Lecture slides with $\LaTeX$

A user friendly template

Pedro Fonseca

ISEG Lisbon School of Economics & Management

April 24, 2020

Contents

Before we start

In this lecture you will see:

- a graph
- some formulas
- some code
- some citations (Kass and Raftery, 1995)

A scatterplot

This picture is in the "figures" folder:

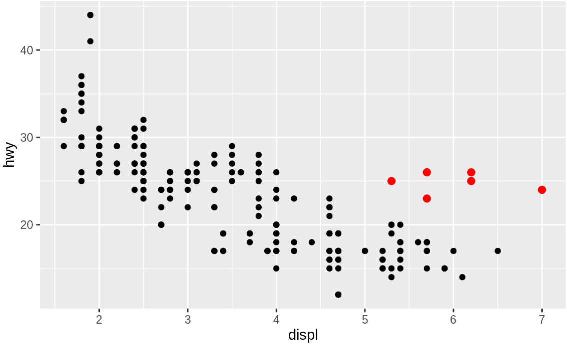

Figure 1: Fuel consumption (Hwy) vs engine size (displ)

Model selection

- Likelihood: $\boldsymbol{x}|\boldsymbol{\theta} \sim \mathcal{M}_k(N, \boldsymbol{\theta})$
- Hypotheses: $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ vs $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$
- Prior: $\pi(\boldsymbol{\theta}|H_0) = 1_{\boldsymbol{\theta}_0}(\boldsymbol{\theta}_0)$ and $\boldsymbol{\theta}|H_1, \boldsymbol{\alpha} \sim \mathrm{Dir}_k(\boldsymbol{\alpha})$
- Marginal density: $m_i(\boldsymbol{x}) = \int_{\Theta_i} f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|H_i)\, d\boldsymbol{\theta}, \quad i = 0, 1$
  - $m_0(\boldsymbol{x}) = \frac{N!}{\prod_{i=1}^{k+1} x_i!} \prod_{i=1}^{k+1} {\theta_{0i}}^{x_i}$
  - $m_1(\boldsymbol{x}) = \frac{N!}{\prod_{i=1}^{k+1} x_i!} \frac{B(\boldsymbol{\alpha}+\boldsymbol{x})}{B(\boldsymbol{\alpha})}$
- Bayes factor:
  $B_{01}(\boldsymbol{x}) = \frac{m_0(\boldsymbol{x})}{m_1(\boldsymbol{x})} = \frac{\prod_{i=1}^{k+1} ({\theta_{0i}}^{x_i}) \prod_{i=1}^{k+1}[\Gamma(\alpha_i)]\Gamma[\sum_{i=1}^{k+1}(\alpha_i+x_i)]}{\Gamma(\sum_{i=1}^{k+1} \alpha_i) \prod_{i=1}^{k+1} \Gamma(\alpha_i+x_i)}$
- For an interesting application see Pericchi and Torres (2011).

Parameter estimation

- Prior: $\boldsymbol{\theta}|\boldsymbol{\alpha} \sim \mathrm{Dir}_k(\boldsymbol{\alpha})$
- Posterior: $\boldsymbol{\theta}|\boldsymbol{\alpha}, \boldsymbol{x} \sim \mathrm{Dir}_k(\boldsymbol{\alpha} + \boldsymbol{x}) \rightarrow$ Mean, mode, credible intervals.
- For an interesting application see Ley (1996).

Model selection

- Likelihood: $x \sim \text{Bin}(N, \theta)$
- Hypotheses: $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$
- Prior: $\pi(\theta|H_0) = 1_{\theta_0}(\theta)$ and $\theta|H_1 \sim \text{Beta(a,b)}$
- Marginal density: $m_i(x) = \int_0^1 f(x|\theta)\pi(\theta|H_i)d\theta,\ i = 0, 1$
  - $m_0(x) = \binom{N}{x}{\theta_0}^x(1-\theta_0)^{N-x}$
  - $m_1(x) = \binom{N}{x}\frac{B(x+a,n-x+b)}{B(a,b)}$
- Bayes factor: $B_{01}(x) = \frac{m_0(x)}{m_1(x)} = \frac{\theta_0(1-\theta_0)^{N-x}\,\Gamma(a)\,\Gamma(b)\,\Gamma(n+a+b)}{\Gamma(a+b)\,\Gamma(n+a-x)\,\Gamma(x+a)}$

Parameter estimation

- Prior: $\theta \sim \text{Beta}(a, b)$
- Posterior: $\theta | x \sim \text{Beta}(a + x, b + N - x) \rightarrow$ Mean, mode, credible intervals

Some python code

This code prints the Fibonacci sequence:

```
nterms = int(input("How many terms? "))
n1, n2 = 0, 1
count = 0
print("Fibonacci sequence:")
while count < nterms:
    print(n1)
    nth = n1 + n2
    n1 = n2
    n2 = nth
    count += 1
}
```

Now some R code

I also included a a verbatim environment with background color:

```
getmode <- function(v) {

  uniqv <- unique(v)

  uniqv[which.max(tabulate(match(v, uniqv)))]

}
```

With this R code you can build a function that calculates the mode.

References

Kass, RE and AE Raftery (1995). Bayes factors. *Journal of the American Statistical Association* **90**(430), 773–795.

Ley, E (1996). On the peculiar distribution of the US stock indexes' digits. *The American Statistician* **50**(4), 311–313.

Pericchi, L and D Torres (2011). Quick anomaly detection by the Newcomb—Benford law, with applications to electoral processes data from the USA, Puerto Rico and Venezuela. *Statistical Science* **26**(4), 502–516.