Machine Learning

EX 1 2. $f(x|\lambda) = \lambda e^{-\lambda x}$ for $\lambda > 0$, $x_i \geq 0$

① Likelihood funktion $L(\lambda|x_i) = \prod_{j=1}^{n} f(x_j|\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_j}$

② Dann
$$\frac{d}{d\lambda_i} L(\lambda | x_i \cdots x_n) = \frac{d}{d\lambda} \lambda^n [e^{-\lambda(x_1 + \cdots + x_n)}] \quad = \lambda e^{-\lambda x_i}$$
$$= \frac{d}{d\lambda} \log(\lambda^n [e^{-\lambda(x_1+\cdots+x_n)}])$$
$$= \frac{d}{d\lambda} \log(\lambda^n) + \log[e^{-\lambda(x_1+\cdots+x_n)}]$$
$$= \frac{d}{d\lambda} n\log(\lambda) - \lambda(x_1 + \cdots + x_n)$$
$$= n \cdot \frac{1}{\lambda} - (x_1 + \cdots + x_n)$$

③ sei $0 = n \cdot \frac{1}{\lambda} - (x_1 + \cdots + x_n)$

Dann $\Leftrightarrow \hat{\lambda} = n/(x_1 + \cdots + x_n)$

D.h estimator $\hat{x}_n \Rightarrow \frac{\sum_{i=1}^{n} x_i}{n}$

1. Gaussian normal distribution $N(x|\mu, \sigma^2) \rightsquigarrow f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

① $\Rightarrow L(\mu, \sigma^2) = \prod_{i=1}^{n} f(x|\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x_i-\mu)^2}{2\sigma^2})$
$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

② Dann
$$\ln L(\mu, \sigma^2) = \ln(2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}\right)$$
$$= \ln(2\pi\sigma^2)^{-\frac{n}{2}} + \ln \cdot \exp\left(-\sum_{i=1}^{n}\left(\frac{x_i-\mu)^2}{2\sigma^2}\right)\right)$$
$$= -\frac{n}{2}\ln 2\pi\sigma^2 - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}$$
$$= -\frac{n}{2}(\ln 2\pi + \ln\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

③ sei
$$\begin{cases} \frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0 \\ \frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \end{cases} \Rightarrow \begin{cases} \sum_{i=1}^{n}(x_i - \mu) = 0 \quad (1) \\ \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \end{cases}$$

$\Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$.

D.h. fur μ gilt $\mu = \frac{1}{\bar{x}}$; $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$

Ex2. $P(\theta) = Dir(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \cdot 1$ mit $B(\alpha) = \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{K} \alpha_i)}$ $\alpha = (\alpha_1 \cdots \alpha_K)$

$P(D|\theta) = Mu(x|n, \theta) = \prod_{k=1}^{K} \theta_k^{N_k}$

Dann $P(\theta|D) = P(D|\theta) \cdot P(\theta) = \prod_{k=1}^{K} \theta_k^{N_k} \theta_k^{\alpha_k - 1} = \prod^{K} \theta_k^{\alpha_k + N_k - 1}$

$= Dir(\theta | \alpha_1 + N_1 + \dots + \alpha_K + N_K)$

$= Dir(\theta | \alpha + x)$