Animation in Statistics: Dynamic Graphics for Statistical Models and Practical Applications

Yihui XIE



School of Statistics, Renmin University of China

Dec 12, 2007

Statistics

Contents

Why Employ Animations?

- Interesting
- Intuitional
- 2 How to Make Animations?
 - R Graphical Devices
 - HTML & JavaScript



- Iterative Algorithms
- Random Experiments
- What's More?
 - Statistical Theories
 - Practical Applications

Interest is the Best Teacher

Nobody likes dull stuff...

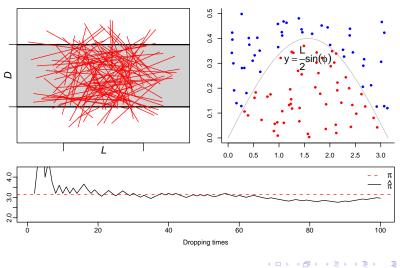
Buffon's Needle

- Buffon's Needle is one of the oldest problems in the field of geometrical probability; It was first stated in 1777.
- It involves dropping a needle on a lined sheet of paper and determining the probability of the needle crossing one of the lines on the page.
- The remarkable result is that the probability is directly related to the value of π .
- From frequency to probability.

Why Employ Animations?

Interesting

Interest is the Best Teacher (cont'd)



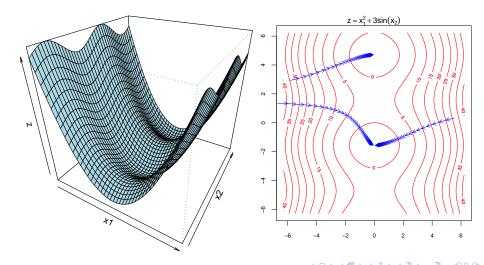
Simulation of Buffon's Needle

School of Statistics

Dec 12, 2007 4 / 16

Two Pictures are Worth 2000 Words

You don't like those abstract theories, do you?



Produce Animations inside R

- The R package grDevices has offered a variety of graphics devices.
- We may just create animations in the Windows graphics devices (Windows) or X Window System graphics devices (Linux) or MacOS X Quartz devices (MacOS X).
- The most critical elements are loops and Sys.sleep().
- It's convenient to produce single image files there are several choices such as PNG, JPEG, BMP, PDF, PS, TEX/LATEX and WMF, etc.

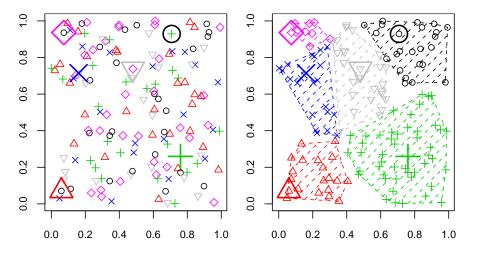
A B K A B K

Produce Animations outside R

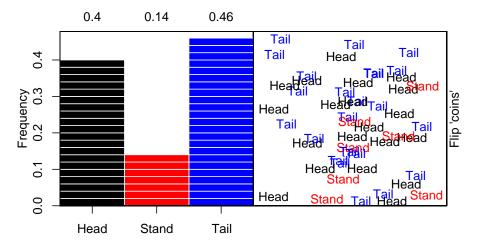
- It's convenient to produce single image files using R there are several choices such as PNG, JPEG, BMP, PDF, PS, TEX/LATEX and WMF, etc.
- We may use JavaScript to animate these image frames quite naive way.
- $\bullet \mid / _ \setminus \mid$
- . . ₀ o O
- Many examples in this talk have been implemented in the R package 'animation'; you may download it from CRAN.
- The package 'animation' has offered both ways of creating animations.
- See buffon.needle(), flip.coin(), kmeans.ani(), boot.iid(), knn.ani(), ...

イロト 不得下 イヨト イヨト 二日

K-Means Cluster Algorithm

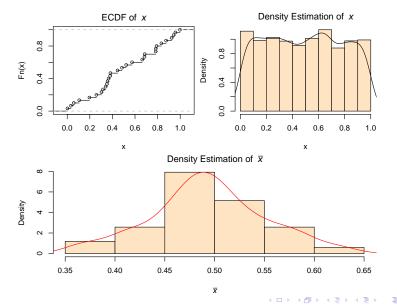


Flipping Coins



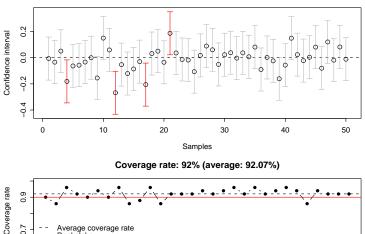
Statistical Theories

Probability Theory: Central Limit Theorem



Yihui XIE

Mathematical Statistics: Confidence Interval



CI: $[\overline{x} - z_{\alpha/2}\sigma/\sqrt{n}, \overline{x} + z_{\alpha/2}\sigma/\sqrt{n}]$

School of Statistics

0.7

Average coverage rate Real alpha

10

5

15

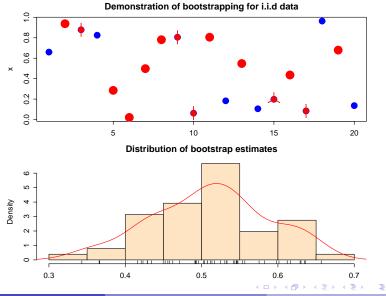
20

Dec 12, 2007 11 / 16

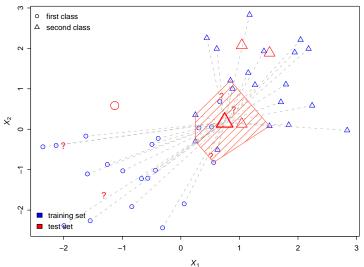
30

25

Machine Learning: Bootstrapping

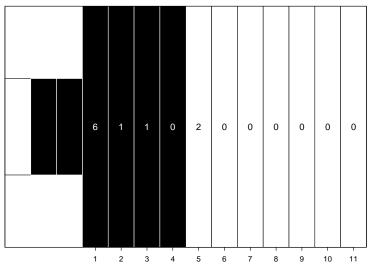


Machine Learning: k-Nearest Neighbor Algorithm



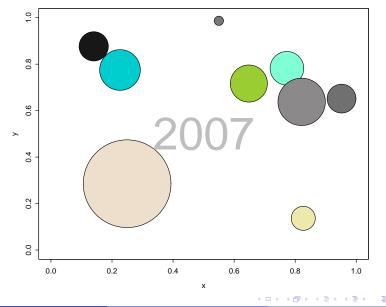
Demonstration for kNN Classification

Simulation of A Chemical Experiment



73

Social-Economic Changes Over Time



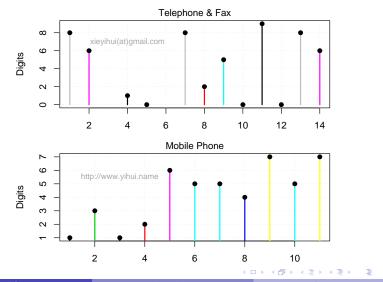
School of Statistics

Dec 12, 2007 15 / 16

Contact Me

Thanks!

A weird name card...



School of Statistics

Dec 12, 2007 16 / 16