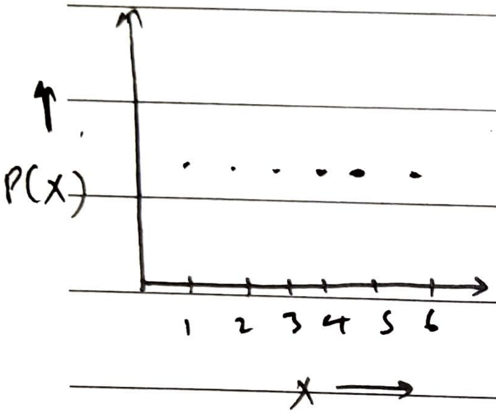
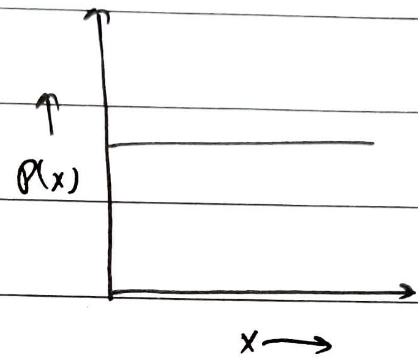
Probability Distributions & statistics

Ex: Throwing a die, $X = \{1, 2, 3, 4, 5, 6\}$





X = Discrete

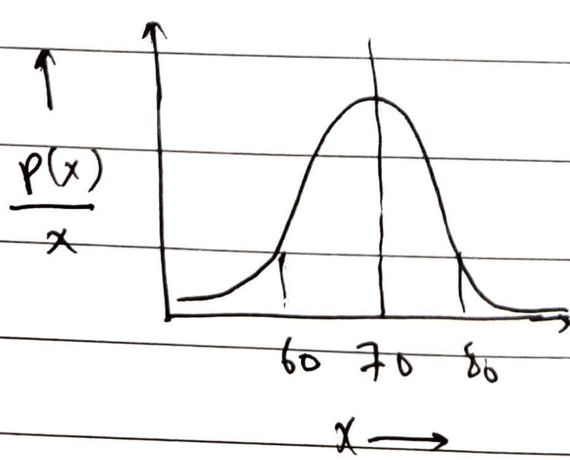
X = Continuous

Normal Distribution :

Here we take 2 factors :-

1. $\mu \rightarrow$ Mean
2. $\sigma \rightarrow$ Standard Deviation

Ex: Physics exam, N students, Avg = 70, Deviation = 10



If, $\frac{P(x)}{x}$ = say $f(x)$

$f(x)$ is <u>Probability density function</u>

If $f(x)$ = say 0.01 at $x = 175$

i.e for a small range, say 174.9 to 175.1, the probability of a person being found in this range is:

$= 0.01 \times$ width

$= 0.01 \times (175.1 - 174.9)$

$= 0.01 \times 2 = 0.002$

So, 0.2% chance of finding a person there.

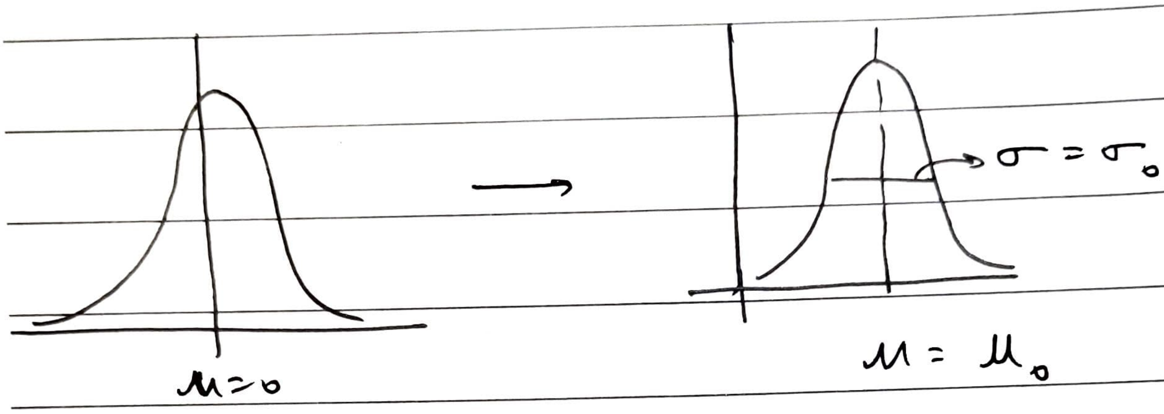
<u>Why do we need Normal distribution?</u>

* We use it when, say we don't have data.

Because it approximates many natural (real life) phenomenons.

✲ Standard Normal Distribution:

$\mu = 0$ & $\sigma = 1$



$$\boxed{X = (X * \sigma_0) + \mu_0}$$

✲ Multivariate Normal Distribution:

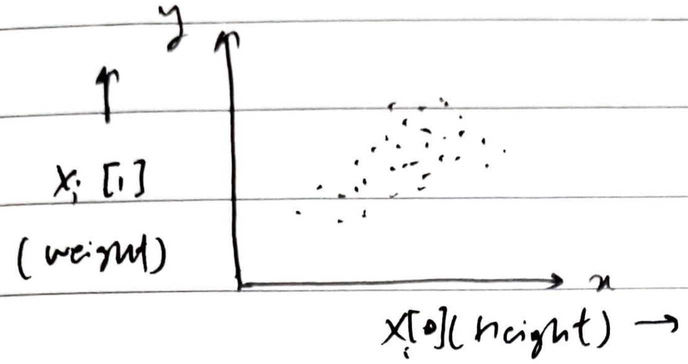
Dataset of monkeys:

$X_1 = [75\text{cm}, 20\text{kg}]$, $X_2 = [72\text{cm}, 21\text{kg}]$

✲ $m = 2$ = features

$$X = \begin{bmatrix} \cdots \\ \cdots \end{bmatrix}_{n \times m}$$

$n \rightarrow$ monkeys $\qquad m \rightarrow$ features



Here, you can see a co-relation between x & y. It is a positive corelation i.e. if height increases, weight will also increase.

☆ This is also a normal distribution with parameters, ① $\mu = [\mu_x, \mu_y]$

μ_x → avg height

μ_y → avg weight

② Spread is defined by std deviation (varience) $= [\sigma_x \;\; \sigma_y]$

③ Covarience matrix : $= \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix}$

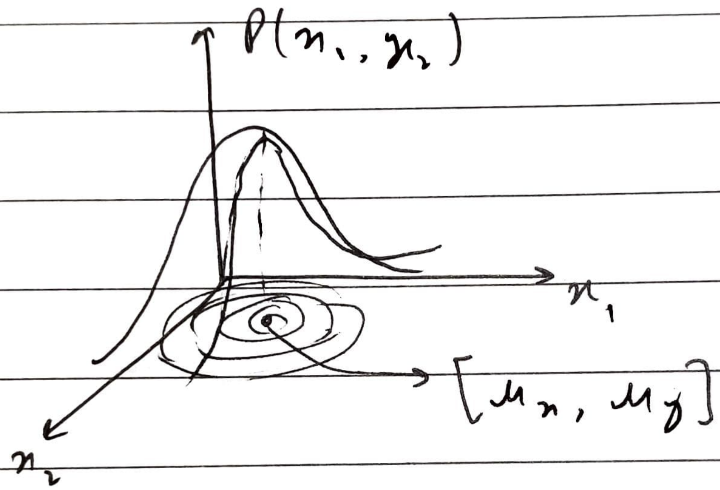
$$\sigma(x,y) = \frac{1}{(n-1)} \times \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$= \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix}$$

$$= \begin{bmatrix} \mathrm{var}(x) & \sigma_{xy} \\ \sigma_{yx} & \mathrm{var}(y) \end{bmatrix}$$

$$= \begin{bmatrix} var(x) & cov(x,y) \\ cov(x,y) & var(y) \end{bmatrix}$$

Probability density graph: (for 2-D)



$X = [x_1 \;...\; x_m]$

$\mu = [u_1 \;...\; u_m]$

column vectors

✡ **Probability density ~~distribution~~:**

$$f(x_1, x_2, \ldots, x_m) = \frac{e^{\left[-\frac{1}{2}(X-\mu)^T \Sigma^{-1} (X-\mu)\right]}}{\sqrt{2\pi^k |\Sigma|}}$$

X & μ are m-dimensional column vectors.

$$\Sigma = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix}$$

$m =$ no. of dimensions / features