1. **Gaussian Estimation**

Let $Y = X + Z$ and $U = X - Z$, where $X$ and $Z$ are i.i.d. $\mathcal{N}(0, 1)$.

(a) Find the joint distribution of $U$ and $Y$.

(b) Find the MMSE of $X$ given the observation $Y$, call this $\hat{X}(Y)$.

(c) Let the estimation error $E = X - \hat{X}(Y)$. Find the conditional distribution of $E$ given $Y$. 
2. Overlapping Normals

As you will see in the lab, a big part of the Kalman filter is “overlapping” two normal distributions. In particular, suppose at the current time step, you have a state $X \sim \mathcal{N}(\mu, \sigma_1^2)$ and an observation $Y \sim X + \mathcal{N}(0, \sigma_2^2)$. The two noises are independent.

(a) Not knowing $Y$, what is your best guess of $X$?

(b) $X$ and $Y$ are jointly Gaussian. Write the vector $[X \ Y]^T$ as an affine transformation of independent unit Gaussians. i.e find $A \in \mathbb{R}^{2 \times 2}$ and $\mu \in \mathbb{R}^2$ such that $[X \ Y]^T = AZ + \mu$ where $Z = [Z_1 \ Z_2]^T$ and $Z_1, Z_2 \text{i.i.d.} \sim \mathcal{N}(0, 1)$.

(c) What is $\mathbb{E}[X \mid Y]$?

(d) Given $Y = y$, the conditional distribution $X \mid Y = y$ turns out to be normal, which is the reason why we can continue using the same process for future time steps. In fact, it turns out $\text{var}(X \mid Y)$ is constant, so it's always equal to its expectation $\mathbb{E}[\text{var}(X \mid Y)] = \mathbb{E}[(X - \mathbb{E}[X \mid Y])^2]$. WLOG assume $\mu = 0$ and show that this is equal to $\frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$. 