Assignment 1

problem 3

3) We are given with the data of tumour volume V mm^3 and T number of years until death.
Let V denote the volume of the tumour in mm^3 and $\hat{\lambda}$ denote the observed rate of death (persons/year).
Given it was observed a linear relation between $\hat{\lambda}$ and V $\Rightarrow$ $\hat{\lambda} = \theta_0 + \theta_1$V
Applying exponential distribution with parameter $\hat{\lambda}$ to find survial time $\Rightarrow$
P(X=t) = $\hat{\lambda}e^{-\hat{\lambda}t}$
Likelihood function: It is defnined as the function L(θ_0,θ_1) = Π Prob(T|$\hat{T}$), where $\hat{T}$ is the computed from exponential distribution
L(θ_0,θ_1) = $\Pi_i\ (\theta_0 + \theta_1 V_i)e^{-(\theta_0+\theta_1 V_i)T_i}$
In order to find the maximum vaue of this function, we need to find the values of θ_0 and θ_1 which would give maximum value for the above function. This is done by setting $\triangledown$L = 0 and evaluating for θ_0,θ_1.

Loss based on maximum likelihood:It is defined as -log(L(θ_0, θ_1))
Loss = -log($\Pi_i\ (\theta_0 + \theta_1 V_i)e^{-(\theta_0+\theta_1 V_i)T_i}$)
Loss = $[\Sigma_i(\theta_0 + \theta_1 V_i)T_i - log(\theta_0 + \theta_1 V_i)]$

Empirical Risk (R): It is defined the loss averaged over the data set.
R = 1/N $\Sigma_i((\theta_0 + \theta_1 V_i)T_i - log(\theta_0 + \theta_1 V_i))$