Problem #1: Problem 4.1 on Page 4-18 of the *Optimization Based Control* notes.

In this problem you were asked to consider a random variable, Z, that is the sum of two independent Gaussian distributed variables, x_1 and x_2:

$$x_1 \sim \mathcal{N}(m_1, \sigma_1^2) \qquad x_2 \sim \mathcal{N}(m_2, \sigma_2^2) \ .$$

You were then asked to show that X is a Gaussian random variable with mean $\bar{z} = m_1 + m_2$ and variance $\sigma_z^2 = \sigma_1^2 + \sigma_2^2$. Note that:

$$\begin{aligned} \bar{z} &= E[z] = E[x_1 + x_2] = E[x_1] + E[x_2] = m_1 + m_2 \\ \sigma_z^2 &= E[(z - \bar{z})(z - \bar{z})^T] = E[(x_1 + x_2 - m_1 - m_2)(x_1 + x_2 - m_1 - m_2)^T] \\ &= E[(x_1 - m_1)(x_1 - m_1)^T + (x_2 - m_2)(x_2 - m_2)^T + (x_1 - m_1)(x_2 - m_2)^T \\ &\quad +(x_2 - m_2)(x_1 - m_1)^T] \\ &= \sigma_1^2 + \sigma_2^2 \end{aligned}$$

Hence,

$$p(z) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{1}{2}\frac{[z-(m_1+m_2)]^2}{\sigma_1^2+\sigma_2^2}} \tag{1}$$

You were also asked to show that, equivalently, $p(z)$ could be represented by the formula:

$$p(z) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left\{ -\frac{(z - x - m_1)^2}{2\sigma_1^2} - \frac{(x - m_2)^2}{2\sigma_2^2} \right\} dx \ . \tag{2}$$

Thus, you need to show that Equations (1) and (2) are equivalent. There are several different procedures to show the equivalence between the two. Here we will use an approach that follows very closely from pages 4-4 and 4-5 of the *Optimization Based Control* class notes.

As before, let $Z = X_1 + X_2$, with X_1 and X_2 being independent Gaussian random variables. Let's compute the joint probability of the events A and B, where:

$$A = \{x_{1l} \le x_1 \le x_{1u}\}, \qquad B = \{z_l \le z \le z_u\} \ .$$

The joint probability of both events A and B occurring is:

$$\begin{aligned} P(A \cap B) &= P(x_{1l} \le x_1 \le x_{1u}, z_l \le x_1 + x_2 \le z_u)\} \\ &= P(x_{1l} \le x_1 \le x_{1u}, (z_l - x_1) \le x_2 \le (z_u - x_1))\} \end{aligned}$$

Now use the Gaussian probability density functions for X_1 and X_2 to derive an explicit expression for this probability.

$$\begin{aligned} P(A \cap B) &= \int_{x_{1l}}^{x_{1u}} \left(\int_{z_l - x_1}^{z_u - x_1} p_{X_2}(x_2)dx_2 \right) p_{X_1}(x_1)dx_1 \\ &= \int_{x_{1l}}^{x_{1u}} \left(\int_{z_l}^{z_u} p_{X_2}(z - x_1)dz \right) p_{X_1}(x_1)dx_1 \\ &\triangleq \int_{z_l}^{z_u} \int_{x_{1l}}^{x_{1u}} p_{Z,X_1}(z, x_1)dx_1 dz \end{aligned}$$

where we have used a change of variables in the second line of the derivation, and we have used the fact that the joint probability of events A and B is equivalent to the double integral over the joint probability density function of the random variables Z and X_1. Hence, we have shown that $p_{Z,X_1} = p_{X_2}(z - x_1)p_{X_1}(x_1)$. Recalling that:

$$p_Z(z) = \int p_{Z,X_1}(z, x_1)dx_1$$

we can see that

$$p_Z(z) = \int_{-\infty}^{\infty} p_{Z,X_1}dx_1 = \int_{\infty}^{\infty} p_{X_2}(z - x_1)P_{X_1}(x_1)dx_1.$$

Now subsitute in the explicit Gaussian formulae for the probability densities into the integral expression:

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(z-x_1-m_2)^2}{2\sigma_2^2}} \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x_1-m_1)^2}{2\sigma_1^2}} dx_1 \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left\{ - \frac{(z - x_1 - m_2)^2}{2\sigma_2^2} - \frac{(x_1 - m_1)^2}{2\sigma_1^2} \right\} dx_1 \end{aligned}$$

Note that this result is slightly different than then one in Equation (2). If we had taken a different ordering of the integrating variables, we would arrive at that exact expression.

Problem 2 Statement: Consider the motion of a particle that is undergoing a random walk n one dimension. We model the position of the particle as $x[k + 1] = x[k] + u[k]$, where x is the position of the particle and u is a white noise process with $\mathbb{E}\{u[i]\} = 0$ and $\mathbb{E}\{u[i]u[j]\} = R_u\delta(i - j)$. We assume that we can measure x subject to additive, zero-mean, Gaussian white noise with covariance 1. Show that the expected value of the particle as a function of k is given by:

$$\mathbb{E}\{x[k]\} = \mathbb{E}\{x[0]\} + \sum_{i=0}^{k-1} \mathbb{E}\{u[i]\} = \mathbb{E}\{x[0]\} \triangleq \mu_x \tag{3}$$

Also show that the covariance $\mathbb{E}\{(x[k] - \mu_x)^2\} = \sum_{i=0}^{k-1} \mathbb{E}\{u^2[i]\} = kR_u$

Solution: By iterating the recursive dynamical relationship it is easy to show that $x[k]$ can be expressed as:

$$x[k] = x[0] + \sum_{i=0}^{k-1} u[i] \tag{4}$$

As expectation is a linear operator, $\mathbb{E}\{x[k]\} = \mathbb{E}\{x[0]\} + \sum_{i=0}^{k-1} \mathbb{E}\{u[i]\}$. However, we are given that $\mathbb{E}\{u[i]\} = 0 \forall i \Rightarrow \boxed{\mathbb{E}\{x[k]\} = \mathbb{E}\{x[0]\} \triangleq \mu_x}$. Given the recursive dynamical equation, we may also directly evaluate the covariance:

$$\mathbb{E}\{(x[k] - \mu_x)^2\} = \mathbb{E}\{(\sum_{i=0}^{k-1} u[i])^2\} = \mathbb{E}\{\sum_{i=0}^{k-1} u[i]^2 + \sum_{i=0}^{k-1} \sum_{j \neq i}^{k-1} u[i]u[j]\} \tag{5}$$

Again using the fact that $\mathbb{E}\{\cdot\}$ is a linear operator, we find that:

$$\mathbb{E}\{(x[k] - \mu_x)^2\} = \sum_{i=0}^{k-1} \mathbb{E}\{u[i]^2\} + \sum_{i=0}^{k-1} \sum_{j \neq i}^{k-1} \mathbb{E}\{u[i]u[j]\} \tag{6}$$

In the above, we are asumming that $x[0] = \mathbb{E}\{x[0]\}$, or equivalently that the covariance on the initial state estimate is zero. However, since the perturbation $u[i]$ is a white-noise process, its' value at two successive times are uncorrelated, thus $\mathbb{E}\{u[i]u[j]\} = 0 \forall i \neq j$, and the double sum evaluates to zero. Then, given $\mathbb{E}\{u[i]u[j]\} = \delta(i-j)R_u$ we find that:

$$\boxed{\mathbb{E}\{(x[k]-\mu_x)^2\} = kR_u} \tag{7}$$

Problem 3 Statement: Let X and Y be two jointly distributed random variable with X scalar and let Y take the value y. Let $\hat{x}$ be an estimate chosen so that:

$$\mathbb{E}(|X-\hat{x}| : Y = y) \leq \mathbb{E}(|X-z| : Y = y) \tag{8}$$

In other words, $\hat{x}$ is chosen to minimize the average value of the *absolute error* between $\hat{x}$ and the actual value taken by X. Show that $\hat{x}$ is the median of the conditional density $p_{x|y}(x|y)$. Note: The median of a continuous density $p_A(a)$ is that value of a, call it α, for which $P(A \leq \alpha) = P(A \geq \alpha)$.

Solution: Via the condition $\mathbb{E}\{|X-\hat{x}| : Y = y) \leq \mathbb{E}\{|X-z| : Y = y\} \forall z$, it is clear that we are searching for a *function* that, given some value of $Y = y$ will return an estimate that minimizes the expected value of the absolute error. This can be written in another way as:

$$\hat{x}(y) = \arg\min_q \mathbb{E}\{|X-q| : Y = y\} \tag{9}$$

We can write the expected value explicitly as:

$$\hat{x}(y) = \arg\min_q \Big(\int_{-\infty}^{q} (q-x)p(x|y)\ dx + \int_{q}^{\infty} (x-q)p(x|y)\ dx \Big) \tag{10}$$

Two integrals are required above to handle the two cases that $q > X = x$ and $q < X = x$, respectively. The advantage of framing the problem in this way is that now we have a standard minimization problem on the variable q. We can take the derivative of this function with respect to the variable q. To do this, we need to know how to take derivatives with respect to an integral when the limits of integration depend on the variable we are varying. *Leibniz' Rule* tells us that:

$$\frac{d}{dx}\Big(\int_{\alpha(x)}^{\beta(x)} f(x,t)\ dt \Big) = \int_{\alpha(x)}^{\beta(x)} \frac{df(x,t)}{dx}\ dt + \frac{d\beta(x)}{dx} f(x,\beta(x)) - \frac{d\alpha(x)}{dx} f(x,\alpha(x)) \tag{11}$$

Applying Leibniz' Rule to our problem (and noting that the integrand of both integrals evaluated at q is equal to zero):

$$\begin{aligned} \frac{d}{dq}\Big(\int_{-\infty}^{q} (q-x)p(x|y)\ dx + \int_{q}^{\infty} (x-q)p(x|y)\ dx \Big) &= 0 \\ \Rightarrow \int_{-\infty}^{q} p(x|y)\ dx - \int_{q}^{\infty} p(x|y)\ dx &= 0 \\ \boxed{\Rightarrow \int_{-\infty}^{q} p(x|y)\ dx = \int_{q}^{\infty} p(x|y)\ dx = 0} \end{aligned} \tag{12}$$

This is telling us that our solution, the minimizer q^* must satisfy: $\int_{-\infty}^{q^*} p(x|y)\ dx = \int_{q^*}^{\infty} p(x|y)\ dx$. Note that this is claiming that the estimate q^* is such that the cumulative distribution of the conditional probability up to the estimate is equal to that after the estimate, i.e. the estimate q^* equally splits the distribution, which defines q^* as the median.