

Pesquisa quantitativa em turismo: os dados gerados são válidos e confiáveis?

DOI: 10.2436/20.8070.01.134

Thiago Coelho Soares

Doutor em Administração pela Universidade Federal de Santa Catarina, Brasil.
Professor do Programa de Pós-Graduação em Administração da Universidade do Sul de Santa Catarina, Brasil.

E-mail: thiago.soares@unisul.br

João Coelho Soares

Mestre em Administração pela Universidade do Sul de Santa Catarina, Brasil.

E-mail: joacsoares88@gmail.com

Sandro Vieira Soares

Doutor em Contabilidade pela Universidade de São Paulo. USP, Brasil.
Professor do Programa de Pós-Graduação em Administração da Universidade do Sul de Santa Catarina, Brasil.

E-mail: sandro.soares@unisul.br

Resumo

O objetivo desta pesquisa é verificar se a geração de dados primários nas publicações nacionais em turismo é válido e confiável. Para tanto se analisou os periódicos da área com Qualis B2 ou superior. O processo de pesquisa foi realizado em três etapas: Delimitação do escopo de análise; Seleção dos artigos; análise de forma quantitativa, visando descrever as principais características dos trabalhos selecionados. Considerando que na pesquisa quantitativa o rigor é caracterizado pela validade e confiabilidade dos dados gerados, se estudou se os artigos apresentavam rigor científico. Procurou-se identificar nos artigos que possuíam pesquisa de campo quantitativa com questionário se foram adotados procedimentos estatísticos para validar e confiar os resultados, tais como: Análise Fatorial; Teoria da resposta ao Item; Análise da consistência interna; Análise da linearidade dos dados; Análise da homoscedasticidade dos dados; Análise de

multicolinearidade dos dados. Os resultados apontaram para uma pequena quantidade de estudos adotando tais técnicas. Desta forma se conclui que o campo da pesquisa em Turismo precisa se preocupar em gerar dados primários válidos e confiáveis.

Palavras-chave: Turismo. Pesquisa Quantitativa. Validade. Confiabilidade. Questionário.

1 INTRODUÇÃO

A análise da produção científica de um país, seja ela relativa a todas as áreas ou especificamente de uma área de pesquisa, é fundamental para indicar os rumos que a ciência está seguindo, seus desdobramentos, evitar possíveis equívocos e corrigir rotas (SANTOS, NETTO, WANG, 2017).

O turismo enquanto atividade econômica passou por longos períodos sendo planejado de forma ineficaz. Isso se deu pela falta de bibliografia e de profissionais qualificados. A desvalorização dos profissionais do turismo é reforçada principalmente na academia, ou seja, no amadurecimento do turismo enquanto ciência, quantidade de pesquisadores de formações diversas escrevendo sobre o turismo. Com o aumento de bibliografias, as formações técnicas e acadêmicas vêm crescendo. O turismo vem conquistando espaço e atenção acadêmica, pública e política, porém ainda existem fragilidades nas questões acadêmicas que influenciam as outras esferas (SEIXAS, CARON, 2018).

A área de turismo necessita estabelecer relações e parâmetros que possibilitem compreender o processo de construção epistemológica e metodológica da pesquisa (ROCHA, COSTA, ANJOS, 2017). Os avanços dos estudos acadêmicos em turismo, de origem multidisciplinar, têm revelado novas possibilidades epistemológicas, teóricas e metodológicas para a consolidação deste campo de pesquisas (COSTA et al., 2018).

A pesquisa desenvolvida por Seixas e Caron (2018) sobre as publicações sobre o turismo identificou um aumento de pesquisadores das áreas de administração, geografia, economia e comunicação publicando sobre turismo, e um aumento bastante significativo nos últimos doze anos de pesquisadores doutores das áreas do turismo. Em relação às abordagens de estudos, identificou-se uma tendência de aumento no planejamento e nos negócios empresariais.

Na área de turismo é preocupante a quantidade de metodologias distintas encontradas. Isto é reflexo da quantidade de pesquisadores de outras áreas e do turismo ser uma área inter e multidisciplinar. É inegável a necessidade, após anos de estruturação científica do turismo, de metodologias serem estabelecidas pela e para a área. Seria prudente a realização de pesquisas sobre seu próprio processo de desenvolvimento (metodologias e teorias), para que as demais pesquisas da área sejam melhores desenvolvidas e mais rapidamente consolidadas (SEIXAS; CARON, 2018). Rocha et al. (2017) consideram relevantes novas pesquisas voltadas para os artigos publicados nas revistas nacionais de turismo, colaborando assim com o conhecimento de novas articulações teóricas e perspectivas metodológicas que contribuam para a evolução científica da área de turismo. Costa et al. (2018) afirma que devem existir reflexões acerca da adequação entre os instrumentos de coleta de dados, objetivos e métodos de pesquisas propostos pelos estudiosos do Turismo. Para Santos et al. (2017) a pesquisa científica brasileira em turismo ainda apresenta algumas características de áreas incipientes.

Para Hair, Harrison e Risher (2018) as ciências sociais estão começando a reconhecer o potencial de métodos analíticos mais complexos, sendo que seus pesquisadores ainda carecem aderir e aplicar análise de dados com maior frequência em

suas pesquisas. Pesquisas nacionais quantitativas, na área do turismo, buscaram, por exemplo, identificar as emoções evocadas em experiências negativas de turismo (MEDEIROS, GOSLING, VERA, 2015) e analisar o uso de instrumentos de coleta de dados nas produções acadêmicas em turismo (COSTA et al., 2018).

No tocante aos métodos de pesquisa quantitativa, estes são utilizados quando se quer medir opiniões, reações, sensações, hábitos e atitudes etc. de um universo (público-alvo) através de uma amostra que o represente de forma estatisticamente comprovada.

Desta forma, decidiu-se pela realização do presente estudo com a seguinte pergunta de pesquisa: a geração de dados primários nas publicações nacionais em turismo é válido e confiável?. Com isso, o objetivo desta pesquisa é verificar se a geração de dados primários nas publicações nacionais em turismo é confiável. Para isto analisou-se os periódicos da área com Qualis B2 ou superior.

2 DESENVOLVIMENTO

Os dados primários coletados necessitam ser inicialmente analisados com propósito de examinar exploratoriamente conforme recomendam Hair, Black, Babin e Anderson (2010). Desta forma, busca-se examinar as características dos dados coletados e se esses permitem a aplicação das técnicas estatísticas a serem utilizadas, atenuando os vieses na análise e interpretação dos resultados. Esta seção apresenta construção de um questionário, sua validação e análise de consistência interna: análise da normalidade dos dados, análise da linearidade dos dados, análise da homoscedasticidade dos dados e análise de multicolinearidade dos dados.

2.1 Questionário

O questionário é uma das formas de coletar os dados primários. A coleta de dados estruturada é feita por meio de questionários elaborados de modo a mostrar o verdadeiro objetivo da pesquisa ao entrevistado. Conforme Malhotra (2012, p. 140), “Em um questionário típico, a maioria das perguntas é do tipo de alternativa fixa, que exige que o entrevistado faça sua escolha em um conjunto predeterminado de respostas”.

Questionários com perguntas fechadas trarão alternativas específicas para que o respondente escolha uma delas. As perguntas abertas são aquelas que permitem liberdade ilimitada de respostas ao respondente. Nelas poderá ser utilizada linguagem própria do respondente. Esta última mais adequada para instrumentos de coleta qualitativos.

Soares e Lima (2018) afirmam que a inserção de legendas ou glossários junto aos questionários facilita o entendimento dos respondentes, tendo em vista que estes estarão sozinhos e sem auxílio do pesquisador no momento de responder ao questionário.

Decorrente da evolução da tecnologia e abrangência da Internet, as pesquisas online através de questionários não apenas continuam a crescer em popularidade e a amadurecer como uma técnica de pesquisa, mas as novas tecnologias também mudaram as metodologias e as possibilidades de pesquisa (EVANS, MATHUR, 2018).

2.2 Análise Fatorial

Antes de se aplicar a técnica de análise fatorial sobre um questionário, deve-se verificar se esta é adequada aos dados a serem analisados. O teste de Kaiser-Meyer-

Olkin (KMO) consiste em verificar o grau de intercorrelações das variáveis e a adequação ao uso da Análise Fatorial (SOARES, 2018).

A classificação quanto ao KMO, sugerida por Fávero, Belfiore, Silva e Chan (2009), pode ser vista na Tabela 1 a seguir. Para a teoria da resposta ao item os valores variam entre 0 (zero) e 1 (um).

Tabela 1 – Estatística Kaiser-Meyer-Olkin (KMO)

Intervalo do teste	Análise fatorial
$KMO \geq 0,9$	Muito boa
$0,8 \leq KMO < 0,9$	Boa
$0,7 \leq KMO < 0,8$	Média
$0,6 \leq KMO < 0,7$	Razoável
$0,5 \leq KMO < 0,6$	Má
$KMO < 0,5$	Inaceitável

Fonte: Fávero *et al.* (2009).

Conforme Hair *et al.* (2010), a carga fatorial de um item que compõe o questionário indica o percentual de covariância existente entre o item e o fator. Desta forma, um item que obtiver um alto valor para carga fatorial indica que ele está altamente associado com outros itens e que eles representam um mesmo constructo, conforme Tabela 2.

Tabela 2 – Carga fatorial: orientações para o tamanho da amostra

Carga Fatorial	Tamanho necessário da amostra para significância de 5%
0,30	350
0,35	250
0,40	200
0,45	150
0,50	120
0,55	100
0,60	85
0,65	90
0,70	60
0,75	50

Fonte: Hair *et al.* (2010).

Assim, a análise fatorial analisa a estrutura das correlações entre muitas variáveis (por exemplo, escores de testes, itens de testes, respostas de questionários) definindo um conjunto de dimensões latentes comuns, chamadas de fatores (SOARES, 2018).

2.3 Teoria da Resposta ao item (TRI)

A teoria de resposta ao item surgiu na década de 1950. Porém, por conta da complexidade dos cálculos matemáticos para a estimação dos parâmetros dos modelos, começou a ser utilizada com mais frequência a partir da década de 1980, com o desenvolvimento de *softwares* (SOARES, 2018).

A TRI corresponde a um conjunto de modelos matemáticos que permitem avanços nos estudos de validade e confiabilidade de instrumentos de coleta de dados por meio de análises aplicadas. Dentre os modelos da TRI, incluem-se: modelo de resposta nominal; modelo normal modificado; modelo de eleição múltipla; modelo de escala de estimação; modelo de crédito parcial; modelo de crédito parcial generalizado; modelo de resposta gradual e modelo contínuo (GOMES *et al.*, 2018).

A TRI indica a proporção da variância dos dados que pode ser considerada comum as variáveis do estudo. Existem dois pressupostos teóricos para este modelo: unidimensionalidade e independência local. (SOARES, 2018). A TRI utiliza modelos probabilísticos para fazer inferências, sendo possível desenvolver uma escala de medida padronizada. Permite fazer comparações entre itens, indivíduos e diferentes populações (BORTOLOTTI *et al.*, 2012)

2.4 Análise da consistência interna

A consistência interna é um indicador estatístico utilizado na verificação da confiabilidade de um instrumento de coleta ou escala de mensuração. O coeficiente Alpha de Cronbach (α) expressa a consistência interna com base na correlação entre os diferentes itens/variáveis mensurados e revela o grau de confiabilidade do instrumento ou escala analisada (CRONBACH, 1951). Essa técnica é uma das mais utilizadas em turismo na identificação do grau de confiabilidade de instrumentos e escalas (HAIR et al., 2010; MALHOTRA, 2012), conforme Tabela 3 a seguir.

Tabela 3 - Critérios de classificação da confiabilidade pelo Alpha de Cronbach

Muito Baixa	Baixa	Moderada	Alta	Muito Alta
$\alpha < 0,30$	$0,30 \leq \alpha < 0,60$	$0,60 \leq \alpha < 0,75$	$0,75 \leq \alpha < 0,90$	$0,90 \leq \alpha$

Fonte: Hair et al. (2010).

Como recomendam Hair et al. (2010), deve-se apresentar critérios de classificação da confiabilidade de instrumentos e escalas com base no Alpha de Cronbach (α). Aconselha-se também, a correlação por variável e o valor de Alpha se o item/variável for excluído. A correlação do item compreende o coeficiente de correlação (R de Pearson) de cada variável em relação as demais; enquanto o Alpha de Cronbach se item for excluído mostra o efeito da variável no resultado alpha total se o referido item/variável for excluído da análise.

2.5 Análise da normalidade dos dados

Uma das suposições preliminares na análise estatística inferencial é a normalidade da distribuição dos dados, que consiste em verificar se os dados coletados estão distribuídos segundo o conceito da curva de distribuição normal, conhecida como Distribuição de Gauss (HAIR et al., 2010). Uma das possibilidades é de utilizar o teste Kolmogorov-Smirnov para analisar se as distribuições subjacentes dos dados coletados têm características de uma distribuição normal (LILLIEFORS, 1967). Esse teste é indicado para amostras de tamanho superior a 50 observações/casos. Assim, as seguintes hipóteses podem ser estabelecidas:

H0 = dados da variável analisada se originem de uma distribuição normal.

H1 = dados da variável analisada não se originam de uma distribuição normal.

Os resultados do teste levam à rejeição da hipótese nula (H0) para todas as variáveis analisadas quando apresentam Sig 0,000 em favor de H1, ou seja, hipótese que

afirma que os dados coletados das variáveis dependentes não se originam de uma distribuição normal.

Segundo Hair et al. (2010) uma outra maneira de analisar a normalidade dos dados é pela análise dos valores de assimetria (skewness) e curtose (kurtosis) dos dados coletados. A distribuição dos dados pode ser considerada normal (ou próxima à normal) se o valor calculado da assimetria está entre 1 e -1 (SWEENEY, WILLIAMS; ANDERSON, 2013).

2.6 Análise da linearidade dos dados

A análise de linearidade visa verificar a existência de associação linear (correlação) entre as variáveis e a intensidade das relações entre as variáveis, o que representa a força real da relação (Hair et al., 2010). Quanto maior a correlação, maior tende a ser a linearidade entre as variáveis analisadas. Sugere-se o uso do coeficiente de correlação bivariada de Pearson para mensurar o grau de associação entre as variáveis analisadas (AAKER, KUMAR, DAY, 2013). Assume-se que valores da correlação podem ser de 1 a -1, sendo 1 relação positiva elevada, -1 relação negativa ou inversa, e 0 a ausência de relação entre as variáveis (HAIR, et al. 2010).

Como propõem Cohen, Cohen, West e Aiken (2003), a Tabela 4 mostra o critério de classificação utilizado na interpretação do nível de correlação entre variáveis analisadas.

Tabela 4 – Critérios de classificação da correlação

Valor de ρ	Correlação
0,10 a 0,30	Baixa
0,30 a 0,50	Moderada
0,50 a 1,0	Elevada

Fonte: Cohen et al. (2003).

2.7 Análise da homoscedasticidade dos dados

A homoscedasticidade dos dados envolve a verificação das relações de dependência entre as variáveis analisadas para conferir se variáveis dependentes exibem níveis iguais de variância em relação às variáveis independentes (HAIR et al., 2010). O teste de Levene para amostras independentes pode ser utilizado por ser amplamente utilizado em pesquisas (MALHOTRA, 2012) e analisa os desvios em relação à média dos grupos para confirmação da homogeneidade. O teste de White é outra opção para testar a homoscedasticidade dos dados. Assim, ao realizar/elaborar o teste, hipóteses devem ser estabelecidas, como:

H0 = a variável dependente não exhibe níveis diferentes de variância em relação à variável independente; ou seja, há homogeneidade na variância.

H1 = a variável dependente exhibe níveis diferentes de variância em relação à variável independente, ou seja, não há homogeneidade na variância.

As variáveis dependentes que apresentaram Sig superior a 0,05 sugerem a rejeição de H0; que afirma não haver homoscedasticidade.

2.8 Análise de multicolinearidade dos dados

A multicolinearidade dos dados consiste em verificar se há colinearidade entre as variáveis independentes em relação às variáveis dependentes. A multicolinearidade é observada quando uma variável pode ser explicada por outras variáveis em análise (HAIR et al., 2010) e acontece quando há alta correlação (relação de dependência) entre duas ou

mais variáveis independentes analisadas. O VIF (Variance Inflation Factor) é um indicador utilizado, frequentemente, no diagnóstico da multicolinearidade entre variáveis; há indicativo de problema de multicolinearidade quando esse valor é superior a 10 (HAIR et al., 2010).

As variáveis mensuradas que apresentam VIF inferior a 10, indicam que não há problema de multicolinearidade nos dados coletados, viabilizando a aplicação da técnica análise de regressão múltipla. Quando existe VIF superior a 10, a variável deve ser excluída na aplicação da técnica análise de regressão múltipla.

3 METODOLOGIA

Tendo em vista o objetivo de pesquisa, desenvolveu-se uma revisão sistemática da literatura, objetivando mapear e analisar pesquisas efetuadas acerca de geração de dados primários em turismo. Para a elaboração do presente artigo, o processo de pesquisa foi realizado em três etapas, conforme Figura 1 e discriminadas na sequência:

Figura 1 – Etapas de pesquisa



Fonte: Elaboração dos autores (2018).

Etapa 1 – Delimitação do escopo de análise: Primeiramente, foram definidos os periódicos de busca a serem utilizados, bem como os termos, sendo estes delineados a

partir de leituras preliminares acerca do tema. Ressalta-se também, a utilização de operadores booleanos dentro da caixa de busca dos sítios eletrônicos dos referidos portais, sendo a estrutura de busca operacionalizada por dois e três grupos de palavras com o operador booleano “AND” dentro do grupo, conforme Quadro 1:

Quadro 2 - Descritores

Descritores em português
Questionário AND “Análise Fatorial”
Questionário AND “Teoria da resposta ao Item”
Questionário AND “Alpha de Cronbach”
Questionário AND “Análise da consistência interna”
Questionário AND “correlação de Person”
Questionário AND “Análise da linearidade dos dados”
Questionário AND “Análise da homoscedasticidade dos dados”
Questionário AND “Análise de multicolinearidade dos dados”

Fonte: Elaboração dos autores (2018).

Definiram-se como periódicos para busca os que possuíam em seu nome “Turismo” e que fosse ao menos Qualis B2 na Área de “Administração Pública e de Empresas, Ciências Contábeis e Turismo: Scielo: revista acadêmica observatório de inovação do turismo; Caderno virtual de turismo (UFRJ); Revista brasileira de pesquisa em turismo; Revista de turismo - visão e ação; e Revista turismo em análise.

Etapa 2 – Seleção dos artigos pela filtragem por meio do *download* dos artigos que estavam disponíveis, de forma gratuita e completos. Nesta etapa observou-se que nem todos cumpriam este requisito. Após a realização desse processo restaram 212 artigos nos periódicos em que o termo questionário aparecia ao longo do artigo.

Etapa 3 – Inicialmente, os artigos foram analisados de forma quantitativa, visando descrever as principais características dos trabalhos selecionados, por meio das seguintes variáveis: (A) Autores; (B) Título do Artigo; (C) Ano de Publicação; (D) Análise Fatorial; (E) Teoria de Resposta ao Item; (F) Alpha de Cronbach; (G) Análise de Consistência Interna; (H) Teste de Kolmogorov-Smirnov; (I) Análise da Normalidade dos Dados; (J) Correlação de Person; (K) Análise da Linearidade dos Dados; (L) Teste de Levene; (M) Análise da Homoscedasticidade dos Dados; (N) Análise de Multicolinearidade dos Dados.

4 ANÁLISE DOS DADOS

Neste item se discutirá o quanto os artigos que se utilizam de questionários em pesquisas em Turismo têm explicitado a preocupação com as características dos dados coletados e se esses dados permitem a aplicação das técnicas estatísticas atenuando os vieses na análise e interpretação dos resultados.

Nas cinco revistas pesquisadas Revista Acadêmica Observatório De Inovação Do Turismo; Caderno Virtual De Turismo (UFRJ); Revista Brasileira De Pesquisa Em Turismo; Revista De Turismo - Visão E Ação E Revista Turismo Em Análise, totalizaram-se 1570 (um mil quinhentos e setenta) artigos, sendo destes 212 (duzentos e doze) artigos com questionários, ou seja, com geração de dados primários. O Quadro a seguir identifica os valores encontrados para as variáveis analisadas nesta pesquisa.

Quadro 2 - Testes encontrados em artigos

Teste	Ocorrências	Percentual
Questionário	212	100,00%
Análise Fatorial	33	15,57%
Teoria da resposta ao Item	0	0,00%
Análise da consistência interna (Alpha de Cronbach)	20	9,43%
Normalidade dos dados	4	1,89%
Análise da linearidade dos dados	3	1,42%
Análise da homoscedasticidade dos dados	1	0,47%
Análise de multicolinearidade dos dados	5	2,36%

Fonte: Elaboração dos autores (2018).

Dos 212 (duzentos e dozes) artigos que possuem questionários, 33 (trinta e três) aplicaram análise fatorial para testar a dimensionalidade de um teste. A análise fatorial serve para, entre outras aplicações, verificar se há apenas um fator dominante (constructo) responsável por responder por todos os demais. Assim, análise fatorial é uma técnica de análise multivariada que é aplicada a um conjunto de variáveis, com o objetivo de identificar quais são as variáveis mais relevantes na decomposição de cada fator, que são não correlacionados.

A análise fatorial é um recurso estatístico inapropriado se a correlação entre as variáveis estiver correlacionada na população (matriz identidade). Para isso se opta pelo teste de esfericidade de Bartlett. Porém nos artigos onde a análise fatorial não foi explorada, não foi apresentado este argumento para a não utilização da análise fatorial. Mondo, Hallmann e Burg (2018) lembram que ao replicar o questionário é importante verificar se este já tenha sido testado e validado.

Nenhum artigo aplicou a teoria de resposta ao item. Segundo Araujo, Andrade e Bortolotti (2012) é que ela leva em consideração o item particularmente, sem relevar os escores totais; portanto, as conclusões não dependem apenas do teste ou questionário, mas de cada item que o compõe.

No que tange a análise da consistência interna (Alpha de Cronbach) foram identificados 20 (vinte) artigos com a utilização da técnica. Para a validação do questionário deve-se utilizar a técnica de Alpha de Cronbach. Com esta técnica é possível estimar a confiabilidade do questionário aplicado em uma pesquisa. O Alpha de Cronbach, segundo Hora, Monteiro e Arica (2010) mede a correlação das respostas por meio da análise das respostas dadas, apresentando uma correlação média entre as perguntas. A correlação superior a 0,8 (zero vírgula oito) significa que o questionário é confiável. O coeficiente Alpha é calculado a partir da variância dos itens individuais e da variância da soma dos itens de cada avaliador de todos os itens de um questionário que utilizem a mesma escala de medição.

O uso do Alpha de Cronbach auxilia na purificação dos dados coletados, isso porque é realizado um novo cálculo do coeficiente de confiabilidade, desconsiderando um único item, e observando o comportamento do mesmo. Ao final, sumariza-se a purificação eliminando todos os itens cuja eliminação contribui para o aumento da confiabilidade da medição.

Quanto a normalidade dos dados, foi observado 4 (quatro) pesquisa que apresentaram tal teste. Os principais testes estatísticos têm como suposição a

normalidade dos dados, que deve ser verificada antes da realização das análises principais.

Existem diferentes testes para estimar a normalidade da distribuição dos dados. Por exemplo, no teste de Kolmogorov-Smirnov um resultado não significativo ($p > 0,05$) indica normalidade. Outros testes de normalidade incluem Anderson-Darling, Cramervon Mises e Shapiro-Wilk. Graficamente, a normalidade pode ser observada a partir de histogramas e Q-Q plots (FIGUEIREDO FILHO; SILVA JUNIOR, 2009).

Quando se tem amostras de tamanhos pequenos (tamanho menor ou igual a 10), não se recomenda utilizar o teste de normalidade dos dados. Isso porque o tamanho influencia no resultado dos testes. Desta forma, sugere-se utilizar estratégia não-paramétrica de análise.

Na análise da linearidade dos dados verifica-se se existe uma relação linear entre as variáveis e se os dados. O método usualmente conhecido para medir a correlação entre duas variáveis é o coeficiente de correlação linear de Pearson. Dos artigos na área de turismo estudados nesta pesquisa, apenas 3 (três) apresentaram ou mencionaram a utilização da técnica. O coeficiente de correlação linear de Pearson entre duas variáveis serve para medir a intensidade da relação linear entre elas, ou seja, a correlação.

Primeiramente, Pinotti e Moretti (2018) afirmam que deve se calcular os coeficientes de determinação de Pearson, mostrando a alternância das variáveis endógenas e indicando a qualidade do modelo ajustado. O coeficiente varia entre 0 e 1 e quanto maior o modelo, mais explicativo este se torna.

Quanto a análise da homoscedasticidade dos dados foi identificado 1 (um) estudo que utilizou a técnica. Uma das suposições para a aplicação da técnica da ANOVA é que a variância é igual em todos os níveis. Assim, a hipótese de homocedasticidade consiste em que os erros E_i são variáveis aleatórias de variância constante (σ^2).

A hipótese de ausência de multicolinearidade entre as variáveis explicativas de um estudo, ou seja, não são correlacionadas corresponde a análise de multicolinearidade dos dados. Nas pesquisas em Turismo observaram-se 5 (cinco) estudos que apresentaram formalmente o resultado do teste.

Para Medeiros et al. (2015) as variáveis da regressão foram extraídas da análise fatorial e, por esta razão, os problemas de multicolinearidade dos dados provenientes da relação entre as variáveis independentes podem ser resolvidos e verificados pela estatística de Durbin-Watson.

Considerando o que foi apresentado neste artigo, se propõe um conjunto de passos para que o levantamento e análise de dados primários em turismo gerem dados válidos e confiáveis.

Desenvolvimento do questionário: tipo de questões adequadas ao objetivo de estudo, estabelecer o público alvo, calcular amostra para nível de confiança e erro amostral aceitável, realizar pré-teste com especialistas.

Procedimentos preliminares à análise fatorial: Identificou-se a presença de missings (dados ausentes); identificação e exclusão de outliers com base no cálculo da distância Mahalanobis; pressuposto estatístico de normalidade; análise da linearidade dos dados; análise da homoscedasticidade; estudo de multicolinearidade.

Análise dos Dados: validação do instrumento por análise fatorial; análise da adequação do instrumento à TRI; análise de consistência interna pelo Alpha de Cronbach.

Em qualquer que seja o trabalho científico, a produção honesta do conhecimento está estreitamente associada ao seu rigor. Na literatura de metodologia da pesquisa quantitativa o rigor é caracterizado pela validade e confiabilidade do estudo, que tem por objetivo reduzir o erro dos valores observados frente aos valores “verdadeiros”.

5 CONSIDERAÇÕES FINAIS

O objetivo desta pesquisa é verificar se a geração de dados primários nas publicações nacionais em turismo é válido e confiável. Para tanto se analisou os periódicos da área com Qualis B2 ou superior. Os resultados apontaram para uma pequena quantidade de estudos com utilização de questionário. Dos que utilizaram questionário uma parcela pequena demonstrou ter utilizados técnicas estatísticas para reduzir o erro dos valores observados, gerando dados válidos e confiáveis.

Existe um espaço para o aprimoramento das pesquisas quantitativas, tanto do ponto de validação das informações quanto da confiabilidade. No que tange a validação, pouco mais de quinze por cento dos estudos envolvendo questionário explicitaram o uso de análise fatorial para validar sua validação e nenhum estudo fez mencionou ter utilizado a teoria da resposta ao item. No que tange a confiabilidade dos dados, menos de dez por cento dos estudos que envolveram questionário mencionaram ter feito análise da consistência interna, pouco mais de dois por cento mencionaram ter realizado análise de multicolinearidade dos dados e menos de dois por cento normalidade, linearidade e homoscedasticidade dos dados.

Os autores esperam que os resultados apresentados possam subsidiar futuras pesquisas quantitativas envolvendo questionário na área de Turismo. Conhecimento mais específico dos aspectos examinados nesta pesquisa pode fornecer orientações importantes para os pesquisadores fazendo com que as pesquisas da área se qualifiquem. Após anos de estruturação científica do turismo, pretende-se contribuir para o amadurecimento das pesquisas quantitativas na área, promovendo levantamento de dados primários válidos e confiáveis.

Espera-se que os resultados deste artigo sirvam como ponto inicial para o desenvolvimento de outras pesquisas. E, por fim, acredita-se que, com o aprofundamento dos estudos, será possível retroalimentar o processo de desenvolvimento científico.

REFERÊNCIAS

AAKER, D. A., KUMAR, V., DAY, G. S.. *Marketing research*. New York, NY: John Wiley & Sons. 2013.

ARAUJO, E. A. C. D., ANDRADE, D. F. D., BORTOLOTTI, S. L. V.. Teoria da resposta ao item. *Revista da Escola de Enfermagem da USP*, 43(spe), 1000-1008. 2009.

BORTOLOTTI, S. L. V., MOREIRA JUNIOR, F. D. J., BORNIA, A. C., SOUSA JÚNIOR, A. D., ANDRADE, D. D.. Avaliação do nível de satisfação de alunos de uma instituição de ensino superior: uma aplicação da Teoria da Resposta ao Item. *Gestão & Produção*, 19(2), 287-302. 2012

COHEN, J., COHEN, P., WEST, S. G., & AIKEN, L. S.. *Applied multiple correlation/regression analysis for the behavioral sciences*. UK: Taylor & Francis. 2003.

COSTA, W. F., TITO, A. L. A., BRUMATTI, P. N. M., DE OLIVEIRA ALEXANDRE, M. L.. Uso de instrumentos de coleta de dados em pesquisa qualitativa: um estudo em produções científicas de turismo. *Turismo, Visão e Ação*, 20(1), 02-28. 2018

CRONBACH, L. J.. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. 1951.

EVANS, J. R., MATHUR, A.. The value of online surveys: a look back and a look ahead. *Internet Research*, 28(4), 854-887. 2018.

FÁVERO, L. P., BELFIORE, P., SILVA, F. D., CHAN, B. L.. *Análise de dados: modelagem multivariada para tomada de decisões*. Rio de Janeiro: Campus/Elsevier. 2009.

FIGUEIREDO FILHO, D. B., SILVA JÚNIOR, J. A. D.. Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). *Revista Política Hoje*, 18(1). 2009.

GOMES, D.E., SANTOS, J.L.G., BORGES, J.W.P., ALVES, M.P., ANDRADE, D.F., ERDMANN, A.L. Teoria da resposta ao item nas pesquisas em saúde pública. *Rev enferm UFPE on line.*, Recife, 12(6):1800-12, jun., 2018.

HAIR, J. F., BLACK, W. C., BABIN, B. J., ANDERSON, R. E.. *Multivariate data analysis: A global perspective*. NJ: Pearson. 2010.

HAIR, J. F., HARRISON, D. E., RISHER, J. J.. Marketing Research in the 21st Century: Opportunities and Challenges. *Revista Brasileira de Marketing*, 17(5), 666-699. 2018.

HORA, H. R. M., MONTEIRO, G. T. R., ARICA, J.. Confiabilidade em questionários para qualidade: um estudo com o Coeficiente Alfa de Cronbach. *Produto & Produção*, 11(2), 85-103. 2010.

LILLIEFORS, H. W.. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318), 399-402. 1967.

MALHOTRA, N. K.. *Pesquisa de marketing: uma orientação aplicada*. Bookman Editora. 2012.

MEDEIROS, S. A., GOSLING, M., VERA, L. A. R.. Emoções em Experiências Negativas de Turismo: um estudo sobre a influência na insatisfação. *Revista Turismo em Análise*, 26(1), 188-215. 2015.

MONDO, T. S.; HALLMANN, G.; BURG, O. A qualidade de serviços em dois bairros turísticos de Florianópolis-SC Brasil: um estudo com o uso do Tourqual. *Revista Iberoamericana de Turismo- RITUR*, v. 8, n. 1, p.4-16, 2018.

PINOTTI, R. C., MORETTI, S. L. A.. Hospitalidade e Intenção de recompra na economia compartilhada: um estudo com equações estruturais em meios de hospedagem alternativos. *Revista Turismo em Análise*, 29(1), 1-18. 2018.

ROCHA, E. S., COSTA, L.A., ANJOS, F. A.. Perspectiva sistêmica nos artigos publicados nos anais dos seminários da anptur. *Revista Turismo, Visão e Ação*, 19(3), 614-631. 2017.

SANTOS, G. E. O., NETTO, A. P., WANG, X.. Análise de citações de periódicos científicos de turismo no Brasil: subsídios para a estimação de indicadores de impacto. *Revista Brasileira de Pesquisa em Turismo*, 11(1), 61-88. 2017.

SEIXAS, C. L., CARON, M. F.. Análise dos Artigos Publicados na Revista Turismo em Análise entre os Anos de 2004 e 2015. *Revista Turismo em Análise*, 29(1). 2018.

SOARES, D. J. M.. *Teoria clássica dos testes e teoria de resposta ao item aplicadas em uma avaliação de matemática básica*. Dissertação (Estatística aplicada a biometria), Universidade Federal de Viçosa, Viçosa, MG. 2018.

SOARES, T. C., LIMA, M. A.. Managerialism nas instituições de ensino superior brasileiras. *Revista Ciências Administrativas*, 24(2). 2018.

SWEENEY, D. J., WILLIAMS, T. A., ANDERSON, D. R. *Estatística aplicada à administração e economia*. São Paulo: CENGAGE Learning [Links]. 2013.

Quantitative research in tourism: are the data generated valid and reliable?

Abstract

The objective of this research is to verify if the generation of primary data in national publications in tourism is valid and reliable. For that, we analyzed the journals of the area with Qualis B2 or higher. The research process was carried out in three stages: Delimitation of the scope of analysis; Selection of articles; quantitative analysis, aiming to describe the main characteristics of the selected works. Considering that in the quantitative research rigor is characterized by the validity and reliability of the generated data, it was studied if the articles presented scientific rigor. We tried to identify in the articles that had quantitative field research with questionnaire if statistical procedures were adopted to valid and reliable results, such as: Factor Analysis; Item Response Theory; Analysis of internal consistency; Analysis of linearity of data; Analysis of the homoscedasticity of the data; Multicollinearity analysis of the data. The results pointed to a small number of studies adopting such techniques. In this way it is concluded that the field of research in Tourism needs to be concerned with generating valid and reliable primary data.

Keywords: *Tourism, Quantitative Research, Validity, Reliability, Questionnaire.*

Artigo recebido em 23/01/2019. Artigo aceito em 13/06/2019.