Cawangan Kedah
Kampus Sungai Petani

e-PROCEEDINGS

of The 5th International Conference on Computing, Mathematics and Statistics (iCMS2021)

4-5 August 2021

Driving Research Towards Excellence

e-Proceedings of the 5th International Conference on Computing, Mathematics and Statistics (iCMS 2021)

Driving Research Towards Excellence

e-ISBN: 978-967-2948-12-4
DOI

Library of Congress Control Number:

Publication by
Department of Mathematical Sciences
Faculty of Computer & Mathematical Sciences
UiTM Kedah

TABLE OF CONTENT

PART 1: MATHEMATICS

	Page
STATISTICAL ANALYSIS ON THE EFFECTIVENESS OF SHORT-TERM PROGRAMS DURING COVID-19 PANDEMIC: IN THE CASE OF PROGRAM BIJAK SIFIR 2020 *Nazihah Safie, Syerrina Zakaria, Siti Madhihah Abdul Malik, Nur Baini Ismail, Azwani Alias Ruwaidiah Idris*	**1**
RADIATIVE CASSON FLUID OVER A SLIPPERY VERTICAL RIGA PLATE WITH VISCOUS DISSIPATION AND BUOYANCY EFFECTS *Siti Khuzaimah Soid, Khadijah Abdul Hamid, Ma Nuramalina Nasero, NurNajah Nabila Abdul Aziz*	**10**
GAUSSIAN INTEGER SOLUTIONS OF THE DIOPHANTINE EQUATION $x^4 + y^4 = z^3$ FOR $x \neq y$ *Shahrina Ismail, Kamel Ariffin Mohd Atan and Diego Sejas Viscarra*	**19**
A SEMI ANALYTICAL ITERATIVE METHOD FOR SOLVING THE EMDEN-FOWLER EQUATIONS *Mat Salim Selamat, Mohd Najir Tokachil, Noor Aqila Burhanddin, Ika Suzieana Murad and Nur Farhana Razali*	**28**
ROTATING FLOW OF A NANOFLUID PAST A NONLINEARLY SHRINKING SURFACE WITH FLUID SUCTION *Siti Nur Alwani Salleh, Norfifah Bachok and Nor Athirah Mohd Zin*	**36**
MODELING THE EFFECTIVENESS OF TEACHING BASIC NUMBERS THROUGH MINI TENNIS TRAINING USING MARKOV CHAIN *Rahela Abdul Rahim, Rahizam Abdul Rahim and Syahrul Ridhwan Morazuk*	**46**
PERFORMANCE OF MORTALITY RATES USING DEEP LEARNING APPROACH *Mohamad Hasif Azim and Saiful Izzuan Hussain*	**53**
UNSTEADY MHD CASSON FLUID FLOW IN A VERTICAL CYLINDER WITH POROSITY AND SLIP VELOCITY EFFECTS *Wan Faezah Wan Azmi, Ahmad Qushairi Mohamad, Lim Yeou Jiann and Sharidan Shafie*	**60**
DISJUNCTIVE PROGRAMMING - TABU SEARCH FOR JOB SHOP SCHEDULING PROBLEM *S. Z. Nordin, K.L. Wong, H.S. Pheng, H. F. S. Saipol and N.A.A. Husain*	**68**
FUZZY AHP AND ITS APPLICATION TO SUSTAINABLE ENERGY PLANNING DECISION PROBLEM *Liana Najib and Lazim Abdullah*	**78**
A CONSISTENCY TEST OF FUZZY ANALYTIC HIERARCHY PROCESS *Liana Najib and Lazim Abdullah*	**89**
FREE CONVECTION FLOW OF BRINKMAN TYPE FLUID THROUGH AN COSINE OSCILLATING PLATE *Siti Noramirah Ibrahim, Ahmad Qushairi Mohamad, Lim Yeou Jiann, Sharidan Shafie and Muhammad Najib Zakaria*	**98**

RADIATION EFFECT ON MHD FERROFLUID FLOW WITH RAMPED WALL TEMPERATURE AND ARBITRARY WALL SHEAR STRESS **106**
Nor Athirah Mohd Zin, Aaiza Gul, Siti Nur Alwani Salleh, Imran Ullah, Sharena Mohamad Isa, Lim Yeou Jiann and Sharidan Shafie

PART 2: STATISTICS

A REVIEW ON INDIVIDUAL RESERVING FOR NON-LIFE INSURANCE **117**
Kelly Chuah Khai Shin and Ang Siew Ling

STATISTICAL LEARNING OF AIR PASSENGER TRAFFIC AT THE MURTALA MUHAMMED INTERNATIONAL AIRPORT, NIGERIA **123**
Christopher Godwin Udomboso and Gabriel Olugbenga Ojo

ANALYSIS ON SMOKING CESSATION RATE AMONG PATIENTS IN HOSPITAL SULTAN ISMAIL, JOHOR **137**
Siti Mariam Norrulashikin, Ruzaini Zulhusni Puslan, Nur Arina Bazilah Kamisan and Siti Rohani Mohd Nor

EFFECT OF PARAMETERS ON THE COST OF MEMORY TYPE CHART **146**
Sakthiseswari Ganasan, You Huay Woon and Zainol Mustafa

EVALUATION OF PREDICTORS FOR THE DEVELOPMENT AND PROGRESSION OF DIABETIC RETINOPATHY AMONG DIABETES MELLITUS TYPE 2 PATIENTS **152**
Syafawati Ab Saad, Maz Jamilah Masnan, Karniza Khalid and Safwati Ibrahim

REGIONAL FREQUENCY ANALYSIS OF EXTREME PRECIPITATION IN PENINSULAR MALAYSIA **160**
Iszuanie Syafidza Che Ilias, Wan Zawiah Wan Zin and Abdul Aziz Jemain

EXPONENTIAL MODEL FOR SIMULATION DATA VIA MULTIPLE IMPUTATION IN THE PRESENT OF PARTLY INTERVAL-CENSORED DATA **173**
Salman Umer and Faiz Elfaki

THE FUTURE OF MALAYSIA'S AGRICULTURE SECTOR BY 2030 **181**
Thanusha Palmira Thangarajah and Suzilah Ismail

MODELLING MALAYSIAN GOLD PRICES USING BOX-JENKINS APPROACH **186**
Isnewati Ab Malek, Dewi Nur Farhani Radin Nor Azam, Dinie Syazwani Badrul Aidi and Nur Syafiqah Sharim

WATER DEMAND PREDICTION USING MACHINE LEARNING: A REVIEW **192**
Norashikin Nasaruddin, Shahida Farhan Zakaria, Afida Ahmad, Ahmad Zia Ul-Saufie and Norazian Mohamaed Noor

DETECTION OF DIFFERENTIAL ITEM FUNCTIONING FOR THE NINE-QUESTIONS DEPRESSION RATING SCALE FOR THAI NORTH DIALECT **201**
Suttipong Kawilapat, Benchlak Maneeton, Narong Maneeton, Sukon Prasitwattanaseree, Thoranin Kongsuk, Suwanna Arunpongpaisal, Jintana Leejongpermpool, Supattra Sukhawaha and Patrinee Traisathit

ACCELERATED FAILURE TIME (AFT) MODEL FOR SIMULATION PARTLY INTERVAL-CENSORED DATA **210**
Ibrahim El Feky and Faiz Elfaki

MODELING OF INFLUENCE FACTORS PERCENTAGE OF GOVERNMENTS' RICE RECIPIENT FAMILIES BASED ON THE BEST FOURIER SERIES ESTIMATOR **217**
Chaerobby Fakhri Fauzaan Purwoko, Ayuning Dwis Cahyasari, Netha Aliffia and M. Fariz Fadillah Mardianto

CLUSTERING OF DISTRICTS AND CITIES IN INDONESIA BASED ON POVERTY INDICATORS USING THE K-MEANS METHOD **225**
Khoirun Niswatin, Christopher Andreas, Putri Fardha Asa OktaviaHans and M. Fariz Fadilah Mardianto

ANALYSIS OF THE EFFECT OF HOAX NEWS DEVELOPMENT IN INDONESIA USING STRUCTURAL EQUATION MODELING-PARTIAL LEAST SQUARE **233**
Christopher Andreas, Sakinah Priandi, Antonio Nikolas Manuel Bonar Simamora and M. Fariz Fadillah Mardianto

A COMPARATIVE STUDY OF MOVING AVERAGE AND ARIMA MODEL IN FORECASTING GOLD PRICE **241**
Arif Luqman Bin Khairil Annuar, Hang See Pheng, Siti Rohani Binti Mohd Nor and Thoo Ai Chin

CONFIDENCE INTERVAL ESTIMATION USING BOOTSTRAPPING METHODS AND MAXIMUM LIKELIHOOD ESTIMATE **249**
Siti Fairus Mokhtar, Zahayu Md Yusof and Hasimah Sapiri

DISTANCE-BASED FEATURE SELECTION FOR LOW-LEVEL DATA FUSION OF SENSOR DATA **256**
M. J. Masnan, N. I. Maha3, A. Y. M. Shakaf, A. Zakaria, N. A. Rahim and N. Subari

BANKRUPTCY MODEL OF UK PUBLIC SALES AND MAINTENANCE MOTOR VEHICLES FIRMS **264**
Asmahani Nayan, Amirah Hazwani Abd Rahim, Siti Shuhada Ishak, Mohd Rijal Ilias and Abd Razak Ahmad

INVESTIGATING THE EFFECT OF DIFFERENT SAMPLING METHODS ON IMBALANCED DATASETS USING BANKRUPTCY PREDICTION MODEL **271**
Amirah Hazwani Abdul Rahim, Nurazlina Abdul Rashid, Abd-Razak Ahmad and Norin Rahayu Shamsuddin

INVESTMENT IN MALAYSIA: FORECASTING STOCK MARKET USING TIME SERIES ANALYSIS **278**
Nuzlinda Abdul Rahman, Chen Yi Kit, Kevin Pang, Fauhatuz Zahroh Shaik Abdullah and Nur Sofiah Izani

PART 3: COMPUTER SCIENCE & INFORMATION TECHNOLOGY

ANALYSIS OF THE PASSENGERS' LOYALTY AND SATISFACTION OF AIRASIA PASSENGERS USING CLASSIFICATION 291
Ee Jian Pei, Chong Pui Lin and Nabilah Filzah Mohd Radzuan

HARMONY SEARCH HYPER-HEURISTIC WITH DIFFERENT PITCH ADJUSTMENT OPERATOR FOR SCHEDULING PROBLEMS 299
Khairul Anwar, Mohammed A.Awadallah and Mohammed Azmi Al-Betar

A 1D EYE TISSUE MODEL TO MIMIC RETINAL BLOOD PERFUSION DURING RETINAL IMAGING PHOTOPLETHYSMOGRAPHY (IPPG) ASSESSMENT: A DIFFUSION APPROXIMATION – FINITE ELEMENT METHOD (FEM) APPROACH 307
Harnani Hassan, Sukreen Hana Herman, Zulfakri Mohamad, Sijung Hu and Vincent M. Dwyer

INFORMATION SECURITY CULTURE: A QUALITATIVE APPROACH ON MANAGEMENT SUPPORT 325
Qamarul Nazrin Harun, Mohamad Noorman Masrek, Muhamad Ismail Pahmi and Mohamad Mustaqim Junoh

APPLY MACHINE LEARNING TO PREDICT CARDIOVASCULAR RISK IN RURAL CLINICS FROM MEXICO 335
Misael Zambrano-de la Torre, Maximiliano Guzmán-Fernández, Claudia Sifuentes-Gallardo, Hamurabi Gamboa-Rosales, Huizilopoztli Luna-García, Ernesto Sandoval-García, Ramiro Esquivel-Felix and Héctor Durán-Muñoz

ASSESSING THE RELATIONSHIP BETWEEN STUDENTS' LEARNING STYLES AND MATHEMATICS CRITICAL THINKING ABILITY IN A 'CLUSTER SCHOOL' 343
Salimah Ahmad, Asyura Abd Nassir, Nor Habibah Tarmuji, Khairul Firhan Yusob and Nor Azizah Yacob

STUDENTS' LEISURE WEEKEND ACTIVITIES DURING MOVEMENT CONTROL ORDER: UiTM PAHANG SHARING EXPERIENCE 351
Syafiza Saila Samsudin, Noor Izyan Mohamad Adnan, Nik Muhammad Farhan Hakim Nik Badrul Alam, Siti Rosiah Mohamed and Nazihah Ismail

DYNAMICS SIMULATION APPROACH IN MODEL DEVELOPMENT OF UNSOLD NEW RESIDENTIAL HOUSING IN JOHOR 363
Lok Lee Wen and Hasimah Sapiri

WORD PROBLEM SOLVING SKILLS AS DETERMINANT OF MATHEMATICS PERFORMANCE FOR NON-MATH MAJOR STUDENTS 371
Shahida Farhan Zakaria, Norashikin Nasaruddin, Mas Aida Abd Rahim, Fazillah Bosli and Kor Liew Kee

ANALYSIS REVIEW ON CHALLENGES AND SOLUTIONS TO COMPUTER PROGRAMMING TEACHING AND LEARNING 378
Noor Hasnita Abdul Talib and Jasmin Ilyani Ahmad

PART 4: OTHERS

ANALYSIS OF CLAIM RATIO, RISK-BASED CAPITAL AND VALUE-ADDED INTELLECTUAL CAPITAL: A COMPARISON BETWEEN FAMILY AND GENERAL TAKAFUL OPERATORS IN MALAYSIA **387**
Nur Amalina Syafiqa Kamaruddin, Norizarina Ishak, Siti Raihana Hamzah, Nurfadhlina Abdul Halim and Ahmad Fadhly Nurullah Rasade

THE IMPACT OF GEOMAGNETIC STORMS ON THE OCCURRENCES OF EARTHQUAKES FROM 1994 TO 2017 USING THE GENERALIZED LINEAR MIXED MODELS **396**
N. A. Mohamed, N. H. Ismail, N. S. Majid and N. Ahmad

BIBLIOMETRIC ANALYSIS ON BITCOIN 2015-2020 **405**
Nurazlina Abdul Rashid, Fazillah Bosli, Amirah Hazwani Abdul Rahim, Kartini Kasim and Fathiyah Ahmad@Ahmad Jali

GENDER DIFFERENCE IN EATING AND DIETARY HABITS AMONG UNIVERSITY STUDENTS **413**
Fazillah Bosli, Siti Fairus Mokhtar, Noor Hafizah Zainal Aznam, Juaini Jamaludin and Wan Siti Esah Che Hussain

MATHEMATICS ANXIETY: A BIBLIOMETRIX ANALYSIS **420**
Kartini Kasim, Hamidah Muhd Irpan, Noorazilah Ibrahim, Nurazlina Abdul Rashid and Anis Mardiana Ahmad

PREDICTION OF BIOCHEMICAL OXYGEN DEMAND IN MEXICAN SURFACE WATERS USING MACHINE LEARNING **428**
Maximiliano Guzmán-Fernández, Misael Zambrano-de la Torre, Claudia Sifuentes-Gallardo, Oscar Cruz-Dominguez, Carlos Bautista-Capetillo, Juan Badillo-de Loera, Efrén González Ramírez and Héctor Durán-Muñoz

CONFIDENCE INTERVAL ESTIMATION USING BOOTSTRAPPING METHODS AND MAXIMUM LIKELIHOOD ESTIMATE

Siti Fairus Mokhtar[1,3], Zahayu Md Yusof[2] and Hasimah Sapiri[3]
Department of Mathematical Sciences, Universiti Teknologi Mara, Malaysia[1]
Institute of Strategic Industrial Decision Modelling (ISDIM), School of Quantitative Sciences Universiti Utara Malaysia, Kedah, Malaysia[2]
School of Quantitative Sciences Universiti Utara Malaysia, Kedah, Malaysia[3]
([1]fairus706@uitm.uitm.edu.my, [2]zahayu@uum.edu.my, [3]hasimah@uum.edu.my)

Abstract: Confidence interval estimation is an important technique to estimate parameter of a population calculated from a sample drawn from the population. The objective of this study is to present the steps to calculate confidence interval using SPSS. The objective of this paper also is to compare confidence interval using maximum likelihood estimate, percentile bootstrap, and bias-corrected and accelerated methods. Bootstrap is not commonly used because this method is complex to calculate. The advantages of bootstrapping are valid for small samples, and it is a convenient tool. The study found that the BCa method produced CIs closer to the desired level of the coverage than the other methods.

Keywords: bootstrapping method**,** confidence interval, maximum likelihood estimate

1. Introduction

According to Petty (2012), confidence interval (CI) is an interval estimate of a parameter of a population (e.g., a mean) calculated from a sample drawn from the population. A confidence interval has an associated confidence level, which is frequency with which a calculated confidence interval is expected to contain the population parameter. Confidence interval estimation plays an important part in statistical inferences about a population parameter and evaluate the accuracy of its estimator. Determining the unknown parameter estimation of confidence interval values is an important guarantee for the success of subsequent data processing. Confidence interval estimation are useful to modelers to build confidence about parameter estimates (Dogan, 2004).

Confidence intervals also provide more information than point estimates (Das, 2019). By establishing a 95% confidence interval using the sample's mean and standard deviation, and assuming a normal distribution as represented by the bell curve. Confidence interval is the upper and lower bound that contains the true mean 95% of the time. Assume the interval is 61 marks to 82 marks. If 100 random samples taken from the population of statistics class students, the mean should fall between 61 and 82 marks in 95% of those samples.

The purpose of this study is to present the steps to calculate confidence interval using SPSS. The objective of this paper also is to calculate and maximum likelihood estimate, percentile bootstrap, and bias-corrected and accelerated methods. In practice, one of the main reasons for using bootstrap methods is uncertainty whether certain assumptions hold, and in most applications the nonparametric bootstrap is used (Wehrens et al., 2000).

Section 1 is the introduction section. Followed by section 2 which provides statistical background and methodology to find the confidence interval. Section 3 explains the real data example. Section 4 explain the findings of the result. Finally, section 5 is the conclusion of this paper.

2. Statistical background

This section provides brief but essential statistical background. Topics covered include the concepts of bootstrap in calculating the confidence intervals.

2.1 Bootstrap methods

A statistical method called bootstrap method was developed by Efron (1979), which can be used to increase the sample size by applying nonparametric resampling. This method involves the extraction of a bootstrap sample of size n with the original sample data. The samples are used to test the statistical characteristics of the unknown distribution, such as mean, variance, standard deviation, and confidence interval (Zhang et al., 2019). Bootstrap is not commonly used because this method is complex to calculate (Doğan, 2017). The advantages of bootstrapping are valid for small samples, and it is a convenient tool.

This paper focuses on the usage of fatigue test data to estimate the confidence interval by drawing samples with replacement from sample data. According to Thai et al., (2013), let B be the number of bootstrap samples to be drawn from the original dataset, a general bootstrap algorithm is:

1.Generate a bootstrap sample by resampling from the data and/or from the estimated model (Sample n elements with replacement from original sample data)

2.Obtain the estimates for all parameters of the model for the bootstrap sample eg. mean, median etc.

3. Repeat steps 1-2 B times to obtain the bootstrap distribution of parameter estimates and then compute mean, standard deviation, and 95% confidence interval of this distribution

In bootstrap sampling the number of replications is very important. Diciccio and Efron (1996) highlighted the importance of using at least 2000 replications while conducting bootstrap resampling. According to (Efron and Tibshirani, 1994), the number of bootstrap draws cannot be less than n^n. Based on empirical research (IBM, 2013), it has been shown that enough draws for conducting tests is B = 1000 measurements for percentile bootstrap method and Bias-Corrected and Accelerated Bootstrap Method (BCa). A schematic description of the steps for estimating confidence intervals using bootstrap formed by Haukoos and Lewis (2005) is shown in Figure 1.

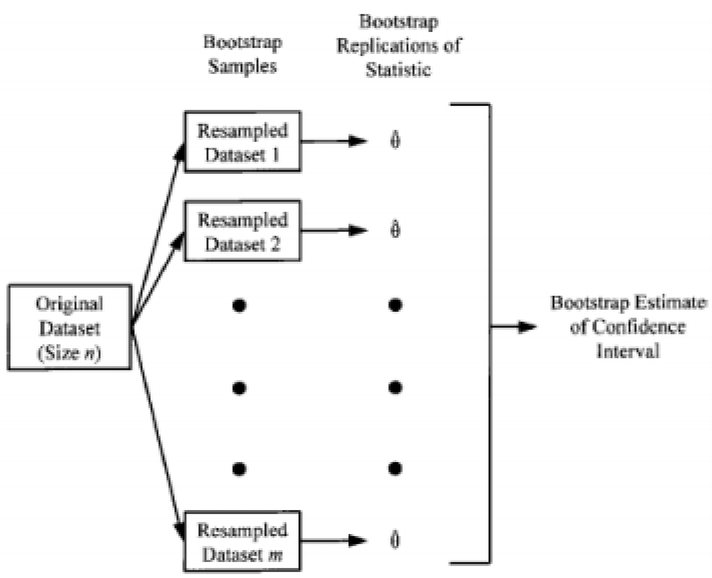

Figure 1: Description of the steps in bootstrapping.

In this paper, there are two types of confidence interval estimation which are MLE and bootstrapping. The confidence interval construction is based on asymptotic normality of the MLE (Kreutz et al., 2013). The different methods available for estimating bootstrap confidence intervals for estimated parameters (Mesabbah et al., 2015). This paper utilizes two: percentile bootstrap method and BCa.

2.1.1 Percentile Bootstrap Method

According to Mesabbah et al., (2015), the bootstrap percentile confidence interval method is based on the quantile of the bootstrap distribution of the parameters estimate. The percentile bootstrap interval is just the interval between the 100×($\frac{\alpha}{2}$) and 100×(α_2) and 100×(1−$\frac{\alpha}{2}$) percentiles of the distribution of θ estimates obtained from resampling, where θ represents a parameter of interest and α is the level of significance (e.g., α = 0.05 for 95% CIs) (Efron, 1982). A bootstrap percentile CI of $\widehat{\theta}$ (an estimator of θ) can be obtained as follows: (1) *B* random bootstrap samples are generated, (2) a parameter estimate is calculated from each bootstrap sample, (3) all B bootstrap parameter estimates are ordered from the lowest to highest, and (4) the CI is constructed as follows, [$\widehat{\theta}$lower limit, $\widehat{\theta}$upper limit]=[$\widehat{\theta}_j^*$, $\widehat{\theta}_k^*$], where $\widehat{\theta}_j^*$ denotes the *j*th quantile (lower limit), and $\widehat{\theta}_k^*$ denotes the *k*th quantile (upper limit); j=[$\frac{\alpha}{2}$×B], k=[(1−$\frac{\alpha}{2}$)×B]. For example, a 95% percentile bootstrap CI with 1,000 bootstrap samples is the interval between the 25th quantile value and the 975th quantile value of the 1,000 bootstrap parameter estimates (Jung et al., 2019).

2.1.2 Bias-Corrected and Accelerated Bootstrap Method (BCa)

To overcome the over coverage issues in percentile bootstrap CIs (Efron and Tibshirani, 1993), the BCa method corrects for both bias and skewness of the bootstrap parameter estimates by incorporating a bias-correction factor and an acceleration factor (Efron, 1987; Efron and Tibshirani, 1993). Equation (1) is the bias-corrected and accelerated formulae. The bias-correction factor $\hat{z}_0$ is estimated as the proportion of the bootstrap estimates less than the original parameter estimate $\widehat{\theta}$,

$$\hat{z}_0 = \phi^{-1}\left(\frac{\#\{\widehat{\theta}^{\cdot} < \widehat{\theta}\}}{B}\right) \tag{1}$$

where ϕ^{-1} is the inverse function of a standard normal cumulative distribution function (e.g., ϕ^{-1} (0.975) = 1.96). So the bias correction bootstrap percentile confidence interval is given by:

$$\left[\widehat{\theta}^{\cdot \alpha_1}, \widehat{\theta}^{\cdot \alpha_2}\right]$$

where α_1 and α_2 are modified quantities of the location of the confidence interval's endpoints. The confidence interval endpoints at significant level 100α% are defined as:

$$\alpha_1 = \phi\left(2\hat{z}_0 + Z^{\alpha/2}\right),$$

and

$$\alpha_2 = \phi\left(2\hat{z}_0 + Z^{1-\alpha/2}\right)$$

where ϕ is the cumulative standard normal distribution (Mesabbah et al., 2015).

2.2 MLE

According to (Doğan, 2017), the traditional confidence interval are computed using the following formula. The 95% confidence intervals are computed using the formula $\bar{x} \pm 1.96 * \left(\frac{s}{\sqrt{n}}\right)$ for large data and

$\bar{x} \pm t_{\alpha/2} * \left(\frac{s}{\sqrt{n}} \right)$ for small data (n less than 30) where $\bar{x}$ is the sample mean, s is the standard deviation and n is the sample size.

3. Real Data Example

The marks data (Mathur & Kaushik, 2014) for 20 students are determined. The data is 20, 19, 17, 18, 17, 17, 17, 17, 17, 18, 18, 19, 17, 19, 18, 17, 19, 18, 19 and 16.8. The histogram in Figure 2 suggest lack of normality. For this reason, the percentile bootstrap and BCa bootstrap confidence interval appears more appropriate.

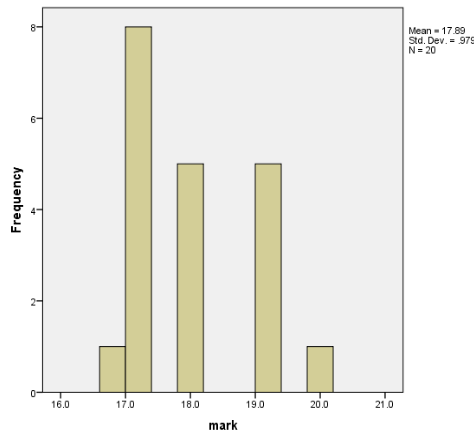

Figure 2: Histogram: Students Marks

4. Findings

The point estimation of the original sample is estimated to be $\hat{\mu} = 17.89$ and standard deviation is $s = 0.9787$. The formula for point estimation and standard deviation is showed below.

$$\hat{\mu} = \frac{20 + 19 + \ldots + 16.8}{20} = \frac{357.8}{20} = 17.89$$

$$s = \sqrt{\frac{1}{20 - 1}\left[6419.24 - \frac{357.8^2}{20}\right]} = 0.9787$$

The following example demonstrates the ease of the bootstrap procedure for using SPSS an extremely user-friendly statistical package with "point and click" commands. The output refers to Table 1 and Table 2.

Analyze -> Descriptive Statistics -> Frequency -> Variable: Mark -> Chart ->Histogram -> Bootstrap ->Number of samples: 1000 ->Percentile/Bias corrected accelerated -> Continue -> OK

Table 1. percentile bootstrap (SPSS Output)

Descriptive Statistics

		Statistic	Bootstrap[a]			
			Bias	Std. Error	95% Confidence Interval	
					Lower	Upper
mark	N	20	0	0	20	20
	Mean	17.890	.001	.214	17.490	18.300
	Std. Deviation	.9787	-.0345	.1195	.7016	1.1760
Valid N (listwise)	N	20	0	0	20	20

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Table 2. BCa bootstrap (SPSS Output)

Descriptive Statistics

		Statistic	Bootstrap[a]			
			Bias	Std. Error	BCa 95% Confidence Interval	
					Lower	Upper
mark	N	20	0	0	.	.
	Mean	17.890	-.012	.212	17.540	18.240
	Std. Deviation	.9787	-.0308	.1133	.8013	1.1014
Valid N (listwise)	N	20	0	0	.	.

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Confidence interval = Point estimate ± Critical value (t) x Standard deviation /$\sqrt{n}$
= Point estimate ± $t_{0.05,19}$ x Standard deviation /$\sqrt{n}$
= 17.89 ± 1.729 x $\frac{0.9787}{\sqrt{20}}$
= 17.89 ± 0.3784
= [17.5116, 18.2684]

The 95% confidence interval is between 17.5116 and 18.2684 through the MLE method. Based on the bootstrap method, the original samples were randomly sampled n=1000 times with replacement, and 1000 bootstrap samples were obtained. The point estimation of the original sample is estimated to be be $\hat{\mu} =$ 17.89, and the 95% confidence interval for percentile bootstrap is between 17.490 and 18.300. For BCa bootstrap, the point estimation of the original sample is estimated to be be $\hat{\mu} = 17.89$, and the 95% confidence interval between 17.540 and 18.240. The original sample data are statistically analyzed through three methods. The comparison of the three methods is summarized in Table 3. The study found that the BCa method produced CIs closer to the desired level of the coverage than the other methods. The mean and standard deviation in Table 1 refers to the difference between MLE, percentile bootstrap and BCa bootstrap.

Table 3. Parameter estimation results of three methods for the sample

Methods	Point estimation		Confidence interval	Interval length
	Estimate	Dev.		
MLE	17.89	0.9787	[17.5116, 18.2684]	0.7568
percentile bootstrap	17.89	0.9787	[17.490, 18.300]	0.8100
BCa	17.89	0.9787	[17.540, 18.240]	0.7000

5. Conclusion

There are three methods to calculate confidence interval. There are MLE, percentile bootstrap and BCa bootstrap. Then computation off bootstrapping methods for mean and median were explained using SPSS. Furthermore, some comparisons were done. Traditional and bootstrapped confidence intervals were compared for mean. The advantages of bootstrapping are assumptions on bootstrap are less restrictive, and more easily checked, than the assumptions on MLE. The bootstrap also can be applied to situations where MLE may be difficult or impossible to find.

Acknowledgment

Thanks are due to the anonymous reviewers for suggestions and helpful comments to improve the manuscript.

References

Das, S. K. (2019). Confidence Interval Is More Informative Than P-Value in Research. *International Journal of Engineering Applied Sciences and Technology*, 04(06): 278–282.

Doğan, C. D. (2017). Applying bootstrap resampling to compute confidence intervals for various statistics with R. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 2017(68): 1–18.

Dogan, G. (2004). Confidence Interval Estimation in System Dynamics Models: Bootstrapping vs. Likelihood Ratio Method. *Proceedings of the 22nd International Conference of the System Dynamics Society.*, 1–38.

Efron, B., and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. (1st ed.). CRC Press.

IBM. (2013). *IBM SPSS Bootstrapping 22*. http://library.uum.edu/services/statistics/SPSS22Manuals/IBM SPSS Bootstrapping.pdf

Jung, K., Lee, J., Gupta, V., and Cho, G. (2019). Comparison of Bootstrap Confidence Interval Methods for GSCA Using a Monte Carlo Simulation. *Frontiers in Psychology*, 10:24–25.

Kreutz, C., Raue, A., Kaschek, D., and Timmer, J. (2013). Profile likelihood in systems biology. *FEBS Journal*, 280(11): 2564–2571.

Mathur, B., and Kaushik, M. (2014). Data Analysis of Students Marks with Descriptive Statistics. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2:1188–1191.

Mesabbah, M., Rashwan, W., and Arisha, A. (2015). An empirical estimation of statistical inferences for system dynamics model parameters. *Proceedings - Winter Simulation Conference*, *2015-Janua*, 686–697.

Petty, M. D. (2012). Calculating and using confidence intervals for model validation. *Fall Simulation Interoperability Workshop 2012*, *2012 Fall SIW*, 37–45.

Thai, H. T., Mentré, F., Holford, N. H. G., Veyrat-Follet, C., and Comets, E. (2013). A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models. *Pharmaceutical Statistics*, 12(3): 129–140.

Wehrens, R., Putter, H., and Buydens, L. M. C. (2000). The bootstrap: A tutorial. *Chemometrics and Intelligent Laboratory Systems*, 54(1): 35–52.

Zhang, M., Liu, X., Wang, Y., and Wang, X. (2019). Parameter distribution characteristics of material fatigue life using improved bootstrap method. *International Journal of Damage Mechanics*, 28(5): 772–793.

2021 ICMS

INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND STATISTICS

e ISBN 978-967-2948-12-4

9 7 8 9 6 7 2 9 4 8 1 2 4