Scientia Africana, Vol. 18 (No. 2), August, 2019. Pp 173-182
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

COMPARATIVE MODELLING OF PARAMETER ESTIMATION METHODS WITH NIGERIAN FINANCIAL SECTOR STATISTICS

Chigozie Kelechi Acha[1]

[1]Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
E-mails- specialgozie@yahoo.com and acha.kelechi@mouau.edu.ng

Received: 17-08-19
Accepted: 24-08-19

ABSTRACT

The study focused on the performance of the parameter estimation methods on financial sector statistics in the Nigerian economy and identified the sampling distribution of the test statistic θ. Three general methods of parameter estimation (least-squares estimation, maximum likelihood estimation and method of moments) were compared using their standard errors. Secondary quarterly data collected from Central Bank of Nigeria Statistical Bulletin 2017 spanning from 1988-2017. Datasets on financial sector statistics were used as the basis for defining the population and the true standard errors. The sampling distribution of the financial sector statistics was found to be a Chi-square distribution and was confirmed using a resampling method. The stability of the test statistic θ was also ascertained. In addition, other diagnostic checks, Akaike Information criteria; Schwart Bayesian Information criterion, Hannan-Quinn Information criterion were used and they confirmed that when the financial sector statistics was resampled, the parametric resampled model (R311) turned out to be the best model under several assessment conditions.

Keywords: Information criteria, resampling, parametric models, parameter estimation, kernel density, quantile-quantile plots

INTRODUCTION

The goal of modeling is to deduce the form of the underlying process that approximates the true models. Once a model is specified with its parameters, and data have been collected, one is in a position to evaluate its goodness of fit and how well it fits the observed data. Goodness of fit is assessed by finding parameter values of a model that best fits the data; a procedure called parameter estimation. There are basically three general methods of parameter estimation: least-squares estimation, maximum likelihood estimation, and method of moments. These parameter estimations have different methods under them and the methods that suit the study will be used, and they are least-squares estimation – ordinary least square method, maximum likelihood estimation – secant method, and method of moments – cue method, then for the kernel density and quantile-quantile plots - quadratic spectral method and also a parametric resampling method. There are also many sectors in the Nigerian economy namely; educational sector, agricultural sector, financial sector and so on. Financial sector is a vital aspect of the Nigerian economy. This is because whether seen from the point of view of export or import, it has fundamental implications for the economy. It therefore becomes expedient for any economy that wishes to grow to pay attention to changes in the financial sector The importance of financial statistics for monetary policy decision making cannot be over-emphasized. Without basic information on monetary, financial and economic developments, and systemic risks, it will be difficult for central banks to assess economic progress. Since the monetary authorities operate in a highly uncertain environment due to disturbances like financial shocks, demand and supply shocks, timely and quality statistical information are necessary for sound monetary policy-making. Also, the major users of money and banking statistics in Nigeria are the policy makers at the CBN, Federal Ministry of Finance, National Planning Commission, the Presidency, the financial sub-sector (banks and other non-bank financial institutions), research institutes, private researchers, universities and investors. The economic and financial statistics give the background information to the status of their citizen's health, education, housing, recreation and other areas of welfare.

Therefore, the purpose of this study is to investigate and understand the parametric methods of the financial sector statistics from the parameter estimation methods under a variety of assessment conditions and to compare it with parametric resampling method in estimating its parameter. Datasets on export, import and gross domestic product (GDP) from Central Bank of Nigeria Statistical Bulletin 2017 from 1988-2017 was used as the basis to define the population and the true standard errors for sound monetary policy-making. The following parameter estimation methods are considered;

1. **Ordinary Least Squares**: To employ this method, its five basic assumptions/conditions (Linear relationship, Multivariate normality, No or little multicollinearity, No auto-correlation and Homoscedasticity) must be satisfied. Also, a note about sample size is important. In many cases, at least 20 cases per independent variable in the analysis. It finds the best fit for a set of data points by minimizing the sum of the offsets or residuals of points from the plotted curve and predict the behavior of dependent variables (*Dougherty, 2002; Gujarati and Porter, 2009; Hill, et.al., 2008; Wooldridge, 2008)*
2. M**aximum likelihood estimation** (**MLE**): There exist a set of distributions indexed by a parameter that could have generated the sample, the **likelihood** is a **function** that associates with each parameter the probability of observing the given sample. It find the parameter values that give the distribution that maximise the probability of observing the data (Rossi, 2018; Hendry and Nielsen, 2007; Chambers, et al.,2012; Ward and Ahlquist, 2018; Press et al.,1992; Myung, 2003; Gourieroux and Monfort, 1995; Kane, 1968)
3. M**ethod of moments.** They **moment conditions** are functions of the model parameters and the data, such that their

Scientia Africana, Vol. 18 (No. 2), August, 2019. Pp 173-182
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

expectation is zero at the parameters' true values. The GMM method then minimizes a certain norm of the sample averages of the **moment conditions** by Pafnuty Chebyshev in 1887**Moments** are a set of **statistical** parameters to measure a distribution. Four methods of **moments** that are commonly used: the first **moment** is the mean, the second central **moment** is the variance, (Variance: Standard deviation is the square root of the variance: an indication of how closely the values are spread about the mean), the third standardized **moment** is the skewness, and the fourth standardized **moment** is the kurtosis. **Moments** about mean are used to ascertain the **moments** of the distribution about mean.

The parameter estimation methods; **ordinary least squares** (**OLS**), M**aximum likelihood estimation** (**MLE**) and M**ethod of moments are** used in estimating the parameters of a statistical model and for fitting a statistical model to data. In this study, resampling will also be considered. **Resampling** is the practice of estimating properties of an estimator by measuring those properties when sampling from an approximating distribution. One standard choice for an approximating distribution is the empirical distribution of the observed data. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of resamples of the observed dataset and of equal size to the observed dataset, each of the resampling methods is obtained by random sampling with replacement. For more details: Efron (2000); Efron and Tibshirani (1993), Gonzalez-Manteiga and Crujeiras (2008), Freedman (1981), Lahiri (2006), Bowman and Shenton (1998), Munkhammar, et.al. (2017), Good (2004), Hall and Maiti (2006), Quenouille (1956), Rossi (2018), Ward and Ahlquist (2018*)*, *Wooldridge (2008)*, Xu (2008), Hall, Lee and Park (2009), Mahiane et.al. (2010), Paparoditis, and Politis, (2005), Andrews and Buchinsky (2000), Andrews and Guggenberger (2009), Tanner (2012), R - Core Team (2017), Pawitan (2000), Mita et al (2012), Mashreghi et al (2016), Hirukawa (2015), Martins-Filho, et al (2016), Torsen, et al (2018), Torsen, et al (2019).

RESEARCH METHODOLOGY

i. The data set x for a classic linear regression model consist of n points $x_1, x_2, \ldots, x_n$, where each x_i is itself a pair, say

$$x_i = (c_i, y_i) \quad 3.1$$

where c_i is a 1xp vector $c_i = (c_{i1}, c_{i2}, \ldots, c_{ip})$ called the covariate vector or predictor, while y_i is a real number called the response. Let μ_i indicate the conditional expectation of ith response y_i given the predictor c_i. The probability structure of the linear model is usually expressed as

$$y_i = c_i\beta + \epsilon_i \quad \forall \quad i=1,2,\ldots,n \quad 3.3$$

the error terms ϵ_i in (3.3) are assumed to be a random sample from an unknown error distribution.

Linear regression model in vector form,

$$y_i = x_i^T\beta + \epsilon_i \quad 3.4$$

where β is a $p\times1$ vector of unknown parameters; the ε_i's are unobserved scalar random variables which account for influences upon the responses y_i from sources other than the explanatory variables x_i; and x_i is a column vector of the *i*th observations of all the explanatory variables.

$F \rightarrow (\beta_1, \beta_2, \ldots, \beta_p) = \epsilon \quad [E_F(\epsilon) = 0]$ 3.5

We want to estimate the regression parameter vector β from the observed data $(c_1, y_1), (c_2, y_2), \ldots, (c_n, y_n)$. let C be the n x p matrix with ith row c_i (the design matrix), and let y be the vector $(y_1, y_2, \ldots, y_p)^T$. then the least squares estimate is the solution to the normal equations;

$$\hat{\beta} = (C^TC)^{-1}C^Ty \qquad 3.6$$

where C is of full rank p. Furthermore, xi's from the original sample was used and the associated yi's were generated by a random draw y| xi from F(xi, $\hat{\theta}$) . That is, N resampled samples of x, call each of them xi* was drawn, where i = 1, 2, … N, and generate the associated y's by random draws from F(xi*, $\hat{\theta}$).

ii. The data set x for a method of moments with a sample of size n is drawn, resulting in the values w_1 ... w_n. For j = 1, …, k

let $\widehat{\mu_1} = \frac{1}{n}\sum_{i=1}^{n} w_i^j$ 3.7

be the *j*-th sample moment, an estimate of μ_i.

iii. The data set x for a maximum likelihood estimate that has a random sample $X_1, X_2, \ldots, X_n$ for which the probability density (or mass) function of each X_i is $f(x_i; \theta)$. Then, the joint probability mass (or density) function of $X_1, X_2, \ldots, X_n$, which we'll (not so arbitrarily) call $\boldsymbol{L(\theta)}$, the likelihood function is defined as

$$L(\theta; y) = f(y_1; \theta)f(y_2; \theta) \ldots f(y_n; \theta) = \prod_{i=1}^{n} f(y_i; \theta) \qquad 3.8$$

Furthermore, resampling can be applied to more general regression to make a statistical inference. Efron and Tibshirani, (1993), suggested resampling algorithm for estimating standard errors and bias from regression models as shown below;

1. Select B independent resampling samples $x^{*1}, x^{*2}, \ldots, x^{*B}$ each consisting of n data values drawn with replacement from x, for estimating a standard error, the number B will ordinarily be in the range 25 – 200.

2. Evaluate the resampling replication corresponding to each resampling sample, $\hat{\theta}^*(b) = s(x^{*b})$ b = 1,2,…,B. 3.9

3. The resampling estimate of standard error is the standard deviation of the resampling (B) replications:

$$\hat{se}_{boot} = \{\sum_{b=1}^{B}[s(\mathbf{x}^{*b}) - s(\cdot)]^2/(B-1)\}^{1/2} \qquad 3.10$$

where $s(\cdot) = \sum_{b=1}^{B} s(\mathbf{x}^{*b})/B$

ANALYSIS AND DISCUSSION

The main aim of this section is to analyze the results in order to make necessary policy deductions from them.

A. Case 1: Parametric Model

➢ The three methods of parameter estimation in regression namely (OLS), (MLE) and (MOM)] on the original data set.

Scientia Africana, Vol. 18 (No. 2), August, 2019. Pp 173-182
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

Test Hypothesis

Ho: Nigerian financial sector does contribute significantly to economic growth.

Equation Estimated:

GDPt = bo + b1IM + b2EX + e

From R-Statistical package the OLS model (R311) obtained is

GDPt = 4.235e+04 + 1.450e+00IM + 1.730e+00EX 4.1

Standard error (2.055e+05) (3.235e-01) (1.841e-01)

From R-Statistical package the MLE model (R312) obtained from the hessian matrix is

GDPt = 322637.3694+ 1.7115308348b_1+ 1.4078314520 b_2 4.2

Standard error (205473.5) (0.323530) (0.184102)

From R-Statistical package the MOM (R313) model obtained is

GDPt = 4.2349e+04 + 1.7302e+00zm_1 + 1.4500e+00zm_2 4.3

Std. Error (9.5280e+04) (2.0010e-01) (3.7148e-01)

B. Case 2: Parametric Resampling Model Tested Hypothesis

Ho: Nigerian financial sector contributes significantly to economic growth.

Equation Estimated:

GDPt = bo + b1IM + b2EX + e

From R-Statistical package the OLS model (B311) obtained when n = 1000 is

GDPt = 5.145e+04 + 1.601e+00EX +1.596e+00IM 4.4

Std. Error (1.585e+05) (1.235e-01) (2.129e-01)

The models in cases 1&2, which are, equation 4.1, 4.2, 4.3, and 4.4 respectively indicate positive relationship between GDP and Nigeria financial sector. The R311 model using OLSM and R312 model using Cue Method (CM) are exactly the same value when approximated but there is a little difference in R313 model using secant method (SM). Moreover, R311 model was the best among the three models when compared in terms of their coefficients of the independent variables (β's) and standard errors. Since the R311 model was the best among the three models, it was further resampled. In fact, comparing in terms of their coefficients of the independent variables of the Parametric models and Parametric resampled models using standard errors to ascertain the best under several conditions. The resampled model (R311) turned out to be better than R311 model.

C. CASE 3: Selection of the Best Model Based on other Evaluation Criteria under several conditions of the original data set gotten from the OLS, MLE, MOM and resampled models of the SLR equation.

Table 3: Selection of the Best Model Based on other Evaluation Criteria

	R^2	R^2_{adj}	AIC	HQIC	SBIC	Convergence
Parametric_{OLS}	0.8956	0.8099	86.8222	90.8659	90.1692	1
Parametric_{MLE}	0.8685	0.7091	88.3971	86.2802	87.3139	1
Parametric_{MOM}	0.7985	0.8096	91.2413	87.4069	86.9804	1
Parametric_{BOOT}	0.9295*	0.9529*	86.9545*	85.4435*	86.7872*	0

It is pertinent to note that when the coefficient of determination (adjusted R^2) is high, it shows that the model is a reasonable fit of the relationship among the variable. It also confirms its efficiency in prediction. The model hence explains 95% of the behavior of financial sector in Nigeria. Since the probability (F-statistic) of 0.00 is less than the chosen 5% significance level we cannot accept the null hypotheses 1 and 2. The alternative hypotheses that the financial sector in Nigerian economy is significant are accepted instead. The best set of parameters found in parametric resampled information criteria and the function minimizes with first argument the vector of parameters over which minimization is to take place and the result must be a scalar. However, among all the Models based on other evaluation criteria, parametric resampled information converges at zero, which indicates successful completion of the resampling analysis compared to others. This follows an important theorem which states that for a given iterative method and its iteration matrix C it is convergent if and only if its spectral radius $p(C)$ is smaller than unity. The spectral radius of a square matrix or a bounded linear operator is the largest absolute value of its eigenvalues, that is, $p(C) \leq 1$. Iteration matrix $p(C)$ is called *convergent* as the number of iterations occurs before the stationary value, (Chernick, 2008 and Lax, 2002).

Scientia Africana, Vol. 18 (No. 2), August, 2019. Pp 173-182
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

D. kernel density and quantile-quantile plots of the original and resampled data set.

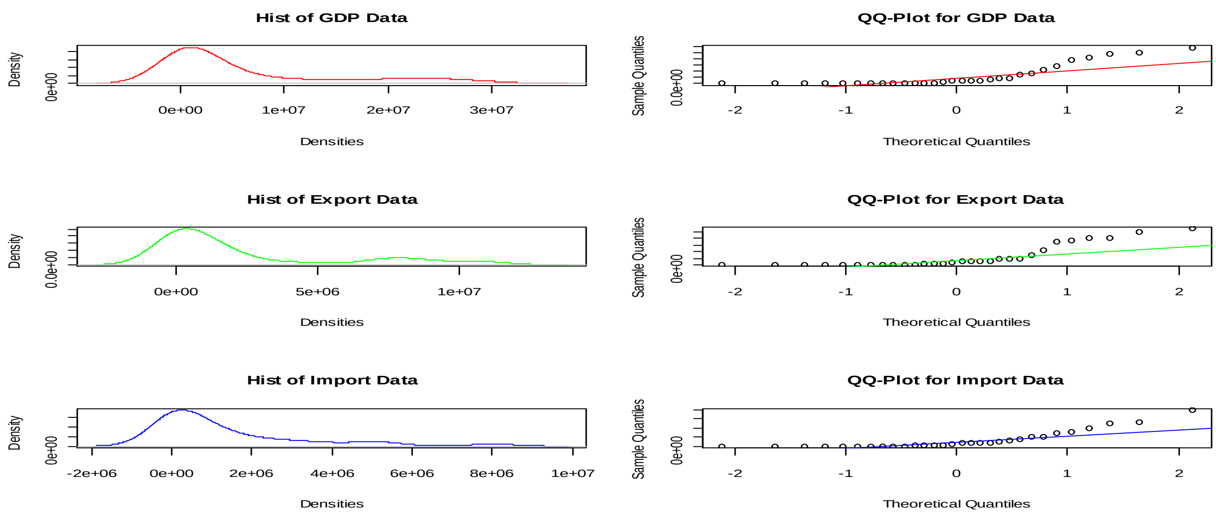

Figure 1: The kernel density and quantile-quantile plots of the original data set from the Nigerian financial sector statistics.

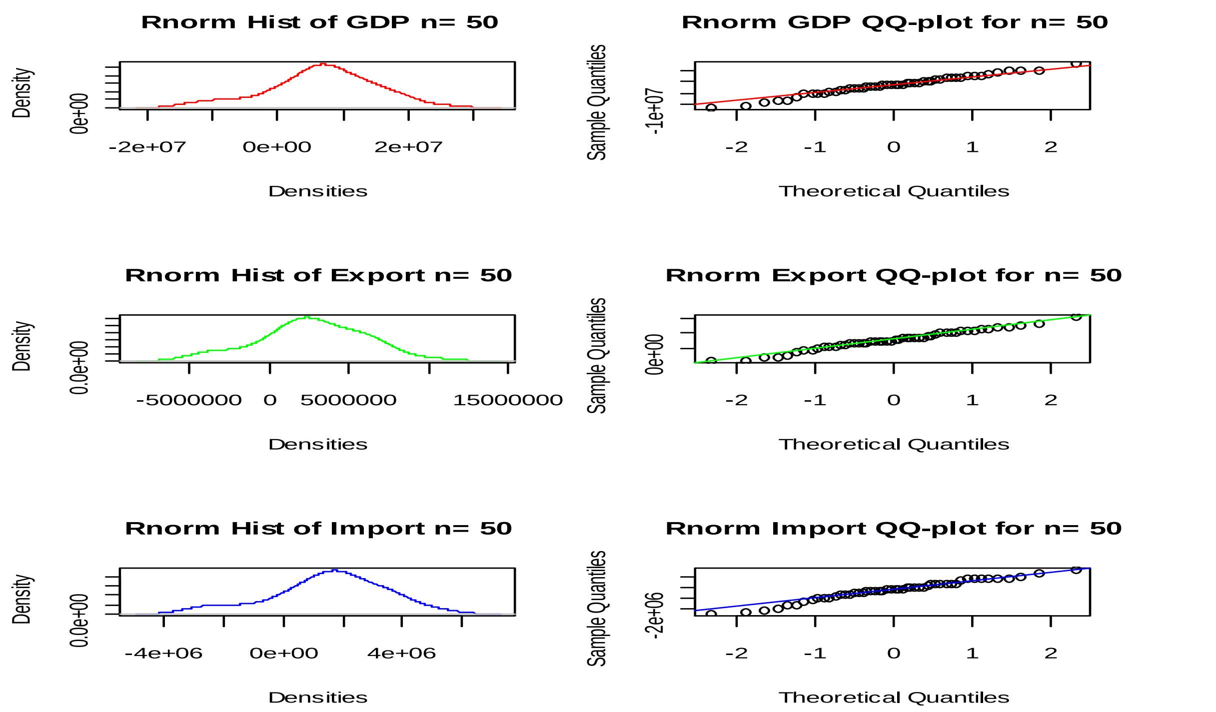

Figure 2: The kernel density and quantile-quantile plots of the resampling data set from the Nigerian financial sector statistics.

- **To identify the sampling distribution of a statistic θ (the mean)**:

Here the shape, center and spread of the kernel density and quantile-quantile plots of the original data set and the original when resampled are being considered. The kernel density and the qq-plots in figures 1 & 2 show that there is variation among the means. It also shows that the distribution possess the following characteristics; It is not symmetric around the mean, It is skewed to the right in small samples and finally, the mean value and the variance are not equal to zero and one respectively, (Chernick , 2008).

Statistical theory shows that any distribution with the following properties is a Chi-square distribution. Therefore, the approximate distribution of the financial sector statistics in Nigeria is a Chi-square distribution and it is denoted by χ^2.

CONCLUSION

This study reveals that the parametric resampled model (R311) turned out to be better than other parametric models under various conditions. Also, the parametric estimation methods, parametric resampled method and the information criteria confirmed that model (R311) is the best. The kernel density and qq plot the financial sector statistics distribution is not symmetric around the mean, skewed to the right in small samples and finally, the mean value and the variance are not equal to zero and one respectively, which is in line with the properties of the Chi-square distribution. Hence, the approximate distribution of the financial sector statistics in Nigeria has been found to be a Chi-square distribution. Apart from the result showing a good fit , the financial sector statistics sampling distribution in Nigeria of the original dataset and also when resampled was ascertained. The study also shows that import and export are strong determinants of economic growth as the nation works toward actualization of number 8^{th} of 17 sustainable development goals (SDGs) to determine human activities and engagement till 2030.

REFERENCES

Andrews DW, Buchinsky M (2000) A three-step method for choosing the number of bootstrap repetitions. Econometrica 68(1):23–51

Andrews DW, Guggenberger P (2009) Incorrect asymptotic size of subsampling procedures based on post-consistent model selection estimators. Journal of Econometrics 152(1):19–27

Bowman, K. O. and Shenton L. R., (1998). "Estimator: Method of Moments", pp 2092–2098, Encyclopedia of statistical sciences, Wiley.

Chambers, R. L., Steel, D. G., Wang, S., Welsh, A. (2012). Maximum Likelihood Estimation for Sample Surveys. Boca Raton: CRC Press. ISBN 978-1-58488-632-7.

Chernick, M. R. (2008) Bootstrap Methods: A guide for practitioners and researchers. Wiley Series in Probability and Statistics, Second Edition.

Dougherty, C. (2002). Introduction to Econometrics (2nd ed.). New York: Oxford University Press. pp. 48–113. *ISBN* *0-19-877643-8*.

Efron, B. (2000). The resampling and modern statistics. J. Am. Statist. Assoc. 95, 1293-1296.

Scientia Africana, Vol. 18 (No. 2), August, 2019. Pp 173-182
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

Efron, B. and Tibshirani, R.J. (1993). An Introduction to the Resampling. Chapman and Hall, New York.

Freedman, D.A. (1981) Resamplingping regression models, Ann. Statist. 6, 1218- 1228.

Gonzalez-Manteiga, W. and Crujeiras, R.M. (2008). A review on goodness-of-fit tests for regression models. Pyrenees International Workshop on Statistics, Probability and Operations Research: SPO 2007, 21-59.

Good, P. (2004). Permutation, Parametric, and Resampling Tests of Hypotheses. 3rd Edition. Springer-Verlag, New York.

Gourieroux, C. and Monfort, A. (1995). Statistics and Econometrics Models. Cambridge University Press. p. 161. ISBN 0-521-40551-3.

Gujarati D, and Porter, D. C. (2009). Basic Econometics (Fifth ed.). Boston: McGraw-Hill Irwin. pp. 55–96. ISBN 978-0-07-337577-9.

Hall, P. and Maiti, T. (2006). On parametric resampling methods for small area prediction. Journal of the Royal Statistical Society. Series B 68, 221-238.

Hall, P. Lee, E.R. and Park, B.U. (2009) Resampling-based penalty choice for the Lasso achieving oracle performance. Statistica Sinica, 19, 449-471.

Hendry, D.F. and Nielsen, B. (2007). Econometric Modeling: A Likelihood Approach. Princeton: Princeton University Press. ISBN 978-0-691-13128-3.

Hill, R. C.; Griffiths, W. E.; Lim, G. C. (2008). Principles of Econometrics (3rd ed.). Hoboken, NJ: John Wiley & Sons. pp. 8–47. ISBN 978-0-471-72360-8.

Hirukawa, M (2015) "Independence of the sample mean and variance for normal distributions: a proof by induction," Pontian Economic Research, 5(1) 1–5.

Kane, E. J. (1968). Economic Statistics and Econometrics. New York: Harper & Row. p. 179.

Lahiri, S. N. (2006). Resampling Methods: A Review. In Frontiers in Statistics (J. Fan and H.L. Koul, editors) 231-265, Imperial College Press, London.

Lax, P. D. (2002), *Functional Analysis*, Wiley-Interscience, ISBN 0-471-55604-1

Mahiane, S. G., Nguema, E.-P. Ndong, Pretorius, C., and Auvert, B. (2010). Mathematical models for coinfection by two sexually transmitted agents: the human immunodeficiency virus and herpes simplex virus type 2 case. Appl. Statist 59, 547-572.

Martins-Filho, C, Yao, F. and Torero, M (2016) "Nonparametric estimation of conditional value-at-risk and expected shortfall based on extreme value theory," Econometric Theory, 34 (1) 1–45.

Mashreghi, Z., Haziza, D., and Léger, C. (2016). A survey of bootstrap methods in finite population sampling. Statistics Surveys, 10, 1-52

Mita N, Jiao J, Kani K, Tabuchi A, Hara H (2012) The parametric and non-parametric bootstrap resamplings for the visual acuity measurement. Kawasaki Journal of Medical Welfare 18:19–28

Munkhammar, J, Mattsson, L. and Rydén J (2017) "Polynomial probability distribution estimation using the method of moments". PLoS ONE 12(4): e0174573. https://doi.org/10.1371/journal.pone.0174573

Myung, I. J. (2003). "Tutorial on Maximum Likelihood Estimation". Journal of mathematics Psychology. **47** (1): 90–100. doi:10.1016/S0022-2496(02) 00028-7.

Papa`roditis, E. and Politis, D. (2005). Resampling hypothesis testing in regression models, Statistics and Probability Letters, 74: 356-365.

Pawitan Y (2000) Computing empirical likelihood from the bootstrap. Statistics & Probability Letters 47(4):337– 345

Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W.T (1992). “Least squares as a maximum likelihood estimator” Numerical Recipes in FORTRAN: The Art of Scientific Computing (2nd ed.). Cambridge: Cambridge University Press. pp. 651–655. ISBN 0-521-43064-X.

Quenouille, M. H. (1956). Notes on bias in estimation. Biometrika 43: 353-360.

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/

Rossi, R. J. (2018). Mathematical Statistics: An Introduction to Likelihood Based Inference. New York: John Wiley & Sons. p. 227. ISBN 978-1-118-77104-4.

Tanner, M, A. (2012) Tools for statistical inference. Springer, New York

Torsen, E. Mwita, P. N and Mungatu, J. K. (2019) “A three-step nonparametric estimation of conditional value-at-risk admitting a location-scale model,” Journal of Statistical and Econometric Methods, 8(4),1-24.ISSN: 1792-6602 (print), 1792-6939 (online).

Torsen, E., Mwita, P. N., and Mungatu, J. K. (2018) “Nonparametric estimation of the error functional of a location-scale model,” Journal of Statistical and Econometric Methods, 7 (4)18-26.

Ward, M. D, and Ahlquist, J. S. (2018). Maximum Likelihood for Social Science: Strategies for Analysis. New York: Cambridge University Press. ISBN 978-1-107-18582-1.

Wooldridge, J. (2008). The simple regression model. Introductory Econometrics: A Modern Approach (4th ed.). Mason, OH: Cengage Learning. pp. 22–67. ISBN 978-0-324-58162-1

Xu, K. (2008). Resampling autoregression under nonstationary volatility, Econometrics Journal, 11: 1-26.