1. (a) Suppose $X,Y,Z$ are binary random variables that take the values 0 or 1 with probability 0.5. Construct an example where $X,Y,Z$ are pairwise independent ( $P_{XY}(x,y) = P_X(x)P_Y(y)$, $P_{YZ}(y,z) = P_Y(y)P_Z(z)$, $P_{XZ}(x,z) = P_X(x)P_Z(z)$ for all $x$, $y$, $z$), but are not mutually independent ( $P_{XYZ}(x,y,z) \neq P_X(x)P_Y(y)P_Z(z)$ for some $x$, $y$, $z$).

(b) Does pairwise independence ensure that


2. (a) The conditional variance of $X$ given $Y$ is defined by

$$Var(X|Y) = E[(X - E(X|Y))^2|Y].$$

Prove that

$$Var(X) = E[Var(X|Y)] + Var(E[X|Y]).$$

(b) A mouse is trapped in the center of a maze. There are three paths that begin at the center. The first path takes the mouse to safety after 2 hours of travel. The second and third paths return the mouse to the center of the maze after 3 and 5 hours of travel respectively. Assume the mouse is equally likely to choose any of the three paths at any time. Find the variance of $X$, the total time the mouse spends in the maze.

3. Suppose that we are designing a digital receiver. The sender will transmit a single bit $H$ that takes the value 0 or 1 with probabilities $P_0$ and $P_1$ respectively. The receiver does not get to see $H$ directly, but instead observes a continuously distributed random vector $\hat{Y}$. $\hat{Y}$ depends on both $H$ as well as random noise that is introduced in the channel that connects the transmitter to the receiver. Let $\hat{H}(\hat{y})$ be the receiver’s guess of what the original bit was given that the receiver observes $\hat{Y} = \hat{y}$. We would like to maximize the probability of the receiver making the correct guess. Therefore, we use the rule

$$P_{\hat{H}|\hat{Y}}(0|\hat{y}) \overset{\hat{H}=0}{\geq} P_{\hat{H}|\hat{Y}}(1|\hat{y}) \overset{\hat{H}=1}{<}$$

where the notation $P_{\hat{H}|\hat{Y}}(0|\hat{y})$ means $P(H = 0|\hat{Y} = \hat{y})$ and the notation means that the receiver picks $\hat{H} = 1$ if the left side is larger, and picks $\hat{H} = 0$ otherwise. Rule (1) is known as a maximum a posterior probability (MAP) rule. It turns out we can simplify (1) in the following way. First, we make use of Bayes rule to write

$$\lim_{\epsilon \to 0} \frac{P(H = 0, \hat{Y} \in B(\hat{y}, \epsilon))}{P(\hat{Y} = B(\hat{y}, \epsilon))} \overset{\hat{H}=0}{\geq} \lim_{\epsilon \to 0} \frac{P(H = 1, \hat{Y} \in B(\hat{y}, \epsilon))}{P(\hat{Y} = B(\hat{y}, \epsilon))}.$$ 

where $B(\hat{y}, \epsilon)$ is a ball centered at $\hat{y}$ with radius $\epsilon$. We apply Bayes rule again to write

$$\lim_{\epsilon \to 0} \frac{P_0P(\hat{Y} \in B(\hat{y}, \epsilon)|H = 0)}{P(\hat{Y} = B(\hat{y}, \epsilon))} \overset{\hat{H}=0}{\geq} \lim_{\epsilon \to 0} \frac{P_1P(\hat{Y} \in B(\hat{y}, \epsilon)|H = 1)}{P(\hat{Y} = B(\hat{y}, \epsilon))}$$

where $P_0, P_1$ are the values of the priors $P_0$ and $P_1$ in rule (2), then it is MAP rule.

The left hand side of (2) is known as the likelihood ratio. Oftentimes, we do not know the “prior” probabilities, that is the probabilities that the sender sends a 1 or 0. In that case, we can assume that $P_0 = P_1$. When we assume $P_0 = P_1$, rule (2) is called a maximum likelihood (ML) test. When we use the values of the priors $P_0$ and $P_1$ in rule (2), then it is MAP rule.
(a) Suppose that the sender sends a signal of size $X = a$ volts to represent the bit $H = 0$ and a signal of $X = b$ volts to represent the bit $H = 1$. Suppose that the receiver receives the signal

$$Y = X + Z$$

where $Z \sim N(0, \sigma^2)$. What is the MAP rule for determining $\hat{H}$? Hint: Start with (2). To simplify your final expression as much as possible, take the log of both sides to eliminate exponential functions. Also put all terms dependent on $y$ on the left side of your final expression, and all terms not dependent on $y$ on the right side of your expression.

(b) In part (a), suppose $\sigma = 2, a = 0, b = 5$, and $P_0 = P_1$. What is the probability of your MAP detection rule making an error?

(c) Suppose the sender sends a vector valued signal $\vec{X} = \vec{a}$ to represent a 0 and $\vec{X} = \vec{b}$ to represent a 1. Suppose that the receiver receives the signal

$$\vec{Y} = \vec{X} + \vec{Z}$$

where $Z \sim N(0, \sigma^2 I)$. What is the MAP rule for determining $\hat{H}$? Simplify your answer as much as possible.

(d) Repeat (c) but now suppose that $Z \sim N(0, K_Z)$, where $K_Z$ is non-singular. Hint: There exists a nonsingular matrix $A$ such that $K_Z = AA^T$. The signal $\vec{W} := A^{-1}\vec{Y}$ contains all of the information that $\vec{Y}$ does because we can recover $\vec{Y}$ from $\vec{W}$. Thus you can construct your detector using $\vec{W}$.

(e) Now suppose that $K_z$ is singular, and in particular $\vec{v}^T K_z \vec{v} = 0$. Also suppose that $\vec{v}^T (\vec{a} - \vec{b}) \neq 0$. What is the MAP rule? Hint: consider the signal $\vec{v}^T \vec{Y}$. 