1. Suppose $X, Y, Z$ are pairwise Jointly Gaussian (JG). In other words the pairs $\{X, Y\}$, $\{Y, Z\}$, and $\{X, Z\}$ are JG. Must the collection $\{X, Y, Z\}$ be JG? If so, prove it. If not, provide a counterexample.

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 sender is trying to send a packet to a receiver in a wireless network. The sender starts in the "ready" state, and whenever the sender is in the "ready" state, three events are equally likely to follow. In the first, the sender pauses for 1 ms and then returns to the "ready" state. In the second, the sender tries to send but is unsuccessful because of a collision. The collision takes 2 ms to take place, after which the sender returns to the "ready" state. In the third event the transmission is successful and takes 3 ms to complete. Find the variance of $X$, the total time it takes to complete the sending of the packet. Note that what happens in each visit to the "ready" state is independent of what happened in previous visits. HINT: Let $Y$ be the time taken for the event that takes places after the initial start at the "ready" state. Convince yourself that the expected amount of time remaining until the packet is finished is always the same from the "ready" state, regardless of how many times we’ve already come back to it.

(c) A combined search engine and ad-network company called Doogle just received an order to display an ad whenever someone makes a specific, very unusual query. Suppose that Doogle receives a query every microsecond, and each query has a $10^{-6}$ probability of being suitable for displaying the ad. What is the variance of $X$, the total time that elapses until Doogle displays the ad.

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 continuous valued random vector $\vec{Y}$. $\vec{Y}$ depends on both $H$ as well as random noise that is introduced in the channel that connects the transmitter to the receiver. Let $\hat{H}(\vec{y})$ be the receiver’s guess of what the original bit was given that the receiver observes $\vec{Y} = \vec{y}$. We would like to maximize the probability of the receiver making the correct guess. Therefore, we use the rule

$$P_{H|\vec{Y}}(0|\vec{y}) = \begin{cases} P_{H|\vec{Y}}(1|\vec{y}) & \hat{H} = 0 \\ P_{H|\vec{Y}}(1|\vec{y}) & \hat{H} = 1 \end{cases}$$

(1)

where the notation $P_{H|\vec{Y}}(0|\vec{y})$ means $P(H = 0|\vec{Y} = \vec{y})$ 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

$$\lim_{\epsilon \to 0} \frac{P(H = 0, \vec{Y} = B(\vec{y}, \epsilon))}{P(\vec{Y} = B(\vec{y}, \epsilon))} = \frac{P_0}{P_1} \lim_{\epsilon \to 0} \frac{P(H = 0, \vec{Y} = B(\vec{y}, \epsilon))}{P(\vec{Y} = B(\vec{y}, \epsilon))}$$

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

$$\lim_{\epsilon \to 0} \frac{P_0 f_{\vec{Y}|H}(\vec{y}|0)}{f_{\vec{Y}}(\vec{y})} = \frac{P_0}{P_1} \lim_{\epsilon \to 0} \frac{P_1 f_{\vec{Y}|H}(\vec{y}|1)}{f_{\vec{Y}}(\vec{y})}$$

$$f_{\vec{Y}|H}(\vec{y}|1) = \frac{P_0}{P_1} f_{\vec{Y}|H}(\vec{y}|0)$$

(2)
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 \( X = a \) to represent the bit \( H = 0 \) and a signal of \( X = b \) 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 \( \sigma = 1, a = 0, b = 4, \) and \( P_0 = 0.6 \). 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 \( \vec{Z} \sim N(0, \sigma^2 I) \). What is the MAP rule for determining \( \hat{H} \)? Simplify your answer as much as possible.

(d) Repeat (c) but now suppose that \( \vec{Z} \sim N(0, \vec{K}_Z) \), where \( \vec{K}_Z \) is non-singular. Hint: There exists a nonsingular matrix \( \vec{A} \) such that \( \vec{K}_Z = \vec{A}\vec{A}^T \). The signal \( \vec{W} := \vec{A}^{-1}\vec{Y} \) 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 \( \vec{K}_Z \) is singular, and in particular \( \vec{v}^T \vec{K}_Z \vec{v} = 0 \). Also suppose that \( \vec{v}^T (\vec{a} - \vec{b}) \neq 0 \). What is the MAP rule? Hint: consider the signal \( \vec{v}^T \vec{Y} \).