Quest Journals
Journal of Research in Applied Mathematics
Volume 8 ~ Issue 3 (2022) pp: 01-06
ISSN(Online) : 2394-0743 ISSN (Print): 2394-0735
www.questjournals.org

Research Paper

Determining the Conformity to Normal Distribution of the Number of Punches in Boxing Competitions

Şenol Çelik
Bingöl University Faculty of Agriculture Biometry and Genetic Department, Bingöl-Turkey

Abstract
This study examined the statistical distribution of boxers' punches per round in heavyweight boxing competitions occurring between 2016-2021 at various times in various places throughout the world. The punches of a total of 16 competing boxers in 12 rounds of 8 randomly chosen matches were counted. A total of 192 recordings were watched, and the average punch per round was found to be 49.58, with a standard deviation of 17.78. The data were found to be normally distributed after normality tests were performed. The probabilities of the punch counts between 30 and 70 were calculated using the average and standard deviation information. The probability density function of the normal distribution was found to be adequate for calculating the probability of punch counts per round.
Keywords: *Normal distribution, boxing, punch, round.*

R*eceived 01 Mar, 2022;* ***R****evised 10 Mar, 2022;* ***A****ccepted 13 Mar, 2022 © The author(s) 2022.*
P*ublished with open access at* ***www.questjournals.org***

I. Introduction

Statistics is the most widely applied of all mathematical disciplines and at the center of statistics lies the normal distribution, known to millions of people as the bell curve, or the bell-shaped curve. The origins of the mathematical theory of probability are justly attributed to the famous correspondence between Fermat and Pascal, which was instigated in 1654 by the queries of the gambling Chevalier de Mer´e [1]. Among the various types of problems they considered were binomial distributions, which today would be described by such sums as:

$$P(x) = \sum_{x=1}^{j} \binom{n}{x} p^x (1-p)^{n-x}$$

As the binomial examples Fermat and Pascal worked out involved only small values of n, they were not concerned with the computational challenge presented by the evaluation of general sums of this type [2].

The origin of normal distribution can be traced to a French mathematician Abraham de Moivre. He had scientific interest in gambling and often acted as a consultant to gamblers to determine probabilities. De Moivre allegedly was studying the probability distribution of coin flips. He was trying to come up with a mathematical expression such as finding a probability of 60 or more tails out of one hundred coin flips. As an answer to this question he derived a bell shaped distribution which is commonly referred as the normal curve. This was a crucial observation as a large number of phenomena follow approximately normal distribution. For example, such variables (phenomena) as height, weight and strength are characterized with normal distribution. That's why it is possible to determine one's weight or height standing compared to others using z score tables. A Belgian astronomer - Lambert Quetelet, was the first one who noticed the link between weight and height distribution and the normal curve [3].

In 1808, Robert Adrain, an American mathematician, debated the validity of the normal distribution, expounding on distributions of measurement errors. His discoveries led to further work in proving Adrien-Marie Legendre's method of least squares. In 1809, without knowledge of Adrain's work, Gauss published his Theory of Celestial Movement. This work presented substantial contributions to the statistics field, including the method of least squares, the maximum likelihood parameter estimation, and the normal distribution. The significance of these contributions is possibly why Gauss is given credit over Adrain in regard to the normal

distribution. In use from 1991 to 2001, the German 10 DM banknote displayed a portrait of Gauss and a graphical display of the normal density function [4].
The aim of this study is to determine the statistical distribution of the number of punches thrown per round in boxing matches and to determine the probabilities of the number of punches according to the appropriate distribution.

II. Material and Method

Material

The material of the study consisted of the punch counts in 12 rounds competed by the boxers presented in Table 1. The punch counts of boxers who had highly successful matches in the world in their high-level matches were recorded by scanning the internet. The punch counts of 16 boxers who played in 8 matches over 12 rounds, were recorded. A total of 192 data were collected.

Table 1. The matches played by the boxers and the number of punches thrown

No and Ref.	Round	1	2	3	4	5	6	7	8	9	10	11	12
1 [5]	Alvarez	36	48	41	58	43	56	63	51	61	56	45	64
	Golovkin	47	54	65	87	67	68	72	82	80	83	75	89
2 [6]	Joshua	21	44	40	39	63	49	68	67	52	63	60	75
	Usyk	30	42	42	34	48	34	53	39	43	51	45	68
3 [7]	Hurd	19	46	46	47	76	90	67	74	75	78	88	94
	Williams	29	72	51	66	60	55	39	69	46	61	64	75
4 [8]	Charlo	18	39	42	53	36	45	46	58	49	49	40	58
	Castano	30	23	40	60	58	55	66	48	54	41	46	65
5 [9]	Haney	48	48	49	41	45	52	58	43	48	49	62	71
	Diaz	11	44	45	46	49	62	58	51	57	35	46	69
6 [10]	Pacquiao	56	60	69	65	76	61	62	72	63	75	78	78
	Ugas	29	45	30	27	34	33	42	40	34	34	32	25
7 [11]	Kovalev	34	49	27	37	33	36	29	38	46	58	40	47
	Ward	20	16	22	25	24	25	32	32	38	35	26	42
8 [12]	Broner	12	20	25	25	24	32	35	36	38	28	29	34
	Santiago	17	29	53	68	56	82	54	63	57	65	58	95

Ref.: References (5-12)

Method

Normal (Gaussian) Distribution

The normal distribution is by far the most important probability distribution. One of the main reasons for that is the Central Limit Theorem (CLT). The importance of this result comes from the fact that many random variables in real life can be expressed as the sum of a large number of random variables and, by the CLT, it can be argued that distribution of the sum should be normal. The CLT is one of the most important results in probability [13].
The normal distribution is defined by the following probability density function, where μ is the population mean and σ^2 is the variance [14].

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

If random variable X follows the normal distribution, then we write:

$$X \sim N(\mu, \sigma^2)$$

A continuous random variable Z is said to be a standard normal random variable, shown as Z∼N(0,1), if its probability density function is given by [13]

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad for\ all\ z\epsilon\mathbb{R}.$$

Figure 3 displays the probability density function of the standard normal random variable.

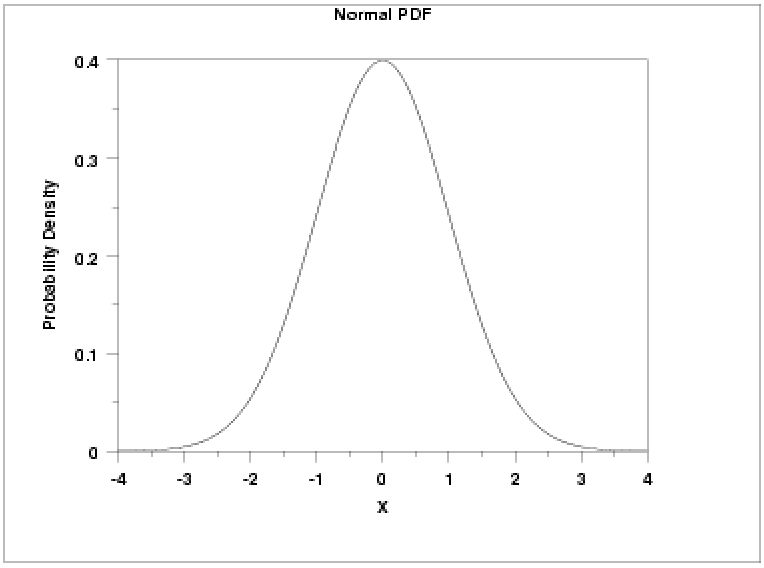

Figure 1. Probability density function of the standard normal variable.

The constants μ and σ included in normal distribution $N(\mu,\sigma^2)$ correspond to the expectation and standard deviation [15]:
E(X)= μ and V(X)=σ^2

The Q-Q plot, or quantile-quantile plot, is a graphical tool to help it assess if a set of data plausibly came from some theoretical distribution like a Normal. A Q-Q plot is a scatterplot created by plotting two sets of quantiles against one another. If both sets of quantiles came from the same distribution, it should be seen the points forming a line that's roughly straight. Here's an example of a Normal Q-Q plot when both sets of quantiles truly come from Normal distributions [16].
The Shapiro-Wilk test is a way to tell if a random sample comes from a normal distribution. The test gives you a W value; small values indicate your sample is not normally distributed. The formula for the W value is[17, 18].

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

x_i are the ordered random sample values
a_i are constants generated from the covariances, variances and means of the sample (size n) from a normally distributed sample.

III. Results

Descriptive statistics of boxers' punches in matches and competition dates are given in Table 2.

Table 2. Descriptive statistics of the number of punches thrown

No	Competitors	$\bar{X}$	s	Competition date
1	Alvarez-Golovkin	51.83-72.42	9.23-12.95	15.09.2018
2	Joshua- Usyk	53.42-44.08	15.49-10.24	24.09.2021
3	Hurd- Williams	66.67-57.25	22.56-13.87	11.05.2019
4	Charlo-Castano	44.42-48.83	10.92-13.48	18.07.2021
5	Haney-Diaz	51.17-47.75	8.56-14.77	05.12.2021
6	Pacquiao-Ugas	67.92-33.75	7.73-6	21.08.2021
7	Kovalev-Ward	39.5-28.08	9-7.74	23.11.2016
8	Broner-Santiago	58.08-49.58	20.64-17.78	21.02.2021

$\bar{X}$: Mean, s: Standard deviation

The data were examined to see if they conformed to normal distribution. For this purpose, the data's mean and standard deviation were calculated. The average punch count over 192 rounds was calculated as 49.58 and the standard deviation was calculated as 17.78. The Shapiro-Wilk test was used to determine normality assumption. The Shapiro-Wilk statistic was 0.99, and since p=0.188>0.05, it was observed that the data were normally distributed. Figures 1 and 2 show the Q-Q plot graph for normal distribution and the graph obtained to determine whether there is an outlier value in the data.

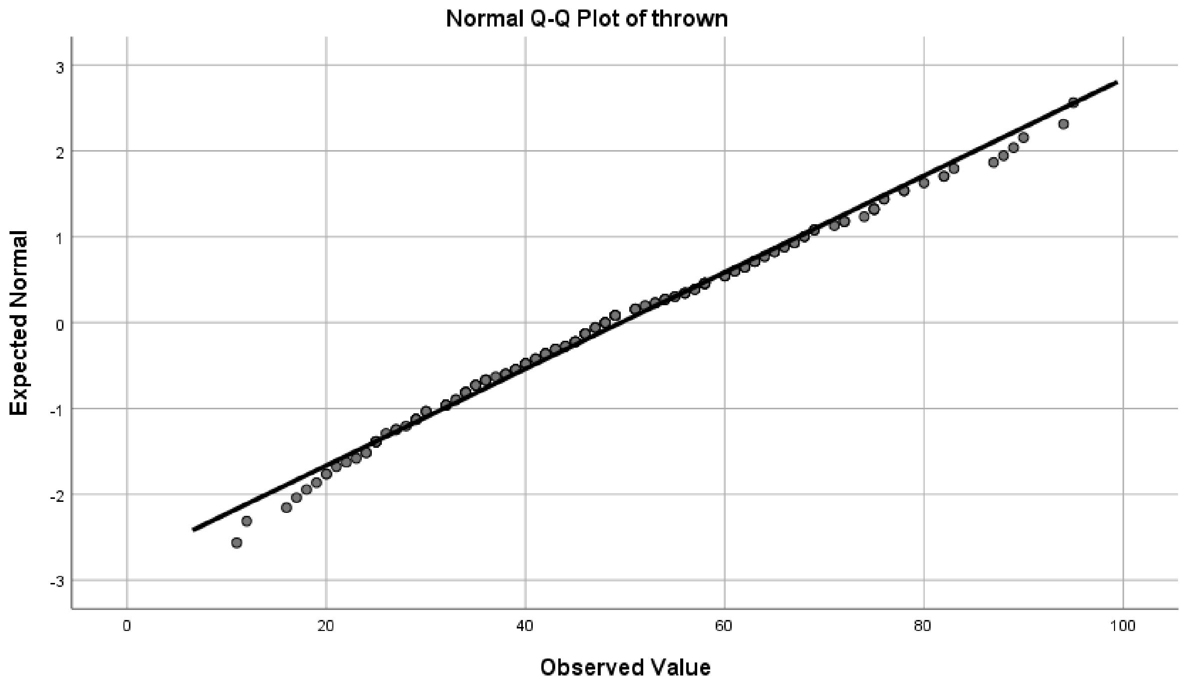

Figure 1. Q-Q Plot-plot

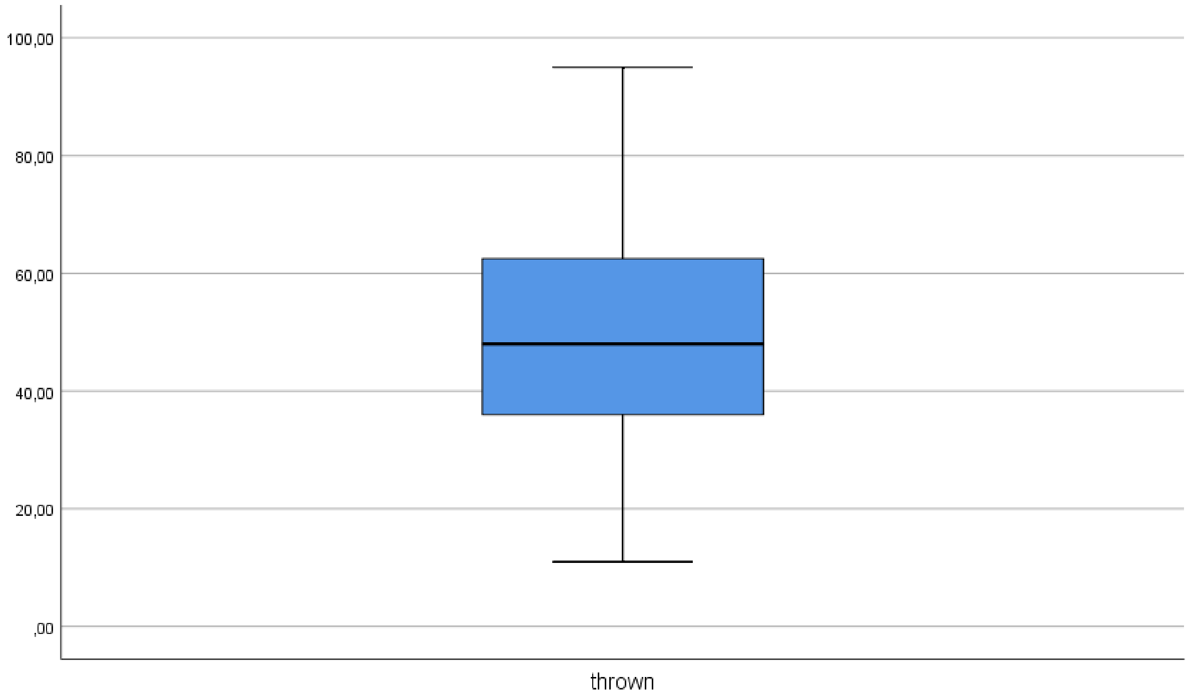

Figure 2. Outlier graph of number of thrown

Figure 1 shows that the data in the Q-Q plot graph were distributed around the line, and that there were no direct significant deviations. Figure 2 shows that there were no extreme values in the data, indicating that the data conformed to normal distribution. The average of the data set suitable for normal distribution was 49.58 and the standard deviation was 17.78. The probability of a boxer throwing more than 30 punches in a round, for example, can be calculated using this information as follows.

$$P\left(X > \frac{30 - 49.58}{17.78}\right) = P(Z > -1.10) = 0.50 + 0.3643 = 0.8643$$

In other words, any boxer has an 86.43% chance of throwing more than 30 punches in a single round.
If the probability of any boxer throwing between 50 and 60 punches per round is calculated,

$$P\left(\frac{50-49.58}{17.78}<X<\frac{60-49.58}{17.78}\right)=P(0.02<Z<0.59)=0.2144$$

is obtained. In other words, there is a 21.44% probability that a boxer will throw 50 to 60 punches per round.
R Studio software is also used to perform necessary calculations. For example, the probability of a boxer throwing more than 30 punches in a round was calculated to be
>pnorm(30, mean=49.5833, sd=17.77619, lower.tail=FALSE)>
[1] 0.864695
Similarly, for probabilities of punching more than 35, 40, 45, 50, 55, 55, 60, 65, and 70, R Studio commands and calculated results are as follows.
>pnorm(35, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.7940014
>pnorm(40, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.7050941
>pnorm(45, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.6017324
>pnorm(50, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.4906491
>pnorm(55, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.380291
>pnorm(60, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.2789405
>pnorm(65, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.1928979
>pnorm(70, mean=49.5833, sd=17.77619, lower.tail=FALSE)
[1] 0.1253725

The probabilities of any boxer throwing more than 30, 35, 40, 45, 50, 55, 55, 60, 65 and 70 punches per round are presented in Table 3.

Table 3. Probability of throwing more than X punches

X >	z>	p
X > 30	Z > -1.10166	0.8643
X > 35	Z > -0.82038	0.7939
X > 40	Z > -0.53911	0.7054
X > 45	Z > -0.25783	0.6026
X > 50	Z > 0.023441	0.4920
X > 55	Z > 0.304717	0.3821
X > 60	Z > 0.585992	0.2776
X > 65	Z > 0.867267	0.1922
X > 70	Z > 1.148542	0.1251

The probabilities of any boxer throwing 30-40, 40-50, 50-60 and 60-70 punches per round are given in Table 4.

Table 4. Probability of X being between certain values

a<X< b	a<Z< b	p
30 <X 40	-1.10 <Z< -0.54	0.1589
40 <X 50	-0.54 <Z< 0.02	0.2134
50 <X< 60	0.02 <Z< 0.59	0.2144
60 <X< 70	0.59 <Z< 1.15	0.1525

IV. Conclusion

It was found in the study that the data on the number of punches thrown by 16 boxers over 12 rounds in 8 high-level boxing matches conformed to normal distribution. A boxer's probability of throwing fewer than 50 punches in a round is 50.94%, which is slightly higher than 50%.

Normal distribution, which is the most widely used and holds a significant place in statistics, is used in a variety of applications in daily life.

References

[1]. David, F.N. 1962. Games, Gods, and Gambling, New York, Hafner Pub. Co.

[2]. Stahl, S. 2006. The Evolution of the Normal Distribution. Mathematics Magazine, 79(2):95-113

[3]. Anonymous, 2022. History of the Normal Distribution.http://www.z-table.com/history-on-normal-distirbution.html

[4]. Musselwhite, D.J., Wesolowski, B.C. 2018. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation Normal Distribution (Editor:Bruce B. Frey). SAGE Publications, Inc., Thousand Oaks.

[5]. https://fightclub.com.tr/boks/canelo-alvarez-gennady-golovkin-macinin-yumruk-istatistikleri-ve-resmi-puanlama-sonuclari/

[6]. OleksandrUsyk vs. Anthony Joshua- CompuBox Punch Stats.https://www.boxingscene.com/oleksandr-usyk-vs-anthony-joshua-compubox-punch-stats--160817

[7]. https://roundbyroundboxing.com/hurd-vs-williams-compubox-punch-stats/

[8]. https://www.boxingscene.com/jermell-charlo-vs-brian-castano-compubox-punch-stats--159209

[9]. https://www.boxingscene.com/devin-haney-vs-joseph-diaz-compubox-punch-stats--162525

[10]. https://www.spin.ph/boxing/punch-stats-clearly-shows-ugas-dominate-pacquiao-in-upset-win-a793-20210822

[11]. https://www.ringnews24.com/2016/11/23/punch-stats-clearly-favored-sergey-kovalev-controversial-decision/

[12]. https://www.sportbible.com/boxing/boxing-boxer-somehow-wins-first-round-on-judges-card-without-landing-a-punch-20210221

[13]. Pishro-Nik. H. 2014. Introduction to Probability. Statistics. and Random Processes. Kappa Research. LLC. ISBN: 0990637204. 9780990637202

[14]. Yau. C. 2014. R Tutorial. With Bayesian Statistics Using Stan. R Tutorial eBook. http://www.r-tutor.com/elementary-statistics/probability-distributions/normal-distribution

[15]. Sugiyama, M. 2016. Introduction to Statistical Machine Learning. Morgan Kaufman is an imprint of Elsevier, USA.

[16]. Ford, C. 2015. Understanding Q-Q Plots. Satistical Research Consultant University of Virginia Library. https://data.library.virginia.edu/understanding-q-q-plots/

[17]. Everitt, B.S., Skrondal, A. 2010. The Cambridge Dictionary of Statistics, Cambridge UniversityPress.

[18]. Vogt, W.P. 2005. Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences.SAGE.