

ELABORAÇÃO DE ÍNDICES MUNICIPAIS DE DESENVOLVIMENTO SOCIOECONÔMICO NO BRASIL

Etapas da Construção de um Índice

Marco Antonio Jorge

Neílson Santos Meneses

Introdução

O processo de elaboração de um índice não é simples. A figura 1 sintetiza as etapas de construção de um índice, as quais são detalhadas no restante do presente capítulo:



Figura 1: Etapas para Construção de um Índice. (Fonte: Scandar Neto ET AL (2008: 8))

Adoção de um Marco Conceitual:

Na medida em que um índice é calculado a partir de um volumoso conjunto de indicadores representativos de diversos temas ou dimensões, estes necessitam ser organizados de forma a garantir a coerência do instrumento. Esta é precisamente a função do marco ordenador, o qual pode “ser uma simples proposta de classificação dos indicadores segundo temas e subtemas” (SCANDAR NETO, 2006: 38), mas, se estiver relacionado a uma determinada concepção teórica ou discussão conceitual, torna-se um marco conceitual.

Dessa forma, o conceito de desenvolvimento sócio-econômico define o marco conceitual a partir do qual deve-se organizar a construção de um índice que tenha por finalidade representar tal fenômeno, permitindo que este instrumento seja melhor formulado. Se às dimensões econômica e social somar-se a dimensão ambiental, tem-se por marco ordenador o desenvolvimento sustentável, embora segundo Scandar Neto (2006) este não seja ainda um conceito preciso.

O fato de que a maioria das pessoas vive nas cidades e de que neste locus se reproduzem as relações sociais fundamenta a necessidade de elaboração de indicadores sociais urbanos. Além disso, é nas cidades que ocorrem as maiores pressões para consumo de recursos naturais – renováveis ou não – bem como a maior geração de poluição e de resíduos sólidos. Dessa forma, faz-se necessário incluir a dimensão ambiental na questão urbana (RIBEIRO, 2004).

Por esta razão, o conceito de qualidade de vida urbana incorpora a dimensão ambiental, além da pujança econômica e da equidade em termos econômicos e sociais (NAHAS ET AL, 2006b e NAHAS, 2002). Nesse sentido, o conceito de qualidade de vida urbana apresenta certa similaridade com o de desenvolvimento sustentável.

Uma maneira alternativa de organizar a construção de indicadores e índices é o enfoque nas carências ou privações enfrentadas por determinada população. Nessa linha, são formulados instrumentos para mensuração da vulnerabilidade social, pobreza ou do déficit social, isto é,

da noção do nível de carência da população no usufruto ou consumo de bens públicos, bens privados e serviços sociais entendidos como básicos e necessários segundo uma agenda de compromissos político-sociais (GUIMARÃES e JANNUZZI, 2004: 10).

A citação traz à tona a ideia de que as carências / privações são relativas, isto é, devem ser confrontadas com algum padrão de mínimo socialmente aceitável, o que, obviamente, depende de juízos de valor. Logo, trata-se de um conceito normativo e que varia ao longo do tempo, bem como de população para população.

Portanto, recolocando a questão em sua ordem lógica, primeiro é necessário se ter clareza sobre o fato, fenômeno ou processo que se quer medir, conhecer ou sobre o qual se vai intervir, para só então buscar os instrumentos que irão informar o seu comportamento (RIBEIRO, s/d: 3).

Em suma, o passo inicial para a construção de qualquer índice consiste na definição de um marco ordenador e, de preferência, atrelado a um conceito, isto é, um marco conceitual. Só então pode-se passar para a etapa seguinte, qual seja, a da seleção dos temas abarcados pelo marco, bem como dos indicadores primários que serão utilizados.

Seleção dos Indicadores Primários:

Alguns princípios devem pautar a escolha das variáveis que constituirão o índice. Dentre estes, pode-se mencionar (SCANDAR NETO, 2006: 46-7):

- i. relevância do indicador dentro da agenda política e social de cada população em dado momento e frente ao marco conceitual do índice;
- ii. validade em representar o conceito abstrato que se propõe mensurar, evitando múltiplas interpretações dos resultados obtidos. É o grau de proximidade entre o conceito e sua medida;
- iii. confiabilidade dos dados empregados na construção do indicador, aí inclusas a credibilidade do método, bem como da fonte produtora da informação;
- iv. grau de cobertura populacional e/ou territorial adequada à abrangência populacional / espacial do índice. Corresponde também à sua capacidade de ser reproduzido em vários espaços geográficos ou para diferentes populações;
- v. periodicidade de atualização das informações utilizadas no cálculo do indicador em prazo curto ou médio, possibilitando a construção de séries históricas que permitam sua comparação intertemporal.

- vi. sensibilidade do indicador para refletir mudanças no fenômeno medido, bem como captar os resultados da implementação de ações de políticas públicas;
- vii. inteligibilidade de sua elaboração, ou seja, do grau de compreensão para os agentes econômicos, aí inclusos formuladores de políticas, gestores públicos, beneficiários dos programas e população em geral, dos procedimentos utilizados em sua construção e
- viii. desagregabilidade da informação, isto é, a possibilidade de replicação do indicador para espaços territoriais e grupos populacionais específicos.

A disponibilidade dos indicadores depende fundamentalmente do grau de cobertura e da regularidade de atualização, as quais, por sua vez, são função do custo e da complexidade necessários à obtenção da informação.

Às características mencionadas Nahas (2002) acrescenta a possibilidade de georreferenciamento dos dados na forma compatível com a abrangência espacial dos indicadores e índices deles derivados.

A despeito da importância dos princípios mencionados, em muitos casos, na prática, é a disponibilidade de dados que acaba por definir a composição dos índices, particularmente no caso daqueles de abrangência intra-municipal:

Os indicadores produzidos no município não são voltados para questões intra-urbanas, desta forma, inexistem precisão quanto ao endereço, rua e número, ou estes são muitas vezes desprezados, ou ainda de qualidade duvidosa, ocasionando baixa precisão intra-urbana (FREI ET AL, 2005: 7).

Dessa forma, a disponibilidade de dados confiáveis acaba, muitas vezes, ficando restrita às informações censitárias, cuja periodicidade de atualização é inadequada para o monitoramento e avaliação de políticas públicas.

Uma vez escolhido o conjunto de variáveis que comporá o índice, surge outro problema: como comparar indicadores tão díspares, como p. ex., o PIB per capita – variável monetária – com a taxa de homicídios, mensurada em número de casos por cem mil habitantes e com o acesso à coleta de lixo, expresso em % de domicílios particulares permanentes ou mesmo da população de uma localidade?

É necessário que aos indicadores seja conferida a propriedade da comensurabilidade, isto é, que todos sejam convertidos a uma base comum. Para tanto, faz-se necessária a transformação destes indicadores.

Técnicas de Transformação de Variáveis:

Para viabilizar a comensurabilidade dos indicadores, três técnicas são normalmente empregadas: a padronização das variáveis (transformação pelo escore z), a transformação 0-1 e a escala de razão. Vamos a elas:

A Padronização das Variáveis (Transformação pelo Escore Z)

Consiste em procedimento aplicado frequentemente em trabalhos estatísticos (vide, p. e.x, JORGE, 2000, cap. 4), no qual subtrai-se de cada observação a média amostral e divide-se o resultado pelo desvio-padrão da mesma, tal que o escore z de cada variável x é

$$Z_i = \frac{X_i - \bar{X}}{s} \quad (1)$$

Onde:

Z_i = escore z da i-ésima variável x

X_i = i-ésima variável x

\bar{X} = média de x

S = desvio-padrão da variável x

Este procedimento, embora amplamente utilizado, possui dois inconvenientes:

- gera valores negativos
- produz z com média zero e desvio-padrão um.

Escore negativos

Não se deve utilizar o módulo matemático como alternativa pois este, ao tornar positivos todos os escores z, não permite diferenciar os piores dos melhores desempenhos, já que um resultado bastante negativo tornar-se-ia altamente positivo quando da aplicação do módulo, falseando seu desempenho real.

Dada a fórmula empregada na padronização, todos os valores que estiverem abaixo da média produzirão **escores negativos**, o que não representa qualquer problema teórico, mas pode causar certa estranheza ou dificultar o entendimento dos resultados quando da divulgação do índice.

Uma maneira de evitar o surgimento de valores negativos é adicionar à fórmula uma constante superior a determinado mínimo o que, além de produzir números estritamente positivos, auxilia na diferenciação do desempenho final mensurado pelo índice.

Como exemplo pode-se mencionar a padronização utilizada no cálculo do IDM – Índice de Desenvolvimento Municipal – elaborado pela Superintendência de Estudos Econômicos e Sociais do Estado da Bahia (SEI), onde cada escore padronizado é obtido de acordo com a seguinte fórmula:

$$Z_i = \left| \left[\frac{X_i - \bar{X}}{s} \right] \times 100 \right| + 5.000 \quad (2)$$

A constante de 5.000, além de produzir escores estritamente positivos, permite maior amplitude de escala no resultado final, evitando a superposição de índices ou o surgimento de resultados muito próximos, o que auxilia a diferenciar o desempenho medido pelo índice (JORGE ET AL, 2010).

Outra implicação da padronização é que os escores gerados possuem média zero e desvio padrão igual a um, o que prejudica a posterior aplicação da análise de componentes principais para definição do peso de cada indicador e/ou dimensão no cálculo do índice final.

Outra observação com relação ao procedimento discutido nesta seção diz respeito aos casos em que os valores mais elevados representam as piores situações. Assim, um valor maior para a taxa de homicídios, p. ex., tem significado diverso de um valor mais elevado para a expectativa média de vida. Para compatibilizar os resultados e evitar confusões de interpretação, no caso da transformação de variáveis socialmente indesejadas deve-se utilizar a fórmula:

$$Z_i = \frac{\bar{X} - X_i}{s} \quad (3)$$

Ainda assim, surgiriam tanto valores positivos como negativos. Novamente, para se contornar tal situação, poder-se-ia utilizar o procedimento expresso em (2) ou empregar uma forma alternativa de transformação,

A Escala de Intervalo Linear (Transformação 0-1)

Este procedimento consiste em subtrair de cada observação o valor mínimo da amostra e dividir o resultado pela diferença entre os valores máximo e mínimo para aquela variável. Dessa forma, a variável transformada v de cada variável original x é:

$$V_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (4)$$

Onde:

V_i = variável transformada da i -ésima variável x

X_i = i -ésima variável x

X_{\max} = valor máximo da variável x

X_{\min} = valor mínimo da variável x

A aplicação deste procedimento resulta em valores obrigatoriamente contidos no intervalo aberto entre zero e um, evitando, portanto, o surgimento de valores negativos e sua conseqüente dificuldade de interpretação. Para variáveis socialmente indesejadas,

pode-se novamente alterar a fórmula para compatibilizar seu resultado diante do de variáveis socialmente positivas, tal como segue:

$$V_i = 1 - \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} = \frac{X_{\max} - X_i}{X_{\max} - X_{\min}} \quad (5)$$

Quanto aos valores obtidos, o máximo será um e o mínimo, zero. Pode-se operar com estes valores ou estabelecer limites – superior e inferior – de modo que para observações cujo valor exceda os limites impostos seja imputado o limite máximo ou mínimo, conforme o caso. Este procedimento é chamado de truncamento e serve para diminuir a influência de valores extremos no resultado final (SCANDAR NETO, 2006: 54), o que é particularmente importante no caso de índices compostos por poucos indicadores.

Não por acaso, esta é a técnica utilizada no cálculo do IDH: para a expectativa média de vida no nascimento foram estabelecidos os limites superior e inferior de 85 e 25 anos. De forma análoga, para o PIB real per capita (corrigido pelo poder de compra) foram definidos como limites superior e inferior US\$ 40.000 e US\$ 100 (PNUD, 1998: 108).

Como o IDH é composto somente por quatro variáveis, um valor bastante destoante para uma delas poderia influenciar significativamente no resultado final. No caso do PIB per capita, para limitar ainda mais este risco, é feita uma correção adicional: os valores são transformados em logaritmos neperianos, reduzindo a contribuição de PIB's muito elevados no desempenho do IDH.

Levantamento feito por Nahas et al (2006a) mostra que mais da metade dos índices existentes no Brasil à época (2005) utilizavam esta técnica para transformação de variáveis. Estes autores mencionam ainda outra técnica para obtenção da comensurabilidade: a escala de razão.

A Escala de Razão

Tal procedimento consiste em dividir cada indicador original por um valor de referência, tal que:

$$V_i = \frac{X_i}{V_r} \quad (6)$$

Onde:

V_i = variável transformada da i -ésima variável x

X_i = i -ésima variável x

V_r = valor de referência

O valor de referência pode ser o valor máximo observado para determinada variável ou uma meta que se pretenda atingir. Caso V_r não seja o máximo ou trate-se de uma meta já atingida, novamente pode-se utilizar o truncamento, atribuindo o valor 1 para V_i nos casos em que $X_i > V_r$. O objetivo é garantir que V_i fique entre zero e um.

Uma vez que todos os componentes do índice tenham sido trazidos para a mesma unidade, deve-se definir a ponderação de cada um deles, ou seja, sua importância no cômputo do resultado final.

Métodos de Aglutinação das Variáveis:

Esta etapa é crucial para a definição do índice, pois retrata a importância atribuída pela população e/ou pelos gestores públicos a cada componente e dimensão do índice na medida em que estabelece pesos a cada um deles.

Tais pesos podem ser definidos arbitrariamente ou com o auxílio de técnicas de análise multivariada.

Arbitrariamente, os pesos podem ser estabelecidos de duas maneiras: pela média aritmética simples, a qual impõe pesos iguais a todos os componentes, ou pela média aritmética ponderada que atribui pesos diferenciados a cada componente.

Média Aritmética Simples e Ponderada

Tecnicamente, a média aritmética simples é representada pela seguinte fórmula (KUME, 1993: 155):

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (7)$$

Onde:

\bar{X} = média de x

X_i = i -ésima variável x

n = número de variáveis

Σ = somatório de todas as i -ésimas variáveis x

Logo, a fórmula acima equivale a:

$$\bar{X} = \frac{1}{n} X_1 + \frac{1}{n} X_2 + \frac{1}{n} X_3 + \dots + \frac{1}{n} X_n \quad (8)$$

o que implica em atribuir exatamente o mesmo peso ($1/n$) para cada componente no cálculo do índice.

Scandar Neto (2006) e Scandar Neto et al (2008) enumeram como propriedades desejáveis da média aritmética simples:

- i. simplicidade de cálculo;
- ii. facilidade de compreensão por não-especialistas;
- iii. carregamento do máximo possível de informação de cada indicador para o índice sintético. Este ponto merece uma consideração adicional: caso fosse estabelecido um peso muito pequeno para uma determinada variável, uma mudança em seu desempenho não seria adequadamente refletida no resultado final do índice. Este ponto será abordado com maior profundidade na próxima seção do trabalho. Por fim,
- iv. minimização da máxima discordância possível entre índices obtidos através de diferentes ponderações. Suponha que o mesmo conjunto de indicadores fosse combinado através de diversas técnicas de aglutinação para a construção de um índice. Assim, a utilização de três técnicas gerará três índices diferentes a partir das mesmas variáveis. Scandar Neto (2006: 80-82) mostra que a média produz um resultado que minimiza as variâncias da combinação extraída do conjunto de variáveis (maior r^2) comparativamente às técnicas de análise multicritério e de análise de componentes principais.

Consulta

Para a autora, a população deve participar ativamente do processo de construção do índice, particularmente na etapa de seleção de indicadores e dimensões, já que tal participação auxilia a captar a percepção da população acerca de determinadas condições, além de conferir maior legitimidade política ao índice, bem como maior controle sobre a distribuição de recursos públicos, nos casos em que esta se baseie nos resultados do índice

Uma crítica que pode ser feita ao emprego da média como método de aglutinação segue na linha de que os índices devem refletir prioridades de pesquisadores, gestores públicos ou da população, isto é, o estabelecimento das ponderações é uma decisão política e, portanto, deve ser conduzido através de amplo processo de **consulta** (NAHAS, 2002).

Este problema pode ser contornado com a aplicação da média aritmética ponderada, onde os pesos dos diversos componentes diferem, refletindo as prioridades dos agentes envolvidos na construção do índice, tal que:

$$\bar{X} = aX_1 + bX_2 + cX_3 + \dots + zX_n \quad (9)$$

Onde:

X = média de x

a ... z = peso de cada indicador

n = número de variáveis

Com isso, pode-se atribuir pesos maiores para indicadores de saúde, p. ex., caso esta dimensão seja a principal prioridade de determinada comunidade. Uma população

bastante vulnerável à violência urbana, por sua vez, pode estabelecer ponderações maiores para os indicadores de segurança pública no cálculo do índice.

Os dois métodos de aglutinação mencionados, porém, estão sujeitos à crítica de que os pesos devem ser obtidos de forma “neutra”, isto é, através de técnicas científicas e não de decisão política. Dessa forma, técnicas de análise multivariada deveriam ser empregadas para estabelecer tais pesos. Neste trabalho serão mencionadas de forma resumida três destas técnicas: análise de componentes principais, análise fatorial e análise multicritério.

Técnicas de Análise Multivariada de Dados

A análise fatorial é uma técnica que abrange um conjunto de métodos estatísticos capazes de reduzir a quantidade de dados originais e encontrar fatores ou dimensões latentes por trás destes. Nesse sentido, a análise de componentes principais é também uma variante de análise fatorial.

Análise Fatorial e Análise de Componentes Principais

De acordo com Hair Jr et al (2005:94):

o propósito geral de técnicas de análise fatorial é encontrar um modo de resumir a informação contida em diversas variáveis originais em um conjunto menor de novas dimensões compostas ou variáveis estatísticas (fatores) com uma perda mínima de informação.

Trata-se também de um conjunto de técnicas de interdependência, isto é, busca-se explorar o potencial explicativo das variáveis ao invés de prever ou analisar o comportamento de uma variável dependente (como no caso da regressão linear múltipla, p. ex.).

Dessa forma, são métodos adequados à construção de um índice caso sejam capazes de reduzir todas as dimensões envolvidas a um único fator. As técnicas operam da seguinte maneira:

A análise de componentes principais produz combinações lineares das variáveis originais, tal que (JORGE, 2000: 141-2):

$$Z_i = a'_1 y \quad (10)$$

Onde:

Z_i = i-ésimo componente principal

a'_1 = ponderação

y = vetor de variáveis originais

Variáveis

Também este tópico, por ser muito técnico, foge ao escopo deste trabalho e não será abordado aqui. Para maior detalhamento vide Hair Jr. et al (2005, cap. 3) ou Scandar Neto (2006).

A ponderação, a grosso modo, é o vetor característico normalizado da matriz de covariância formada a partir da matriz ortogonal das variáveis originais, ao passo em que os componentes principais são combinações lineares dessas **variáveis**.

A análise fatorial, por sua vez, produz fatores ou dimensões que estão por trás dos dados originais e a partir dos quais estes poderiam surgir. Assim, a análise fatorial busca explicar a correlação ou variância comum entre as variáveis originais (comunalidades) enquanto a análise de componentes principais foca a variância amostral total. A figura 2 ajuda a diferenciar as duas técnicas:



Figura 2: Tipos de Variância na Matriz Fatorial. (Fonte: Hair Jr. et al (2005: 101)

A área hachurada representa a variância extraída pela técnica enquanto a área em branco representa a variância perdida. Nota-se que a variância total da amostra pode ser dividida em três partes: a variância comum ao conjunto de variáveis, a variância específica de cada variável e a variância do erro decorrente de componentes aleatórios ou falhas de medida.

Na análise de componentes principais procura-se obter fatores que expliquem a variância amostral total. O primeiro fator ou componente extraído seria, então, aquele que explica a maior parte da variância total e pode ser entendido como o melhor resumo possível dos dados originais, já que implica em menor perda de poder explicativo. Já o segundo fator ou componente a ser extraído é aquele que explica a maior parte da variância remanescente, depois de removido dos dados o efeito do primeiro fator. Aqui a técnica impõe uma restrição: o segundo fator deve ser ortogonal ao primeiro.

Na análise fatorial, por sua vez, busca-se obter fatores que captam o que existe de comum nos dados, ou seja, o foco desta técnica é a variância comum.

Sob certas condições, ambas as técnicas produzem resultados similares: basta que o número de variáveis seja superior a trinta ou que as comunalidades (variância compartilhada) sejam superiores a 0,60 para a maioria das variáveis (HAIR JR. ET AL, 2005: 99-101). Como se pode observar, a multicolinearidade não é um problema neste modelo, pelo contrário, é desejável.

Em princípio, apesar de ambas as técnicas gerarem resultados similares sob estas condições, a análise de componentes principais parece mais adequada à construção de índices do que a análise fatorial na medida em que busca obter fatores produzidos pelas variáveis e não encontrar dimensões subjacentes ou latentes dos dados. Devemos nos lembrar que “os índices surgem das variáveis e não o contrário” (JORGE, 2000: 142). Além disso, requer menos pressupostos que o modelo de análise fatorial. Assim, Jorge

(2000) e Scandar Neto (2006) utilizam a análise de componentes principais para a construção de índices. O Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE), no entanto, emprega a análise fatorial no cálculo do Índice de Desenvolvimento Municipal (IDM) dos municípios cearenses (IPECE, 2008). Rodríguez Rodríguez (1988: 104) também relata que, à época de seu estudo, dentre ambas, esta última técnica era a mais difundida entre os pesquisadores espanhóis.

No caso da construção de um índice sintético, porém, ambas as técnicas possuem características potencialmente limitantes, dentre as quais pode-se destacar:

i. em termos ideais, a análise deveria produzir um único fator, o qual seria já o índice, mas nada garante que somente um fator seja gerado. No caso de haver mais de um fator ou componente principal toma-se o primeiro como representativo do índice, mas isto pode implicar em significativa perda de informação, caso a variância explicada por ele seja um percentual pequeno da variância total (componentes principais) ou da variância comum (análise fatorial);

ii. a técnica não gera uma solução única, ou seja, um mesmo conjunto de variáveis originais pode produzir fatores com cargas (pesos) diferenciados para cada variável;

iii. podem surgir pesos com sinal negativo, já que estes são definidos em função das interrelações entre os dados e não com base conceitual. Dessa forma, podem surgir resultados contraintuitivos que dificultarão a interpretação do fator. A título de exemplo, Jorge (2000) tinha como objetivo elaborar um índice de adequação setorial a políticas de emprego, que seria uma função da geração de postos de trabalho qualificados, isto é, com registro em carteira, maior tempo de permanência no emprego e maior nível educacional. O emprego da análise de componentes principais resultou em dois fatores, ou seja, não foi possível criar o índice. Pior: um deles atribuiu sinal negativo ao nível de registro em carteira. Assim, de acordo com este fator, seriam melhor ranqueados os setores que gerassem maior número de empregos sem carteira assinada (JORGE, 2000: 143-4). Como interpretar este resultado?

iv. outra característica particularmente problemática é a instabilidade dos fatores criados, já que, a cada novo experimento novos fatores podem ser gerados a partir das variáveis originais. Em outros termos: ainda que fossem produzidos fatores baseados nas mesmas variáveis, os pesos de cada variável poderiam ser diferentes. Dessa forma, a cada atualização a composição do índice seria diferente, o que praticamente inviabilizaria comparações intertemporais dos resultados, permitindo apenas comparações transversais, isto é, em um determinado ponto do tempo;

v. por fim, como já mencionado, os pesos das variáveis dependem de suas correlações, sendo definidos pela técnica. Assim, indicadores que apresentam as maiores variâncias terão maior peso no cálculo do índice. Esta determinação “científica” das ponderações pode ser vista como um ponto positivo – justamente o argumento em defesa da escolha das técnicas de análise mul-

tivariada –, mas pode ser vista também como um ponto negativo por alguns, visto que não leva em conta as preferências e prioridades da população e dos gestores na determinação dos pesos.

Análise Multicritério

Trata-se de uma técnica que objetiva

enfrentar situações nas quais um decisor, atuando com racionalidade, deveria resolver um problema em que vários objetivos, nem sempre bem conformados e por vezes até conflitantes, deveriam ser alcançados simultaneamente (GOMES; ARAYA e CARIGNANO, 2004 apud SCANDAR NETO, 2006:65)

Assim, é adequada a situações nas quais necessita-se tomar decisões objetivas, de forma imparcial, em busca da escolha da alternativa ótima (JANNUZZI, 2005), ou seja, é uma técnica que deve ser aplicada a problemas de otimização simultânea de diferentes funções.

Este é precisamente o caso de fenômenos multifacetados como o desenvolvimento sócio-econômico, cuja mensuração deve representar uma série de variáveis / indicadores. Suponha que no caso do desenvolvimento municipal cada cidade apresente melhor performance em um ou alguns indicadores. Qual será, então, o município melhor posicionado em termos do marco conceitual referenciado?

Multicritério

Existem diferentes abordagens para aplicação da análise multicritério: escola americana, escola francesa e métodos interativos (JANNUZZI, 2005). Neste trabalho será focado o método PROMETHEÉ II da escola francesa que parece ser o mais difundido. Vide p. ex. Scandar Neto (2006) e Silvério, Ferreira e Rangel (2007).

Para chegar à resposta, a análise **multicritério** segue alguns passos: o primeiro deles consiste em estabelecer um critério para comparação do desempenho municipal com relação a um determinado atributo (indicador). Para tanto, existem dois tipos de critérios gerais (SILVÉRIO, FERREIRA e RANGEL, 2007:3). Scandar Neto (2006), p. ex., faz uso do verdadeiro critério ou critério usual que consiste em

atribuir 1 para o município (alternativa) cujo desempenho seja superior ao de outro e 0 no caso contrário. Assim, o critério é sempre aplicado a pares de alternativas. Dessa forma, no segundo passo ordena-se os municípios frente ao conjunto de variáveis que compõem o índice.

O passo seguinte consiste em calcular, para cada município, os fluxos positivos e negativos de sua posição frente aos outros, considerando-se o conjunto de atributos para se obter o resultado líquido. Assim, uma vez computado o resultado líquido, o método permite estabelecer rankings.

Uma vantagem do método é a possibilidade de se atribuir pesos aos diferentes critérios, incorporando assim as preferências da população, dos gestores, enfim dos agentes envolvidos na construção do índice.

No caso do método PROMETHEÉ II, com a utilização do verdadeiro critério para estabelecimento das preferências, duas limitações surgem quando o objetivo é a construção de um índice (SCANDAR NETO, 2006):

- i. quando da não atribuição de pesos aos critérios, a técnica considera apenas o posto do município em cada indicador e não o seu valor;
- ii. ao atribuir valores binários para a comparação par a par, o índice gerado pelo método terá uma distribuição próxima da distribuição normal, mas esta não é necessariamente a situação dos dados originais. Porém,

se o objetivo desta síntese, além de produzir um ordenamento, é também ter um caráter descritivo, espera-se que (as) distribuições dos indicadores originais e dos indicadores sintéticos tenham formas parecidas (SCANDAR NETO, 2006: 79-80)

Assim, embora o emprego de outros critérios, bem como a definição de pesos diferenciados para cada critério possa abrandar o problema, parece-nos que a análise multicritério seja mais útil para ordenar os resultados de diversos índices construídos com auxílio de outras técnicas do que para estabelecer as ponderações de um índice específico.

Apresentação e Avaliação dos Resultados:

Uma vez estabelecidas as ponderações para o conjunto de componentes do índice, procede-se o seu cálculo. Resta, então, a última etapa do processo que consiste na apresentação e na interpretação dos resultados obtidos.

Embora a figura 1 tenha sido reproduzida tal como se encontra em Scandar Neto (2006) e Scandar Neto et al (2008), parece-nos mais razoável inverter a ordem dos dois últimos retângulos, tendo em vista que a avaliação dos resultados pode ser facilitada ou não pela maneira com que estes são trazidos a público, ou seja, a interpretação depende da forma de apresentação dos resultados.

Esta última deve valorizar sobretudo o entendimento pelo público, visto que os índices representam fenômenos complexos, alguns cujo marco conceitual não está ainda definido de forma totalmente consensual. Somando-se a isso a dificuldade da população em lidar com dados estatísticos, este requisito torna-se de difícil alcance, mas de crucial importância, já que a bagagem cultural e informacional de gestores, formuladores e dos estratos mais bem organizados da sociedade coloca tais grupos em posição de vantagem para lidar com os índices, o que poderia possibilitar a manipulação política ou econômica dos resultados, caso estes não sejam compreendidos pelo grande público (NAHAS, 2002: 24).

Scandar Neto (2006) observa que a apresentação dos resultados deveria ter por objetivo uma descrição dos componentes do índice (variáveis, dimensões) e não apenas posicioná-lo em relação a outros desempenhos para a construção de rankings, sejam estes municipais, estaduais ou mesmo internacionais. Em outros termos, trata-se de apresentar não apenas o índice sintético, mas também os indicadores que o compõem.

Isto ajudaria, inclusive, a abrandar a dicotomia entre painéis de indicadores e índices sintéticos, a ser abordada no próximo capítulo. Antes, porém, cabe chamar a atenção para um possível problema decorrente da representação gráfica do tipo radar, como a da figura 3:



Figura 3: Representação Gráfica do Índice DNA Brasil 2004. (Fonte: UNICAMP, NEPP (2004) apud Scandar Neto (2006: 86).

No DNA Brasil cada indicador é transformado com base na escala de razão, descrita na seção 2.3. Um mérito desta forma de representação é a inclusão de todos os indicadores em uma única figura, o que facilita a percepção das principais forças e fraquezas de uma dada localidade (respectivamente a menor e maior distância para a situação pretendida, representada pelo limite do círculo). Apesar do mérito, há um problema: se o índice equivale à área da figura interna em proporção à figura externa (o círculo), ele variará conforme a ordem de disposição dos indicadores. Para ilustrar, tome-se um exemplo (SCANDAR NETO, 2006: 86-8): imagine-se um índice composto por quatro indicadores, cujos valores estão representados na tabela abaixo:

Tabela 1: Indicadores Hipotéticos e seus Valores

INDICADOR	VALOR
A	8
B	6
C	5
D	3

Fonte: Scandar Neto (2006: 87)

Imagine-se ainda que a situação pretendida para cada indicador seja equivalente a dez. Uma representação do tipo radar, obedecendo à sequência A-B-C-D dos indicadores (Sequência 1) resulta na figura 4.1:

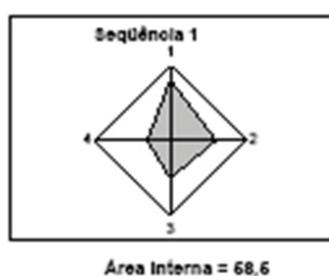


Figura 4.1

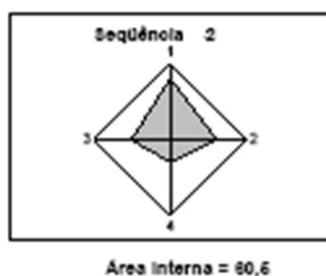


Figura 4.2

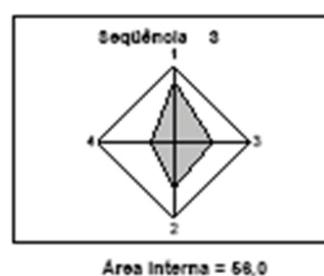


Figura 4.3

Fonte: Scandar Neto (2006: 87)

A área interna do polígono da figura 4.1 equivale a 58,5. Se pegarmos os mesmos indicadores e ordenarmos na sequência A-B-D-C (Sequência 2), como representado na figura 4.2, formaremos um polígono de área igual a 60,5(!). O mesmo ocorre com uma terceira forma de ordenação – sequência A-C-B-D (Sequência 3) – mostrada na figura 4.3: o polígono formado possui agora área 56,0.

Assim, a ordem de disposição dos indicadores determina o valor final do índice, ou seja, há um sério problema de consistência dos resultados.

A solução proposta pelo autor para sanar o problema consiste em tomar-se um quadrado de lado igual a um e dividi-lo em j retângulos iguais e paralelos, sendo j o número de variáveis que compõem o índice. Cada variável pode, então, ser expressa na escala de razão, o que produzirá um histograma com j barras de altura igual a 1 e base $1/j$. Interessante nesta solução proposta por Scandar Neto é que, independentemente da maneira como as barras sejam ordenadas, a área da figura interna formada será sempre a mesma, tal que:

$$S = \frac{1}{j}X_1 + \frac{1}{j}X_2 + \frac{1}{j}X_3 + \dots + \frac{1}{j}X_j = \frac{1}{j} \sum_{i=1}^j X_i \quad (11)$$

que nada mais é do que a fórmula da média aritmética simples vista anteriormente. Assim, gera-se um resultado estável, independentemente da ordenação das variáveis e equivalente à média dos indicadores transformados.

Resolvido o problema resta tratar de uma questão de fundo: o que deve ser mostrado: o índice sintético ou o conjunto de indicadores que o compõem? Ou, ainda, de forma mais incisiva: é válido o cálculo de um índice sintético para retratar fenômenos complexos e multifacetados como o desenvolvimento sócio-econômico ou deve-se analisar um painel de indicadores relativos a um determinado marco conceitual?