

MEASURES OF SKEWNESS AND KURTOSIS *R Programming*

10 of June 2013

Al-Ahmadgaid B. Asaad
alstated@gmail.com

Skewness and kurtosis in R are available in the moments package (to install a package, click [here](#)), and these are:

- Skewness - `skewness`; and,
- Kurtosis - `kurtosis`.

Example 1. Mirra is interested on the elapse time (in minutes) she spends on riding a tricycle from home, at Simandagit, to school, MSU-TCTO, Sanga-Sanga for three weeks (excluding weekends). She obtain the following data: 19.09, 19.55, 17.89, 17.73, 25.15, 27.27, 25.24, 21.05, 21.65, 20.92, 22.61, 15.71, 22.04, 22.60, and 24.25. Compute and interpret the skewness and kurtosis.

```
> time <- c(19.09, 19.55, 17.89, 17.73, 25.15, 27.27, 25.24,
+          21.05, 21.65, 20.92, 22.61, 15.71, 22.04, 22.60, 24.25)
> library(moments)
> skewness(time)
[1] -0.01565162
>
> kurtosis(time)
[1] 2.301051
```

Interpretation: The skewness here is -0.01565162. This value implies that the distribution of the data is slightly skewed to the left or negatively skewed. It is skewed to the left because the computed value is negative, and is slightly, because the value is close to zero. For the kurtosis, we have 2.301051 implying that the distribution of the data is platykurtic, since the computed value is less than 3. Graphical illustration of the data is in Figure 1. Figure 1 confirms the numerical findings above, it is clear that the histogram is slightly skewed to the left, and is platykurtic. Below is the codes of the said figure,

```
> library(ggplot2)
> qplot(time, geom = 'histogram', binwidth = 2) + xlab('Time')
```

Example 2. Simulate 10000 samples from a normal distribution with mean 55, and standard deviation 4.5, then compute and interpret the skewness and kurtosis, and plot the histogram.

```
> #Simulation
> n.sample <- rnorm(n = 10000, mean = 55, sd = 4.5)
>
> #Skewness and Kurtosis
```

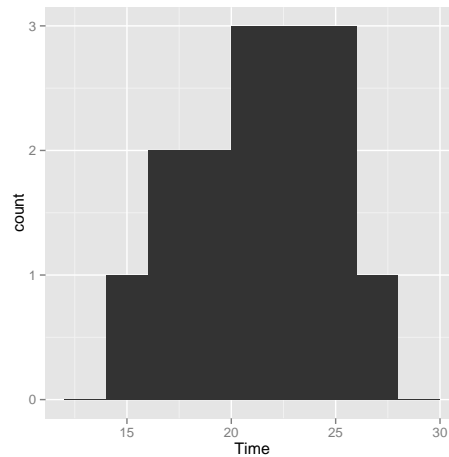


Figure 1. Histogram of the Time Elapsed

```
> library(moments)
> skewness(n.sample)
[1] -0.008525844
>
> kurtosis(n.sample)
[1] 2.96577
>
> #Histogram
> library(ggplot2)
> datasim <- data.frame(n.sample)
> ggplot(datasim, aes(x = n.sample), binwidth = 2) +
>   geom_histogram(aes(y = ..density..), fill = 'red',
>   alpha = 0.5) +
>   geom_density(colour = 'blue') +
>   xlab(expression(bold('Simulated Samples')))) +
>   ylab(expression(bold('Density'))))
```

Interpretation: The skewness of the simulated data is -0.008525844. This concludes that the data are close to bell shape but slightly skewed to the left. The computed kurtosis is 2.96577, which means the data is mesokurtic. Figure 2 is the histogram of the simulated data with empirical PDF.

References

Asaad, Abubakar S. (2011). *Simplified Biostatistics*. Manila: Rex Book Store, Inc..

Labels

Descriptive Statistics, R, Tutorial

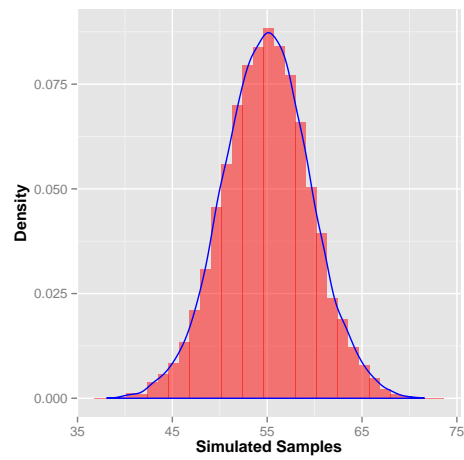


Figure 2. Histogram of the Simulated Data