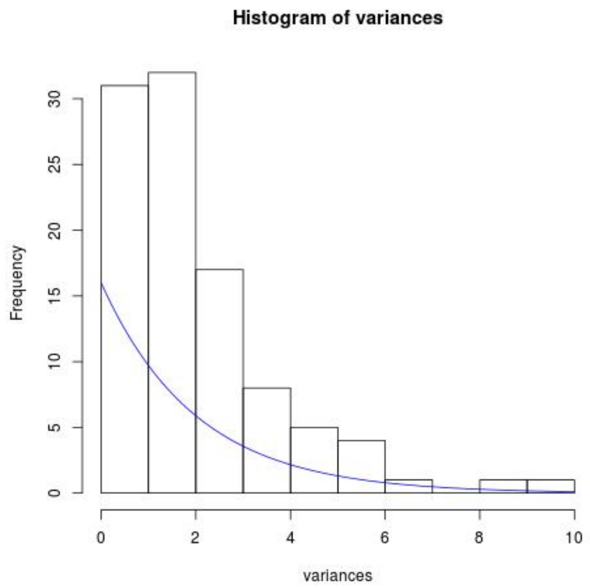
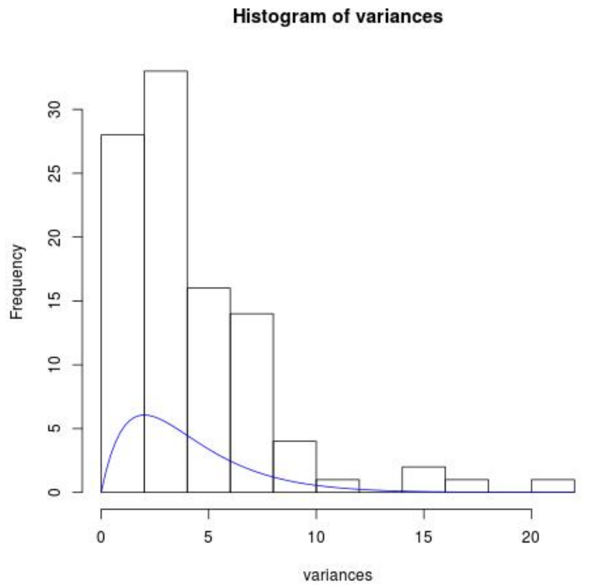
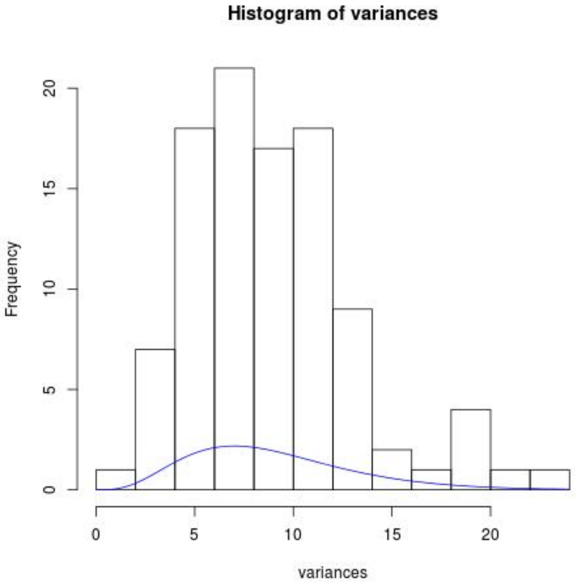
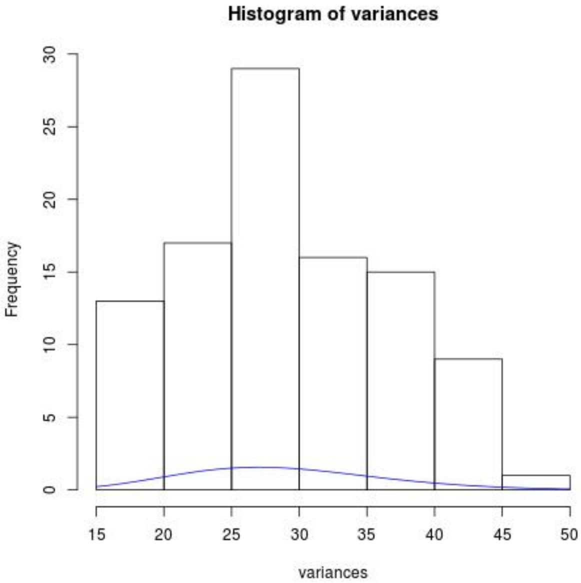
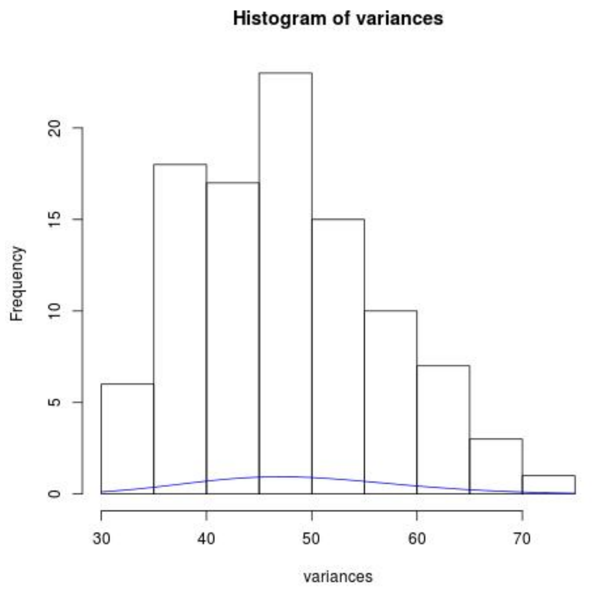
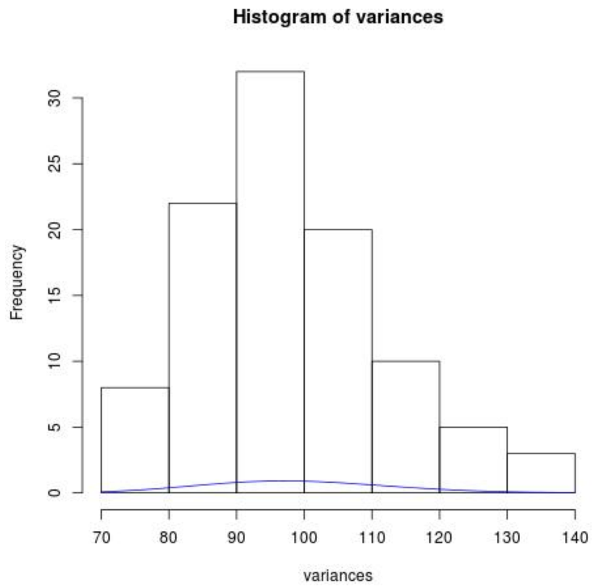
<u>Distribution of Sample Variance</u>

DA-A09

Sample Size (n)	Degrees of freedom (sampleSize-1)	pdf (i.e. histogram) of distribution of sample variance $\frac{(n-1)S^2}{\sigma^2}$ (here σ = 1)
3	2	
5	4	

10	9	Histogram of variances Frequency variances 0 5 10 15 20
30	29	Histogram of variances Frequency variances 15 20 25 30 35 40 45 50

50	49	Histogram of variances Frequency 0 5 10 15 20 30 40 50 60 70 variances
100	99	Histogram of variances Frequency 0 5 10 15 20 25 30 70 80 90 100 110 120 130 140 variances

Observations

1. Distribution of sample variance from a population that is normally distributed is following a **chi - squared** distribution.
2. Degrees of freedom : sample size – 1

3. As the degrees of freedom increases, the Chi Square distribution **approaches a normal distribution**.
   (skew is decreasing as the degrees of freedom is increasing.)
4. The mean of a Chi Square distribution is its degrees of freedom

Let **X** be the sample mean; *n* be the sample size
Now sample variance of population is given by:
$S^2 = \Sigma(Xi - \mathbf{X})^2/(n-1)$

So S^2 is a **sum of *n* independent chi-square random variables**. That's because we have assumed that $X_1, X_2, \ldots, X_n$ are observations of a random sample of size n from the normal distribution $N(\mu, \sigma^2)$.

It follows that

$$\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

-[1]

Since the variance is being 'estimated' for the population, 1 degree of freedom is lost (for each parameter estimated in certain chi – squared random variables)

Ref : https://onlinecourses.science.psu.edu/stat414/node/174 (proof of [1])