

Conclusão

O objetivo desta tese foi o de sugerir um procedimento inferencial para análise fatorial, mas precisamente estabelecer um critério para a significância para as cargas fatoriais, utilizando-se das técnicas de reamostragem *bootstrap* e *jackknife*, isto é, um critério para a significância de cargas fatoriais.

A análise fatorial é um tipo de processo destinado essencialmente a redução e à sumarização dos dados.

O *bootstrap* e o *jackknife* constituem técnicas *CIS* (Computer Intensive Statistics), que especificam o modelo da distribuição por amostragem empírica das estatísticas consideradas no estudo.

Estas metodologias de reamostragem não paramétricas são particularmente úteis quando é inviável pela metodologia tradicional se estabelecer a distribuição teórica de amostragens das estimativas investigadas, bem como seus parâmetros característicos. Esta dificuldade reside no fato de em muitas situações o comportamento regular dos parâmetros descritivos da distribuição por amostragem da variável aleatória em foco não se evidenciar tal facilmente ao pesquisador, variando de caso para caso ou quando a complexidade da natureza da estimativa analisada é um complicador para o estudo de sua regularidade via métodos tradicionais, como análise assintótica e por máxima verossimilhança.

As técnicas *CIS* funcionam como processos alternativos de estudos do comportamento probabilístico de variáveis aleatórias de interesse apelando para especificação das distribuições por amostragem através da força bruta da computação, simulando o universo de todas as estimativas possíveis para o estimador e calculando seu erro-padrão, seu viés e seu EMQ com base na experiência. Com estas estatísticas, é possível se obter intervalos de confiança e testes de hipótese saindo, então, do terreno puramente descritivo no que se refere à análise fatorial.

Para realizar o *link* da análise fatorial com as técnicas *bootstrap* e *jackknife*, o pesquisador elaborou dois algoritmos e destes algoritmos estabeleceu um programa em linguagem *R 2.1.1* e dois programas em linguagem de programação *SAS Versão 8*.

Os resultados *bootstrap* e *jackknife* se mostraram a princípio muito eficazes, pois revelaram erro-padrão e viés muito baixo para os estudos de casos considerados.

Os principais achados da pesquisa ou o balanço geral dos resultados fornecidos pelo *Método TIAF* com as bases testadas são:

- *A regularidade no comportamento geral das distribuições por amostragem das variáveis aleatórias cargas fatoriais de componentes principais reamostradas nos estudos de casos não se evidenciou: para cada exemplo se mostrou um resultado próprio;*
- *As teorias existentes sobre o comportamento do erro-padrão e do viés entre o bootstrap e o jackknife não se sustentaram em todos os estudos das estatísticas cargas fatoriais das componentes principais 1 e 2;*
- *Em todas as bases testadas, a precisão, o viés e o **EMQ** se revelaram satisfatórios, em torno de zero;*
- *As cargas fatoriais têm probabilidade de serem estatisticamente não significantes, mesmo se tratando de grandes amostras, nas estatísticas cargas fatoriais de componentes principais 1 e 2;*
- *A vantagem do critério de significância de cargas fatoriais exposto nesta tese é que o número de variáveis analisadas e o fator específico em exame são considerados no processo de significância das cargas fatoriais. As reamostragens são realizadas levando em consideração todas as variáveis do problema (quanto mais completa for a lista de variáveis melhor o teste de significância da análise fatorial) e para um fator específico F_j gerado em cada amostra bootstrap ou jackknife conforme seu poder de explicação dentro do modelo fatorial.*
- *Não houve uma regularidade no ajuste ou na aderência do modelo normal às estimativas de cargas fatoriais analisadas. O método jackknife foi onde à aderência ao modelo de probabilidade normal mais se evidenciou;*

- *O método jackknife foi o que teve, também, um saldo positivo maior com relação ao desempenho em relação ao método bootstrap para os casos analisados neste estudo;*
- *A análise da convergência no uso da técnica bootstrap indicou que com um $B=200$ já se tem a **Regularidade Estatística dos Resultados**. Contudo, para otimizar ao máximo a qualidade das distribuições amostrais resultantes ainda sugere-se um $B=1000$.*
- *Os resultados comparativos dos Métodos Tradicionais Hair e Anderson e o TIAF serviram para corroborar a tese de que o procedimento indutivo aqui proposto **tem eficácia bem satisfatória**, podendo-se afirmar que supera, principalmente no uso do valor-p, a eficácia do método tradicional.*
- *Concluindo, o método inferencial para significância de cargas fatoriais proposto nesta tese é uma alternativa para a garantia da qualidade dos resultados da análise fatorial melhor dos que os existentes atualmente, pois além de levar em consideração vários fatores relevantes, tem comparativamente desempenho melhor que os outros.*

O TIAF, segundo a opinião do autor, precisa ainda ser testado com maior número de aplicações ou base de dados. Fica sugerida a continuidade deste trabalho por outros pesquisadores.

Seguindo a mesma metodologia deste trabalho-aplicação de técnicas de reamostragem: bootstrap e jackknife-, pretende-se como trabalho futuro desenvolver um procedimento inferencial para a análise de conglomerados. Segundo Aaker (2001), é difícil avaliar a qualidade do processo de agrupamento. Não existem testes estatísticos padrão para garantir que o resultado não seja puramente aleatório. Apesar das tentativas de construção de vários testes para a confiabilidade estatística dos conglomerados, nenhum procedimento totalmente comprovado está disponível. A ausência de testes adequados provém da dificuldade de especificação de hipóteses nulas realísticas. O valor do critério de medida, a legitimidade do resultado, a aparência de uma hierarquia natural -

quando for empregado um método não hierárquico- e a confiabilidade de testes de divisão de amostra, tudo isto oferece informação útil. Entretanto, é difícil saber exatamente quais conglomerados são muito parecidos e quais objetos são difíceis de serem inseridos. Geralmente, é difícil selecionar um critério e programa de agrupamento por meio de outra referência que não a disponibilizada.

Fica sugestionado o desenvolvimento de estudos para identificação de procedimentos inferências para as demais técnicas de análise multivariadas.