MIAMI UNIVERSITY

OXFORD, OH • EST. 1809

Central Limit Theorem and Different Estimators

College of Engineering and Computing

ECE 345: Engineering Statistics Final Project

Professor: Dr. Zhou

December 1, 2022

Ethan Barnes barnese6@miamioh.edu
Deniz Misirlioglu misirld@miamioh.edu
Jared Rozell rozellj@miamioh.edu
Riley Taylor taylo550@miamioh.edu

A) Approximate Theoretical PDF Using Arithmetic Mean

First we solved for the random variable X in the equation $Y = X^2$. This allows us to find the pdf of x with respect to y. Solving the equation gives us:

$$Y = X^2$$

$$X = \pm\sqrt{Y}$$

Then we solved for the expectation of Y:

$$E(A_n) = E(Y) = \int_0^\infty y f_Y(y) dy$$

$$E(Y) = \int_0^\infty y \frac{1}{\sigma\sqrt{2\pi y}} e^{-\frac{1}{2}(\frac{y}{\sigma^2})} dy = \sigma_x^2$$

$$V(Y) = \int_0^\infty (y - \sigma^2)^2 \frac{1}{\sigma\sqrt{2\pi y}} e^{-\frac{1}{2}(\frac{y}{\sigma^2})} dy = 2\sigma_x^4$$

The pdf of x with respect to y will result in the following:

$$f_x(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \mu = 0$$

$$f_x(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, x^2 = y$$

The complete derived equation:

$$f_x(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y}{2\sigma^2}}, y > 0$$

According to the central limit theorem the approximate theoretical pdf of An is shown as follows:

$$f(A_n) = \frac{1}{\sqrt{2\pi}\sigma_{A_n}} e^{-\frac{(A_n - \mu_{A_n})^2}{2\sigma_{A_n}^2}}, x^2 = y$$

Using the following values of n = 2, 3, 16, 144:

$$f(A_2) = N(\sigma_X^2, \sigma_X^4), y > 0$$

$$f(A_3) = N(\sigma_X^2, \frac{2\sigma_X^4}{3}), y > 0$$

$$f(A_{16}) = N(\sigma_X^2, \frac{\sigma_X^4}{8}), y > 0$$

$$f(A_{144}) = N(\sigma_X^2, \frac{\sigma_X^4}{72}), y > 0$$

B) True PDF of A_n for $n = 2$ Without Approximation

$$f(A_n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{((\frac{1}{2}(y_1+y_2))-\sigma^2)^2}{2\sigma^4}}, y > 0$$

Calculating the bounds so we can solve for the pdf:

$$y > 0$$

$$z - y > 0$$

$y < z$ and $y > 0$, resulting in $0 > y > z$

$A_2 = \frac{1}{2}(y_1 + y_2)$, transforming into $Z = X + Y$

$$f_z(z) = \int \int_{x+y=z} f_{xy}(x, y) dx dy$$

Since x and y are independent:

$$f_z(z) = \int \int_{x+y=z} f_x(x) f_y(y) dx dy$$

Resulting in a convolution integral of:

$$f_z(z) = \int f_y(y) f_y(z - y) dy$$

Implementing the bounds:

$$\int_0^z f_y(y) f_y(z - y) dy$$

$$f_y(y) = \frac{1}{\sqrt{2\pi y}\sigma} e^{-\frac{y}{2\sigma^2}}, y > 0$$

$$f(A_2) = f(z) = \int_0^z \frac{1}{\sqrt{2\pi y}\sigma} e^{-\frac{y}{2\sigma^2}} \frac{1}{\sqrt{2\pi(z-y)}\sigma} e^{-\frac{(z-y)}{2\sigma^2}} dy$$

$$= \int_0^z \frac{1}{\sqrt{2\pi y}\sigma} \frac{1}{\sqrt{2\pi(z-y)}\sigma} e^{-\frac{z}{2\sigma^2}} dy$$

$$= e^{-\frac{z}{2\sigma^2}} \int_0^z \frac{1}{\sqrt{2\pi y}\sigma} \frac{1}{\sqrt{2\pi(z-y)}\sigma} dy$$

$$\boxed{= \frac{e^{-\frac{z}{2\sigma^2}}}{2\sigma^2}}$$

C) Matlab Simulation for the PDF of A_n

Standard Deviation (n=2) = 0.9328

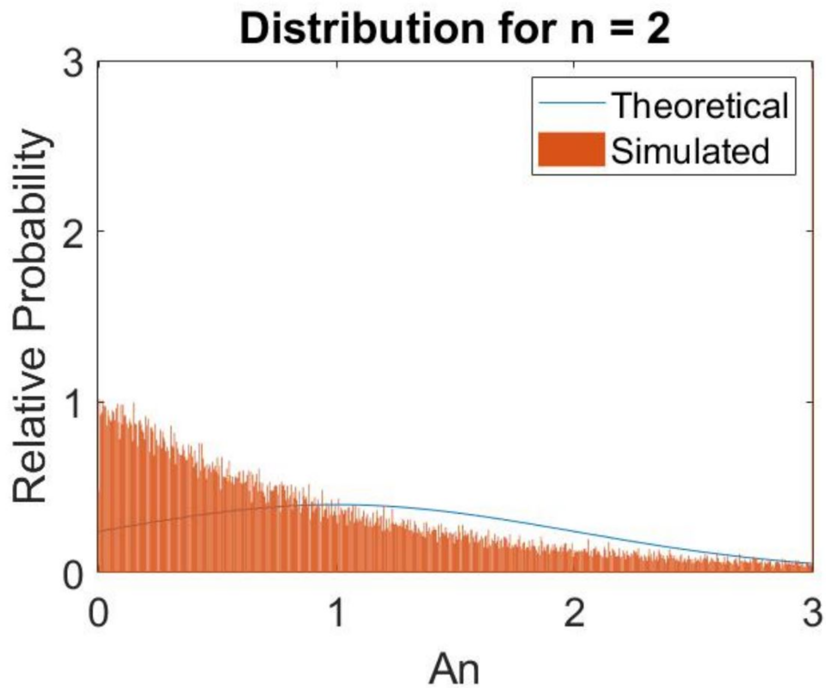

Figure 1: Simulated and Theoretical graph for $n = 2$

Standard Deviation (n=3) = 0.7645

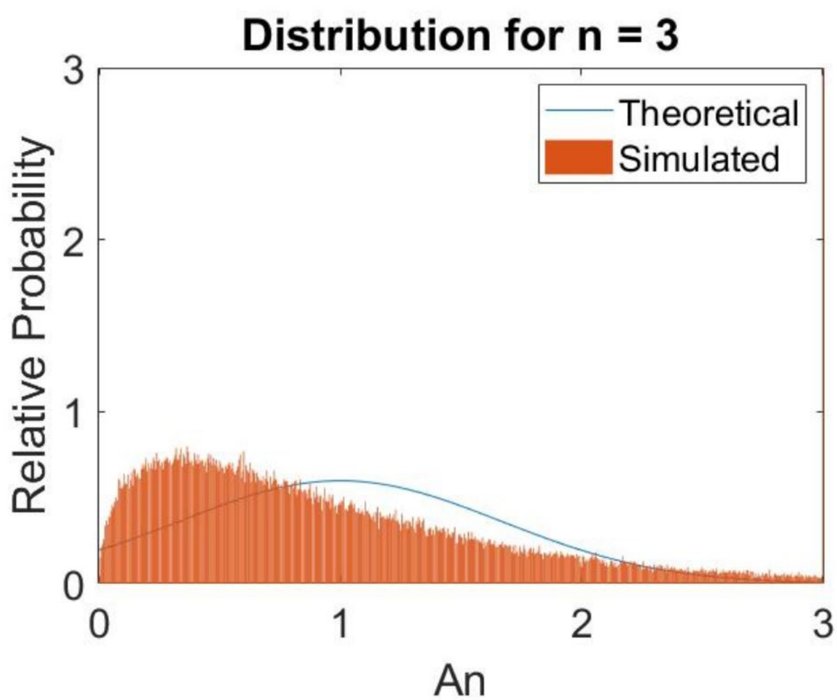

Figure 2: Simulated and Theoretical graph for $n = 3$

Standard Deviation (n=16) = 0.3295

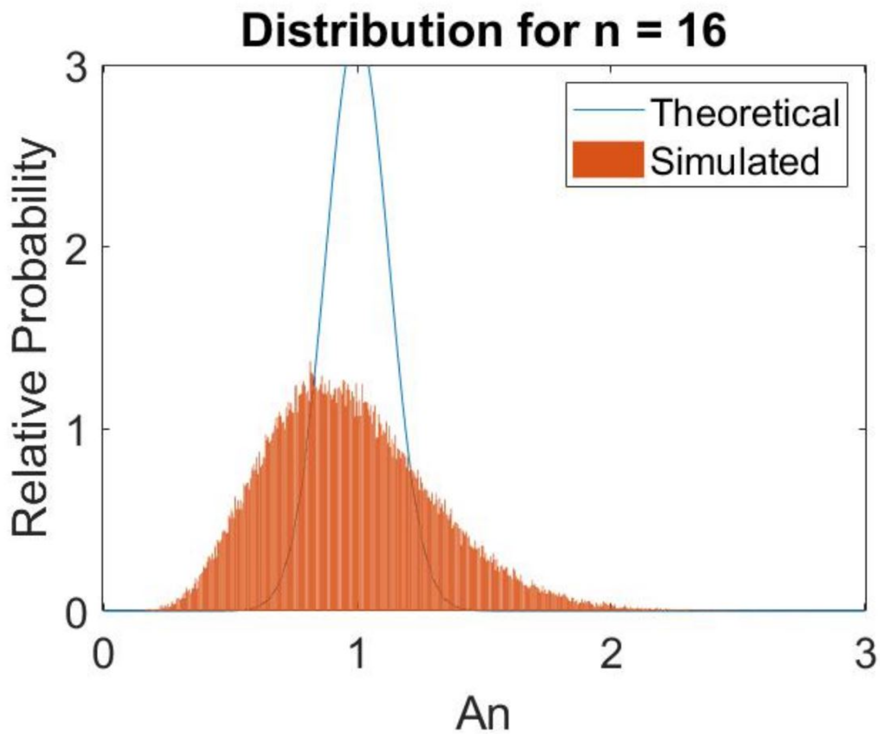

Figure 3: Simulated and Theoretical graph for $n = 16$

Standard Deviation (n=144) = 0.1100

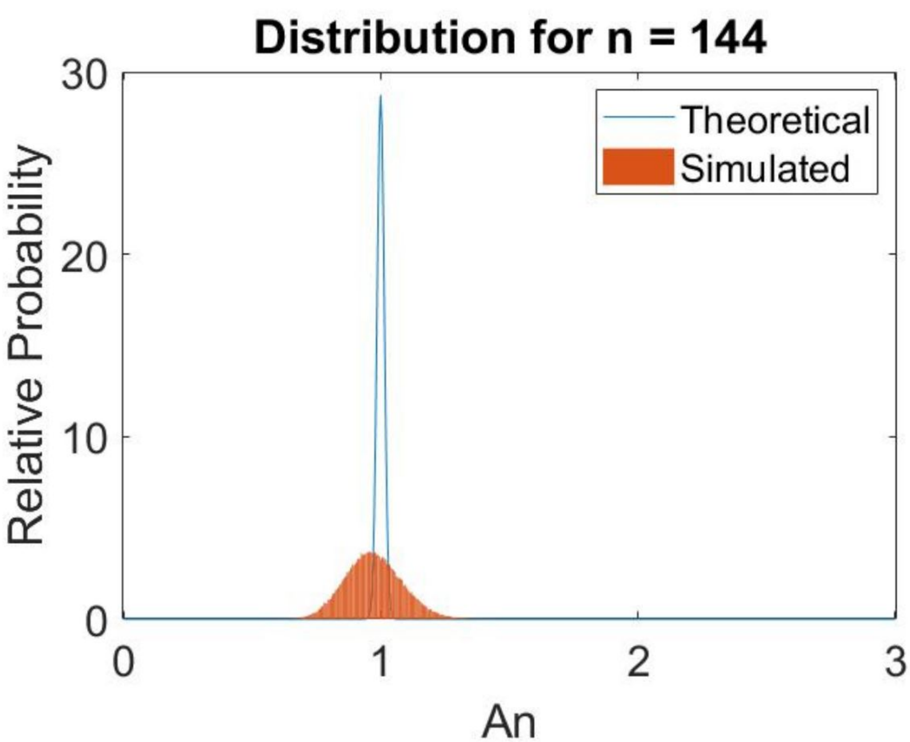

Figure 4: Simulated and Theoretical graph for $n = 144$

D) Comparing the Three PDFs

The theoretical distribution is normal and its mean is slightly larger than 1. The other two distributions appear to be a decaying exponential. The simulated distribution follows the shape of the true PDF, but is more biased towards 0. CLT is used to find the theoretical PDF, this approximation is most likely the reason behind the difference in shape.

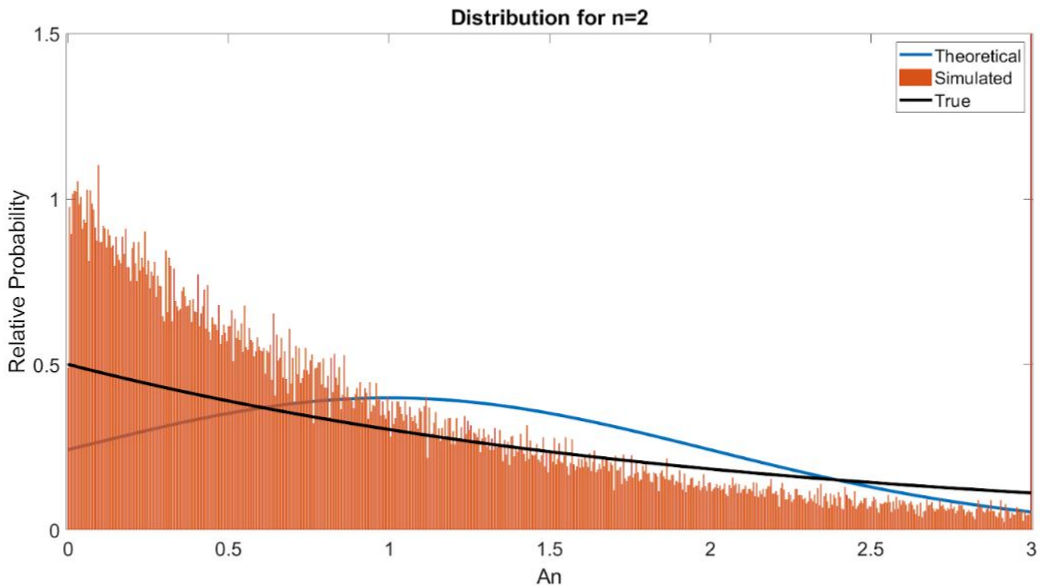

Figure 5: Comparing three PDFs using $n = 2$

The distribution for both theoretical and simulated curves gets narrower as n increases. The main difference between the theoretical and simulated curves is the deviation. A higher deviation can be seen in the simulated PMF.

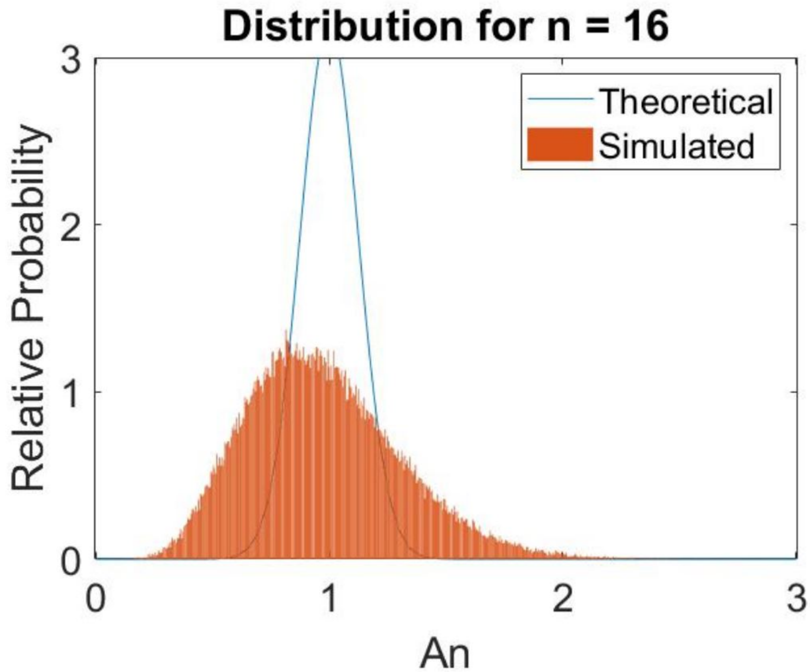

Figure 6: Simulated and Theoretical graph for $n = 16$

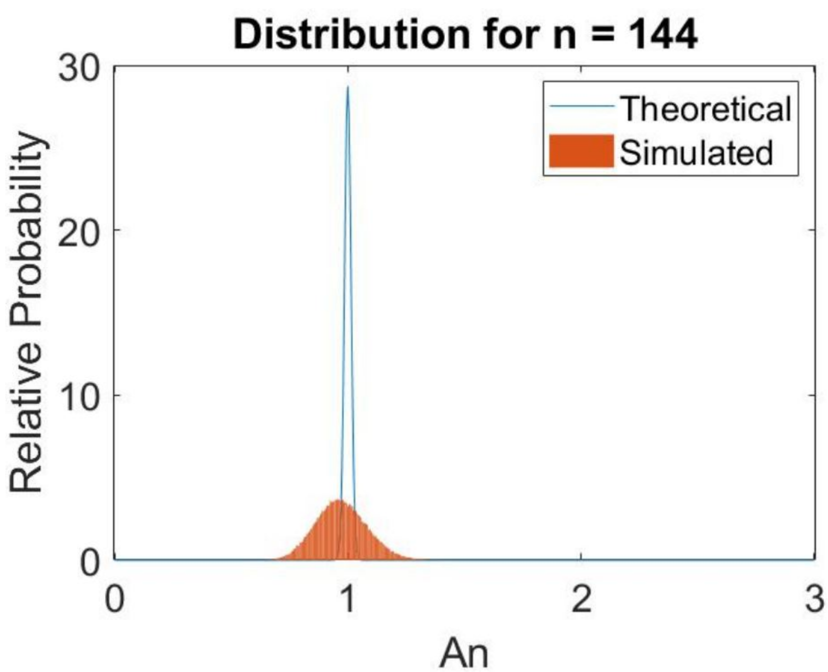

Figure 7: Simulated and Theoretical graph for $n = 144$

E) Approximate Theoretical PDF Using Geometric Mean

To approximate the theoretical PDF, we set G_n equivalent to the sigma product of $(x_i^2)^{\frac{1}{n}}$. Natural logging both sides, we then can transform the sigma product to a summation of x^2, and move the $(1/n)$ to the front. Substituting $\ln(G_n)$ as q, and $\ln(x^2)$ as b, we can simplify the equation and compute the definite infinite integral of $\int_{-\infty}^{\infty} G_n f_{G_n}(x_i) dx_i$.

$$G_n = (\prod_{i=1}^{n} x_i^2)^{\frac{1}{n}}$$

$$\ln G_n = \frac{1}{n} \sum_{i=1}^{n} (\ln x_i^2)$$

$$\ln(G_n) = q, \ln x^2 = b$$

$$q = \frac{1}{n} \sum_{i=1}^{n} b_i = E(b)$$

$$E(G_n) = \int_{-\infty}^{\infty} G_n f_{G_n}(x_i) dx_i$$

$$[\int_0^{\infty} y^{\frac{1}{n}} \frac{1}{\sqrt{2\pi y}\sigma} e^{\frac{-y}{2\sigma^2}} dy]^n$$

Using substitutions to mimic the integral table $x_1 = \frac{1}{n} - \frac{1}{2}$, $A = \frac{1}{2\sigma^2}$, $B = \frac{1}{\sqrt{2\pi}\sigma}$

$$[B \int_0^{\infty} y^{x_1} y^{\frac{-1}{2}} e^{-Ay} dy]^n$$

Simplifying further:

$$[B \int_0^{\infty} y^{x_1} e^{-Ay} dy]^n$$

We referred to the integral tables to find a similar integral to ours and found a Gamma result.

$$\int_0^{\infty} x^n e^{-ax}\, dx = \begin{cases} \dfrac{\Gamma(n+1)}{a^{n+1}} & (n > -1,\ \mathrm{Re}(a) > 0) \\ \\ \dfrac{n!}{a^{n+1}} & (n = 0, 1, 2, \ldots,\ \mathrm{Re}(a) > 0) \end{cases}$$

$$\overline{G_n} = [\frac{\frac{1}{\sqrt{2\pi}\sigma}\Gamma(\frac{1}{n} - \frac{1}{2} + 1)}{(\frac{1}{2\sigma^2})^{\frac{1}{n} - \frac{1}{2} + 1}}]^n$$

$$= [\frac{\frac{1}{\sqrt{2\pi}\sigma}\Gamma(\frac{1}{n} + \frac{1}{2})}{\frac{1}{(2\sigma^2)^{\frac{1}{n} + \frac{1}{2}}}}]^n$$

To remove Gamma from our equation we used the substitution $\Gamma(N) = (N-1)!$ to get a simplified equation.

$$[(2\sigma^2)^{1+\frac{n}{2}}((\frac{1}{n} - \frac{1}{2})!)(\frac{1}{\sqrt{2\pi}\sigma})]^n$$

We used Desmos to carry out the calculations. As you can see from the table below, our means match up with our Figure 8.

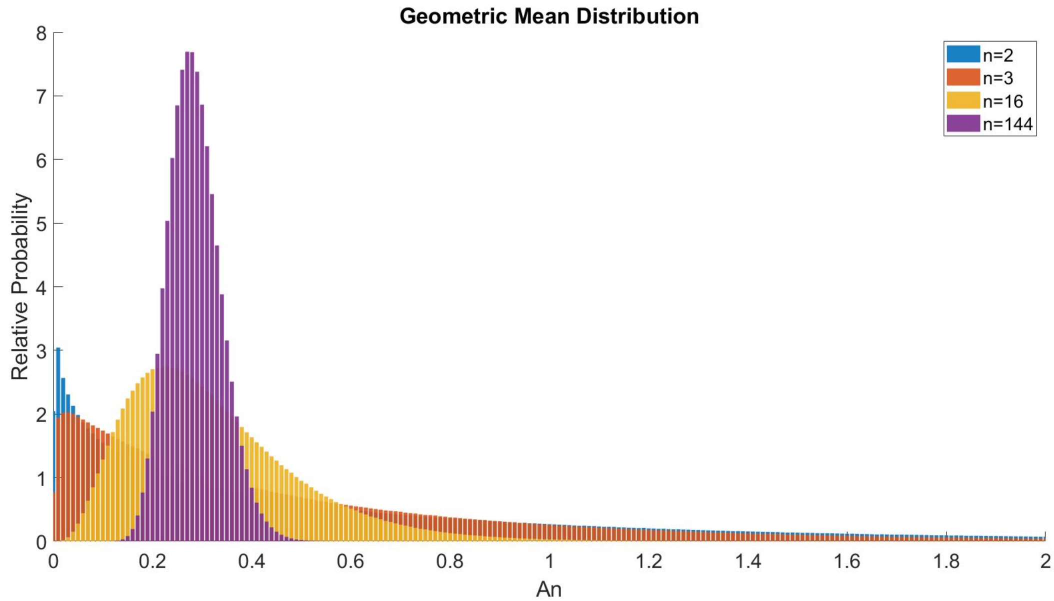

Figure 8: Simulated and Theoretical Graph for Geometric Mean

Standard Deviation and Expectation for G_n				
	n = 2	**n = 3**	**n = 16**	**n = 144**
Sim Standard Deviation	0.7714	0.5538	0.1744	0.0527
Sim Expectation	0.002	0.026	0.222	0.274
Theoretical Stand Dev	0.637	0.517	0.324	0.286

Appendix

Matlab part C using X^2

```
%%% Matlab part C using X^2. %%%
clear all
sz = 0.01;
x=-3:sz:3;
y = [0:(.01/2):3];
n = 16;

% theoretical pdf of standard norm.
y1=normpdf(x, 0, 1);

% theoretical pdf of Y
fA = normpdf(y, 1, 2/n);
plot(y,fA,'LineWidth', 3);
hold on

%fy = 1./(sqrt(2*pi*y)).*exp(-y./2);

samp=100000;
for samples = 1:samp
    for index = 1:n

        y1 = y1/sum(y1); % Make sure probabilites add up to 1.
        cp = [0, cumsum(y1)];
        r = rand;
        ind = find(r>cp, 1, 'last');
        randx = x(ind);
        %https://www.mathworks.com/matlabcentral/answers/506654-select-random-
        %number-from-an-array-with-probabilities
        xsamp(index)=randx;
    end

    An(samples)=(1/n)*sum(xsamp.^2);

end
standardDev = std(An);
% simulated pdf of Y
[y3]=hist(An,y)/(samp*(sz/2));
bar(y,y3)

z=0:.01:3;
A2B = exp(-z/2)/2;
if n==2
plot(z,A2B,'K','LineWidth', 3)
end
title("Distribution for n=" + n)
ylabel('Relative Probability')
xlabel('An')
legend('Theoretical','Simulated','True')
set(gca,'FontSize',18)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Matlab part E using X^2

```
clear all
sz = 0.01;
x=-30:sz:30;
y = [0:sz:30];
n = [2 3 16 144];
binfac = 1;
yg = [-30:sz*binfac:30];
samp=1000000;

% theoretical pdf of standard norm.
y1=normpdf(x, 0, 1);

% theoretical pdf of Y
%fA = normpdf(y, 1, 2/n);
%plot(y,fA,'LineWidth', 3);
hold on
for i = 1:4
for samples = 1:samp
    ysamp=randn([1 n(i)]).^2;
    An(samples)=(prod(ysamp).^(1/n(i)));

end
standardDev = std(An);
% simulated pdf of An
[y3]=hist(An,yg)/(samp*(sz*binfac));
b1 = bar(yg,y3);
b1.FaceAlpha = 0.9;
end
z=0:.01:3;
A2B = exp(-z/2)/2;
if n==2
    plot(z,A2B,'K','LineWidth', 3)
end

%ylim([0, 1.5])
xlim([0, 2])
title("Geometric Mean Distribution")
ylabel('Relative Probability')
xlabel('An')
legend('n=2','n=3','n=16','n=144')
set(gca,'FontSize',18)
```

Responsibilities and Work			
Ethan Barnes	**Deniz Misirlioglu**	**Jared Rozell**	**Riley Taylor**
All Matlab	A and B	PDF matlab	Coding in LaTeX
PDF comparison	part E	.	Theoretical Part A
Part E	Gamma Explanation	.	Part E

Note: Ethan, Deniz, and Riley all worked together on most parts