Example 4: Brain and Weight Data

Contents

1 Introduction **1**

References **2**

1 Introduction

Table 1 presents the brain weight (in grams) and the body weight (in kilograms) of 28 animals (this sample was taken from larger data sets in Weisberg (1980) and Jerison (1973)). It is to be investigated whether a larger brain is required to govern a heavier body. A clear picture of the relationship between the logarithms (to the base 10) of these measurements is shown in Figure 1. This logarithmic transformation was necessary because plotting the original measurements would fail to represent either the smaller or the larger measurements. Indeed, both original variables range over several orders of magnitude. A linear fit to this transformed data would be equivalent to a relationship of the form

$$\hat{y} = \hat{\Theta}'_2 x^{\hat{\Theta}_1}$$

between brain weight (y) and body weight (x).

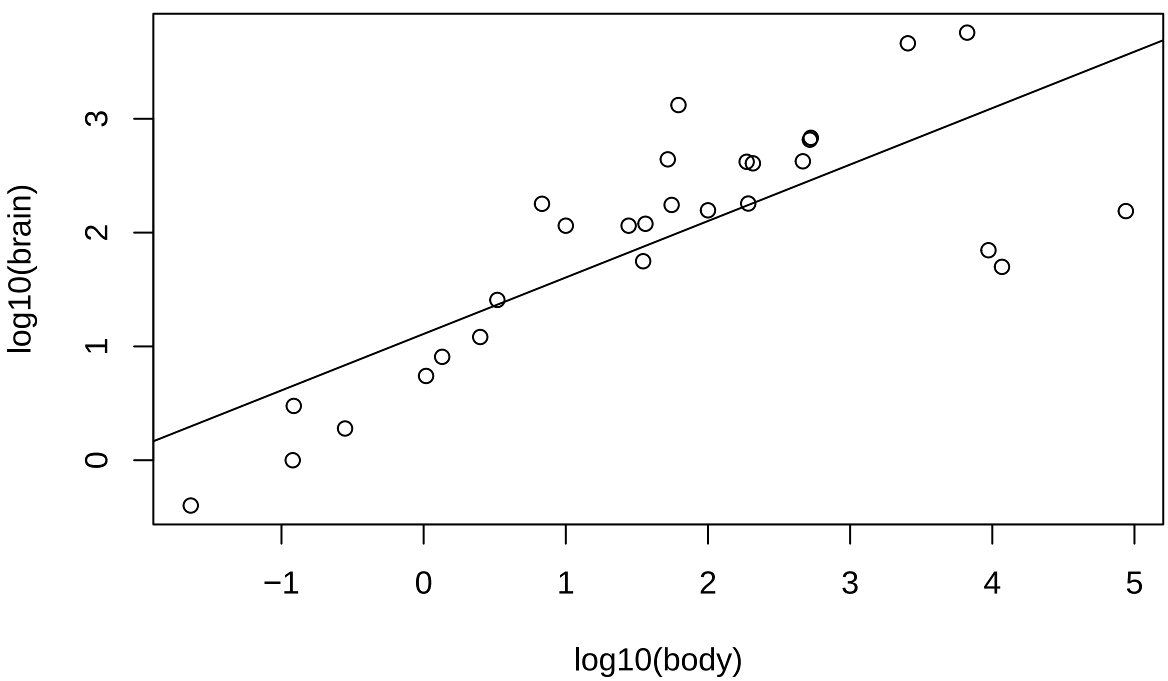

Figure 1: Logarithmic brain weight versus logarithmic body weight for 28 animals

Looking at Figure 1, it seems that this transformation makes things more linear. Another important advantage of the log scale is that the heteroscedasticity disappears.

Table 1: Table 7. Body and Brain Weight for 28 Animals

	Body Weight (kg)	Brain Weight (g)
Mountain beaver	1.350	8.1
Cow	465.000	423.0
Grey wolf	36.330	119.5
Goat	27.660	115.0
Guinea pig	1.040	5.5
Dipliodocus	11700.000	50.0
Asian elephant	2547.000	4603.0
Donkey	187.100	419.0
Horse	521.000	655.0
Potar monkey	10.000	115.0
Cat	3.300	25.6
Giraffe	529.000	680.0
Gorilla	207.000	406.0
Human	62.000	1320.0
African elephant	6654.000	5712.0
Triceratops	9400.000	70.0
Rhesus monkey	6.800	179.0
Kangaroo	35.000	56.0
Golden hamster	0.120	1.0
Mouse	0.023	0.4
Rabbit	2.500	12.1
Sheep	55.500	175.0
Jaguar	100.000	157.0
Chimpanzee	52.160	440.0
Rat	0.280	1.9
Brachiosaurus	87000.000	154.5
Mole	0.122	3.0
Pig	192.000	180.0

The LS fit is given by

$$\log \hat{y} = 0.49601 \log x + 1.10957$$

(dashed line in Figure 1). The standard error associated with the slope equals 0.0782, and that of the intercept term is 0.1794. In Section 3, we explained how to construct a confidence interval for the unknown regression parameters. For the present example, n = 28 and p = 2, so one has to use the 97.5% quantile of the t-distribution with 26 degrees of freedom, which equals 2.0555. Using the LS results, a 95% confidence interval for the slope is given by t0.3353; 0.65673. The RLS yields the solid line in Figure 1, which is a fit with a steeper slope:

$$\log \hat{y} = 0.75092 \log x + 0.86914$$

References

Jerison, Harry J. 1973. "Introduction to the Basic Vertebrate Radiation." In *Evolution of the Brain and Intelligence*, edited by Harry J. Jerison, 97. Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-385250-2.50012-2.

Weisberg, Sanford. 1980. *Applied Linear Regression*. Vol. 528. John Wiley & Sons.