Date: ___.___.___

Representation Learning

Q3(a) **Binomial**

$$\mathcal{L}(x;p) = \prod_{i=1}^{N} {}^{n}C_{x_i}\, p^{x_i} (1-p)^{n-x_i}$$

$$l(x;p) = \sum_{i=1}^{N} \log\left({}^{n}C_{x_i}\, p^{x_i} (1-p)^{n-x_i}\right)$$

$$= \sum_{i=1}^{N} \left[\log {}^{n}C_{x} + x_i \log p + (n-x_i)\log(1-p)\right]$$

$$\min_{p} \left(l(x;p)\right) \Rightarrow \frac{dl}{dp} = 0$$

$$\Rightarrow \sum_{i=1}^{N} \left[\frac{x_i}{p} + \left(\frac{n-x_i}{1-p}\right)(-1)\right] = 0$$

$$\frac{\sum_{i=1}^{N} x_i}{p} - \frac{nN}{(1-p)} + \frac{\sum_{i=1}^{N} x_i}{(1-p)} = 0$$

$$\sum_{i=1}^{N} x_i (1-p) = \left(nN - \sum_{i=1}^{N} x_i\right)(p)$$

$$\sum_{i=1}^{N} x_i - \cancel{\sum_{i=1}^{N} x_i p} = nNp - \cancel{\sum_{i=1}^{N} x_i p}$$

$$p = \frac{\sum_{i=1}^{N} x_i}{n.N}$$

apsara

Date: ____.____.____

b) $f(x;\lambda) = \dfrac{\lambda^{x} e^{-\lambda}}{x!}$

$L(x;\lambda) = \prod_{i=1}^{N} \dfrac{\lambda^{x_i} e^{-\lambda}}{x_i!}$

$l(x;\lambda) = \sum_{i=1}^{N} \log\left(\dfrac{\lambda^{x_i} e^{-\lambda}}{x_i!}\right)$

$= \sum_{i=1}^{N} \left[x_i \log\lambda - \lambda - \log(x_i!)\right]$

$\dfrac{\partial l}{\partial \lambda} = 0$

$\Rightarrow \sum_{i=1}^{N} \left[\dfrac{x_i}{\lambda} - 1\right] = 0$

$\lambda = \dfrac{\sum_{i=1}^{N} x_i}{N}$

c) $f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \in R \\ 0 & x \notin R \end{cases}$

$L(x;\lambda) = \prod_{i=1}^{N} f(x;\lambda)$

$l(x;\lambda) = \sum_{i=1}^{N} \log(\lambda e^{-\lambda x_i})$

$= \sum_{i=1}^{N} [\lambda - \lambda x_i]$

$$\frac{\partial l}{\partial \lambda} = 0$$

$$\Rightarrow \sum_{i=1}^{N}\left[\frac{1}{\lambda} - x_i\right] = 0$$

$$\frac{N}{\lambda} = \sum_{i=1}^{N} x_i$$

$$\lambda = \frac{N}{\sum_{i=1}^{N} x_i}$$

d) $f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$$l(x; \mu, \sigma) = \sum_{i=1}^{N}\left[\log\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$\frac{\partial l}{\partial \mu} = 0$$

$$\Rightarrow \sum_{i=1}^{N} \frac{2(x_i-\mu)}{2\sigma^2} = 0$$

$$\mu = \sum_{i=1}^{N} \frac{x_i}{N}$$

$$\frac{\partial l}{\partial \sigma} = 0$$

$$\Rightarrow \sum_{i=1}^{N}\left[\frac{1}{\sigma} - \frac{(x_i-\mu)^2}{2\sigma^3}(-2)\right] = 0$$

$$\Rightarrow \sum_{i=1}^{N} \frac{(x_i-\mu)^2}{\sigma^2} = N$$

$$\boxed{\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}}$$

apsara

Date: __.__.__

e) $f(x;\mu,b) = \frac{1}{2b} e^{-|x-\mu|/b}$

$L(x;\mu,b) = \prod_{i=1}^{N} \frac{1}{2b} e^{-|x-\mu|/b}$

$l(x;\mu,b) = \sum_{i=1}^{N} \left[\log\frac{1}{2b} - \frac{|x_i-\mu|}{b} \right]$

$\frac{\partial l}{\partial \mu} = 0$,

$\Rightarrow \frac{\partial l}{\partial \mu} = \frac{1}{b} \sum_{i=1}^{N} \frac{\partial |x_i-\mu|}{\partial \mu}$

$\because \frac{\partial |x|}{\partial x} = \frac{\partial \sqrt{x^2}}{\partial x} = x(x^2)^{-1/2} = \frac{x}{|x|} = sgn(x)$

$\Rightarrow \frac{\partial l}{\partial \mu} = \sum_{i=1}^{N} \frac{sgn(x_i-\mu)}{b} = 0$

If N is odd & $\underline{x}$ is sorted

$\sum_{i=1}^{N} sgn(x_i-\mu) = 0$ when $\mu_{MLE} = x_{\frac{N+1}{2}}$

If N is even and $\underline{x}$ is sorted

$\sum_{i=1}^{N} sgn(x_i-\mu) = 0$ when

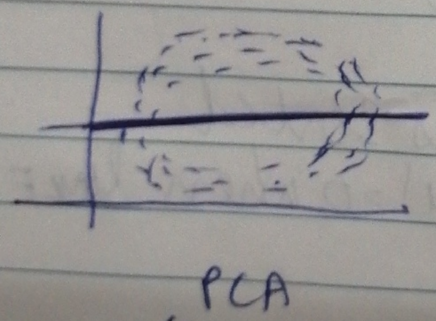
$\mu_{MLE} = \frac{\left(x_{N/2} + x_{\frac{N}{2}+1}\right)}{2}$

$$\frac{\partial l}{\partial b} = 0$$

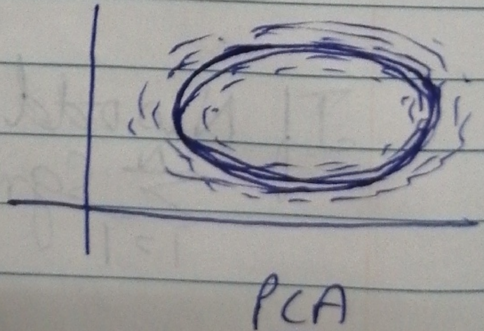
$$\Rightarrow \sum_{i=1}^{N}\left(\frac{-1}{b}\right) + \sum_{i=1}^{N}\frac{|x_i - \mu_{MLE}|}{b^2} = 0$$

$$\therefore \boxed{b = \sum_{i=1}^{N}\frac{|x_i - \mu_{MLE}|}{N}}$$

Q2 ① The Standard PCA always finds linear principal components to represent the data in lower dimension. Sometimes we need non-linear principal components.



PCA (linear)



PCA (nonlinear)

② Suppose a simple case with 3 independent variables x_1, x_2, x_3 and the output y and suppose now that $x_3 = y$. and so you should be able to get a 0 error model.

Now suppose that in the training set the

apsara

Date:___.___.___

Variation of y is very small and so also the variation of x_3.

Now if you run PCA and you decide to select only 2 variables you will obtain a combination of x_1 and x_2. So the information of x_3 that was the only variable able to explain y is lost.