Regression

1- Renewable quantile regression for streaming datasets

By:
Wang, KN (Wang, Kangning) [1] ; Wang, HW (Wang, Hongwei) [1] ; Li, SM (Li, Shaomin) [2] , [3]

Volume
235
Article Number
107675
DOI
10.1016/j.knosys.2021.107675
Published
JAN 10 2022
Early Access
NOV 2021
Indexed
2021-11-21
Document Type
Article

Abstract

Streaming data analysis has drawn much attention, where large amounts of data arrive in streams. Because limited memory can only store a small batch of data, fast analysis without access to the historical data is necessary. Quantile regression has been widely used in many fields because of its robustness and comprehensiveness. However, in the streaming data environment, it is challenging to implement quantile regression by the conventional methods, because they are all based on the assumption that the memory can fit all the data. To fix this issue, this paper proposes a novel online renewable quantile regression strategy, in which the resulting estimator is renewed with current data and summary statistics of historical data. Thus, it is computationally efficient, and not storage-intensive. What is more, the theoretical results also confirm that the proposed estimator is asymptotically equivalent with the oracle estimator calculated using the entire data together. Numerical experiments on both synthetic and real data verify the theoretical results and illustrate the good performance of the new method. (C) 2021 Elsevier B.V. All rights reserved.

Keywords
Author Keywords
Streaming data environmentQuantile regressionOnline updating learning
Keywords Plus
SELECTION

Regression

2- Gaussian Process Regression for Materials and Molecules

By:
Deringer, VL (Deringer, Volker L.) [5] ; Bartok, AP (Bartok, Albert P.) [6] , [7] ; Bernstein, N (Bernstein, Noam) [1] ; Wilkins, DM (Wilkins, David M.) [2] ; Ceriotti, M (Ceriotti, Michele) [3] , [4] ; Csanyi, G (Csanyi, Gabor) [8]

Volume
121
Issue
16
Page
10073-10141
DOI
10.1021/acs.chemrev.1c00022
Published
AUG 25 2021
Early Access
AUG 2021
Indexed
2021-09-06
Document Type
Review

Abstract

We provide an introduction to Gaussian process regression (GPR) machinelearning methods in computational materials science and chemistry. The focus of the present review is on the regression of atomistic properties: in particular, on the construction of interatomic potentials, or force fields, in the Gaussian Approximation Potential (GAP) framework; beyond this, we also discuss the fitting of arbitrary scalar, vectorial, and tensorial quantities. Methodological aspects of reference data generation, representation, and regression, as well as the question of how a data-driven model may be validated, are reviewed and critically discussed. A survey of applications to a variety of research questions in chemistry and materials science illustrates the rapid growth in the field. A vision is outlined for the development of the methodology in the years to come.

Keywords
Keywords Plus
DENSITY-FUNCTIONAL THEORYPHASE-CHANGE MATERIALSPOTENTIAL-ENERGY SURFACESMACHINE LEARNING-MODELSX-RAY SPECTROSCOPYAB-INITIOAMORPHOUS-CARBONINTERATOMIC POTENTIALSELECTRON-DENSITYCOMBINING EXPERIMENTS

3- NONPARAMETRIC REGRESSION USING DEEP NEURAL NETWORKS WITH RELU ACTIVATION FUNCTION

By:
Schmidt-Hieber, J (Schmidt-Hieber, Johannes) [1]

Volume
48
Issue
4
Page
1875-1897
DOI
10.1214/19-AOS1875
Published
AUG 2020
Indexed
2020-09-01
Document Type
Article

Abstract

Consider the multivariate nonparametric regression model. It is shown that estimators based on sparsely connected deep neural networks with ReLU activation function and properly chosen network architecture achieve the minimax rates of convergence (up to log n-factors) under a general composition assumption on the regression function. The framework includes many well-studied structural constraints such as (generalized) additive models. While there is a lot of flexibility in the network architecture, the tuning parameter is the sparsity of the network. Specifically, we consider large networks with number of potential network parameters exceeding the sample size. The analysis gives some insights into why multilayer feedforward neural networks perform well in practice. Interestingly, for ReLU activation function the depth (number of layers) of the neural network architectures plays an important role, and our theory suggests that for nonparametric regression, scaling the network depth with the sample size is natural. It is also shown that under the composition assumption wavelet estimators can only achieve suboptimal rates.

Keywords
Author Keywords
Nonparametric regressionmultilayer neural networksReLU activation functionminimax estimation riskadditive modelswavelets
Keywords Plus
MULTILAYER FEEDFORWARD NETWORKSUNIVERSAL APPROXIMATIONBOUNDSINEQUALITIESERROR

Regression

4- Molecular and cellular mechanisms of liver fibrosis and its regression

By:
Kisseleva, T (Kisseleva, Tatiana) [1] ; Brenner, D (Brenner, David) [2]

Volume
18
Issue
3
Page
151-166
DOI
10.1038/s41575-020-00372-7
Published
MAR 2021
Early Access
OCT 2020
Indexed
2020-11-17
Document Type
Review

Abstract

Chronic liver injury leads to liver inflammation and fibrosis, through which activated myofibroblasts in the liver secrete extracellular matrix proteins that generate the fibrous scar. The primary source of these myofibroblasts are the resident hepatic stellate cells. Clinical and experimental liver fibrosis regresses when the causative agent is removed, which is associated with the elimination of these activated myofibroblasts and resorption of the fibrous scar. Understanding the mechanisms of liver fibrosis regression could identify new therapeutic targets to treat liver fibrosis. This Review summarizes studies of the molecular mechanisms underlying the reversibility of liver fibrosis, including apoptosis and the inactivation of hepatic stellate cells, the crosstalk between the liver and the systems that orchestrate the recruitment of bone marrow-derived macrophages (and other inflammatory cells) driving fibrosis resolution, and the interactions between various cell types that lead to the intracellular signalling that induces fibrosis or its regression. We also discuss strategies to target hepatic myofibroblasts (for example, via apoptosis or inactivation) and the myeloid cells that degrade the matrix (for example, via their recruitment to fibrotic liver) to facilitate fibrosis resolution and liver regeneration.
Chronic liver injury leads to liver inflammation and fibrosis, through which activated myofibroblasts in the liver secrete extracellular matrix proteins that generate the fibrous scar. This Review summarizes studies of the molecular mechanisms underlying liver fibrosis and its reversibility.

Keywords

Regression

Keywords Plus

HEPATIC STELLATE CELLSMESENCHYMAL STEM-CELLSPLASMINOGEN-ACTIVATOR INHIBITORSINUSOIDAL ENDOTHELIAL-CELLSWEIGHT HYALURONAN PROMOTESMARROW-DERIVED FIBROCYTESRENIN-ANGIOTENSIN SYSTEMTGF-BETAGROWTH-FACTORNONALCOHOLIC STEATOHEPATITIS

Regression

5- SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes

By:
Elosua-Bayes, M (Elosua-Bayes, Marc) [1] ; Nieto, P (Nieto, Paula) [1] ; Mereu, E (Mereu, Elisabetta) [1] ; Gut, I (Gut, Ivo) [1] , [2] ; Heyn, H (Heyn, Holger) [1] , [2]
(provided by Clarivate)

Volume
49
Issue
9
Article Number
e50
DOI
10.1093/nar/gkab043
Published
MAY 21 2021
Early Access
FEB 2021
Indexed
2021-06-18
Document Type
Article

Abstract

Spatially resolved gene expression profiles are key to understand tissue organization and function. However, spatial transcriptomics (ST) profiling techniques lack single-cell resolution and require a combination with single-cell RNA sequencing (scRNA-seq) information to deconvolute the spatially indexed datasets. Leveraging the strengths of both data types, we developed SPOTlight, a computational tool that enables the integration of ST with scRNA-seq data to infer the location of cell types and states within a complex tissue. SPOTlight is centered around a seeded non-negative matrix factorization (NMF) regression, initialized using cell-type marker genes and non-negative least squares (NNLS) to subsequently deconvolute ST capture locations (spots). Simulating varying reference quantities and qualities, we confirmed high prediction accuracy also with shallowly sequenced or small-sized scRNA-seq reference datasets. SPOTlight deconvolution of the mouse brain correctly mapped subtle neuronal cell states of the cortical layers and the defined architecture of the hippocampus. In human pancreatic cancer, we successfully segmented patient sections and further fine-mapped normal and neoplastic cell states. Trained on an external single-cell pancreatic tumor references, we further charted the localization of

clinical-relevant and tumor-specific immune cell states, an illustrative example of its flexible application spectrum and future potential in digital pathology.

[GRAPHICS]

Keywords
Keywords Plus
NONNEGATIVE MATRIX FACTORIZATIONGENE-EXPRESSIONSEQ

Regression

6- Using Predictions and Marginal Effects to Compare Groups in Regression Models for Binary Outcomes

By:
Long, JS (Long, J. Scott) [1] , [2] ; Mustillo, SA (Mustillo, Sarah A.) [3]
(provided by Clarivate)

Volume
50
Issue
3
Page
1284-1320
Special Issue
SI
DOI
10.1177/0049124118799374
Published
AUG 2021
Indexed
2021-08-04
Document Type
Article

Abstract

Methods for group comparisons using predicted probabilities and marginal effects on probabilities are developed for regression models for binary outcomes. Unlike approaches based on the comparison of regression coefficients across groups, the methods we propose are unaffected by the scalar identification of the coefficients and are expressed in the natural metric of the outcome probability. While we develop our approach using binary logit with two groups, we consider how our interpretive framework can be used with a broad class of regression models and can be extended to any number of groups.

Keywords
Author Keywords
logistic regressionprobabilitiesmarginal effectsgroup differencesinteractionsprobit
Keywords Plus
PROBIT COEFFICIENTSLOGITOBESITYCHOICELEVEL

Regression

7- Nexus between air pollution and NCOV-2019 in China: Application of negative binomial regression analysis

By:
Iqbal, W (Iqbal, Wasim) [1] ; Tang, YM (Tang, Yuk Ming) [2] , [3] ; Chau, KY (Chau, Ka Yin) [3] ; Irfan, M (Irfan, Muhammad) [4] , [5] ; Mohsin, M (Mohsin, Muhammad) [6]
(provided by Clarivate)

Volume
150
Page
557-565
DOI
10.1016/j.psep.2021.04.039
Published
JUN 2021
Early Access
MAY 2021
Indexed
2021-06-11
Document Type
Article

Abstract

On a global scale, the epidemic of the novel coronavirus (NCOV-2019) has become a major issue that is seriously harming human health and impairing the environment's quality. The current study examines the association between air pollution and NCOV-2019 in China, where cases of NCOV-2019 are correlated with deaths in public databases with data on air pollution tracked at multiple locations in different provinces of China. A negative binomial regression (NBR) model was applied to examine the difference between the number of people infected with NCOV-2019 and the number of deaths in China. The findings show that, after population density regulation, there is a positive connection between air pollutants concentration (particularly nitrogen dioxide) and the number of NCOV-2019 cases and deaths. Furthermore, PM2.5 is the key cause of NCOV-2019 cases and deaths in China. The results indicate that a 1% increase in the average of PM2.5 was correlated with an increase of 11.67 % in NCOV-2019 cases and a rise of 18 % in NCOV-2019 deaths. We concluded that a slight rise in air pollution has caused the number of NCOV-2019 cases and deaths to increase dramatically. This research provides a basis for future policies affected by this pandemic in terms of health and pollution.

Keywords

Regression

Author Keywords
COVID-19Air pollutionPM25SO2PM10NO2O3Negative binomial regression
Keywords Plus
TERM EXPOSURECOVID-19POISSONDESIGN

Regression

8- Ecological footprint, economic complexity and natural resources rents in Latin America: Empirical evidence using quantile regressions

By:
Alvarado, R (Alvarado, Rafael) [1] , [2] ; Tillaguango, B (Tillaguango, Brayan) [3] , [4] ; Dagar, V (Dagar, Vishal) [5] ; Ahmad, M (Ahmad, Munir) [6] ; Isik, C (Isik, Cem) [7] ; Mendez, P (Mendez, Priscila) [8] ; Toledo, E (Toledo, Elisa) [9]
(provided by Clarivate)

Volume
318
Article Number
128585
DOI
10.1016/j.jclepro.2021.128585
Published
OCT 10 2021
Early Access
AUG 2021
Indexed
2021-10-02
Document Type
Article

Abstract

This paper analyses the environmental degradation associated with the ecological footprint in Latin America. The studied region has an important feature of less industrialization and are having high advantage of biodiversity. As ecological footprint is a comprehensive indicator of environmental deterioration because it assesses the impact of all human activities on nature. This research examines the effect of economic complexity and the natural resources rents on the per capita ecological footprint in Latin America. To assess the sensitivity of the parameters and to capture the impact of the growth in the structure of the industrial sector for the region, the study includes globalization, inequality, internal credit and trade as additional covariates. The theoretical framework comprises the Environmental Kuznets Curve (EKC) and for the methodology the quantile regression approach, proposed by Canay (2011), Powell (2016) and Machado and Santos Silva (2019), has been used in this study. The study found that the impact of economic complexity and the natural resources rents is heterogeneous throughout the distribution on the ecological footprint. It is also observed that the level of inequality for the utilization of region-specific climatic distribution of natural resources, increases the ecological footprint in the lower quantiles, while in the upper quantiles, it decreases the same. This study provides a new approach to analyze the degradation of environment in countries with high dependence on natural resources rents and highincome inequality.

Regression

Keywords
Author Keywords
Ecological footprintEconomic complexityNatural resources rentsQuantile regressionEnvironmental kuznets curve (EKC)
Keywords Plus
ENVIRONMENTAL KUZNETS CURVECO2 EMISSIONSGROWTHHYPOTHESIS

9- On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data

By:
Imai, K (Imai, Kosuke) [1] , [2] ; Kim, IS (Kim, In Song) [3]
(provided by Clarivate)

Volume
29
Issue
3
Page
405-415
Article Number
PII S1047198720000339
DOI
10.1017/pan.2020.33
Published
JUL 2021
Indexed
2021-06-17
Document Type
Article

Abstract

The two-way linear fixed effects regression (2FE) has become a default method for estimating causal effects from panel data. Many applied researchers use the 2FE estimator to adjust for unobserved unit-specific and time-specific confounders at the same time. Unfortunately, we demonstrate that the ability of the 2FE model to simultaneously adjust for these two types of unobserved confounders critically relies upon the assumption of linear additive effects. Another common justification for the use of the 2FE estimator is based on its equivalence to the difference-in-differences estimator under the simplest setting with two groups and two time periods. We show that this equivalence does not hold under more general settings commonly encountered in applied research. Instead, we prove that the multi-period difference-in-differences estimator is equivalent to the weighted 2FE estimator with some observations having negative weights. These analytical results imply that in contrast to the popular belief, the 2FE estimator does not represent a design-based, nonparametric estimation strategy for causal inference. Instead, its validity fundamentally rests on the modeling assumptions.

Regression

Keywords
Author Keywords:
difference-in-differenceslongitudinal datamatchingunobserved confoundingweighted least squares

Regression

10- Generalizing hierarchical and variation partitioning in multiple regression and canonical analyses using the rdacca.hp R package

By:
Lai, JS (Lai, Jiangshan) [1] , [2] ; Zou, Y (Zou, Yi) [3] ; Zhang, JL (Zhang, Jinlong) [4] ; Peres-Neto, PR (Peres-Neto, Pedro R.) [5] , [6]
(provided by Clarivate)

Volume
13
Issue
4
Page
782-788
DOI
10.1111/2041-210X.13800
Published
APR 2022
Indexed
2022-04-18
Document Type
Article

Abstract

Canonical analysis, a generalization of multiple regression to multiple-response variables, is widely used in ecology. Because these models often involve many parameters (one slope per response per predictor), they pose challenges to model interpretation. Among these challenges, we lack quantitative frameworks for estimating the overall importance of single predictors in multi-response regression models. Here we demonstrate that commonality analysis and hierarchical partitioning, widely used for both estimating predictor importance and improving the interpretation of single-response regression models, are related and complementary frameworks that can be expanded for the analysis of multiple-response models. In this application, we (a) demonstrate the mathematical links between commonality analysis, variation and hierarchical partitioning; (b) generalize these frameworks to allow the analysis of any number of predictor variables or groups of predictor variables as in the case of variation partitioning; and (c) introduce and demonstrate the implementation of these generalized frameworks in the R package rdacca.hp.

Regression

Keywords
Author Keywords
averaging over orderingsCCAcommonality analysisconstrained ordinationdb-RDAexplained variationRDArelative importance
Keywords Plus
RELATIVE IMPORTANCEDOMINANCE ANALYSISPREDICTORS

Regression

11- How external debt led to economic growth in South Asia: A policy perspective analysis from quantile regression

By:
Mohsin, M (Mohsin, Muhammad) [1] ; Ullah, H (Ullah, Hafeez) [2] ; Iqbal, N (Iqbal, Nadeem) [3] ; Iqbal, W (Iqbal, Wasim) [4] ; Taghizadeh-Hesary, F (Taghizadeh-Hesary, Farhad) [5]
(provided by Clarivate)

Volume
72
Page
423-437
DOI
10.1016/j.eap.2021.09.012
Published
DEC 2021
Early Access
OCT 2021
Indexed
2021-12-12
Document Type
Article

Abstract

The study analyzes the relationship between external debt and economic growth in the South Asian region. The panel ordinary least square (OLS), fixed effect, Quantile regression, and robust output regression were used to analyze the World Bank data from 2000 to 2018. South Asian countries, Afghanistan, Bangladesh, Bhutan, India, Pakistan, Sri Lanka, Maldives, and Nepal, were included in the assessment. The analysis exhibited that external debt has a negative impact, and on the other hand, external debt stock has a positive impact on economic growth. The robust regression analysis substantiated the findings and yielded total external debt and external debt service impact of 39% and 31%, respectively. The study also showed that gross capital formation and trade openness have a positive effect on economic growth. Moreover, compared to domestic debt, Threshold analysis reveals that the external debt becomes a drag on growth and instigates a more substantial adverse effect on growth (due to the rising indebtedness of a country). Thus, the study serves as a base for policymakers and government officials to upsurge economic growth while reducing the foreign debt of the economy. Even though the presence of high borrowing costs, better institutional quality can help alleviate the adverse impact of external borrowing on growth. (C) 2021 Economic Society of Australia, Queensland. Published by Elsevier B.V.

Regression

Keywords
Author Keywords
Total external debtExternal debt stockExternal debt serviceEconomic growthCurrent account deficitTrade opennessGross capital formation
Keywords Plus
FINANCIAL DEVELOPMENTSECURITYNEXUS

Regression

12- Asymmetric impact of temperature on COVID-19 spread in India: Evidence from quantile-on-quantile regression approach

By:
Irfan, M (Irfan, Muhammad) [1] , [2] ; Razzaq, A (Razzaq, Asif) [3] , [4] ; Suksatan, W (Suksatan, Wanich) [5] ; Sharif, A (Sharif, Arshian) [6] ; Elavarasan, RM (Elavarasan, Rajvikram Madurai) [7] , [8] ; Yang, CX (Yang, Chuxiao) [1] , [2] ; Hao, Y (Hao, Yu) [1] , [2] , [9] , [10] , [11] ; Rauf, A (Rauf, Abdul) [12]
(provided by Clarivate)

Volume
104
Article Number
103101
DOI
10.1016/j.jtherbio.2021.103101
Published
FEB 2022
Indexed
2022-03-17
Document Type
Article

Abstract

The emergence of new coronavirus (SARS-CoV-2) has become a significant public health issue worldwide. Some researchers have identified a positive link between temperature and COVID-19 cases. However, no detailed research has highlighted the impact of temperature on COVID-19 spread in India. This study aims to fill this research gap by investigating the impact of temperature on COVID-19 spread in the five most affected Indian states. Quantile-on-Quantile regression (QQR) approach is employed to examine in what manner the quantiles of temperature influence the quantiles of COVID-19 cases. Empirical results confirm an asymmetric and heterogenous impact of temperature on COVID-19 spread across lower and higher quantiles of both variables. The results indicate a significant positive impact of temperature on COVID-19 spread in the three Indian states (Maharashtra, Andhra Pradesh, and Karnataka), predominantly in both low and high quantiles. Whereas, the other two states (Tamil Nadu and Uttar Pradesh) exhibit a mixed trend, as the lower quantiles in both states have a negative effect. However, this negative effect becomes weak at middle and higher quantiles. These research findings offer valuable policy recommendations.

Keywords
Author Keywords
TemperatureCOVID-19TransmissibilityQuantile-on-quantile regressionIndia

Regression

Keywords Plus
CO2 EMISSIONSMORTALITYCHINANEXUS

Regression

13- Estimating the change in soccer's home advantage during the Covid-19 pandemic using bivariate Poisson regression

By:
Benz, LS (Benz, Luke S.) [1] ; Lopez, MJ (Lopez, Michael J.) [2]

DOI
10.1007/s10182-021-00413-9
Early Access
JUL 2021
Indexed
2021-08-01
Document Type
Article; Early Access

Abstract

In wake of the Covid-19 pandemic, 2019-2020 soccer seasons across the world were postponed and eventually made up during the summer months of 2020. Researchers from a variety of disciplines jumped at the opportunity to compare the rescheduled games, played in front of empty stadia, to previous games, played in front of fans. To date, most of this post-Covid soccer research has used linear regression models, or versions thereof, to estimate potential changes to the home advantage. However, we argue that leveraging the Poisson distribution would be more appropriate and use simulations to show that bivariate Poisson regression (Karlis and Ntzoufras in J R Stat Soc Ser D Stat 52(3):381-393, 2003) reduces absolute bias when estimating the home advantage benefit in a single season of soccer games, relative to linear regression, by almost 85%. Next, with data from 17 professional soccer leagues, we extend bivariate Poisson models estimate the change in home advantage due to games being played without fans. In contrast to current research that suggests a drop in the home advantage, our findings are mixed; in some leagues, evidence points to a decrease, while in others, the home advantage may have risen. Altogether, this suggests a more complex causal mechanism for the impact of fans on sporting events.

Keywords
Author Keywords
Bivariate PoissonSoccerHome advantageCovid-19
Keywords Plus
MODELENGLISHSCORESSKILLSPORT

Regression

14- The heterogeneity of renewable energy consumption, carbon emission and financial development in the globe: A panel quantile regression approach

By:
Khan, H (Khan, Hayat) [1] ; Khan, I (Khan, Itbar) [2] ; Binh, TT (Truong Tien Binh) [1]
(provided by Clarivate)

Volume
6
Page
859-867
DOI
10.1016/j.egyr.2020.04.002
Published
NOV 2020
Indexed
2021-01-18
Document Type
Article

Abstract

The present study examines the heterogeneity of renewable energy consumption, Carbon dioxide emission and financial development in the global panel of 192 countries. Panel quantile regression has been used for tickling distributional and unobserved individual heterogeneity. The findings indicate that our variables in the model on each others are heterogeneous across quantiles. More specifically, the effect of renewable energy consumption on carbon emission is negative while financial development has increasing influence on carbon emission. Carbon emission decreases the use of renewable energy while financial development positively affects renewable energy consumption. The increasing effect of carbon emission and renewable energy consumption on financial development has also been found. Finally, the current study findings give important recommendations to policy makers. (C) 2020 The Authors. Published by Elsevier Ltd.

Keywords
Author Keywords
Renewable energy consumptionFinancial developmentCarbon emissionGlobal panelPanel quantile regression
Keywords Plus
ENVIRONMENTAL KUZNETS CURVEFOREIGN DIRECT-INVESTMENTUNIT-ROOT TESTSECONOMIC-GROWTHCO2 EMISSIONSELECTRICITY CONSUMPTIONDIOXIDE EMISSIONSTRADEIMPACTCOINTEGRATION

Regression

15- The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation

By:
Chicco, D (Chicco, Davide) [1] ; Warrens, MJ (Warrens, Matthijs J.) [2] ; Jurman, G (Jurman, Giuseppe) [3]
(provided by Clarivate)

Article Number
e623
DOI
10.7717/peerj-cs.623
Published
JUL 5 2021
Indexed
2021-07-30
Document Type
Article

Abstract

Regression analysis makes up a large part of supervised machine learning, and consists of the prediction of a continuous independent target from a set of other predictor variables. The difference between binary classification and regression is in the target range: in binary classification, the target can have only two values (usually encoded as 0 and 1), while in regression the target can have multiple values. Even if regression analysis has been employed in a huge number of machine learning studies, no consensus has been reached on a single, unified, standard metric to assess the results of the regression itself. Many studies employ the mean square error (MSE) and its rooted variant (RMSE), or the mean absolute error (MAE) and its percentage variant (MAPE). Although useful, these rates share a common drawback: since their values can range between zero and +infinity, a single value of them does not say much about the performance of the regression with respect to the distribution of the ground truth elements. In this study, we focus on two rates that actually generate a high score only if the majority of the elements of a ground truth group has been correctly predicted: the coefficient of determination (also known as R-squared or R-2) and the symmetric mean absolute percentage error (SMAPE). After showing their mathematical properties, we report a comparison between R-2 and SMAPE in several use cases and in two real medical scenarios. Our results demonstrate that the coefficient of determination (R-squared) is more informative and truthful than SMAPE, and does not have the interpretability limitations of MSE, RMSE, MAE and MAPE. We therefore suggest the usage of R-squared as standard metric to evaluate regression analyses in any scientific domain.

Keywords
Author Keywords

Regression

RegressionRegression evaluationRegression evaluation ratesCoefficient of determinationMean square errorMean absolute errorRegression analysis

Keywords Plus

ABSOLUTE ERROR MAESAMPLE-SIZEAPPROXIMATIONACCURACYR-2

Regression

16- Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression

By:
Feikin, DR (Feikin, Daniel R.) [1] ; Higdon, MM (Higdon, Melissa M.) [2] ; Abu-Raddad, LJ (Abu-Raddad, Laith J.) [4] ; Andrews, N (Andrews, Nick) [5] ; Araos, R (Araos, Rafael) [6] ; Goldberg, Y (Goldberg, Yair) [7] ; Groome, MJ (Groome, Michelle J.) [9] , [10] ; Huppert, A (Huppert, Amit) [8] ; O'Brien, KL (O'Brien, Katherine L.) [1] ; Smith, PG (Smith, Peter G.) [11] ;
(provided by Clarivate)

Volume
399
Issue
10328
Page
924-944
DOI
10.1016/S0140-6736(22)00152-0
Published
MAR 2022
Indexed
2022-04-02
Document Type
Review

Abstract

Background Knowing whether COVID-19 vaccine effectiveness wanes is crucial for informing vaccine policy, such as the need for and timing of booster doses. We aimed to systematically review the evidence for the duration of protection of COVID-19 vaccines against various clinical outcomes, and to assess changes in the rates of breakthrough infection caused by the delta variant with increasing time since vaccination.

Methods This study was designed as a systematic review and meta-regression. We did a systematic review of preprint and peer-reviewed published article databases from June 17, 2021, to Dec 2, 2021. Randomised controlled trials of COVID-19 vaccine efficacy and observational studies of COVID-19 vaccine effectiveness were eligible. Studies with vaccine efficacy or effectiveness estimates at discrete time intervals of people who had received full vaccination and that met predefined screening criteria underwent full-text review. We used random-effects meta-regression to estimate the average change in vaccine efficacy or effectiveness 1-6 months after full vaccination.

Findings Of 13 744 studies screened, 310 underwent full-text review, and 18 studies were included (all studies were carried out before the omicron variant began to circulate widely). Risk of bias, established

Regression

using the risk of bias 2 tool for randomised controlled trials or the risk of bias in non-randomised studies of interventions tool was low for three studies, moderate for eight studies, and serious for seven studies. We included 78 vaccine-specific vaccine efficacy or effectiveness evaluations (Pfizer-BioNTech-Comirnaty, n=38; Moderna-mRNA-1273, n=23; Janssen-Ad26.COV2.S, n=9; and AstraZeneca-Vaxzevria, n=8). On average, vaccine efficacy or effectiveness against SARS-CoV-2 infection decreased from 1 month to 6 months after full vaccination by 21.0 percentage points (95% CI 13 .9-29.8) among people of all ages and 20.7 percentage points (10. 2-36.6) among older people (as defined by each study, who were at least 50 years old). For symptomatic COVID- 19 disease, vaccine efficacy or effectiveness decreased by 24.9 percentage points (95% CI 13.4-41.6) in people of all ages and 32.0 percentage points (11. 0-69.0) in older people. For severe COVID-19 disease, vaccine efficacy or effectiveness decreased by 10 .0 percentage points (95% CI 6.1-15.4) in people of all ages and 9.5 percentage points (5.7-14.6) in older people. Most (81%) vaccine efficacy or effectiveness estimates against severe disease remained greater than 70% over time.

Interpretation COVID-19 vaccine efficacy or effectiveness against severe disease remained high, although it did decrease somewhat by 6 months after full vaccination. By contrast, vaccine efficacy or effectiveness against infection and symptomatic disease decreased approximately 20-30 percentage points by 6 months. The decrease in vaccine efficacy or effectiveness is likely caused by, at least in part, waning immunity, although an effect of bias cannot be ruled out. Evaluating vaccine efficacy or effectiveness beyond 6 months will be crucial for updating COVID-19 vaccine policy.

Regression

17- The asymmetric effects of renewable energy consumption and trade openness on carbon emissions in Sweden: new evidence from quantile-on-quantile regression approach

By:
Adebayo, TS (Adebayo, Tomiwa Sunday) [1] ; Rjoub, H (Rjoub, Husam) [2] ; Akinsola, GD (Akinsola, Gbenga Daniel) [3] ; Oladipupo, SD (Oladipupo, Seun Damola) [4]

(provided by Clarivate)
Volume
29
Issue
2
Page
1875-1886
DOI
10.1007/s11356-021-15706-4
Published
JAN 2022
Early Access
AUG 2021
Indexed
2021-08-15
Document Type
Article

Abstract

With the passage of time, the continued burning of fossil fuels is proving to be one of the world's most serious issues. In response, the current research aims to assess the critical linkage between carbon emissions and renewable energy, trade openness, and economic growth in Sweden utilizing a dataset from 1965 to 2019. The study applied the novel quantile-on-quantile regression (QQ) approach to assess this relationship. The main objectives are to address the following questions: (i) What are the effects of trade openness on CO2 emissions in each quantile? (ii) Does renewable energy consumption mitigates CO2 emissions in each quantile? What is the impact of economic growth on CO2 emissions in each quantile? The outcomes from the QQ approach revealed that at low and medium quantiles (0.1-0.6), the effect of trade openness on CO2 emissions is negative. Furthermore, at lower and higher quantiles (0.1-0.90) of combination of renewable energy consumption and CO2 emissions, the effect of renewable energy consumption on CO2 emissions is negative. Finally, at majority of the quantiles, the effect of economic growth on CO2 emissions is negative. Moreover, the present study applied the quantile regression (QR) approach as a robustness check. The findings of the QR validate the findings of the QQR approach. The study proposes that policy-makers in Sweden should place greater emphasis on raising public awareness regarding the issues of renewable energy since it mitigates environmental degradation.

Regression

Keywords
Author Keywords
CO2 emissionsEconomic growthTrade opennessRenewable energy consumptionQuantile-on-quantile regressionSweden
Keywords Plus
CO2 EMISSIONSECONOMIC-GROWTHNEXUS

18- The moderating role of renewable and non-renewable energy in environment-income nexus for ASEAN countries: Evidence from Method of Moments Quantile Regression

By:
Anwar, A (Anwar, Ahsan) [1] ; Siddique, M (Siddique, Muhammad) [1] ; Dogan, E (Dogan, Eyup) [2] , [5] ; Sharif, A (Sharif, Arshian) [3] , [4]
(provided by Clarivate)

Volume
164
Page
956-967
DOI
10.1016/j.renene.2020.09.128
Published
FEB 2021
Indexed
2021-01-05
Document Type
Article

Abstract

A vast body of studies estimates the impact of energy consumption on the environment. A typical empirical study either use aggregate energy consumption or apply conventional econometric techniques in modelling the nexus of energy, income and environment. To correct these gaps, the objective of the study is to use renewable and non-renewable energy consumption in analyzing energy-incomeenvironment nexus, and to apply the novel Method of Moments Quantile Regression for ASEAN countries. The outcomes indicate that non-renewable energy consumption stimulate carbon emissions across all quantiles (10th to 90th), the value of the 10th quantile is 0.257 which rises to 0.501 till 90th quantile. Whereas, the renewable energy consumption leads to a decrease in CO2 emissions across all the quantiles (10th to 90th) but this association is statistically insignificant at higher quantiles from 60th to 90th. The empirical outcomes also verify the presence of the environmental Kuznets curve relationship, which is statistically significant from the middle (30th) to higher (90th) quantiles. Moreover, the finding of panel estimation approaches (FMOLS, DOLS, FE-OLS) also verify the existence of the EKC hypothesis in ASEAN countries. Their finding also describes that 1% increase in non-renewable energy consumption increase CO2 emission by 0.29%, 0.26% and 0.30% whereas 1% increase in the usage of renewable energy reduces CO2 emission by 0.17%, 0.15% and 0.17% in case of FMOLS, DOLS and FE-OLS respectively. The empirical results conclude that the government should encourage and subsidize the sources of green energy to tackle environmental degradation. More policy implications are further discussed in the study.

Regression

Keywords
Author Keywords
Renewable energy consumptionNon-renewable energy consumptionEKC hypothesisMethod of moments quantile regressionASEAN
Keywords Plus
CARBON-DIOXIDE EMISSIONSECONOMIC-GROWTH NEXUSFOREIGN DIRECT-INVESTMENTGREENHOUSE-GAS EMISSIONSCO2 EMISSIONSKUZNETS CURVEELECTRICITY CONSUMPTIONFINANCIAL DEVELOPMENTEMPIRICAL-EVIDENCEDECOMPOSITION ANALYSIS