جــامـــعـــة دمـــشـــق
كـليـة الهـنـدسـة المـدنيـة

تـصـمـيــم بـمـعـونــة الـحــاســب
Computer Aided Design

Lecture 4

Statistics & Optimization

2/12/2025

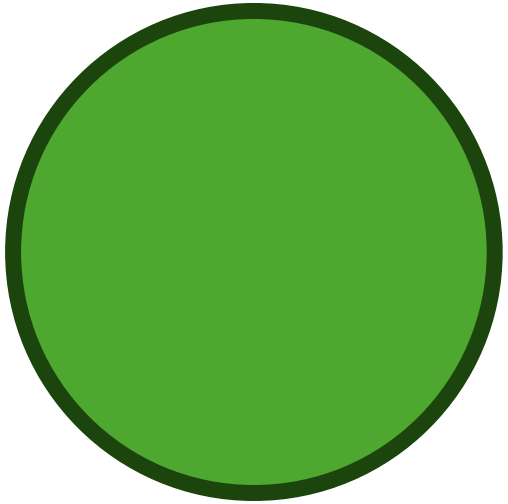
Wael Darwich

Damascus University

Civil Engineering Faculty

Last Lecture

- Loops Continued…
- Examples
- Notes `inside`:
  1. Debugging
  2. Performance
  3. Numerical Tolerance
  4. `on_circle`
- Figures



```
dist2 == r2
```

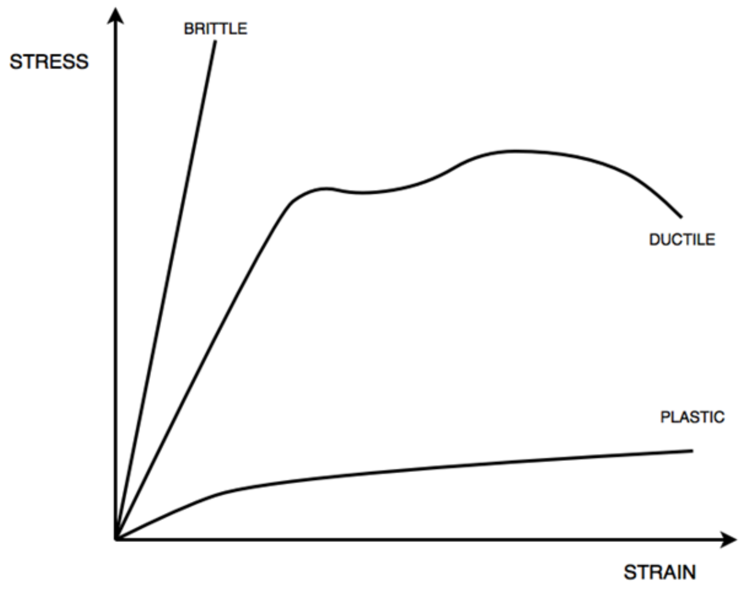
```
dist2 <= r2 + 1E-6 && dist2 >= r2 - 1E-6
```

Content

- Statistics
  - Graphics
- Regression and Curve Fitting
  - Linear
  - Higher order
  - Example
- Optimization
- Projectile Motion
  - Analytical solution
  - Numerical solution
- Other Examples
- Layout Optimization

Statistics

- Data (collected from measurements)
  - e.g. section dimensions, concrete compressive strength, soil density...
- Mathematical model
  - e.g. Young's modulus $= \frac{\sigma}{\varepsilon} = \frac{F/A}{dL/L}$



Data Analysis

- S275 yield strength specimens $f_y = 275\ MPa$
- `[328.0 307.1 305.8 292.1 305.7 325.7 343.3 326.8 307.5 314.3]`
- [**292.1** 305.7 305.8 307.1 307.5 314.3 325.7 326.8 328.0 **343.3**]
- Min/Max
- Sort

Data Analysis

- S275 yield strength specimens $f_y = 275\ MPa$
- `[328.0 307.1 305.8 292.1 305.7 325.7 343.3 326.8 307.5 314.3]`
- [**292.1** 305.7 305.8 307.1 307.5 314.3 325.7 326.8 328.0 **343.3**]

• Range:	max – min value	= 051.2000	range
• Mean:	the average	= 315.6300	mean
• Median:	middle number	= 310.9000	median
• Mode:	most frequent	= 292.1000	mode

Data Analysis

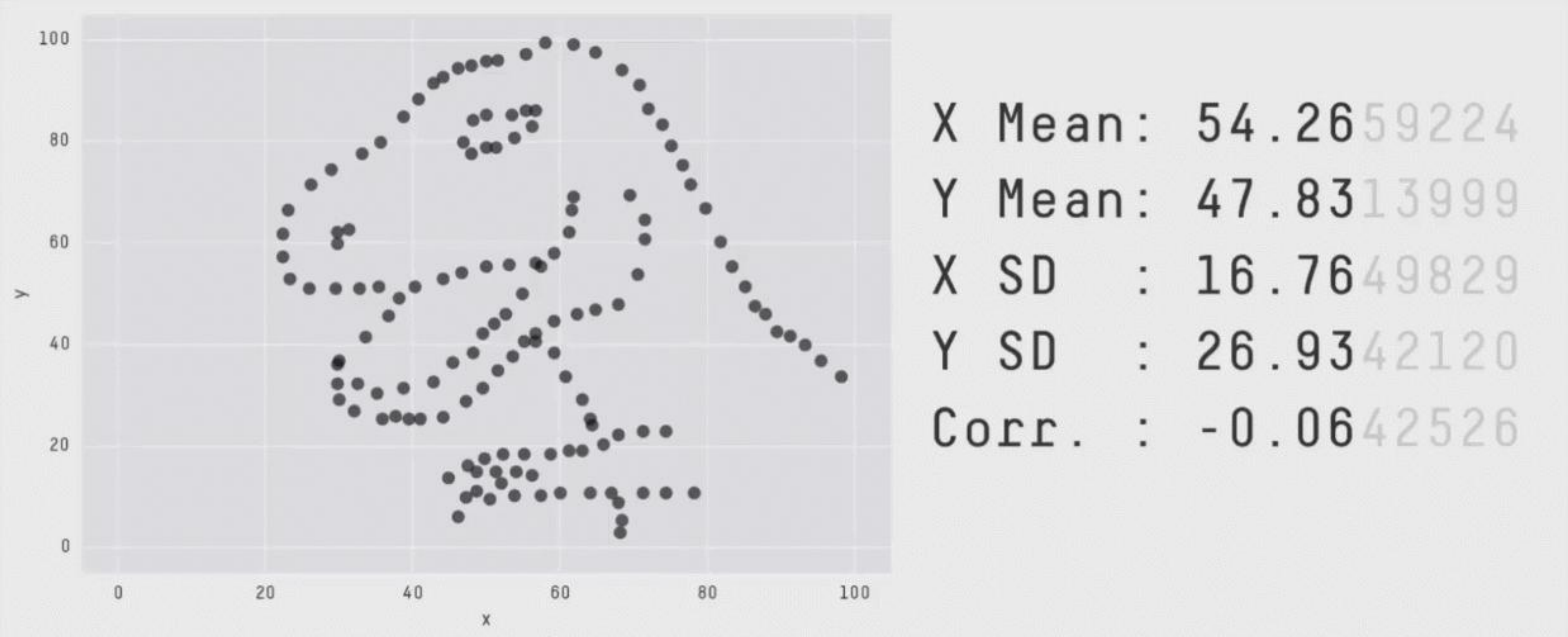
- S275 yield strength specimens $f_y = 275\ MPa$
- `[328.0 307.1 305.8 292.1 305.7 325.7 343.3 326.8 307.5 314.3]`
- `[`**`292.1`**` 305.7 305.8 307.1 307.5 314.3 325.7 326.8 328.0 `**`343.3`**`]`

• Range:	max – min value	`= 051.2000`	`range`
• Mean:	the average	`= 315.6300`	`mean`
• Median:	middle number	`= 310.9000`	`median`
• Mode:	most frequent	`= 292.1000`	`mode`
• Variance:	dispersion of data	`= 226.0379`	`var`
• Standard deviation:	$= \sqrt{\text{Variance}}$	`= 015.0346`	`std`

Data Analysis

- $x \ = [x_1, x_2 \dots \ x_N]$
- Range = Max - Min
- $\bar{x} = \ \frac{\sum_{i=1}^{N} x_i}{N}$
- $s_N = \ \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}}$
- Never trust summary statistics alone;
- Always visualize your data

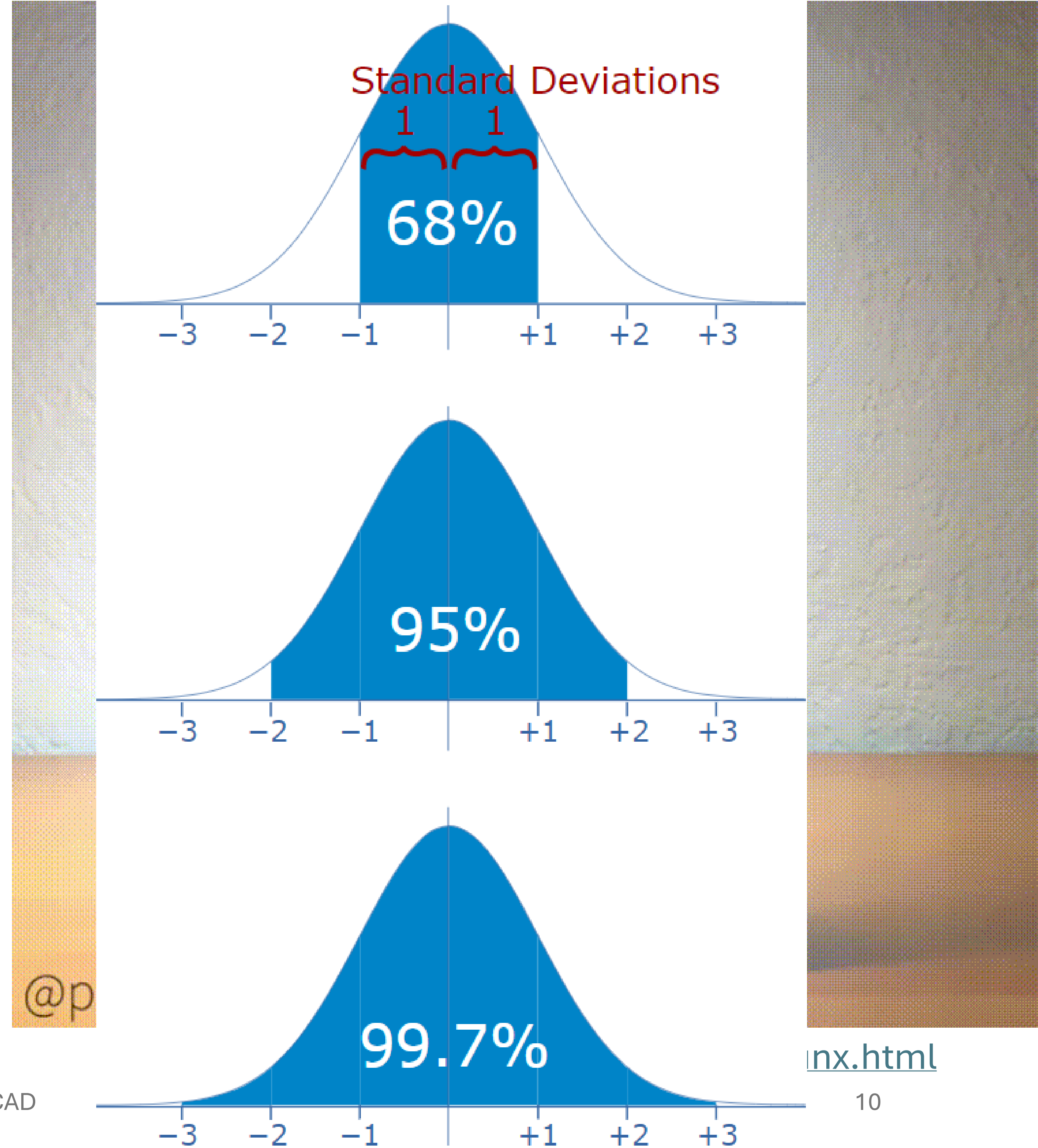
Data Visualization



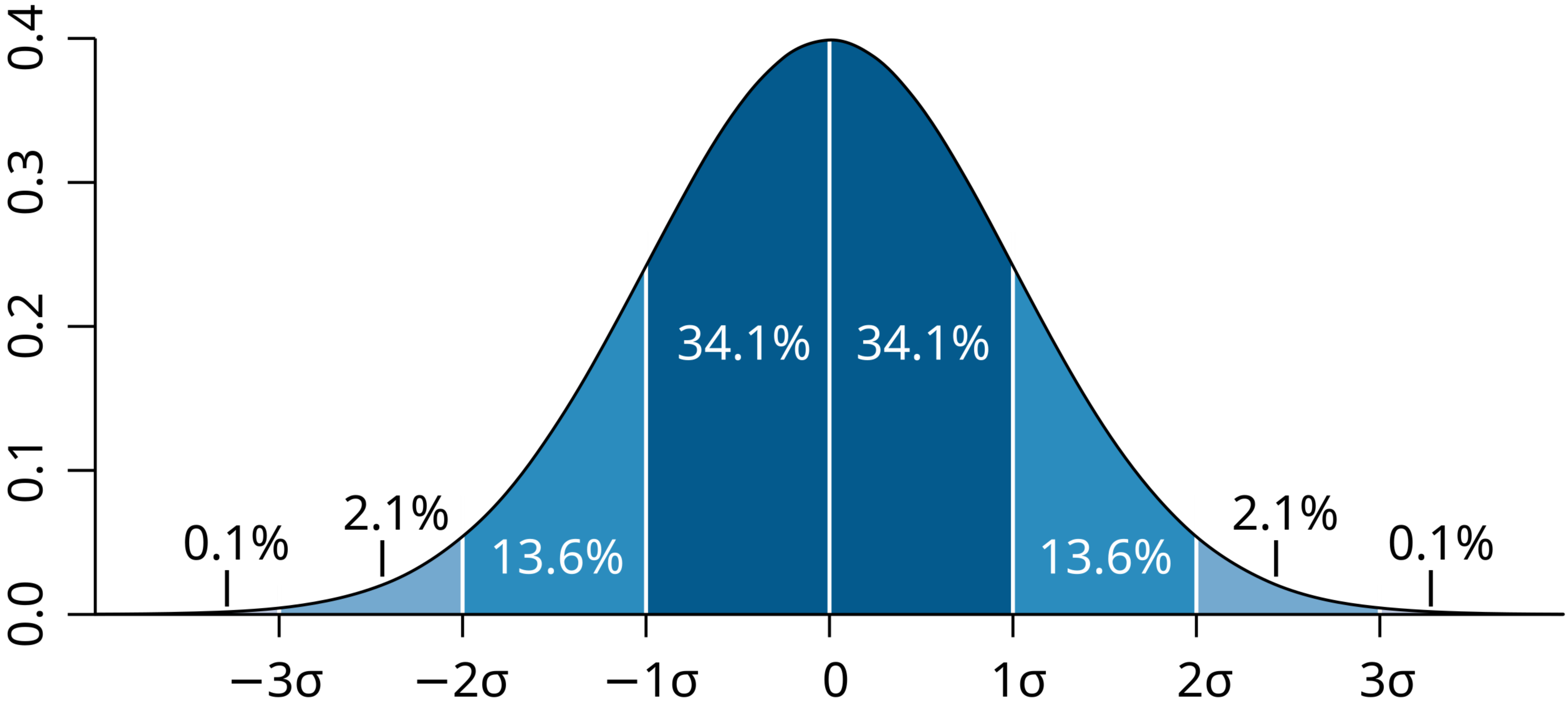
https://www.autodesk.com/research/publications/same-stats-different-graphs

Normal Distribution

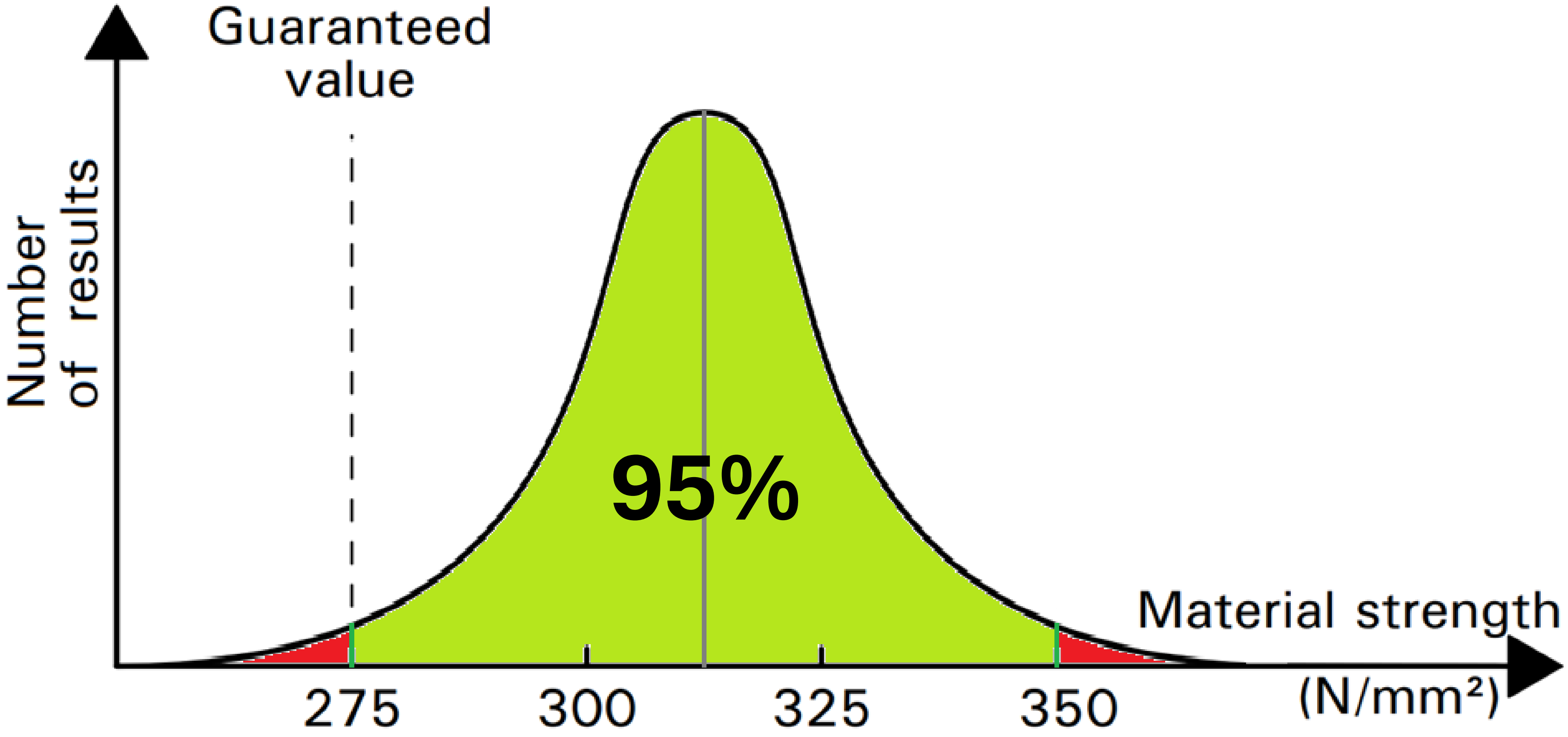
- Bell curve
- Symmetry about the centre
- mean = median = mode



nx.html


0.4
0.3
0.2
0.1
0.0
34.1%
34.1%
13.6%
13.6%
2.1%
2.1%
0.1%
0.1%
−3σ
−2σ
−1σ
0
1σ
2σ
3σ

S275



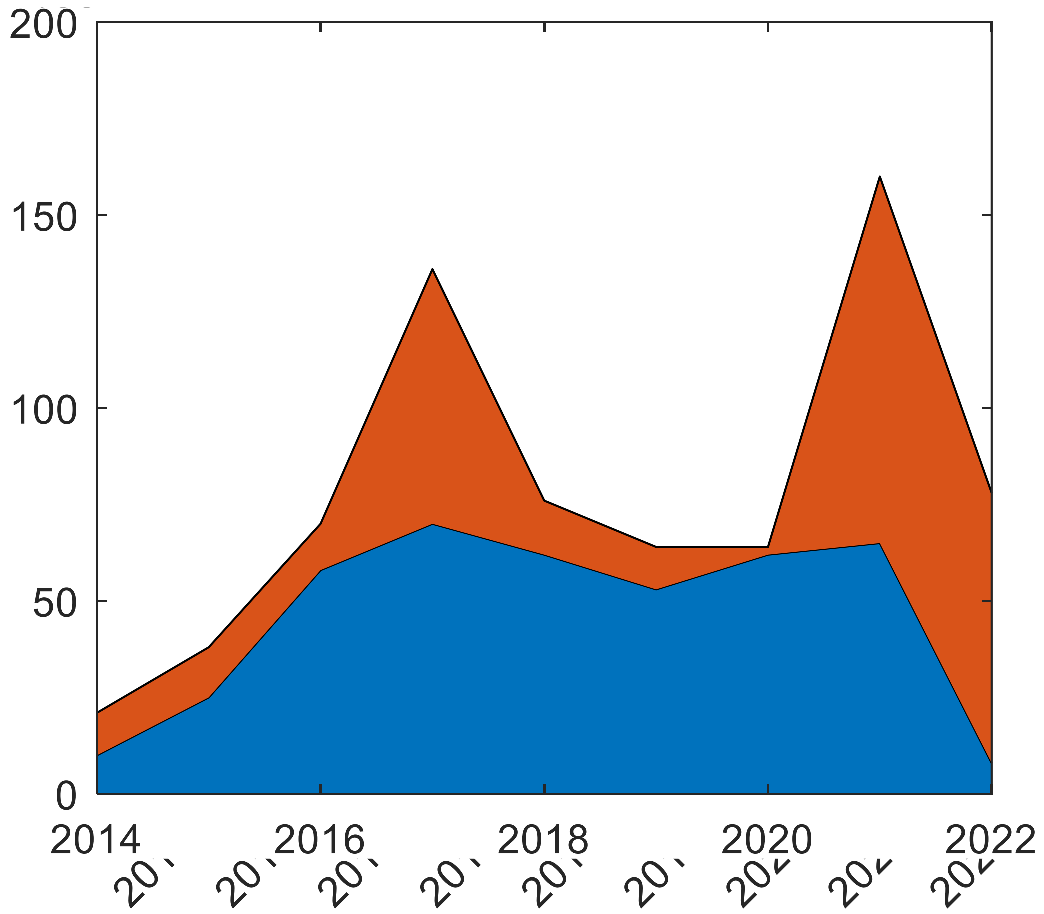
In this case the number of results falling below the guaranteed minimum is very small

Graphics

- Bar
- Area
- Scatter
- Histogram

Bar & Area

```
year = 2014:2022;
count = [10 25 58 70 62 53 62 65 8];
bar(year, count)

counts = [10 11; 25 13; 58 12;
          70 66; 62 14; 53 11;
          62  2; 65 95;  8 70];
bar(year, counts)

area(year, count)
area(year, counts)
```

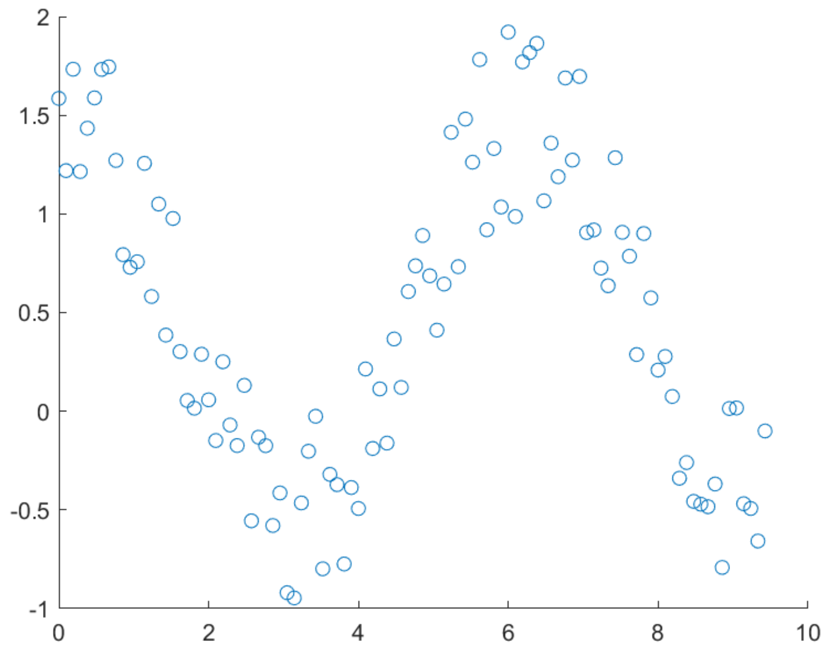


Random Data

```
rand(1)          0.3692   between 0 and 1
rand(3)          0.1112    0.2417    0.1320
                 0.7803    0.4039    0.9421
                 0.3897    0.0965    0.9561

                 0.5752    0.3532
rand(3, 2)       0.0598    0.8212
                 0.2348    0.0154

rand(1, 10)
    0.0430    0.1690    0.6491    0.7317    0.6477    0.4509    0.5470    0.2963    0.7447    0.1890
```

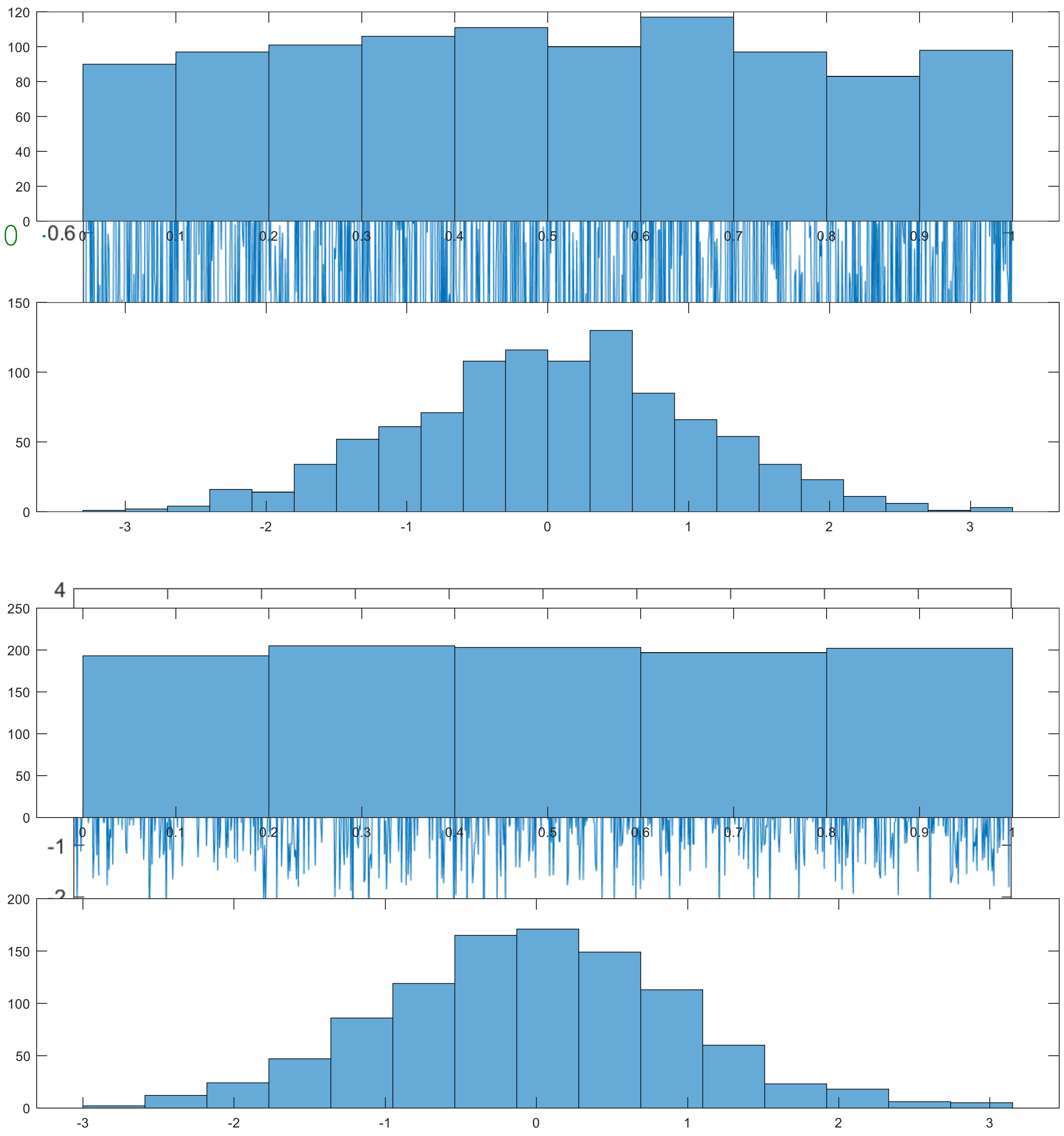
Scatter

```
x = linspace(0, 3*pi, 100);
y = cos(x) + rand(1, 100);
scatter(x, y)
```

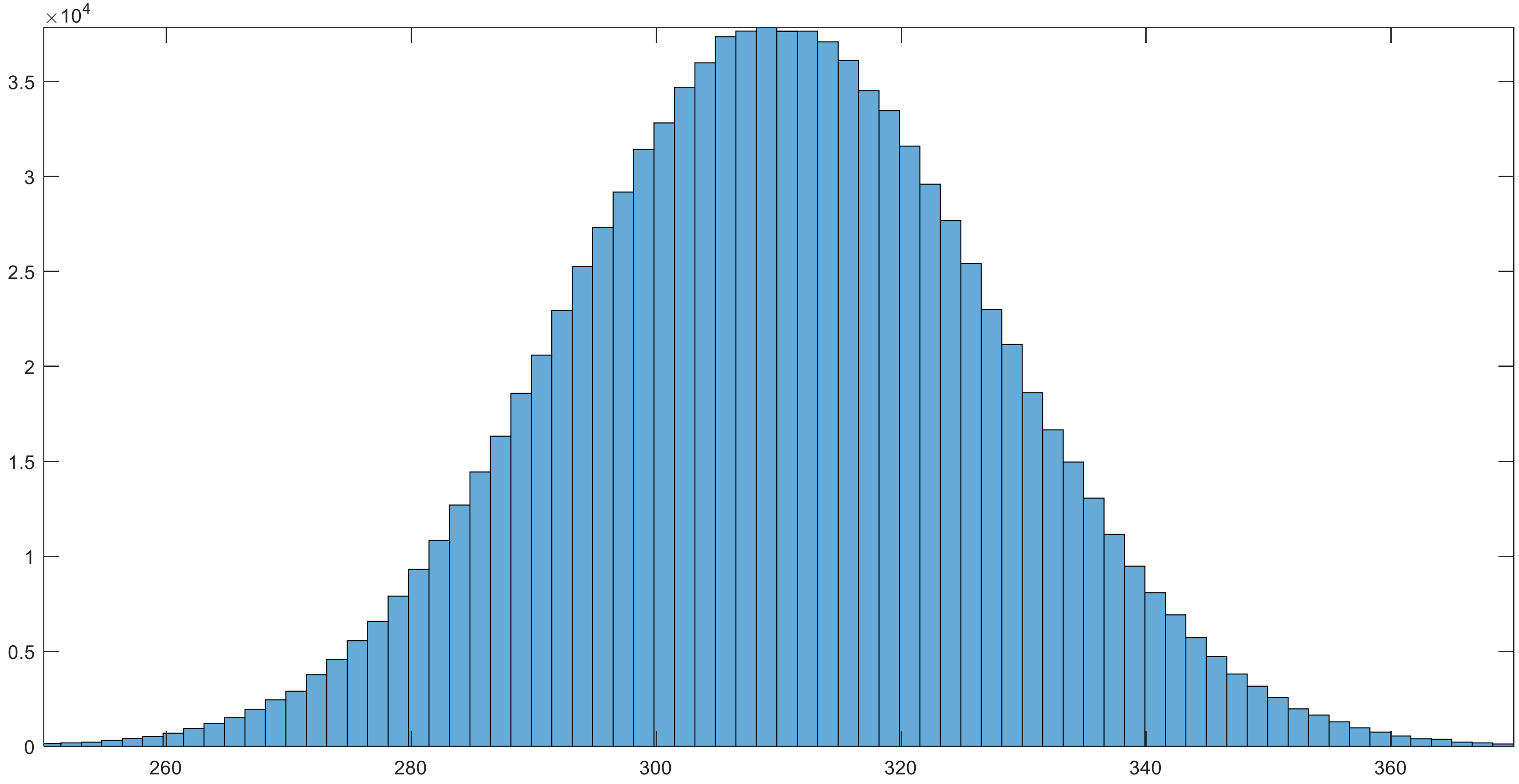


Histogram

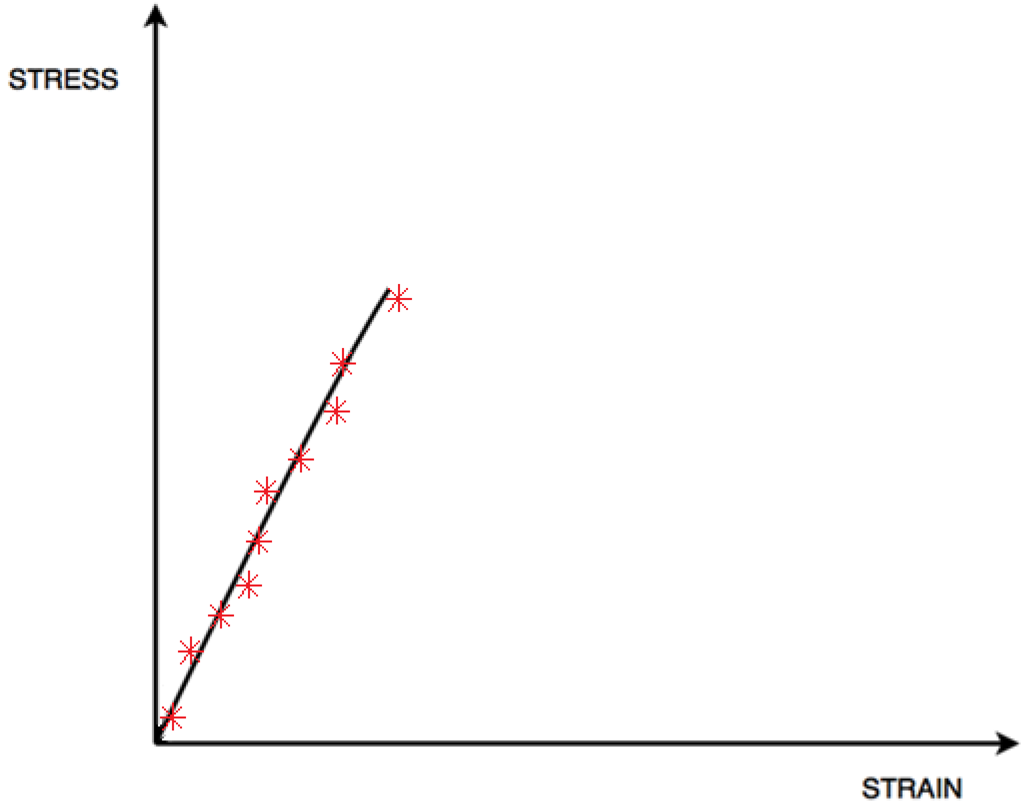
```
% rand: Uniformly distributed pseudorandom numbers 0 - 1
min(rand(1, 1e6))
max(rand(1, 1e6))

% randn: Normally distributed pseudorandom numbers
max(randn(1, 1e6))
min(randn(1, 1e6))

plot(rand(1000,1))
plot(randn(1000,1))

histogram(rand(1000,1))
histogram(randn(1000,1))

histogram(rand(1000,1), 5) % change bins number
histogram(randn(1000,1), 15)
```



S275

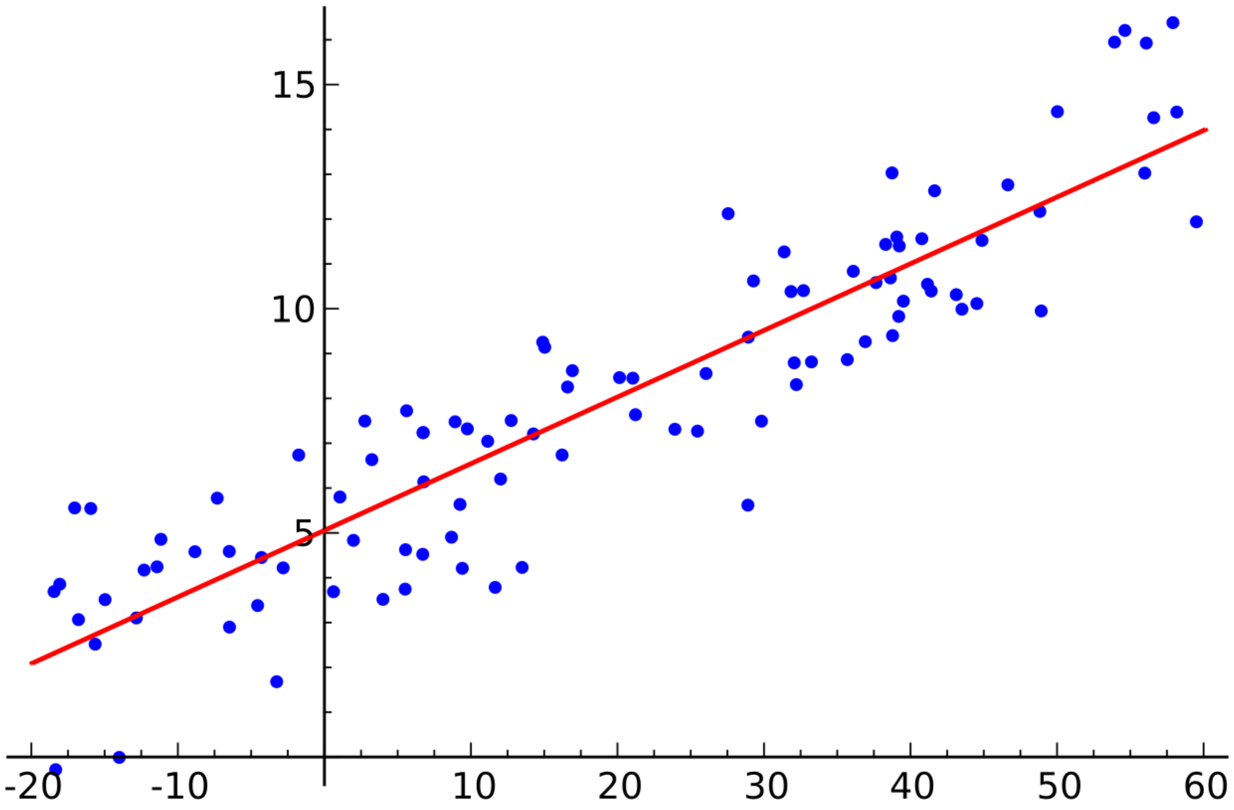


Curve Fitting



Regression Analysis

- Find relationships between variables
  - Linear regression
  - minimizes sum of squared differences between true data and that line
- Power, exponential, sine…
- Single vs multiple regression



Regression Analysis

- n points

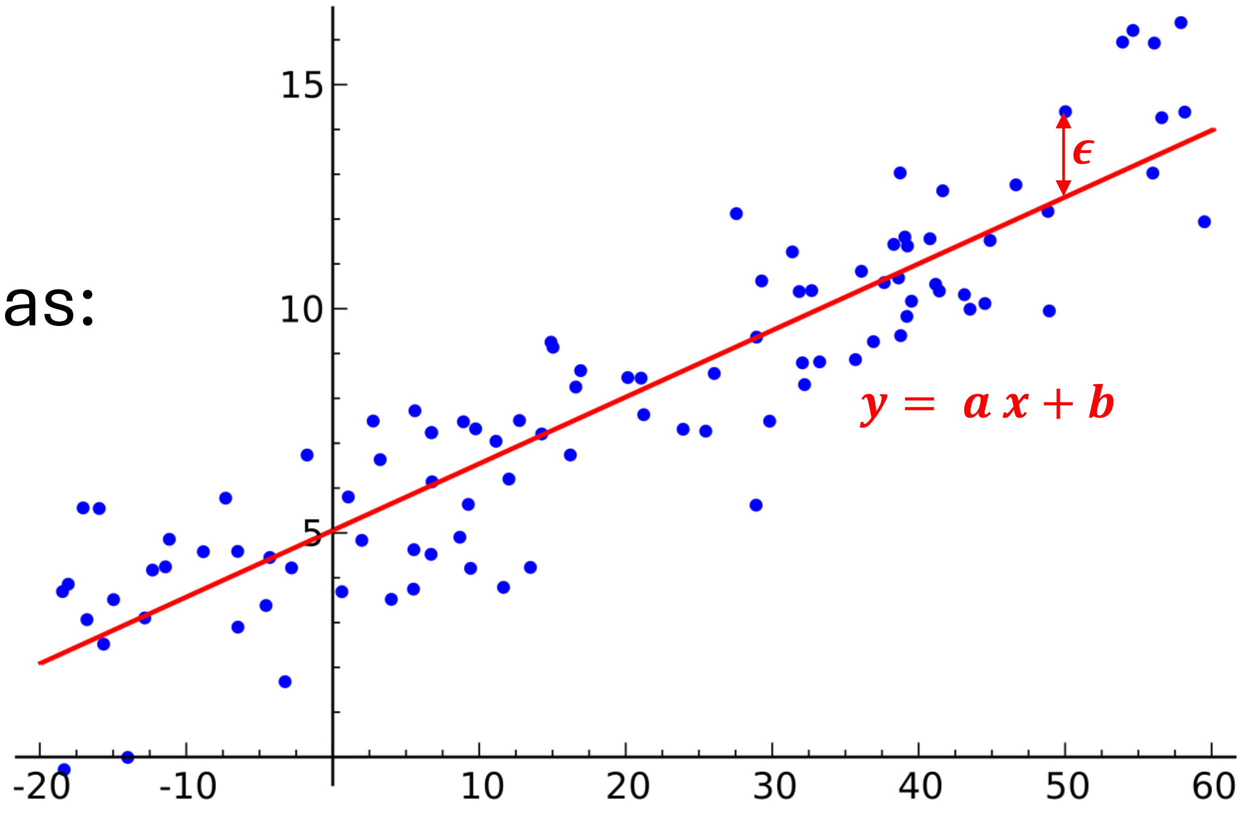
  $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$

- for linear regression each point has:

$$y = a\,x + b$$
$$z = a\,x + by + c$$

Due to error:

$$y = \beta_0 + \beta_1\,x + \varepsilon$$

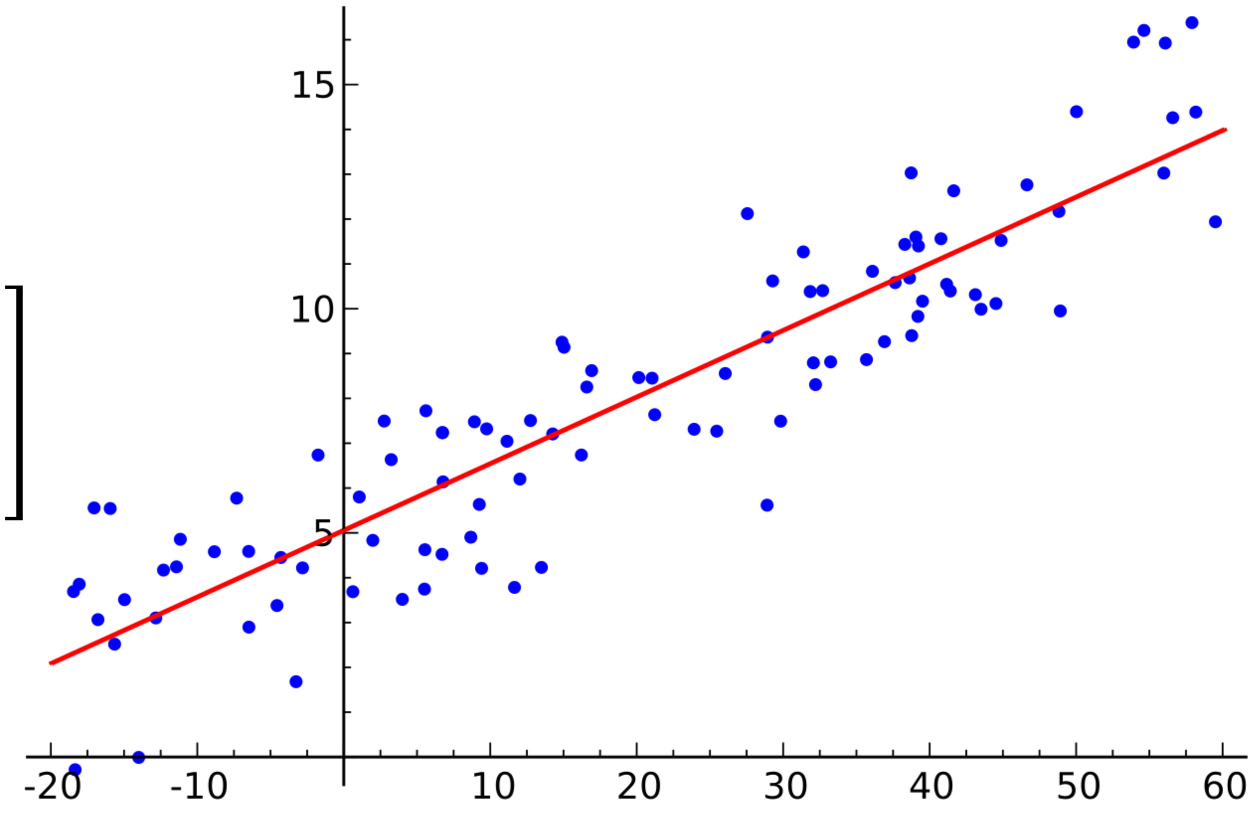


Regression Analysis

For *n* points: $y = \beta_0 + \beta_1 x + \varepsilon$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad x = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$y = \beta x + \varepsilon$$



Find β to minimize n errors $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n$

$$\varepsilon_i = y_i - \beta x_i$$

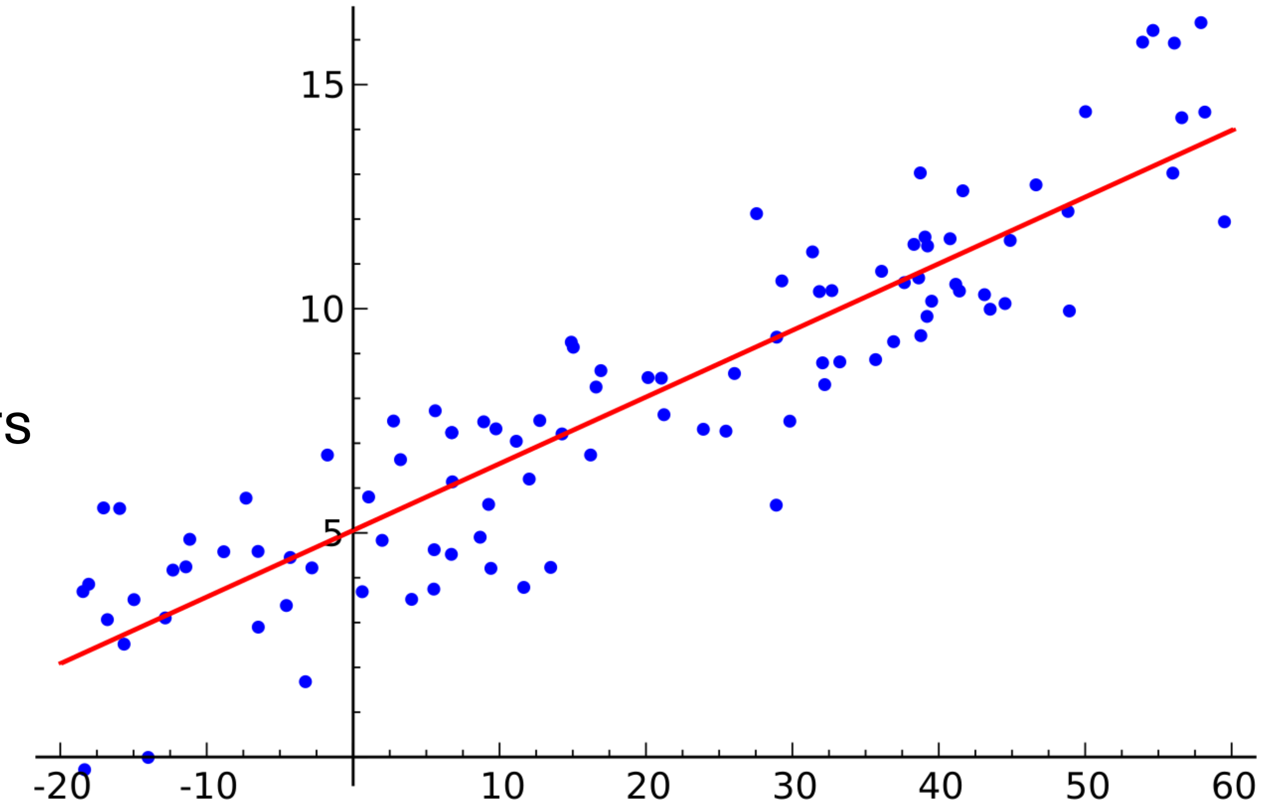
Least Squares Method

$$\varepsilon_i = y_i - \beta\, x_i$$

- Minimize all errors
- Minimize both +tive and –tive errors
- Penalize larger errors more
- Minimize sum of squares

$$error = \sum_{i=1}^{n} \varepsilon_i^2$$

$$= \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$



Least Squares Method

Minimize error:

$$error\ =\ \sum_{i=1}^{n}(y_i - (\beta_0 + \beta_1 x_i))^2 = \sum_{i=1}^{n} y_i^2 + (\beta_0^2 + \beta_1^2 {x_i}^2 + 2\beta_0\beta_1 x_i) - 2(y_i\beta_0 + y_i x_i \beta_1)$$

$$= \sum_{i=1}^{n} y_i^2 + \beta_0^2 + \beta_1^2 {x_i}^2 + 2\beta_0\beta_1 x_i - 2y_i\beta_0 - 2y_i x_i \beta_1$$

Least Squares Method

Minimize error:

$$error = \sum_{i=1}^{n} y_i^2 + \beta_0^2 + \beta_1^2 {x_i}^2 + 2\beta_0\beta_1 x_i - 2y_i\beta_0 - 2y_i x_i \beta_1$$

$$\frac{\partial_{error}}{\partial_{\beta_0}} = \sum_{i=1}^{n} 2\beta_0 + 2\beta_1 x_i - 2y_i = 2\sum_{i=1}^{n} \beta_0 + \beta_1 x_i - y_i$$

$$\frac{\partial_{error}}{\partial_{\beta_1}} = \sum_{i=1}^{n} 2\beta_1 {x_i}^2 + 2\beta_0 x_i - 2y_i x_i = 2\sum_{i=1}^{n} x_i(\beta_1 x_i + \beta_0 - y_i)$$

Least Squares Method

$$\frac{\partial_{error}}{\partial_{\beta_0}} = 2\sum_{i=1}^{n} \beta_0 + \beta_1 x_i - y_i = 0 \qquad \sum_{i=1}^{n} \beta_0 + \sum_{i=1}^{n} \beta_1 x_i = \sum_{i=1}^{n} y_i \qquad n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\frac{\partial_{error}}{\partial_{\beta_1}} = 2\sum_{i=1}^{n} x_i(\beta_1 x_i + \beta_0 - y_i) = 0 \qquad \beta_1 \sum_{i=1}^{n} {x_i}^2 + \beta_0 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

Least Squares Method

$$n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i y_i$$

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum {x_i}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum {x_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

Higher Order

- Higher order (j):

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \ldots + \beta_j x^j + \varepsilon$$

$$y = \sum_{k=0}^{j} \beta_k x^k + \varepsilon$$

$$error = \sum_{i=1}^{n} \varepsilon_i^2$$

$$= \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

$$= \sum_{i=1}^{n} (y_i - \sum \beta_k x_i^k)^2$$

Higher Order

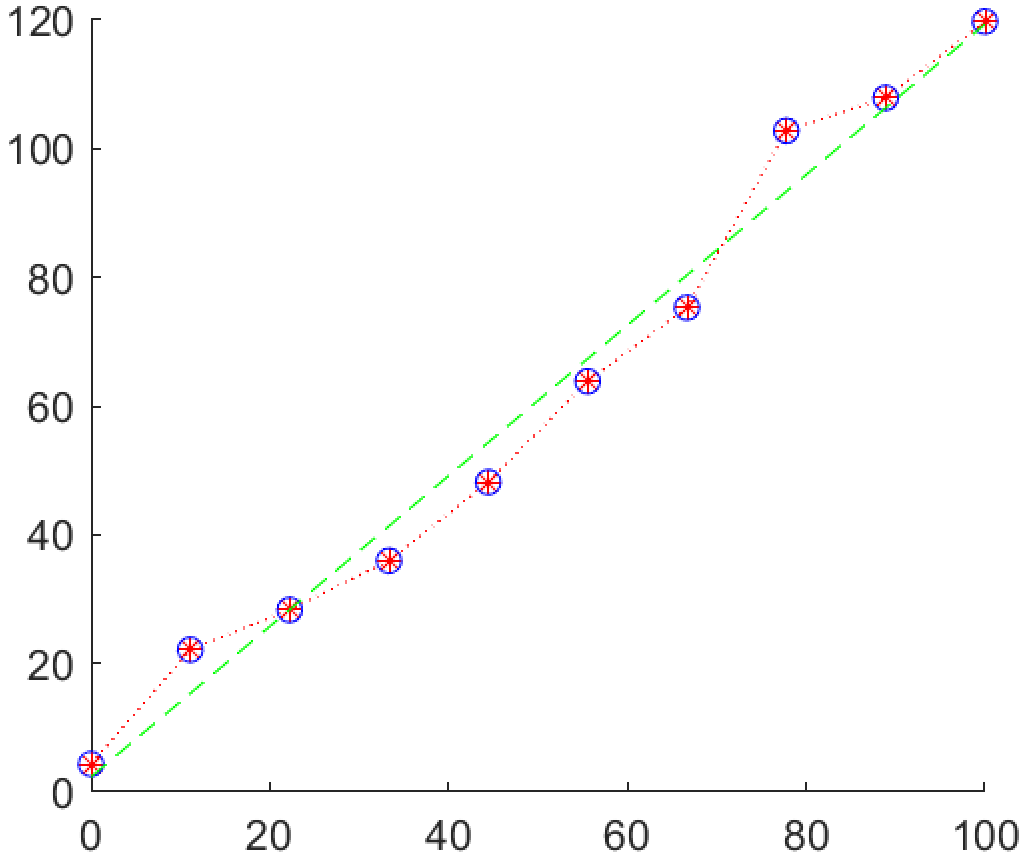
$$\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum {x_i}^2\end{bmatrix}^{-1}\begin{bmatrix}\sum y_i \\ \sum x_i y_i\end{bmatrix}$$

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \vdots\\ \beta_j\end{bmatrix} = \begin{bmatrix} n & \sum x_i & \sum x_i^2 & \dots & \sum x_i^j \\ \sum x_i & \sum x_i^2 & \sum x_i & \dots & \sum x_i^{j+1} \\ \sum x_i^2 & \sum x_i^2 & \sum x_i^2 & \dots & \sum x_i^{j+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum x_i^j & \sum x_i^{j+1} & \sum x_i^{j+2} & \dots & \sum x_i^{j+j}\end{bmatrix}^{-1}\begin{bmatrix}\sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \\ \vdots \\ \sum x_i^j y_i\end{bmatrix}$$

Example

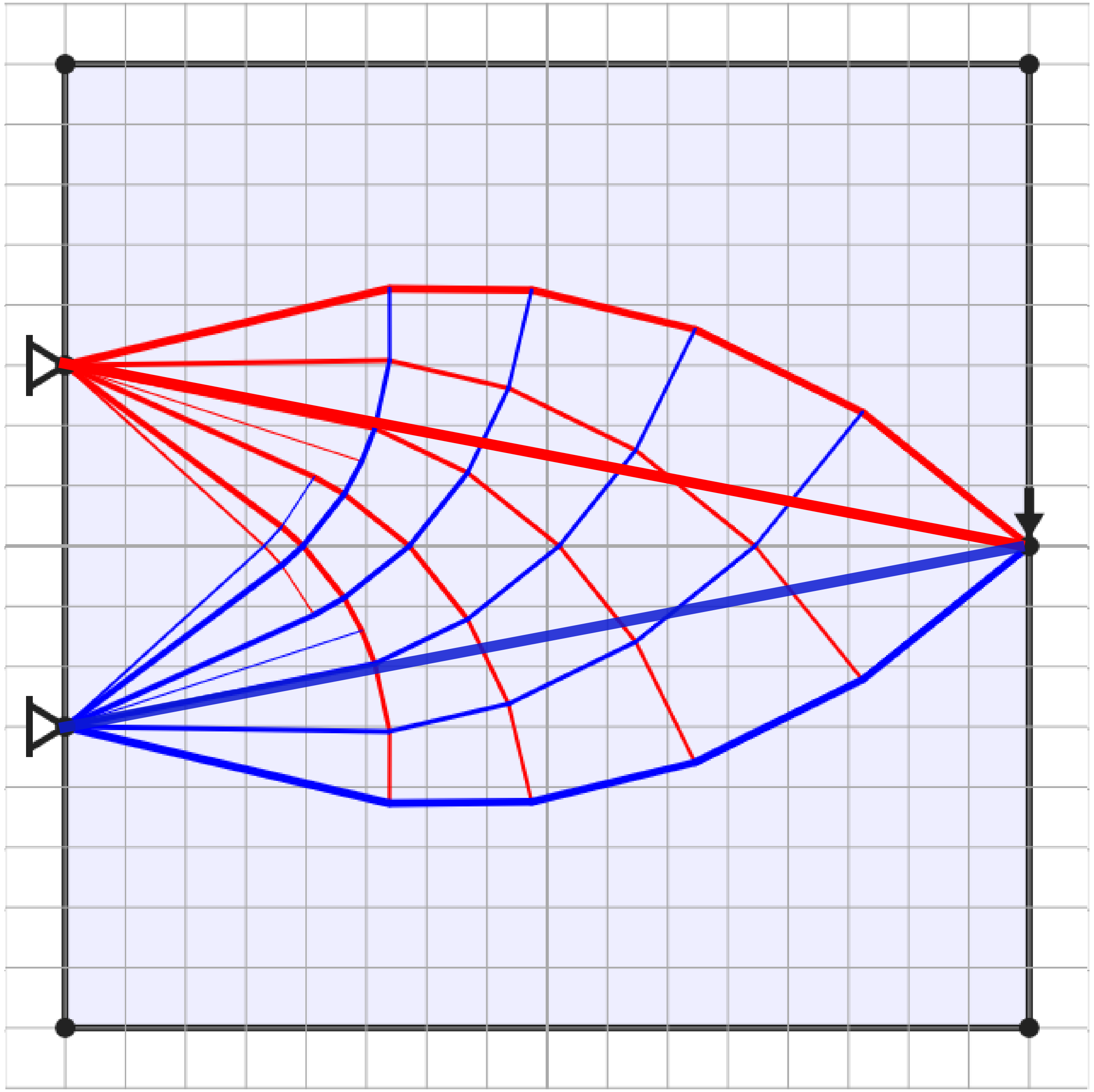
```
a = 1.2; b = 2;
points = 10;
x = linspace(0, 100, points); y = a * x + b + rand (1, points)*20-10;

plot(x, y, 'r:*')
scatter(x, y, 'b')

n = length(x);
sigma_x = sum(x); sigma_x2 = sum(x.^2);
sigma_y = sum(y); sigma_xy = sum(x.*y);

Beta = [n sigma_x; sigma_x sigma_x2] \ [sigma_y; sigma_xy];
     = [2.3389;   1.1701]

p = polyfit(x, y, 1);

y2 = p(1) * x + p(2);
y2 = polyval(p, x);

plot(x, y2, 'g--')
```

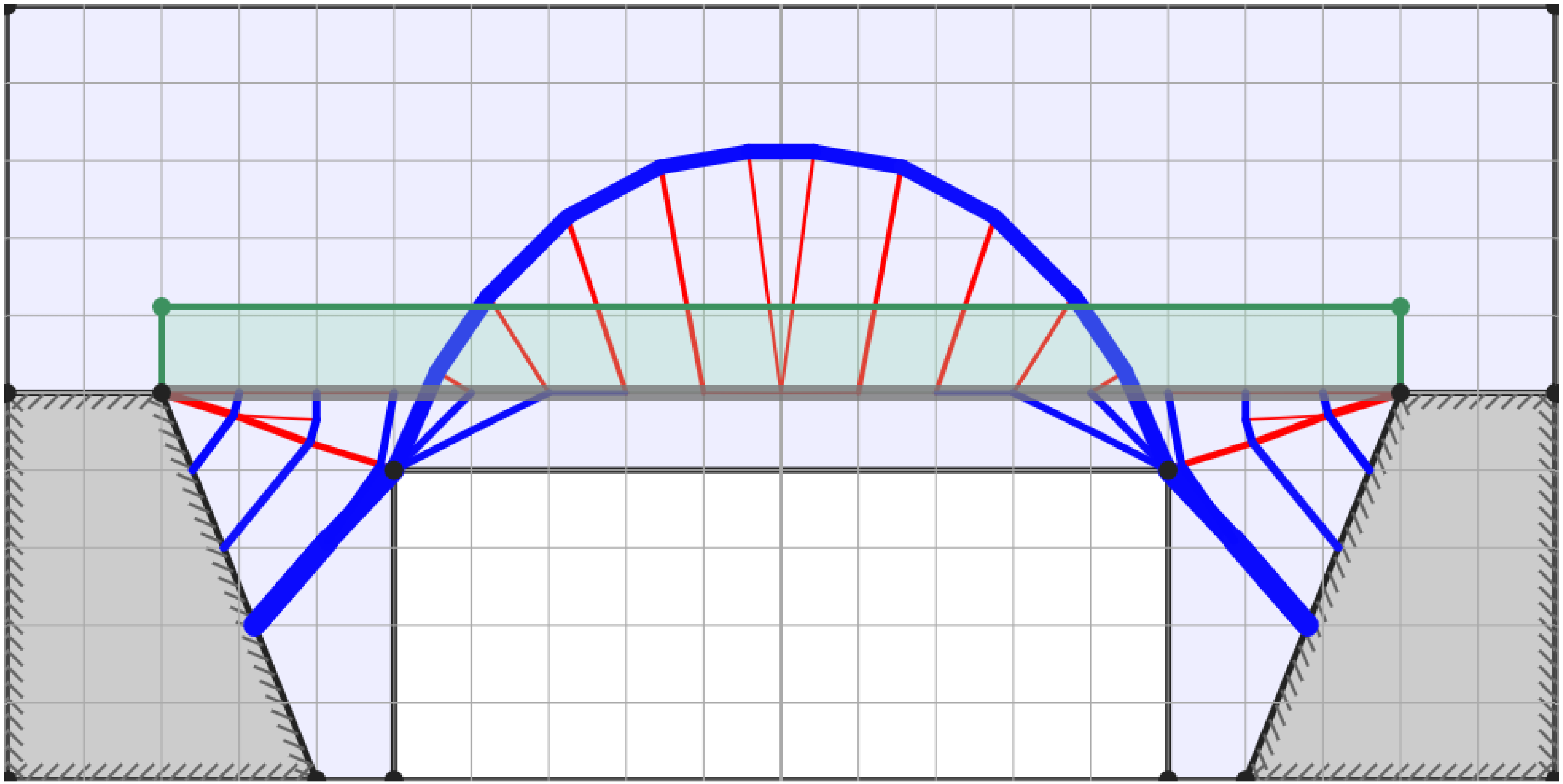
$$\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum {x_i}^2 \end{bmatrix}^{-1} \begin{bmatrix}\sum y_i \\ \sum x_i y_i\end{bmatrix}$$



Optimization

- Optimum design
- Minimize material
- Truss example
- Minimize cost / maximize gain
- Minimize build time





Mathematical Optimization

- Mathematical programming
- Objective function
  - E.g. most efficient design
  - Maximization, minimization
- Variables
  - Continuous, e.g. geometry
  - Discrete, e.g. section or material selection (integer, binary)
- Constraints

Maximize Revenue

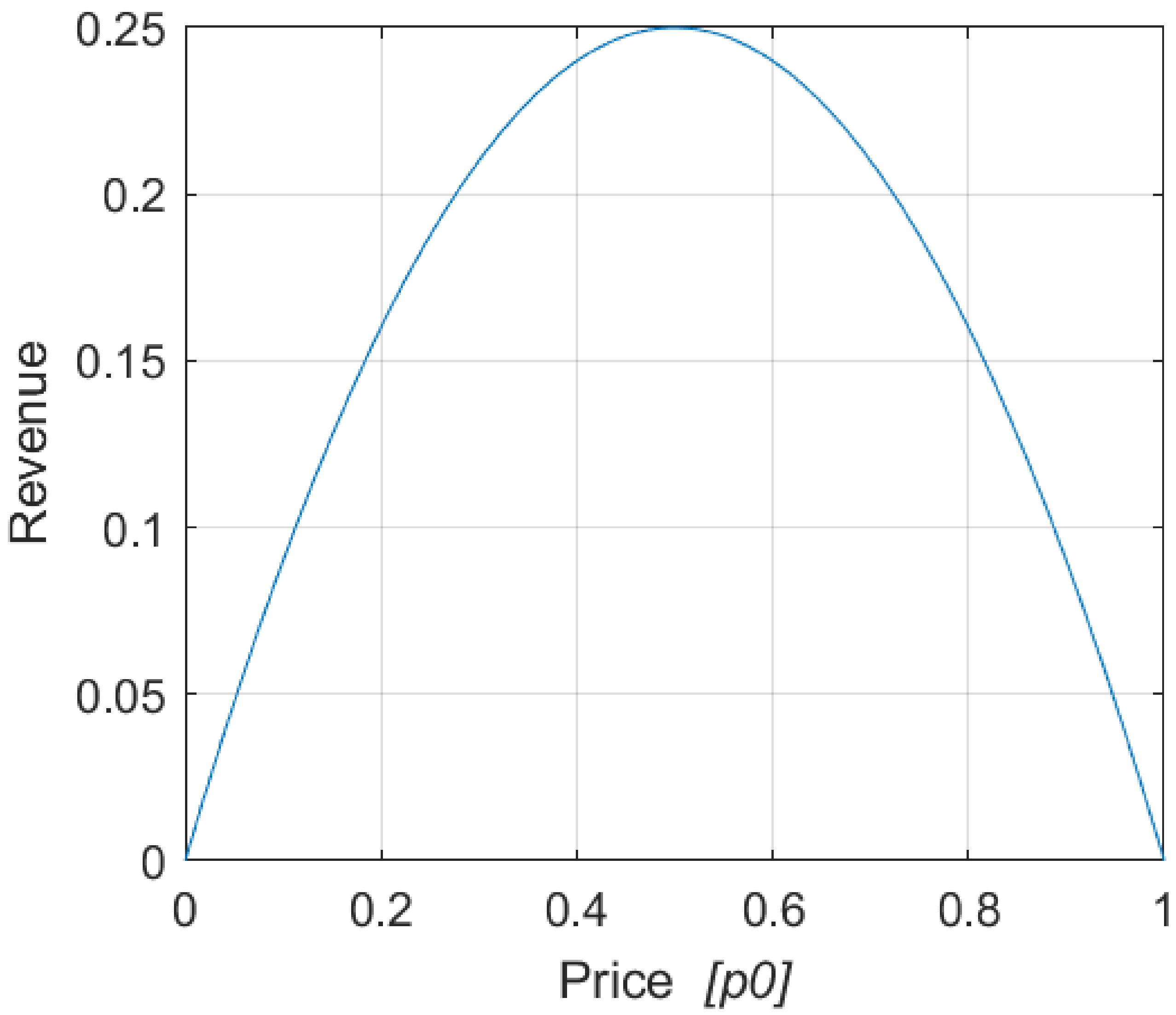
- A company is studying the pricing of a product.
- The expected number of people $n(p)$ that would buy the product is linearly decreasing with the product price p.
  - If the price is p_0, then nobody would buy it.
  - If the product was free, then a number n_0 of people would want it.
- What should be the product price to maximize total revenue?

Maximize Revenue

- Revenue = unit price × number of units
  - Unit price p (variable)
  - Number of units :
    - $n = 0$ at $p = p_0$
    - $n = n_0$ at $p = 0$
- p in the range $0 - p_0$
- n in the range $n_0 - 0$
- Revenue = $n \times p$

Numerical Solution

```
price = linspace(0, 1);
quantity = linspace(1, 0);

revenue = price .* quantity;

clf
plot(price, revenue)
grid on
xlabel('Price {\it [p0]}')
ylabel(Revenue')
```



Analytical Solution

- Revenue = unit price × number of units
  - Unit price p (variable)
  - Number of units :
    - $n = 0$ at $p = p_0$
    - $n = n_0$ at $p = 0$
  - $n = n_0 - p \frac{n_0}{p_0}$
- Revenue = $n \times p$

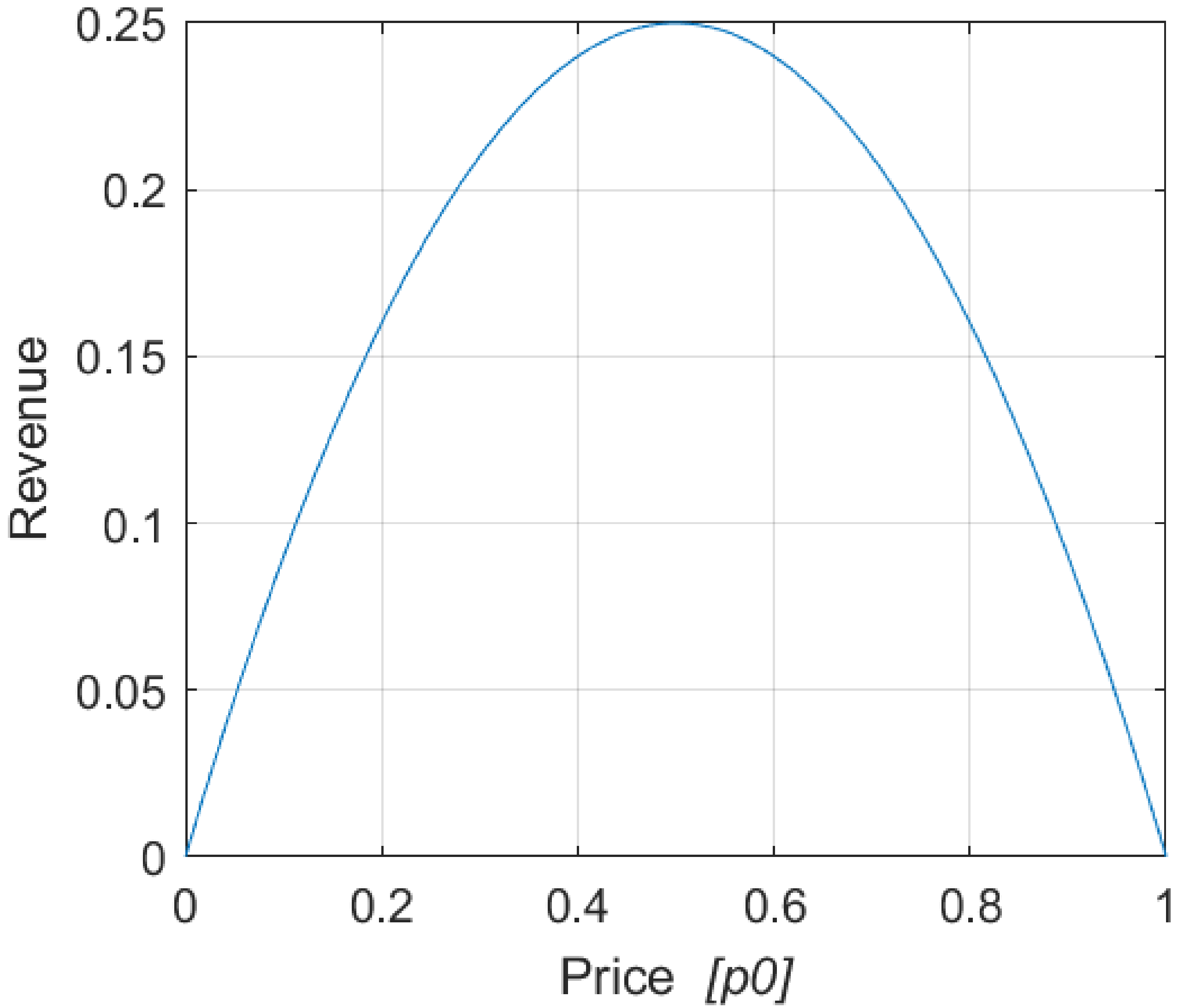
$$= p.n_0 - p^2 \frac{n_0}{p_0}$$

Analytical Solution

- Revenue = $n \times p$

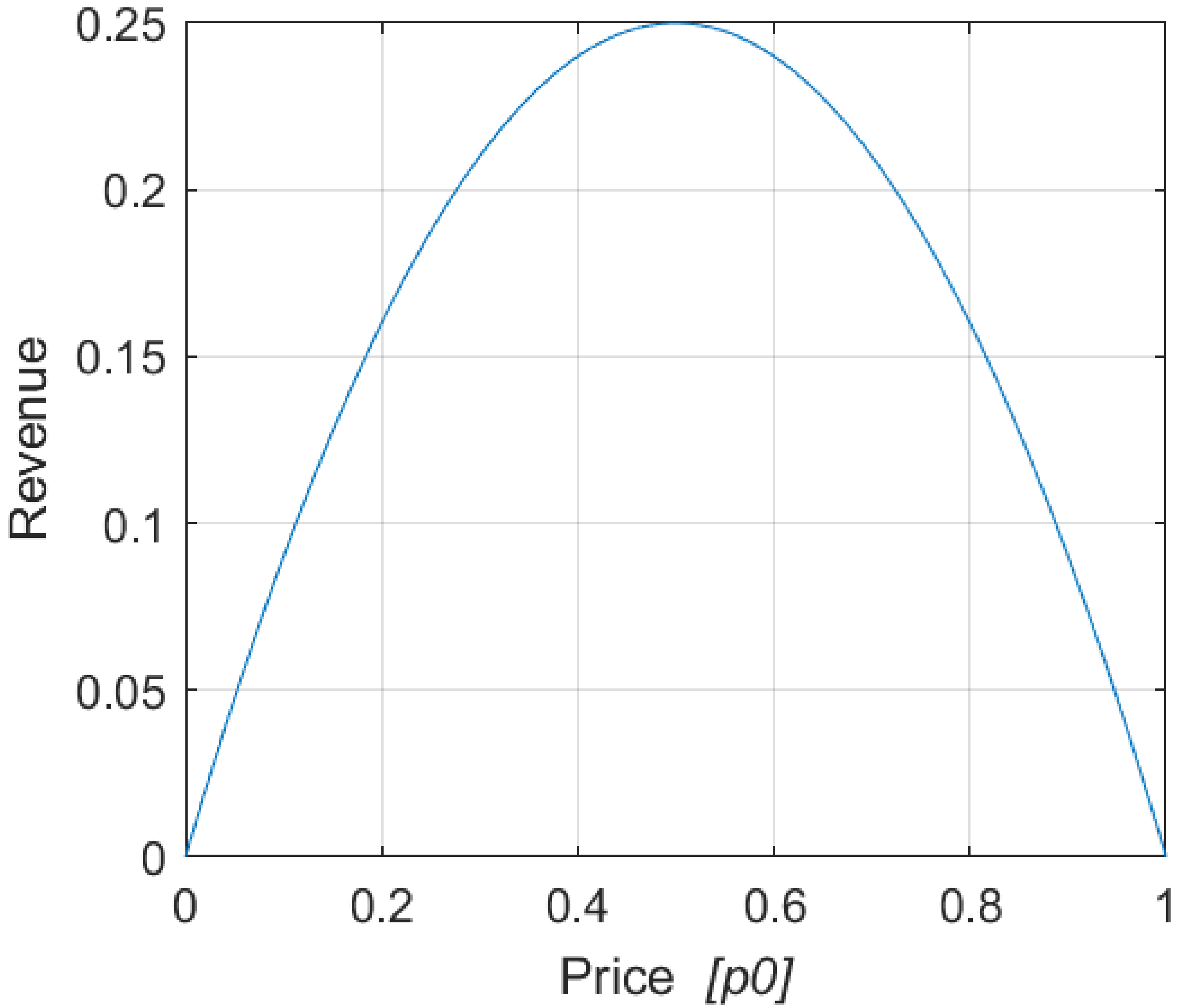
$$f(p) = p.n_0 - p^2 \frac{n_0}{p_0}$$

- $f' = n_0 - 2.\frac{n_0}{p_0}.p$
- $f' = 0$
- $p = n_0 \,.\, 0.5 \,.\, \frac{p_0}{n_0} = \frac{p_0}{2}$



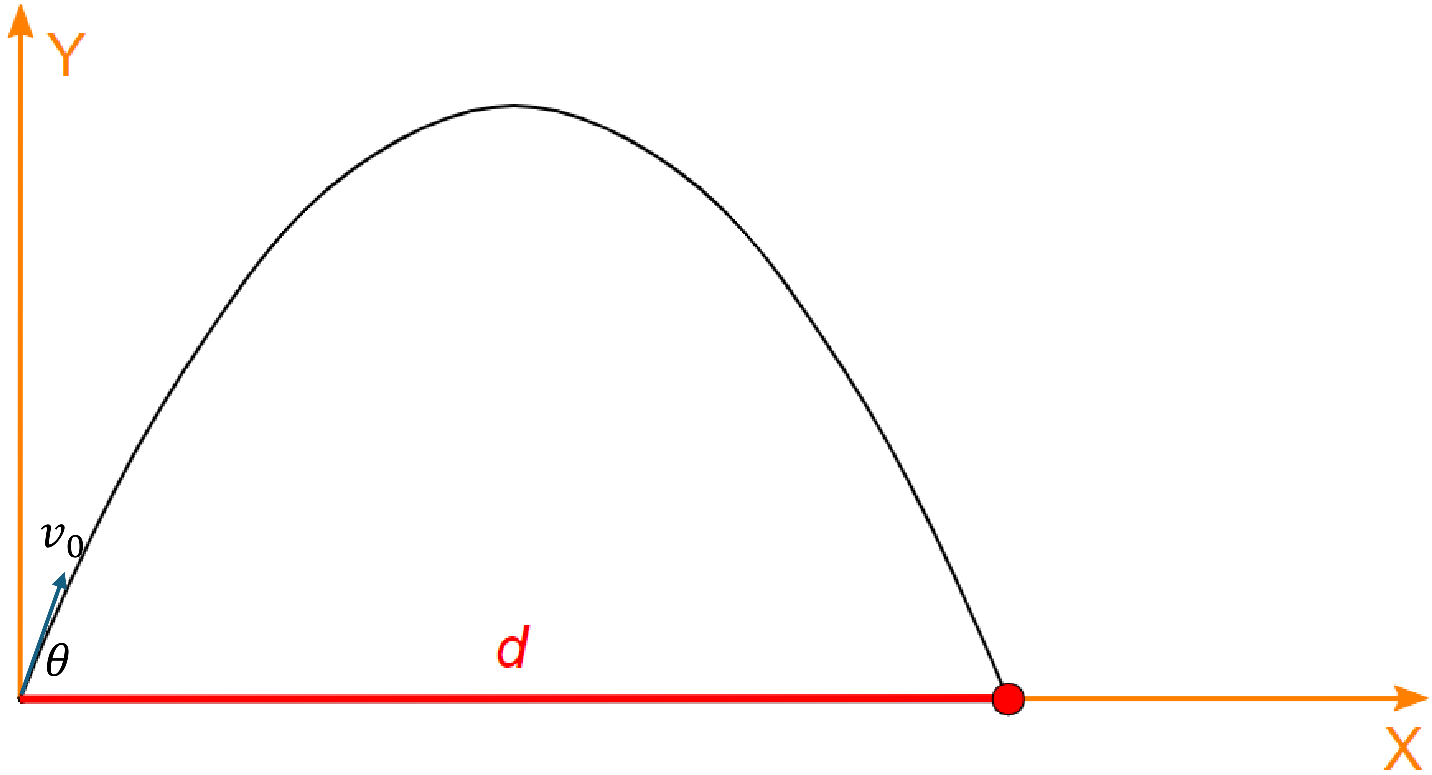
Analytical Solution

- $f(\mathrm{x})$
- $f'(\mathrm{x})$ maximum/minimum
- $f''(\mathrm{x})$:
  - < 0 maximum
  - > 0 minimum
  - $= 0$ inflection



Projectile Motion

- Acceleration
  - $a_x = 0$
  - $a_y = -g$
- Velocity
  - $v_{0x} = v_0 . \cos(\theta)$
  - $v_{0x} = v_0 . \sin(\theta)$
- Displacement
  - $x = v_0 . t . \cos(\theta)$
  - $y = v_0 . t . \sin(\theta) - 0.5 . g . t^2$
- Find maximum horizontal displacement



https://en.wikipedia.org/wiki/Projectile_motion

Projectile Motion

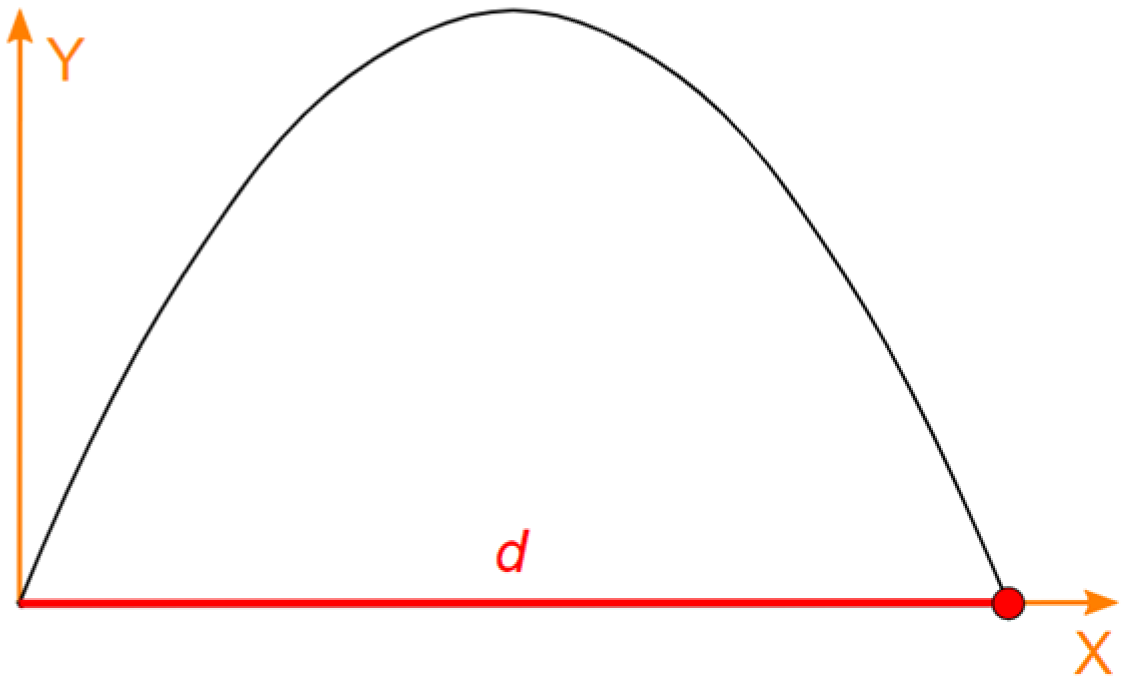
$$x = v_0.t.\cos(\theta)$$
$$y = v_0.t.\sin(\theta) - 0.5.g.t^2$$
$$y = 0$$
$$0.5.g.t^2 = v_0.t.\sin(\theta)$$

$$t = 2.v_0.\frac{\sin(\theta)}{g}$$

$$d = v_0.t.\cos(\theta)$$

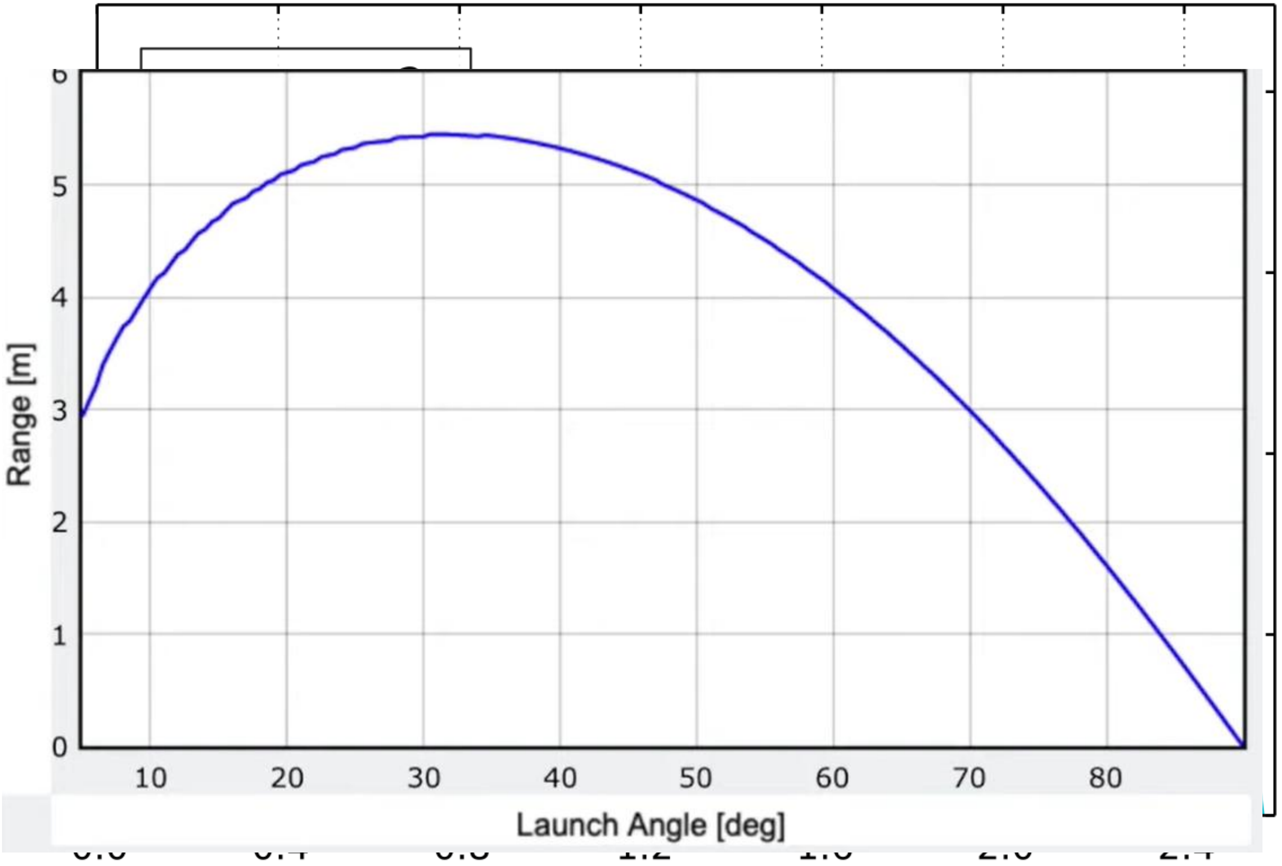
$$d = v_0.2.v_0.\frac{\sin(\theta)}{g}.\cos(\theta)$$

$$d = \frac{v_0^2}{g}\sin(2\theta)$$
$$\sin(2\theta) = 1 \Rightarrow \theta = 45°$$



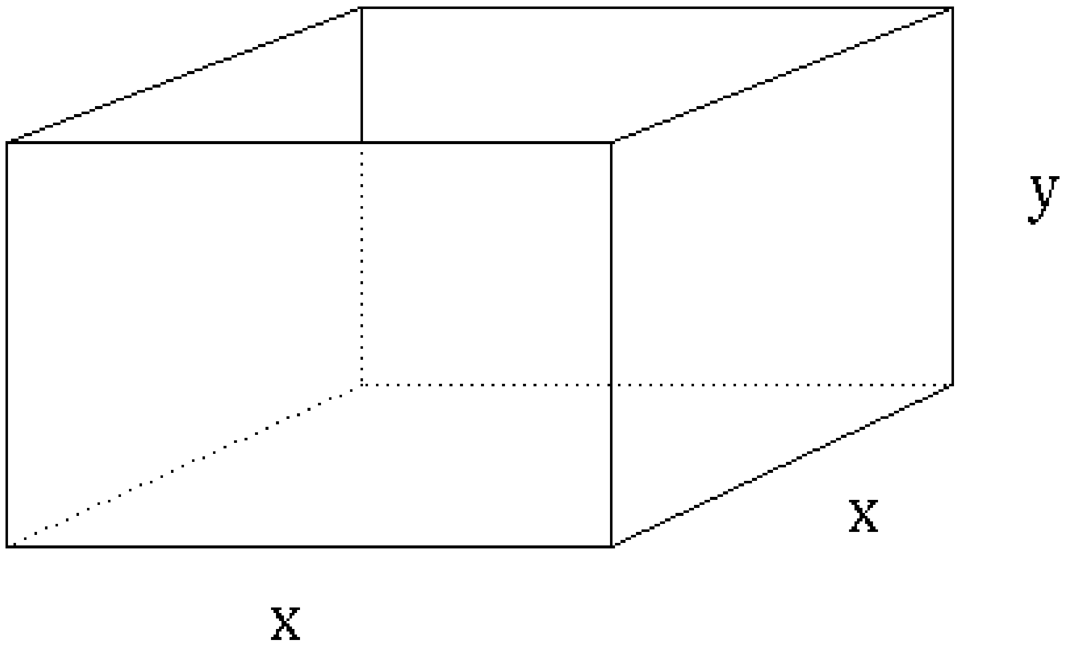
Projectile Motion

- Include air resistance
- Drag force proportional to the speed squared, or higher order
- Analytical solution
- Numerical solution



Examples

- An open rectangular box
- Square base
- To be made from 48 m^2 of material
- What dimensions will result in a box with the largest possible volume?



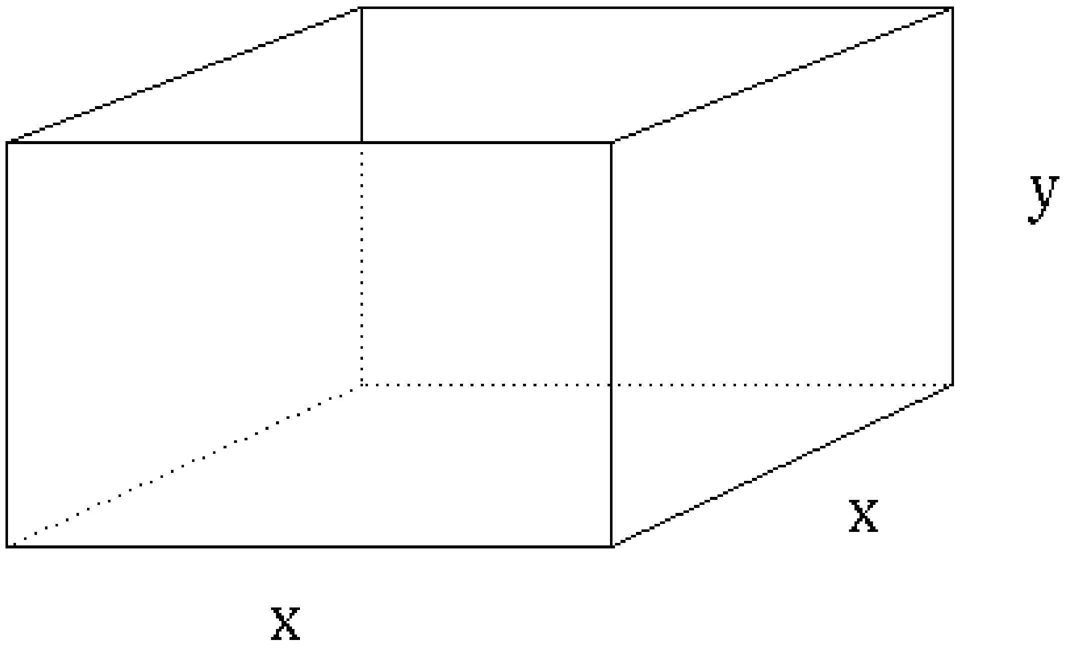
Examples

- Variables x and y
- Objective V(x, y)
- Constraint Area = 48

$$x^2 + 4.x.y = 48$$

$$y = \frac{48 - x^2}{4.x}$$

$$y = \frac{12}{x} - \frac{x}{4}$$



Examples

$$Volme = x^2 \times y$$

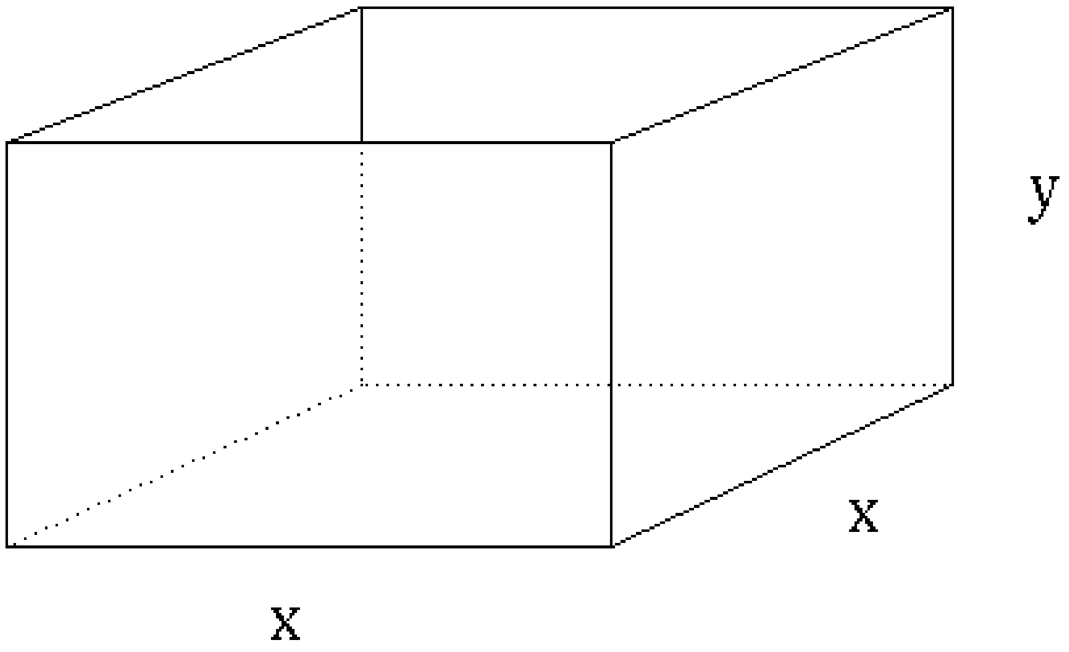
$$= x^2 \times \left(\frac{12}{x} - \frac{x}{4}\right)$$

$$= 12x \ - 0.25x^3$$

$$Volme' = 12 \ - 0.75x^2 = 0$$

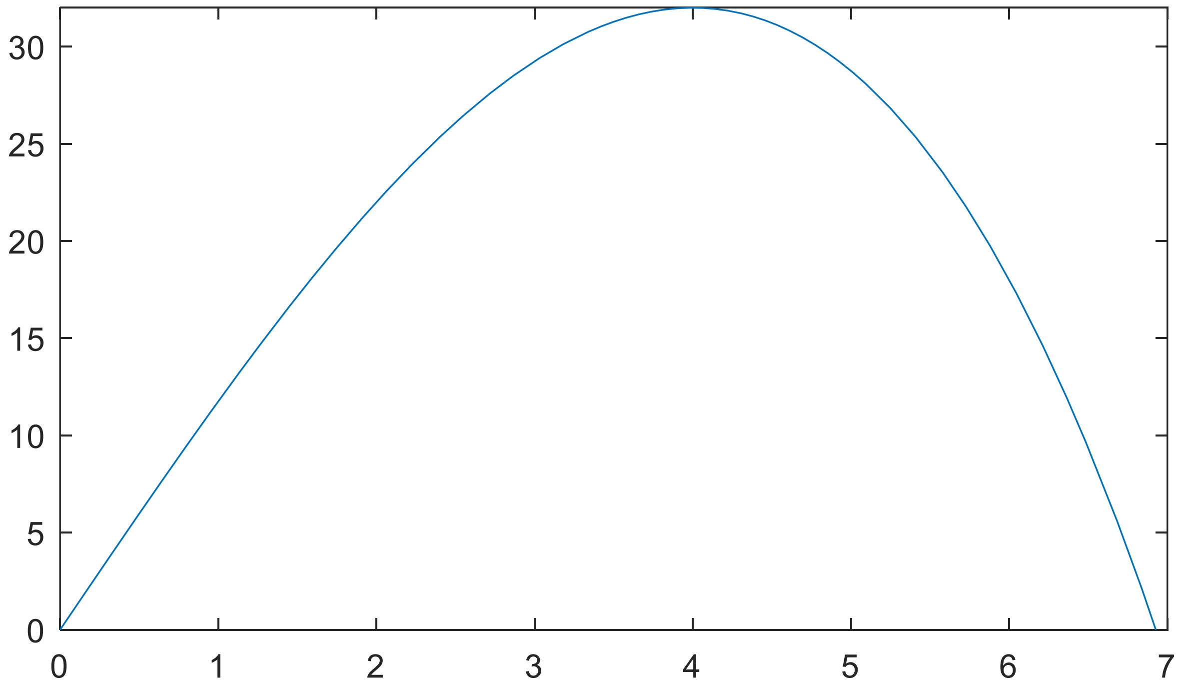
$$x^2 = 16$$

$$x = 4, \qquad V = 32$$



Examples

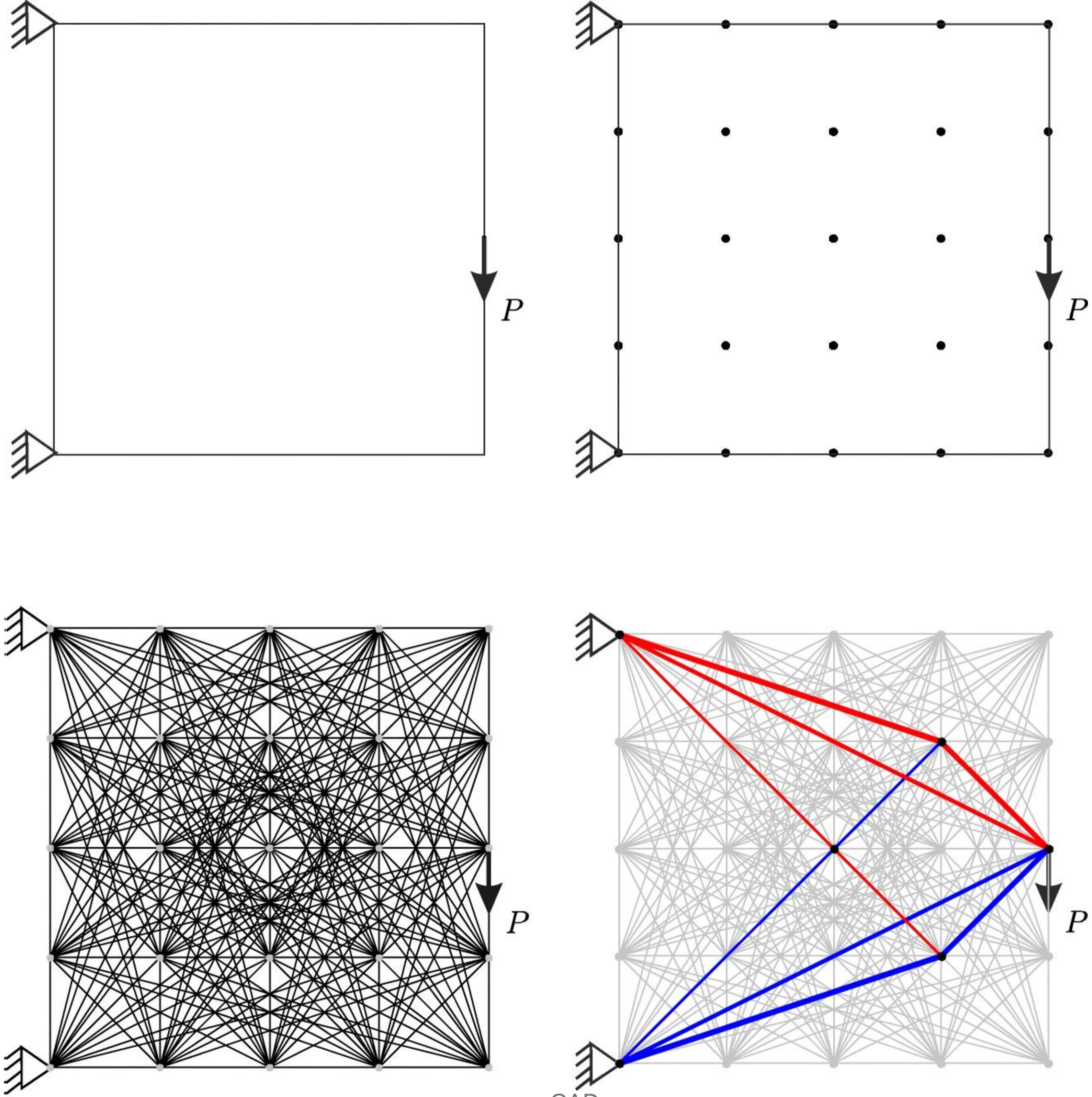
```
fplot(@(x) 12*x - 0.25*x.^3,[0 10])
```

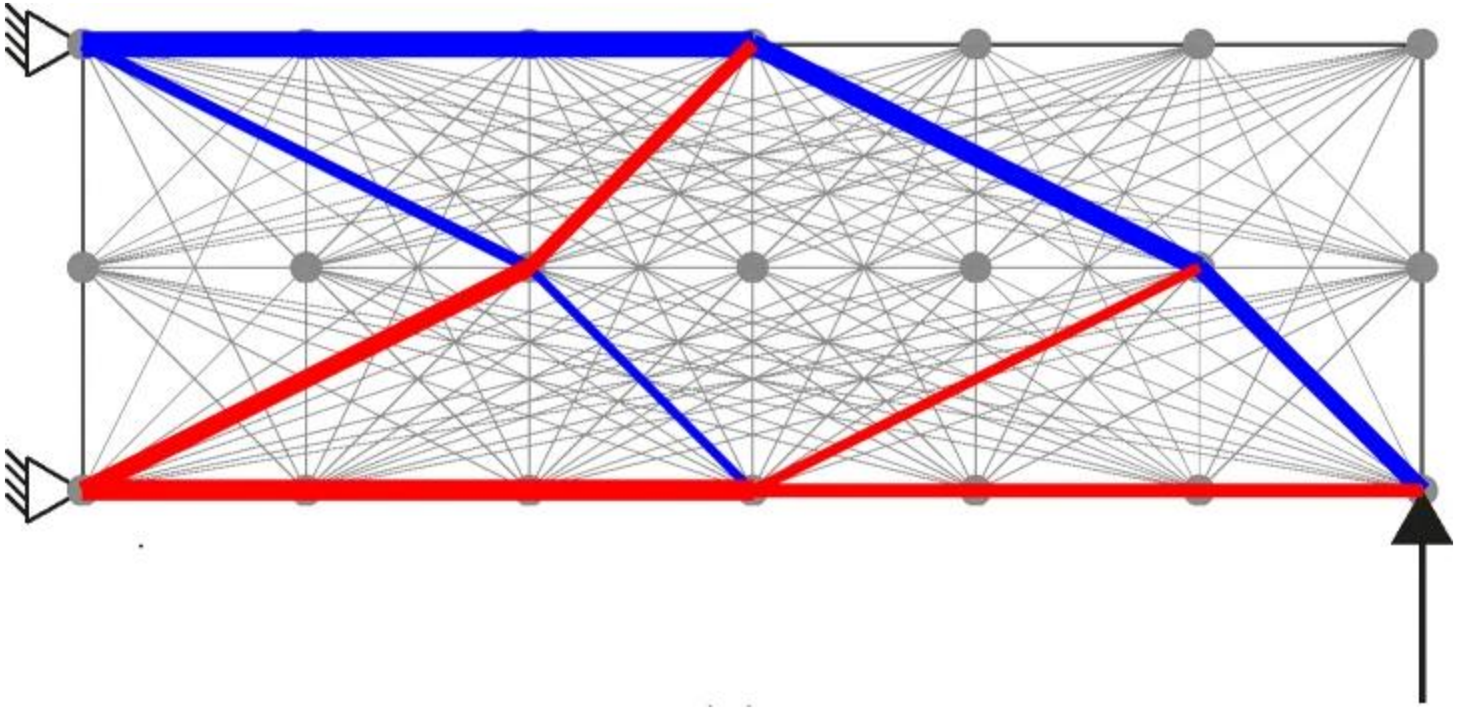



```
syms x
volume = 12*x - 0.25*x.^3;
df = diff(volume, x);
critical_points = solve(df == 0, x);
```

```
critical_points =

-4
 4
```

Layout Optimization


P
P
P
P



Layout Optimization

- Objective
  - Minimize Volume
- Variables
  - Area of truss members
- Constraints
  - Sum of all forces at each node = 0

Layout Optimization

https://layopt.com/