Lec 1

Genral Review of Basic Statistic

Mean, Variance and Standared Deviation

for a one-dimensinal variable X_i

where:

$$X_i = [X_1, X_2, X_3 \,\text{------}\, X_n]$$

i: can be expressed as time or space index.

* The Mean of a certain Data is:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Mean: is the average of Data (Numbers) ~~whe~~ which calculate the "central" value of a set of data.

Variance: Second moment about the mean.

$$Var = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$

Standard Deviation: or (Avarge error)

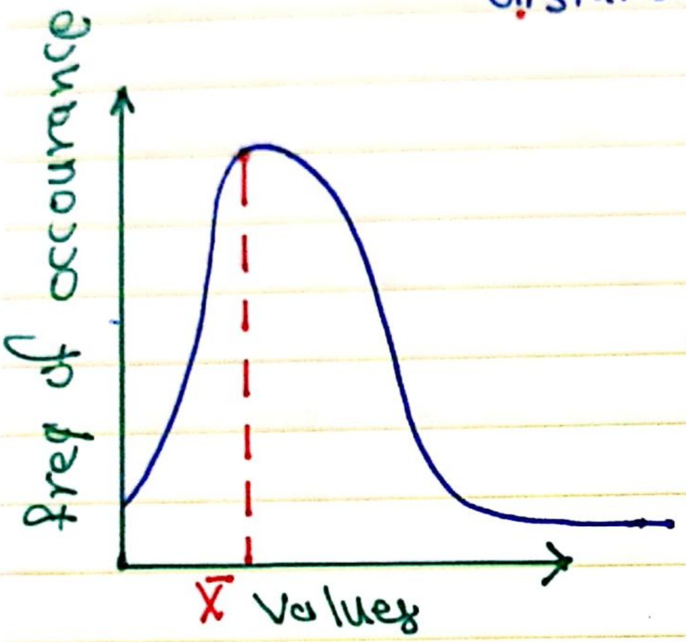
The standard Deviation of a Set values is the average error of a set values or total amount of error distributed evently to each value

$$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{n}}$$

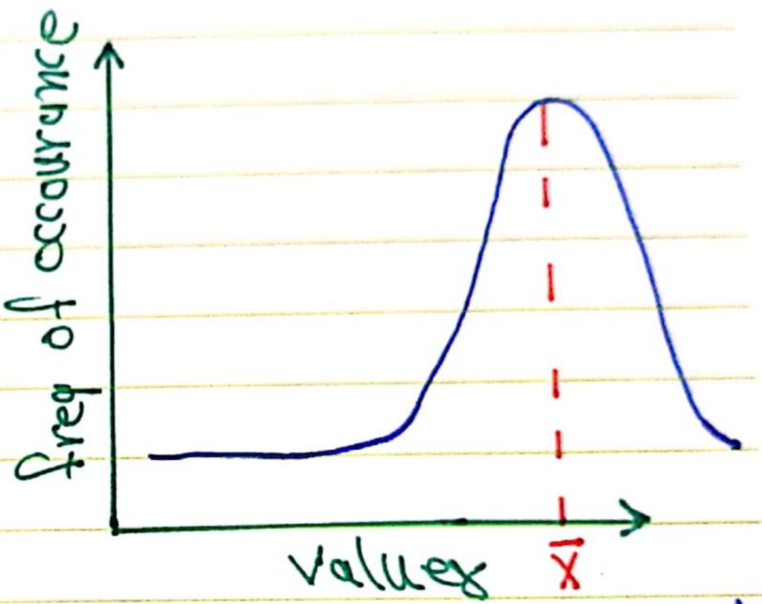
* Deviation from the mean

$$X_i' = X_i - \bar{X}$$

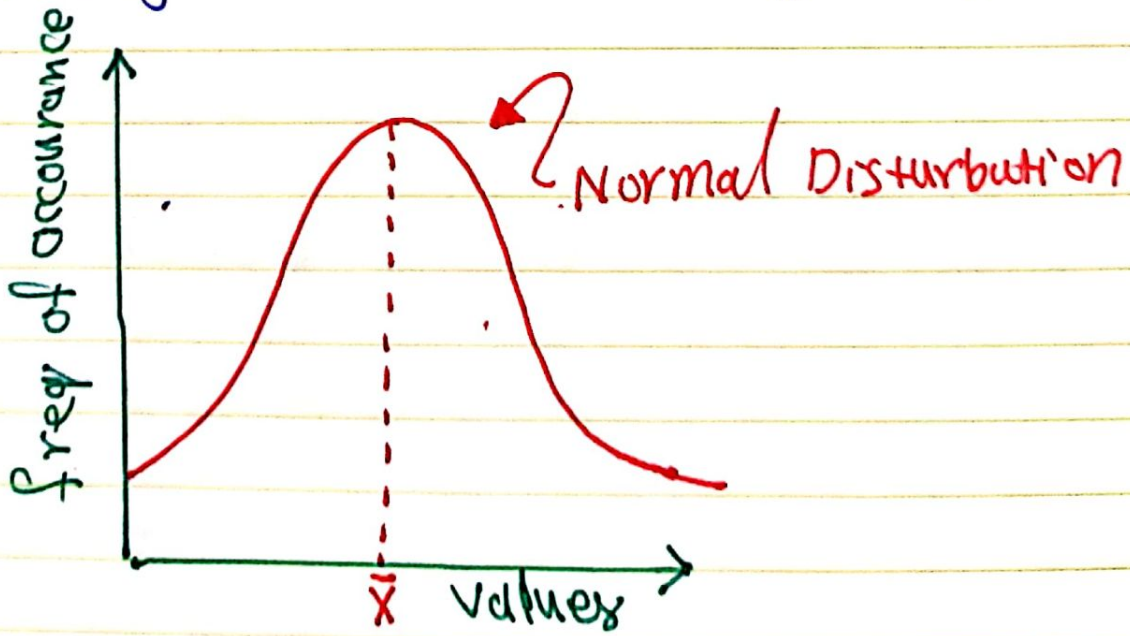
Skewness : indicates degree of a symmetry of the disturbution about the mean.



Positively skewed



Negatively skewed



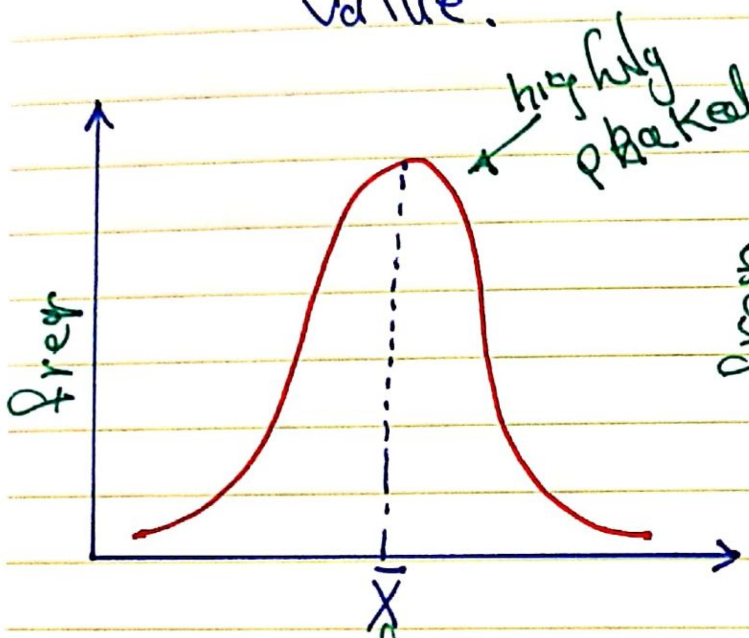
"Zero Skewed"

For many Types of atmospheric data have a normal disturbutio like (temperature, solar radiation ....)

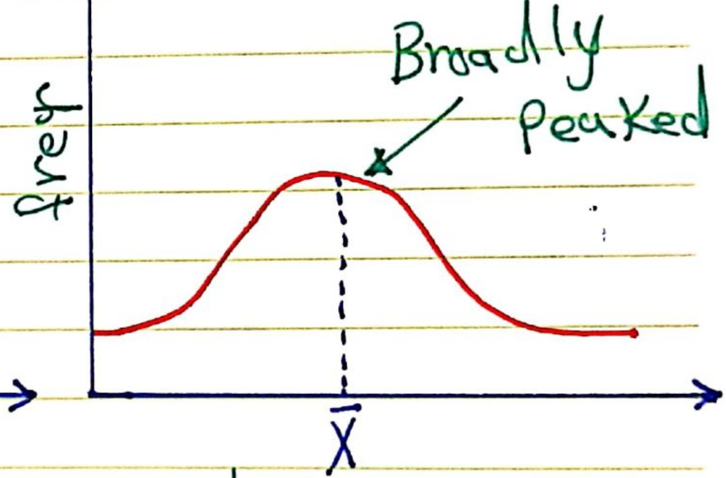
i.e have a zero sekwenss

where as othe atmospheric or meteorological data such as percipittion which typically has a postively skewed frequently disturbution, especially in arid climates like Iraq and suadi arabia, which characterized by relatively low mean Pericpitio but a few intense rain events.

Kurtosis: indicates to degree to which precipitation is peaked near mean value.



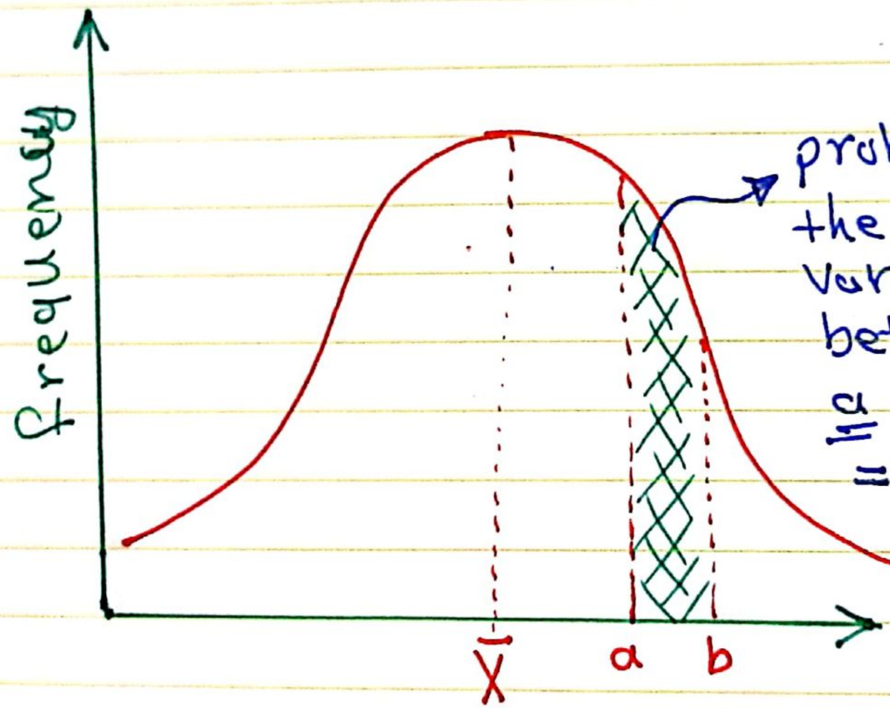
High Kurtosis



Low Kurtosis

Normal Distrubution

Can be express the probability that a selected variable fall between $\underline{\underline{a}}$ and $\underline{\underline{b}}$ on a frequency distribution plot, or histogram.



probability of the selected variable between $\underline{\underline{a}}$ and $\underline{\underline{b}}$

$$= \int_a^b f(x)\,dx$$

f(x): is referred to as the probability density function.

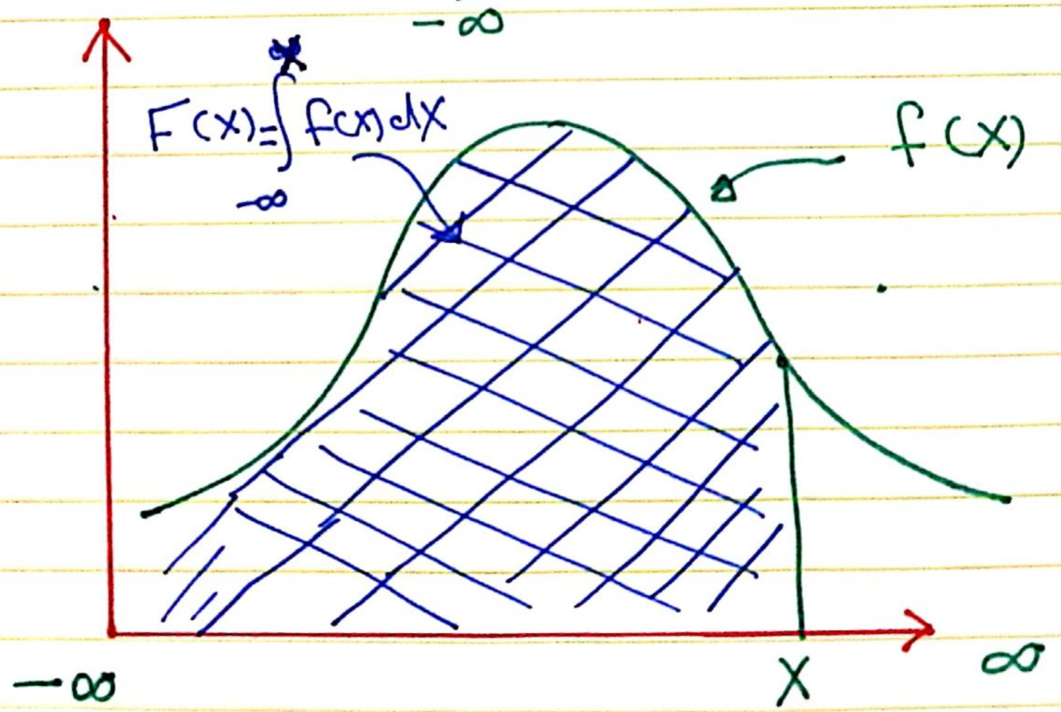
considering the whole range of values the probability must be equal to 1.

i.e

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

Cumulative distribution function (f(x)) can be defined as a variable of the probability that a variable assumes a value less than X.

$$f(X) = \int_{-\infty}^{\infty} f(x)\,dx .$$



The most common distribution which fits geographysical data is the normal or Gaussian distribution. More commonly Known as the "bell curve".

for M = population mean

σ = Standerd deviation

The bell curve or normal distribution is:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{X-M}{\sigma}\right)^2\right]$$

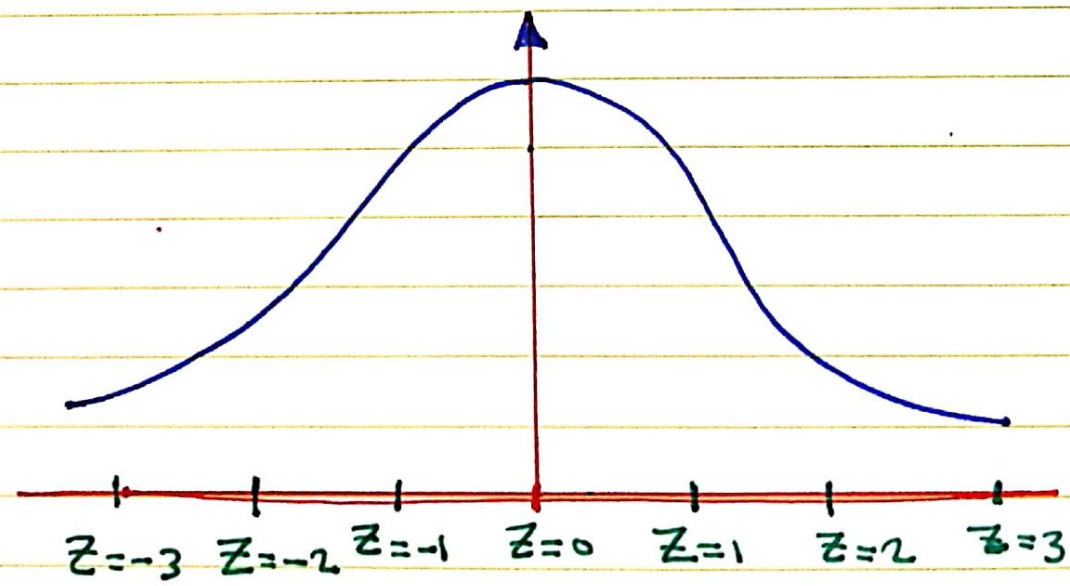
if we consider a normilized variable

$$Z = \frac{X-M}{\sigma}$$

$$F(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} Z^2\right]$$

Cumulative Probability distribution:

$$f(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} Z^2\right] dZ$$



Most geophysical fields are normally distributed, but ~~it~~ it is always a good idea to plot a histogram along with the normal distribution just to make sure it is a good assumption for whatever data set you're working with.