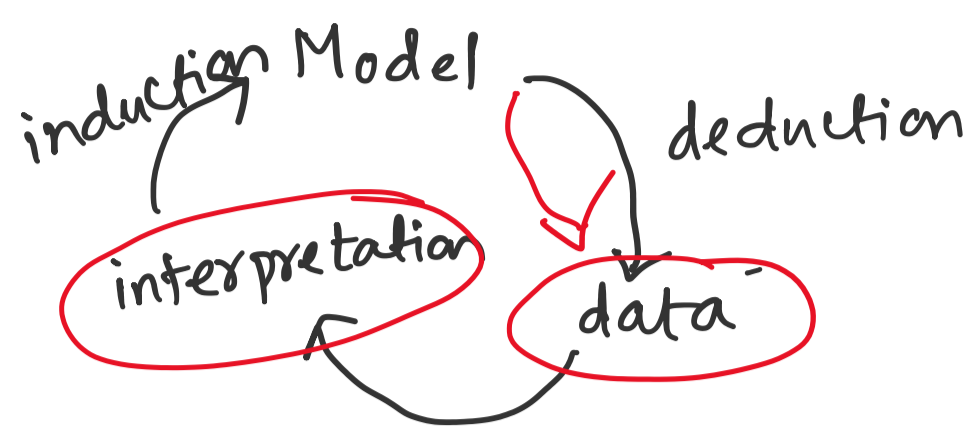


- discuss no lecture days and potential assignment releases
- Q: what is the closed loop iterative process for experimental design
- student questions: No

### Recap from lec 01



3 key things to keep in mind:

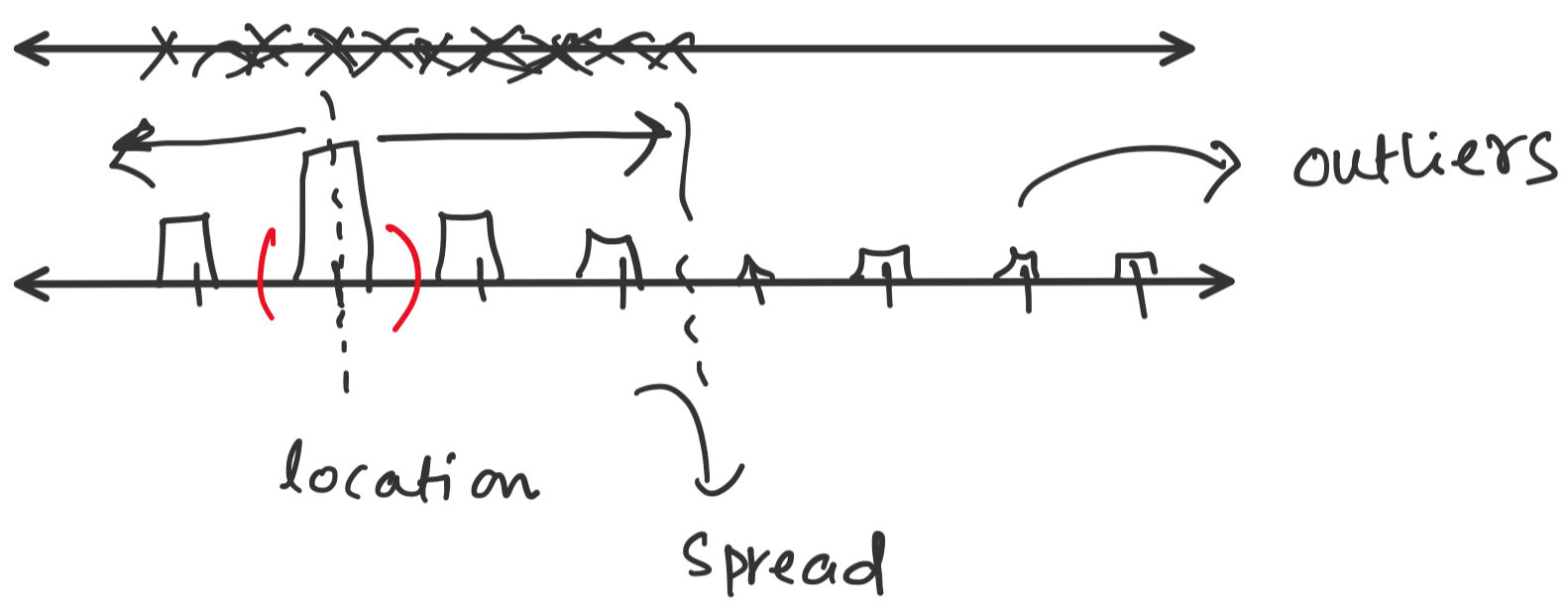
- 1) complexity
- 2) experimental errors
- 3) correlation vs causation

### distribution:

1. weather changes from day to day
2. physical characteristics (wt/height)

- single variable distributions
- multi variate distributions (eg: Temperature, seasons)

A distribution is quantified using 1) location 2) spread



outliers  $\leftarrow$  far away from both the location and also its spread of the distribution.

population  $\rightarrow$   $N$  is large

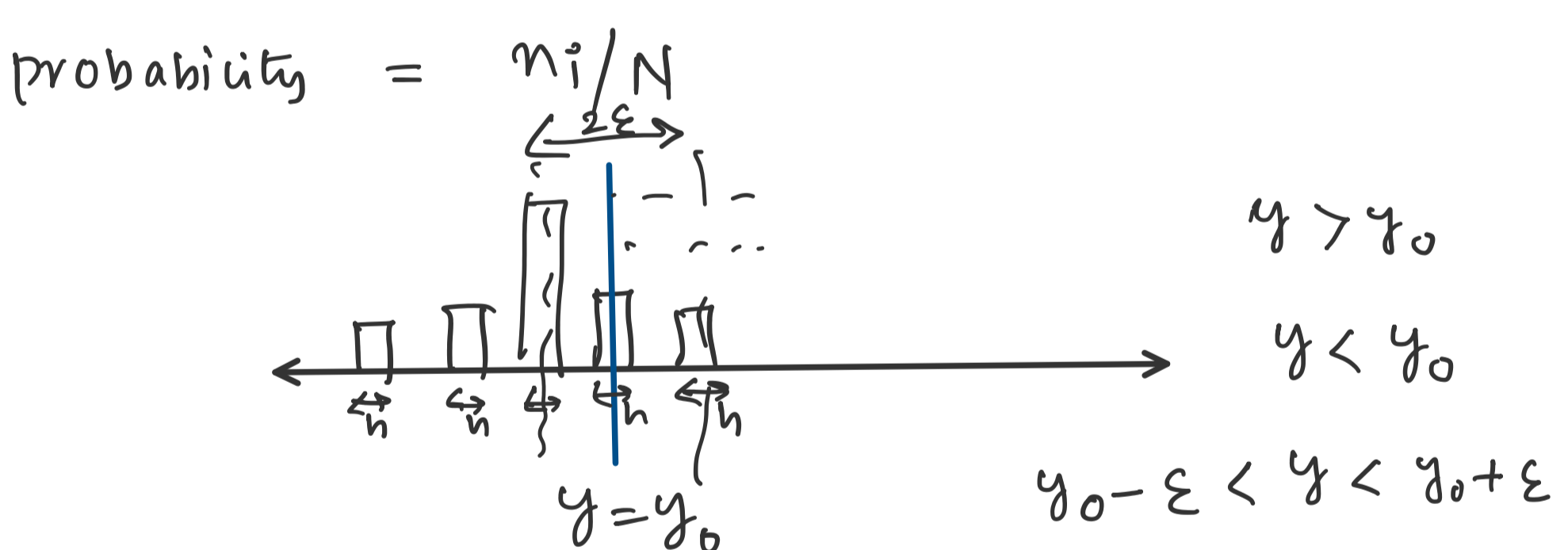
sample  $\rightarrow$   $N$  is small ( $n$ )

few observations ( $y_i \quad i \in 1, 2, \dots, n$ )

$$\text{location} = \sum y_i / n$$

$$\text{mean parameter } \mu = \frac{\sum y_i}{N}$$

random sample  $\rightarrow$  all the things you can sample have equal chance



$$\text{probability density} = P/h$$

$$\text{probability} = \left(\frac{P}{h}\right) \times h$$

$$\text{mass} = \int \times \text{volume}$$

Continuous distribution are  $h \rightarrow 0$

$\rightarrow$  auto correlation: consecutive samples are not random

spread: variance and standard deviation.

$y_i \rightarrow$  random variable from a distribution

$$\text{parameters} \left\{ \begin{array}{l} \mu = \mathbb{E}[y] \quad (\text{expected value}) \\ \sigma^2 = \mathbb{E}[y - \mu]^2 = \frac{\sum (y - \mu)^2}{N} \end{array} \right.$$

$$\bar{y} = \frac{\sum y}{n}$$

$$\begin{aligned} s^2 &= \frac{\sum (y_i - \bar{y})^2}{n-1} \quad (\text{linear combination rule} \\ &\quad + \text{loss degree of freedom}) \\ &= \frac{\sum (y_i - \bar{y})^2}{\text{dof}} \end{aligned}$$

Median  $\rightarrow$  middle value of your sorted sample  
(less influence by outliers)