Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 248 (2007) 480–489
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Stochastic models of kleptoparasitism

G.E. Yates, M. Broom*

Mathematics Department, University of Sussex, Mantell Building, Falmer, Brighton BN1 9RF, UK

Received 16 January 2007; received in revised form 9 May 2007; accepted 9 May 2007
Available online 17 May 2007

Abstract

In this paper, we consider a model of kleptoparasitism amongst a small group of individuals, where the state of the population is described by the distribution of its individuals over three specific types of behaviour (handling, searching for or fighting over, food). The model used is based upon earlier work which considered an equivalent deterministic model relating to large, effectively infinite, populations. We find explicit equations for the probability of the population being in each state. For any reasonably sized population, the number of possible states, and hence the number of equations, is large. These equations are used to find a set of equations for the means, variances, covariances and higher moments for the number of individuals performing each type of behaviour. Given the fixed population size, there are five moments of order one or two (two means, two variances and a covariance). A normal approximation is used to find a set of equations for these five principal moments. The results of our model are then analysed numerically, with the exact solutions, the normal approximation and the deterministic infinite population model compared. It is found that the original deterministic models approximate the stochastic model well in most situations, but that the normal approximations are better, proving to be good approximations to the exact distribution, which can greatly reduce computing time.

Keywords: Food stealing; Equilibrium distribution; Cumulant generating function; Markov process; Parasitism

1. Introduction

There are many biological situations in which a population is divided amongst a number of sites. These sites may be physical locations, such as patches of food, or may categorise the population in some way, perhaps by their activity or whether they have a particular disease or not. The rates at which individuals leave each site depend on the current location of the individual in question, and are also often dependent on the location of other members of the population.

A feature of interest in such situations is the expected proportion of the population at each site at equilibrium. This is often calculated using a deterministic model, which assumes that the population is large. Individual movements in a small population have a much larger effect on transition rates than in a large population, and can also result in large differences between the expected proportion of the population on each site given by the two models. As a result of this, stochastic models are generally more appropriate when considering small populations. The use of stochastic models also enables the variance of the number of individuals at each site to be calculated, which is not possible using deterministic models.

Examples of such models include:

(1) The SIS epidemic. This is one of the simplest epidemic models, in which individuals are either susceptible or infective, and once they recover from the disease they are immediately susceptible again (see Weiss and Dishon, 1971; Nasell, 1996, 1999; Kryscio and Lefèvre, 1989). The deterministic and stochastic models in this case give very different results. The deterministic model either converges to a stable mixture of infective and susceptible individuals, or the disease becomes extinct, depending upon parameter values. For the stochastic model the stationary distribution is always the trivial one where the disease is extinct. Extinction may take a long time to happen, and we

---

*Corresponding author. Tel.: +44 1273 877243; fax: +44 1273 678097.
E-mail address: m.broom@sussex.ac.uk (M. Broom).

0022-5193/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2007.05.007

may be interested in interim behaviour, so that the quasi-stationary distribution (conditional on non-extinction) is considered.

(2) Coagulation–fragmentation processes. The model population contains N individuals who are grouped into clusters of various sizes, where possible events are the merging of two groups into one and the separation of a group into two smaller groups. These processes have many applications including animal grouping; see Durrett et al. (1998, 1999) and their references. The long-term behaviour of these processes again differs between the deterministic and stochastic cases, but the deterministic model is generally a better approximation to the stochastic one than in the SIS epidemic.

(3) The ideal free distribution (Fretwell and Lucas, 1970) which describes the distribution of animals among a number of patches of a resource, such as prey or mates (see, for example, Jackson et al., 2004; Hugie and Grand, 1998; Houston and McNamara, 1988; Ruxton and Humphries, 2003; Yates and Broom, 2005). Deterministic models predict the optimal distribution of individuals between patches of different quality. Stochastic models generally predict undermatching (a lower proportion of individuals on the higher quality site than predicted by the deterministic model) and this is often what is observed in real populations, see for example Kennedy and Gray (1993). It should be noted that Hamilton (2002) combined the concept of the ideal free distribution and food stealing, and modelled individuals moving between two food patches where stealing was also possible.

We focus in this paper on the modelling of food stealing or *kleptoparasitism*. Many authors have observed various animals stealing food from others. For example Brockmann and Barnard (1979) reviewed literature from the previous 40 years and found such occurrences among a variety of bird species. Such kleptoparasitic behaviour can be intraspecific, when food is stolen from members of the same species, or interspecific, when food is stolen from members of a different species. Although kleptoparasitism is particularly well documented among birds, it occurs in many species. For example, Vollrath (1979) conducted both field and laboratory research into the theridiid spider, which steals from two other spider species, while Homer et al. (2002) have observed the behaviour in hyenas.

Recently a series of game-theoretic models of kleptoparasitic behaviour (e.g. Broom and Ruxton, 2003; Luther and Broom, 2004; Broom et al., 2004) have been developed investigating the strategic choices of individuals in a variety of circumstances. They took as their basis the paper Ruxton and Moody (1997) where behaviour was completely deterministic in character, due to the large population assumed and the lack of decisions of the individuals involved, and its refinement in Broom and Ruxton (1998). Indeed, conditional on individuals' decisions, all of these models were essentially deterministic. In this paper we consider the refinement of Ruxton and Moody (1997) in Broom and Ruxton (1998) where no decisions are allowed by individuals, and introduce a stochastic version of this model, comparing the large population situation to the non-deterministic stochastic version.

2. The models

We first recall the structure of the basic model from Ruxton and Moody (1997) and Broom and Ruxton (1998).

They considered a population where the density of individuals is P. This population is divided into three different activities; searching for food items, which we label state S_1 (state S in the original papers), handling a food item, state S_2 (originally state H), and involved in an aggressive contest, state S_3 (originally state A). We shall label the number of individuals in S_i as X_i for $i = 1, 2$ and the number of *fighting pairs* in S_3 as X_3. Thus if the total population size is n, then

$$X_1 + X_2 + 2X_3 = n.$$

The number of available food items per unit area is given by f. Individuals were able to search an area v_f for food in unit time, so that the rate at which individual searchers found food (and so moved from state S_1 to S_2) was $\lambda_{12} = v_f f$. At the end of handling, the handler resumed searching. It was assumed that food items take a time to handle drawn from an exponential distribution with mean t_h, so that individuals move from state S_2 to S_1 at rate $\lambda_{21} = 1/t_h$. Individuals also searched for handlers, being able to search an area of size v_h per unit time. When a searcher encountered a handler, it challenged for the food item and a fight ensued, and one individual from each of S_1 and S_2 moved to the fighting state S_3, this occurred at rate $\lambda_{13} = v_h$ per pair of individuals. Contest times were drawn from an exponential distribution with mean time $t_a/2$. At the end of a contest, the winner started handling the food, and the loser resumed searching, so that a fighting pair splits, one of each moving from S_3 to the states S_1 and S_2, at rate $\lambda_{31} = 2/t_a$. These are summarised in Fig. 1.

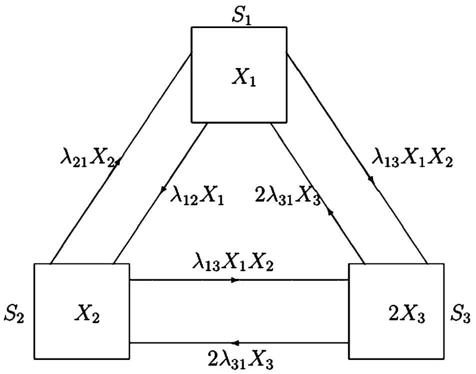

Fig. 1. The stochastic kleptoparasitism model. The sites are indicated by the boxes, transition rates are given on the arc with the arrow in the appropriate direction between the states. Both rates into S_3 are identical, as are those leaving S_3. Individuals enter state S_3 in pairs, one each from S_1 and S_2. Similarly, pairs leave S_3 simultaneously, with one going to each of S_1 and S_2.

2.1. The deterministic model

We further define the proportion of individuals in state S_i as z_i for all i and thus if the total population size is n, then $z_1 = X_1/n$, $z_2 = X_2/n$ and $z_3 = 2X_3/n$. The transition rates are as in Fig. 2.

We can thus find the large population transition rates λ'_{ij} in terms of λ_{ij} as follows: The transition rate from S_1 to S_2 is given by $\lambda_{12}X_1 = \lambda_{12}(nz_1) = n(\lambda'_{12}z_1)$ for large $n \Rightarrow \lambda'_{12} = \lambda_{12}$. Similarly $\lambda'_{21} = \lambda_{21}$, while $2\lambda_{31}X_3 = 2\lambda_{31}(nz_3/2) = n(\lambda'_{31}z_3)$ for large $n \Rightarrow \lambda'_{31} = \lambda_{31}$ and $\lambda_{13}X_1X_2 = \lambda_{13}(nz_1)(nz_2) = n(\lambda'_{13}z_1z_2)$ for large $n \Rightarrow \lambda'_{13} = n\lambda_{13}$.

Following Broom and Ruxton (1998) we see that for an essentially infinite population, the proportions of the population in each site satisfy the differential equations:

$$\frac{\mathrm{d}z_1}{\mathrm{d}t} = \lambda'_{21}z_2 + \lambda'_{31}z_3 - \lambda'_{12}z_1 - \lambda'_{13}z_1z_2, \tag{1}$$

$$\frac{\mathrm{d}z_2}{\mathrm{d}t} = \lambda'_{12}z_1 + \lambda'_{31}z_3 - \lambda'_{21}z_2 - \lambda'_{13}z_1z_2, \tag{2}$$

$$\frac{\mathrm{d}z_3}{\mathrm{d}t} = 2\lambda'_{13}z_1z_2 - 2\lambda'_{31}z_3. \tag{3}$$

Note that Eq. (3) can be obtained from a linear combination of Eqs. (1) and (2). At equilibrium $\mathrm{d}z_1/\mathrm{d}t = \mathrm{d}z_2/\mathrm{d}t = \mathrm{d}z_3/\mathrm{d}t = 0$, and substituting $z_3 = 1 - z_1 - z_2$ into Eqs. (1) and (2) and rearranging gives

$$(\lambda'_{12} + \lambda'_{31})z_1 + (\lambda'_{31} - \lambda'_{21})z_2 + \lambda'_{13}z_1z_2 - \lambda'_{31} = 0,$$

$$(\lambda'_{31} - \lambda'_{12})z_1 + (\lambda'_{21} + \lambda'_{31})z_2 + \lambda'_{13}z_1z_2 - \lambda'_{31} = 0.$$

Therefore

$$z_1 = \frac{\lambda'_{21}}{\lambda'_{12}}z_2, \tag{4}$$

while z_2 solves

$$\lambda'_{13}\lambda'_{21}z_2^2 + \lambda'_{31}(\lambda'_{12} + \lambda'_{21})z_2 - \lambda'_{12}\lambda'_{31} = 0. \tag{5}$$

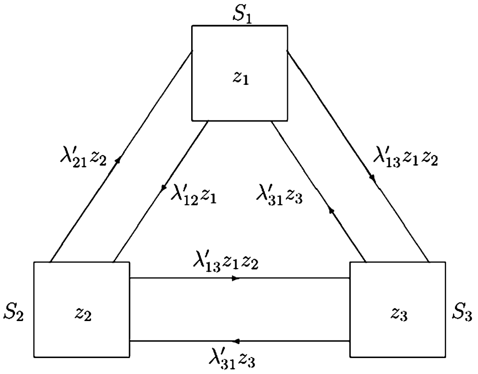

Fig. 2. The deterministic kleptoparasitism model. The states are indicated by the boxes, transition rates are given on the arc with the arrow in the appropriate direction between the states per individual; thus the actual transition rate of movement in a large population of size n is $n\times$ this individual rate. Both rates into S_3 are identical, as are those leaving S_3. Individuals enter state S_3 in pairs, one each from S_1 and S_2.

Since the coefficients of z_2^2 and z_2 in this equation are positive, z_2 is given by the positive root, meaning that

$$z_2 = \frac{-\lambda'_{31}(\lambda'_{12} + \lambda'_{21}) + \sqrt{{\lambda'_{31}}^2(\lambda'_{12} + \lambda'_{21})^2 + 4\lambda'_{12}\lambda'_{13}\lambda'_{21}\lambda'_{31}}}{2\lambda'_{13}\lambda'_{21}}.$$

Note that z_2 is the *handling ratio*, which is directly related to the food consumption rate per individual. In fact this consumption rate is $z_2/t_h = \lambda'_{21}z_2$.

2.2. The stochastic kleptoparasitism model

Following the description of transitions at the start of this section, we can see that the transition rates for this model are as shown in Fig. 1. The numbers of individuals in S_1 and S_2 are X_1 and X_2, respectively, while X_3 gives the number of pairs in S_3. There are only four possible movements. Recall that the population size n is fixed, and that $X_1 + X_2 + 2X_3 = n$.

The probability that $X_1(t) = x_1$ and $X_2(t) = x_2$ is denoted by $p_{x_1,x_2}(t)$. The Kolmogorov forward equations for $p_{x_1,x_2}(t)$ for this process are

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}p_{x_1,x_2}(t) \\ &= \lambda_{13}(x_1+1)(x_2+1)p_{x_1+1,x_2+1}(t) \\ &\quad + \lambda_{31}(n - x_1 - x_2 + 2)p_{x_1-1,x_2-1}(t) \\ &\quad + \lambda_{12}(x_1+1)p_{x_1+1,x_2-1}(t) + \lambda_{21}(x_2+1)p_{x_1-1,x_2+1}(t) \\ &\quad - [\lambda_{13}x_1x_2 + \lambda_{31}(n - x_1 - x_2) + \lambda_{12}x_1 \\ &\quad + \lambda_{21}x_2]p_{x_1,x_2}(t).\end{aligned} \tag{6}$$

$\forall x_1, x_2 \in \{0, 1, \ldots, n\}$, where by convention $p_{x_1,x_2} = 0$ if any of $x_1 < 0$, $x_2 < 0$ or $x_1 + x_2 > n$ occur. In addition, since the number of individuals in S_3 must be even, $p_{x_1,x_2}(t) = 0$ if $n - x_1 - x_2$ is odd. At equilibrium $(\mathrm{d}/\mathrm{d}t)p_{x_1,x_2}(t) = 0\ \forall\ x_1, x_2$ and these equations become a set of linear equations for the joint equilibrium distribution of X_1 and X_2.

2.2.1. The number of distinct states of the stochastic model

Under the model, $n - x_1 - x_2$ must be even; the number of states for which $X_1 = x_1$ depends on whether $n - x_1$ is odd or even. If $n - x_1$ is odd, there are $(n - x_1 + 1)/2$ such states, otherwise there are $(n - x_1)/2 + 1$ states. Since $(n - x_1 + 1)/2 = (n - (x_1 + 1))/2 + 1$, the number of states when n is even is

$$\begin{aligned}\frac{n}{2} + 1 + \sum_{i=1}^{n/2} 2\left(\frac{n-2i}{2} + 1\right) &= \frac{n}{2} + 1 + \sum_{i=1}^{n/2} n - 2i + 2 \\ &= \frac{n}{2} + 1 + (n+2)\frac{n}{2} \\ &\quad - \frac{n}{2}\left(\frac{n}{2} + 1\right) \\ &= \left(\frac{n}{2} + 1\right)^2.\end{aligned}$$

If n is odd the total number of states is

$$\begin{aligned}\sum_{k=1}^{(n+1)/2} 2\left(\frac{n-(2k-1)}{2}+1\right) &= (n+3)\frac{n+1}{2}-2\sum_{k=1}^{(n+1)/2} k\\ &= (n+3)\frac{n+1}{2}\\ &\quad -\frac{n+1}{2}\left(\frac{n+1}{2}+1\right)\\ &= \frac{(n+1)(n+3)}{4}.\end{aligned}$$

In either case the number of states is of order $n^2/4$.

2.2.2. Equations for the moments of X_1 and X_2

It is not practical to solve the equations for p_{x_1,x_2}. Instead we find equations for the principal moments of the process, namely the means, variances and the covariance of the random variables X_1 and X_2 (from which the moments involving X_3 can also be found). We begin by obtaining a partial differential equation for the cumulant generating function of X_1 and X_2, from which equations for the above moments can be derived. The cumulant generating function is denoted $K(s_1,s_2)$ and is defined to be $\log M(s_1,s_2)$, where $M(s_1,s_2)$ is the moment generating function. From this it is possible to obtain equations for the cumulants of X_1 and X_2. The cumulant $k_{i,j}$ is the coefficient of $s_1^i s_2^j/i!j!$ in the series expansion of $K(s_1,s_2)$. The partial differential equation for $K(s_1,s_2)$ is

$$\begin{aligned}&n\lambda_{31}(\mathrm{e}^{s_1+s_2}-1)+[\lambda_{12}(\mathrm{e}^{s_2-s_1}-1)+\lambda_{31}(1-\mathrm{e}^{s_1+s_2})]\frac{\partial K}{\partial s_1}\\ &\quad+[\lambda_{21}(\mathrm{e}^{s_1-s_2}-1)+\lambda_{31}(1-\mathrm{e}^{s_1+s_2})]\frac{\partial K}{\partial s_2}\\ &\quad+\lambda_{13}(\mathrm{e}^{-(s_1+s_2)}-1)\left(\frac{\partial^2 K}{\partial s_1\partial s_2}+\frac{\partial K}{\partial s_1}\frac{\partial K}{\partial s_2}\right)=0. \qquad (7)\end{aligned}$$

This is derived in Appendix A.

Differentiating $K(s_1,s_2)$ with respect to s_1 i times and s_2 j times and setting $s_1=s_2=0$ in the function obtained gives the (i,j) cumulant of X_1 and X_2.

For example,

$$\left.\frac{\partial K}{\partial s_1}\right|_{s_1=s_2=0}=E(X_1),\quad \left.\frac{\partial^2 K}{\partial s_1^2}\right|_{s_1=s_2=0}=Var(X_1)\quad\text{and}$$

$$\left.\frac{\partial^2 K}{\partial s_1\partial s_2}\right|_{s_1=s_2=0}=Cov(X_1,X_2).$$

Applying this to Eq. (7) gives the following set of equations:

$$n\lambda_{31}-\lambda_{31}\mu_1-\lambda_{31}\mu_2-\lambda_{13}\sigma_{12}-\lambda_{13}\mu_1\mu_2=0, \qquad (8)$$

$$\lambda_{12}\mu_1-\lambda_{21}\mu_2=0, \qquad (9)$$

$$\begin{aligned}&n\lambda_{31}-\lambda_{31}\mu_1+(\lambda_{21}-\lambda_{31})\mu_2-(\lambda_{12}+\lambda_{31})\sigma_1^2+(\lambda_{21}-\lambda_{31})\sigma_{12}\\ &\quad-\lambda_{13}k_{2,1}-\lambda_{13}\mu_1\sigma_{12}-\lambda_{13}\mu_2\sigma_1^2=0, \qquad (10)\end{aligned}$$

$$\lambda_{12}\mu_1+\lambda_{21}\mu_2-\lambda_{12}\sigma_1^2-\lambda_{21}\sigma_2^2+(\lambda_{12}+\lambda_{21})\sigma_{12}=0, \qquad (11)$$

$$\begin{aligned}&n\lambda_{31}+(\lambda_{12}-\lambda_{31})\mu_1-\lambda_{31}\mu_2-(\lambda_{21}+\lambda_{31})\sigma_2^2+(\lambda_{12}-\lambda_{31})\sigma_{12}\\ &\quad-\lambda_{13}k_{1,2}-\lambda_{13}\mu_2\sigma_{12}-\lambda_{13}\mu_1\sigma_2^2=0. \qquad (12)\end{aligned}$$

These five equations contain the seven unknowns μ_1, μ_2, σ_1^2, σ_2^2, σ_{12}, $k_{2,1}$ and $k_{1,2}$, where $\mu_i=E(X_i)$, $\sigma_i^2=Var(X_i)$, $\sigma_{12}=Cov(X_1,X_2)$, $k_{2,1}=E[(X_1-\mu_1)^2(X_2-\mu_2)]$ and $k_{1,2}=E[(X_1-\mu_1)(X_2-\mu_2)^2]$. Further equations can be obtained, but the number of unknowns would also be increased.

It is possible to obtain expressions for μ_1 and σ_{12} in terms of μ_2 from the first two of these equations. These are

$$\mu_1=\frac{\lambda_{21}}{\lambda_{12}}\mu_2 \qquad (13)$$

and

$$\sigma_{12}=-\frac{\lambda_{21}}{\lambda_{12}}\mu_2^2-\frac{\lambda_{31}(\lambda_{12}+\lambda_{21})}{\lambda_{13}\lambda_{12}}\mu_2+\frac{n\lambda_{31}}{\lambda_{13}} \qquad (14)$$

and are derived in Appendix B.

We now show two results relating these moments.

Result 1. The signs of $nz_1-\mu_1$ and $nz_2-\mu_2$ are the same.

The proof follows from the fact that expression (4) for z_1, with the substitutions $\lambda'_{12}=\lambda_{12}$ and $\lambda'_{21}=\lambda_{21}$ can be rearranged to give

$$\lambda_{12}z_1-\lambda_{21}z_2=0.$$

Multiplying this equation by n and subtracting (9) gives

$$\lambda_{12}(nz_1-\mu_1)=\lambda_{21}(nz_2-\mu_2).$$

Since λ_{12} and λ_{21} are both positive, it follows that $nz_1-\mu_1$ and $nz_2-\mu_2$ have the same sign.

Thus the number of searchers and handlers is either both overestimated or both underestimated by the deterministic model.

Result 2. The signs of $nz_2-\mu_2$ and σ_{12} are the same (and so are also the same as that of $nz_1-\mu_1$).

Making the substitutions $\lambda'_{12}=\lambda_{12}$, $\lambda'_{13}=n\lambda_{13}$, $\lambda'_{21}=\lambda_{21}$ and $\lambda'_{31}=\lambda_{31}$ into Eq. (5) and multiplying by n gives

$$\lambda_{13}\lambda_{21}(nz_2)^2+\lambda_{31}(\lambda_{12}+\lambda_{21})nz_2-n\lambda_{12}\lambda_{31}=0.$$

Substituting the expression for μ_1 into (8) and multiplying by λ_{12} gives

$$n\lambda_{12}\lambda_{31}-\lambda_{31}(\lambda_{12}+\lambda_{21})\mu_2-\lambda_{12}\lambda_{13}\sigma_{12}-\lambda_{13}\lambda_{21}\mu_2^2=0.$$

Adding these two equations gives

$$\lambda_{13}\lambda_{21}((nz_2)^2-\mu_2^2)+\lambda_{31}(\lambda_{12}+\lambda_{21})(nz_2-\mu_2)-\lambda_{12}\lambda_{13}\sigma_{12}=0.$$

$$\Rightarrow (nz_2-\mu_2)(\lambda_{13}\lambda_{21}(nz_2+\mu_2)+\lambda_{31}(\lambda_{12}+\lambda_{21}))=\lambda_{12}\lambda_{13}\sigma_{12}.$$

This shows that the signs of $nz_2-\mu_2$ and σ_{12} are the same, since the λ_{ij}s and μ_2 are all positive.

Thus if X_1 and X_2 are negatively correlated, as we would naively assume should usually be the case since the total number of individuals in the three categories is fixed, then this would indicate that the deterministic model would underestimate the number of searchers and handlers.

2.2.3. The normal approximation

If one assumes that X_1 and X_2 are bivariate normal, then $k_{2,1}$ and $k_{1,2}$ are both 0. This method is described by Whittle (1957) and has been applied more recently by Nasell (2003) with the stochastic logistic model, of which the SIS epidemic is an example. Using this approximation it is possible to derive expressions for $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ in terms of $\hat{\mu}_2$, and a quartic equation for $\hat{\mu}_2$, where the ^ above each variable denotes the normal approximation. These are derived in Appendix B.

The variances are given by

$$\begin{aligned}\hat{\sigma}_1^2 = {} & \frac{\lambda_{21}^2}{\lambda_{12}^2}\hat{\mu}_2^2 + \frac{2\lambda_{21}(\lambda_{31}-\lambda_{21})}{\lambda_{12}\lambda_{13}}\hat{\mu}_2 + \frac{\lambda_{21}-\lambda_{31}}{\lambda_{13}} + \frac{2\lambda_{21}^2 - 3\lambda_{21}\lambda_{31} + \lambda_{31}^2}{\lambda_{13}^2} - \frac{\lambda_{21}\lambda_{31}(n+1)}{\lambda_{12}\lambda_{13}} + \frac{\lambda_{21}\lambda_{31}(\lambda_{21}-\lambda_{31})}{\lambda_{12}\lambda_{13}^2} \\ & + \frac{\lambda_{31}((2\lambda_{21}-\lambda_{31})n + \lambda_{12} + \lambda_{31}) - \lambda_{12}\lambda_{21}}{\lambda_{13}(\lambda_{12}+\lambda_{31}+\lambda_{13}\hat{\mu}_2)} + \frac{\lambda_{31}n}{\lambda_{12}+\lambda_{31}+\lambda_{13}\hat{\mu}_2} + \frac{(\lambda_{31}-\lambda_{21})(3\lambda_{21}\lambda_{31} + 2\lambda_{12}\lambda_{21} - \lambda_{31}^2 - \lambda_{12}\lambda_{31})}{\lambda_{13}^2(\lambda_{12}+\lambda_{31}+\lambda_{13}\hat{\mu}_2)} \\ & + \frac{\lambda_{21}\lambda_{31}^2(n+1)}{\lambda_{12}\lambda_{13}(\lambda_{12}+\lambda_{31}+\lambda_{13}\hat{\mu}_2)} + \frac{\lambda_{21}\lambda_{31}^2(\lambda_{31}-\lambda_{21})}{\lambda_{12}\lambda_{13}^2(\lambda_{12}+\lambda_{31}+\lambda_{13}\hat{\mu}_2)}\end{aligned} \tag{15}$$

and

$$\begin{aligned}\hat{\sigma}_2^2 = {} & \hat{\mu}_2^2 + \frac{2(\lambda_{31}-\lambda_{12})}{\lambda_{13}}\hat{\mu}_2 + \frac{\lambda_{12}-\lambda_{31}}{\lambda_{13}} + \frac{2\lambda_{12}^2 - 3\lambda_{12}\lambda_{31} + \lambda_{31}^2}{\lambda_{13}^2} - \frac{\lambda_{12}\lambda_{31}(n+1)}{\lambda_{13}\lambda_{21}} + \frac{\lambda_{12}[\lambda_{31}((2\lambda_{12}-\lambda_{31})n + \lambda_{21} + \lambda_{31}) - \lambda_{12}\lambda_{21}]}{\lambda_{13}(\lambda_{21}\lambda_{12} + \lambda_{31}\lambda_{12} + \lambda_{21}\lambda_{13}\hat{\mu}_2)} \\ & + \frac{\lambda_{12}\lambda_{31}(\lambda_{12}-\lambda_{31})}{\lambda_{13}^2\lambda_{21}} + \frac{\lambda_{12}\lambda_{31}n}{\lambda_{21}\lambda_{12} + \lambda_{31}\lambda_{12} + \lambda_{21}\lambda_{13}\hat{\mu}_2} + \frac{\lambda_{12}(\lambda_{31}-\lambda_{12})(3\lambda_{31}\lambda_{12} + 2\lambda_{21}\lambda_{12} - \lambda_{31}^2 - \lambda_{21}\lambda_{31})}{\lambda_{13}^2(\lambda_{21}\lambda_{12} + \lambda_{31}\lambda_{12} + \lambda_{21}\lambda_{13}\hat{\mu}_2)} \\ & + \frac{\lambda_{12}^2\lambda_{31}^2(n+1)}{\lambda_{13}\lambda_{21}(\lambda_{21}\lambda_{12} + \lambda_{31}\lambda_{12} + \lambda_{21}\lambda_{13}\hat{\mu}_2)} + \frac{\lambda_{12}^2\lambda_{31}^2(\lambda_{31}-\lambda_{12})}{\lambda_{13}^2\lambda_{21}(\lambda_{21}\lambda_{12} + \lambda_{31}\lambda_{12} + \lambda_{21}\lambda_{13}\hat{\mu}_2)},\end{aligned} \tag{16}$$

while the quartic equation for $\hat{\mu}_2$ is

$$\begin{aligned} & 2\lambda_{13}^3\lambda_{21}^2(\lambda_{12}+\lambda_{21})\hat{\mu}_2^4 + \lambda_{13}^2\lambda_{21}(3\lambda_{31}(\lambda_{12}^2+\lambda_{21}^2) \\ & \quad + 2\lambda_{12}\lambda_{21}(\lambda_{12}-\lambda_{13}+\lambda_{21}+5\lambda_{31}))\hat{\mu}_2^3 + \lambda_{13}[\lambda_{31}(\lambda_{12}+\lambda_{21})(\lambda_{31}(\lambda_{12}^2+\lambda_{21}^2) \\ & \quad + \lambda_{21}\lambda_{12}(3(\lambda_{12}+\lambda_{21}) - 2\lambda_{13}(n+2) + 8\lambda_{31})) - 2\lambda_{12}^2\lambda_{13}\lambda_{21}^2]\hat{\mu}_2^2 \\ & \quad + \lambda_{12}\lambda_{31}[2(\lambda_{12}+\lambda_{21})^2\lambda_{31}^2 - \lambda_{13}\lambda_{31}((\lambda_{12}^2+\lambda_{21}^2)(n+1) + 2\lambda_{12}\lambda_{21}(3n+2)) \\ & \quad + \lambda_{31}(\lambda_{12}+\lambda_{21})^3 - \lambda_{12}\lambda_{13}\lambda_{21}(\lambda_{12}+\lambda_{21})(2n+3) + 2n\lambda_{12}\lambda_{13}^2\lambda_{21}]\hat{\mu}_2 \\ & \quad + -\lambda_{12}^2\lambda_{31}n(\lambda_{31}(\lambda_{12}+\lambda_{21})(\lambda_{12}-\lambda_{13}+\lambda_{21}+2\lambda_{31}) - 2\lambda_{12}\lambda_{13}\lambda_{21}) \\ & \qquad = 0.\end{aligned} \tag{17}$$

These equations are of course rather cumbersome, but can be used to find numerical results, as we do in the following section.

3. Numerical evaluations

Tables 1 and 2 show the results of evaluations of both the stochastic and deterministic models, along with the normal approximation to the stochastic model, for a variety of parameter values. The λ_{ij}'s were randomly generated; $\lambda_{12}, \lambda_{21}, \lambda_{31} \sim U(0, 50)$ and $\lambda_{13} \sim U(0, 2.5)$, the distributions chosen to ensure that the transition rates were, on average, of similar order (λ_{13} being associated with the only quadratic transition). Table 1 shows μ_1, μ_2, σ_1^2, σ_2^2, σ_{12}, $k_{2,1}$ and $k_{1,2}$, while Table 2 shows μ_3, σ_3^2, σ_{13} and σ_{23}.

Since $\hat{\mu}_2$ satisfies a quartic equation, the normal approximation gives four values for $\hat{\mu}_2$. Only one of the solutions of Eq. (17) was a plausible value for $\hat{\mu}_2$ for each set of parameters used here. The normal approximation is very accurate and all of the estimates of σ_3^2, σ_{13} and σ_{23} are fairly good. In particular, the normal approximation agrees more closely with the stochastic model than does the deterministic model for each parameter set.

The deterministic model, stochastic model and its normal approximation were evaluated for 400 sets of parameters with the results summarised in Table 3. In each case the population size was either 10, 20, 30, 40 or 50. For the first 200 sets (Run 1 in Table 3) λ_{12}, λ_{21} and $\lambda_{31} \sim U(0, n)$ and $\lambda_{13} \sim U(0, 3)$ so that the transition rates were of similar order. The covariance between X_1 and X_2, σ_{12}, is negative for each of these sets of parameters. In the last 200 parameter sets (Run 2 in Table 3) $\lambda_{12} \sim U(0.5, 1.5)$, $\lambda_{13}, \lambda_{21} \sim U(10^6, 1.0001 \times 10^{10})$ and $\lambda_{31} \sim U(0, 0.1)$. These distributions were chosen to give parameter sets for which σ_{12} may be positive. If σ_{12} is to be positive it must be when μ_2 is small, since expression (14) for σ_{12} in terms of μ_2 is decreasing in μ_2. Also, $n\lambda_{31}/\lambda_{13}$ must be larger than $(\lambda_{21}/\lambda_{12})\mu_2^2 + (\lambda_{31}(\lambda_{12}+\lambda_{21})/\lambda_{13}\lambda_{12})\mu_2$. They were chosen as a result of carrying out some evaluations for which σ_{12} was close to zero. When the normal approximation provided more than one set of values for the cumulants, the set which gave the means closest to those given by the deterministic model were used.

From Table 1 and Run 1 in Table 3 we can see that generally the deterministic model underestimates the number of individuals in states S_1 and S_2, i.e. that

Table 1
The mean, variance and covariance of X_1 and X_2 for a variety of transition rates

Model	λ_{12}	λ_{13}	λ_{21}	λ_{31}	μ_1	μ_2	σ_{12}	σ_1^2	σ_2^2	$k_{2,1}$	$k_{1,2}$
S	26.39	1.297	23.80	11.81	13.4338	14.8958	−2.78	10.92	11.81	−2.06	−1.98
D	26.39	1.297	23.80	11.81	13.3769	14.8327					
N	26.39	1.297	23.80	11.81	13.4348	14.8968	−2.83	10.88	11.76		
S	13.50	2.380	13.96	3.281	7.1911	6.9542	−0.58	6.59	6.39	−0.52	−0.53
D	13.50	2.380	13.96	3.281	7.1562	6.9204					
N	13.50	2.380	13.96	3.281	7.1931	6.9561	−0.61	6.56	6.36		
S	17.47	2.300	32.69	30.01	20.9940	11.2195	−3.47	14.51	9.37	−1.89	−2.63
D	17.47	2.300	32.69	30.01	20.9123	11.1758					
N	17.47	2.300	32.69	30.01	20.9950	11.2200	−3.51	14.47	9.33		
S	49.44	1.516	35.76	9.751	10.7328	14.8386	−2.13	9.19	11.89	−1.72	−1.57
D	49.44	1.516	35.76	9.751	10.6853	14.7730					
N	49.44	1.516	35.76	9.751	10.7338	14.8400	−2.18	9.15	11.84		
S	46.02	1.274	46.33	4.207	10.0455	9.9783	−1.25	8.79	8.74	−1.05	−1.06
D	46.02	1.274	46.33	4.207	9.9983	9.9314					
N	46.02	1.274	46.33	4.207	10.0470	9.9798	−1.29	8.75	8.70		
S	44.00	0.2067	34.01	21.19	19.7018	25.4890	−9.16	12.62	13.64	−3.16	−1.40
D	44.00	0.2067	34.01	21.19	19.6698	25.4476					
N	44.00	0.2067	34.01	21.19	19.7019	25.4890	−9.17	12.61	13.63		
S	23.63	1.273	16.87	40.05	16.0989	22.5499	−5.91	11.88	14.27	−3.39	−2.38
D	23.63	1.273	16.87	40.05	16.0499	22.4813					
N	23.63	1.273	16.87	40.05	16.0992	22.5503	−5.94	11.85	14.25		
S	49.66	0.1327	2.763	13.47	2.5746	46.2746	−2.33	2.45	4.41	−2.10	1.79
D	49.66	0.1327	2.763	13.47	2.5735	46.2538					
N	49.66	0.1327	2.763	13.47	2.5746	46.2746	−2.33	2.44	4.42		
S	17.15	0.9662	10.15	0.9526	4.6964	7.9353	−0.42	4.45	7.22	−0.40	−0.38
D	17.15	0.9662	10.15	0.9526	4.6734	7.8964					
N	17.15	0.9662	10.15	0.9526	4.6979	7.9379	−0.45	4.43	7.18		
S	5.771	0.2995	19.09	28.51	35.3951	10.7001	−7.02	12.16	8.58	1.53	−4.44
D	5.771	0.2995	19.09	28.51	35.3468	10.6855					
N	5.771	0.2995	19.09	28.51	35.3952	10.7001	−7.03	12.18	8.56		

$n = 50$ in all cases. The letters in the Model column correspond to each of the models. S stands for stochastic, D for deterministic and N for normal approximation. The values for the deterministic model in the 'mean' columns are nz_1 and nz_2, respectively. The stochastic model and normal approximation are described in Sections 2.2 and 2.2.3, respectively.

$nz_1 < \mu_1$ and $nz_2 < \mu_2$. Thus there is a consistent bias when using the large population approximation, when the values of the transition rates are broadly comparable. Associated with this, the covariance of X_1 and X_2 was negative in each case (as we may expect because there is a total of n individuals to divide between these categories). However, the covariance between X_1 and X_2 was positive in 180 out of the 200 cases from Run 2 in Table 3 (which used values evaluated in order to try to make this occur). It follows from Results 1 and 2 in Section 2.2 that $nz_1 > \mu_1$ and $nz_2 > \mu_2$ in these cases, and so sometimes the more usual result above does not occur. Note that, as can be seen, the bias is very small and the large population approximation predicts the expectation of the number of individuals in these states well (although the variance can be quite large). The use of deterministic models for the various previous works on kleptoparasitism such as Broom and Ruxton (1998) thus seems reasonable.

The other two covariances were not positive for any of the parameter sets investigated. Thus the number in the fighting category S_3 seems always negatively correlated with the numbers in either of the other categories. It was found that $\sigma_1^2 < \sigma_2^2 \iff \mu_1 < \mu_2$ and $\sigma_1^2 > \sigma_2^2 \iff \mu_1 > \mu_2$ for all cases evaluated. This again makes sense; the larger the expected number on a given site, the larger the magnitude of the variance of this number. We have not been able to prove the generality of these results.

For every set of parameters tested $\mu_i \leqslant \hat{\mu}_i$, $i = 1, 2$ and $\sigma_{12} > \hat{\sigma}_{12}$. $\hat{\mu}_i$ and $\hat{\sigma}_{12}$ are the estimates of μ_i and σ_{12} under the normal approximation. Thus the normal approximation seems to always overestimate μ_1 and μ_2 (although this overestimate is very small). It has, again, not been possible to prove the generality of this result.

4. Discussion

This paper considers both a deterministic and a stochastic model for kleptoparasitism, along with a normal approximation to the stochastic model. The deterministic model is that of Broom and Ruxton (1998), for which explicit solutions for the equilibrium proportion of the population on each site can be found. The forward Kolmogorov equations were derived for the equilibrium

Table 2
The mean and variance of the number of individuals on site 3, along with the covariances of the number on site 3 with the numbers of individuals on sites 1 and 2, for a variety of transition rates

Model	λ_{12}	λ_{13}	λ_{21}	λ_{31}	μ_3	σ_{13}	σ_{23}	σ_3^2
S	26.39	1.297	23.80	11.81	21.6704	−8.14	−9.02	17.16
D	26.39	1.297	23.80	11.81	21.7904			
N	26.39	1.297	23.80	11.81	21.6685	−8.05	−8.93	16.98
S	13.50	2.380	13.96	3.281	35.8547	−6.01	−5.81	11.82
D	13.50	2.380	13.96	3.281	35.9235			
N	13.50	2.380	13.96	3.281	35.8508	−5.95	−5.75	11.70
S	17.47	2.300	32.69	30.01	17.7865	−11.04	−5.90	16.94
D	17.47	2.300	32.69	30.01	17.9119			
N	17.47	2.300	32.69	30.01	17.7850	−10.96	−5.82	16.78
S	49.44	1.516	35.76	9.751	24.4286	−7.06	−9.76	16.81
D	49.44	1.516	35.76	9.751	24.5417			
N	49.44	1.516	35.76	9.751	24.4263	−6.98	−9.66	16.64
S	46.02	1.274	46.33	4.207	29.9763	−7.54	−7.49	15.03
D	46.02	1.274	46.33	4.207	30.0702			
N	46.02	1.274	46.33	4.207	29.9732	−7.46	−7.41	14.87
S	44.00	0.2067	34.01	21.19	4.8092	−3.46	−4.48	7.94
D	44.00	0.2067	34.01	21.19	4.8826			
N	44.00	0.2067	34.01	21.19	4.8091	−3.44	−4.46	7.90
S	23.63	1.273	16.87	40.05	11.3511	−5.97	−8.36	14.33
D	23.63	1.273	16.87	40.05	11.4688			
N	23.63	1.273	16.87	40.05	11.3506	−5.91	−8.31	14.22
S	49.66	0.1327	2.763	13.47	1.1508	−0.12	−2.08	2.20
D	49.66	0.1327	2.763	13.47	1.1727			
N	49.66	0.1327	2.763	13.47	1.1507	−0.11	−2.09	2.20
S	17.15	0.9662	10.15	0.9526	37.3683	−4.02	−6.79	10.81
D	17.15	0.9662	10.15	0.9526	37.4301			
N	17.15	0.9662	10.15	0.9526	37.3642	−3.97	−6.73	10.70
S	5.771	0.2995	19.09	28.51	3.9048	−5.13	−1.55	6.69
D	5.771	0.2995	19.09	28.51	3.9677			
N	5.771	0.2995	19.09	28.51	3.9047	−5.14	−1.52	6.67

$n = 50$ in all cases. The letters in the Model column correspond to each of the models. S stands for stochastic, D for deterministic and N for normal approximation. The value for the deterministic model in the 'mean' column is nz_3. The stochastic model and normal approximation are described in Sections 2.2 and 2.2.3, respectively.

Table 3
A summary of numerical evaluations of the deterministic and stochastic models and the normal approximation

Characteristic	Run 1	Run 2
Total evaluations	200	200
$\sigma_{12} > 0$	0	180
$\mu_1 < \mu_2$ and $\sigma_1^2 > \sigma_2^2$	0	0
$\mu_2 < \mu_1$ and $\sigma_2^2 > \sigma_1^2$	0	0
$\mu_1 > \mu_3$	0	0
$\sigma_{13} > 0$	0	0
$\sigma_{23} > 0$	0	0
Multiple possible sets of parameters under normal approximation	7	196

The numbers in the Run 1 and Run 2 columns give the number of times each characteristic was observed in that set of evaluations. $n = 10, 20, 30, 40$ or 50. For Run 1 $\lambda_{12}, \lambda_{21}, \lambda_{31} \sim U(0, n), \lambda_{13} \sim U(0, 3)$ and for Run 2 $\lambda_{12} \sim U(0.5, 1.5)$, $\lambda_{13}, \lambda_{21} \sim U(10^6, 1.0001 \times 10^{10})$, $\lambda_{31} \sim U(0, 0.1)$.

distribution of the stochastic model. It is not practical to solve these explicitly, but numerical solutions have been obtained for many sets of parameters.

A partial differential equation for the cumulant generating function was derived for the stochastic model. A set of five equations for the means, variances and covariance of the number of individuals on sites 1 and 2 has been obtained from this. These equations contain seven unknowns, since they also contain the cumulants $k_{2,1}$ and $k_{1,2}$. The normal approximation reduces the number of unknowns, since $k_{2,1}$ and $k_{1,2}$ are assumed to be zero under this approximation. These equations are not linear, and the solution is given in the form of a quartic equation for μ_2 and expressions for μ_1, σ_1^2, σ_2^2 and σ_{12} in terms of μ_2.

The stochastic model generally gave larger numbers of individuals on sites 1 and 2 than the deterministic model. This corresponds to a larger number of searchers and handlers, and a smaller number of individuals fighting. Thus, the deterministic approximation will often overestimate the number of individuals involved in a contest, and we may expect to see less conflict than is predicted by our models, especially in the case of small populations, where the differences with the infinite population approximation will be largest. One possible reason for this is as follows: if the number of individuals on S_1 and S_2 in equilibrium were equal, then any movement between the two would reduce the product X_1X_2, and so the rate of movement towards the fighting state S_3 would generally be reduced by this variability. This argument may still hold when X_1 and X_2 are of similar order. The differences between stochastic and deterministic models were always small, however, indicating the validity of the use of the deterministic model for these kind of systems. The difference between these values is related to the sign of the covariance between the numbers of individuals on sites 1 and 2 as shown in Section 3. There were some cases for which this covariance is positive, meaning that the deterministic model gives the larger numbers of individuals on sites 1 and 2. These were associated when the numbers in S_1 and S_2 were very uneven (in particular X_2 was small), so the above argument may be reversed, and variability may increase X_1X_2 on average.

The normal approximation performs well, particularly for estimating the means, where these were a lot closer to the means given by the stochastic model than were those given by the deterministic model. In fact the differences between the actual means and the normal approximations were almost zero. The approach of using Normal approximations to simplify the analysis of such systems seems a possible way forward. For any such system of equations we would have a multivariate normal random variable of dimension k, the total number of sites. Although the quartic equation and expressions for the variances are long, they can be evaluated numerically much more quickly than the equilibrium

Kolmogorov equations can be solved. This will be especially true if, as in our kleptoparasitism model, we are interested in the distribution of the number of individuals involved in a particular behaviour, for an intermediately sized population (for a very small population, we could find numerical solutions to the Kolmogorov equations themselves).

The deterministic nature of earlier game-theoretic models of kleptoparasitism can thus be considered justified. These models, starting with Broom and Ruxton (1998), showed that stealing is optimal under certain conditions even though this inevitably reduced the food intake of the population. Stealing in particular was profitable when fights were not costly, and food was hard to find. One important prediction was that there should be a step change in behaviour when food availability crossed a certain threshold, so that small ecological effects would have a dramatic effect on the behaviour of the population and so on the food consumption rate. The potential for kleptoparasitism can thus have a serious effect on the fitness of a species and on its evolution; it should be noted that many species find much of their food through kleptoparasitism (see for example Brockmann and Barnard 1979).

The kleptoparasitism model that is our focus here is one of several that considers competition over food in various scenarios. Another such model is the producer–scrounger game (for example see Caraco and Giraldeau, 1991; Dubois and Giraldeau, 2005). The producer–scrounger model starts from rather different assumptions to the kleptoparasitism models, for example animals search for static food clumps which contain several food items with negligible handling times and are either producers or scroungers; our models assume single food items which require a non-zero handling time once discovered, and individuals search for both food and stealing opportunities. The dynamic structure of the kleptoparasitism models, with individuals being tracked between the different states, is also absent. The solutions from both models are influenced by similar factors such as effectiveness of foraging and the effect of population density, although both of these are introduced in rather different ways. Thus the models can be thought of as complementary. Another model of similar type is described in Auger et al. (2006) which is a deterministic predator–prey model. Individuals are tracked between states as in our model, so in this sense it resembles our model more than the producer–scrounger models. The contests for food are different to ours, and involve individuals changing defensive strategies but not challenging ones. An interesting attempt to unify some of the different models (but not Auger et al., 2006) is described by Vahl (2006) who examines the assumptions, both explicit and implicit, of the key food competition models. It would be of interest to develop similar stochastic versions of some of these other related models, and make similar comparisons as we do here.

Appendix A. Derivation of Eq. (7)

The first step in deriving this equation is to multiply Eq. (6) by $\exp\{s_1x_1+s_2x_2\}$ and add over all values of X_1 and X_2. This gives

$$\begin{aligned}\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}&\exp\{s_1x_1+s_2x_2\}\frac{\mathrm{d}}{\mathrm{d}t}p_{x_1,x_2}(t)\\ =&\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}\exp\{s_1x_1+s_2x_2\}[\lambda_{13}(x_1+1)(x_2+1)\\ &\times p_{x_1+1,x_2+1}(t)+\lambda_{31}(n-x_1-x_2+2)p_{x_1-1,x_2-1}(t)\\ &+\lambda_{12}(x_1+1)p_{x_1+1,x_2-1}(t)+\lambda_{21}(x_2+1)p_{x_1-1,x_2+1}(t)\\ &-[\lambda_{13}x_1x_2+\lambda_{31}(n-x_1-x_2)+\lambda_{12}x_1+\lambda_{21}x_2]\\ &\times p_{x_1,x_2}(t)].\end{aligned}\tag{18}$$

At equilibrium, the left-hand side of this equation reduces to 0, while each sum on the right can be expressed in terms of $M(s_1,s_2)$, $\partial M/\partial s_1$, $\partial M/\partial s_2$ and $\partial^2 M/\partial s_1\partial s_2$. The first term on the right-hand side is equivalent to

$$\begin{aligned}\sum_{x_1=0}^{n-2}\sum_{x_2=0}^{n-x_1-2}&\lambda_{13}(x_1+1)(x_2+1)p_{x_1+1,x_2+1}\exp\{s_1x_1+s_2x_2\}\\ &=\lambda_{13}\sum_{x_1=1}^{n-1}\sum_{x_2=1}^{n-x_1}x_1x_2p_{x_1,x_2}\exp\{s_1(x_1-1)+s_2(x_2-1)\}.\end{aligned}$$

Since

$$M(s_1,s_2)=\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}\exp\{s_1x_1+s_2x_2\}p_{x_1,x_2},$$

this is equivalent to

$$\lambda_{13}\exp\{-(s_1+s_2)\}\frac{\partial^2 M}{\partial s_1\partial s_2}.$$

The second term on the right-hand side of Eq. (18) is equivalent to

$$\begin{aligned}\sum_{x_1=1}^{n-1}\sum_{x_2=1}^{n-x_1}&\lambda_{31}(n-x_1-x_2+2)p_{x_1-1,x_2-1}\exp\{s_1x_1+s_2x_2\}\\ &=\lambda_{31}\sum_{x_1=0}^{n-2}\sum_{x_2=0}^{n-x_1-2}(n-x_1-x_2)p_{x_1,x_2}\\ &\quad\times\exp\{s_1(x_1+1)+s_2(x_2+1)\}.\end{aligned}$$

Since $n-x_1-x_2=0$ when $x_1+x_2=n$ and $p_{x_1,x_2}=0$ when $x_1+x_2=n-1$, the sum above is equivalent to

$$\begin{aligned}\lambda_{31}\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}&(n-x_1-x_2)p_{x_1,x_2}\exp\{s_1(x_1+1)+s_2(x_2+1)\}\\ &=\lambda_{31}\exp\{s_1+s_2\}\left(nM-\frac{\partial M}{\partial s_1}-\frac{\partial M}{\partial s_2}\right).\end{aligned}$$

The third term on the right-hand side of Eq. (18) is equivalent to

$$\sum_{x_1=0}^{n-1}\sum_{x_2=1}^{n-x_1}\lambda_{12}(x_1+1)p_{x_1+1,x_2-1}\exp\{s_1x_1+s_2x_2\}$$

$$=\lambda_{12}\sum_{x_1=1}^{n}\sum_{x_2=0}^{n-x_1}x_1p_{x_1,x_2}\exp\{s_1(x_1-1)+s_2(x_2+1)\}$$

$$=\lambda_{12}\frac{\partial M}{\partial s_1}.$$

Similarly

$$\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}\lambda_{21}(x_2+1)p_{x_1-1,x_2+1}\exp\{s_1x_1+s_2x_2\}=\lambda_{21}\frac{\partial M}{\partial s_2}$$

and

$$\sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1}[\lambda_{13}x_1x_2+\lambda_{31}(n-x_1-x_2)+\lambda_{12}x_1+\lambda_{21}x_2]p_{x_1,x_2}$$
$$\times\exp\{s_1x_1+s_2x_2\}$$
$$=\lambda_{13}\frac{\partial^2M}{\partial s_1\partial s_2}+\lambda_{31}\left(nM-\frac{\partial M}{\partial s_1}-\frac{\partial M}{\partial s_2}\right)+\lambda_{12}\frac{\partial M}{\partial s_1}+\lambda_{21}\frac{\partial M}{\partial s_2}.$$

Inserting each of these expressions into Eq. (18) gives the following partial differential equation for $M(s_1,s_2)$:

$$n\lambda_{31}(e^{s_1+s_2}-1)M+[\lambda_{12}(e^{s_2-s_1}-1)+\lambda_{31}(1-e^{s_1+s_2})]\frac{\partial M}{\partial s_1}$$
$$+[\lambda_{21}(e^{s_1-s_2}-1)+\lambda_{31}(1-e^{s_1+s_2})]\frac{\partial M}{\partial s_2}$$
$$+\lambda_{13}(e^{-(s_1+s_2)}-1)\frac{\partial^2M}{\partial s_1\partial s_2}=0. \tag{19}$$

Since $K(s_1,s_2)=\log M(s_1,s_2)$,

$$\frac{\partial K}{\partial s_i}=\frac{1}{M}\frac{\partial M}{\partial s_i},\quad i=1,2\quad\text{and}$$

$$\frac{\partial^2K}{\partial s_1\partial s_2}=\frac{1}{M}\frac{\partial^2M}{\partial s_1\partial s_2}-\frac{1}{M^2}\frac{\partial M}{\partial s_1}\frac{\partial M}{\partial s_2}.$$

Thus

$$\frac{\partial M}{\partial s_i}=M\frac{\partial K}{\partial s_i},\quad i=1,2$$

and

$$\frac{\partial^2M}{\partial s_1\partial s_2}=M\left(\frac{\partial^2K}{\partial s_1\partial s_2}+\frac{1}{M^2}\frac{\partial M}{\partial s_1}\frac{\partial M}{\partial s_2}\right)$$

$$=M\left(\frac{\partial^2K}{\partial s_1\partial s_2}+\frac{\partial K}{\partial s_1}\frac{\partial K}{\partial s_2}\right).$$

These expressions can be substituted into Eq. (19) to give Eq. (7).

Appendix B. Derivation of the expressions for μ_1, σ_1^2, σ_2^2 and σ_{12} in terms of μ_2 and the quartic equation for μ_2 under the normal approximation

Firstly, expressions for μ_1 and σ_{12} are derived from Eqs. (8) and (9). Since neither equation includes the terms $k_{2,1}$ and $k_{1,2}$, the expressions for $\hat{\mu}_1$ and $\hat{\sigma}_{12}$ are identical to those for μ_1 and σ_{12}, except that μ_2 is replaced by $\hat{\mu}_2$ in each. It is clear from Eq. (9) that

$$\mu_1=\frac{\lambda_{21}\mu_2}{\lambda_{12}}.$$

Substitution of this expression for μ_1 into Eq. (8) gives

$$n\lambda_{31}-\lambda_{31}\frac{(\lambda_{12}+\lambda_{21})}{\lambda_{12}}\mu_2-\frac{\lambda_{13}\lambda_{21}}{\lambda_{12}}\mu_2^2-\lambda_{13}\sigma_{12}=0,$$

which can be rearranged to give

$$\sigma_{12}=-\frac{\lambda_{21}}{\lambda_{12}}\mu_2^2-\frac{\lambda_{31}(\lambda_{12}+\lambda_{21})}{\lambda_{13}\lambda_{12}}\mu_2+\frac{n\lambda_{31}}{\lambda_{13}}.$$

Setting $k_{2,1}=k_{1,2}=0$ in Eqs. (10)–(12), gives

$$n\lambda_{31}-\lambda_{31}\hat{\mu}_1+(\lambda_{21}-\lambda_{31})\hat{\mu}_2-(\lambda_{12}+\lambda_{31})\hat{\sigma}_1^2+(\lambda_{21}-\lambda_{31})\hat{\sigma}_{12}$$
$$-\lambda_{13}\hat{\mu}_1\hat{\sigma}_{12}-\lambda_{13}\hat{\mu}_2\hat{\sigma}_1^2=0, \tag{20}$$

$$\lambda_{12}\hat{\mu}_1+\lambda_{21}\hat{\mu}_2-\lambda_{12}\hat{\sigma}_1^2-\lambda_{21}\hat{\sigma}_2^2+(\lambda_{12}+\lambda_{21})\hat{\sigma}_{12}=0 \tag{21}$$

and

$$n\lambda_{31}+(\lambda_{12}-\lambda_{31})\hat{\mu}_1-\lambda_{31}\hat{\mu}_2-(\lambda_{21}+\lambda_{31})\hat{\sigma}_2^2+(\lambda_{12}-\lambda_{31})\hat{\sigma}_{12}$$
$$-\lambda_{13}\hat{\mu}_2\hat{\sigma}_{12}-\lambda_{13}\hat{\mu}_1\hat{\sigma}_2^2=0. \tag{22}$$

When the above expressions for $\hat{\mu}_1$ and $\hat{\sigma}_{12}$ are substituted into Eq. (20) the following equation involving $\hat{\mu}_2$ and $\hat{\sigma}_1^2$ is obtained:

$$-(\lambda_{12}+\lambda_{31}+\lambda_{13}\hat{\mu}_2)\hat{\sigma}_1^2+\frac{\lambda_{13}\lambda_{21}^2}{\lambda_{12}^2}\hat{\mu}_2{}^3+\frac{\lambda_{21}(\lambda_{21}\lambda_{31}+2\lambda_{12}\lambda_{31}-\lambda_{12}\lambda_{21})}{\lambda_{12}^2}\hat{\mu}_2{}^2$$
$$-\frac{\lambda_{31}(\lambda_{13}\lambda_{21}(n+1)+\lambda_{12}\lambda_{13}+\lambda_{21}^2+\lambda_{12}\lambda_{21})-\lambda_{31}^2(\lambda_{12}+\lambda_{21})-\lambda_{12}\lambda_{13}\lambda_{21}}{\lambda_{12}\lambda_{13}}\hat{\mu}_2+\frac{\lambda_{31}(\lambda_{13}+\lambda_{21}-\lambda_{31})n}{\lambda_{13}}=0.$$

The coefficient of $\hat{\sigma}_1^2$ in this equation is negative, since λ_{12} λ_{13} and λ_{31} are all positive and $\hat{\mu}_2\geqslant 0$. This means that the equation can be rearranged, to give an expression for $\hat{\sigma}_1^2$ in terms of $\hat{\mu}_2$. Expression (15) is the partial fraction expansion of this.

Similarly, substituting the expressions for $\hat{\mu}_1$ and $\hat{\sigma}_{12}$ into Eq. (22) gives

$$
\begin{aligned}
&-\left(\lambda_{21}+\lambda_{31}+\frac{\lambda_{13}\lambda_{21}}{\lambda_{12}}\hat{\mu}_2\right)\hat{\sigma}_2^2+\frac{\lambda_{13}\lambda_{21}}{\lambda_{12}}\hat{\mu}_2^3+\frac{\lambda_{12}\lambda_{31}+2\lambda_{21}\lambda_{31}-\lambda_{12}\lambda_{21}}{\lambda_{12}}\hat{\mu}_2^2\\
&\quad-\frac{\lambda_{31}(\lambda_{12}\lambda_{13}(n+1)+\lambda_{12}^2+\lambda_{12}\lambda_{21}+\lambda_{13}\lambda_{21})-\lambda_{31}^2(\lambda_{12}+\lambda_{21})-\lambda_{12}\lambda_{13}\lambda_{21}}{\lambda_{12}\lambda_{13}}\hat{\mu}_2+\frac{\lambda_{31}(\lambda_{12}+\lambda_{13}-\lambda_{31})n}{\lambda_{13}}=0.
\end{aligned}
$$

Since all of the λ_{ij}'s are positive and $\hat{\mu}_2 \geqslant 0$, this equation can be rearranged to give $\hat{\sigma}_2^2$ in terms of $\hat{\mu}_2$. Expression (16) is the partial fraction expansion of this.

Finally, the quartic equation for $\hat{\mu}_2$ is obtained by inserting the expressions for $\hat{\mu}_1$, $\hat{\sigma}_{12}$, $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ into Eq. (21). When all terms on the left-hand side of this equation are placed over a common denominator, we have the left-hand side of Eq. (17) as the numerator and $\lambda_{12}\lambda_{13}(\lambda_{12}+\lambda_{31}+\lambda_{13}\hat{\mu}_2)(\lambda_{12}\lambda_{21}+\lambda_{12}\lambda_{31}+\lambda_{13}\lambda_{21}\hat{\mu}_2)$ as the denominator. This denominator can be removed, since it is always positive.

References

Auger, P., Kooi, B.W., Bravo de la Parra, R., Poggiale, J.-C., 2006. Bifurcation analysis of a predator-prey model with predators using hawk and dove tactics. J. Theor. Biol. 238, 597–607.

Brockmann, H.J., Barnard, C.J., 1979. Kleptoparasitism in birds. Anim. Behav. 27, 487–514.

Broom, M., Luther, R.M., Ruxton, G.D., 2004. Resistance is useless?—extensions to the game theory of kleptoparasitism. Bull. Math. Biol. 66, 1645–1658.

Broom, M., Ruxton, G.D., 1998. Evolutionarily stable stealing: game theory applied to kleptoparasitism. Behav. Ecol. 9, 397–403.

Broom, M., Ruxton, G.D., 2003. Evolutionarily stable kleptoparasitism: consequences of different prey types. Behav. Ecol. 14, 23–33.

Caraco, T., Giraldeau, L.A., 1991. Social foraging—producing and scrounging in a stochastic environment. J. Theor. Biol. 153, 559–583.

Dubois, F., Giraldeau, L.A., 2005. Fighting for resources: the economics of defense and appropriation. Ecology 86, 3–11.

Durrett, R., Granovsky, B.L., Gueron, S., 1998. The steady-state distributions of coagulation–fragmentation processes. J. Math. Biol. 37, 1–27.

Durrett, R., Granovsky, B.L., Gueron, S., 1999. The equilibrium behaviour of reversible coagulation–fragmentation processes. J. Theor. Probab. 12, 447–474.

Fretwell, S.D., Lucas, J.H.J., 1970. On territorial behaviour and other factors influencing habitat distribution in birds. I. Theoretical development. Acta Biotheoret. 19, 16–36.

Hamilton, I.M., 2002. Kleptoparasitism and the distribution of unequal competitors. Behav. Ecol. 13, 260–267.

Homer, O.P., Wachter, B., East, M.L., Hofer, H., 2002. The response of spotted hyaenas to long-term changes in prey populations: functional response and interspecific kleptoparasitism. J. Animal Ecol. 71, 236–246.

Houston, A.I., McNamara, J.M., 1988. The ideal free distribution when competitive abilities differ—an approach based on statistical-mechanics. Animal Behav. 36, 166–174.

Hugie, D.M., Grand, T.C., 1998. Movement between patches, unequal competitors and the ideal free distribution. Evol. Ecol. 12, 1–19.

Jackson, A.L., Humphries, S., Ruxton, G.D., 2004. Resolving the departures of observed results from the ideal free distribution with simple random movements. J. Animal Ecol. 73, 612–622.

Kennedy, M., Gray, R.D., 1993. Can ecological theory predict the distribution of foraging animals—a critical review of experiments on the ideal free distribution. OIKOS 68, 158–166.

Kryscio, R.J., Lefèvre, C., 1989. On the extinction of the SIS stochastic logistic epidemic. J. Appl. Probab. 27, 685–694.

Luther, R.M., Broom, M., 2004. Rapid convergence to an equilibrium state in kleptoparasitic populations. J. Math. Biol. 48, 325–339.

Nasell, I., 1996. The quasi-stationary distribution of the closed endemic SIS model. Adv. Appl. Probab. 28, 895–932.

Nasell, I., 1999. On the quasi-stationary distribution of the stochastic logistic epidemic. Math. Biosci. 156, 21–40.

Nasell, I., 2003. Moment closure and the stochastic logistic model. Theoret. Popul. Biol. 62, 159–168.

Ruxton, G.D., Humphries, S., 2003. Multiple ideal free distributions of unequal competitors. Evol. Ecol. Res. 1, 635–640.

Ruxton, G.D., Moody, A.L., 1997. The ideal free distribution with kleptoparasitism. J. Theor. Biol. 186, 449–458.

Vahl, W.K., 2006. Interference competition among foraging waders. Ph.D. Thesis, University of Groningen.

Vollrath, F., 1979. Behaviour of the kleptoparasitic spider argyrodes elevatus (araneae, theridiidae). Animal Behav. 27, 515–521.

Weiss, G.H., Dishon, M., 1971. On the asymptotic behavior of the stochastic and deterministic models of an epidemic. Math. Biosci. 11, 261–265.

Whittle, P., 1957. On the use of the normal approximation in the treatment of stochastic processes. J. Roy. Statist. Soc. Ser. B 19, 268–281.

Yates, G.E., Broom, M., 2005. A stochastic model of the distribution of unequal competitors between resource patches. J. Theor. Biol. 237, 227–237.