Normal Distribution and Empirical Rule

Student's Name

Institution

Normal Distribution and Empirical Rule

Normal distribution and empirical rule is most often discussed together due to the nature of the relationship they share. According to Rees (2018), in statistics, the normal distribution is one of the most significant probability distributions since it has the capacity to fit as many natural phenomena as possible. In other words, it is a probability function that can illustrate how differently the variable values are distributed within a given dataset. The most common natural phenomena represented best by the former include blood pressure, size of things resulting from machine production, marks on a specific test, IQ scores, and errors in measurements. All of the latter are occasionally represented in a normal distribution curve, which is calculated from the normal equation, especially for continuous probability distributions. The normal equation is as presented below:

$$Y = \{1/ [\sigma * sqrt(2\pi)]\} * e^{-(x - \mu)2/2\sigma 2}$$

Where σ represents the standard deviation (SD), e is considered an approximation of 2.718, whereas π indicates 3.14. μ is the data, meanwhile X represents a normal but random variable. It is, however, important to note that the equation above is considered the probability density function for any normal distribution of a given dataset (Rees, 2018). The primary factors recorded in the equation is the mean and the standard deviation. These are the functions that determine the graphs resulting from the distribution of data (D’Agostino, 2017). Examples of a normal curve are presented below. Both illustrate a certain degree of standard deviation. Consequently, the common characteristic that can be picked from them is that they are bell-shaped as well as symmetrical.

Figure 1: Normal Distribution Curve with a smaller SD

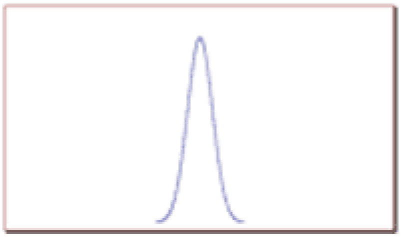
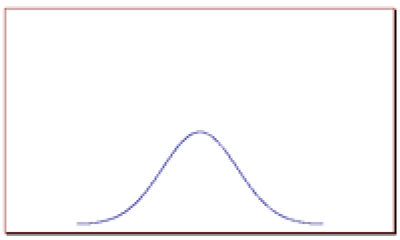


Figure 2: Normal Distribution Curve with a larger SD

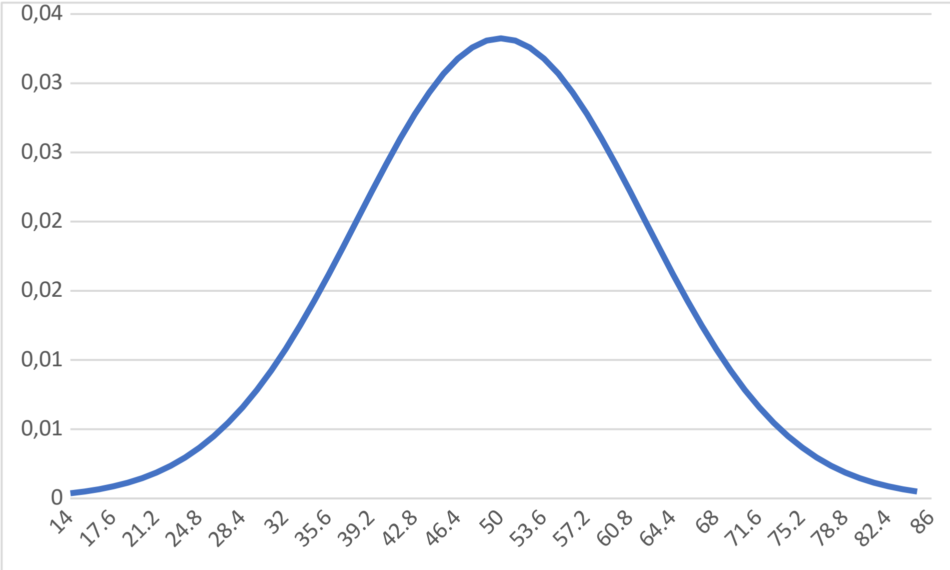


The diagrams above do not illustrate actual data and are used purposely to demonstrate the impact of SD on the standard curve. The table below represents data collected from a test marks whose mean was 50 and had a standard deviation of 12.

	Probability Density
14	0.000369
15.2	0.000496
16.4	0.00066
17.6	0.000868
18.8	0.001132
20	0.001461
21.2	0.001866
22.4	0.002361
23.6	0.002956
24.8	0.003665
26	0.004499
27.2	0.005468
28.4	0.006579
29.6	0.007837
30.8	0.009243
32	0.010793
33.2	0.012477
34.4	0.014281

35.6	0.016182
36.8	0.018154
38	0.020164
39.2	0.022174
40.4	0.024141
41.6	0.026021
42.8	0.027769
44	0.029339
45.2	0.030689
46.4	0.031782
47.6	0.032587
48.8	0.033079
50	0.033245
51.2	0.033079
52.4	0.032587
53.6	0.031782
54.8	0.030689
56	0.029339
57.2	0.027769
58.4	0.026021
59.6	0.024141
60.8	0.022174
62	0.020164
63.2	0.018154
64.4	0.016182
65.6	0.014281
66.8	0.012477
68	0.010793
69.2	0.009243
70.4	0.007837
71.6	0.006579
72.8	0.005468
74	0.004499
75.2	0.003665
76.4	0.002956
77.6	0.002361
78.8	0.001866
80	0.001461
81.2	0.001132
82.4	0.000868
83.6	0.00066
84.8	0.000496

Figure 3: Normal Curve for Test Marks (Mean = 50, SD = 12)



The relationship between normal distribution and the empirical rule is that the former is a statement about the other one. The rule states that on a normal distribution curve where mean = median = mode, about 68 percent of the represented data will lie within one SD of the mean. At the same time, 95 percent of the same will lie within two SD of the mean, whereas 99.7 will lie with the three SD of the mean (Johnson & Bhattacharyya, 2018). In other words, the empirical rule is used to find out where specific data is to come up with final reports on the same.

For instance, from the data used in the construction of the normal curve, the following can be deduced from it using the empirical rule:

From the mean of 50 and SD of 12, about 68% of students will record test marks in the intervals below:

50 +/- 1(12) = [38, 62].

Alternatively, 95% will score within the following intervals:

50 +/- 2(12) = [26, 74].

This means that 99.7% will record marks which fall within the intervals calculated below:

50 +/- 3(12) = [14, 86]

References

D'Agostino, R. B. (2017). Tests for the normal distribution. In *Goodness-of-fit-techniques* (pp. 367-420). Routledge.

Johnson, R. A., & Bhattacharyya, G. K. (2018). *Statistics: principles and methods*. Wiley.

Rees, D. G. (2018). *Essential statistics*. Chapman and Hall/CRC.