

DATA557 - ex07

Will Wright

2/27/2019

Applied Statistics and Design of Experiments Exercise 7 February 27, 2019

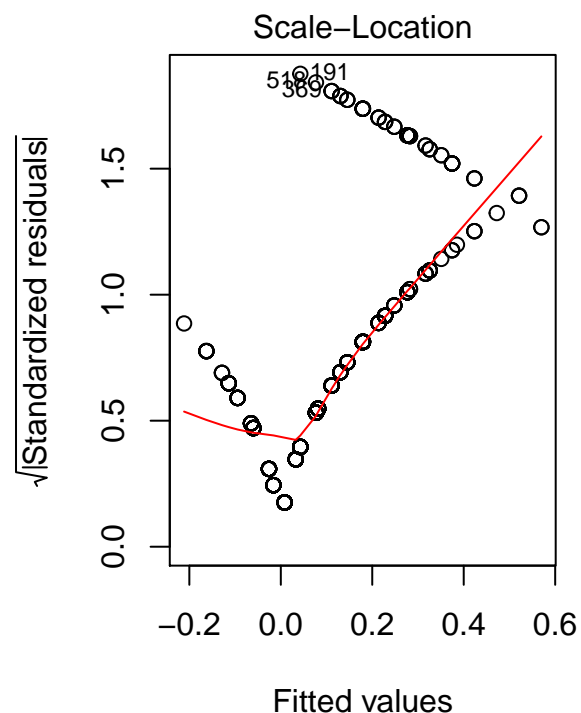
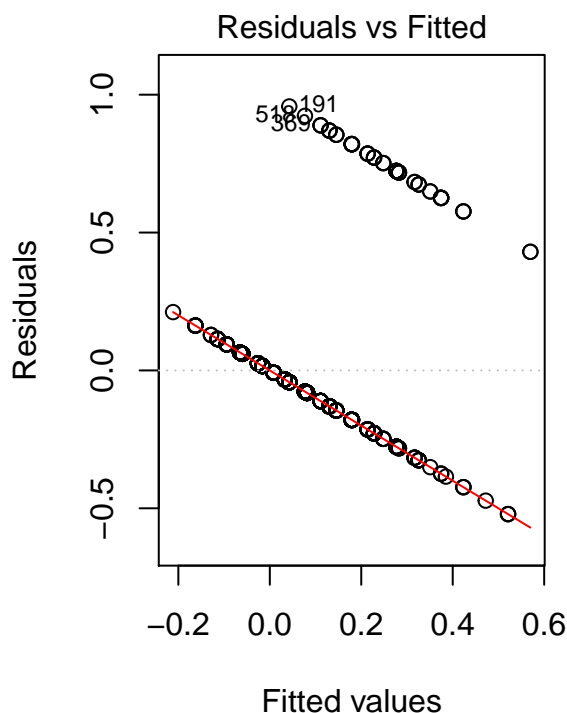
For these exercises you may work individually or in groups. There are no uploads required for this exercise.

Part 1.

Q1. Fit a linear regression model to the fev data with the smoking indicator as the response variable, and with age and male gender, as well as their interaction, as predictor variables. Assess the assumptions of this model. Compare the default SEs and the robust SEs for the coefficient estimates. Is there evidence for an interaction between age and gender?

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.3580340	0.0535707	-6.683391	0.0000000
age	0.0488354	0.0052167	9.361307	0.0000000
male	0.0924192	0.0748240	1.235154	0.2172196
age:male	-0.0145868	0.0072263	-2.018560	0.0439431

	(Intercept)	age	male	age:male
(Intercept)	0.0035059	-0.0004195	-0.0035059	0.0004195
age	-0.0004195	0.0000519	0.0004195	-0.0000519
male	-0.0035059	0.0004195	0.0066296	-0.0007835
age:male	0.0004195	-0.0000519	-0.0007835	0.0000953



Q2. Fit a logistic regression model corresponding to the linear regression model in Q1.1. Compare the results from the linear and logistic models.

Part 2.

Data set: “Teeth.csv”

Variables:

EXTR: number of teeth extracted due to periodontal disease PDALL: measure of disease severity (average pocket depth) GENDER: M/F AGE: age of patient (yrs)

Q3. Use Poisson regression to assess the association between PDALL and EXTR adjusting for age and gender. Compare default and robust SEs. Is there evidence from this model that age or gender are associated with EXTR? Interpret the exponentiated coefficient of gender in the model. Compare the association between EXTR and PDALL with and without adjusting for age and gender.