# STOR 565 Homework

1. Let $X$ have a $\mathcal{N}(\mu, \sigma^2)$ distribution. Show that $\mathbb{E}X = \mu$.

2. Let $Z \sim \mathcal{N}(0, 1)$. Use the CDF method to find the density of $X = aZ + b$.

3. Let $X \in \mathbb{R}^k$ be a random vector and $A \in \mathbb{R}^{r \times k}$. Use the definition of expected value, variance, and linear algebra to establish the following.

   a. $\mathbb{E}(AX) = A\,\mathbb{E}X$

   b. $\text{Var}(X)$ is symmetric and non-negative definite

   c. $\text{Var}(X)_{ij} = \text{Cov}(X_i, X_j)$

   d. $\text{Var}(AX) = A\,\text{Var}(X)A^t$

4. Let $(X, Y) \in \mathbb{R} \times \{0, 1\}$ be a random predictor-response pair. Suppose that $Y$ has prior probabilities $\pi_1 = \mathbb{P}(Y = 1)$ and $\pi_0 = \mathbb{P}(Y = 0)$, and that $X$ is continuous with marginal density $f$ and class conditional densities $f_0$ and $f_1$.

   a. Derive an expression for the Bayes rule $\phi^*(x)$ in terms of the logarithm of the ratio $\pi_1 f_1(x)/\pi_0 f_0(x)$.

Suppose that $f_1$ is $\mathcal{N}(\mu_1, \sigma^2)$ and that $f_0$ is $\mathcal{N}(\mu_0, \sigma^2)$ where $\mu_1 > \mu_0$.

   b. Using the result of part (a), find an expression for the Bayes rule $\phi^*(x)$ in terms of the parameters $\pi_0$, $\pi_1$, $\mu_0$, $\mu_1$, and $\sigma^2$.

   c. What is the form of the rule in part (b) when $\pi_1 = 1/2$? Explain why this makes intuitive sense.

   d. Suppose for simplicity that $\mu_1 = u$ and $\mu_0 = -u$ for some $u > 0$. What form does the Bayes rule take when $u$ increases (tends to infinity), and in particular, how does the rule depend on $\pi_1$ versus $\pi_0$? A informal but clear answer is fine.

5. Let $(X_1, Y_1), \ldots, (X_n, Y_n) \in \mathbb{R} \times \{0, 1\}$ be a labeled set of real observations.

   a. Give an estimate of the probability that $Y = 0$. What does the law of large numbers say about the limiting behavior of your estimate as $n$ gets very large?

b. Write the sample mean $\hat{\mu}_0$ of the zero-labeled observations using indicator functions.

c. Write the sample variance $\hat{\sigma}_0^2$ of the zero-labeled observations using $\hat{\mu}_0$ and indicator functions.

6. Consider the setting of linear discriminant analysis in which the class-conditional densities $f_0$ and $f_1$ have the multivariate normal form $f_k = \mathcal{N}(\mu_k, \Sigma_k)$.

a. Using the expression for the multivariate normal density, show that the discriminant functions $\delta_k(x) = \log(\pi_k \, f_k(x))$ have the form

$$\delta_k(x) = -\frac{1}{2} x^t \Sigma_k^{-1} x + \langle x, \Sigma_k^{-1} \mu_k \rangle - \frac{1}{2} \left[ \log(2\pi)^d \pi_k^{-2} \det(\Sigma_k) + \mu_k^t \Sigma_k^{-1} \mu_k \right]$$

b. Show that when $\Sigma_0 = \Sigma_1 = \Sigma$ the decision boundary $B = \{x : \delta_1(x) = \delta_0(x)\}$ has the form

$$B \;=\; \{x : x^t \, \Sigma^{-1}(\mu_1 - \mu_0) + (c_0 - c_1) = 0\}$$

where $c_0, c_1$ are real valued constants, and argue that this set is a hyperplane.

7. In a previous homework you showed that any symmetric matrix $A \in \mathbb{R}^{d \times d}$ can be written in the form $A = \Gamma D \Gamma^t$, where $\Gamma \in \mathbb{R}^{d \times d}$ is an orthogonal matrix and $D = \text{diag}(\lambda_1, \ldots, \lambda_d)$ is a diagonal matrix with $D_{ii}$ equal to the $i$th eigenvalue of $A$ and all other entries equal to zero. Suppose that $A$ is non-negative definite, so that each $\lambda_i \geq 0$. Define $A^{1/2} = \Gamma D^{1/2} \Gamma^t$ where $D^{1/2} = \text{diag}(\lambda_1^{1/2}, \ldots, \lambda_d^{1/2})$. Show that $A^{1/2}$ is symmetric and satisfies $A^{1/2} A^{1/2} = A$.

8. Let $X \sim \mathcal{N}_d(\mu, \Sigma)$ and let $Y = \Sigma^{1/2} Z + \mu$ where $Z \sim \mathcal{N}_d(0, I)$.

(a) Show that $\mathbb{E}Y = \mathbb{E}X$ and that $\text{Var}(Y) = \text{Var}(X)$. Hint: Use the definition of $\mathcal{N}_d(\mu, \Sigma)$ and basic properties of multivariate means and variances.

(b) Fix $v \in \mathbb{R}^d$. Find the distributions of the random variable $v^t X$.