

Stat 471: Lecture 1

MATLAB/OCTAVE

Moo K. Chung
mchung@stat.wisc.edu

September 3, 2003

1. Statistics is concerned with the science of uncertainty. Computational statistics is a collection of techniques that have a strong focus on the exploitation of computing in the creation of new statistical methodology (Wegman, 1988).
2. Cheap computation enabled us to store and process massive data so that we often collect the data first and then design study to gain information. The data sets these days is very large and high-dimensional.
3. Data types: scalar, vector, matrix, tensor.
Observed value = true value + measurement error.

$$Y_{ijk} = \mu + e_{ijk},$$

where ijk is in certain index set. For instance, if $Y_i, i \in R$, Y_i is a continuous observation (function) such as temperature. In the case of high dimensional index set, $Y_i \in R^n$, Y_i becomes n -dimensional data. Tensorial is a generalization of matrix data. Think a tensor as a high dimensional array with complicated relationship among elements. For data with the large size of index set, computation-intensive statistical methods (Efron and Tibshirani, 1991) are more practical than the traditional design oriented data analysis.

4. Computer languages: MATLAB or Octave will be used as the language of instruction. GNU Octave is a free version of MATLAB (www.octave.org) developed by the department of chemical engineering, WISC. Splus or R are programming language mainly used by statisticians. It shares similar language

structure with LISP. If you don't have access to any computer or programming language, create one in the statistics computer lab located in the first floor. Read first <http://www.cs.wisc.edu/csl/doc/faq/started/index.html>.

5. Type either octave or *matlab* from the *linux* prompt to run the programs. For octave, you may possibly use text editor such as Emacs to save your work.
6. Uniform random numbers $U \sim Unif(0, 1)$. Using U , generate 1000 uniform random numbers $X \sim Unif(-1, 5)$.

```
>> u = rand(1000,1);
>> size(u)
ans =
      1000         1
>> x = 4*u-1;
>> rand
ans =
      0.4048
>> rand
ans =
      0.6271
>> s=rand('state')
s =
      0.8598
      0.1398
      .....
>> rand('state',s)
>> rand
ans =
      0.4048
```

Use help command to see the grammar. ex.
`help rand.`