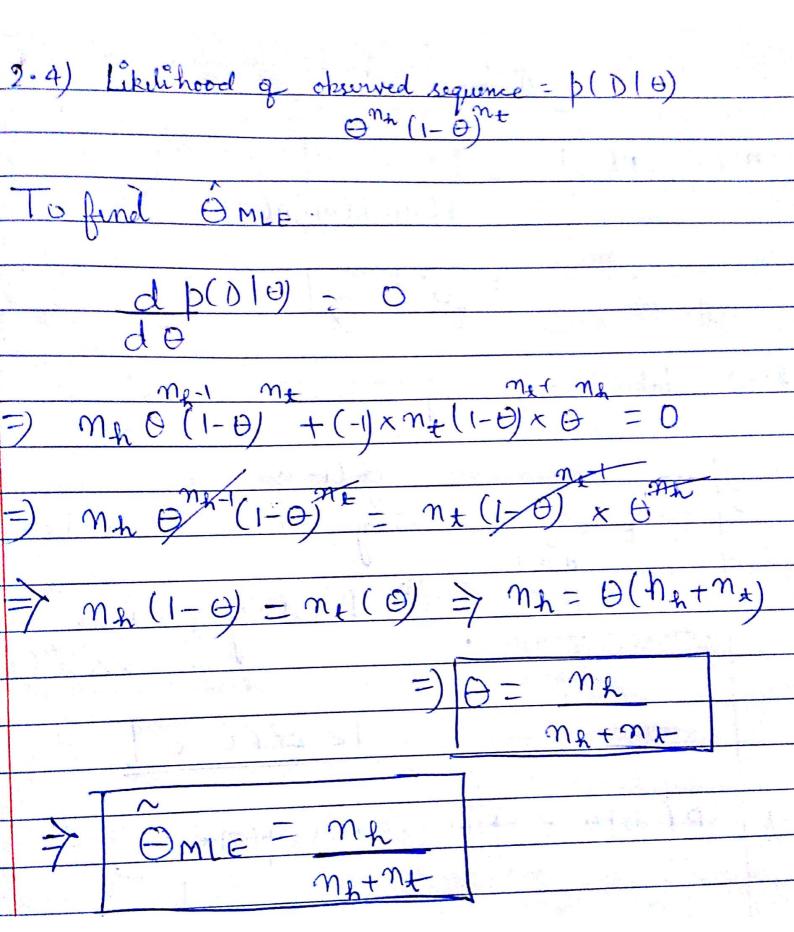
4.3.1) Variance on q Beta distribution = (arb) (Arbei) mean on of Beta destribution. a+b. Let $1-m=1-\frac{a}{a+b}\cdot\frac{a+b-a}{a+b}\cdot\frac{b}{a+b}$ Let n = a+b $\Rightarrow U = \frac{m(1-m)}{(\eta+1)}$ Solving for a & b we get: $A = m^{\eta}, \qquad b = (1-m)^{\eta},$ $\Rightarrow \text{Beta}(m, n) = \frac{1}{B(a, b)} = \frac{1}{B(a, b)} = \frac{1}{B(a, b)} = \frac{1}{B(a, b)}$ $\Rightarrow \frac{1}{B(a, b)} = \frac{1}{B(a,$

4.3.2 m @ lies in the stange (0,1) and Beta distribution is defined on [0,1] hence Beta distribution to Novies in the stange (0,0), on which Gramm distribution is defined. Hence we choose (1) mine Prival for 10.

4.3.3)
Approximating the distribution by mobilent mass at mode. $b(\theta_1, \dots \theta_d | D) \approx b(\theta_1, \dots \theta_d | m_{\text{map}}, m_{\text{map}}).$ without Approximation postinor distribution, $b(\theta_1, \dots \theta_d | D) \approx b(\theta_1, \dots \theta_d | m_{\text{map}}, m_{\text{map}})$

P(D10) = P((H, H, T) 0) 2.1 P(H). P(H). P(T) - \(\theta^2 \big(+ \theta \big) \) Likelihood of seeing 2 Heads and a Tail as calculated above is = 2.2) 02 (1- 0) No. of ways of obtaining 2 Hearts and a Tail are 7 Probability of 2 Heads and a Tail
is 3 X 02 (1-9) HITH T HH 302(1-4) P(D/0)= $b(H) \cdot \cdot \cdot b(H) \times b(T) \times \cdot \cdot \cdot \cdot - \cdot$ mg times

One (1-0) W + punts



3.1) posterior
$$\alpha$$
 (habited \times priore

$$= \beta(D|\theta) \times \beta(\theta)$$

$$= \theta^{h-1}(1-\theta)^{1-1} \times \theta^{m_h}(1-\theta)^{m_t}$$

$$= \theta^{m_h+h-1}(1-\theta)$$

$$\Rightarrow \text{ posterior is a Beta distribution with parameter (h+m_h, t+m_t)}$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

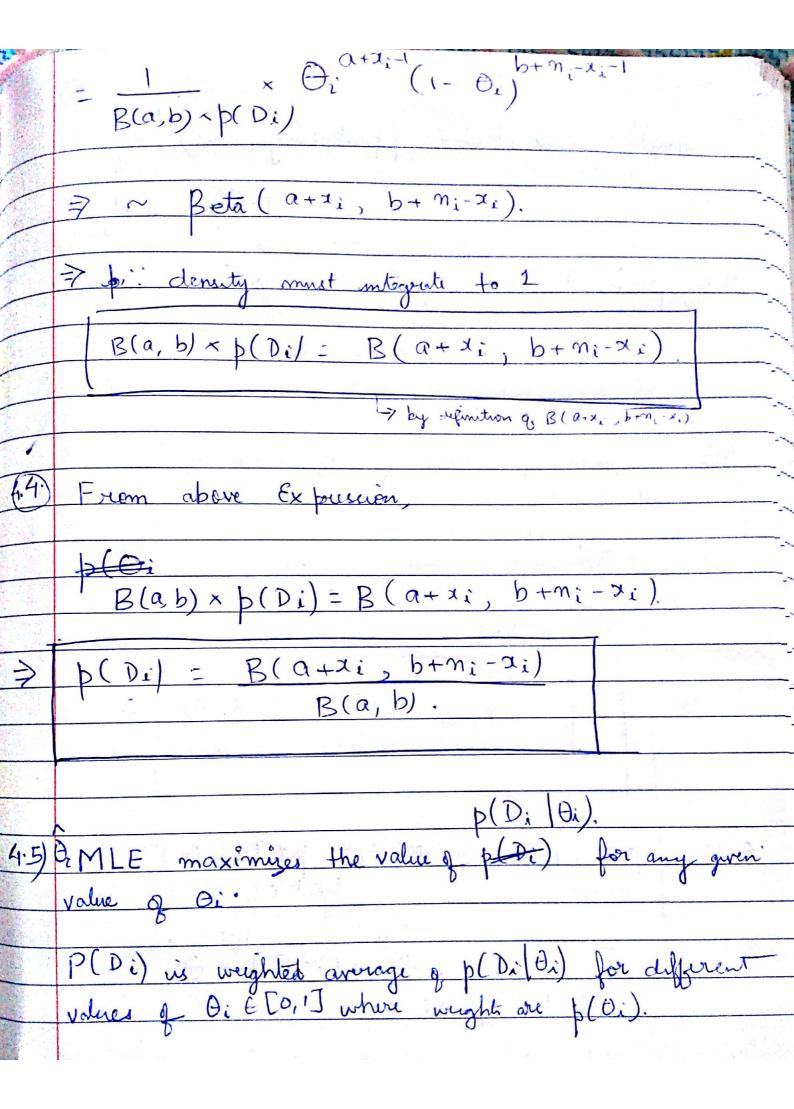
$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

$$\Rightarrow \text{ Posterior } P(\theta|D) \sim \text{ Beta } (m_h+h, t+m_t).$$

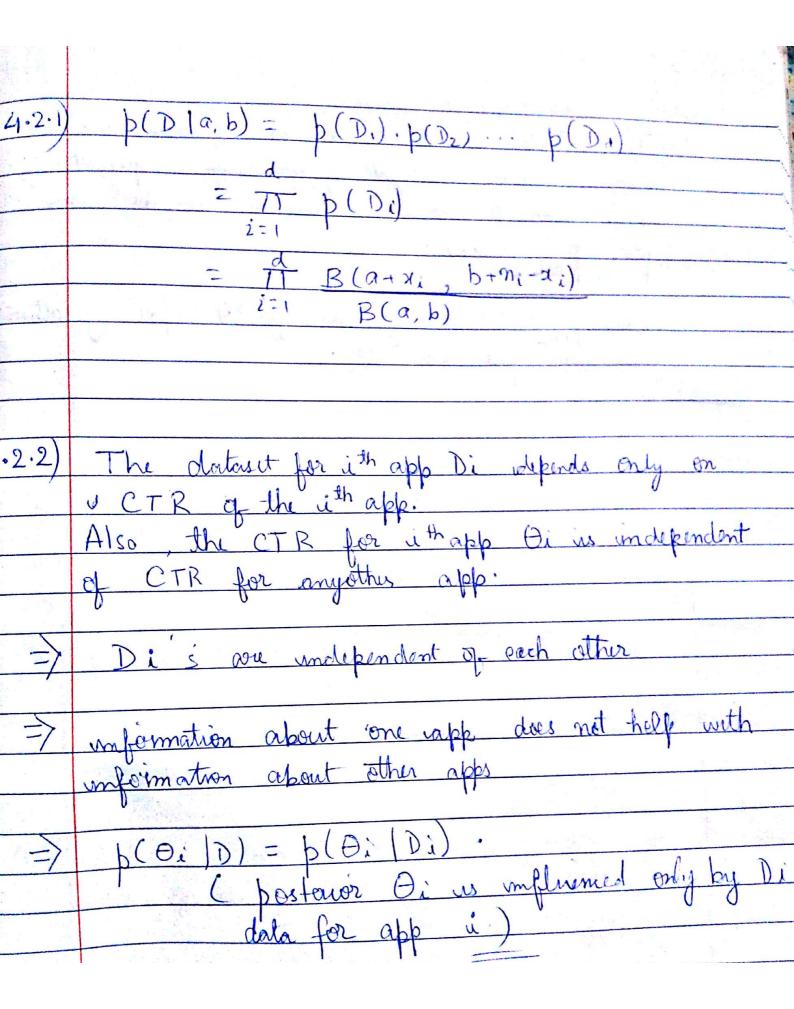
3.3) They converge to 8, the adress value of forebability of Head. As we get more clata the effect of data on the posterior increases and the effect excited by prior electrones.
3.4) MLF give an imbiases estimate of to.
3.4) MLE gives an imbiases estimate of O. MAP and poslower mean are braid owing to the prior which assumes a distribution.
by a which assumes a distribution
3.5) MLE. since we have small date, MAP of
hasterian MODF would give an external biased
under our prior. Since the con is fair MLE
is our best bet.
Barrier Carlotter Carlotte

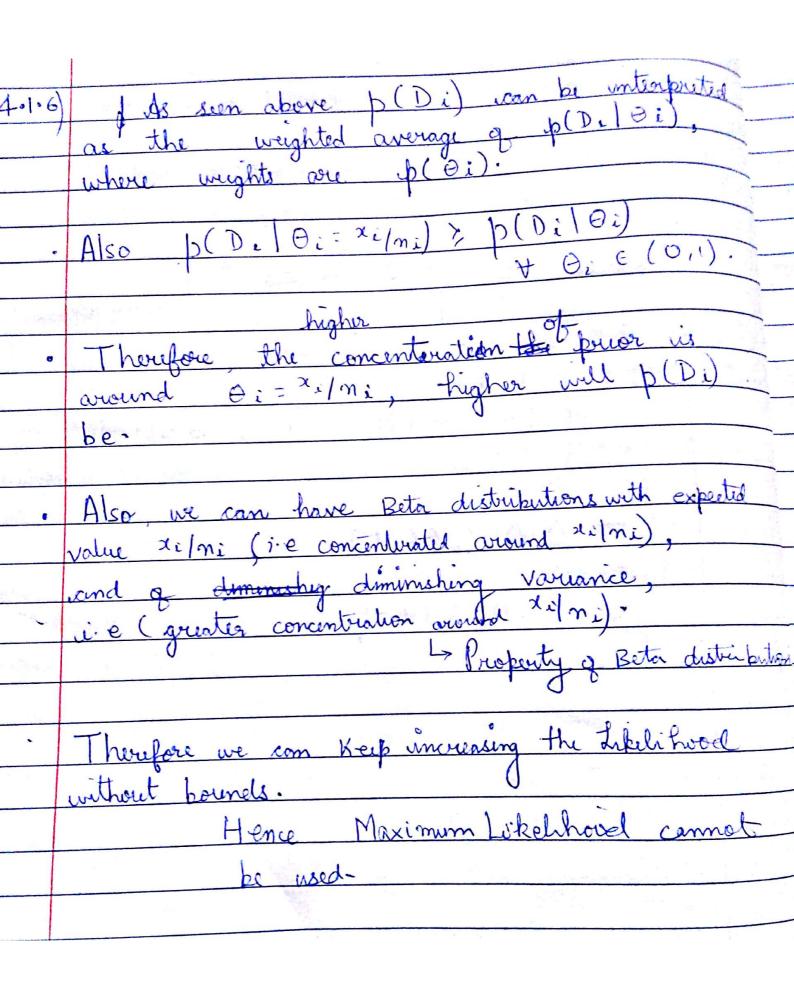
(1- propositing of click) i.i) $b(D_i | \theta_i) = (brob. q click) \times (brob. q non-click)$ $(\theta_i) \times (1-\theta_i)$ 4.2) Sum of prob. of diff values of $\theta_i = 1$. $b(\theta_i) = 1$ $\int_{B(a,b)}^{a-1} \frac{b-1}{B(a,b)} d\theta = 1$ $\theta_i (1-\theta_i) d\theta_i = B(a, b)$ $|b(\Theta_i|P_i) = |b(\Theta_i) \times |b(D_i|\Theta_i)|$ $|b(D_i) \times |b(D_i|\Theta_i)|$ $= \frac{a-1}{x} \frac{b-1}{x} \frac{x_i}{(1-\theta_i)} \frac{m_i-x_i}{x}$ B(a,b) × p(Di)



The wighted average $p(D_i)$ tokes accumes maximum value when, weight of maximum component is I and 0 otherwise which is the cest for $p(D_i)$.

Hence $p(D_i)$ is larger than $p(D_i)$ for any other prior we put on $p(D_i)$.





4.23 posterior for app i: p(0, 10): p(0, 10, posterior ~ Beta (a+xi, b+Mi-xi) > posterior Mean = a+xi - a+xi a+b+n; a+b+n; $MAP = \frac{(a+x_i)-1}{(a+x_i)+(b+m_i-x_i)-2} = \frac{(a+x_i)-1}{(a+b+m_i-2)}$ a+x,-1 a+b+mi-2 Vor siturio SD = $Var = \frac{ab}{(a+b+1)} - \frac{(a+x_i)(b+m_i-x_i)}{(a+b+m_i)}$ $\frac{(a+b)^2(a+b+1)}{(a+b+m_i)} = \frac{(a+b+m_i)(b+m_i-x_i)}{(a+b+m_i)}$ > For App 1 MAP = 6.47 +50-1 (6.47+50)+(11814+6+10000-50) -L = 0.49%

56-47+ 1181.4+10000-50 0.0050 974 0.50474-/ posterior SD 6.47×5001181.4 (6.47+50) x (1181.4+10000-50) SD= (6.47+1181.4+10000)2 (6.47+1181.4+10000+1) 0.000 6699 0.06699 00

4.2.4	MAP	Posterior Mean	Costerior SU
Aph!	0.5%	0.49%	0.07%
App 2	0.78.1.	0.791.	0.061.
Abb3	0.3./_	0.3.1	0.02%
NEWSCHOOL BOOK BY A SECTION OF	0.42-1.	0.5%	0.191.
App 4	0.459%	0.54.1.	0.21%
App 5 App 6	0.5441.	0.62/	0.221
Aft			