Q 2.6.1

$y_i \in \{c_1, c_2, \dots c_{nj}\}$

we have to classify y_i such that prosterior is max.
or c_i here.

$$y_i^* = \arg\max P(c_i / w_i)$$

$$y_i^* = \arg\max_{c_i} \underbrace{P(w/c_i)}_{\text{Bernoulli}} \cdot \underbrace{\text{Prior}}_{\text{Beta}(c_i)}$$

$$\text{Bernoulli}(w/c_i) = c_i^{w} (1-c_i)^{(1-w)}$$

$$\text{Beta}(c_i) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} \cdot c_i^{\alpha-1} \cdot (1-c_i)^{\beta-1}$$

$$\text{Likelihood} = \arg\max_{c_i} \prod_{i=1}^{n} \left[c_i^{w} (1-c_i)^{(1-w)}\right] \left| \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma\beta} \cdot c_i^{\alpha-1} (1-c_i)^{\beta-1}\right.$$

taking log to Convert $\prod \longrightarrow \Sigma$

$$LL = \Sigma \ln\left(c_i^{w} (1-c_i)^{1-w}\right) + \ln\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} \cdot c_i^{\alpha-1} (1-c_i)^{\beta-1}\right.$$

taking derivative & putting equal to zero

$$\Sigma \frac{\partial}{\partial c_i} \ln(c_i^{w} (1-c_i)^{1-w}) = \frac{1}{c_i} \Sigma w_i - \frac{1}{1-c_i} \Sigma(1-w_i)$$

Scanned by CamScanner

Q2.6.1 contd

$$\frac{\partial}{\partial c_i} \log\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} \cdot c_i^{\alpha-1}(1-c_i)^{\beta-1}\right)$$

$$= \frac{\partial}{\partial c_i} \log\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}\right) + \frac{\partial}{\partial c_i} \log c_i^{\alpha-1} \cdot (1-c_i)^{\beta-1}$$

$$= 0 + \frac{\alpha-1}{c_i} - \frac{\beta-1}{1-c_i}$$

Now we put $\frac{\partial LL}{\partial c_i} = 0$

$$0 = \frac{1}{c_i}\sum w_i - \frac{1}{1-c_i}\sum(1-w_i) + \frac{\alpha-1}{c_i} + \frac{\beta-1}{(1-c_i)}$$

$$c_i\left[\sum(1-w_i) + \beta-1\right] = (1-c_i)\left[\sum w_i + \alpha-1\right]$$

Scanned by CamScanner

$$c_i \left[\sum (1 - w_i) + \sum w_i + \beta - 1 + \alpha - 1 \right] = \sum_{i=1} w_i + \alpha - 1$$

$$c_i \left[\underset{\downarrow}{\sum 1} + \beta + \alpha - 2 \right] = \sum w_i + \alpha - 1$$

$N \rightarrow$ total N.

$\because$ $\sum w_i$ denote word to be ~~fond~~ found

we can say N^f

$$\boxed{c_i = \frac{N^f + \alpha - 1}{N + \beta + \alpha - 2}}$$

$\rightarrow$ smoothing.

Q2.6.3. Problem with ML estimates

① ML estimates does not incorporate any prior knowledge and thus dont generate estimate of certainity of results.

② ML estimates just represent proportions. i.e μ.

~~MAP.~~

MAP.

~~we have prior beleifs, without seeing the data~~ the

→ we have prior beleifs about μ, even before observing the data.

→ with Baye's law we convert these prior beleifs to posterior ~~Prop~~ Probability. $P(\mu/X)$

→ MAP estimates are ~~reported~~ produced by maximizing the ~~poste~~ ~~log~~ likelihood func.

and hence we can inject out prior belifs in to new estimates.

Scanned by CamScanner