---

title: "Underdispersion in Macroparasite Models"

author:

- affiliation: UCD

email: nmross\@ucdavis.edu

footnote: Corresponding author

name: Noam Ross

date: "August 17, 2015"

output:

pdf\_document:

dev:

- pdf

fig\_caption:

- yes

fig\_height:

- 6

fig\_width:

- 11

keep\_tex:

- yes

template:

- elsarticle.latex

csl: ecology-letters.csl

address:

- address: |

Department of Environmental Science and Policy

University of California-Davis

1 Shields Avenue Davis, CA 95616 USA

code: UCD

documentclass: elsarticle

layout: review,12pt

linenumbers: yes

bibliography: references.bib

abstract: ""

---

```{r get\_chunks, echo=FALSE, message=FALSE, warning=FALSE, cache=FALSE}

library(knitr)

opts\_knit$set(eval.after = "fig.cap")

opts\_chunk$set(echo=FALSE, cache=TRUE, message=FALSE, warning=FALSE)

```

# Abstract

In multi-infection, or macroparasite, models of disease, distribution of

infections among individuals are critical to host-pathogen dynamics. Traditionally

these distributions are modeled using Negative Binomial distributions. Here

I show that the Conway-Maxwell Poisson (CMP) distribution [@Conway1962], provides a better approximation of the evolving distribution of infections in classic macroparasite models. I also introduce an R package, \*\*cmp\*\*, for fitting this distribution to data.

BIG THINGS FOR OUTLINING AND WRITING TOMORROW

- Underdispersion can take place in the classical model in the absence of density dependence

- In and near the unstable parameter space, even at the edge of the stable parameter space.

-

- Traditionally analyze the stable parameter space [@Pugliese... proved where this exists, @Kretzschmar showed this was where mean/var grew with mean]

- Other space is important as transient dynamics are the norm [@Hastings...], in the case of diseases driving organisms extinct locally (bats, frogs)

- CMP fits the under-distributed data better, and somewhat in over-dispersed data, as well.

- Stochastic dynamics reduce this effect

# Introduction

Aggregation of parasites within hosts is a very common phenomenon [@Shaw1995]. So much so that it has been described as a "law" of parasite ecology [@Poulin2007a]. Aggregation describes the phenomenon of a small number of hosts housing larger fraction of the parasite population than would be expected by change. It is characterized by the distribution of parasite loads having a variation larger than its mean, or overdispersion. This arises from heterogeneities such as individual variation in host susceptibility and exposure to parasites [@Keymar1979; @Anderson1982], and "clumping" of parasites that colonize hosts together [@Pugliese1998]. Even in the absence of such heterogeneities, basic host-macroparasite models exhibit variance/mean ratios above one [@Adler1992].

In rare cases, parasite loads may be underdispersed, with parasites more evenly distributed among hosts. Forces driving underdispersion include density-dependent parasite mortality, nonlinear parasite-induced host mortality [@Barbour2000]. Underdispersed parasite disributions have been found in oxyuroid parasites of cockroaches [@Muller-Graf2001; Zervos1988], in which sexual competition appears to induce density dependence. There are also a variety of examples in fish populations: @Burn1980 found nearly all Deretrema parasites in flashlight fish (\*Anomalops katopteron\*) were in groups of 2. In many cases, parasite numbers greater than one are rare [@Donnelly1994; Adlard1994; @Kennedy1979]. Often, this density dependence is driven by limited space in the host [@Uebelacker1978].

In parasite populations in hosts are most frequently modelled using the negative binomial distribution. It has performed very well for field data [@Pacala1998].

The negative binomial distribution has been used to represent overdisperson in

the @anderson1978 model. It has proven an excellent empirical distribution in

the field [@Pacala1998]. It represents over-dispersed populations well.

@Anderson1978's classic macroparasite reduced an infinite-dimensional system of equations to a tractable system of two equations by making the assumption that parasite distributions were constant and conformed to negative binomial distributions. @Kretzschmar1993a showed that, at equilibrium, such reduced models are equivalent in terms of stability, even if the exact form of equilibrium distributions varies some from the negative-binomial. However, the distribution of infections in the infinite system changes over time, as does its variance/mean ratio [@Adler1992], and thus the distribution.

@Kretzschmar1993b showed that the variance-mean ratio was important to system stability; it must increase with the mean for there to be host-parasite equilibrium. This suggests why under-dispersed populations are rare - they lead to extinction of these populations.

@Anderson1978 suggest the binomial distribution to represent these underdisperesed

populations. However, this has the drawback of an absolute maximum number of infections, above which the probability is zero. In addition, there are

practical problems a discrete parameters that preclude using many numeric solvers.

Fungal diseases can be similar to macroparasites in that they are load-dependent

[@Briggs2010; Langwig2015]. They are capable of causing local extinctions in host

populations [@Briggs2010; @Fisher2012; @James2015; Langwig2015].

It is theorized that one reason underdispersal is rare because system stability depends on overdispersal [@Adler1992, @Kretzschamr1993a]. The stable distributions

of standard host-parasite models, even in the absence of individual variation,

are over-dispersed.

Characterize how the distributions of infections in a host population vary under the transient dynamics of an epidemic, as simulated using an individual-based model

- Show how these distributions differ from Poisson or Negative-Binomial assumptions in classic multi-infection models over time

- Demonstrate how a Conway-Maxwell-Poisson distribution proves a better characterization of the distributions generated from the IBM.

- Show the same for an empirical counts of canker infection on oaks from Meentemeyer et al.’s 2012 Sonoma data set.

Demonstrate and document an R package for fitting such distributions.

## The Conway-Maxwell-Poisson distribution

The Conway-Maxwell Poisson (CMP) [@Conway1962; @Shmueli2005; @Lynch2014] distibution is a distribution that can be used to model discrete count data. It has the form:

$$P(X = x) = \frac{\lambda^x}{(x!)^\nu} \frac{1}{Z(\lambda, \nu)}$$

where

$$Z(\lambda, \nu) = \sum\_{j=0}^\infty \frac{\lambda^j}{(j!)^\nu}$$

As in the Poisson distribution, $\lambda$ is an encounter rate. Unlike the Poisson

distribution, the $\nu$ term modifies it. When $\nu < 1$, the distribution is overdispersed; it has a fatter tail in a Poisson, and its variance is greater than its mean. When $\nu > 1$, the distribution is underdispersed. Its tail is thinned, and its variance is less than its mean. If $\nu = 1$, the CMP distribution is a Poisson distribution. If $\nu = \infty$, the CMP distribution

is the same as a Bernoulli distribution.

The CMP distribution intuitively describes the processes that generate parasite distributions. A Poisson process represents the accumulation of parasites, while the censorship of the tail, driven by the $\nu$ term, represents the density-dependent processes that operate most strongly in hosts with high parasite loads.

## The `cmp` package

CMP distributions, along with functions are implemented in the `cmp` package [@Ross2015c, dx.doi...TODO].

- Fit via max likelihood, KLD

- Standard functions

- High-speed and quiality implementation, using approximations where approriate, defining error bounds.

The `cmp` package contains density, probability, quantile, and random number functions

for the CMP distribution, as as well as fitting these distributions to data via

maximum likelihood and Kullback–Leibler divergence.

# Models

Two versions of the model:

I use two versions of a model closely derived from @Anderson1978. This first, determistic, continuous version, consists of an infinite-dimensional system of ODEs representing the population in classes of infection level. The second is its equivalent individual-based model, which includes demographic stochasticity. The IBM consists of a set of rate equations which are solved via Gillespies [-@Gillespie1978] stochastic simulation algorithm (SSA).

An infinite-dimensional model

$$\begin{aligned}

N\_0 &= r \sum\_{i = 0}^\infty N\_i \left(1 - \frac{\sum\_{i = 0}^\infty N\_i}{K}\right) - (d + \Lambda) N\_0 \\

N\_i &= Lambda N\_{i-1} - Lambda N\_i - \alpha i^p - mu i N\_i + mu (i + 1) N\_{i+1} \\

\Lambda = \lambda\_ex + \lambda \sum\_{i = 1}^\infty i N\_i \\

$$

In the individual-based model, the host population is reprented as a discrete set of individuals, each with a discrete number of infections. Births, deaths, new infections and loss of infections (recovery) in each individual $j$ with infections $i\_j$

$$\begin{aligned}

r\_{j, \, birth} &= \max \left\{ r(1-N/K), \, 0 \right\} \\

r\_{j, \, death} &= \alpha j\_i + d \\

r\_{j, \, infection} &= \Lambda \\

r\_{j, \, recovery} &= \mu j\_i

\Lambda = \lambda\_ex + \lambda \sum\_{j = 1}^N i\_j \\

\end{aligned}$$

The rates $\alpha, \my, \lambda, \lambda\_ex, K$ and $d$ represent the same rates

as in the ODE model.

| Parameter | Symbol | Value |

|-----------|--------|-------|

| birth rate | $r$ | `r format(parms$r)` |

| carrying capacity | K | `r format(parms$K)` |

| intrinsic host mortality rate | $d$ | `r format(parms$d)` |

| intrinsic pathogen mortality rate | $\mu$ | `r format(parms$mu)` |

| additional mortality per infection | $\alpha$ | `r format(parms$alpha)` |

| nonlinear component of additional mortality per infection | $p$ | `r format(parms$alpha\_power)` |

| contact rate | $\lambda$ | `r format(parms$lambda)` |

| external spore arrival rate | $\lambda\_{ex}$ | `r format(parms$lambda\_ex)` |

| initial host population | $N\_0$ | `r format(macro\_state[1])` |

| initial pathogen population | $P\_0$ | `r format(macro\_state[2])` |

| time period | $T$ | `r format(parms$time\_max)` |

Table: Parameters for the host-pathogen models and management problem

# Results

Underdispersion occurs in the transient dynamics of simulations where hosts are driven to extinction, as well as those with stable equilibria near the edge of the stable parameter space. Figure 1 shows the dynamics of host and parasite populations, and variance-to-mean-ratios of parasite distributions among hosts of the system across a range of host growth rates ($r$) in the infinite ODE system.

At low growth rates, where the population is driven to extinction, the parasite population goes through a period of underdispersion as the population declines to zero.

At medium growth rates, where there is a stable but strongly suppressed host population at equilibrium, the population goes through a period of underdispersion followed by overdispersion, and at highest growht rates, goes through a period of overdispersion before settling on an overall overdispersed equilibrium.

mortality. It shows the relative KLD values for the negative binomial and CMP

distributions with respect to a baseline Poisson distribution.

For the nonlinear case (Figure 3), the variance/mean ratio remains below one for the entire

period. In this case, the CMP distribution outperforms the negative binomial and

Poisson distributions over the whole period. The variance/mean ratio of the distribution has a minimum

at the peak of the epidemic. This corresponds to the peak advantage for the CMP distribution.

Mean dynamics differ in the stochastic version of the model, with greater host and

parasite populations than the deterministic model. The reason for this can be

seen in Figure X, which compares the evolution of distributions over time. In

the deterministic model, the tail of the distribution extends far beyond that of

the stochastic averages, because it can contain populations below one. With nonlinear

mortality, this tail can drive a greater portion of the dynamics.

Hence in the stochastic case, the CMP distributon provides a better fit over a larger

course of the simulation.

The distribution of infections among hosts is better approximated by the CMP distribution than the negative binomial. As may be expected, it performs better in approximating the transient under-dispersed population than the negative binomial, which performs no better than a Poisson in this case. However, it CMP also better approximates the transient and equilibrium over-dispersed parasite distribution in the stable case.

Stochastic

```{r fig1}

```

```{r fig2}

```

# Discussion

Underdispersion can can occur in host-parasite models even in the absence of parasite density dependence, though this density dependence strengthens this effect. It occurs under transient dynamics in both unstable and stable portions of the parameter space of classic representations macroparasite models, in both mean-field ODE and individual-based models.

Neither weak overdispersion nor underdispersion can be counted as evidence of the forces of density dependence or heterogeneity.