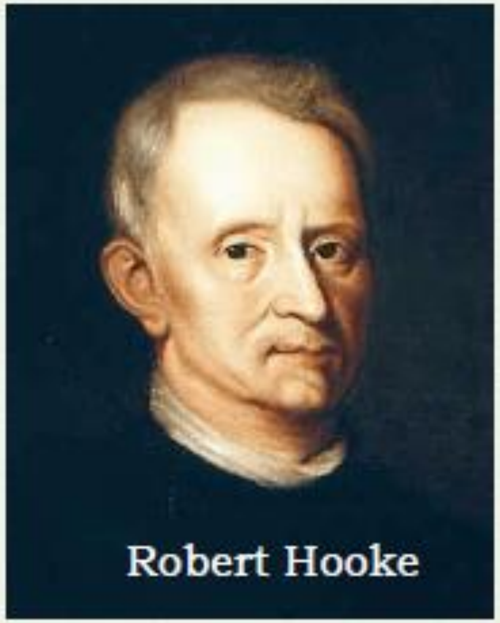
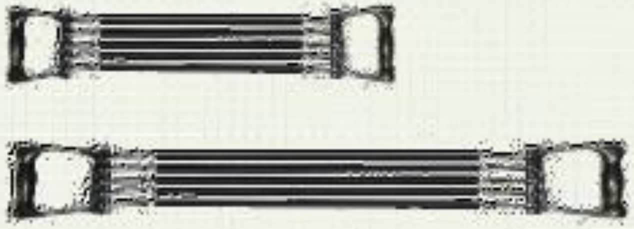
A Common Pattern in Science and Engineering

Hypothesis: Hooke's Law for Springs

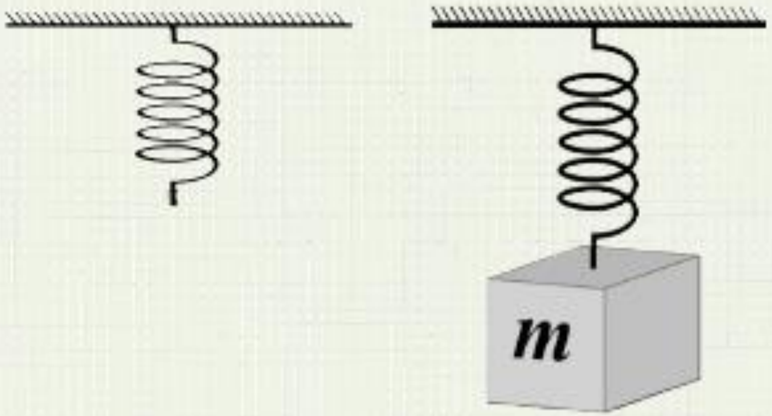


Robert Hooke

"The power of any springy body is in the same proportion with the extension."



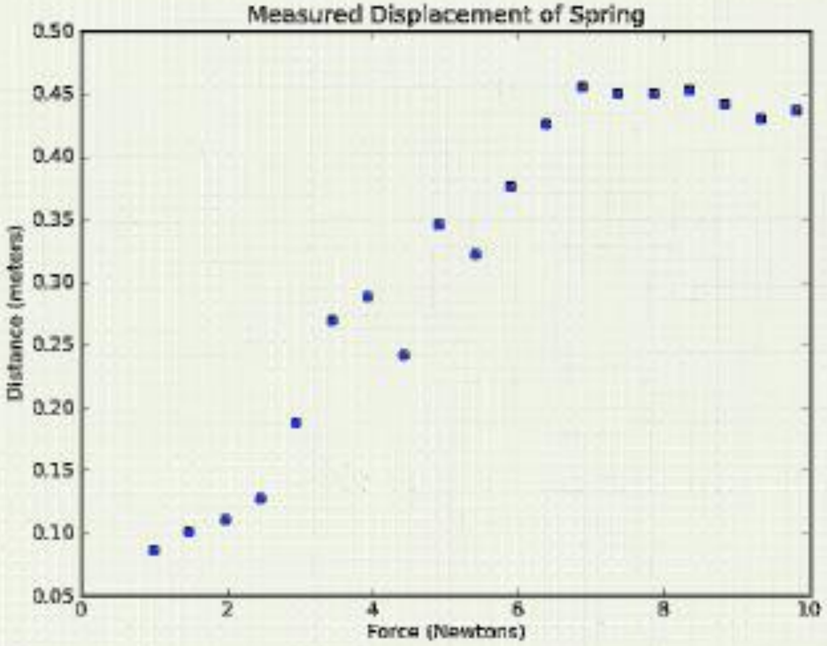
Experiment: apply force, measure extension



Process Observations

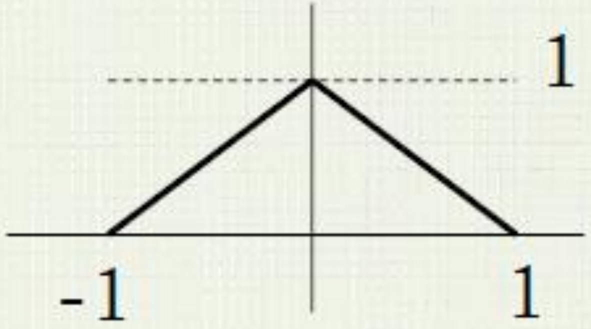
Distance(m)	Mass(kg)
0.0865	0.1
0.1015	0.15
0.1106	0.2
0.1279	0.25
…	
0.4416	0.9
0.4304	0.95
0.437	1.0

Hypothesis: linear relationship between F and x

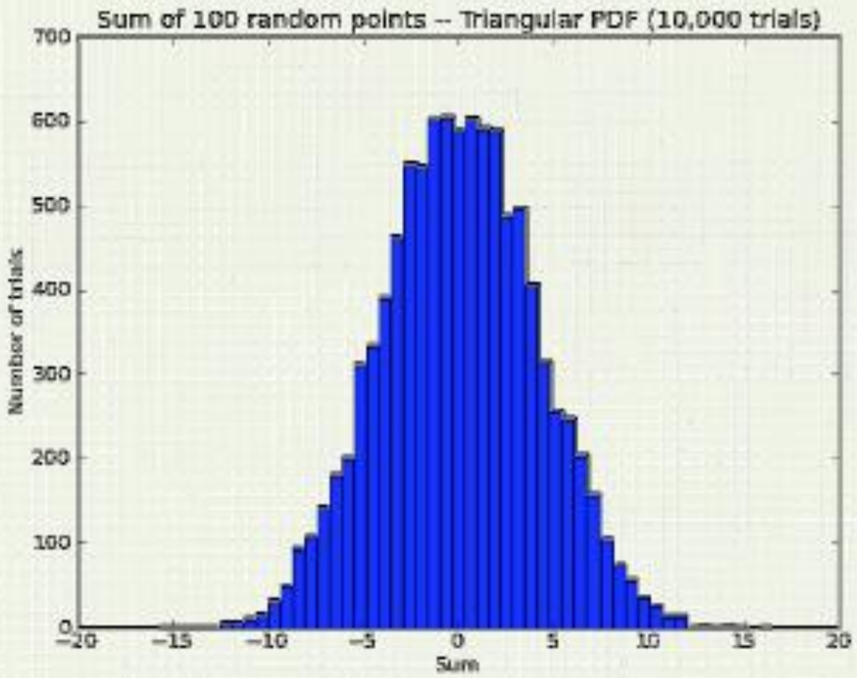


$$\text{observation}_i = \text{prediction}_i \pm \text{error}$$

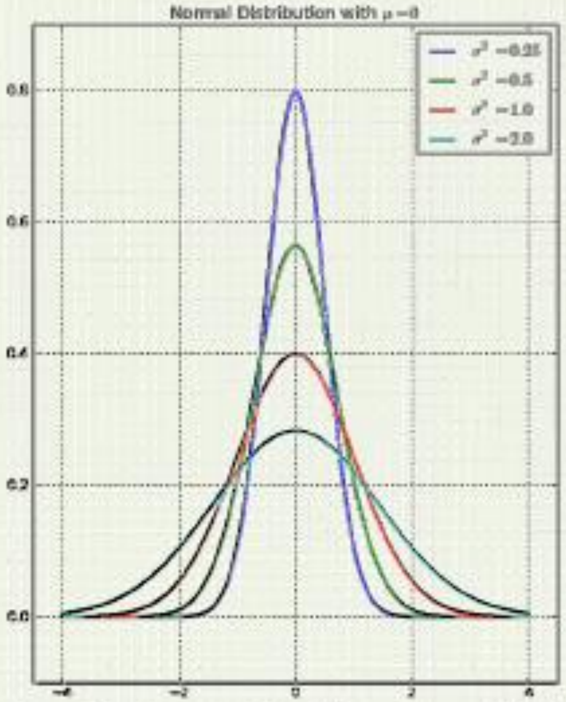
Accumulation of many small random errors



Accumulation of many small random errors



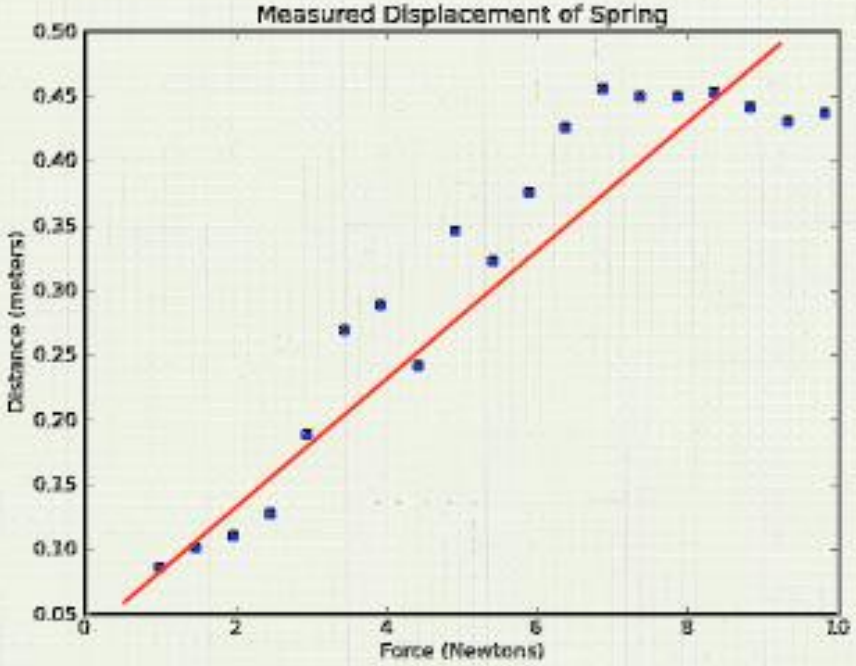
The Normal distribution



Gaussian model for observed errors

So when observation errors are due to the accumulation of many small random perturbations:

Have: observations, Want: most likely line



Log Likelihood

$$\text{Maximize} \prod_{i=0}^{len(obs)-1} L_{err}(obs_i - pred_i)$$

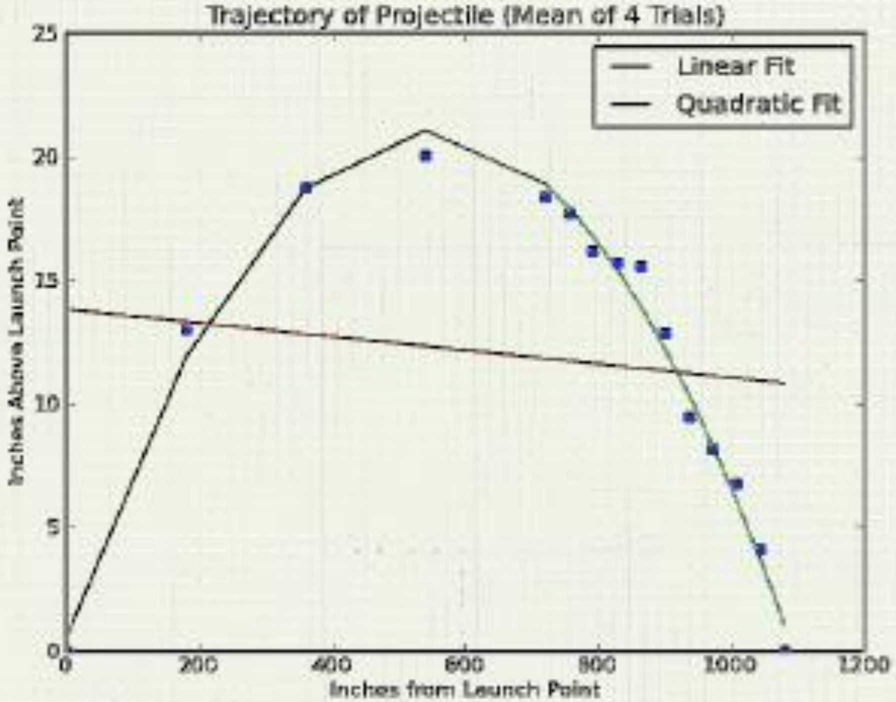
Least Squares

pylab.polyfit(xvals, yvals, degree)

```
# find a, b that minimize
# sum((yvals-(a*xvals + b))**2)
a,b = pylab.polyfit(xvals,yvals,1)


# find a, b, c that minimize
# sum((yvals-(a*xvals**2 + b*xvals + c))**2)
a,b,c = pylab.polyfit(xvals,yvals,2)
```

Measuring "goodness" of fit



R^2: Coefficient of Determination

Fraction of variability *not* explained by model

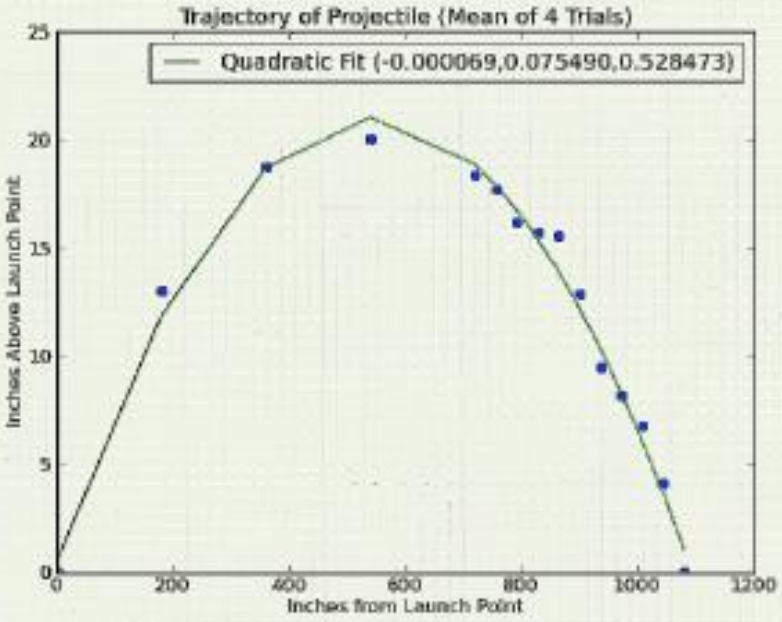
Fraction of variability explained by model

Using a model: How thick a shield?

Use the model to make predictions when experiments are impractical or inadvisable.

Want to know the speed of the arrow as it reaches the target.

Arrow speed when reaching target



$$y = (-.000069)x^2 + (.0755)x + .528$$