Homework 3 - Question 2a

Riley Smith

05 December 2016

Data Summary

	M	SD	Min	Max	NAs
id	1186	631.2	16	2355	0
lonely1	0.41	0.49	0	1	0
lonely2	0.28	0.45	0	1	0
incadq1	3.19	0.88	1	4	0
health1	2.9	0.7	0	4	0
age1	70.88	6.23	59	86	0
educ1	13.17	1.98	6	19	0

Question 2.a.

Table 2: Variable Name Assignments for Analysis

Original Variable Name	Label in Model
lonely2	Y
lonely1	x1
incadq1	x2
health1	x3

Table 3: Fitting generalized (binomial/logit) linear model: Y ~ x1 + x2 + x3

	Estimate	Std. Error	z value	Pr(>|z|)
x1	1.711	0.283	6.048	0.000000001471
x2	0.04948	0.1566	0.3159	0.752
x3	-0.4606	0.1989	-2.316	0.02056
(Intercept)	-0.6602	0.7479	-0.8827	0.3774

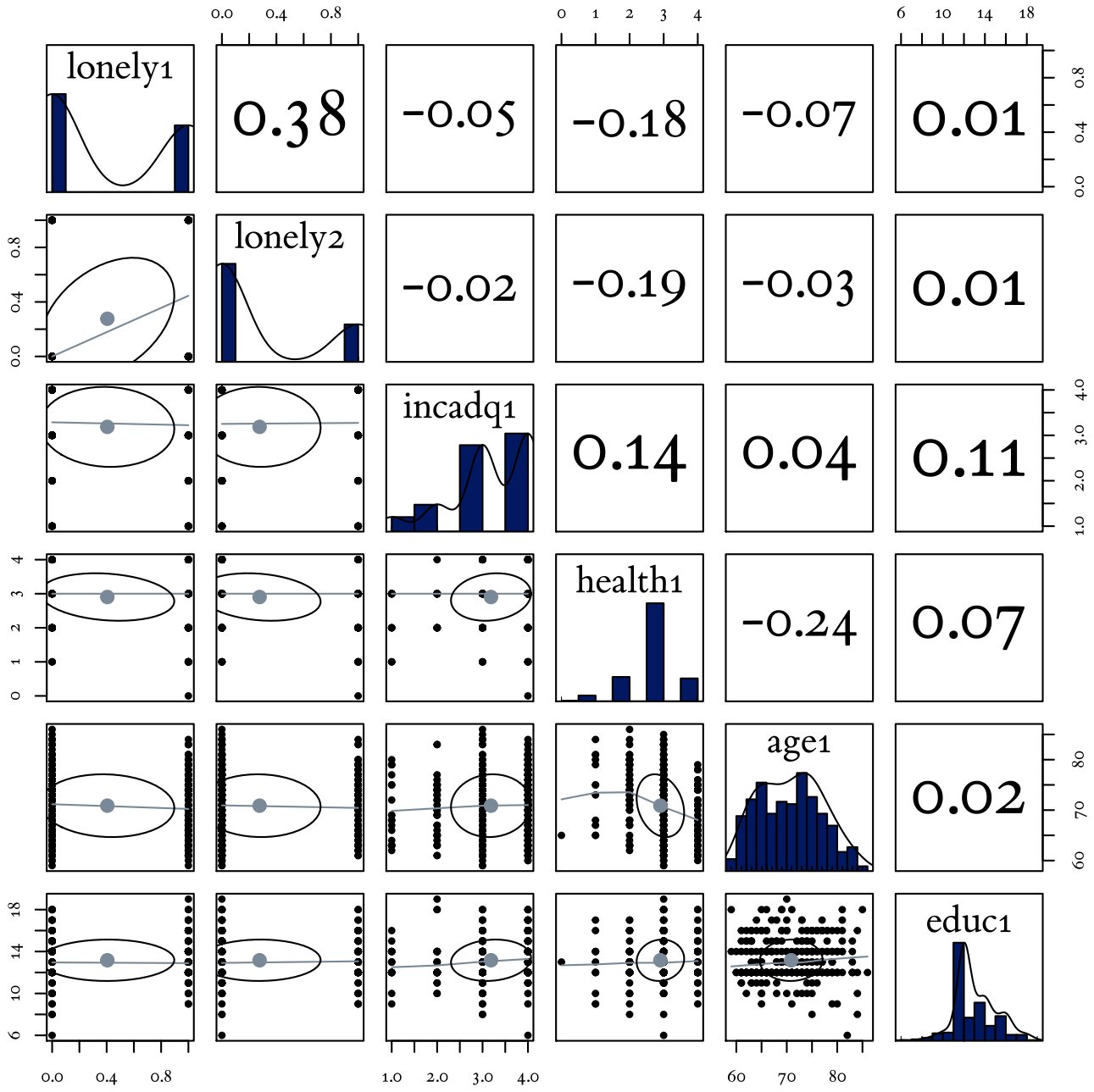

Figure 1: Bivariate correlations (upper), implied bivariate loess model curves (lower), and individual variable distributions (diagonal)

Table 4: Analysis of Deviance Table

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			310	366.8	
x1	1	45.68	309	321.1	1.392e-11
x2	1	0.001243	308	321.1	0.9719
x3	1	5.544	307	315.5	0.01854

Table 5: Likelihood Ratio χ^2

Log Likelihood	*df*	χ^2	*p*
-183.4	1	51.23	0

Table 6: Logistic Regression Coefficients (β) & Coresponding Confidence Intervals (*CI*)

		CI_β	
	β	2.5 %	97.5 %
(Intercept)	-0.6602	-2.1408	0.8026
x1	1.7114	1.167	2.2795
x2	0.0495	-0.2547	0.3617
x3	-0.4606	-0.8596	-0.0765

Table 7: Logistic Regression Odds Ratios (Φ) & Coresponding Confidence Intervals (*CI*) [1]

		CI_Φ	
	Φ	2.5 %	97.5 %
(Intercept)	0.5168	0.1176	2.2314
x1	5.5368	3.2125	9.7718
x2	1.0507	0.7751	1.4357
x3	0.6309	0.4233	0.9263

Note:

[1] Confidence intervals are based on the logistic regression model's profiled log-likelihood function, rather than the standard errors

Table 8: Additional Logistic Regression Model Fit Statistics

	Estimate	Degrees of Freedom
Null Deviance	366.76	310
Residual Deviance	315.53	307
AIC	323.53	

Lagged Logistic Regression Summary

A lagged logistic regression model was tested to investigate whether baseline reported income adequacy and health predicted loneliness at Time-2 after controlling for loneliness at Time-1. In line with expectations based on bivariate correlations among the outcome variable and each predicter included in the model (see Figure 1), income adequecy reported at baseline did not significantly relate to Time-2 reported loneliness ($b = 0.05, SE = 0.16, OR = 1.05, p = 0.752$), whereas both Time-1 reported loneliness and baseline health significantly predicted loneliness at Time-2. Specifically, respondents higher in loneliness at Time-1 were more likely to have higher levels of Time-2 loneliness relative to respondents with lower levels of loneliness at either timepoint, $b = 1.71, SE = 0.28, OR = 5.54, p < 0.001$. In contrast, respondents with higher baseline health scores were more likely to have lower Time-2 reported loneliness relative to respondents with lower baseline health scores and higher Time-2 reported loneliness, $b = -0.46, SE = 0.2, OR = 0.63, p < 0.05$. Overall, the model accounted for 14 of the variance in reported Time-2 loneliness (*McFadden's pseudo-*$R^2 = 0.14$).