

EPISOUTH-MADRID 16 June 2009: PRACTICE 3

1. First, we are going to create two functions that are not implemented in R: kurtosis and skewness

```
source("fskewness.R")  
source("fkurtosis.R")
```

2. Now, obtained the following summary statistics for variables in the database

| Estimate/variable | Sex | Literacy | Members | Explica | Age | Case status |
|--------------------|-----|----------|---------|---------|-----|-------------|
| Mean | | | | | | |
| Median | | | | | | |
| Variance | | | | | | |
| Standard deviation | | | | | | |
| Kurtosis | | | | | | |
| Skewness | | | | | | |

Syntax example to age:

```
mean(dat$age)  
median(dat$age)  
var(dat$age)  
sd(dat$age)  
kurtosis(dat$age)  
skewness(dat$age)
```

3. Focusing on age, do you think that the parameters fit with a normal distribution?
4. What transformation can help to make the variable more Gaussian?
5. Look at the values of the variable "family members", what distribution can be appropriate to describe this variable?