Lecture 9 Cogsci 109

Wed. Oct. 17, 2007
How to create colormaps, meshgrid
application, and computing basic
statistics I (central tendency)

Outline for today

- Announcements
- Matlab demonstrations - loading data, then…
  - ☐ **Creating and applying color maps**
  - ☐ **Creating *pcolor* and *surf* plots**
    - How to input data
    - How to make the axes scale appropriately - *meshgrid* command

Outline for today (II)

- Super-/sub-sampling
  - □ **Definitions**
  - □ **Examples**
  - □ **When you use one vs. another**
- Computing basic statistics
  - □ **Looking at data quantitatively**
  - □ **Central tendency - Mean, median, mode**
    - Definitions
    - distributions
    - Examples from last year's surveys

Announcements

- Hw2
  - grade distributions
  - In future - turn in all info at once, or if not, the first version is a placeholder - and must have a clear page stating such, second is the one that will be graded, so it must be complete

- Hw3
  - questions?
  - Formatting details - grouping and organization

Announcements II

- Midterm is coming up somewhere in the next couple of weeks
  - Anything lectured on, presented in section, in the reading and on homework is fair game
  - Will be part multiple choice, part short answer
  - Any calculations will be simple
  - Programming questions will be fairly qualitative
  - A practice midterm will be posted
    - You should study this carefully

Matlab demonstrations

- Loading data, then
  - Creating and applying color maps
  - Creating pcolor and surf plots
    - Methods of Inputting data
    - *meshgrid* command
      - Scaling axes appropriately

Super-/sub-sampling

- ## Super-sampling (up-sampling)

$$y(n) = \begin{cases} x\left(\dfrac{n}{N}\right), n = 0, \pm N, \pm 2N \\ 0 \quad, \quad otherwise \end{cases}$$
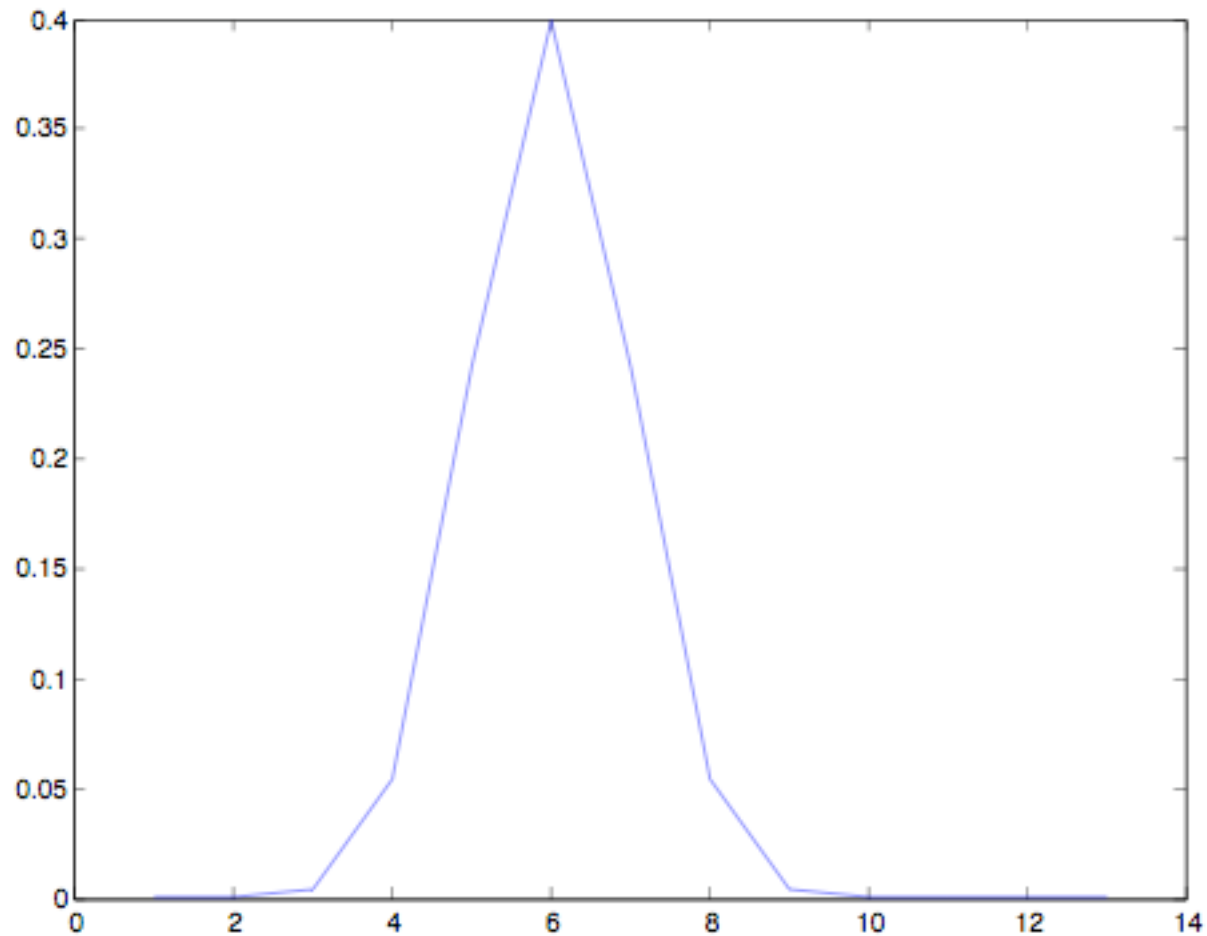
- ## Sub-sampling (down-sampling)
  - **Take every M-th sample**
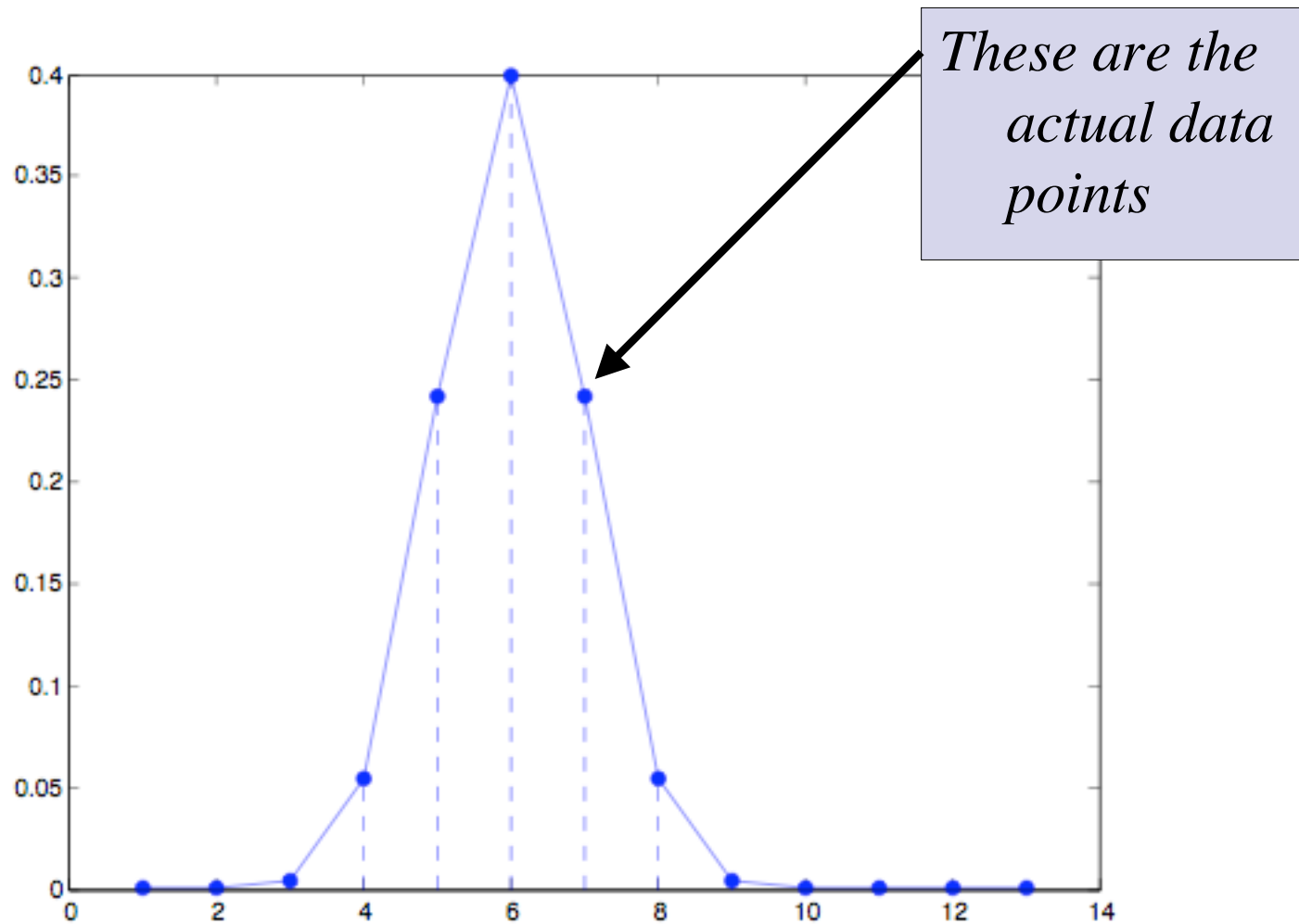
$$y(n) = x(Mn)$$

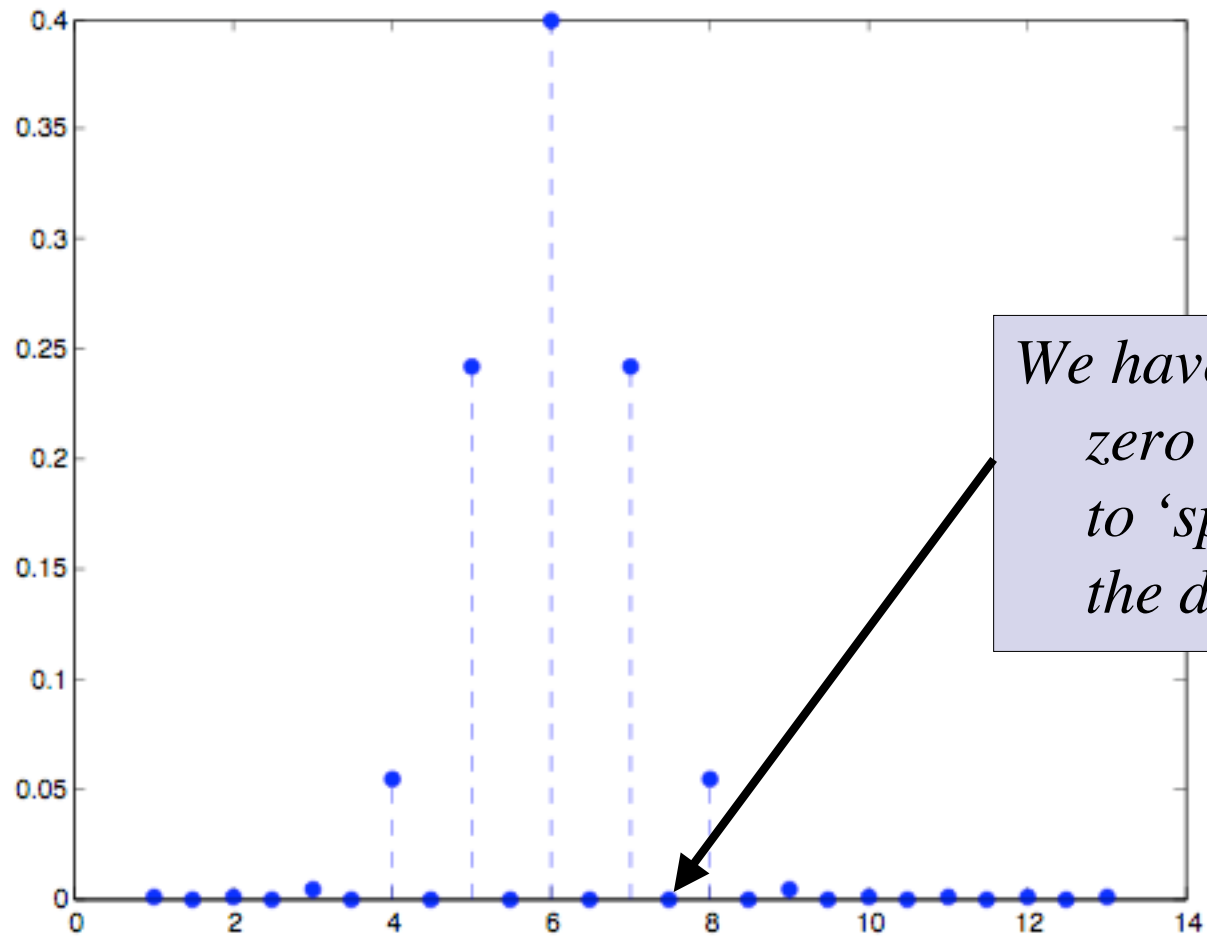 - **1000Hz sample rate becomes 100Hz sample rate if we down sample by a factor of 10**

What does super-sampling look like?

What does super-sampling look like?

These are the actual data points

Example: Up-sampling

We have added zero data points to 'space out the data'

Why add only zeros?

- If you want temporal matching of events, but you had different sample rates for different recording instruments
  - **i.e. - data matching...**
    - 200Hz movement sensors (Polhemus Liberty 6dof sensors)
    - 50Hz simulation VR update
    - 1kHz EEG recording
  - **How would we compare all this data in an event-related setting?**
    - One way - Up-sample or down-sample to make sample rates match
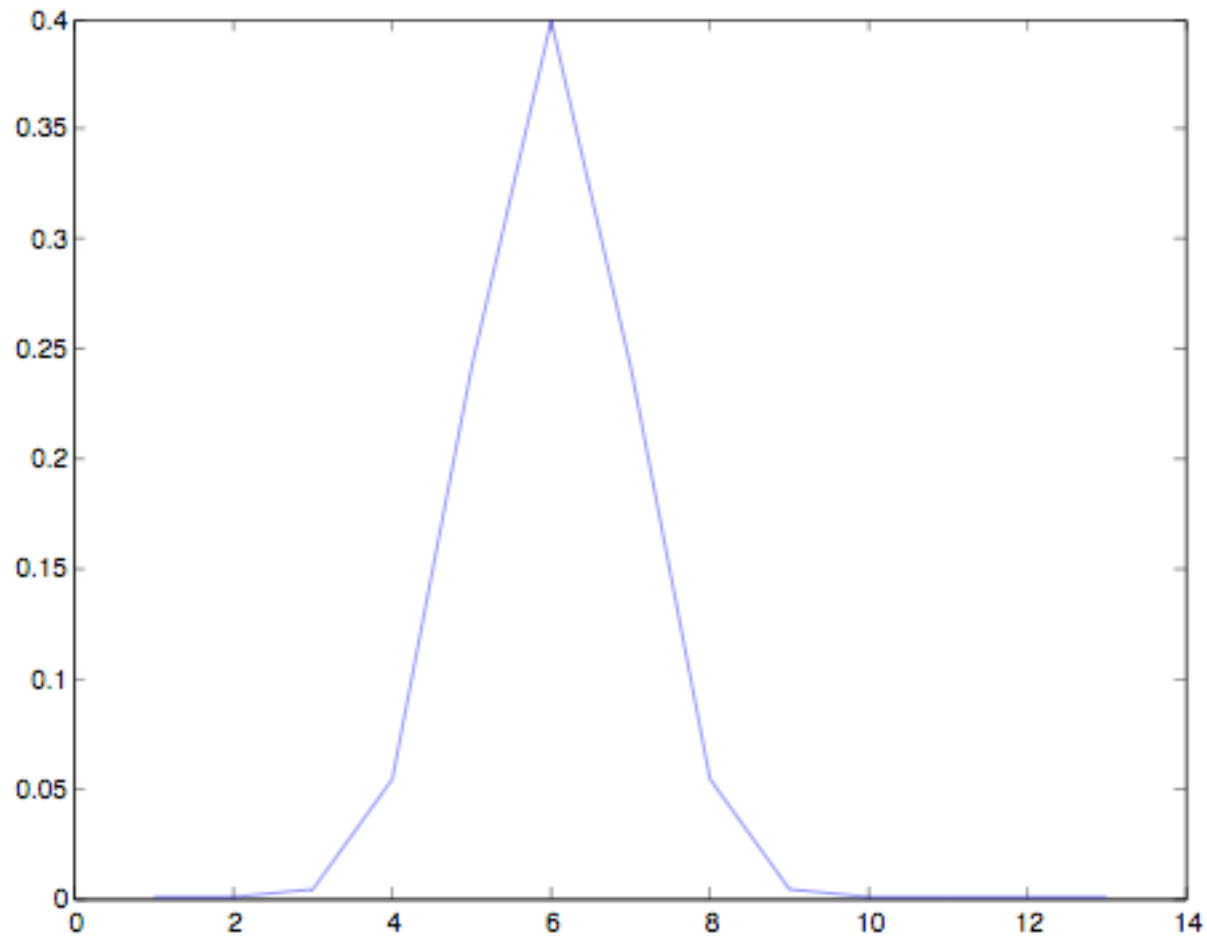   - Another way - interpolate (we'll cover this in a later lecture)

Are there other ways to do this?

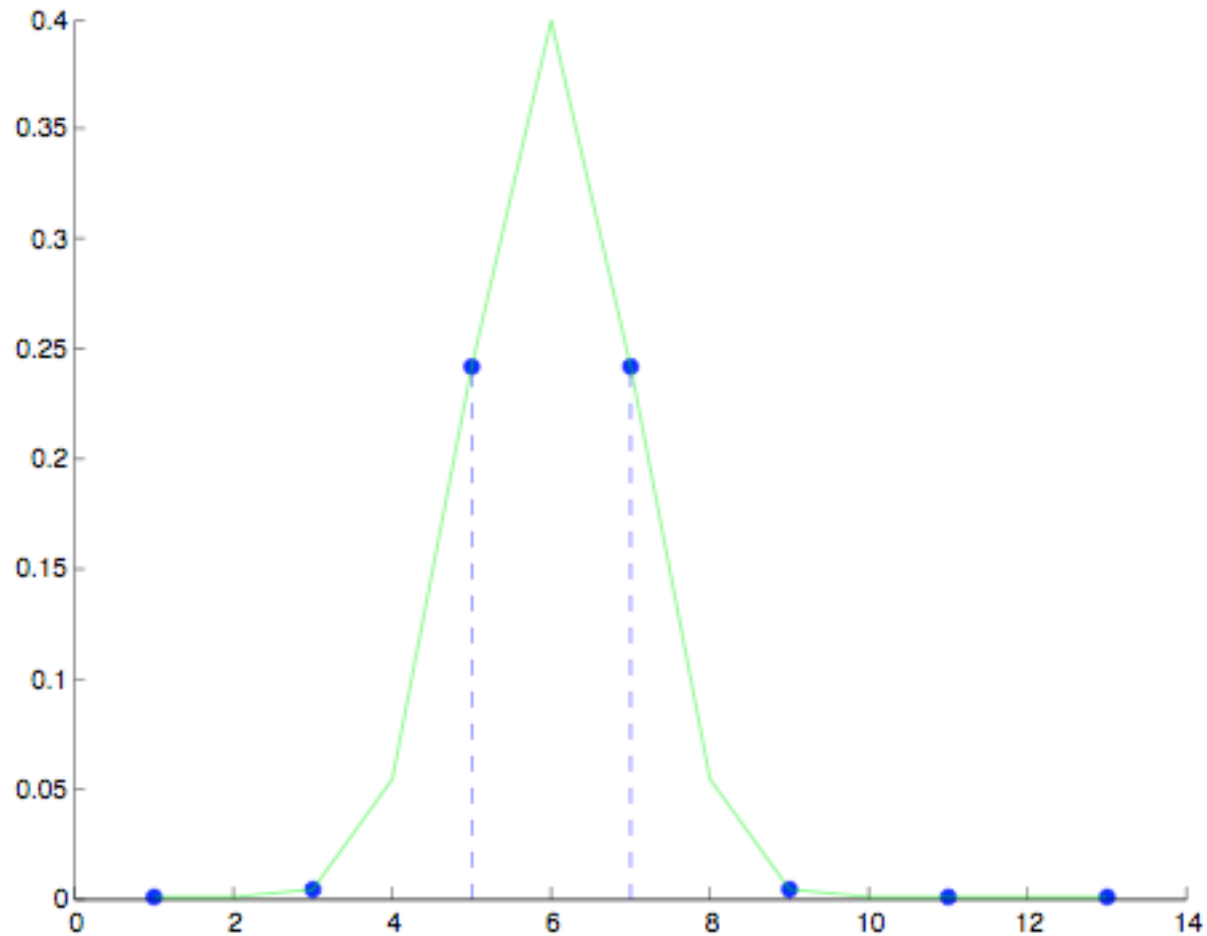- We could also interpolate the points instead of adding zeros…but that involves data fitting which we'll discuss soon
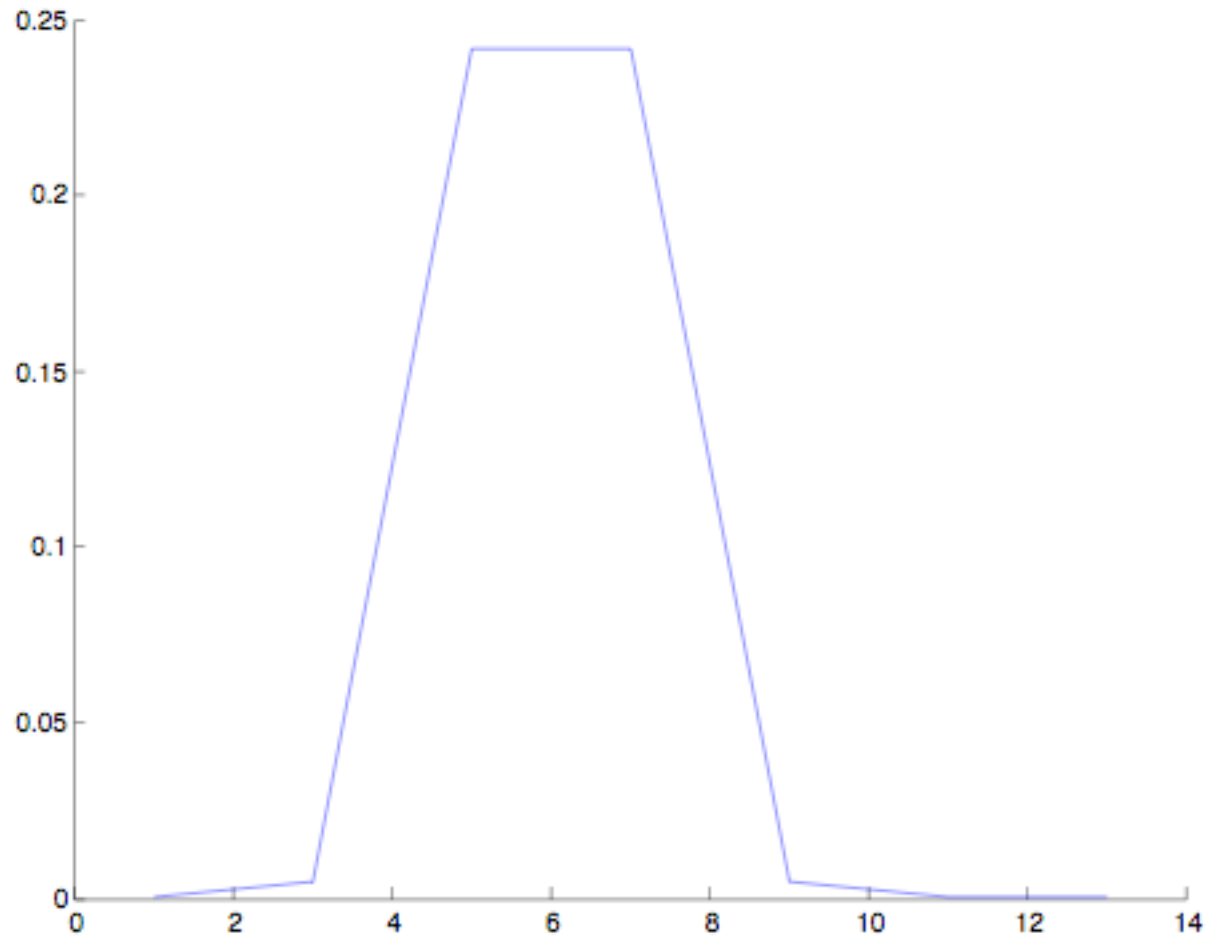
What does sub-sampling look like?

What does sub-sampling look like?

Then the sub-sampled data looks like

Quantitatively: Another way to 'look' at data

- How do we look at data quantitatively and extract meaningful information?

- Computing basic statistics is a start

  - Mode
  - Mean
  - Median
  - Standard Deviation
  - Variance
  - Covariance
  - Correlation

Central Tendency - *Mode*

- Most common number of a distribution
- Tells you which value has the highest frequency
- What if there are ties?
  - **More than one mode!**
  - **Which of the following is the mode?**

$$\{1,2,2,2,2,2,2,3,4,5,6,7,8,8,8,9,9\}$$

  - **2**
- Matlab help: *help mode*

Central Tendency - *Mean*

- Think of it as similar to a balance point
- 'Expected value'
- Computed by the following
  - Sum all scores
  - Divide that sum by the number of scores
- Here's the formula:

$$M = \left(\sum_{i=1}^{N} x_i\right)\Big/N$$

- And an example:

$$\{1.0, 1.0, 2.0, 3.0, 4.0, 4.0, 4.0, 4.0, 8.0, 8.0, 8.0, 8.0, 8.0, 8.0, 9.0, 0.0, 0.0, 0.0\}$$

$$N = 18$$

$$\sum x_i = 80.0$$

$$M = \left(\sum x_i\right)\Big/N = 80.0/18 = 4.4$$

How to compute the mean in Matlab

- Function : *mean()*
- *Side note about matlab help and capitals*
- Example computation…
- Matlab help : *help mean*

Median

- The middle number of a distribution when the numbers have been ordered (sorted)
- Each score is counted separately, so if you have repeating scores such as 50 and 50, each one becomes part of the count
- Order the scores from low to high or high to low
- Count from both ends to the middle position

Median

- If **odd** number of scores, there will be one median
  - Example: *Find the median*

$$\{1,2,3,10,50\}$$
$$Median = 3$$

- If an **even** number of scores, count to the two closest to the middle (ie count from low towards high, high towards low) and take their average (add them up and divide by two)
  - Example: *Find the median*

$$\{1,2,2,3,3,4\}$$
$$2,3$$
$$Median = (2+3)/2 = 2.5$$

How to compute the median in matlab

- Function *median()*
- Example
- Matlab help: *help median*

How are they related?

- If you have a…
  - **Normal distribution,**
    - Mean=Median=Mode
  - **Symmetric distribution**
    - Median = Mean
  - **Skew distribution**
    - Median towards the body, mean towards the tail
      - **+skew: mean>median**
      - **-skew: mean<median**
- But this doesn't seem to be saying everything…

The mean isn't everything!
These all have the same mean