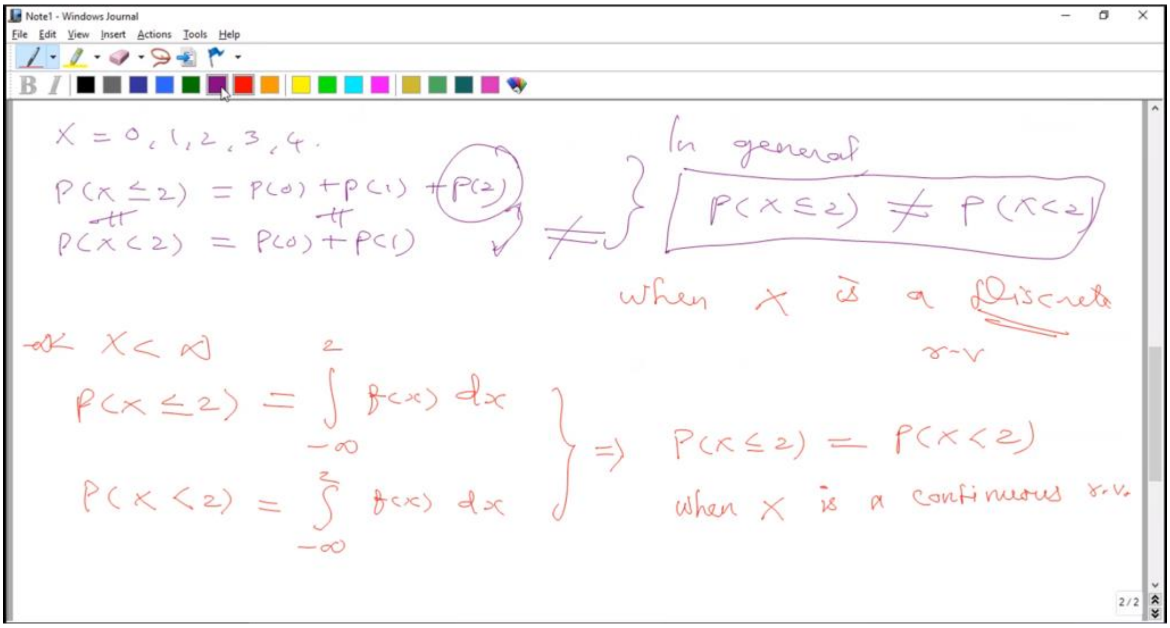
Major Difference between Discrete and Continuous Random variable



Two dimensional random Variables

Our study of random variables and their probability distributions in the preceding sections was restricted to one-dimensional sample spaces, in that we recorded outcomes of an experiment as values assumed by a single random variable.

There will be situations, however, where we may find it desirable to record the simultaneous outcomes of several random variables.

For example, we might measure the amount of precipitate P and volume V of gas released from a controlled chemical experiment, giving rise to a two-dimensional sample space consisting of the outcomes (p, v), or we might be interested in the hardness H and tensile strength T of cold-drawn copper, resulting in the outcomes (h, t).

Two dimensional random Variables

In a study to determine the likelihood of success in college based on high school data, we might use a threedimensional sample space and record for each individual his or her aptitude test score, high school class rank, and grade-point average at the end of freshman year in college.

For example, if an 18-wheeler is to have its tires serviced and X represents the number of miles these tires have been driven and Y represents the number of tires that need to be replaced, then $p(30000, 5)$ is the probability that the tires are used over 30,000 miles and the truck needs 5 new tires.

Joint probability distribution

Definition

The function $p(x, y)$ is a joint probability distribution or probability mass function of the discrete random variables X and Y if

1. $p(x, y) \geq 0$, *for all* (x, y),
2. $\sum_x \sum_y p(x, y) = 1$,
   where $p(x, y) = P(X = x, Y = y)$.

Variance:

Variance(x) → E(x^2) – [E(x)]^2
Variance(x) → 1/n (∑x^2) - x^2
Variance(x) → ∑ (xi-x)^2

Covariance :

Covariance(x,y) → E(xy) – [E(xy)]^2
Covariance(x,y)   → 1/n (∑ xy) -   xy

Karl Pearson correlation coefficient → Covariance(x,y) / SDx SDy

Spearman Rank Correlation → 1 – (∑ di^2 / n(n^2-1))
di → Rank of X - Rank of Y

$$r(X, Y) = 1 - \frac{6 \Sigma d_i^2}{n(n^2 - 1)}$$

Repeated Rank Correlation→ 1 – 6 [∑ di^2 + 1/12 m1(m1^2-1) + 1/12 m2(m2^2-1) +..+] / n(n^2-1)

Partial and Total correlation

Partial and Multiple Correlation

Let us consider the example of yield of rice in a firm. It may be affected by the type of soil, temperature, amount of rainfall, usage of fertilizers etc. It will be useful to determine how yield of rice is influenced by one factor or how yield of rice is affected by several other factors. This is done with the help of partial and multiple correlation analysis.

The basic distinction between multiple and partial correlation analysis is that in the former, the degree of relationship between the variable Y and all the other variables $X_1, X_2, ..., X_n$ taken together is measured, whereas, in the later, the degree of relationship between Y and one of the variables $X_1, X_2, ..., X_n$ is measured by removing the effect of all the other variables.

Partial correlation

Partial correlation coefficient provides a measure of the relationship between the dependent variable and other variable, with the effect of the rest of the variables eliminated. If there are three variables X_1, X_2 *and* X_3, there will be three coefficients of partial correlation, each studying the relationship between two variables when the third is held constant. If we denote by $r_{12.3}$, that is, the coefficient of partial correlation X_1 *and* X_2 keeping X_3 constant, it is calculated as

Multiple Correlation

In multiple correlation, we are trying to make estimates of the value of one of the variable based on the values of all the others. The variable whose value we are trying to estimate is called the dependent variable and the other variables on which our estimates are based are known as independent variables.

The coefficient of multiple correlation with three variables X_1, X_2 *and* X_3 are $R_{1.23,}$ $R_{2.13}$ and $R_{3.21}$. $R_{1.23,}$ is the coefficient of multiple correlation related to X_1 as a dependent variable and X_2, X_3 as two independent variables and it can be expressed in terms of $r_{12,}$ r_{23} and r_{13} as

Y	X	
Yield of crops	chemical fertilizers	$\Longrightarrow$ Partial correlation
Yield of crops	chemical fertilizers, effect of pesticides, manures	$\Longrightarrow$ Total correlation

Partial correlation:

the degree of relationship bt Y and $X_1, X_2, \dots X_n$ is measured by removing the effect of all other variables

one of the variables

Multiple correlation:

The degree of relationship between Y and all other variables $X_1, X_2 \cdots X_n$ taken together.

Bt 1 and 2 by keeping 3→constant

Partial & Multiple correlation

Partial correlation

x, y, z

Partial correlation coefficient of x & y, controlling z

$$r_{xy\cdot z} \Rightarrow \frac{r_{xy} - r_{xy}\, r_{yz}}{\sqrt{1-r^2_{xz}}\,\sqrt{1-r^2_{yz}}} \quad \text{(or)} \quad r_{12\cdot 3} \Rightarrow \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{1-r_{12}^2}\,\sqrt{1-r_{23}^2}}$$

$b_{xy} = b_{yx}$

regression coefficient } not symmetric

correlation coefficient } symmetric

$r_{21} = r_{12}$

Multiple correlation

$$R_{1(23)} \Rightarrow \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}\, r_{23}\, r_{31}}{1 - r_{23}^2}}$$

$x_1 \Rightarrow$ dependent

$x_2, x_3 \Rightarrow$ independent

$$R_{2(31)} \Rightarrow \sqrt{\frac{r_{23}^2 + r_{21}^2 - 2r_{23}\, r_{21}\, r_{31}}{1 - r_{31}^2}}$$

$x_2 \Rightarrow$ dependent

$x_3, x_1 \Rightarrow$ independent

V ' 12

PROPERTIES OF MULTIPLE CORRELATION COEFFICIENT

The following are some of the properties of multiple correlation coefficients:

1. Multiple correlation coefficient is the degree of association between observed value of the dependent variable and its estimate obtained by multiple regression,

2. Multiple Correlation coefficient lies between 0 and 1.

3. If multiple correlation coefficient is 1, then association is perfect and multiple regression equation may said to be perfect prediction formula.

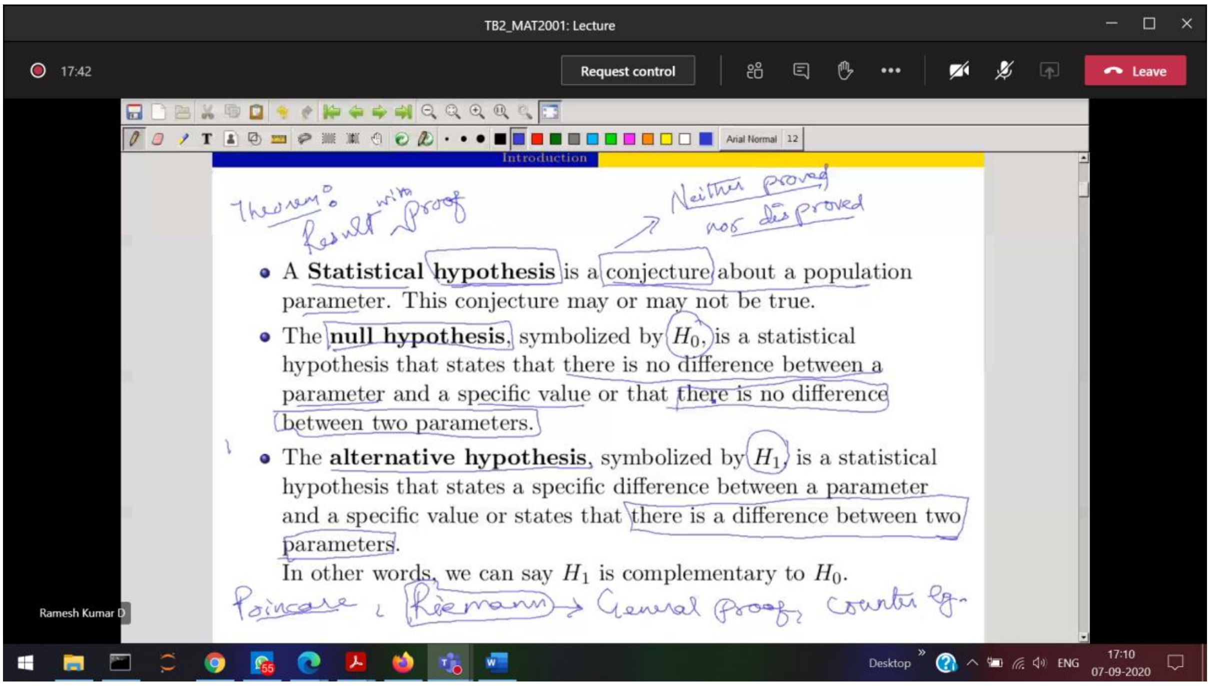
4. If multiple correlation coefficient is 0, dependent variable is uncorrelated with other independent variables. From this, it can be concluded that multiple regression equation fails to predict the value of dependent variable when values of independent variables are known.

5. Multiple correlation coefficient is always greater or equal than any total correlation coefficient. If $R_{1.23}$ is the multiple correlation coefficient than $R_{1.23} \geq r_{12}$ or r_{13} or r_{23} and

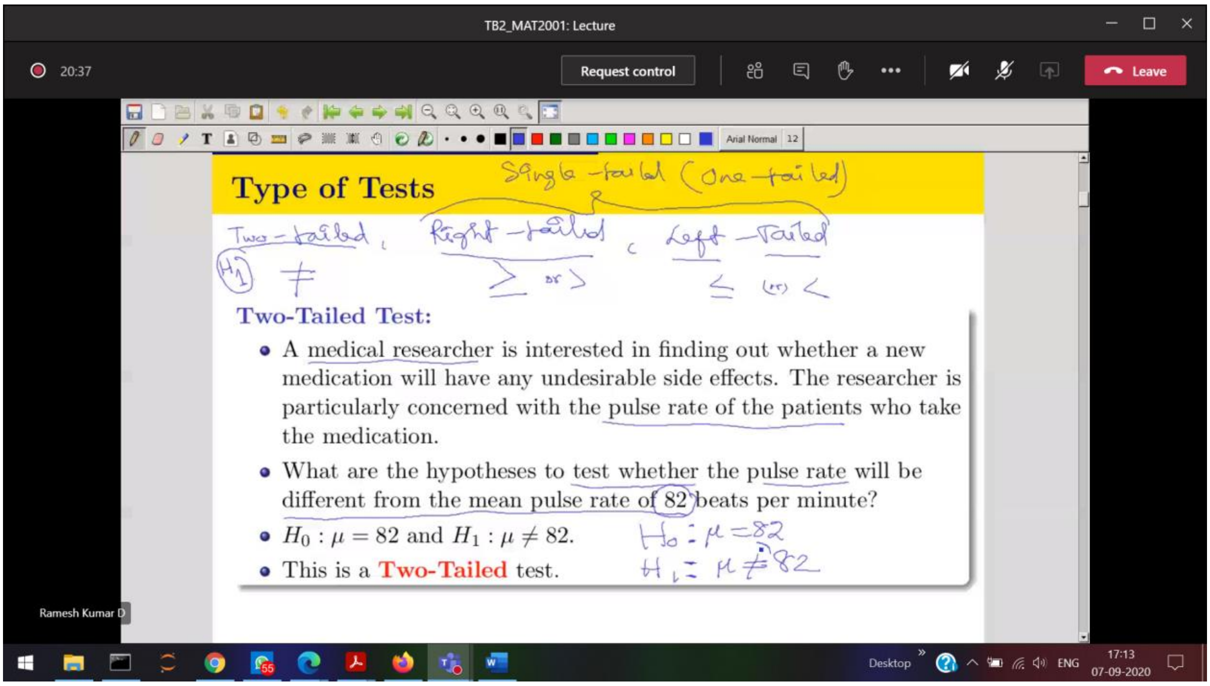
6. Multiple correlation coefficient obtained by method of least squares would always be greater than the multiple correlation coefficient obtained by any other method.

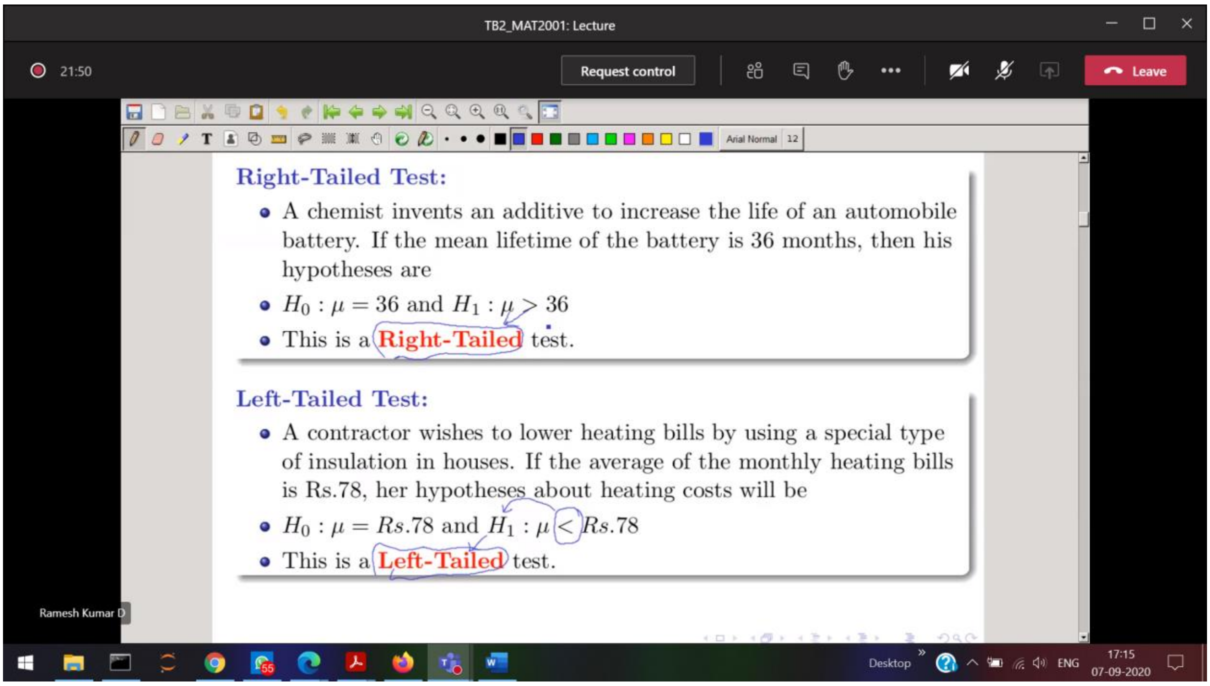
Binomial Distribution

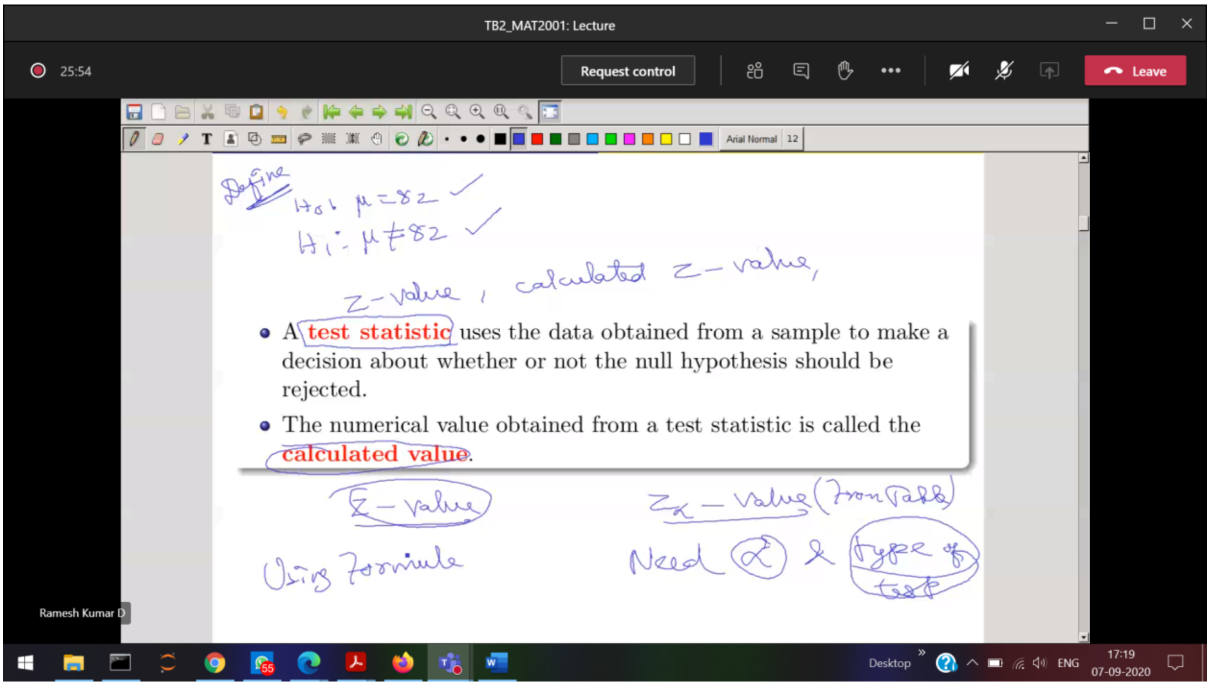
Question : Whatever asking to find out → success

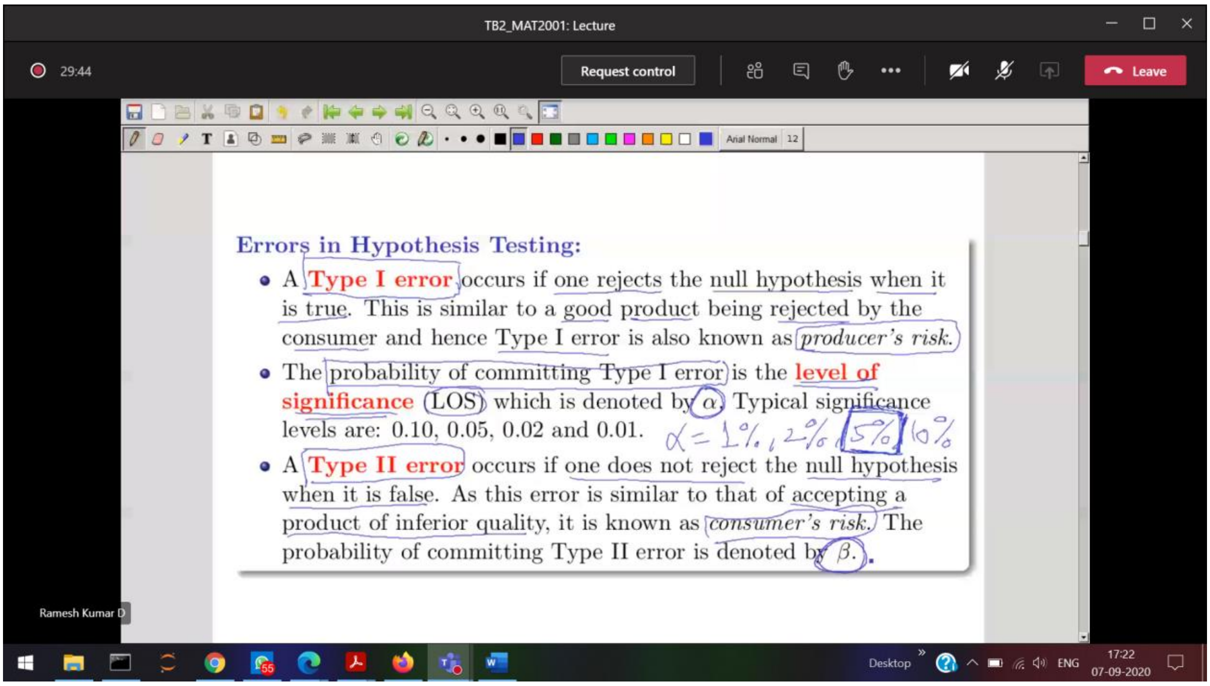


H1 complement of H0




TB2_MAT2001: Lecture
21:50
Request control
Leave
Right-Tailed Test:
A chemist invents an additive to increase the life of an automobile battery. If the mean lifetime of the battery is 36 months, then his hypotheses are
$H_0: \mu = 36$ and $H_1: \mu > 36$
This is a Right-Tailed test.
Left-Tailed Test:
A contractor wishes to lower heating bills by using a special type of insulation in houses. If the average of the monthly heating bills is Rs.78, her hypotheses about heating costs will be
$H_0: \mu = Rs.78$ and $H_1: \mu < Rs.78$
This is a Left-Tailed test.
Ramesh Kumar D
17:15
07-09-2020


TB2_MAT2001: Lecture
25:54
Request control
Leave
Define
$H_0: \mu = 82$ ✓
$H_1: \mu \neq 82$ ✓
z-value, calculated z-value,
A test statistic uses the data obtained from a sample to make a decision about whether or not the null hypothesis should be rejected.
The numerical value obtained from a test statistic is called the calculated value.
Z-value
Using Formula
Z_α – value (from Table)
Need (α) & type of test
Ramesh Kumar D
17:19
07-09-2020

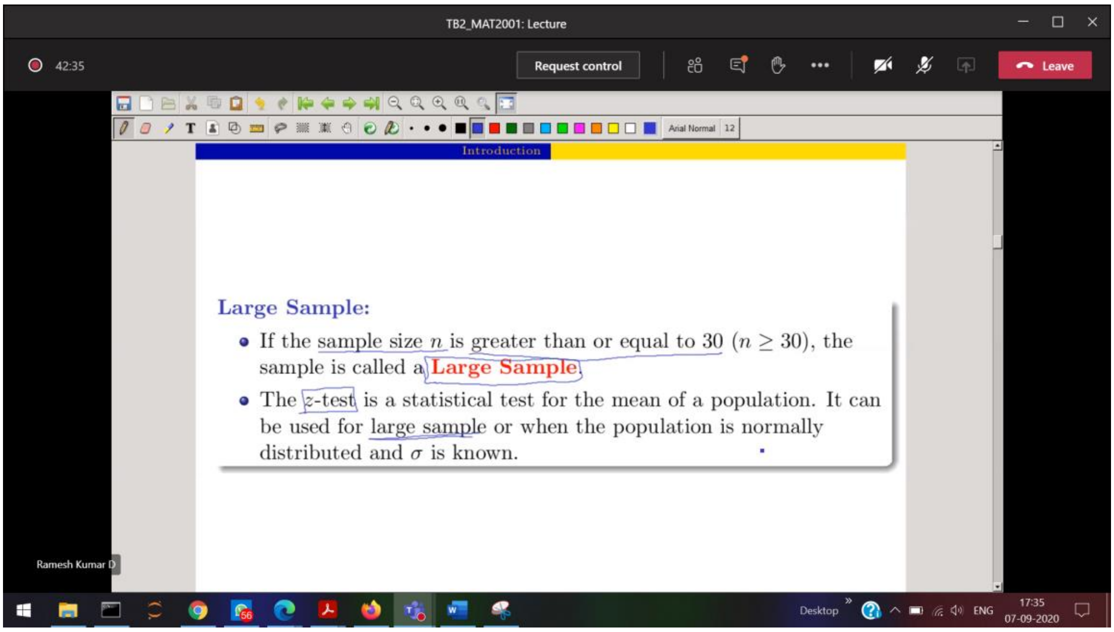


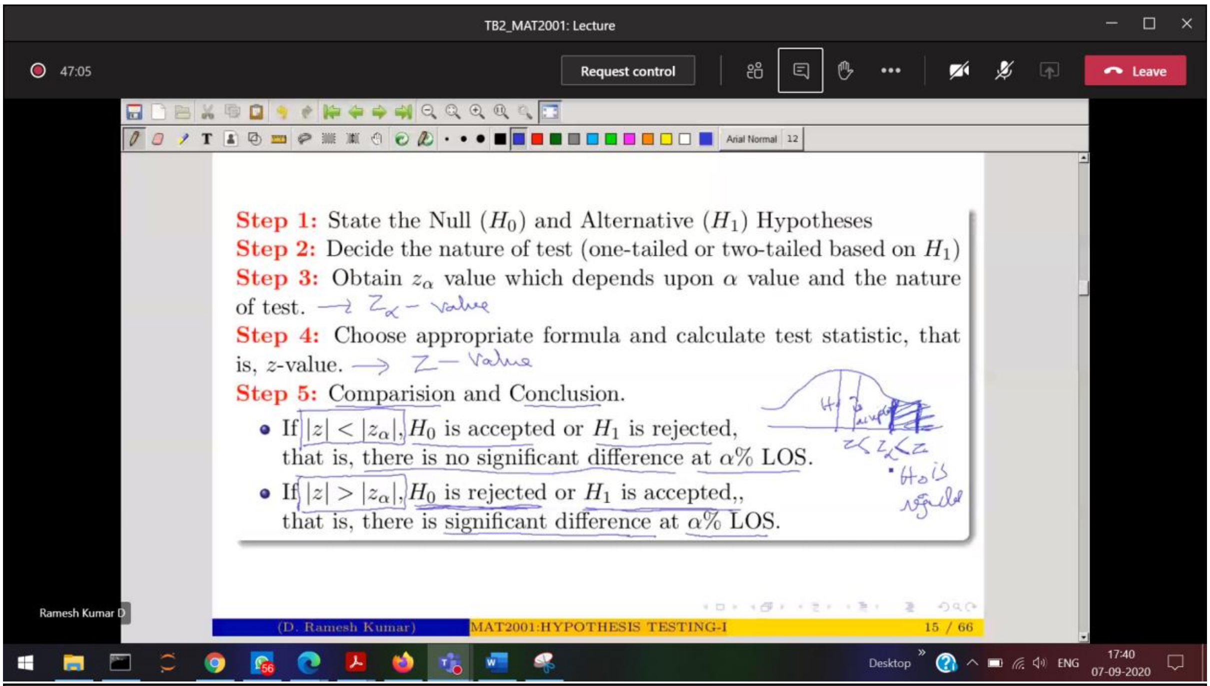
Type-1 and Type-2 are complement to each-other

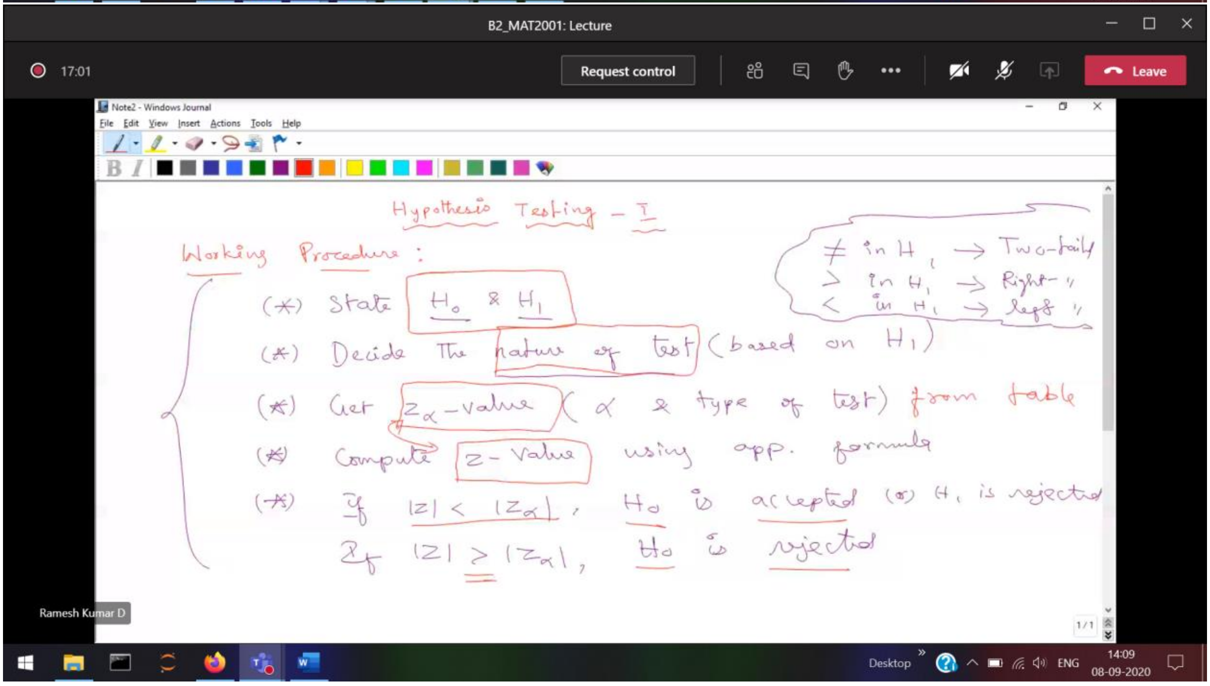
The critical values for some standard LOS's are given in the following table:

Table 1: → LOS

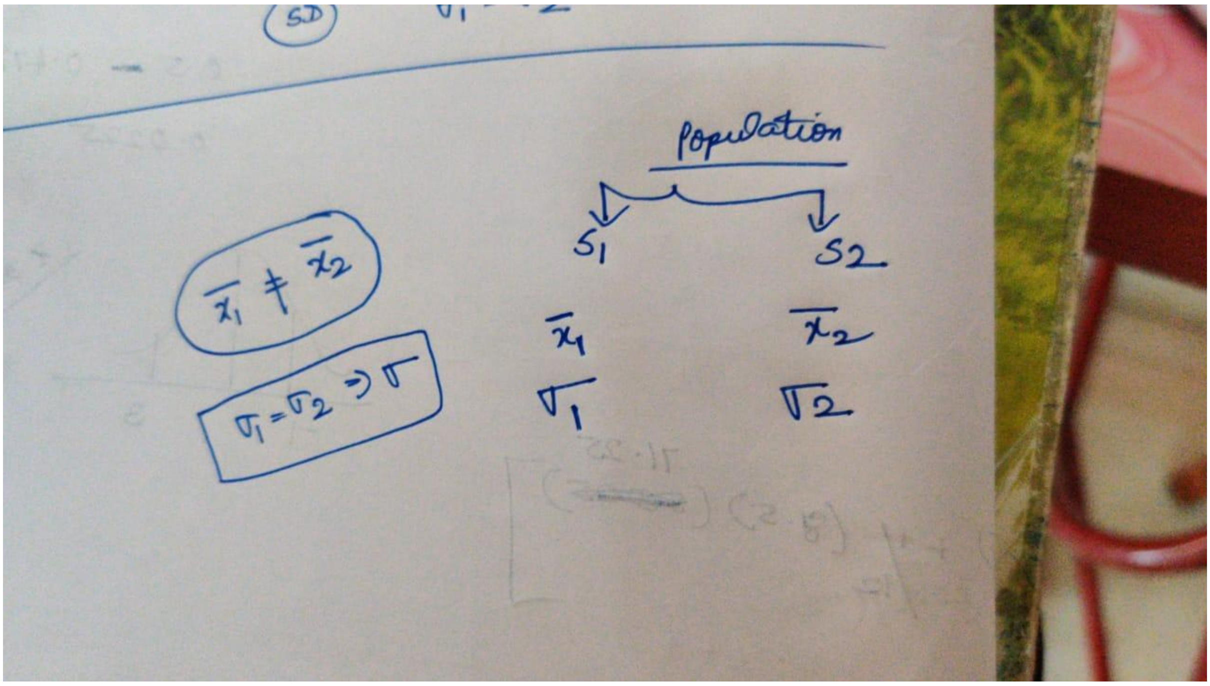
Type of Test	$\alpha = 1\%(0.01)$	$\alpha = 2\%(0.02)$	$\alpha = 5\%(0.05)$	$\alpha = 10\%(0.1)$
Two-Tailed	$\lvert z_\alpha \rvert = 2.58$	$\lvert z_\alpha \rvert = 2.33$	$\lvert z_\alpha \rvert = 1.96$	$\lvert z_\alpha \rvert = 1.645$
Right-Tailed	$z_\alpha = 2.33$	$z_\alpha = 2.055$	$z_\alpha = 1.645$	$z_\alpha = 1.28$
Left-Tailed	$z_\alpha = -2.33$	$z_\alpha = -2.055$	$z_\alpha = -1.645$	$z_\alpha = -1.28$







Z basically denotes area



The SD's are same, but the means should not be the same,

Differentiation

PRODUCT RULE

If u and v are two functions of x, then the derivative of the product uv is given by...

$$\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$$

Don't miss...

Later in this section:
Quotient Rule

In words, this can be remembered as:

> "The derivative of a product of two functions is the first times the derivative of the second, plus the second times the derivative of the first."

u*v = uv'+vu'

QUOTIENT RULE

(A **quotient** is just a fraction.)

If u and v are two functions of x, then the derivative of the quotient $\frac{u}{v}$ is given by...

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$$

In words, this can be remembered as:

> "The derivative of a quotient equals bottom times derivative of top minus top times derivative of the bottom, divided by bottom squared."

u/v = (vu' – uv') / v^2

Binomial formulae

n is positive

(i) $(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \dots\dots + \frac{n(n-1)(n-2)\dots\dots(n-r+1)}{r!}x^r + \dots\dots$

(ii) $(1-x)^n = 1 - nx + \frac{n(n-1)}{2!}x^2 - \dots\dots + \frac{n(n-1)(n-2)\dots\dots(n-r+1)}{r!}(-x)^r + \dots\dots$

n is negative

(iii) $(1-x)^{-n} = 1 + nx + \frac{n(n+1)}{2!}x^2 + \frac{n(n+1)(n+2)}{3!}x^3 + \dots\dots + \frac{n(n+1)\dots\dots(n+r-1)}{r!}x^r + \dots\dots$

(iv) $(1+x)^{-n} = 1 - nx + \frac{n(n+1)}{2!}x^2 - \frac{n(n+1)(n+2)}{3!}x^3 + \dots\dots + \frac{n(n+1)\dots\dots(n+r-1)}{r!}(-x)^r + \dots\dots$

(v) $(1+x)^{-1} = 1 - x + x^2 - x^3 + \dots\dots \infty$

(vi) $(1-x)^{-1} = 1 + x + x^2 + x^3 + \dots\dots \infty$

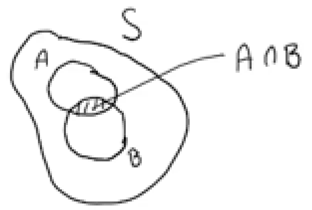
(vii) $(1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + \dots\dots \infty$

(viii) $(1-x)^{-2} = 1 + 2x + 3x^2 + 4x^3 + \dots\dots \infty$

(ix) $(1+x)^{-3} = 1 - 3x + 6x^2 - \dots\dots\dots \dots\dots \infty$

(x) $(1-x)^{-3} = 1 + 3x + 6x^2 + \dots\dots\dots \dots\dots \infty$

Independent random variables



$P(A) = 0.5$

$P(B) = 0.6$

$P(A \cap B) = 0.3$

A, B are independent

$0.5 \cdot 0.6 = 0.3$

$$P(A \cap B) = P(A)P(B)$$

X, Y

$$A = \{X \le a\} \quad B = \{Y \le b\}$$

$$A \cap B = \{X \le a, Y \le b\}$$

$$P(X \le a, Y \le b) = P(X \le a) \cdot P(Y \le b)$$

$$\boxed{F(a,b) = F_X(a) \cdot F_Y(b)} \leftarrow$$

X and Y are two random variables. We say that X and Y are *independent* random variables if the events $\{X \le a\}$ and $\{Y \le b\}$ are independent events for all real numbers a, b.

Thus, X and Y are independent random variables if and only if

$$F(x, y) = F_X(x)F_Y(y)$$

for all x, y.

The joint p.d.f of the random variable (X, Y) is given by $f(x, y) = Kxy\, e^{-(x^2 + y^2)}$, $x > 0, y > 0$. Find the value of K and also prove that X and Y are independent.

To prove: X and Y are independent

i.e., To prove: $\boxed{f(x)\, f(y) = f(x, y)}$

Conditional probability of Joint probability density function

$$f_{Y/X}\left(\frac{y}{x}\right). \qquad f\left(\frac{y}{x}\right) = \frac{f(x, y)}{f(x)}$$

gives

Joint probability density function(x,y) in the marginal density function(x)

1. $(a + b)^2 = a^2 + 2ab + b^2 = (-a - b)^2$
2. $(a - b)^2 = a^2 - 2ab + b^2$
3. $(a - b)(a + b) = a^2 - b^2$
4. $(a + b + c)^2 = a^2 + b^2 + c^2 + 2ab + 2bc + 2ca$
5. $(a + b - c)^2 = a^2 + b^2 + c^2 + 2ab - 2bc - 2ca$
6. $(a - b + c)^2 = a^2 + b^2 + c^2 - 2ab - 2bc + 2ca$
7. $(-a + b + c)^2 = a^2 + b^2 + c^2 - 2ab + 2bc - 2ca$
8. $(a - b - c)^2 = a^2 + b^2 + c^2 - 2ab + 2bc - 2ca$
9. $(a + b)^3 = a^3 + b^3 + 3ab(a + b)$
10. $(a - b)^3 = a^3 - b^3 - 3ab(a - b)$
11. $a^3 + b^3 = (a + b)^3 - 3ab(a + b)$
    $= (a + b)(a^2 - ab + b^2)$
12. $a^3 - b^3 = (a - b)^3 + 3ab(a - b)$
    $= (a - b)(a^2 + ab + b^2)$
13. $a^3 + b^3 + c^3 - 3abc = (a + b + c)(a^2 + b^2 + c^2 - ab - bc - ca)$

if $a + b + c = 0$ then $a^3 + b^3 + c^3 = 3abc$

Conditional density functions

*) Conditional density function of X with given Y } $\int_{x=0}^{1} \frac{f_{X,Y}(x,y)}{\boxed{f_Y(y)}} \, dx$

*) Conditional density function of Y with given X } $\int_{y=0}^{2} \frac{f_{X,Y}(x,y)}{\boxed{f_X(x)}} \, dy$