

## PREPARING THE DATA

### Define your RNG

Select a RNG to use in your analysis.

1. Implement this RNG.
2. Test the correctness of your RNG using at least two tests.

Remember that you can define your own functions in R like this:

```
My_RNG <- function(x, seed=7){  
  #Do something  
  return(x)  
}
```

You can find a source of RNG here: [http://wiki.fib.upc.es/sim/index.php/Main\\_Page/en](http://wiki.fib.upc.es/sim/index.php/Main_Page/en)

*Clue: are 3 or more sub-streams following the same distribution?*

### Simulate your data.

The structure of the table must be similar to the presented next.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	..	Factor 10
Individual 1							
Individual 2							
Individual m							

1. Define, for each factor (from 1 to 5) a distribution. For the factors 6 to 10 define a function that uses the previous variables, as an example  $F6 = F1 + 2F3$ .
2. Define an answer variable that will be composed by a function that combines a subset of the previous factors plus a normal distribution (with some parameters you know).

### Obtain an expression to generate new data.

Imagine that you don't know nothing regarding this dataset. You need to explore it because you want to define an expression to obtain new data for your DOE (you want to detect the possible relations and the interactions between the factors).

1. Explore the possible relations of all the factors and the answer variable with the PCA.
2. Describe what you find on this analysis and if it is coherent with the knowledge you have from the data.
3. Propose an expression to generate new data (new values using a subset of the factors).

## DOE

Now you have an expression to generate new data (a simple method to obtain new values depending on the factors).

1. Define a DOE to explore the relations between the different factors you have.
2. Detect and analyze the interactions.

*Remember*

- *Set the objectives.*
- *Select the process variables.*
- *Define an experimental design.*
- *Execute the design.*
- *Check that the data are consistent with the experimental assumptions.*
- *Analyze and interpret the results, detect effects of main factors and interactions.*