All answers should be justified unless otherwise stated.

You may consult a single, double-sided sheet of paper with notes. Apart from that, you may not look at books, notes, electronic devices etc.

You have 120 minutes. There are 6 questions, of varying credit. The questions are of varying difficulty, so avoid spending too long on any one question.

Good luck!
Problem 1 (5 points)

A continuous random variable $X$ is uniformly distributed in $[-1, 1]$. Compute the density of $Y = X^3$.

Problem 2 (5 points)

Suppose that the rv $X$ is continuous and has a strictly increasing CDF $F_X(x)$. Consider another rv $Y = f(X)$ where the function $f$ is defined as $f(x) = F_X(x)$. Show that $Y$ is uniformly distributed in the interval 0 to 1.

Problem 3 (8 points)

$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is a jointly Gaussian random vector with zero mean and covariance matrix $K_X = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$.

Find a matrix $B \in \mathbb{R}^{2 \times 2}$ such that $Y = BX$ is a jointly Gaussian vector with i.i.d. entries.

Problem 4 (8 points)

Consider a binary detection problem:

$$X = \begin{cases} -1 & \text{w.p. } 1/2 \\ +1 & \text{w.p. } 1/2 \end{cases}$$

and $Y = Xa + Z$ is the observation random vector. Here, $a \in \mathbb{R}^n$ and $Z \sim \mathcal{N}(0, K_Z)$, where $K_Z$ is an $n \times n$ covariance matrix which is singular. Given $K_Z$, choose a vector $a$, subject to the constraint $\|a\| = 1$, to minimize the error probability of the MAP detection rule. (Hint: no computation needed.)

Problem 5 (22 points)

$X, Z_1, Z_2$ are mutually independent $\mathcal{N}(0, 1)$ random variables. Let:

$$Y_1 = X + Z_1$$

$$Y_2 = X + Z_2$$

be two noisy observations about $X$.

(a) (4 points) Compute the mean and covariance matrix of the random vector $[X, Y_1, Y_2]^T$. Is it jointly Gaussian?

(b) (2 points) Are $Y_1$ and $Y_2$ independent?
(c) (2 points) Are $Y_1$ and $Y_2$ independent conditional on $X$?

(d) (5 points) Compute the conditional distribution of $X$ given $Y_1 = y_1$.

(e) (5 points) Compute the conditional distribution of $X$ given $Y_1 = y_1$ and $Y_2 = y_2$.

(f) (2 points) From your answers in parts (d) and (e), what can you say about the value of the additional observation $Y_2 = y_2$ for estimating $X$?

(g) (2 points) Suppose now $Z_2$ is correlated with $Z_1$ (but still $\mathcal{N}(0, 1)$ and still independent of $X$). Compared to your answer to part (f), would you say that the value of the additional observation $Y_2 = y_2$ is now: 1) increased, 2) decreased, or 3) the same? (No justification needed.)

**Problem 6 (32 points)**

Consider the handwritten digit recognition problem. As in the homeworks, each image is a random vector $X \in \mathbb{R}^{784}$. We are trying to recognize whether an image represents a 0 or a 2.

Suppose we model $X \sim \mathcal{N}(a_i, \Sigma_i)$ if digit $i$ is written, $i = 0, 2$. The vector $a_i$ is what the image would have looked if there were no noise, but there is additive noise. We assume as a prior that the image can be equally likely to be a 0 or a 2.

(a) (2 points) Suppose we are given training data for both digits. Explain how you would estimate $a_0$ and $a_2$.

(b) (2 points) Explain how you would estimate $\Sigma_0$ and $\Sigma_2$ based on the data.

(c) (3 points) Now we know $a_0$ and $a_2$ and we also find out the noise is actually white noise, i.e., $\Sigma_0 = \Sigma_2 = \sigma^2 I$. For a test image $X$, find the MAP detection rule for deciding whether it is a 0 and a 2.

(d) (2 points) Sketch the decision regions in the observation space.

(e) (2 points) Find a low-dimensional sufficient statistic for the detection problem.

(f) (2 points) Compute the error probability of the MAP detection rule.

(g) (6 points) To improve the detection performance, the person is asked to write the same digit three times. Find the MAP detection rule based on the three images, and compute the resulting error probability. Compare the error probability with that in part (d). You can assume that the additive noise is independent across the three images.

(h) Suppose the probabilistic model for image for digit 0 remains the same, but there are two ways of writing a 2, represented by vectors $b_2$ and $c_2$ respectively. The resulting image for the digit 2 is modeled as equally likely to be one of the two vectors plus white Gaussian noise with mean 0 and covariance $\sigma^2 I$.

(i) (4 points) Conditional on the digit 2 is written, is the distribution of the image $X$ jointly Gaussian? Explain.

(ii) (4 points) Assume we know $a_0, b_2, c_2$ and $\sigma^2$. Find the MAP detection rule for deciding whether a test image is a 0 or a 2.
(iii) (3 points) Find a 2-dimensional sufficient statistic for this detection problem.

(iv) (4 points) Suppose we don’t know $b_2, c_2$ and $\sigma^2$ but we can estimate both the mean vector and the covariance of the conditional distribution of the image $X$ given that digit 2 is written. Is that enough information to find $b_2, c_2$ and $\sigma^2$?