Department of mathematical Sciences

Examination paper for ST2304 Statistical modelling for biologists and biotechologists

Examination date: 25th May 2021

Examination time (from-to): 09:00 – 13:00

Permitted examination support material: All support material is allowed

Academic contact during examination: Bob O'Hara
Phone: 915 54 416

Technical support during examination: Orakel support services
Phone: 73 59 16 00

OTHER INFORMATION

- If a question is unclear/vague – make your own assumptions and specify in your answer the premises you have made. Only contact academic contact in case of errors or insufficiencies in the question set.
- **Saving:** Answers written in Inspera are automatically saved every 15 seconds. If you are working in another program remember to save your answer regularly.
- **Cheating/Plagiarism:** The exam is an individual, independent work. Examination aids are permitted. All submitted answers will be subject to plagiarism control. *Read more about cheating and plagiarism here.*
- **Notifications:** If there is a need to send a message to the candidates during the exam (e.g. if there is an error in the question set), this will be done by sending a notification in Inspera. A dialogue box will appear. You can re-read the notification by clicking the bell icon in the top right-hand corner of the screen. All candidates will also receive an SMS to ensure that nobody misses out on important information. Please keep your phone available during the exam.
- **Weighting:** Weighting of the questions is given for each question.

ABOUT SUBMISSION

- **Your answer will be submitted automatically when the examination time expires and the test closes**, if you have answered at least one question. This will happen even if you do not click “Submit and return to dashboard” on the last page of the question set. You can reopen and edit your answer as long as the test is open. If no questions are answered by the time the examination time expires, your answer will not be submitted.
- **Withdrawing from the exam:** If you wish to submit a blank test/withdraw from the exam, go to the menu in the top right-hand corner and click “Submit blank”. This can not be undone, even if the test is still open.
- **Accessing your answer post-submission:** You will find your answer in Archive when the examination time has expired.

Lizards in the Gym

Researchers were interested in how exercise affects stress in Australian painted dragon lizards, and whether male lizards with different colouration (which indicates different life history strategies) responded differently to exercise.

The researchers measured the size of each lizard before and after a period when lizards were given different levels of exercise. Morph was measured in two ways: head colour (red, orange, yellow, or blue) and whether the individual had a bib.

The response used in this data was the change in body size, which they measure as Snout-Ventral Length (SVL) from before to after the experiment.

i Lizards in the Gym

Researchers were interested in how exercise affects stress in Australian painted dragon lizards, and whether male lizards with different colour morphs (which indicates different life history strategies) responded differently to exercise.

The researchers measured the size of each lizard before and after a period when lizards were given different levels of exercise. Morph was measured in two ways: head colour (red, orange, yellow, or blue) and whether the individual had a bib.

The response used in this data was the change in body size, which they measure as Snout-Ventral Length (SVL), which was measured before and after the experiment.

1 What sort of model would you use for this data?
Select one alternative:

- Generalised Linear Model assuming another distribution
- ✗ Linear Model, assuming a Normal distribution
- Generalised Linear Model assuming a Poisson distribution
- Generalised Linear Model assuming a Binomial distribution

Maximum marks: 1

2 Briefly explain your answer to the last question.

Fill in your answer here

Format

Data is continuous, and probably homoscedastic (2 marks)

or

Is a reasonable default (1 mark)

Words: 0

Maximum marks: 2

Lizards in the Gym

The first model looked at is this:

```
SVL ~ Exercise + HeadCol + HasBib
```

Exercise is a factor with three levels: *Control*, *Low*, *High*
HeadCol is a factor with one level for each head colour: *Blue*, *Orange*, *Red*, *Yellow*.
HasBib is a factor with levels *Yes* and *No*.

The model gave the following summary:

```
Call:
lm(formula = SVLDiff ~ Exercise + HeadCol + HasBib, data = Data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.23089 -0.11475 -0.01475  0.09791  0.33267

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.124239   0.058651   2.118  0.03958 *
ExerciseLow   -0.012659   0.052487  -0.241  0.81049
ExerciseHigh  -0.100257   0.053707  -1.867  0.06832 .
HeadColORANGE -0.023780   0.063067  -0.377  0.70786
HeadColRED    -0.009491   0.061202  -0.155  0.87744
HeadColYELLOW -0.068705   0.061112  -1.124  0.26674
HasBibYES      0.143093   0.047718   2.999  0.00436 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1494 on 46 degrees of freedom
Multiple R-squared:  0.2694,    Adjusted R-squared:  0.1741
F-statistic: 2.827 on 6 and 46 DF,  p-value: 0.01993
```

3 What is the estimated difference between a lizard with an *Yellow* coloured head and one with a *Blue* coloured head?
(the Blue coloured head is the level at the intercept)

Select one alternative:

- -0.013
- 0.143
- 0.124
- -0.024
- ✗ -0.069
- -0.100
- -0.009

(note: you may have been asked a slightly different question: these were randomised)

Maximum marks: 1

Lizards in the Gym

The first model looked at is this:

```
SVL ~ Exercise + HeadCol + HasBib
```

Exercise is a factor with three levels: *Control*, *Low*, *High*
HeadCol is a factor with one level for each head colour: *Blue*, *Orange*, *Red*, *Yellow*.
HasBib is a factor with levels *Yes* and *No*.

The model gave the following summary:

```
Call:
lm(formula = SVLDiff ~ Exercise + HeadCol + HasBib, data = Data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.23089 -0.11475 -0.01475  0.09791  0.33267

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.124239   0.058651   2.118  0.03958 *
ExerciseLow   -0.012659   0.052487  -0.241  0.81049
ExerciseHigh  -0.100257   0.053707  -1.867  0.06832 .
HeadColORANGE -0.023780   0.063067  -0.377  0.70786
HeadColRED    -0.009491   0.061202  -0.155  0.87744
HeadColYELLOW -0.068705   0.061112  -1.124  0.26674
HasBibYES      0.143093   0.047718   2.999  0.00436 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1494 on 46 degrees of freedom
Multiple R-squared:  0.2694,    Adjusted R-squared:  0.1741
F-statistic: 2.827 on 6 and 46 DF,  p-value: 0.01993
```

4 What is the *predicted* SVL for a lizard with an *Orange* head at low exercise and without a bib?
Select one alternative:

- ✗ 0.088
- 0.043
- 0.258
- 0.102
- 0.244
- 0.199

(note: you may have been asked a slightly different question: these were randomised)

Maximum marks: 2

Lizards in the Gym

The first model looked at is this:

```
SVL ~ Exercise + HeadCol + HasBib
```

Exercise is a factor with three levels: *Control*, *Low*, *High*
HeadCol is a factor with one level for each head colour: *Blue*, *Orange*, *Red*, *Yellow*.
HasBib is a factor with levels *Yes* and *No*.

The model gave the following summary:

```
Call:
lm(formula = SVLDiff ~ Exercise + HeadCol + HasBib, data = Data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.23089 -0.11475 -0.01475  0.09791  0.33267

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.124239   0.058651   2.118  0.03958 *
ExerciseLow   -0.012659   0.052487  -0.241  0.81049
ExerciseHigh  -0.100257   0.053707  -1.867  0.06832 .
HeadColORANGE -0.023780   0.063067  -0.377  0.70786
HeadColRED    -0.009491   0.061202  -0.155  0.87744
HeadColYELLOW -0.068705   0.061112  -1.124  0.26674
HasBibYES      0.143093   0.047718   2.999  0.00436 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1494 on 46 degrees of freedom
Multiple R-squared:  0.2694,    Adjusted R-squared:  0.1741
F-statistic: 2.827 on 6 and 46 DF,  p-value: 0.01993
```

5 How much of the variation is explained by this model (as a percentage, to the nearest whole number)? **27** %.

Maximum marks: 1

6 Based on this model and the results presented, what can you say about how and whether exercise affect SVL?

Fill in your answer here

Format

1 for saying there is no effect or it is inconclusive, 1 for reporting effect size (at worst, it is statistically significant).

2 marks for justifying their answer, e.g. quoting confidence intervals or p- values to say effects significant, and results suggest more exercise gives a larger effect.

Words: 0

Maximum marks: 4

7 What other statistical methods could you use to decide if there was an effect of exercise, and how would they show an effect (or no effect)?

Fill in your answer here

Format

1 for raising model comparison

1 for ANOVA (half for BIC)
plus
up to 2 for explaining that would want to see big F-ratio/small p-value

OR give one mark for suggesting confidence intervals.

Words: 0

Maximum marks: 4

The researchers also wanted to know if the effects of exercise was different for the different morphs, so they included interactions between exercise and morph, and exercise and whether the lizard had a bib. The model can be written like this:

```
SVL ~ Exercise*(HeadCol + HasBib)
```

They fitted the model, and compared the model with the main effects with the model where Exercise interacts with head colour, and also with whether the lizard has a bib :

```
mod.Interactions <- lm(SVLDiff ~ Exercise*(HeadCol + HasBib), data=Data)
anova(mod.MainEffects, mod.Interactions)
```

Which gave the output below:

```
Analysis of Variance Table

Model 1: SVLDiff ~ Exercise + HeadCol + HasBib
Model 2: SVLDiff ~ Exercise * (HeadCol + HasBib)
  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     46 1.02669
2     38 0.67152  8   0.35517 2.5123 0.0268 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

To remind you:

Exercise is a factor with three levels: *Control*, *Low*, *High*

HeadCol is a factor with one level for each colour morph (*Blue*, *Orange*, *Red*, *Yellow*)

HasBib is a factor with levels Yes and No.

8 Is this problem, of whether there are interaction effects, exploratory or confirmatory?

Select one alternative:

- Exploratory
- ✗ Confirmatory

Maximum marks: 1

9 Explain briefly why you think it it exploratory or confirmatory.

Fill in your answer here

Format

Up to 2 marks: It is test a specific hypothesis (albeit a complex one!)

Half a mark (at most) for "because it used ANOVA".

Words: 0

Maximum marks: 2

10 Does this suggest that the effect of exercise is different for different morphs (i.e. lizards with different head colours or with/without a bib)? Quote the statistics you use to come to this conclusion.

Fill in your answer here

Format

1 mark for "yes"

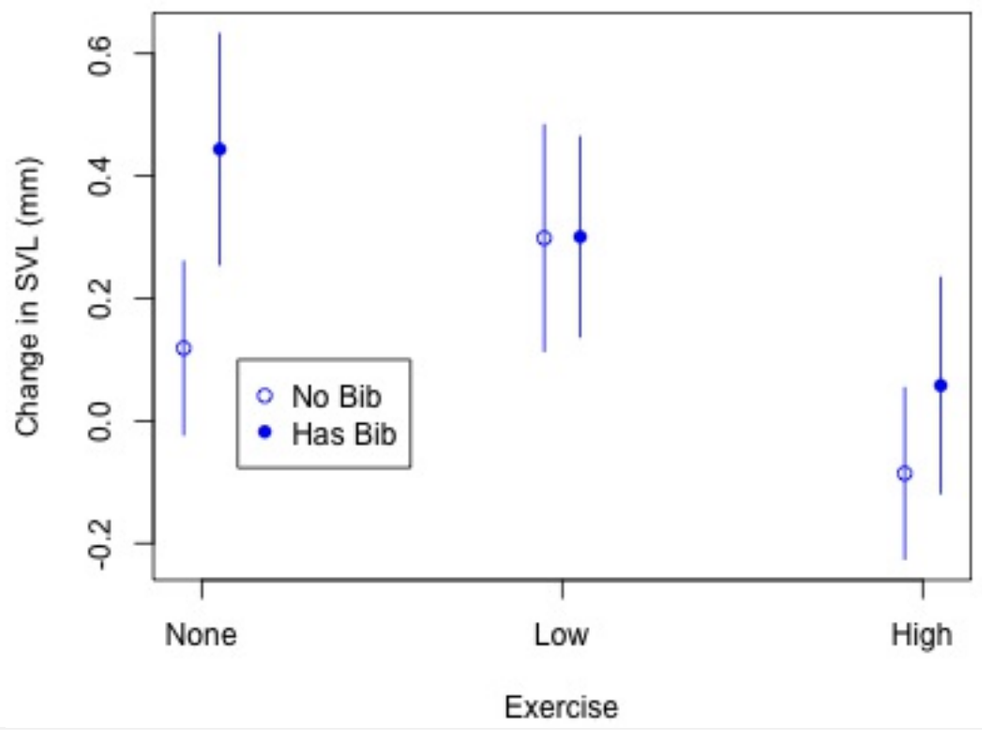
+ 1 mark for weak evidence etc.

2 marks for quoting statistics, such as estimates, p-value and F statistic

Words: 0

Maximum marks: 4

11



The estimates of the means for the blue morph for different exercises and whether the lizard has a bib are plotted here (bars are 95% confidence intervals).

How does the effect of exercise vary depending on whether the lizard has a bib?

Fill in your answer here

Format

(note that there are a couple of ways to approach answering this question)
1 mark: with no exercise, increase in SVL higher if has bib.
1 mark: more similar with exercise.
1 mark at high exercise, not much change in size (i.e. don't grow)
1 mark: with bib, exercise makes lizards smaller

Up to 2 marks: using CIs to assess uncertainty

Words: 0

Maximum marks: 4

The next few questions are short and (slightly) random.

12 Which of these is the best description of a maximum likelihood estimate?

Select one alternative:

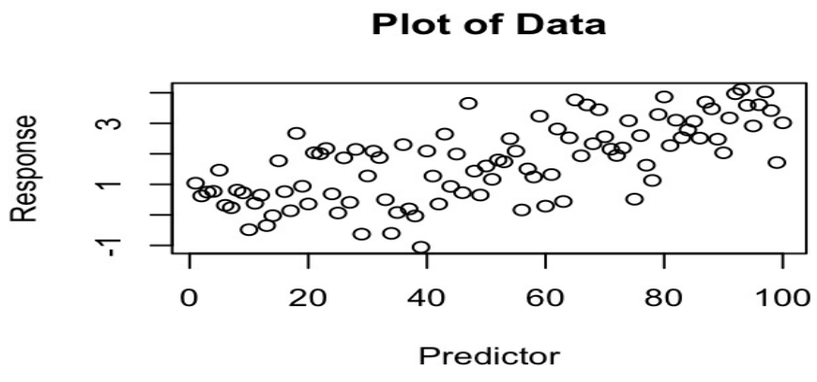
- The most likely value of the parameter
- The estimate of the data that makes the parameter most likely
- ✗ The value of the parameter that makes the data most likely
- The estimate most likely to to be true

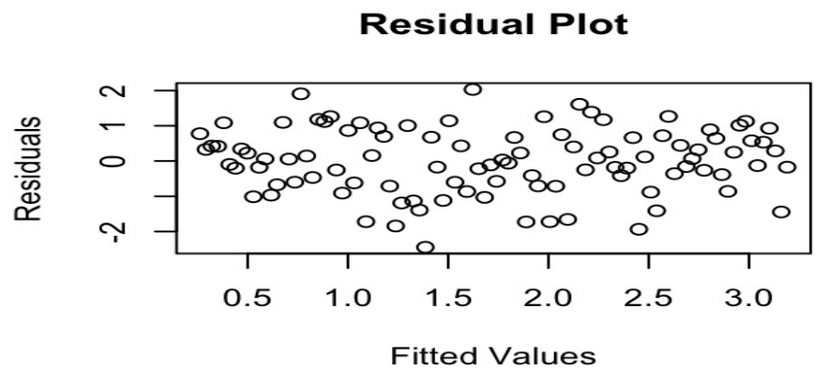
Maximum marks: 1

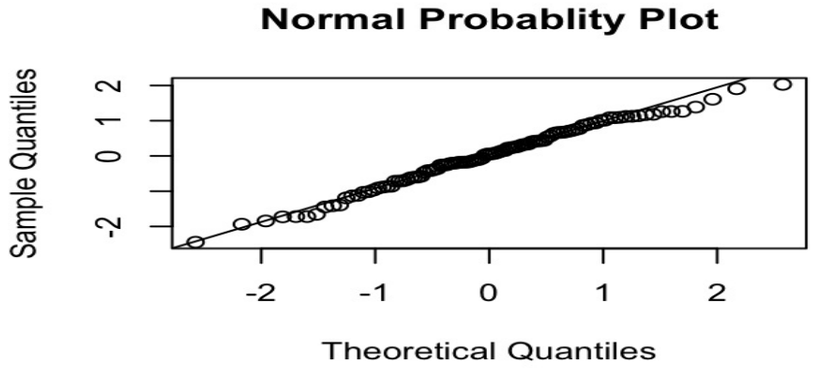
13 Which of these statements about a 95% confidence interval is correct?
Select one alternative:

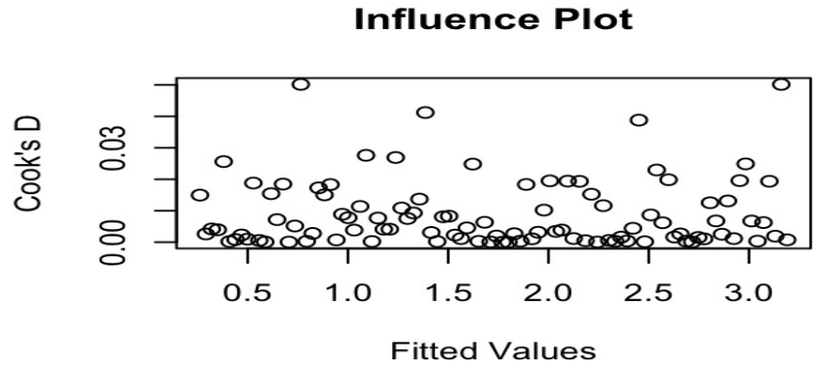
- I am 95% certain to get this answer wrong
- If we repeated the data collection, there is a 95% probability that the new estimate would be true
- If we repeated the data collection, there is a 95% probability of getting an estimate of the parameter within the limits
- ✗ If we repeated the data collection, there is a 95% probability that the true value of the parameter will be within the new confidence interval.

Maximum marks: 1



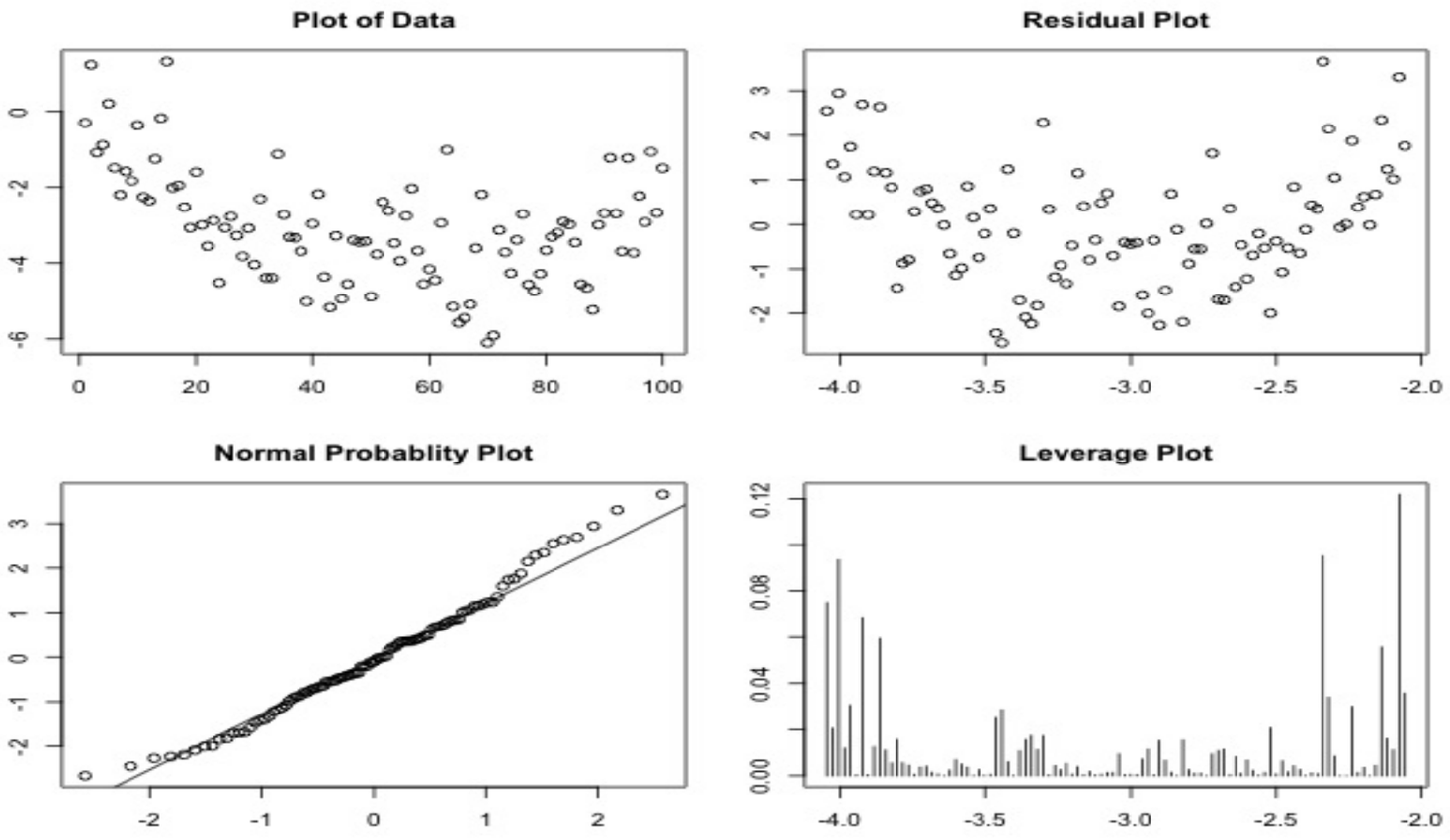






This is a plot of the data and some plots for assessing model fit, for a data set.

14



This is a plot of the data and some plots for assessing model fit, for a data set.

Which aspects of the model do you think are problematic?
Select one or more alternatives:

- There are outliers
- The model seems fine
- The residuals are not normally distributed
- X The relationship between X and Y is not linear
- The variance of the residuals changes (i.e. they are heteroscedastic)
- Some points are influential

(note: you may have been asked a slightly different question:
the data sets were randomised, all with different problems)

Maximum marks: 2

15 Based on your assessment, how could you improve the model?
(if you think the model is fine, just say that)

Format

Outliers: Check outlier, possibly remove

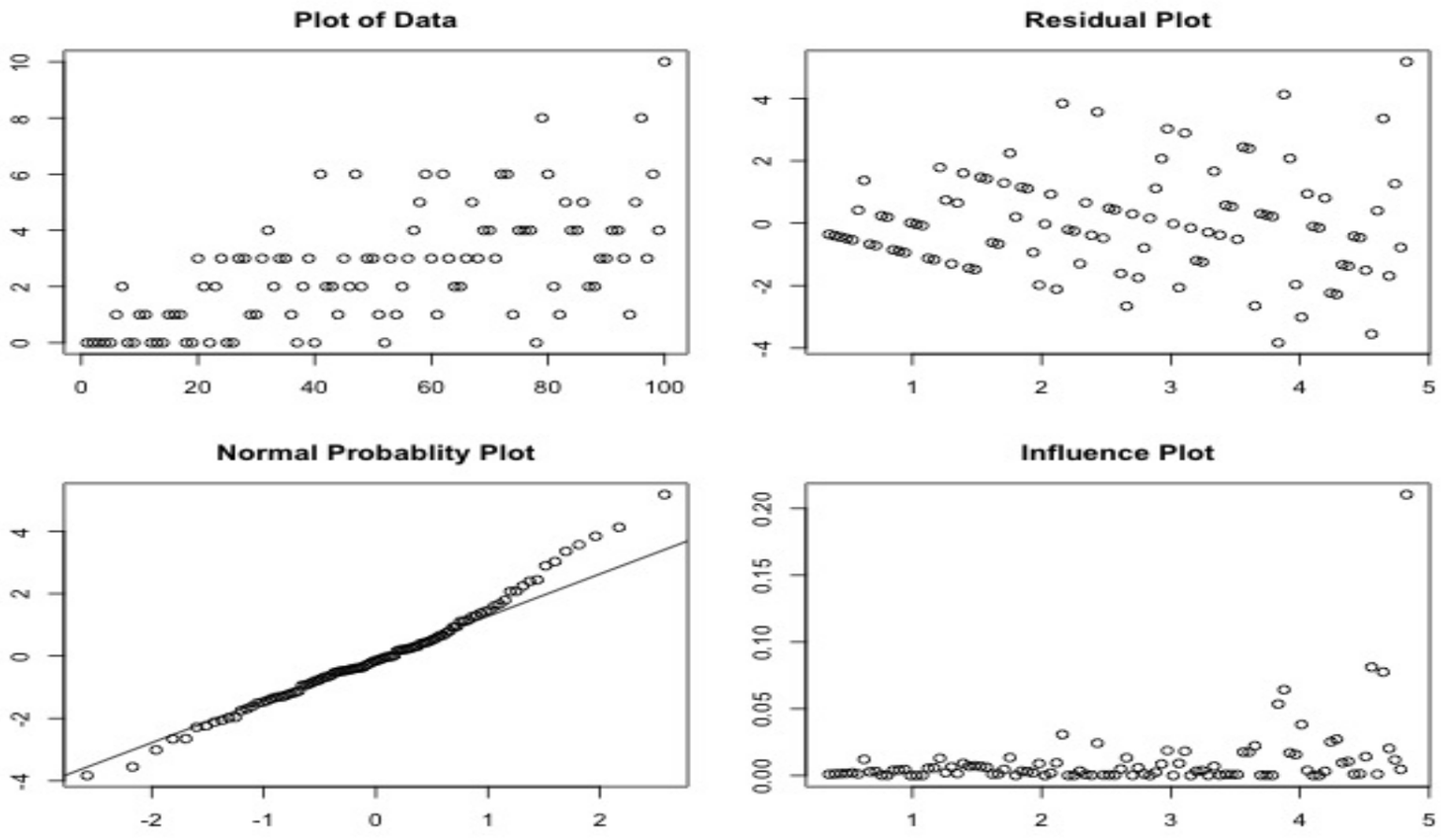
Heteroscedasticity: Box-Cox transformation (or something else, even if we haven't covered it, e.g. GLS, GLM)

Non-linear: Box-Cox transformation or polynomials

Words: 0

Maximum marks: 2

16



Here is another data set. Which aspects of the model do you think are problematic?
Select one or more alternatives:

- [x] The variance of the residuals changes (i.e. they are heteroscedastic)
- [ ] The model seems fine
- [ ] There are outliers
- [x] Some points are influential
- [x] The residuals are not normally distributed
- [ ] The relationship between X and Y is not linear

(note: you may have been asked a slightly different question:
the data sets were randomised, all with different problems.
They were all actually generated from a GLM)

Maximum marks: 2

17 Based on your assessment, how could you improve the model?
(if you think the model is fine, just say that)

Format

For all:
2 marks: try a GLM (Poisson or possibly Binomial)
+ 2 marks if comment on check type of data.

Alternatives (max 3 marks?):

Poisson: Box-Cox
Binomial 1: Add quadratic term
Binomial 2: Might be OK, or add quadratic.

Words: 0

Maximum marks: 4

As part of quality assurance at Camelot, King Arthur has demanded a report into the work of his knights. As part of the report, you are looking at whether they have been carrying out enough quests. (yes, I know you are biologists, but it was either this or the cleaning out the Dragon pits, a task even Hercules declined).

Questing can be affected by a variety of things, such as how long the knight has been at the round table and whether the knight spent their time on the Grail Quest (which was very important but mean they couldn't go on other quests). Here you have the following variables:

- **Time**: Length of time in service to King Arthur. We will use a natural log transformation of time.
- **Grail**: Have taken part in the Grail Quest (factor with 2 levels: "No", "Yes")
- **Tournaments**: Number of Tournaments won
- **Weapon**: Has a holy weapon (factor with 2 levels: "Normal", "Holy")
- **Armour**: Has magical armour (factor with 2 levels: "Normal", "Magic")
- **Retinue**: Has a retinue, (i.e. a group who follow him around mostly trying to help (factor with 2 levels: "No", "Yes").

Your task is to help work out what makes a knight a successful questor, and identify any exceptionally good or bad knights.

Questing is measured as the number of successful quests the knight has been on.

18 Is this problem exploratory or confirmatory?

Select one alternative:

- Confirmatory
- X Exploratory

Maximum marks: 1

19 Your supervisor (Merlin) has told you to model the quests with a Poisson distribution. Does this make sense? Justify your answer.

Fill in your answer here

Format

One mark: yes.

2 marks: count data, can have "rate" of questing (or Poisson approximation to binomial?)

No marks for "because if I don't, I will be turned into a frog", or similar. Sorry.

Words: 0

Maximum marks: 3

All combinations of models were fitted, and different statistics calculated to compare the models. The summaries of the best models are below: all of the others were much worse by any criterion.

Table 1: Number of parameters ('Pars'), Deviance, AIC, and BIC for models with Deviance below 1100.

Model	Pars	Deviance	AIC	BIC
Time + Retinue	2	91.0	95.0	102.1
Time + Grail + Retinue	3	85.7	91.7	102.3
Time + Weapon + Retinue	3	90.0	96.0	106.5
Time + Armour + Retinue	3	90.2	96.2	106.8
Time + Tournaments + Retinue	3	90.7	96.7	107.3
Time + Grail + Weapon + Retinue	4	83.7	91.7	105.8
Time + Grail + Tournaments + Retinue	4	85.5	93.5	107.5
Time + Grail + Armour + Retinue	4	85.5	93.5	107.6
Time + Weapon + Armour + Retinue	4	88.9	96.9	111.0
Time + Tournaments + Weapon + Retinue	4	89.7	97.7	111.8
Time + Tournaments + Armour + Retinue	4	89.9	97.9	112.0
Time + Grail + Weapon + Armour + Retinue	5	83.3	93.3	111.0
Time + Grail + Tournaments + Weapon + Retinue	5	83.5	93.5	111.0
Time + Grail + Tournaments + Armour + Retinue	5	85.2	95.2	112.8
Time + Tournaments + Weapon + Armour + Retinue	5	88.7	98.7	116.3
Time + Grail + Tournaments + Weapon + Armour + Retinue	6	83.1	95.1	116.3

20 Which statistic is best to use to compare these models, if we want to explain why one or another knight it best?
Select one alternative:

- AIC
- Deviance
- Number of Parameters
- ✗ BIC

Maximum marks: 1

21 Which is the best model?

Select one alternative:

- Time + Weapon + Armour + Retinue
- Time + Weapon + Retinue
- Time + Grail + Tournaments + Retinue
- Time + Tournaments + Weapon + Armour + Retinue
- Time + Grail + Tournaments + Armour + Retinue
- ? Time + Grail + Weapon + Retinue
- X Time + Retinue
- Time + Grail + Tournaments + Weapon + Retinue
- Time + Tournaments + Retinue
- Time + Tournaments + Retinue
- Time + Tournaments + Armour + Retinue
- ? Time + Grail + Retinue
- Time + Armour + Retinue
- Time + Grail + Weapon + Armour + Retinue
- Time + Grail + Tournaments + Weapon + Armour + Retinue
- Time + Grail + Armour + Retinue

Note: The ?'s are the answer(s) if someone suggested using AIC, and in that case was also marked correctly

Maximum marks: 2

22 Why do you think this is the best model? Which statistics did you use to come to this conclusion and why? Are there any other models you might consider?

Fill in your answer here

Format

1 mark for "it has the lowest BIC." (or equivalent), 1 mark for citing it.

2 marks for citing similar models (e.g. "model Z has and BIC that is only 1 higher, but is a bit simpler, so we could also consider that").

The marks are, of course, adjusted accordingly for people using AIC.

Words: 0

Maximum marks: 2

After some discussion, it was decided to use the following model:

```
Time + Tournaments + Retinue
```

even though it is not the best model. Fitting the model gave the following summary:

```
Call:
glm(formula = NQuests ~ Time + Grail + Retinue, family = "poisson",
    data = KnightData)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.2724  -0.8722  -0.0305   0.5432   3.3272

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.06384    0.10868   0.587    0.557
Time         1.20672    0.05424  22.247   <2e-16 ***
GrailYes     0.11404    0.04940   2.308    0.021 *
RetinueYes  -0.50997    0.05061 -10.077   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1131.89  on 249  degrees of freedom
Residual deviance:  269.06  on 246  degrees of freedom
AIC: 1093.7

Number of Fisher Scoring iterations: 5
```

23 Which link function was used in this model?

Select one alternative:

- identity
- ✗ log
- cloglog
- logit
- probit

Maximum marks: 1

24 Based on this model, explain (briefly) what sort of knight you think is best for a quest.

Fill in your answer here

Format

2 marks for right model (1 if partially right), i.e. been knight a long time,
done grail quest, but doesn't have a retinue
+ 1 mark for quoting stats
+ 1 mark for assessing size or doing more interpretation
+ 1 mark for assessing size (e.g. grail effect small)

Notes (hard points, for which extra marks would be awarded):
1. The time effect is >1, so it is not just that a knight that has been in service twice as long has done twice as many quests: it is more than that. This suggests knights get better.
2. Even though grail questing takes time, knight who have been on it have done more quests. This is probably because they were better knights anyway, so would have done more quests if they had not been hunting the grail.

Words: 0

Maximum marks: 5

Merlin has narrowed his interest in on two knights for their performance review:

- Sir Robin: has been a knight for 5 years, has a retinue and said he went on the grail quest. He has completed 0 quests.
- Sir Dagonet: has been a knight for 2 years, does not has a retinue and did not go on the grail quest. He has completed 8 quests.

Merlin wants to know if either is performing below expectations. So first he wants you to calculate how many quests you would expect them to have performed.

(hint: Time was log transformed, using natural logs, so log(5) = 1.61 and log(2) = 0.69).

25 What is the prediction on the link scale for the number of quests for Sir Robin? **1.61** (answer to 2 decimal places). **Calculation: 0.06 + 1.21*log(5) + 0.11*1 - 0.51*1**

What is the prediction for the actual number of quests of Sir Robin?

5.0 (answer to 1 decimal places) **Calculation: exp(1.61)**

Maximum marks: 4

26 What is the prediction on the link scale for the number of quests for Sir Dagonet? **0.9** (answer to 2 decimal places) **Calculation: 0.06 + 1.21*log(2) + 0.11*0 - 0.51*0**

What is the prediction for the actual number of quests of Sir Dagonet? **2.5** **Calculation: exp(0.9)**

(answer to 1 decimal place)

(I calculated this as 2.46, so I allowed a bit of rounding leeway)

Maximum marks: 4

27 If you compare the expected number of quests that you have calculated and the actual number they completed (**Sir Robin: 0, Sir Dagonet: 8**), what can you say about how well these knights performed? What would you recommend to Merlin to help the knights' performance, and the performance of the Round Table??

Fill in your answer here

Sir Robin (for a short documentary , see here: https://www.youtube.com/watch?v=jYFefppqEtE)
1 mark: He is underperforming (massively)
1 mark: Ditching his retinue would help a bit, and getting magical armour & weapons.
1 mark: he's useless (or summary to that effect)

Sir Dagomat
1 mark: over-performing
1 mark: but not by much
1 mark: Get better armour/weapon
1 mark: keep Sir Dagomat, feed Sir Robin to the fishes.

I was flexible with this, but was looking for ability to assess the data and interpret it.

Maximum marks: 6