DATA.ML.200, Exercise set 2.

Question 1.

$n \in [0, \cdots, 9]$

$$s_n = \begin{cases} -1 & 0 \leq n < 5 \\ 1 & 5 \leq n < 10 \end{cases}$$

Optimal decision rule given in lecture slides:

Reject H_0, accept H_1 if

$$\sum_{n=0}^{N-1} x_n s_n > \gamma' = \sigma^2 ln(\gamma) + \frac{1}{2} \sum_{n=0}^{N-1} s_n^2$$

In the case of given example, the decision boundary takes the form:

$$\gamma' = \sigma^2 ln(\gamma) + \frac{1}{2} \sum_{n=0}^{N-1} s_n^2 = \sigma^2 ln(\gamma) + \frac{1}{2} \sum_{n=0}^{N-1} 1 = \sigma^2 ln(\gamma) + 5$$

Likelihood ratio test statistic can be simplified to:

$$\sum_{n=0}^{N-1} x_n s_n = \sum_{n=0}^{4} x_n s_n + \sum_{n=5}^{9} x_n s_n = -\sum_{n=0}^{4} x_n + \sum_{n=5}^{9} x_n = \sum_{n=0}^{4} (x_{9-n} - x_n)$$

Therefore the test can be formalized as:

Reject H_0, accept H_1 if

$$\sum_{n=0}^{4} (x_{9-n} - x_n) > \sigma^2 ln(\gamma) + 5$$

Natural choice for γ is 1. Then the test takes the form of

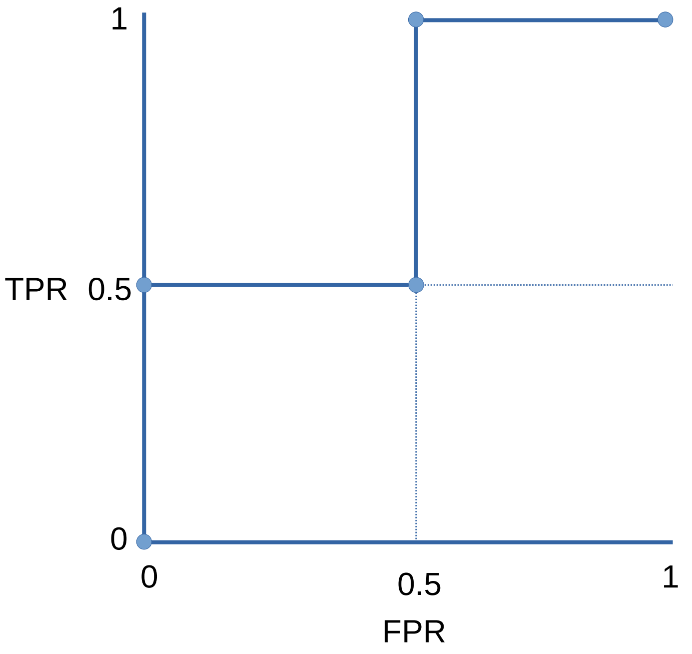
$$\sum_{n=0}^{4} (x_{9-n} - x_n) > 5.$$

Question 2.

	Pred	**True**
P1	0.8	1
P3	0.7	0
P2	0.5	1
P4	0.1	0

ROC-curve parameters, when classification rule is defined as: 1 if pred > γ.

Treshold γ	TPR	FPR
0.0	1	1
0.1	1	1/2
0.5	1/2	1/2
0.7	1/2	0
0.8	0	0

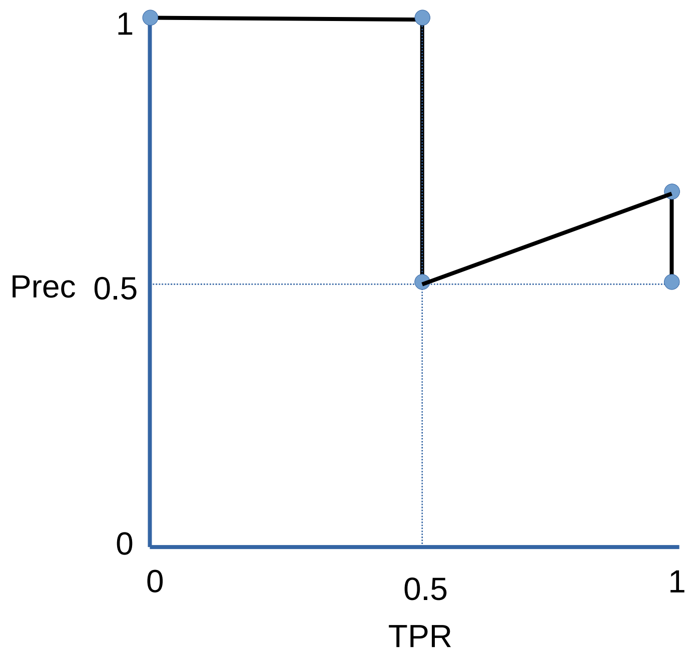


ROC-curve.
AUC-score = 0.75

Question 3.

Precision-recall-curve parameters, when classification rule is defined as: 1 if pred > γ.

Treshold γ	TPR	Prec
0.0	1	1/2
0.1	1	2/3
0.5	1/2	1/2
0.7	1/2	1
0.8	0	1



Precision-recall-curve.
AUPRC-score = 19/24 = 0.792