Problem 1. A prize is hidden behind one of three doors A, B, and C. The contestant picks a
door, say A, but it is left closed. The host opens door C and shows that there is no prize behind
it. Should the contestant change his mind and select door B?

Formulate this problem as Bayes inference. What is the prior probability for the position of
the prize before door C was opened? What is the posterior probability after door C has been
opened? What should the contestant’s decision rule be?

Problem 2. Suppose you have an image $I$ and want to decide whether it is a Sea Bass $S$ or a
Flounder $F$. Formulate this as a decision problem in terms of distributions $P(I|\omega) \& P(\omega)$, where
$\omega \in \{S, F\}$. Define the loss function, the risk, the Bayes rule and the Bayes risk. Write down the
loss function if all errors are weighted equally.

Now suppose we have a set of labelled samples $\{(I_\mu, \omega_\mu) : \mu = 1, ..., N\}$ (i.e. each of the $N$
images is labelled as being Sea Bass or Flounder). Define the empirical risk and say how it relates
to the risk. Describe at least two strategies for obtaining a decision rule to label new images as
Sea Bass or Flounder. Describe the difference between generalization and memorization.

Problem 3.
Suppose the feature space is a two-dimensional plane and the two likelihood functions
$p(x, y|\omega_1), p(x, y|\omega_2)$
are Gaussians with means $\mu_1 = (2, 8), \mu_2 = (8, 2)$, and identical variances $\sigma_1^2 = \sigma_2^2 = 1$. The prior
probabilities are $P(\omega_1) = 4P(\omega_2)$. Derive the Bayes rule and Bayes risk for this problem assuming
that the loss function weights all errors equally. Draw the decision boundary in the two-dimensional
plane.

Now repeat the problem but changing the variances of the likelihood functions, so that $\sigma_1^2 = 4$
and $\sigma_2^2 = 16$.

Problem 4. Define the errors for the two class discrimination problem. Derive, and show
graphically, the form of the ROC curve if the data is generated by two Gaussian distributions with
the same variance $\sigma^2$: $P(x|\mu_1, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu_1)^2/(2\sigma^2)}$ and $P(x|\mu_2, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu_2)^2/(2\sigma^2)}$. 

1
Problem 5. Let \( p_1(x) \) and \( p_2(x) \) be two one-dimensional Gaussian distributions \( p_1(x) = N(\mu_1, \sigma_1^2) \) and \( p_2(x) = N(\mu_2, \sigma_2^2) \). Compute the entropy of \( p_1(x) \), the difference in the entropies of \( p_1(x) \) and \( p_2(x) \), and the Kullback-Leibler divergence between \( p_1(x) \) and \( p_2(x) \).

Problem 6. Use the Maximum Entropy principle to derive a distribution \( P(x|\psi) \) defined for \( 0 \leq x < \infty \) with constraint \( \sum_x xP(x) = \psi \).

What is the distribution for \( M \) samples \( x_1, \ldots, x_M \) independently and identically distributed from \( P(x|\psi) \)? What is the sufficient statistic for this distribution? Use Maximum Likelihood to estimate the parameter \( \psi \) of the distribution.