MULTIVAR STATS IN ECOL AND GENETICS — EXERCISES, SHEET 3

1. A beak of a Darwin finch has been found. Use the data in the table FinchesSulloway.txt to optimize the method of predicting the body length, the wing length and the tail length of the bird by its beak dimensions BeakH, LBeakL and N-UBkL. Also take transformations of and interactions between these variables into account. How much does the precision of the prediction change when it is known on which island the beak was found.

2. Go back to exercise 3 of sheet 2. There is an .xls file among the downloaded files which contains all the RIKZ data. In the 4th table of this file, the ShannonW index of biodiversity is computed for all sample sites. Explore how ShannonW depends on the parameters week, angle1, angle2, exposure, salinity, temperature, NAP, penetrability, grain-size, humus, chalk, sorting1 and Beach. Consider also combinations and interactions of these parameters.

3. Once more the RIKZ data: In the same data where ShannonW is given you also find a column "Richness" which is just the number of species found at each sample site. How does Richness depend on week, angle1, angle2, exposure, salinity, temperature, NAP, penetrability, grain-size, humus, chalk, sorting1 and Beach? Fit Poisson and quasipoisson models.

4. Collet[1] and Venables and Ripley[2] report an experiment to investigate the toxicity of trans-cyphemethrin to the tobacco budworm *Heliothis virescens.* Batches of 20 male and 20 female moths were exposed to different doses of trans-cyphemethrin for three days. The following table shows the numbers of dead or knocked down moths.

dose $[\mu g]$	1	2	4	8	16	32
male	1	4	9	13	18	20
female	0	2	6	10	12	16

How does the probability of a moth to be killed or knocked out depend on the dose and on the moth's sex? Fit a logistic regression model. Check whether the model fit can be improved by rescaling the dose in an appropriate way. Present your final results graphically with 95% confidence bands.

5. The data file `TbDeerAndBoar.txt` contains data from a survey of Vicente et al.[3], see also Zuur et al.[4]. Boars and deers on 32 farms in Spain were tested for tubercolosis (Tb) and for the parasite *Elaphostrongylus cervi*. The table contains the numbers of sampled individuals and the numbers of positively tested individuals on each farm. The other variables in the table describe the habitat: the percentage of open land, pine and scrubs plantation, density of quercus plants, density of quercus trees, a wild boar abundance index, a red deer abundance index, the size of the habitat in ha, and whether the habitat was fenced (1) or not (0).
How does the risk of a red deer / wild boar to be infected with tubercolosis / *Elaphostrongylus cervi* depend on the other variables?

---

[1] Collet, D. (1991) *Modelling Binary Data*. Chapman & Hall, London.

[2] Venables, W.N., Ripley, B.D (2002) *Modern Applied Statistics with S, 4th ed.* Springer, New York.

[3] Vicente, J., Höfle, U., Garrido, J.M., Fernandez-de-Mera, I.G., Juste, R., Barralb, M., Gortarzar, C. (2006) Wild boar and red deer display high prevalences of tubercolosis-like lesions in Spain. *Veterinary Research* **37**: 107–119

[4] Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M. (2009) *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.