**The `;` operator in stats (e.g. P(x; o) -** the **;** splits the variables on the left side from the parameters on the right side

- **Variable -** represents an argument passed in that depends on the data we have
- **Parameter -** a value that we already know or choose irrespective of what data we have

**Method of Moments -** Method of **estimating population parameters**

- **Population Parameter -** a number that describes something about an **entire** group or population
  - **Statistic -** an approximation of a population parameter based on a **subsample**
- We can estimate moments using this formula ----------> $\widehat{\mu}_j = \frac{1}{n}\sum_{i=1}^{n} w_i^j$
  - In this method, we don't need to adjust for the reduced degrees of freedom
  - This method is often **biased**, however, it has a relatively **very low complexity**
- **Bias -** The error between our statistic (estimate of parameter), and the actual parameter

https://en.wikipedia.org/wiki/Variational_message_passing

https://en.wikipedia.org/wiki/Expectation_propagation

**Expectation Propagation (EP) -** it is an **iterative approach** used to **approximate probability distributions**

- It approximates **intractable probability distributions**, by **minimising $D_{KL}$**(p||q)

https://en.wikipedia.org/wiki/Variational_Bayesian_methods

**Variational Bayesian Methods -** Approaches to approximating intractable p.ds by minimising the reverse $D_{KL}$(q||p)

- They estimate a complex network **p**, with a simpler network **q**, minimising information loss

**Belief Propagation (sum-product message passing) -** A message-passing algorithm used on Bayesian networks, that estimates the marginal distribution of a node **x**, without summing over every combination of every other variable for some value of **x == i**

**Bayesian Network -** graph which models events with probabilities to model an **intractably complex system** with a DAG

- It is a **DAG - Directed, acyclic graph**
  - There is some cause and effect taking place
- Nodes are connected with probabilities of going from one node to the other
- They are used for **inference -** deriving new knowledge from existing knowledge
- It is updated using Bayesian inference, updating your prior beliefs to get new posterior beliefs
- They can represent **conditionally dependent** and **conditionally independent** relationships
  - **Conditionally Independent -** if two nodes A and B point to C, we don't need to know A to know B
  - **Conditionally Dependent -** if two nodes A and B point to C, we need to know A to know C or vice versa
- It can be used to **simulate scenarios**, and see how they develop