Final Syllabus Stats

1. **Normal Dist.**
Find the value of z according to this formula if X is given. And then find area by seeing corresponding value of z.

$$Z = \frac{X - \mu}{\sigma}$$

Normal dist in reverse:

- If area is above then write + with value of z.
- If area is below then write – with value of z.

Binomial:

- If less than X: find prob. till X-0.5 e.g X=30 then P(x<=29.5)
- If less than equal to X: Find prob till X+0.5 e.g X=30 then P(x<=30.5)
- If greater than X: then X+0.5 e.g, X=60 then P(x>=60.5)
- If at least X: then X-0.5 e.g, X=60 then P(x>=59.5)
- If in between X1 and X2: X1-0.5 and X2+0.5 e.g, X1=40 and X2=60 P(39.5<=X<=60.5)
- Exactly X: make X1=X-0.5 and X2=X+0.5 e.g, X=50 then find P(49.5<=X<=50.5)

2. **Linear Regression:**

- Co-eff. Of Correlation "r":

$$r = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sqrt{[\sum x^2 - \frac{(\sum x)^2}{n}][\sum y^2 - \frac{(\sum y)^2}{n}]}}$$

Or $r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 . \sum(y-\bar{y})^2}}$

- **Y on X:**
Y=$\alpha + \beta x$
$\alpha = \bar{y} - \beta\bar{x}$

$$\beta_{yx} = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{[\sum x^2 - \frac{(\sum x)^2}{n}]}$$

Or $\beta_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$

- **X on Y:**
x=$\alpha + \beta y$
$\alpha = \bar{x} - \beta\bar{y}$

$$\beta_{xy} = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{[\sum y^2 - \frac{(\sum y)^2}{n}]}$$

Or $\beta_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(y-\bar{y})^2}$

- **Properties:**
  Rxy=Ruv
  Rxy=Ryx
  Rxy=$\sqrt{\beta_{xy} * \beta_{yx}}$

- **SSE:**
  $\sum e^2 = \sum(y-\hat{y})^2$
- **Co-eff. Of determination:**

$$R^2 = \frac{\sum(y-\hat{y})^2}{\sum(y-\bar{y})^2}$$

It is percentage that x% times our estimation is right.

3. **Testing and Hypothesis:**
   1. **General Procedure of Testing Mean (for large data):**
      If n>=30. And population is known.

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

And

$$Z = \frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}}$$

X bar = Sample Mean.
$\mu = Population\ Mean$
$\sigma = Population\ S.D$
$\sigma^2 = Pop.Variance$
S^2= Sample Variance
s^2= Unbiased Sample Variance
N= Population
N=Sample Size

Note:

if there is discussion of equality in question then put this in H0.

If there is not discussion of equality in question then put in H1.

2. **Testing of Mean for Small Data:**
   If pop. Is unknown and n<30.
   Apply t-test.

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$$

And s=$\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$ and calculate s according to this formula in 5^{th} step calculation.

3. **Testing of Proportion: (don't apply t-test in this)**
   P= Proportion of Population

q=1-p
n= Sample Size
In this we change mu to P and P0. In 1st step

$Z = \frac{X-np}{\sqrt{npq}}$ for 5th step calculation

4. **Testing of Difference of Two Means:**
Use mu1 and mu2 in hyp.
Z=$\frac{(\overline{x1}-\overline{x2})}{\sqrt{\frac{s1^2}{n1}+\frac{s2^2}{n2}}}$

5. **Testing of Difference of Two Proportions:**
Use P1 and P2 in hyp.
Z=$\frac{(\widehat{P1}-\widehat{P2})}{\sqrt{\frac{\widehat{P1}\widehat{q1}}{n1}+\frac{\widehat{P2}\widehat{q2}}{n2}}}$

$\widehat{P1} = X1/n1$
$\widehat{P2} = X2/n2$
$\widehat{q1} = 1 - \widehat{P1}$
$\widehat{q2} = 1 - \widehat{P2}$