The normal distribution application: systolic blood pressure

Svetlana Svistova[1*] and *Tatiana* Nikitina[2]

[1]MIREA, 78 Vernadsky Avenue, 119454, Moscow, Russia
[2]Moscow Economic School, 1a Zaitsevo, 143390, Russia

Abstract. The present study established that sample data for 50 participants were approximately normally distributed using visual analysis. The distributions categorised by respective levels of gender, smoking condition, and across levels of age group showed approximate normality with varying standard deviations. From our research we conduct that smokers have higher probability of high blood pressure and hypertension, males had a higher systolic blood pressure than females, however, smokers had higher blood pressure than non-smokers.

1 Introduction

We decided to do this research to see if smoking affects systolic blood pressure with the use of The Normal Distribution. Three pieces of research caught our attention, one of them was "Factors influencing blood pressure classification for adults: Gender differences" [1], another being "Mechanism of injury, Glasgow Coma Scale, age, and systolic blood pressure: a new trauma scoring system to predict mortality in trauma patients" [2] and final being "Using decomposition analysis to identify modifiable racial disparities in the distribution of blood pressure in the United States" [3]. The Normal Distribution is considered one of the most fundamental statistical distributions. It can describe several biological variables such as heights of individuals, bone densities, and intelligence quotient, to name a few [4]. Simple observable factors such as heights of individuals and consequently their weights, a coin toss many times (say 100 or more), when graphed with the frequencies on the y-axis and the scores on the x-axis, would form a rough bell-shape.

2 Methods

The sampling technic was to invite University staff and family members of students to participate in our research. As they were more likely to respond, this meant that our sampling technic was convenience sampling. Participants were invited through emails to participate in the study; this leads to a coverage error as samples should be sufficiently more significant and unbiased. Another coverage error was the age group 45 or older; due to their age, they have other medical issues that have no connection with smoking but splitting the group 45 or older into multiple groups would have given us a more in-depth answer to our research.

* Corresponding author: svistova.s@mail.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

Seventy participants were invited, but only fifty showed up; this is a non-response error as 2/7 of all selected participants did not turn up. The present participants were explained what was needed from them and given informed consent to sign, showing agreement to participate in the study. Information collected would not be shared with any third parties, and neither will the information be used out of the scope of the present research. The participants were asked to indicate whether they had smoked within the last three months (yes or no), to indicate their ages and gender, as the essential demographic information of the participants' systolic blood pressure was then taken using a sphygmomanometer.

2.1 Data and analysis

The fifty participants' data was collected.

The histogram of systolic blood pressure appeared to be approximately normally distributed except for the right-hand skewness. The average blood pressure (systolic) was 123.8 mmHg, as seen in figure 1. The highest frequency was in the bracket between 120 mmHg and 128 mmHg, since these two limits had the highest bars in the histogram.

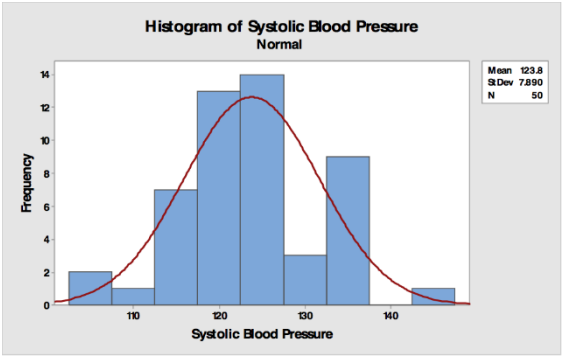

Fig. 1. Histogram of systolic blood pressure.

The superimposed histograms of systolic blood pressure categorised by gender showed that the distribution for males was less skewed and resembled a normally distributed variable than the distribution of systolic blood pressure for the females. The females' systolic blood pressure was skewed to the left and appeared flattered at the peak than the systolic blood pressure for males. The mean for males was also more extensive than the mean for females, and the spread of the systolic blood pressure for males being narrower than for females see figure 2.

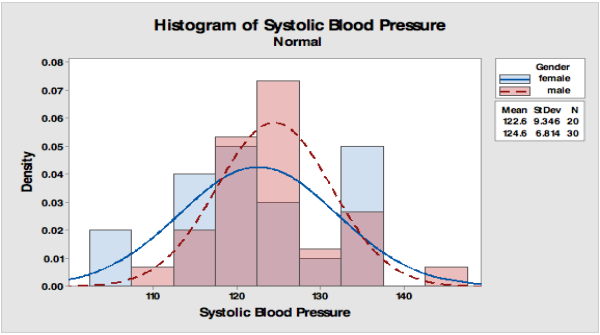

Fig. 2. Systolic blood pressure by gender.

When categorised according to smoking conditions, the distribution of smokers is narrower than the distribution of non-smokers. The peak of the distribution of systolic blood pressure for the non-smokers was also higher than the peak of the distribution of smokers. The average systolic blood pressure for smokers was also higher than the average systolic blood pressure for non-smokers. The analysis showed that for both smokers and non-smokers, the distribution of systolic blood pressure was approximately distributed evenly. The mean for smokers was higher than the average systolic blood pressure for non-smokers.

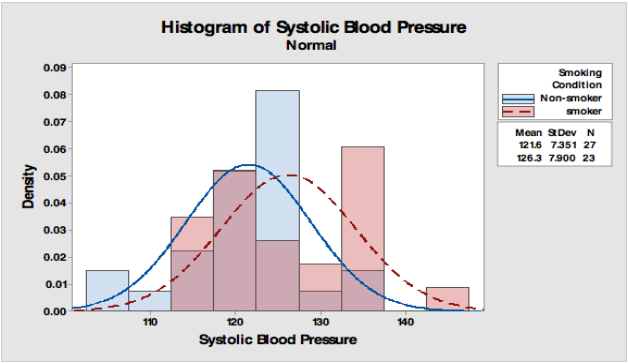

Fig. 3. Systolic blood pressure by smoking condition.

Blood pressure was finally categorised based on the age group of the participants, using a box and whisker plot. Figure 4 shows that as the age increased, the median systolic blood pressure also increased. Individuals belonging to the age group 18-26 had the lowest median systolic blood pressure. However, the distribution in this group was skewed upwards than downwards, as seen in the long upper tail than the lower tail. The lower tail for the age group 27-35 was longer than the right tail, notwithstanding, with a potential outlier variable at the low end. However, the median in this group did not appear to be any higher than the age group 18-26 years. Anyhow, the group of 45 years or older had a longer right tail than the left tail, also with the highest median. Overall, there was an outlier point in the systolic blood pressure, with the data being generally symmetrical about the median.

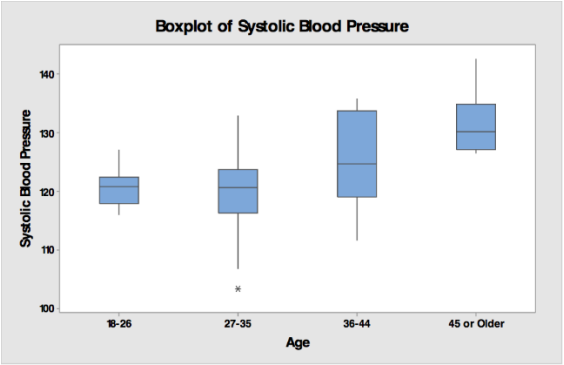

Fig. 4. Box and whisker plot of systolic blood pressure by age-group.

2.2 Summary of Statistics

The graphical analysis gave a description of the visually observable features of the data. Summary statistics provide quantitative statistical descriptive values and hence more objective in approach. First, the systolic blood pressure based on gender is analysed and

provided as in table 2 below. The sample means that systolic blood pressure for females was distributed with mean = 122.6 mmHg, standard deviation = 9.4 mmHg, while mean for males was 124.6 mmHg, standard deviation = 6.8 mmHg. The lower and upper quartile for the female and male average systolic blood pressure were respectively 117.0 mmHg, 120.1 mmHg and 131.0 mmHg, 127.2 mmHg. See table 1.

Table 1. Summary statistics by Age.

Gender	Mean (mmHg)	Std. Dev. (mmHg)	Lower Quartile	Upper Quartile	N
Female	122.6	9.4	117.0	131.0	20
Male	124.6	6.8	120.1	127.2	30
Total	123.8	7.9	119.3	127.8	50

Non-smokers had mean = 121.6 mmHg, standard deviation = 7.4 mmHg with lower quartile being lower for non-smokers 118.2 mmHg, and upper quartile being 125.8 mmHg. While Smokers had a higher average systolic blood pressure than non-smokers: mean = =126.3 mmHg, standard deviation = 7.9 mmHg, and have 120.0 mmHg as lower quartile and upper quartile being significantly higher for smokers than non-smokers with 133.2 mmHg. 27 (54%) of the participants were smokers, while 23 (46%) were non-smokers. Table 2 gives the summary statistics for systolic blood pressure for smoking conditions.

Table 2. Systolic blood pressure for smoking condition.

Smoking Condition	Mean (mmHg)	Std. Dev.(mmHg)	Lower Quartile	Upper Quartile	N
Non-smoker	121.6	7.4	118.2	125.8	27
Smoker	126.3	7.9	120	133.2	23
Total	123.8	7.9	119.3	127.8	50

Across ages, the highest average systolic blood pressure by smoking condition was for the age group 45 years or older with the mean = 131.7 mmHg, standard deviation = 5.1 mmHg, and a lower and upper quartile being 127.0 mmHg, 134.8 mmHg respectively. The age group 27-35 years had recorded systolic blood pressure of mean = 116.9 mmHg, standard deviation = 7.7 mmHg with a lower and upper quartile being 116.9 mmHg, 122.9 mmHg. See table 3 below for the summary statistics.

Table 3. Summary statistics by age group.

Age	Mean (mmHg)	Std. Dev (mmHg)	Lower Quartile	Upper Quartile	N
18-26	120.5	3.1	118.0	122.3	12
27-35	119.3	7.7	116.9	122.9	14
36-44	124.9	8.2	118.9	133.8	13
45>	131.7	5.1	127.0	134.8	11
Total	123.8	7.9	119.3	127.8	50

3 Results

The overall distribution of systolic blood pressure was found to be approximately normally distributed with the sample mean of 123.8 mmHg, standard deviation = 7.9 mmHg. The American Heart Association (AHA) defined in 2017 new guidelines for defining high blood pressure for different categories of individuals. The blood pressure readings were defined as high (HBP) if the systolic blood pressure is 130-139 mmHg or the diastolic blood pressure is 80-89 mmHg [5]. AHA guidelines were that High Blood Pressure (HBP) was between 130-139 mmHg, hypertension was above 140 mmHg, while the hypertensive crisis was anything more than 180 mmHg. Further analysis was taken as the data was calculated and compared for the probabilities of these conditions for all participants, smokers and non-smokers.

3.1 All participants

X ~ N (123.8, 7.9), where X is systolic blood pressure for the sample of 50 participants.

i. Probability of having High Blood Pressure

$$\mathrm{P(HBP)} = \mathrm{P}(130 < \mathrm{X} < 139) = \mathrm{P}\left(\frac{130 - 123.8}{7.9} < \mathrm{Z} < \frac{139 - 123.8}{7.9}\right) = 0.20$$

Based on the sampled data, 20% of the adults have a probability of having high pressure.

ii. Hypertension is defined as systolic blood pressure above 140 mmHg or higher. The proportion of the population with hypertension is given as:

$$\mathrm{P(Hypertension)} = \mathrm{P}(\mathrm{X} > 140) = \mathrm{P}\left(\mathrm{Z} > \frac{140 - 123.8}{7.9}\right) = 0.0202$$

The proportion of the population with a probability of hypertension is 2.02%, according to the sampled data collected from 50 participants.

iii. Hypertensive crisis is systolic blood pressure above 180 mmHg. Based on the sample, the proportion of the population with the Hypertensive crisis is;

$$\mathrm{P}(\mathrm{X} > 180) = \mathrm{P}\left(\mathrm{Z} > \frac{180 - 123.8}{7.9}\right) = \mathrm{P}(\mathrm{Z} > 7.111) = 0.000$$

Hypertensive crisis cases are rare, with the probability of occurrence is approximately 0.000 for 50 participants that were sampled.

3.2 Smokers

X ~ N (126.3, 7.9), where X is systolic blood pressure for smokers.

i. Probability of having High Blood Pressure (HBP) for smokers

$$\mathbf{P(HBP) = P(130 < X < 139) = P\left(\frac{130 - 126.3}{7.9} < Z < \frac{139 - 126.3}{7.9}\right) = 0.266}$$

Smokers have a 6.6% more likely probability of having High Blood Pressure than among all participants.

ii. Probability of having High Blood Pressure for smokers

$$\mathbf{P(Hypertension) = P(X > 140) = P\left(Z > \frac{140 - 126.3}{7.9}\right) = 0.0414}$$

Smokers have a two times higher probability of having Hypertension than among all participants.

iii. Probability of having Hypertensive crisis for smokers

$$\mathbf{P(X > 180) = P\left(Z > \frac{180 - 126.3}{7.9}\right) = P(Z > 6.797) = 0.000}$$

Hypertensive crisis is rare; even smokers have a probability of approximately 0.000 for 27 participants that were sampled.

3.3 Non-smokers

X ~ N (121.6, 7.35), where X is systolic blood pressure for non-smokers.

i. Probability of having High Blood Pressure (HBP) for non-smokers

$$\mathbf{P(HBP) = P(130 < X < 139) = P\left(\frac{130 - 121.6}{7.35} < Z < \frac{139 - 121.6}{7.35}\right) = 0.118}$$

Non-smokers have a two times smaller probability of having High Blood Pressure than smokers.

ii. Probability of having hypertension for non-smokers

$$\mathbf{P(Hypertension) = P(X > 140) = P\left(Z > \frac{140 - 121.6}{7.35}\right) = 0.00615}$$

Non-smokers have almost seven times more negligible probability of having hypertension than smokers.

iii. Probability of having Hypertensive crisis for non-smokers

$$\mathbf{P(X > 180) = P\left(Z > \frac{180 - 121.6}{7.35}\right) = P(Z > 7.946) = 0.000}$$

Non-smokers have the same probability of having a hypertensive crisis as Smokers as both have a probability of occurrence is approximately 0.000.

4 Discussion

The applications of The Normal Distribution to the real world provide a solid basis for analysis and research, especially in statistical inference. The present study established that sample data for 50 participants were approximately normally distributed using visual analysis. The distributions categorised by respective levels of gender, smoking condition, and across levels of age group showed approximate normality with varying standard deviations.

While the study provides some empirical evidence for the normality of systolic blood pressure, the limitation of using a convenience sample instead of a probabilistic one limits the validity. However, in investigating systolic blood pressure and the normal distribution, future studies should consider other variables such as body mass index, which has been

studied to affect blood pressure. Our research also had a tiny scale to analyse all the data. All participants were living in the same country and were all working in the same city. For accurate and more in-depth research, these factors need to be considered to make this fundamental research more accurate.

From our research we conduct that smokers have higher probability of high blood pressure and hypertension, males had a higher systolic blood pressure than females, however, smokers had higher blood pressure than non-smokers. This result resonated with the findings of Linneberg on the effects of smoking on systolic blood pressure [6]. The findings also showed that individuals at least 45 years of age had higher systolic blood pressure than individuals in the age group 18-26 years and 27-35 years.

References

1. Choi, Jin Yi, Eun Kyoung Yun, E.J. Yeun, E.S. Jeong, International journal of nursing practice **25(3)**, e12706 (2019)
2. Baghi, Iraj, L. Shokrgozar, M.R. Herfatkar, K.N. Ehsan, Z.M. Amiri, Trauma monthly **20(3)** (2015)
3. Basu, Sanjay, A. Hong, A. Siddiqi, American journal of epidemiology **182(4)**, 345-353 (2015)
4. Ahsanullah, Mohammad, B.MG. Kibria, M. Shakil, *"Normal Distribution" Normal and Student' Standard Distributions and Their Applications* (Atlantis Press, Paris, 2014)
5. *New ACC/AHA High Blood Pressure Guidelines Lower Definition of Hypertension* (American college of cardiology, 2017*)* https://www.acc.org/latest-in-cardiology/articles/2017/11/08/11/47/mon-5pm-bp-guideline-aha-2017
6. Linneberg, Allan, R.K. Jacobsen, Tea Skaaby, A.E. Taylor, M.E. Fluharty et al., Circulation: Cardiovascular Genetics **8(6)**, 832-841 (2015)