Theoretical Population Biology xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

The parent–offspring probability when sampling age-structured populations

Hans J. Skaug

University of Bergen, Department of Mathematics, P.O. Box 7803, N-5020 Bergen, Norway

ARTICLE INFO

Article history:
Received 30 January 2017
Available online xxxx

Keywords:
Close-kin
Population dynamics
Sampling probability
Stable age distribution

ABSTRACT

We consider two individuals sampled from an age-structured population, and derive the probability that these have a parent–offspring relationship. Such probabilities play an important role in the recently proposed close-kin mark-recapture methods. The probability is decomposed into three terms. The first is the probability of the parent being alive, the second term involves the mechanism by which individuals are sampled, and the third term is a contribution from the observed age of the parent. A stable age distribution in the population is assumed, and we provide an expression for how this distribution is perturbed by the information that an individual has given birth at a particular time point in the past or in the future. Calculations are performed from the perspective of the offspring, but we also make comparison to the situation where the perspective is put on the parent. Although the resulting probabilities are the same, the actual calculations differ, due to the asymmetry of a parent–offspring relationship.

© 2017 Published by Elsevier Inc.

UNCORRECTED PROOF

1. Introduction

Demographic aspects of age-structured population dynamics are well studied (e.g. Keyfitz and Caswell, 2006, Chpt. 5). For instance, the expected number of living kin of a given individual has been determined for various kinship categories (Goodman et al., 1974; Keyfitz and Caswell, 2006, Chpt. 15). In the present paper we focus on the parent–offspring probability, by which we mean the probability that two sampled individuals have a parent–offspring relationship. The existing literature focuses on human demographics, and does not address complications commonly arising when sampling from wildlife populations. Firstly, sampling probabilities may depend on the age, or some other property of the animal, such as length-dependent selectivity in fisheries. Secondly, age may not always be observable, which causes additional difficulties. The present paper addresses these issues for parent–offspring relationships in particular.

Our study of parent–offspring probabilities is motivated by a generalization of mark-recapture estimation, referred to as "close-kin mark-recapture" (CKMR), which has recently been suggested (Bravington et al., 2016b, a; Skaug, 2001). These methods build on the fact that kinship in large populations can be inferred from genetics (Anderson and Garza, 2006; Skare et al., 2009). The "recapture" of an individual now means the presence of a close relative in the sample. CKMR provides information about the same demographic parameters as ordinary mark-recapture, and hence increases the amount of information that can be extracted from mark-recapture experiments based on genetic tagging. Moreover, CKMR is applicable also when recapture of the same individual is impossible, such as when there is only a single sampling occasion. The joint probability distribution of the genetic profiles in a sample is very complicated, due to the unobserved genealogy of the sampled individuals. As a solution Skaug (2001) suggested a pseudo-likelihood approach in which only pairwise comparisons of individuals are considered. A key component of the pseudo-likelihood are sampling probabilities of the type studied in the present paper.

Denote by i and j two individuals which have been sampled from a population. Bravington et al. (2016b) derive the probability that i and j have a parent–offspring relationship, expressed in terms of the sampling mechanism and demographic parameters such as population size, survival and birth rates. By convention we treat i as the (putative) parent and j as the offspring. For simplicity, we consider only the female part of the population in most of the paper, and hence refer to i as the mother and j as the daughter.

The approach of Bravington et al. (2016b) is based on their notion of i's "reproductive output" at the time point when j was born. In the present paper we follow Skaug (2001) and derive the mother–daughter probability in terms of the probability that the (true) mother of j is alive at the time point when i was sampled. Expressions for the latter have been worked out for general age-dependent mortality and birth rates under the assumption of a stable age distribution (Keyfitz and Caswell, 2006, Chpt. 15). We refer to this method as being offspring-centric because it starts out from the fact that j has been sampled at a given time point, and then asks for the probability that i is her mother. The approach

E-mail address: skaug@math.uib.no.

https://doi.org/10.1016/j.tpb.2017.09.001
0040-5809/© 2017 Published by Elsevier Inc.

of Bravington et al. (2016b), on the other hand, is parent-centric, because it starts out from the fact that i has been sampled, and asks for the probability that j is her daughter. We show that the two approaches are fundamentally different, and that the parent-centric approach is easier to apply when the ages of i and j are known. The parent-centric approach is less general, on the other hand, in that it makes stronger assumptions than the offspring-centric method.

The rest of the paper is organized as follows. Section 2 lays out a few classical results about the stable age distribution, and generalizes these. Section 3 develops formulae for parent–offspring probabilities, conditionally on the age of the offspring. Section 4 does the same when we condition also on the age of the parent. In Section 5 we make comparison to the parent-centric approach by Bravington et al. (2016b), and in Section 6 we discuss the consequences of having both sexes in the population. Finally, Section 7 provides a more general discussion.

2. Demographic framework and age distributions

This section summarizes and extends a few classical results from mathematical demography about stable age distributions (Goodman et al., 1974; Keyfitz and Caswell, 2006, Chpt. 5 and 6). Recall that we only consider the female part of the population.

Let $l(a)$ be the survival function, i.e. the probability that a given individual survives to age a, viewed from the point of birth. Further, denote by $\beta(a)$ the birth rate at age a, i.e. the probability that a given female gives birth (to a female) in the short time interval $[t, t + dt]$ is $\beta(a)dt$. We assume that birth and death are independent processes, i.e. the events that a given (living) individual dies in $[t, t + dt]$ is stochastically independent from it giving birth in the same interval. This rules out for instance a situation where individuals die immediately after having given birth.

A population in which the functions $l(a)$ and $\beta(a)$ do not change with time nor vary among individuals, will settle into a stable age distribution (Keyfitz and Caswell, 2006, Chpt. 5). Further, the population will grow at a rate r which is given as the solution to the "characteristic equation",

$$\int_0^\infty \beta(a) l(a) e^{-ra} da = 1. \tag{1}$$

The stable age distribution has density

$$f(a) = \bar{\beta} l(a) e^{-ra}, \qquad a \geq 0, \tag{2}$$

where

$$\bar{\beta} = \left\{ \int_0^\infty l(a) e^{-ra} da \right\}^{-1} \tag{3}$$

is the average birth rate in the population at any given time point. Result (1)–(3) can be found in (6.1.2) and (5.1.1) in Keyfitz and Caswell (2006). The age distribution among females giving birth has density (Goodman et al., 1974)

$$g(a) = f(a) \frac{\beta(a)}{\bar{\beta}} = \beta(a) l(a) e^{-ra}, \qquad a \geq 0, \tag{4}$$

where r is the intrinsic growth rate determined by (1). In this formula, the age a of the mother refers to the point in time when the birth is taking place. We will need a generalization of (4), in which we allow the birth to have taken place u years in the past. By convention, negative values of u correspond to births that will take place in the future, which in probabilistic terms means that we are conditioning on a future event. Fig. 1 illustrates the meaning of u, as well as the setting in which the following theorem is going to be used.

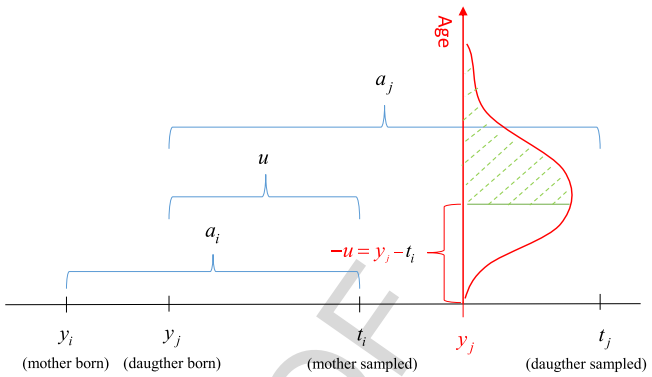

Fig. 1. Illustration of notation in a situation (black) where the mother was sampled before the daughter, but after the daughter was born. The situation where the mother was sampled before the daughter was born ($t_i < y_j$) is shown in red. The density $g(a)$ of the stable age distribution for the birth-giving cohort at y_j is also shown (vertical layout). The shaded area (green) of the density shows the part of the birth-giving cohort that was born before t_i. This sub-cohort plays a role in the proof of Theorem 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Theorem 1. *The stable age distribution among females (currently alive) who gave birth u years ago has density*

$$g(a|u) = \frac{g(a-u)}{\mathcal{P}(u)} \begin{cases} l(a)/l(a-u), & u > 0 \ (past) \\ 1, & u \leq 0 \ (future), \end{cases} \tag{5}$$

for $a \geq \max(0, u)$, where

$$\mathcal{P}(u) = \begin{cases} \int_0^\infty l(a+u)/l(a) g(a) da, & u > 0 \\ \int_{-u}^\infty g(a) da, & u \leq 0, \end{cases} \tag{6}$$

is a normalizing constant.

Proof. Let t be the time point to which a refers (current time), and let y be the time point when the birth is taking place, such that $u = t - y$. We refer to the group of females giving birth at y as the "birth-giving cohort". Our starting point in both cases ($u > 0$ or $u \leq 0$) is that the age distribution in the birth-giving cohort (at y) is given by (4). Consider first the case $u > 0$ ($t > y$), which is the situation depicted in black in Fig. 1. If there was no mortality in the birth-giving cohort during (y, t) its age distribution at t would simply be $g(a-u)$ (properly normalized) for $a > u$. The age specific probability of surviving from y to t is $l(a)/l(a-u)$, and hence we obtain $g(a|u) \propto g(a-u)l(a)/l(a-u)$ for $a > u$. By changing the variable of integration, $a' = a-u$, we find the normalizing constant

$$\int_u^\infty g(a-u) l(a)/l(a-u)\, da = \int_0^\infty g(a') l(a'+u)/l(a')\, da' = \mathcal{P}(u),$$

which proves the first case in (5).

For $u \leq 0$ (i.e. $t \leq y$) we place ourselves at time y and look back to time t (red y_j back to t_i in Fig. 1) to see which part of the birth-giving cohort was born before time t. This sub-cohort (shaded green area in Fig. 1) consists exactly of those individuals being of age $y - t = -u$ or more at time y. By definition there have been no deaths in this sub-cohort during the interval (t, y), and hence its age distribution at t had density proportional to $g(a-u)$, for $a > 0$. Again, changing the variable of integration we find that the normalizing constant must be

$$\int_0^\infty g(a-u)\, da = \int_{-u}^\infty g(a')\, da' = \mathcal{P}(u), \quad u \leq 0.$$

This completes the proof.

UNCORRECTED PROOF

Download English Version:

https://daneshyari.com/en/article/5760553

Download Persian Version:

https://daneshyari.com/article/5760553

Daneshyari.com