**SGN-2506 Introduction to Pattern Recognition, Fall 2011, Exam 13-Dec-2011 / Jari Niemi**

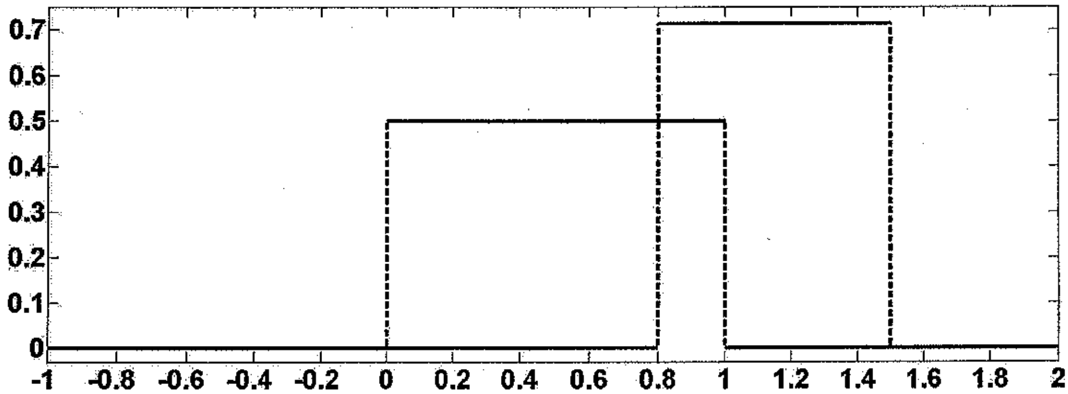
**Perform all the problems 1-5. Each of them is worth of 6 points. No literature. Needed formulas are given. Use of your own calculator allowed. <u>All computational steps must be shown and explained briefly.</u> You can have this exam question paper also after the exam.**

Problems:

**1.** Consider a one-dimensional two-class classification problem with equal priors and continuous uniform class-conditional densities. In the figure below, it is shown these class-conditional densities multiplied by the priors. Based on the figure (approximate the needed quantities just by looking at it)

a) find the decision regions for the Bayes minimum error rate classifier, and (3 p.)

b) compute the Bayes (minimum) error (rate). (3 p.)




**Help:** $P(\omega|x) = \frac{P(\omega)p(x|\omega)}{p(x)}$, Bayes error = $P(error)$ = $\int_{-\infty}^{\infty} P(error|x)p(x)dx$, where $P(error|x)$is the probability that a given feature x is misclassified by the Bayes minimum error rate classifier. Remember that the probability density function of the one-dimensional continuous uniform distribution is of the form

$$p_{\text{uniform}}(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b, \\ 0, & \text{otherwise,} \end{cases}$$ where $a$ and $b$ are constants.

**2.** The two confusion matrices below correspond to two classifiers.

$$\begin{bmatrix} 50 & 0 & 0 \\ 0 & 47 & 1 \\ 0 & 3 & 49 \end{bmatrix} \qquad \begin{bmatrix} 50 & 0 & 2 \\ 0 & 50 & 2 \\ 0 & 0 & 46 \end{bmatrix}$$

a) How many classes are there in total in the underlying classification task? (1 p.)

**Please, turn over →**

b) Approximate the test errors for both classifiers. (2 p.)

c) Discuss the differences in the classification performance between the two classifiers. (3 p.)

**Help:** Test error $\approx \frac{\sum_{i=1}^{c}\sum_{j=1}^{c} s_{ij} - \sum_{i=1}^{c} s_{ii}}{\sum_{i=1}^{c}\sum_{j=1}^{c} s_{ij}}$, where $s_{ij}$ is the entry of the confusion matrix at the position $i, j$.

**3.** The following training data is obtained (mixed sampling) from two classes, $\omega_1$ and $\omega_2$, respectively:

$$D_1 = \{[1,2]^T, [0,1]^T, [0,3]^T\}, \quad D_2 = \{[2,-1]^T, [3,-1]^T, [3,1]^T, [4,0]^T, [4,1]^T, [4,3]^T\}.$$

a) Draw the data into a two-dimensional coordinate system using different markers for $D_1$ and $D_2$. (2 p.)

b) Find (by looking at your picture) a line (and its equation) which separates the training samples $D_1$ and $D_2$ perfectly. (2 p.)

c) Interpret the line found in Item b) as a linear discriminant boundary and classify the test point $[0.75, 1.5]^T$ with it (based on the Bayes minimum error rate classification principle). Explain. (2 p.)

**4.** Use the data in Problem 3 as the training data and classify the point $[0.75, 1.5]^T$ with the nearest neighbor classifier (based on the Bayes minimum error rate classification principle). Explain. Distances are Euclidean, that is, $||\mathbf{x}|| = \sqrt{\mathbf{x}^T\mathbf{x}}$. (6 p.)

**5.** Write an essay (approximately 200-600 words) on the topic *Maximum likelihood classifier.* (6 p.)

**Instruction:** Answer (at least) to the following: What assumptions must be done on the classification problem in order to be able to apply the maximum likelihood principle along with the Bayes minimum error rate classification rule?, How the maximum likelihood estimation is used in the training of a classifier?, What error sources may occur?, How the maximum likelihood training differs from the other training strategies studied in the course?