

[9] studbookR: Data files

Frank PG Princee

Draft 2016-12-29

Introduction

The **studbookR** package has been developed to analyse studbook data in conjunction with the *Population Management Library (PML)* software. This does not mean, however, that use of this package is restricted to data files which have been created by PML. This vignette describes the “easiest” method i.e. following the directory (folder) structure and data/file format of PML.

Data folders

The PML software creates data files of studbook project [studbook] in the following path of the users’ home directory (~ symbol): `~/pml/projects/[studbook]/export/`.

File format

The analyses import data from tab delimited ASCII files in which text is embedded in double quotation marks. The first data line must contain column headers. Data files can contain comment lines which are marked with a hash (#) symbol e.g. *PML* stores the population view settings at the beginning of each data file.

Data files have the extension “.csv”.

Data files

The data files that are used in the *studbookR* analyses are:

File name	Analyses
age-[YYYYMMDD]-[sex]	age distribution
birth_interval-female	interbirth interval and quantitative genetics
birth_season-[sex]	birth season, moon phase at birth
birth-[sex]	births between censuses
census-[sex]	census data
death_season-[sex]	death season
death-[sex]	deaths between censuses
emigration-[sex]	emigration between censuses
fecundity-[sex]	age class based fecundity rates, matrix projection
firstbreeding-[sex]	age at first reproduction and quantitative genetics
immigration-[sex]	immigration between censuses
kmfecundity-[sex]	fecundity rates based on age in days
kmsurvival-[sex]	survival rates based on age in days
lethaleqv-[parents]-[sex]	lethal equivalents
lifespan-[sex]	lifespan, longevity

File name	Analyses
lifespan_model-[sex]	quantitative genetics
lifetable-[sex]	reproductive value, stable age class, generation time
litter_season-all	litter season, litter survival and quantitative genetics
littersize	litter size
mortality-[sex]	age class based mortality rates, matrix projection
neonatal	neonatal mortality between censuses
pedigree	quantitative genetics
reproductive-[sex]	reproductive lifespan

[YYYYMMDD] = census data; [sex] = male, female, unknown or all; [parents] = inbreeding of parents: none, inbred, dam, sire, both, single

The file *litter_season-all.csv* contains a duplicate record of each litter (for sire and dam) to be used in quantitative genetic analysis. The seasonality and litter survival analyses use only records for females.

Data fields

All dates are in YYYY-MM-DD (ISO format); latitude is in decimal degrees i.e. DD.MMMMMM.

File age

Field	Type	Description/remarks
X	integer	age class
Y	integer	number of individuals in age class

Files birth, census, death, emigration, immigration and neontal

Field	Type	Description/remarks
annual	text	census date. Header name refers to interval
COUNT	integer	number of individuals

File birth_interval

Field	Type	Description/remarks
animal ^a	text	studbook ID of individual
YEAR	integer	year of birth (=cohort)
G	integer	(maternal) generation
F	real	inbreeding coefficient
REARING	text	type of rearing: parent, foster, hand or unknown
INTERVAL	integer	days since previous birth

^a The field name “animal” is reserved in the MCMCglmm animal model.

File birth_season

Field	Type	Description/remarks
DAM	text	studbook ID of dam
DATE	text	date of birth/hatching offspring
LOCATION	text	location of birth
PHASE	real	moon phase
SEGMENT	integer	moon segment
LATITUDE	real	latitude of location
PHOTOPERIOD	real	daylength at date of birth at latitude
HOURS	real	total daytime hours since conception at latitude
G	integer	(maternal) generation of dam
F	real	inbreeding coefficient of dam
DAM_AGE	integer	Age of dam at date of birth (in days)

File death_season

Field	Type	Description/remarks
DATE	text	date of death
LOCATION	text	location of death
PHASE	real	moon phase
SEGMENT	integer	moon segment
LATITUDE	real	latitude of location
PHOTOPERIOD	real	daylength at date of death at latitude

File fecundity

Field	Type	Description/remarks
CLASS	integer	age class (starts at 0)
N	real	individuals “at risk” for reproduction
BRED	integer	individuals that did reproduce
B	integer	offspring (of same sex as parents)
OBSERVED	real	observed fecundity rate (m_x)
MEAN	real	bootstrap mean m_x
VAR	real	bootstrap variance of mean m_x
MEDIAN	real	bootstrap median m_x
LOWER	real	bootstrap 2.5 percentile of m_x
UPPER	real	bootstrap 97.5 percentile of m_x

File firstage_model

Field	Type	Description/remarks
animal ^a	text	studbook ID of individual
DAM	text	studbook ID of dam; NA if wild or unknown
SEX	text	male or female
YEAR	integer	year of birth (=cohort)

Field	Type	Description/remarks
G	integer	maternal generation
F	real	inbreeding coefficient
REARING	text	type of rearing: parent, foster, hand or unknown
LOCATION	text	location of birth
AGE	integer	age at first reproduction

^a The field name “animal” is reserved in the MCMCglmm animal model.

File kmfecundity

The file *kmfecundity-[sex].csv* contains recurrent data.

Field	Type	Description/remarks
ID	text	studbook ID of individual
BIRTHDATE	text	date of birth of offspring
STARTAGE	integer	age at start of view (left truncation)
TIME	integer	age at first birth or time since previous birth
EVENT	integer	0=censored, 1= not censored
BIRTHSITE	text	location of birth of individual ID
G	integer	maternal generation
F	real	inbreeding coefficient
INBRED	text	is inbred? yes or no
KIDS	integer	cumulative number of offspring
LOCATION	text	location of birth of offspring
REARING	text	type of rearing: parent, foster, hand or unknown
SEX	text	male or female
TAXON	text	taxon e.g. subspecies

The last record of each individual has EVENT=0 (censored) and KIDS=-1

File kmsurvival

Field	Type	Description/remarks
ID	text	studbook ID of individual
BIRTHDATE	text	date of birth
STARTAGE	integer	age at start of view (left truncation)
TIME	integer	age at death or lost-to-follow-up
EVENT	integer	0=censored, 1= not censored
BIRTHSITE	text	location of birth of individual ID
BREEDER	text	is breeder? yes or no
DAM	text	studbook ID of dam
DAM_AGE	integer	age of dam at birth of individual
DAM_F	real	inbreeding coefficient of dam
F	real	inbreeding coefficient
G	integer	maternal generation of individual
INBRED	text	is inbred? yes or no
LITTER_ID	integer	litter ID in which individual was born

Field	Type	Description/remarks
LITTER_SIZE	integer	number of litter mates
LOCATION	text	last location
REARING	text	type of rearing: parent, foster, hand or unknown
SEX	text	male or female
SIRE_F	real	inbreeding coefficient of sire
TAXON	text	taxon e.g. subspecies

File lethaleqv

Field	Type	Description/remarks
F	real	inbreeding coefficient category
BIRTHS	integer	number of births
SURVIVED	integer	number that survived until specified age (e.g. 30 days)

File lifespan

Field	Type	Description/remarks
animal ^a	text	studbook ID of individual
DAM	text	studbook ID of dam; NA if wild or unknown
SEX	text	male, female or unknown
YEAR	integer	year of birth (=cohort)
G	integer	maternal generation
F	real	inbreeding coefficient
REARING	text	type of rearing: parent, foster, hand or unknown
LOCATION	text	location of death
AGE	integer	age at death

^a The field name “animal” is reserved in the MCMCglmm animal model.

File lifetable

Field	Type	Description/remarks
CLASS	integer	age class (starts at 0)
QX	real	mortality rate
PX	real	survival rate
LX	real	survivorship
MID_LX	real	mid survivorship
EX	real	life expectancy
MX	real	fecundity rate
LXMX	real	product of $l_x m_x$
VX	real	Fisher’s reproductive value
VXS	real	Reproductive value at stable age
CX	real	Proportion of individuals at stable age

File litter_season

Field	Type	Description/remarks
animal	text	studbook ID of individual
SEX	text	male, female or unknown
LITTER_ID	integer	unique ID of litter produced by ID
DATE	text	date of litter
MALES	integer	males in litter
FEMALES	integer	females in litter
UNKNOWN	integer	unknown (and abnormal) sex in litter
D_MALES	integer	neonatal male deaths
D_FEMALES	integer	neonatal female deaths
D_UNKNOWN	integer	neonatal deaths of unknown sex
STILLBIRTH	integer	stillbirths (including death at day of birth)
LOCATION	text	location of litter
ANGLE	real	angle of moon
PHASE	real	moon phase
SEGMENT	real	moon segment
LATITUDE	real	latitude of location
PHOTOPERIOD	real	daylength at date of birth at latitude
HOURS	real	total daytime hours since conception at latitude
G	integer	(maternal) generation of animal
F	real	inbreeding coefficient of animal
AGE	integer	age of animal at litter date

^a The field name “animal” is reserved in the MCMCglmm animal model.

Files littersize and reproductive lifespan

The file *littersize.csv* can include more records with litter counts than *litter_season.csv*, as the latter may have been filtered on e.g. hemisphere.

Field	Type	Description/remarks
DATA	integer	age or litter size

File mortality

Field	Type	Description/remarks
CLASS	integer	age class (starts at 0)
N	real	individuals at risk
D	integer	individuals that died
OBSERVED	real	observed mortality rate (q_x)
MEAN	real	bootstrap mean q_x
VAR	real	bootstrap variance of mean q_x
MEDIAN	real	bootstrap median q_x
LOWER	real	bootstrap 2.5 percentile of q_x
UPPER	real	bootstrap 97.5 percentile of q_x

File pedigree

Field	Type	Description/remarks
ID	text	studbook ID of individual
DAM	text	studbook ID of dam (NA if unknown or wild)
SIRE	text	studbook ID of sire (NA if unknown or wild)
SEX	integer	1 = male, 2 = female, 3 = unknown (required by kinship2)
BIRTHSITE	text	location of birth
BIRTHDATE	text	date of birth
DEATHDATE	text	date of death or lost-to-follow-up
CENSORED	integer	0 = censored (alive or missing), 1 = died
F	real	inbreeding coefficient
G	integer	(maternal) generation