

Chapter 2

Grouping Demographic Data

In the discrete-time approach, demographic rates are estimated from grouped data built using the “Lexis diagram” (Fig. 2.1), that was originally designed for human demography. This diagram plots three types of information simultaneously: the date of birth of an individual, its age and the time. The life history of each individual is represented in the plane “time $\times$ age” by a linear segment joining the entry and the exit of the individual (e.g. from birth to death). The paternity of the Lexis diagram is multiple [37]. Main contributions were provided by demographers of the nineteenth century (Karl Becker and Wilhelm Lexis in Germany, and Abraham Verweij in The Netherlands). Enhancements were then proposed from 1960 by the French demographer Rolland Pressat [38, 39].

Data grouping consists in building a set of discrete time-and-age cells in the Lexis diagram, then in splitting the individual segments into these cells. Split data are then aggregated over all the animals present in each of the cells to constitute the “grouped data” that are finally used to calculate demographic parameters. For a given cell, typical grouped data are in the following form:

- n : number of animals present at the beginning of the cell;
- T : time of presence of these n animals in the cell;
- m : number of events (e.g. parturitions or deaths) observed in the cell.

Several approaches are available to build the cells, which lead to different results. It is important to understand these approaches well in order to avoid misinterpretations. This chapter presents a selection of approaches, taking the example of annual data grouping (i.e. 1-year cells). The same principles apply to any smaller or larger cells.

2.1 Cohort Splitting

A cohort of animals is composed of individuals born during the same period. Figure 2.1 presents an example of an annual cohort of five animals born in year 2002 and monitored until December 2005.

2.1.1 Horizontal and Vertical Cells

In a Lexis diagram, life histories of a cohort can be split into “horizontal” or “vertical” cells:

- Horizontal cells are delimited by exact ages. For instance, a cell ranging from 1 to 2 years in exact age represents events occurring between ages 365 and 730 days;
- Vertical cells are delimited by past ages. At a given time t , the past age i represents animals having exact ages between i and $i + 1$ at t . An annual past age corresponds to the number of full years already lived by the animal. For instance, an animal having a past age of 2 years has an exact age comprised between 2 and 3 years. In the forthcoming text, by default, ages will correspond to exact ages. The text will specify when past ages are used.

In the Lexis diagram, cells are parallelograms, with the exception of the first vertical cell which is a triangle. When time and age have a same scale, a horizontal cell spans two time units and one age unit, and inversely for a vertical cell. By convention, we label the successive cells from 0 to L (L represents the terminal cell). Figure 2.2 presents the cohort 2002 splitting, and Table 2.1 the corresponding grouped mortality data. In this example, horizontal cell labels range from “0” to “2” (cell “3” is incomplete and therefore not considered) and vertical cell labels from “0” to “3”.

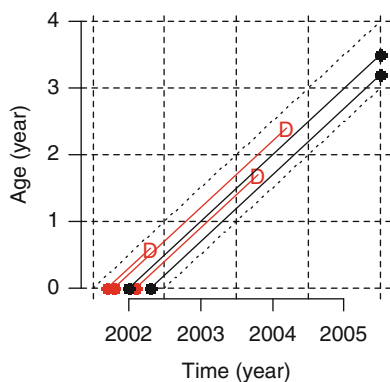


Fig. 2.1 Lexis diagram representing life histories of a cohort of five animals born in 2002 and monitored for four consecutive years. Three animals (red lines) of the cohort died (D) before the end of 2005

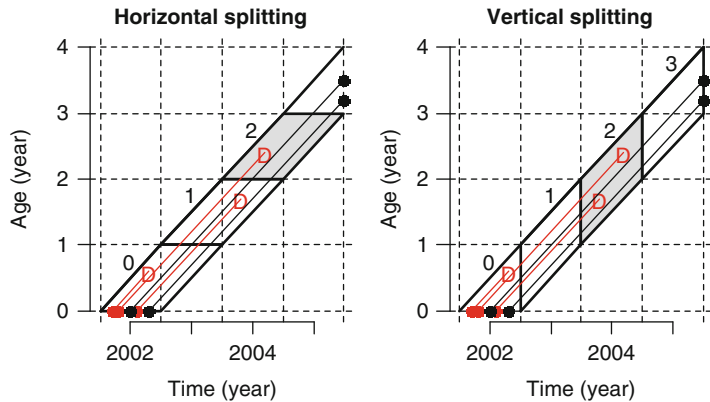


Fig. 2.2 Cohort 2002 annual splitting

Table 2.1 Grouped data corresponding to cohort 2002 annual splitting for calculation of mortality rates

Cell	Type of data splitting					
	Horizontal			Vertical		
	<i>n</i>	<i>T</i> (year)	<i>m</i>	<i>n</i>	<i>T</i> (year)	<i>m</i>
0	5	4.6	1	5	2.2	1
1	4	3.6	1	4	4.0	0
2	3	2.3	1	4	2.9	2
3	–	–	–	2	2.0	0

n number of animals present at the beginning of the cell, *T* time of presence of the animals in the cell, *m* number of natural deaths in the cell

Rates calculated horizontally and vertically have different interpretations. In the example, annual horizontal cell “2” accounts for events occurring between 2 and 3 years old and the mortality probability in this cell is $p = 1/3 = 0.33$ (from Table 2.1). In contrast, vertical cell “2” accounts for events occurring between past ages 1 and 2 years and the mortality probability is $p = 2/4 = 0.50$. The same distinction applies to any cell, in particular cell “0”. A common confusion is to consider that mortality estimated in vertical cell “0” is the mortality from birth to 1 year old, while it only accounts for death between birth and the end the year. Choosing the type of splitting (horizontal vs. vertical) and the size of the cells depends on the study objectives, the data available and, hence, on the survey method used. With ABM data, both splittings can be carried out regardless of the cell size. This is not in general the case for data collected with other methods (see for instance [40]). When very small (e.g. weekly) cells are built, both types of cells yield roughly the same results. This is illustrated further in the handbook.

Fig. 2.3 Cohort 2002 annual horizontal splitting with cells 1 and 2 aggregated into cell 1+

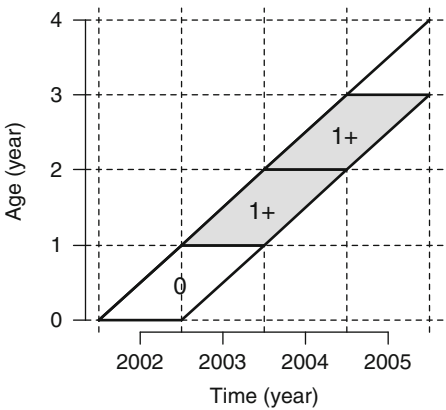


Table 2.2 Grouped mortality data corresponding to cohort 2002 annual horizontal splitting after aggregating cells 1 and 2 into cell 1+ (for which $m = 1 + 1$, $n = 4 + 3$, and then $p = 2/7$)

Cell	n	T	m
0	5	4.6	1
1+ (= 1 + 2)	7	5.9	2

2.1.2 Aggregation of Cells

Demographic parameters can be calculated for a given cell, but also more globally over groups of cells. The procedure simply consists in summing the grouped data over the cells involved in the aggregation. For instance, in the cohort 2002 example, we may find interesting to calculate the average mortality for animals older than 1 year. To do so, we can add horizontal cells “1” and “2”, which leads to cell “1+” (Fig. 2.3 and Table 2.2) and mortality probability $p = 2/7 = 0.29$.

2.2 Pseudo-cohort Splitting

2.2.1 Columns of Cells

Analyzing data over entire cohorts require long-term data. For instance, splitting life histories of cows requires more than 10 years of monitoring, which is very rare in tropical farming systems. For overcoming this problem, a common approach is to group the data by columns of cells (Fig. 2.4). Such columns have been referred to as “pseudo-cohorts” (eg. [41]) since, assuming low or moderate rates variations between years, rates calculated through a column of ages can be considered representative of the rates that would be observed through a true cohort if this cohort

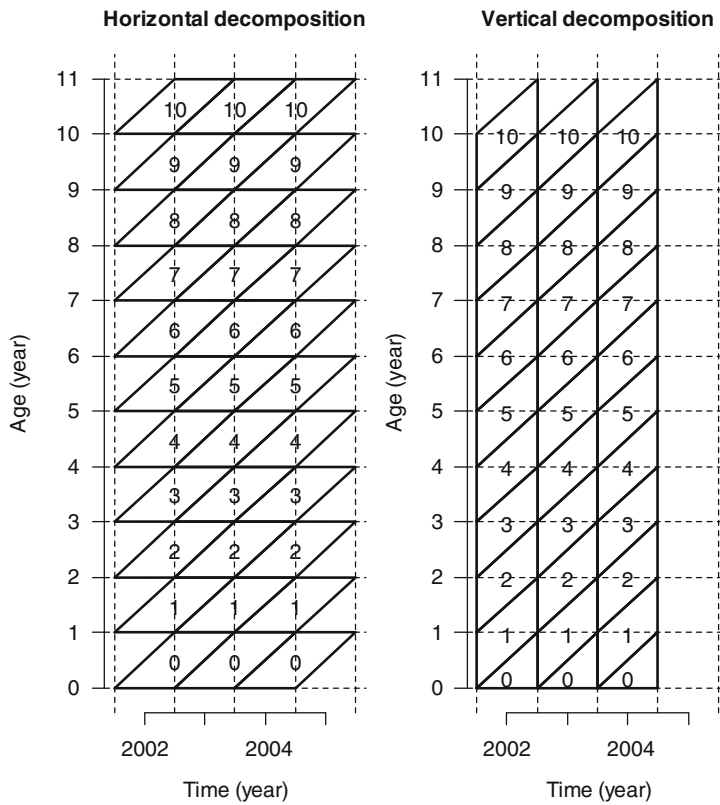
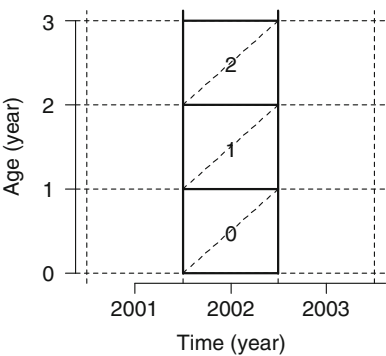


Fig. 2.4 Three columns of annual cells. The horizontal splitting extends from 2002 to 2005, and the vertical splitting from 2002 to 2004

Fig. 2.5 Column of square annual cells (year 2002). Age and time are crossed perpendicularly. Each cell is composed of two half-cells corresponding to different cohorts (Separated by dashed diagonal lines)



could be monitored. In practice, building columns of horizontal or vertical cells is a very flexible approach when analyzing livestock demographic data. The package presented in Chap. 5 implements such type of column grouping from LASER data.

2.2.2 Other Types of Cells

Splitting pseudo-cohorts can also be achieved by delimitating time and age perpendicularly (Fig. 2.5). This generates columns of square cells (each containing two half horizontal or vertical cells of different cohorts) or of rectangular cells when several square cells are aggregated. Furthermore, square or rectangular splitting can be approximated by aggregating smaller horizontal or vertical cells (see examples in Chap. 6).

Calculation of Demographic Parameters in Tropical
Livestock Herds

A discrete time approach with LASER animal-based
monitoring data

Lesnoff, M.; Lancelot, R.; Moulin, C.-H.; Messad, S.;
Juanès, X.; Sahut, C.

2014, X, 99 p. 33 illus., 23 illus. in color., Hardcover
ISBN: 978-94-017-9025-3