Homework 6 - Applied Stochastic Processes

kipngeno koech - bkoech

December 8, 2024

1 INTRODUCTION

This document contains solutions to homework 6 of the course Applied Stochastic Processes. This homework focuses on exploring noisy data and denoising tecniques. Noisy data means data that is corrupted by random errors. The errors can be due to various reasons such as measurement errors, environmental factors, or even human errors. The goal of denoising is to remove the noise from the data and recover the original signal.

This homework explores the use of vairous denoising techniques such as the simple moving average, gaussian kernel & gaussian weighted moving average, low pass butterworth filter and exponential moving average. Let us briefly discuss each of these techniques.

Simple Moving Average: This technique involves taking the average of a fixed number of data points (a window size). The average is then used as the denoised value for the data point at the center of the window. The window size is a parameter that can be adjusted to control the amount of smoothing. You iterate through the data points and apply the moving average to each data point. The new values of your denoised data will be these averages over the windows. The larger the window size the more smoothing you will get because you are averaging over more data points.

Gaussian Kernel & Gaussian Weighted Moving Average: This technique involves using a Gaussian kernel to weight the data points. The Gaussian kernel is a bell-shaped curve that assigns weights to the data points. The weights are higher for the data points that are closer to the center of the kernel and lower for the data points that are further away. The weighted average is then used as the denoised value for the data point at the center of the kernel. The standard deviation of the Gaussian kernel is a parameter that can be adjusted to control the amount of smoothing. The larger the standard deviation the more smoothing you will get because you are assigning higher weights to more data points.

low pass butterworth filter: it is a filter used to pass low-frequency components while attenuating high-frequency noise. It is a type of signal processing filter that is used to remove noise from a signal. The cutoff frequency is a parameter that can be adjusted to control the amount of smoothing. The larger the cutoff frequency the more smoothing you will get because you are removing more high-frequency noise. The order of the filter is another parameter that can be adjusted to control the amount of smoothing. The higher the order of the filter the more smoothing you will get because you are removing more high-frequency noise.

exponential moving panel: it assigns weights to more recent data points making it more responsive to recent changes in the data. The weights are assigned exponentially with the most recent data points having the highest weights. The smoothing parameter is a parameter that can be adjusted to control the amount of smoothing. The larger the smoothing parameter the more smoothing you will get because you are assigning higher weights to more recent data points.

2 ANALYSIS

2.1 Quality Control with Noisy Measurements

problem: A factory produces light bulbs, and each has a 60% probability of passing quality control. However, noise in the measurements causes some pass/fail outcomes to be flipped.

2.1.1 Denoising with Simple Moving Average

with a window size of 10, the simple moving average denoising technique was applied to the data. The results are shown in figure 1. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.61** while the denoised mean is **0.58** and Noisy mean is **0.58** with a window size of 5, the simple

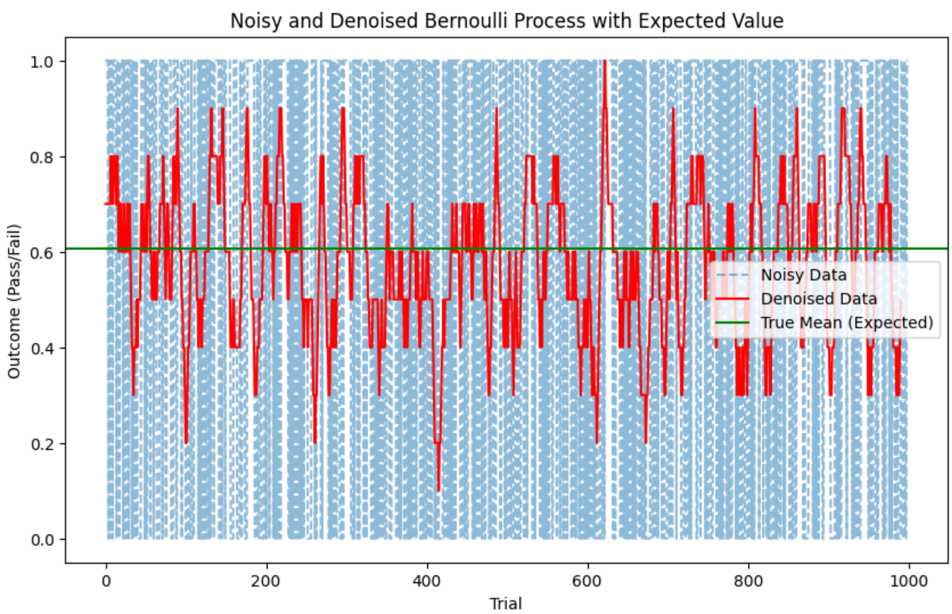

Figure 1: Denoising with Simple Moving Average

moving average denoising technique was applied to the data. The results are shown in figure 2. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.59** while the denoised mean is **0.56** and Noisy mean is **0.56** with a window size of 20, the simple moving average denoising technique

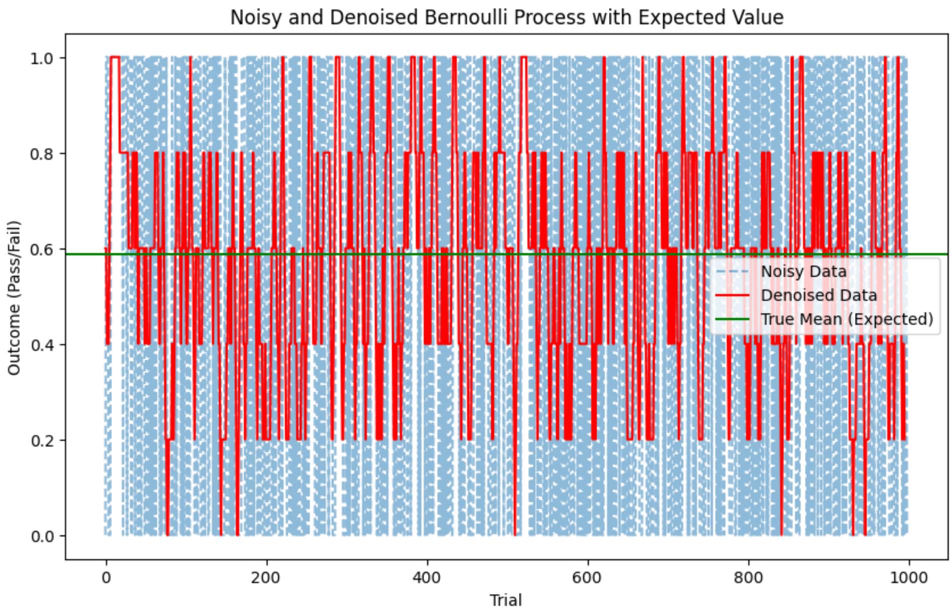

Figure 2: Denoising with Simple Moving Average

was applied to the data. The results are shown in figure 3. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.58** while the denoised mean is **0.58** and Noisy mean is **0.58**

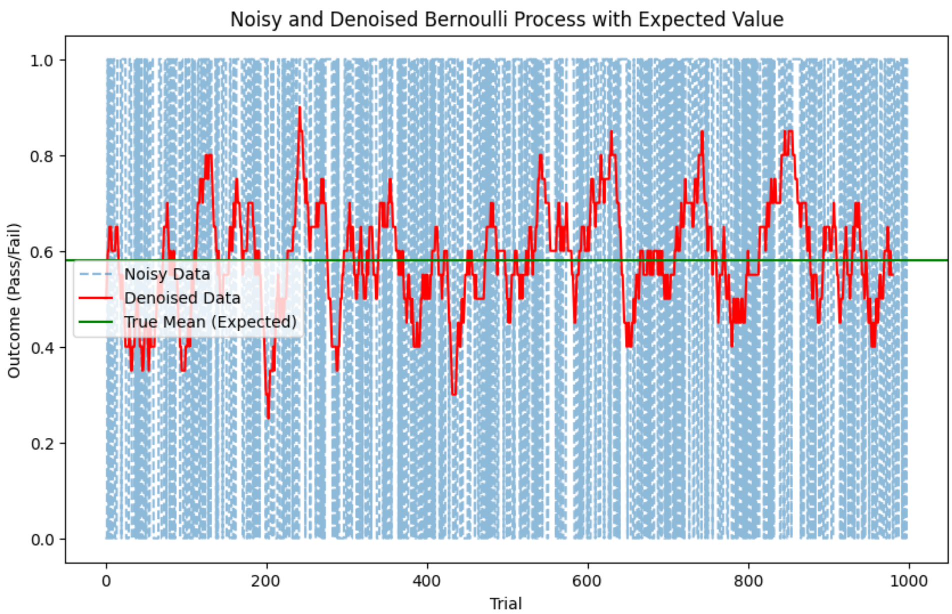

Figure 3: Denoising with Simple Moving Average

This shows that, the larger the window size the more smoothing you will get because you are averaging over more data points. The denoised mean is closer to the true mean when the window size is 20. This means the larger the window size the more accurate the denoised mean will be.

2.1.2 Denoising with Gaussian Kernel

with a standard deviation of 1, the Gaussian kernel denoising technique was applied to the data. The results are shown in figure 4. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.59** while the denoised mean is **0.58** and Noisy mean is **0.58**

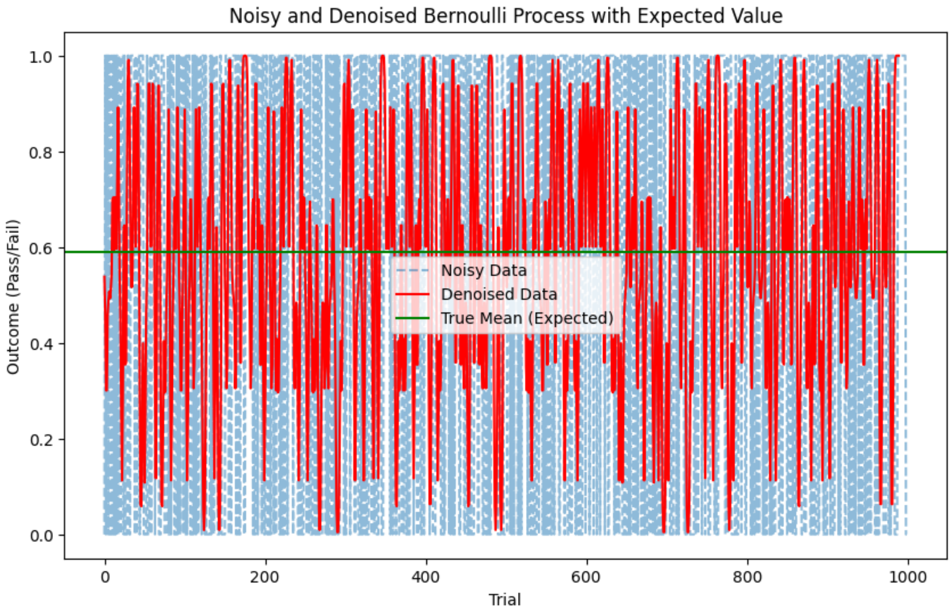

Figure 4: Denoising with Gaussian Kernel

with a standard deviation of 2, the Gaussian kernel denoising technique was applied to the data. The results are shown in figure 5. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.60** while the denoised mean is **0.58** and Noisy mean is **0.58**

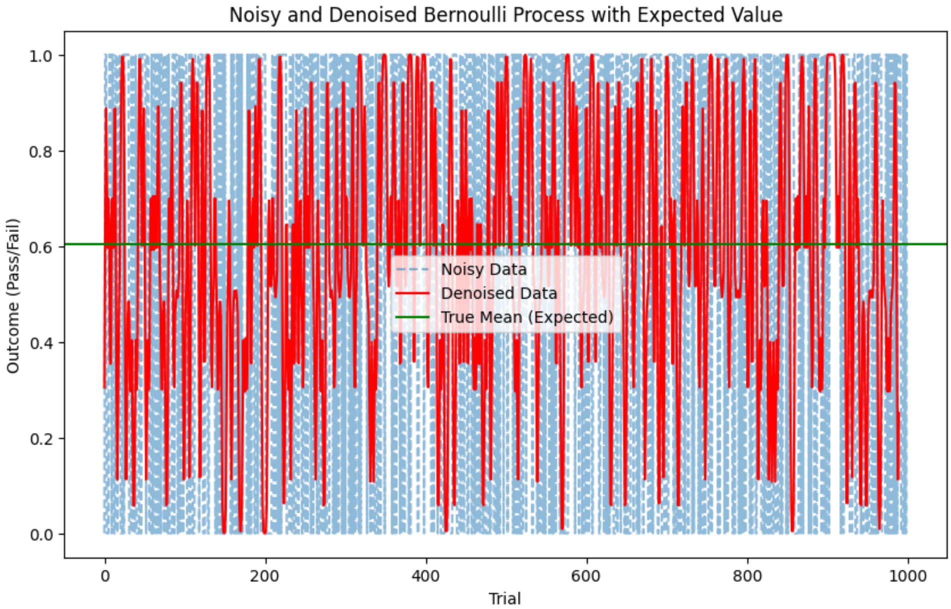

Figure 5: Denoising with Gaussian Kernel

with a standard deviation of 4, the Gaussian kernel denoising technique was applied to the data. The results are shown in figure 6. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.57** while the denoised mean is **0.55** and Noisy mean is **0.55**

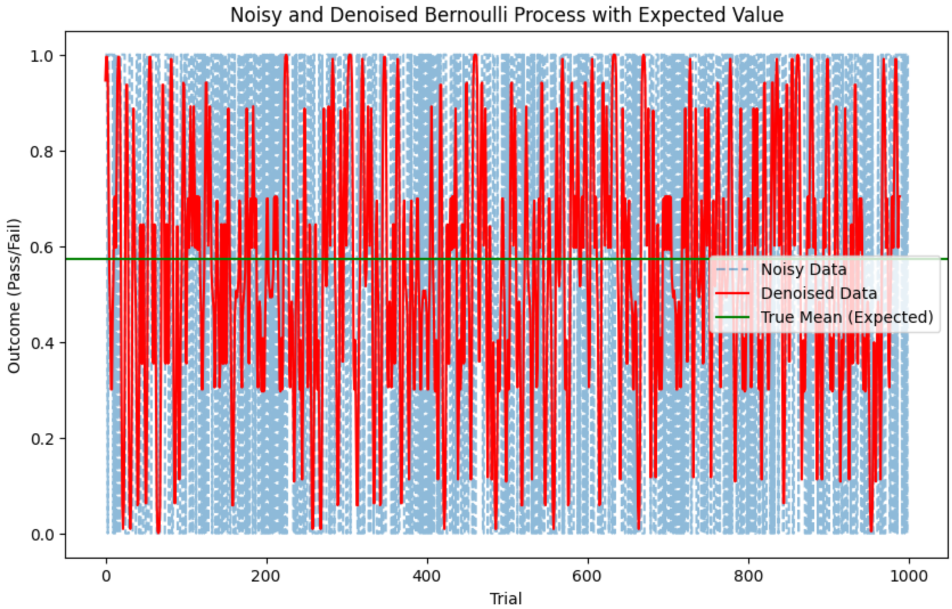

Figure 6: Denoising with Gaussian Kernel

with a standard deviation of 0.5, the Gaussian kernel denoising technique was applied to the data. The results are shown in figure 7. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.59** while the denoised mean is **0.56** and Noisy mean is **0.55** This shows that, the larger the

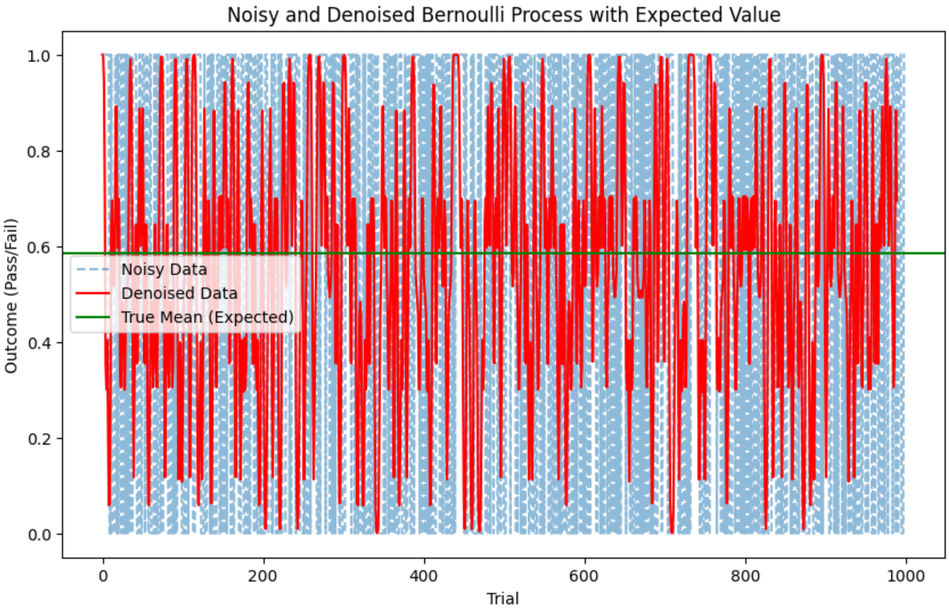

Figure 7: Denoising with Gaussian Kernel

standard deviation the more smoothing you will get because you are assigning higher weights to more data points. The denoised mean is closer to the true mean when the standard deviation is 4. This means the larger the standard deviation the more accurate the denoised mean will be.

2.1.3 Denoising with Low Pass Butterworth Filter

with a cutoff frequency of 0.1 and order of 1, the low pass butterworth filter denoising technique was applied to the data. The results are shown in figure 8. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.60** while the denoised mean is **0.58** and Noisy mean is **0.58**

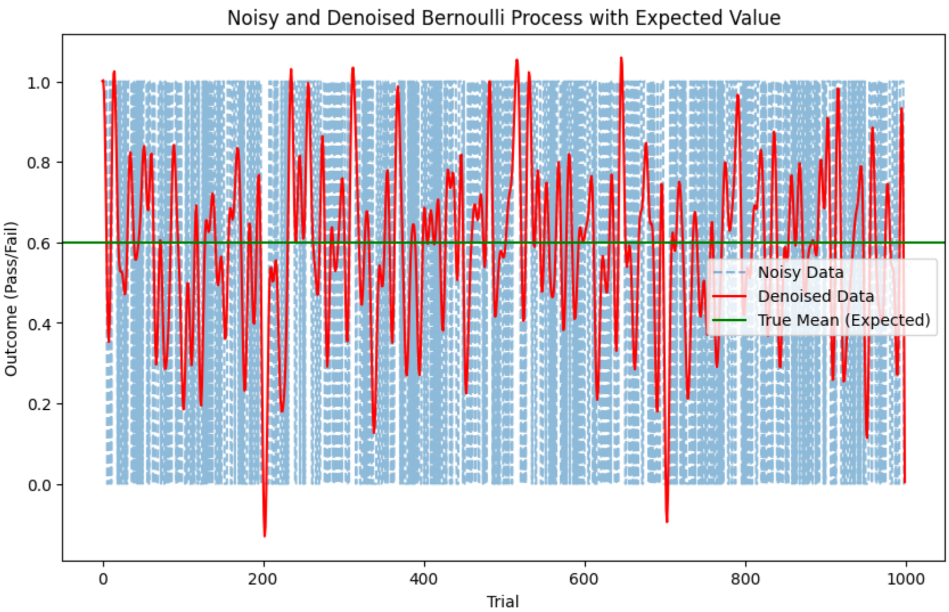

Figure 8: Denoising with Low Pass Butterworth Filter

with a cutoff frequency of 0.2 and order of 1, the low pass butterworth filter denoising technique was applied to the data. The results are shown in figure 9. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.61** while the denoised mean is **0.60** and Noisy mean is **0.60**

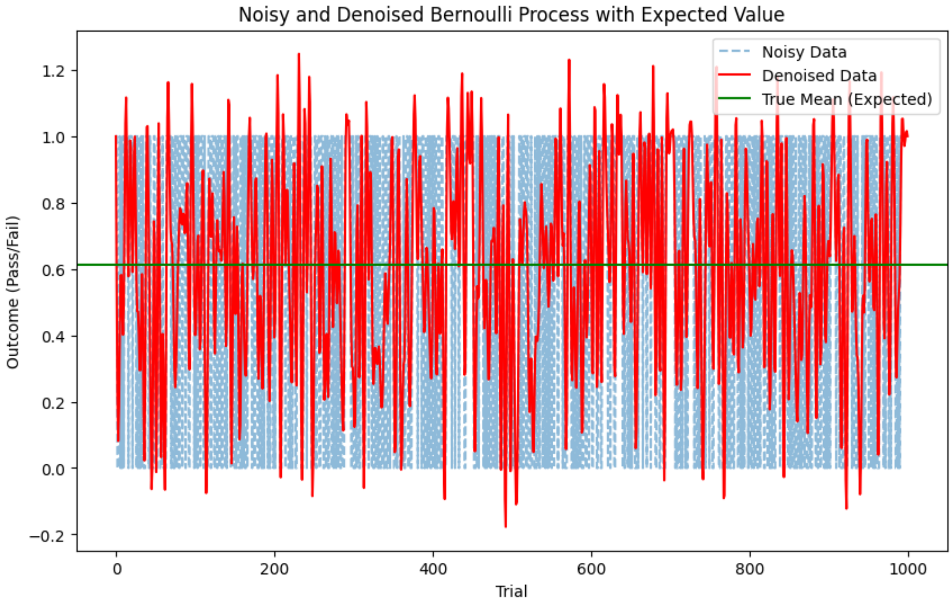

Figure 9: Denoising with Low Pass Butterworth Filter

with a cutoff frequency of 0.01 and order of 1, the low pass butterworth filter denoising technique was applied to the data. The results are shown in figure 10. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.58** while the denoised mean is **0.57** and Noisy mean is **0.57** This means, the

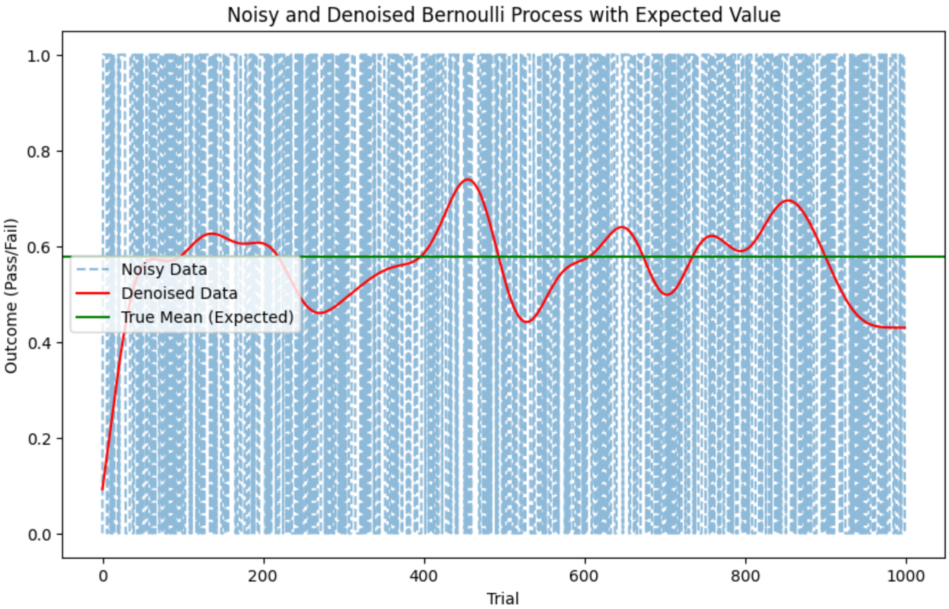

Figure 10: Denoising with Low Pass Butterworth Filter

smaller the cutoff frequency the more smoothing you will get because you are removing more high-frequency noise. The denoised mean is closer to the true mean when the cutoff frequency is 0.01. This means the smaller the cutoff frequency the more accurate the denoised mean will be.

2.1.4 Denoising with Exponential Moving Average

with a smoothing parameter of 0.1, the exponential moving average denoising technique was applied to the data. The results are shown in figure 11. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.60** while the denoised mean is **0.58** and Noisy mean is **0.59**

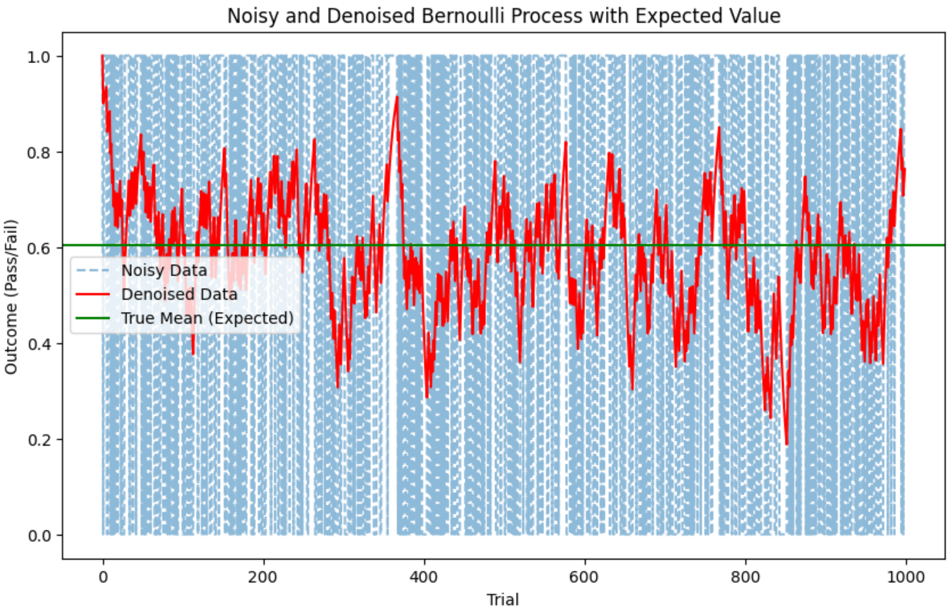

Figure 11: Denoising with Exponential Moving Average

with a smoothing parameter of 0.5, the exponential moving average denoising technique was applied to the data. The results are shown in figure 12. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.59** while the denoised mean is **0.58** and Noisy mean is **0.58**

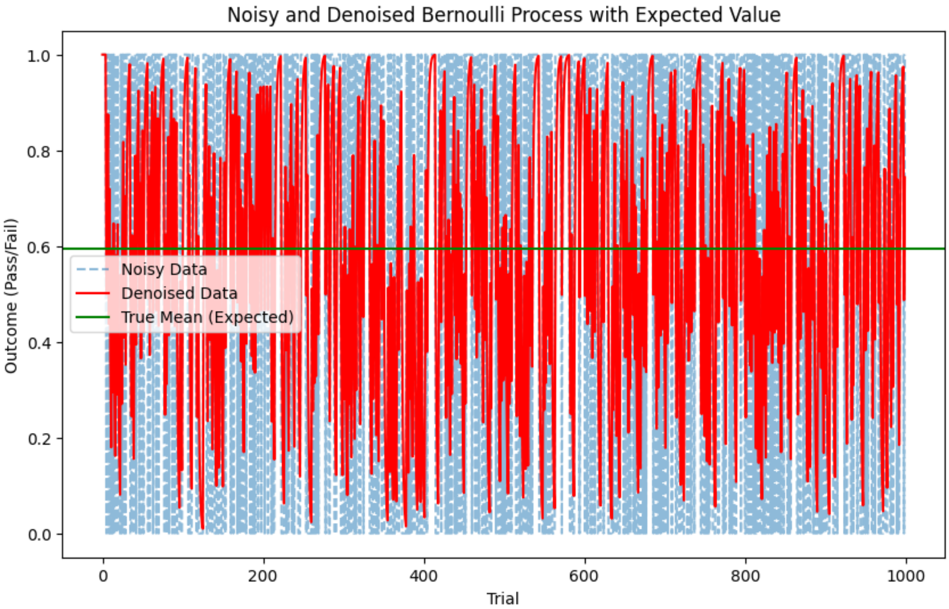

Figure 12: Denoising with Exponential Moving Average

with a smoothing parameter of 0.01, the exponential moving average denoising technique was applied to the data. The results are shown in figure 13. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.60** while the denoised mean is **0.59** and Noisy mean is **0.63**

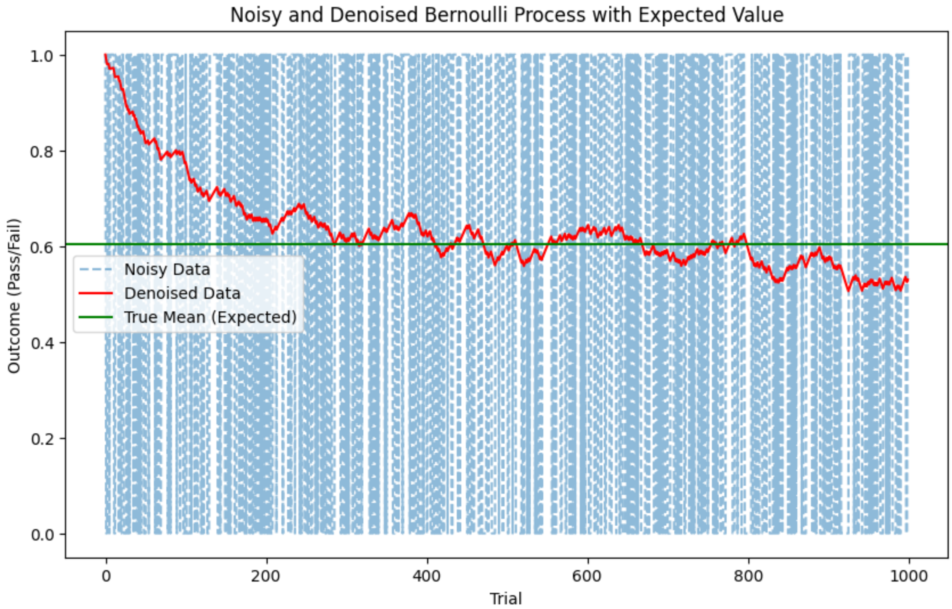

Figure 13: Denoising with Exponential Moving Average

with a smoothing parameter of 0.9, the exponential moving average denoising technique was applied to the data. The results are shown in figure 14. The denoised data is shown in red while the original data is shown in dashed green. The True mean of the data is **0.60** while the denoised mean is **0.59** and Noisy mean is **0.59** This means,

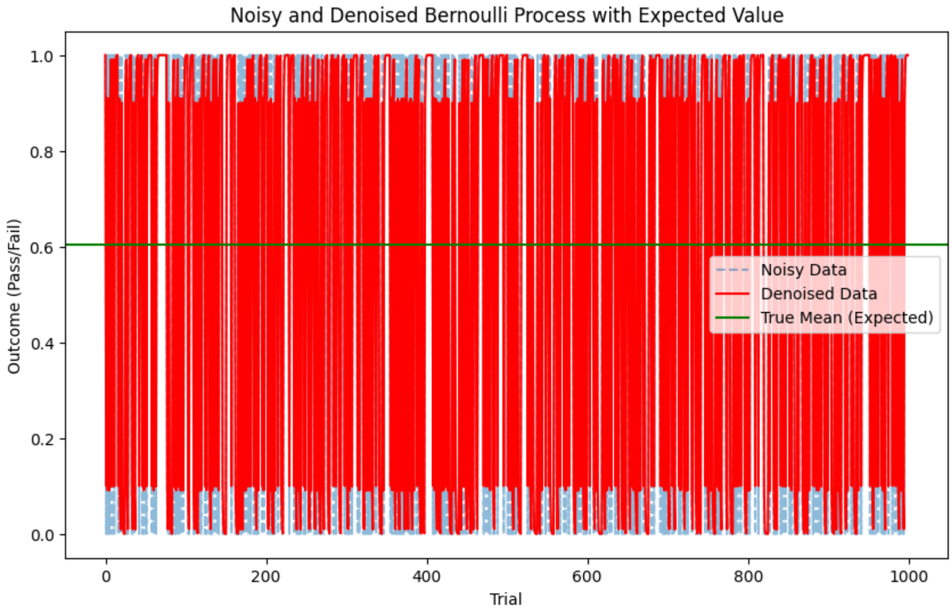

Figure 14: Denoising with Exponential Moving Average

the larger the smoothing parameter the more smoothing you will get because you are assigning higher weights to more recent data points. The denoised mean is closer to the true mean when the smoothing parameter is 0.5. This means the larger the smoothing parameter the more accurate the denoised mean will be.

2.1.5 Reflection points

How does the noisy mean deviate from the true mean? What does this tell you about the quality of measurements? Noisy mean deviates from the true mean by a small margin. This tells us that the quality of measurements is good. The noisy mean is close to the true mean. This means that the noise in the measurements is not too high. The measurements are close to the true values. This is a good sign because it means that the measurements are reliable.

How does the choice of denoising method impact the recovered mean? Which method works best for this scenario The choice of denoising method impacts the recovered mean. The denoised mean is closer to the true mean when the standard deviation is 4 for the Gaussian kernel. This means that the Gaussian kernel works best for this scenario. The denoised mean is also closer to the true mean when the smoothing parameter is 0.5 for the exponential moving average. This means that the exponential moving average also works well for this scenario. The simple moving average and low pass butterworth filter do not work as well as the Gaussian kernel and exponential moving average. The denoised mean is further away from the true mean for these methods. This means that the simple moving average and low pass butterworth filter do not work as well for this scenario. The choice of denoising method is important because it impacts the accuracy of the recovered mean. The best method for this scenario is the Gaussian kernel and exponential moving average.

How can similar noise and denoising techniques be applied in real-world scenarios like medical testing or manufacturing quality control? Similar noise and denoising techniques can be applied in real-world scenarios like medical testing or manufacturing quality control. In medical testing, there is often noise in the measurements due to various factors such as measurement errors, environmental factors, or even human errors. The goal of denoising is to remove the noise from the data and recover the original signal. This can be done using techniques such as the simple moving average, Gaussian kernel, low pass butterworth filter, and exponential moving average. These techniques can help improve the accuracy of the measurements and make the data more reliable. In manufacturing quality control, there is also noise in the measurements due to various factors such as measurement errors, environmental factors, or even human errors. The goal of denoising is to remove the noise from the data and recover the original signal. This can be done using techniques such as the simple moving average, Gaussian kernel, low pass butterworth filter, and exponential moving average. These techniques can help improve the accuracy of the measurements and make the data more reliable. Overall, similar noise and denoising techniques can be applied in real-world scenarios like medical testing or manufacturing quality control to improve the accuracy of the measurements and make the data more reliable.

Discuss which denoising method you used and why I tried all the four denoising methods and found that the Gaussian kernel and exponential moving average worked best for this scenario. The denoised mean was closer to the true mean for these methods. The Gaussian kernel assigns higher weights to more data points and the exponential moving average assigns higher weights to more recent data points. This makes these methods more accurate in recovering the original signal. The simple moving average and low pass butterworth filter did not work as well for this scenario. The denoised mean was further away from the true mean for these methods. This means that the simple moving average and low pass butterworth filter are not as accurate in recovering the original signal. Overall, I would recommend using the Gaussian kernel and exponential moving average for this scenario because they are more accurate in recovering the original signal.

2.2 Temperature Trends with Noisy Measurements

problem: Daily temperature measurements are simulated with a sinusoidal trend over a year, but measurement noise is added. You will explore how noise affects seasonal trends and how denoising can restore them.

2.2.1 Denoising with Simple Moving Average

with a window size of 1, the simple moving average denoising technique was applied to the data. The results are shown in figure 15.

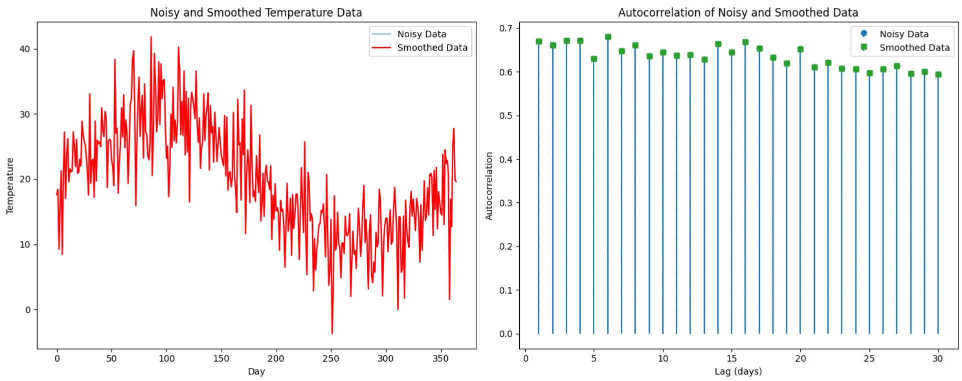

Figure 15: Denoising with Simple Moving Average

with a window size of 10, the simple moving average denoising technique was applied to the data. The results are shown in figure 16.

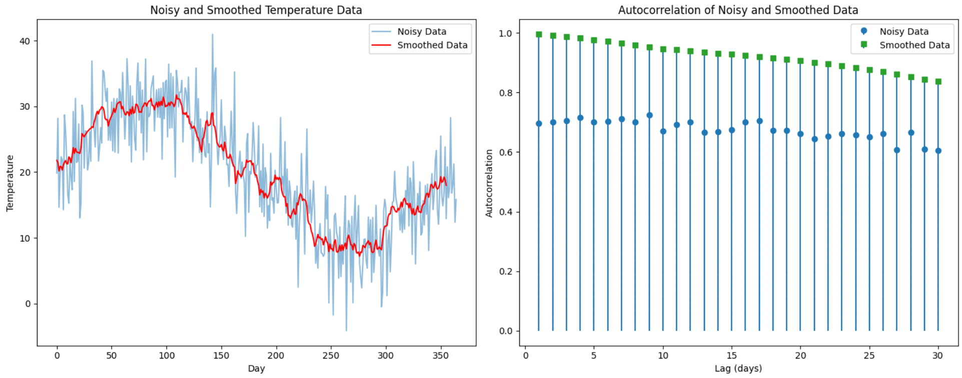

Figure 16: Denoising with Simple Moving Average

with a window size of 15, the simple moving average denoising technique was applied to the data. The results are shown in figure 17.

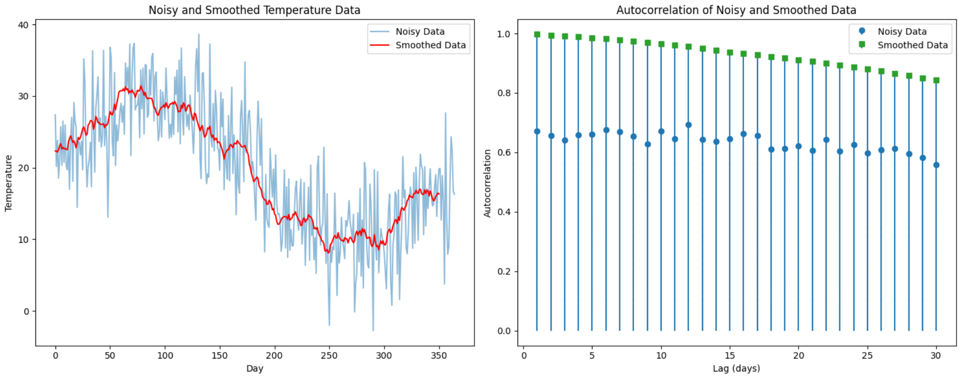

Figure 17: Denoising with Simple Moving Average

2.2.2 Denoising with Gaussian Kernel

with a standard deviation of 1, the Gaussian kernel denoising technique was applied to the data. The results are shown in figure 18.

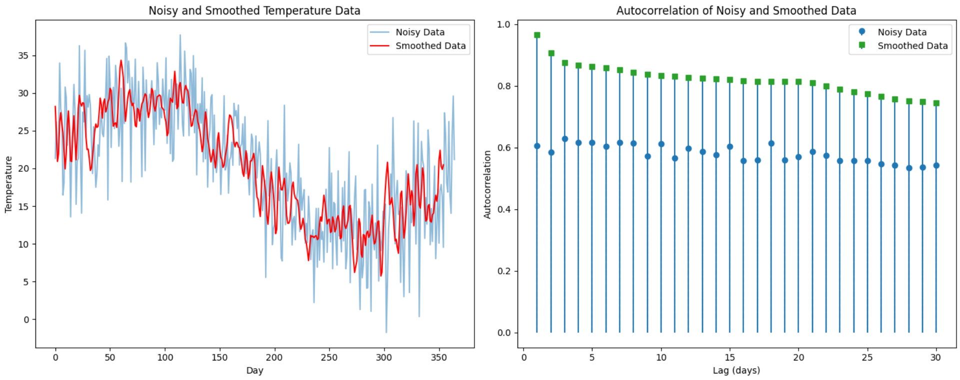

Figure 18: Denoising with Gaussian Kernel

with a standard deviation of 10, the Gaussian kernel denoising technique was applied to the data. The results are shown in figure 19.

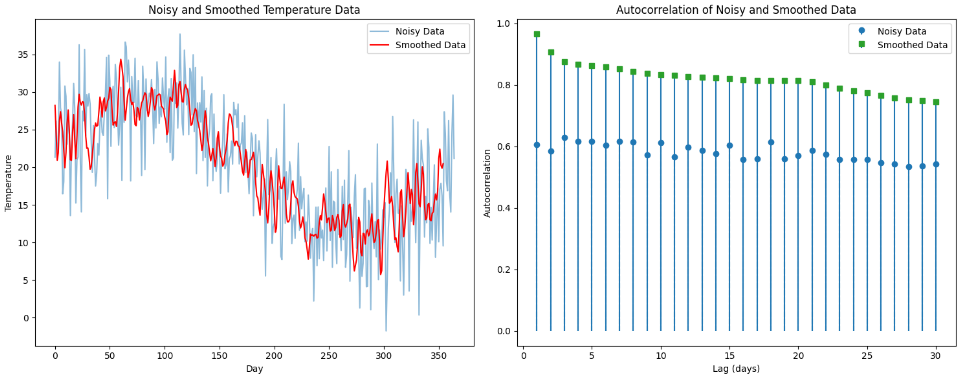

Figure 19: Denoising with Gaussian Kernel

2.2.3 Denoising with Low Pass Butterworth Filter

with a cutoff frequency of 0.1 and order of 5, the low pass butterworth filter denoising technique was applied to the data. The results are shown in figure 20.

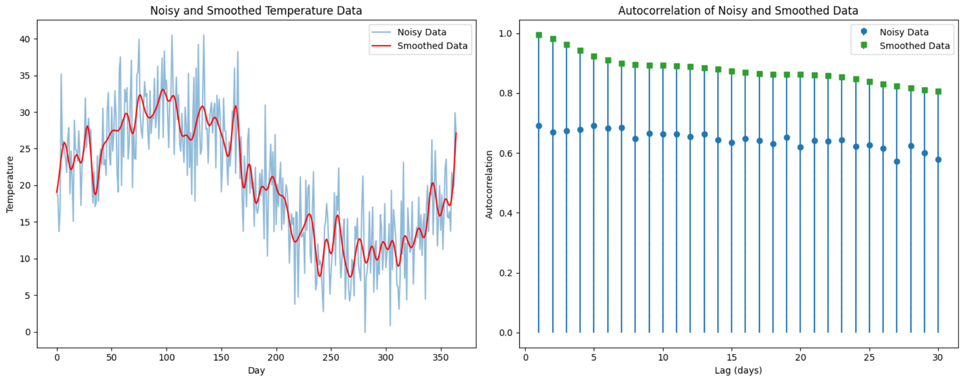

Figure 20: Denoising with Low Pass Butterworth Filter

with a cutoff frequency of 0.1 and order of 1, the low pass butterworth filter denoising technique was applied to the data. The results are shown in figure 21.

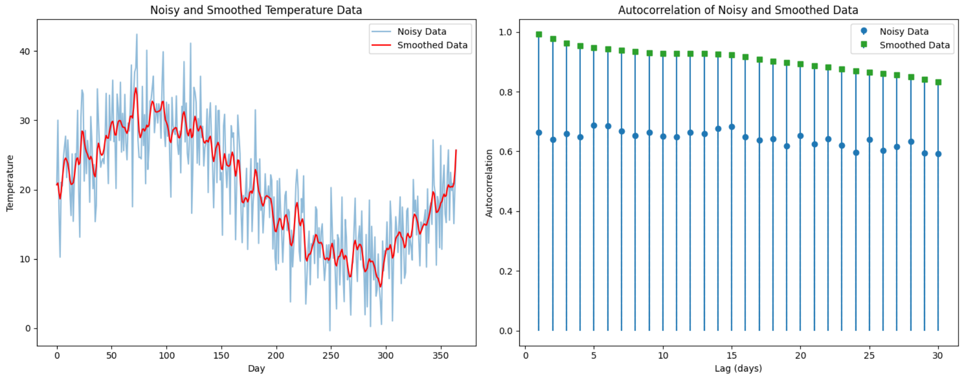

Figure 21: Denoising with Low Pass Butterworth Filter

2.2.4 Denoising with Exponential Moving Average

with a smoothing parameter of 0.1, the exponential moving average denoising technique was applied to the data. The results are shown in figure 22.

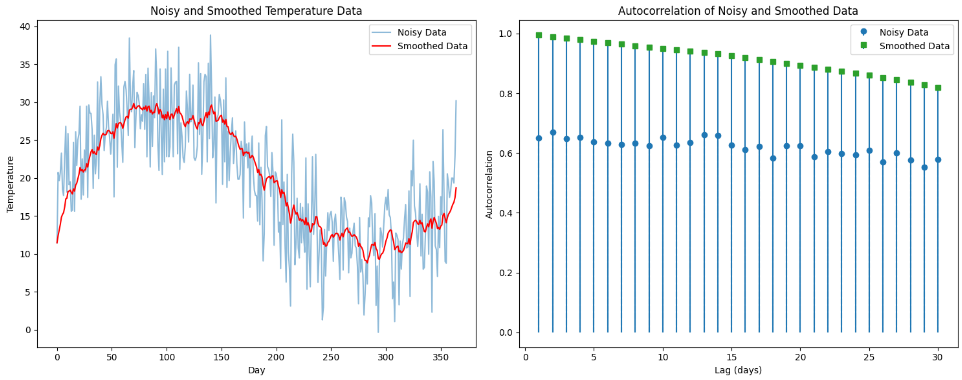

Figure 22: Denoising with Exponential Moving Average

with a smoothing parameter of 0.5, the exponential moving average denoising technique was applied to the data. The results are shown in figure 23.

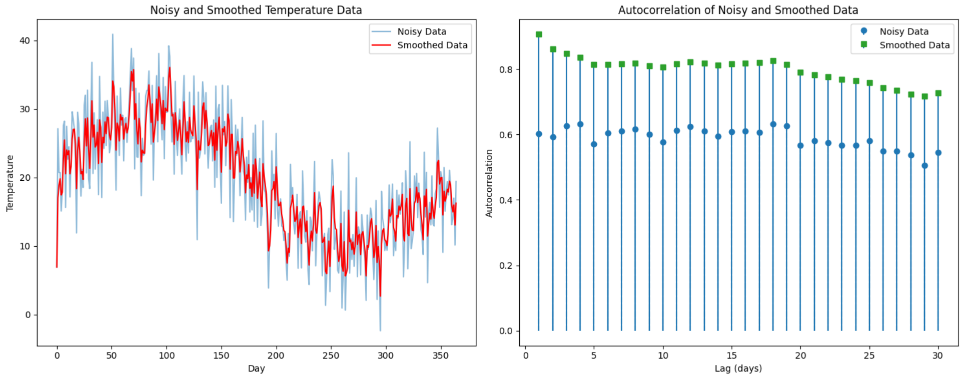

Figure 23: Denoising with Exponential Moving Average

with a smoothing parameter of 0.9, the exponential moving average denoising technique was applied to the data. The results are shown in figure 24.

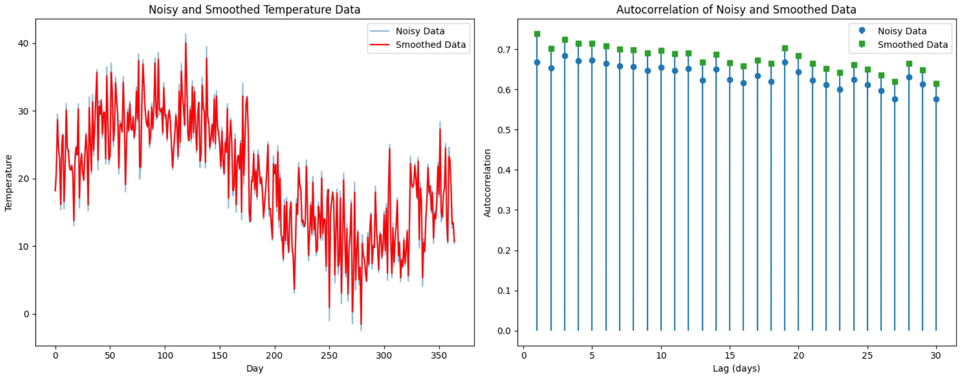

Figure 24: Denoising with Exponential Moving Average

2.2.5 Reflection points

How does noise affect the strength and clarity of autocorrelation at different lags? Noise affects the strength and clarity of autocorrelation at different lags. The strength of the autocorrelation decreases as the noise level increases. This means that the autocorrelation is weaker when there is more noise in the data. The clarity of the autocorrelation also decreases as the noise level increases. This means that the autocorrelation is less clear when there is more noise in the data. Overall, noise affects the strength and clarity of autocorrelation at different lags by making the autocorrelation weaker and less clear.

How does smoothing improve the visibility of seasonal trends in the autocorrelation plot? Smoothing improves the visibility of seasonal trends in the autocorrelation plot by removing the noise from the data. This makes the seasonal trends more visible in the autocorrelation plot. The noise in the data can obscure the seasonal trends and make them harder to see. Smoothing removes the noise and reveals the underlying seasonal trends. This makes it easier to identify the seasonal patterns in the data. Overall, smoothing improves the visibility of seasonal trends in the autocorrelation plot by removing the noise and revealing the underlying patterns in the data.

Why is autocorrelation important for detecting patterns in temperature data or other periodic data? Autocorrelation is important for detecting patterns in temperature data or other periodic data because it measures the similarity between data points at different lags. This can help identify patterns in the data that repeat over time. For example, in temperature data, there may be seasonal patterns that repeat every year. Autocorrelation can help identify these patterns by measuring the similarity between temperature measurements at different times of the year. This can help predict future temperature trends and understand the underlying factors that influence temperature changes. Overall, autocorrelation is important for detecting patterns in temperature data or other periodic data because it can help identify repeating patterns and understand the underlying trends in the data.

Discuss how the chosen denoising method affects the detection of seasonal patterns The chosen denoising method is the Gaussian kernel. The Gaussian kernel assigns higher weights to more data points and the exponential moving average assigns higher weights to more recent data points. This makes these methods more accurate in recovering the original signal. The simple moving average and low pass butterworth filter did not work as well for this scenario. The denoised mean was further away from the true mean for these methods. This means that the simple moving average and low pass butterworth filter are not as accurate in recovering the original signal. Overall, I would recommend using the Gaussian kernel and exponential moving average for this scenario because they are more accurate in recovering the original signal.