I. Bayesian Networks

In Figure 3.4, evaluate P(R), P(S) and P(W).

II. Parametric Classification

- Given two normal distributions $p(x|C_1) \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $p(x|C_2) \sim \mathcal{N}(\mu_2, \sigma_2^2)$ and $P(C_1)$ and $P(C_2)$, calculate the Bayes' discriminant points analytically.
- Give an example dataset with 10 data points, where each class has 5 data points. Using the equation derived above, calculate the Bayes' discriminant points.

III. Multivariate Regression

- Let say we have two variables x_1 and x_2 and we want to make a quadratic fit using them, namely

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1 x_2 + w_4 (x_1)^2 + w_5 (x_2)^2 \quad (1)$$

  Given sample data X, how can we find optimum w_i's?
- Give an example dataset with 10 data points. Using the equation derived above, calculate optimum w_i's for this dataset (Do not solve the equation system).

IV. Feature Selection

- Give an example feature selection run on 5 feature dataset, where the best feature subset found by Forward Selection and Backward Elimination is $\{F_2, F_3, F_5\}$. On the other hand, arrange the error rates of the Feature subsets so that the global optimum is not $\{F_2, F_3, F_5\}$.
- Propose a floating search algorithm, and explain it using the above data.

V. Feature Extraction

In Principal Component Analysis, w_k should be the eigenvector of Σ with the k'th largest eigenvalue. Prove this by induction.

VI. Histogram and Naive Estimators

Given the following dataset, plot histogram and naive estimate for bin length h = 1.

Dataset: 0.5, 0.8, 0.9, 1.7, 2.4, 3.6, 5.2, 6.1, 6.3, 7.2