Problem 1

Lemma: Let $P_1, P_2 \in \mathcal{P}$ be two PMFs such that $\rho(P_1, P_2) = 2\delta$

Then

$$M_n(P_1, P_2) \geq \Phi(\delta) \sum_{x \in \mathcal{X}^n} \min \left( P_1(x), P_2(x) \right)$$

$$\geq \frac{\Phi(\delta)}{2} \exp \left( -n \, D(P_1 \parallel P_2) \right)$$

Proof: As shown in class (also Duchi, Prop. 13.3)

$$M_n \geq \Phi(\delta) \inf_{P} P(\psi(X_i^n) \neq V), \text{ where } V \in \{1, 2\} \text{ is a random index.}$$

Writing out the probability on the right-hand side (assuming $P_V(1) = P_V(2) = \frac{1}{2}$)

we have

$$M_n \geq \Phi(\delta) \left( \frac{1}{2} (P_1(\psi(X_i^n) = 1) + P_2(\psi(X_i^n) = 2)) \right)$$

Part(a) of the problem: We assume that $\psi$ is the Neyman-Pearson test, as stated in the question. Let $A \subseteq \mathcal{X}^n$ be such that $P_i(A) = P_2(A), \quad x \in A$.

$$P_1(\psi(X_i^n) = 1) + P_2(\psi(X_i^n) = 2) = \sum_{x \in \mathcal{X}^n} P_1(x) + \sum_{x \in \mathcal{X}^n} P_2(x)$$

$$= \sum_{x \in \mathcal{X}^c} \min(P_1, P_2) + \sum_{x \in A} \min(P_1, P_2)$$

$$= \sum_{x \in \mathcal{X}} \min(P_1, P_2)$$

This implies inequality (1) in the statement of Lemma.
Part (b) of the problem

We have \[ \sum_{x \in \mathcal{X}} (\max(p_1, p_2) + \min(p_1, p_2)) = \sum_{x \in \mathcal{X}} (p_1 + p_2) = 2 \]
and thus \[ \sum_{x \in \mathcal{X}} \max(p_1, p_2) = 2 - \sum_{x \in \mathcal{X}} \min(p_1, p_2) \]

Next

\[ \sum_{x} \min(p_1, p_2) \geq \left( 2 - \sum_{x} \min(p_1, p_2) \right) \sum_{x} \min(p_1, p_2) \]

\[ = \sum_{x} \max(p_1, p_2) \sum_{x} \min(p_1, p_2) \]

Cauchy-Schwarz\textsuperscript{2}

\[ \left( \sum_{x} \sqrt{\min(p_1, p_2) \max(p_1, p_2)} \right)^2 = \left( \sum_{x} \sqrt{p_1(x) p_2(x)} \right)^2 \]

\[ = \exp \left( 2 \log \sum_{x} \sqrt{p_1(x) p_2(x)} \right) \]

\[ = \exp \left( 2 \log \sum_{x} p_1(x) \sqrt{\frac{p_2(x)}{p_1(x)}} \right) \overset{\text{Jensen}}{\geq} \exp \left( \sum_{x} p_1(x) \log \frac{p_2(x)}{p_1(x)} \right) = \exp(-D(p_1 \parallel p_2)) \]

This proves inequality (2) and concludes the proof of the lemma.

Part (c) Both parts are actually proved above. Let us show the first equality: We use the set $A$ as defined in Part (a).

\[ \|P - Q\|_{TV} = \frac{1}{2} \sum_{x \in \mathcal{X}} |P(x) - Q(x)| \]

\[ \sum_{x \in \mathcal{X}} \min(P, Q) + \frac{1}{2} \sum_{x \in \mathcal{X}} |P - Q| = \sum_{x \in A} Q + \frac{1}{2} P - \frac{1}{2} Q + \sum_{x \in A^c} P + \frac{1}{2} Q - \frac{1}{2} P \]

\[ = \frac{1}{2} \sum_{x \in \mathcal{X}} (P + Q) = 1. \]

The second inequality is shown in (3) above on this page; the relation with the Hellinger distance follows by definition.
Problem 2

(a) \( \Delta = |\mu_1(0) - \mu_2(0)| = C_5 \)

(b) \( f_1(x, y) = f(x) f_1(y|x) = \varphi(y) \), where \( \varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \), \(-\infty < t < \infty\)

\[ \int f_2(x, y) = f_2(y|x) = \varphi(y - \mu_2(x)) \]

(c) This is a straightforward calculation:

\[
D(Z_1 \parallel Z_2) = \int \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-m_1)^2}{2}} \ln \left( e^{-\frac{(x-m_1)^2}{2} + \frac{(x-m_2)^2}{2}} \right) \right) dx
\]

\[ = \int \frac{1}{\sqrt{2\pi}} (-(m_1^2 - x^2 + 2xm_1 - m_1^2 + x^2 - 2xm_2 + m_2^2)) dx \]

\[ = \frac{1}{2} (-m_1^2 + 2m_1^2 - 2m_1m_2 + m_2^2) = \frac{1}{2} (m_1 - m_2)^2. \]

(d) Now compute \( D(P_1 \parallel P_2) \), where \( P_1 \sim N(0, 1) \), \( P_2 \sim N(\mu_2, 1) \):

using (a) and (c), we obtain

\[
D(P_1 \parallel P_2) = \int \left( \frac{1}{2} \left[ C(\delta - t)^2 - 0 \right] \right) dt = \frac{C^2 \delta^3}{6}
\]

(e) Take \( \delta = (6 \log 2 / C^3 n)^\frac{1}{3} \), then

\[
\mu_n \geq \Delta \exp(-n D(P_1 \parallel P_2)) = C \left( \frac{6 \log 2}{C^3 n} \right)^\frac{1}{3} \exp(-n \frac{C^2 \delta^3}{6} \frac{6 \log 2}{C^3 n})
\]

\[ = \left( \frac{C}{n} \right)^\frac{1}{3} \left( \frac{6 \log 2}{2} \right)^\frac{1}{3} > \left( \frac{C}{n} \right)^\frac{1}{3} \]

(f) A full solution will not be provided (it would be too much of a detour to write it); a scheme that achieves the same order of minimax risk is obtained by linear regression.