DSCI 552 Assignment 3 - Theory Questions

Cheng-Ruei Chiu

Q1 perform MLE to find the Parameters of a Gaussian Distribution

$$P(x \mid \mu, V) = \frac{1}{\sqrt{2\pi V}} e^{-\frac{(x-\mu)^2}{2V}}$$

$$P(X \mid \mu, V) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi V}} e^{-\frac{(xi-\mu)^2}{2V}}$$

take the log of the equation

$$\log P(x \mid \mu, V) = \log \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi V}} e^{\frac{-(x_i-\mu)^2}{2V}}$$

$$= \sum_{i=1}^{n} \log \left(\frac{1}{\sqrt{2\pi V}} + e^{\frac{-(x_i-\mu)^2}{2V}} \right)$$

$$= \sum_{i=1}^{n} \left(\log \left(1 - \sqrt{2\pi V} \right) + \log(e) \cdot \left(-\frac{(x_i-\mu)^2}{2V} \right) \right)$$

$$= \sum_{i=1}^{n} \left(-\log \sqrt{2\pi V} + \left(-\frac{1}{2} \frac{(x_i-\mu)^2}{V} \right) \right)$$

$$= \sum_{i=1}^{n} \left(-\frac{1}{2} \log(2\pi V) - \frac{1}{2} \frac{(x_i-\mu)^2}{V} \right)$$

$$= -\frac{n}{2} \log(2\pi V) - \frac{1}{2V} \sum_{i=1}^{n} (x_i-\mu)^2$$

Find the Maximum estimate of μ

$$\frac{d}{d\mu} \log P(x \mid \mu, V) = \frac{d}{d\mu} \left(-\frac{1}{2V} \sum_{i=1}^{n} (x_n-\mu)^2 \right)$$

$$= \sum_{i=1}^{n} \left(-\frac{1}{2V} (x_i-\mu)^2 \right)$$

$$= \sum_{i=1}^{n} \left(-\frac{1}{2V} \times 2(x_i-\mu) \cdot (-1) \right)$$

$$0 = \frac{1}{V} \sum_{i=1}^{n} (x_i-\mu)$$

$$0 = \sum_{i=1}^{n} (x_i-\mu)$$

$$0 = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \mu$$

$$0 = \sum_{i=1}^{n} x_i - n\mu$$

$$n\mu = \sum_{i=1}^{n} x_i$$

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

Find the Maximum of V

$$\frac{d}{dV}\log P(x|\mu,V) = \frac{d}{dV}\left(-\frac{n}{2}\log(2\pi V) - \frac{1}{2V}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= \frac{d}{dV}\left(-\frac{n}{2}\log(2\pi V)\right) + \frac{d}{dV}\left(-\frac{1}{2V}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= -\frac{n}{2}\cdot\frac{d}{dV}\cdot\log(2\pi V) + \frac{d}{dV}\left(-\frac{1}{2V}\right)\cdot\sum_{i=1}^{n}(x_i-\mu)^2)$$

$$= -\frac{n}{2}\cdot\frac{1}{2\pi V}\times 2\pi + \sum_{i=1}^{n}\left(\frac{1}{2V^2}\times(x_i-\mu)^2\right)$$

$$= -\frac{n}{2V} + \frac{1}{2V^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$0 = \frac{1}{2V}\left(-n + \frac{1}{V}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$nV = \sum_{i=1}^{n}(x_i-\mu)^2$$

$$V = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

Q2

2) Given the following statistics, what is the probability that a man has a particular disease in a town if he has been tested positive from a home testing kit

- One percent of men have the disease
- 90% of men who have the disease test positive on the home kit
- 8% of men who use the kit will have false positives.

From the information we know that

$P(D) = 0.01$

$P(T|D) = 0.9$

$P(T') = 0.08$

We want to get

$$P(D|T) = \frac{P(T|D)\cdot P(D)}{P(T)}$$

And we know that

$$P(T) = P(T|D)\cdot P(D) + P(T|D')\cdot P(D')$$

$$= 0.9 \times 0.1 + 0.08 \times 0.99$$

$$= 0.0882$$

So

$$P(D|T) = \frac{0.9 \times 0.01}{0.0882} = 0.102$$

Thus, the probablity of a man has disease in town and has tested positive is 10.2%