

## DATATION PAR LE CARBONE 14

*Ce texte vise à étudier le procédé de datation au carbone 14. Dans un premier temps, nous nous focalisons sur l'étude de la désintégration d'un atome isolé. Ceci nous permet de modéliser ensuite l'occurrence des phénomènes de désintégration survenant dans un organisme mort. Nous abordons dans une troisième partie le principe de la datation proprement dite.*

### 1. DÉSINTÉGRATION D'UN ATOME DE CARBONE 14

**1.1. Description de l'Atome.** D'un point de vue physique, l'atome de carbone 14 est un isotope radioactif du carbone 12. Son noyau est composé de six protons et de huit neutrons, à la différence du carbone 12 qui ne compte que six neutrons. Cet élément est présent en quantité constante dans l'atmosphère car en régénération perpétuelle sous l'action des rayons cosmiques sur les atomes d'azote présents dans l'atmosphère. Dans son livre intitulé *La physique du quotidien*, Istvan Berkès, explique ce phénomène de la façon suivante :

*“La Terre est bombardée en permanence par des particules très énergétiques venant du cosmos. Ce rayonnement cosmique est composé notamment de protons très rapides. Les noyaux des atomes présents dans l'atmosphère “explorent” littéralement sous le choc de ces protons très énergétiques et, parmi les fragments, on trouve des neutrons très rapides. Ces neutrons rapides peuvent à leur tour réagir avec des noyaux d'azote de la haute atmosphère. Lors du choc, tout se passe comme si un neutron rapide éjectait un des protons d'un des noyaux d'azote et prenait sa place pour former un noyau comptant six protons et huit neutrons. Ce noyau est un isotope particulier du carbone, le carbone 14, qui est radioactif : en émettant un électron et une particule non observable, l'antineutrino, il se décompose en un noyau d'azote comptant sept protons et sept neutrons. La période ou demi-vie du carbone 14 est 5570 ans. Comme le rayonnement cosmique bombarde la Terre depuis longtemps, un équilibre s'établit entre la création et la décomposition du carbone 14 : il y a autant de production que de décomposition si bien que la teneur en carbone 14 de tous les organismes vivants reste identique au cours du temps. Ce carbone s'oxyde en dioxyde de carbone qui se mélange à celui de l'atmosphère, à celui dissous dans l'eau, etc... et sera métabolisé par les plantes et à travers elles par tous les organismes vivants. Dans chaque gramme de carbone de l'atmosphère ou des organismes vivants, les atomes de carbone sont en très grandes majorités des atomes de carbone 12, mais, il y a  $6,8 \cdot 10^{10}$  atomes de carbone 14.”*

**1.2. Modélisation de la Désintégration.** Dans la suite de ce texte, nous ne cherchons pas à décrire outre mesure le processus de régénération du carbone 14 dans l'atmosphère ni son absorption dans la chaîne du vivant. En revanche, nous visons à modéliser précisément le phénomène de désintégration de l'atome.

Dans ce contexte, voici comment les auteurs d'un ouvrage de physique à destination du grand public explique au lecteur le phénomène radioactif :

*“La probabilité qu'un atome radioactif, présent dans une source donnée à un instant  $t$ , se désintègre durant un court intervalle de temps  $dt$  suivant cet instant  $t$  est proportionnelle à la l'intervalle  $dt$  considéré.”*

Essayons d'interpréter ce résultat en termes probabilistes. Désignons pour cela par  $T$  la variable

modélisant la durée de vie de l'atome, c'est-à-dire, la durée séparant sa création de sa désintégration. Nous supposons  $T$  définie sur un espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$  et à valeurs dans  $[0, +\infty[$ . Traduisant les propos précédents, nous sommes tentés d'écrire pour tout  $t \geq 0$  et pour tout  $h$  petit :

$$(1.1) \quad \mathbb{P}\{T > t\} > 0 \Rightarrow \mathbb{P}\{T \in ]t, t+h] | T > t\} = \lambda h + o(h),$$

pour une certaine constante  $\lambda > 0$ .

En désignant par  $F$  la fonction de répartition de la variable aléatoire  $T$ , nous déduisons de (1.1) que pour tout  $t \geq 0$  :

$$(1.2) \quad \lim_{h \rightarrow 0, h > 0} \frac{F(t+h) - F(t)}{h} = \lambda(1 - F(t)).$$

La relation (1.2) montre que la fonction de répartition est dérivable à droite et que la dérivée à droite, notée  $F'_d$  vérifie :

$$(1.3) \quad \forall t \geq 0, F'_d(t) = \lambda(1 - F(t)).$$

L'équation différentielle (1.3) s'intègre aisément sous la forme suivante :

$$(1.4) \quad \forall t \geq 0, [(F - 1) \exp(\lambda \cdot)]'_d(t) = 0.$$

Dans le cas où la fonction de répartition  $F$  est continue, le théorème des accroissements finis montre que  $F$  s'écrit :

$$(1.5) \quad \forall t \geq 0, F(t) = 1 + (F(0) - 1) \exp(-\lambda t).$$

Mais, la continuité de  $F$  implique dans ce cas  $F(0) = 0$  et donc  $\forall t \geq 0, F(t) = 1 - \exp(-\lambda t)$ . Nous en déduisons le principe suivant :

**Proposition 1.1.** *Il existe une unique loi diffuse solution du problème de modélisation (1.1). Cette loi est la loi exponentielle de paramètre  $\lambda$ .*

En revanche, si jamais  $F$  est discontinue en certains points de  $\mathbb{R}$ , rien ne garantit plus le passage de (1.4) à (1.5), de sorte qu'il n'est plus possible d'identifier la loi de  $T$  avec une distribution exponentielle de paramètre  $\lambda$ .

En fait, comme l'indique la proposition suivante, l'équation (1.1) admet dans ce cas une infinité de solutions :

**Proposition 1.2.** *Il existe une infinité de lois non diffuses solutions au problème de modélisation (1.1). Par exemple, pour tout  $C \in ]0, 1[$ , la loi de fonction de répartition  $F$  donnée par :*

$$F(0) = C, \quad \forall t > 0, F(t) = 1 + (C - 1) \exp(-\lambda t),$$

*est solution du problème (1.1).*

**Remarque 1.3.** *Une construction analogue à celle exhibée dans la proposition 1.2 permet de montrer l'existence d'une infinité de lois, solutions au problème (1.1), admettant une masse de Dirac en un point strictement positif donné.*

**1.3. Eventuelles Discontinuités de  $F$ .** Les propositions 1.1 et 1.2 soulèvent la question suivante : est-il crédible d'un point de vue pratique de supposer que la loi de  $T$  est diffuse ?

Commençons par examiner le signe de  $F(0)$ . Admettre que  $F(0)$  est strictement positif revient à supposer, qu'avec probabilité non nulle, l'atome de carbone 14 se désintègre instantanément au moment de son apparition. Cette possibilité manque pour le moins de signification d'un point de vue pratique. Comment distinguer un atome d'azote stable durant une période donnée d'un atome d'azote frappé, au cours de cette même période, par un neutron, transformé en noyau de carbone à huit neutrons, et libérant dans le même instant un électron pour redevenir azote ?

Tout se passe donc, d'un point de vue probabiliste, comme si nous pouvions supposer  $F(0) = 0$ .

Est-il finalement possible que la loi de  $T$  admette des masses de Dirac en des réels strictement positifs? Supposer que  $F$  est discontinue en certains points de  $]0, +\infty[$ , revient à distinguer, pour la loi  $\mathbb{P}_T$ , une famille au plus dénombrable de réels strictement positifs, ou encore à privilégier pour le processus de désintégration une famille au plus dénombrable d'instants non nuls.

D'un point de vue purement pratique, le seul instant pour lequel un traitement particulier pourrait être réservé serait éventuellement l'instant 0. Mais, rien n'indique dans la description physique de la désintégration donnée dans la Sous-section 1.2 l'existence de points singuliers non nuls, différenciés des autres, pour le processus de désintégration.

Autrement dit, tout conduit à accepter le caractère diffus de la loi de  $T$ .

**1.4. Loi Exponentielle.** Des discussions précédentes, nous affirmons :

**Théorème 1.4.** *La loi de modélisation de la durée de vie d'un atome de Carbone 14 est une loi exponentielle.*

Mais, il est bien connu que la loi exponentielle caractérise les phénomènes sans mémoire :

**Proposition 1.5.** *Étant donnée une v.a.  $S$  à valeurs dans  $[0, +\infty[$ , vérifiant  $\mathbb{P}\{S > 0\} > 0$ , et :*

$$(1.6) \quad \forall t > 0, s > 0, \mathbb{P}\{S > t\} > 0 \Rightarrow \mathbb{P}\{S > t + s | S > t\} = \mathbb{P}\{S > s\},$$

*$S$  suit une loi exponentielle.*

*Réciproquement, toute v.a.  $S$  de loi exponentielle vérifie (1.6).*

Autrement dit, la modélisation retenue de la durée de vie d'un atome de carbone 14 conduit, et même revient, à supposer qu'un atome de carbone 14 ne vieillit pas.

Dans ce contexte, il est assez tentant de revenir à l'explication physique donnée dans la Sous-section (1.2). Bien entendu, la constante  $\lambda$  suggère l'absence de mémoire du noyau de l'atome.

Néanmoins, nous savons que les solutions de (1.1) peuvent ne pas être de type exponentiel (cf. Proposition 1.2). La description (1.1) n'est donc pas caractéristique du non-vieillessement.

L'obstacle essentiel tient au  $o(h)$ . Dit crûment, le  $o(h)$  pourrait très bien vieillir !

Une façon très forte de contredire cette possibilité consiste à supposer, en remplacement de (1.1), qu'il existe une application  $\varepsilon$  de  $\mathbb{R}_+$  dans  $\mathbb{R}$ , valant 0 en 0, et continue en 0, telle que :

$$(1.7) \quad \forall t, h > 0, \mathbb{P}\{T > t\} > 0 \Rightarrow \mathbb{P}\{T \in ]t, t + h] | T > t\} = \lambda h + h\varepsilon(h),$$

Dans ce cas, nous affirmons en effet :

**Proposition 1.6.** *Supposons que  $F(0) = 0$  et que la condition (1.7) soit satisfaite, alors  $F$  suit une loi exponentielle de paramètre  $\lambda$ .*

Il est possible d'affaiblir la condition (1.7). Essentiellement, nous pouvons modéliser le "non-vieillessement de  $o(h)$ " en supposant qu'il existe une application  $\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , valant 0 en 0, et continue en 0, contrôlant uniformément le  $o(h)$  :

**Proposition 1.7.** *Supposons que  $F(0) = 0$  et que la condition suivante soit vérifiée :*

$$\begin{aligned} \forall \varepsilon > 0, \exists \delta > 0, \forall t, h > 0, h < \delta, \mathbb{P}\{T > t\} > 0 \\ \Rightarrow |\mathbb{P}\{T \in ]t, t + h] | T > t\} - \lambda h| \leq \varepsilon h. \end{aligned}$$

*Alors,  $F$  suit une loi exponentielle de paramètre  $\lambda$ .*

1.5. **Demi-Vie.** D'un point de vue probabiliste, la période ou demi-vie de l'atome n'est rien d'autre que la médiane de la loi de  $T$ . Un simple calcul montre que la médiane de la loi exponentielle de paramètre  $\lambda$  est :

$$\tau = \frac{\ln(2)}{\lambda}.$$

En particulier, en retenant que la période de l'atome de carbone 14 est 5570 ans, on en déduit que la constante  $\lambda$  vaut :

$$\lambda = 3,94 \cdot 10^{-12} \text{ s}^{-1}.$$

Le paramètre  $\lambda$  est alors appelé *constante de radioactivité*.

## 2. SUPERPOSITION DE PHÉNOMÈNES SANS VIEILLISSEMENT

La description faite par I. Berkès de la régénération perpétuelle du carbone 14 dans l'atmosphère et de son incorporation dans la chaîne du vivant nous apprend que la teneur en carbone 14 dans les organismes vivants est constante, égale à  $6,8 \cdot 10^{10}$  atomes de carbone 14 par gramme de carbone.

A l'opposée, le métabolisme disparaissant, le carbone 14 présent dans un organisme mort n'est plus renouvelé et se désintègre petit à petit en atomes d'azote.

Dans cette perspective, cette partie vise à étudier le processus stochastique décrivant les désintégrations successives de carbone 14 survenant dans une masse macroscopique de matière devenue inerte. L'organisme mort contient alors un très grand nombre de particules de carbone 14, et nous ferons l'hypothèse, raisonnable, qu'elles se désintègrent indépendamment les unes des autres.

La loi de modélisation de la durée de vie d'un atome de carbone 14 étant, d'après la première partie, de type exponentiel, tout se passe comme si tous les atomes de carbone 14 présents au moment de la mort de l'organisme étaient créés à cet instant même.

Dans ce contexte, nous allons démontrer le principe suivant :

**Théorème 2.1.** *Les désintégrations successives d'atomes de carbone 14 survenant dans un organisme mort forment un processus ponctuel assimilable à un processus de Poisson.*

Pour démontrer le Théorème 2.1, nous visons à étudier, sans référence explicite au cas du carbone 14, la nature du processus induit par la superposition de phénomènes sans vieillissement. L'application à notre modèle n'est donnée que dans la suite du texte.

**2.1. Résultats Théoriques.** Commençons par rappeler un résultat bien connu sur la statistique d'ordre d'un échantillon de loi exponentielle :

**Proposition 2.2.** *Soient  $n \geq 1$ , et  $\tau_1, \dots, \tau_n$ ,  $n$  variables aléatoires indépendantes de même loi exponentielle de paramètre  $\lambda > 0$ . Alors, en appelant  $(\tau_{(1)}, \dots, \tau_{(n)})$  la statistique d'ordre de  $(\tau_1, \dots, \tau_n)$ , les variables  $\tau_{(1)}, \tau_{(2)} - \tau_{(1)}, \dots, \tau_{(n)} - \tau_{(n-1)}$  sont indépendantes et suivent des lois exponentielles dont les paramètres sont respectivement  $n\lambda, (n-1)\lambda, \dots, \lambda$ .*

Étant donnés  $N \geq 1$  et  $(\tau_j^N)_{1 \leq j \leq N}$  une famille de v.a.i.id de loi exponentielle de paramètre  $\lambda/N$ , la proposition 2.2 nous permet d'affirmer que :

**Lemme 2.3.** *Pour tout  $k \geq 1$  et toute fonction  $f$  borélienne bornée de  $\mathbb{R}^k$  dans  $\mathbb{R}$  :*

$$\lim_{N \rightarrow +\infty} \mathbb{E}f(\tau_{(1)}^N, \dots, \tau_{(k)}^N - \tau_{(k-1)}^N) = \mathbb{E}f(\tau_1, \dots, \tau_k),$$

où  $\tau_1, \dots, \tau_k$  sont  $k$  v.a.i.id de loi exponentielle de paramètre  $\lambda$ .

**Preuve.** Il suffit de voir que la densité du vecteur  $(\tau_{(1)}^N, \dots, \tau_{(k)}^N - \tau_{(k-1)}^N)$  est donnée, d'après la Proposition 2.2, par :

$$(x_1, \dots, x_k) \in \mathbb{R}^k \mapsto \lambda^k \prod_{i=0}^{k-1} \left(1 - \frac{i}{N}\right) \prod_{i=1}^k \mathbf{1}_{]0, +\infty[}(x_i) \exp\left(-\lambda \sum_{i=1}^k \left(1 - \frac{i-1}{N}\right) x_i\right).$$

Il en résulte que :

$$\begin{aligned} \mathbb{E}f(\tau_{(1)}^N, \dots, \tau_{(k)}^N - \tau_{(k-1)}^N) &= \lambda^k \prod_{i=0}^{k-1} \left(1 - \frac{i}{N}\right) \int_{\mathbb{R}^k} f(x_1, \dots, x_k) \\ &\quad \times \prod_{i=1}^k \mathbf{1}_{]0, +\infty[}(x_i) \exp\left(-\lambda \sum_{i=1}^k \left(1 - \frac{i-1}{N}\right) x_i\right) dx_1 \dots dx_k. \end{aligned}$$

Le théorème de Convergence dominée permet aisément de conclure (en minorant  $1 - k/N$  par  $1/2$  pour  $N$  grand).  $\square$

En désignant par  $(T_j)_{j \geq 1}$  les instants de saut d'un processus de Poisson de paramètre  $\lambda$ , nous en déduisons la proposition suivante :

**Proposition 2.4.** *Pour tout  $k \geq 1$  et toute fonction  $f$  borélienne bornée de  $\mathbb{R}^k$  dans  $\mathbb{R}$  :*

$$\lim_{N \rightarrow +\infty} \mathbb{E}f(\tau_{(1)}^N, \dots, \tau_{(k)}^N) = \mathbb{E}f(T_1, \dots, T_k).$$

**Preuve.** Il suffit de remarquer que les variables aléatoires  $T_1, T_2 - T_1, \dots, T_k - T_{k-1}$  sont indépendantes de loi exponentielle de paramètre  $\lambda$  et d'appliquer le Lemme 2.3.  $\square$

On en déduit le théorème suivant :

**Théorème 2.5.** *Étant donné  $\lambda > 0$ ,  $N \geq 1$  et une famille  $(\tau_j^N)_{1 \leq j \leq N}$  de v.a.i.i.d de loi exponentielle de paramètre  $\lambda/N$ , la statistique d'ordre  $(\tau_{(1)}^N, \dots, \tau_{(N)}^N)$  vérifie :*

$$(\tau_{(1)}^N, \dots, \tau_{(k)}^N) \xrightarrow[\text{Loi}]{} (T_1, T_2, \dots, T_k), \text{ lorsque } N \rightarrow +\infty,$$

où  $(T_j)_{j \geq 1}$  désigne les instants de saut d'un processus de Poisson de paramètre  $\lambda$ .

**2.2. Interprétation du Théorème 2.5.** Supposons qu'un grand nombre de phénomènes surviennent indépendamment les uns des autres au bout de temps régis par des lois exponentielles de même paramètre supposé très petit (i.e. au bout de temps de très grande espérance). Alors, la suite des instants d'occurrence de ces phénomènes est assimilable à une suite d'instants de saut d'un processus de Poisson.

Cette idée peut être confirmée en introduisant le nombre de phénomènes survenus entre 0 et  $t$ , i.e. en définissant la variable  $S_t^N$  par :

$$S_t^N = \sum_{k=1}^N \mathbf{1}_{[0, t]}(\tau_k^N) = \sum_{k=1}^N \mathbf{1}_{[0, t]}(\tau_{(k)}^N).$$

Alors, pour tout  $n \geq 1$ ,  $0 < t_1 < \dots < t_n$ , et  $0 \leq k_1 \leq \dots \leq k_n \leq N - 1$  :

$$\mathbb{P}\{S_{t_1}^N = k_1, \dots, S_{t_n}^N = k_n\} = \mathbb{P}\{\tau_{(k_1)}^N \leq t_1 < \tau_{(k_1+1)}^N, \dots, \tau_{(k_n)}^N \leq t_n < \tau_{(k_n+1)}^N\},$$

avec la convention  $\tau_{(0)}^N = 0$ .

Par application de la Proposition 2.4, il est alors clair que :

$$\begin{aligned} \lim_{N \rightarrow +\infty} \mathbb{P}\{S_{t_1}^N = k_1, \dots, S_{t_n}^N = k_n\} &= \mathbb{P}\{T_{k_1} \leq t_1 < T_{k_1+1}, \dots, T_{k_n} \leq t_n < T_{k_n+1}\} \\ &= \mathbb{P}\{N_{t_1} = k_1, \dots, N_{t_n} = k_n\}, \end{aligned}$$

où  $(N_t)_{t>0}$  désigne un processus de Poisson de paramètre  $\lambda > 0$ .

Autrement dit, pour  $N$  très grand, le processus  $(S_t^N)_{t \geq 0}$  modélisant le nombre de phénomènes survenus entre 0 et  $t$  est assimilable à un processus de Poisson.

**2.3. Application Pratique.** Dans la pratique  $N$  est choisi très grand, mais n'est jamais infini ! En particulier, les passages à la limite dans les égalités précédentes sont en réalité des approximations : celles-ci sont d'autant plus valables que  $N$  est grand par rapport à l'indice  $k$  de la variable  $\tau_{(k)}^N$ . A l'inverse, si l'on vient à étudier des quantités définies à partir de variables  $\tau_{(k)}^N$  d'indices élevés, l'approximation risque d'être mauvaise.

Ce phénomène risque de se produire dans le cas d'approximations de la forme :

$$\mathbb{P}\{S_{t_1}^N = k_1, \dots, S_{t_n}^N = k_n\} \sim \mathbb{P}\{N_{t_1} = k_1, \dots, N_{t_n} = k_n\},$$

s'appuyant sur des grandes valeurs de  $k_1, k_2, \dots, k_n$ .

De fait, si  $t_n$  est petit, le vecteur  $(N_{t_1}, \dots, N_{t_n})$  charge essentiellement les petites valeurs de  $k_1, \dots, k_n$ , c'est-à-dire les valeurs assurant une bonne qualité de l'approximation. Dit autrement, les deux probabilités se ressemblent sur des événements dont la probabilité est élevée sous la loi limite, et par conséquent, dont la probabilité est également élevée sous la loi à approcher. Il est alors vraisemblable que la réalisation du vecteur  $(S_{t_1}^N, \dots, S_{t_n}^N)$  soit assimilable à une réalisation du vecteur  $(N_{t_1}, \dots, N_{t_n})$ .

A l'inverse, si  $t_n$  est grand, le vecteur  $(N_{t_1}, \dots, N_{t_n})$  charge avec une probabilité non négligeable les grandes valeurs de  $k_n$ , valeurs pour lesquelles l'approximation risque d'être mauvaise. Il est alors à craindre que la réalisation du vecteur  $(S_{t_1}^N, \dots, S_{t_n}^N)$  ne s'éloigne d'une réalisation typique du vecteur  $(N_{t_1}, \dots, N_{t_n})$ .

Nous faisons donc la conclusion suivante : en pratique, l'approximation par le processus de Poisson est essentiellement valable pour des valeurs de temps raisonnables. Nous ne détaillerons pas ici la signification mathématique du mot raisonnable et ne chercherons en particulier pas à déterminer jusqu'à quel instant l'approximation peut être considérée comme valable.

**2.4. Visualisation Informatique.** Nous visons maintenant à visualiser une partie de ce phénomène à l'aide de l'outil informatique. Pour cela, nous comptons procéder selon l'algorithme suivant :

**Algorithme 2.6.** *L'algorithme suivant permet de tracer la loi empirique d'un 1000-échantillon de loi  $S_1^N$  :*

- (1) *Ecrire une fonction desintegration( $N, \lambda$ ) simulant la réalisation d'une v.a. de loi  $S_1^N$ . Pour cela, on pourra simuler un  $N$ -échantillon de loi exponentielle de paramètre  $\lambda/N$  et compter les réalisations inférieures à 1.*
- (2) *Ecrire une fonction echantillon( $N, \lambda$ ) simulant la réalisation d'un 1000-échantillon de loi  $S_1^N$ . Pour cela, on pourra exécuter 1000 fois la fonction desintegration.*
- (3) *Tracer l'histogramme des valeurs obtenues pour  $N = 10^4$  et  $\lambda = 1$ .*
- (4) *Comparer l'histogramme avec les poids de la loi de Poisson de paramètre 1.*

La figure 1 en Annexe détaille les résultats du point 4.

**2.5. Application à la Désintégration du Carbone 14.** Rappelons que la constante de radioactivité, telle que nous l'avons calculée dans la première partie, est donnée par  $\lambda = 3.92 \cdot 10^{-12} \text{ s}^{-1}$ . Rappelons également, d'après le texte d'I. Berkès, qu'à la mort d'un organisme, on compte dans un gramme de carbone (contenant les isotopes 12, 13 et 14 du carbone) environ  $N = 6,8 \cdot 10^{10}$  atomes de carbone 14. Dans ce contexte, la constante  $\lambda N$  est donc de l'ordre de  $0,27 \text{ s}^{-1}$ .

En conclusion, les désintégrations des atomes de carbone 14 survenant à la suite de la mort d'un organisme vivant forment un processus assimilable, à l'échelle humaine, à un processus de Poisson de paramètre 0,27.

### 3. DATATION AU CARBONE 14

Nous abordons dans cette dernière partie le principe de datation au carbone 14. Dans le manuel de physique cité au début de ce texte, son fonctionnement est résumé de la façon suivante :

*“ Dater un échantillon par le radiocarbone consiste à mesurer sa teneur actuelle en carbone 14 et à la comparer à celle qu'il avait lors de sa formation. ”*

**3.1. Datation par Loi des Grands Nombres.** Étant donné un gramme de matière d'origine organique, imaginons que les physiciens et les chimistes soient capables d'en déterminer avec précision la teneur en atomes de carbone 14. Désignons alors par  $N$  le nombre estimé d'atomes dans le morceau en question et appelons  $t$  la durée nous séparant de sa mort.

A la mort de l'organisme, référencée à partir de maintenant comme l'instant 0 (de sorte que  $t$  désigne également l'instant présent), nous savons que le gramme de carbone comptait environ  $N_0 = 6,8 \cdot 10^{10}$  atomes de carbone 14. La mesure réalisée par les physiciens nous indique que  $N_0 - N$  particules se sont alors désintégrées.

En supposant une fois encore que les particules de carbone 14 au sein du gramme étudié se désintègrent indépendamment les unes des autres, nous pouvons écrire :

$$(3.8) \quad N_0 - N = \sum_{i=1}^{N_0} \mathbf{1}_{\{r_i \leq t\}},$$

où  $(r_i)_{1 \leq i \leq N_0}$  est une famille de v.a. indépendantes de loi exponentielle de paramètre  $\lambda$  modélisant les différents instants de désintégration des atomes présents dans le gramme d'origine organique.

Mais, en divisant chaque membre de (3.8) par  $N_0$ , nous déduisons que :

$$\frac{1}{N_0} \sum_{i=1}^{N_0} \mathbf{1}_{\{r_i \leq t\}} = 1 - \frac{N}{N_0}.$$

Le membre de gauche apparaît de fait comme une moyenne empirique de  $N_0$  v.a. indépendantes de loi de Bernoulli de paramètre  $(1 - \exp(-\lambda t))$ .

La constante  $N_0$  étant grande, la loi des grands nombres nous permet alors d'affirmer :

$$(3.9) \quad \exp(-\lambda t) \approx \frac{N}{N_0}.$$

Soit encore :

$$(3.10) \quad t \approx \frac{\ln(N_0) - \ln(N)}{\lambda}.$$

**3.2. Détermination de la Teneur.** Se pose en pratique la question de la détermination de la teneur en carbone 14 du morceau de matière d'origine organique. Dans la même référence que précédemment, le mode expérimental est décrit de la façon suivante :

*“ La mesure de la teneur en carbone 14 actuelle de l'échantillon est évaluée par comptage des désintégrations sur une période de trois jours. ”*

Rappelons dans ce contexte que  $N$  désigne le nombre d'atomes de carbone 14 actuellement présents

dans l'échantillon. Le nombre  $n$  de désintégrations ayant lieu entre l'instant  $t$  et l'instant  $t + \delta$  ( $\delta = 3$  jours selon les conditions expérimentales précédentes) est donné par :

$$n \equiv \sum_{i=1}^{N_0} \mathbf{1}_{\{t < r_i \leq t + \delta\}}.$$

En utilisant le caractère sans mémoire de la loi exponentielle, il est possible de montrer que le couple  $(N, n)$  suit la même loi que le couple :

$$\left( N, \sum_{i=1}^N \mathbf{1}_{\{\rho_i \leq \delta\}} \right),$$

où  $(\rho_i)_{i \geq 1}$  désigne une famille de v.a.i.id de loi exponentielle de paramètre  $\lambda$  indépendante de la v.a.  $N$ . De sorte que tout se passe comme si nous étions ramenés au paragraphe précédent avec pour situation initiale  $N$  atomes au lieu de  $N_0$ . De fait, un raisonnement heuristique fondé sur la loi des grands nombres assure que :

$$(3.11) \quad \frac{n}{N} \approx 1 - \exp(-\lambda\delta),$$

soit encore  $N \approx n(1 - \exp(-\lambda\delta))^{-1}$ .

**3.3. Limites de la Méthode.** Pour que la loi des grands nombres puisse être appliquée au contexte précédent, et que l'approximation ci-dessus soit acceptable, il est nécessaire que  $N$  soit suffisamment grand.

Nous cherchons désormais à quantifier cette restriction théorique, afin d'exhiber un seuil de validité de la méthode de datation.

Commençons par rappeler que le théorème central limite assure que :

$$\lim_{m \rightarrow +\infty} \mathbb{P} \left\{ \left| \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{\{\rho_i \leq \delta\}} - (1 - \exp(-\lambda\delta)) \right| \geq 2m^{-1/2}\sigma \right\} < 0,05,$$

où  $\sigma^2$  désigne la variance d'une loi de Bernoulli de paramètre  $(1 - \exp(-\lambda\delta))$ .

Autrement dit, à partir d'un certain rang  $m_0$  déterministe ( $m_0 \approx 1000(\lambda\delta)^{-1}$  par application du théorème de Berry-Essen), pour tout  $m \geq m_0$  :

$$\mathbb{P} \left\{ \left| \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{\{\rho_i \leq \delta\}} - (1 - \exp(-\lambda\delta)) \right| \geq 2m^{-1/2}\sigma \right\} < 0,1.$$

De la condition  $N$  grand, nous formulons l'hypothèse que  $\mathbb{P}\{N \geq m_0\}$  est très proche de 1. De fait, au regard de l'indépendance de la v.a.  $N$  et de la suite  $(\rho_i)_{i \geq 1}$ , il est raisonnable d'admettre l'inégalité suivante :

$$(3.12) \quad \mathbb{P} \left\{ \left| \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{\rho_i \leq \delta\}} - (1 - \exp(-\lambda\delta)) \right| \geq 2N^{-1/2}\sigma \right\} \lesssim 0,1,$$

soit encore :

$$(3.13) \quad \mathbb{P} \left\{ \left| \frac{n}{N} - (1 - \exp(-\lambda\delta)) \right| \geq 2N^{-1/2}\sigma \right\} \lesssim 0,1.$$

Désignons par  $\mathcal{E} \equiv N^{-1}|N - n(1 - \exp(-\lambda\delta))^{-1}|$  l'erreur relative entre  $N$  et son estimateur. L'inégalité (3.13) s'écrit :

$$(3.14) \quad \mathbb{P} \left\{ \mathcal{E} \geq 2N^{-1/2}\sigma(1 - \exp(-\lambda\delta))^{-1} \right\} \lesssim 0,1.$$

Rappelons que la variance  $\sigma^2$  s'écrit :

$$\sigma^2 = \exp(-\lambda\delta)(1 - \exp(-\lambda\delta)) \approx \lambda\delta.$$

De fait, il est acceptable de déduire de (3.14) la nouvelle inégalité :

$$(3.15) \quad \mathbb{P}\{\mathcal{E} \geq 2(\lambda\delta N)^{-1/2}\} \lesssim 0, 1.$$

De (3.9), nous tirons par ailleurs que la probabilité que  $N$  diffère de  $N_0 \exp(-\lambda t)$  de façon excessive est très petite. Ceci nous conduit, au regard de (3.15), à accepter l'approximation suivante :

$$(3.16) \quad \mathcal{E} \approx 2(N_0\lambda\delta \exp(-\lambda t))^{-1/2} = 2(N_0\lambda\delta)^{-1/2} \exp(\lambda t/2).$$

Aussi, pour obtenir une erreur relative d'au plus 5% (une telle erreur induisant une erreur de datation de l'ordre de 500 ans), la durée  $t$  doit être au plus de l'ordre de :

$$t_{\max} \equiv \frac{2}{\lambda} \left[ \frac{1}{2} \ln(N_0\lambda\delta) - \ln(40) \right].$$

Pour  $\delta = 3$  jours, on trouve  $t_{\max} \approx 30\,000$  ans, et pour  $\delta = 10$  jours  $t_{\max} \approx 40\,000$  ans.

Le  $N_{\max}$  associé (indexé par max, mais correspondant en réalité à un minimum) est de l'ordre de  $N_0 \exp(-\lambda t_{\max}) = 1600(\lambda\delta)^{-1}$ . Pour  $\delta = 3$  jours,  $N_{\max} \approx 1,6 \cdot 10^9$ , pour  $\delta = 10$  jours,  $N_{\max} \approx 4,7 \cdot 10^8$ . Dans ces deux cas, l'approximation de  $N_{\max}$  par  $N_0 \exp(-\lambda t_{\max})$  peut être justifiée *a posteriori* : l'intervalle de confiance au niveau 0,95 donné par le TCL dans l'heuristique (3.9) est de l'ordre de  $2N_0^{1/2} \approx 5 \cdot 10^5$ . Dans ce contexte, l'application du TCL dans (3.12) est également validée après coup.

En pratique, la quantité d'échantillon analysée se situe généralement autour de 0,1 à 3 g de carbone. Il est alors communément admis que la méthode de datation au radiocarbone s'applique pour des échantillons d'au plus 50 000 ans.

*Fin du texte.*

#### 4. PROPOSITIONS DE DÉVELOPPEMENTS

- (1) Le candidat pourra détailler les preuves de certaines propositions de la Section 1.

La preuve de la Proposition 1.5 pourrait être l'occasion, le cas échéant, de citer quelques modèles connus supposés être sans vieillissement.

Concernant les Propositions 1.6 et 1.7, le candidat pourra expliquer en quoi l'hypothèse  $F(0) = 0$  est nécessaire.

- (2) Le candidat pourra expliquer pourquoi la fonction de répartition  $F$  introduite dans la Partie 1 peut être supposée continue. Dans ce cadre, il veillera à rappeler les liens entre la régularité de  $F$  et les propriétés de la loi de  $T$ .

- (3) Dans la Partie 2, Sous-section 2.1, le candidat pourra détailler tout ou partie des preuves présentées. Dans ce contexte, le candidat commencera par rappeler ce qu'est un processus de Poisson ainsi que ses propriétés de base.

- (4) Le candidat pourra implémenter l'algorithme 2.6 donnée dans la Sous-section 2.5. Une attention toute particulière sera portée à la modélisation de la loi exponentielle. Un test statistique pourrait alors confirmer les observations graphiques.

Le candidat veillera aussi à expliquer pourquoi cette vérification graphique n'est que partielle. Il pourra éventuellement réfléchir aux moyens possibles pour améliorer cette procédure de vérification.

A l'inverse, le candidat pourra montrer que l'hypothèse poissonnienne est mise en défaut dans certains cas, comme cela est expliqué dans la Sous-section 2.3.

- (5) Dans la Sous-section 3.1, le candidat pourra rappeler ce qu'est la loi des grands nombres. Il serait également pertinent de construire un ou plusieurs intervalles de confiance pour l'âge de décès, en prenant en compte les valeurs de  $t_{\max}$  exhibées en fin de texte.

Une illustration numérique serait alors la bienvenue.

- (6) Dans la Sous-section 3.2, le candidat pourra expliquer l'identité en loi annoncée. Il pourra en particulier donner la loi de la variable  $n$  conditionnellement à la variable  $N$ .

- (7) Dans la Sous-section 3.3, le candidat pourra détailler l'utilisation du théorème central limite ainsi que le sens de l'approximation dans (3.16).

Le candidat pourra également commenter l'ordre de grandeur de  $N_{\max}$  au regard des approximations utilisées, et expliquer la phrase :

*“Dans ces deux cas, l'approximation de  $N_{\max}$  par  $N_0 \exp(-\lambda t_{\max})$  peut être justifiée a posteriori : l'intervalle de confiance au niveau 0,95 donné par le TCL dans l'heuristique (3.9) est de l'ordre de  $2N_0^{1/2} \approx 5.10^5$ . Dans ce contexte, l'application du TCL dans (3.12) est également validée après coup.”*

- (8) Il serait bienvenu de comparer les durées maximales de 30 000 et 40 000 ans obtenues dans la Sous-section 3.3 avec la durée conventionnellement retenue de 50 000 ans. On pourra s'interroger en particulier sur le choix du paramètre  $\delta$  et sur l'influence de la masse de l'échantillon (on pourra illustrer ce point avec une représentation graphique de l'évolution de  $t_{\max}$ ).

On pourra également réfléchir dans ce contexte aux éventuelles difficultés expérimentales de la méthode de comptage.

- (9) En guise de conclusion, il serait également bon d'estimer la précision de la méthode de datation. Une attention toute particulière sera alors portée à la remarque :

*“[...] pour obtenir une erreur relative d'au plus 5% (une telle erreur induisant une erreur de datation de l'ordre de 500 ans), [...]”.*

## 5. ANNEXE

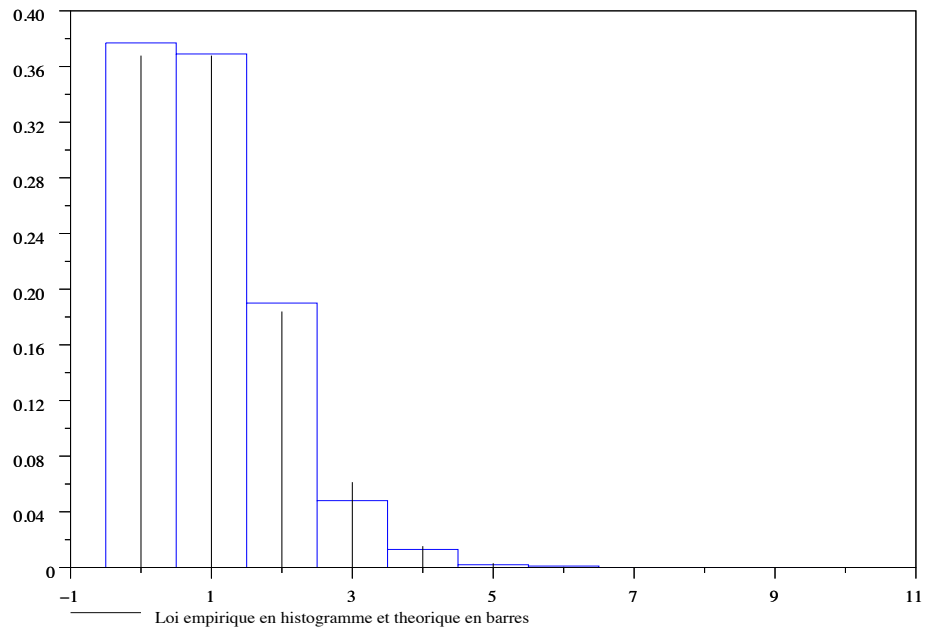


FIGURE 1. COMPARAISON LOI EMPIRIQUE ET LOI DE POISSON (1)