Stat 3202 Lab 6

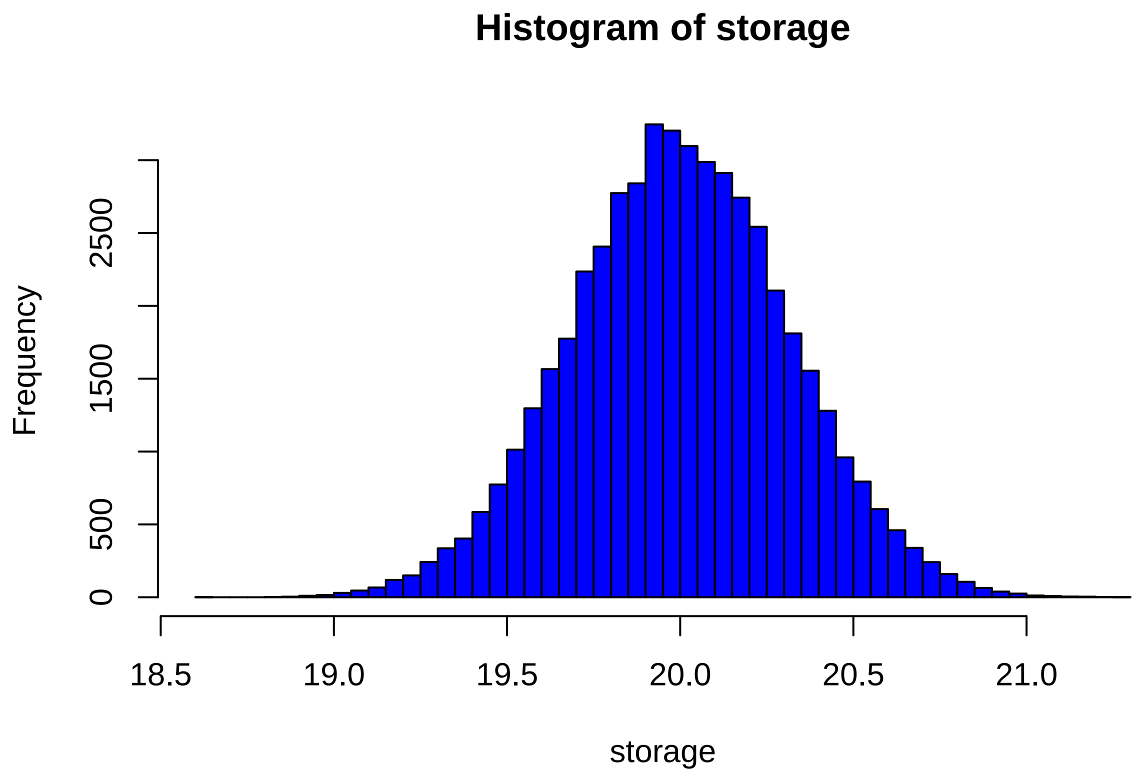
Nathan Johnson.9254

2023-02-27

Problem 1: Part a: $N(\mu, \frac{\sigma^2}{n})$.

Part b:

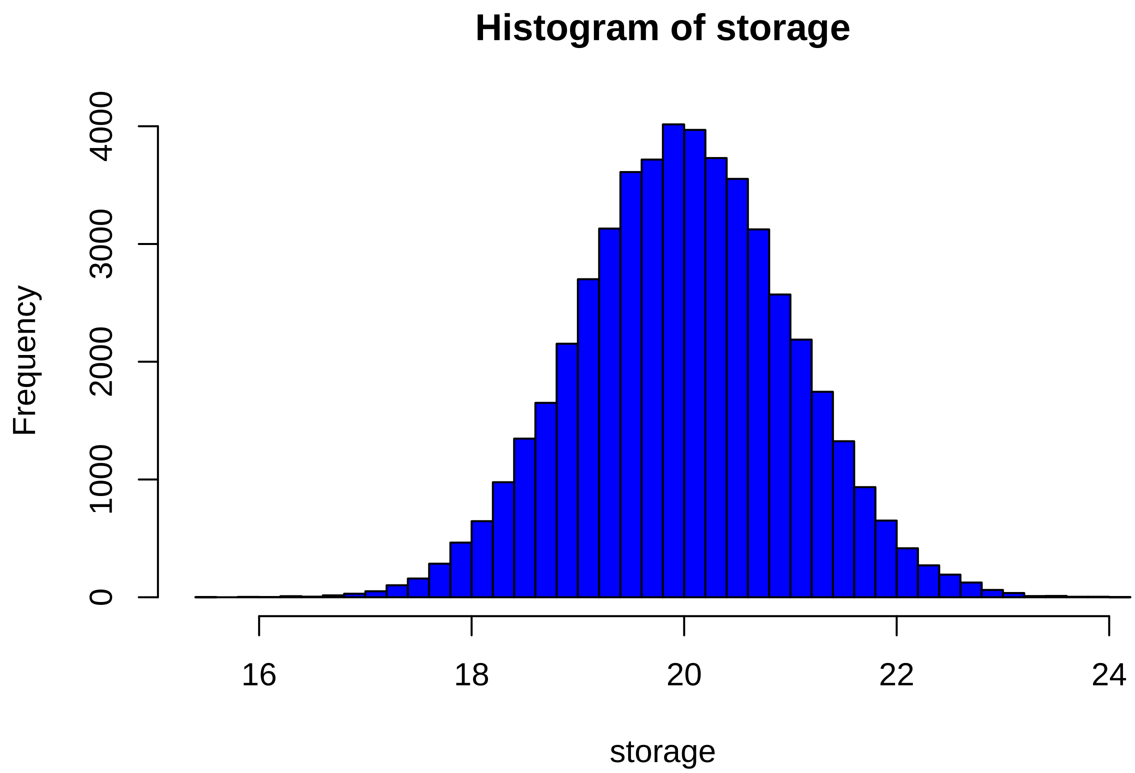
```
set.seed(1)

mc = 50000

n = 40
mu = 20
x = sqrt(4)

storage = c()

for(i in 1:mc) {
  samp = rnorm(n, mu, x)
  storage[i] = mean(samp)
}

hist(storage, col = "blue", breaks = 50)
```



$N(20, \frac{2}{\sqrt{40}})$.

Part c:

```
set.seed(1)

mc = 50000

n = 4
mu = 20
x = sqrt(4)

storage = c()

for(i in 1:mc) {
  samp = rnorm(n, mu, x)
  storage[i] = mean(samp)
}

hist(storage, col = "blue", breaks = 50)
```

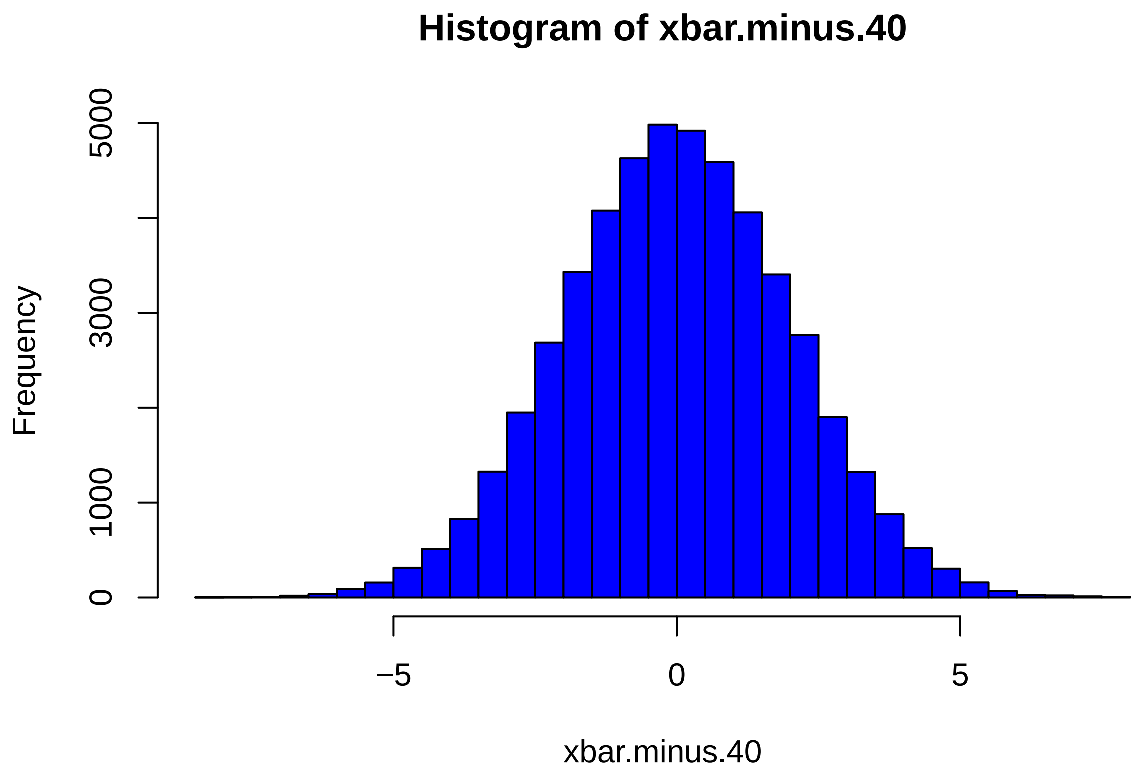


$N(20, \frac{2}{\sqrt{4}})$.

Part d: The sample sizes didn't impact the mean, but did impact the variation of the sampling distribution of $\bar{x}$.

Part e:

```
set.seed(1)

mc = 50000

n = 9
mu = 40
x = sqrt(36)

xbar.minus.40 = c()

for(i in 1:mc) {
  sample = rnorm(n, mu, x)

  xbar.minus.40[i] = mean(sample - 40)
}

hist(xbar.minus.40, col = "blue", breaks = 50)
```



i: Yes, it does look normal. **ii.**

```
mean(xbar.minus.40)
```

```
## [1] -0.0002237961
```

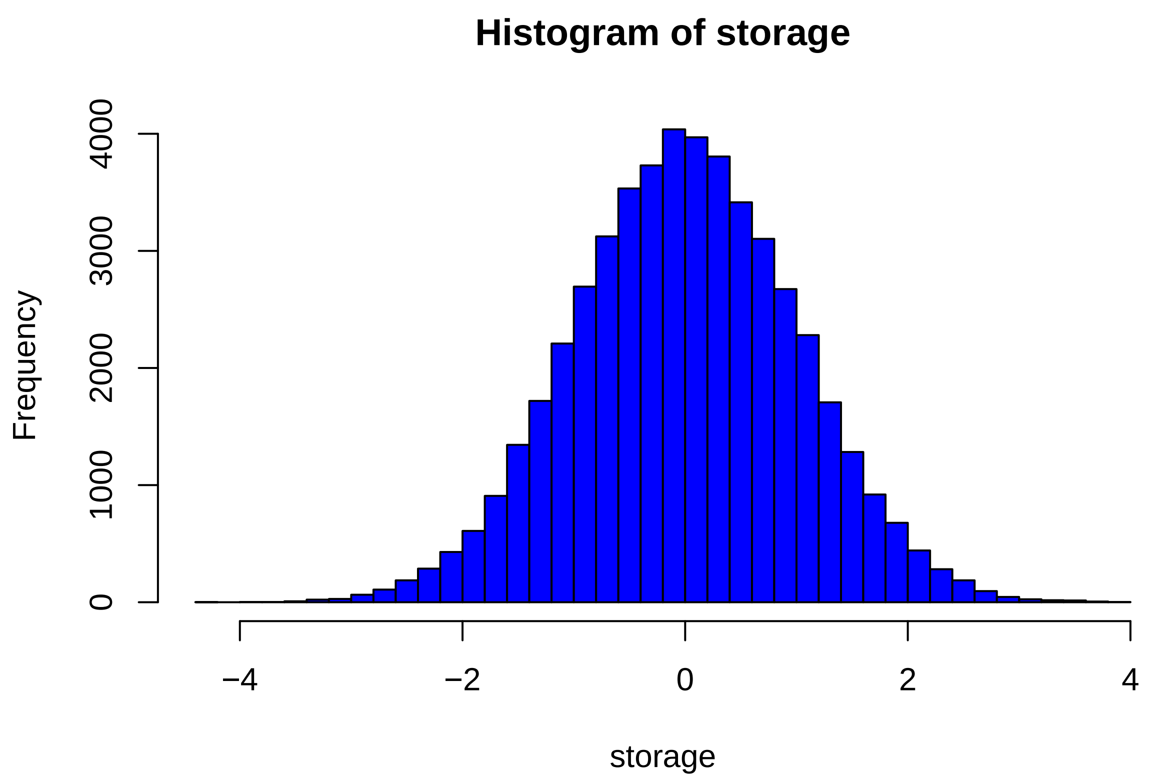
```
var(xbar.minus.40)
```

```
## [1] 3.983718
```

Part f:

```
set.seed(1)

mc = 50000
n = 9

mu = 40
sigma = sqrt(36)

storage = c()

for(i in 1:mc) {
  sample = rnorm(n, mu, sigma)
```

```
  storage[i] = ((mean(sample) - mu) / (sigma/sqrt(n)))
}

hist(storage, col = "blue", breaks = 50)
```




```
mean(storage)
```

```
## [1] -0.0001118981
```

```
var(storage)
```

```
## [1] 0.9959294
```

i. Yes, it is nearly $N(0, 1)$. **ii.** Mean = -0.0001118981. Variance = 0.9959294.

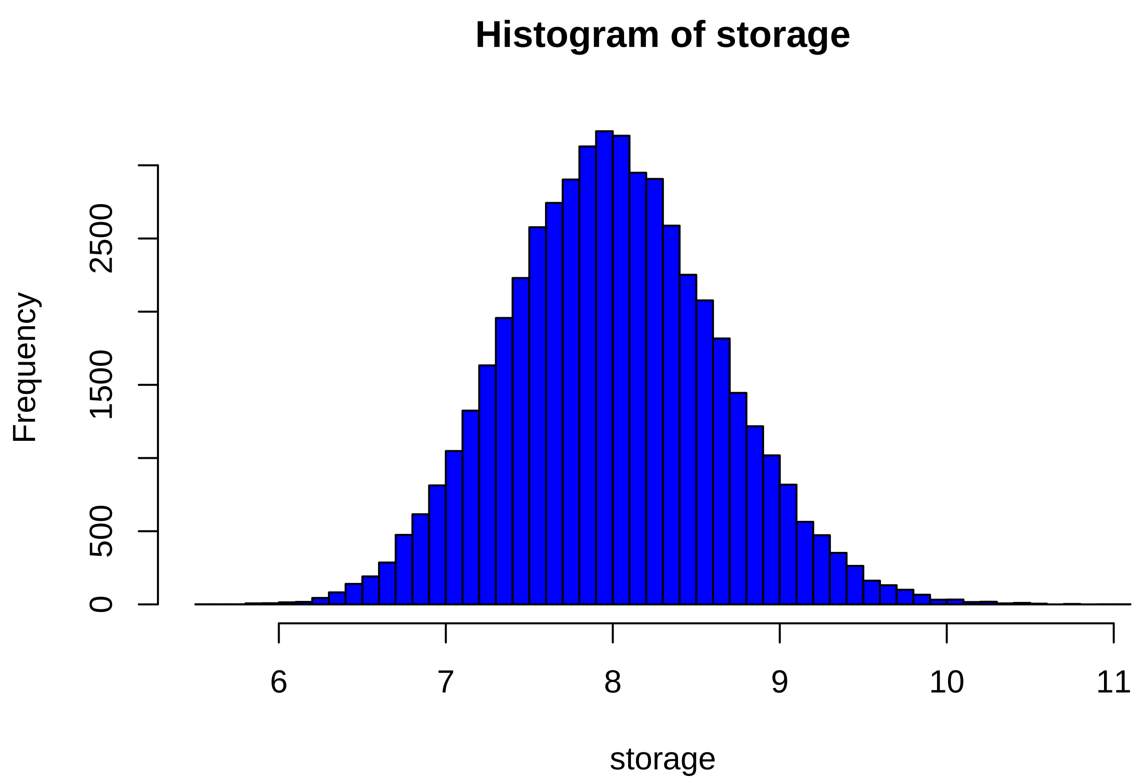
Problem 2: Part a: The CLT is if $x_1, ..., x_n$ are iid from a distribution with finite mean $/mu$ and finite variance σ^2, then the sampling distribution of $\bar{x}$ is approximately, for large enough n, normal.

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim (approx)N(0, 1)$$

Part b: Approximately normal with sample mean $\bar{x} = \lambda$

Part c:

```
set.seed(1)

mc = 50000

n = 40
mu = 8

storage = c()

for(i in 1:mc) {
  samp = rchisq(n, mu)
  storage[i] = mean(samp)
}

hist(storage, col = "blue", breaks = 50)
```

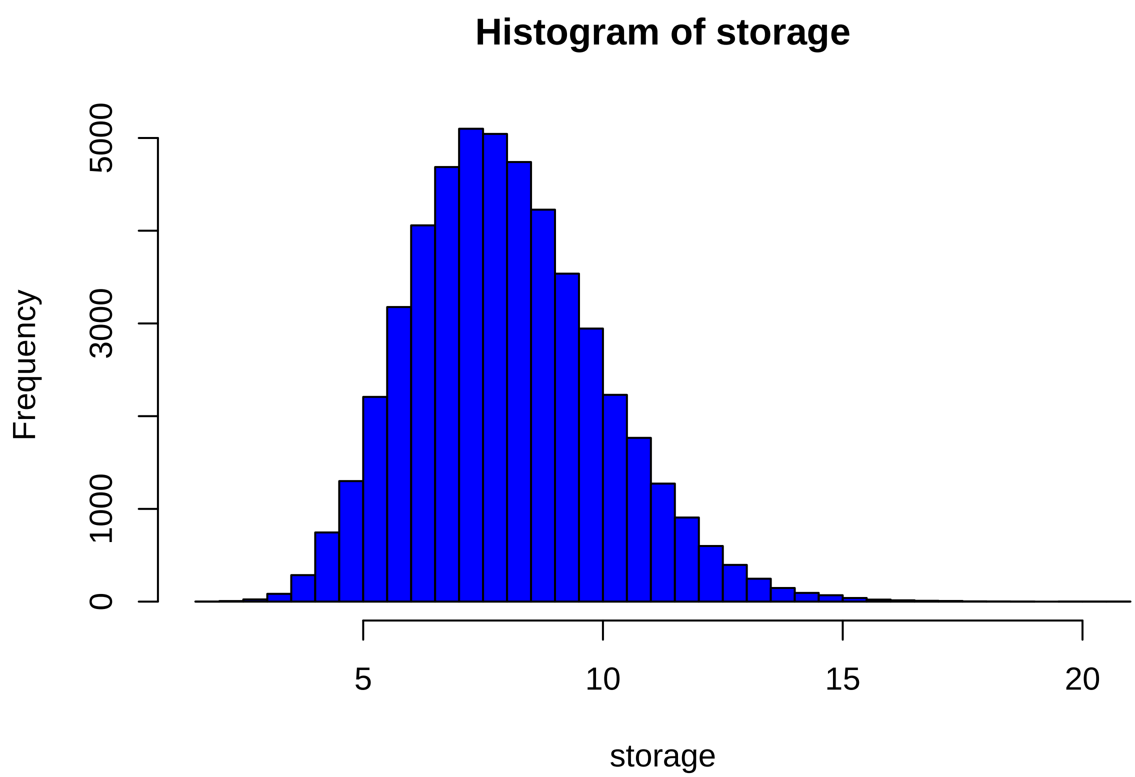


Yes, it looks normal.

Part d:

```
set.seed(1)

mc = 50000

n = 4
mu = 8
```

```r
storage = c()

for(i in 1:mc) {
  samp = rchisq(n, mu)
  storage[i] = mean(samp)
}

hist(storage, col = "blue", breaks = 50)
```



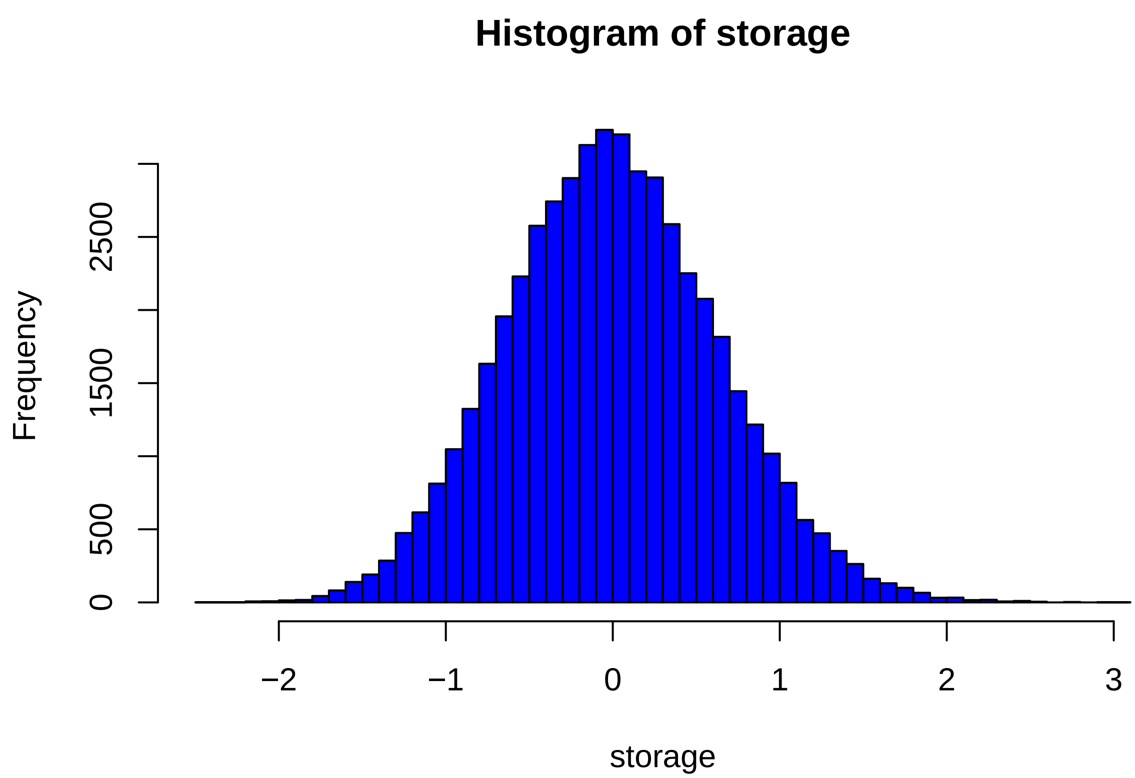
No because λ can't be less than 0 due to the bounds of the chi-square distribution. So it is right skewed.

Part e: Because part d has less samples so it's variance is larger.

Part f:

```r
set.seed(1)

mc = 50000

n = 40
mu = 8

storage = c()

for(i in 1:mc) {
  samp = rchisq(n, mu)
  storage[i] = mean(samp) - 8
```

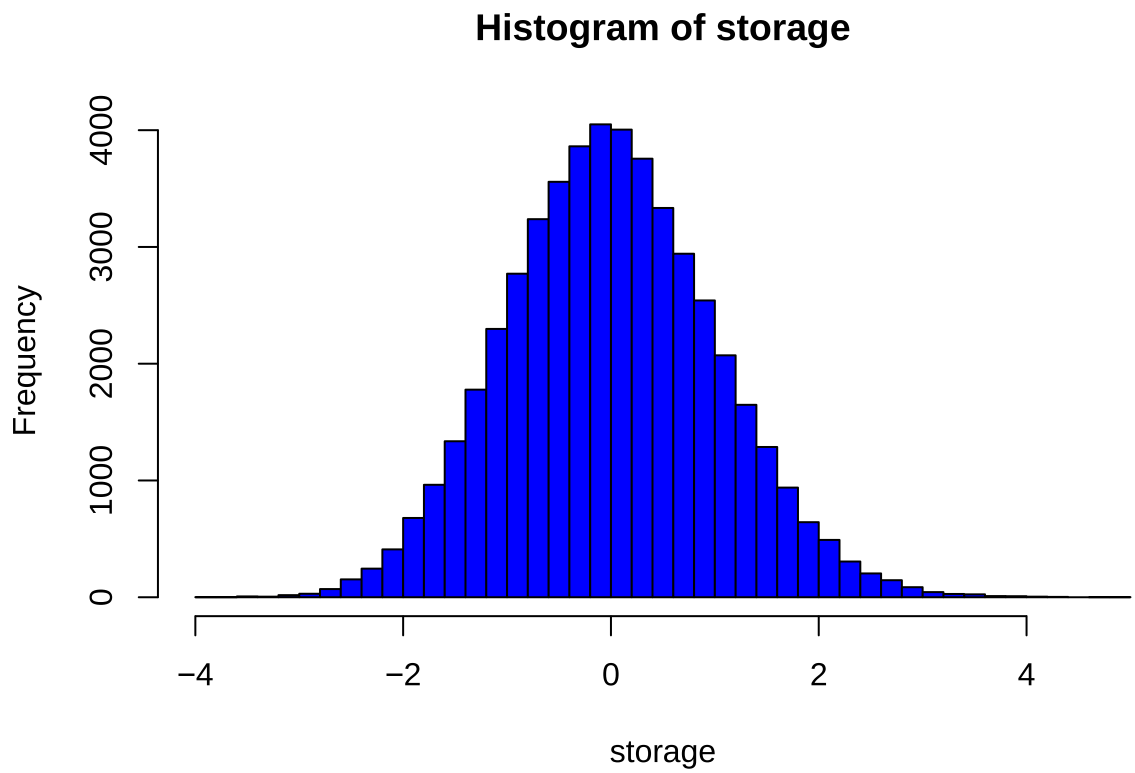
```
}

hist(storage, col = "blue", breaks = 50)
```



Yes, it does look normal.

Part g:

```
set.seed(1)

mc = 50000

n = 40
mu = 8
s = sqrt(16)

storage = c()

for(i in 1:mc) {
  samp = rchisq(n, mu)
  storage[i] = ((mean(samp) - 8)/(s/(sqrt(n))))
}

hist(storage, col = "blue", breaks = 50)
```




```
mean(storage)
```

```
## [1] -0.0006903592
```

```
var(storage)
```

```
## [1] 1.008031
```

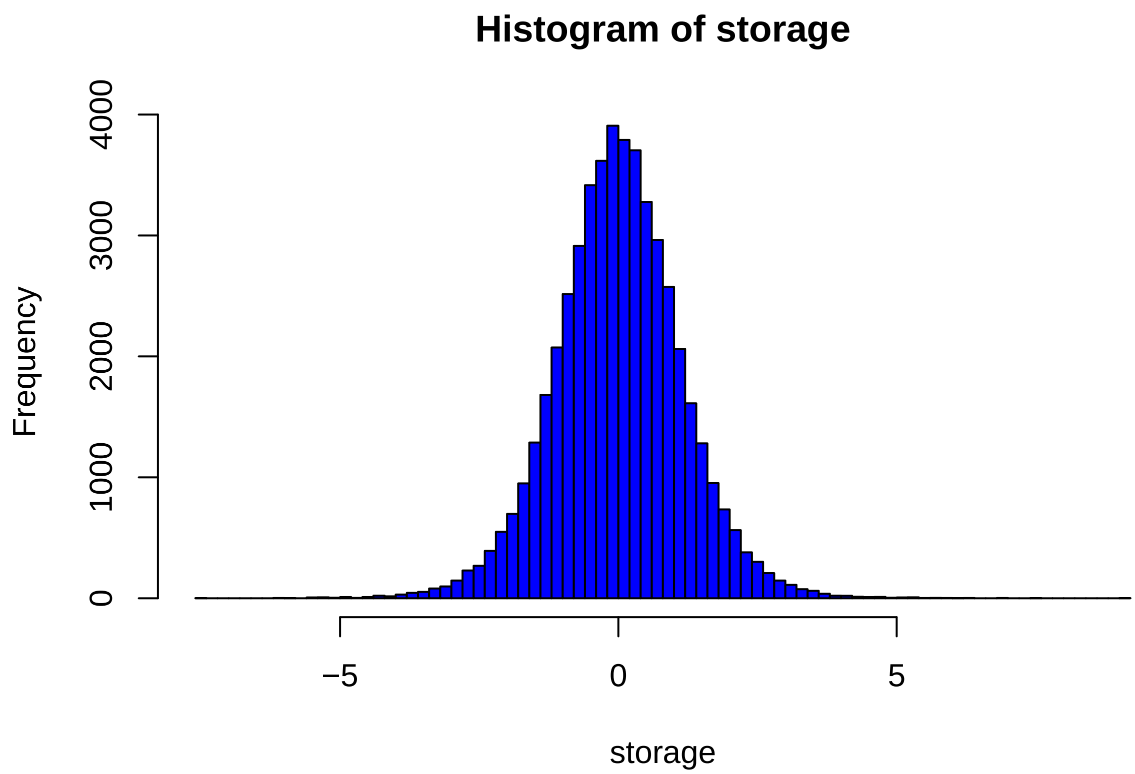
The sampling distribution looks normal. The mean is -0.0006903592 and the variance is 1.008031 which are both pretty close to 0 and 1.

Problem 3: Part a: $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$

Part b:

```
set.seed(1)

mc = 50000

n = 9
mu = 40
x = sqrt(36)

storage = c()
```

```r
for(i in 1:mc) {
  samp = rnorm(n, mu, x)
  s = sqrt(sum((samp - mean(samp))^2)/(n-1))
  storage[i] = (mean(samp) - 40)/(s/sqrt(n))
}

hist(storage, col = "blue", breaks = 60)
```

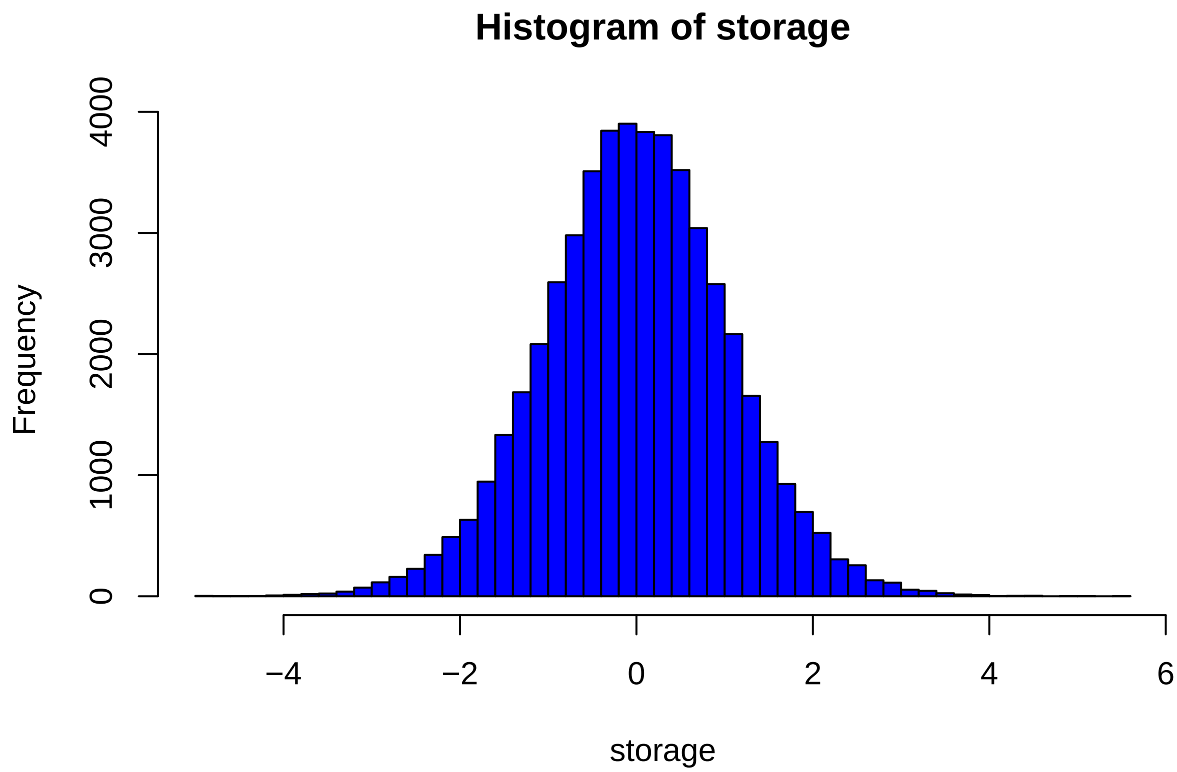


Yes, it looks normal, but if we substitute in σ for s, then the distribution "stops" when less than -4 and above 4. Where the t distribution "stops" at about -5 and 5.

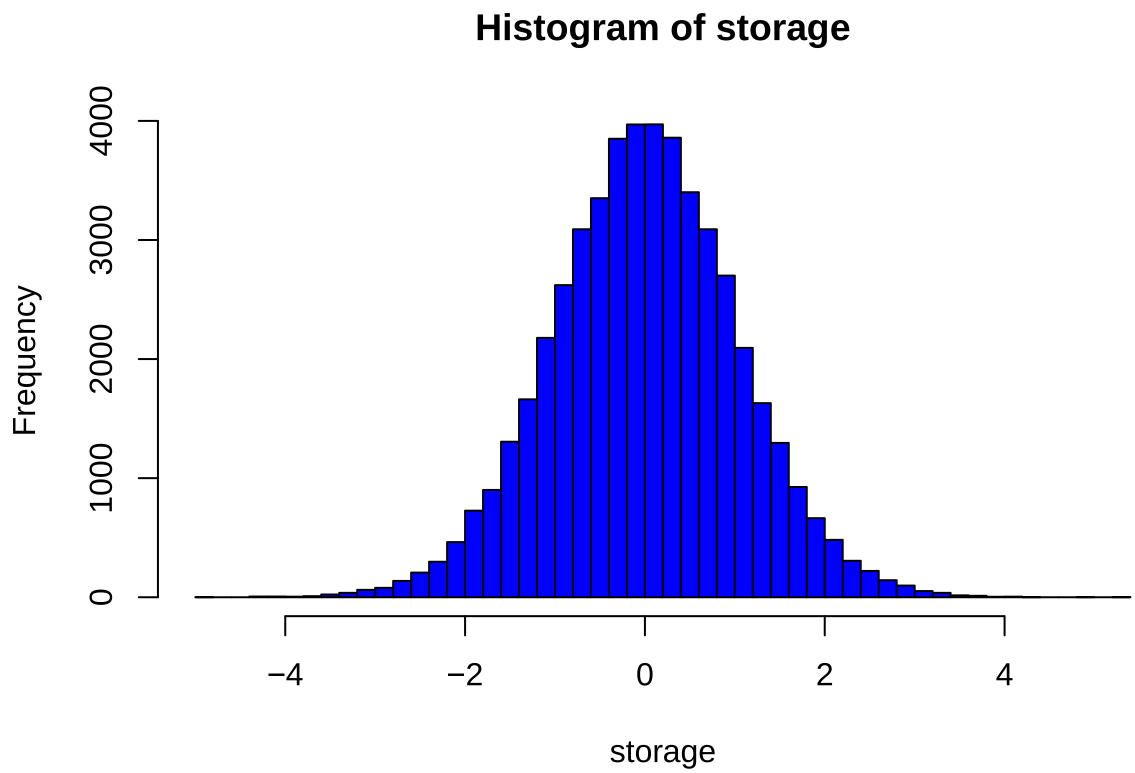
Part c:

```r
set.seed(1)

mc = 50000

n = 20
mu = 40
x = sqrt(36)

storage = c()

for(i in 1:mc) {
  samp = rnorm(n, mu, x)
  s = sqrt(sum((samp - mean(samp))^2)/(n-1))
```

```
  storage[i] = (mean(samp) - 40)/(s/sqrt(n))
}

hist(storage, col = "blue", breaks = 50)
```




```
set.seed(1)

mc = 50000

n = 30
mu = 40
x = sqrt(36)

storage = c()

for(i in 1:mc) {
  samp = rnorm(n, mu, x)
  s = sqrt(sum((samp - mean(samp))^2)/(n-1))
  storage[i] = (mean(samp) - 40)/(s/sqrt(n))
}

hist(storage, col = "blue", breaks = 50)
```



The sample standard deviation gets smaller and smaller.