Perching Habits of Lizards

Arden Miller

Executive Summary

It was found that there were significant differences in the perching habits of the grahami and opalinus lizards. Both species preferred shaded perches but this preference was stronger for the opalinus lizards. This preference for shaded perches was strongest at mid-day for both species. It was also found that both species preferred perches that were less than 5 feet off the ground but again this preference was stronger for the opalinus lizards. The grahami lizards have a preference for perches that are less than 2 inches in diameter whereas the opalinus lizards have a slight preference for perches more than 2 inches in diameter. Both species were most apt to be observed perching at mid-day.

1 The Data

The data consisted of observations made on the daytime perching habits of two species of lizards: *grahami* (G) and *opalinus* (O). The data was collected by observing occupied sites (perches) and recording the relevant information: species involved (O or G), time of day (early, mid-day, or late), whether the perch was sunny or shaded, height of perch (< 5 or ≥ 5 feet), and diameter of perch (≤ 2 or > 2 inches).

	Perch		Time of Day					
			Early		Mid-day		Late	
S	D	H	G	O	G	O	G	O
sunny	≤ 2	< 5	20	2	8	1	4	4
		≥ 5	13	0	8	0	12	0
	> 2	< 5	8	3	4	1	5	3
		≥ 5	6	0	0	0	1	1
shaded	≤ 2	< 5	34	11	69	20	18	10
		≥ 5	31	5	55	4	13	3
	> 2	< 5	17	15	60	32	8	8
		≥ 5	12	1	21	5	4	4

2 Perching Habits of Lizards

The analysis of this data indicated that the species of lizard was related to each of the other 4 factors. This means that the different species had different preferences (or in some cases strengths of preference) with respect to the time of day they are most apt to perch, whether they prefer sunny or shaded perches, the diameter of perch they prefer, and the height of perch they prefer. It was also found that the time of day was related to the preference for sunny or shaded perches and the diameter of the perch was related to the height of the perch. On this basis it is sensible to divide the data by species and to collapse the data for each species into one 2 way table for time of day and sunny or shaded perches and a second 2 way table for perch diameter and perch height. These tables are given in Table 1.

From Table 1 we can see the grahami lizards have a preference for perches that are less than 2 inches in diameter (66% of grahami sightings) and are less than 5 feet off the ground (59% of grahami sightings). They seem to have a fairly strong aversion to perches that are both larger than 2 inches in diameters and more than 5 feet off the ground (only 10% of grahami sightings). Opalinus lizards would seem to have a stronger preference (than the grahami lizards) for perches that are less than 5 feet off the ground (83% of opalinus sightings). They also appear to have a slight preference for perches that are more than 2 inches in diameter (55% of opalinus sightings).

Both species of lizards prefer shaded perches to sunny perches although this preference is stronger for the opalinus lizards than for the grahami lizards (89% of opalinus sightings and 79% of grahami sightings involved shaded perches). For both species this preference for shaded perches is stronger at mid-day than either early in the day or late in the day. For grahami lizards at mid-day 91% of sightings were shaded whereas early in the day 67% were shaded and late in the day 66% were shaded. For opalinus lizards at mid-day 97% of sightings were shaded whereas early in the day 86% were shaded and late in the day 76% were shaded. We can also see that both species of lizard were most apt to be observed perching at mid-day (52% of all grahami sightings and 48% of all opalinus sightings).

Table 1: Collapsed tables

	Perch Height	Perch diameter ≤ 2	Perch diameter > 2		Time of Day Early	Time of Day Mid-day	Time of Day Late
grahami	< 5	153	102	sunny	47	20	22
	≥ 5	132	44	shaded	94	205	43
opalinus	< 5	48	62	sunny	5	2	8
	≥ 5	12	11	shaded	32	61	25

Statistical Appendix

I started by fitting the log-linear (Poisson regression) model that contained all effects up to 3 factor interactions. Since none of the 3 factor interactions were significant, I dropped these and fitted the model that contained main effects and 2 factor interactions. The analysis of deviance table for this model is

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(Chi)
NULL			47	737.5553	
ss	1	242.6905	46	494.8648	0.0000000
dim	1	28.3879	45	466.4769	0.0000001
height	1	49.5892	44	416.8877	0.0000000
time	2	98.5273	42	318.3605	0.0000000
species	1	165.7469	41	152.6136	0.0000000
ss:dim	1	3.5737	40	149.0399	0.0587016
ss:height	1	0.9450	39	148.0949	0.3310045
ss:time	2	47.9689	37	100.1260	0.0000000
ss:species	1	6.4724	36	93.6537	0.0109564
dim:height	1	16.1171	35	77.5365	0.0000595
dim:time	2	2.1326	33	75.4039	0.3442815
dim:species	1	17.1220	32	58.2819	0.0000351
height:time	2	1.1492	30	57.1328	0.5629368
height:species	1	20.6264	29	36.5064	0.0000056
time:species	2	11.4526	27	25.0537	0.0032591

Based on this table I decided to drop the height:time, dim:time, ss:height, and ss:dim interactions. It is debatable whether ss:dim should be dropped or not. I decided to drop it in the interests of simplifying the analysis - if there is a relationship between ss and dim it will be small compared with some of the other relationships. Therefore, the model I have selected indicates that species is related to all the other factors and in addition ss is related to time and dim is related to height. The association graph for this model is:

We can see from this graph that if we fix species then dim and height become independent of of ss and time. This indicates that for each species we can collapse the table into a 2 way table for ss and time and a 2 way table for dim and height. I used the raw data rather than the fitted values to produce these collapsed tables since I thought that it would be easier for a non-statistician to understand if the tables contained the actual counts. The same conclusions would have been reached had the fitted values been used.

The output from `dummy.coef` was used to identify how the different factors were related. It indicates that the `ss:time` interaction is having the biggest effect on the fitted model. The conclusion given in the report can be obtained from this output. For example the positive coefficient for level `M` of `time` compared to 0 for `E` and a negative coefficient for `L` indicates more lizards are observed at mid-day than early or late.

```
>  dummy.coef(lizards.glm3)
$"(Intercept)":
 (Intercept)
    3.529169

$ss:
 shade        sun
     0 -0.7319271

$dim:
 less       more
    0 -0.4275722

$height:
 less       more
    0 -0.1666996

$time:
 E          L         M
 0 -0.8334299 0.7690016

$species:
 G         O
 0 -1.149561

$"ss:time":
 shadeE sunE shadeL     sunL shadeM     sunM
      0    0      0 0.171698      0 -1.61862

$"ss:species":
 shadeG sunG shadeO       sunO
      0    0      0 -0.8696537

$"dim:height":
 lessless moreless lessmore   moremore
        0        0        0 -0.6305421

$"dim:species":
 lessG moreG lessO    moreO
     0     0     0 0.733571

$"height:species":
```

```
 lessG moreG lessO     moreO
     0     0     0 -1.084465

$"time:species":
 EG LG MG EO         LO          MO
  0  0  0  0 0.6880784 -0.09231916
```

A plot of Pearson residuals versus the linear predictors and an index plot of deviance residuals are given in Figure 1. Neither indicates any problem with the fitted model. Plots of Cook's distances and deviance changes are given in Figure 2. The Cook's distance plot indicates that observation 39 (grahami, mid-day, shaded, dim>2, height<5) is having an unusually large influence on the fitted model. If this point is dropped and the model is refit we find that the biggest changes occur for the coefficients for `dim` and `dim:height`. These still have the same signs as before but now the preference for perches of diameter less than 2 inches is stronger and the interaction between diameter and height is not as strong. Therefore dropping this observation would only have a minor effect on our conclusions.

```
$dim:
 less       more
    0 -0.7137586

$"dim:height":
 lessless moreless lessmore   moremore
        0        0        0 -0.3796044
```

Since there is no evidence that this observation is an error, the results for the fitted model using all the data are used in the report.

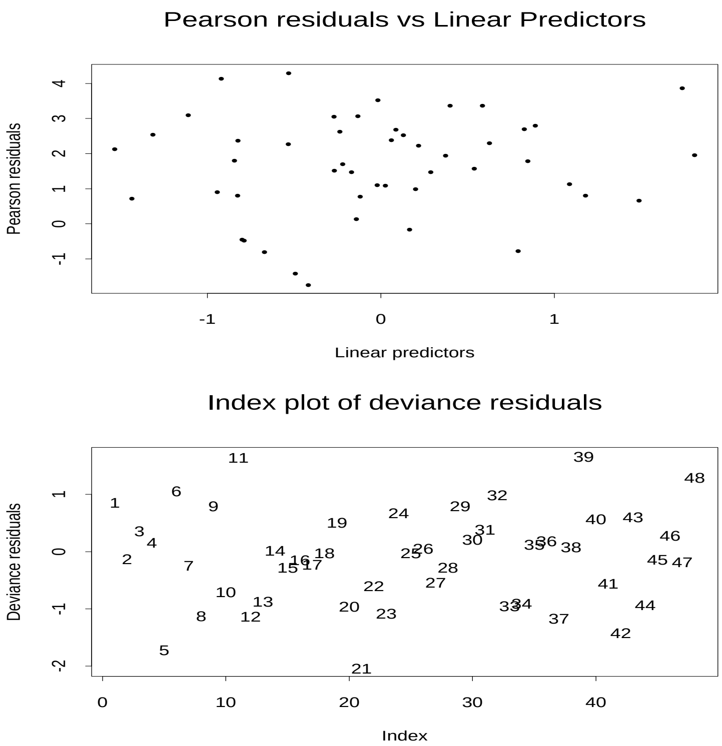

Figure 1: Residual plots

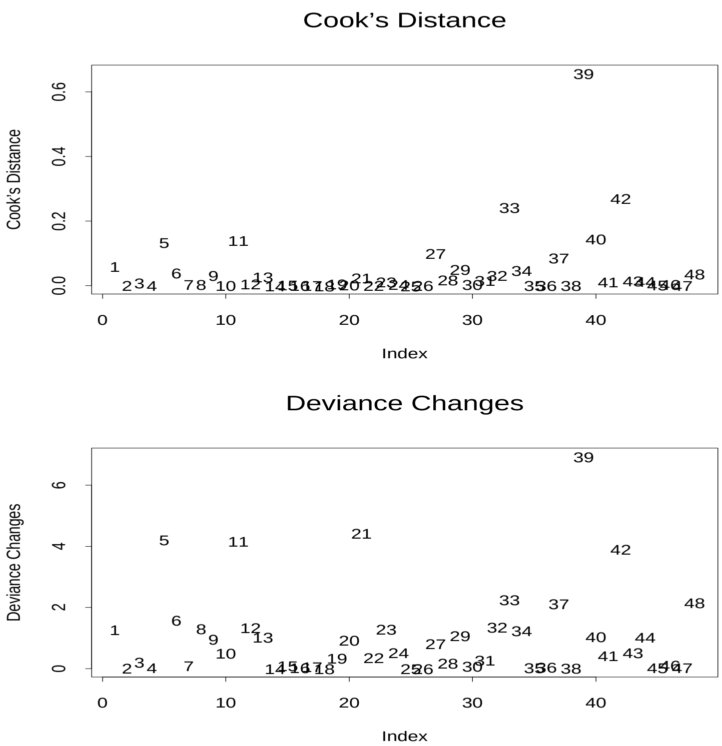

Figure 2: Cook's distances and deviance changes

475.330 Assignment 4: Marking Guide

This assignment asks the students to analyse data collected on the perching habits of 2 types of lizards. They were asked to identify the factors that were related to each other.

As usual their report should be divided into 2 parts: a report on their findings that clearly explains what they found and a statistical appendix that explains how they analysed the data. The marks should be split as follows:

Report	13 marks
Statistical Appendix	7 marks
Total	20 marks

Report on the Lizard Perching Habits (13 Marks)

This part of the assignment should summarise their findings on the perching habits of the 2 species of lizards. They should indicate the types of perches each species prefers and compare the preferences of the 2 species. They should also discuss whether any of the factors used to describe perches are related (e.g. the time of day is related to the lizard's preferences to sunny or shaded perches).

- 5 marks for a generally well laid out, coherent report. The reader should not have to search for the important parts among a lot of details. Give: (1) 5 marks for a clear, precise report, (2) 3 marks if the report would be difficult for a non-statistician to understand, and (3) 1 mark if it would be difficult for anyone to understand.
- 8 marks for a sensible discussion about the perching habits of the lizards. Look for a reasonable discussion of the important factors that affect the perching habits. They should compare the habits for the 2 species and identify and discuss which of the factors that were used to describe the perches are related.
  - Most students should do a Poisson regression where they have included species as a factor with 2 levels. A few may have done a separate Poisson regression for each species and then compared the results. The first approach is preferred but the second way is okay. They should come to essentially the same conclusions using either approach.
  - Another possibility is to do a logistic regression using the species (proportion of "G" lizards for example as the response) – this will not identify relationships that do not involve species. If they do this in addition to Poisson regression analysis that is fine (it will help identify how the two species differ). However, deduct a couple of marks if they only do this analysis.
  - All the other factors are related to species which indicates that the 2 species do have different preferences when it comes to perches. In addition the "time of day"

by “sunny or shaded” interaction and “diameter” by “height” interaction are important. One good way they could make comparisons is to consider “time of day” by “sunny or shaded” combinations for each species and then the “diameter” by “height” combinations for each species.

- They should find that both species prefer shaded perches but this preference is more pronounced for the “O” lizards than the “G” lizards. For both types of lizard the preference for shade is greatest at mid-day. Both types of lizard are most apt to be perching at mid-day (this preference is more pronounced for the “G” lizards.
- Both species prefer perches that are less than 5 feet high but this preference is more pronounced for the “O” lizards. The “O” lizards show no real preferences with respect to the diameter of the perch but the “G” lizards have a preference for perches that less than 2 inches in diameter.
- One observation (species G, mid-day, shaded, D=> 2, H=< 5) is influential. If this point is dropped some of the significant interactions become less pronounced. They should comment on this and discuss how their conclusion would change if this point were removed – they should not simply drop this point
- Some students may use interactions plots to present their results. This is fine (but not required) but these plots should be accompanied by an explanation.
- Some student will use association graphs (a line graph where lines connect factors that are involved in an active interaction). Again such a graph is useful but not absolutely necessary and should be accompanied by an explanation.
- Some students will use the Dissimilarity Index (indicates the proportion of points that needs to be shifted in order for the data to fit the model exactly) in the discussion of their analysis.

Statistical Appendix (7 Marks)

This appendix should outline the reasons that they came to the conclusions they presented in the first part of the analysis. They are not required to give a detailed account of everything they did and they most certainly should not just have a whole bunch of computer output. Rather they should present a coherent account of the parts of their analysis that led them to their model for savings. For this assignment since all the variables are factors they only need to do a few diagnostics: index plot of deviance residuals, Cook’s distance plot, and a plot of deviance changes are useful.

Give: (1) 7 marks for a well documented account (2) 5-6 marks if they leave some important aspects unexplained or if they include lots of unimportant details (3) 3-4 marks if some of the conclusions they made are clearly not supported by their analysis (4) 2 or less marks if they just present output without reasons for their conclusions or if their reasons are not sensible.