



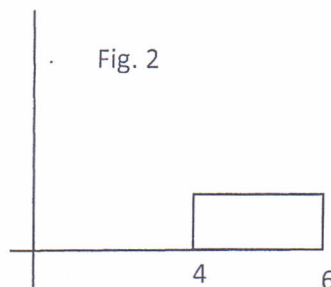
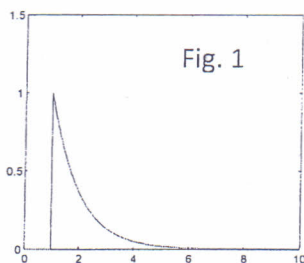
Two pages exam.

Answer the following questions, assume any missing data

**Q1)** Consider a piece of data collected over the course of 14 days where the features are Outlook, Temperature, Humidity, Wind and the outcome variable is whether Golf was played on the day. Now, your job is to build a predictive model which takes in above 4 parameters and predicts whether Golf will be played on the day. We'll build a decision tree to do that. (15 points)

Day	Outlook	Temperature	Humidity	Wind	Play Golf
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

**Q2)** Assume we want to discriminate three types of fish (A, B, C) using the fish weight  $w$ . Fish A has a weight pdf of  $\exp(-(t-1))U(t-1)$  as shown in Fig. 1 where  $U(t)$  is a unit step function, Fish B has a weight pdf of normal distribution of mean 2 and a standard deviation of .5, and fish C has a uniform distribution as shown in Fig. 2. Assume the prior probabilities are  $\frac{1}{4}$ ,  $\frac{1}{2}$ ,  $\frac{1}{4}$  respectively. Find the Bayes classifier for the three types of fish and the threshold between classes. If the caught fish has  $w=2.5$ , what is the classifier output. (15 points)



Q3) Assume we have a training data for two classes.

(20 points)

$w_1 = [(1, 1), (2, 3), (3, 3), (4, 5)]$ ,  $w_2 = [(4, 4), (5, 3), (5, 5), (6, 4)]$

- a) Draw the scattering plot of the training data
- b) Find the Principal components
- c) Find and plot the projection of the training data on the principal components
- d) Find the projection of the training data if only one principal vector is used to represent the data
- e) Find LDA axes
- f) Find and plot the projection of the training data on the LDA

Good luck

*Sherif Kishk*