Exercise 5.

Group name: TAU03E

Task 1

Random variables are defined as thigns that get different values each time observed.

a) No. Equation is not random variable since equation has certain value for x where it is true. There is no randomness in solution

b) Yes. Security camera picture is probably different each time it is taken. Thus, it can be a random variable

c) Yes. Daily stock price varies on multiple factors and variables.

d) No. All the variables are known for the mass index BMI. Thus, it cannot be a random variable.

e) Yes. One variable now is unknown for the mass index BMI, which leads to the fact it can be modeled with probability distribution and result is dependent on the unknown variable.

f) No. Distribution of students is not a random variable.

e) No. Physics theorem validity is not a random variable since it is objective.

g) No. Mathematical hypothesis are also objective matters. However they can be studied and tested using random variables but the hypothesis validity is not a random variable.

Task 2

a)

Let's show that $\hat{\mu} = \mu$ as $E\left(\hat{\mu}\right) = E\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) \Rightarrow \ \hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} E\left(x_i\right)$. Now as $E\left(x_i\right) = \mu$ therefore $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} \mu = \frac{N\mu}{N} = \mu$.

b)

(We assume that that the equation is missing a minus sign by accident in the assignment).

Let's show that $\hat{\sigma^2} \neq \sigma^2$ as
$\hat{\sigma^2} = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \hat{\mu}\right)^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i^2 - 2X_i\hat{\mu} + \hat{\mu^2}\right) = \frac{1}{N}\left(\sum_{i=1}^{N} X_i^2 - 2\hat{\mu}\sum_{i=1}^{N} X\right.$
.

Now it's known that $\sum_{i=1}^{N} X_i = n \cdot \frac{1}{n} \sum_{i=1}^{N} X_i = n\hat{\mu}$. Thus
$\hat{\sigma^2} = \frac{1}{N}\left(\sum_{i=1}^{N} X_i^2 - 2n\hat{\mu^2} + n\hat{\mu^2}\right) = \frac{1}{N}\left(\sum_{i=1}^{N} X_i^2 - n\hat{\mu^2}\right)$.

Let's take the expectation of the estimator to show that it's biased:

$$E\left(\hat{\sigma^2}\right) = E\left(\frac{1}{N}\left(\sum_{i=1}^{N} X_i^2 - n\hat{\mu^2}\right)\right) = \frac{1}{N}\left(\sum_{i=1}^{N} E\left(X_i^2\right) - nE\left(\hat{\mu^2}\right)\right) = E\left(X_i^2\right) -$$

.

Now as $Var\left(X\right) = E\left(X^2\right) - E(X)^2 \Rightarrow \sigma^2 = E\left(X^2\right) - \mu^2 \Rightarrow E\left(X^2\right) = \sigma^2 + \mu^2$.

Also $Var\left(\hat{\mu}\right) = Var\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) = \frac{\sigma^2}{n}$ and thus,
$Var\left(\hat{\mu}\right) = E\left(\hat{\mu}^2\right) - E(\hat{\mu})^2 \Rightarrow E\left(\hat{\mu}^2\right) = \frac{\sigma^2}{n} + \mu^2$.

Hence, $E\left(X_i^2\right) - E\left(\hat{\mu^2}\right) = \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2 = \left(1 - \frac{1}{n}\right)\sigma^2 \neq \sigma^2$

c)

Now that it's known that $E\left(\hat{\sigma^2}\right) = \left(1 - \frac{1}{n}\right)\sigma^2 = \frac{n-1}{n}\sigma^2$. To get the unbiased estimator, we just have to multiply the $E\left(\hat{\sigma^2}\right)$ by $\frac{n}{n-1}$

as $\frac{n}{n-1}E\left(\hat{\sigma^2}\right) = \left(\frac{n}{n-1}\right)\left(\frac{n-1}{n}\right)\sigma^2 = \sigma^2$.

Therefore, the unbiased estimator is
$\left(\frac{N}{N-1}\right)\sigma^2 = \left(\frac{N}{N-1}\right)\left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(X_i - \hat{\mu}\right)^2 = \left(\frac{1}{N-1}\right)\sum_{i=1}^{N}\left(X_i - \hat{\mu}\right)^2$.

Task 3

a)

Conditional distribution of a multivariate normal distribution with two dimensions is just a normal distribution.

We can calculate the mean of the conditional distribution $\mu_C = \mu_2 + \Sigma_{2,1}\Sigma_1^{-1}(x_1 - \mu_1)$, where μ_2 is the mean of the second random variable and $\Sigma_{2,1}$ the covariate matrix of both variables, Σ_1^{-1} the inverse of the variance of the first random variable and μ_1 the mean of the first variable.

By inserting the values we get

$$\mu_C = 8 + 3 \cdot 4^{-1} \cdot (10 - 4)$$
$$\mu_C = 12.5$$

as the mean of the conditional distribution.

The conditional variance is calculated with $\Sigma_C = \Sigma_2 - \Sigma_{2,1}\Sigma_1^{-1}\Sigma_{2,1}^T$. Let's calculate.

$$\Sigma_C = 4 - 3 \cdot 4^{-1} \cdot 3$$
$$\Sigma_C = 1.75$$

Thus, we get $\mathcal{N} \sim (12.5, 1.75)$ as the conditional distribution $p(x_2|x_1 = 10)$.

b)

We know that the variance of the distribution $p(x_2|x_1)$ is defined as $\Sigma_C = \Sigma_2 - \Sigma_{2,1}\Sigma_1^{-1}\Sigma_{2,1}^T$, where Σ_2 stands for the variance of X_2, $\Sigma_{2,1}$ the covariance between X_1 and X_2 and Σ_1 the variance of X_1.

In terms of statistical coefficients, variance and covariance are defined based on all possible values of X_1 and X_2, i.e. all possible observations. Now, because none of the terms of Σ_C rely solely on a single observation of the random variable X_1, we can say that the observed value of x_1 in the conditional variable does not affect the variance of the conditional distribution.

In case of $p(x_2|x_1 = 1)$, $p(x_2|x_1 = 2)$ and $p(x_2|x_1 = 3)$, we also know that the covariance matrix is the same for all values of x_1 and x_2. Thus, Σ_C does not change with different observations of x_1.

Task 4

a)

Let's define a 6-dimensional gaussian process to estimate the sales of a bakery X_1 as count of products bought throughout the year.

Sales are affected by the weekday $X_2 \in \{1, 2, 3, 4, 5, 6, 7\}$, by the binary value indicating whether date is a holiday or not $X_3 \in \{0, 1\}$, by weather X_4 (0 = sunny, 1 = cloudy, 2 = rainy), by outside temperature X_5 and by time of the year X_6 (month, e.g. 1 = january, 12 = december).

Mean function is a constant, that is the average sales in the previous year.

Suitable covariance functions could be the locally periodic function, squared exponential function or a mixture of them to take into account the periodic characteristics of the data and to provide smooth changes. Functions were found from https://www.cs.toronto.edu/~duvenaud/cookbook/ .

b)

Because conditional distribution $p(x_3|x_1 = 3, x_2 = 6)$ requires knowing the mean and covariance of x_1 and x_2. Let's find the covariance matrices first by using the covariance function. To use the covariance matrix, we know that $x_1 : t = 1$ and $x_2 : t = 2$. Inputting these into the covariance function we get following.

$$\Sigma_{1,2} = \begin{pmatrix} cov(1,1) & cov(1,2) \\ cov(2,1) & cov(2,2) \end{pmatrix}$$
$$\Sigma_{1,2} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$$

Inverse of it is:

$$\Sigma_{1,2}^{-1} = \frac{1}{5 \cdot 5 - 2 \cdot 2} \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix}$$
$$\Sigma_{1,2}^{-1} = \frac{1}{21} \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix}$$

Now let's compute the cross-covariance matrix between X_1, X_2 and X_3.

$$\Sigma_{1,2,3} = \begin{pmatrix} cov(3,1) \\ cov(3,2) \end{pmatrix}$$
$$\Sigma_{1,2,3} = \begin{pmatrix} 0.25 & 2 \end{pmatrix}$$

Then we know that mean is a constant $\mu = 5$.

Let's compute the mean of the conditional distribution.

$$\mu_C = \mu_2 + \Sigma_{1,2,3} \Sigma_{1,2}^{-1} (x_1 - \mu_1)$$
$$\mu_C = 5 + \begin{pmatrix} 0.25 & 2 \end{pmatrix} \frac{1}{21} \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} (\begin{pmatrix} 3 \\ 6 \end{pmatrix} - \begin{pmatrix} 5 \\ 5 \end{pmatrix})$$
$$\mu_C \approx 5.714$$

Thus, we get $\mu_C = 5.714$.

Let's compute the variance with the following:

$$\Sigma_C = \Sigma_3 - \Sigma_{1,2,3} \Sigma_{1,2}^{-1} \Sigma_{1,2,3}^T$$
$$\Sigma_C = 5 - \begin{pmatrix} 0.25 & 2 \end{pmatrix} \frac{1}{21} \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} 0.25 \\ 2 \end{pmatrix}$$
$$\Sigma_C \approx 4.128$$

Thus, we get $\Sigma_C = 4.128$.

After calculating the mean and variance with the given conditional distribution's equations, we get the following parameters for the distribution $p(x_3 | x_1 = 3, x_2 = 6)$:

$\mu_C = 5.714$ and $\Sigma_C = 4.128$.