



ORIGINAL RESEARCH ARTICLE / ÖZGÜN ARAŞTIRMA MAKALESİ



Copyright@Author(s) - Available online at dergipark.org.tr/en/pub/igusbd.
Content of this journal is Licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

Investigating Violations of Parametric Test Assumptions in Tourism Research

Turizm Araştırmalarında Parametrik Test Varsayımlarının İhlallerinin İncelenmesi

¹Duygu EREN

²Eda ÖZGÜL KATLAV

³Şule AYDIN

¹Prof., Nevşehir Hacı Bektaş Veli
University, Faculty of Tourism,
Department of Tourism Management,
Nevşehir, Türkiye.
✉ deren@nevsehir.edu.tr

²Assoc. Prof., Nevşehir Hacı Bektaş
Veli University, Faculty of Tourism,
Department of Tourism Management,
Nevşehir, Türkiye.
✉ edaozgul@nevsehir.edu.tr

²Prof., Nevşehir Hacı Bektaş Veli
University, Faculty of Tourism,
Department of Tourism Management,
Nevşehir, Türkiye
✉ suleaydin@nevsehir.edu.tr

Geliş/Received: 11.07.2025

Kabul/Accepted: 15.12.2025

Abstract

Aim: This study evaluates the extent to which researchers adhere to parametric testing assumptions in quantitative studies published in SSCI-indexed tourism journals.

Method: Using document analysis, 239 empirical articles published between 2020 and 2024 in 19 journals were systematically reviewed with a focus on the statistical methods employed and on whether basic parametric assumptions were acknowledged, tested, and reported.

Results: The findings show a predominant use of SEM/CB-SEM and PLS-SEM. Data-type assumptions were mentioned in 94% of cases, whereas normality and sample-size assumptions were addressed far less frequently: only 43% of tests reported checking normality and 29% reported assessing sample size. Normality was not mentioned in 51% of tests and explicitly stated as “not needed” in 6%, while sample-size assumptions were not mentioned in 58% and claimed to be unnecessary in 13% of cases. When all parametric assumptions were considered together, only 24% of studies reported them comprehensively, whereas 57% did not mention them and 18% explicitly stated that such assumptions were not needed; all of these “not needed” cases employed PLS-SEM.

Conclusion: These results highlight a need for greater methodological rigor in tourism research and underscore the importance of systematically considering parametric assumptions to strengthen the validity and reliability of quantitative findings in the field.

Keywords

Parametric test assumptions, violation, normal distribution, sample size, measure type, tourism research

Jel Code: C18, C12, Z30

Öz

Amaç: Bu çalışma, SSCI kapsamında taranan turizm dergilerinde yayımlanan nicel araştırmalarda araştırmacıların parametrik test varsayımlarına ne ölçüde uyduğunu değerlendirmeyi amaçlamaktadır.

Yöntem: Belge analizi yöntemiyle, 2020–2024 yılları arasında 19 dergide yayımlanan 239 ampirik makale; kullanılan istatistiksel yöntemler ve temel parametrik varsayımların kabul edilme, test edilme ve raporlanma düzeyleri açısından sistematik olarak incelenmiştir.

Bulgular: Çalışmalarda ağırlıklı olarak SEM/CB-SEM ve PLS-SEM yöntemlerinin tercih edildiği saptanmıştır. Veri türü varsayımları %94 oranında belirtilmesine karşın; normallik (%43) ve örneklem büyüklüğü (%29) varsayımları düşük düzeyde kontrol edilip raporlanmıştır. Normallik varsayımı %51 oranında hiç anılmamış, %6'sında "gerekli değil" olarak ifade edilmiştir. Örneklem büyüklüğü varsayımına %58'inde yer verilmemiş, %13'ünde ise gerekli olmadığı belirtilmiştir. Tüm parametrik varsayımlar birlikte değerlendirildiğinde; çalışmaların yalnızca %24'ünde varsayımlar kapsamlı biçimde raporlanmış, %57'sinde hiç anılmamış ve %18'inde gerekli olmadığı yönünde açık ifade kullanılmıştır. "Gerekli değil" ifadesi yer alan çalışmaların tamamında PLS-SEM yöntemi kullanılmıştır.

Sonuç: Bu sonuçlar, turizm araştırmalarında yöntemsel titizliğin artırılması gereğine işaret etmekte ve parametrik varsayımların sistematik biçimde dikkate alınmasının, nicel bulguların geçerlilik ve güvenilirliğini güçlendirmedeki önemini vurgulamaktadır.

Anahtar Kelimeler

Parametrik test varsayımları, ihlal, normal dağılım, örneklem büyüklüğü, ölçek türü, turizm araştırması

Jel Sınıflandırması: C18, C12, Z30

Introduction

All research typically involves a literature review, research hypotheses or questions, research design and methodology, data collection and analysis, and interpretation and writing. According to Abrami, Cholsky, and Gordon (2000), data analysis is only one but an important part of this process, and we should consider the following research equation to guide our own work as well as critique the research of others.

Theory+ good design+ proper analysis= an excellent investigation

Similarly, Gürbüz and Şahin (2018) stated that research should have at least three important characteristics: 1) conducting it with a concrete purpose to learn something new; 2) systematically collecting data; and 3) analyzing and interpreting the meaning of the collected data with appropriate techniques. To achieve the purpose and answer the research questions, researchers develop hypotheses, which they then test these hypothesis using statistical analysis. However, in order to obtain reliable and accurate results, researchers should answer the following questions about data analysis before analyzing the data (Choi, 2005; Alpar, 2014; Tabachnick and Fidell, 2014, p. 94; Forstmeier, Wagenmakers, and Parker, 2017, p. 1950-1965).

1. Is the sample size of the study adequate?
 - a. What is the research question?
 - b. Should the sample size be determined by the primary outcome of interest?
2. Are the statistics appropriate?
 - a. Is the measured variable expected to have a normal distribution?
 - b. What were the measures of centrality and spread?
 - c. What inferential tests have been used?
3. How were common problems dealt with when they arose?
 - a. Was the P-value adjusted for multiple comparisons?
 - b. Were case losses, case withdrawals and outliers dealt with appropriately?
 - c. If there were no differences between groups, was a power analysis performed, a definition of the smallest detectable difference based on the sample size of the study, or a calculation of the sample size required based on the observed frequencies?
 - d. If no events were observed, was there a description of the maximum possible frequency based on the sample size of the study?

Answering these questions requires a good understanding of some basic statistical principles and assumptions. This is because most statistical techniques require one or more assumptions to be met, or if a technique is found to be robust to violations of an assumption, the assumption should not be violated to an extreme degree (Hoekstra, Kiers and Johnson, 2012). When assumptions are violated, using statistical techniques is a serious problem in data analysis (Olsen, 2003; Choi, 2005). Violation of assumptions can seriously affect Type I and Type II errors and can lead to overestimation or underestimation of inferential measures and effect sizes (Osborne and Waters, 2002). Therefore, researchers should examine the assumptions of statistical analyses in general before deciding which statistical analysis technique to use.

In general, basic assumptions should be made before deciding on the appropriate analysis technique for research. The first of these assumptions concerns the nature of the data. Clearly, the application of statistical analyses depends on whether the data are metric or non-metric. Researchers use different statistical tests for non-metric and metric data to answer the same research questions. More specifically, researchers use non-parametric tests for non-metric data and parametric tests for metric data. It is therefore essential for researchers to understand the type of data generated in their studies.

The second basic assumption is that the distribution is normal. Parametric tests undoubtedly provide more accurate results than non-parametric tests, but they rely on a general assumption of normality as well as some specific assumptions associated with each test. Serious violations of the assumption of normality can bias the results of parametric tests (Verma and Abdel-Salam, 2019). Errors in analysis can lead to misinterpretations of data and erroneous conclusions (Olsen, 2003). The sample size and sampling method used are among the basic assumptions for statistical analysis. However, each analysis requires specific assumptions for its application. For example, equality of variances for a t-test, linearity for regression, randomness of error terms, etc.

Many statistical software programs are now available to analyse both non-metric and metric data. These software programs provide researchers with excellent convenience in their analysis. These programs are easy to use and learn. The primary issue arises when researchers solely concentrate on mastering these programs, neglecting to learn about statistics or data analysis. They do not automatically check the assumptions of the analysis. However, they do provide outputs for testing assumptions about statistical tests. Based on these outputs, researchers can decide on the appropriate analysis technique to achieve accurate and valid results. However, researchers need to know how to use and interpret these outputs.

Hoekstra et al. (2012) conducted a study of 30 researchers to specifically examine researchers' understanding of the assumptions of statistical tests and their behaviour in checking these assumptions. This study assessed researchers' understanding of the basic assumptions of statistical tests and their checking behaviors through articles published in journals. This study examined journal articles published for this purpose, taking into account the fundamental assumptions of statistical analysis such as sample size, data type, and normal distribution.

Basic Assumptions of Parametric Analyses

Parametric analyses were the first widely used analyses to come to the fore in the development of modern statistical inference techniques. Parametric tests rely on specific assumptions about the underlying population, based on the premise that population parameters are essential for interpreting research results (Doğan and Doğan, 2014). The assumptions for parametric analyses can be listed as follows (Kurtuluş, 2006; Alpar, 2014; Doğan and Doğan, 2014).

- The units included in the sample should be selected randomly.
- Sample units should be selected independently.
- The theoretical probability distribution of the population from which the sample units are selected should be known.
- Data must be measured on at least an interval scale.

If the above assumptions are not met, non-parametric analyses should be used. Parametric analyses are statistically more efficient and powerful than non-parametric analyses in all cases. However, Kurtuluş (2006) states that this judgment is only valid if it does not violate any of the assumptions made for the use of parametric analysis. These assumptions are explained below.

Assumptions About Type of Data

We classify data into two categories: non-metric and metric. Non-metric data is also called qualitative data, while metric data is also called quantitative data. We further classify non-metric data into nominal and ordinal categories.

Non-metric data is a categorical measurement that is commonly referred to as 'categorical' data. On the other hand, metric data, always associated with a scale measure, is also referred to as scale data. We can measure metric data using two different types of scales: interval and ratio (Verma and Abdel-Salam, 2019).

These data are known to everyone and are used in research. Nominal data is data where observations are simply categorised or counted. For example: gender (male/female), eye colour (black, brown, blue, green), place of birth, religious belief, political opinion, etc. Ordinal data is where categories are ranked. If data can be sorted into an order based on its different characteristics, it is ordinal data. For example, level of education (primary, secondary, post-secondary), income (low, middle, high), quality (poor to excellent), attitude (strongly disagree to strongly agree) (Verma and Abdel-Salam, 2019; Gürbüz and Şahin, 2018; Alpar, 2014; Mason, et al., 2000). Interval data are obtained by measuring a phenomenon on a scale where each location is equidistant from the others. On the scale used to measure these data, the distance between the measurements is meaningful, standardised and equivalent. The only problem with this scale is that there is no real zero on the scale, so it is not possible to interpret it in terms of a real zero point. For example, thermometers or intelligence tests (1-10), etc.

At this point, we should highlight the confusion in the literature. While some sources (Verma and Abdel-Salam, 2019; Gürbüz and Şahin, 2018; Alpar, 2014; Mason et al., 2000) refer to the data obtained with Likert-type scales as ordinal data, others (Burns and Veeck, 2020; Nakip, 2003; Abrami, Cholmsky and Gordon, 2000) refer to them as interval data. Given the frequent use of Likert scales in the social sciences, what is the level of acceptance for these data? If the data are ordinal, the studies conducted so far have violated at least one basic assumption. If the data are interval, there seems to be no problem with one of the basic assumptions. However, different scales used in the same study may have different ratings (e.g., 5-point, 7-point, or 9-point, etc.), and researchers treat data from scales with different ratings as having the same rating and use them in their analyses. This leads to a violation of the analysis, even if there is no violation of a basic assumption. Finally, ratio data measured on the ratio scale has a real and meaningful zero point and is equidistant. For example, the difference between 30 and 40 is the same as the difference between 60 and 70. All kinds of mathematical operations can be performed on this type of data. Examples of ratio data are weight, height, distance, salary, etc. (Verma and Abdel-Salam, 2019; Gürbüz and Şahin, 2018; Alpar, 2014; Nakip, 2003; Abrami, Cholmsky and Gordon, 2000; Mason et al., 2000).

It is commonly known that parametric statistics are calculated for metric data and non-parametric statistics are used for non-metric data. Violation of these assumptions can lead to misleading results. A simple example of this is the calculation of central training measures. If the data are nominal, we calculate the mode as an appropriate measure of central tendency, and if the data are ordinal, we calculate the median. Since both nominal and ordinal data are not metric, non-parametric statistics (mode and median) are used. On the other hand, if the data are metric (interval or ratio), parametric statistics such as mean and standard deviation should be used. However, the normality assumption must hold for the calculation of parametric statistics for metric data. If normality is violated, non-parametric statistics such as median and quartile deviation should be used (Verma and Abdel-Salam, 2019).

Assumptions About Sample Size

The number of participants needed for a sample is one of the most important decisions in a research process. In general, the decision on sample size is a compromise between theoretical perfection and practical feasibility. The first assumption about sample size is actually a complete count. This is due to the absence of sampling error in a complete census, as the study includes all individuals in the population (Tabachnick and Fidell, 2014, p. 43; Burns and Veeck, 2020). However, in the case of large groups, a complete count is expensive and time-consuming. Instead, it is preferable to collect information that reflects the group's general characteristics by interviewing only a few of the large group members. The data can be collected cheaper, easier, and in less time. This is the idea behind sampling.

Any research's results rely on selecting an appropriate number of samples for the study. Researchers are always in a dilemma about how much sample is sufficient for the study. A small sample may lead to false results, while a large sample may lead to the unnecessary use of additional resources such as time and money. A small sample may not be able to detect an effect in the study, whereas in a large sample, even a small effect may be significant (Verma and Abdel-Salam, 2019; Mason et al., 2000). Goodhue, Lewis and Thompson (2012, p.990-1000) emphasise that basic statistics tell us that the smaller the sample, the less confident we can be about our estimates. Therefore, it is necessary to keep the sample size at the optimum point in order to conduct the research in a sound manner (Nakip, 2003).

Mason et al., (2000) state that determining the correct sample size depends on three factors. These factors are: the desired level of confidence; the margin of error that the researcher is willing to tolerate; and the variability of the population under study. Verma and Abdel-Salam (2019) state that in survey studies, the estimation of population

characteristics is mainly concerned with precision and confidence coefficients. Therefore, violations of assumptions in such studies affect these two criteria. For example, in the case of extreme violations of the normality assumption, the precision in the estimation of population parameters may not be as claimed, or the confidence coefficient in interval estimates may not be consistent with the results. Similarly, if we use parametric statistics to estimate population characteristics based on non-metric data, the results may be completely implausible. Some of the assumptions may be less serious, but they may still have a negative impact on the results. Excessive violation of certain assumptions can lead to completely misleading results (Verma and Abdel-Salam, 2019).

The sample size plays a big role in determining the power ($1-\beta$) of a statistical inference test. Power is based on the effect size, the significance level (α), and the sample size. Type II error is the chance of not rejecting the null hypothesis when it is actually false. For any given alpha level, increasing the sample size will always increase the power of the statistical test. In other words, sample size enhances statistical power by reducing sampling error. As the sample size increases, researchers must decide whether the power is too high. As the sample size increases, excessive power leads to the statistical significance of smaller effects, such as correlations, and at very large sample sizes, nearly all effects become significant. Researchers should always be aware that sample size can affect statistical tests by making them either insensitive (for small sample sizes) or overly sensitive (for very large sample sizes) (Hair et al., 2019, p. 145-150).

Burns and Veeck (2020) highlighted the importance of selecting the sample using random sampling methods; larger samples yield higher accuracy and a smaller margin of error. However, in practice, non-random sampling methods are often used instead of random sampling methods. In these sampling methods, units have no chance of being selected for the sample. Therefore, statistically, it can be said that the ability of the sample to represent the population is low (Burns and Veeck, 2020; Nakip, 2003). Therefore, Nakip (2003) points out that it is not correct to generalise the results obtained from this sample, especially since the sample obtained through convenience sampling is not sufficiently representative of the population.

Large sample sizes are necessary in quantitative research, particularly when using a large number of variables. This is because increasing the sample size is the simplest and quickest way to reduce the cumulative effect of sampling errors between variables. However, when performing multivariate and complex analyses on the data, a large sample is necessary (Nakip, 2003). It can be said that there are different approaches in this respect. For example, the number of variables, factor loadings, etc.

Assumptions About Normality

Sir Francis Galton said of the normal distribution that “if the Greeks had known it, they would have personalised it and deified it.” This quote really describes the importance of the normal distribution. The normal distribution is a statistical tool that helps people visualise and measure natural variability between objects or events and is a key concept in statistical inference (Abrami, Cholmsky, and Gordon, 2000; Tabachnick and Fidell, 2013, p. 688-689). The normal distribution, which applies to almost all parametric tests, is one of the basic assumptions for making reliable inferences about the population.

According to Alpar (2014), the reliability of parametric analyses is reduced when the normality assumption is not met, and there are three options in this case. The first is to continue the analysis assuming that the data are normally distributed, which is a violation and is strongly discouraged. The second option involves transforming the variables that potentially violate the normality assumption. The third is to use non-parametric tests.

Researchers categorize their various approaches to assessing normality as either graphical or statistical. Statistical measures provide a guide to variables that show significant departures from normality, while normal probability plots provide a visual representation of the shape of the distribution. These two types of presentations complement each other in selecting appropriate transformations. Graphical methods enable the assessment of normality without requiring complex calculations (Hair et al., 2019). Graphical methods include histograms, root and leaf plots, normal Q-Q plots, P-P plots and box plots. Normality can be tested using the Shapiro-Wilk test or the Kolmogorov-Smirnov test. The Shapiro-Wilk test is more appropriate for small samples ($n \leq 50$) (it can be used for sample sizes up to 2000 observations), while the Kolmogorov-Smirnov test is used for larger samples. One of the weaknesses of these tests is that they give significant results for large samples, even with small deviations from normality (Verma and Abdel-Salam, 2019).

Perfect normality of data occurs when skewness and kurtosis are ‘zero’, which is rare in practice (Hair et al., 2019; Verma and Abdel-Salam, 2019; Gürbüz and Şahin, 2019; Alpar, 2014; Abrami, Cholmsky and Gordon, 2000). Values above or below zero represent deviations from normality. It can be said that the relevant variable deviating from full

normality is approximately normal, and parametric tests can be applied within the range of ± 1 skewness and kurtosis (Hair et al., 2019; Verma and Abdel-Salam, 2019; Gürbüz and Şahin, 2018). However, some studies suggest that we can accept these thresholds up to about 2 or 3.

In recent years, there has been a widespread belief that partial least squares (PLS) has advantages that make it more suitable than other statistical estimation techniques such as regression or covariance-based structural equation modelling (CB-SEM). These advantages include small sample sizes, non-normal distributions, and non-metric levels of measurement (Hair et al., 2019; Hair, Ringle and Sarstedt 2011; Fornell and Bookstein, 1982, p.440-442). These beliefs have led to the widespread adoption of PLS. However, Goodhue, Lewis and Thompson (2012) found that PLS has no advantage over other techniques for non-normal data or small sample sizes. The authors noted that path estimates suffer from increasingly large standard deviations due to decreasing sample sizes and a reduction in statistical power. They also noted that small or moderate deviations from normality are not significant, but extreme cases of non-normality show significant losses (Goodhue, Thompson and Lewis, 2013; Goodhue, Lewis and Thompson, 2012, p.992-1001). However, it is evident that this area requires further research.

Methodology

The purpose of this study is to examine researchers' understanding of the basic assumptions of statistical tests and their behaviour in testing these assumptions through articles published in journals. The study only looked at violations of the assumptions of data type, sample size, and normal distribution. A study must check and meet these basic assumptions. Violation of these assumptions can make the results of the study unreliable or even wrong. A significant methodological gap in tourism research is the absence of studies that systematically examine whether the parametric assumptions underlying widely used quantitative techniques such as regression, ANOVA, SEM, and PLS-SEM (e.g. normality, sample size, data type) are actually met. This gap creates a risk that the quantitative findings accumulated in the tourism literature rest on largely invisible methodological problems, including assumption violations and inflated error rates. In tourism and hospitality, recent methodological discussions have highlighted misapplications of advanced quantitative techniques such as SEM and PLS-SEM, as well as design issues related to sample size, response bias, and unobserved heterogeneity, all of which are closely connected to how statistical assumptions are specified and interpreted in empirical research (Nunkoo, 2018). Our study examines articles published in tourism journals indexed in the SSCI to reveal how these assumptions are addressed. In doing so, it offers a critical, evidence-based contribution to tourism methodology and provides concrete recommendations for more rigorous reporting and oversight by editors, reviewers, and researchers. Furthermore, the study's scope does not encompass the assumptions required for each individual analysis. In order to achieve this objective, studies conducted in the last five years in the field of tourism and published in a journal included in the SSCI index were analysed on the basis of the basic assumptions.

The empirical corpus of this study consists of quantitative articles published in SSCI-indexed tourism journals. Journals were first identified using the Clarivate Analytics Journal Citation Reports, focusing on outlets whose primary scope is tourism and closely related fields, and all issues published between 2020 and 2024 were screened. We applied the following inclusion criteria: (a) the article reported an empirical, quantitative study; (b) at least one parametric statistical technique was used (e.g. t test, ANOVA, regression, SEM, PLS-SEM); and (c) the manuscript was written in English. Conceptual and theoretical papers, qualitative studies, mixed-methods articles that did not employ parametric tests, editorials, book reviews, non-English manuscripts, and articles without full-text access were excluded. Using this convenience sampling approach, and restricting the corpus to freely available full-text articles, a total of 239 articles from 19 SSCI-indexed tourism journals were retained for analysis. These studies were then examined to determine whether analyses were conducted to check whether the relevant parametric assumptions were met and how the results of these checks were presented.

The final set of articles represents a convenience sample of SSCI-indexed tourism journals to which full-text access was available through institutional subscriptions. This approach enabled an in-depth examination of recent practice in high-impact outlets but may introduce an accessibility bias, as journals or articles that are less easily accessible could be underrepresented. The aim was therefore not to construct an exhaustive census of all tourism research, but to provide an analytically rich snapshot of how parametric test assumptions are currently treated in leading tourism journals.

The study employed a directed content analysis to examine how statistical assumptions were addressed in the selected articles. The coding scheme was developed a priori, drawing on the methodological literature on parametric tests and their assumptions. For each article, we coded whether the following assumptions were: (1) explicitly mentioned and checked, (2) not mentioned at all, or (3) explicitly stated as "not needed":

- Data type (e.g. use of interval/ratio data for parametric tests),
- Normality of the distribution of variables or residuals,
- Sample size adequacy for the statistical technique employed, and
- The overall treatment of assumptions across the analyses reported.

The initial coding protocol was piloted on a subset of articles, during which categories and decision rules were refined to improve clarity and consistency. Following this calibration phase, two independent coders with training in quantitative methods applied the coding scheme to the full sample of 239 articles. Inter-coder reliability was assessed using Cohen's kappa for each main category, with all coefficients exceeding commonly accepted thresholds for substantial agreement. Discrepancies were resolved through discussion until consensus was reached. This procedure ensured that the classification of how assumptions were handled was both systematic and replicable.

Results

Descriptive statistics of the articles analysed in the study are presented in Table 1.

Table 1. Statistics of the analyzed articles

Years	f	%
2020	51	21
2021	44	18
2022	63	26
2023	56	23
2024	65	3
Total	239	100

In Table 1, a total of 239 studies were examined in which parametric analyses were applied in 19 journals indexed in the SSCI index between 2020 and 2024. The fact that the number of articles examined in 2024 is lower than in other years is due to the fact that the number of articles published in the first quarter of 2024 is lower compared to the whole year and the full text of these articles is not accessible.

Table 2. Types of analyses used in the articles reviewed

Multivariate analyzes in articles	f*
Structural Equation Model (SEM/ CB-SEM)	103
Multilevel Structural Equation Modeling (MSEM)	5
Partial Least Squares Structural Equation Modeling (PLS-SEM)	84
Regression	11
Hierarchical Regression	6
Ordinary Least Squares Regression	3
PROCESS- Mediating-Moderating Effect	18
ANOVA	14
MANOVA	4
ANCOVA	2

*The number of analyses in the table is higher than the number of studies examined. This is due to the use of more than one parametric analysis in some studies.

In Table 2 above, it can be seen that SEM and PLS-SEM analyses are more frequently used in the studies. While the reasons for using these analyses were explained in 91 studies, they were not mentioned in the other studies. Of the articles that did give an explanation, 40 were studies that used PLS-SEM analyses. In 25 of these studies it was emphasized that PLS-SEM was used because the assumptions of the parametric test were not met. In the others, it is emphasized that this method is popular or has recently been widely used.

Explanations of the parametric test assumptions in the articles are given in Table III.

Table 3. Statistics of parametric assumptions

Assumptions	Mentioned f	Mentioned %	Not-Mentioned f	Not-Mentioned %	Not Needed f	Not Needed %
Data Type	227	%94	12	%5	-	-
Normality	102	%43	123	%51	14*	%6
Sample Size	70	%29	139	%58	30*	%13
All Assumption	58	%24	137	%57	44*	%18
*It was highlighted that in some of the studies that performed PLS SEM analysis, the normality and sample size assumptions did not need to be met. Some PLS-SEM studies mentioned that they applied this analysis because the assumptions were not met after considering these two assumptions. However, in some studies it was stated that PLS SEM was applied even though all assumptions were met.						

An analysis of Table 3 shows that only 58 out of 239 studies met the assumptions of measurement type, normality and sample size. These assumptions were violated in about half of the remaining studies. In 18% of the articles analysed, it was stated that parametric test assumptions did not need to be met. All of these studies used PLS-SEM analysis. However, the articles did not clearly state why these assumptions did not need to be met.

Analyzing the type of measurement in the studies revealed that 227 of them mentioned it. The studies generally used Likert scales. 97 of the articles used a 5-point Likert scale, 111 used a 7-point scale, 2 used a 6-point scale, and 1 used a 10-point scale. 28 of them employed both the 5, 7 and 6 point Likert scales. It was noted that there was no mention of how the measurement of different variables in the measurement tool with different levels of measurement affects the analyses between variables, or how these levels of measurement are transformed. Despite the literature's debate over whether the Likert scale is intermittent or proportional, researchers generally accept it as an intermittent scale. In 12 of the analysed articles, it is unclear what the measurement level is. Whatever the analysis, it's important not to specify the measurement level or scale used in the studies.

The articles analysed used different methods to test for conformity to the normal distribution. The articles tested the normal distribution using kurtosis and skewness, the Kolmogorov-Smirnov test, Mardia's coefficients, and histogram plots. Some studies stated that the data showed a normal distribution without specifying which test was used to analyse normality. The table reveals that 51% of the analysed articles failed to explain how they met the assumption of normal distribution. In 44 of the studies using PLS-SEM, it was stated that the data were not normally distributed and that there was no need to test this assumption.

When analysing the assumption about the sample size, almost all studies describe in detail the characteristics of the sample selected and the reasons for the choice of sampling methods. However, 70 studies provide an explanation of the sample size calculation method. These studies found that the sample size calculation relied on G*Power, reference studies from the literature, and specific multiples of the number of questions in the measurement tool. A few studies relied on the EFA Bartlett's test to determine the adequacy of the sample for analysis.

However, when looking at the sample size in the study, the research sample included a minimum of 70 and a maximum of 4074 people. The average sample size in other studies is between 300 and 400 people. In 62 studies that used PLS-SEM and reported sample size, the sample size was 200 or more. More than half of these have a sample size of 500 or more. However, in the same studies, researchers applied PLS-SEM analysis without specifying whether there was a parametric test violation in terms of measurement type or normal distribution. Do researchers choose this type of analysis because it is popular? Furthermore, in certain articles that employ multivariate analyses such as SEM, PLS-SEM, CB-SEM, and PROCESS, the researcher(s) do not need to formulate parametric test assumptions when providing goodness of fit values.

Conclusion

Scientific research in any field seeks answers to a number of questions. The data collected to answer these questions is tested using statistical analysis. To do this, researchers should first ask and check whether some basic assumptions are met in order to decide on the appropriate test statistic. These basic assumptions are: data type, normal distribution, and sample size. If these assumptions are not met, the use of statistical techniques is considered to be a serious problem in data analysis. Failure to meet these assumptions also affects the reliability of the analysis and reduces the credibility of the research results.

This study examined researchers' understanding of these assumptions and their control behavior through articles published in journals. We analysed articles published in the field of tourism in journals indexed in the SSCI over the last five years for this purpose. As a result of this review, it was observed that SEM/CB-SEM and PLS-SEM were generally used in the articles. It was found that more than half of the articles did not mention whether these basic assumptions were met or not. What does it mean that the articles do not mention whether the basic assumptions are met or not? Does it mean that these assumptions are met or not? Approximately one fifth of the articles stated that it was not necessary to meet the assumptions in question, and all of these articles used PLS-SEM analysis. First of all, the most violated or ignored assumption in the analysed articles is the sample size. Second, we found violations or ignored the assumption of normal distribution. The data type is the least violated or most mentioned parametric test assumption in the analyzed articles.

There may be a variety of reasons why researchers did not check or violate the parametric test assumption in the articles analysed. The first of these reasons may be the researchers' lack of knowledge about the assumption. Hoekstra, et al. (2012) found that researchers were relatively uninformed about when and how to check whether data violated the assumptions of statistical tests. Another reason could be the researchers' lack of knowledge on how to verify the violation of an assumption. A third reason could be the absence of a known alternative in the event of an apparent violation of an assumption. In all three cases, it seems important to consider how to improve the statistical training of researchers. The fourth reason could be the lack of concern for potential assumption violations (e.g., PLS-SEM). The literature claims that such analyses do not need to check and fulfil these assumptions, but other studies disagree. In this case, how should researchers proceed, and which result should they consider? Generally, the reviewed studies advise researchers to employ this analysis, even if the assumptions are not satisfied. However, it is clear that more research is needed to decide this question.

But should it be considered a violation of the studies not to mention whether the assumptions are met or not? What conclusions can be drawn from this situation? Do researchers not know or care about the assumptions? Or does not mentioning them mean that the assumptions are already met? Analysing a study with so many questions can be confusing for the reader. In this respect, I think it would not be wrong to say that the actors involved in the publication process, namely editors and reviewers, have a duty. As suggested by Bakker and Wicherts (2011), asking authors to explain how they check whether assumptions are violated would force researchers at both ends of the publication process to show more awareness of this important issue.

Theoretical Implications

This study examined how researchers understood the basic assumptions of parametric tests and how they behaved when testing these assumptions through articles published in journals. In the literature, there are studies on violations of assumptions in analyses. These studies, however, either examined each assumption separately or considered the specific assumptions of a single type of analysis. However, Hoekstra et al. (2012) conducted a quantitative study with 30 researchers to examine the basic assumptions of parametric analyses. This study examined all the basic assumptions of parametric analyses through published articles. This study addresses a significant gap in the existing literature on this topic.

Negative consequences may arise from failing to verify the fulfilment of parametric test assumptions or failing to provide relevant information. First of all, violations of assumptions may reduce the reliability of research results. Therefore, problems may arise in the interpretation and generalization of research findings. Furthermore, serious violations of assumptions may lead to a weakening or questioning of the research's basic theoretical assumptions. In this case, the research's theoretical contribution may become questionable. This can lead to a loss of credibility and criticism of scientific knowledge.

The results of this study are theoretically important because understanding why assumptions are violated can lead to a deeper understanding of the relevant field of science and contribute to the development of better-designed research models for future research. Violations of assumptions in studies, or failure to provide information that these assumptions have been checked, can set a negative example for readers and lead to further errors. Researchers can contribute to the correct and healthy progress of science by verifying or verifying these assumptions and increasing their awareness of this issue.

This study examines the basic assumptions of parametric analysis. However, each parametric analysis has its own special assumptions. In the future, another study could potentially verify the violation of these special assumptions. This study analyzed articles in field of tourism published in SSCI-indexed journals between 2020 and 2024. A more comprehensive study covering all years and all journals can be carried out in the future. Researchers in a large sample

group can conduct quantitative studies that measure the level of knowledge of parametric test assumptions by providing datasets.

Practical Implications

Positivism asserts that science advances through the accumulation of truths or the elimination of falsehoods (e.g., Karl Popper). To achieve this, positivism employs highly structured research methods and contributes to knowledge production by analyzing quantified quantitative data using statistical methods (Gürbüz & ahin, 2018; Demir, 2000). Whether we consider the principle of verifiability or the principle of falsifiability of positivism, in both cases, we need accurate and sound results for the progress of science. The research must first fulfill the basic assumptions of the statistical analysis to obtain accurate results.

The study's results revealed that many of the analysed articles violated the basic assumptions. In the conclusion section of the study, it is stated that these violations may occur for various reasons. Expert educators should offer serious courses in research methods and quantitative analysis techniques to researchers at the master's and doctoral level to prevent these violations and ensure the correct progress of science. Expert trainers are scientists who are competent in the field of quantitative analysis and statistics and who can apply statistical techniques to the social sciences. Furthermore, encouraging researchers to attend training courses on research methods and quantitative analysis could be another recommendation from the study. The study recommends offering these trainings at no cost to encourage a wider participation, or providing financial support to researchers who attend.

Making researchers aware of the assumptions of parametric analysis and the significance of meeting these assumptions is also crucial. This can be achieved through training on the subject. The actors involved in the publication process are another important issue. Journal editors and reviewers should ensure that the reviewed studies meet the basic assumptions of parametric analyses and report the results. In addition, it can be communicated to the authors that checks on these assumptions should be included in the journal's writing rules and that relevant information should be provided in the study.

One of the key findings of the research relates to PLS-SEM models, which have been widely used by researchers in recent times. Some studies assert that PLS-SEM models are non-parametric, allowing for the disregard of parametric test assumptions. On the other hand, some studies assert that PLS-SEM models are parametric, necessitating the fulfilment of assumptions. This situation confuses us as researchers. However, researchers may prefer PLS-SEM models either because they cannot meet the assumptions of parametric tests and want to run parametric tests or because these models are very popular in research. Are PLS-SEM models parametric or nonparametric? There is a need for more research to elucidate the rationale behind disregarding assumptions in PLS-SEM models. More data is required to better understand the practical applications of PLS-SEM models, and researchers should conduct further studies to determine the conditions under which these models are parametric or nonparametric.

References

- ABRAMI, P. C., CHOLMSKY, P., and GORDON, R., (2000). Statistical analysis for the social sciences: An alternative approach. United States of America: A Pearson Education Company.
- ALPAR, R. (2014). Uygulamalı istatistik ve geçerlik-güvenirlik. Ankara: Detay Yayıncılık.
- BAKKER, M., and WICHERTS, J. M. (2011). The (mis)reporting of statistical results in psychology journals. Behavior Research Methods, 43, 666–678.
- BURNS, A.C., and VEECK, A., (2020). Marketing research (9th edition). New York: Pearson Education.
- CHOI, P. T., (2005). Statistics for the reader: what to ask before believing the results. Canadian Journal of Anesthesia, 52 (6), pp. 1-5
- DEMİR, Ö. (2000). Bilim felsefesi. Ankara: Vadi Yayınları.
- DOĞAN, İ. and DOĞAN, N. (2014). Parametrik olmayan istatistiksel yöntemler. Ankara: Detay Yayıncılık.
- FORNELL, C. and BOOKSTEIN, F.L., (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory, Journal of Marketing Research, 19 (4), Special Issue on Causal Modeling, pp. 440-452.
- FORSTMEIER, W., WAGENMAKERS, E-J., and PARKER, T. H., (2017). Detecting and avoiding likely false-positive findings – a practical guide. Biological Reviews, 92, pp. 1941-1968.
- GOODHUE, D.L., LEWIS, W., and THOMPSON, R., (2012). Does pls have advantages for small sample size or non-normal data?, MIS Quarterly, 36 (3), pp. 981-1001.
- GOODHUE, D.L., THOMPSON, R., AND LEWIS, W., (2013). Why you shouldn't use PLS: four reasons to be uneasy about using PLS in analyzing path models, 46th Hawaii International Conference on System Sciences, 07-10 January 2013, Wailea, HI, USA.

- GÜRBÜZ, S. and ŞAHİN, F., (2018). Sosyal bilimlerde araştırma yöntemleri. Ankara: Seçkin Yayınevi.
- HAIR, J. F., RINGLE, C. M., and SARSTEDT, M., (2011). PLS-SEM: Indeed a silver bullet, *Journal of Marketing Theory and Practice*, 19 (2), pp. 139-152.
- HAIR, J.F., BLACK, W.C., BABIN, B.J., and ANDERSON, R.E., (2019). *Multivariate data analysis* (8th edition). United Kingdom: Cengage Learning EMEA.
- HOEKSTRA, R., KIERS, H. A. L., and JOHNSON, A., (2012). Are assumptions of well-known statistical techniques checked, and why (not)?, *Frontiers in Psychology*, 3, pp. 1-9.
- KURTULUŞ, K. (2006). *Pazarlama araştırmaları* (8. Baskı). İstanbul: Literatür Yayıncılık.
- MASON, R. D., LIND, D. A., and MARCHAL, W. G., (2000). *Statistical Techniques in business and economics* (10th edition). Singapore: McGraw-Hill Companies, Inc.
- NAKIP, M. (2003). *Pazarlama Araştırmaları*. Ankara: Seçkin Yayıncılık.
- NUNKOO, R. (2018). The state of research methods in tourism and hospitality. In R. Nunkoo (Ed.), *Handbook of research methods for tourism and hospitality management* (pp. 3–23). Cheltenham, UK: Edward Elgar.
- OLSEN, C. H., (2003). Review of the use of statistics in infection and immunity, *Infection and Immunity*, 71 (12). pp. 6689–6692.
- OSBORNE, J. W. And WATERS, E., (2002). Four assumptions of multiple regression that researchers should always test, *Practical Assessment, Research, and Evaluation*, 8 (1), pp. 1-5.
- TABACHNICK, B., G., FIDELL. L. S. (2013), *Using Multivariate Statistics*, Sixth Edition. by Pearson Education,, Link: http://ndl.ethernet.edu.et/bitstream/123456789/27657/1/Barbara%20G.%20Tabachnick_2013.pdf
- TABACHNICK, B., G., FIDELL. L. S (2014), *Using Multivariate Statistics*, Sixth Edition. by Pearson Education, Link: https://api.pageplace.de/preview/DT0400.9781292034546_A24616694/preview-9781292034546_A24616694.pdf
- VERMA, J.P. and ABDEL-SALAM, G. A-S., (2019). *Testing statistical assumptions in research*. United States of America: John Wiley & Sons, Inc.

Özet

Çalışma, turizm alanında gerçekleştirilen nicel araştırmalarda parametrik test varsayımlarına uyumun ne ölçüde sağlandığını ortaya koymayı amaçlamaktadır. Özellikle SSCI indeksli turizm dergilerinde yayımlanan makaleler üzerinde odaklanan araştırma, alan yazında istatistiksel analizlerin güvenilirliği ve geçerliliği açısından kritik bir boşluğa dikkat çekmektedir. Parametrik testler, doğru yorumlanabilmeleri için belirli varsayımların sağlanmasına ihtiyaç duymaktadır. Ancak bu varsayımlar sağlanmadığında, elde edilen bulguların bilimsel güvenilirliğini zayıflatılabileceği gibi, turizm araştırmalarının metodolojik yönden zayıf kalmasına neden olabilir. Bu nedenle çalışmada bilimsel dergilerde bu varsayımlara ne kadar dikkat edildiğinin ortaya konmasını amaçlamaktadır. Araştırmanın verileri, 2020–2024 yılları arasında yayımlanmış ve açık erişim olarak sunulmuş 239 makaleden elde edilmiştir. Makaleler, SSCI kapsamındaki 19 farklı turizm dergisinden seçilmiştir. İçerik analizi yöntemi ile yürütülen incelemede, her bir çalışmada kullanılan istatistiksel yöntemler, bu yöntemlerin dayandığı temel parametrik test varsayımları ve bu varsayımlara ilişkin raporlama düzeyleri detaylı biçimde değerlendirilmiştir. İnceleme sürecinde, özellikle örneklem büyüklüğü, normallik varsayımı, varyans homojenliği ve ölçek türü gibi temel kriterler dikkate alınmıştır.

Elde edilen bulgular, turizm araştırmalarında yapısal eşitlik modellemesi (SEM) tekniklerinin yaygın olarak kullanıldığını göstermektedir. Özellikle kovaryans temelli SEM (CB-SEM) ve kısmi en küçük kareler yöntemi (PLS-SEM) en sık tercih edilen analiz yöntemleri olmuştur. Bununla birlikte, incelenen makalelerin yarısından fazlasında parametrik testlerin temel varsayımlarına ya hiç değinilmediği ya da yalnızca yüzeysel bir şekilde bahsedildiği tespit edilmiştir. En sık ihmal edilen unsurların başında örneklem büyüklüğüne ilişkin değerlendirme ve normallik testleri gelmektedir. Dikkat çekici bir bulgu olarak, incelenen çalışmaların yaklaşık beşte biri varsayımların gerekli olmadığını açıkça ifade etmiştir. Bu grubun tamamında PLS-SEM yönteminin tercih edilmiş olması, bu yöntemin varsayımlardan bağımsız çalışabildiği algısının araştırmacılar arasında yaygın olduğunu göstermektedir.

Bu sonuçlar, turizm araştırmalarında yöntemsel titizliğin güçlendirilmesi gerektiğine işaret etmektedir. Parametrik test varsayımlarının göz ardı edilmesi, sonuçların bilimsel değerini azaltılabileceği gibi, politika yapıcılar ve sektör temsilcileri açısından yanlış çıkarımlara yol açabilir. Dolayısıyla araştırmacıların, istatistiksel analizlerde yalnızca uygun yöntemi seçmekle kalmayıp, bu yöntemlerin gerektirdiği koşulları titizlikle kontrol etmeleri ve raporlamaları büyük önem taşımaktadır. Sonuç olarak, bu çalışma, turizm alanında bilimsel araştırma kalitesini artırmak amacıyla, parametrik test varsayımlarına ilişkin farkındalığın güçlendirilmesi gerektiğini vurgulamaktadır. Özellikle lisansüstü eğitimde ve akademik araştırma projelerinde, istatistiksel varsayımlar konusunda daha kapsamlı metodoloji eğitimleri verilmesi önerilmektedir. Bu sayede, hem araştırma bulgularının güvenilirliği artacak hem de turizm literatüründe daha sağlam bilimsel temellere dayalı çalışmaların sayısı çoğalacaktır.