2 Recognising statistical dependency

We start this chapter with an explanation of pseudoreplication. Ignoring pseudoreplication can have serious consequences for the conclusions based on the statistical analyses. In our experience, nearly every ecological study has some form of pseudoreplication. We will present three examples in this chapter that will help you recognise the presence of pseudoreplication. In later chapters we will provide solutions.

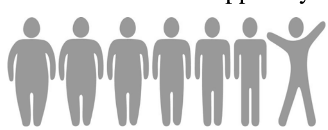
Prerequisites for this chapter: Knowledge of R, multiple linear regression, and Poisson GLM is required. Familiarity with GAM and mixed-effects models is recommended. In later chapters we will explain GAM and linear mixed-effects models in more detail.

2.1 Pseudoreplication

The term pseudoreplication has a somewhat scary aura surrounding it. It goes back to a publication by Hurlbert (1984) and a series of follow-up papers, sometimes shaming authors who ignored pseudoreplication. The debate whether pseudoreplication is still a problem, 30 years after being introduced, can sometimes still be heated (Oksanen 2001; Freeberg and Lucas 2009; Hurlbert 2004).

So what is pseudoreplication? Hurlbert (1984) defined it as: '… the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent'. Hurlbert (1984) then went on to define three forms of pseudoreplication in the context of ecological field experiments (where experiments can be manipulated). Millar and Anderson (2004) provide a more modern discussion on pseudoreplication, and use examples from fisheries science. There is no need to know the names of all the different forms of pseudoreplication; at the end of the day it means that the observations on the response variable are not independent and this aspect is ignored during the statistical analysis. Let's present some examples.

A classical example of pseudoreplication is as follows. Suppose you investigate whether a certain diet for weight loss is effective. A possible approach is to sample the weight of a person a couple of time before he starts a diet, while he is doing the diet, and after the diet. Let's say seven measurements per person in total. And suppose this is done for 100 people. So we have 700 weight observations. Because the same person is measured seven times it would be wrong to assume that we have 700 independent observations to test for a diet effect. Pretending that we have 700 independent observations results in standard errors that are too small.



The correct analysis approach is a mixed-effects model; see, for example, Pinheiro and Bates (2000).

Zuur et al. (2013) used a turnstone (*Arenaria interpres*) data set (originally published in Fuller et al. (2013)). A flock of birds on the beach was identified. From this flock one focal bird was selected and observed for 30 seconds. The number of times it raised its head (to look for danger) was recorded. Then a second bird from the *same* flock was selected and the ecologists counted again the number of 'heads-up'.

This process was repeated until the flock flew away. A large number of flocks were sampled in the same way. It would be wrong to assume that the observations from the same flock are independent. If a fox is watching the birds from a nearby dune then most likely every observation on the number of heads-up from that flock will be very high! So we have pseudoreplication. The correct analysis approach is a generalised linear mixed-effects model using flock as random effect. See Zuur et al. (2014) for a fully worked-out solution.

Sick et al. (2014) investigated social strategies throughout the course of the day in chacma baboons (*Papio ursinus*). The data were also used in Ieno and Zuur (2015) and Zuur et al. (2016a). Data were collected from 60 baboons from two troops in the Tsaobis Leopard Park in Namibia. An individual baboon was followed for 1 h (focal hour), during which its grooming and dominance interactions were recorded, including the identity of the baboon being groomed (receiver). Multiple observations of the same baboon in a focal hour was the first source of dependency. During the 6-month sampling period, each baboon was repeatedly sampled, which is another source of dependency. To increase complexity, the receiver represents another level of dependency. Dealing with the pseudoreplication requires a mixed effects model with a two-way nested and crossed random effect.

Reed et al. (2011) analysed long-term survey data for three endangered waterbirds endemic to the Hawaiian Islands, the Hawaiian moorhen (*Gallinula chloropus sandvicensis*), Hawaiian coot (*Fulica alai*), and Hawaiian stilt (*Himantopus mexicanus knudseni*). The data set consists of annual (winter) counts covering a time span of 1956–2007. Each time series represented one species on one island: coot and stilt numbers on Oahu, Maui, and Hawaii, and moorhen numbers on Oahu and Kauai. Bird counts of a species in a particular year on a particular island are likely to depend on the counts in the previous year for the same island, but potentially also on the counts for other islands, and counts of other bird species. So we have temporal (and potentially spatial) pseudoreplication.