

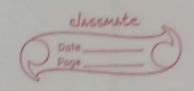
Ouestion 5 Griven D = {]; y; } = 1, what is the log likelihood function of Mi's) guien D? Cruien PrCYi=K|Xi)=Mike-Mi Let us consider a datapoid Ela;, y;) & D Then P(Yi = yilxi) = mi e -h. tinding log (Hi e - Pi = log (µi) + log (e-li)-log(g. = y log (pi) - pi - log (y!) This function has harameter pri and observable data yis Log likelihood of the entire dataset is a summation of the log of each individual datapoint

 $f(\mu_i) = \sum_{i=1}^{\infty} (y_i \log \mu_i) - \mu_i - \log (y_i)$

truce justification of log hi = w'x th

The parameterization of pic is justified due to the fact that pin is the mean of the Cousan Distribution because of which it must always lee 7.0.

Log pic is valid because pic will rever be regative while log(pi) to go can be regative.



	mean and horaneter
	As such hi is aluque a replit and de loss de
	lon (u) = uTx - th
	As such pri is always a valid min as long as long as
	The state of the s
3	Write down the objective function for Parison regression
	We aim to maximize log likelihood of data and we learned log (pi) = wtxi+b is a valid pararela
	learned log (hi) = with is a valid harmely
	n a management
	Objective function = \(\frac{1}{2} \) (y i log (\mu_i) - \mu_i - log (g; !)
	i-T
	$= \sum_{i=1}^{n} (y_i(w_{x_i}^{\dagger} + b) - e^{w_{x_i}^{\dagger} + b} - log(y_i)$
	$\tilde{\lambda} = 1$
	We are trying to maxinize wrt. wfb.
4	Compute gradient
	J. S. Cyilog (*)
	With greifect to W WIXING
	$\nabla f = \frac{1}{2} \sum_{i=1}^{\infty} y_i w_i x_i + y_i b - e e - \log(y_i)$
	$\frac{\partial w}{\partial w} = \frac{1}{2} \frac{1}{2$
	With respect to w $ \nabla f = \frac{1}{2} \sum_{i=1}^{n} y_i w^T x_i + y_i b - e e -log(g_i!) $ $ = \sum_{i=1}^{n} (y_i x_i - x_i e^{w^T x_i + b}) $
	$\tilde{\lambda} = 1$
	With respect to b
	n b wto n l
	Tfi=d Z y. w'xityb-ee-loglyit
	db n=1 Tyth
	$ \nabla f_b = \frac{\partial}{\partial b} \sum_{i=1}^{n} y_i \omega^{\dagger} x_i + y_i b - e^b e^{\omega^{\dagger} x_i} - \log(y_i) $ $ = \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - e^{\omega^{\dagger} x_i + b}) $
	$\hat{\lambda} = 1$

Classiate Company

To marrial objective function, we need to constant what want b using geradient descent similar to oustron I hart 3 of the assignment N = step sign $W_{new} = W_{old} - N & V_{fw}$ $W_{new} = W_{old} - N$

A) Continue

1 p+ x tu 0

0. .c - .r W =

1=1

9-dutixtury

9-16 01 5- 13