Asian Journal of Probability and Statistics

10(2): 48-58, 2020; Article no.AJPAS.63324

ISSN: 2582-0230

Improving Research through Avoiding Common Statistical Errors: The Case of Piosphere

Eahsan Shahriary[1*], Thomas E. Gill[1,2], Richard P. Langford[2], Musa Hussein[2], William L. Hargrove[3] and Peter Golding[4]

[1]*Environmental Science and Engineering Program, University of Texas at El Paso, El Paso, Texas 79968, USA.*
[2]*Department of Geological Sciences, University of Texas at El Paso, El Paso, Texas 79968, USA.*
[3]*Center for Environmental Resource Management, University of Texas at El Paso, El Paso, Texas 79968 USA.*
[4]*College of Engineering, University of Texas at El Paso, El Paso, Texas 79968 USA.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2020/v10i230244
Editor(s):
(1) Dr. Manuel Alberto M. Ferreira, Lisbon University, Portugal.
Reviewers:
(1) Mohamed Hannabou, University Sultan Moulay Slimane, Morocco.
(2) A'Qilah Ahmad Dahalan, National Defence University of Malaysia, Malaysia.
Complete Peer review History: http://www.sdiarticle4.com/review-history/63324

Original Research Article

Received: 02 October 2020
Accepted: 07 December 2020
Published: 18 December 2020

Abstract

For many years scientists studied the piosphere concept- a grazing gradient around a natural/artificial watering point. As is the case for other kinds of ecological studies, the method of statistical analyses applied in many publications is not always appropriate. We note there are many statistical errors and misapplication of data analysis techniques. We reviewed 875 piosphere-related publications between 1915-2018 to find the common statistical methods and common statistical errors in the design of the study, data analyses, presentation of results, and interpretation of study findings. One-way ANOVA, multiple linear regression, Pearson correlation coefficient, permutational multivariate analysis of variance, canonical correspondence analysis, and mean were the most frequent statistical methods applied. Seventy-one common statistical errors in piosphere publications were found. The most common errors were not choosing the proper or appropriate statistical techniques, not checking the assumptions and diagnostics of statistical methods, partial and wrong interpretation of results, and not using informative figures and tables to help readers. Negligence to the proper application of statistics by

**Corresponding author: E-mail: eshahriary@miners.utep.edu;*

researchers results in inaccurate interpretation and spurious conclusions. It is recommended researchers seek advice from statisticians at the early stages of research to save resources, time, and labor and to provide increased trust in recommendations and findings.

Keywords: Common statistical errors; piosphere research; data analyses; assumptions; interpretation.

1 Introduction

Since the first publication about the role of water in sheep grazing [1] and later introduction of the term *piosphere* by Lange [2], there are many piosphere publications, including research articles, Ph.D. dissertation, book chapters, MS theses, reports, review articles, BS theses, plenary sessions, viewpoints, and certificates. In the recent work *Bibliometric analysis of piosphere research: implications for conservation and scientific collaboration*, the authors [3] encountered insufficient/wrong application of statistical methods, which prompted the work described here.

Gottfried Achenwall [4] coined the term "Statistik," which was introduced as "statistics" into the English language by John Sinclair [5]. Statistics uses mathematics and probability and has two subsets being descriptive statistics and inferential statistics [6]. Statistics concerns collecting data from a study population and analyzing sample data, interpreting, visualizing, and generalizing the results to a target population [7].

Statistics play an important role in essentially all forms of scientific and technical research. Reviews of statistical applications have been made, for example, in agricultural research [8], medicine [9], engineering [10], environmental science [11], wildlife research [12], forensic studies [13], economics and business [14], anthropology [15], physics [16], social science [17], applied statistics and mathematics [18], ecology [19], biostatistics [20] and biology [21].

There is a consensus about the misuse of statistical methods and common statistical errors [22]. Misuse of statistical methods wastes resources, time, and labor and is the cause of incorrect conclusions and recommendations. Inappropriate use of statistical methods in piosphere research has not been discussed and challenged, contrary to other scientific disciplines. In these different disciplines, researchers have tried to shed lights on some crucial aspects of the applications of statistical methods, for example, wildlife management [23], crop science [24], rangeland science [25], ecology [26], medicine [27], biology [28] and engineering [29]. There is no specific protocol regarding statistical analysis of piosphere data (which can be used by other disciplines). Except for one article about applying a logistic growth curve equation for piosphere data analysis [30] and another article about using regression models to predict spatial patterns of cattle behavior [31], we have not find publications about the appropriate use of statistics in piosphere research. After reviewing piosphere publications, this paper provides a brief review of the common statistical errors in reporting piosphere research, including errors in the design of the study, statistical analyses, presentation of results, and interpretation of study findings.

2 Methodology

Using Google Scholar and keywords: piosphere, water point, watering point, borehole, and waterhole, the database of 875 publications between 1915 -2018 was prepared [3]. We reviewed 875 publications (672 research articles, 74 Ph.D. dissertation, 40 book chapters, 38 MS theses, 32 reports, 14 review articles, two BS theses, one plenary session, one viewpoint, and one certificate). We extracted the information on common statistical methods and common statistical errors in the design of the study, statistical analyses, presentation of results, and interpretation of study findings. We analyzed and summarized the extracted data. This process is depicted visually in Fig. 1.

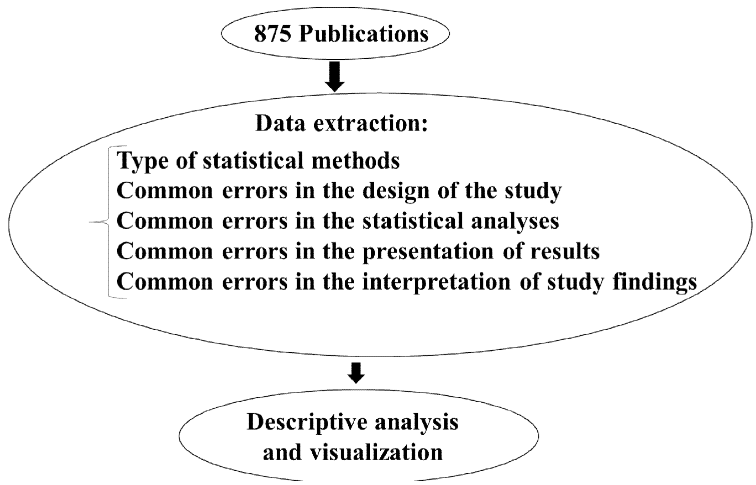

Fig. 1. Step-by-step diagram for data extraction of statistical methods and errors in piosphere research between 1915-2018.

There are many ways to classify statistical methods according to the objective of data analysis. For example, tests to look at differences, tests to look at relationships, and tests for data explorations [32]. Another case is univariate, bivariate, and multivariate statistical methods [33]. We modified the classification proposed by Šmilauer and Lepš [34] based on the number of response(s) and predictor(s) variables (Fig. 2) and classified the statistical methods of the 875 reviewed publications. There are six groups comprising group I (no predictor, one response variable), group II (one predictor, one response variable), group III (> 1 predictor, one response variable), group IV (No predictor, > 1 response variable), group V (one predictor, > 1 response variable), and group VI (> 1 predictor, > 1 response variable).

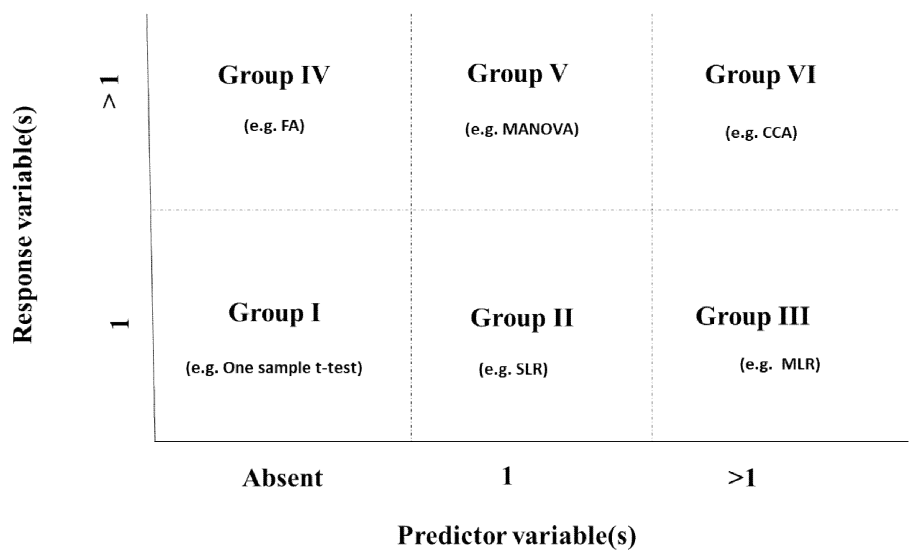

Fig. 2. Classification of statistical methods (modified after Šmilauer & Lepš [34]). Readers may refer to the list of acronyms for assistance as needed (Table 6)

3 Results

3.1 Number of statistical methods

The total number of statistical methods in 875 piosphere publications for the period of 1915 to 2018 is presented in Table 1. We have classified the statistical methods during 1915 to 2018 to six groups: group I (675, 77.14%), group II (299, 34.17 %), group III (532, 60.80%), group IV (418, 47.77%), group V (23, 2.62%) and group VI (50, 5.71 %). The most frequent statistical methods were mean (317, 36.22%), SD (104, 11.88%), One-way (ANOVA) (213, 24.34%), Kruskal-Wallis test (37, 4.22%), MLR (200, 22.85%), multiple logistic regression (55, 6.28%), PCC (72, 8.22%), PCA (55, 6.28%), PERM (13, 1.48%), MANOVA (5, 0.57%), CCA (28, 3.2 %) and RDA (8, 0.91%) (Table 1).

Table 1. The total number and most frequent statistical methods used in 875 piosphere publications between 1915 -2018. Readers may refer to the list of acronyms for assistance as needed (Table 6).

Groups	Total	Most frequent methods
I (No predictor, one response variable)	(675,77.14%)	mean (317, 36.22%), SD (104, 11.88%), χ^2 (75, 8.75%), CV (49, 5.6%), one sample t-Test (39, 4.45%)
II (one predictor, one response variable)	(299,34.17%)	one-way (ANOVA) (213, 24.34%), Kruskal-Wallis test (37, 4.22%), SLR (21, 2.4%), OLS (16, 1.82%)
III (> 1 predictor, one response variable)	(532,60.80%)	MLR (200, 22.85%), multiple logistic regression (55, 6.28%), GLM (49, 5.6%), repeated measures ANOVA (32, 3.65%), GLMs (25, 2.85%), GLMM (20, 2.28%)
IV (No predictor, > 1 response variable)	(418,47.77%)	PCC (72, 8.22%), PCA (55, 6.28%), DCA (30, 3.42%), cluster analysis (30, 30.42%)
V (one predictor, > 1 response variable)	(23, 2.62%)	PERM (13, 1.48%), MANOVA (5, 0.57%), SAR (3, 0.34%), MRBP (2, 0.22%)
VI (> 1 predictor, > 1 response variable)	(50, 5.71%)	CCA (28, 3.2%), RDA (8, 0.91%), RLQ analysis (5, 0.57%), pCCA (2, 0.22%), HMMs (2, 0.22%), SEM (2, 0.22%), DCCA (1, 0.11%), CAP (1, 0.11%)

3.2 Common errors in the design of the study

The design of the study is the most vital part leading to proper and reliable research [35]. Errors in this stage nullify posterior efforts and phases. First and foremost, we need to design a protocol with the research objectives and the expected outcomes. Researchers should consider the different types of study; hypothesis-driven and data-driven. If the study is hypothesis-driven, scientific and statistical hypotheses should be defined and formulated [35]. After proper exploratory analysis, data-driven statistical and scientific hypotheses could be defined and developed [36]. One of the most important things at this stage is setting a priori, sample size calculation, estimating the power of the test, effects of treatment (fixed and random), and minimizing type I error or false positive (rejecting the true null hypothesis) and type II error or false-negative (failing to reject the false null hypothesis) [37]. Researchers should consider replication, randomization, and control (if needed) to stop biased studies and provide accurate and valid results [38]. Sampling (with replacement and without replacement) and randomization are essential in inferential statistics [39]. If we use randomization and control treatment, we can specify the effects of treatments [40]. Common errors in the design of the study are mentioned in Table 2.

3.3 Common errors in the statistical analyses

Scherrer [20] discussed the importance of the type of data (nominal, ordinal, interval, and ratio) and the study's objective in selecting the statistical data analysis method. Unfortunately, many piosphere research publications have not followed the assumptions for ordinary methods (e.g., simple linear regression) and advanced methods (e.g., canonical correspondence analysis). For example, when data are discrete, we cannot use a t-test and mean comparison; we should use median comparison and a nonparametric test [41]. Researchers should notice that each statistical method has assumption(s) which should be met to obtain reliable and robust results. For example, Sokal & Rohlf [42] stated four assumptions of analysis of variance (ANOVA), including independence, homogeneity of variances, normality, and additivity. Another vital consideration regards the choice of model fitting is that good linear models should be checked using a plot of residuals [43] (Fig. 3). Dowdy and her colleagues [44] explained the relationship among the type of multiple comparison procedure, power, and type I error, which should be considered before data analysis to control Type II and Type I errors. Common errors of statistical analyses are presented in Table 3.

Table 2. Common errors in the design of the study

No	Common statistical errors
1	Not determining the objective of the study and the expected outcome.
2	Not reporting the sample size.
3	Not estimating power (effect of sample size) based on the pilot study.
4	Not stating a clear null hypothesis.
5	Not reporting the number of replications.
6	Not reporting the selection of pseudoreplication.
7	Not reporting the use of randomization.
8	Not controlling the confounding variable(s).
9	Not reporting the use of control treatment.
10	Not collecting the proper data to achieve the objectives.
11	Not choosing a priori.
12	Not using the representative sample.
13	Not mentioning the variables, unit of measurement, and if they are fixed or random.
14	Biased sampling.

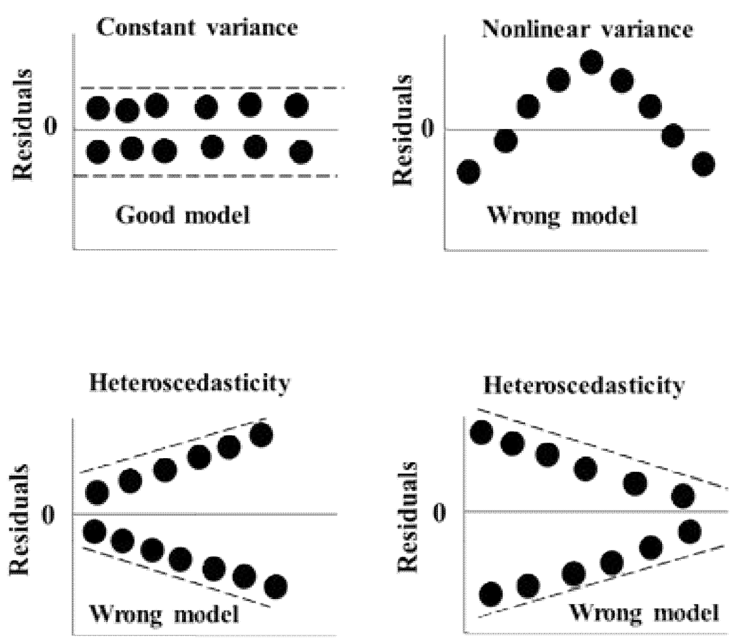

Fig. 3. Residual plots of good and wrong models

Table 3. Common errors in the statistical analysis

No	Common statistical errors
1	Not choosing the proper statistical methods according to the type of data (e.g., Not using GLMs to analyze count data).
2	Not choosing the proper statistical methods according to research hypothesis and objective of study.
3	Increasing the Type II error by choosing the wrong multiple comparisons.
4	Not using Bonferroni correction when performing simultaneous tests.
5	Choosing the wrong multiple comparison methods.
6	Not checking the targeted statistical method's assumption(s) (e.g., normality, collinearity, interaction).
7	Not undertaking remedial actions when assumptions are violated.
8	Not using Fisher Exact test when Chi-square is not a valid test.
9	Not using new methods such as the generalized linear mixed model instead of data transformation.
10	Not checking the diagnostics after data analysis to assure the selected models (e.g., residuals).
11	Not performing the nonparametric test when assumption (s) is/are violated.
12	Not using data transformation (e.g., centering, standardizing).
13	Not using autocorrelation when data collected over time.
14	Not checking the compatibility of data with the statistical method.
15	Paying no attention to the independence and dependence of variables of interest in data analysis.
16	Not checking the influential points (outliers, inliers, and leverages) in regression analysis.
17	Not reporting the software(s), packages, and procedures.
18	Not running exploratory data analysis.
19	Not determining the ad-hoc vs. post-hoc comparisons.
20	Not explaining the selection of matrix (correlation vs. covariance) to calculate principal components in PCA.
21	Not performing DCA (to find the gradient length) before choosing and running CCA/RDA and pCCA/pRDA.
22	Focusing on the main effects when the interaction is significant.
23	Not using permutation tests to perform a significance test.
24	Not reporting the selection of prior probabilities when using Bayesian statistical inference.
25	Try to test multiple hypotheses with the same dataset.
26	Not reporting how researchers treated the missing data.
27	Not using resampling methods to validate the models.

3.4 Common errors in the presentation of results

In line with the adage, "a picture is worth a thousand words," a good graphical presentation saves text and makes the length of a publication shorter. Proper presentation of statistical results through illustration is as important as other stages of research [45]. Proper use of statistical terms is crucial. One such example is the misuse of error and residual terms. Error is (true value – observed value) (Fig. 4a) but residual is (observed value – predicted value) [6] (Fig. 4b). Well prepared tables and figures are the best ways to provide information on results to readers. Murray [46] mentioned that about 50 % of the presentation of statistical analysis results has errors. Table 4 shows the common errors in the presentation of the statistical analysis results.

3.5 Common errors in the interpretation of study findings

Researchers should keep in mind that they should not make conclusions that cannot be supported by data or pick only part of the data to conveniently support their ideas. For example, it is vital to understand the difference between statistical significance and biological significance [47] and finding statistical non-significance does not mean biological non-significance on every occasion. It can be due to choosing the wrong statistical test or small sample size (inflation of Type II error) or not checking the assumptions, diagnostics, and remedies. Researchers should be careful in interpreting their findings. Table 5 shows the common statistical errors of the interpretation of the study findings.

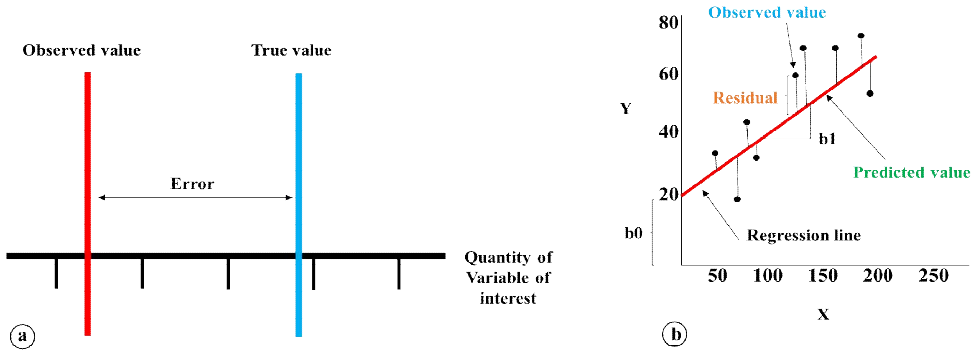

Fig. 4. Graphical presentation of error (a) and residual terms (b)

Table 4. Common errors in the presentation of results

No	Common statistical errors
1	Choosing the wrong graphs that are not related to the type of data and objective of the presentation.
2	Not using informative figures and tables to help readers.
3	Mixing to concepts of errors (true value – observed value) and residuals (observed value – predicted value).
4	Not Reporting the confidence interval in graphs and tables.
5	Not reporting the true exact P-value instead of (P<0.05), (P<0.01), and (P>0.05) with confidence interval and sample size.
6	Using standard error instead of standard deviation.
7	Not reporting estimates with the rate of error.
8	Use of control and treatment comparison instead of across treatment studies.
9	Making irrelevant conclusions from poor experiments or small datasets.
10	Not reporting confidence interval and prediction interval of regression models.
11	Using a confidence interval instead of standard error.
12	Not mentioning if the used test is one-tailed or two-tailed.
13	Anscombe's quartet (error of summary statistics).
14	Not using mean ± standard deviation to describe normal data.
15	Not using median (positional average) for non-normal, discrete, or skewed data.
16	Wrong interpretation of confidence interval.
17	Using range when there is/are outlier(s).
18	Not reporting the α level.
19	Not reporting prior and posterior probabilities.
20	Not reporting coefficient of determination, correlation coefficient, AIC, and BIC.

Table 5. Common errors in the interpretation of study findings

No	Common statistical errors
1	Reporting statistical non-significant results as no biological difference.
2	Making generalization which is beyond the collected data (inductive method).
3	Partial/poor/wrong interpretation of results.
4	Reporting statistical non-significance without checking the Type II error.
5	Reporting statistical non-significance without checking the confounding.
6	Not considering statistical non-significance as inconclusive.
7	Considering statistical significance as biological significance.
8	Concluding causation with correlation.
9	Interpreting prediction by regression beyond data range.
10	Reporting conclusion without having enough data or doing required statistical analysis.

4 Discussion

This work's objective was to determine and discuss common statistical errors in piosphere publications to help researchers improve analysis and correctly conduct and report meaningful research. Investigators who fully consider the errors in statistics discussed in this article will enhance the quality of piosphere research and, by association, most research in ecology and much other science and engineering fields. Statistical errors cause incorrect interpretations and tentative conclusions.

We recommend that researchers discuss their research project with statisticians throughout their discourse, from the first step (design of the study) to the last step (writing up) of their projects. To improve research quality, scientists could apply advanced statistical methods such as modern experimental design, Bayesian models, multivariate analysis, GLMS, GAM, GLMM, meta-analysis, and likelihood models. Increasing knowledge of the available and currently-used statistical software and packages is fundamental in the proper use of statistics. For example, SAS uses Type III sum of squares (SS) (orthogonal) [48]. Still, function *aov* in package 'stats' in R by default uses Type I SS (sequential) [49], and if in R the calculation based on Type III SS (orthogonal) is needed package 'car' and function *Anova* should be used [50]. Researchers should know and study the help document/manual of each software package and procedure to prevent common errors.

Table 6. List of acronyms and their meanings.

Acronym	Full word
AIC	Akaike information criterion
ANOVA	Analysis of variance
BIC	Bayesian information criterion
CAP	Canonical analysis of principal coordinates
CCA	Canonical correspondence analysis
CV	Coefficient of variation
DCA	Detrended correspondence analysis
DCCA	Detrended canonical correspondence analysis
FA	Factor analysis
GAM	Generalized additive models
GLM	General linear model
GLMM	Generalized linear mixed models
GLMs	Generalized linear model
HMMs	Hidden Markov models
MANOVA	Multivariate analysis of variance
MLR	Multiple linear regression
MRBP	Multi-response permutation procedure
OLS	Ordinary least squares
PCA	Principal component analysis
PCC	Pearson correlation coefficient
pCCA	Partial canonical correspondence analysis
PERM	Permutational multivariate analysis of variance
RDA	Redundancy analysis
RLQ	R-mode linked to Q-mode
SAR	Spatial simultaneous autoregressive
SD	Standard deviation
SEM	Structural equation modeling
SLR	Simple linear regression
$\chi 2$	Chi-squared test

5 Conclusions

Researchers (before submission of the final version of any publication or report) and reviewers should review documents multiple times, four times at least. The first will be from the biological/disciplinary point of view; the second is from a statistical analysis point of view, the third from the standpoint of language and writing quality, and the fourth, regarding the target journal's format and style. As Marcel Proust [51] says, "error, by force of contrast, enhances the triumph of truth." The proper use of statistics improves research findings and recommendations, provides improvements in conservation and management strategies for piosphere and other ecological investigations, and ensures stable and sound consideration of future research directions.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Jardine JT. Grazing sheep on range without water. The National Wool Grower. 1915;5:7-10.

[2] Lange RT. The piosphere: sheep track and dung patterns. Journal of Range Management. 1969;22:396-400.

[3] Shahriary E, Gill TE, Langford R. Bibliometric analysis of piosphere research: implications for conservation and scientific collaboration, International Journal of Sustainable Development & World Ecology; 2020.
DOI: 10.1080/13504509.2020.1821256.2020.

[4] Droesbeke JJ, Tassi P. Histoire de la statistique. Presses universitaires de France. Paris. 1990;12.

[5] Peters WS. Counting for something: statistical principles and personalities. Springer, New York, USA. 2011;297.

[6] Dodge Y. Statistique: dictionnaire encyclopédique. Springer Science & Business Media. 2007;653.

[7] Binder DA, Roberts G. Design-and model-based inference for model parameters. In Handbook of Statistics. Elsevier, Amsterdam, The Netherlands. 2009;29B(33-54):666.

[8] Mead R. Statistical methods in agriculture and experimental biology. Chapman and Hall/CRC, Boca Raton, FL, USA. 2017;488.

[9] Hoffman JI. Biostatistics for medical and biomedical practitioners. Academic Press, London, UK. 2015;717.

[10] Montgomery DC, Runger GC. Applied statistics and probability for engineers. Wiley, Hoboken, NJ, USA. 2013;832.

[11] Al-Karkhi AF, Alqaraghuli WA. Applied Statistics for Environmental Science with R. Elsevier, Amsterdam, Netherlands. 2019;221.

[12] Hooten MB, Johnson DS, McClintock BT, Morales JM. Animal movement: statistical models for telemetry data. CRC Press, Boca Raton, USA. 2017;320.

[13] Aitken C, Taroni F. Statistics and the evaluation of evidence for forensic scientists. John Wiley & Sons, Chichester, England. 2020;248.

[14] McClave JT, Benson PG, Sincich T. Statistics for business and economics. Pearson, Boston, MA, USA. 2017;888.

[15] Obertova Z, Stewart A, Cattaneo C. Statistics and probability in forensic anthropology. Academic Press, London, UK. 2020;696.

[16] Roe BP. Probability and statistics in experimental physics. Springer Science & Business Media, Berlin, Germany. 2012;264.

[17] Warne RT. Statistics for the social sciences: A general linear model approach. Cambridge University Press, UK. 2017;601.

[18] Panaretos VM. Statistics for mathematicians. Compact Textbook in Mathematics. Birkhäuser, Basel, Switzerland. 2016;196.

[19] Hobbs NT, Hooten MB. Bayesian models: a statistical primer for ecologists. Princeton University Press, Princeton, USA. 2015;320.

[20] Scherrer B, Biostatistique V1, Gaëtan Morin, Boucherville, Québec, Canada. 2008;816.

[21] Van Emden HF. Statistics for terrified biologists. Wiley-Blackwell, Hoboken, NJ, USA. 22019;360.

[22] Good PI, Hardin JW. Common errors in statistics (and how to avoid them). Wiley, Hoboken, NJ, USA. 2012;352.

[23] Steidl RJ, Hayes JP, Schauber E. Statistical power analysis in wildlife research. The Journal of Wildlife Management. 1997;61:270-279.

[24] Scott SJ, Jones RA, Williams W. Review of data analysis methods for seed germination. Crop Science. 1984;24:1192-1199.

[25] Batabyal AA, Godfrey EB. Rangeland management under uncertainty: a conceptual approach. Journal of Range Management. 2002;55:12–15.

[26] Zuur AF, Ieno EN. A protocol for conducting and presenting results of regression-type analyses. Methods in Ecology and Evolution. 2016;7:636-645.

[27] Moodie EEM, Krakow EF. Precision medicine: Statistical methods for estimating adaptive treatment strategies. Bone Marrow Transplant. 2020;55:1890–1896.

[28] Cumming G, Fidler F, Vaux DL. Error bars in experimental biology. The Journal of Cell Biology. 2007;177:7-11.

[29] Kang DG. Comparison of statistical methods and deterministic sensitivity studies for investigation on the influence of uncertainty parameters: Application to LBLOCA. Reliability Engineering & System Safety. 2020;203:107082.

[30] Graetz RD, Ludwig JA. A method for the analysis of piosphere data applicable to range assessment. The Rangeland Journal. 1976;2:126-136.

[31] Senft RL, Rittenhouse LR, Woodmansee RG. The use of regression models to predict spatial patterns of cattle behavior. Journal of Range Management. 1983;36:553-557.

[32] Dytham C. Choosing and using statistics: a biologist's guide. Wiley-Blackwell, Oxford, UK. 2011;320.

[33] Denis JD. Univariate, Bivariate, and Multivariate Statistics Using R: Quantitative Tools for Data Analysis and Data Science. Wiley, Hoboken, NJ, USA. 2020;366.

[34] Šmilauer P, Lepš J. Multivariate analysis of ecological data using CANOCO 5. Cambridge University Press, UK. 2014;374.

[35] Creswell JW, Creswell JD. Research design: qualitative, quantitative, and mixed methods approaches. SAGE publications, Thousand Oaks, CA, USA. 2018;304.

[36] Kutz JN. Data-driven modeling & scientific computation: methods for complex systems & big data. Oxford University Press, UK. 2013;656.

[37] Ruxton G, Colegrave N. Experimental design for the life sciences. Oxford University Press, UK. 2017;208.

[38] Casella G. Statistical design. Springer Science & Business Media, New York, USA. 2008;307.

[39] Sahu PK, Pal SR, Das AK. Estimation and inferential statistics. Springer, New Delhi, India. 2015;317.

[40] Bate ST, Clark RA. The design and statistical analysis of animal experiments. Cambridge University Press, UK. 2014;326.

[41] Sprent P, Smeeton NC. Applied nonparametric statistical methods. Chapman and Hall/CRC. Boca Raton, FL, USA. 2016;542.

[42] Sokal RR, Rohlf FJ. Biometry. 3rd ed. W.H. Freeman and Company, New York. 1995;880.

[43] Kutner MH, Nachtsheim CJ, Neter J, Li W. Applied linear statistical models. McGraw-Hill Irwin, New York, USA. 2013;1396.

[44] Dowdy S, Wearden S, Chilko D. Statistics for research. Wiley-Interscience; 3rd edition. 2011;640.

[45] Gustavii B. How to write and illustrate a scientific paper. Cambridge University Press, UK. 2017;210.

[46] Murray GD. Statistical aspects of research methodology. British journal of surgery. 1991;78:777-781.

[47] Krebs CJ. Ecological methodology. Menlo Park, CA, USA: Benjamin/Cummings. 1999;624.

[48] O'Rourke N, Hatcher L, Stepanski EJ. A step-by-step approach to using SAS for univariate & multivariate statistics. SAS Institute, Cary, NC, USA. 2005;548.

[49] Kabacoff R. R in Action: Data analysis and graphics with R. Manning Publications, Shelter Island, NY, USA. 2015;608.

[50] Fox J, Weisberg S. An R companion to applied regression. SAGE Publications, Thousand Oaks, CA, USA. 2019;608.

[51] Proust M. À la recherche du temps perdu. Premier Volume: Du côté de chez Swann. (Initialement publié:1913). Gallimard, Paris, France. 2002;676.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/63324