Generalized Linear Models*

Logistic Regression

Tam Ly

March 5, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

2 Data

3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i|\pi_i \sim \text{Bern}(\pi_i) \tag{1}$$
$$\text{logit}(\pi_i) = \alpha + \beta_1 \times \text{gender}_i + \beta_2 \times \text{education}_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta_1 \sim \text{Normal}(0, 5.02) \tag{4}$$
$$\beta_2 \sim \text{Normal}(0, 6.34) \tag{5}$$

*Code and data are available at: https://github.com/atn-ly/politics.

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

4 Results

Our results are summarized in Table 1.

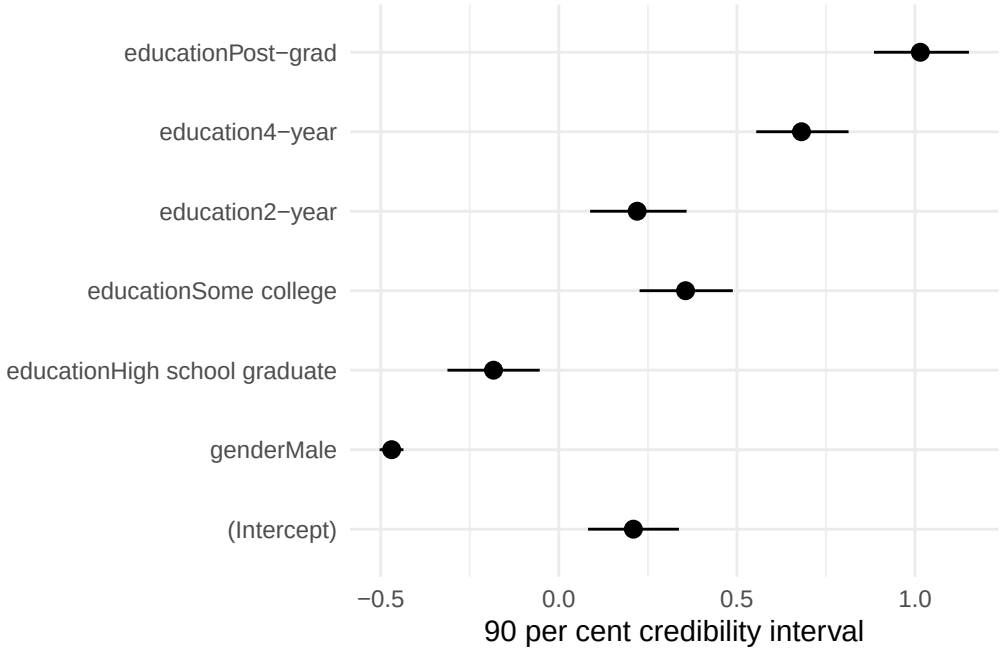


Figure 1: Explanatory models of flight time based on wing width and wing length

5 Discussion

Table 1: Explanatory models of flight time based on wing width and wing length

	First model
(Intercept)	0.21
	(0.07)
genderMale	−0.47
	(0.02)
educationHigh school graduate	−0.18
	(0.08)
educationSome college	0.36
	(0.08)
education2-year	0.22
	(0.08)
education4-year	0.68
	(0.08)
educationPost-grad	1.01
	(0.08)
Num.Obs.	43 554
R2	0.046
Log.Lik.	−28 357.925
ELPD	−28 364.9
ELPD s.e.	58.1
LOOIC	56 729.7
LOOIC s.e.	116.3
WAIC	56 729.7
RMSE	0.48

Appendix

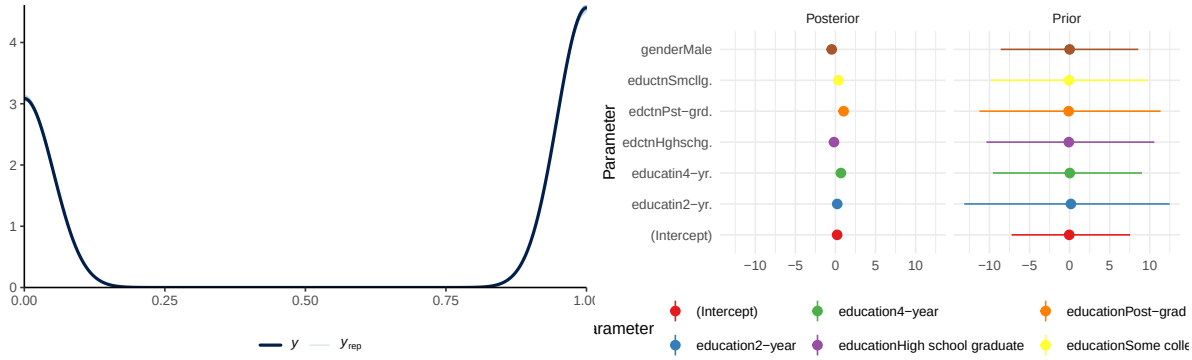
A Additional data details

B Model details

B.1 Posterior predictive check

In Figure 2a we implement a posterior predictive check. This shows...

In Figure 2b we compare the posterior with the prior. This shows...



(a) Posterior prediction check

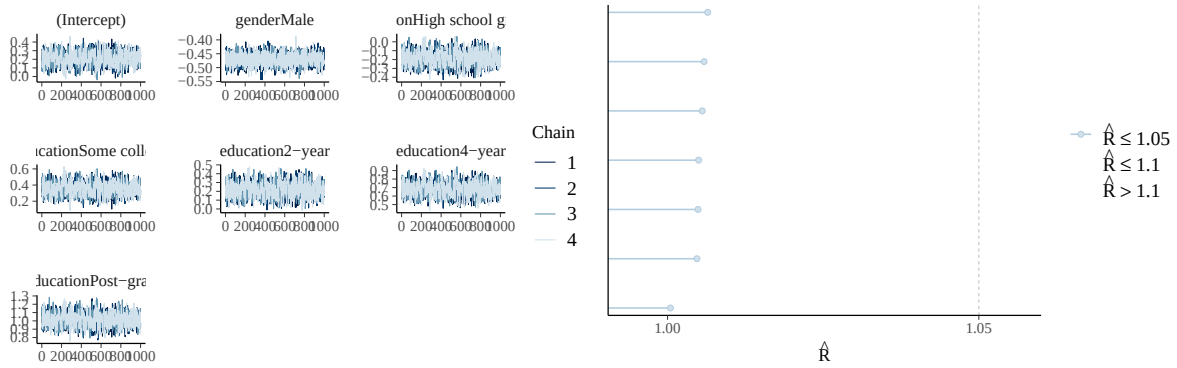
(b) Comparing the posterior with the prior

Figure 2: Examining how the model fits, and is affected by, the data

B.2 Diagnostics

Figure 3a is a trace plot. It shows... This suggests...

Figure 3b is a Rhat plot. It shows... This suggests...



(a) Trace plot

(b) Rhat plot

Figure 3: Checking the convergence of the MCMC algorithm

References

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.