1. **Trace distance of probability distributions:** Let \( p, q \in \mathcal{P}(\Sigma) \) be probability distributions. In today’s lecture, we defined \( \|p - q\|_1 := \sum_{x \in \Sigma} |p(x) - q(x)| \).

(a) Show that
\[
\frac{1}{2} \|p - q\|_1 = \max_{A \subseteq \Sigma} \left( \sum_{x \in A} p(x) - \sum_{x \in A} q(x) \right).
\]

Do you recognize this as the probability theory analog of a formula that we proved for quantum states?

(b) Let \( p \in \mathcal{P}(\Sigma \times \Sigma) \) be a joint probability distribution of two random variables \( X \) and \( Y \), with marginal distributions \( p_X \) and \( p_Y \). Prove that:
\[
\frac{1}{2} \|p_X - p_Y\|_1 \leq \Pr(X \neq Y)
\]

2. **Fidelity between two classical-quantum states:** Let \( p \in \mathcal{P}(\Sigma) \). Show that the fidelity between two states of the form \( \rho = \sum_{x \in \Sigma} p(x) |x\rangle \langle x| \otimes \rho_x \) and \( \sigma = \sum_{x \in \Sigma} p(x) |x\rangle \langle x| \otimes \sigma_x \) is given by the average of the fidelities \( F(\rho_x, \sigma_x) \), namely
\[
F(\rho, \sigma) = \sum_{x \in \Sigma} p(x) F(\rho_x, \sigma_x).
\]

We used this identity in the lecture.

3. **On the definition of quantum codes:** The definition of an \((n, R, \delta)\)-quantum code in the lecture was perhaps surprising. Why did we not simply demand that \( F(D[\mathcal{E}[\rho^{\otimes n}]], \rho^{\otimes n}) \geq 1 - \delta \)? Argue that this does not correspond to a reliable compression protocol.

*Hint: In last week’s coin flip scenario, what would be the (classical) analog of the condition that \( D[\mathcal{E}[\rho^{\otimes n}]] \approx \rho^{\otimes n} \)?*

4. **Converse of Schumacher’s theorem:** In this problem you can prove part B of Schumacher’s theorem in case we only gave a sketch in class. Let \( \rho \in \mathcal{D}(\mathcal{X}) \), \( \delta \in (0, 1) \), and \( R < H(\rho) \).

(a) Show that there exists a function \( f(n) \) such that \( f(n) \to 0 \) and \( \text{Tr}[P \rho^{\otimes n}] \leq f(n) \) for every orthogonal projection \( P \in \mathcal{L}(\mathcal{X}^{\otimes n}) \) of rank \( \leq 2^n R \).

(b) Show that there exist \((n, R, \delta)\)-quantum codes for \( \rho \) for at most finitely many \( n \).

*Hint: Each Kraus operator of \( \mathcal{D} \mathcal{E} \) has rank \( \leq 2^n R \). Use the formula for the channel fidelity from class and estimate each term using the Cauchy-Schwarz inequality \( |\text{Tr}[AB]| \leq ||A||_2 ||B||_2 \) after having inserted a suitable projection.*