Homework 2

Kevin Ha

September 12, 2021

Problem 1

Problem 1.1

Given $p = 0.4$ and $P(N \leq 8.25)$:

```
pbinom(q = 8.25, size = 20, prob = 0.4)
```

```
## [1] 0.5955987
```

```
pbinom(q = 8.25, size = 30, prob = 0.4)
```

```
## [1] 0.09401122
```

```
pbinom(q = 8.25, size = 50, prob = 0.4)
```

```
## [1] 0.0002305229
```

```
pbinom(q = 8.25, size = 75, prob = 0.4)
```

```
## [1] 1.826106e-08
```

```
pbinom(q = 8.25, size = 100, prob = 0.4)
```

```
## [1] 5.431127e-13
```

Problem 1.2

Laplace theorem equation: $P(N \leq x) \approx \Phi\left(\frac{x-np}{\sqrt{np(1-p)}}\right)$ `laplace_calculate` is a function that calculates $\Phi\left(\frac{x-np}{\sqrt{np(1-p)}}\right)$ and takes a parameter `num_trials` :

```
pnorm(laplace_calculate(num_trials = 20))
```

```
## [1] 0.5454243
```

```
pnorm(laplace_calculate(num_trials = 30))
```

```
## [1] 0.08112525
```

```
pnorm(laplace_calculate(num_trials = 50))
```

```
## [1] 0.0003470073
```

```
pnorm(laplace_calculate(num_trials = 75))
```

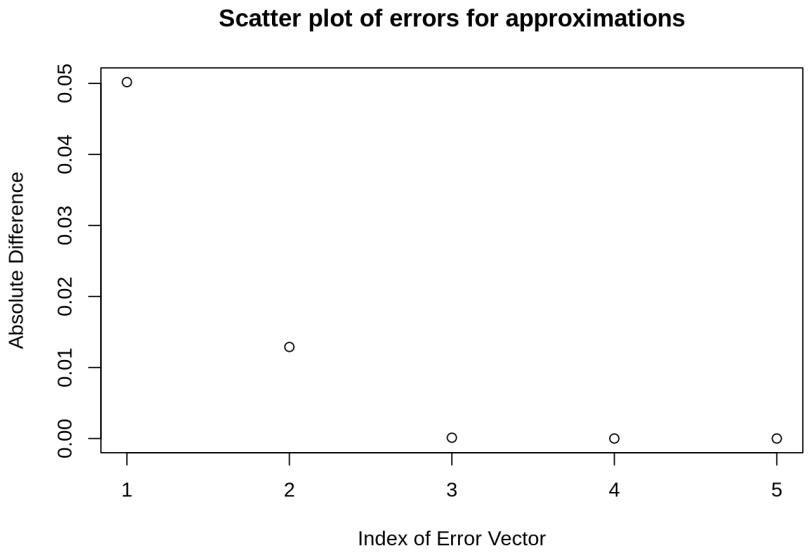
```
## [1] 1.475701e-07
```

```
pnorm(laplace_calculate(num_trials = 100))
```

```
## [1] 4.557597e-11
```

Problem 1.3

```
## [1] 5.017440e-02 1.288597e-02 1.164844e-04 1.293090e-07 4.503286e-11
```



Problem 1.4

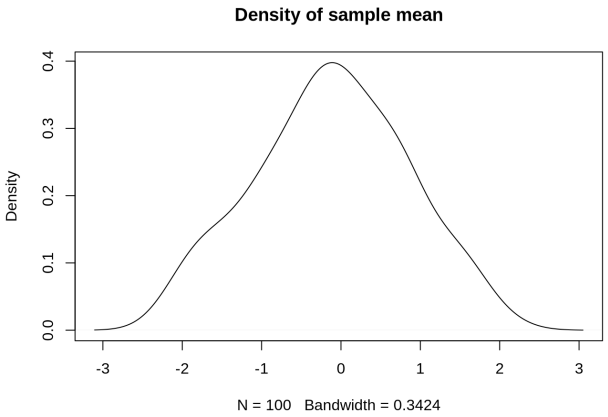
Based on the scatter plot above, it can be said that the Laplace theorem becomes more accurate to the actual probability as the number of trials increases. Thus the larger the trials, the better the normal distribution can be used to approximate the binomial distribution.

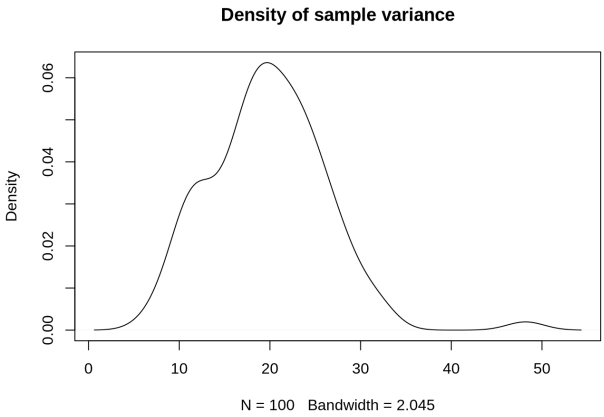
Problem 2

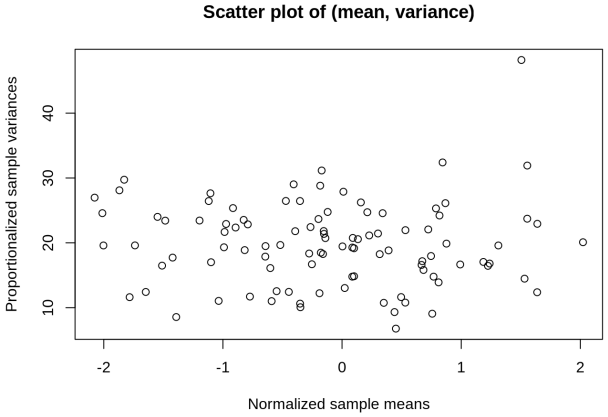
Given $X \sim N(2, 3^2)$

Problem 2.1 - 2.4

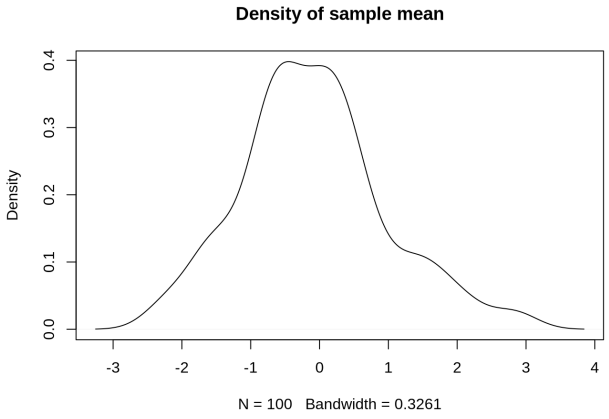
For $n = 20$

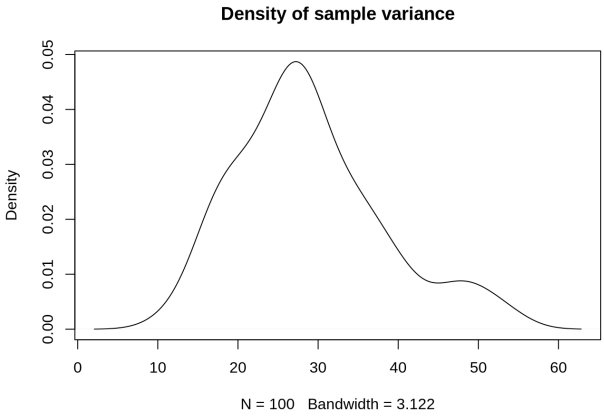


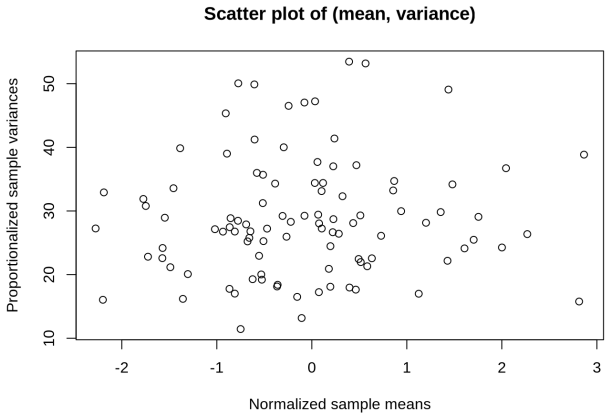




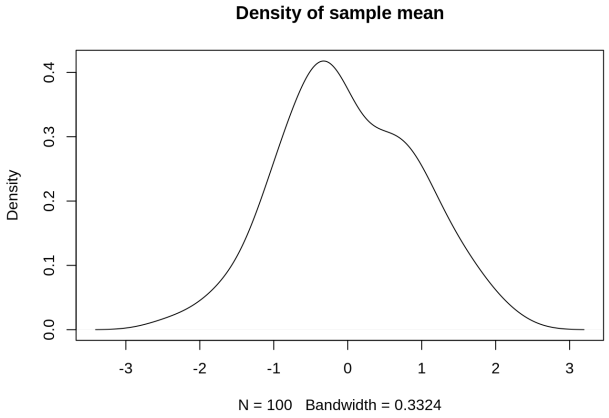
For $n = 30$

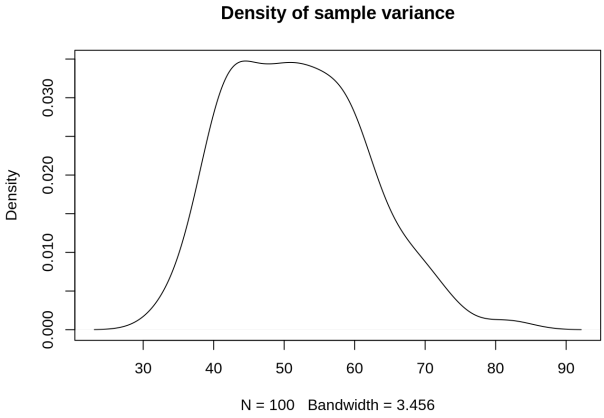

Density of sample mean
Density
N = 100 Bandwidth = 0.3261

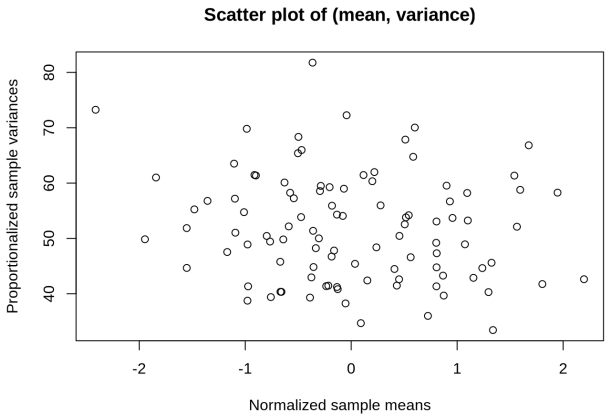

Density of sample variance
Density
N = 100 Bandwidth = 3.122


Scatter plot of (mean, variance)
Proportionalized sample variances
Normalized sample means

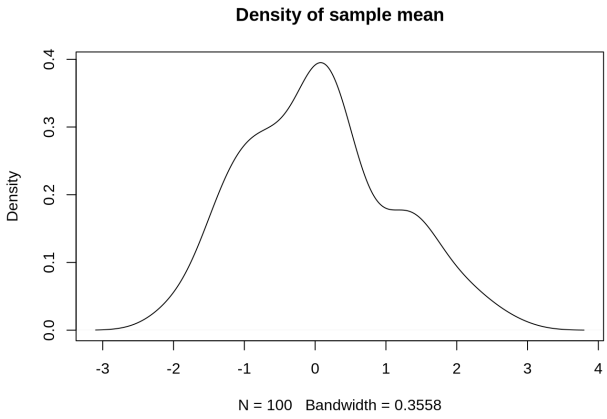
For $n = 50$

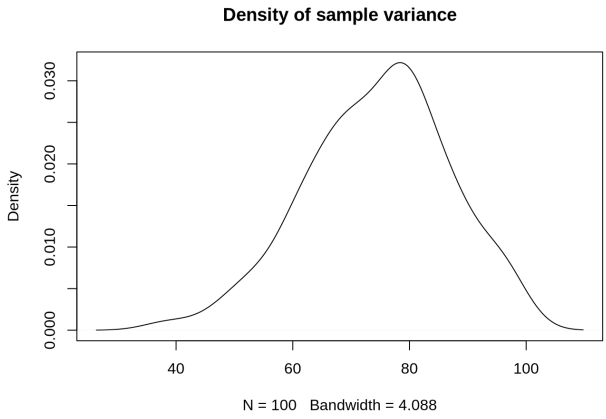

Density of sample mean
Density
N = 100 Bandwidth = 0.3324

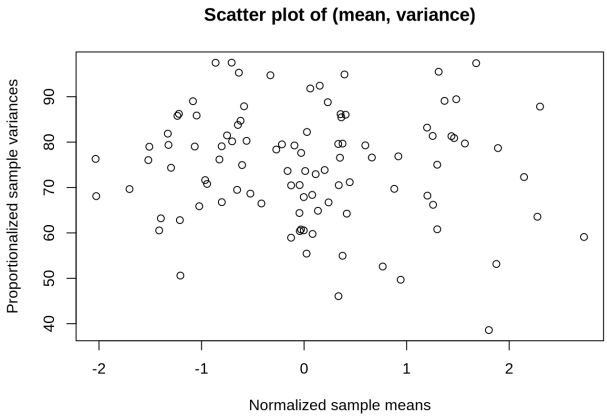

Density of sample variance
Density
N = 100 Bandwidth = 3.456


Scatter plot of (mean, variance)
Proportionalized sample variances
Normalized sample means

For $n = 75$







Problem 2.5

From the plots of the sample mean and sample variance made above, we can say that for a SRS, the sample mean follows a **normal distribution** while the sample variance follows a χ^2 **distribution.**

Problem 2.6

Based on the scatter plots generated from (sample mean, sample variance) it shows that the two have a weak correlation which proves that the two are **mutually independent.**

Problem 3

A binomial distribution can be interpreted alternatively as the sum of n Bernoulli random variables. Each Bernoulli random variable has a mean defined as p. Using this fact, we can state that $Y_1, \ldots, Y_n$ to be a series of independent Bernoulli random variables with $E(Y_i) = p$

Since the random binomial random variable, N is the sum of these, it can be said that:

$N = Y_1 + \ldots + Y_n$

To find the expected value of N, the expected value of the sum of all Bernoulli random variables must be derived:

$E(N) = E(Y_1 + \ldots + Y_n)$

However, using the property of expectations $E[X + Y] = E[X] + E[Y]$, the above can be changed to:

$E(N) = E(Y_1) + \ldots + E(Y_n)$

Since all the expected values of each Bernoulli random variable is p, the equation above is the same as adding p for n times.

$E(N) = p + \ldots + p = np$

Problem 4

Given that $T \sim \tau_n$, we know that the random variable T has a t-distribution and is thus also a continuous random variable.

To find the expected value of a continous random variable, the following formula will be used:
$E[X] = \int_{-\infty}^{\infty} x * f(x)\, dx$

and $f(x)$ will be the probability density function of T: $f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n+1}{2})}(1 + \frac{x^2}{n})^{-(\frac{n+1}{2})}$

So, $E[X] = \int_{-\infty}^{\infty} x* \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n+1}{2})}(1 + \frac{x^2}{n})^{-(\frac{n+1}{2})}\ dx$

Simplifying further, $E[X] = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n+1}{2})}\ \int_{-\infty}^{\infty} x * (1 + \frac{x^2}{n})^{-(\frac{n+1}{2})}$

Note that the function $f(x) = x * (1 + \frac{x^2}{n})^{-(\frac{n+1}{2})}$ is an odd function since $-f(x) = f(-x)$

When taking the integral of an odd function, the result will always be 0.

Thus, $E[X] = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n+1}{2})} * 0 = 0$