(23) a) one vs all multi dans dansfication using logistic segression splits me detaret into bina multiple binary classification problem; where in each dossification one dans is aneidered. separately us. It all the rost of the dance. For example if we have 3 days men, C1 V8 [C2, C3] (a va [c3, a] c3 vs [c1, c2] will be me classifiers. Thus for At classificaes N- class instances, we will have N binary classificaes. The prediction of danes are made with the model which is most confident. b) one vs one multi class classification using Jugistic sugression splits dataset into multiple binary darsification problems; where in each classification one class is considered ve the classes are generaled. For example, if we have 3 clarkes then, el 18 ca i ignore c3 C2 V8 C3, ignore C1 C1 V8 C3, ignore C2 will be me dossifiers. Thus for N-class instances, we will have N(N-1) binary dessifiers.



belongs to some family of curve as poisson distribution

Now, if are con show that both gamma and poisson distribution belong to exponential family from the above fact will be proved true.

O proving that gamma distribution belongs to exponential

natural exponential family: exp (exp (exp) + c(1,0)

gamma distribution: + f(x) = pn x e px where

 $\chi_{1a,\beta} > 0$

taking log on both sides,

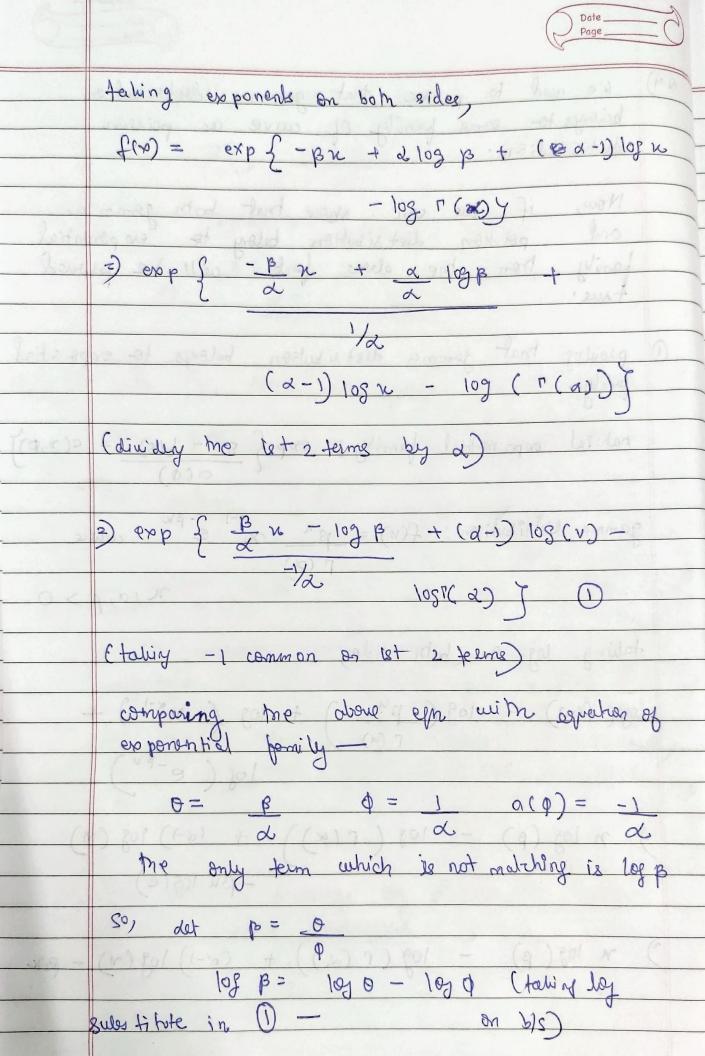
log (froz) = 109 (12m) + 109 (2m2-1) +

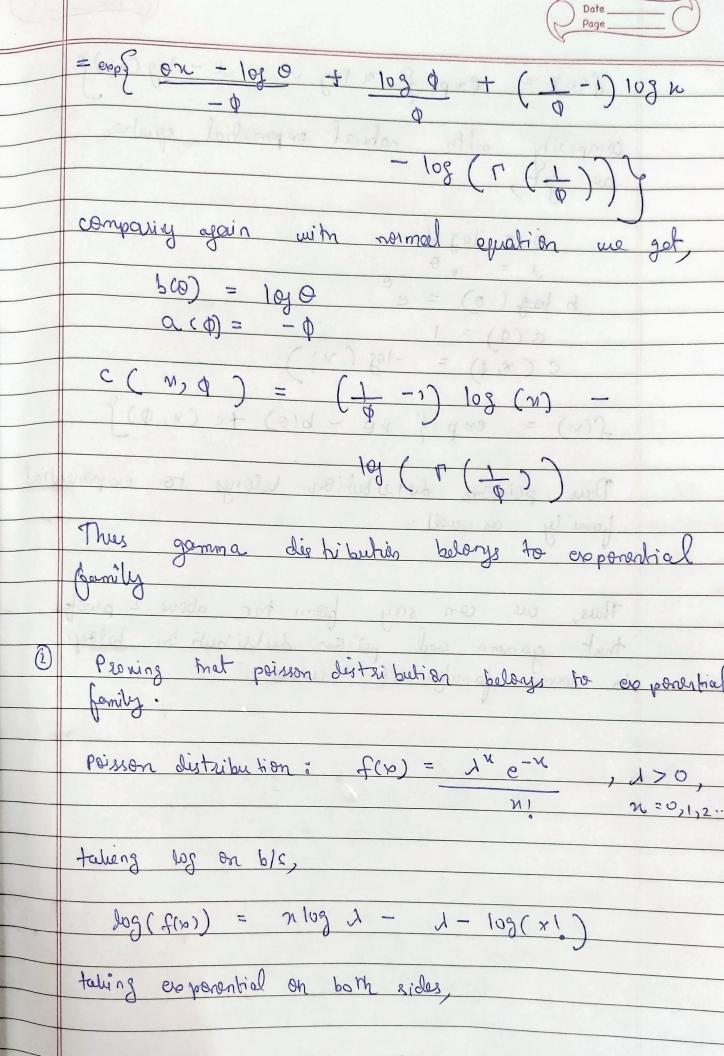
log (e-Bn)

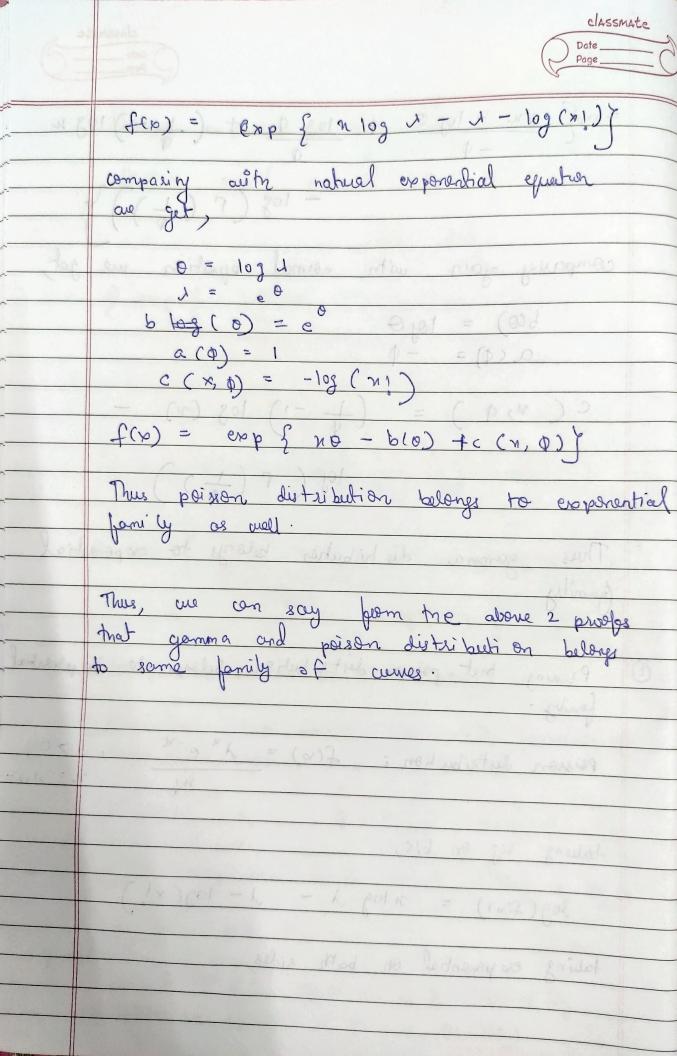
2) n log (p) - 10g (r(x)) + (d-1) 10g (n) - BN 10g(e)

) n log(B) - log(r(Z)) + (d-1) log(m) - pr

1) wi st









07)a) F score is used for besting accuracy of binary classification. It is calculated using precarion and Procesion = true positive four positive + pube positive Recall true positive + rue positive + pale regative Fi sure is me hormanic mean of precesion and reall. FB is a more general F sure) F= a + (1-2) $FB = (B^2+1)PR \quad \text{where} \quad f^2 \text{ peression}$ $B^2P+R \quad R^2 \text{ secall}$ where $d \in [0,1]$ and $g^2 \in [0...0]$ $F_5 = (25+1) PR = 26 PR$ 25P + P 25P + P

corresponding value of
$$\alpha = 25 = 1 - \alpha$$

) 25 d =

2) 2= 0.04

07) b) FB = (1+p2) x precosion x recall (B2 × procession) + recall are can require me abone as -FB = (B2 +1) B² + -1 precision Secold now we can see if B> 1 men FB will give more weight to recall. Thus are can say that whon pie changed from I to 5 men reall will be given more emphasis



k means will perform a good job when duster are spherical in shape and also that the data can be dustored i.e. in the data the point must be dose to each other of similar dusty 2 points must be for away in case of différent clusters. Computsi. Conamance mateix will give us information about individual dustey and not me dissimilarity of between 2 dusters. Also conscionce matrix gives no information about similarity of data points. They me key points for which I means would have performed well one not satisfied.