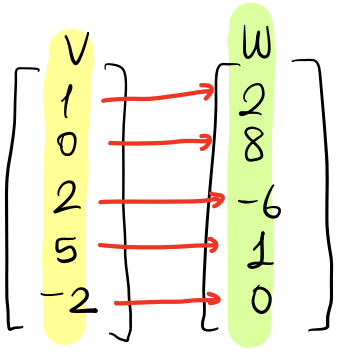
1) Dot Product (2 vectors, 2 matrices, 2 tensors, 2 signals or 2 images)



$$V^T W = 1*2 + 0*8 + 2*-6 + 5*1 + -2*0$$

Dot product is only possible when the dimensions are same.

$$nd\ V = nd\ W$$

Dot Product in 2D (Matrix) has the same logic

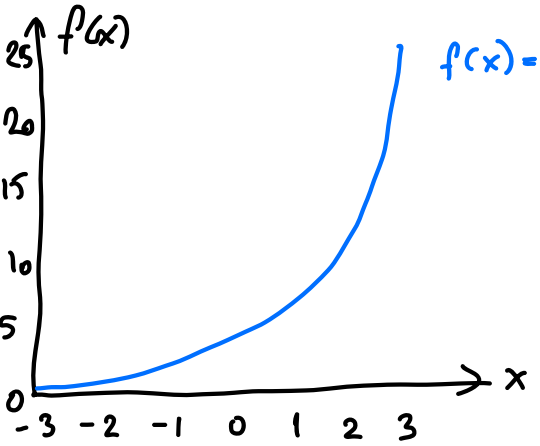
2) Matrix Multiplication

M rows, N columns $\rightarrow$ $M \times N$

$$[M \times N] \times [N \times K] = [M \times K]$$

$$\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0.a + 1.c & 0.b + 1.d \\ 2.a + 3.c & 2.b + 3.d \end{bmatrix}$$

3) Softmax



$$\sigma i = \frac{e^{z_i}}{\sum e^z}$$

$z = \{1, 2, 3\}$

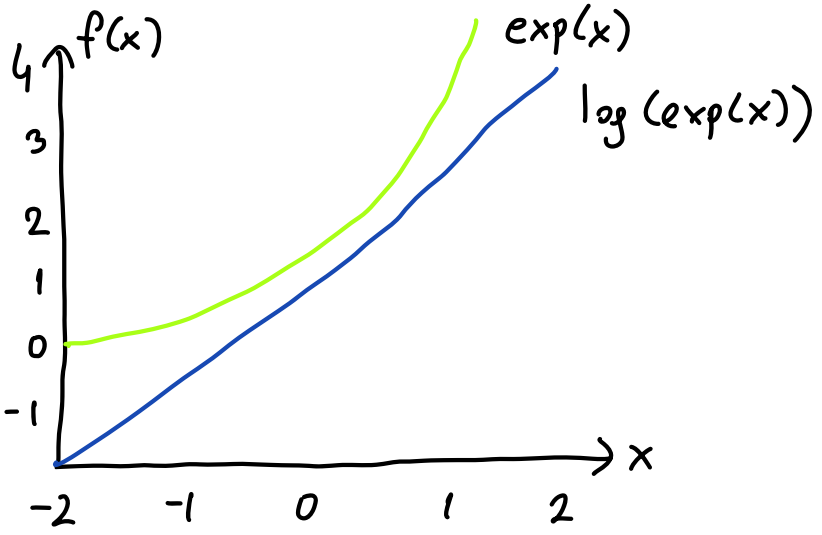
$e^z = \{2.72, 7.39, 20.01\}$

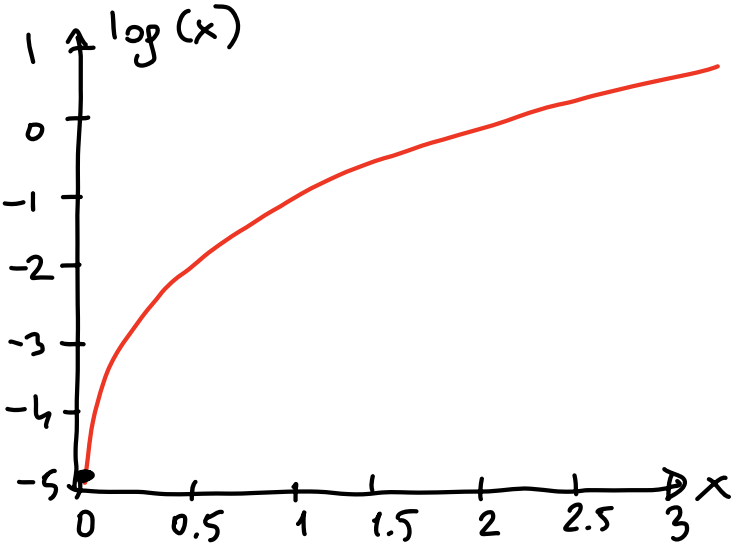
$\sum_{e^z} = 30.19$

$\sigma = \{.09, .24, .67\}$

4) Logarithm : The inverse of "Natural Exponential"

"i^x"





log is a "monotonic" function of x

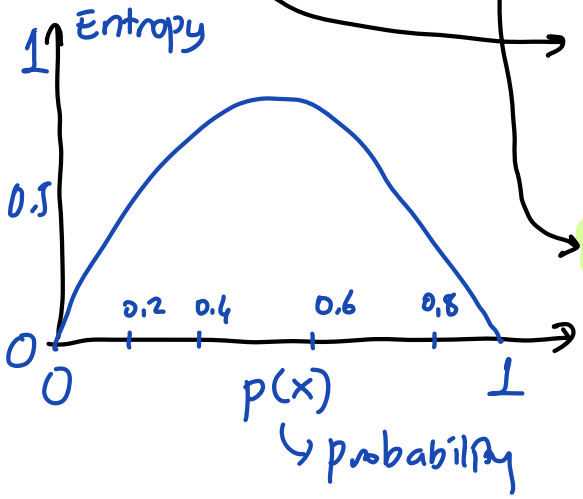
↗ ↗

↙ ↙

x log(x)

"nonmonotonic" = ∿∿

5) Entropy and Cross-Entropy



$H(p) = -\sum p \log(p)$ 1 probability distribution

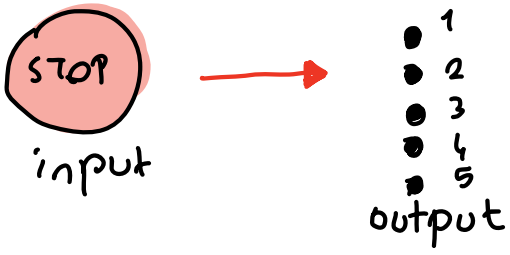
$H(q)$ $-\sum p \log(q)$ 2 " "

↳ $q = (1-p)$ " "

: Because it's rather yes 1 or no 0

6) min finds the smallest value, arg-min finds where the smallest value occurs.
max finds the highest value, arg-max finds " " " "

$\min\{1, 0, -1, 2\} = -1$; $\arg\min\{1, 0, -1, 2\} = 3$



1) Squirrel : p=0
2) speed limit sign : p = 0.5
3) Stop sign : p = 0.8
4) Tomato : p = 0.1
5) Car : p = .05

argmax M(y) = 3

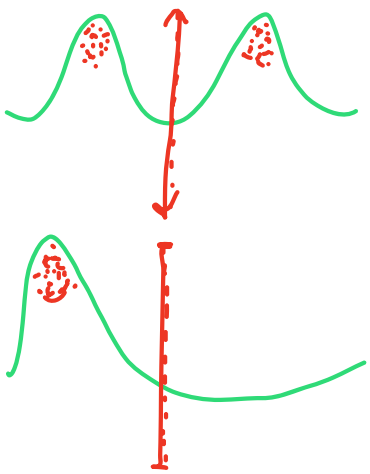
np.argmax(pvec) → 2 (because python index starts with "0".

7) Mean and Variance → (varies...)

↙ average
sum / len(n)

↘ $\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ "mean center"

Failure Scenario



8) Random Sampling – Sampling Variability

→ Take lots of samples. (true population mean)

8) Seed Function (Reproducible randomness through SEEDING)

When you set a random seed, it ensures that the random numbers generated by the algorithm are the same everytime you run the code. This is important for debugging + understanding the model behavior as it allows you to recreate the exact same conditions in diff. runs.

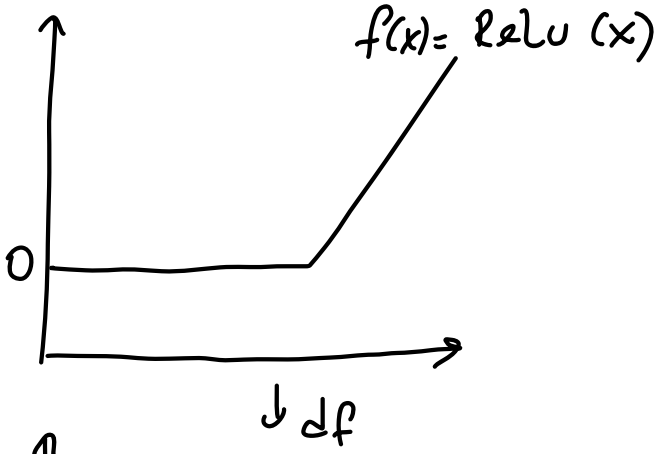
9) T-TEST

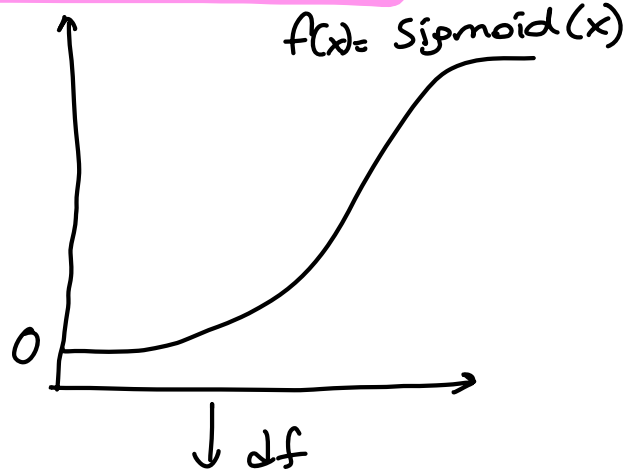
H_A = 🔵 H_B = 🟢 $H_0 = H_A == H_B$ (equally well)

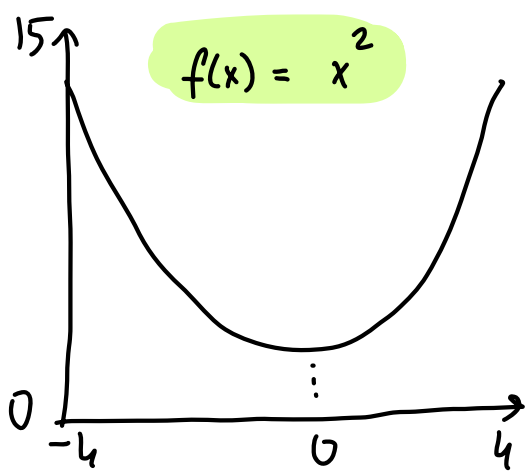
$$t_K = \frac{\bar{x} - \bar{y}}{s/\sqrt{n}}$$

T-test = a statistical hypothesis test used to determine if there is a significant difference between the means of two groups or populations. Used generally when you have a small sample size and want to assess whether the means of 2 groups are different.

10) Derivatives → find the smallest error for the best solution.







→ $df = 2x$