Problem 1

Yes. Since all data was drawn from a distribution parameterized by θ^0, and we are measuring the likelihood of the data given those parameters (i.e. maximum likelihood), then since the loss is the log-likelihood of the data, the data will be most likely at the parameters θ^0 which parameterize the distributions. As such, θ^0 will be an optimum.

Problem 2

Problem 3

Problem 4

Problem 5

Problem 6