## Empirical Model Building I: Objectives:

By the end of this class you should be able to:

- find the equation of the "best fit" line for a linear model
- explain the criteria for a "best fit" line
- linearize exponential and power models.
- use plots to determine if a linear, exponential or power model fits a given data set.

Palm, Section 5.5

Download file FnDiscovery.mat
and load into MATLAB

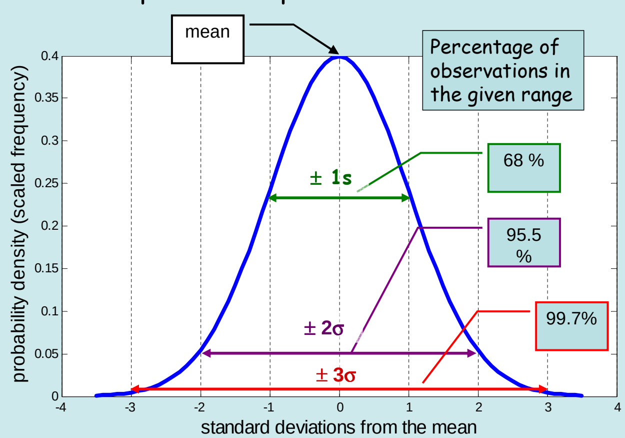
## Review Exercise (pairs)

The compressive strength of samples of a specific type of cement can be modeled by a normal distribution with a mean of 6000 kilograms per square centimeter and a standard deviation of 100 kilograms per square centimeter

- What is the variance of this compressive strength?
- What percent of a batch of cement samples is expected to be greater than 5800 kg/cm$^2$?

Adapted from: Montgomery, Runger and Hubele,
*Engineering Statistics*, 2nd Ed., Wiley (2001)

## Expected Proportions for known σ

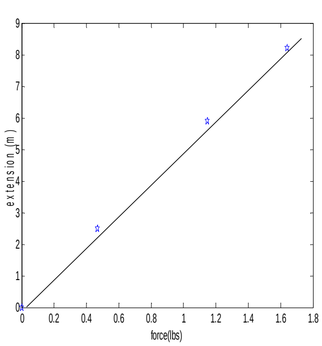



## Modeling Spring Lengthening

What type of model (equation) would you use to represent the data below? Why?

| **Force** (lb) | Spring **Length** Increase (in.) |
|---|---|
| 0 | 0 |
| 0.47 | 2.5 |
| 1.15 | 5.9 |
| 1.64 | 8.2 |

## Identifying an Empirical Model




- plot data (see points at left)
- looks linear
- equation: F = mL + b (y = mx + b)
- Here we are trying to determine an **empirical** model (an equation that matches the data shape but is not from theory)
- .
- There is also a **mechanistic** (theoretical) model for this system from Hook's law (F = k L)
- Note L = the change in length.
- Can you estimate the coefficients of this model?
- We will look at this more later
- m ~ 5, b ~ 0

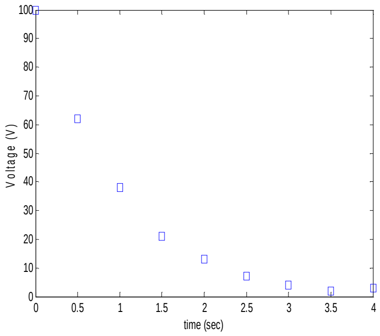
## Capacitor Discharge Data

Plot this data.

Can you fit a linear model to it?

| **t = Time (s)** | **V = Voltage (V)** |
|---|---|
| 0 | 100 |
| 0.5 | 62 |
| 1 | 38 |
| 1.5 | 21 |
| 2 | 13 |
| 2.5 | 7 |
| 3 | 4 |
| 3.5 | 2 |
| 4 | 3 |

## Capacitor Discharge Example




Commands to plot Graph

```
>> plot(t, V, 's')
>> xlabel('time(sec)')
>> ylabel('Voltage(V)')
```

Notice:

- Curve is smooth and "**monotonic**"
- but not linear
- a monotonic curve is one that either is constantly increasing or constantly decreasing but never reverses direction for positive x

- We need another model
- Options (see following slide)

## Three Basic Two-parameter Models
(colored equations down on a blank paper)

- Linear $$y = mx + b$$
- Power $$y = bx^m$$
- Exponential $$y = be^{mx}$$ or $$y = b10^{mx}$$

## Remember basic Log/ln rules

- Multiplication:
  log(ab) = log(a) + log(b)
  or ln(ab) = ln(a) + ln(b)

- Powers:
  $\log(a^m) = m \log(a)$
  or $\ln(a^m) = m \ln(a)$

For the power and exponential models you wrote down :
Take the log of both sides and simplify

## Transformation of Models
(have students try, then do on board)

| power law model | exponential |
|---|---|
| $y = bx^m$<br>$\log(y) = \log(b) + m\log(x)$<br><br>For: Y = log(y)<br>X = log(x)<br>B = log (b)<br>Eqn becomes: Y = B + mX<br><br>If the power law model applies a log-log plot should be linear | $y = b10^{mx}$<br>$\log(y) = \log(b) + m\ x$<br><br>For Y = log(y)<br>B = log (b)<br><br>Eqn becomes: Y = B + mx<br><br>If exponential model applies a semilogy plot should be linear |

## A few notes on MatLAB commands

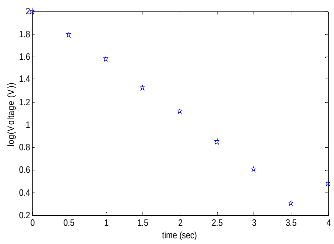
| Log transformations: | Also remember: |
|---|---|
| >> log(x)<br>➔ the natural log of x (i.e., ln(x))<br><br>>> log10(x)<br>➔ the base 10 log of x<br><br>This pattern is typical for many programs | >> semilogy(x,y, ...)<br>➔ creates a plot with a log scale on the y axis<br><br>>> loglog(x,y, ...)<br>➔ creates a plot with log scales on both axes |

## Try these transformations of capacitor discharge data (follow the leader exercise)

Exponential Model – if model is exponential I should get a straight line when plotting log(V) vs t

```
>> logV = log10(V)
>> plot(t, logV, 'p')
>> xlabel('time (sec)'); ylabel('log(Voltage (V))')
```

% notice the y-axis must be labeled as the log of voltage.




This looks to be the correct model, the data lines up nicely on a straight line

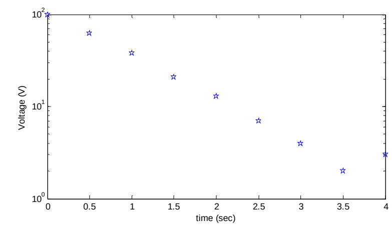
## Can also plot using the semilogy command

Commands (putting graph on a new figure)

```
>> figure
>> semilogy(t, V, 'p')
>> xlabel('time (sec)'); ylabel('Voltage (V)')
```

Notice we <u>do not</u> use the log in the y axis title this time. The y-axis is in the actual variable, the spacing on the axis is based on a the log of the value.

Notice the difference in the look of the y-axis. Otherwise this graph is identical to the earlier plot

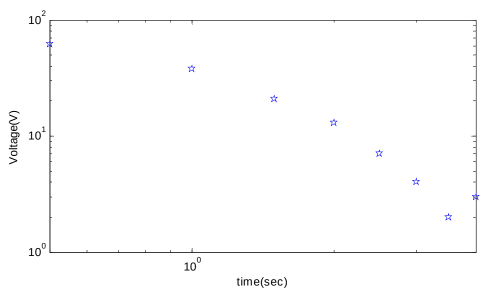



## for comparison look at this data on other log-log plot

Commands (putting graph on a new figure)

```
>> figure
>> loglog(t, V, 'p')
>> xlabel('time (sec)'); ylabel('Voltage (V)')
```

This graph is clearly not straight – the exponential model is our best option.




## Function Discovery: 2-parameter models
(Try to find a plot that makes the data look linear)

| **Model** | **Equation** | **Linerized equation** | **Plot (command)** |
|---|---|---|---|
| Linear | $y = mx + b$ | $y = mx + b$ | linear (plot) |
| Exponential | $y = be^{mx}$ | $\ln(y) = \ln(b) + mx$ | semilog (semilogy) |
| | $y = b10^{mx}$ | $\log(y) = \log(b) + mx$ | |
| Power | $y = bx^m$ | $\log(y) = \log(b) + m\log(x)$ | log-log (loglog) |

## Handout: Function Discovery for monotonic functions

- Summary table (previous slide)
  - plus fitting models (we will cover next time)
- Assessing function type:
  - standard x-y plot
  - does data line up on a straight line? ➔ linear model
  - is the data monotonic for positive x?
  - ➔could be power law or exponential
  - ➔plot semilog and log-log plots.
    - linear on semilog ➔ exponential model
    - linear on loglog ➔ power law model
- Table of MATLAB commands for reference
- Try the three problems (see following slides)
- Fitting models ➔ what we need to do next

## Aside: base 10 vs. base e exponential models ➔ Different by a constant

| <u>base 10 model</u> | <u>base e model</u> |
|---|---|
| $y = b10^{m_1x}$ | $y = be^{m_2x}$ |
| $\log(y) = \log(b) + \log(10^{m_1x})$ | $\log y = \log(b) + \log(e^{m_2 x})$ |
| $\log(y) = \log(b) + m_1x \log(10)$ | $\log(y) = \log(b) + m_2x \log(e)$ |
| $\log(10) = 1$ | $\log(e) = 0.4343$ |
| $\log(y) = \log(b) + m_1x$ | $\log(y) = \log(b) + 0.4343\, m_2x$ |

therefore:

$$m_1x = 0.4343m_2x$$

$$m_1 = 0.4343\, m_2$$

## Exercise:

Determine the likely model form for:

1. Deflection of a cantilever beam (F, d)
2. x1 vs. y1
3. x2 vs y2

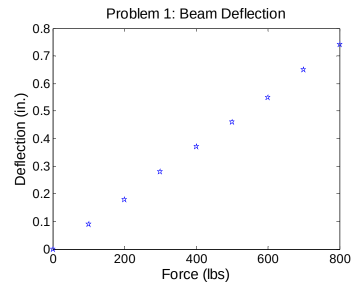
Data vectors Available

- On handout
- In MATLAB data file FnDiscovery.mat

## Problem 1: Deflection of a cantilever beam

This problem uses the variables F & d in FnDiscovery.mat

```
>> plot(F,d,'p')
>> xlabel 'Force (lbs)'
>> ylabel 'Deflection (in.)'
>> title 'Problem 1: Beam
   Deflection'
```




- data points are in a straight line

➔ the model is linear

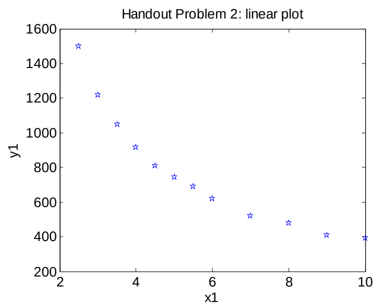
➔ d = mF + b

## Handout Problem 2

This problem uses variables x1, y1

standard x-y plot

```
>> plot(x1,y1,'p')
>> xlabel 'x1', ylabel 'y1'
>> title 'Handout Problem
   2: linear plot'
```

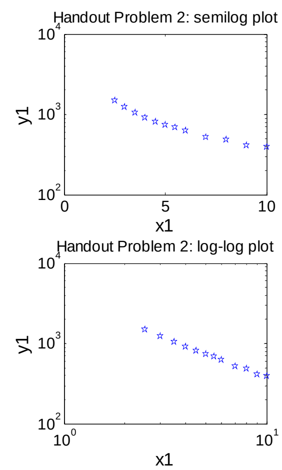



Result is monotonic but clearly non-linear

Need to try semilog and log-log plots

## Handout Problem 2 (cont.)

```
>> subplot(2,1,1)
>> semilogy(x1,y1,'p')
>> xlabel 'x1', ylabel 'y1'
>> title 'Handout Problem 2: semilog plot'
>>
>> subplot(2,1,2)
>> loglog(x1,y1,'p')
>> xlabel 'x1', ylabel 'y1'
>> title 'Handout Problem 2: log-log plot'
```




The log-log plot is linear ➔ power model applies

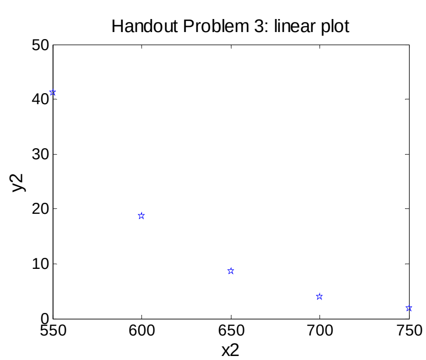
$y1 = b(x1)^m$

## Handout Problem 3

This problem uses variables x2, y2

standard x-y plot

```
>> plot(x2,y2,'p')
>> xlabel 'x2', ylabel 'y2'
>> title 'Handout Problem 3:
   linear plot'
```

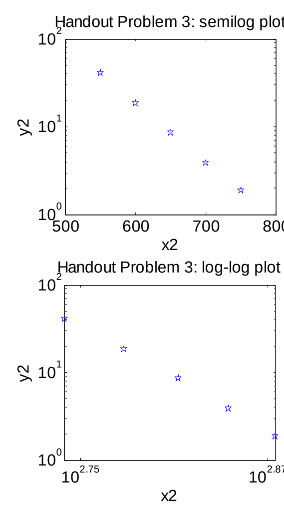



Result is monotonic but clearly non-linear

Need to try semilog and log-log plots

## Handout Problem 3 (cont.)

```
>> subplot(2,1,1)
>> semilogy(x2,y2,'p')
>> xlabel 'x2', ylabel 'y2'
>> title 'Handout Problem 3: semilog plot'
>>
>> subplot(2,1,2)
>> loglog(x2,y2,'p')
>> xlabel 'x2', ylabel 'y2'
>> title 'Handout Problem 3: log-log plot'
```




Both plots are fairly linear. However, I would choose the semilog plot for two reasons. (1) exponential models are generally preferred to power and (2) there is some possible curvature in the log-log plot.

➔ exponential model: $y = b\,e^{mx}$