η sa plus grande racine.
- La fonction d'échelle W associée à ψ est telle que

$$\int_0^\infty e^{-\lambda x} W(x) dx = \frac{1}{\psi(\lambda)}, \quad \lambda > \eta.$$

- Rq : En général, on ne sait pas calculer explicitement W .

Notations

- On note

$$\psi(\lambda) = \lambda - \int_0^\infty (1 - e^{-\lambda r}) \Lambda(dr), \quad \lambda > 0.$$

- Cette fonction est convexe et vérifie $\psi(0) = 0$: on note η sa plus grande racine.
- La fonction d'échelle W associée à ψ est telle que

$$\int_0^\infty e^{-\lambda x} W(x) dx = \frac{1}{\psi(\lambda)}, \quad \lambda > \eta.$$

- Rq : En général, on ne sait pas calculer explicitement W .

Propriétés de X

Proposition (Lambert 2010)

- Pour $t \geq 0$, la *loi de $X(t)$* est donnée par

$$\mathbb{P}(X(t) = 0) = 1 - \frac{W'(t)}{bW(t)},$$

$$\mathbb{P}(X(t) = n) = \left(1 - \frac{1}{W(t)}\right)^{n-1} \frac{W'(t)}{bW(t)^2}, \quad n \in \mathbb{N}^*.$$

- $\mathbb{P}(\text{Ext}) = \mathbb{P}\left(X(t) \xrightarrow[t \rightarrow \infty]{} 0\right) = 1 - \eta/b$.
- If $m > 1$ (cas surcritique), $\eta > 0$, et conditionnellement à Ext^c ,

$$e^{-\eta t} X(t) \xrightarrow[t \rightarrow \infty]{p.s.} \mathcal{E}(\psi'(\eta)).$$

η est le **paramètre malthusien** de la population.

Propriétés de X

Proposition (Lambert 2010)

- Pour $t \geq 0$, la *loi de $X(t)$* est donnée par

$$\mathbb{P}(X(t) = 0) = 1 - \frac{W'(t)}{bW(t)},$$

$$\mathbb{P}(X(t) = n) = \left(1 - \frac{1}{W(t)}\right)^{n-1} \frac{W'(t)}{bW(t)^2}, \quad n \in \mathbb{N}^*.$$

- $\mathbb{P}(\text{Ext}) = \mathbb{P}\left(X(t) \xrightarrow[t \rightarrow \infty]{} 0\right) = 1 - \eta/b$.
- If $m > 1$ (cas surcritique), $\eta > 0$, et conditionnellement à Ext^c ,

$$e^{-\eta t} X(t) \xrightarrow[t \rightarrow \infty]{p.s.} \mathcal{E}(\psi'(\eta)).$$

η est le **paramètre malthusien** de la population.

Propriétés de X

Proposition (Lambert 2010)

- Pour $t \geq 0$, la *loi de $X(t)$* est donnée par

$$\mathbb{P}(X(t) = 0) = 1 - \frac{W'(t)}{bW(t)},$$

$$\mathbb{P}(X(t) = n) = \left(1 - \frac{1}{W(t)}\right)^{n-1} \frac{W'(t)}{bW(t)^2}, \quad n \in \mathbb{N}^*.$$

- $\mathbb{P}(\text{Ext}) = \mathbb{P}\left(X(t) \xrightarrow[t \rightarrow \infty]{} 0\right) = 1 - \eta/b$.
- If $m > 1$ (cas surcritique), $\eta > 0$, et conditionnellement à Ext^c ,

$$e^{-\eta t} X(t) \xrightarrow[t \rightarrow \infty]{p.s.} \mathcal{E}(\psi'(\eta)).$$

η est le **paramètre malthusien** de la population. 

Propriétés de X

Proposition (Lambert 2010)

- Pour $t \geq 0$, la *loi de $X(t)$* est donnée par

$$\mathbb{P}(X(t) = 0) = 1 - \frac{W'(t)}{bW(t)},$$

$$\mathbb{P}(X(t) = n) = \left(1 - \frac{1}{W(t)}\right)^{n-1} \frac{W'(t)}{bW(t)^2}, \quad n \in \mathbb{N}^*.$$

- $\mathbb{P}(\text{Ext}) = \mathbb{P}\left(X(t) \xrightarrow[t \rightarrow \infty]{} 0\right) = 1 - \eta/b$.
- If $m > 1$ (cas surcritique), $\eta > 0$, et conditionnellement à Ext^c ,

$$e^{-\eta t} X(t) \xrightarrow[t \rightarrow \infty]{p.s.} \mathcal{E}(\psi'(\eta)).$$

η est le *paramètre malthusien* de la population. 

Propriétés de X

Proposition (Lambert 2010)

- Pour $t \geq 0$, la *loi de $X(t)$* est donnée par

$$\mathbb{P}(X(t) = 0) = 1 - \frac{W'(t)}{bW(t)},$$

$$\mathbb{P}(X(t) = n) = \left(1 - \frac{1}{W(t)}\right)^{n-1} \frac{W'(t)}{bW(t)^2}, \quad n \in \mathbb{N}^*.$$

- $\mathbb{P}(\text{Ext}) = \mathbb{P}\left(X(t) \xrightarrow[t \rightarrow \infty]{} 0\right) = 1 - \eta/b$.
- If $m > 1$ (cas surcritique), $\eta > 0$, et conditionnellement à Ext^c ,

$$e^{-\eta t} X(t) \xrightarrow[t \rightarrow \infty]{p.s.} \mathcal{E}(\psi'(\eta)).$$

η est le **paramètre malthusien** de la population.

- 1 Le modèle
- 2 Propriétés sans mutation
- 3 Spectre des fréquences**

Processus clonal

- On appelle **processus clonal**, noté X_C , le processus décrivant la famille des individus de même type que l'ancêtre.

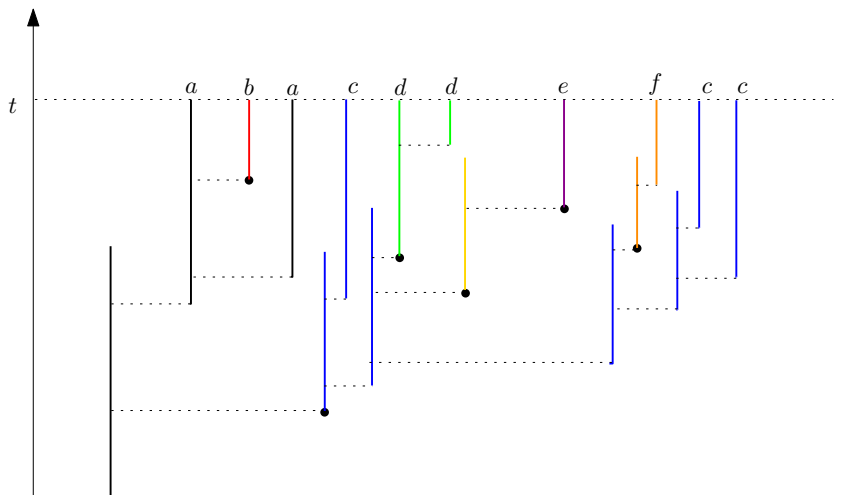


FIGURE: En noir, le processus clonal.

Processus clonal

- On appelle **processus clonal**, noté X_c , le processus décrivant la famille des individus de même type que l'ancêtre.
- Par la propriété d'amincissement des processus de Poisson, X_c se comporte comme X en "remplaçant" b par $b(1 - p)$.
- On peut alors définir son exposant de Laplace ψ_c , de plus grande racine η_c et de fonction d'échelle W_c .

Processus clonal

- On appelle **processus clonal**, noté X_c , le processus décrivant la famille des individus de même type que l'ancêtre.
- Par la propriété d'amincissement des processus de Poisson, X_c se comporte comme X en "remplaçant" b par $b(1 - p)$.
- On peut alors définir son exposant de Laplace ψ_c , de plus grande racine η_c et de fonction d'échelle W_c .

Spectre allélique moyen

- Pour $t > a \geq 0$ et $i \geq 1$, on veut calculer $\mathbb{E}[M_t^{i,a}]$.
- Soit $M_t^{i,da}$ le nombre de familles de taille i et apparues entre les dates $t - a$ et $t - a + da$.
- Conditionnellement à $X(t - a)$, $M_t^{i,da}$ est la somme de $X(t - a)$ v.a. i.i.d. distribuées comme le nombre d'atomes dans l'intervalle $[0, da]$ d'un PP de paramètre $bp\mathbb{P}(X_c(a) = i)$ (propriété de branchement).
- On a alors

$$\mathbb{E} [M_t^{i,a}] = \int_0^a \mathbb{E}[X(t - r)] \cdot bp\mathbb{P}(X_c(r) = i) dr$$

Spectre allélique moyen

- Pour $t > a \geq 0$ et $i \geq 1$, on veut calculer $\mathbb{E}[M_t^{i,a}]$.
- Soit $M_t^{i,da}$ le nombre de familles de taille i et apparues entre les dates $t - a$ et $t - a + da$.
- Conditionnellement à $X(t - a)$, $M_t^{i,da}$ est la somme de $X(t - a)$ v.a. i.i.d. distribuées comme le nombre d'atomes dans l'intervalle $[0, da]$ d'un PP de paramètre $bp\mathbb{P}(X_c(a) = i)$ (propriété de branchement).
- On a alors

$$\mathbb{E} \left[M_t^{i,a} \right] = \int_0^a \mathbb{E}[X(t - r)] \cdot bp\mathbb{P}(X_c(r) = i) dr$$

Spectre allélique moyen

- Pour $t > a \geq 0$ et $i \geq 1$, on veut calculer $\mathbb{E}[M_t^{i,a}]$.
- Soit $M_t^{i,da}$ le nombre de familles de taille i et apparues entre les dates $t - a$ et $t - a + da$.
- **Conditionnellement à $X(t - a)$** , $M_t^{i,da}$ est la somme de $X(t - a)$ v.a. i.i.d. distribuées comme le nombre d'atomes dans l'intervalle $[0, da]$ d'un PP de paramètre $bp\mathbb{P}(X_c(a) = i)$ (**propriété de branchement**).
- On a alors

$$\mathbb{E} \left[M_t^{i,a} \right] = \int_0^a \mathbb{E}[X(t - r)] \cdot bp\mathbb{P}(X_c(r) = i) dr$$

Spectre allélique moyen

- Pour $t > a \geq 0$ et $i \geq 1$, on veut calculer $\mathbb{E}[M_t^{i,a}]$.
- Soit $M_t^{i,da}$ le nombre de familles de taille i et apparues entre les dates $t - a$ et $t - a + da$.
- Conditionnellement à $X(t - a)$, $M_t^{i,da}$ est la somme de $X(t - a)$ v.a. i.i.d. distribuées comme le nombre d'atomes dans l'intervalle $[0, da]$ d'un PP de paramètre $bp\mathbb{P}(X_c(a) = i)$ (**propriété de branchement**).
- On a alors

$$\mathbb{E} [M_t^{i,a}] = \int_0^a \mathbb{E}[X(t - r)] \cdot bp\mathbb{P}(X_c(r) = i) dr$$

Spectre allélique moyen

- Pour $t > a \geq 0$ et $i \geq 1$, on veut calculer $\mathbb{E}[M_t^{i,a}]$.
- Soit $M_t^{i,da}$ le nombre de familles de taille i et apparues entre les dates $t - a$ et $t - a + da$.
- Conditionnellement à $X(t - a)$, $M_t^{i,da}$ est la somme de $X(t - a)$ v.a. i.i.d. distribuées comme le nombre d'atomes dans l'intervalle $[0, da]$ d'un PP de paramètre $bp\mathbb{P}(X_c(a) = i)$ (**propriété de branchement**).
- On a alors

$$\mathbb{E} \left[M_t^{i,a} \right] = \int_0^a \mathbb{E}[X(t - r)] \cdot bp\mathbb{P}(X_c(r) = i) dr$$

Spectre allélique moyen

Proposition (R. 11)

(i) Pour $a < t$ et $i \geq 1$,

$$\mathbb{E} \left[M_t^{i,a} \right] = \frac{p}{b(1-p)} \int_0^a W'(t-r) \left(1 - \frac{1}{W_c(r)} \right)^{i-1} \frac{W'_c(r)}{W_c^2(r)} dr$$

(ii) De plus, pour $i \geq 1$,

$$\lim_{t \rightarrow \infty} e^{-\eta t} \mathbb{E} [M_t^{i,a}] = \frac{\eta}{b} \cdot \frac{pJ^{i,a}}{(1-p)} \cdot \frac{1}{\psi'(\eta)}$$

avec

$$J^{i,a} := \int_0^a e^{-\eta r} \left(1 - \frac{1}{W_c(r)} \right)^{i-1} \frac{W'_c(r)}{W_c^2(r)} dr.$$

Spectre allélique moyen

Proposition (R. 11)

(i) Pour $a < t$ et $i \geq 1$,

$$\mathbb{E} \left[M_t^{i,a} \right] = \frac{p}{b(1-p)} \int_0^a W'(t-r) \left(1 - \frac{1}{W_c(r)} \right)^{i-1} \frac{W'_c(r)}{W_c^2(r)} dr$$

(ii) De plus, pour $i \geq 1$,

$$\lim_{t \rightarrow \infty} e^{-\eta t} \mathbb{E} [M_t^{i,a}] = \frac{\eta}{b} \cdot \frac{pJ^{i,a}}{(1-p)} \cdot \frac{1}{\psi'(\eta)}$$

avec

$$J^{i,a} := \int_0^a e^{-\eta r} \left(1 - \frac{1}{W_c(r)} \right)^{i-1} \frac{W'_c(r)}{W_c^2(r)} dr.$$

Comportement asymptotique

Proposition (R. 11)

Conditionnellement à la non-extinction de X ,

(i)

$$e^{-\eta t} M_t \xrightarrow[t \rightarrow \infty]{} \frac{p}{1-p} J \mathcal{E} \quad \text{p.s.}$$

(ii)

$$e^{-\eta t} M_t^{i,a} \xrightarrow[t \rightarrow \infty]{} \frac{p}{1-p} J^{i,a} \mathcal{E} \quad \text{p.s.}$$

(iii)

$$\frac{M_t^{i,a}}{M_t} \xrightarrow[t \rightarrow \infty]{} \frac{J^{i,a}}{J} \quad \text{p.s.}$$

où $J := \eta \int_0^\infty e^{-\eta u} \ln(W_c(u)) du$ et \mathcal{E} est une v.a. $\text{Exp}(\psi'(\eta))$.

On prouve ce résultat en utilisant la théorie des processus CMJ comptés avec caractéristiques aléatoires.

Comportement asymptotique

Proposition (R. 11)

Conditionnellement à la non-extinction de X ,

(i)

$$e^{-\eta t} M_t \xrightarrow[t \rightarrow \infty]{} \frac{p}{1-p} J \mathcal{E} \quad \text{p.s.}$$

(ii)

$$e^{-\eta t} M_t^{i,a} \xrightarrow[t \rightarrow \infty]{} \frac{p}{1-p} J^{i,a} \mathcal{E} \quad \text{p.s.}$$

(iii)

$$\frac{M_t^{i,a}}{M_t} \xrightarrow[t \rightarrow \infty]{} \frac{J^{i,a}}{J} \quad \text{p.s.}$$

où $J := \eta \int_0^\infty e^{-\eta u} \ln(W_c(u)) du$ et \mathcal{E} est une v.a. $\text{Exp}(\psi'(\eta))$.

On prouve ce résultat en utilisant la théorie des **processus CMJ comptés avec caractéristiques aléatoires**.

Perspectives

- Obtenir les lois de M_t , $M_t^{i,a}$?
- Obtenir des informations sur le nombre d'individus de type k , *i.e.* le nombre d'individus avec k mutations de plus que l'ancêtre.

Perspectives

- Obtenir les lois de M_t , $M_t^{i,a}$?
- Obtenir des informations sur le nombre d'individus de type k , *i.e.* le nombre d'individus avec k mutations de plus que l'ancêtre.