How can I overlay a normal curve on a histogram in R?

Authored by

stats writer

June 29, 2024

Overlaying a normal curve on a histogram in R refers to the process of superimposing a bell-shaped curve over a graph that displays the distribution of a dataset. This technique is commonly used to visually compare the distribution of the dataset to a normal distribution. In R, this can be achieved by using the "curve" function to plot the normal curve on the same graph as the histogram. This allows for a better understanding of the data and can help in identifying any deviations from a normal distribution. The overlaying of a normal curve on a histogram is a useful tool for data analysis and presentation, particularly in the field of statistics.

Overlay Normal Curve on Histogram in R (2 Examples)

Often you may want to overlay a normal curve on a histogram in R.

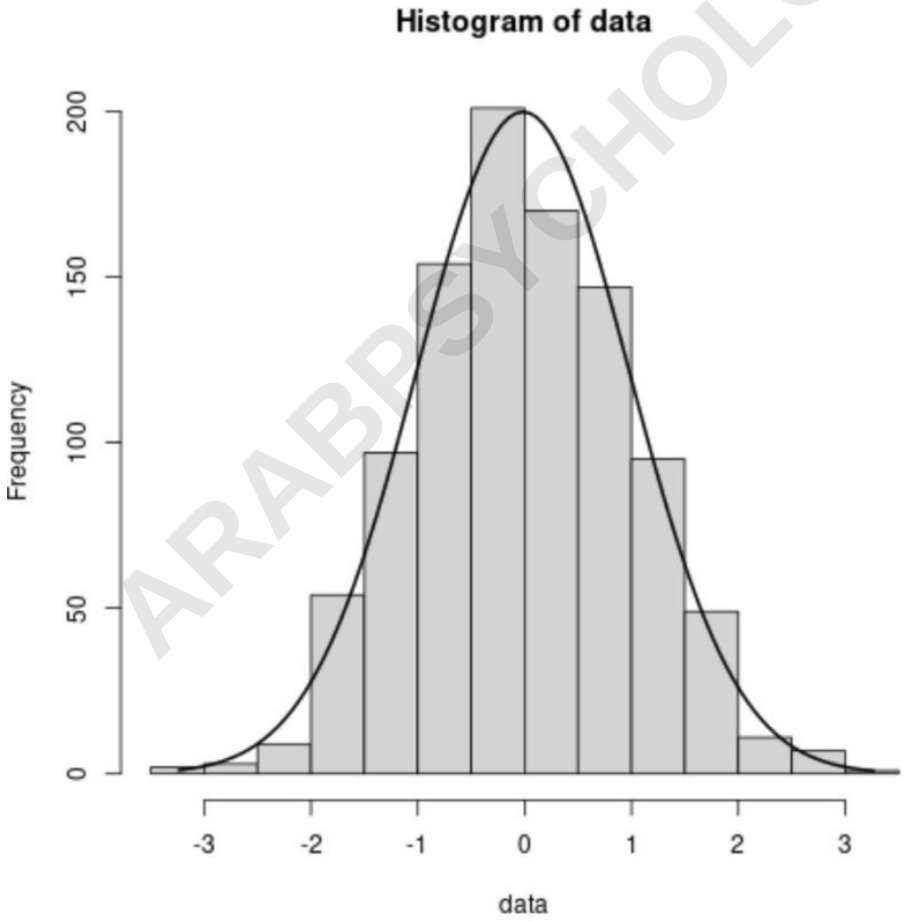
The following examples show how to do so in base R and in .

Example 1: Overlay Normal Curve on Histogram in Base R

We can use the following code to create a histogram in base R and overlay a normal curve on the histogram:

```
#make this example reproducible
set.seed(0)

#define data
data <- rnorm(1000)

#create histogram
hist_data <- hist(data)
```

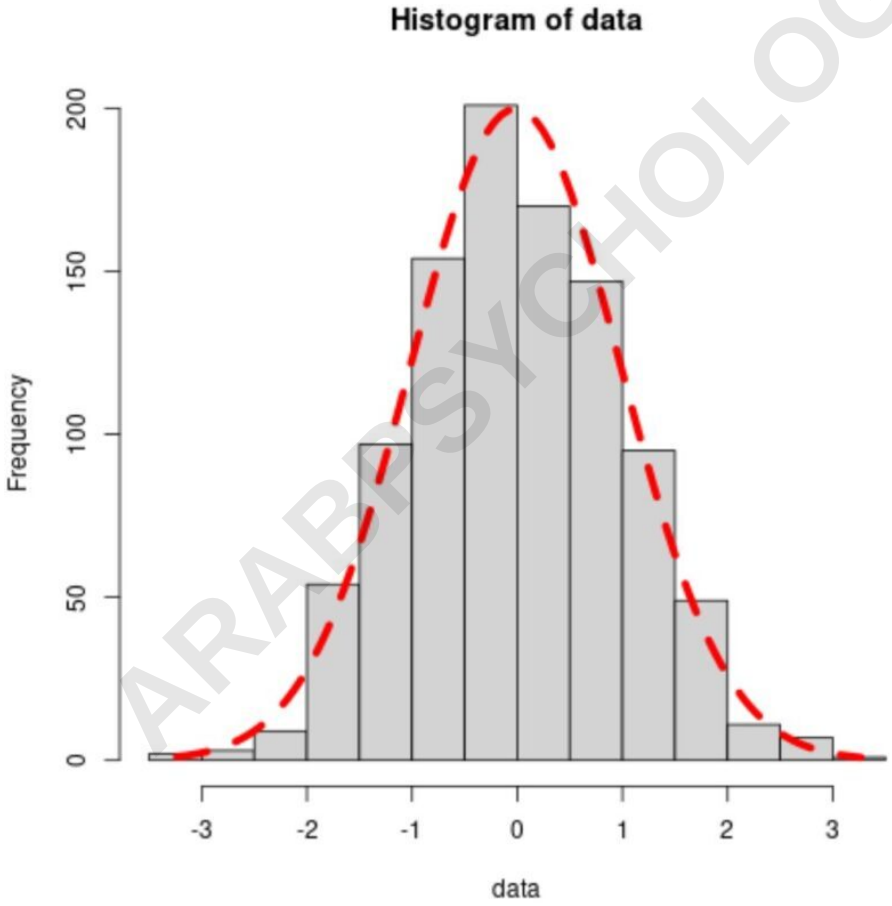
```
#define x and y values to use for normal curve
x_values <- seq(min(data), max(data), length = 100)
y_values <- dnorm(x_values, mean = mean(data), sd = sd(data))
y_values <- y_values * diff(hist_data$mids) * length(data)

#overlay normal curve on histogram
lines(x_values, y_values, lwd = 2)
```



The black curve in the plot represents the normal curve.

Feel free to use the col, lwd, and lty arguments to modify the color, line width, and type of the line, respectively:

```
#overlay normal curve with custom aesthetics
lines(x_values, y_values, col='red', lwd=5, lty='dashed')
```

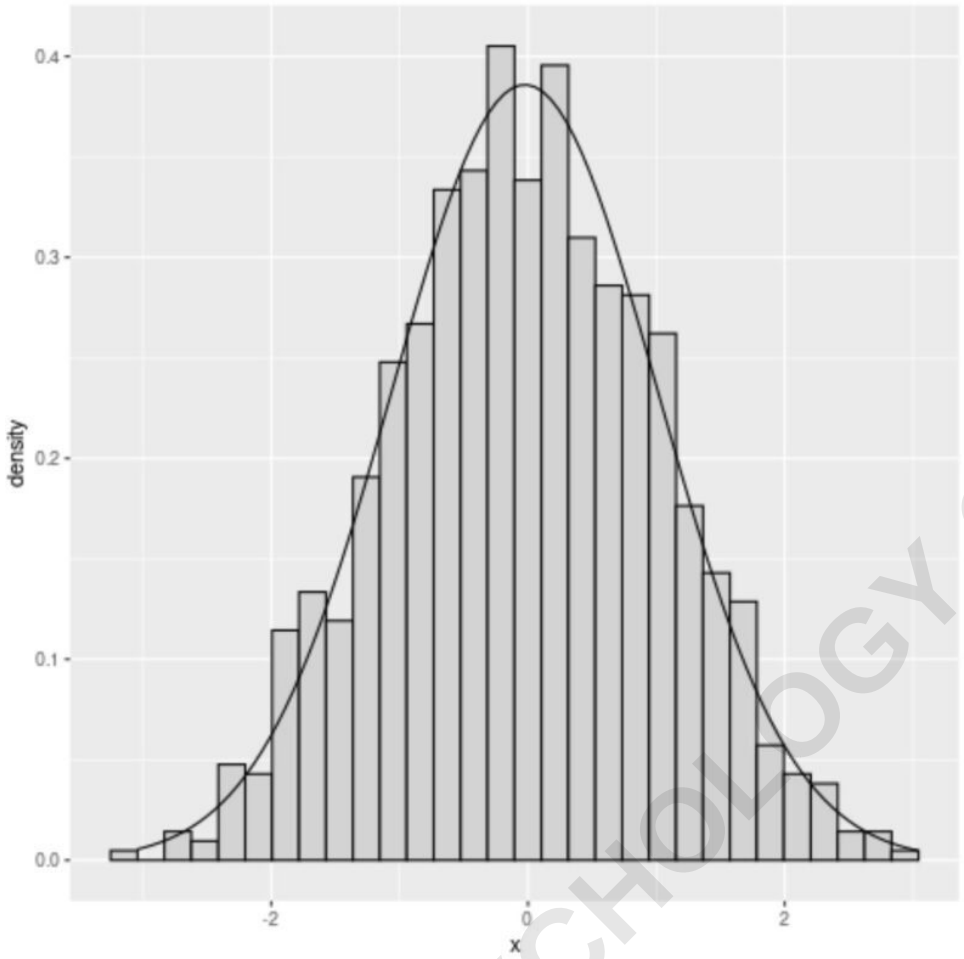


Example 2: Overlay Normal Curve on Histogram in ggplot2

We can use the following code to create a histogram in

ggplot2 and overlay a normal curve on the histogram:

```
library(ggplot2)

#make this example reproducible
set.seed(0)

#define data
data <- data.frame(x=rnorm(1000))

#create histogram and overlay normal curve
ggplot(data, aes(x)) +
geom_histogram(aes(y = ..density..), fill='lightgray',
col='black') +
stat_function(fun = dnorm, args =
list(mean=mean(data$x), sd=sd(data$x)))
```

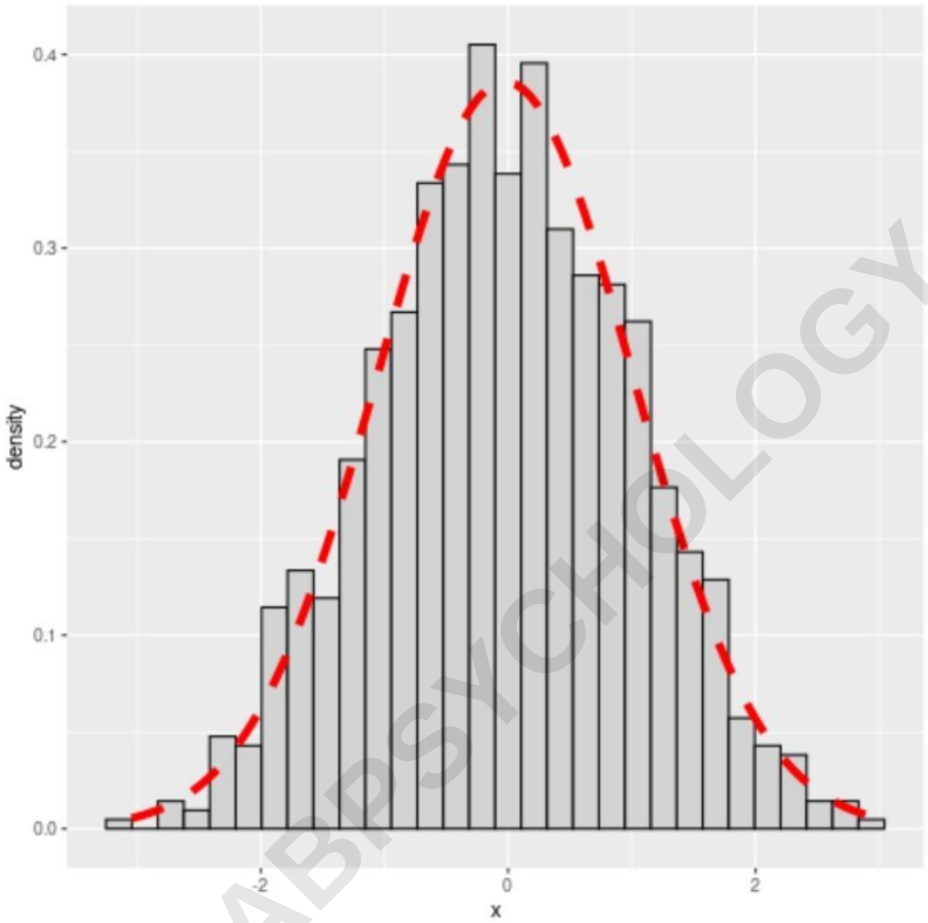


The black curve in the plot represents the normal curve.

Feel free to use the col, lwd, and lty arguments to modify the color, line width, and type of the line, respectively:

```
#overlay normal curve with custom aesthetics
ggplot(data, aes(x)) +
geom_histogram(aes(y = ..density..), fill='lightgray', col='black') +
```

```
stat_function(fun = dnorm, args = list(mean=mean(data$x), sd=sd(data$x)), col='red', lwd=2, lty='dashed'))
```



Note: You can find the complete documentation for stat_function .

Additional Resources