Measurement and Error Analysis

Prepared by: Jonathan Boylan

Group Members:

Jedreck Aquissa

Samuel Duval

Jacob Owens

Axcell Vargas

Aaron Weingarten

Section #: Group 2B

Performed on: 9/2/2021

Due on: 10/7/2021

Contents

Abstract

In this 2-part experiment, measurements were taken on the diameter of 100 glass spheres and the length, width, and height of 20 wooden blocks using English Vernier Calipers (I.L.E = 0.001 in). From the data, the means, standard deviations, and limits of errors were calculated in order to determine the best estimate and uncertainties of the corresponding measurements taken on the spheres and blocks. The results show that the diameter of the glass sphere was 0.570 ± 0.004 in and the length, width, and height of the wooden block was 3.029 ± 0.004 in, 0.859 ± 0.006 in, and 0.623 ± 0.004 in, respectively. Using the measurements of the wooden block and error propagation, the volume of the block was found to be 1.622 ± 0.014 in^3.

1 Introduction

When something is measured, a quantitative determination of a fixed physical variable or characteristic of a system, there is no way to determine if that measurement represents that variable's true value. Given the inevitability of error and uncertainty, a measurement can only be used to estimate what the true value is. To express it mathematically, we use this expression for a measurement:

$$x = x_m \pm \delta x \tag{1}$$

where x_m is the mean of a sample of measurements, and δx is the error in the measurement. The error parameter creates a possible range in which the true value will lie. As more measurements are taken, the confidence in the best value will increase and the error will narrow down into a smaller range. [1]

1.1 Mean, Mode, and Standard Deviation

Given a numerical data set, important values to consider while evaluating the data are the mean, mode, and standard deviation. These values help determine the distribution of the data and how confident one can be in the results.

The mean of a data set is indicated by the symbol μ for populations ($N > 30$) and x_m for samples ($n \leq 30$), and is determined using the equations:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2}$$

$$x_m = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{3}$$

where x_i is the data at index i, and N/n is the total number of elements in the set. The mean is an indicator of the average value in the set and is used as the best estimate of the true value of the distribution.

The mode, or most common value, of the set is found by counting like data points and is the number with the highest tally. Like the mean, the mode can be related to the value where most data points tend; however, that is not always the case.

The standard deviation, denoted σ for populations or S for samples, of the set is determined using the equations:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2} \tag{4}$$

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - x_m)^2} \tag{5}$$

The standard deviation is indicative of the spread of the data. The more precise the data is around one value, the lower the standard deviation will be.

1.2 Limits of error

The limit of error of a measurement determines the range in which the true value is likely to lie. In general, the limit of error is calculated as a function of the instrumental limits of error (I.L.E) and the statistical limits of error (S.L.E). The I.L.E is determined by the measurement device used, and is normally given by the manufacturer. If no I.L.E is given, a basic rule is to use the smallest division on the device's output scale.

The S.L.E is determined in a more statistical manner. The general equation to determine the S.L.E for a population is as follows:

$$S.L.E = \frac{3\sigma}{\sqrt{N}} \tag{6}$$

With the I.L.E and S.L.E, the overall limit of error (L.E) can be calculated, and will be done using the root-square-sum (RSS) method:

$$L.E = \sqrt{(I.L.E)^2 + (S.L.E)^2} \tag{7}$$

1.3 Error propagation

When an equation has two or more values with uncertainties, how the uncertainties affect each other must be taken into consideration. To determine the new uncertainty of an equation, there is a general formula that gives the best estimate of what that uncertainty should be, called the Root of the Sum of Squares Uncertainty, or RSS. [2] For an equation, R, with n independent variables, x, the new uncertainty is given by the RSS:

$$\delta R = \sqrt{\sum_{i=1}^{n} (\delta x_i \frac{\partial R}{\partial x_i})^2} \tag{8}$$

For the case where the equation R consists only of variables multiplied by each other, as in the volume equation $V = l * w * h$, the general formula can be simplified down by division to the expression:

$$\frac{\delta V}{V} = \sqrt{(\frac{\delta l}{l})^2 + (\frac{\delta w}{w})^2 + (\frac{\delta h}{h})^2} \tag{9}$$

1.4 Data rejection and Chauvenet's criterion

When data is being measured, there is a chance that the measurement taken will differ significantly from the other observations. These are generally known as outliers of the data. In these cases, it is common practice to ignore these data points from the statistical analysis because they will grossly affect the mean and standard deviation. This is called data rejection. To determine which data points to reject, the normalized deviation of the data point, or z-score, is compared to some threshold, and then rejected if that threshold is passed. The z-score of a data point (in a population) is found using the equation:

$$z = \frac{x - \mu}{\sigma} \tag{10}$$

Chauvenet's criterion, named after mathematician William Chauvenet, is one method to determine the threshold to reject data against. Using the number of values in the set (N), Chauvenet's criterion rejects data with a z-score corresponding to a confidence level of $1 - \frac{1}{2N}$ or greater. [3] For example, Chauvenet's criterion for a set size of $N = 20$ gives a threshold value of 2.24 and a set size of $N = 100$ gives a value of 2.81. Simply put, data 2.24 standard deviations from the mean, and data 2.81 standard deviations from the mean would be rejected in data sets of size 20 and 100, respectively.

2 Experimental Procedure

The experiment involved 100 glass spheres separated into groups of 10 into 10 labeled containers, and 20 labeled wooden blocks. The diameter of all 100 glass spheres and the length, width, and height of all 20 wooden blocks was measured using English Vernier Calipers with an instrumental limit of error (I.L.E) of 0.001 in. The measurements were logged into a notebook, then later re-entered into a digital CSV spreadsheet for analysis. Outliers of the data were rejected using Chauvenet's criterion. This resulted in 1 data point removed from both the block length and block width data sets, and no change in the sphere diameter and block height data sets.

3 Results and Discussion

3.1 100 Sphere Measurements

100 sphere measurements, shown in Table A-1, were taken and the results formed a distribution as shown in Figure 1.

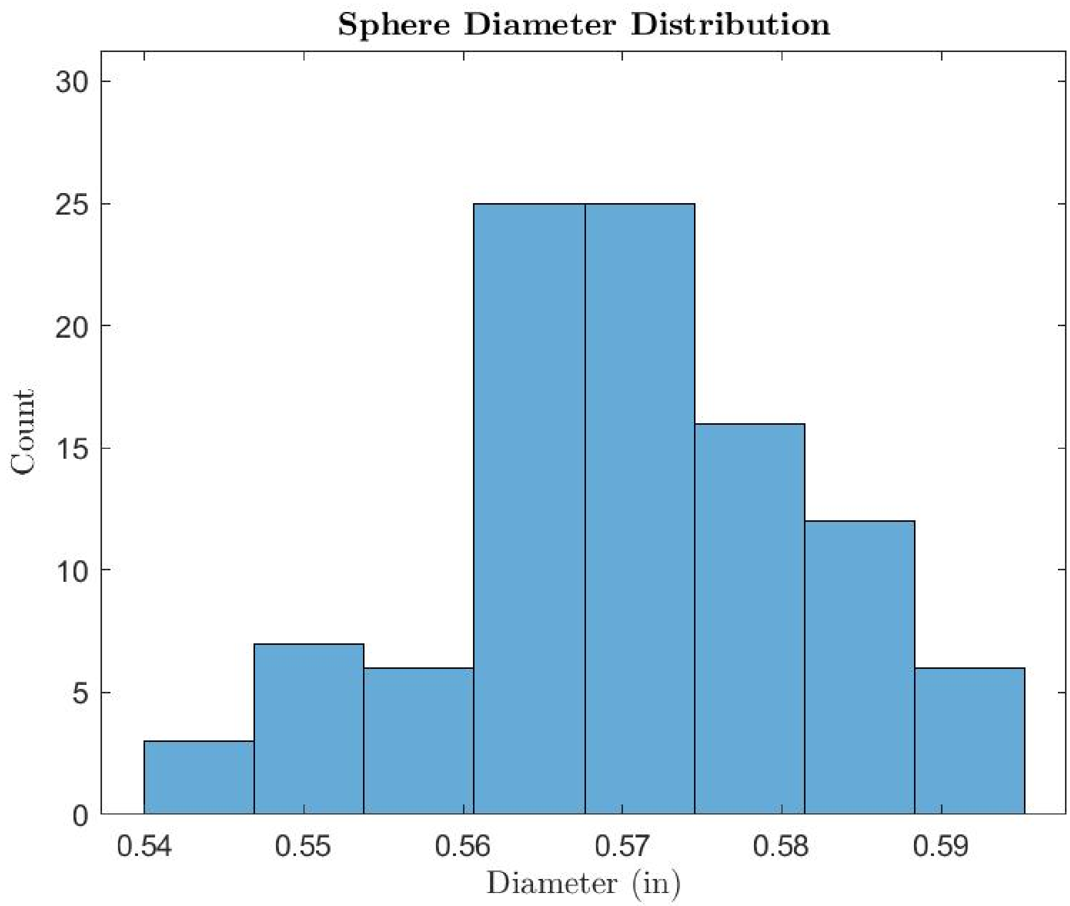


Figure 1: Sphere Diameter Distribution

Using Equations (2) and (4), the mean and standard deviation are found to be:

$$\mu = 0.5700 \text{ in}$$
$$\sigma = 0.0117 \text{ in}$$

By counting the most common value, the mode was determined to be 0.5710 in.

Knowing the standard deviation, the statistical limit of error can be calculated. Inputting the standard deviation and number of elements into Equation (6) returns the statistical limit of error for the sphere diameter measurements:

$$S.L.E = \frac{3*0.0117}{\sqrt{100}} = 0.0035 \text{ in}$$

The I.L.E for English Vernier calipers and the S.L.E of the measurements can be used to find the overall limit of error using Equation (7), which comes out to:

$$L.E = \sqrt{0.001^2 + 0.0035^2} = 0.0036 \text{ in}$$

The overall limit of error represents the range in which the true value is most likely to lie. Combining the mean and L.E found above, the diameter of the glass spheres can be estimated to:

$$d = 0.570 \pm 0.004 \text{ in}$$

Diameter Measurement Results		
Quantity	**Symbol**	**Value**
Average Diameter	μ	0.5700 in
Standard Deviation	σ	0.0117 in
Statistical Limit of Error	$S.L.E$	0.0035 in
Overall Limit of Error	$L.E$	0.0036 in
Diameter of Sphere	d	0.570 ± 0.004 in

3.2 20 Block Measurements

After the 20 blocks were measured and lengths, widths, and heights logged, the resulting distributions were as shown in Figure 2.

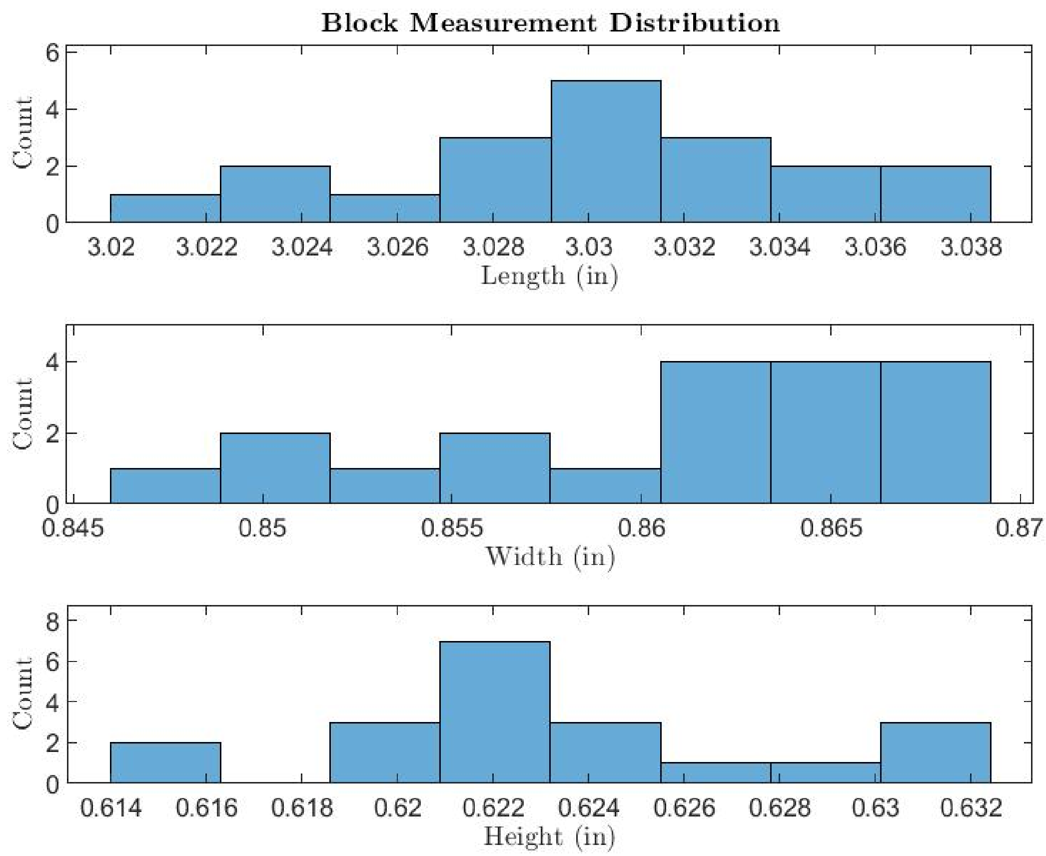


Figure 2: Block Length, Width, and Height Distributions

Calculating the means and standard deviations in the same manner as in 3.1 gave the following results:

$$\bar{x}_l = 3.0300 \text{ in} \quad \bar{x}_w = 0.8604 \text{ in} \quad \bar{x}_h = 0.6233 \text{ in}$$
$$S_l = 0.0048 \text{ in} \quad S_w = 0.0069 \text{ in} \quad S_h = 0.0049 \text{ in}$$

where $\bar{x}$ is the mean, S is the standard deviation, and the subscripts l, w, and h denote length, width, and height respectively.

Using the standard deviations, the S.L.E's can be found, again using Equation (6), and come out to:

$$S.L.E_{length} = 0.0033 \text{ in}$$
$$S.L.E_{width} = 0.0047 \text{ in}$$
$$S.L.E_{height} = 0.0033 \text{ in}$$

Then, using a I.L.E of 0.001 in, the overall limits of error come out to:

$$L.E_{length} = 0.0035 \text{ in}$$
$$L.E_{width} = 0.0048 \text{ in}$$
$$L.E_{height} = 0.0035 \text{ in}$$

which gives the results for the length, width, and height as:

$$l = 3.030 \pm 0.004 \text{ in}$$
$$w = 0.860 \pm 0.005 \text{ in}$$
$$h = 0.623 \pm 0.004 \text{ in}$$

To determine the best estimate of the volume of the blocks, the best estimates for each the length, width, and height of the block are used and produces a volume of:

$$V_{best} = l_{best} * w_{best} * h_{best} = \bar{x}_l * \bar{x}_w * \bar{x}_h = 1.6250 \text{ in}^3$$

The uncertainty of the volume is calculated as described in 1.3 using Equation (9). The uncertainty, only taking into consideration the I.L.E (0.001 in), comes out to:

$$\delta V_{ILE} = V_{best}\sqrt{\frac{0.001}{l_{best}}^2 + \frac{0.001}{w_{best}}^2 + \frac{0.001}{h_{best}}^2} = 0.0033 \text{ in}^3$$

Considering the overall limit of error (L.E) instead of just he I.L.E changes the uncertainty of the volume to:

$$\delta V = 0.0129 \text{ in}^3$$

Overall, the best estimate of the volume of the block based on the measurements gathered is:

$$V = 1.625 \pm 0.013 \text{ in}^3$$

Block Measurement Results		
Quantity	**Symbol**	**Value**
Average Block Length	$\bar{x}_l$	3.0300 in
Average Block Width	$\bar{x}_w$	0.8604 in
Average Block Height	$\bar{x}_h$	0.6233 in
Length Standard Deviation	S_l	0.0048 in
Width Standard Deviation	S_w	0.0069 in
Height Standard Deviation	S_h	0.0049 in
Length Statistical Limit of Error	$S.L.E_{length}$	0.0033 in
Width Statistical Limit of Error	$S.L.E_{width}$	0.0047 in
Height Statistical Limit of Error	$S.L.E_{height}$	0.0033 in
Length Overall Limit of Error	$L.E_{length}$	0.0035 in
Width Overall Limit of Error	$L.E_{width}$	0.0048 in
Height Overall Limit of Error	$L.E_{height}$	0.0035 in
Length of Block	l	3.030 ± 0.004 in
Width of Block	w	0.860 ± 0.005 in
Height of Block	h	0.623 ± 0.004 in
Volume of Block	V	1.625 ± 0.013 in^3

4 Conclusion

In conclusion, the measurements of the glass spheres gave a best estimate of the diameter to be 0.570 ± 0.004 in. This result was calculated from the mean and standard deviation of the measurements' distribution. In the same manner, the best estimate of the length, width, and height of the wooden block were determined to be 3.030 ± 0.004 in, 0.860 ± 0.005 in, and 0.623 ± 0.004 in, respectively. Using error propagation, the best estimate of the volume of the wooden block was calculated as 1.625 ± 0.013 in^3.

References

[1] D. D. Campbell, "Lecture 1: Measurements and uncertainties," 2021.

[2] ——, "Lecture 2: Error propagation and statistical analysis," 2021.

[3] L. Lin and P. Sherman, "Cleaning data the chauvenet way," 2009.

Appendix A - Raw Data

Table A-1: Sphere Measurement Raw Data

Diameter Measurements for 100 Glass Spheres					
Row	**Diameter (in)**	**Diameter (in)**	**Diameter (in)**	**Diameter (in)**	**Diameter (in)**
1	0.550	0.581	0.545	0.575	0.575
2	0.571	0.561	0.559	0.542	0.575
3	0.580	0.547	0.590	0.575	0.563
4	0.563	0.562	0.571	0.595	0.547
5	0.580	0.562	0.568	0.552	0.586
6	0.569	0.590	0.566	0.571	0.581
7	0.561	0.595	0.571	0.573	0.580
8	0.584	0.586	0.583	0.574	0.569
9	0.565	0.569	0.571	0.587	0.562
10	0.583	0.580	0.585	0.568	0.558
11	0.570	0.568	0.569	0.564	0.573
12	0.577	0.550	0.576	0.592	0.574
13	0.582	0.571	0.567	0.589	0.573
14	0.565	0.565	0.573	0.587	0.564
15	0.551	0.557	0.566	0.544	0.561
16	0.570	0.566	0.585	0.562	0.572
17	0.550	0.575	0.574	0.568	0.567
18	0.567	0.584	0.554	0.577	0.571
19	0.560	0.564	0.579	0.557	0.563
20	0.563	0.582	0.562	0.575	0.567

Table A-2: Block Measurement Raw Data

Length, Width, and Height Measurements for 20 Blocks			
Block	**Length (in)**	**Width(in)**	**Height (in)**
1	3.015	0.867	0.624
2	3.031	0.863	0.631
3	3.029	0.857	0.621
4	3.035	0.849	0.614
5	3.025	0.855	0.627
6	3.037	0.860	0.619
7	3.020	0.867	0.623
8	3.031	0.864	0.616
9	3.032	0.865	0.619
10	3.038	0.864	0.62
11	3.023	0.849	0.623
12	3.031	0.862	0.631
13	3.031	0.869	0.630
14	3.023	0.869	0.621
15	3.033	0.862	0.624
16	3.034	0.853	0.625
17	3.027	0.847	0.623
18	3.027	0.837	0.622
19	3.032	0.862	0.621
20	3.031	0.864	0.632

Appendix B - Matlab Code

The following pages in this appendix contain the Matlab code used for all of the calculations in this report. The code beings on the next page.

Table of Contents

Measurement and Error Analysis MATLAB

Jonathan Tyler Boylan 9/11/2021 EML3012C Measurements Lab

```
clc
clear
format compact
```

Load Data

```
data = readtable('LabData.csv'); % import data csv as table

sphere_data = data.Diameters';

block_length_data = rmmissing(data.Lengths');
block_width_data = rmmissing(data.Widths');
block_height_data = rmmissing(data.Heights');
```

Calculate Means, Modes, and Standard Deviations

```
sphere_avg = mean(sphere_data);
sphere_std = std(sphere_data,1); % Normalize by N
sphere_mode = mode(sphere_data);

length_avg = mean(block_length_data);
length_std = std(block_length_data);
length_mode = mode(block_length_data);

width_avg = mean(block_width_data);
width_std = std(block_width_data);
width_mode = mode(block_width_data);
```

```
height_avg = mean(block_height_data);
height_std = std(block_height_data);
height_mode = mode(block_height_data);
```

Normalizations

```
sphere_z_data = (sphere_data - sphere_avg)./sphere_std;

length_z_data = (block_length_data - length_avg)./length_std;
width_z_data = (block_width_data - width_avg)./width_std;
height_z_data = (block_height_data - height_avg)./height_std;
```

Data Rejection

```
% Chauvenet's criterion for n = 100 and n = 20
chauvenet_100 = 2.807;
chauvenet_20 = 2.241;

sphere_data = sphere_data(abs(sphere_z_data) < chauvenet_100); % 0
 removed

block_length_data = block_length_data(abs(length_z_data) <
 chauvenet_20); % 1 removed
block_width_data = block_width_data(abs(width_z_data) <
 chauvenet_20); % 1 removed
block_height_data = block_height_data(abs(height_z_data) <
 chauvenet_20); % 0 removed
```

Recalculate Means and Standard Deviations

```
sphere_avg = mean(sphere_data);
sphere_std = std(sphere_data,1); % Normalize by N

length_avg = mean(block_length_data);
length_std = std(block_length_data);

width_avg = mean(block_width_data);
width_std = std(block_width_data);

height_avg = mean(block_height_data);
height_std = std(block_height_data);
```

Limits of Error

```
ILE = 0.001;
SLE = @(std,n) 3*std/sqrt(n);

sphere_SLE = SLE(sphere_std,length(sphere_data));
sphere_LE = norm([ILE sphere_SLE]);
```

```
length_SLE = SLE(length_std,length(block_length_data));
length_LE = norm([ILE length_SLE]);

width_SLE = SLE(width_std,length(block_width_data));
width_LE = norm([ILE width_SLE]);

height_SLE = SLE(height_std,length(block_height_data));
height_LE = norm([ILE height_SLE]);
```

Block Volume

```
volume_best = length_avg * width_avg * height_avg;
volume_uncert = volume_best * norm([length_LE/length_avg width_LE/
width_avg height_LE/height_avg]);
```

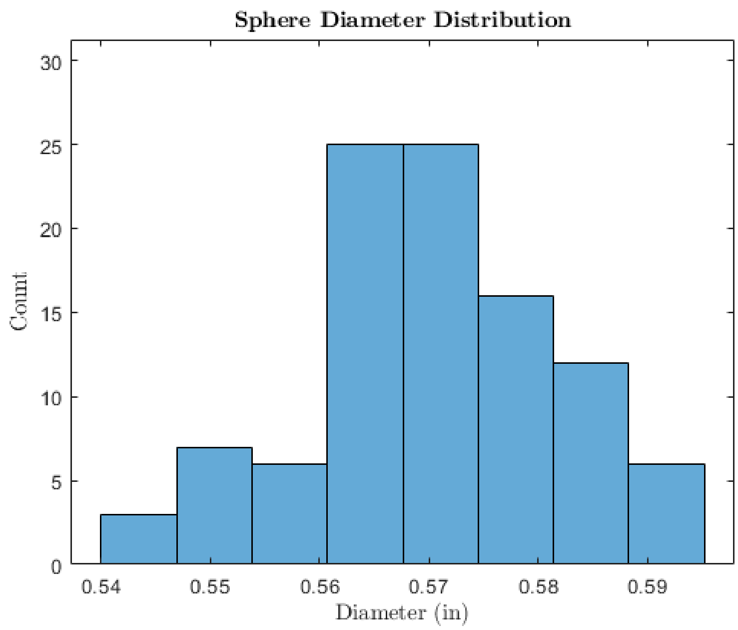
Gaussian Normal Distribution Function

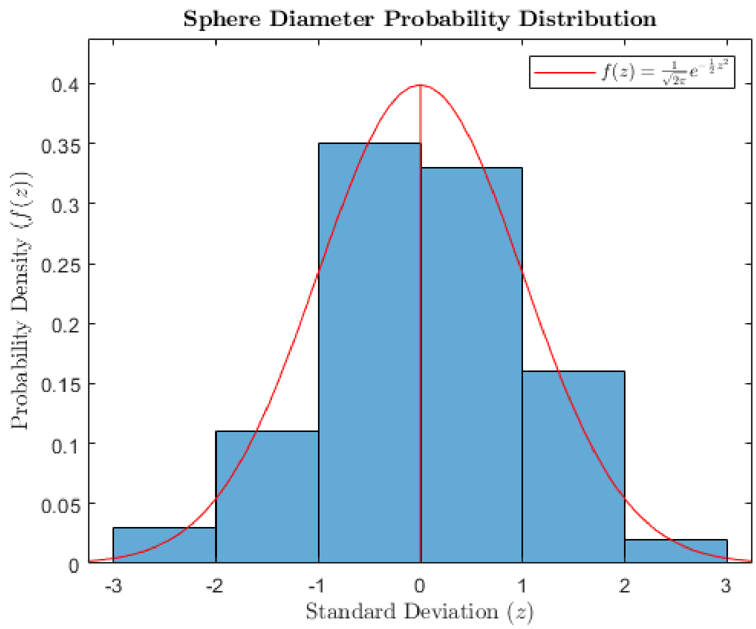
```
gaussian = @(Z) (1/(sqrt(2*pi))) .* exp(-Z.^2/2);
```

Sphere Histograms

```
% Figure 1: Diameter measurement distribution
fig1 = figure(1);
n_bins = 8; %  Number of histogram bins

% Create distribution histogram
h1 = histogram(sphere_data,n_bins);
ylim([0 max(h1.Values)*1.25]) % Add margin to y-axis

% Using LaTeX for font and text formatting
title('\textbf{Sphere Diameter Distribution}','Interpreter','latex')
xlabel('Diameter (in)','Interpreter','latex')
ylabel('Count','Interpreter','latex')

% Save first figure to jpg
saveas(gcf, 'SphereHistogram1.jpg');

% Figure 2: Diameter probability distribution
fig2 = figure(2);

% Create probability density histogram
h2 = histogram(sphere_z_data,-3:3,'Normalization','pdf');
ylim([0 max(h2.Values)*1.25]) % Add margin to y-axis
xlim([-3.25 3.25]) % add margin to x-axis

% Add Gaussian Normal Distribution Function to plot
Z = -5:0.1:5; % graph 5 st. devs

hold on
p2 = plot(Z,gaussian(Z),'r');
plot([0 0],[0 gaussian(0)], 'r'); % Plot mean
```

```
legend(p2,'$f(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}
{2}z^2}$','Interpreter','latex'); % Add legend

title('\textbf{Sphere Diameter Probability
 Distribution}', 'Interpreter','latex')
xlabel('Standard Deviation ($z$)','Interpreter','latex')
ylabel('Probability Density ($f(z)$)','Interpreter','latex')

% Save second figure to jpg
saveas(gcf,'SphereHistogram2.jpg');
```



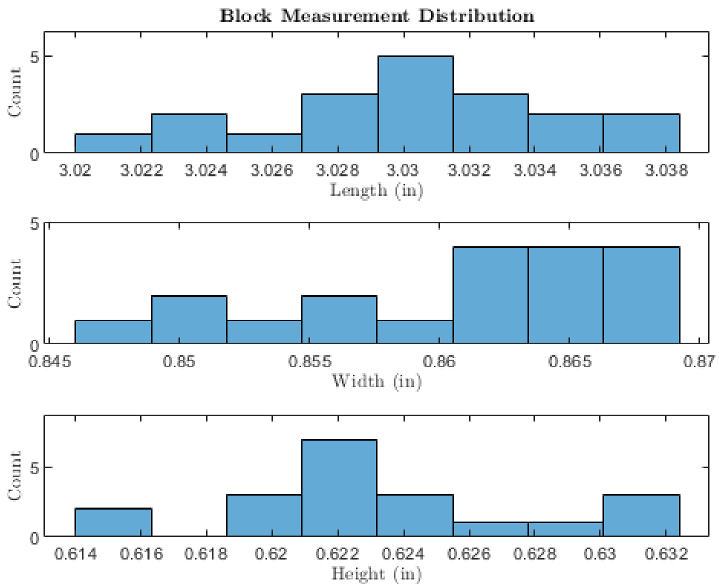


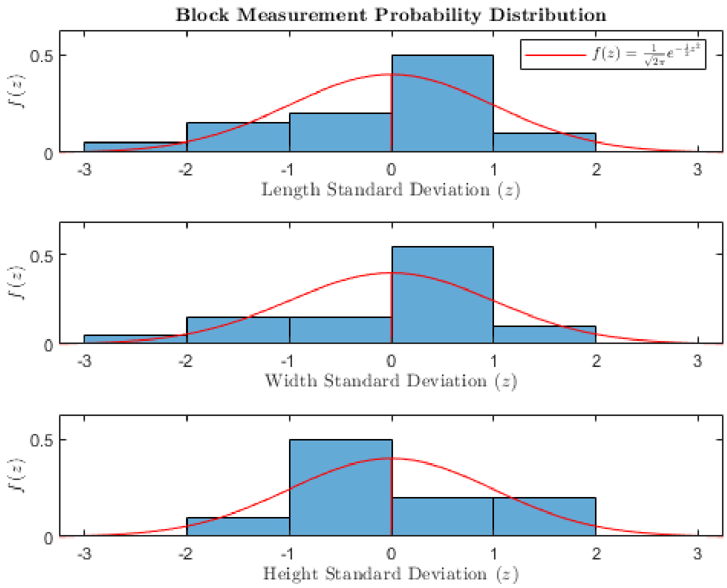
Block Histograms

```
% Figure 3: Block measurement distributions
fig3 = figure(3);

n_bins = 8; % Number of bins

% 3 Histograms: Length, Width, Height
% Length
subplot(3,1,1)
h1_length = histogram(block_length_data,n_bins);
ylim([0 max(h1_length.Values)*1.25]) % Add margin to y-axis

title('\textbf{Block Measurement Distribution}','Interpreter','latex')
ylabel('Count', 'Interpreter','latex')
xlabel('Length (in)','Interpreter','latex')

% Width
subplot(3,1,2)
h1_width = histogram(block_width_data,n_bins);
ylim([0 max(h1_width.Values)*1.25])

ylabel('Count', 'Interpreter','latex')
xlabel('Width (in)','Interpreter','latex')
```

```
% Height
subplot(3,1,3)
h1_height = histogram(block_height_data,n_bins);
ylim([0 max(h1_height.Values)*1.25])

ylabel('Count', 'Interpreter','latex')
xlabel('Height (in)','Interpreter','latex')

% Save figure to jpg
saveas(gcf,'BlockHistogram1.jpg')


% Figure 4: Block probability density
fig4 = figure(4);

% Length
subplot(3,1,1)
h2_length = histogram(length_z_data,-3:3,'Normalization','pdf');
ylim([0 max(h2_length.Values)*1.25]) % Add margin to y-axis
xlim([-3.25 3.25]) % add margin to x-axis

% Plot Gaussian
hold on
p4 = plot(Z,gaussian(Z),'r');
plot([0 0],[0 gaussian(0)], 'r'); % Plot mean
legend(p4,'$f(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}
{2}z^2}$','Interpreter','latex'); % Add legend

title('\textbf{Block Measurement Probability
 Distribution}', 'Interpreter','latex')
xlabel('Length Standard Deviation ($z$)','Interpreter','latex')
ylabel('$f(z)$','Interpreter','latex')

% Width
subplot(3,1,2)
h2_width = histogram(width_z_data,-3:3,'Normalization','pdf');
ylim([0 max(h2_width.Values)*1.25]) % Add margin to y-axis
xlim([-3.25 3.25]) % add margin to x-axis

% Plot Gaussian
hold on
p4 = plot(Z,gaussian(Z),'r');
plot([0 0],[0 gaussian(0)], 'r'); % Plot mean

xlabel('Width Standard Deviation ($z$)','Interpreter','latex')
ylabel('$f(z)$','Interpreter','latex')

% Height
subplot(3,1,3)
h2_height = histogram(height_z_data,-3:3,'Normalization','pdf');
ylim([0 max(h2_height.Values)*1.25]) % Add margin to y-axis
xlim([-3.25 3.25]) % add margin to x-axis

% Plot Gaussian
```

```
hold on
p4 = plot(Z,gaussian(Z),'r');
plot([0 0],[0 gaussian(0)], 'r'); % Plot mean

xlabel('Height Standard Deviation ($z$)','Interpreter','latex')
ylabel('$f(z)$','Interpreter','latex')

% Save second figure to jpg
saveas(gcf,'BlockHistogram2.jpg');
```





Published with MATLAB® R2020a

Appendix C - Additional Plots

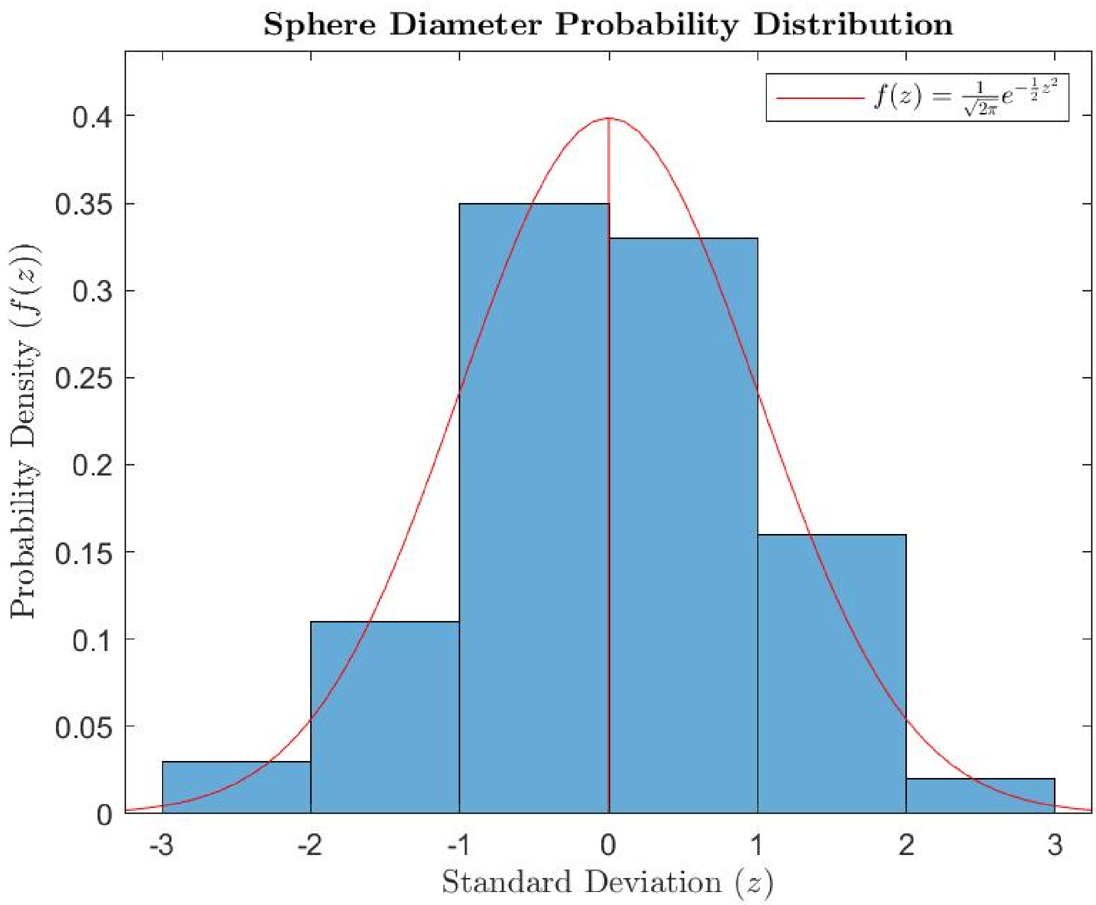


Figure C-1: Sphere Diameter Probability Distribution

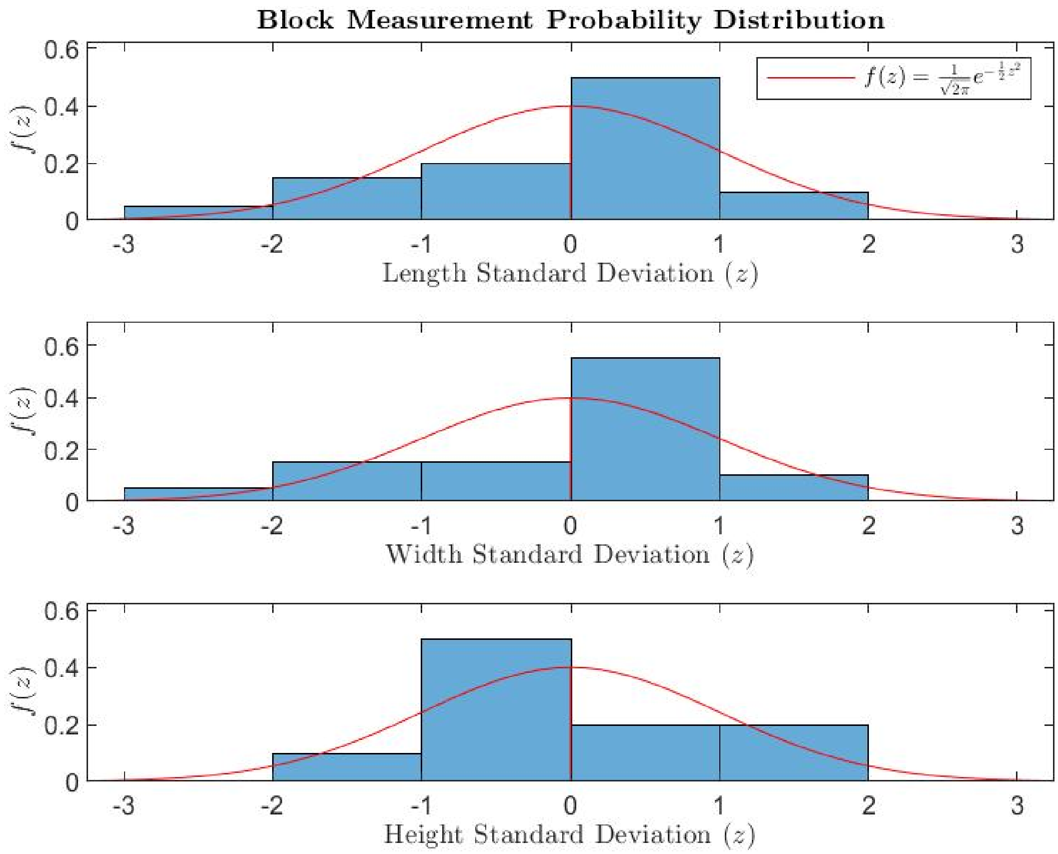


Figure C-2: Block Length, Width, and Height Probability Distribution