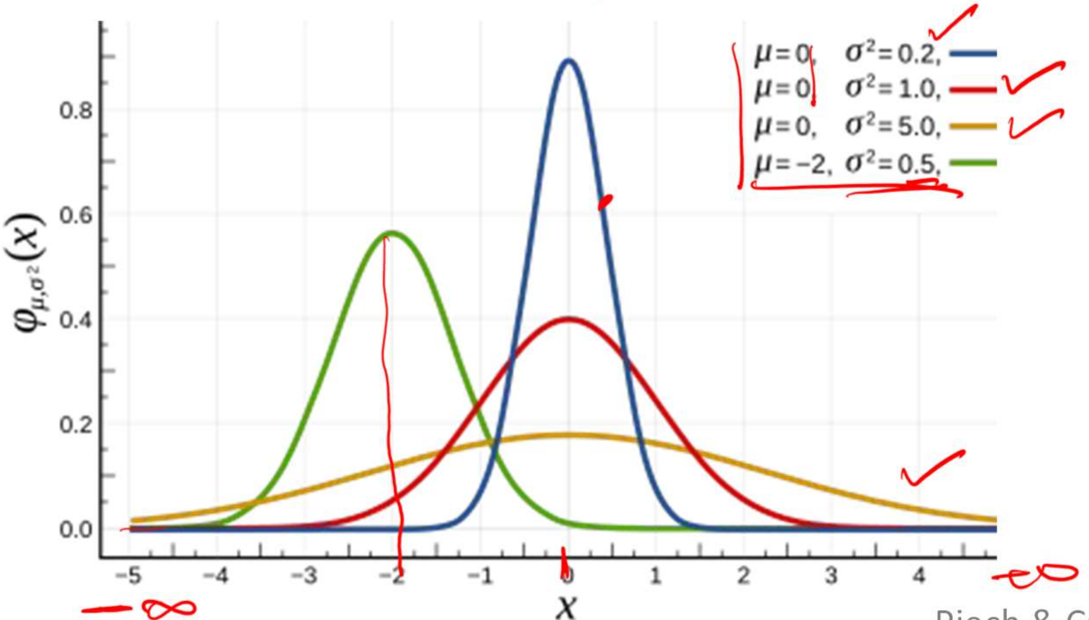
Normal (Gaussian) Random Variable

Support:
$(-\infty, \infty)$

mean variance

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

density



Normal (Gaussian) Random Variable

Support:
$(-\infty, \infty)$

$$X \sim \mathcal{N}(\underbrace{\mu}_{}, \sigma^2)$$

mean: μ — variance: σ^2

PDF:

$$f(X = x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Anatomy of a The Normal PDF

distance to the mean
(makes the PDF symmetric
around the mean)

$$f(X = x) = \underbrace{\frac{1}{\sigma\sqrt{2\pi}}}_{\text{a constant}} e^{\frac{-\overbrace{(x-\mu)^2}^{\text{distance to the mean}}}{2\sigma^2}}$$

a constant:
makes the integral
over all possible
outcomes sum to 1

...normalized by
the variance

Expected value of a normal distribution

verify that μ is the expected value of

$x \sim N(\mu, \sigma^2)$

$E([x-\mu]) = E(x) - \mu$

$$\int_{-\infty}^{\infty} (x-\mu) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \left. \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma\,(\sigma^2)} \right|_{-\infty}^{+\infty} = 0$$

$E[(x-\mu)] = 0 \Rightarrow E(x) = \mu$

Variance

$$E((X-\mu)^2)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^2 e^{-(x-\mu)^2/(2\sigma^2)}dx$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}y^2 e^{-(y)^2/(2)}dy=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(y)(ye^{-(y)^2/(2)})dy$$

u

v du

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left[\left(-ye^{-y^2/2}\right)_{-\infty}^{\infty}-\int_{-\infty}^{\infty}-e^{-y^2/2}dy\right]$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-y^2/2}dy=\frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi}=\sigma^2$$

$$\int udv=uv-\int vdu$$

$$\int ye^{-y^2/2}dy=-e^{-y^2/2}$$

Properties

If $X \sim N(\mu, \sigma^2)$ and if $Y = aX + b$, then   a & b are scalars.

Top property

$Y \sim N(a\mu + b, a^2\sigma^2)$ Y also follows a Normal Distribution.

Let F_Y be the cumulative density of Y

$F_Y = P(Y \le y)$ $f_Y = \frac{\partial}{\partial y} F(y)$

$F_X = P(X \le x)$ $f_X = \frac{\partial}{\partial x} F_X(x)$

Let $a > 0$

$P(Y \le y) = P(aX + b \le y) = P\left(X \le \frac{y-b}{a}\right) \Rightarrow \frac{\partial}{\partial y} F_Y(y) = \frac{\partial}{\partial y} F_X\left(\frac{y-b}{a}\right)$

$f_Y(y) = \frac{\partial}{\partial x} F_X(x) \frac{\partial}{\partial y}\left[\frac{y-b}{a}\right]$

$= f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}$

$f_X\left(\frac{y-b}{a}\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a} - \mu\right)^2}{2\sigma^2}}$

Properties

- Median = mean (why?)
- Because of symmetry of the pdf about the mean

$$f_X\left(\frac{y-b}{a}\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-(\mu a+b))^2}{2\sigma^2 \cdot a^2}}$$

- Mode = mean – can be checked by setting the first derivative of the pdf to 0 and solving, and checking the sign of the second derivative.

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a} = \frac{1}{\sqrt{2\pi}\sigma a} e^{-\frac{(y-(\mu a+b))^2}{2a^2\sigma^2}}$$

$$\Rightarrow Y \sim N(\mu a+b;\ a^2\sigma^2) \text{ if } a > 0$$

$a < 0$

$$F_Y(y) = P(Y \le y) = P(aX+b \le y) = P\left(X \geq \frac{y-b}{a}\right)$$

$$= 1 - F_X\left(\frac{y-b}{a}\right)$$

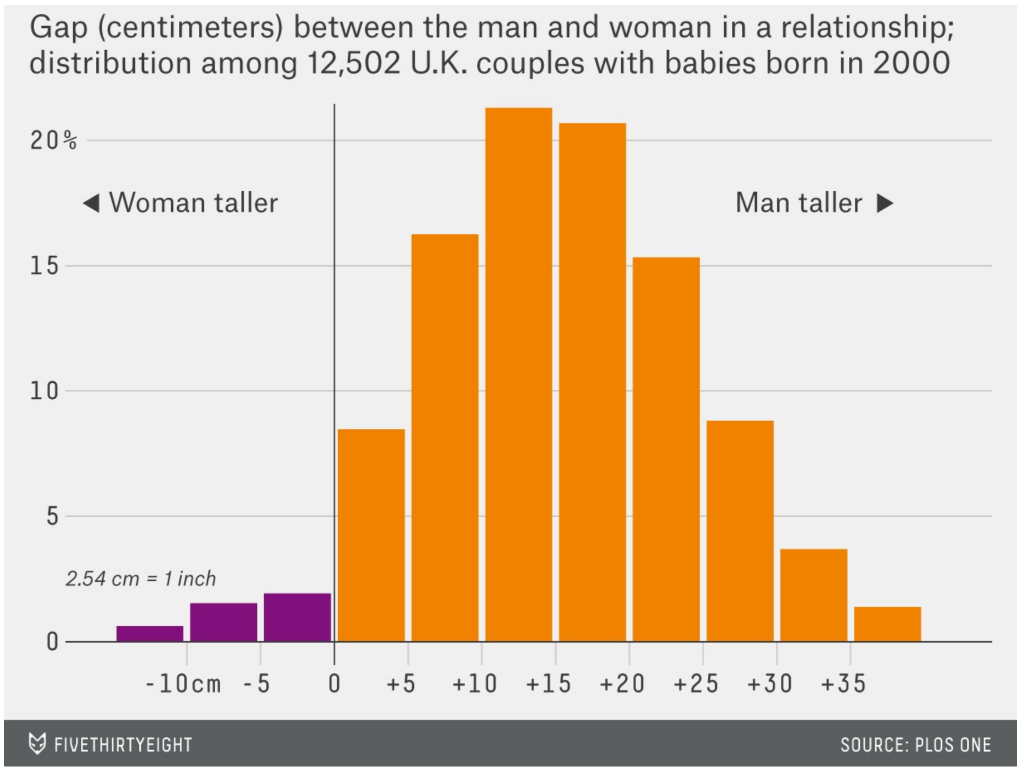
$$Y \sim N(a\mu+b;\ \sigma^2 a^2)$$

Carl Friedrich Gauss (1777-1855)

- German mathematician
- Sort-of invented the normal distribution
- Also astronomer, geologist, physicist
- Super influential in a lot of fields

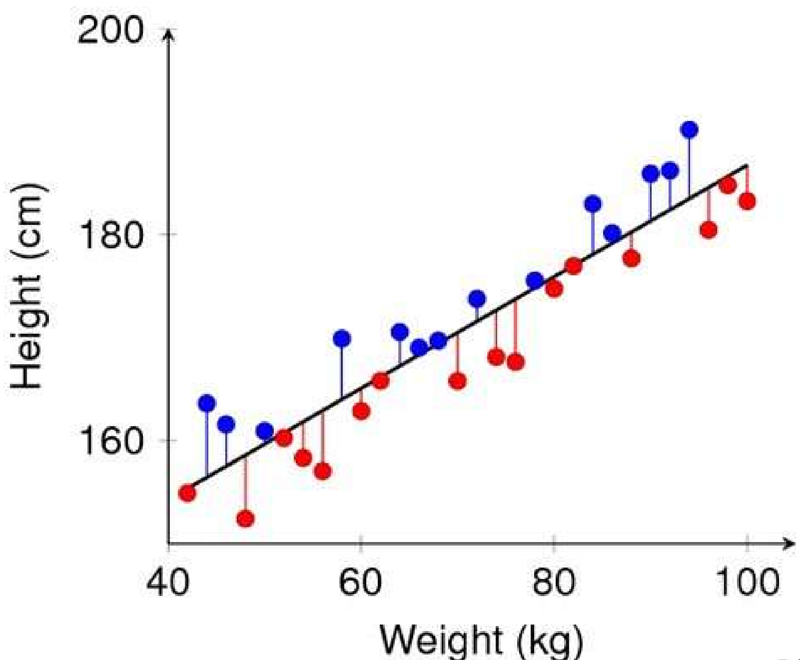
Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.



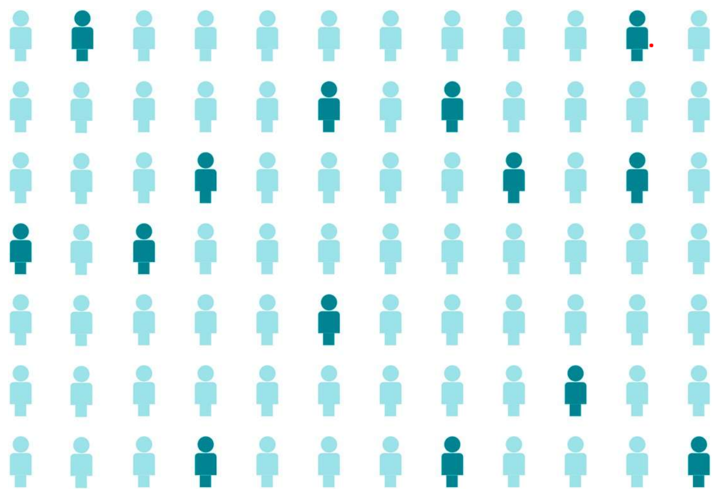
Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.
- A lot of noise in the world is Normal
    - E.g. random errors in measurements, residuals in linear regression



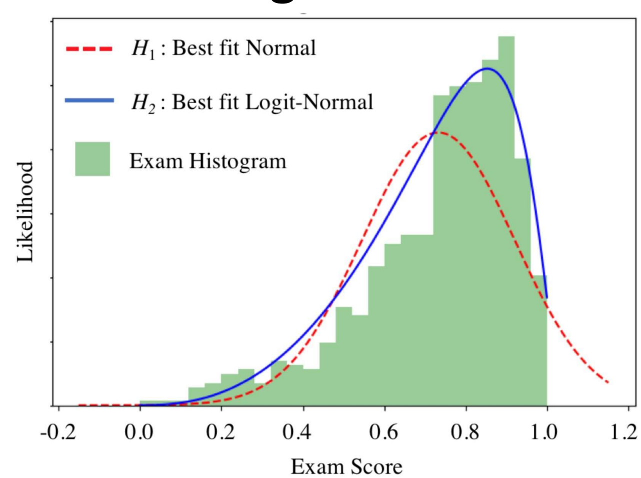
Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.
- A lot of noise in the world is Normal
  - E.g. random errors in measurements, residuals in linear regression
- The sum of many random variables often looks Normal (spoilers)
- Sample means are distributed normally – important for statistics



Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.
- A lot of noise in the world is Normal
  - E.g. random errors in measurements, residuals in linear regression
- The sum of many random variables often looks Normal (spoilers)
- Sample means are distributed normally – important for statistics
- Even things that aren't Normal might fit a normal-related distribution



Why the Normal?

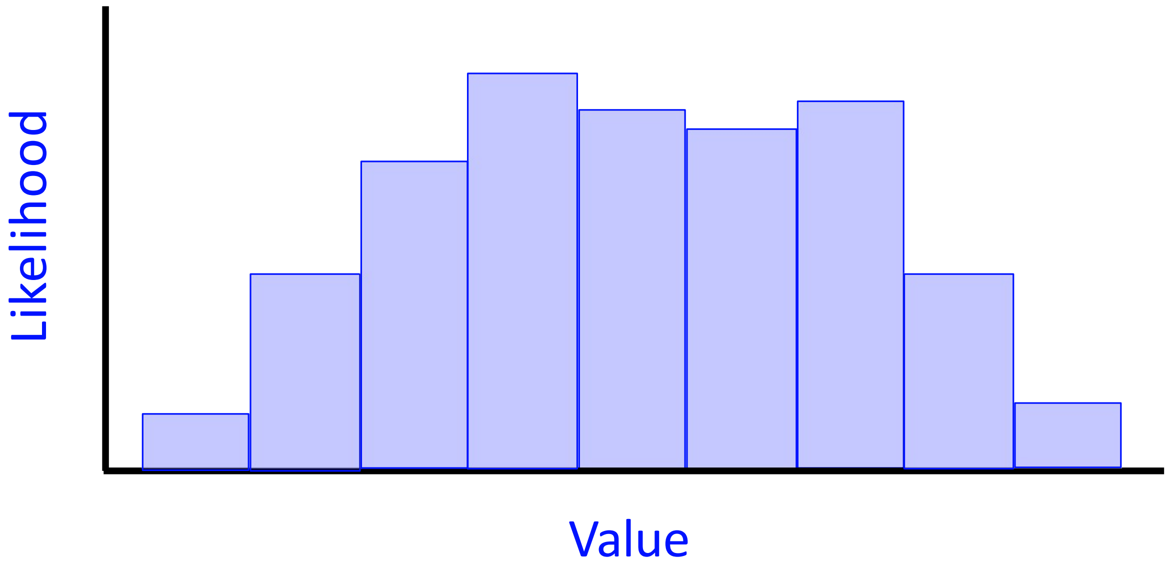
- Common for natural phenomena: human height, weight, shoe sizes, etc.
- A lot of noise in the world is Normal
  - E.g. random errors in measurements, residuals in linear regression
- The sum of many random variables often looks Normal (spoilers)
- Sample means are distributed normally – important for statistics
- Even things that aren't Normal might fit a normal-related distribution

People also just assume things are normally distributed a lot.

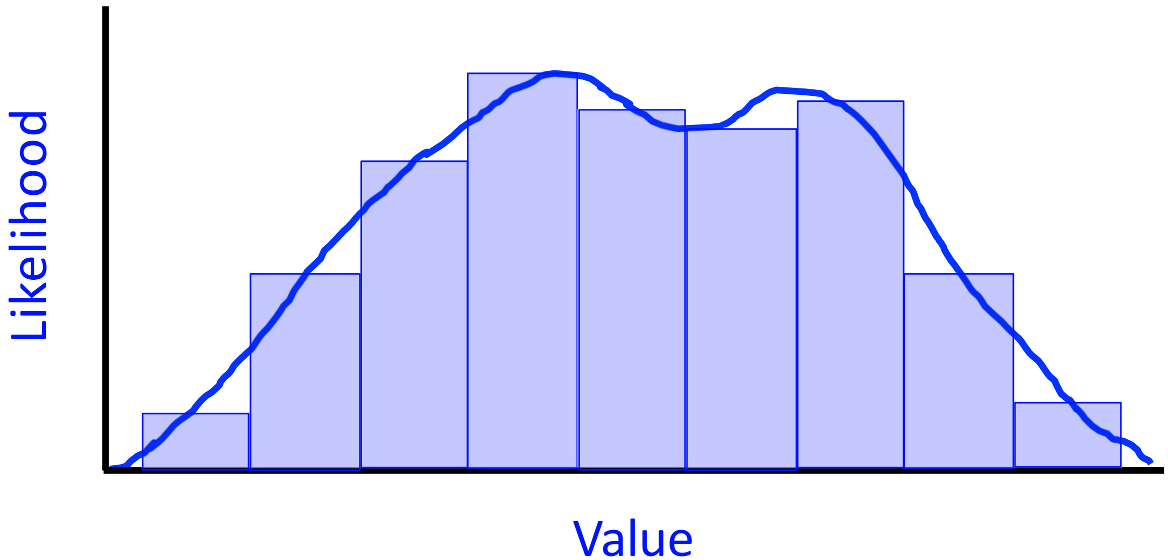
- They can do this in part because the Normal is so common
- But there's a deeper reason to it…


Ockham's razor
Shaving your hypothesis since 14th Century
Ockham
AMAZING!

When We Fit Models To Data, We Try To Keep It Simple

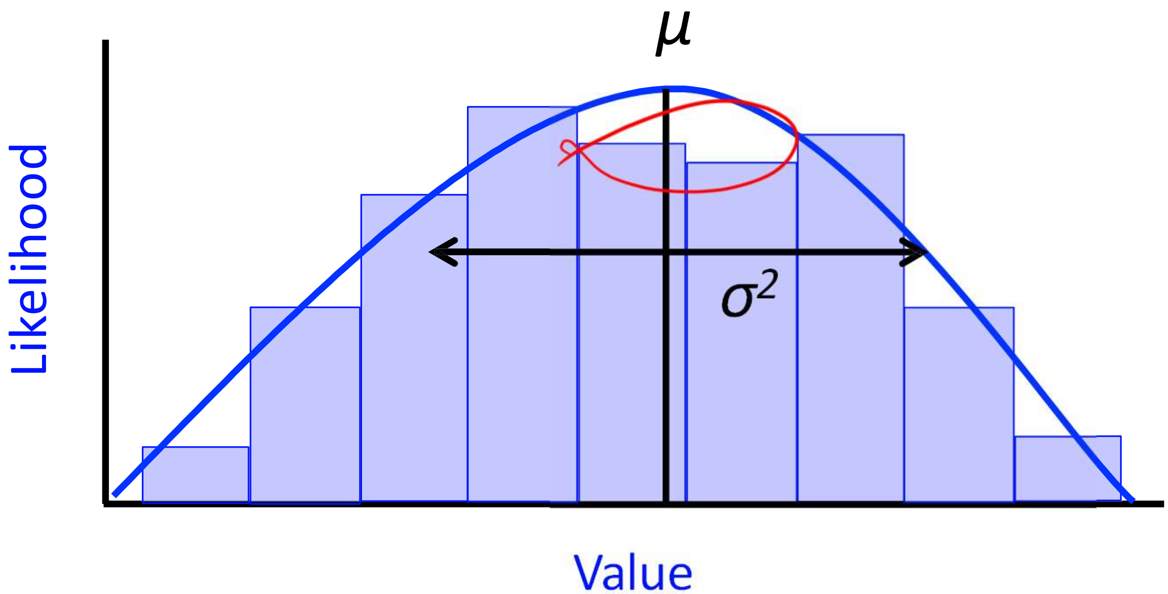


When We Fit Models To Data, We Try To Keep It Simple



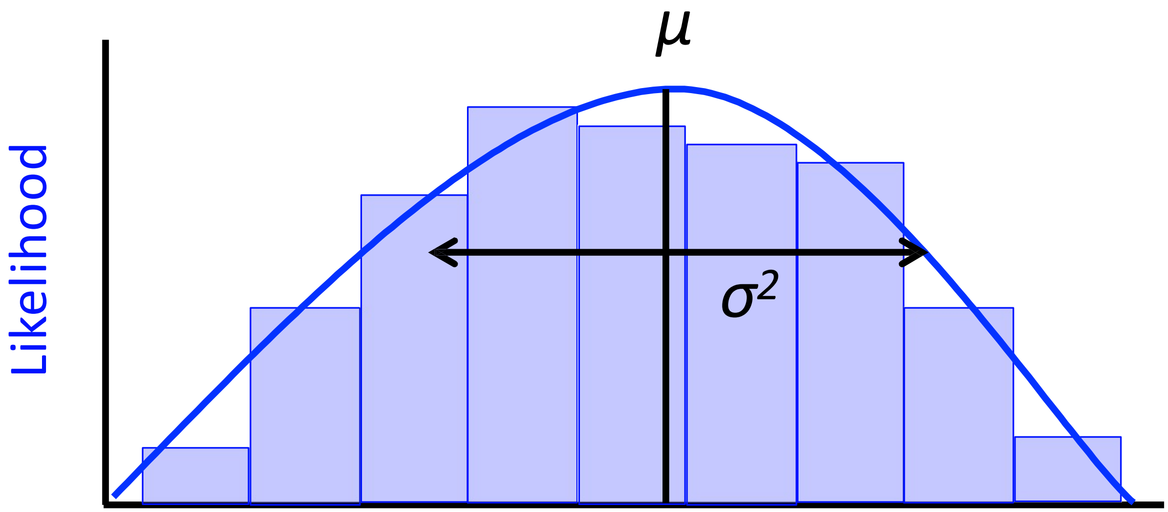
This curve fits the data well, but does it really represent the distribution?
Or is it “overfit”, so that the curve captures too much of the noise?

When We Fit Models To Data, We Try To Keep It Simple



This curve fits the data about as well, but appears to overfit less.
We could say that this simpler distribution makes fewer assumptions.
The formal concept for this idea is **entropy**

When We Fit Models To Data, We Try To Keep It Simple



For a fixed mean and variance, the unique distribution that maximizes the entropy is the normal distribution.

https://medium.com/mathematical-musings/how-gaussian-distribution-maximizes-entropy-the-proof-7f7dcb2caf4d
https://statproofbook.github.io/P/norm-maxent.html