

# Sur les distributions rencontrées en parasitologie

Gaston Pichon,<sup>a</sup>

Christian Mullon<sup>a</sup>

## Introduction : la Filariose de Bancroft, prototype des relations hôte-parasites

Sir Patrick Manson, en mettant pour la première fois en évidence le rôle des Arthropodes dans la transmission des parasites, a posé en 1877 les problèmes théoriques fondamentaux de l'entomologie médicale à l'occasion de travaux sur la filariose.

- celui de la dualité entre finalisme et périodicité ;
- celui de la coaptation temporelle entre le parasite et son vecteur : où vont les microfilaires quand elles disparaissent du sang périphérique ?
- celui de la périodicité des microfilaires : constitue-t-elle un avantage sélectif pour promouvoir une plus grande transmission ?

Nous souhaitons montrer ici que ces problèmes ont une traduction statistique propre qui en permet une approche originale et donne quelques éléments de solutions.

La pénétration des parasites dans le vecteur n'est pas exempte de risques : dents pharyngiennes, membrane péritrophique, coagulation du sang, éjection dans l'intestin postérieur, imperméabilité de la paroi stomacale, lyse ou encapsulation (mélánisation) dans les muscles thoraciques, etc, enfin, possibilité de mort du vecteur, indépendamment ou non du parasitisme. Tout ce cursus se traduit par des observations de distributions de parasites très irrégulières au point qu'on a pu les déclarer imprévisibles, voire aberrantes ; on s'attend en effet à trouver une variance de l'ordre de la moyenne comme dans les lois de Poisson et on trouve expérimentalement une variance de l'ordre du carré de la moyenne. En fait une expérimentation conduite par l'un d'entre nous (G. P., avec J. Prod'hon et F. Rivière) portant sur plusieurs milliers de moustiques (appartenant à 3 espèces) a montré que la prise des microfilaires suit une loi de probabilité simple, la loi géométrique. Nous avons également montré que la même loi s'adapte aux données publiées précédemment et portant sur d'autres couples vecteur/parasites.

Les caractéristiques de la distribution géométrique, celles de sa généralisation, la loi binomiale négative ainsi que leur application en parasitologie conduisent à

<sup>a</sup>LIA ORSTOM 70-74 route d'Aulnay 93140 BONDY

mettre l'accent sur deux paramètres fondamentaux de ces distributions, à savoir leur dispersion

$$d = \frac{\text{variance}}{\text{moyenne}}$$

et le coefficient

$$k = \frac{\text{moyenne}^2}{\text{variance} - \text{moyenne}}$$

Nous présentons ici un logiciel de traitement des distributions parasitaires, PARADIS, basé sur ces constatations.

Il est intéressant d'expliquer ces distributions par la formation de files d'attente dans les capillaires. Nous présentons un second logiciel de simulation de la formation de files d'attente dans les capillaires, CINEFIL.

## 1 Les distributions en parasitologie

L'unité d'observation est l'hôte : chaque hôte est caractérisé par le nombre de parasites qu'il héberge. On obtient ainsi des distributions du type :



Figure 1: forme de distributions de parasites.

où l'on note que 15 hôtes n'ont pas de parasites, 12 en ont 1, 9 en ont 2, etc

## 2 Le problème de la sur-dispersion

Tous les calculs statistiques sur de telles distributions apparaissent sensibles à une de leurs caractéristiques, à savoir le petit nombre d'hôtes ayant un grand nombre de parasites. Cette caractéristique se quantifie par la dispersion

$$d = \frac{\text{variance}}{\text{moyenne}}$$

qui est évaluée directement à partir de la moyenne et de la variance observées, calculées sur les observations dont on dispose. On peut tester l'égalité de

$$\chi^2 = (n - 1) \frac{s^2}{m}$$

(où  $n$  est le nombre d'hôtes de l'expérience) à la valeur nominale du  $\chi^{(2)}$  à  $(n - 1)$  degrés de liberté.

Si l'on représente la distribution spatiale des parasites, on obtient selon les cas de dispersion, des schémas du type :

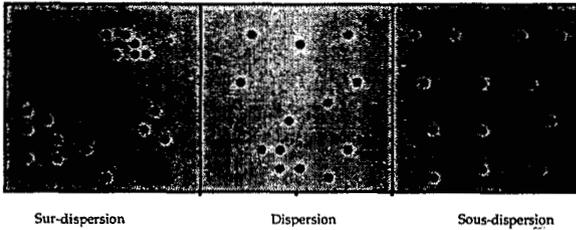


Figure 2: dispersion.

La surdispersion est caractérisée par la constitution d'agrégats, et par conséquence de zones vides ou presque. La sous-dispersion est souvent caractéristique d'une population de prédateurs.

En parasitologie, on observe généralement une surdispersion significative. Seules exceptions les distributions à moyenne très faibles (dans ce cas, la loi binomiale négative et la loi de Poisson sont indiscernables).

### 3 Les principales lois d'ajustement théorique et leurs propriétés

Rappelons brièvement les formules, les caractéristiques et les hypothèses propres aux principales lois de distribution utilisées pour modéliser des populations, Rappelons que l'on pose :

$$d = \frac{\text{variance}}{\text{moyenne}}$$

et

$$k = \frac{\text{moyenne}^2}{\text{variance} - \text{moyenne}}$$

#### 3.1 Loi de Poisson : $P(\lambda)$

La loi de Poisson est utilisée pour modéliser des événements indépendants ayant une espérance constante ; en parasitologie, cela se traduit par l'hypothèse que la position d'un parasite est indépendante de celle d'un de ses congénères.

La loi de Poisson dépend d'un paramètre  $\lambda$ . Sa formule est :

$$P(n) = e^{-\lambda} \frac{\lambda^n}{n!}$$

Ses caractéristiques sont : moyenne = variance =  $\lambda$ ,  $d = 1$ ,  $k = \text{infini}$

### 3.2 Loi géométrique : $G(r)$

La loi géométrique est utilisée pour modéliser des événements successifs affectés d'un facteur constant ; en parasitologie, cela se traduit par l'hypothèse, qu'un certain nombre d'agrégats ayant été constitués, la probabilité qu'un nouvel individu aille se fondre dans l'un d'eux dépend simplement de leur effectif. Elle dépend d'un paramètre  $r$ . Sa formule est :

$$P(n) = (1 - r)r^n$$

Ses caractéristiques sont :

$$\text{moyenne} = \frac{r}{1 - r}$$

$$\text{variance} = \frac{r}{(1 - r)^2}$$

$$d = \frac{1}{1 - r}$$

$$k = 1$$

### 3.3 Loi binomiale négative : $B_n(m, k)$

La loi binomiale négative est une généralisation de la loi géométrique. Elle dépend de deux paramètres  $m$  et  $k$ . Sa formule est :

$$P(n) = CG_k^n (1 - p)^k p^n$$

où

$$CG_k^n = \frac{(k + n)!}{k!(n + 1)!}$$

et

$$p = \frac{m}{m + k}$$

Ses caractéristiques sont :

$$\text{moyenne} = m$$

$$\text{variance} = m + \frac{m^2}{k}$$

$$d = 1 + \frac{m}{k}, k = k$$

Lorsque  $k$  est un entier, cette loi est connue sous le nom de loi de Pascal : en théorie des jeux, c'est la loi du nombre de succès consécutifs. En parasitologie, cette loi correspond à deux hypothèses : celle de la contagion, et celle de l'hétérogénéité.

Le second paramètre,  $k$ , mesure l'écart de la variance à celle qu'on attendrait d'une loi de Poisson. C'est le prototype des distributions parasitaires. Suivant le couple hôte-parasite considéré,  $k$  varie de 0.1 à 5.0.

La loi géométrique est un cas particulier de la loi binomiale négative, où  $k = 1$ .

## 4 Le problème des faux négatifs

Dans les conditions naturelles, il convient de ne pas tenir compte de la classe des hôtes sans parasite, car elle risque d'être composite : âge trop jeune de l'hôte, résistance génétique au parasitisme, non anthropophilie.

Il faut utiliser une loi binomiale négative tronquée. Pour une endémie de filariose de Bancroft donnée, l'ajustement (sans tenir compte de cette classe) à plusieurs foyers donne une valeur constante de  $k$ . Ceci permet, en extrapolant à une loi binomiale complète, d'estimer, à l'aide d'abaques, la proportion de faux négatifs, c'est à dire le nombre d'hôtes pour lesquels on n'a pas trouvé de parasites alors qu'ils en hébergent (Pichon *et al*, 1977).

En fait, la constante  $k=0.3$  paraît caractériser le parasitisme par microfilaires chez l'homme. Dans le cas de la filariose de Bancroft, elle s'applique aux foyers de Polynésie, d'Afrique, des Comores, d'Inde. Ce même paramètre semble caractéristique des populations de *Loa loa* et de *Mansonella perstans* au Congo (Noireau, Pichon, 1991).

## 5 Importance de l'approche statistique

Les paramètres  $d$  et  $k$  sont en fait caractéristiques de la dynamique épidémiologique. On montre ainsi qu'ils sont liés au comportement sexuels des parasites : un faible coefficient  $k$  est favorable à la rencontre des sexes chez un même hôte lorsque la population parasitaire approche de l'éradication.

D'autre part, l'ajustement par la loi binomiale négative permet de quantifier un paramètre épidémiologique fondamental, souvent négligé, voire nié : le coefficient de pathogénicité parasitaire qui dans les conditions naturelles peut atteindre un important ordre de grandeur alors que par contre, dans les conditions d'élevage, la mortalité due au parasitisme peut paradoxalement être plus faible que chez les témoins : du danger de l'extrapolation labo-conditions naturelles ! La comparaison des parasites chez le vecteur à intervalles connus (par exemple, avant et après le repas sanguin des moustiques dans un foyer) permet grâce à l'hypothèse d'une loi de distribution connue d'estimer le facteur létal dû au parasitisme sur le vecteur. Si celui-ci est important, le parasite jouant un rôle limitant pour sa propre prolifération, on aura plutôt intérêt à entreprendre une lutte antivectorielle. La stratégie inverse utilisant une chimioprophylaxie antiparasitaire risque d'avoir pour effet une recrudescence de la population vectorielle et de la transmission.

Très généralement, la probabilité de réussite d'un parasite pour franchir une étape de son évolution chez le vecteur n'est pas indépendante du nombre de parasites en présence. On parle de limitation si le rendement parasitaire diminue avec leur nombre et de facilitation dans le cas contraire. La facilitation, observée en Afrique de l'Ouest (transmission par *Anopheles*), produit une répartition très focalisée et traduit une situation instable ; par contre, la limitation observée en Polynésie (transmission par *Aedes*), s'accompagne d'une grande stabilité, d'une répartition remarquablement homogène malgré l'insularité, et d'une difficulté exceptionnelle pour le combattre. A chaque étape où se produit une nouvelle réduction parasitaire, la distribution résultante suit une loi binomiale négative dont le coefficient  $k$  peut être modifié.

## 6 Le logiciel PARADIS : analyse statistique des distributions

PARADIS regroupe l'ensemble des méthodes statistiques utilisées pour l'analyse des distributions de parasites. Il permet en particulier :

- ajustement à une loi binomiale négative, entière ou tronquée à zéro ;
- estimation des paramètres ;
- estimation des faux négatifs ;
- calculs des coefficients de dispersion ;
- calcul du coefficient  $k$  commun à plusieurs échantillons.

Deux cas extrêmes de la loi binomiale négative, la loi de Poisson et la série logarithmique de Fisher, sont traités. En outre la loi de Polya est incluse ; elle pourrait s'appliquer à des parasites subissant une multiplication asexuée chez l'hôte.

Le logiciel PARADIS utilise un micro-ordinateur de type IBM-PC avec carte graphique. Son emploi est totalement interactif.

```
PARADIS 1.3 : DIStribution de PARAsites. (G.Fichon,Orstom_LIA,1990)
-----
Modèles : Loi Binomiale Négative + Loi B.N.tronquée
Donnée : charge parasitaire (entière) d'un hôte
ou numération (entiere) dans un échantillon donné (volume de sang,snip..)

1.Généralités,Fichiers,Saisie des données :.....[G]
-----
2.Traitement : méthode du maximum de vraisemblance
-----
Distribution complète :.....[C]
Distribution tronquée : (accepte les dist.complètes).....[T]
k commun a plusieurs distributions complètes:.....[K]
Essai d'autres paramètres sur un histogramme observé:.....[P]
3.Divers
-----
Calcul commun de k par méthode des moments
pour plusieurs échantillons de faible effectif.....[E]
Essai d'autres modèles sur un histogramme observé:.....[M]
Revoir un histogramme :.....[R]
Commandes DOS :.....[D]
Sortir :.....[Q]
```

Figure 3: Menu principal de PARADIS.

```
GENERALITES,FICHIERS,SAISIE DES DONNEES
-----
Généralités sur la Binomiale Négative,Bibliographie,etc....[G]
Structure des fichiers de distribution des fréquences:.....[S]
Saisie directe d'une distribution de fréquences :.....[F]
Fichier de données brutes-->distribution de fréquences :...[B]
Retour au menu principal :.....[R]
```

Figure 4 : Sous menu des utilitaires.

On notera le chapitre qui concerne les généralités sur la loi binomiale négative et qui constitue en fait la documentation en ligne du logiciel.

```

Distribution B.N. complète Méthode du maximum de vraisemblance.
-----
Détail des calculs : [o/N]
Nom du fichier :
Démø : Exemple de Bliss et Fisher
Valeur observée maximale : 7
Taux de winsorisation <[1] :

Distribution observée :

n: 150 xmax : 7 t1 :172 t2:536 t3:2170
moy :1.147 var:2.2736 ecty :1.5079
k1 : 1.1667 ( 0.37039)
k2 : 0.9923 ( 0.27512)
*k : 1.0245 ( 0.27609)

X2: 2.490219 ddl: 4 -->Prob :0.646388 ns
Test sur le 3e moment :
T:-1.553092 sT:2.031765 P:0.777687 ns
Test sur le 2e moment (calcul par k2) :
U:-0.198056 sU:0.301520 P:0.744365 ns
Voir l'histogramme : [O/n]
    
```

Figure 5 : estimation de la loi binomiale négative sur le fichier Bliss-Fisher.

$k_1$  et  $k_2$  sont des estimations approchées (méthode des moments et de  $P_0$ ) de la valeur la plus vraisemblable  $*k$ .

$\chi^2$ ,  $T$  et  $U$  sont trois tests d'adéquation de la loi théorique et de la distribution empirique.

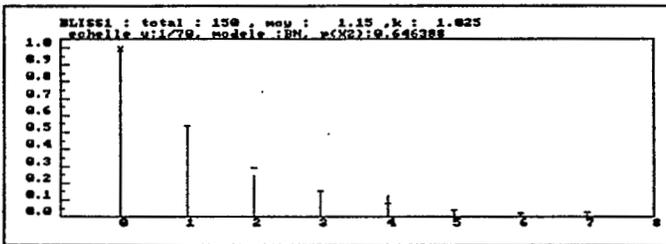


Figure 6 : exemple de Bliss-Fisher, distribution empirique et distribution théorique.

Les premiers paramètres,  $w_1$   $w_2$   $w_3$  décrivent la méthode des moments utilisée pour l'estimation.

Les résultats sont :  $n_0$  = nombre de faux négatifs et  $TD$  = taux de détection. Sur le schéma ci-dessus,  $x$  représente la fréquence théorique des faux négatifs.

```

Binomiale négative tronquée
-----
Voir détail des calculs : [o/N] :
Nom du fichier      :
Démos : [B]rass ou [s]ampford :
BRASS.BNT
Valeur observée maximale : 12
Taux de winsorisation <[1] :
moyenne arithmétique :      3.94  ecart-type :      2.372
moyenne géométrique :      3.24  lambda :      1.331
Méthodes approchées de calcul de w et k :
méth.moments: w0 :0.605488  k0 : 5.707452
méth.Brass : w1 :0.599379  k1 : 5.538540
méthode GPDS: w2 :0.563513  k2 : 4.759365
Méthode du maximum de vraisemblance :
Seuil : 1.00000e-05
Iteration : 4
*w:0.564668  *k:4.853858
vw:0.003004  vk:1.261278  cov:0.060606
sw:0.054808  sk:1.123066

p0:0.062408  n0 :22.630972  n:340  nn:362.630972  TD:0.937592
m:3.991176  M:3.742096

```

Figure 7 : loi binomiale négative tronquée ; données de Brass.

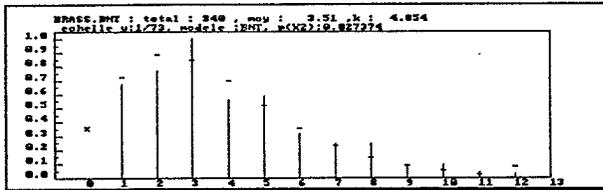


Figure 8 : loi théorique et distribution empirique.

## 7 Le logiciel CINEFIL : simulation des relations hôtes-parasites

D'où peut provenir cette conclusion expérimentale qui veut que les distributions rencontrées en parasitologie suivent une loi binomiale négative ?

### 7.1 Le modèle probabiliste

Une explication de type mathématique et reposant sur des hypothèses physiologiques est la suivante : le phénomène résulte de la conjonction de deux processus stochastiques :

- le dénombrement des microfilaires arrivant à l'entrée du lit capillaire suit une loi de Poisson ;

- le nombre de traversées du lit capillaire par des microfilières isolées suit une loi de Poisson ;
- dans les capillaires où les microfilières sont ingérées par leurs vecteurs, le dépassement n'est pas possible.

On démontre alors que les microfilières constituent des files d'attente dont la taille suit une loi géométrique, cas particulier remarquable de la loi binomiale négative, prototype des relations hôte-parasite. En fait, la démonstration n'est valable que dans cette situation de "guichet" bien définie.

## 7.2 Le modèle "événementiel"

On souhaite pouvoir modifier les hypothèses ci-dessus en décrivant différemment le comportement de microfilières ; mais alors, dans presque tous les cas, le phénomène ne semble pas soluble mathématiquement. Une simulation informatique est donc nécessaire. Reposant sur un "modèle événementiel", le logiciel CINEFIL permet de visualiser le comportement des microfilières et de déduire leur distribution dans un grand nombre de situations.

- On peut ainsi vérifier que l'hypothèse de départ est parasitologiquement réaliste en contrôlant que la distribution résultante ne s'éloigne pas trop d'une loi binomiale négative ;
- On peut aussi suivre l'évolution de différents paramètres du système ;
- Le recours à l'animation donne une impression qualitative du phénomène.

## 7.3 Application à diverses situations hôtes-parasites

Le logiciel CINEFIL, initialement connu pour vérifier le comportement des microfilières, est aisément transformable en outil d'exploration parasitologique.

L'analogie entre un capillaire et un hôte est envisageable, dans la mesure où, comme un capillaire, un hôte peut-être considéré comme le siège d'un flux de parasites, auxquels il impose et dont il subit certaines contraintes, et dans lequel les parasites interagissent.

Dans CINEFIL, chaque microfilière est dotée d'un petit nombre d'attributs qui suffit pour produire un comportement à la fois diversifié et complexe ; on peut donc l'utiliser dans un contexte plus général pour explorer les conséquences de divers comportements :

- de l'hôte : survie en fonction du fardeau parasitaire, réaction contre les parasites, hôtes de sensibilité différente... ;
- du parasite : rôle de l'état sexué, "over-crowding" et autres effets densité-dépendants, comportement grégaire ("togetherness" de May, 3) ou individualiste, sensibilité périodique à certains stimulus ou à des agents anti-parasitaires ("barrières oxygène" d'Hawking, 2).

## 7.4 Exploitations

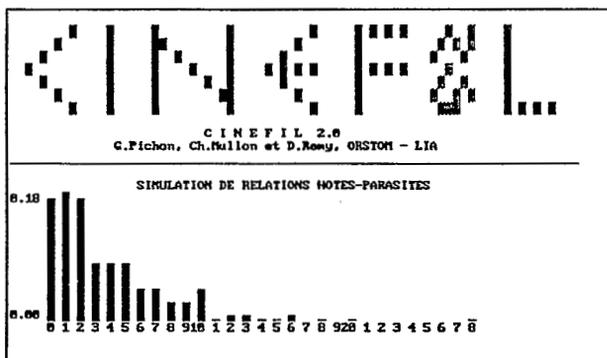


Figure 9 : écran de présentation de CINEFIL.

Avant les simulations proprement dites, un didacticiel décrit les distributions en parasitologie, les processus stochastiques et les formations de files d'attente qui se produisent.

```
INITIALISATION: SAISIE DES PARAMETRES
(Valeurs par défaut: [ ]
-Nombre de capillaires (< 5)           :[ 5]
-Nombre de microfilaries (< 400)       :[ 100]
-Modification de la longueur des capillaires :[O/N]
Intervalles moyens :
-Traversee individuelle des capillaires o5 :[30.000]
-A l'exterieur d'un capillaire oa(>o5)   :[31.000]
[C]ontinuer ou [A]utre choix :
```

Figure 10 : CINEFIL, menu principal.

Un des intérêts de CINEFIL est de proposer toute une gamme d'options sur le comportement des microfilaraires dans les capillaires. Ainsi le comportement normal correspond à un choix aléatoire du capillaire et le comportement misofilaire correspond à choisir le capillaire le moins parasité, etc.

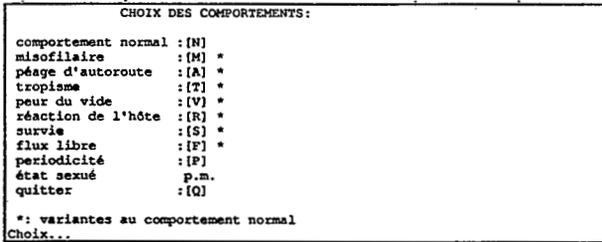


Figure 11 : CINEFIL, options.

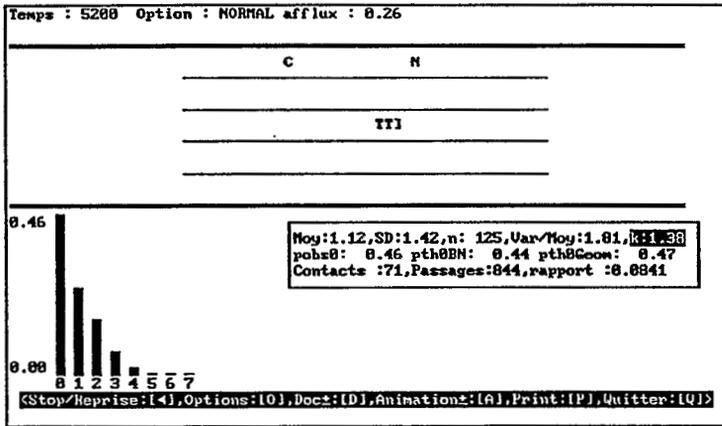


Figure 12 : L'écran de CINEFIL pendant l'animation.

A tout moment, il est possible de débrayer l'animation pour accélérer le processus. On peut aussi faire apparaître une brève documentation sur les hypothèses mises en jeu.

## Conclusion

Nous avons montré que sur un échantillon assez important de distributions hôtes/parasites publiés, on observe que la valeur du coefficient  $k=1$  occupe une position centrale entre les extrêmes  $k=0,1$  et  $k=5,0$ . Cette valeur est aussi celle que l'on obtient, à moins de 2% près, pour nos dénombrements de microfilaraires.

Nous avons suggéré une explication de ce phénomène en termes dynamiques : par la constitution de files d'attente.

Une autre explication en termes évolutionnistes est la suivante :

La loi géométrique occupe également une position particulière dans la famille des lois binomiales négatives : pour une moyenne donnée, c'est celle dont l'entropie est la plus grande (figure 13). L'entropie, ou incertitude est mesurée en bits par la formule de Shannon :

$$H = \sum_i p_i \log_2(p_i)$$

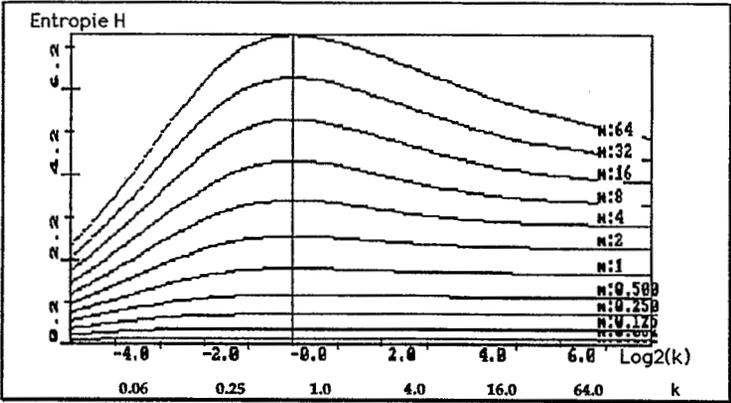


figure 13 : entropie en fonction de k pour différentes valeurs de la moyenne d'une loi binomiale négative.

On peut l'interpréter comme une mesure du degré d'indifférenciation, de non-coévolution d'un couple hôte/parasite. Toute modification du comportement de l'un d'eux ne peut entraîner qu'une diminution de cette incertitude vers une plus faible ou une plus grande valeur de k. C'est ce que l'on obtient avec le logiciel CINEFIL : si on modifie les hypothèses de départ, les valeurs de k résultantes sont inférieures ou supérieures à 1. En théorie de l'information, si on code par A, B, et C des infestations faibles, moyennes ou fortes, une distribution géométrique se traduira par un message peu informatif, du type *CCACBBABCBA*, d'après lequel la lettre suivante est peu prévisible. La première solution conduira à des messages du type *AAAAACAABA* tandis que l'autre tendra vers des messages *BBABBBBCBB*.

$k < 1$  serait donc une adaptation à un environnement instable. Cette distribution pourrait supporter de fortes diminutions de la population parasitaire sans l'éradiquer, grâce à une plus grande probabilité de rencontre des sexes différents chez un même hôte, et de trop grandes proliférations pourraient être régulées par l'élimination des hôtes hyperparasités. Inversement, la seconde stratégie ( $k > 1$ ) serait une adaptation à un environnement stable. Mais des fluctuations catastrophiques exceptionnelles de l'environnement seraient fatales pour des populations poissonniennes ayant perdu toute adaptabilité, ce qui expliquerait que  $k$  ne dépasse pas certaines limites (5.0) dans les populations naturelles.

L'écart de  $k$  à 1 serait donc un indice de co-évolution entre l'hôte et le parasite.

## Références bibliographiques

- R.M. May : (1977) *Parasitology*, 75, 259-76.
- C Mullon, G Pichon : (1990) Problèmes statistiques de la grande variabilité, dans *Seminfor 4 Le transfert d'échelle*, ORSTOM.
- G Pichon : (1988) Transmission des filarioses : du modèle descriptif au modèle heuristique, dans *Seminfor 2 La modélisation, aspects pratiques et méthodologie*, ORSTOM.
- G. Pichon, J. Prod'hon, F. Rivière : (1980) *Cahier ORSTOM Série Entomologie Médicale et Parasitologie*, 1, 24-47.