How do literacy and age of a marriage affect family size?

Maryam Ansari & Alexia Xu

January 27, 2025

Contents

```
##   age ageMarried monthsSinceM pregnancies children sons region literacy
## 1  43     22to25          242           3        3    2  lt10k      yes
## 2  32     22to25          124           1        1    0  lt10k      yes
## 3  22     15to18           59           1        1    1  lt10k      yes
## 4  28     22to25           63           1        1    0  lt10k      yes
## 5  30     15to18          169           2        2    2  lt10k      yes
## 6  37     18to20          226           2        2    1  lt10k      yes
```

```
##
##  lt10k lisbon  porto   20k+ 10-20k
##   3502    470    160    583    433
```

Introduction

Checklist

- Strong and relevant motivation as to why it is important to study this particular research question
- Summaries of the main results from three peer-reviewed articles on same or relevant topic
- Make sure that articles talk about which variables have been shown to be related to the response variables in your data.

Actual Text

Methods

Checklist

- Choice of GLM model that will be fit
- Justification for model based on data and research goal
- Process for decision on the appropriateness of the model.

Planning

- Categorical: age, ageMarried, region
- Binary: Literacy
- Response: children -> count variable
- Poisson should be best since response is a count variable?
- Maybe neg binomial?

Fitting the model:

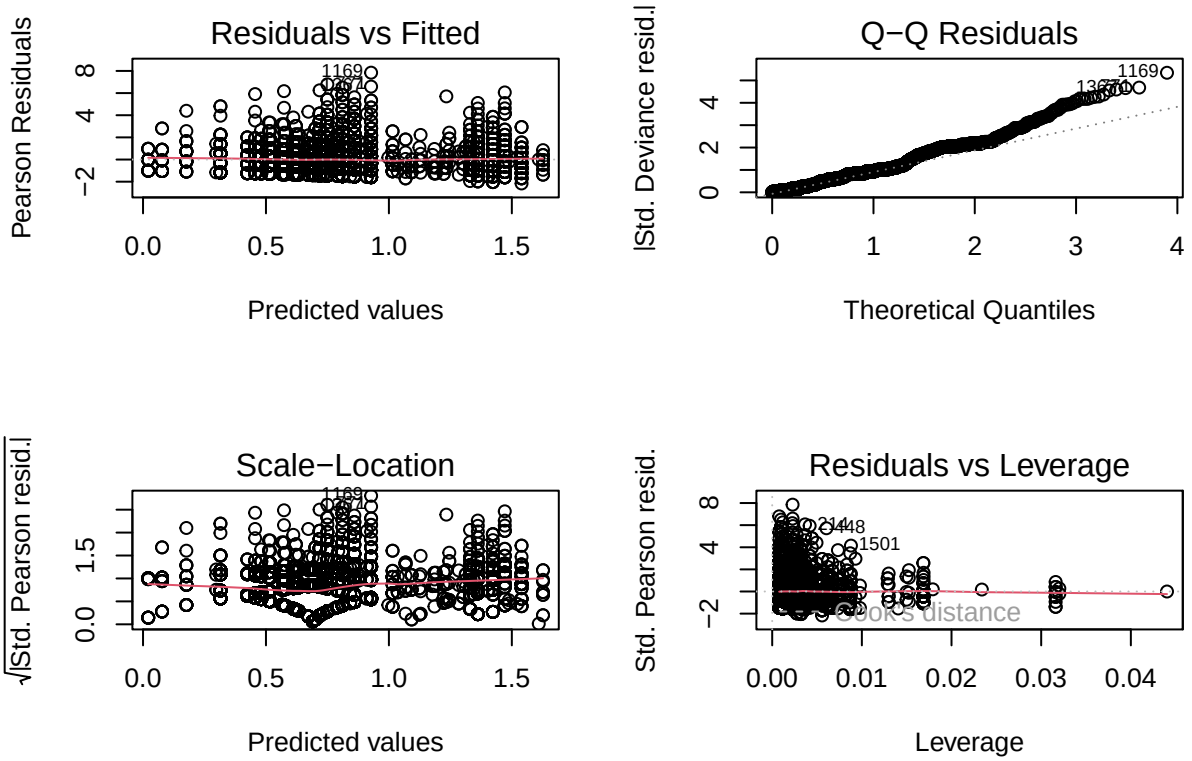
Checking for overdispersion:

```
## Residual Deviance: 6047.284
```

```
## Degrees of Freedom: 5136
```

```
## Overdispersion Ratio: 1.177431
```

ChatGPT interpretation of overdispersion ratio - If the overdispersion ratio is close to 1, the Poisson model is a good fit for the data, as the variance roughly equals the mean (an assumption of the Poisson distribution). - A ratio of 1.177 is slightly greater than 1, suggesting mild overdispersion (the variance is slightly greater than the mean).



Residual Analysis:

Actual Text

Results

Checklist

- Statistical summaries of the data and variables being considered
- Discussion of the data distributions and their preliminary implications regarding the research question.
- Presentation of model results (e.g. estimated coefficients, standard errors, CIs) and evidence supporting inclusion/exclusion of predictors.
- Discussion of the process and results involved in developing a suitable model that addresses the research goal

Actual Text

Conclusion

Checklist

- A formal statistical interpretation of a coefficient of at minimum one of the most important or relevant predictors.
- A general summary of what your model as a whole says about the research question.
- A comparison to the literature results summarized in the introduction section, commenting on what is similar and what is different.

Actual Text

References