MLE :

Maximum Likelihood Estimation :-

→ In Statistics, MLE is a method of estimating the parameters of probability distribution by maximizing a likelihood function, so that under the assumed statistical model is most probable. It is Generally a function defined over the Sample Space.

→ It determines Value for the parameters of a mode. Parameter Values are found such that they maximize the likelihood that the process described by the model.

① Eg: Suppose that x is a discrete Random Variable with a following Probability mass function where $0 < \theta < 1$ is a Parameter. The following 10 independent Observations.

X	0	1	2	3
$P(x)$	$\frac{2\theta}{3}$	$\frac{\theta}{3}$	$\frac{2(1-\theta)}{3}$	$\frac{1-\theta}{3}$

PMF

θ is Parameter.

from such a distribution (3, 0, 2, 1, 3, 2, 1, 0, 2, 1)

what is the maximum likelihood Estimation for θ?

Sol:

Sol: Since the Sample is $(3, 0, 2, 1, 3, 2, 1, 0, 2, 1)$, the likelihood is $L(\theta) = P(x=3)\, P(x=0)\, P(x=2)\cdot$

$$P(x=1)\, P(x=3)\, P(x=2)$$

$$P(x=1)\, P(x=0)\, P(x=2)$$

$$P(x=1)$$

Substituting from the probability distribution given above

$$L(\theta) = \prod_{i=1}^{n} P(x_i|\theta) = \left(\frac{2\theta}{3}\right)^2 \left(\frac{\theta}{3}\right)^3 \left(\frac{2(1-\theta)}{3}\right)^3 \left(\frac{1-\theta}{3}\right)^2$$

Clearly the likelyhood function $L(\theta)$ is not easy to maximize

$$\left[\begin{array}{l} f'(x) = 0 \\ f''(x) < 0 \quad \begin{array}{l} \nearrow \text{maxima} \\ \searrow \text{minima} \end{array} \end{array}\right] \text{ recall}$$

$d(uv) = uv' + vu'$ differentiation is very complex

So, apply log

$L(\theta) = \log L(\theta)$

$\left\{ \begin{array}{l} \log a^m = m \log a \\ \log mn = \log m + \log n \end{array} \right.$

$$2\left[\log\left(\frac{2}{3}\right) + \log\theta\right] + 3\left[\log\left(\frac{1}{3}\right) + \log\theta\right]$$

$$+ 3\left[\log\left(\frac{2}{3}\right) + \log(1-\theta)\right] + 2\left[\log\left(\frac{1}{3}\right) + \log(1-\theta)\right]$$

constants

$l(\theta) = C + 5\log\theta + 5\log(1-\theta)$

where c is a constant which does not depend on θ.

It can be seen that log likelihood function is easier to maximize. Compare to likelihood function.

$$l(\theta) = C + 5\log\theta + 5\log(1-\theta)$$

$$\frac{d}{d\theta} l(\theta) = 0 + \frac{5}{\theta} + \frac{5}{1-\theta}(0-1) \quad \left[\frac{d}{dx}(\theta=1)\right]$$

$$\left[\frac{d}{dx}(\log x) = \frac{1}{x}\right] \quad \frac{d\, l(\theta)}{d\theta} = \frac{5}{\theta} - \frac{5}{1-\theta} \Rightarrow \frac{d}{d\theta} l(\theta) = 0$$

$$\frac{dl(\theta)}{d\theta} = 0 \Rightarrow \frac{5}{\theta} - \frac{5}{1-\theta} = 0$$

$$\frac{5(1-\theta) - 5\theta}{\theta(1-\theta)} = 0$$

$$5 - 5\theta - 5\theta = 0$$

$$\left[\frac{d}{dx}(\log x) = 1/x\right]$$

$$10\theta = 5$$

$$\boxed{\theta = 0.5}$$

Scanned with OKEN Scanner

* Stochastic process:

Mathematically, a Stochastic process is a Set of random Variables $\{x_t\}$ or $\{x(t)\}$ depending on Some real Parameter like time t. These are also Known as random processes or random functions.

[Random process/stochastic process are Same]

Eg: 1) Queuing

2) unbiased die

X_n Outcome of n^{th} throw

$\{X_n \mid n \geq 1\}$
$\downarrow$
family of Random Variables

$X_1 = 1$ to 6

$X_2 = 1$ to 6

3) X_n is maximum number shown in first n rows.

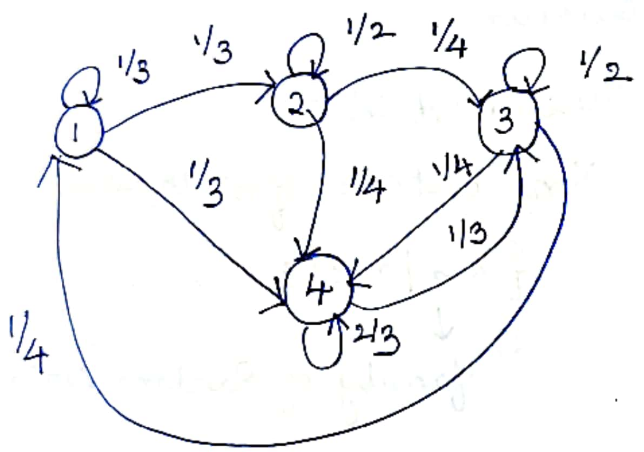
2/2/2021

(i) Determine if the following transition matrix is ergodic markov chain?

$$\begin{array}{c|c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 0 & 1/2 & 1/4 & 1/4 \\ 1/4 & 0 & 1/2 & 1/4 \\ 0 & 0 & 1/3 & 2/3 \end{bmatrix} \end{array}$$

Present States → Probabilities

Sol: The transition diagram for the given transition matrix.



✓ In the above transition diagram, State 1 can go to any other State directly except to State 3.

→ To go to State 3, one must go from State 1 to State 2 and then to State 3.

→ It is Possible for State 2 to directly to all other States except to State 1.

→ now from State 2 one can reach State 1 in two ways. one way is from State 3 to State 1 and the other way is from State 4 to State 3 and then State 1.

→ From State 4 it is Possible to go to State 3 and state 4. except to State 1 and State 2.

☆ { Hence the given Matrix is an Ergodic Markov Chain }

② A training process is considered as two State markov chain. If it rains, it is considered to be in State 0 and it does not rain, the chain is in the state of 1. The transition Probability of the markov chain is defined by $P = \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}$. find the Probability that it will rain for 3 days from today assuming that it is raining today. Assume that the mutual probabilities of State 0 are or State 1 as 0.4 and 0.6 respectively.

Sol:

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0.6 & 0.4 \\ 1 & 0.2 & 0.8 \end{array}$$ → 2 State markov chain

$2^2 = 4$

$$P^2 = \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}$$

$$= \begin{bmatrix} 0.44 & 0.56 \\ 0.28 & 0.72 \end{bmatrix}$$

$P^2 \cdot P \Rightarrow \begin{bmatrix} 0.44 & 0.56 \\ 0.28 & 0.72 \end{bmatrix} \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}$

$P^3 = P^2 \cdot P$

[$\because P^{(3)} \rightarrow$ 3 days]

$$= \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0.376 & 0.624 \\ 1 & 0.312 & 0.688 \end{array}$$

The probability that it will rain on 3^{rd} day given that it will rain today is 0.376

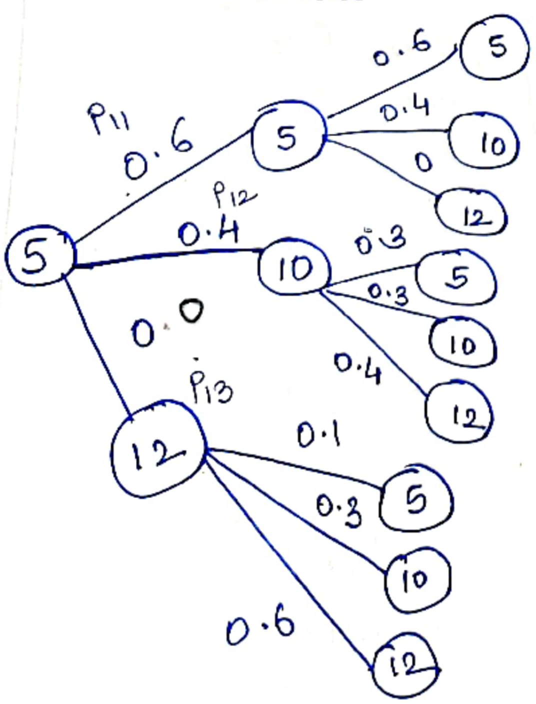
(3) The number of units of an item that are
3/2/21 withdrawn from inventory on a day-to-day basis is a markov chain process in which requirements for tomorrow depend on today's requirements. A one-day-transition matrix is given below.

$$\begin{array}{cc} & \text{Tomorrow} \\ & \begin{array}{ccc} 5 & 10 & 12 \end{array} \\ \text{Today} \begin{array}{c} 5 \\ 10 \\ 12 \end{array} & \begin{bmatrix} 0.6 & 0.4 & 0.0 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.3 & 0.6 \end{bmatrix} \end{array}$$

(i) Construct a tree diagram showing inventory requirements on two consecutive days.

(2) Develop a two-day transition matrix.

Sol: (I) The inventory requirements on two consecutive days is represented by a tree diagram as shown below



tree diagram (must be Same line)

$P_{11}^{(2)} = (0.6)(0.6) + (0.4)(0.3) + 0(0.1)$

$= 0.36 + 0.12 + 0$

(2) 2days.

$P_{11}^{(2)} = 0.48$

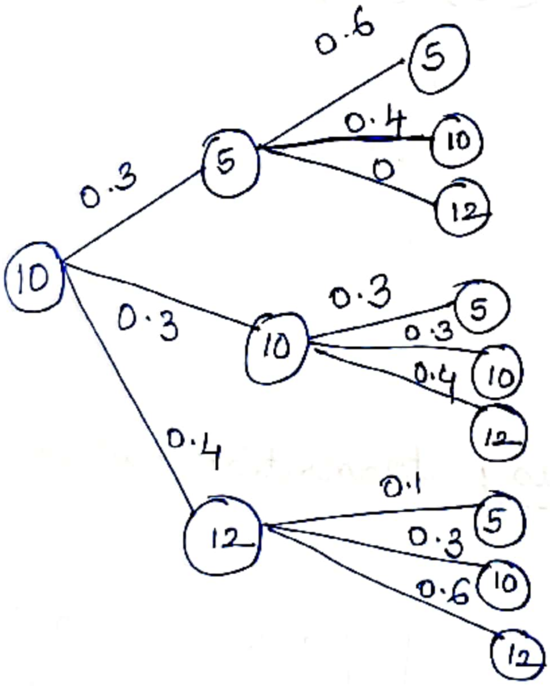
$P_{12}^{(2)} = (0.6)(0.4) + (0.4)(0.3) + 0(0.3)$

$= 0.24 + 0.12 + 0$

$P_{12}^{(2)} = 0.36$

$P_{13}^{(2)} = (0.6)(0) + (0.4)(0.4) + 0(0.6)$

$P_{13}^{(2)} = 0.16$



$P_{23}^{(2)} = (0.3)(0) + (0.3)(0.4) + (0.4)(0.6)$

$= 0 + 0.12 + 0.24$

$P_{23}^{(2)} = 0.36$

$P_{21}^{(2)} = (0.3)(0.6) + (0.3)(0.3) + (0.4)(0.1)$

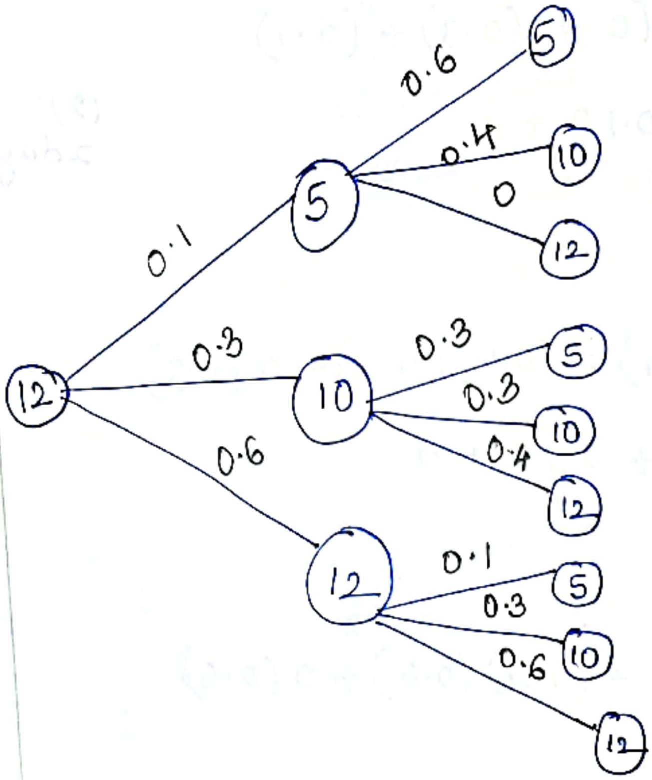
$= 0.18 + 0.9 + 0.04$

$P_{21}^{(2)} = 0.31$

$P_{22}^{(2)} = (0.3)(0.4) + (0.3)(0.3) + (0.4)(0.3)$

$= 0.12 + 0.9 + 0.12$

$P_{22}^{(2)} = 0.33$



$$P_{31}^{(2)} = (0.31)(0.6) + (0.3)(0.3) + (0.6)(0.1)$$

$$P_{31}^{(2)} = 0.21$$

$$P_{32}^{(2)} = 0.31$$

$$P_{33}^{(2)} = 0.48$$

(II) Develop a <u>two-day</u> transition matrix

$$P^2 = \begin{bmatrix} 0.6 & 0.4 & 0.0 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.3 & 0.6 \end{bmatrix} \begin{bmatrix} 0.6 & 0.4 & 0.0 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.3 & 0.6 \end{bmatrix}$$

$$= \begin{bmatrix} 0.48 & 0.36 & 0.16 \\ 0.31 & 0.33 & 0.36 \\ 0.21 & 0.31 & 0.48 \end{bmatrix}$$

* Stochastic Matrix :

→ A Stochastic Matrix is a random matrix with non-negative elements and unit row Sums.

must be Square matrix

≥ 0

* Regular Matrix :

→ A Stochastic Matrix P is Said to be regular if all the entries of Some Power P^m are Positive.

* Not Regular Matrix:

→ A Stochastic Matrix P is not regular if one 1 occurs in principle Mean diagonal.

Eg: (i) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}_{2\times 3}$ → Not a Square matrix ∴ It is not Stochastic

(ii) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times 2}$ → Matrix is Square Matrix with non negative entries and Sum of the elements in each row is equal to (=) 1

∴ The matrix is Stochastic.

(iii) $\begin{bmatrix} 0 & 1 \\ 1/3 & 1/4 \end{bmatrix}_{2\times 2}$ → The matrix is Square matrix but Sum in each row is not equal to 1. So it is not Stochastic.

4) $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ Stochastic

5) $\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$ The matrix is not Stochastic because it contains negative elements.

6) $\begin{bmatrix} 0 & 2 \\ 1/4 & 1/4 \end{bmatrix}$ The matrix is Square but Sum in each ~~in~~ row is not equal to 1

$\therefore$ It is not Stochastic process

Q. check whether the following Markov chain is Regular or Not.

(i) $P = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 1/2 \end{bmatrix}_{4\times4}$

(ii) $P = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$

$$P^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

$$P^3 = P^2 \cdot P \Rightarrow \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix} = P$$

$P^3 = P$

$P^4 = P^3 \cdot P = P^2$

$P^6 = P^2$

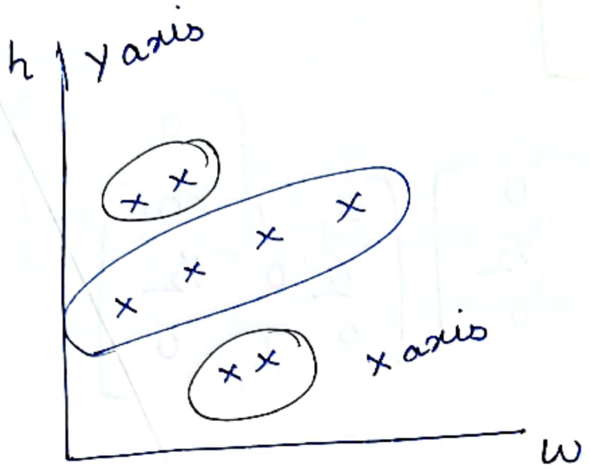
$P^5 = P$

$$\begin{Bmatrix} P^{2n} = P^2 \\ P^{2n+1} = P \end{Bmatrix}$$

$\therefore$ Not a Regular Matrix

4/1/2021

★ Linear Regression: Geometrical Interpretation



usually height ↑ weight ↑

Exceptions : Outliers

Obesity very thin

Plot the Graph for majority (max) Points.

xx

minimize the errors: