Exercise 1

1) likelyhood function $= \prod_{i=1}^{n} \frac{\mu_i^k}{k!} \exp(-\mu_i)$

$\log(\text{likelyhood function}) = L = \sum_{i=1}^{n} [k \log(\mu_i) - \log(k!) - \mu_i]$

2) Poisson distribution belongs to discrete probability distribution. So can use μ_i to represent the expected number of events happened per unit time. So μ_i is continuous and non-negative. Firstly, because it is continuous, we can let $w^T x_i + b$ (this is also continuous) relate to μ_i. Secondly, μ_i is non-negative, so we need use a mathmatical way to express this limitation. $\log(\mu_i)$ is resonable because true number in logarithmic function is non-negative. Above all, we can let $\log(\mu_i) = w^T x_i + b$

3) Know it is known that $\log(\mu_i) = w^T x_i + b$.

so,

$$\hat{w} = \arg\max_{w} L = \arg\max_{w} \sum_{i=1}^{n} [k \log(w^T x_i + b) - \log(k!) - w^T x_i - b]$$

likelyhood function used the multiply of probability that Y_i equals to k subject to x_i. For each input x_i, we want probability is high (this means high accuracy). If each probability is high, then the multiply of probability is high. Therefore, $\hat{w} = \arg\max_{w} L$ can gives resulted parameter maximizing the lable accuracy.

4) $$\frac{dL}{dw} = \sum_{i=1}^{n} \left(\frac{K x_i}{w^T x_i + b} - x_i \right)$$

$$\frac{dL}{db} = \sum_{i=1}^{n} \left(\frac{K}{w^T x_i + b} - 1 \right)$$

Input : $X \in R^{n \times d}$, $Y \in R^{n}$, $w_0 = 0_d$, $b_0 = 0$, max_pass $\in N$, $\eta > 0$, tol > 0

output : w, b

for $t = 1, 2, \ldots,$ max_pass do

$w_t \leftarrow w_{t-1} - \eta\, dw$

$b_t \leftarrow b_{t-1} - \eta\, db$

if $\|w_t - w_{t-1}\| \leq$ tol then

break

$w \leftarrow w_t$, $b \leftarrow b_t$