MATH35600 Assessed Practical 2019

You should work in groups of 3 for this assignment, handing in one report at the end. It is worth 20% of the marks for this unit.

The data loaded with the command

```
grouse <- read.table("https://people.maths.bris.ac.uk/~sw15190/TOI/harrier.dat")
```

give the consumption rate of grouse by hen harriers (per day per hen harrier) against the density of grouse (per square kilometre) from a managed grouse moor. The data are from an earlier version of a paper by Smout et al. which you can find here:
`https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0010761`

The data are highly variable, but show an increase in consumption rate of grouse as the grouse density increases. Ecological theory predicts that the expected rate of consumption, F, should be related to the density of grouse, N, via the following relationship:

$$F = \frac{\alpha N^m}{1 + \alpha \tau N^m} \tag{1}$$

where α, τ and m are positive parameters. α is interpreted as being related to the encounter rate between prey (grouse) and predators (hen harriers) at low prey densities. τ relates to the length of time it takes a predator to handle a prey item, while m allows the model to reflect the fact that the predators may not actively search for prey when they are at low densities, thereby lowering the consumption rate disproportionately at low prey densities.

Your aims are

1. to help the biologists who gathered the data to turn (1) into a reasonable statistical model, and to provide them with approximate confidence intervals for m and τ, in particular.
2. to appropriately visualize the uncertainty in the expected consumption rate as a function of grouse density.
3. to assess whether the data provide evidence that a model like (1) is better supported by the data than a simpler model such as $F = \beta N$ (where β is a parameter).

For part 1) you will obviously need to augment (1) with a distributional assumption relating the predicted F at any N to the corresponding observed consumption rate. Initially it is probably sensible to assume a normal distribution with constant variance, but this is unlikely to be completely satisfactory, so you should also consider other possibilities. Maximum likelihood estimation is a reasonable framework to use for this assignment, although for part 2) you might also find bootstrapping to be helpful. Before maximizing any log likelihoods numerically, take care to find initial values for the parameters that are not completely absurd.

What to hand in: You should write, as a group, a concise report of no more than 5 sides of A4 (normal margins $\geq$ 10pt font). The report should be accompanied by an appendix containing well structured, clearly commented R code for performing the analysis. Both the report and the appendix should start with a title and the names of all the group participants.

The report should contain 2 parts:

1. A one page report for the biologist providing clear, concise and properly justified answers to the 3 points listed above.
2. A technical report suitable for a statistician explaining the analysis, allowing them to understand what you have done, what you have concluded and why. There should be enough detail for them to judge the appropriateness of the approach. The report should include appropriate plots. The main body of the report should ideally include no R code (technicalities should be explained with maths, if necessary).

One report (as a pdf file) and one appendix (plain text file is best) per group should be emailed to `simon.wood@bristol.ac.uk` with the subject `MA35600 Grouse` followed by your surnames, by **12 noon, on Thursday 21 March 2019**.

Mark scheme guidance

First class marks will be awarded for work that could be passed on to the scientists who gathered the data essentially without modification. That is to say the statistics is appropriate and clearly explained, the conclusions appropriately drawn and any limitations are discussed fairly.

Upper second class marks will be awarded for work that could be passed on to the scientists who gathered the data, after a round of revision correcting some errors of presentation, interpretation or statistics that are relatively minor.

Lower second class marks will be awarded to work that has some more substantial flaws of presentation, interpretation or statistical reasoning which would require some more work to correct.

Third class marks will be awarded for work that contains some indication of substantive understanding and engagement, but contains more serious errors and misunderstandings.