Digital Signal Processing

Statistics, Probability and Noise

Part 1

Today's Key Points

- Input signals often contain both the desired information and some level of noise.

- Statistics are used to characterize these signals
  - Mean, Standard Deviation, Variance

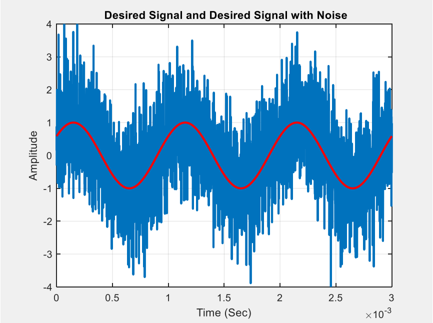
- The signal to noise ratio (SNR) is used to compare the signal level to the noise level

Signal Visualization
The Domain of a Signal

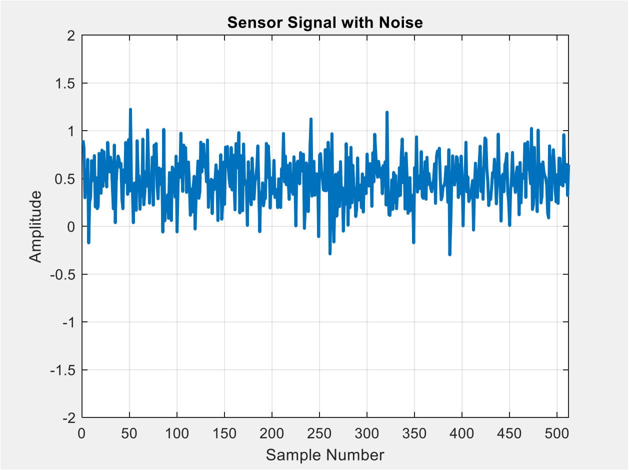
- On a graph the dependent value is on the y-axis
- The independent value on the x-axis

- The independent value (x-axis) is sometimes described as the "domain" of the signal
  - Time, Spatial, Frequency for example

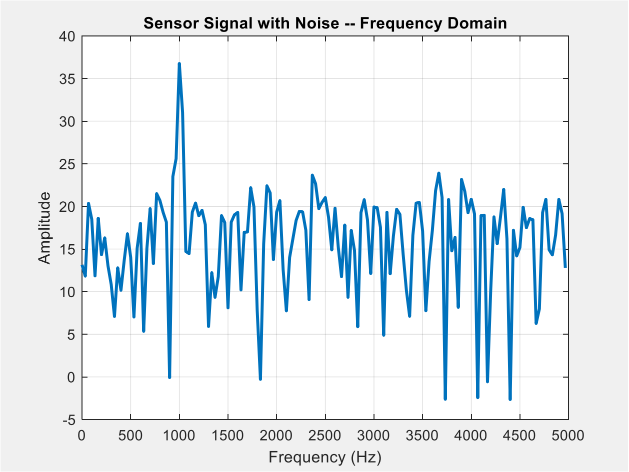
- In many cases we'll use the sample number on the x-axis.
  - We'll still refer to this as the time domain



Signal Domains

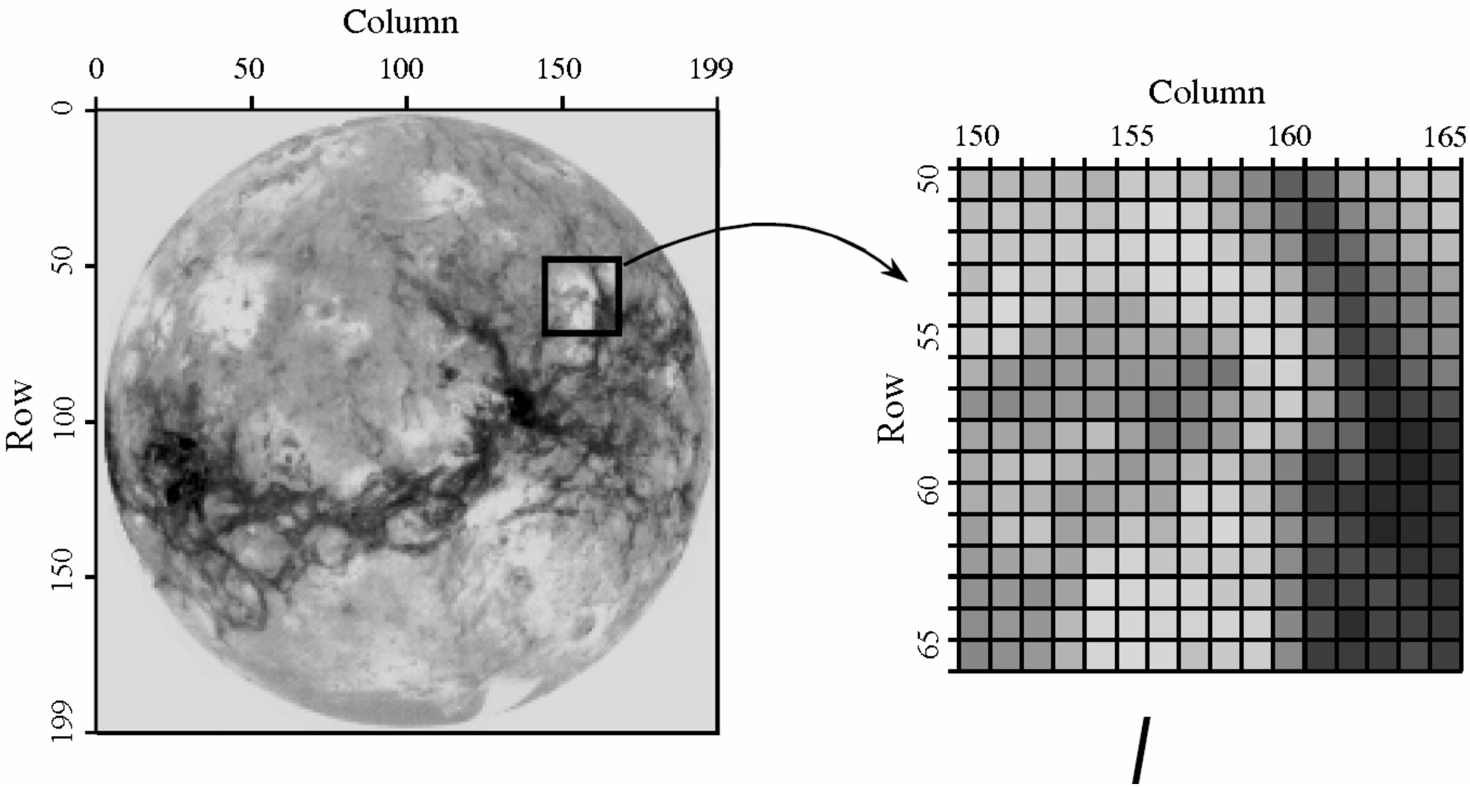


Time Domain – The independent value is time or samples



Frequency Domain – The independent value is Frequency

Signal Domains



Spatial Domain – The independent value is distance

Introduction to Noise

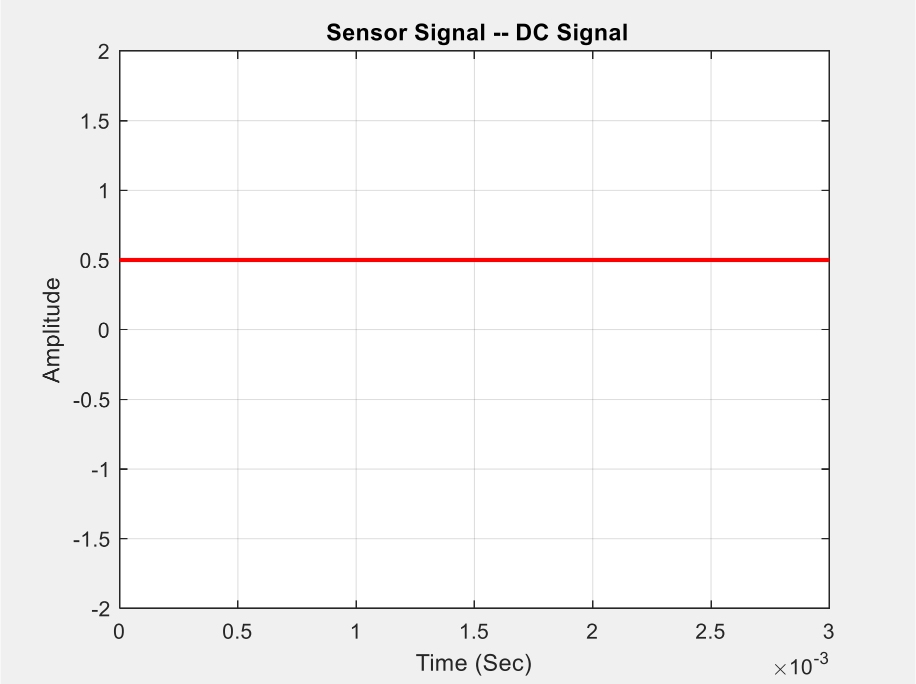
- Noise may come from many different sources

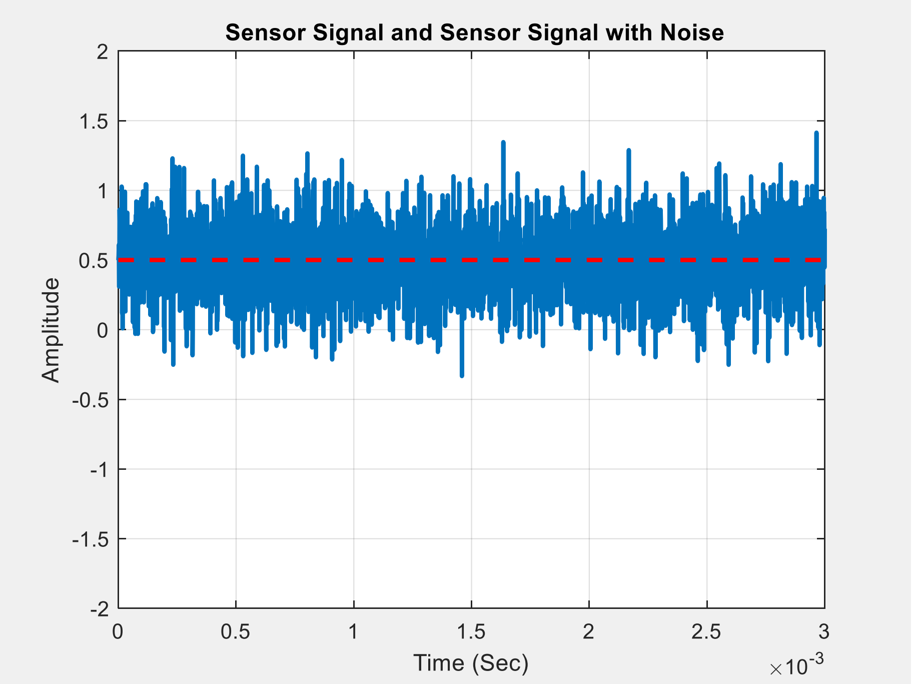
- External Noise
  - Noise inherent on the incoming signal from the analog world
  - For example noise from temperature sensors, voltage sources, strain gauges, microphones, etc..

- Internal Noise
  - Noise added by the Analog to Digital conversion process
  - Noise added by digital calculations in software or hardware

- Noise is random, or at least assumed to be random

Signal and Noise Examples
Sensor Signal



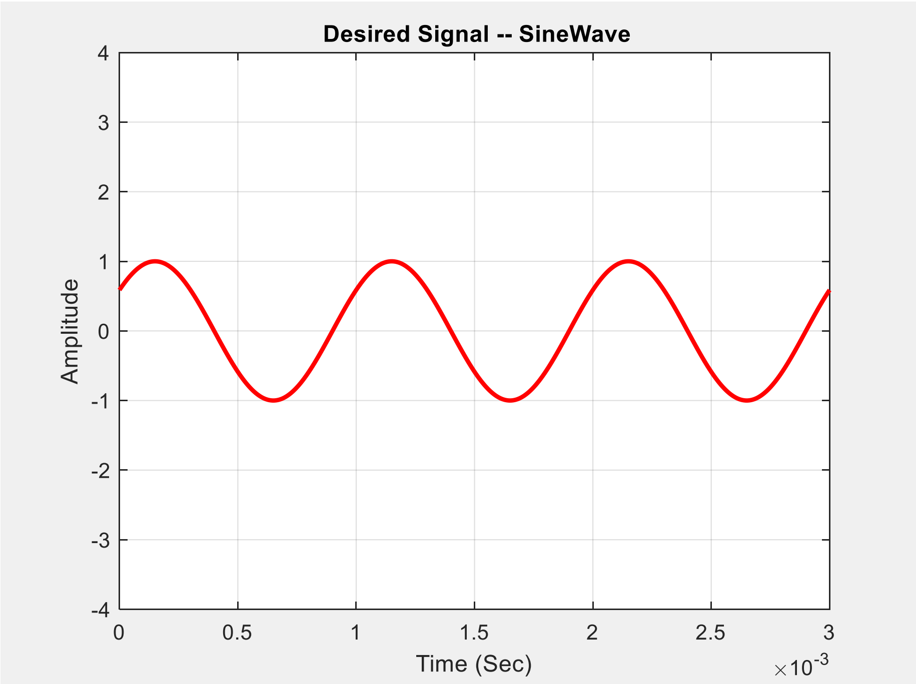


Sensor Signal

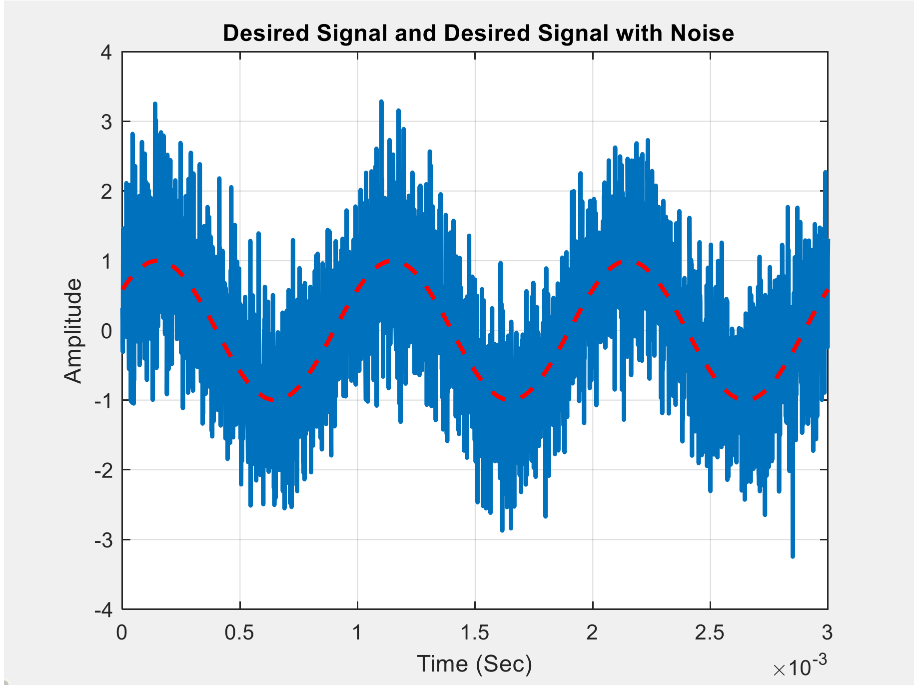
Fixed DC Value = 0.5V

Sensor Signal with Noise

Signal and Noise Example
Sinewave Signal



Desired Signal
1 kHz Sinewave



Desired Signal with Noise

Using Statistics to Characterize Signals and Noise

- Common Statistics Used to Describe Noise and Signals
  - Mean, Standard Deviation, Variance
  - Each statistic has its own purpose and its own properties

- A set of measurements of signal can be displayed using histograms to help visualize how much random noise is present.

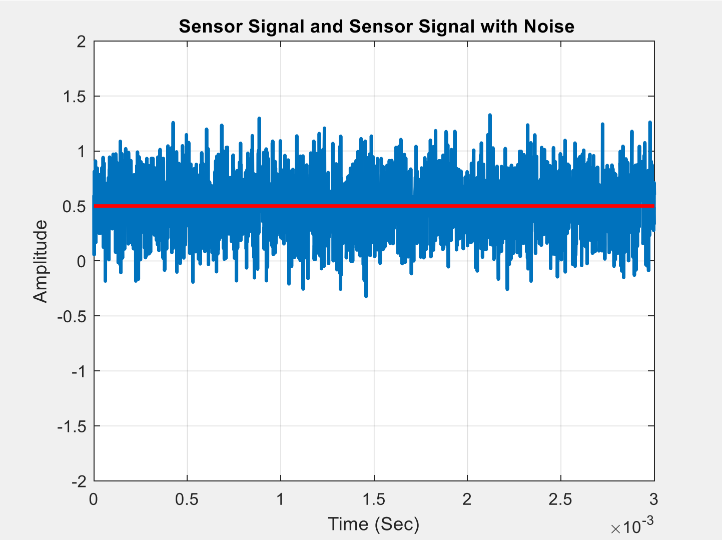
Mean

- The sample mean is the average of all the sample values.

$$\hat{\mu} = \frac{1}{N}\sum_{i=0}^{N-1} x_i$$

This is an estimate of the true mean

- In electronics it is the DC value of the signal



Mean = 0.5

Variance

- The sum of the squared difference of the signal from the mean divided by $N - 1$.

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \hat{\mu})^2$$

- Represents the <u>*power*</u> of the signal variations around the mean

Standard Deviation

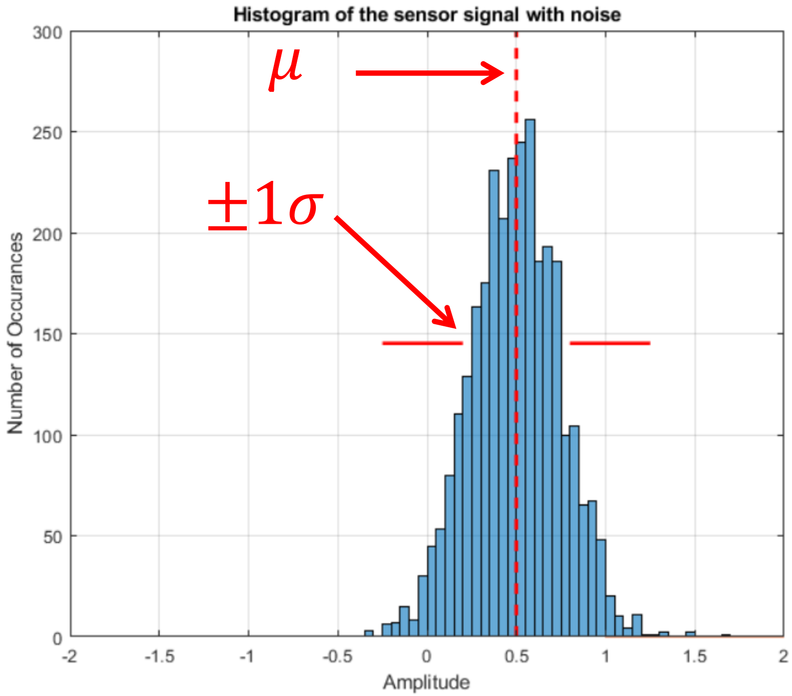
- Standard deviation is the square root of the variance

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \mu)^2}$$

- A measure of the how far the signal varies from the mean
  - Has the same units as the signal (e.g. volts)
  - Akin to the RMS value of an electrical signal

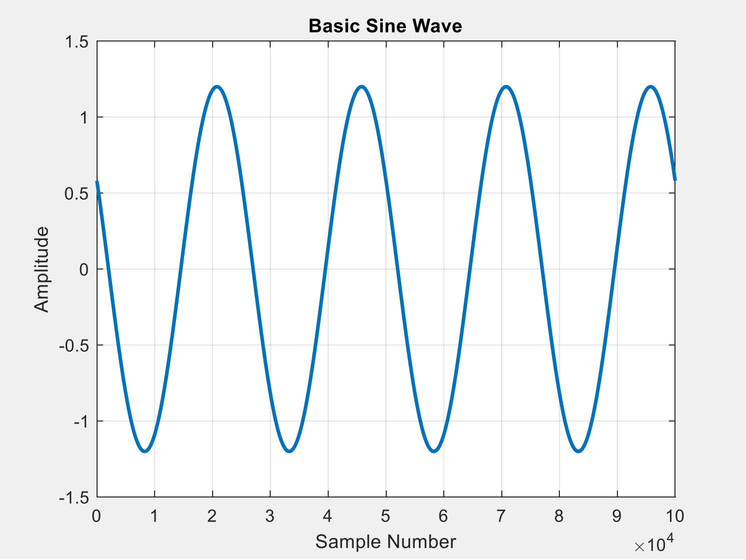
Histograms

- Some statistical parameters can be estimated from a histogram of the data
  - Mean
  - Standard Deviation



Electrical Analogy

- If an electrical signal has only an AC component, then the RMS (root mean square) of that signal is the same as its standard deviation.



$$V_{peak} = 1.2V$$

$$V_{RMS} = \frac{V_{peak}}{\sqrt{2}} = 0.8485$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \mu)^2}$$

Examples of Common Signals and Their Standard Deviations

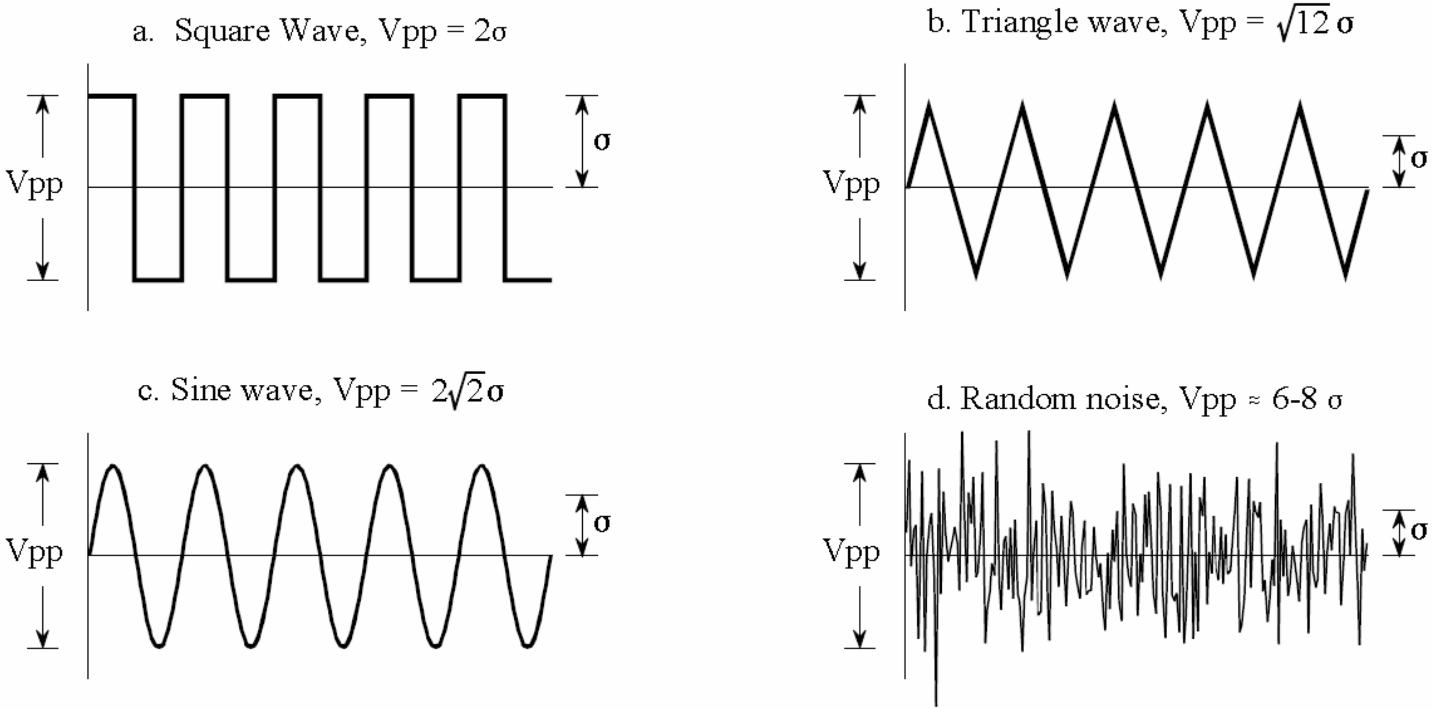
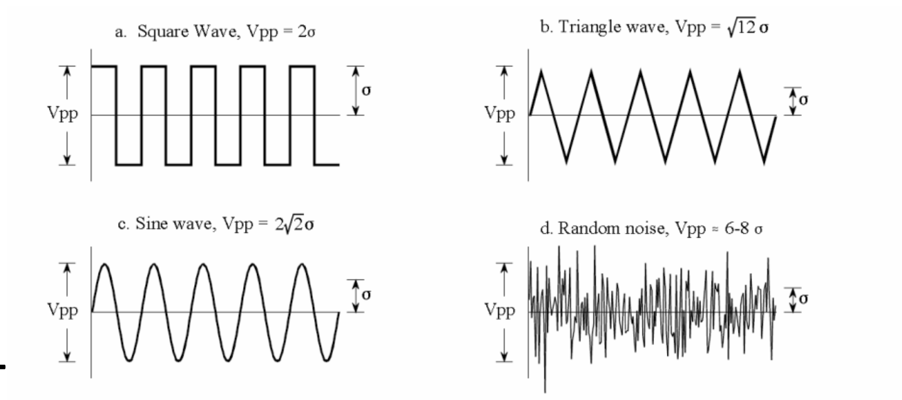


FIGURE 2-2
Ratio of the peak-to-peak amplitude to the standard deviation for several common waveforms. For the square wave, this ratio is 2; for the triangle wave it is $\sqrt{12} = 3.46$; for the sine wave it is $2\sqrt{2} = 2.83$. While random noise has no *exact* peak-to-peak value, it is *approximately* 6 to 8 times the standard deviation.

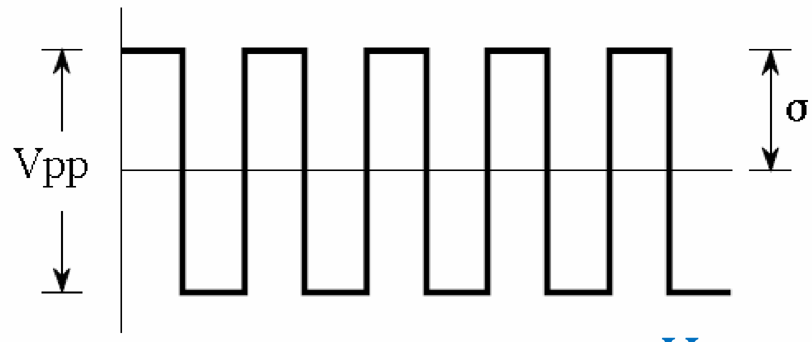
In Class Problem:

- What is the power in the fluctuation for each of these signals?
- A) Square wave with $V_{pp} = 2$
- B) Sine wave with $V_{pp} = 2.828$
- C) Triangular wave with $V_{pp} = 3.464$
- D) Random noise with $V_{pp} = 7$



In Class Problem:

- A) Square Wave with Vpp = 2



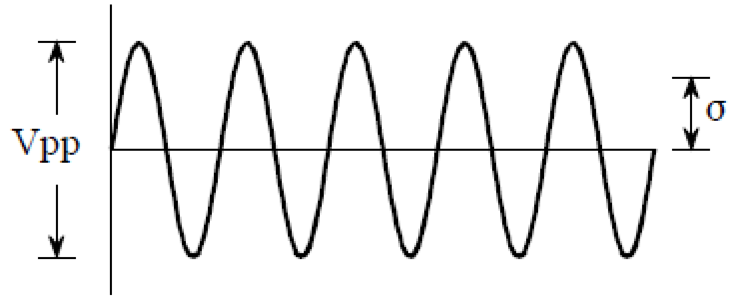
$$V_{pp} = 2\sigma \qquad \sigma = \frac{V_{pp}}{2}$$

$$\sigma = \frac{2}{2} = 1 \qquad \sigma^2 = 1^2 = 1$$

In Class Problem:

- B) A Sine Wave with $Vpp = 2.828$

c. Sine wave, Vpp = $2\sqrt{2}\sigma$

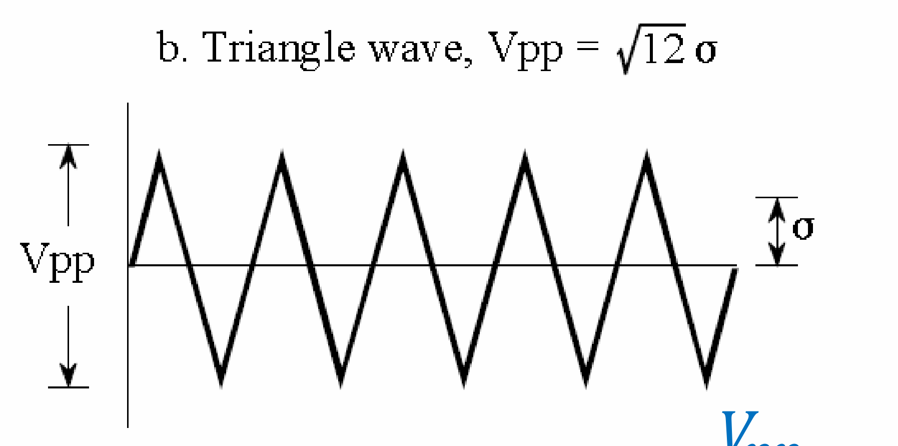


$$V_{pp} = 2\sqrt{2}\sigma \qquad \sigma = \frac{V_{pp}}{2\sqrt{2}}$$

$$\sigma = \frac{2.828}{2\sqrt{2}} = 1 \qquad \sigma^2 = 1^2 = 1$$

In Class Problem:

- C) Triangle Wave with $Vpp = 3.464$

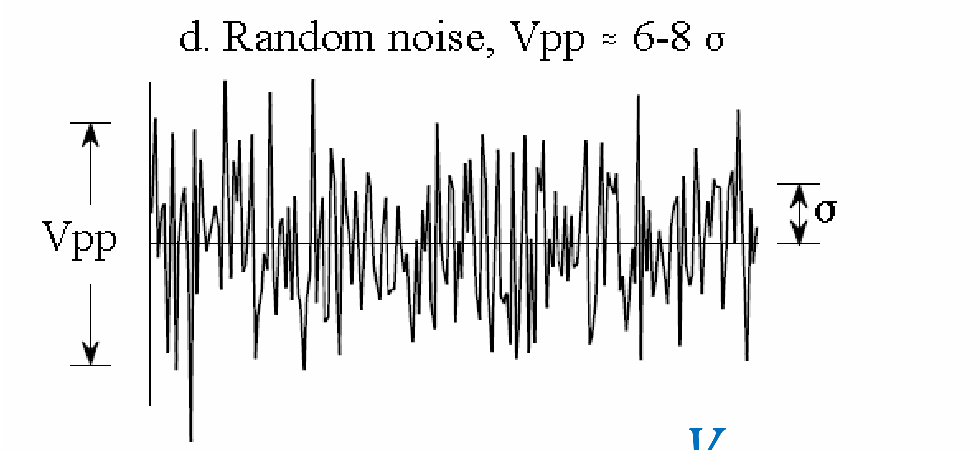


$$V_{pp} = \sqrt{12}\sigma \qquad \sigma = \frac{V_{pp}}{\sqrt{12}}$$

$$\sigma = \frac{3.464}{\sqrt{12}} = 1 \qquad \sigma^2 = 1^2 = 1$$

In Class Problem:

- D) Random Noise with $Vpp\ =\ 7$



$$V_{pp} = 6\sigma\ to\ 8\ \sigma \qquad \sigma \sim \frac{V_{pp}}{7}$$

$$\sigma = \frac{7}{7} = 1 \qquad \sigma^2 = 1^2 = 1$$

Computing Statistics on a Sample

- When computing $\hat{\sigma}$ or $\hat{\mu}$ we use a set of values that are a sample of N values $x = 6.3, 4.5, 8.9, 10.2, \ldots$

$$\hat{\mu} = \frac{1}{N}\sum_{i=0}^{N-1} x_i$$

Sum up all the samples

Divide by N

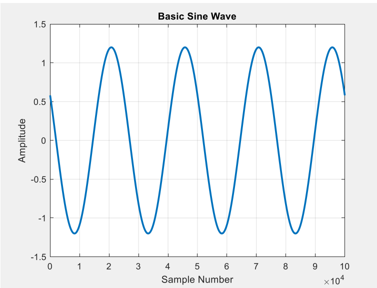
$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=0}^{N-1} (x_i - \hat{\mu})^2$$

Find $\hat{\mu}$

Sum up all the samples - $\hat{\mu}$ squared

Divide by N-1

Compute the SD using MATLAB



Compute the Variance Using all the Samples Recursively

Compute the mean

Compute the variance

```
%  Compute the standard deviation using the standard equation using all the
%  samples at once

%  Compute an estimate of the mean

N = length( basicSineWave );
meanSum = 0;
for i = 1:N
    meanSum = meanSum + basicSineWave(i);
end

meanEstimate = meanSum / N;

varSum = 0;
for i = 1:N
    varSum = varSum + (basicSineWave(i) - meanEstimate )^2;
end

varEstimate = varSum / (N-1);
stdEstimate = sqrt(varEstimate);

resultStr = sprintf('Variance Estimate -- %2.4f\nSD Estimate - %2.4f', varEstimate,
disp(resultStr)
```

$$\boxed{\mu = \frac{1}{N}\sum_{i=0}^{N-1} x_i}$$

$$\boxed{\sigma^2 = \frac{1}{N-1}\sum_{i=0}^{N-1} (x_i - \mu)^2}$$

Compute the standard deviation

```
Variance Estimate -- 0.7200
SD Estimate - 0.8485
```

Computed SD matches RMS value of 0.8485

$$V_{RMS} = \frac{1.2}{\sqrt{2}} = 0.8485$$

Mean and SD Calculations With Limited Precision

- A problem occurs when the mean is much greater than the variation.
- The difference $(x_i - \mu)$ is a very small number and round off error can occur

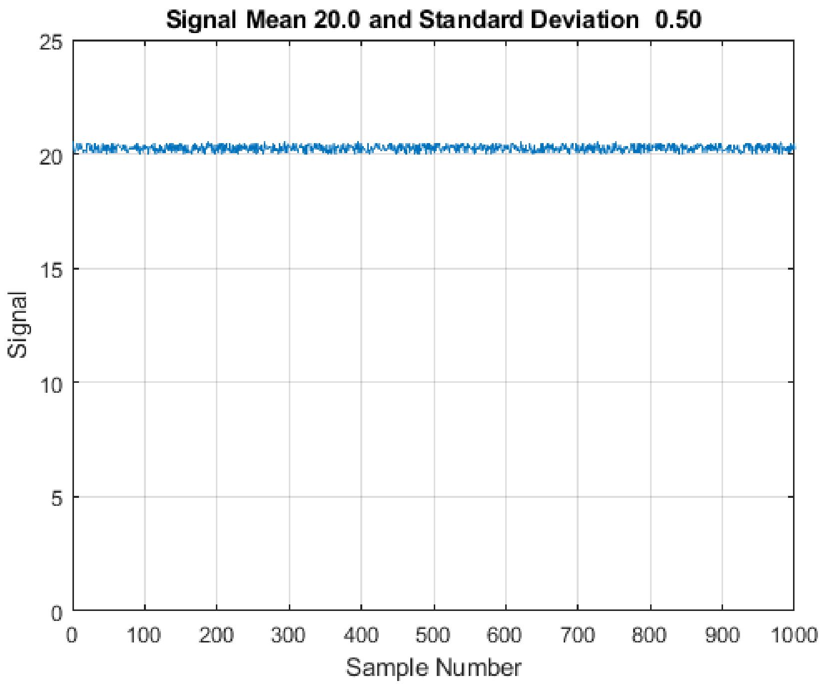
$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \hat{\mu})^2$$

Potential Round off error when this difference is small

- Occurs when the mean is much greater than the standard deviation

Mean and SD Calculations With Limited Precision

- When the mean is much larger than the standard deviation then there are small differences between the mean and each sample



$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \hat{\mu})^2$$

Potential Round off error when this difference is small

Issues with Mean and SD Calculations – Efficiency

- What happens if we get one more sample?
- Or we want to keep updating variance for each new sample?

Must save all the samples and recompute the mean for each new sample x_i

☹

Must recompute σ^2 including for each new sample x_i

$$\mu = \frac{1}{N}\sum_{i=0}^{N-1} x_i$$

$$\sigma^2 = \frac{1}{N-1}\sum_{i=0}^{N-1} (x_i - \mu)^2$$

- May require significant memory storage and computation time!

Computing Running Variance A Better Way

- Refactor the variance equation as

$$\sigma^2 = \frac{1}{N-1}\left[\sum_{i=0}^{N-1} x_i^2 - \frac{1}{N}\left(\sum_{i=0}^{N-1} x_i\right)^2\right]$$

$$\sigma^2 = \frac{1}{N-1}\left[sum\ of\ squares - \frac{sum^2}{N}\right]$$

- Keep a running value of sum of squares, the mean and the total number of samples

Alternative Calculation Using Running Statistics

$$\sigma^2 = \frac{1}{N-1}\left[\sum_{i=0}^{N-1} x_i^2 - \frac{1}{N}\left(\sum_{i=0}^{N-1} x_i\right)^2\right]$$

Need to store only N, (sum of squares) and sum

$$\sigma^2 = \frac{1}{N-1}\left[sum\ of\ squares - \frac{sum^2}{N}\right]$$

- More computationally efficient, requires less memory

MATLAB Example for Calculating Running Statistics

Compute the estimate using the running variance formula

$$\boxed{N = N + 1}$$

$$\boxed{Sum = \sum_{i=0}^{N-1} x_i}$$

$$\boxed{SS = \sum_{i=0}^{N-1} x_i^2}$$

$$\boxed{\sigma^2 = \frac{1}{N-1}\left[\sum_{i=0}^{N-1} x_i^2 - \frac{1}{N}\left(\sum_{i=0}^{N-1} x_i\right)^2\right]}$$

```
N = 0;
sumOfValues = 0;
sumSquareOfValues = 0;

varData = zeros( length( basicSineWave ), 1);

for i = 1:length( basicSineWave)

    N = N + 1;
    sumOfValues = sumOfValues + basicSineWave(i);
    sumSquareOfValues = sumSquareOfValues + basicSineWave(i)^2;

    runningVar = (  sumSquareOfValues - ( sumOfValues^2 / N ) )/ (N-1);
    varData(i) = runningVar;

end

runningVarEstimate = runningVar;
runningStdEstimate = sqrt( varEstimate );
```

In Class Problem: Running Statistics

- If you have N-1 samples and a new sample x_i is acquired, how many calculations (multiply, add, divide, square root) are required to compute the new variance using :

  1) The standard calculation

  2) The running variance calculation

- What impact does this have on calculation time?

The Standard Calculation

- Recompute the mean

  $N - 1$ additions, 1 divide

$$\mu = \frac{1}{N}\sum_{i=0}^{N-1} x_i$$

- Recompute the variance

$$\sigma^2 = \frac{1}{N-1}\sum_{i=0}^{N-1} (x_i - \mu)^2$$

$N - 1$ subtractions, N multiplies, $N - 1$ more additions, 1 divide

$3N - 2$ additions (and subtractions), N multiplies, 2 divides

The Running Variance

$$\sigma^2 = \frac{1}{N-1}\left[\sum_{i=0}^{N-1} x_i^2 - \frac{1}{N}\left(\sum_{i=0}^{N-1} x_i\right)^2\right]$$

Running Sum of Squares

1 multiply (square), 1 addition

Running Sum

1 addition

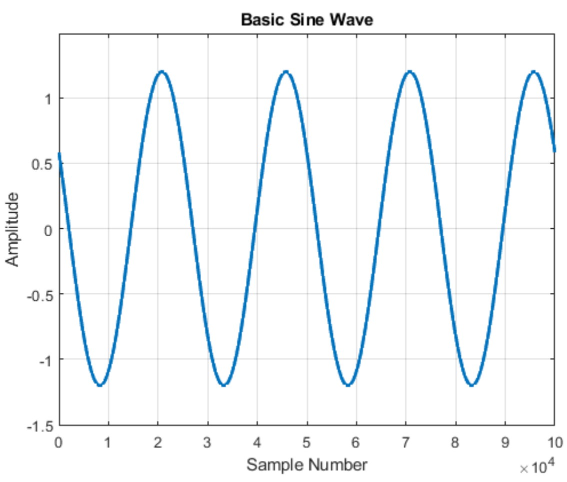
Running Variance

1 multiply (square), 1 divide, 1 addition, 1 divide

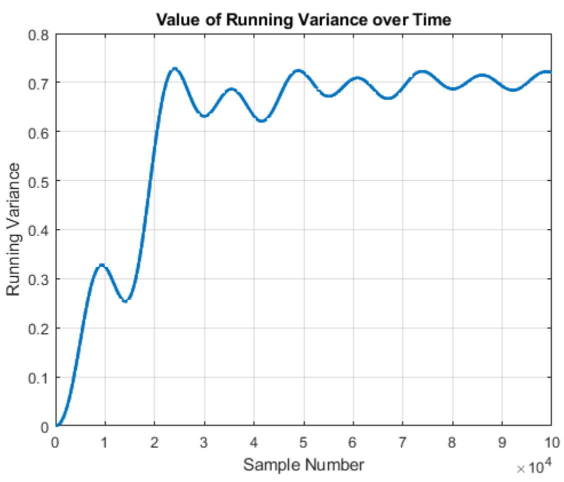
Total = 2 multiplies, 3 additions, 2 divides

Why Do We Care?

- In the breathing rate detection system the variance will be used to determine the strength of a signal
    - Need to compute a running variance efficiently



Running variance of a sinusoidal signal



Signal to Noise Ratio

- We’ve defined how to quantify a signal in terms of mean and variance
- Let’s apply this to describing how clean or noisy a signal is.

Signal to Noise Ratio (SNR)

- Crudely --
  - We'll call the part we want "signal of interest"
  - We'll call the parts we don't want "noise"

- In many cases we are interested knowing how large the signal of interest is <u>relative</u> to the noise in the signal.

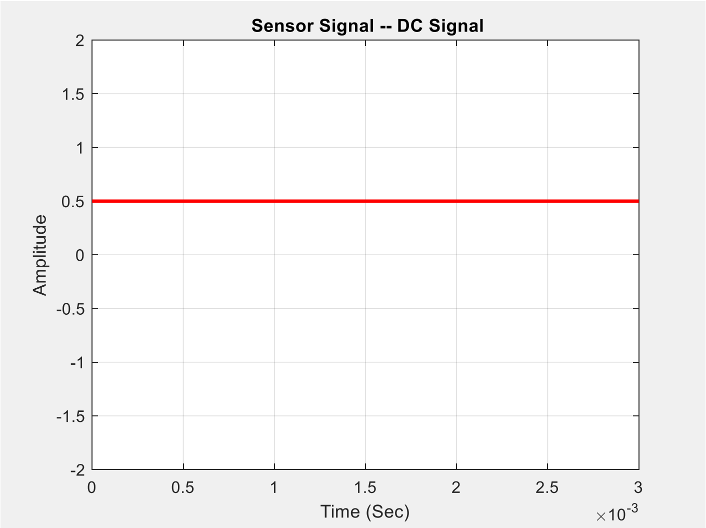
- We express this as a ratio of the signal to the noise or SNR

$$SNR = \frac{Level\ of\ the\ Signal\ of\ Interest}{Level\ of\ the\ Noise}$$

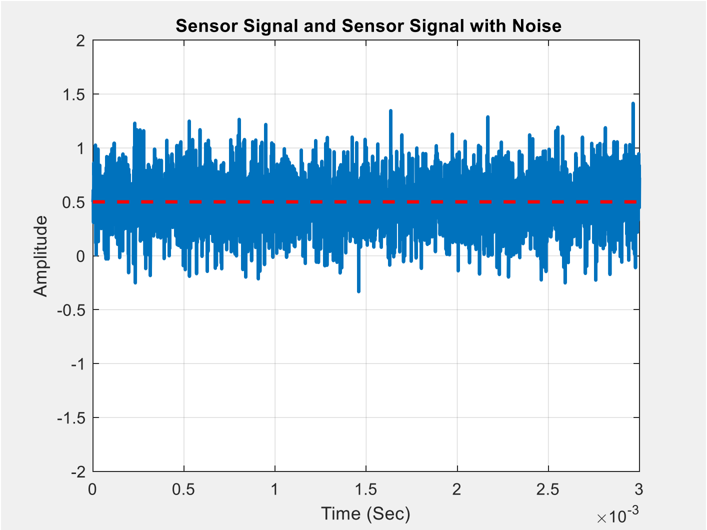
SNR has no units. It is a ratio

Signal to Noise Ratio

- In <u>some</u> cases, the <u>mean</u> describes what is being measured
- The standard deviation represents the noise.



Sensor Signal Alone
Fixed DC Value = 0.5V



Sensor Signal with Noise
$\mu = 0.5, \sigma = 0.25$

Signal to Noise Ratio (SNR)

- The text describes the SNR as the mean μ divided by the standard deviation σ.
- This is the definition often used in imaging

Signal Amplitude μ

SNR as a numerical ratio $SNR = \frac{\mu}{\sigma}$

Noise Amplitude σ

Power Signal to Noise Ratio

- In communications, the ratio of the signal power to the noise power defines the signal to noise ratio

$$SNR = \frac{P_{signal}}{P_{noise}}$$

Signal Power μ^2

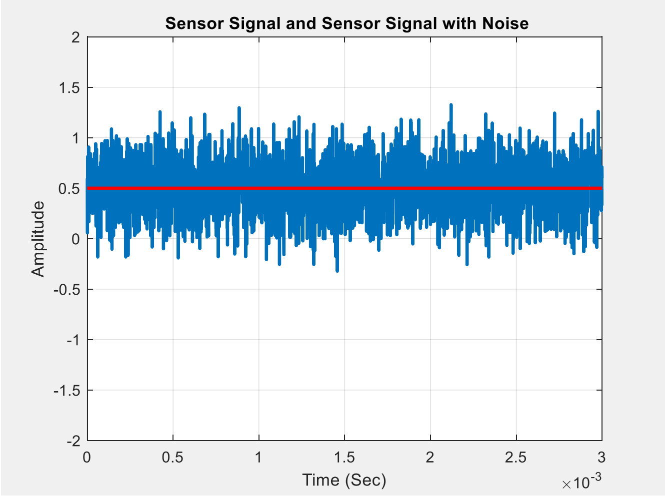
Power SNR $$SNR = \frac{\mu^2}{\sigma^2}$$

Noise Power σ^2

SNR Example

- A signal from a sensor has a mean value $\mu = 0.5$ and a standard deviation of $\sigma = 0.25$
- What is the SNR?

$$SNR = \frac{\mu}{\sigma} = \frac{0.5}{0.25} = 2$$



Sensor Signal with Noise
$\mu = 0.5, \sigma = 0.25$

Signal to Noise Ratio in Decibels

- Often the SNR is expressed in decibels

$$SNR = \frac{\mu}{\sigma} = 2$$ SNR as a numerical ratio

$$SNR_{dB} = 20 \times \log_{10}\frac{\mu}{\sigma} = 6.02\ \mathrm{dB}$$ SNR in decibels

---

$$Power\ SNR = \frac{\mu^2}{\sigma^2} = 4$$ SNR as a numerical ratio

$$Power\ SNR_{dB} = 10 \times \log_{10}\frac{\mu^2}{\sigma^2} = 6.02\ \mathrm{dB}$$ SNR in decibels

- We'll use $SNR = \mu/\sigma$ unless specifically called out

Coefficient of Variation

- Another metric is the coefficient of variation (CV) expressed in %

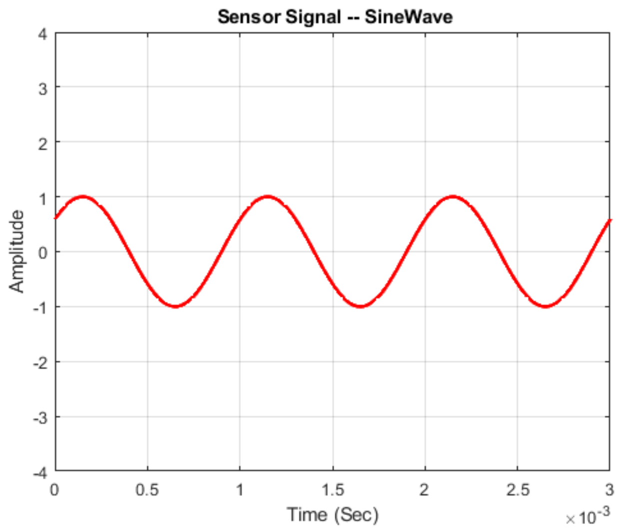
$$CV = \frac{\sigma}{\mu} \times 100$$

- Example: A signal has a mean of 6 lumens and a standard deviation of .18 lumens. Calculate the CV in %

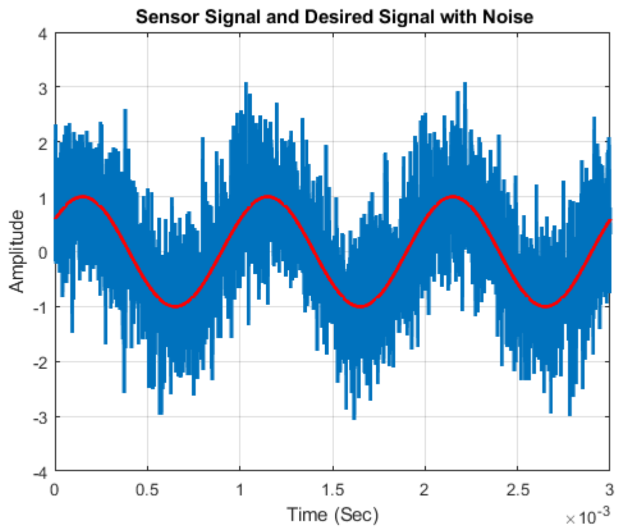
$$CV = \frac{\sigma}{\mu} \times 100 = \frac{.18}{6} \times 100 = 3.0\%$$

What about a Sinusoidal Signal?

- In <u>some</u> cases, the <u>standard deviation</u> best describes what is being measured
- The standard deviation represents the noise.



Sensor Signal Alone
Sinusoid $\sigma = 0.707$V



Sensor Signal with Noise
$\sigma_{signal} = 0.707, \sigma_{noise} = 0.75$

Signal to Noise Ratio (SNR)

- In this case use the ratio of the power of the signal to the power of the noise for SNR

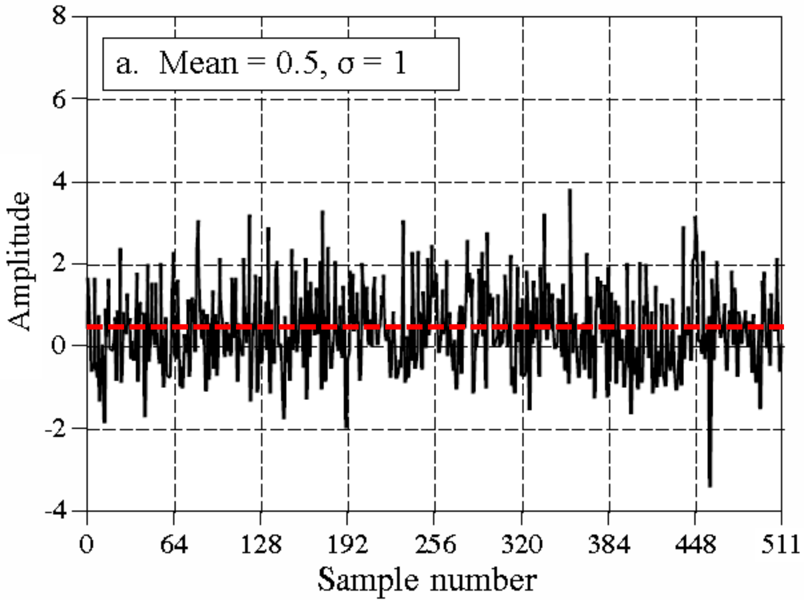
Signal Power

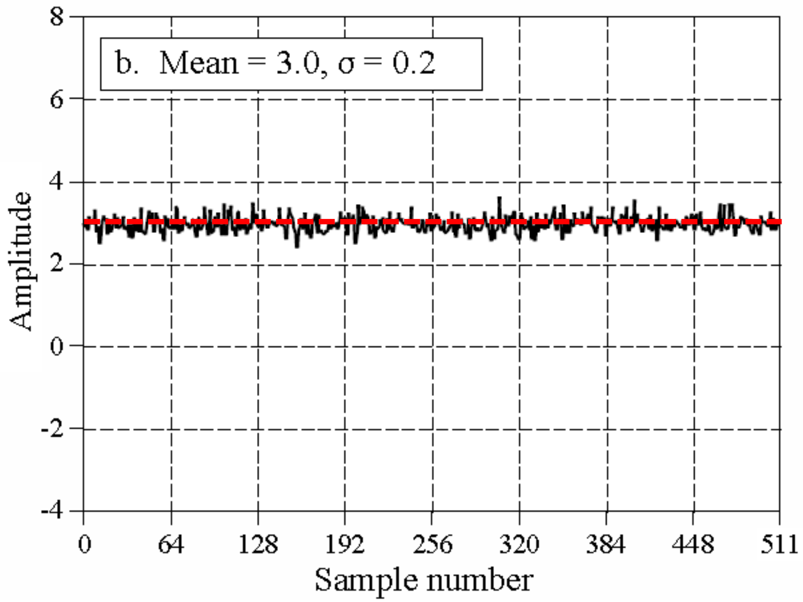
SNR as a numerical ratio $$SNR = \frac{\sigma_{signal}^2}{\sigma_{noise}^2}$$

Noise Power

SNR in decibels $$SNR_{dB} = 10\log_{10}\left(\frac{\sigma_{signal}^2}{\sigma_{noise}^2}\right)$$

SNR Examples





$$SNR = \frac{0.5}{1} = .5$$

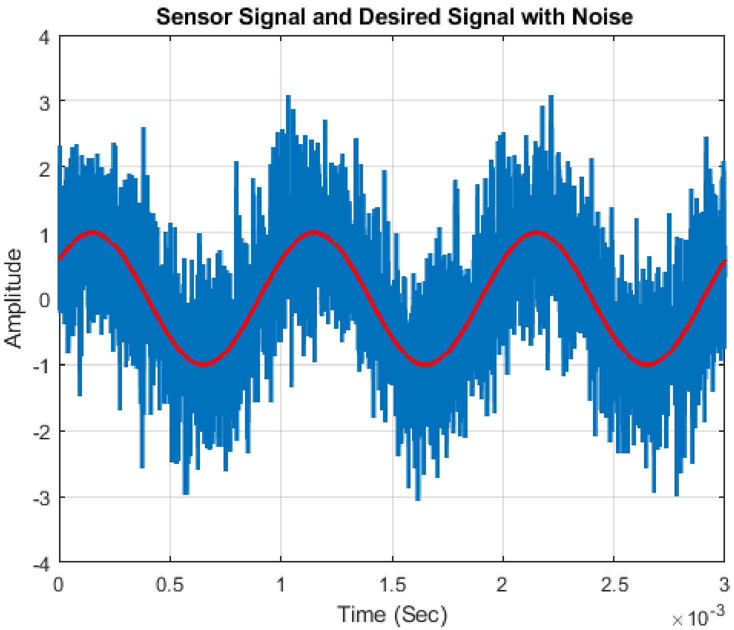
$$SNR = \frac{3.0}{.2} = 15$$

$$SNR_{dB} = 20\log_{10} .5 = -6\ dB$$

$$SNR_{dB} = 20\log_{10} 15 = 23.52$$

SNR Examples

Sensor Signal with Noise
$\sigma_{signal} = 0.707, \sigma_{noise} = 0.75$



$$SNR = \sigma_{signal}^2 / \sigma_{noise}^2$$

$$SNR = \left(\frac{0.707^2}{0.75^2}\right) = 0.89$$

$$SNR_{dB} = 10 \log_{10} 0.89 = -.51\ dB$$

In Class Problem

- The mean of a signal with noise is 3 volts
- The standard deviation of the signal is .35 volts
- Calculate the SNR as a numerical ratio and in decibels
- Calculate the coefficient of variation in %

In Class Problem

- The mean of a signal with noise is 3 volts
- The standard deviation of the signal is .35 volts
- Calculate the SNR as a numerical ratio and in decibels
- Calculate the coefficient of variation in %

$$SNR = \frac{\mu}{\sigma} = \frac{3}{.35} = 8.571$$

$$SNR_{dB} = 20 \times \log_{10} SNR = 20 \times \log_{10} 8.571$$

$$SNR_{dB} = 18.66 \text{ dB}$$

$$CV = \frac{\sigma}{\mu} \times 100 = \left(\frac{.35}{3}\right) 100 = 11.67\%$$