Data Analysis in Evol/Evol - HW Week6

Jing-Yi Lu

Reference R code see: github.com/jingyilu/Data-analysis-ecoevo

1. Confidence limits on proportions are asymmetrical. Can you explain this?

Proportions follows binomial distribution, which is a discrete distribution. Thus, we add a continuity correction when inferring confidence limits by the approximation of normal distribution to reflect the nature of asymmetry of binomial distribution.

2. P.493, q.18

a. Large sample size could decrease the variance within groups (error mean squares) and increase the power. Large sample size also leads to robustness to deviation from normality because of central limit theorem (large sampling will approach normal distribution).
b. Balanced design can lead to the robustness to departures from equal variance in each population.

3. P.494, q.22 (use R)

a. 0.00017
b. 0.00025
c. Repeatability = 0.60.
d. The femur length has higher repeatability (0.75), while the head width is more affected by measurement error.

4. P.499, q.32

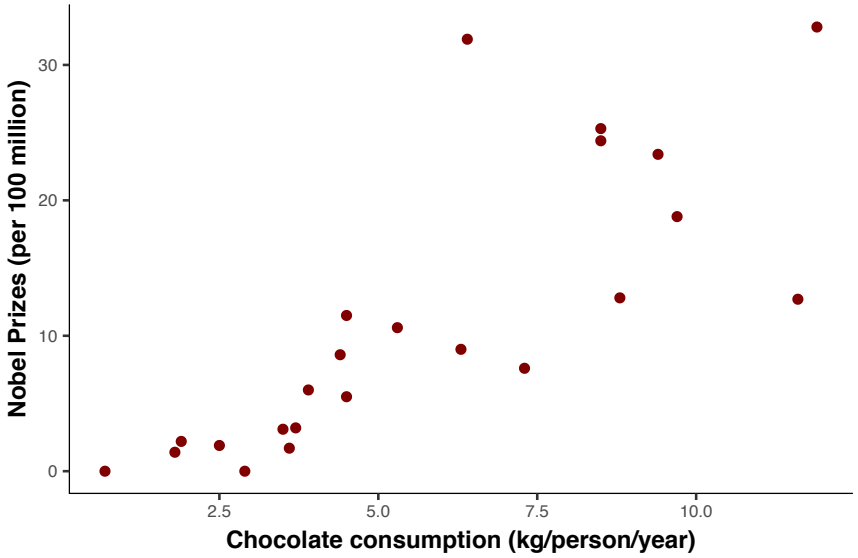
a. The boxplot depicts the patterns better. Firstly, bar graph has large unmeaningful areas and blurred the comparison of relative amounts of ratio. Secondly, the boxplot represents the distribution of the original data better with the quartiles and outliers.
b. Kruskal-Wallis (nonparametric) test. The boxplot indicates that both "No CFS" and "No seizures" groups are strongly deviated from normal distribution. Also, the variance of each group does not look similar.

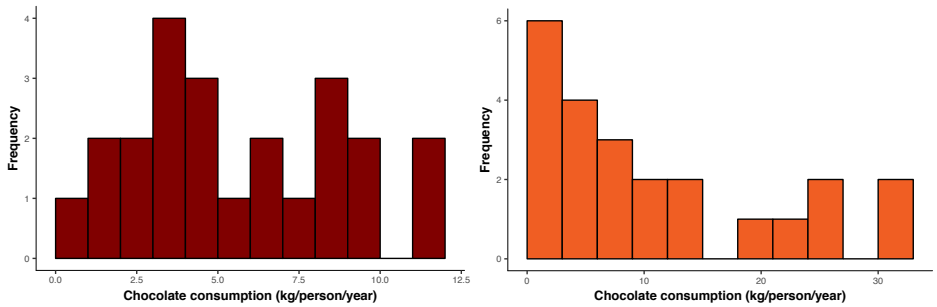
5. P.526, q.4

a. $r \sim 0$ **b.** $r \sim 0.7$ **c.** $r \sim 0.5$ **d.** $r \sim 0$

6. P.533, q.26

a. The measurements do not meet with the assumption of **bivariate normal distribution**.





b. H_0: The population correlation coefficient $\rho = 0$.
H_1: The population correlation coefficient $\rho \neq 0$.
$r = 0.792$, $t = 5.938$, d.f. = 21, *P*-value = $6.808 \times 10^{-6} < 0.05$
We can reject the null hypothesis that the population correlation coefficient $\rho = 0$.

c. The result suggests a positive correlation between chocolate consumption and number of Nobel Prizes among different countries. However, this correlation does not lead to direct causality. We cannot say that the chocolate consumption will increase the probability of winning a Nobel Prize. To be considered as a national priority (if the goal of the country is to win more Nobel Prizes), more research should be done to disentangle the potential confounding factors.

Using the library APE, your phylogeny, and the two columns of trait data you have made, compute phylogenetic independent contrasts using the PIC function. Positivize the first column (X contrast values), changing the sign of the Y value when the original X is negative. Plot the Y versus X contrasts.

Phylogeny: 17 (out of 23 in total) species of *Hemiboea* and 2 species of close-related *Briggsia*.
Morphology data: Length of involucre (bracts) and length of corolla (petals).