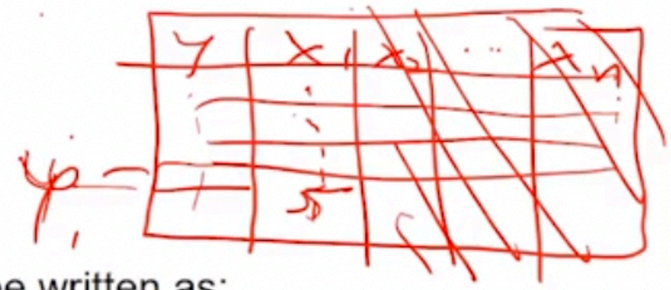
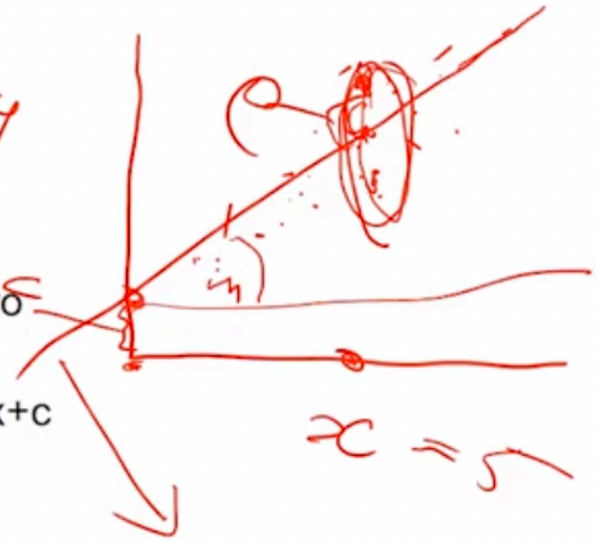


Linear Regression Theory

1. The general form of linear regression model may be written as:

   $Y = m1x1 + m2x2 + m3x3 + \ldots + mnxn + C + e$

   Y = Regressand / dependent variable
   Xi = are the independent / explanatory variables
   e = random / stochastic error term

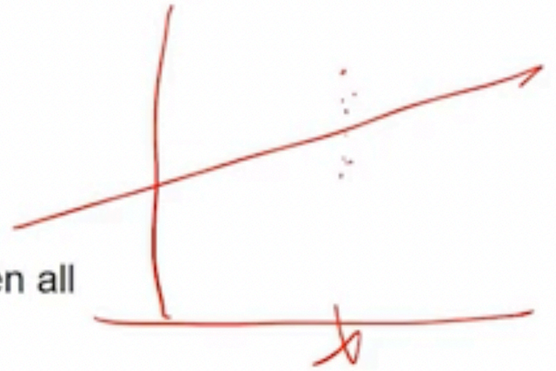


2. The equation above is known as the population / true model. It consist of two components
   a. Deterministic component (m1x1+m2x2+….+ mnxn +C). Let this be mx+c
   b. Non-deterministic component e

3. The expression mx+c is conditional mean of y, E(yi|xi), i.e. expected value of y for given value of xi

4. Thus the expression y = mx +c states that an individual value of yi is equal to the mean of the population of which it is a member +/- the random error term ‘e’

Linear Regression Theory



5. In the equation 'C' is known as the intercept i.e. the default value of Y when all independent parameters are 0

6. The m1…mn are known as slope coefficients reflecting the change in y for one unit change in corresponding independent variable

7. In regression analysis the primary objective is to explain the average behavior of Y, that is, how average Y responds to changes in the values of the X variables. An individual Y value will hover around its mean value

$y = m_1 x_1 + m_2 x$

8. The error term 'e' is a catchall for all those variables that cannot be introduced in the model for a variety of reasons

$y = m_1 x_1 + m_2 x_2 + \cdots$

9. The term 'Linear' can be interpreted in two ways –
   a) Linearity in the Variables - the variables x are raised to power 1
   b) **Linearity in the Parameters** – the coefficients m are raised to power 1, x can be raised to any power

$y = m_1 x_1^2 + m_2 x_2^{-2}$

Linear Regression Theory

10. The second interpretation of linearity is that the conditional expectation of Y, E(Y | Xi), is a linear function of the parameters, the m's; it may or may not be linear in the variable X. In this interpretation $E(Y \mid X_i) = \beta_1 + \beta_2 X_i^2$