MATH 4432 Mini-Project 1: Sleep Analysis for Mammalian Species

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Haojie WANG
Department of Civil and Environmental Engineering, HKUST

1. Introduction

Sleep is an essential factor for the health of mammalian species. To study the influence factors which affect mammals' sleep, a mammal sleep dataset is systematically analyzed through this project. By utilizing multiple linear regression, this project is dedicated to exploring the interrelationships between sleep, ecological and constitutional variables for mammalian species statistically.

2. Data Preprocessing and analysis

Data Preprocessing

- Records which have missing values are removed for the benefit of accurate fitting. Among all 62 records, 20 records are removed.

Variable Correlation Analysis

Through the correlation analysis, several relationships can be found:

- SWS ('slowWaveSleep') and DS ('dreamSleep') are negatively correlated with all other variables.
- All constitutional variables(B for 'body'; Br for 'brain'; L for 'life') are positively correlated with one another, so are ecological variables(G for 'gestation'; P for 'predation'; SE for 'sleepExposure' D for 'danger').

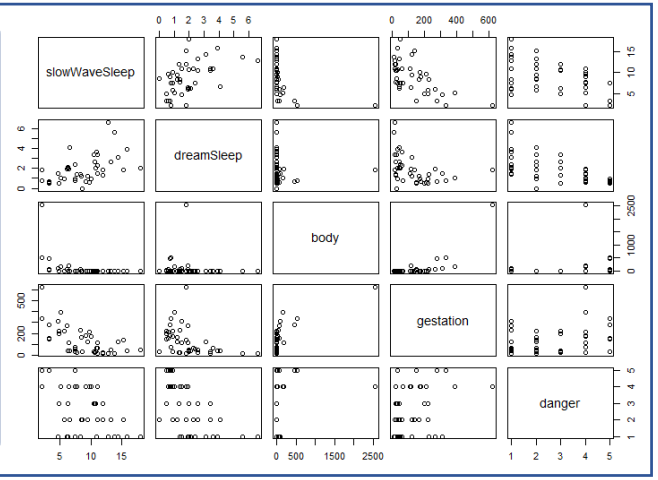

Fig.1 Scatterplots of selected predictors

3. Model Design

Methodology

- multiple linear regression

Variable Selection & Model Design

To determine the variables that predict the sleep variability, a lot of attempts have been carried out (e.g., interaction term, non-linear transformation of predictors) to find better models, in which SWS and DS are predicted separately using constitutional and ecological variables. Hence, two candidate multiple regression equations are selected for SWS and DS, respectively.

- For SWS:

I. $SWS = 14.37 - (0.53 \pm 0.19)logB - (0.615 \pm 0.55)logG - (0.93 \pm 0.32)D$

II. $SWS = 11.9 - (0.68 \pm 0.14)logB - (0.97 \pm 0.32)D$

- For DS:

III. $DS = 5.58 - (0.44 \pm 0.12)D - (0.57 \pm 0.16)D$

IV. $DS = 9.65 - (2.04 \pm 0.48)D - (1.49 \pm 0.3)logG + (0.34 \pm 0.1)D{:}\,logG$

4. Prediction on Mammals Sleep Dataset

Leave One Out Cross Validation (LOOCV) is used in this part to choose models.

Through testing, the MSE for Eq. I, II, III & IV are 7.71, 7.55, 1.11, 0.89. Hence, we choose Eq. II for SWS and Eq. IV for DS.

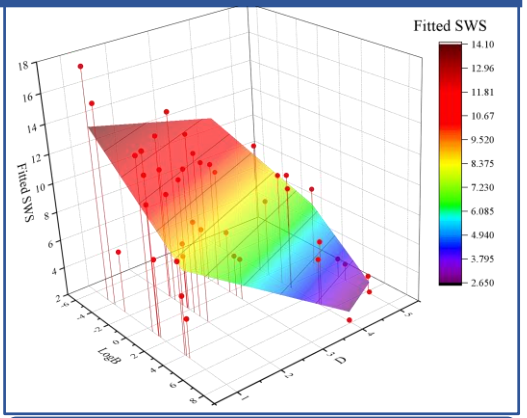

Fig.1 Fitted SWS regression plane and real SWS

5. Model Uncertainty Quantification

Bootstrap analyses are conducted to estimate the accuracy of our chosen linear regression models.

Thee standard errors of 1,000 bootstrap estimates for the intercept and slope terms are computed for both linear regression models:

- For SWS: $SWS = \beta_0 + \beta_1\, logB + \beta_2 D$

The bootstrap estimate for $SE(\hat{\beta}_0)$ is 1.02, for $SE(\hat{\beta}_1)$ is 0.13 and for $SE(\hat{\beta}_2)$ is 0.31, which are smaller than results of standard formula.

- For DS: $DS = \beta_0 + \beta_1 D + \beta_2 logG + \beta_3 D{:}\,logG$

The bootstrap estimate for $SE(\hat{\beta}_0)$ is 1.87, for $SE(\hat{\beta}_1)$ is 0.57, for $SE(\hat{\beta}_2)$ is 0.38, for $SE(\hat{\beta}_0)$ is 0.11, which are larger than results of standard formula.

6. Analysis & Conclusion

By analyzing the correlations and multiple linear regression for SWS and DS, several conclusions can be drawn from this project:

- Predators often need a great deal of sleep as they face less danger, while species subject to heavy predation often sleep less.
- Slow-wave sleep is negatively correlated with a factor related to body weight and size, which indicates that large species can hardly have long deep sleeping.
- Dream sleep is associated with a factor related to predatory danger, which suggests that dream sleep are disadvantageous in prey species.
- Both constitutional and ecological influences are important predictors for the amount and type of sleep which a mammalian specie need.

7. References

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). New York: springer.

8. Contact Information

Haojie WANG (20371572)
PhD Student
Department of Civil and Environmental Engineering, HKUST
Tel: (852) 5422-9876
Email: hwangbw@ust.hk or h.wang@connect.ust.hk