Random Samples

Contents

We will get random samples of continuous distributions for different sample sizes.

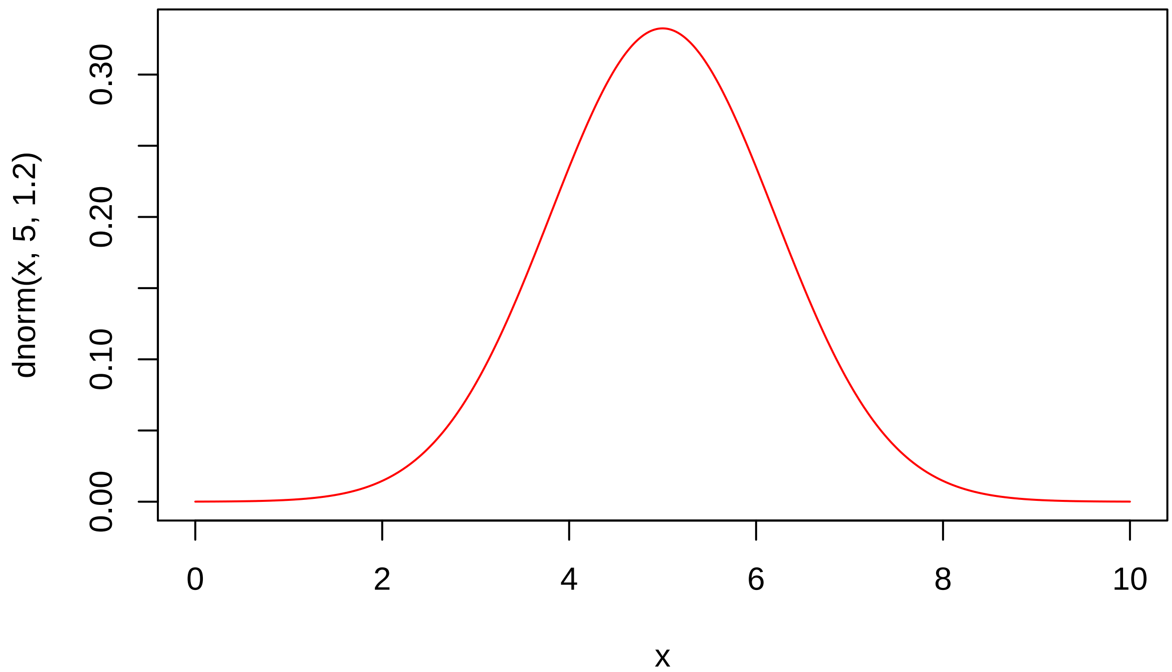
Normal Distibution

Consider the Normal distribution $N(\mu = 5, \sigma = 1.2)$. We will get random samples from this distribution and check out the histograms for each of these samples. We expect that as the sample size increases, the sample distribution will provide an increasingly improved approximation of the the true distribution.

Density of $N(\mu, \sigma)$.

Plot the density curve of $N(\mu, \sigma)$,

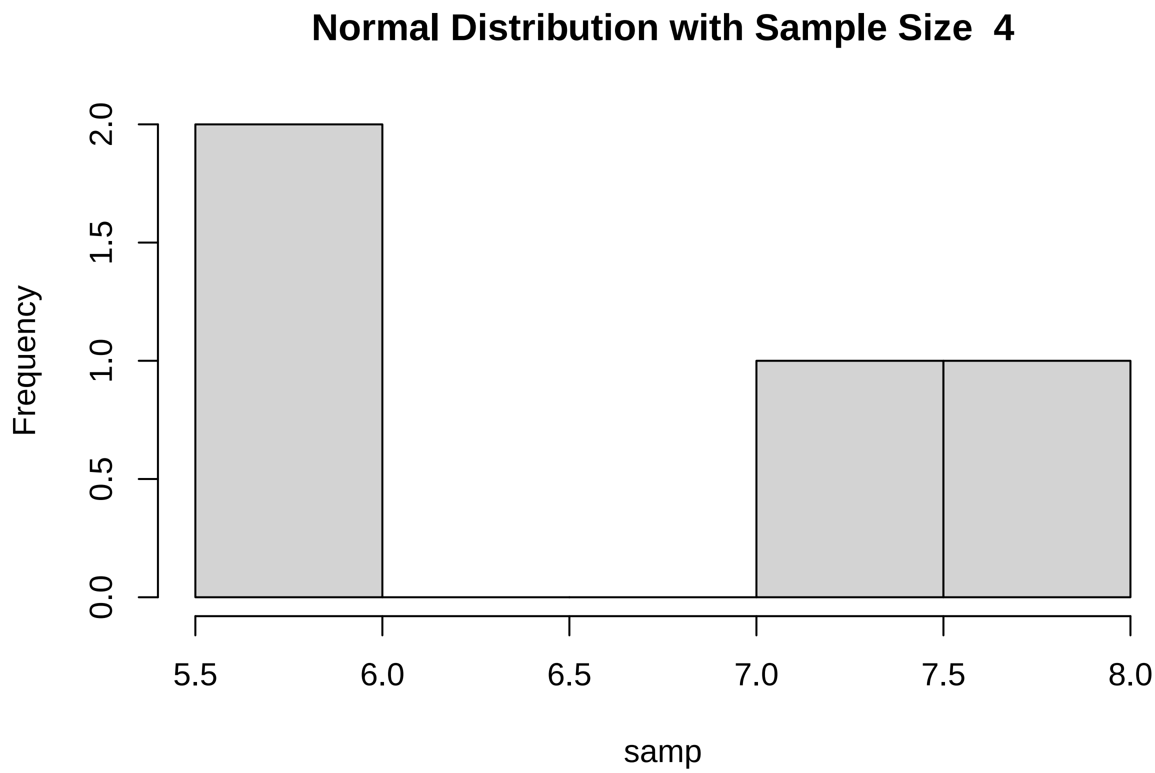
```
x <- seq(0, 10, length.out=1000)
plot(x, dnorm(x, 5, 1.2), type="l", col="red")
```



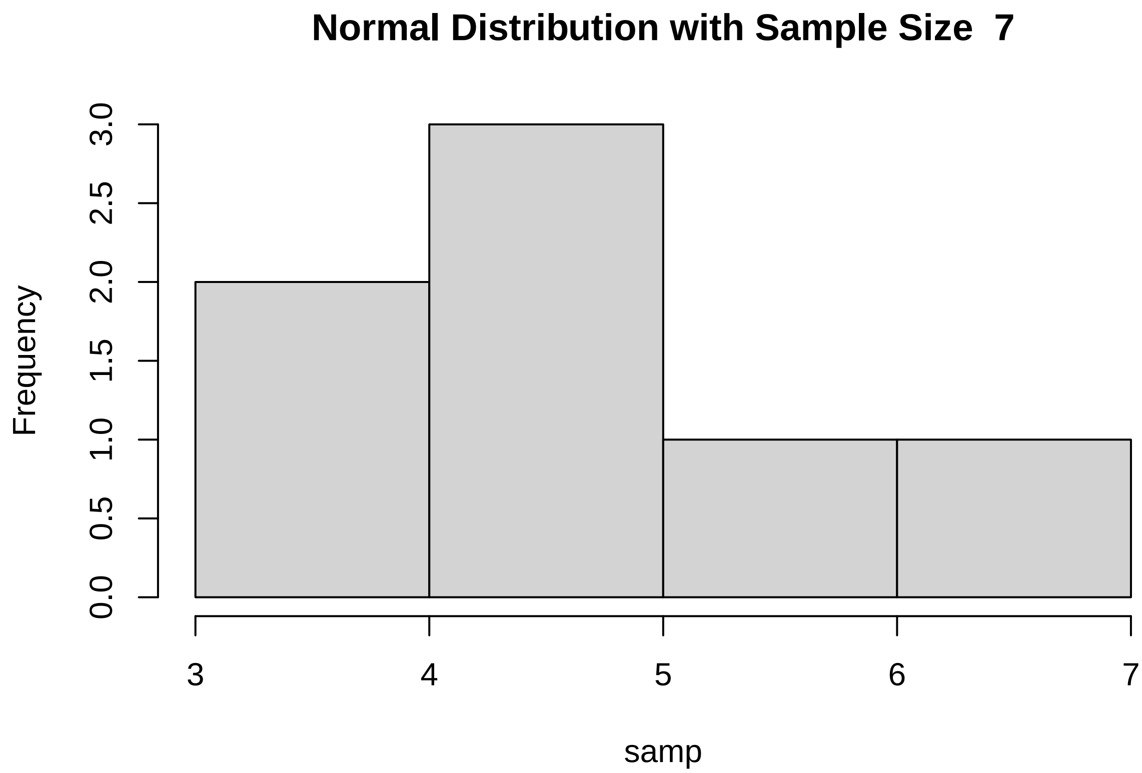
Sample distributions

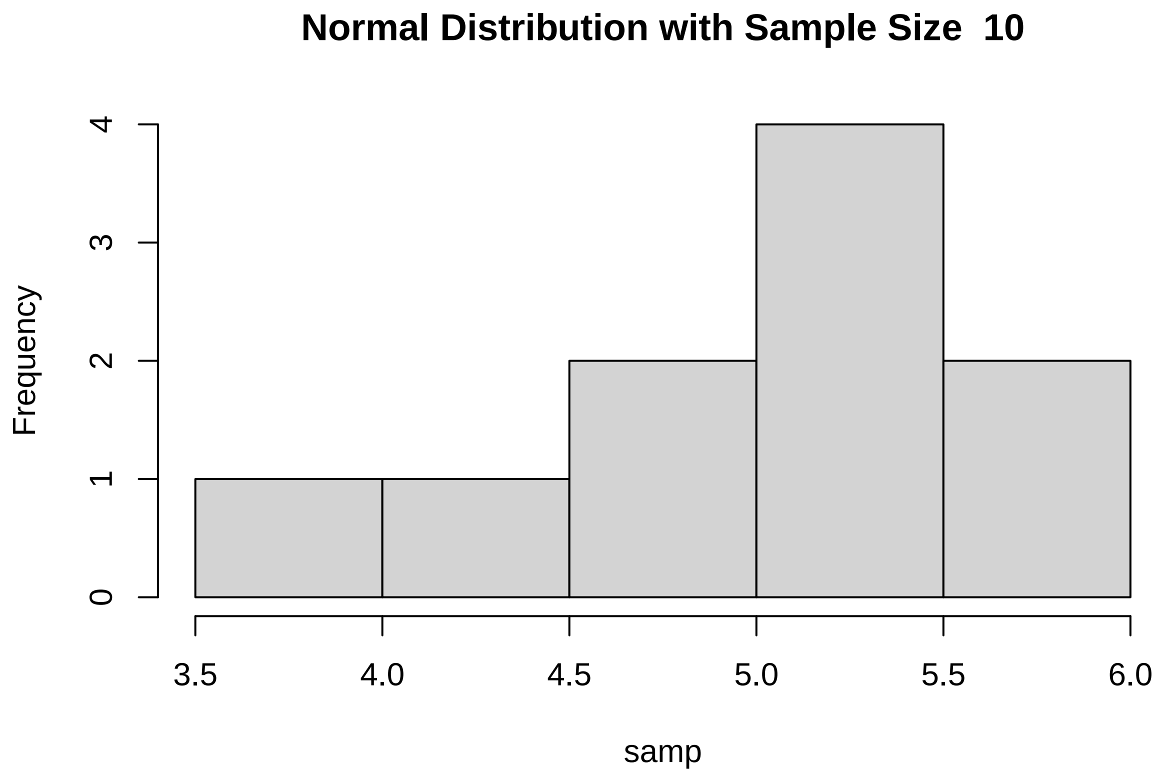
For the sample sizes $4, 7, 10, 15, 20, 30, 40, 80, 1000$ get random samples of respective size, and plot the histograms for each of these samples.

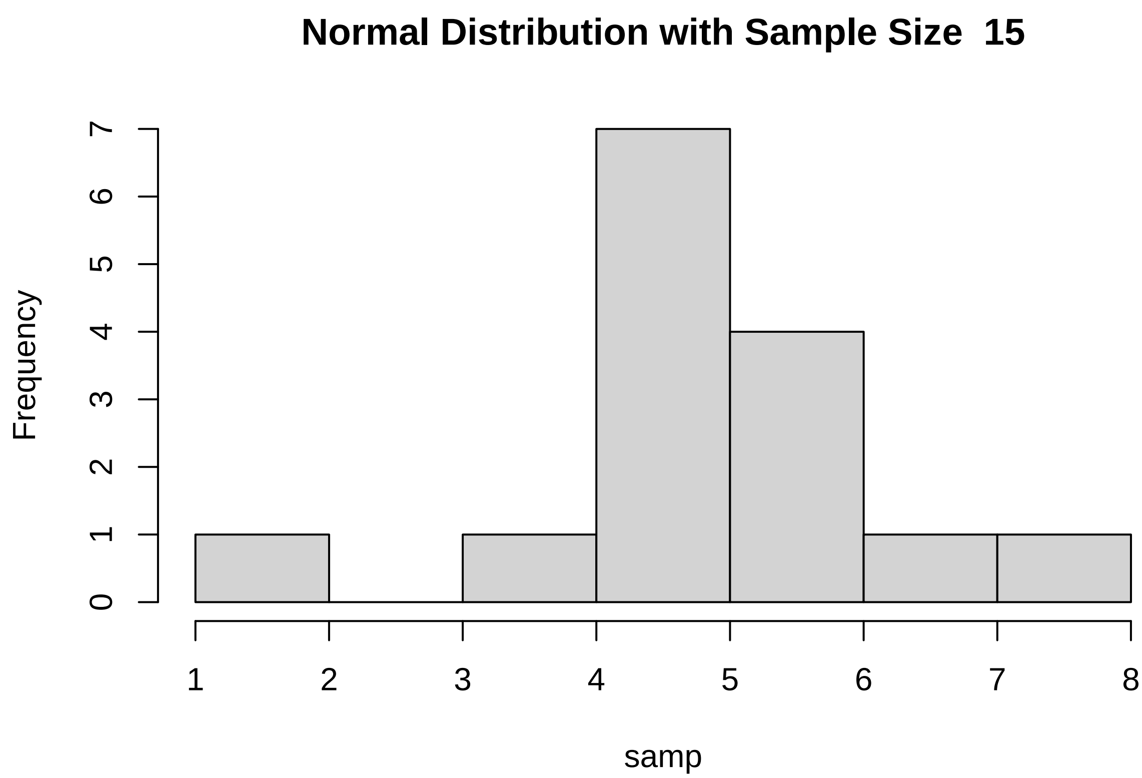
```
for (size in c(4, 7, 10, 15, 20, 30, 40, 80, 1000)) {
  samp <- rnorm(size, 5, 1.2)
  hist(samp, main=paste("Normal Distribution with Sample Size ", size))
}
```

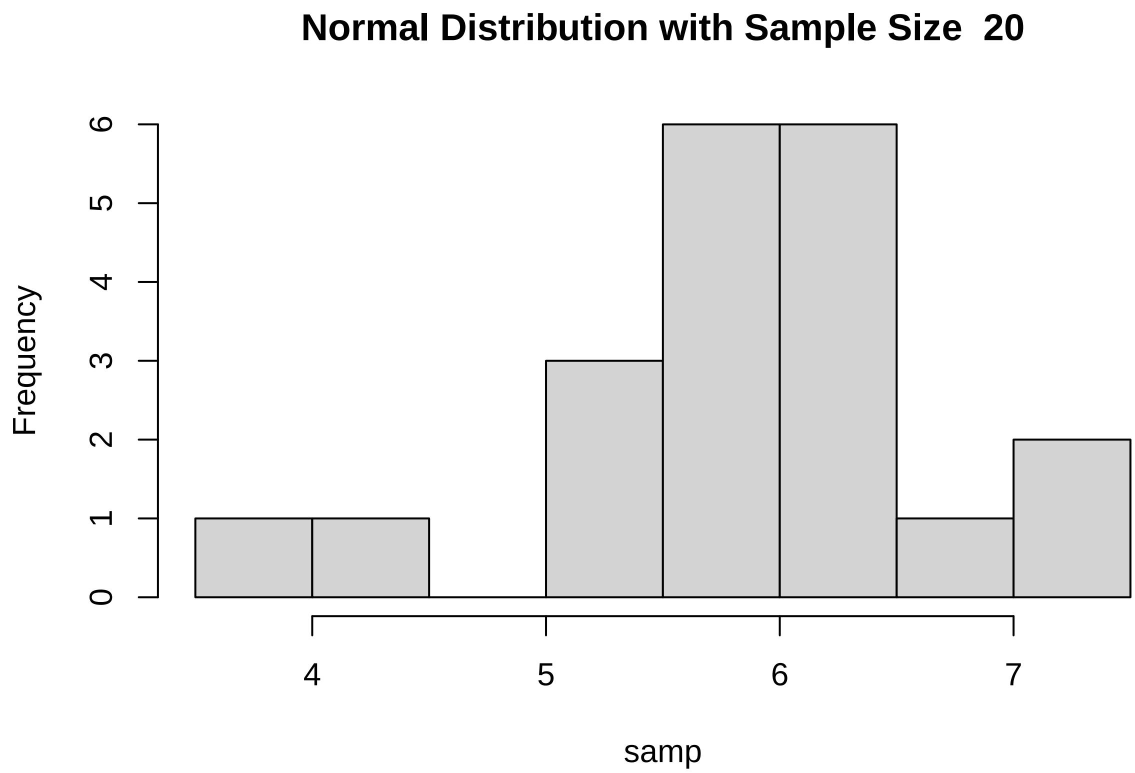


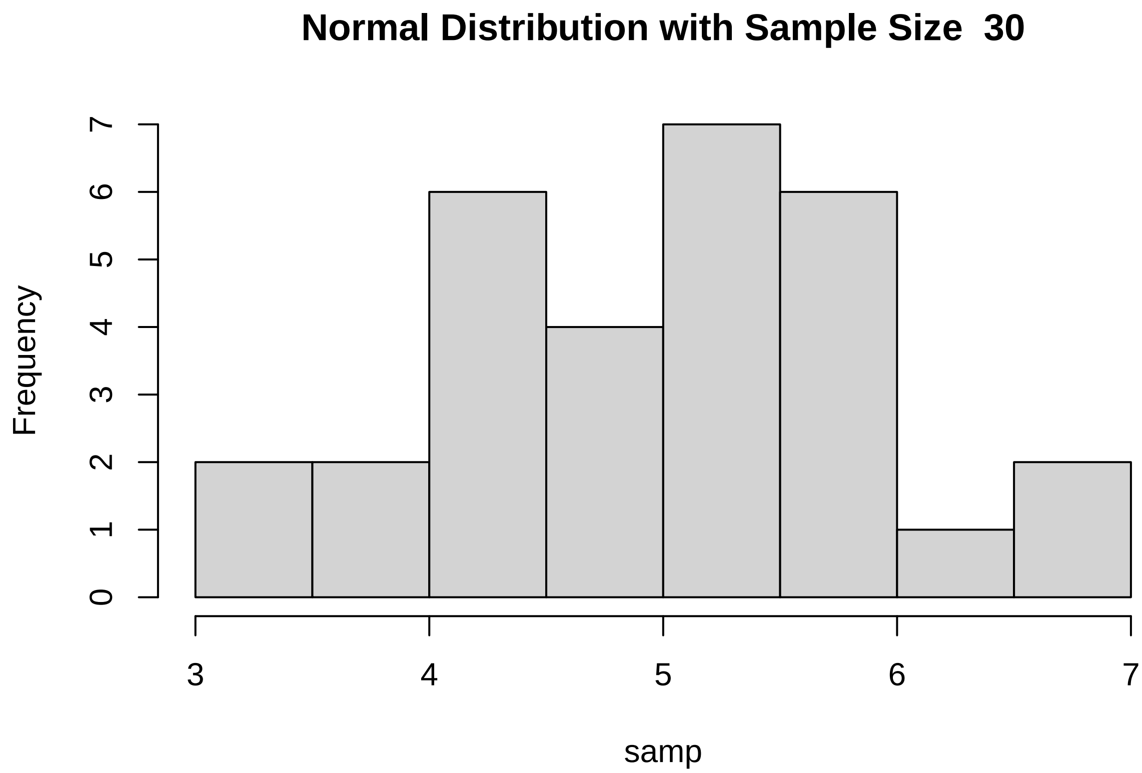
Normal Distribution with Sample Size 4
Frequency
0.0
0.5
1.0
1.5
2.0
5.5
6.0
6.5
7.0
7.5
8.0
samp

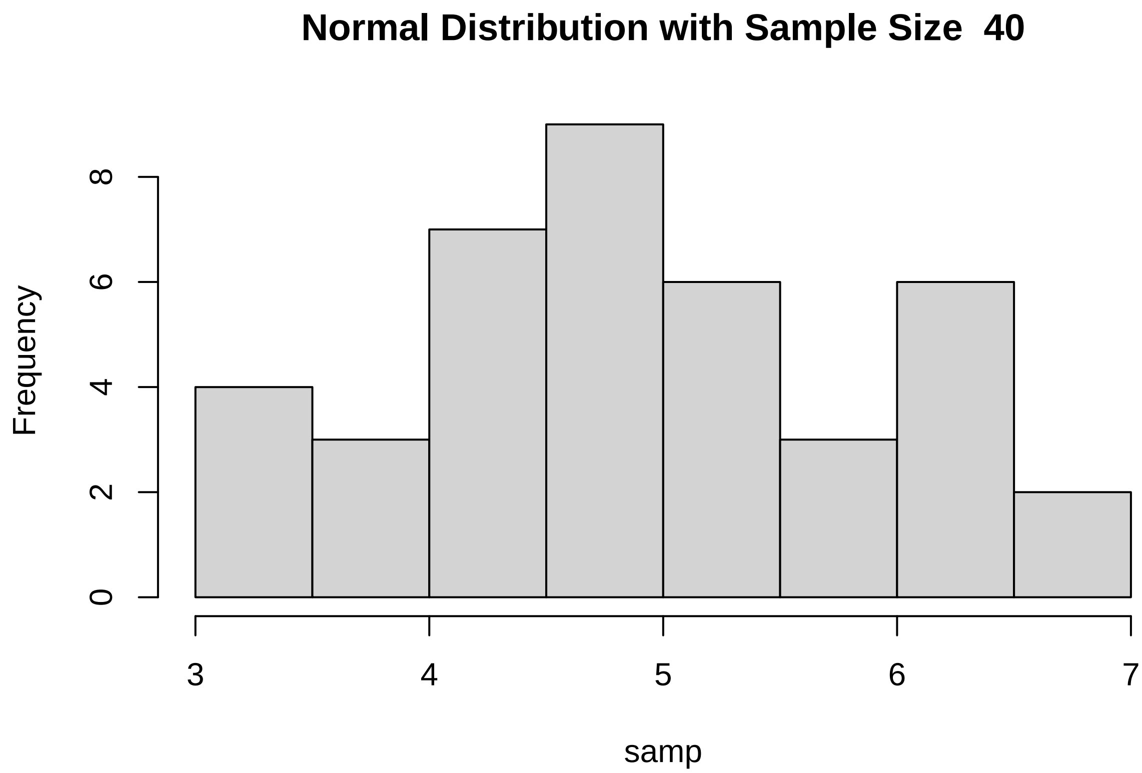

Normal Distribution with Sample Size 7
Frequency
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3
4
5
6
7
samp

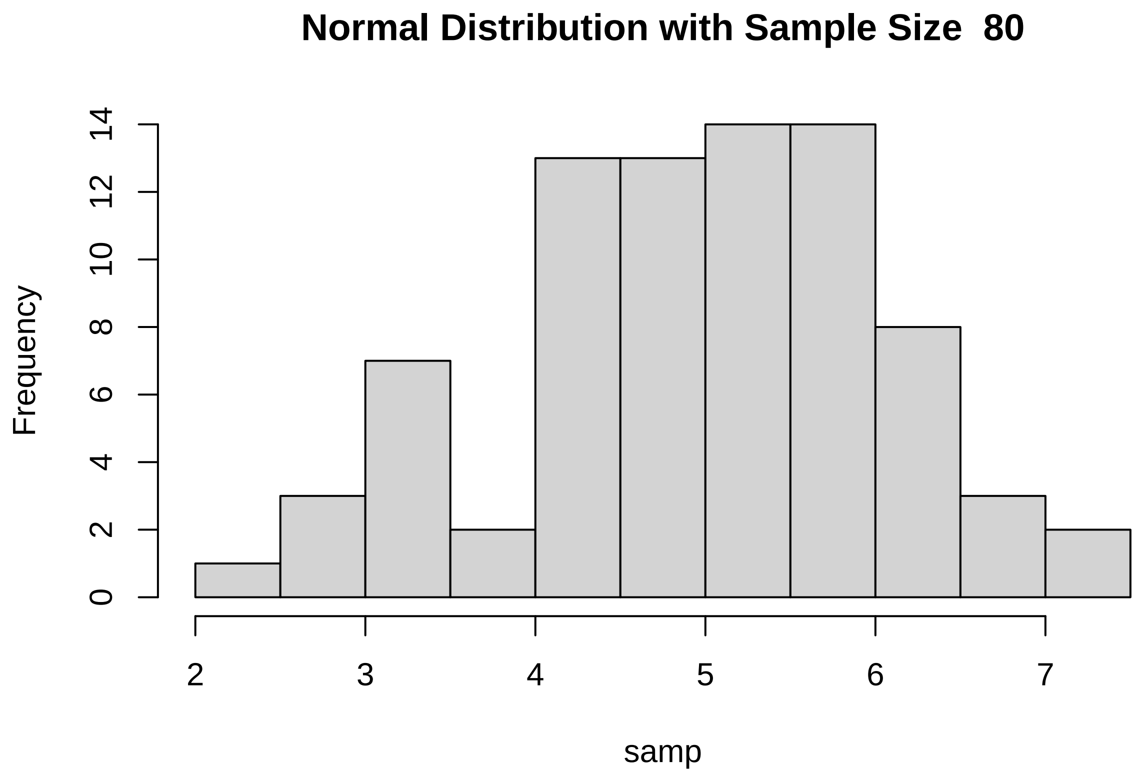

Normal Distribution with Sample Size 10
Frequency
0
1
2
3
4
3.5
4.0
4.5
5.0
5.5
6.0
samp

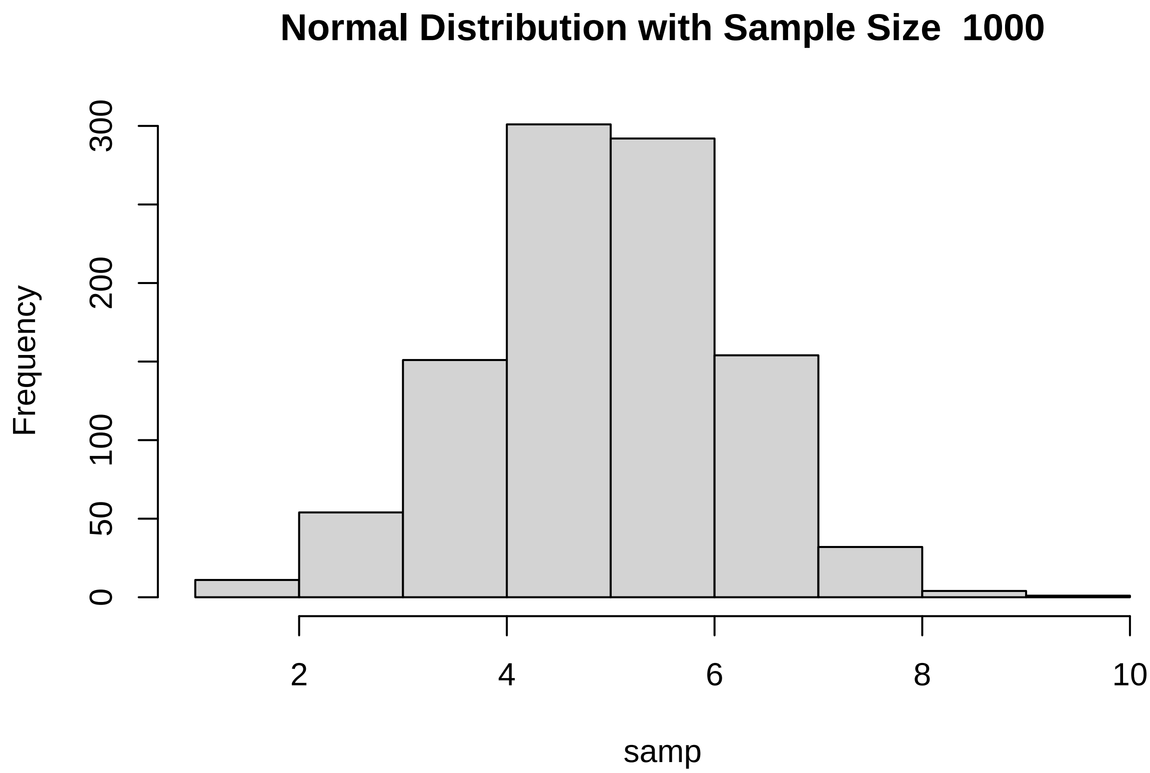

Normal Distribution with Sample Size 15
Frequency
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7
8
samp


Normal Distribution with Sample Size 20
Frequency
0
1
2
3
4
5
6
4
5
6
7
samp


Normal Distribution with Sample Size 30
Frequency
0
1
2
3
4
5
6
7
3
4
5
6
7
samp


Normal Distribution with Sample Size 40
Frequency
0
2
4
6
8
3
4
5
6
7
samp


Normal Distribution with Sample Size 80
Frequency
0
2
4
6
8
10
12
14
2
3
4
5
6
7
samp



What do you notice? Answer: Only at sample sizes 40 and 80, the shape starts resembling the normal distribution's density function. However, the deviations are still noticeable, and the histogram only becomes close at sample size 1000.

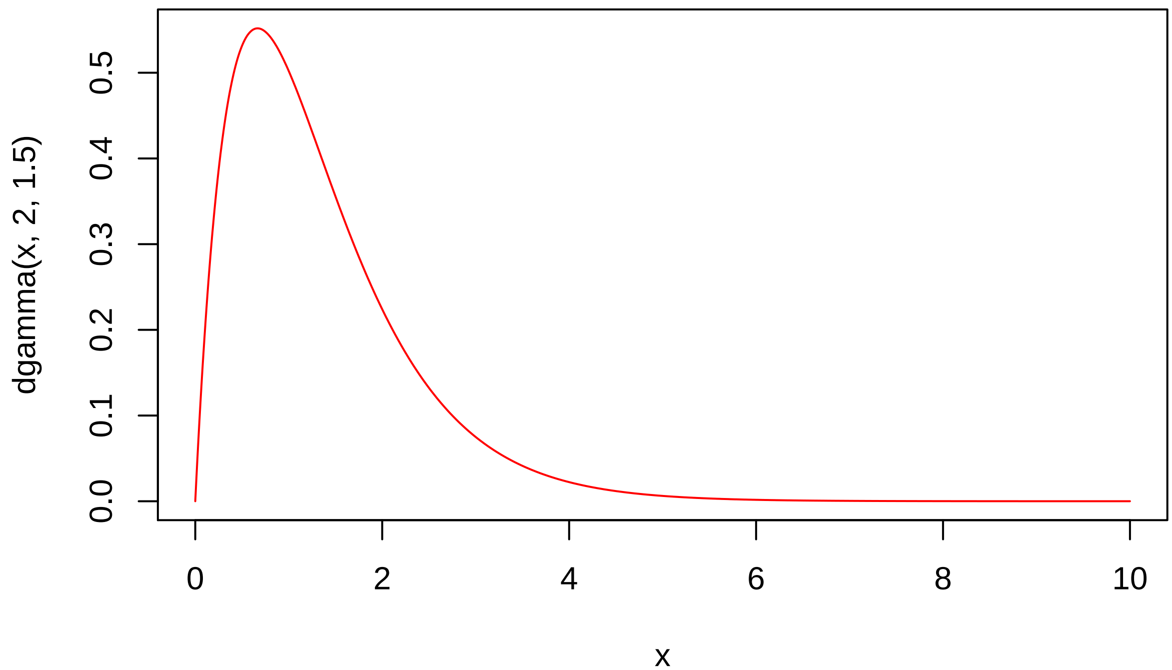
Gamma Distribution

Consider the Gamma distribution Gamma($\alpha = 2, \beta = 1.5$). We will get random samples from this distribution and check out the histograms for each of these samples. We expect that as the sample size increases, the sample distribution will provide an increasingly improved approximation of the the true distribution.

Density of Gamma($\alpha = 2, \beta = 1.5$).

Plot the density curve of Gamma($\alpha = 2, \beta = 1.5$),

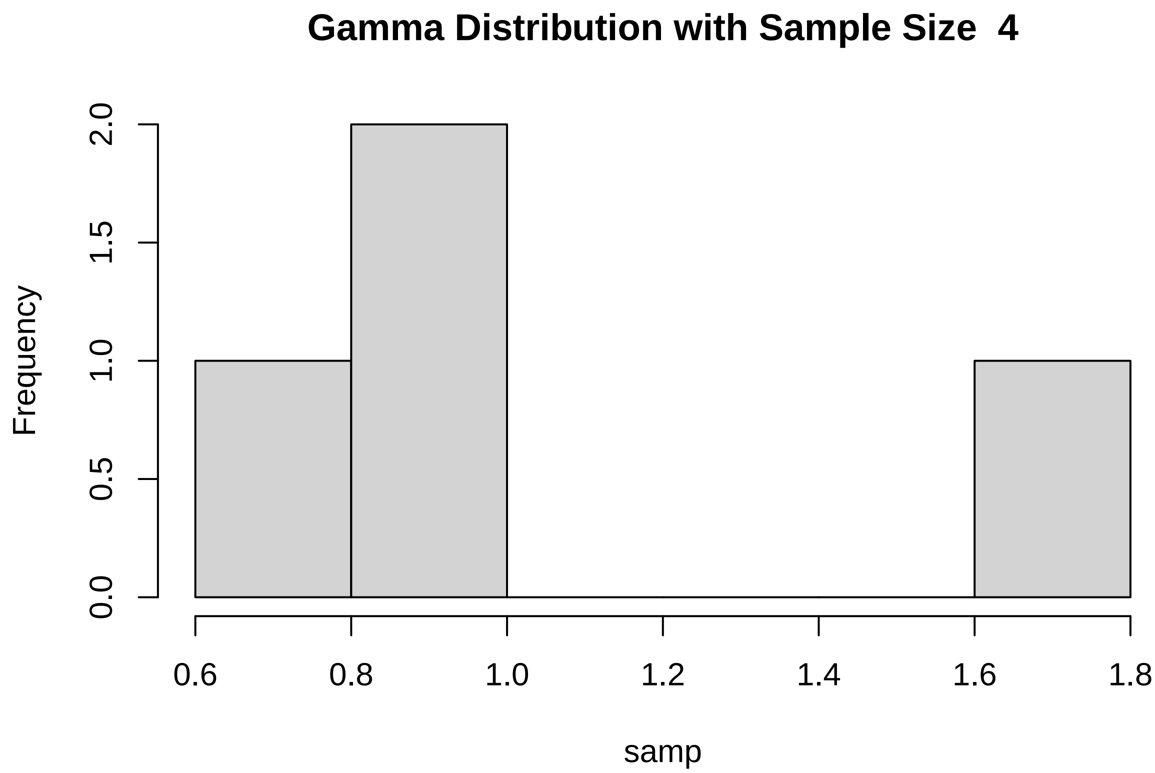
```
x <- seq(0, 10, length.out=1000)
plot(x, dgamma(x, 2, 1.5), type="l", col="red")
```

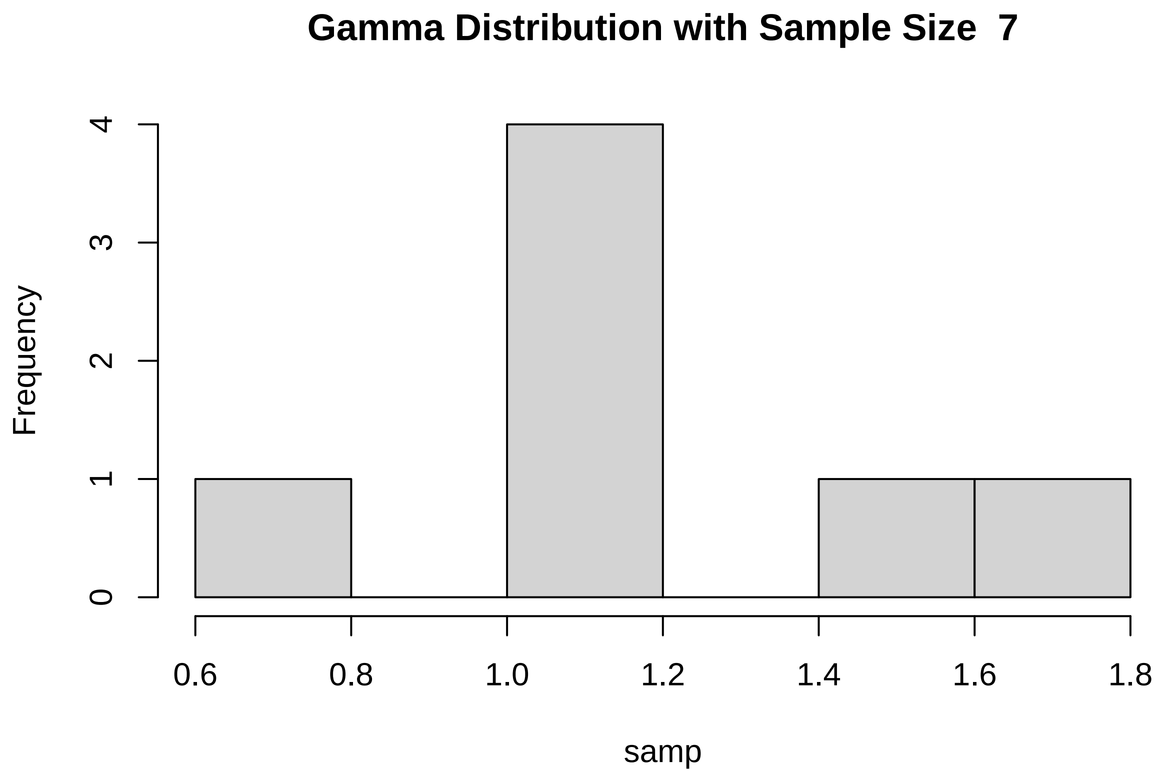


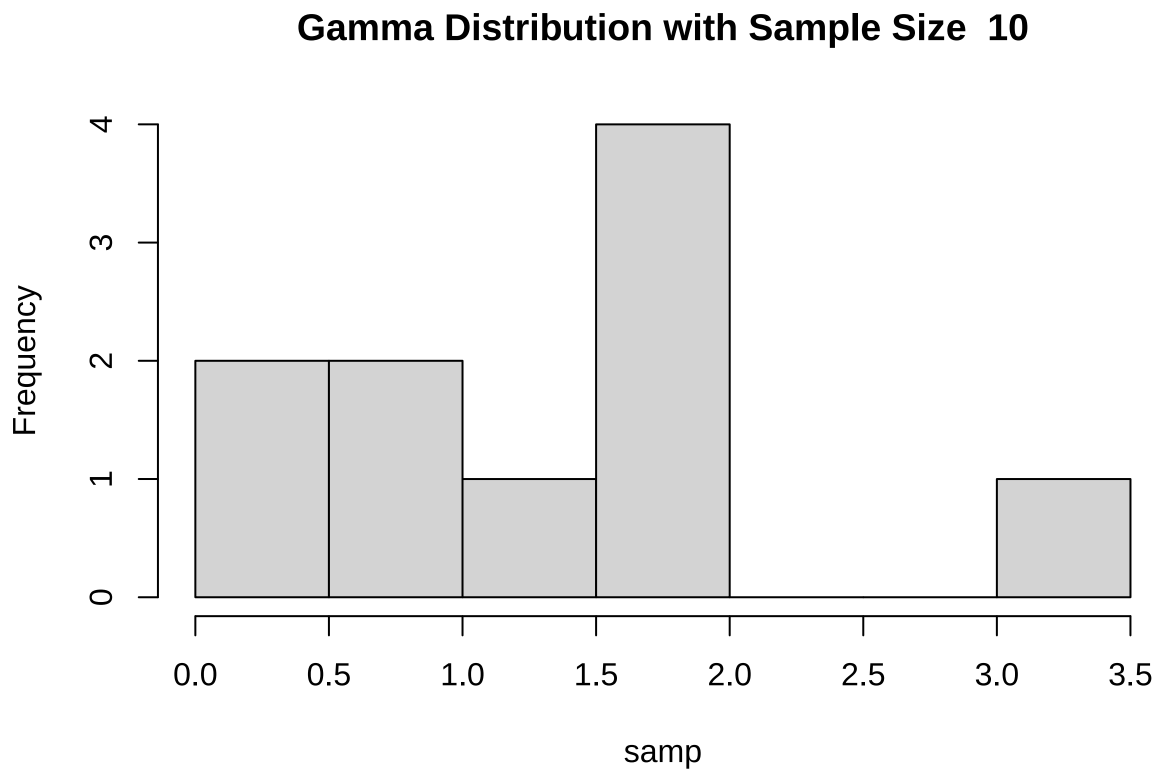
Sample distributions

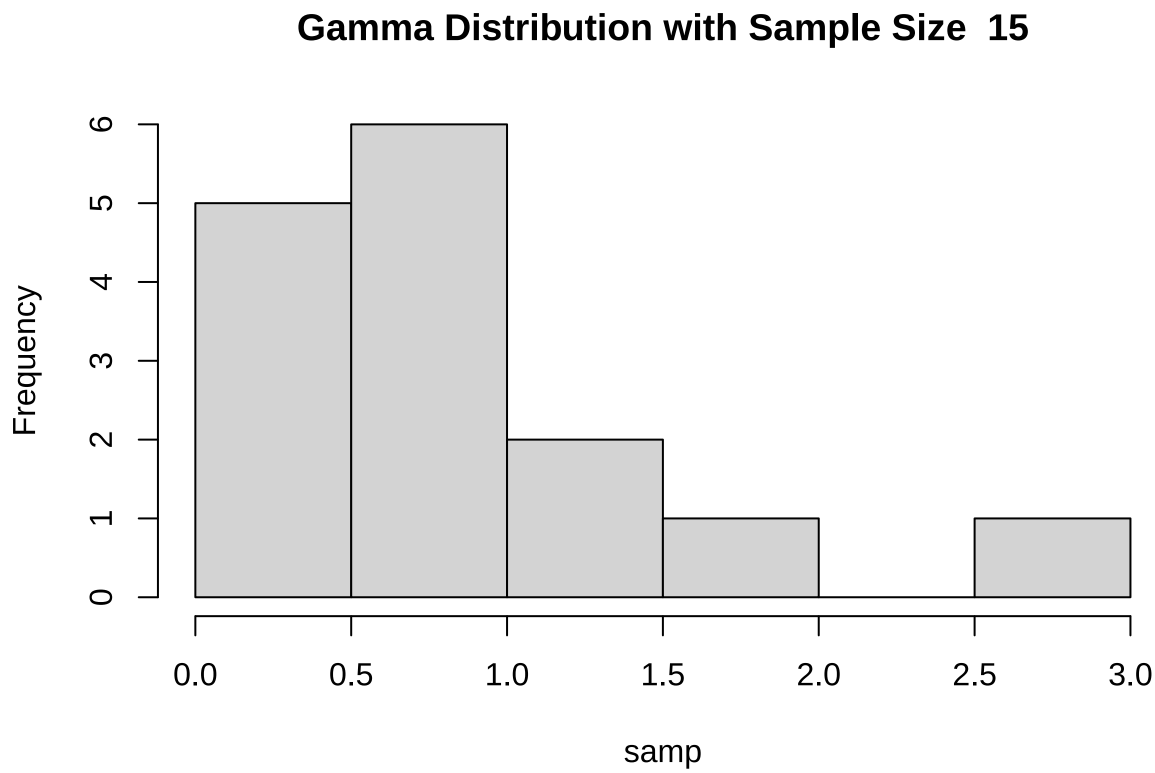
For the sample sizes $4, 7, 10, 15, 20, 30, 40, 80, 1000$ get random samples of respective size, and plot the histograms for each of these samples.

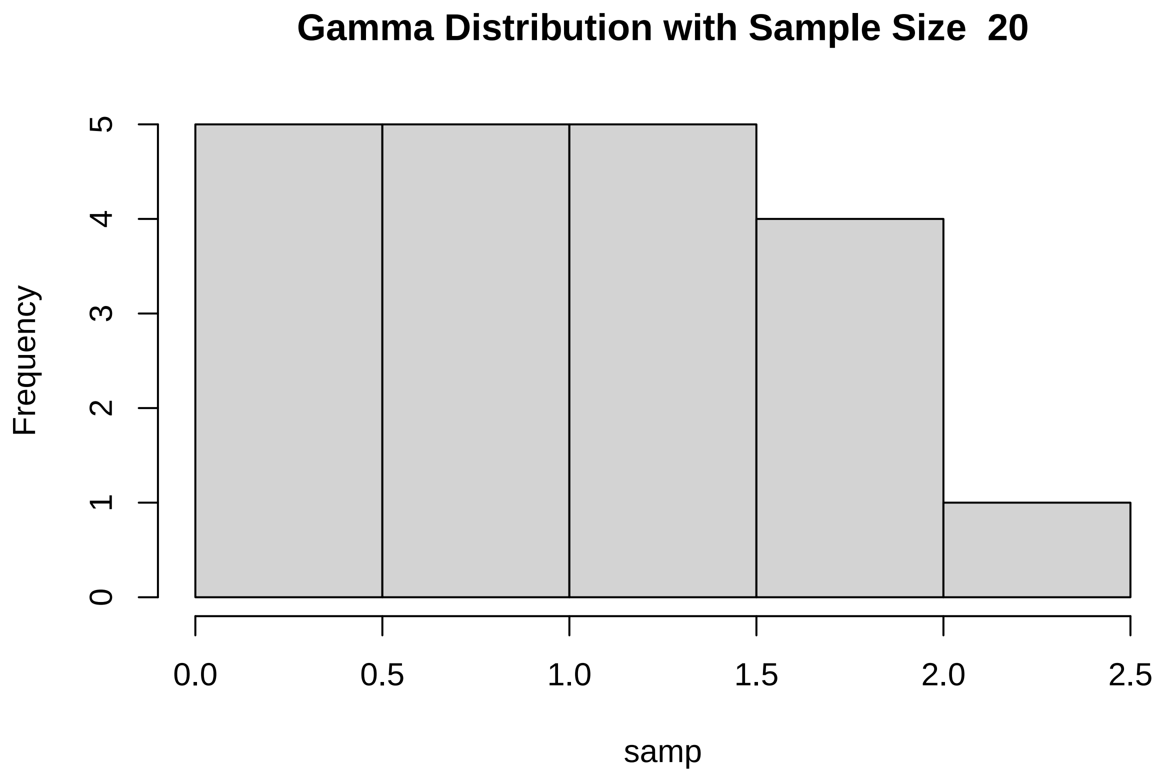
```
for (size in c(4, 7, 10, 15, 20, 30, 40, 80, 1000)) {
  samp <- rgamma(size, 2, 1.5)
  hist(samp, main=paste("Gamma Distribution with Sample Size ", size))
}
```

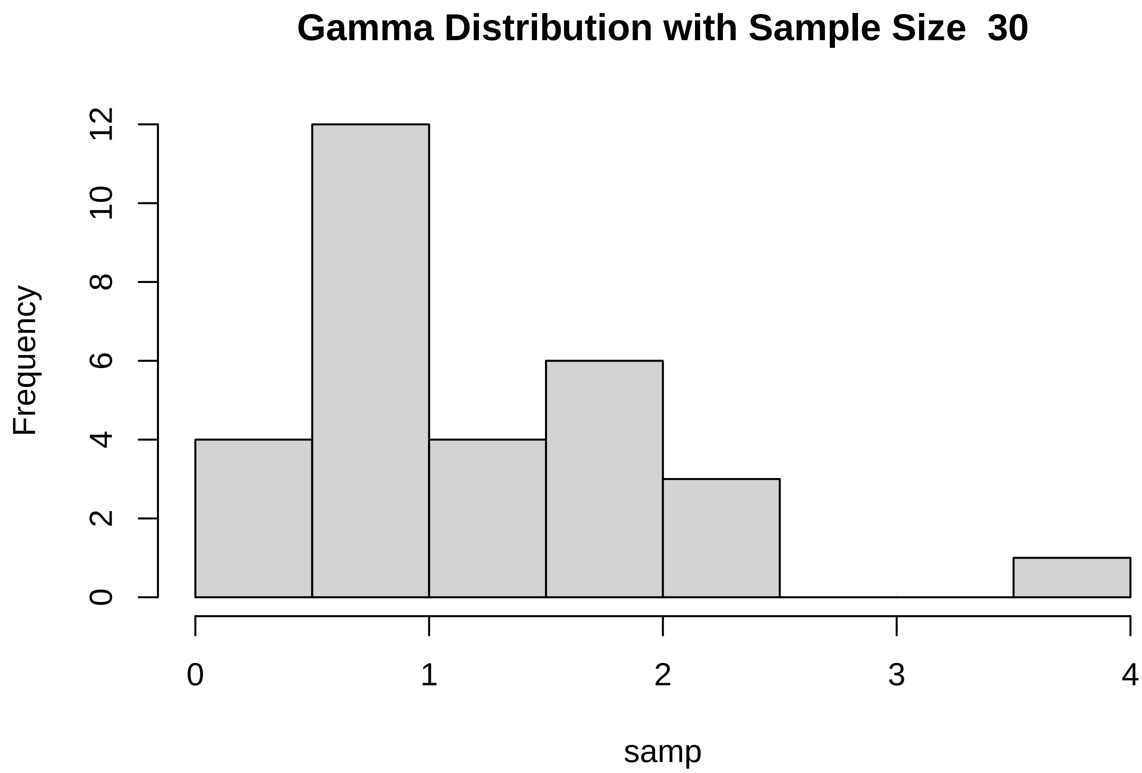

Gamma Distribution with Sample Size 4
Frequency
0.0
0.5
1.0
1.5
2.0
0.6
0.8
1.0
1.2
1.4
1.6
1.8
samp

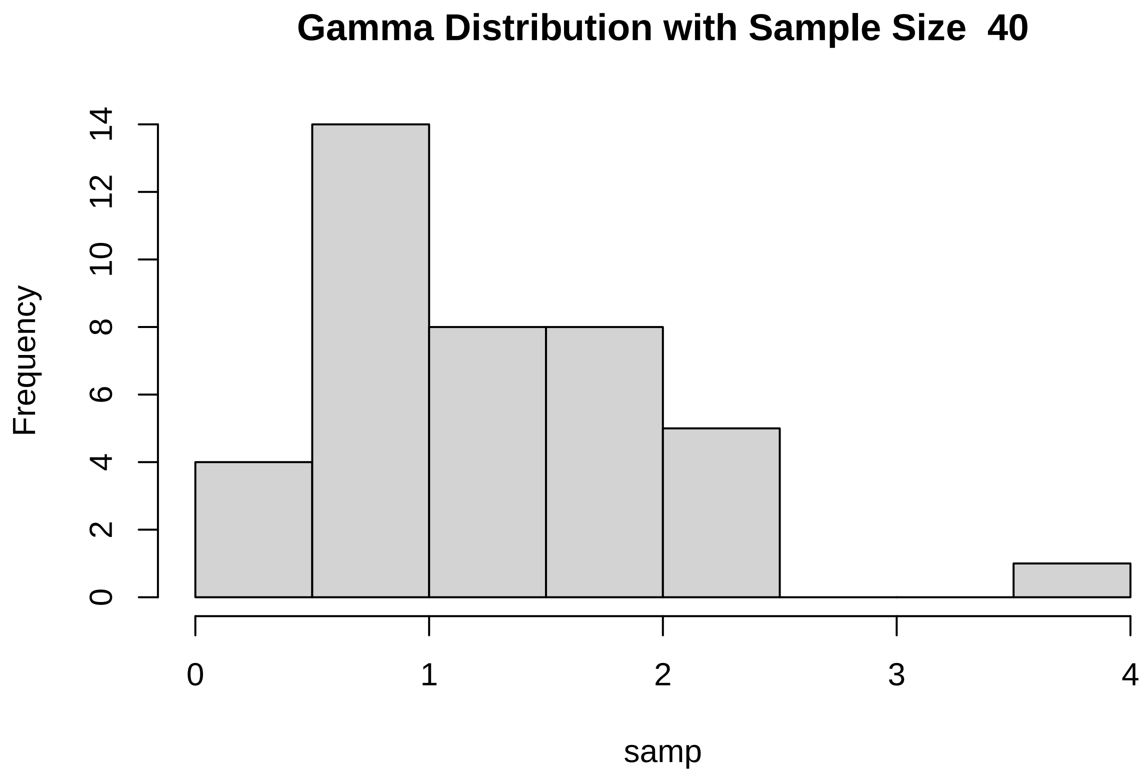

Gamma Distribution with Sample Size 7
Frequency
0
1
2
3
4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
samp

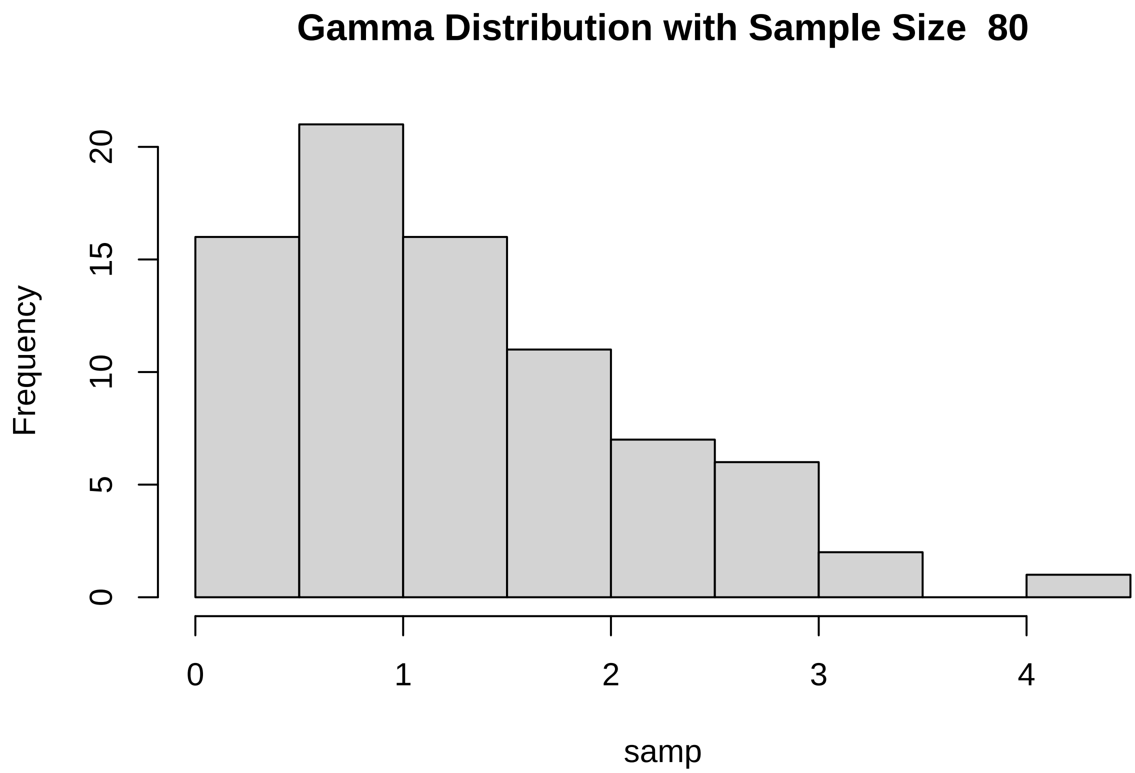

Gamma Distribution with Sample Size 10
Frequency
0
1
2
3
4
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
samp

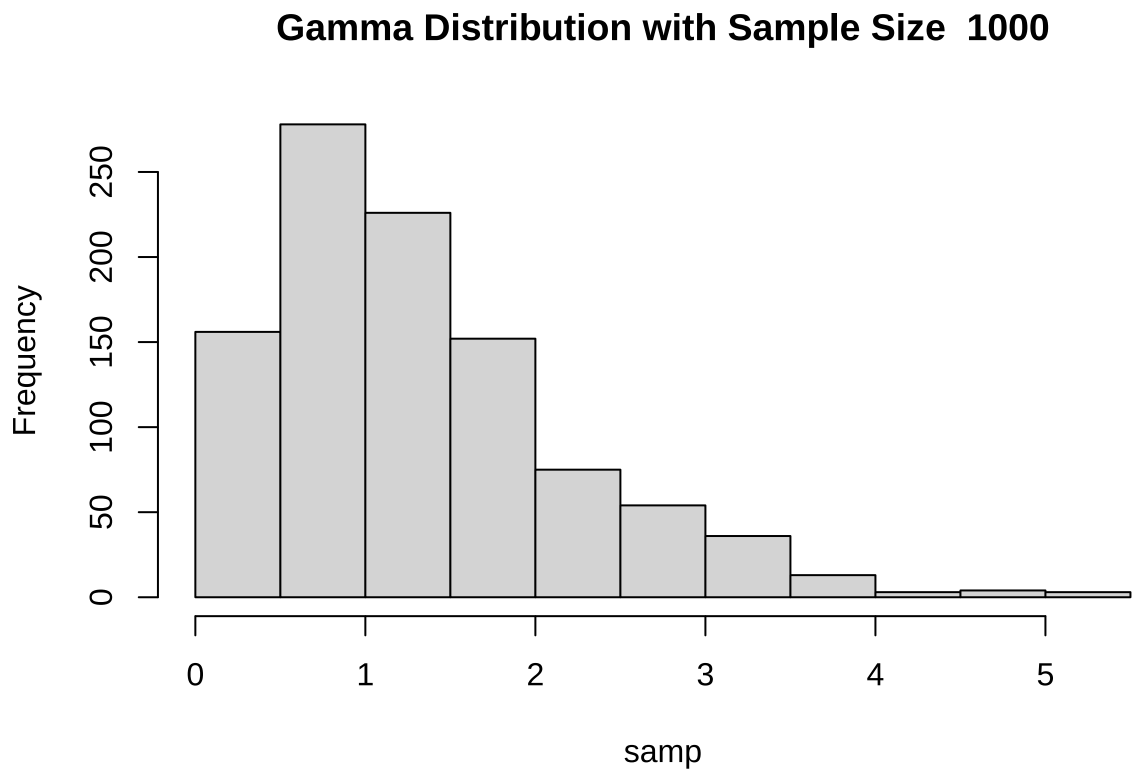

Gamma Distribution with Sample Size 15
Frequency
0
1
2
3
4
5
6
0.0
0.5
1.0
1.5
2.0
2.5
3.0
samp


Gamma Distribution with Sample Size 20
Frequency
0
1
2
3
4
5
0.0
0.5
1.0
1.5
2.0
2.5
samp


Gamma Distribution with Sample Size 30
Frequency
0
2
4
6
8
10
12
0
1
2
3
4
samp


Gamma Distribution with Sample Size 40
Frequency
0
2
4
6
8
10
12
14
0
1
2
3
4
samp


Gamma Distribution with Sample Size 80
Frequency
0
5
10
15
20
0
1
2
3
4
samp



What do you notice? Answer: the higher the sample size, the better is the approximation of the distribution. At size 1000, the approximation looks very close to the gamma distribution.