Data Analysis in Evol/Evol - HW Week8

Jing-Yi Lu

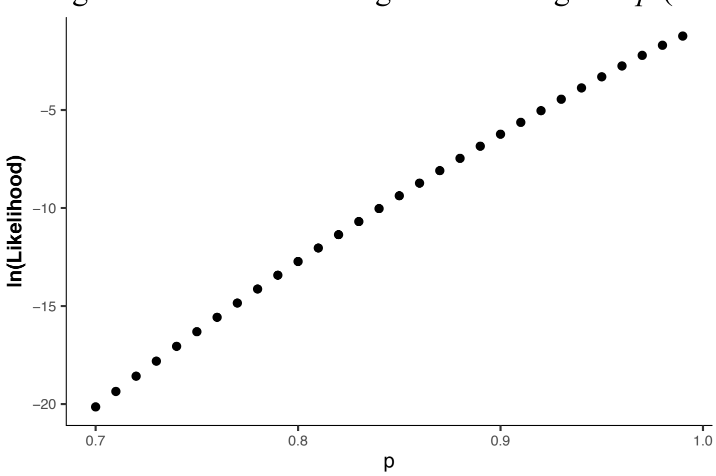
Reference R code see: github.com/jingyilu/Data-analysis-ecoevo

1. Send code you used to calculate the phylogeny statistics in previous homework

```
# Packages
library(tidyverse); library(ape); library(ggplot2)
# Read phylogenetic tree file into R
Hemiboea <- read.nexus("Hemiboea4.0.root.tre", tree.names = "Hemiboea")
# Read morphological data into R
Hemiboea_morph <- read.csv("Hemiboea_morph.csv", header = T)
# Force biforcation
Hemiboea.bi <- multi2di(Hemiboea)
# Create vectors for calculating phylogenetic independent contrasts
Corolla.length <- Hemiboea_morph$Corolla_length; Involucre.length <- Hemiboea_morph$Involucre_length
names(Corolla.length) <- names(Involucre.length) <- Hemiboea_morph$Species #Give names for matching tips
# Calculate Phylogenetic Independent Contrasts (PIC)
pic.Corolla.length <- pic(Corolla.length, Hemiboea.bi); pic.Involucre.length <- pic (Involucre.length, Hemiboea.bi)
# Positivise PIC values (using function "case_when" as logical operators)
Contrast <- as.data.frame(cbind(pic.Corolla.length, pic.Involucre.length))
Contrast.positivise <- Contrast %>%
mutate(pic.Involucre.length = case_when(pic.Corolla.length<0 ~ -pic.Involucre.length,
                                        pic.Corolla.length>=0 ~ pic.Involucre.length),
       pic.Corolla.length = case_when(pic.Corolla.length<0 ~ -pic.Corolla.length,
                                        pic.Corolla.length>=0 ~ pic.Corolla.length))
# Plot Contrast (detailed adjustment for graph making neglected here.)
Hemiboea.contrast.cor.inv <- Contrast.p %>%
  ggplot(aes(x = pic.Corolla.length, y = pic.Involucre.length)) +
  geom_point(color = "#800000")
```

2. P.677 q.15

a. The log-likelihood curve of log-likelihood against p (0.7-0.99):



b. The maximum likelihood estimate $\hat{p}$: 0.99.

c. From the log-likelihood-p curve, the difference between maximum(ln(likelihood)-ln(likelihood) exceeds 1.92 when p = 0.95 (the difference = 2.08).

Therefore, the 95% confidence interval for p is (0.95, 1).

3. Using the formula for a normal distribution, write down the likelihood of the mean and ln(Likelihood) of the mean, assuming the variance, $\sigma^2 = 1$. Draw a ln(likelihood) surface for μ = -3 through μ = +3 and indicate the 2 units of support.

Probability density function of a normal distribution: $\frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

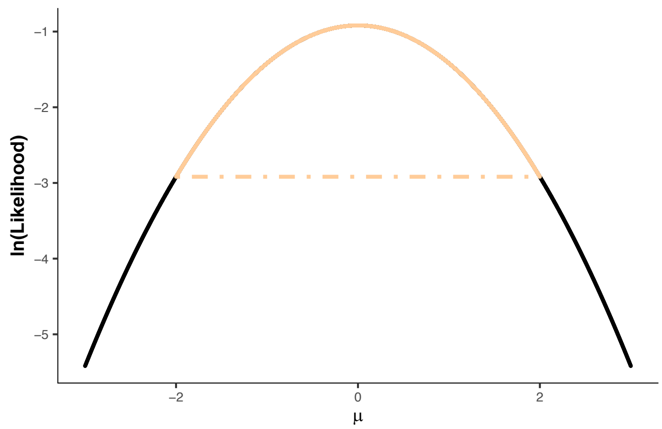
as $\frac{1}{\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2}\right)$ when $\sigma^2 = 1$

Likelihood($\{x\}$) = Pr($\mu|\{x\}$) = $(2\pi)^{-n/2} exp\left[\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right]$

$$\ln(\text{Likelihood}(\{x\})) = -\frac{n}{2}ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2 = -\frac{n}{2}ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2 - \frac{n}{2}(\bar{x}-\mu)^2$$

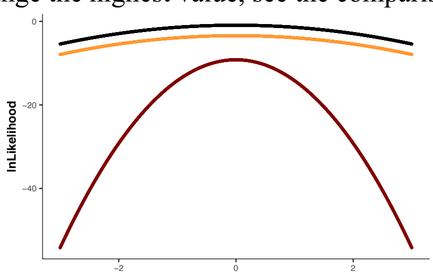
$$= -\frac{n}{2}ln(2\pi) - \frac{n}{2}s^2 - \frac{n}{2}(\bar{x}-\mu)^2$$

As there are no observed set of $\{x\}$ given, I draw the ln(likelihood) surface for a distribution with $x = 0$, $s^2 = 0$, n = 1:



The 2 units of support is indicated by the light orange dot-dashed line.

*Note: from the formula above, the s^2 will not affect the relative lnL, while n will affect quantitatively but not change the highest value, see the comparison plot:



The black curve is the same from the above figure. Orange curve with higher variation = 5; red curve with sample number = 5.

4. Using the function rasterToPoints plot annual precipitation against elevation.